UNIVERSITÄT PASSAU
Fakultät für Informatik und Mathematik

Dissertation

Eine Anwendung der Invariantentheorie auf das Korrespondenzproblem lokaler Bildmerkmale

Dipl.-Inf. (Univ.) Thomas Stadler

Eingereicht zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für Informatik und Mathematik
der Universität Passau

Betreuer:
Prof. Dr. Martin Kreuzer

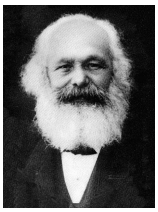
Oktober 2015

Heinrich Heine[1]

Anfangs wollt ich fast verzagen,
und ich glaubt, ich trüg es nie.
Und ich hab es doch getragen,
aber fragt mich nur nicht, wie.

[1]Bildquelle: http://de.wikipedia.org/wiki/Heinrich_Heine vom 21.03.2013.

Zusammenfassung



Karl Marx[2]

Jeder Mensch und jedes Buch läßt sich auf drei Seiten zusammenfassen, und diese drei Seiten lassen sich auf zwei Zeilen reduzieren.

Als sich in der ersten Hälfte des 19. Jahrhunderts zunehmend mehr bedeutende Mathematiker mit der Suche nach Invarianten beschäftigten, konnte natürlich noch niemand vorhersehen, dass die Invariantentheorie mit Beginn des Computerzeitalters in der Bildverarbeitung bzw. dem Rechnersehen ein äußerst fruchtbares Anwendungsgebiet finden wird. In dieser Arbeit wird eine neue Anwendungsmöglichkeit der Invariantentheorie in der Bildverarbeitung vorgestellt. Dazu werden lokale Bildmerkmale betrachtet. Dabei handelt es sich um die Koordinaten einer Polynomfunktion bzgl. einer geeigneten Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, die die zeitintegrierte Sensorinputfunktion auf lokalen Pixelfenstern bestmöglich approximiert. Diese Bildmerkmale werden in vielen Anwendungen eingesetzt, um Objekte in Bildern zu erkennen und zu lokalisieren. Beispiele hierfür sind die Detektion von Werkstücken an einem Fließband oder die Verfolgung von Fahrbahnmarkierungen in Fahrerassistenzsystemen. Modellieren lässt sich die Suche nach einem Muster in einem Suchbild als Paar von Stereobildern, auf denen lokal die affinlineare Gruppe $\mathrm{AGL}_2(\mathbb{R})$ operiert. Will man also feststellen, ob zwei lokale Pixelfenster in etwa Bilder eines bestimmten dreidimensionalen Oberflächenausschnitts sind, ist zu klären, ob die Bildausschnitte durch eine Operation der Gruppe $\mathrm{AGL}_2(\mathbb{R})$ näherungsweise ineinander übergeführt werden können. Je nach Anwendung genügt es bereits, passende Untergruppen G von $\mathrm{AGL}_2(\mathbb{R})$ zu betrachten. Dank der lokalen Approximation durch Polynomfunktionen induziert die Operation einer Untergruppe G eine Operation auf dem reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Damit lässt sich das Korrespondenzproblem auf die Frage reduzieren, ob es eine Transformation $T \in G$ gibt so, dass $p \approx q \circ T$ für die zugehörigen Approximationspolynome $p, q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ gilt. Mit anderen Worten, es ist zu klären, ob sich p und q näherungsweise in einer G-Bahn befinden, eine typische Fragestellung der Invariantentheorie. Da nur lokale Bildausschnitte betrachtet werden, genügt es weiter, Untergruppen G von $\mathrm{GL}_2(\mathbb{R})$ zu betrachten. Dann erhält man sofort auch die Antwort für das semidirekte Produkt $\mathbb{R}^2 \rtimes G$. Besonders interessant für Anwendungen ist hierbei die spezielle orthogonale Gruppe $G = \mathrm{SO}_2(\mathbb{R})$ und damit insgesamt die eigentliche Euklidische Gruppe $\mathbb{R}^2 \rtimes \mathrm{SO}_2(\mathbb{R})$. Für diese Gruppe und spezielle Pixelfenster ist das Korrespondenzproblem bereits gelöst. In dieser Arbeit wird das Problem in eben dieser Konstellation ebenfalls gelöst, allerdings auf elegante Weise mit Methoden der Invariantentheorie.

[2]Bildquelle: https://de.wikipedia.org/wiki/Karl_Marx vom 21.09.2015.

Der Ansatz, der hier vorgestellt wird, ist aber nicht auf diese Gruppe und spezielle Pixelfenster begrenzt, sondern leicht auf weitere Fälle erweiterbar. Dazu ist insbesondere zu klären, wie sich sogenannte fundamentale Invarianten von lokalen Bildmerkmalen, also letztendlich Invarianten von Polynomfunktionen, berechnen lassen, d.h. Erzeugendensysteme der entsprechenden Invariantenringe. Mit deren Hilfe lässt sich die Zugehörigkeit einer Polynomfunktion zur Bahn einer anderen Funktion auf einfache Weise untersuchen.

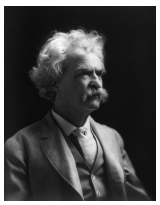
Neben der Vorstellung des Verfahrens zur Korrespondenzfindung und der dafür notwendigen Theorie werden in dieser Arbeit Erzeugendensysteme von Invariantenringen untersucht, die besonders „schöne" Eigenschaften besitzen. Diese schönen Erzeugendensysteme von Unteralgebren werden, analog zu Gröbner-Basen als Erzeugendensysteme von Idealen, SAGBI-Basen genannt („Subalgebra Analogs to Gröbner Bases for Ideals"). SAGBI-Basen werden hier insbesondere aus algorithmischer Sicht behandelt, d.h. die Berechnung von SAGBI-Basen steht im Vordergrund. Dazu werden verschiedene Algorithmen erarbeitet, deren Korrektheit bewiesen und implementiert. Daraus resultiert ein Software-Paket zu SAGBI-Basen für das Computeralgebrasystem ApCoCoA, dessen Funktionalität in diesem Umfang in keinem Computeralgebrasystem zu finden sein wird. Im Zuge der Umsetzung der einzelnen Algorithmen konnte außerdem die Theorie der SAGBI-Basen an zahlreichen Stellen erweitert werden.

Inhaltsverzeichnis

Abbildungsverzeichnis

KAPITEL 1

Einleitung



Mark Twain[3]

Jeder Mensch mit einer neuen Idee ist ein Spinner, bis die Idee Erfolg hat.

Zu Beginn dieser Arbeit wollen wir passenderweise ein Bild in den Mittelpunkt stellen, das seit 2010 in einem Blog eines Mathematiklehrers unter dem Titel „Mathematik ist ... - wie dieses Bild" zu finden ist.

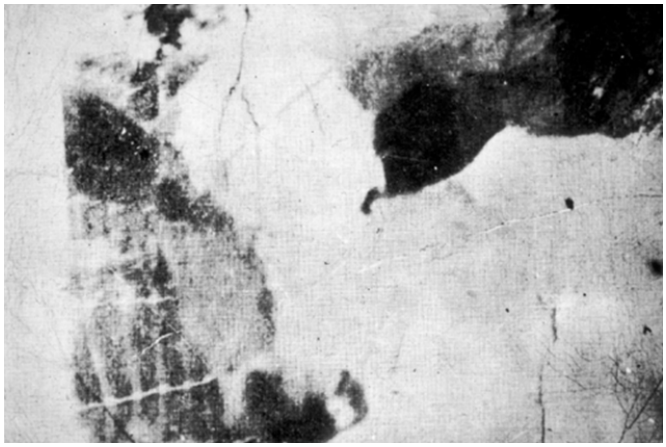


Abbildung 1.1: Mathematik ist ... - wie dieses Bild.[4]

Aber was hat dieses Bild nun mit Mathematik zu tun? Der Mathematiklehrer Jan-Martin Klinge vergleicht in seinem Blog nicht das, was hier abgebildet ist, mit Mathematik - auch wenn viele das vielleicht anders sehen mögen, sondern er vergleicht den „Erkennungsprozess", der nötig ist, um überhaupt irgendetwas auf diesem Bild zu sehen, mit dem Verstehensprozess in der Mathematik. Gerade für viele Schüler erscheint die Mathematik undurchschaubar zu sein, bis eine Thematik verstanden wurde, ab dann ist dieses Thema plötzlich ganz einfach und macht

[3]Bildquelle: http://en.wikipedia.org/wiki/Mark_Twain vom 25.04.2015.
[4]Bildquelle: http://halbtagsblog.de/schule/mathematik-ist-wie-dieses-bild/ vom 10.12.2012

kaum noch Probleme. Genauso verhält es sich mit diesem Bild: Viele erkennen hier lange Zeit gar nichts (und wir wollen an dieser Stelle auch nicht verraten, was zu erkennen ist), aber sobald man einmal erkannt hat, was hier zu sehen ist, wird man es immer wieder auf einen Blick erkennen. Sicher beschreibt die Problematik, die hinter diesem Bild steckt, auf analoge Weise sehr treffend das Auf und Ab während der Entstehung einer Disserationsschrift, aber das ist nicht der Grund, warum dieses Bild hier an den Anfang gestellt wurde. Es beschreibt auch äußerst eindrucksvoll, dass es selbst für das menschliche Auge nicht immer einfach ist, in einem Bild etwas zu erkennen. Wie viel schwerer muss es also für einen Computer sein, auch nur annähernd dasselbe zu leisten? Die Probleme lassen sich also durchaus vergleichen: Um eine Computer „sehen" zu lassen, muss man sich überlegen, wie man aus einer Ansammlung von ganzen Zahlen, üblicherweise zwischen 0 und 255, den sogenannten Grauwerten[5], Objekte erkennen kann.

Seit jeher ist die Erkennung, Identifizierung und Lokalisierung von Objekten eines der zentralen Themen der digitalen Bildverarbeitung. So gibt es eine Vielzahl an Anwendungen mit hohem Praxisbezug, die im Kern genau diese Problematik beinhalten. Typische Beispiele sind (siehe auch Abbildung 1.2):

- Lokalisierung und Identifizierung von Werkstücken auf einem Fließband (vgl. hierzu z.B. [Pis02]), wobei hier eventuell auch im Rahmen der Qualitätssicherung der anschließende Vergleich des lokalisierten Objekts mit einem sogenannten „Masterpiece" eine Rolle spielen kann (vgl. [Pis02]). Es ist also möglich, dass neben der Frage, *wo* ein Objekt zu finden ist, auch die Frage nach der Güte der Übereinstimmung mit einem Muster im Vordergrund steht. Beide Probleme werden auch unter dem Begriff **Transformationspassung** zusammengefasst, der sich beispielsweise auch Pisinger in [Pis02] gewidmet hat.
- Erkennung und Verfolgung von Objekten, wie z.B. von Verkehrszeichen, Fahrbahnmarkierungen, Personen, Fahrzeugen, usw. in Fahrerassistenzsystemen (siehe z.B. [Haa00]),
- Erkennung von Schriftzeichen (siehe z.B. [Bis11]),

um nur einige wenige Anwendungen zu nennen. Gerade der Bereich Fahrerassistenzsysteme ist zur Zeit hochaktuell, aber auch im Bereich der automatisierten Fertigung gibt es natürlich noch zahlreiche weitere Anwendungen, in denen die Erkennung von Objekten im Mittelpunkt steht. Jedes Objekt innerhalb eines Bildes besteht aus einzelnen Punkten. Wenn man so will, sind also Punkte die kleinstmöglichen „Objekte", die sich erkennen und zuordnen lassen. Punkte in unterschiedlichen Bildern, die Abbilder ein und desselben dreidimensionalen Punktes auf einem Objekt der realen Welt sind, werden **korrespondierende Punkte** genannt. Die Detektion solcher korrespondierender Punkte, wie in Abbildung 1.2 (unten) dargestellt, hat ebenfalls einen hohen Stellenwert in der Praxis, wie z.B. für die 3D-Rekonstruktion (vgl. hierzu [Sta07]). Damit ein Computer all diese Aufgaben leisten kann, ist es notwendig, das Zitat des berühmten amerikanischen Fotografen Ansel Adams (1902–1984) gleichsam als Aufforderung zu verstehen, der wir im Laufe der Arbeit auch noch nachkommen wollen:

> „Ein Foto wird meistens nur angeschaut - selten schaut man in es hinein."

Das Szenario des Erkennens von Objekten in Bildern lässt sich vereinfacht als Paar von **Stereobildern** modellieren, wie sie beispielsweise in Abbildung 1.2 (unten) dargestellt sind; zwei Bilder einer Szene aus verschiedenen Blickwinkeln. Eines der beiden Bilder können wir dabei als **Musterbild**, das andere als **Suchbild** betrachten, d.h. das im Musterbild zu sehende „Muster", also ein bestimmtes Objekt, wie z.B. ein Werkstück, ein Schriftzeichen, ein Verkehrszeichen oder

[5] In dieser Arbeit werden ausschließlich Grauwertbilder betrachtet. Farbbilder werden ggf. zuvor in Grauwertbilder konvertiert (siehe dazu [FDFH92], S. 585 ff.).

einfach nur ein Punkt, soll im Suchbild wiedergefunden werden. Ein Prozess der für den Menschen häufig „auf einen Blick", also in der Regel sehr einfach und schnell erledigt ist. Wir wollen nun im Folgenden die Geometrie der Stereobilder näher beleuchten, um die Problemstellung genauer erfassen zu können.

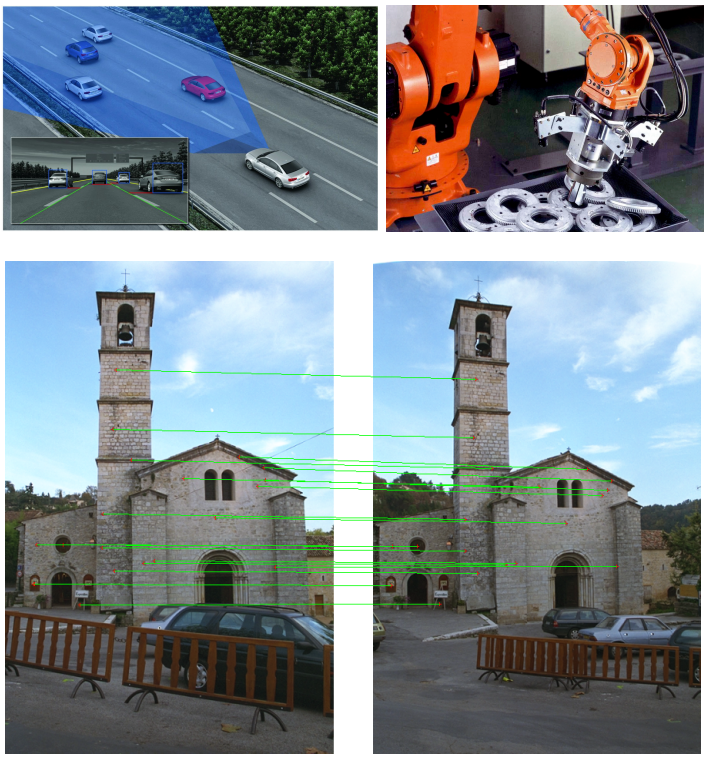


Abbildung 1.2: Beispiele von Anwendungen der digitalen Bildverarbeitung: Fahrerassistenzsysteme, Erkennung von Werkstücken am Fließband, Suche nach korrespondierenden Punkten in Stereobildern.[6]

1.1 Die Geometrie der Stereobilder

Stereobilder können grundsätzlich auf verschiedene Arten entstehen: Es kann sich um ein und dieselbe Kamera handeln, die eine bestimmte Szene aus unterschiedlichen Perspektiven abbildet[7] oder um zwei (baugleiche) Kameras, die unterschiedlich positioniert sind und ein und dieselbe Szene abbilden. Geometrisch sind diese beiden Szenarien jedoch nicht zu unterscheiden und beide resultieren in zwei Bildern einer Szene, wobei natürlich manche Teile des einen Bildes im anderen verdeckt sein können, was eine Objekterkennung nicht gerade einfacher werden lässt, aber auch nicht unmöglich machen sollte.

[6] Bildquellen:

(1) Audi Adaptive Cruise Control:
http://www.motoreport.de/assistenzsysteme-i-adaptive-cruise-control-acc-im-audi-q5-video/ vom 30.07.2015.

(2) http://www.pressebox.de/pressemitteilung/profactor-gmbh/3D-PROMPT-Bildverarbeitung-PlugPlay-fuer-den-Roboter/boxid/207375 vom 30.07.2015.

(3) Im Original von http://www.robots.ox.ac.uk/ vom 16.09.2014, nach Punktzuordnung aus [Sta07].

[7] Wir nehmen hier stets eine Bildentstehung durch eine Kamera an (vgl. hierzu auch Abschnitt 9.1). Andere Arten der Bildentstehung, wie z.B. durch Scannen, Röntgen, Computer-Tomographie o.ä. werden nicht betrachtet.

Für die geometrische Modellierung wollen wir nun von einem Zwei-Kamera-Szenario ausgehen, was man auch als **Stereosystem** bezeichnet, d.h. insbesondere werden wir voraussetzen, dass die Kamerazentren C und C' verschieden sind. Auch hier kann man zwei Arten der Anordnung unterscheiden: Die Kameras können aus unterschiedlichen Positionen auf ein bestimmtes Objekt zentriert ausgerichtet sein, d.h. die optischen Achsen (vgl. hierzu Abschnitt 9.1.1) schneiden sich in einem Konvergenzpunkt, oder die optischen Achsen verlaufen parallel. Eine Anordnung mit parallelen optischen Achsen wird auch ein **achsenparalleles Stereosystem** genannt. Die verschiedenen Anordnungen sind in Abbildung 1.3 skizziert, links der allgemeine Fall, rechts eine achsenparallele Anordnung. Die Geometrie, die diese Anordnungen beschreibt, wird als **Epipolargeometrie**, manchmal auch als **Stereogeometrie**, bezeichnet. Die Verbindungsgerade der beiden optischen Zentren C und C' heißt die **Basislinie** oder **Stereobasis** bzw. kurz die **Basis** des Stereosystems.

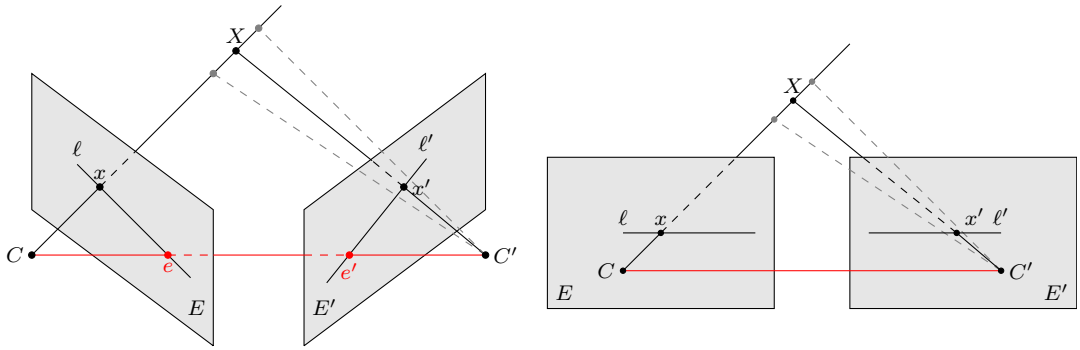


Abbildung 1.3: Darstellungen zur Epipolargeometrie

Als mathematisches Modell für die Kameras wollen wir das für unsere Zwecke auch vollkommen ausreichende, einfache Lochkameramodell verwenden (vgl. hierzu Abschnitt 9.1.1). Hierbei wird durch Koordinatentransformationen $\tau, \tau' : \mathbb{R}^3 \to \mathbb{R}^3$ zunächst ein Punkt $X \in \mathbb{R}^3$ im sogenannten Weltkoordinatensystem in das jeweilige kartesische Kamerakoordinatensystem mit Ursprung im Kamerazentrum abgebildet. Beide Transformationen sind Bewegungen im euklidischen Raum $\mathbb{R}^3$, also Elemente der eigentlichen Euklidischen Gruppe $\mathrm{Iso}_3^+(\mathbb{R})$. Die Bildebenen E bzw. E' werden üblicherweise als projektive Ebenen aufgefasst, d.h. im jeweiligen Koordinatensystem mit der Ebene $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_3 = 1\}$ identifiziert (näheres hierzu in Abschnitt 9.1.1). Die Abbildung $\mathrm{pr} : \mathbb{R}^3 \setminus \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_3 = 0\} \to \mathbb{R}^3$, definiert durch

$$\mathrm{pr}(x_1, x_2, x_3) = \left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1\right),$$

ist die Zentralprojektion am Ursprung auf die Ebene $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_3 = 1\}$. Um zu unterscheiden, in welchem Kamerakoordinatensystem wir uns befinden, schreiben wir pr_C bzw. $\mathrm{pr}_{C'}$ für die Projektion im Kamerakoordinatensystem bzgl. C bzw. C'. Somit erhalten wir die beiden Bildpunkte $\widehat{x} = \mathrm{pr}_C \circ \tau(X)$ und $\widehat{x}' = \mathrm{pr}_{C'} \circ \tau'(X)$ des Objektpunkts X in den Bildebenen E bzw. E'. Diese werden nun mittels einer Koordinatensystemtransformation $\iota : E \to \mathbb{R}^2$ bzw. $\iota' : E' \to \mathbb{R}^2$ in zweidimensionale Punkte im Bildkoordinatensystem umgerechnet. Die beiden Abbildungen ι und ι' kodieren die sogenannten intrinsischen Kameraparamter. Da wir baugleiche Kameras voraussetzen, können wir also $\iota = \iota'$ annehmen. Die Komposition der drei genannten Abbildungen ergibt die Kameraabbildungen (vgl. im Detail Definition 9.1.1):

$$\kappa = \iota \circ \mathrm{pr}_C \circ \tau \quad \text{und} \quad \kappa' = \iota \circ \mathrm{pr}_{C'} \circ \tau'.$$

Die beiden Kameraabbildungen liefern die zweidimensionalen Punkte $x = \kappa(X)$ und $x' = \kappa'(X)$ im Bildkoordinatensystem, die **korrespondierende Punkte**. Eine Kameraabbildung ist offensichtlich nicht umkehrbar, da die Projektion pr nicht injektiv ist, was auch in Abbildung 1.3 gut zu erkennen ist. Das von der zweiten Kamera aufgenommene Bild dieses Strahls ist eine Gerade in der zweiten Bildebene, die zu x korrespondierende **Epipolargerade** ℓ' (vgl. [HZ06], S. 241). Analog ist ℓ die zu x' korrespondierende Epipolargerade. Beide Epipolargeraden liegen in einer Ebene, die von den Kamerazentren C und C' sowie dem Objektpunkt X aufgespannt wird. Diese Ebene bezeichnet man entsprechend als **Epipolarebene** (vgl. [HZ06], S. 241). Die Basislinie schneidet beide Epipolargeraden in der jeweiligen Bildebene in den Punkten e und e', die als **Epipole** bezeichnet werden (vgl. [HZ06], S. 240).

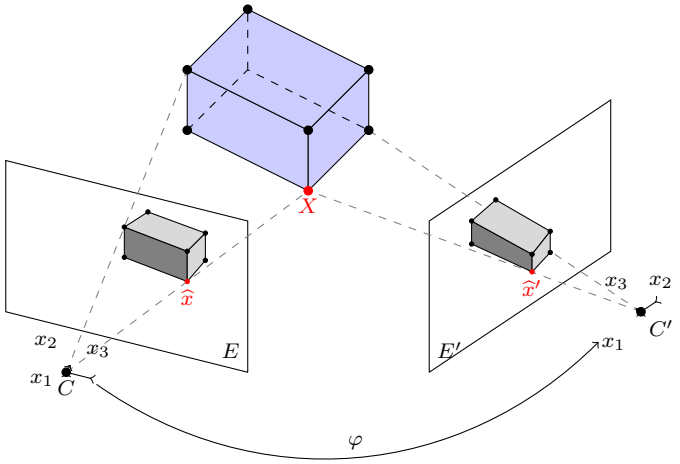


Abbildung 1.4: Schematische Darstellung der durch eine Bewegung bedingten Änderung der Perspektive.

Wir wollen die Beziehung zwischen korrespondierenden Bildpunkten x und x' noch aus einem anderen Blickwinkel beleuchten, indem wir die mathematische Beschreibung der Transformationen angeben, die ein Bild eines Stereobildpaares in das andere überführen (vgl. die schematische Darstellung in Abbildung 1.4). Dazu werden wir nun zur Vereinfachung ohne Einschränkung annehmen, dass das Weltkoordinatensystem und das Kamerakoordinatensystem bzgl. des Zentrums C übereinstimmen. Mittels einer Bewegung $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ kann der Wechsel vom Koordinatensystem bzgl. C in das Koordinatensystem bzgl. C' vollzogen werden (siehe auch Abbildung 1.4), d.h. insbesondere gilt $\varphi(C) = C'$ und $\varphi(E) = E'$ Unter der obigen Annahme, dass das Koordinatensystem bzgl. C mit dem Weltkoordinatensystem zusammenfällt, gilt weiter $\tau = \mathrm{id}_{\mathbb{R}^3}$ und $\tau' = \varphi$. Somit wird ein beliebiger Punkt $X \in \mathbb{R}^3$, für den wir ohne Einschränkung annehmen können, dass weder X noch $\varphi(X)$ in der Ebene $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_3 = 0\}$ - betrachtet im jeweiligen Koordinatensystem - enthalten sind, auf $\widehat{x} = \mathrm{pr}_C(X)$ in die Bildebene E und auf $\widehat{x}' = \mathrm{pr}_{C'} \circ \varphi(X)$ in die Bildebene E' projiziert. Insbesondere erhalten wir dann für den Punkt $\widehat{x} = (x_1, x_2, 1)$ in der Bildebene E den Bildpunkt $\widehat{x}' = \mathrm{pr}_{C'} \circ \varphi(x_1, x_2, 1)$ in der Bildebene E'. Wegen $\mathrm{pr}_{C'}|_{E'} = \mathrm{id}_{E'}$ gilt $\mathrm{pr}_{C'} \circ \varphi|_E = \varphi|_E$. Fassen wir E bzw. E' als projektive reelle Ebene $\mathbb{P}^2_{\mathbb{R}}$ auf, so ist mit anderen Worten $\varphi|_{\mathbb{P}^2_{\mathbb{R}}} : \mathbb{P}^2_{\mathbb{R}} \to \mathbb{P}^2_{\mathbb{R}}$ eine **projektive Abbildung** (vgl. [HZ06] , Definition 2.9, S. 32), die die korrespondierenden Punkte $\widehat{x} \in E$ und $\widehat{x}' \in E'$ aufeinander abbildet. Wir setzen zur besseren Lesbarkeit nun $\psi := \varphi|_{\mathbb{P}^2_{\mathbb{R}}}$. Dann gibt es weiter eine reguläre Matrix $\mathcal{H} \in \mathrm{GL}_3(\mathbb{R})$ der Form

$$\mathcal{H} = \begin{pmatrix} \mathcal{B} & t \\ v & \lambda \end{pmatrix} \tag{1.1.1}$$

mit $\psi(x_1, x_2, 1) = \mathcal{H} \cdot (x_1, x_2, 1)^{\mathrm{tr}}$ und $t, v \in \mathbb{R}^2$, $\mathcal{B} \in \mathrm{GL}_2(\mathbb{R})$ sowie $\lambda \in \mathbb{R}$ (vgl. [HZ06], S. 41). Wechseln wir auch hier für einen kurzen Moment die Blickrichtung, indem wir uns vorstellen, dass wir nur eine Kamera betrachten, die eine bestimmte sich bewegende Szene zu unterschiedlichen Zeitpunkten abbildet, was sich geometrisch vom Szenario mit zwei Kameras nicht unterscheidet, dann beschreibt die Abbildung ψ die Bewegung der realen Welt in der Bildebene. Abbildung 1.4 versucht diesen Sachverhalt darzustellen. Die Umrechnung in Bildkoordinaten mittels ι wollen wir zur Vereinfachung vernachlässigen, da eine Anwendung von ι keine Änderung an der geometrischen Anordnung bewirkt.

Damit können wir festhalten, dass Stereobilder durch projektive Transformationen ineinander übergeführt werden können. Auch in Abbildung 1.4 ist es gut zu erkennen, dass korrespondierende Flächen (in der Abbildung ist ein Beispiel dunkelgrau markiert) durch eine projektive Transformation auseinander hervorgehen. Die Menge der projektiven Abbildungen $\psi : \mathbb{P}^2_{\mathbb{R}} \to \mathbb{P}^2_{\mathbb{R}}$ bilden mit der Komposition als Verknüpfung eine Gruppe, die zweidimensionale **projektive lineare Gruppe** $\mathrm{PGL}_2(\mathbb{R})$ (vgl. [HZ06], Abschnitt 2.3, S. 32 ff.). Etwas anders ausgedrückt kann man also sagen, dass die projektive lineare Gruppe auf den Stereobildern *operiert*.

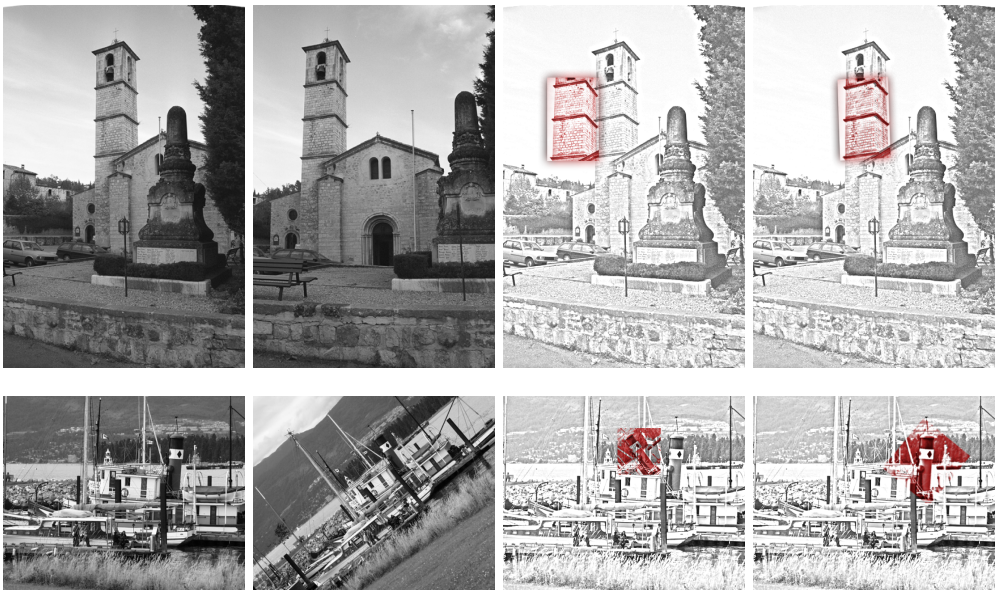


Abbildung 1.5: Schematische Darstellung der Wirkungsweise geometrischer Transformationen auf Stereobilder: Die ersten beiden Bilder jeder Zeile sind das betrachtete Stereobildpaar. Diese beiden Bilder werden übereinander gelegt, wobei nur ein bestimmter Ausschnitt (rot) betrachtet wird. Dieser Ausschnitt lässt sich näherungsweise in Deckung bringen mit dem korrespondierenden Ausschnitt im ersten Bild.[8]

Die meisten Anwendungen betrachten jedoch nur lokale, eng begrenzte Gebiete in Bildern und untersuchen deren Transformationen. Abbildung 1.5 versucht anschaulich einen gewissen Eindruck davon zu vermitteln, die in der Praxis betrachteten Gebiete sind jedoch wesentlich kleiner und umfassen nur wenige Pixel. Für lokale Bildbereiche ist die Annahme gerechtfertigt, dass der abgebildete Oberflächenausschnitt planar ist. Deshalb ist es ausreichend, sich auf eine Untergruppe der projektiven linearen Gruppe zu beschränken: auf die zweidimensionale **affine lineare Gruppe** $\mathrm{AGL}_2(\mathbb{R})$. Je nach Anwendungsfall kann es sogar genügen, sich auf spezielle Untergruppen von $\mathrm{AGL}_2(\mathbb{R})$ zu beschränken. Fasst man $\mathrm{AGL}_2(\mathbb{R})$ als Matrixgruppe auf, so sind

[8] Bildquellen: http://www.robots.ox.ac.uk/ vom 16.09.2014.

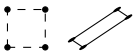
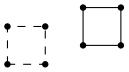
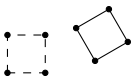
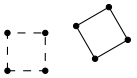
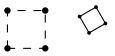
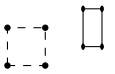
die Elemente dieser Gruppe aufbauend auf Gleichung (1.1.1) von der Form

$$\begin{pmatrix} \mathcal{A} & t \\ 0 \quad 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{R})$$

mit einer invertierbaren Matrix $\mathcal{A} \in \mathrm{GL}_2(\mathbb{R})$ und einem Vektor $t \in \mathbb{R}^2$. Da sich die Bildebene auf kanonische Weise mit der reellen Ebene $\mathbb{R}^2$ identifizieren lässt, werden wir die Elemente von $\mathrm{AGL}_2(\mathbb{R})$ als Transformationen $T : \mathbb{R}^2 \to \mathbb{R}^2$ in der euklidischen Ebene auffassen. Somit gilt:

$$\mathrm{AGL}_2(\mathbb{R}) = \{T : \mathbb{R}^2 \to \mathbb{R}^2 : \text{ Es gibt } \mathcal{A} \in \mathrm{GL}_2(\mathbb{R}),\ t \in \mathbb{R}^2 \text{ mit } T(x) = \mathcal{A} \cdot x + t\}.$$

Die Elemente der Gruppe $\mathrm{AGL}_2(\mathbb{R})$, also die affinen Transformationen der Ebene, werden auch als **geometrische Transformationen** oder **geometrische Bildoperationen** bezeichnet (vgl. [HZ06], Abschnitt 2.4, S. 37 ff.). Je nach Wahl der Parameter $\mathcal{A}$ und t ergeben sich somit unterschiedliche Untergruppen und Teilmengen von $\mathrm{AGL}_2(\mathbb{R})$, die unterschiedliche Wirkungen auf lokale Bildbereiche nach sich ziehen. In der folgenden Tabelle sind die bedeutendsten Transformationen aufgelistet und deren jeweilige Wirkung grob skizziert.

Parameter		Art der Transformation	Gruppe	Wirkung
$\mathcal{A} \in \mathrm{GL}_2(\mathbb{R})$	$t \in \mathbb{R}^2$	Affine Transformation	$\mathrm{AGL}_2(\mathbb{R})$	
$\mathcal{A} = \mathcal{I}_2$	$t \in \mathbb{R}^2$	Translation um den Vektor t	$\mathrm{Trans}_2(\mathbb{R})$	
$\mathcal{A} \in \mathrm{SO}_2(\mathbb{R})$	$t = 0$	Rotation	$\mathrm{SO}_2(\mathbb{R})$	
$\mathcal{A} \in \mathrm{O}_2(\mathbb{R})$	$t = 0$	Rotation oder Drehspiegelung	$\mathrm{O}_2(\mathbb{R})$	
$\mathcal{A} \in \mathrm{O}_2(\mathbb{R})$	$t \in \mathbb{R}^2$	Bewegung bzw. Kongruenzabbildung	$\mathrm{Iso}_2(\mathbb{R})$	
$\mathcal{A} \in \mathrm{SO}_2(\mathbb{R})$	$t \in \mathbb{R}^2$	Eigentliche Bewegung	$\mathrm{Iso}_2^+(\mathbb{R})$	
$\mathcal{A} = \alpha \cdot \mathcal{B}$ mit $\mathcal{B} \in \mathrm{SO}_2(\mathbb{R})$	$t \in \mathbb{R}^2$	Ähnlichkeitstransformation mit Skalierungsfaktor $\alpha \in \mathbb{R} \setminus \{0\}$	$\mathrm{Sim}_2(\mathbb{R})$	
$\mathcal{A} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}$	$t \in \mathbb{R}^2$	Skalierung in x bzw. y-Richtung um den Faktor s_x bzw. s_y		
$\mathcal{A} = \begin{pmatrix} 1 & s_x \\ s_y & 1 \end{pmatrix}$	$t \in \mathbb{R}^2$	Scherung in x bzw. y-Richtung um den Faktor s_x bzw. s_y		

1.2 Die Bildverarbeitung als natürliches Anwendungsgebiet der Invariantentheorie

Wie wir also gerade gesehen haben, operieren je nach Anwendungsfall unterschiedliche Gruppen auf den Stereobildern. Damit verbunden ist eine Operation der Gruppe auf unterschiedliche „Elemente" der Bilder, wie z.B. auf die Bildpunkte. Die Bildverarbeitung bzw. das Rechnersehen ist somit geradezu ein natürliches Anwendungsgebiet der sogenannten **Invariantentheorie** (siehe insbesondere Kapitel 6). Die Invariantentheorie selbst ist eine vergleichsweise junge mathematische Disziplin, die in der ersten Hälfte des 19. Jahrhunderts zarte Knospen bildete und

um 1900 ihre erste Blütezeit erleben durfte. Dennoch kann die Invariantentheorie in dieser verhältnismäßig kurzen Zeit auf eine äußerst wechselhafte Geschichte zurückblicken (siehe dazu Abschnitt 6.1). So rasant ihr Aufstieg im 19. Jahrhundert war, so schnell geriet sie bis Mitte des 20. Jahrhunderts auch wieder in Vergessenheit. Ihr Erwachen aus diesem „Dornröschenschlaf" in den letzten gut 30 Jahren verdankt sie letztendlich dem Computerzeitalter, da nun Dank Einsatz von Computern Probleme lösbar waren, die zuvor als nahezu unlösbar erschienen.

Die heutzutage gern als „moderne Algebra" bezeichnete Neuausrichtung der klassischen Algebra in den letzten Hundert Jahren geht im Wesentlichen zurück auf Erkenntnisse, die der Invariantentheorie entsprungen sind. Hier sind besonders die bedeutenden Sätze von David Hilbert zu nennen, die er in Arbeiten zur Invariantentheorie bewiesen hatte, dort aber eher „Randnotizen" bildeten: Der Nullstellensatz, der Basissatz und der Syzygiensatz. Die seit den 1980er Jahren stattfindende zweite Blütezeit der Invariantentheorie führte auch dazu, dass die Invariantentheorie in den unterschiedlichsten mathematischen Disziplinen Einzug hielt und verschiedenste Anwendungen hervorbrachte. Gregor Kemper und Harm Derksen nennen in [DK02] beispielsweise neben der Bildverarbeitung unter anderem folgende weitere Anwendungsgebiete der Invariantentheorie:

- Projektive Geometrie (siehe z.B. [Stu08]),
- Berechnung von Kohomologie-Ringen (siehe z.B. [AM04]),
- Berechnung von Galois-Gruppen (siehe z.B. [Sta73] und [GK00]),
- Lösung algebraischer Gleichungssysteme unter Berücksichtigung von Symmetrien (siehe z.B. [Stu08]),
- Dynamische Systeme (siehe z.B. [GG99] und [GL98]) mit Anwendungen in Chemie, Physik und Ingenieurswissenschaften (siehe z.B. [JMS84], [CH92] oder [CT95]),
- Graphentheorie (siehe z.B. [ACG96] und [Thi08]),
- Kombinatorik (siehe z.B. [Sta79a] und [Sta79b]),
- Codierungstheorie (siehe z.B. [Slo77]),
- Materialwissenschaften (siehe z.B. [Has80] und [Hel93]).

Eine wahre Fundgrube für Anwendungen der Invariantentheorie in der Bildverarbeitung ist das 1992 erschienene Buch *Geometric Invariance in Computer Vision* ([MZ92]) von Joseph L. Mundy und Andrew Zisserman. Aber bevor wir uns ein paar Beispiele von Anwendungen in der Bildverarbeitung betrachten, wollen wir kurz erklären, worum es in der Invariantentheorie geht. Dazu betrachten wir eine affine Varietät V über einem Körper K und eine Gruppe G, die auf V regulär durch $(a, v) \mapsto a(v)$ für alle $a \in G$ und $v \in V$ operiert. In unseren Fällen wird es sich stets um sogenannte linear algebraische Gruppen handeln. Durch $f^a(v) := f(a^{-1}(v))$ wird nun eine Operation auf dem Koordinatenring von V induziert. Eine Funktion $f \in K[V]$ mit $f^a = f$ für alle $a \in G$ wird als **invariant** bezeichnet und die Menge

$$K[V]^G = \{f \in K[V] : \ f^a = f \text{ für alle } a \in G\},$$

die im Mittelpunkt der Invariantentheorie steht, heißt der **Invariantenring** von G bzgl. V. In den meisten Fällen werden wir sehr einfache affine Varietäten, nämlich n-dimensionale K-Vektorräume, betrachten, deren Koordinatenring isomorph zum Polynomring $K[x_1, \ldots, x_n]$ ist, d.h. $x_1, \ldots, x_n$ stehen für die Koordinaten der Vektoren $v \in V$ bzgl. einer Basis von V. Eines der zentralen Themen der Invariantentheorie ist die Suche nach einem endlichen Erzeugendensystem von $K[V]^G$, sofern dieses existiert. Laut dem Hilbertschen Endlichkeitssatz ist dies für eine

spezielle Klasse von Gruppen, den sogenannten linear reduktiven Gruppen, stets gewährleistet. Eine weitere Frage, die für viele Anwendungen von zentraler Bedeutung ist, lässt sich ebenfalls mit Invariantentheorie gut beantworten: Gibt es für zwei Elemente $v, w \in V$ ein $a \in G$ mit $w = a(v)$? Mit anderen Worten, wie lässt sich feststellen, ob sich v und w in einer **Bahn** bzw. in einem **Orbit** befinden? Die Fragen nach Erzeugendensystemen und der Klassifikation in Bahnen wird uns auch vordergründig beschäftigen. Nun aber wollen wir in vereinfachter Form ein paar Beispiele bestehender Anwendungen der Invariantentheorie in der Bildverarbeitung angeben, um eine Vorstellung von den Möglichkeiten, die die Invariantentheorie der Bildverarbeitung bietet, zu erhalten. Diese Beispiele sind zum Teil auch in [MZ92] zu finden sowie in [DK02] als Konzepte enthalten.

Beispiel 1.2.1. (Beispiele von Anwendungen der Invariantentheorie in der Bildverarbeitung)

a) Wir betrachten n Punkte auf einem Objekt, wie z.B. den hier dargestellten Kirchturm, wobei wir an dieser Stelle nicht näher auf die Detektion dieser Punkte eingehen wollen. Diese Punkte sollen sich näherungsweise in einer Ebene befinden, die wie in Abbildung 1.6 nicht senkrecht zur Bildebene verläuft. Die n Punkte sollen weiter so gewählt sein, dass

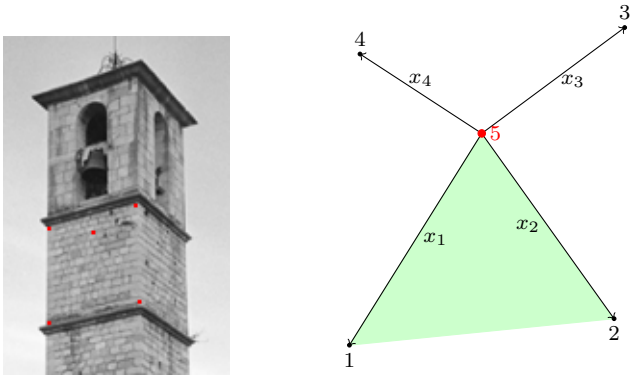


Abbildung 1.6: Beispiel für die Operation der affinen linearen Gruppe auf Punktmengen.

keine drei Bildpunkte kollinear sind. Wir wählen nun einen Punkt explizit aus, ohne Einschränkung soll dies der n-te Punkt sein, und bilden jeweils den Verbindungsvektor vom n-ten Punkt zu allen anderen Punkten. Seien $x_1, \ldots, x_{n-1} \in \mathbb{R}^2$ diese $n-1$ Verbindungsvektoren. Wie wir wissen, operiert die affine lineare Gruppe $\mathrm{AGL}_2(\mathbb{R})$ auf diesem Bild, also auch auf diesen $n-1$ Vektoren. Genauer liegt hier eine Operation der Gruppe $\mathrm{AGL}_2(\mathbb{R})$ auf dem $\mathbb{R}$-Vektorraum $V = (\mathbb{R}^2)^{n-1}$ vor.

Ein klassisches Resultat der Invariantentheorie besagt, dass der Invariantenring $\mathbb{R}[V]^{\mathrm{SL}_2(\mathbb{R})}$ der speziellen linearen Gruppe erzeugt wird von den Determinanten der Form $\det(x_i, x_j)$ mit $1 \leq i < j \leq n-1$ (vgl. Theorem 6.3.3). Anschaulich interpretiert sind also die Flächen der Parallelogramme, und damit der Dreiecke, die von den Vektoren x_i und x_j aufgespannt werden, unter der Gruppe $\mathrm{SL}_2(\mathbb{R})$ invariant. Offensichtlich sind diese Flächen auch translationsinvariant. Hält man ein Dreieck fest, z.B. das von x_1 und x_2 aufgespannte Dreieck, so erhält man durch

$$f_{i,j} = \frac{\det(x_i, x_j)}{\det(x_1, x_2)} = \frac{x_{i,1}x_{j,2} - x_{i,2}x_{j,2}}{x_{1,1}x_{2,2} - x_{1,2}x_{2,1}}, \quad 1 \leq i \leq 2 \text{ und } 3 \leq j \leq n-1$$

rationale Funktionen in den $2(n-1)$ Unbestimmten $x_{1,1}, x_{1,2}, \ldots, x_{n-1,1}, x_{n-1,2}$, die zusätzlich skalierungsinvariant sind, d.h. die $f_{i,j}$ sind rationale Invarianten unter der Operation der $\mathrm{AGL}_2(\mathbb{R})$. Will man nun wissen, ob zwei Konfigurationen mit n Punkten in

unterschiedlichen Bildern denselben Oberflächenausschnitt zeigen, sind nur diese Funktionen an den gemessenen Vektoren auszuwerten. Stimmen alle (näherungsweise) überein, handelt es sich um denselben Ausschnitt. Problematisch dabei ist nur, dass die Punkte jeweils in derselben Reihenfolge vorliegen müssen.

b) Nehmen wir nun an, wir hätten ebenfalls n besondere Punkte $x_1, \ldots, x_n \in \mathbb{R}^2$ in einem Bild auf eine bestimmte Art und Weise erhalten, die ein bestimmtes Objekt, wie z.B. ein Werkstück auf einem Fließband, festlegen.

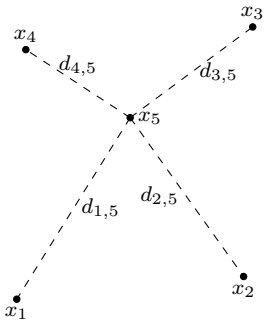


Abbildung 1.7: Ansatzweise Darstellung der gegenseitigen Abstände.

Diese Punkte sollen auch hier wieder Bildpunkte von n dreidimensionalen Punkten sein, die sich in einer Ebene befinden. Nun wollen wir die Übereinstimmung dieses Werkstücks mit einem Masterpiece aus einem anderen Bild des Werkstücks feststellen, d.h. wir nehmen also an, wir hätten dieselben n Punkte auch in einem anderen Bild erkannt. Wir wollen hier Skalierungen vernachlässigen, was für das Szenario Fließband auch gerechtfertigt ist. Was sich nun leicht bestimmen lässt, sind die (quadrierten) gegenseitigen Abstände der Punkte

$$d_{i,j} := (x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2$$

für $1 \leq i < j \leq n$. In Abbildung 1.7 ist dies ansatzweise skizziert. Offensichtlich sind diese Abstände translationsinvariant. Ein weiteres klassisches Result der Invariantentheorie besagt, dass die gegenseitigen Abstände $d_{i,j}$ den Invariantenring $\mathbb{R}[x_{1,1}, x_{1,2}, \ldots, x_{n,1}, x_{n,2}]^{\mathrm{Iso}_n(\mathbb{R})}$ der Euklidische Gruppe erzeugen (siehe Theorem 6.4.8). Leider liegt auch hier dasselbe Problem vor: Um die beiden Punktkonfigurationen miteinander vergleichen zu können, müssten die Punkte jeweils in derselben Ordnung vorliegen. Auch hier kann die Invariantentheorie Abhilfe schaffen, zumindest theoretisch. Die Abstände $d_{i,j}$ liegen auf einer affinen Varietät $X \subseteq \mathbb{R}^{\binom{n}{2}}$, auf der die symmetrische Gruppe S_n operiert (vgl. [DK02], 5.10.2, S. 233). Die Invarianten von $\mathbb{R}[X]^{S_n}$ lassen sich mittels Graphentheorie berechnen, denn die Punktkonfiguration lässt sich als gewichteter Graph auffassen, dessen Gewichte genau die Abstände $d_{i,j}$ sind (vgl. Abbildung 1.7). Berechnet man nun die Invarianten von $\mathbb{R}[X]^{S_n}$, so sind diese Invarianten nur an den Stellen $d_{i,j}$ auszuwerten, um die Übereinstimmung der Punktkonfigurationen festzustellen. Ein Problem bleibt allerdings auch hier: Die Berechnung der Invarianten von $\mathbb{R}[X]^{S_n}$ ist bisher nur für $n \leq 5$ bekannt und auch für diese n teilweise sehr aufwendig (siehe [DK02], 5.5, S. 220 f.).

c) Ein in der Bildverarbeitung wohl bekannter Ansatz zur Erkennung von Objekten oder Formen sind sogenannte **Momente** (siehe [Jäh05], 19.3, S. 548 ff.). Dazu betrachten wir vereinfacht ein Grauwertbild als Funktion $g : D \to \mathbb{R}$ auf einem kompakten Definitionsbereich $D \subseteq \mathbb{R}^2$, der Fläche des betrachteten Objekts. Dann ist für $i, j \in \mathbb{N}$ das Integral

$$m_{i,j} := \int_D x^i y^j g(x, y) dx dy$$

das **Moment $(i + j)$-ten Grades**. Ist $g : S \to \mathbb{R}$ mit $S \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ ein digitales Grauwertbild, so ergeben sich die Momente durch

$$m_{i,j} = \sum_{(x,y) \in S} x^i y^j g(x, y).$$

Besonders einfach wird die Berechnung von Momenten auf Binärbildern, da hier nur über die Pixel in S summiert wird, die zu einem Objekt gehören, das erkannt werden will. So ist

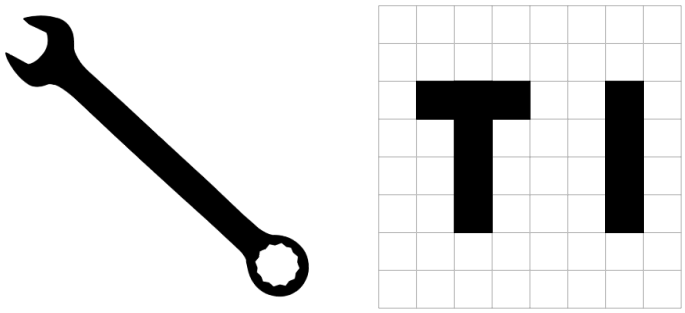


Abbildung 1.8: Beispiele für die Anwendung von Momenten.

$m_{0,0}$ auf Binärbildern nur die Fläche des Objekts. Momente werden häufig auf Binärbildern eingesetzt, z.B. zur Formerkennung oder Texterkennung (vgl. Abbildung 1.8). Dabei erlauben sie eine vollständige Formerkennung (siehe [Rei93]), die um so detaillierter wird, je mehr Momente höherer Ordnung zur Erkennung herangezogen werden. Natürlich sieht man sofort, dass sich bei einer affin-linearen Transformation $T \in \mathrm{AGL}_2(\mathbb{R})$ des Bildes g auch die Momente entscheidend ändern. Deshalb ist das Interesse an invarianten Größen in den Momenten nur natürlich (siehe [Jäh05] oder [TC92]). Den Punkt $(\overline{x}, \overline{y}) \in \mathbb{R}^2$ mit $\overline{x} = \frac{m_{1,0}}{m_{0,0}}$ und $\overline{y} = \frac{m_{0,1}}{m_{0,0}}$ bezeichnet man in Analogie zur Mechanik als Schwerpunkt. Durch Verschiebung in den Schwerpunkt erhält man translationsinvariante Momente

$$\mu_{i,j} := \int_D (x - \overline{x})^i (y - \overline{y})^j g(x, y) dx dy,$$

die als **zentrale Momente** bezeichnet werden (vgl. [Jäh05], S. 549). Diese zentralen Momente erzeugen jedoch nicht den Invariantenring bzgl. der Translationsgruppe, sie sind lediglich sogenannte **separierende Invarianten**. Die zentralen Momente bis zum Grad 2 lauten:

$$\mu_{0,0} = m_{0,0} \qquad \mu_{1,0} = 0 \qquad \mu_{0,1} = 0$$

$$\mu_{2,0} = m_{2,0} - \frac{m_{1,0}^2}{m_{0,0}} \qquad \mu_{1,1} = m_{1,1} - \frac{m_{1,0} m_{0,1}}{m_{0,0}} \qquad \mu_{0,2} = m_{0,2} - \frac{m_{0,1}^2}{m_{0,0}}$$

Translationsinvariante Momente eignen sich beispielsweise besonders gut, um in einer Bildfolge Veränderungen erkennen zu können, und das unabhängig von einer Erschütterung der Kamera. Auch für die Texterkennung eignen sie sich sehr gut, so lassen sich z.B. die in Abbildung 1.8 dargestellten Buchstaben „T" und „I" mit zentralen Momenten gut unterscheiden. Mit den zentralen Momenten zweiter Ordnung ist eine genauere Analyse der Form des Objekts möglich, sie bilden den sogenannten **Trägheitstensor** (vgl. [Jäh05], 19.3.3, S. 550 f.), eine reelle symmetrische Matrix $J = \begin{pmatrix} \mu_{2,0} & -\mu_{1,1} \\ -\mu_{1,1} & \mu_{0,2} \end{pmatrix}$. Der Eigenvektor zum minimalen Eigenwert von J gibt dabei die Orientierung des Objekts an, d.h. in der Richtung dieses Eigenvektors hat das Objekt seine größte Ausdehnung. Mit den Eigenwerten von J lässt sich außerdem die **Extrinsität** des Objekts bestimmen; sie ist nahe 0 bei runden Objekten und nahe 1 bei geradlinigen Objekten. Indem man die zentralen Momente vom Grad 2 oder höher durch entsprechende Potenzen der Momente vom Grad 0 teilt, erhält man skalierungs- und translationsinvariante Momente

$$\overline{\mu}_{i,j} := \frac{\mu_{i,j}}{\mu_{0,0}^{(2+i+j)/2}}, \quad i + j \geq 2.$$

Wir werden in dieser Arbeit u.a. einen Algorithmus (siehe Algorithmus 7.2 aus [DK02], Algorithmus 4.1.9) vorstellen, der rotationsinvariante Momente berechnen kann. Mit Hilfe dieses Algorithmus erhält man beispielsweise für die Momente zweiten Grades die Erzeuger

$$f_1 = \overline{\mu}_{2,0} + \overline{\mu}_{0,2} \qquad\qquad f_2 = \overline{\mu}_{2,0}\overline{\mu}_{0,2} - \overline{\mu}_{1,1}^2$$

des Invariantenrings $\mathbb{R}[\overline{\mu}_{2,0}, \overline{\mu}_{1,1}, \overline{\mu}_{0,2}]^{\mathrm{SO}_2(\mathbb{R})}$. Damit erhalten wir folgende translations-, skalierungs- und rotationsinvariante Größen in den Momenten bis zum Grad 2:

$$g_1 = \frac{1}{m_{0,0}^3} \cdot (m_{0,0}(m_{2,0} + m_{0,2}) - (m_{1,0}^2 + m_{0,1}^2))$$
$$g_2 = \frac{1}{m_{0,0}^5} \cdot (m_{0,0}(m_{2,0}m_{0,2} - m_{1,1}^2) + 2m_{0,1}m_{1,0}m_{1,1} - m_{0,1}^2 m_{2,0} - m_{1,0}^2 m_{0,2})$$

Natürlich lassen sich auf analoge Weise mit den Momenten höherer Ordnung weitere Invarianten erzeugen. ◁

1.3 Problemstellung

Wir haben bereits gesehen, dass zumindest lokal die geometrischen Transformationen auf den Bildern operieren. Je nach Anwendungsfall kann eine Teilmenge von $\mathrm{AGL}_2(\mathbb{R})$ ausreichend sein, wie z.B. die Menge der Ähnlichkeitstransformationen $\mathrm{Sim}_2(\mathbb{R})$, wenn man an das Szenario eines Fließbands denkt. Hat man es nur mit sehr „sanften" Transformationen zu tun, d.h. unterscheiden sich die beiden Bilder nur geringfügig, so bietet die Theorie des **optischen Flusses** verschiedene effiziente Lösungsansätze (vgl. [Han10], Kapitel 7). Die Theorie des optischen Flusses wird dominiert von folgenden Methoden und Arbeiten: der Verfolgung markanter Punkte nach Chris Harris und Mike Stephens ([HS88]), dem **Lucas-Kanade Featuretracker** nach Bruce Lucas und Takeo Kanade ([LK81]) sowie dem **Horn-Schunk-Verfahren** nach Berthold Horn und Brian Schunk ([HS81]). All diese Verfahren bekommen aber bei „größeren" Translationen, Rotationen oder Skalierungen schnell Probleme. Hier sind andere Ansätze gefragt. Die gängigen Verfahren zur Erkennung von Objekten in Bildern lassen sich in drei Kategorien einteilen (vgl. [Pis02], S. 7):

- Regionenbasierte Verfahren, wie z.B. das „Template Matching", dem sich auch der Lucas-Kanade-Featuretracker bedient, oder momentenbasierten Verfahren (siehe Beispiel 1.2.1).
- Konturbasierte Verfahren, wie z.B. Verfahren, die die Codierung der Objektberandung durch Fourierdeskriptoren verwenden (siehe [Jäh05], 19.4, S. 551 ff.).
- Punktbasierte Verfahren

Punktbasierte Verfahren, wie sie z.B. Pisinger in [Pis02] für die Transformationspassung oder Haas in [Haa00] für Fahrerassistenzsysteme verwenden, bauen auf ein und demselben Prinzip auf: Auf der lokalen Rekonstruktion der Sensorinputfunktion aus digitalen Grauwertbildern (siehe [Don09]). Dies erlaubt eine Betrachtung der jeweiligen Probleme im Subpixelbereich und ermöglicht damit eine viel genauere Lösung. Wie wir nun sehen werden, eröffnet diese Herangehensweise neue Möglichkeiten, Invariantentheorie in der Bildverarbeitung einzusetzen. Für die Erkennung von Objekten ist es nach wie vor das Ziel, die Übereinstimmung von Such- und Musterbild festzustellen. Dazu ist im Grunde „nur" zu untersuchen, ob es eine Transformation $T \in \mathrm{AGL}_2(\mathbb{R})$ gibt so, dass T das Suchbild näherungsweise in das Musterbild überführt. In jedem Fall ist also zunächst zu klären, wie sich Bilder vergleichen lassen (siehe dazu z.B.

[Pis02]). Vereinfacht wollen wir die beiden Bilder als Funktionen $f, g : D \to \mathbb{R}$ mit kompaktem Definitionsbereich $D \subseteq \mathbb{R}^2$ betrachten, wobei f das Suchbild und g das Musterbild sei. Da wir eventuell nur einen Ausschnitt des Musterbildes betrachten, sei $B \subseteq D$ mit $\lambda^2(B) > 0$, wobei λ^2 das zweidimensionale Lebesgue-Maß bezeichnen soll. Auf B gilt für eine geometrische Transformation $T \in \mathrm{AGL}_2(\mathbb{R})$ genau dann $f \circ T = g$, wenn $\int_B (f \circ T(x) - g(x))^2 d\lambda^2(x) = 0$ ist. Die am besten „passende" Transformation T, für die $f \circ T$ und g auf B möglichst gut übereinstimmen, lässt sich durch Lösen des nicht-linearen Optimierungsproblems

$$\inf_{T \in \mathrm{AGL}_2(\mathbb{R})} \int_B (f \circ T(x) - g(x))^2 d\lambda^2(x)$$

bestimmen (vgl. [Pis02], S. 3). Um dieses Optimierungsproblem zu lösen, werden „markante Punkte" bzw. „markante Strukturen" verwendet, die Pisinger **lokale Stützstrukturen** nennt. Auch regionen- oder konturbasierte Verfahren bedienen sich ihrer Bezeichnung entsprechender lokaler Stützstrukturen. Sind P_f bzw. P_g endliche, nicht-leere Mengen „markanter Punkte" aus den beiden Bildern, so ist nun

$$\inf_{T \in \mathrm{AGL}_2(\mathbb{R})} \sum_{x \in P_g} d(x, T(P_f))^2 \tag{1.3.1}$$

zu bestimmen (vgl. [Pis02], S. 6), wobei d die euklidische Metrik auf $\mathbb{R}^n$ bezeichnet und wie üblich $d(x, A) = \inf\{d(x, y) : \ y \in A\}$ der Abstand vom Punkt x zur Menge $A \subseteq \mathbb{R}^n$ ist. Dieses Konzept ähnelt der von Donner in [Don97] vorgestellten „Prototyp-Passung". Da auf natürliche Weise durch die Bildgröße die zulässigen Transformationen eine abgeschlossene und beschränkte Teilmenge von $\mathrm{AGL}_2(\mathbb{R})$ bilden, ist die Existenz einer optimalen Transformation für Gleichung (1.3.1) gewährleistet (vgl. [Don97]). Zu klären ist noch die Frage, was Pisinger als „markante Punkte" bzw. lokale Stützstrukturen verwendet. Die Herangehensweise ähnelt dabei der Extraktion markanter Punkte nach [HS88]. Diese lokalen Stützstrukturen finden in vielen Anwendungen Verwendung, wie z.B. zur Erkennung von Werkstücken am Fließband (siehe [Pis02]) oder in Fahrerassistenzsystemen (siehe [Haa00]). Sie erwiesen sich dabei als äußerst effizient und effektiv und waren den gängigen Verfahren überlegen (siehe z.B. [Pis02] oder [Haa00])

Zur Klärung der Frage nach lokalen Stützstrukturen wollen wir der Aufforderung von Ansel Adams nun nachkommen und tatsächlich *in* Bilder hineinsehen. Zunächst wird für jedes Pixel in Such- und Musterbild, jeweils mit Ausnahme der Pixel am Rand, z.B. ein 3×3-Rechteckgitter betrachtet, das dieses Pixel als Zentrum hat. Auf diesem Rechteckgitter wird anschließend die sogenannte **zeitintegrierte Sensorinputfunktion** (vgl. Abschnitt 9.1.2) rekonstruiert. An dieser Stelle wird also gleichsam wirklich tiefer *in* das Bild eingetaucht. Diese Rekonstruktion ist durch eine Polynomfunktion aus $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ möglich, wobei n so zu wählen ist, dass die Dimension von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ die Kardinalität der betrachteten lokalen Pixelfenster, wie z.B. 3×3-Fenstern, nicht übersteigt. Polynomfunktionen erwiesen sich laut Pisinger für die gegebene Situation als das Mittel der Wahl. Auf diese Weise erhält man für jedes Pixel in Such- und Musterbild eine Polynomfunktion aus $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Selbstverständlich ist die Tatsache, dass ein einzelnes Pixel zu einer Punktzuordnung kaum ausreichen kann, da einfach zu wenig Information in einem Pixel vorhanden ist.

Aus diesem Grund ist es notwendig auch die Umgebung zu betrachten. Da die Dimension des Rekonstruktionsraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ die Anzahl der Pixel in der Umgebung nicht übersteigt, handelt es sich letztendlich um eine Datenkompression.

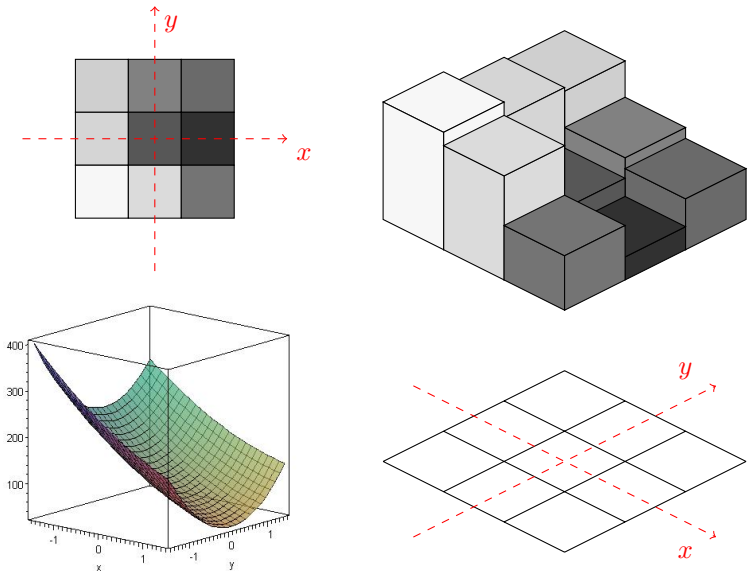


Abbildung 1.9: Darstellung der Approximation der zeitintegrierten Sensorinputfunktion durch Polynomfunktionen.

In Abbildung 1.9 ist das Vorgehen schematisch für 3×3-Rechtecke dargestellt. Dabei betrachten wir die Ecke eines Sims im oben abgebildeten Kirchturm. Diese neun Pixel sind in Abbildung 1.9 oben links zu sehen. Mit anderen Worten handelt es sich bei dieser Darstellung der neun Pixel um ein dreidimensionales Histogramm, wobei die dritte Dimension durch die Grauschattierung vorliegt. Dies lässt sich aber tatsächlich als dreidimensionales Histogramm veranschaulichen, indem man an jedem Pixel entsprechend dem jeweiligen Grauwert ein entsprechend hohen Quader errichtet (siehe Abbildung 1.9 (rechts)). Für 3×3-Gitter ist der zu verwendende Rekonstruktionsraums $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$. Der Graph des Rekonstruktionspolynoms für dieses 3×3-Gitter ist in Abbildung 1.9 links unten dargestellt. Schön zu sehen ist, dass der Graph in etwa genau so verläuft, wie das dreidimensionale Histogramm.

Die Polynomfunktion, die zur Rekonstruktion verwendet wird, lässt sich eindeutig als Linearkombination bzgl. einer Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ darstellen, was eine Identifikation der Polynomfunktion alleine durch ihre Koeffizienten bzgl. der gewählten Basis möglich macht, den sogenannten **Orthonormalkoeffizienten**. Diese Koeffizienten werden auch als **lokale Bildmerkmale** bezeichnet (vgl. Abschnitt 9.2). Wegen $\dim_{\mathbb{R}}(\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})) = 6$ korrespondiert im Falle von 3×3-Pixelfenstern beispielsweise das Zentralpixel folglich mit einem Koeffizientenvektor aus $\mathbb{R}^6$. Diesen Vektor nennt man auch einen **Merkmalsvektor** und den $\mathbb{R}$-Vektorraum $\mathbb{R}^6$ den **Merkmalsraum**. Um korrespondierende Punkte in beiden Bildern finden zu können, wäre es natürlich möglich, nun alle Merkmalsvektoren des Musterbildes mit allen Merkmalsvektoren des Suchbildes zu vergleichen, was aber viel zu aufwendig werden würde. Aus diesem Grund sucht man im Musterbild zunächst nach besonders „aussagekräftigen" Merkmalsvektoren, d.h. man vergleicht die Merkmalsvektoren des Musterbildes untereinander, was ein „Vergleichswerkzeug" erfordert. Pisinger verwendet zum Vergleich der Merkmalsvektoren ein Ähnlichkeitsmaß, das auf der **Mahalanobis-Distanz** basiert. Mit Hilfe dieses Ähnlichkeits-

maßes lassen sich einige wenige besonders „markante" Merkmalsvektoren für das Musterbild extrahieren, abhängig davon, was man als signifikant bezeichnen will. Diese wenigen Merkmalsvektoren will man nun im Suchbild wiederentdecken, d.h. man ist nun an einem effizienten Vergleich von Merkmalsvektoren zweier Bilder interessiert.

Die Operation von $\mathrm{AGL}_2(\mathbb{R})$ auf den Bildern induziert also eine Operation von $\mathrm{AGL}_2(\mathbb{R})$ auf dem $\mathbb{R}$-Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Wünschenswert ist es damit natürlich, Invarianten für interessante Untergruppen G von $\mathrm{AGL}_2(\mathbb{R})$ zu kennen. Genau das wollen wir in dieser Arbeit tun. Ist weiter $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ eine Polynomfunktion des Musterbildes repräsentiert durch einen Merkmalsvektor und $q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ eine Polynomfunktion des Suchbildes, so lautet die entscheidende Frage: Gibt es eine Transformation $T \in \mathrm{AGL}_2(\mathbb{R})$ so, dass q und $p \circ T$ möglichst gut übereinstimmen, d.h. so, dass $q \approx p \circ T$ gilt innerhalb eines gewünschten Toleranzbereichs? Auch das ist eine klassische Fragestellung der Invariantentheorie, denn mit anderen Worten ist zu untersuchen, ob p und q näherungsweise in einer Bahn liegen. Im Verfahren von Pisinger reduziert sich diese Frage im Kern auf die Frage: Gibt es eine Rotation $R \in \mathrm{SO}_2(\mathbb{R})$ mit $q \approx p \circ R$? Dazu benötigen wir nur die Invarianten von $\mathbb{R}[\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})]^{\mathrm{SO}_2(\mathbb{R})}$, die wir unter anderem in dieser Arbeit präsentieren werden.

1.4 Aufbau und Ergebnisse der Arbeit

Als Ziel dieser Arbeit wollen wir einen Ansatz präsentieren, mit dem sich das Korrespondenzproblem für lokale Bildmerkmale mit Methoden der Invariantentheorie lösen lässt. Wir wollen also die Frage behandeln, wann zwei lokale Bildmerkmale in einer Bahn bzgl. der Operation einer Gruppe sind. Dieses Korrespondenzproblem ist im speziellen Falle von 3×3-Pixelfenstern und der Operation der speziellen orthogonalen Gruppe mit größerem Aufwand bereits gelöst (siehe [Pis02]). Unser Ansatz ermöglicht eine einfache und elegante Lösung für diesen Spezialfall und ist außerdem leicht auf weitere Fälle erweiterbar, in denen der Lösungsansatz aus [Pis02] an seine Grenzen stößt. Der Aufbau dieser Arbeit ergibt sich auf kanonische Art und Weise aus dem Aufeinandertreffen der beiden Disziplinen „Invariantentheorie" und „Bildverarbeitung". Obwohl beide nicht als besonders „nah verwandt" anzusehen sind, ist die Bildverarbeitung nahezu ein natürliches Anwendungsgebiet der Invariantentheorie, wie wir im letzten Abschnitt bereits sehen konnten. Aus diesem Grund gliedern wir die Arbeit in drei größere Teile, die wiederum aus mehreren Kapiteln bestehen. Im ersten Teil behandeln wir in kurzer Form die benötigten algebraischen Grundlagen, im zweiten Teil widmen wir uns der Invariantentheorie und im dritten Teil wenden wir die Methoden der Invariantentheorie auf das Korrespondenzproblem lokaler Bildmerkmale an.

Nach der Einleitung und Motivation in die Problemstellung des Korrespondenzproblems in Stereobildern der vorangegangenen Abschnitte, betrachten wir im ersten Kapitel des ersten Teils einige ausgewählte Themen der Computeralgebra. Dieses Kapitel dient insbesondere dazu, zunächst in kurzer und prägnanter Form die verwendeten Begriffe und Notationen aus diesem Bereich einzuführen, insbesondere diejenigen Begriffe, die in der Literatur nicht einheitlich verwendet werden, wie z.B. Terme und Monome. Darüber hinaus präsentieren wir in diesem Kapitel einfache Anwendungen der Computeralgebra, die an verschiedenen Stellen der Arbeit mehrfach vorkommen. Insbesondere behandeln wir im dritten Abschnitt die Berechnung torischer Ideale, die für die späteren SAGBI-Basen von zentraler Bedeutung sind. In Kapitel 3 führen wir grundlegende Konzepte der Algebraischen Geometrie an, die wir im Folgenden an verschiedenen Stellen benötigen. Dies wird sogleich in Kapitel 4 über lineare algebraische Gruppen der Fall sein. Bei linearen algebraischen Gruppen handelt es sich um Gruppen, die

gleichzeitig die Struktur einer affinen Varietät besitzen. Im ersten Abschnitt dieses Kapitels geben wir neben zahlreichen Beispielen linearer algebraischer Gruppen eine Definition sowie erste Eigenschaften dieser Gruppen an. Im nächsten Abschnitt betrachten wir Gruppenoperationen, bevor wir im dritten Abschnitt lineare Darstellungen von linearen algebraischen Gruppen in K-Vektorräumen behandeln. Zuletzt stellen wir spezielle lineare algebraische Gruppen vor, sogenannte linear reduktive Gruppen. Lineare Reduktivität ist für die späteren Berechnungen von Invariantentringen von zentraler Bedeutung. Viele der klassischen Untergruppen von $\mathrm{GL}_n(K)$, darunter auch viele für Anwendungen interessante, sind linear reduktiv.

Bei der zentralen algebraischen Struktur des folgenden zweiten Teils, dem **Invariantenring**, handelt es sich nicht nur, wie es der Name bereits vermuten lässt, um einen Ring, sondern auch um eine Unteralgebra einer K-Algebra. Im Zusammenhang mit Unteralgebren ist es nur konsequent, sich mit sogenannten **SAGBI-Basen** zu beschäftigen, ein Erzeugendensystem von Unteralgebren mit besonders „schönen" Eigenschaften, analog den Gröbner-Basen für Ideale, was auch das Akronym SAGBI erklärt: **S**ubalgebra **A**nalogs to **G**röbner **B**ases for **I**deals. Mit SAGBI-Basen ist es möglich, besonders effizient mit Invarianten zu „hantieren", aber sie können auch von theoretischem Nutzen sein, wie diverse Beweise in der Invariantentheorie zeigen. Wir werden SAGBI-Basen, die 1989/1990 eingeführt wurden, ausführlich in Kapitel 5 diskutieren und dabei insbesondere auch feststellen, dass die Analogie zu Gröbner-Basen durchaus ihre Grenzen hat. Der Schwerpunkt liegt in diesem Kapitel auf der algorithmischen Betrachtung von SAGBI-Basen, was insgesamt zu einem Software-Paket für das Computeralgebrasystem ApCoCoA führte, dessen Beschreibung in Anhang C zu finden ist. Bislang gibt es nur sehr wenige Computeralgebrasysteme, die Funktionen für SAGBI-Basen bereitstellen, und wenn, dann nur in geringem Umfang, wie z.B. im Computeralgebrasystem Singular. Unser Software-Paket ist deutlich umfangreicher und bietet wesentlich mehr Funktionalität. Nach der Definition und grundlegenden Eigenschaften von SAGBI-Basen im ersten Abschnitt, präsentieren wir im zweiten Abschnitt einen Divisionsalgorithmus (siehe Algorithmus 5.2 mit Theorem 5.2.6) für Unteralgebren. In dieser Form fehlte ein derartiger Algorithmus noch in der einschlägigen Literatur, wie z.B. in [KR05]. Hierbei kommen die bereits erwähnten torischen Ideale zum Einsatz. Der Divisionsalgorithmus bildet die spätere Grundlage für die Berechnung von SAGBI-Basen. Zuvor behandeln wir im dritten Abschnitt die SAGBI-Normalform und reduzierte SAGBI-Basen. Dabei konnten wir insbesondere Rechenregeln für die Normalform (siehe Satz 5.3.5) beweisen bzw. korrigieren, die in [KR05] fehlerhaft und unbewiesen Teil eines Tutorials sind, und als Anwendung einen Unteralgebra-Mitgliedschaftstest (siehe Satz 5.3.7) beweisen, der sich bei Existenz einer endlichen SAGBI-Basis auch effizient umsetzen lässt (siehe Korollar 5.3.8). Der erste Teil des folgenden Abschnitts ist der SAGBI-Prozedur (siehe Pozedur SAGBI, Seite 94, mit Theorem 5.4.4) zur Berechnung von SAGBI-Basen gewidmet. Dabei terminiert die Prozedur genau dann, wenn eine endliche SAGBI-Basis der betrachteten Unteralgebra existiert. Der zweite Teil des letzten Abschnitts beschäftigt sich mit dem homogenen Fall (bzgl. der Standardgraduierung). In [RS90] ist der homogene Fall nur sehr kurz am Ende des Artikels erwähnt. Wir greifen die in [KR05] im Rahmen eines Tutorials erwähnte homogene SAGBI-Prozedur (siehe Prozedur HomSagbi, Seite 98, mit Theorem 5.4.13) auf, zeigen auf, wie sich diese implementieren lässt und beweisen die Korrektheit dieser Prozedur. Zum Abschluss dieses Kapitels behandeln wir als letzten Teil des vierten Abschnitts Grad-beschränkte SAGBI-Basen. Auch diese werden in [RS90] nur am Rande erwähnt. Wir präsentieren einen Algorithmus zur Berechnung Grad-beschränkter SAGBI-Basen (siehe Algorithmus 5.6 mit Korollar 5.4.15) und beweisen Charakterisierungen Grad-beschränkter SAGBI-Basen (siehe Satz 5.4.16). Mit Hilfe Grad-beschränkter SAGBI-Basen ist es möglich, im homogenen Fall einen effizienten Unteralgebra-Mitgliedschaftstest anzugeben (siehe Korollar 5.4.18).

Das folgende sechste Kapitel behandelt die Theorie der Invarianten. Wir geben im ersten Abschnitt einen kurzen historischen Überblick über die zwar kurze, aber äußerst bewegte Geschichte der Invariantentheorie. Im zweiten Abschnitt führen wir die Definition eines Invariantenrings sowie erste Beispiele und Eigenschaften von Invariantenringen an. Insbesondere beinhaltet dieser Abschnitt HILBERTs Endlichkeitssatz, der besagt, dass der Invariantenring bzgl. linear reduktiver Gruppen stets endlich viele Erzeuger, sogenannte fundamentale Invarianten, besitzt. In den folgenden beiden Abschnitten stellen wir klassische Invariantenringe vor, zuerst die Invarianten der speziellen und allgemeinen linearen Gruppe, die in einer Anwendung des letzten Abschnitts bereits erwähnt wurden (siehe Beispiel 1.2.1). Der nächste Abschnitt thematisiert sogenannte Vektorinvarianten der speziellen orthogonalen bzw. orthogonalen Gruppe und der Euklidischen Gruppe. Bei Vektorinvarianten handelt es sich um Invarianten von m Punkten im K^n. Für die Vektorinvarianten der speziellen orthogonalen Gruppe hat David RICHMAN in [Ric89] einen interessanten Beweis geliefert, der allerdings Einiges an Vorarbeit erfordert. Diese Vorarbeit eröffnet jedoch eine durchaus spannende Sichtweise auf Polynomringe. Wir haben uns aus Interesse etwas intensiver mit diesem umfangreichen Artikel befasst und dessen Inhalt neu aufgerollt, besser strukturiert und Ungenauigkeiten sowie kleinere Fehler korrigiert, was wir dem Leser nicht vorenthalten wollen. Um andererseits den Rahmen im sechsten Kapitel nicht zu sprengen, wird die Theorie hinter den Resultaten von RICHMAN jedoch in Anhang A präsentiert und nur die wichtigsten Aussagen sind Inhalt von Kapitel 6. Das Resultat über die spezielle orthogonale Gruppe $\mathrm{SO}_n(K)$, das RICHMAN allgemein bewiesen hat, lässt sich mit den späteren Algorithmen in konkreten Fällen auch berechnen. Bevor im sechsten Abschnitt die Hilbertreihe von Invariantenringen behandeln, stellen wir im fünften Abschnitt homogene Parametersysteme sowie primäre und sekundäre Invarianten vor. In manchen Situationen lässt sich die Hilbertreihe bereits ohne Kenntnis fundamentaler Invarianten berechnen, z.B. für endliche Gruppen, was als Molien-Formel bekannt ist. Diese Kenntnis der Hilbertreihe ohne Wissen um die fundamentalen Invarianten, ermöglicht andererseits aber, eben diese zu berechnen.

Womit wir beim siebten Kapitel angelangt sind, in dem wir die Invariantentheorie aus algorithmischer Sicht betrachten. Dazu stellen wir im ersten Abschnitt den Reynolds-Operator vor, mit dessen Hilfe ein effizienter Algorithmus zur Berechnung fundamentaler Invarianten von linear reduktiven Gruppen angegeben werden kann. Das als DERKSEN-Algorithmus bekannte Verfahren stellen wir im zweiten Abschnitt vor. Darüber hinaus zeigen wir, wie es mit Hilfe der Hilbertreihe möglich ist, fundamentale Invarianten zu berechnen. Aufbauend darauf stellen wir eine weitere, weniger effiziente Art der Berechnung vor, die nur auf Linearer Algebra beruht und ohne Reynolds-Operator auskommt. Diese Berechnungsmethode wird in der Literatur meist nur am Rande erwähnt. Wir werden hier einen entsprechenden Algorithmus vorstellen und dessen Korrektheit beweisen (siehe Algorithmus 7.5 mit Theorem 7.2.14). Diese Methode ist in leicht modifizierter Form Bestandteil des Computeralgebrasystems CoCoA. Nach kleineren Anpassungen konnten wir diese Implementierung für unsere Zwecke nutzen. Eine Implementation des DERKSEN-Algorithmus ist im Computeralgebrasystem Magma zu finden. Nach der algebraischen Sicht blicken wir im achten Kapitel aus geometrischer Sicht auf die Invariantentheorie, indem wir den sogenannten algebraischen Quotienten betrachten. Insbesondere bedeutsam ist die Trennungseigenschaft, die wir im zweiten Abschnitt des Kapitels behandeln. Ist der algebraische Quotient geometrisch, so besagt die Trennungseigenschaft, dass der algebraische Quotient die Bahnen trennt. Genau diese Eigenschaft werden wir im weiteren Verlauf ausnutzen.

Der dritte und letzten Teil steht im Zeichen der Bildverarbeitung. Im ersten Kapitel dieses Teils erklären wir, was Bilder mit Polynomfunktionen zu tun haben. Dazu behandeln wir im ersten Abschnitt den Bildentstehungsprozess. Nachdem wir in kurzer Form das Lochkameramodell einführen, thematisieren wir die mathematische Modellierung von Bildern. Wir zeigen dabei

auf, wie sich aufbauend auf einem Sensorarray ein digitales Bild ergibt. Im zweiten Abschnitt des neunten Kapitels leiten wir die neben den Invariantenringen zentralen Objekte dieser Arbeit her, die lokalen Bildmerkmale. Wir beschäftigen uns dazu mit der Rekonstruktion der Sensorinputfunktion, genauer mit der zeitintegrierten Sensorinputfunktion. Diese lässt sich auf geeignet gewählten Pixelfenstern durch Polynomfunktionen rekonstruieren, was die Lösung von Problemen der Bildverarbeitung im Subpixelbereich ermöglicht. Die Bestimmung des Rekonstruktionspolynoms ist am effizientesten durch Orthogonalentwicklung bzgl. einer geeigneten Orthonormalbasis möglich. Wir führen hier aus, welche Orthonormalbasen dafür in Frage kommen und wie man diese bestimmen kann. Die Koordinaten eines Rekonstruktionspolynoms bzgl. einer geeigneten Orthonormalbasis werden lokale Bildmerkmale genannt. Abschließend legen wir noch dar, wie sich diese lokalen Bildmerkmale effizient berechnen lassen. Dies ist allein durch Polynomauswertungen möglich. Da die Bildmerkmale lokal betrachtet werden, sind sie außerdem insbesondere translationsinvariant.

Damit ist der Bogen zur Invariantentheorie gespannt. Dank der Rekonstruktion der zeitintegrierten Sensorinputfunktion durch Polynomfunktionen operieren die Gruppen, die auf den Bildern operieren, auch auf den Polynomfunktionen. Insbesondere für Anwendungen interessant ist die Operation der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$. Im zehnten Kapitel beschreiben wir zuerst allgemein, wie sich fundamentale Invarianten von Polynomfunktionen mit Hilfe der in Kapitel 7 vorgestellten Methoden berechnen lassen. Dies führen wir dann explizit durch für die spezielle orthogonale und die orthogonale Gruppe, und zwar jeweils für die Polynomvektorräume $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit $n \in \{1, 2, 3\}$. Dies sind unter anderem genau die Vektorräume, die wir in späteren Beispielen betrachten werden. Eine analoge Berechnung für größere Parameter n ist natürlich kein Problem. Die Berechnungen führen wir bzgl. der Standardtermbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ durch, was aber keine essentielle Einschränkung bedeutet. Dennoch gehen wir im letzten Abschnitt dieses Kapitels kurz auf die Berechnung fundamentaler Invarianten bzgl. einer gewählten Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ ein. Im letzten Kapitel stellen wir dann unser Verfahren (siehe Algorithmus 11.1) zur Korrespondenzfindung lokaler Bildmerkmale vor, d.h. einen Algorithmus, mit dem sich entscheiden lässt, ob zwei lokale Bildmerkmale verschiedener Bilder ein und denselben Oberflächenausschnitt repräsentieren. Dieser Ansatz hat insbesondere den Vorteil, dass er relativ einfach implementiert werden und auf andere Gruppen sowie Pixelfenster erweitert werden kann. In [Pis02] ist beispielsweise ein Algorithmus für den Spezialfall der speziellen orthogonalen Gruppe und 3×3-Pixelfenster vorhanden, allerdings erscheint es nahezu unmöglich, dieselbe Idee für größere Pixelfenster oder andere Gruppen zu verfolgen. Für unseren Ansatz bedeutet das prinzipiell keine Probleme. Im zweiten Abschnitt des letzten Kapitels werden wir die Funktionalität des Algorithmus anhand mehrerer Beispiele demonstrieren. Auch wenn ein Nachweis der Praxistauglichkeit durch Anwendung in einer praktischen Anwendungssituation noch aussteht, zeigt sich in den ersten Beispielen durchaus, dass diese Methode erfolgreich sein kann. Problematisch sind allerdings größere Skalierungen. Im letzten Abschnitt blicken wir deshalb noch etwas weiter und stellen einen Ansatz vor, der nachweislich aus rotationsinvarianten Größen skalierungsinvariante erzeugen kann. Leider erwies sich dieser Ansatz in direkter Umsetzung als nicht besonders gut. Dennoch lassen sich vielleicht Teile davon nutzen. In [Pis02] werden beispielsweise für den erwählten Spezialfall manche dieser Skalierungsinvarianten genutzt, um ein Distanzmaß zu erzeugen, das zwar nicht skalierungsinvariant, jedoch robuster gegenüber Skalierungen ist. Diese Idee, Teile der erzeugten Skalierungsinvarianten zu nutzen, um eine gewisse Robustheit zu erzielen, könnte in der Praxis auch mit unserem Verfahren Erfolg haben. Darüber hinaus werfen wir noch einen Blick auf einen anderen „Problemfall" aus der Praxis, nämlich Beleuchtungsschwankungen. Wir stellen noch einen Ansatz vor, wie man lokale Bildmerkmale in gewissem Sinne „standardisieren" kann. Damit ließe sich eine höhere Robustheit gegenüber Beleuchtungsschwankungen erzielen.

Teil I

Algebraische Grundlagen

KAPITEL 2

Computeralgebra

Anton Bruckner[9]

Wer hohe Türme bauen will, muss lange beim Fundament verweilen.

Wie es das Zitat von Anton Bruckner passend auf den Punkt bringt, ist ein mehr oder weniger ausführlicher Blick auf die nötigen Grundlagen unverzichtbar. Das aus algebraischer Sicht notwendige Fundament wollen wir in den nächsten drei Kapiteln nur kurz umreißen. Selbstverständlich können und wollen wir nicht alle Grundlagen angeben. Insbesondere werden wir die nötigen grundlegenden Begriffe der Algebra als bekannt voraussetzen. Als Literatur zu grundlegenden Themen der Algebra sind beispielsweise die Lehrbücher von Gerd Fischer ([Fis08]), von Siegfried Bosch ([Bos06]) oder von Christian Karpfinger und Kurt Meyberg ([KM10]) zu nennen. Außerdem wollen wir die typischen Grundlagen der Computeralgebra, wie sie beispielsweise das Buch [KR00] von Martin Kreuzer und Lorenzo Robbiano absteckt, als bekannt voraussetzen. Wir werden uns generell in erster Linie, besonders was die Notation betrifft, an diesem Buch orientieren. Die Grundlagen der Computeralgebra sind natürlich auch in anderen „Standardlehrbüchern", wie z.B. [BW98] von Thomas Becker und Volker Weispfenning, zu finden.

Dennoch wollen wir im ersten Abschnitt dieses Kapitels in kurzer Form den Rahmen aus Sicht der Computeralgebra für uns festlegen. Dies ist insbesondere deshalb notwendig, da manche grundlegenden Begriffe in der Literatur durchaus unterschiedlich behandelt werden. Man denke hier beispielsweise an die kontroverse Diskussion über Terme und Monome. Wir werden also in diesem Kapitel auch die grundlegendsten Notationen aus dem Bereich der Computeralgebra angeben, aber längst nicht alle. Aus diesem Grund sei an dieser Stelle auch auf das Symbolverzeichnis verwiesen (siehe Seite 309), in dem insbesondere auch grundlegende Mengen definiert sind. In den weiteren Abschnitten werden wir ausgewählte Themen der Computeralgebra kurz präsentieren, die entweder über die allgemeinen Grundlagen hinausgehen oder an vielen zentralen Stellen der Arbeit zu finden sind.

[9] Bildquelle: http://de.wikipedia.org/wiki/Anton_Bruckner vom 21.03.2013.

2.1 Grundlegende Begriffe und Konzepte

Da wir Polynomringe nur über Körpern betrachten werden, verzichten wir auf eine allgemeine Einführung dieser. Sei also K ein Körper. Dann sind **Polynome** in einer Unbestimmten x formale Ausdrücke der Form $\sum_{i=1}^{k} \alpha_i x^i$ mit **Koeffizienten** $\alpha_1, \ldots, \alpha_k \in K$. Die einzelnen Summanden eines Polynoms $f = \sum_{i=1}^{k} \alpha_i x^i$ heißen **Monome** von f. Ein Monom der Form x^i wird **Term** genannt. Die Menge aller Polynome in einer Unbestimmten x mit Koeffizienten in K wird mit $K[x]$ bezeichnet und heißt der **Polynomring** in der Unbestimmten x über K. Sind $x_1, \ldots, x_n$ Unbestimmte, so definieren wir induktiv die K-Algebra

$$K[x_1, \ldots, x_n] = (K[x_1, \ldots, x_{n-1}])[x_n].$$

Wir nennen diese Algebra den **Polynomring** in den $n \in \mathbb{N}_+$ Unbestimmten $x_1, \ldots, x_n$ über K. **Terme** in $K[x_1, \ldots, x_n]$ sind nun formale Potenzprodukte der Form $x_1^{a_1} \cdots x_n^{a_n}$ mit Exponenten $a_1, \ldots, a_n \in \mathbb{N}$. Insbesondere ist also $1 = x_1^0 \cdots x_n^0$ ein Term. Die Menge aller Terme in $K[x_1, \ldots, x_n]$ bilden ein multiplikatives Monoid mit neutralem Element $1 = x_1^0 \cdots x_n^0$, das wir mit $\mathbb{T}^n$ bezeichnen. Die Abbildung $\log : \mathbb{T}^n \to \mathbb{N}^n$ mit $x_1^{a_1} \cdots x_n^{a_n} \mapsto (a_1, \ldots, a_n)$ heißt der **Logarithmus** auf $\mathbb{T}^n$ und bildet das Monoid $\mathbb{T}^n$ isomorph auf das Monoid $\mathbb{N}^n$ ab. Analog zu Polynomringen in einer Unbestimmten heißen formale Produkte der Form $\alpha \cdot t$ mit $\alpha \in K \setminus \{0\}$ und $t \in \mathbb{T}^n$ **Monome** in $K[x_1, \ldots, x_n]$. Ein **Polynom** f in dem Polynomring $K[x_1, \ldots, x_n]$ ist eine endliche formale Summe von Monomen mit paarweise verschiedenen Termen oder Null, d.h. ein Polynom $f \neq 0$ in $K[x_1, \ldots, x_n]$ ist also von der Form $f = \sum_{i=1}^{k} \alpha_i t_i$ mit paarweise verschiedenen Termen $t_1, \ldots, t_k \in \mathbb{T}^n$ und **Koeffizienten** $\alpha_1, \ldots, \alpha_k \in K \setminus \{0\}$. Die Menge $\{t_1, \ldots, t_k\} \subseteq \mathbb{T}^n$ der paarweise verschiedenen Terme von f heißt der **Träger** von f und wird mit $\mathrm{Supp}(f)$ bezeichnet. Ist ein Term $t \in \mathbb{T}^n$ von der Form $t = x_1^{a_1} \cdots x_n^{a_n}$ mit $a_1, \ldots, a_n \in \mathbb{N}$, so heißt die natürliche Zahl $\deg(t) := a_1 + \ldots + a_n$ der **Grad** von t. Betrachten wir alle Terme eines Polynoms $f \neq 0$ in $K[x_1, \ldots, x_n]$, so wird die natürliche Zahl

$$\deg(f) := \max\{\deg(t) : \ t \in \mathrm{Supp}(f)\}$$

der **Grad** von f genannt. Für $d \in \mathbb{N}$ bezeichnen wir mit

$$K[x_1, \ldots, x_n]_d = \{f \in K[x_1, \ldots, x_n] : \ \deg(t) = d \text{ für alle } t \in \mathrm{Supp}(f)\}$$

die Menge aller homogenen Polynome vom Grad d. Dadurch wird $K[x_1, \ldots, x_n]$ zu einem $\mathbb{N}$-graduierten Ring. Die Graduierung mit $\deg(x_i) = 1$ für alle $i \in \{1, \ldots, n\}$ heißt die **Standardgraduierung** von $K[x_1, \ldots, x_n]$. Somit ist der Polynomring $K[x_1, \ldots, x_n]$ die direkte Summe $\bigoplus_{d \geq 0} K[x_1, \ldots, x_n]_d$ mit $P_0 = K$. Weiter beinhaltet folglich die Menge $K[x_1, \ldots, x_n]_{\leq d} := \bigoplus_{0 \leq i \leq d} K[x_1, \ldots, x_n]_i$ alle Polynome vom Grad $\leq d$. Damit lässt sich jedes Polynom $f \in P$ auf eindeutige Weise in seine **homogenen Komponenten** zerlegen, d.h. es gilt $f = \sum_{d \geq 0} f_d$ mit Polynomen $f_d \in P_d$ (vgl. [KR00], Definition 1.7.1, S. 76 f.). Die homogene Komponente mit dem höchsten Grad wird auch die **Gradform** von f genannt und mit $\mathrm{DF}(f)$ bezeichnet.

Lassen sich in univariaten Polynomringen Terme noch anhand ihres Grades ordnen, so ist dies in multivariaten Polynomringen sicher nicht mehr möglich. Die Lösung dieses „Problems" sind die **Termordnungen**: Eine Termordnung σ ist dabei eine totale Ordnungsrelation auf der Menge $\mathbb{T}^n$ der Terme mit folgenden Zusatzeigenschaften:

(i) Seien $t_1, t_2, t_3 \in \mathbb{T}^n$. Ist $t_1 <_\sigma t_2$, so gilt $t_1 t_3 <_\sigma t_2 t_3$.

(ii) Für alle $t \in \mathbb{T}^n$ gilt $1 <_\sigma t$.

Eine Termordnung σ ist also eine Wohlordnung, d.h. jede nicht-leere Teilmenge $S \subseteq \mathbb{T}^n$ besitzt ein bzgl. σ minimales Element. Im Folgenden sind typische Beispiele für Termordnungen mit ihrer jeweiligen Notation aufgelistet, die in dieser Arbeit eine Rolle spielen:

Lexikographische Termordnung (Lex**):** Für zwei Terme $t_1, t_2 \in \mathbb{T}^n$ gilt genau dann $t_1 \geq_{\text{Lex}} t_2$, wenn entweder $t_1 = t_2$ gilt oder der erste Eintrag ungleich 0 in $\log(t_1) - \log(t_2)$ positiv ist.

Graduiert-lexikographische Termordnung (DegLex**:)** Für zwei Terme $t_1, t_2 \in \mathbb{T}^n$ gilt genau dann $t_1 \geq_{\text{DegLex}} t_2$, wenn entweder $\deg(t_1) > \deg(t_2)$ gilt oder $\deg(t_1) = \deg(t_2)$ und $t_1 \geq_{\text{Lex}} t_2$ ist.

Umgekehrt graduiert-lexikographische Termordnung (DegRevLex**):** Für zwei Terme t_1, t_2 in $\mathbb{T}^n$ gilt genau dann $t_1 \geq_{\text{DegRevLex}} t_2$, wenn entweder $t_1 = t_2$ oder $\deg(t_1) > \deg(t_2)$ gilt oder $\deg(t_1) = \deg(t_2)$ und der *letzte* Eintrag ungleich 0 in $\log(t_1) - \log(t_2)$ *negativ* ist.

Da sich Terme somit auch in multivariaten Polynomringen ordnen lassen, können wir in jedem Polynom $f \neq 0$ in $K[x_1, \ldots, x_n]$ von einem bzgl. einer Termordnung σ größten Term reden. Dieser größte Term wird **Leitterm** von f genannt und mit $\text{LT}_\sigma(f)$ bezeichnet. Mit anderen Worten, für ein Polynom $f \neq 0$ in $K[x_1, \ldots, x_n]$ gilt $\text{LT}_\sigma(f) = \max_\sigma\{t : \ t \in \text{Supp}(f)\}$. Das Monom αt in $f \neq 0$ mit $t = \text{LT}_\sigma(f)$ heißt entsprechend das **Leitmonom** von f und wird mit $\text{LM}_\sigma(f)$ bezeichnet. Der Koeffizient $\alpha \in K \setminus \{0\}$ heißt der **Leitkoeffizient** von f und wird mit $\text{LC}_\sigma(f)$ bezeichnet. Gilt $\text{LC}_\sigma(f) = 1$ für ein Polynom $f \neq 0$ in $R[x_1, \ldots, x_n]$, so heißt f **normiert**.

Aus der universellen Eigenschaft von Polynomringen (vgl. [KR00], Satz 1.1.12) folgt, dass es für alle $\alpha_1, \ldots, \alpha_n \in K$ einen eindeutig bestimmten Ring-Homomorphismus $\psi : K[x_1, \ldots, x_n] \to K$ gibt mit $\psi|_K = \text{id}_K$ und $\psi(x_i) = \alpha_i$ für alle $i \in \{1, \ldots, n\}$. Somit gilt $\psi(f) = f(\alpha_1, \ldots, \alpha_n)$ für ein Polynom $f \in K[x_1, \ldots, x_n]$ und $\alpha_1, \ldots, \alpha_n \in K$. Damit können wir jedes Polynom $f \in K[x_1, \ldots, x_n]$ als Funktion $K^n \to K$ betrachten. Wir reden in diesem Zusammenhang von der zu f gehörenden **Polynomfunktion**. Das Funktional, das jedem Polynom durch Auswertung seine Polynomfunktion zuweist, wird mit eval bezeichnet und heißt das **Auswertungsfunktional**. Genauer ist das Auswertungsfunktional $\text{eval} : K[x_1, \ldots, x_n] \to \text{Abb}(K^n, K)$ definiert durch

$$f \mapsto ((\alpha_1, \ldots, \alpha_n) \mapsto f(\alpha_1, \ldots, \alpha_n)).$$

Halten wir ein spezielles Element $(\alpha_1, \ldots, \alpha_n) \in K^n$ fest, so schreiben wir die Abbildung, die alle Polynome an der Stelle $(\alpha_1, \ldots, \alpha_n)$ auswertet, als Abbildung $\text{eval}_{(\alpha_1,\ldots,\alpha_n)} : K[x_1, \ldots, x_n] \to K$, definiert durch

$$\text{eval}_{(\alpha_1,\ldots,\alpha_n)}(f) = f(\alpha_1, \ldots, \alpha_n),$$

und nennen auch sie das **Auswertungsfunktional an der Stelle** $(\alpha_1, \ldots, \alpha_n)$. Mit

$$\mathscr{P}(K^n, K) = \{f : K^n \to K : \text{ es gibt ein } p \in K[x_1, \ldots, x_n] \text{ mit } f = \text{eval}(p)\}$$

bezeichnen wir die **Menge aller Polynomfunktionen** von $K^n \to K$. Analog zum Polynomring bezeichnen wir mit

$$\mathscr{P}_d(K^n, K) = \{f : K^n \to K : \text{ es gibt ein } p \in K[x_1, \ldots, x_n]_d \text{ mit } f = \text{eval}(p)\}$$

bzw. mit $\mathscr{P}_{\leq d}(K^n, K)$ die Menge aller Polynomfunktionen vom Grad $d \in \mathbb{N}$ bzw. vom Grad $\leq d$.

An vielen zentralen Stellen dieser Arbeit werden uns Ideale in Polynomringen begegnen. Ein paar besondere Ideale sowie erste spezielle Anwendungen sollen hier nun kurz vorgestellt werden. Dazu sei $P = K[x_1, \ldots, x_n]$ der Polynomring über K in den Unbestimmten $x_1, \ldots, x_n$

und σ eine Termordnung auf $\mathbb{T}^n$. Eine besondere Rolle z.B. bei SAGBI-Basen (siehe Kapitel 5) spielen Ideale, die nur von Termen erzeugt werden. Derartige Ideale werden als **monomiale Ideale** bezeichnet (siehe [KR00], Definition 1.3.1, S. 41). Monomiale Ideale sind stets endlich erzeugt, d.h. für jedes monomiale Ideal I gibt es eine endliche Teilmenge $M \subseteq \mathbb{T}^n$ mit $I = \langle M \rangle$ (vgl. [KR00], Korollar 1.3.6, S. 43). Außerdem besitzt jedes monomiale Ideal I ein bzgl. Inklusion minimales Erzeugendensystem bestehend aus Termen, welches eindeutig bestimmt ist (vgl. [KR00], Satz 1.3.11, S. 45). Dieses eindeutig bestimmte Erzeugendensystem heißt das **minimale monomiale Erzeugendensystem** des monomialen Ideals. Ein besonderes monomiales Ideal ist das **Leittermideal** eines Ideals $I \subseteq P$. Dabei handelt es sich um das von den Leittermen der Polynome aus I erzeugte Ideal, das mit $\mathrm{LT}_\sigma(I)$ bezeichnet wird (vgl. [KR00], Definition 1.5.4, S. 61 f.). Bekanntermaßen spricht man dann bei einer Menge $\{f_1, \ldots, f_s\} \subseteq I \setminus \{0\}$ von einer σ-Gröbner-Basis von I, wenn

$$\mathrm{LT}_\sigma(I) = \langle \mathrm{LT}_\sigma(f) : \ f \in I \setminus \{0\} \rangle = \langle \mathrm{LT}_\sigma(f_1), \ldots, \mathrm{LT}_\sigma(f_s) \rangle$$

gilt. Die Menge $\mathrm{LT}_\sigma\{I\} := \{\mathrm{LT}_\sigma(f) : \ f \in I \setminus \{0\}\}$ aller Leitterme von Polynomen aus I ist ein **Monoideal** von $\mathbb{T}^n$, d.h. eine Teilmenge M des multiplikativen Monoids $\mathbb{T}^n$ mit der Eigenschaft, dass aus $t \in M$ und $t' \in \mathbb{T}^n$ mit $t \mid t'$ auch $t' \in M$ folgt. Auch mit diesem Monoideal lassen sich Gröbner-Basen charakterisieren: Die Menge $\{f_1, \ldots, f_s\} \subseteq I \setminus \{0\}$ ist genau dann eine σ-Gröbner-Basis von I, wenn die Menge $\{\mathrm{LT}_\sigma(f_1), \ldots, \mathrm{LT}_\sigma(f_s)\}$ das Monoideal $\mathrm{LT}_\sigma\{I\}$ erzeugt.

Eine weitere interessante Klasse von Idealen, die wir hier angeben wollen, sind **homogene Ideale** (bzgl. der Standardgraduierung). Ein Ideal $I \subseteq P$ heißt **homogen**, wenn $I = \bigoplus_{d \geq 0}(I \cap P_d)$ gilt, d.h. wenn für alle $f \in I$ auch die homogenen Komponenten in I enthalten sind (vgl. [CLO07], Kap. 8, §3, Definition 1, S. 379). Ein Ideal ist genau dann homogen, wenn es von homogenen Polynomen erzeugt wird (vgl. [CLO07], Kap. 8, §3, Theorem 2, S. 380). Ein Erzeugendensystem eines homogenen Ideals, das aus homogenen Polynomen besteht, bezeichnen wir kurz nur als **homogenes Erzeugendensystem**. Nun ist weiter bekannt, dass jedes homogene Erzeugendensystem eines homogenen Ideals ein bzgl. Inklusion **minimales Erzeugendensystem** enthält. Darüberhinaus sind alle nicht redundanten homogenen Erzeugendensysteme minimal und enthalten dieselbe Anzahl an Elementen (vgl. [KR05], Satz 4.1.22, 26). Zur Berechnung eines minimalen Erzeugendensystems eines homogenen Ideals gibt es verschiedene Ansätze, die in [KR05], Abschnitt 4.6, S. 102–108, zu finden sind.

2.2 Einfache Anwendungen

Nachdem wir im letzten Abschnitt einige grundlegende Begriffe vorgestellt haben, wollen wir in diesem Abschnitt kurz auf ebenso grundlegende Anwendungen eingehen, die uns in dieser Arbeit mehrfach begegnen werden. Wie zuletzt sei $P = K[x_1, \ldots, x_n]$ ein Polynomring über einem Körper K in den Unbestimmten $x_1, \ldots, x_n$ und σ eine Termordnung auf $\mathbb{T}^n$. Ist I ein Ideal in P und $G = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine σ-Gröbner-Basis von I, dann gibt es für alle Polynome $f \in P$ ein eindeutig bestimmtes, bzgl. der Ersetzungsregel $\xrightarrow{G}$ irreduzibles Polynom $f_G \in P$ mit $f \xrightarrow{G} f_G$. Dieses Polynom f_G ist unabhängig von der speziellen Wahl der σ-Gröbner-Basis (vgl. [KR00], Satz 2.4.7, S. 113) und heißt die **Normalform** von f bzgl. σ, I. Es wird mit $\mathrm{NF}_{\sigma,I}(f)$ oder kurz mit $\mathrm{NF}_I(f)$ bezeichnet (vgl. [KR00], Definition 2.4.8, S. 113). Mit Hilfe der Normalform lässt sich ein **Ideal-Mitgliedschaftstest** angeben (vgl. [KR00], Satz 2.4.10).

Satz 2.2.1. (Ideal-Mitgliedschaftstest)
Seien $I = \langle g_1, \ldots, g_k \rangle$ und $J = \langle h_1, \ldots, h_l \rangle$ endlich erzeugte Ideale in P.

a) Seien $f_1, f_2 \in P$. Genau dann gilt $f_1 - f_2 \in I$, wenn $\mathrm{NF}_{\sigma,I}(f_1) = \mathrm{NF}_{\sigma,I}(f_2)$ gilt. Insbesondere ist also ein Polynom $f \in P$ genau dann in I enthalten, wenn $\mathrm{NF}_{\sigma,I}(f) = 0$ gilt.

b) Es ist genau dann $I \subseteq J$, wenn $\mathrm{NF}_{\sigma,J}(g_i) = 0$ für alle $i \in \{1, \ldots, k\}$ gilt.

c) Es gilt genau dann $I = J$, wenn $\mathrm{NF}_{\sigma,J}(g_i) = 0$ für alle $i \in \{1, \ldots, k\}$ und $\mathrm{NF}_{\sigma,I}(h_j) = 0$ für alle $j \in \{1, \ldots, l\}$ gilt.

d) Gilt $I \subseteq J$ und $\mathrm{LT}_\sigma\{J\} \subseteq \mathrm{LT}_\sigma\{I\}$, so folgt $I = J$.

Eine weitere wichtige Anwendung von Gröbner-Basen ist die **Elimination**. Sei dazu im Folgenden $L \subseteq \{x_1, \ldots, x_n\}$ eine Teilmenge der Menge der Unbestimmten. Mit $\widehat{P}$ bezeichnen wir den Polynomring $K[x_i |\ x_i \notin L]$ in den Unbestimmten $\{x_1, \ldots, x_n\} \setminus L$. Analog bezeichnen wir die Menge der Terme in den Unbestimmten $\{x_1, \ldots, x_n\} \setminus L$ mit $\widehat{\mathbb{T}}$. Eine Termordnung σ auf $\mathbb{T}^n$ heißt eine **Eliminationsordnung** für L, wenn für jedes Polynom $f \in P \setminus \{0\}$ mit $\mathrm{LT}_\sigma(f) \in \widehat{P}$ auch $f \in \widehat{P}$ gilt (vgl. [KR00], Definition 3.4.1, S. 196). Das folgende Beispiel einer Eliminationsordnung spielt auch für viele Anwendungen eine große Rolle (vgl. [KR00], Definition 1.4.10, S. 52)

Beispiel 2.2.2. (Eliminationsordnung Elim(L))
Sei $j \in \{1, \ldots, n-1\}$, sei $L = \{x_1, \ldots, x_j\}$ und seien $t_1, t_2 \in \mathbb{T}^n$ mit $t_1 = x_1^{a_1} \cdots x_n^{a_n}$ und $t_2 = x_1^{b_1} \cdots x_n^{b_n}$ für $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{N}$. Wir definieren wie folgt eine vollständige Ordnung σ auf $\mathbb{T}^n$: Es gilt genau dann $t_1 \geq_\sigma t_2$, wenn einer der folgenden Fälle eintritt:

(i) $a_1 + \ldots + a_j > b_1 + \ldots + b_j$,

(ii) $a_1 + \ldots + a_j = b_1 + \ldots + b_j$ und $t_1 \geq_{\mathrm{DegRevLex}} t_2$.

Dann ist σ eine Eliminationsordnung für L im Sinne der obigen Definition. Wir bezeichnen diese Eliminationsordnung mit Elim(L). ◁

Ist $I \subseteq P$ ein Ideal, so heißt das Ideal $I \cap \widehat{P}$ in $\widehat{P}$ das **Eliminationsideal** von I bzgl. der Menge L (vgl. [KR00], Definition 3.4.1, S. 196). Durch Elimination werden aus einem Ideal I also gleichsam diejenigen Polynome „herausgefiltert", die keine Unbestimmten aus der Menge L enthalten. Diese Polynome bilden das Eliminationsideal. Ist $L = \{x_1, \ldots, x_j\}$ für $j \in \{1, \ldots, n-1\}$, so ist Lex ein Beispiel für eine Eliminationsordnung für L. Außerdem lässt sich zeigen, dass die Einschränkung $\widehat{\sigma} = \sigma|_{\widehat{\mathbb{T}}}$ einer Termordnung σ auf $\widehat{\mathbb{T}}$ wieder eine Termordnung ist (vgl. [KR00], Satz 3.4.4). Verwendet man nun Eliminationsordnungen zur Berechnung von Gröbner-Basen, so lassen sich Eliminationsideale berechnen (vgl. [KR00], Theorem 3.4.5).

Theorem 2.2.3. (Berechnung von Eliminationsidealen)
Sei $I \subseteq P$ ein Ideal in P, sei $L \subseteq \{x_1, \ldots, x_n\}$ eine Menge von Unbestimmten und sei σ eine Eliminationsordnung für L.

a) Es gilt $\mathrm{LT}_{\widehat{\sigma}}(I \cap \widehat{P}) = \mathrm{LT}_\sigma(I) \cap \widehat{P}$.

b) Sei G eine σ-Gröbner-Basis von I. Dann ist $\widehat{G} = G \cap \widehat{P}$ eine $\widehat{\sigma}$-Gröbner-Basis von $I \cap \widehat{P}$.

c) Sei G die reduzierte σ-Gröbner-Basis von I. Dann ist $\widehat{G} = G \cap \widehat{P}$ die reduzierte $\widehat{\sigma}$-Gröbner-Basis von $I \cap \widehat{P}$.

Mit Hilfe der Elimination lässt sich beispielsweise der Durchschnitt zweier Ideale berechnen (vgl. [KR00], Satz 3.4.6, S. 198). Will man den Durchschnitt einer beliebigen Anzahl an Idealen berechnen, kann man rekursiv vorgehen. Eine weitere Möglichkeit, den Durchschnitt mehrerer

endlich erzeugter Ideale simultan mit Hilfe von Elimination zu berechnen, ist in [KR00], Satz 3.4.8, S. 199 enthalten. Elimination ist aber nicht das einzige „Werkzeug" zur Berechnung des Durchschnitts von Idealen. Auch mit Hilfe sogenannter **Syzygien** lässt sich der Durchschnitt zweier Ideale berechnen (vgl. [KR00], Satz 3.2.3, S. 162). Mit Syzygien lässt sich auch das **Quotientenideal** $I :_P J = \{f \in P : \ f \cdot J \subseteq I\}$ von I durch J berechnen (vgl. [KR00], Lemma 3.2.13 und [KR00], Satz 3.2.15, S. 167). Quotientenideale können aber ebenfalls mittels Elimination berechnet werden (vgl. [KR00], Satz 3.4.9, S. 199). Da die i-te Potenz J^i von $J \subseteq P$ ein Ideal in P ist, können wir natürlich auch Quotientenideale der Form $I :_P J^i$ berechnen. Die Vereinigung

$$\bigcup_{i \in \mathbb{N}} I :_P J^i = \{f \in P : \ J^i \cdot f \subseteq I \text{ für ein } i \in \mathbb{N}\}$$

aller derartigen Quotientenideale heißt die **Saturierung** von I durch J in P. Sie wird mit $I :_P J^\infty$ bezeichnet. Ist $f \in P$ ein Polynom mit $J^i \cdot f \subseteq I$ für ein $i \in \mathbb{N}$, so sieht man an der Definition bereits, dass $J^{i+1} \cdot f \subseteq I$ gilt, genauer gilt sogar $J^i \cdot f \subseteq J^{i+1} \cdot f \subseteq I$. Lässt man diesen Exponenten i nun immer größer werden, so wird man feststellen, dass irgendwann ein Punkt erreicht ist, an dem sich die berechneten Quotientenideale nicht mehr ändern (vgl. [KR00], Satz 3.5.9). Dank der Beziehung $I :_P J^{i+1} = (I :_P J^i) :_P J^1$ erhalten wir sofort eine naive Art der Berechnung der Saturierung von I durch J: Wir müssen lediglich so lange iterativ Quotientenideale durch J berechnen, bis der stationäre Punkt erreicht ist. Natürlich gibt es aber auch noch andere Wege, um die Saturierung zu berechnen (vgl. [KR00], Theorem 3.5.13, S. 217). Eine unmittelbare Anwendung der Saturierung ist ein Radikal-Mitgliedschaftstest, d.h. wir können mit Hilfe der Saturierung entscheiden, ob ein Polynom in dem Radikal eines Ideals enthalten ist (vgl. [KR00], Korollar 3.5.15). Eine weitere Anwendung werden wir später kennenlernen: Sogenannte torische Ideale lassen sich ebenfalls mittels Saturierung berechnen (vgl. Abschnitt 2.3).

Nach diesen ersten Anwendungen wollen wir uns nun K-Algebra-Homomorphismen zuwenden. Dazu sei $P' = K[y_1, \ldots, y_m]$ ein weiterer Polynomring sowie $I \subseteq P$ und $I' \subseteq P'$ echte Ideale in P bzw. P'. Sei weiter $\Phi : P'/I' \to P/I$ der durch Polynome $f_1, \ldots, f_m \in P$ via der Zuordnung $y_j + I' \mapsto f_j + I$ für alle $j \in \{1, \ldots, m\}$ definierte K-Algebra-Homomorphismus. Sei außerdem stets Q der Polynomring $K[y_1, \ldots, y_m, x_1, \ldots, x_n]$. Die folgenden Aussagen gelten natürlich insbesondere auch für den Fall $I = \langle 0 \rangle$ oder $I' = \langle 0 \rangle$. Eine im weiteren Verlauf tragende Rolle wird das folgende Ideal in Q spielen, das wir aus diesem Grund auch als Definition angeben.

Definition 2.2.4. (Diagonalideal)
Seien $f_1, \ldots, f_m$ Polynome in P, die einen K-Algebra-Homomorphismus $\Phi : P'/I' \to P/I$ mittels $\Phi(y_j + I') = f_j + I$ für $j \in \{1, \ldots, m\}$ definieren. Das Ideal $\langle y_1 - f_1, \ldots, y_m - f_m \rangle$ in Q heißt das **Diagonalideal** von Φ und wird mit $\triangle_\Phi$ bezeichnet.

Mit Hilfe von Elimination und dem Diagonalideal lässt sich nun der Kern von Φ berechnen (vgl. [KR00], Satz 3.6.2, S. 227).

Satz 2.2.5. (Kerne von Algebra-Homomorphismen)
Sei $\Phi : P'/I' \to P/I$ ein K-Algebra-Homomorphismus, der mit Polynomen $f_1, \ldots, f_m \in P$ definiert ist durch $\Phi(y_j + I') = f_j + I$ für $j \in \{1, \ldots, m\}$. Sei $J = IQ + \triangle_\Phi$. Dann ist $\mathrm{Ker}(\Phi)$ *das Bild des Ideals $J \cap P'$ in P'/I'.*

Als Anwendung der Berechnung des Kerns wollen wir nun zeigen, wie sich die algebraischen Relationen zwischen gegebenen Polynomen $f_1, \ldots, f_m \in P$ finden lassen, d.h. wie sich das

sogenannte *Implizitisierungsproblem* lösen lässt. Dazu benötigen wir den Begriff des Relationenideals. Dieses Ideal spielt auch im späteren Kapitel über SAGBI-Basen (siehe Kapitel 5) sowie in der Invariantentheorie eine zentrale Rolle.

Definition 2.2.6. (Ideal der algebraischen Relationen)
Seien $f_1, \ldots, f_m \in P$. Das Ideal $\{h \in K[y_1, \ldots, y_m] : h(f_1, \ldots, f_m) = 0\}$ in P' heißt das **Ideal der algebraischen Relationen von** $\mathcal{G} := (f_1, \ldots, f_m)$ oder kurz das **Relationenideal von** $\mathcal{G}$ und wird mit $\mathrm{Rel}(\mathcal{G})$ oder $\mathrm{Rel}(f_1, \ldots, f_m)$ bezeichnet.

Wie man an der Definition bereits erkennen kann, ist das Relationenideal genau der Kern eines K-Algebra-Homomorphismus. Somit lässt sich das Relationenideal effizient berechnen (vgl. [KR00], Korollar 3.6.3).

Korollar 2.2.7. (Implizitisierung)
Seien $f_1, \ldots, f_m$ Polynome in P und sei $\Phi : P' \to P$ der durch $\Phi(y_j) = f_j$ für $j \in \{1, \ldots, m\}$ definierte K-Algebra-Homomorphismus. Dann ist das Ideal der algebraischen Relationen von $f_1, \ldots, f_m$ der Kern von Φ, d.h. es gilt $\mathrm{Rel}(f_1, \ldots, f_m) = \mathrm{Ker}(\Phi) = \triangle_\Phi \cap P'$.

Das Bild $\mathrm{Im}(\Phi)$ von $\Phi : P'/I' \to P/I$ ist eine endlich erzeugte K-Unteralgebra von P/I, nämlich genau die K-Unteralgebra von P, die von $f_1+I, \ldots, f_m+I$ als K-Algebra erzeugt wird. Ist σ eine Eliminationsordnung für $\{x_1, \ldots, x_n\}$ und $G = \{g_1, \ldots, g_s\}$ die reduzierte σ-Gröbner-Basis von $J := IQ + \triangle_\Phi$, so gilt $\mathrm{Im}(\Phi) \cong K[y_1, \ldots, y_m]/\langle g_1, \ldots, g_s\rangle$ und Φ ist genau dann surjektiv, wenn G für alle $i \in \{1, \ldots, n\}$ Elemente der Form $x_i - h_i$ mit Polynomen $h_i \in P'$ enthält (vgl. [KR00], Satz 3.6.6, S. 228 f.). Die Mitgliedschaft zur Unteralgebra $\mathrm{Im}(\Phi)$ lässt sich effizient mit Hilfe der Normalform feststellen (vgl. [KR00], Korollar 3.6.7, S. 230). Einen weiteren Mitgliedschaftstest für Unteralgebren mit Hilfe von SAGBI-Basen werden wir in Kapitel 5 kennenlernen.

Satz 2.2.8. (Unteralgebra-Mitgliedschaftstest)
Seien $f_1, \ldots, f_m \in P \setminus \{0\}$ und sei $S = K[f_1, \ldots, f_m]$ die K-Unteralgebra von P, die von $f_1, \ldots, f_m$ erzeugt wird. Sei $\Phi : P' \to P$ der durch $f_1, \ldots, f_m$ definierte K-Algebra-Homomorphismus und sei σ eine Eliminationsordnung für $\{x_1, \ldots, x_n\}$.

a) Ein Polynom $g \in P$ ist genau dann in S enthalten, wenn $\mathrm{NF}_{\sigma,\triangle_\Phi}(g) \in P'$ gilt.

b) Für ein Polynom $g \in S$ liefert $h := \mathrm{NF}_{\sigma,\triangle_\Phi}(g)$ wegen $g = h(f_1, \ldots, f_m)$ eine explizite Darstellung von g als Element von S.

2.3 Torische Ideale

Von zentraler Bedeutung für die Berechnung von SAGBI-Basen (siehe Kapitel 5) sowie für Anwendungen von SAGBI-Basen ist das Lösen von linearen diophantischen Gleichungssystemen der Form

$$(\mathcal{S}) \qquad \begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

mit einer Koeffizientenmatrix $\mathcal{A} = (a_{ij}) \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ und einem Tupel $(b_1, \ldots, b_m) \in \mathbb{Z}^m$, sofern das System $(\mathcal{S})$ lösbar ist. Durch Einführen einer zusätzlichen Variablen z lässt sich jedes diophantische Gleichungssystem der Form $(\mathcal{S})$ in ein homogenes lineares diophantisches

Gleichungssystem

$$(\mathcal{S}') \qquad \begin{array}{rcccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & - & b_1 z & = & 0 \\ & & & & & & \vdots & & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & - & b_m z & = & 0 \end{array}$$

umformen. Somit können wir uns im Folgenden ohne Einschränkung auf homogene lineare diophantische Gleichungssysteme mit Koeffizientenmatrix $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ beschränken. Es gibt viele bekannte Verfahren, ein Erzeugendensystem der Lösungsmenge $\mathcal{L}(\mathcal{A}) \subseteq \mathbb{Z}^n$ eines homogenen linearen diophantischen Gleichungssystems zu berechnen, beispielsweise mit Hilfe der Hermitschen Normalform, auf die hier nicht näher eingegangen wird. Wir wollen an dieser Stelle nur auf die einschlägige Literatur verweisen, wie z.B. das Buch [Coh93] von Henri Cohen. Sehr viel schwieriger ist es, die *nicht-negativen* Lösungen des Gleichungssystems zu berechnen, d.h. die Menge $\mathcal{L}_+(\mathcal{A}) := \mathcal{L}(\mathcal{A}) \cap \mathbb{N}^n$. Aber genau diese Lösungsmenge ist für die späteren Anwendungen von zentraler Bedeutung, spielt sie doch eine besondere Rolle bei der Berechnung von SAGBI-Basen. Natürlich sind wir auch an einem möglichst effizienten Algorithmus zur Berechnung von $\mathcal{L}_+(\mathcal{A})$ interessiert. Einen solchen Algorithmus wollen wir hier nun vorstellen. Dazu benötigen wir den Begriff eines **torischen Ideals**.

Sei weiterhin K ein Körper und $P = K[x_1, \ldots, x_n]$ ein Polynomring über K in den Unbestimmten $x_1, \ldots, x_n$. Seien $y_1, \ldots, y_m$ weitere Unbestimmte und sei $L = K[y_1, \ldots, y_m, y_1^{-1}, \ldots, y_m^{-1}]$ ein Laurent-Polynomring über K in den Unbestimmten $y_1, \ldots, y_m$. Elemente von L der Form $y_1^{i_1} \cdots y_m^{i_m} \in L$ mit $i_1, \ldots, i_m \in \mathbb{Z}$ werden **erweiterte Terme** genannt. Die Menge aller erweiterten Terme in L wird mit $\mathbb{E}^m$ bezeichnet. Damit lässt sich aus jedem Spaltenvektor von $\mathcal{A}$ ein eindeutig bestimmter erweiterter Term erzeugen. Auf diese Weise lassen sich torische Ideale wie folgt definieren (vgl. [KR05], Definition 6.1.1, S. 352).

Definition 2.3.1. (Torisches Ideal)
Sei $\mathcal{A} = (a_{ij}) \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ und seien $t_1, \ldots, t_n \in \mathbb{E}^m$ erweiterte Terme mit $t_j = y_1^{a_{1j}} \cdots y_m^{a_{mj}}$ für alle $j \in \{1, \ldots, n\}$. Sei $\Phi : P \to L$ der durch $\Phi(x_j) = t_j$ für alle $j \in \{1, \ldots, m\}$ definierte K-Algebra-Homomorphismus. Dann heißt das Ideal $\mathrm{Ker}(\Phi)$ in P das zur Matrix $\mathcal{A}$ bzw. zum Tupel $(t_1, \ldots, t_n)$ gehörige **torische Ideal** und wird mit $I(\mathcal{A})$ bezeichnet.

Torische Ideale sind sogenannte **binomiale Ideale**, d.h. Ideale, die von **echten Binomen** erzeugt werden (vgl. [KR05], Satz 6.1.3). Ein **Binom** ist dabei ein Polynom $f \in P$ von der Form $f = \alpha t - \alpha' t'$ mit verschiedenen Termen $t, t' \in \mathbb{T}^n$ und Koeffizienten $\alpha, \alpha' \in K \setminus \{0\}$. Ist $f \in P$ von der Form $f = t - t'$ mit verschiedenen Termen $t, t' \in \mathbb{T}^n$, so heißt das Binom f **unitär**. Sind $t, t' \in \mathbb{T}^n$ zusätzlich teilerfremd, so heißt $f = t - t'$ ein **echtes Binom** (vgl. u.a. [KR05], Definition 6.1.2, S. 353). Für eine Teilmenge $S \subseteq P$ wird die Menge aller unitären Binome in S mit $\mathrm{UB}(S)$ und die Menge aller echten Binome in S mit $\mathrm{PB}(S)$ bezeichnet. Durch Multiplikation mit einem Term $(y_1 \cdots y_m)^d$ für genügend großes $d \in \mathbb{N}$ erhält man zu jedem erweiterten Term $t \in \mathbb{E}^m$ einen Term in $K[y_1, \ldots, y_m]$. Für einen erweiterten Term $t \in \mathbb{E}^m$ bezeichnen wir den kleinsten derartigen Exponenten d mit $\tau(t)$. Mit dem Polynomring $Q = K[x_1, \ldots, x_n, y_1, \ldots, y_m]$ erhalten wir sofort erste Möglichkeiten der Berechnung torischer Ideale (vgl. [KR05], Satz 6.1.3).

Satz 2.3.2. (Berechnung torischer Ideale)
Seien $t_1, \ldots, t_n \in \mathbb{E}^m$ erweiterte Terme und sei $\tau(t_j) \in \mathbb{N}$ für $j \in \{1, \ldots, n\}$ die kleinste natürliche Zahl mit $t_j \cdot (y_1 \cdots y_m)^{\tau(t_j)} \in K[y_1, \ldots, y_m]$. Sei weiter $I \subseteq P$ das zu $(t_1, \ldots, t_n)$ gehörige torische Ideal und sei $J \subseteq Q$ das binomiale Ideal in Q, das erzeugt wird von der Menge $\left\{(y_1 \cdots y_m)^{\tau(t_1)} \cdot (x_1 - t_1),\ \ldots,\ (y_1 \cdots y_m)^{\tau(t_n)} \cdot (x_n - t_n)\right\}$.

a) Es gilt $I = (J :_Q \langle y_1 \cdots y_m \rangle^\infty) \cap P$.

b) Sei z *eine weitere Unbestimmte und sei* G *eine Gröbner-Basis des Ideals* $J + \langle y_1 \cdots y_m z - 1 \rangle$ *in* $Q[z]$ *bzgl. einer Eliminationsordnung für* $\{y_1, \ldots, y_m, z\}$. *Dann wird das torische Ideal* I *von* $G \cap P$ *erzeugt.*

Diese Berechnungsmöglichkeiten torischer Ideale sind allerdings beide nicht besonders effizient, jedoch gibt es noch einen effizienteren Algorithmus, den wir nun vorstellen wollen. Dazu benötigen wir eine besondere Form der Darstellung für Tupel ganzer Zahlen. Zunächst lässt sich jede ganze Zahl $a \in \mathbb{Z}$ in der Form $a = \max\{a, 0\} - \max\{-a, 0\}$ schreiben. Um die Notation zu vereinfachen setzen wir $a^+ := \max\{a, 0\}$ und $a^- := \max\{-a, 0\}$ für $a \in \mathbb{Z}$. Diese Schreibweise lässt sich auf natürliche Weise auf Tupel ganzer Zahlen erweitern: Ist $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ so schreiben wir analog a^+ bzw. a^- für die Tupel $(a_1^+, \ldots, a_n^+)$ bzw. $(a_1^-, \ldots, a_n^-)$. Auch für ein Tupel $a \in \mathbb{Z}^n$ gilt offensichtlich

$$a = a^+ - a^- \qquad (*)$$

Somit gilt $a^+ = a^-$ genau dann, wenn $a = 0$ gilt. Weiter bezeichnen wir mit $\mathbf{x}^a$ den zum Tupel $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ gehörigen erweiterten Term $x_1^{a_1} \cdots x_n^{a_n} \in \mathbb{E}^n$. Man beachte, dass für jede ganze Zahl $a \in \mathbb{Z}$ die Zahlen a^+ und a^- natürliche Zahlen sind. Somit ist $\mathbf{x}^{a^+}$ bzw. $\mathbf{x}^{a^-}$ für ein Tupel $a \in \mathbb{Z}^n$ ein Term in $\mathbb{T}^n$.

Betrachten wir nun die Abbildung $\varrho' : \mathbb{Z}^n \to P$ definiert durch $\varrho'(a) = \mathbf{x}^{a^+} - \mathbf{x}^{a^-}$. Dann ordnet diese Abbildung jedem Tupel aus $\mathbb{Z}^n \setminus \{0\}$ ein Binom aus P zu. In [KR05], Satz 6.1.4, wird gezeigt, dass für jedes $v \in \mathcal{L}(\mathcal{A})$ das Polynom $\varrho'(v)$ ein echtes Binom in $I(\mathcal{A})$ ist. Allerdings ist das so nicht ganz korrekt; das Tupel $v = 0$ muss ausgeschlossen werden, da dieses unter ϱ' kein Binom liefert. Schränkt man die Abbildung ϱ' also auf die Lösungsmenge $\mathcal{L}(\mathcal{A}) \setminus \{0\}$ ein, so erhalten wir eine Abbildung $\varrho : \mathcal{L}(\mathcal{A}) \setminus \{0\} \to \mathrm{PB}(I(\mathcal{A}))$, d.h. mit anderen Worten gilt $\varrho'|_{\mathcal{L}(\mathcal{A}) \setminus \{0\}} = \varrho$ (vgl. [KR05], Satz 6.1.4). Aufbauend auf dieser Abbildung erhalten wir folgenden Zusammenhang zwischen der Lösungsmenge $\mathcal{L}(\mathcal{A}) \subseteq \mathbb{Z}^n$ eines homogenen linearen diophantischen Gleichungssystems mit Koeffizientenmatrix $\mathcal{A}$ und dem zu $\mathcal{A}$ gehörigen torischen Ideal $I(\mathcal{A})$.

Satz 2.3.3. *Sei* $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$. *Dann wird das torische Ideal* $I(\mathcal{A})$ *erzeugt von den echten Binomen der Menge* $\{\varrho(v) : \ v \in \mathcal{L}(\mathcal{A}) \setminus \{0\}\} \subseteq \mathrm{PB}(I(\mathcal{A}))$.

Beweis: Analog zu [KR05], Satz 6.1.4. □

Ist also ein Erzeugendensystem der Lösungsmenge $\mathcal{L}(\mathcal{A})$ bekannt, ist es unter Verwendung der Abbildung ϱ möglich, beliebig viele Elemente des torischen Ideals $I(\mathcal{A})$ zu berechnen. Allerdings ist das noch kein zufriedenstellendes Ergebnis, denn wünschenswert wäre es natürlich, ein Erzeugendensystem von $I(\mathcal{A})$ angeben zu können und das, wie bereits erwähnt, auf möglichste effiziente Weise. Die naheliegende Vermutung, dass die Bilder der Tupel eines Erzeugendensystems von $\mathcal{L}(\mathcal{A})$ unter ϱ das Ideal $I(\mathcal{A})$ erzeugen, ist leider nicht richtig. Wir werden allerdings sehen, dass es mit Hilfe eines Erzeugendensystems $V = \{v_1, \ldots, v_r\} \subseteq \mathbb{Z}^n$ von $\mathcal{L}(\mathcal{A})$ dennoch möglich ist, das Ideal $I(\mathcal{A})$ zu bestimmen. Das von den Bildern der Erzeuger $v_1, \ldots, v_r \in \mathbb{Z}^n$ erzeugte Ideal $I_V := \langle \varrho(v_1), \ldots, \varrho(v_r) \rangle$ in P wird das zu V gehörige **Gitterideale** genannt (vgl. [KR05], Definition 6.1.5, S. 355). Aus der Definition der Gitterideale folgt unmittelbar, dass I_V stets eine Teilmenge von $I(\mathcal{A})$ ist (vgl. Satz 2.3.3). Weiterhin lässt sich folgender Zusammenhang zwischen dem torischen Ideal $I(\mathcal{A})$ und einem speziellen Gitterideal, einem zu einem Erzeugendensystem von $\mathcal{L}(\mathcal{A})$ gehörenden Gitterideal, zeigen (vgl. [KR05], Theorem 6.1.9).

Theorem 2.3.4. *Sei $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ und $V \subseteq \mathbb{Z}^n$ eine endliche Teilmenge der Lösungsmenge $\mathcal{L}(\mathcal{A})$. Genau dann gilt $I(\mathcal{A}) = I_V :_P \langle x_1 \cdots x_n \rangle^\infty$, wenn V ein Erzeugendensystem des $\mathbb{Z}$-Moduls $\mathcal{L}(\mathcal{A})$ ist.*

Aufbauend auf diesem Theorem lässt sich ein effizienter Algorithmus (siehe Algorithmus 2.1) zur Berechnung torischer Ideale formulieren, der auf einem Erzeugendensystem von $\mathcal{L}(\mathcal{A})$ aufbaut. Wie wir zu Beginn erwähnt haben, stellt dessen Berechnung kein besonders großes Problem dar.

Algorithmus 2.1 : Berechnung torischer Ideale

Input : $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$
Result : Ein Ideal I in P mit $I = I(\mathcal{A})$.

```
1 H := ∅;
2 Berechne ein Erzeugendensystem V = {v_1, ..., v_r} von L(A);
3 for i = 1 → r do
4 |   Berechne h := ϱ(v_i);
5 |   H := H ∪ {h};
6 Bilde das Gitterideal I_V := ⟨H⟩;
7 Berechne die Saturierung I := I_V :_P ⟨x_1 ··· x_n⟩^∞;
8 return I;
```

Die Korrektheit von Algorithmus 2.1 ergibt sich unmittelbar aus dem letzten Theorem. Eine Implementation dieses Algorithmus ist beispielsweise im Computeralgebra-System CoCoA (siehe [RAB15]) zu finden. Torische Ideale lassen sich dort mit der CoCoA-Funktion `Toric` berechnen. Zu beachten ist, dass diese Funktion echte Terme als Eingabe erwartet oder anders ausgedrückt, eine Matrix $\mathcal{A}$, die keine Nullspalte enthält. Bevor wir nun mit der Berechnung von $\mathcal{L}_+(\mathcal{A})$ fortfahren, wollen wir eine leichte Varianten dieses Algorithmus präsentieren, die wir später mehrmals benötigen werden. Wir betrachten dazu keine erweiterten, sondern „normale" Terme $t_1, \ldots, t_n \in K[y_1, \ldots, y_m]$. Dann lässt sich natürlich auch hier eine Matrix $\mathcal{A}$ aufstellen, deren Spaltenvektoren die Logarithmen der Terme $t_1, \ldots, t_n$ sind. In dieser Situation erhalten wir also eine Matrix $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{N})$, also mit nicht-negativen Einträgen. Das zu $\mathcal{A}$ bzw. $(t_1, \ldots, t_n)$ gehörige torische Ideal $I(t_1, \ldots, t_n) \subseteq P$ ist dann also nichts anderes als das Relationenideal $\mathrm{Rel}(t_1, \ldots, t_n)$. Mit anderen Worten, es ist also möglich, das Relationenideal eines Tupels von Termen als Alternative zu Korollar 2.2.7 effizient mit Algorithmus 2.1 zu berechnen. Das lässt sich wie folgt als Algorithmus festhalten.

Algorithmus 2.2 : Berechnung von $\mathrm{Rel}(t_1, \ldots, t_n)$

Input : Terme $t_1, \ldots, t_n \in K[y_1, \ldots, y_m]$ mit $t_i \neq 1$ für alle $i \in \{1, \ldots, n\}$
Result : Das Relationenideal $\mathrm{Rel}(t_1, \ldots, t_n) \subseteq P$

```
1 Schreibe die Logarithmen log(t_1), ..., log(t_n) als Spaten in eine Matrix A ∈ Mat_{m,n}(ℕ);
2 Berechne mit Algorithmus 2.1 das torische Ideal I(A);
3 return I(A);
```

Die Korrektheit und Endlichkeit des Algorithmus sind natürlich offensichtlich. Dieser sehr einfache Algorithmus ist Teil des ApCoCoA-Pakets `sagbi.cpkg` (siehe Anhang C, Seite 306). Nach diesem kurzen „Ausflug", wollen wir uns nun betrachten, wie sich die Menge der nicht-negativen Lösungen, also die Menge $\mathcal{L}_+(\mathcal{A}) = \mathcal{L}(\mathcal{A}) \cap \mathbb{N}^n$ effektiv bestimmen lässt. Dazu werden die

Elemente von $\mathcal{L}_+(\mathcal{A})$ partiell wie folgt geordnet: Sind $u = (u_1, \ldots, u_n)$ und $v = (v_1, \ldots, v_n)$ Elemente von $\mathcal{L}_+(\mathcal{A})$, so gilt

$$u \preceq v \quad :\Longleftrightarrow \quad u_i \leq v_i \text{ für alle } i \in \{1, \ldots, n\}.$$

Gilt $u \preceq v$ und $u_i < v_i$ für ein $i \in \{1, \ldots, n\}$, so schreiben wir $u \prec v$. Offensichtlich ist $\mathcal{L}_+(\mathcal{A})$ durch diese partielle Ordnung wohlgeordnet, d.h. es gibt in $\mathcal{L}_+(\mathcal{A}) \setminus \{0\}$ bzgl. $\preceq$ minimale Elemente. Die Menge aller bzgl. der Ordnung $\preceq$ minimalen Elemente in $\mathcal{L}_+(\mathcal{A}) \setminus \{0\}$ wird die **Hilbertbasis** von $\mathcal{L}_+(\mathcal{A})$ genannt. Wie es die Bezeichnung auch erwarten lässt, erzeugt die Hilbertbasis auch die Menge $\mathcal{L}_+(\mathcal{A})$, d.h. jedes Element von $\mathcal{L}_+(\mathcal{A})$ lässt sich als $\mathbb{N}$-Linearkombination von Elementen aus der Hilbertbasis schreiben (vgl. [KR05], Satz 6.1.12). Um also die Lösungsmenge $\mathcal{L}_+(\mathcal{A})$ beschreiben zu können, reicht es, die Hilbertbasis von $\mathcal{L}_+(\mathcal{A})$ zu kennen. Zur effizienten Berechnung der Hilbertbasis wird die **Lawrence-Liftung** der Matrix $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ verwendet. Dabei handelt es sich um die Matrix

$$\overline{\mathcal{A}} = \begin{pmatrix} \mathcal{A} & 0 \\ \mathcal{I}_n & \mathcal{I}_n \end{pmatrix} \in \mathrm{Mat}_{m+n,2n}(\mathbb{Z}).$$

Seien $w_1, \ldots, w_n$ weitere Unbestimmte und sei $R = K[x_1, \ldots, x_n, w_1, \ldots, w_n]$. Dann ist das zur Lawrence-Liftung $\overline{\mathcal{A}}$ gehörige torische Ideal $I(\overline{\mathcal{A}})$ ein Ideal in R. Auch dieses torische Ideal wird natürlich von Binomen erzeugt, genauer von Binomen der Form $\mathbf{x}^a\mathbf{w}^b - \mathbf{x}^b\mathbf{w}^a$ mit $a, b \in \mathbb{N}^n$ (vgl. [KR05], Satz 6.1.15). Zwischen den Lösungsmengen $\mathcal{L}(\mathcal{A})$ und $\mathcal{L}(\overline{\mathcal{A}})$ bzgl. beider Matrizen $\mathcal{A}$ und $\overline{\mathcal{A}}$ besteht ein offensichtlicher Zusammenhang: Ein Tupel $v \in \mathbb{Z}^n$ ist genau dann ein Element von $\mathcal{L}(\mathcal{A})$, wenn $(v, -v) \in \mathbb{Z}^{2n}$ ein Element von $\mathcal{L}(\overline{\mathcal{A}})$ ist. Aber auch zwischen den beiden torischen Idealen $I(\mathcal{A})$ und $I(\overline{\mathcal{A}})$ besteht ein Zusammenhang. Die Elemente aus $\mathrm{PB}(I(\mathcal{A}))$ lassen sich bijektiv auf $\mathrm{PB}(I(\overline{\mathcal{A}}))$ abbilden (vgl. [KR05], Satz 6.1.15). Diese Tatsache liefert unmittelbar eine bijektive Abbildung $\lambda : \mathcal{L}_+(\mathcal{A}) \to \mathrm{PB}(I(\overline{\mathcal{A}}))$ definiert durch $\lambda(v) = \mathbf{x}^v - \mathbf{w}^v$ (vgl. [KR05], Satz 6.1.15). Mit diesen Vorbereitungen lässt sich ein einfacher Algorithmus zur Berechnung der Hilbertbasis von $\mathcal{L}_+(\mathcal{A})$ angeben, der auf der Berechnung torischer Ideale sowie reduzierter Gröbner-Basen beruht und dessen Korrektheit aus [KR05], Theorem 6.1.17, S. 361 folgt. Zudem folgt aus diesem Theorem die Endlichkeit der Hilbertbasis.

Algorithmus 2.3 : Berechnung der Hilbertbasis

Input : $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$
Result : Hilbertbasis $H \subseteq \mathbb{N}^n$ von $\mathcal{L}_+(\mathcal{A}) = \mathcal{L}(\mathcal{A}) \cap \mathbb{N}^n$.
1 Berechne die Lawrence-Liftung $\overline{\mathcal{A}} \in \mathrm{Mat}_{m+n,2n}(\mathbb{Z})$ von $\mathcal{A}$;
2 Berechne mit Algorithmus 2.1 das torische Ideal $I(\overline{A}) \subseteq K[x_1, \ldots, x_n, w_1, \ldots, w_n]$;
3 Berechne eine reduzierte Gröbner-Basis G von $I(\overline{\mathcal{A}})$;
4 $H := \{u \in \mathbb{N}^n |\ \mathbf{x}^u - \mathbf{w}^u \in G\}$;
5 **return** H;

Die Berechnung der Hilbertbasis einer Matrix $\mathcal{A}$ ist beispielsweise in CoCoA als Funktion vorhanden. Der CoCoA-Befehl `HilbertBasis` lässt im Gegensatz zu dem Befehl `Toric` beliebige Matrizen mit ganzzahligen Einträgen als Eingabe zu.

KAPITEL 3

Algebraische Geometrie

David Hilbert[10]

Im großen Garten der Geometrie kann sich jeder nach seinem Geschmack einen Strauß pflücken.

An vielen Stellen dieser Arbeit werden wir grundlegende Begriffe und Resultate aus dem Bereich der Algebraischen Geometrie benötigen, die hier in kurzer und kompakter Form zusammengefasst sind. Für einen ausführlicheren Einstieg in die faszinierende Welt der Algebraischen Geometrie sei auf die bekannten Bücher [Kun85] oder [Kun97] von Ernst Kunz, [CLO07] von David Cox, John Little und Donald O'Shea oder auf einzelne Kapitel aus [KR00] und [KR05] von Martin Kreuzer und Lorenzo Robbiano verwiesen.

In diesem Abschnitt sei K ein Körper und L ein Erweiterungskörper von K. Für $n \in \mathbb{N}_+$ heißt die Menge $\{(a_1, \ldots, a_n) : \ a_1, \ldots, a_n \in K\}$ der n-dimensionale **affine Raum** über K und wird mit $\mathbb{A}_K^n$ bezeichnet. Wir werden uns hier ausschließlich im affinen Raum über K bzw. L bewegen und im ersten Unterabschnitt ganz spezielle Teilmengen von $\mathbb{A}_K^n$ bzw. in der Regel von $\mathbb{A}_L^n$ betrachten.

3.1 Affine algebraische K-Varietäten

Der Einheitskreis in der reellen affinen Ebene $\mathbb{A}_{\mathbb{R}}^2$ lässt sich bekanntlich unter anderem in folgendem Sinne als Menge $S_1(0) := \{(x, y) \in \mathbb{R}^2 : \ x^2 + y^2 = 1\}$ von Punkten beschreiben. Mit dem Polynom $f := x^2 + x^2 - 1 \in \mathbb{Q}[x, y]$ bedeutet das also, dass jeder Punkt des Einheitskreises eine reelle Nullstelle des Polynoms f ist, also eines einzigen Polynoms. Sind $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$ Polynome über K in den Unbestimmten $x_1, \ldots, x_n$, so bilden diese m Polynome ein **algebraisches Gleichungssystem** über K. Sind alle Polynome $f_1, \ldots, f_m$ linear, so handelt es sich bei dabei um ein lineares Gleichungssystem. Klar ist, dass die Polynome $f_1, \ldots, f_m$ ein Ideal erzeugen. Also legen die Polynome eines algebraischen Gleichungssystems ein Ideal in $K[x_1, \ldots, x_n]$ fest. Andererseits folgt aus dem Hilbertschen Basissatz (vgl. [KR00],

[10] Bildquelle: http://de.wikipedia.org/wiki/David_Hilbert vom 06.04.2015.

Theorem 2.4.6, S. 113), dass jedes Ideal im Polynomring $K[x_1, \ldots, x_n]$ endlich erzeugt ist. Sind $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$ die Erzeuger eines Ideals I in $K[x_1, \ldots, x_n]$, so lässt sich mit diesen Polynomen ebenfalls ein algebraisches Gleichungssystem aufstellen. Ist $(a_1, \ldots, a_n) \in \mathbb{A}_L^n$ eine Lösung dieses Systems, so gilt natürlich auch $f(a_1, \ldots, a_n) = 0$ für alle Polynome $f \in I$. Wir können somit ein algebraisches Gleichungssystem mit dem durch $f_1, \ldots, f_m$ erzeugten Ideal in $K[x_1, \ldots, x_n]$ identifizieren. Der Lösungsmenge eines zu einem Ideal $I \subseteq K[x_1, \ldots, x_n]$ gehörenden Gleichungssystems wird dabei ein besonderer Name zugewiesen (vgl. [Kun97], Definition 3.3, S. 25).

Definition 3.1.1. (Nullstellenmenge)
Sei $I \subseteq K[x_1, \ldots, x_n]$ ein Ideal. Die Menge $\{(a_1, \ldots, a_n) \in \mathbb{A}_L^n : f(a_1, \ldots, a_n) = 0 \text{ für alle } f \in I\}$ heißt die **Nullstellenmenge** von I in $\mathbb{A}_L^n$ und wird mit $\mathcal{Z}_L(I)$ bezeichnet.

Die Nullstellenmenge eines Ideals $I = \langle f_1, \ldots, f_m \rangle \subseteq K[x_1, \ldots, x_n]$ ist also genau die Lösungsmenge des durch die Polynome $f_1, \ldots, f_m$ gegebenen algebraischen Gleichungssystems. Die Wahl des Erzeugendensystems von I spielt dabei für die Nullstellenmenge keine Rolle (vgl. [CLO07], Kap. I, § 4, Satz 4), d.h. gilt außerdem $I = \langle g_1, \ldots, g_\ell \rangle$ mit Polynomen $g_1, \ldots, g_\ell \in K[x_1, \ldots, x_n]$, so folgt $\mathcal{Z}_L(\langle f_1, \ldots, f_m \rangle) = \mathcal{Z}_L(\langle g_1, \ldots, g_\ell \rangle)$. Insbesondere kann man dann natürlich „schöne" Erzeugendensysteme wie z.B. Gröbner Basen oder reduzierte Gröbner Basen verwenden. Mit diesen Vorbereitungen können wir den zentralen Begriff dieses Abschnitts nun in folgendem Sinne einführen.

Definition 3.1.2. (Affine algebraische K-Varietät)
Sei $V \subseteq \mathbb{A}_L^n$ eine Teilmenge von $\mathbb{A}_L^n$.

a) Gibt es ein Ideal I in $K[x_1, \ldots, x_n]$ mit $V = \mathcal{Z}_L(I)$, so heißt $V \subseteq \mathbb{A}_L^n$ eine **affine (algebraische) K-Varietät**. Der Körper K heißt der **Definitionskörper** und L der **Koordinatenkörper** von V. Die Punkte der Menge $V \cap \mathbb{A}_K^n$ werden **K-rationale Punkte** von V genannt.

b) Sei V eine affine K-Varietät. Eine affine K-Varietät $U \subseteq \mathbb{A}_L^n$ mit $U \subseteq V$ heißt eine **affine K-Untervarietät** von V.

Gemäß dieser Definition ist natürlich auch $\mathbb{A}_L^n$ selbst eine affine K-Varietät. Auch die leere Menge ist eine affine K-Varietät, nämlich für $I = K[x_1, \ldots, x_n]$. Eine affine K-Varietät ist also durch das Ideal I bzw. das durch I festgelegte algebraische Gleichungssystem bestimmt. Man nennt I deshalb auch das **definierende Ideal** von V bzw. ein zu I gehöriges algebraisches Gleichungssystem $(\mathcal{S})$, das durch Erzeuger des Ideals I gegeben ist, ein **definierendes Gleichungssystem** von V.

Beispiel 3.1.3.

a) Sei $I = \langle x^2 + y^2 - 1 \rangle$ das Ideal in $\mathbb{Q}[x, y]$ und sei $L = \mathbb{R}$. Dann ist $\mathcal{Z}_L(I)$ der eingangs erwähnte Einheitskreis um den Nullpunkt.

b) Sei $I = \langle xy \rangle \subseteq \mathbb{Q}[x, y]$ und $L = \mathbb{R}$. Dann ist $\mathcal{Z}_L(I)$ das Koordinatenkreuz in der reellen Ebene.

c) Sei $I = \langle x^3 - x^2 y - 1 \rangle \subseteq \mathbb{Q}[x, y]$ und $L = \mathbb{R}$. Dann ist $\mathcal{Z}_L(I)$ der Graph der rationalen Funktion $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = \frac{x^3 - 1}{x^2}$.

d) Sei $I = \langle z^2 - y^2 - x^2 \rangle \subseteq \mathbb{Q}[x, y]$ und $L = \mathbb{R}$. Dann ist $\mathcal{Z}_L(I)$ der um die z-Achse rotierende Doppelkegel. ◁

Die Darstellung einer affinen K-Varietät als Nullstellenmenge eines algebraischen Gleichungssystems nennt man auch eine **implizite Darstellung**. Wie aus der Linearen Algebra wohlbekannt ist, lassen sich die Lösungsmengen linearer Gleichungssysteme auch durch eine **Parametrisierung** angeben. Für manche affinen K-Varietäten ist dies ebenfalls möglich, auch für nicht-lineare Varietäten. Betrachtet man beispielsweise den Einheitskreis in der reellen Ebene, so lässt sich dieser durch die Parametrisierung $\left\{\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) : t \in \mathbb{R}\right\}$ beschreiben. Wie man bei genauerer Betrachtung erkennen kann, handelt es sich dabei aber nicht um den gesamten Kreis, denn der Punkt $(-1, 0)$ lässt sich auf diese Weise nicht beschreiben.

Im Folgenden wollen wir zunächst mengentheoretische Operationen auf K-Varietäten behandeln. Dazu blicken wir kurz zurück auf die Skizze des Graphen von f aus Teil c) in Beispiel 3.1.3. In dieser Skizze sind genau genommen zwei Varietäten enthalten: Das Koordinatenkreuz und der Graph von f. Dargestellt in dieser Skizze ist also die Vereinigung beider Varietäten; der Durchschnitt beider Varietäten enthält nur den Punkt $(1, 0) \in \mathbb{A}^2_{\mathbb{R}}$. Beide resultierenden Mengen sind jeweils wieder affine Varietäten, wie uns der folgende Satz zeigt (vgl. [Kun85], Regeln 1.8 und [Kun97], Satz 1.4).

Satz 3.1.4. (Vereinigung und Durchschnitt von Varietäten)
Sei $(V_\lambda)_{\lambda\in\Lambda}$ eine Familie affiner K-Varietäten in $\mathbb{A}^n_L$ und sei $(I_\lambda)_{\lambda\in\Lambda}$ eine Familie von Idealen in $K[x_1, \ldots, x_n]$ mit $V_\lambda = \mathcal{Z}_L(I_\lambda)$ für alle $\lambda \in \Lambda$.

a) Ist Λ endlich, also ohne Einschränkung $\Lambda = \{1, \ldots, \ell\}$ mit $\ell \in \mathbb{N}_+$, so ist $V_1 \cup \ldots \cup V_\ell$ eine affine K-Varietät und es gilt $\bigcup\limits_{i=1}^{\ell} V_i = \mathcal{Z}_L\left(\bigcap\limits_{i=1}^{\ell} I_i\right) = \mathcal{Z}_L\left(\prod\limits_{i=1}^{\ell} I_i\right)$.

b) Der Durchschnitt $\bigcap_{\lambda\in\Lambda} V_\lambda$ ist eine affine K-Varietät und es gilt $\bigcap\limits_{\lambda\in\Lambda} V_\lambda = \mathcal{Z}_L\left(\sum\limits_{\lambda\in\Lambda} I_\lambda\right)$.

Aus diesem Satz folgt unmittelbar, dass das System der affinen K-Varietäten in $\mathbb{A}^n_L$ eine Topologie auf $\mathbb{A}^n_L$ ist (vgl. [Qv13]).

Definition 3.1.5. (Zariski-Topologie)
Diese Topologie heißt die **K-Topologie** oder **Zariski-Topologie** bzgl. K auf $\mathbb{A}^n_L$.

Die affinen K-Varietäten sind in dieser Topologie also gerade die abgeschlossenen Mengen. Benannt wurde diese Topologie nach dem US-amerikanischen Mathematiker Oscar Zariski[11]. Der Abschluss einer Teilmenge $V \subseteq \mathbb{A}^n_L$ wird auch als **Zariski-Abschluss** bezeichnet. Nicht nur endliche Vereinigungen und beliebige Durchschnitte von Varietäten ergeben wieder Varietäten, sondern auch das kartesische Produkt von Varietäten ist eine Varietät (vgl. [Kun97], Satz 1.5).

Satz 3.1.6. (Produkt von Varietäten)
Seien $V \subseteq \mathbb{A}^n_L$ und $W \subseteq \mathbb{A}^m_L$ zwei affine K-Varietäten. Dann ist $V \times W$ eine affine K-Varietät in $\mathbb{A}^n_L \times \mathbb{A}^m_L = \mathbb{A}^{n+m}_L$.

Wir wollen nun unseren Blick nicht vom Ideal aus richten, sondern umgekehrt zunächst beliebige Teilmengen $V \subseteq \mathbb{A}^n_L$ betrachten. Ist $(a_1, \ldots, a_n) \in V$ ein Punkt dieser beliebigen Menge, so gibt es unter Umständen viele Polynome $f \in K[x_1, \ldots, x_n]$ mit $f(a_1, \ldots, a_n) = 0$. Die Menge aller Polynome, die an jedem Punkt von V verschwinden, bilden ein Ideal, das sogenannte Verschwindungsideal (vgl. [Kun97], Definition 3.1, S. 25).

[11] Oscar Zariski, geboren als Ascher Zaritsky am 24. April 1899 in Kobryn, Weißrussland, und gestorben am 4. Juli 1986 in Brookline, Massachusetts, USA.

Definition 3.1.7. (Verschwindungsideal)
Für $V \subseteq \mathbb{A}_L^n$ heißt $\{f \in K[x_1, \ldots, x_n] : \ f(a_1, \ldots, a_n) = 0 \text{ für alle } (a_1, \ldots, a_n) \in V\}$ das **Verschwindungsideal** von V in $K[x_1, \ldots, x_n]$. Es wird mit $\mathcal{I}(V)$ bezeichnet.

Für die Spezialfälle $V = \emptyset$ und $V = \mathbb{A}_L^n$ gilt also $\mathcal{I}(V) = K[x_1, \ldots, x_n]$ und $\mathcal{I}(V) = \{0\}$. Letzteres jedoch nur, wenn L ein unendlicher Körper ist. Zudem ist das Verschwindungsideal stets ein Radikalideal, d.h. es gilt $\mathcal{I}(V) = \sqrt{\mathcal{I}(V)}$ (vgl. [Kun97], 3.4, S. 25). Weiter gilt stets

$$V = \mathcal{Z}_L(\mathcal{I}(V)), \tag{3.1.1}$$

d.h. eine affine K-Varietät $V \subseteq \mathbb{A}_L^n$ ist genau die Nullstellenmenge des Verschwindungsideals von V. Durch die Bildung des Verschwindungsideals bzw. der Nullstellenmenge haben wir einen Zusammenhang zwischen den Idealen des Polynomrings $K[x_1, \ldots, x_n]$ und den affinen K-Varietäten aus $\mathbb{A}_L^n$ erhalten. Wir bezeichnen mit $\mathrm{Id}(K[x_1, \ldots, x_n])$ die Menge aller Ideale in $K[x_1, \ldots, x_n]$ und mit $\mathrm{Aff}_{L/K}^n$ die Menge aller affinen K-Varietäten in $\mathbb{A}_L^n$. Dann erhalten wir zwei Abbildungen

$$\begin{array}{rrcl} \mathcal{Z}: & \mathrm{Id}(K[x_1, \ldots, x_n]) & \to & \mathrm{Aff}_{L/K}^n \\ & I & \mapsto & \mathcal{Z}_L(I) \\ \mathcal{I}: & \mathrm{Aff}_{L/K}^n & \to & \mathrm{Id}(K[x_1, \ldots, x_n]) \\ & V & \mapsto & \mathcal{I}(V) \end{array}$$

die den Zusammenhang zwischen Idealen und affinen K-Varietäten beschreiben. Die Abbildung $\mathcal{I}$ ist injektiv. Da jedes Verschwindungsideal ein Radikalideal ist, ist das Bild der Abbildung $\mathcal{I}$ die Menge $\mathrm{Rad}(K[x_1, \ldots, x_n])$ aller Radikalideale in $K[x_1, \ldots, x_n]$. Außerdem ist die Abbildung $\mathcal{I}$ „inklusions-umkehrend", d.h. für affine K-Varietäten $V, W \subseteq \mathbb{A}_L^n$ gilt genau dann $V \subseteq W$, wenn $\mathcal{I}(V) \subseteq \mathcal{I}(W)$ gilt (vgl. [Kun97], 3.4). Da der Polynomring $K[x_1, \ldots, x_n]$ Noethersch ist, also insbesondere jede aufsteigende Kette von Idealen irgendwann stationär wird (vgl. [KR00], Satz 2.4.5), folgt aus diesem Satz, dass sich diese Tatsache im Falle von Varietäten gerade umgekehrt verhält, d.h. jede absteigende Kette von affinen K-Varietäten wird irgendwann stationär (vgl. [Kun97], 3.4). Die schöne Eigenschaft der Injektivität weist die Abbildung $\mathcal{Z}$ nicht auf, was verschiedene Ursachen haben kann, wie die folgenden Beispiele andeuten (vgl. [CLO07], S. 170).

Beispiel 3.1.8.

a) Seien $I_1 = \langle x \rangle$ und $I_2 = \langle x^2 \rangle$ Ideale in $K[x]$. Dann gilt $I_1 \neq I_2$, aber $\mathcal{Z}_L(I_1) = \{0\}$ und $\mathcal{Z}_L(I_2) = \{0\}$.

b) Seien $I_1 = \langle 1 + x^2 + y^2 \rangle$ und $I_2 = \langle 1 + x^2 + y^4 \rangle$ Ideale in $\mathbb{Q}[x, y]$. Dann gilt $I_1 \neq I_2$ und, da die Polynome $1 + x^2 + y^2$ und $1 + x^2 + y^4$ keine reellen Nullstellen besitzen, $\mathcal{Z}_{\mathbb{R}}(I_1) = \mathcal{Z}_{\mathbb{R}}(I_2) = \emptyset$. ◁

Da $\mathcal{Z}$ nicht injektiv ist, gilt der zu Gleichung (3.1.1) analoge Zusammenhang $I = \mathcal{I}(\mathcal{Z}_L(I))$ im Allgemeinen nicht. Wir werden nun sehen, wann dies der Fall ist. Der nächste Satz wird uns zeigen, dass über einem algebraisch abgeschlossenen Körper ein algebraisches Gleichungssystem nur dann keine Lösung hat, wenn das zugehörige Ideal der ganze Polynomring ist (vgl. [KR00], Theorem 2.6.13, S. 139).

Satz 3.1.9. (Schwache Version des Hilbertschen Nullstellensatz)
Sei L algebraisch abgeschlossen und I ein echtes Ideal in $K[x_1, \ldots, x_n]$. Dann gilt $\mathcal{Z}_L(I) \neq \emptyset$.

Über $L = \mathbb{C}$ beispielsweise besitzt somit jedes algebraische Gleichungssystem, deren Polynome nicht ganz $K[x_1, \ldots, x_n]$ erzeugen, eine Lösung. Die Lösbarkeit lässt sich dabei ganz einfach feststellen: Denn $\mathcal{Z}_L(I)$ ist genau dann leer, wenn $I = K[x_1, \ldots, x_n]$ gilt, und das ist wiederum genau dann der Fall, wenn $1 \in I$ gilt. Wie sich die Mitgliedschaft von 1 in I feststellen lässt, haben wir bereits in Satz 2.2.1 gesehen. Ist L nicht algebraisch abgeschlossen, so ist folgende Implikation immer noch korrekt: Gilt $1 \in I$, so folgt $\mathcal{Z}_L(I) = \emptyset$. Die Umkehrung gilt im Allgemeinen jedoch nicht, wie wir in dem letzten Beispiel gesehen haben. Dort gilt weder $1 \in I_1$ noch $1 \in I_2$, obwohl $\mathcal{Z}_{\mathbb{R}}(I_1)$ und $\mathcal{Z}_{\mathbb{R}}(I_2)$ beide leer sind. Kurzzeitig könnte man auch die Hoffnung haben, dass die Abbildung $\mathcal{Z}$ über einem algebraisch abgeschlossenen Körper L vielleicht injektiv ist. Allerdings spricht Teil a) des obigen Beispiels leider eine andere Sprache. Der Grund ist recht einfach: Jede Potenz eines Polynoms hat dieselbe Nullstellenmenge wie das Polynom selbst. Der starke Hilbertsche Nullstellensatz sagt uns immerhin, dass für einen algebraisch abgeschlossenen Körper L das der einzige Grund ist, warum $\mathcal{Z}$ nicht injektiv ist (vgl. [Kun85], Satz 3.7).

Theorem 3.1.10. (Starker Hilbertscher Nullstellensatz)
Sei L algebraisch abgeschlossen und sei I ein Ideal in $K[x_1, \ldots, x_n]$. Dann gilt:

$$\mathcal{I}(\mathcal{Z}_L(I)) = \sqrt{I}.$$

Aus dem starken Hilbertschen Nullstellensatz folgt sofort, dass die Abbildungen $\mathcal{I}$ und $\mathcal{Z}$ bijektiv und invers zueinander sind, sofern L algebraisch abgeschlossen ist und man sich auf die Radikalideale beschränkt (vgl. [CLO07], S. 177). Bisher wussten wir bereits, dass $V = \mathcal{Z}_L(\mathcal{I}(V))$ für jede affine K-Varietät V gilt. Nun folgt zusätzlich, dass auch $I = \mathcal{I}(\mathcal{Z}_L(I))$ gilt, falls I ein Radikalideal ist und L ein algebraisch abgeschlossener Körper. Aus dem starken Hilbertschen Nullstellensatz und den bisherigen Ergebnissen ergeben sich unmittelbar weitere Folgerungen (vgl. z.T. [Kun97], 3.4).

Korollar 3.1.11. *Sei L algebraisch abgeschlossen.*

a) Seien $V, W \subseteq \mathbb{A}_L^n$ affine K-Varietäten. Dann gilt

$$\mathcal{I}(V \cup W) = \mathcal{I}(V) \cap \mathcal{I}(W) = \sqrt{\mathcal{I}(V) \cdot \mathcal{I}(W)} \text{ und } \mathcal{I}(V \cap W) = \sqrt{\mathcal{I}(V) + \mathcal{I}(W)}.$$

b) Seien $I, J \subseteq K[x_1, \ldots, x_n]$ Ideale. Genau dann gilt $\mathcal{Z}_L(I) = \mathcal{Z}_L(J)$, wenn $\sqrt{I} = \sqrt{J}$ gilt.

Die mengentheoretische Operation auf Varietäten, die wir bisher noch nicht betrachtet hatten, ist die Differenz affiner K-Varietäten, die im Allgemeinen keine affine K-Varietät ist (vgl. z.B. [CLO07], S. 194). Mit Hilfe des Zariski-Abschlusses ist die Behandlung der Differenz nun möglich und wir erhalten erneut eine Korrespondenz zur Idealtheorie, nun zu den Quotientenidealen (vgl. [CLO07], Kap. 4, § 4, Theorem 7).

Theorem 3.1.12. *Seien $V, W \subseteq \mathbb{A}_L^n$ affine K-Varietäten und seien $I, J \subseteq P := K[x_1, \ldots, x_n]$ Ideale mit $V = \mathcal{Z}_L(I)$ und $W = \mathcal{Z}_L(J)$. Dann gilt $\overline{V \setminus W} \subseteq \mathcal{Z}_L(I :_P J)$. Ist L algebraisch abgeschlossen und I ein Radikalideal, so gilt $\overline{V \setminus W} = \mathcal{Z}_L(I :_P J)$.*

Für die Verschwindungsideale affiner K-Varietäten folgt unmittelbar (vgl. [CLO07], Kap. 4, § 4, Korollar 8):

$$\mathcal{I}(V) :_P \mathcal{I}(W) = \mathcal{I}(V \setminus W)$$

Da jede absteigende Kette von affinen K-Varietäten irgendwann stationär wird, ist $\mathbb{A}_L^n$ ein Noetherscher topologischer Raum. Ebenso ist jede affine K-Varietät $V \subseteq \mathbb{A}_L^n$ versehen mit der

Relativtopologie ein Noetherscher topologischer Raum. Da zudem jede irreduzible Komponente eines topologischen Raums abgeschlossen ist, sind die irreduziblen Komponenten einer affinen K-Varietät $V \subseteq \mathbb{A}_L^n$ selbst affine K-Varietäten. Damit besitzt eine affine K-Varietät V nur endlich viele irreduzible Komponenten (vgl. [Kun97], Kap. I, Korollar 4.8). Die Zerlegung einer affinen K-Varietät $V = V_1 \cup \ldots \cup V_m$ in ihre irreduziblen Komponenten $V_1, \ldots, V_m$ heißt auch **minimale Zerlegung**. Sie ist bis auf die Reihenfolge der Vereinigung eindeutig bestimmt (vgl. [CLO07], Kap. 4, § 6, Theorem 4). Aber wie sieht man einer affinen K-Varietät an, dass sie irreduzibel ist? Hilfe bietet auch hier die Idealtheorie (vgl. [CLO07], Kap. 4, § 5, Satz 3 und Korollar 4).

Satz 3.1.13.

a) Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät. Genau dann ist V irreduzibel, wenn ihr Verschwindungsideal $\mathcal{I}(V)$ ein Primideal ist.

b) Sei L algebraisch abgeschlossen. Dann induzieren die Abbildungen $\mathcal{Z}$ und $\mathcal{I}$ eine eindeutige Korrespondenz zwischen den irreduziblen affinen K-Varietäten in $\mathbb{A}_L^n$ und den Primidealen in $K[x_1, \ldots, x_n]$.

3.2 Der affine Koordinatenring und Morphismen affiner K-Varietäten

Nachdem wir uns in den vorangegangenen Abschnitten mit Varietäten beschäftigt haben, wollen wir nun Abbildungen zwischen Varietäten in den Fokus rücken und die für uns nötigen Begriffe und Aussagen hier bereit legen. Dazu sei wie bisher K ein Körper und L ein Erweiterungskörper von K. Zunächst definieren wir Abbildungen zwischen affinen K-Varietäten (vgl. [Kun85], Kap. III, §3, S. 72 f.).

Definition 3.2.1. (K-Morphismus)
Seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten. Eine Abbildung $\varphi : V \to W$ heißt ein **K-Morphismus** oder **K-regulär**, falls es Polynome $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$ gibt mit

$$\varphi(a_1, \ldots, a_n) = (f_1(a_1, \ldots, a_n), \ldots, f_m(a_1, \ldots, a_n))$$

für alle $(a_1, \ldots, a_n) \in V$. Man sagt auch, dass das Tupel $(f_1, \ldots, f_m) \in (K[x_1, \ldots, x_n])^m$ die Abbildung φ **repräsentiert**. Die Menge $\Gamma_\varphi = \{(v, \varphi(v)) : v \in V\} \subseteq V \times W$ heißt der **Graph** von φ.

Die Polynome $f_1, \ldots, f_m$ sind dabei also so zu wählen, dass für alle $v \in V$ der Punkt $\varphi(v)$ Element der Nullstellenmenge des W definierenden Ideals ist. Außerdem ist jeder K-Morphismus $\varphi : V \to W$ mit affinen K-Varietäten $V \subseteq \mathbb{A}_L^n$ sowie $W \subseteq \mathbb{A}_L^m$ die Einschränkung eines K-Morphismus $\tilde{\varphi} : \mathbb{A}_L^n \to \mathbb{A}_L^m$.

Beispiel 3.2.2.

a) Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät. Mit $f_i := x_i$ für alle $i \in \{1, \ldots, n\}$ ist die Identität $\mathrm{id} : V \to V$ ein K-Morphismus.

b) Sei $m \leq n$ und seien $i_1, \ldots, i_m \in \{1, \ldots, n\}$ mit $i_1 < \ldots < i_m$. Mit den Polynomen $f_1 := x_{i_1}, \ldots, f_m := x_{i_m} \in K[x_1, \ldots, x_n]$ ist $\pi_{(i_1,\ldots,i_m)} : \mathbb{A}_L^n \to \mathbb{A}_L^m$ definiert durch

$$\pi_{(i_1,\ldots,i_m)}(a_1, \ldots, a_n) = \pi_{(i_1,\ldots,i_m)}(f_{i_1}(a_1, \ldots, a_n), \ldots, f_{i_m}(a_1, \ldots, a_n)) = (a_{i_1}, \ldots, a_{i_m})$$

ein K-Morphismus. Er heißt die **Projektion** auf die Koordinaten $(i_1, \ldots, i_m)$. ◁

Nachdem erwartungsgemäß auch die Identität ein K-Morphismus ist, können wir den Begriff eines K-Isomorphismus angeben (vgl. [Kun97], Kap. IV, §1, S. 85).

Definition 3.2.3. (K-Isomorphismus)
Seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten. Ein K-Morphismus $\varphi : V \to W$ heißt ein **K-Isomorphismus**, wenn es einen K-Morphismus $\psi : W \to V$ gibt mit $\varphi \circ \psi = \mathrm{id}_W$ und $\psi \circ \varphi = \mathrm{id}_V$.

Erste Eigenschaften von K-Morphismen lassen sich schnell festhalten. Beispielsweise ist jeder K-Morphismus stetig bzgl. der Zariski-Topologie (vgl. [Kun97], Kap. IV, §1, S. 85). Somit sind die Urbilder abgeschlossener Mengen unter K-Morphismen wieder abgeschlossen (vgl. [LS09], S. 9). Auch der Graph eines K-Morphismus ist abgeschlossen im Produkt der Varietäten. Mit etwas anderen Worten können wir also folgenden Satz festhalten.

Satz 3.2.4. *Seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten und sei $\varphi : V \to W$ ein K-Morphismus.*

a) Für alle affinen K-Untervarietäten U von W ist die Urbildmenge $\varphi^{-1}(U)$ eine affine K-Untervarietät von V.

b) Der Graph Γ_φ von φ ist eine affine K-Untervarietät von $V \times W$.

Somit sind insbesondere die Urbildmengen von Punkten aus W unter K-Morphismen affine K-Untervarietäten von V. Diese Untervarietäten erhalten einen eigenen Namen.

Definition 3.2.5. (Faser)
Seien $V \subseteq \mathbb{A}_L^n$ sowie $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten und sei $\varphi : V \to W$ ein K-Morphismus. Die Urbildmenge $\varphi^{-1}(w) = \{v \in V : \ \varphi(v) = w\}$ eines Punktes $w \in W$ heißt die **Faser** des Morphismus φ über w.

Sei $V \subseteq \mathbb{A}_L^n$ im Folgenden eine affine K-Varietät. Wir wollen nun zunächst K-Morphismen der Form $\varphi : V \to L$ betrachten. Die Menge aller K-Morphismen dieser Form bezeichnen wir mit $K[V]$, sie heißt der affine Koordinatenring von V.

Definition 3.2.6. (Affiner Koordinatenring)
Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät. Dann heißt die Menge $K[V]$ aller K-Morphismen der Form $\varphi : V \to L$ der **affine Koordinatenring** von V.

Mit der komponentenweisen Addition und Multiplikation wird auch die Menge $K[V]$ zu einem kommutativen Ring mit Eins. Weiter ist $K[V]$ genau dann ein Integritätsbereich, wenn V irreduzibel ist (vgl. [CLO07], Kap. 5, § 1, Satz 4). Ein Polynom in $K[x_1, \ldots, x_n]$ zu finden, das ein $\varphi : V \to L$ repräsentiert, ist kein großes Problem; genauer gibt es sehr viele Polynome in $K[x_1, \ldots, x_n]$, die φ repräsentieren. Durch

$$f \sim g \quad :\Longleftrightarrow \quad f - g \in \mathcal{I}(V)$$

ist eine Äquivalenzrelation auf $K[x_1, \ldots, x_n]$ gegeben (vgl. [CLO07], Kap. 5, § 1, Satz 2 und § 2, Satz 2), d.h. zwei Polynome $f, g \in K[x_1, \ldots, x_n]$ repräsentieren genau dann denselben K-Morphismus $\varphi : V \to L$, wenn $f - g \in \mathcal{I}(V)$ gilt. Die Äquivalenzklassen bzgl. dieser Äquivalenzrelation bezeichnen wir mit $\overline{f}$ für ein Polynom $f \in K[x_1, \ldots, x_n]$ und die Menge aller Äquivalenzklassen mit $K[x_1, \ldots, x_n]/\mathcal{I}(V)$. Mit den üblichen wohldefinierten Verknüpfungen auf den Äquivalenzklassen wird die Menge $K[x_1, \ldots, x_n]/\mathcal{I}(V)$ der Äquivalenzklassen zu einer affinen, d.h. endlich erzeugten, K-Algebra, die isomorph zu $K[V]$ ist (vgl. [Kun97], Kap. I, §5, S. 36).

Satz 3.2.7. (Darstellung des affinen Koordinatenrings)
Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät. Dann gilt $K[V] \cong K[x_1, \ldots, x_n]/\mathcal{I}(V)$.

Dank dieses Isomorphismus können wir jeden K-Morphismus $\varphi : V \to L$ eindeutig durch eine Äquivalenzklasse $\overline{f}$ identifizieren, wobei $f \in K[x_1, \ldots, x_n]$ ein Polynom ist, das φ repräsentiert.

Beispiel 3.2.8. (i-te Koordinatenfunktion)
Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät und sei $f = x_i \in K[x_1, \ldots, x_n]$ für ein $i \in \{1, \ldots, n\}$. Dann ist die durch f repräsentierte Polynomfunktion $\overline{x}_i : V \to L$ definiert durch $\overline{x}_i(a_1, \ldots, a_n) = a_i$. Sie heißt die **$i$-te Koordinatenfunktion** auf V. ◁

Aufgrund des Isomorphismus $K[V] \cong K[x_1, \ldots, x_n]/\mathcal{I}(V)$ erzeugen also die Koordinatenfunktionen den Ring $K[V]$, d.h. jede Polynomfunktion $\varphi \in K[V]$ lässt sich als K-Linearkombination von Produkten der Koordinatenfunktionen $\overline{x}_1, \ldots, \overline{x}_n$ schreiben. Die Elemente von $K[V]$ werden auch **reguläre Funktionen** auf V genannt (vgl. [Kra85], S. 230). Ist $V \subseteq \mathbb{A}_L^n$ eine affine L-Varietät, so bezeichnen wir analog mit $L[V]$ die Menge aller L-Morphismen der Form $\varphi : V \to L$, für die gilt:

$$L[V] \cong L[x_1, \ldots, x_n]/\sqrt{\mathcal{I}(V) \cdot L[x_1, \ldots, x_n]}.$$

Der affine Koordinatenring $K[V]$ ist außerdem eine sogenannte **reduzierte** affine K-Algebra, d.h. für alle $\varphi \in K[V]$ mit $\varphi^n = 0$ folgt $\varphi = 0$ (vgl. [Kun85], Kap. I, § 3, 3.12). Außerdem lässt sich zeigen, dass jede reduzierte affine K-Algebra isomorph zu einem Koordinatenring einer affinen K-Varietät ist (vgl. [Kun85], Kap. I, § 3, 3.12). Laut [Kun97], Kap. I, § 5, 5.6 ist der affine Koordinatenring des Produkts zweier affiner K-Varietäten isomorph zu dem assoziierten reduzierten Ring des Tensorprodukts der beiden Koordinatenringe. Da aus [Spr80], Lemma 1.5.2, folgt, dass dieses Tensorprodukt stets reduziert ist, erhalten wir folgende Variante der Aussage aus [Kun97].

Satz 3.2.9. (Koordinatenring des Produkts)
Seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ zwei affine K-Varietäten. Dann gilt $K[V \times W] \cong K[V] \otimes_K K[W]$.

Da der affine Koordinatenring $K[V]$ einer affinen K-Varietät V insbesondere ein kommutativer Ring mit Eins ist, können wir natürlich auch Ideale in $K[V]$ betrachten. Damit lässt sich analog Nullstellenmenge und Verschwindungsideal von Idealen bzw. von Untervarietäten definieren.

Definition 3.2.10. Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät, $I \subseteq K[V]$ ein Ideal und sei $U \subseteq V$ eine affine K-Untervarietät von V.

a) Die Menge $\mathcal{Z}_V(I) := \{v \in V : \ \varphi(v) = 0 \text{ für alle } \varphi \in I\}$ heißt die **Nullstellenmenge** von I auf V.

b) Das Ideal $\mathcal{I}_V(U) := \{\varphi \in K[V] : \ \varphi(u) = 0 \text{ für alle } u \in U\}$ heißt das **Verschwindungsideal** von U in $K[V]$.

Wir wollen zuerst ein Beispiel betrachten (vgl. [CLO07], S. 239).

Beispiel 3.2.11. Sei $V = \mathcal{Z}_{\mathbb{R}}(z - x^2 - y^2)$ eine affine $\mathbb{Q}$-Varietät.

a) Sei $I = \langle \overline{x} \rangle \subseteq \mathbb{Q}[V]$. Dann gilt $U := \mathcal{Z}_V(I) = \{(0, u, u^2) : \ u \in \mathbb{R}\}$ und $U \subseteq V$. Uegen $U = \mathcal{Z}_{\mathbb{R}}(z - x^2 - y^2, x)$ ist U selbst eine affine $\mathbb{Q}$-Varietät. Somit ist U also eine $\mathbb{Q}$-Untervarietät von V.

b) Sei $U = \{(1, 1, 2)\} \subseteq V$. Dann gilt $\mathcal{I}_V(U) = \langle \overline{x - 1}, \overline{y - 1} \rangle$. ◁

Die Operationen $\mathcal{Z}_V$ und $\mathcal{I}_V$ weisen ähnliche Eigenschaften auf wie die Operationen $\mathcal{Z}$ und $\mathcal{I}$. Folgender Satz fast einige davon zusammen, u.a. anderem auch eine Eigenschaft, die wir in obigem Beispiel bereits gesehen haben: Die Nullstellenmenge eines Ideals $I \subseteq K[V]$ ist stets eine Untervarietät von V (vgl. [CLO07], Kap. 5, § 4, Satz 3).

Satz 3.2.12. *Sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät.*

a) Für jedes Ideal $I \subseteq K[V]$ ist $\mathcal{Z}_V(I)$ eine K-Untervarietät von V.

b) Für jede affine K-Untervarietät $U \subseteq V$ ist $\mathcal{I}_V(U)$ ein Ideal in $K[V]$.

c) Ist $I \subseteq K[V]$ ein Ideal, dann gilt $I \subseteq \sqrt{I} \subseteq \mathcal{I}_V(\mathcal{Z}_V(I))$.

d) Ist $U \subseteq V$ eine affine K-Untervarietät von V, dann gilt $W = \mathcal{Z}_V(\mathcal{I}_V(U))$.

Auch der Hilbertsche Nullstellensatz und die daraus resultierenden Folgerungen lassen sich übertragen. Dabei ist ein Ideal $I \subseteq K[V]$ genau dann ein Radikalideal, wenn das Ideal

$$\{f \in K[x_1, \ldots, x_n] : \overline{f} \in I\}$$

ein Radikalideal in $K[x_1, \ldots, x_n]$ ist (vgl. [CLO07], Kap. 5, § 4, Satz 4). Damit nun zum Hilbertschen Nullstellensatz in $K[V]$ und seinen Folgerungen (vgl. [Kun97], Kap. I, Satz 5.2 sowie 5.8).

Satz 3.2.13. (Hilbertscher Nullstellensatz in $K[V]$ und seine Folgerungen)
Sei L algebraisch abgeschlossen und sei $V \subseteq \mathbb{A}_L^n$ eine affine K-Varietät.

a) Für jedes Ideal $I \subseteq K[V]$ gilt $\mathcal{I}_V(\mathcal{Z}_V(I)) = \sqrt{I}$.

b) Die Zuordnungen $U \mapsto \mathcal{I}_V(U)$ und $I \mapsto \mathcal{Z}_V(I)$ definieren inklusions-umkehrende Bijektionen zwischen den Mengen der affinen K-Untervarietäten von V und den Radikalidealen in $K[V]$. Die Zuordnungen sind zudem invers zueinander.

c) Bei der Zuordnung $U \mapsto \mathcal{I}_V(U)$ entsprechen die irreduziblen K-Untervarietäten von V eineindeutig den Primidealen in $K[V]$.

d) Die irreduziblen Komponenten von V entsprechen eineindeutig den minimalen Primidealen in $K[V]$.

e) Die Punkte von V entsprechen eineindeutig den maximalen Idealen des Koordinatenrings $L[V]$ und die Punkte einer affinen K-Untervarietät von V entsprechen eineindeutig den maximalen Idealen in $L[x_1, \ldots, x_n]$, die $\mathcal{I}_V(W) \cdot L[x_1, \ldots, x_n]$ umfassen.

Betrachten wir nun wieder einen K-Morphismus $\varphi : V \to W$, wobei $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ jeweils affine K-Varietäten sind, der durch ein Tupel $(f_1, \ldots, f_m) \in (K[x_1, \ldots, x_n])^m$ repräsentiert wird. Folgender Satz stellt uns nun einen Zusammenhang her zwischen K-Morphismen und K-Algebra-Homomorphismen (vgl. [CLO07], Kap. 5, § 4, Satz 8).

Satz 3.2.14. (Koordinatenabbildung)
Seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ zwei affine K-Varietäten.

a) Sei $\varphi : V \to W$ ein K-Morphismus. Dann ist $\varphi^* : K[W] \to K[V]$ definiert durch

$$\varphi^*(\overline{g}) = \overline{g} \circ \varphi$$

*ein wohldefinierter K-Algebra-Homomorphismus. Die Abbildung φ^* heißt die zu φ gehörige **Koordinatenabbildung**.*

b) Sei $\Phi : K[W] \to K[V]$ ein K-Algebra-Homomorphismus mit $\Phi(\overline{a}) = \overline{a}$ für alle $a \in K$. Dann gibt es einen eindeutig bestimmten K-Morphismus $\varphi : V \to W$ mit $\varphi^* = \Phi$.

Für jeden K-Morphismus $\varphi : V \to W$ hat die Koordinatenabbildung $\varphi^* : K[W] \to K[V]$ die Eigenschaft, dass sie konstante Funktionen aus $K[W]$ auf eine konstante Funktion aus $K[V]$ mit demselben Wert abbildet.

Bemerkung 3.2.15. Sei $\overline{a} \in K[W]$ für ein $a \in K$, d.h. $\overline{a}$ ist die konstante Funktion $\overline{a} : W \to L$, die jedem Punkt aus W den konstanten Wert $a \in K$ zuordnet, so ist $\varphi^*(\overline{a}) = \overline{a} \circ \varphi : V \to L$ konstant mit $(\overline{a} \circ \varphi)(v) = a$ für alle $v \in V$.

Insbesondere folgt aus dem letzten Satz, dass durch $\varphi \mapsto \varphi^*$ eine Bijektion zwischen allen K-Morphismen $V \to W$ und allen K-Algebra-Homomorphismen $K[W] \to K[V]$ definiert ist (vgl. [Kun97], Kap. IV, §1, S. 85). Damit entsprechen die K-Isomorphismen eineindeutig den K-Algebra-Ismorphismen. Gibt es einen Isomorphismus $K[V] \cong K[W]$, der die Identität ist auf konstanten Funktionen, so werden V und W als **isomorph** bezeichnet (vgl. [CLO07], Kap. 5, §4, Theorem 9). Aus der folgenden Bemerkung geht hervor, wie man aus einem gegebenen K-Algebra-Homomorphismus Φ einen K-Morphismus φ mit $\varphi^* = \Phi$ konstruieren kann.

Bemerkung 3.2.16. (Konstruktion von K-Morphismen aus K-Algebra-Homomorphismen) Seien $I \subseteq K[x_1, \ldots, x_n]$ und $J \subseteq K[y_1, \ldots, y_m]$ Ideale und sei der K-Algebra-Homomorphismus $\Phi : K[y_1, \ldots, y_m]/J \to K[x_1, \ldots, x_n]/I$ definiert durch $\Phi(\overline{y}_j) \to \overline{f}_j$ für alle $j \in \{1, \ldots, m\}$ mit Polynomen $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$. Setze $V := \mathcal{Z}_L(I)$ und $W := \mathcal{Z}_L(J)$. Dann ist $\varphi : V \to W$ definiert durch

$$\varphi(a_1, \ldots, a_n) = (f_1(a_1, \ldots, a_n), \ldots, f_m(a_1, \ldots, a_n))$$

der eindeutig bestimmte K-Morphismus mit $\varphi^* = \Phi$.

Über die Koordinatenabbildung eines K-Morphismus lassen sich noch folgende Eigenschaften festhalten (vgl. [Kun97], Kap. III, §3, S. 73).

Satz 3.2.17. *Seien $V \subseteq \mathbb{A}^n_L$ und $W \subseteq \mathbb{A}^m_L$ affine K-Varietäten und sei $\varphi : V \to W$ ein K-Morphismus mit Koordinatenabbildung $\varphi^* : K[W] \to K[V]$.*

*a) Genau dann ist φ^* injektiv, wenn $\mathrm{Im}(\varphi)$ dicht in W ist, also wenn $\overline{\mathrm{Im}(\varphi)} = W$ gilt. In diesem Fall nennt man φ **dominant**.*

*b) Genau dann ist φ^* surjektiv, wenn $\varphi(V)$ eine K-Untervarietät von W und $\varphi : V \to \varphi(V)$ ein K-Isomorphismus ist. In diesem Fall nennt man φ eine **abgeschlossene Einbettung** oder **Immersion** von V in W.*

Die Koordinatenabbildung ermöglicht nun eine algorithmische Betrachtung von K-Morphismen. Mit ihrer Hilfe lassen sich bestimmte Varietäten mit Methoden der Computeralgebra berechnen. Wir wollen uns nun Bilder und Urbilder von affinen K-Varietäten unter K-Morphismen etwas näher betrachten. Das folgende Beispiel zeigt uns leider sofort, dass das Bild einer affinen K-Varietät unter einem K-Morphismus im Allgemeinen keine affine K-Varietät ist.

Beispiel 3.2.18. Sei $V = \mathcal{Z}_{\mathbb{R}}(\langle xy - 1 \rangle) \subseteq \mathbb{A}^2_{\mathbb{R}}$, d.h. die affine $\mathbb{Q}$-Varietät V ist der Graph der reellen Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ definiert durch $f(x) = \frac{1}{x}$. Sei $\pi_2 : \mathbb{A}^2_{\mathbb{R}} \to \mathbb{A}^1_{\mathbb{R}}$ mit die $\pi((a_1, a_2)) = a_2$ die Projektion auf die zweite Komponente. Dann gilt $\pi_2(V) = \mathbb{A}^1_{\mathbb{R}} \setminus \{0\}$. Da $\mathbb{A}^1_{\mathbb{R}} \setminus \pi_2(V) = \{0\}$ abgeschlossen ist, ist $\pi_2(V)$ offen bzgl. der Zariski-Topologie und somit keine affine $\mathbb{Q}$-Varietät. ◁

Allerdings ist offensichtlich der Abschluss des Bildes einer affinen K-Varietät unter einem K-Morphismus eine affine K-Varietät. Der nächste Satz liefert uns einen Ansatz zur Berechnung des Abschlusses des Bildes einer K-Varietät, indem man die Berechnung auf die zu einem K-Morphismus gehörige Koordinatenabbildung zurückführt (vgl. [Kun85], Kap. III, §3, S. 73).

Satz 3.2.19. (Bilder von K-Varietäten)
Sei L algebraisch abgeschlossen, seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten und sei $\varphi : V \to W$ ein K-Morphismus. Dann gilt für alle K-Untervarietäten $U \subseteq V$ in $K[W]$:

$$\mathcal{I}_W(\overline{\varphi(U)}) = (\varphi^*)^{-1}(\mathcal{I}_V(U)).$$

Für Urbilder affiner K-Varietäten unter K-Morphismen sieht die Situation besser aus. Wie wir bereits wissen, sind diese wegen der Stetigkeit von K-Morphismen stets affine K-Varietäten. Insbesondere lassen sich damit Fasern von K-Morphismen berechnen (vgl. [Kra85], AI.2, Bemerkung 1).

Satz 3.2.20. (Urbilder von K-Varietäten)
Sei L algebraisch abgeschlossen, seien $V \subseteq \mathbb{A}_L^n$ und $W \subseteq \mathbb{A}_L^m$ affine K-Varietäten und sei $\varphi : V \to W$ ein K-Morphismus. Dann gilt für alle K-Untervarietäten $U \subseteq W$:

$$\varphi^{-1}(U) = \mathcal{Z}_V(\langle \varphi^*(f) : \ f \in \mathcal{I}_W(U) \rangle).$$

3.3 Affine Varietäten

Bislang haben wir affine Varietäten als spezielle Teilmengen des affinen Raumes $\mathbb{A}_L^n$ kennengelernt. Allerdings lässt sich das Konzept der affinen Varietäten auch auf andere Menge anwenden. Wir wollen in diesem Abschnitt den Begriff der affinen K-Varietät auf eine allgemeinere Ebene heben, so wie ihn beispielsweise Hans-Peter Kraft in dem Buch [Kra85] von Anfang an eingeführt hat. Dazu sei K ein nicht-endlicher Körper und L ein Erweiterungskörper von K. Die K-Algebra aller K-Morphismen auf einer Menge X bezeichnen wir analog mit $K[X]$. Diese Menge kann natürlich auch leer sein. Typische Mengen, die wir betrachten werden, sind endlichdimensionale Vektorräume, wie z.B. $\mathrm{Mat}_n(K)$ oder $\mathscr{P}_{\leq n}(K^2, K)$, und bestimmte Teilmengen davon.

Kraft bezeichnet die in unserem Sinne definierten affinen K-Varietäten als abgeschlossene Untervarietäten und führt dann den Begriff der affinen Varietäten für beliebigere Mengen ein. Allerdings erschien uns der Begriff „Untervarietäten" in diesem Kontext als nicht ganz passend und eher verwirrend, da man intuitiv Teilmengen von affinen Varietäten damit assoziieren würde, die selber wieder affine Varietäten sind; eben genau so, wie wir oben Untervarietäten eingeführt haben. Wir werden daher im Gegensatz zu *affinen K-Varietäten* von *affinen Varietäten* reden, allerdings nach dieser Definition auf diesen Unterschied nur mehr dann eingehen, wenn es angebracht ist, da affine K-Varietäten Spezialfälle im Sinne der folgenden Definition sind (vgl. [Kra85], AI.1.6, S. 233).

Definition 3.3.1. (Affine Varietät)
Eine Menge X zusammen mit einer K-Algebra $K[X]$ heißt **affine Varietät**, falls es eine affine K-Varietät $V \subseteq \mathbb{A}_L^n$ (für eine geeignete Zahl $n \in \mathbb{N}$) und eine bijektive Abbildung $\varphi : X \to V$ gibt, die $K[V]$ mit $K[X]$ identifiziert, d.h. der durch $f \mapsto f \circ \varphi$ definierte K-Algebra-Homomorphismus $\varphi^* : K[V] \to K[X]$ ist ein Isomorphismus.

Ist insbesondere X bereits eine affine K-Varietät von $\mathbb{A}_L^n$ für ein $n \in \mathbb{N}$, so ist X auch im Sinne dieser Definition eine affine K-Varietät. Dazu ist lediglich $\varphi = \mathrm{id}$ zu setzen. Die Topologie auf X wird durch φ induziert (vgl. [Kra85], AI.1.6, S. 233).

Bemerkung 3.3.2. Die K-Algebra $K[X]$ wird analog der affine **Koordinatenring** von X genannt. Die **Zariski-Topologie** auf X wird als die durch φ induzierte Topologie definiert. Sie ist dabei unabhängig von φ. Für eine Teilmenge $M \subseteq K[X]$ sind die Mengen der Form

$$\mathcal{Z}_X(M) = \{x \in X : \ f(x) = 0 \text{ für alle } f \in M\}$$

die abgeschlossenen Teilmengen von X.

In dem Sinne der letzten Definition lassen sich nun auch andere Mengen mit der Struktur einer affinen K-Varietät versehen, was wir im weiteren Verlauf an verschiedenen Stellen benötigen werden.

Beispiel 3.3.3.

a) Sei $X = \mathrm{Mat}_n(L)$ die Menge aller $n \times n$-Matrizen mit Einträgen aus L. Dann ist X isomorph zu $V = \mathbb{A}_L^{n^2}$, d.h. es gibt eine bijektive Abbildung $\varphi : X \to V$. Wegen $\mathcal{I}(V) = \langle 0 \rangle$ können wir $K[V]$ mit dem Polynomring

$$K[x_{1,1}, x_{1,2}, \ldots, x_{1,n}, \ldots, x_{n,1}, \ldots, x_{n,n}]$$

identifizieren. Elemente von $K[X]$, also Abbildungen auf X mit Werten in L, sind z.B. die Determinante oder die Spur. Wie sich leicht zeigen lässt, ist $\varphi^* : K[V] \to K[X]$ definiert durch $f \mapsto f \circ \varphi$ ein Ringisomorphismus, d.h. wir können auch $K[X]$ mit $K[x_{1,1}, \ldots, x_{n,n}]$ identifizieren. Dann lassen sich die K-Morphismen Determinante und Spur durch die Polynome

$$\det := \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot \prod_{i=1}^{n} x_{i\sigma(i)} \text{ und Spur} := x_{11} + x_{22} + \ldots + x_{nn}$$

in $K[x_{1,1}, \ldots, x_{n,n}]$ repräsentieren.

b) Sei $X = \mathscr{P}_{\leq 2}(L^2, L)$ die Menge aller Polynomfunktionen auf L vom Grad ≤ 2. Dann ist X ein L-Vektorraum der Dimension 6, d.h. X ist isomorph zu $V := L^6$. Sei B eine Basis von X und sei $\kappa_B : X \to V$ die Koordinatenabbildung bzgl. B. Dann ist κ_B ein Isomorphismus. Wegen $\mathcal{I}(V) = \langle 0 \rangle$ können wir $K[V]$ mit $K[x_1, \ldots, x_6]$ identifizieren. Wie sich auch hier leicht nachweisen lässt, ist $\varphi^* : K[V] \to K[X]$ definiert durch $f \mapsto f \circ \kappa_B$ ein Ringisomorphismus, womit wir analog $K[X]$ mit $K[x_1, \ldots, x_6]$ identifizieren können.

Sei z.B. $B = (1, y, x, xy, y^2, x^2)$ und $g \in \mathscr{P}_{\leq 2}(L^2, L)$, d.h. $g : L^2 \to L$ ist definiert durch

$$(x, y) \mapsto a_1 \cdot 1 + a_2 \cdot y + a_3 \cdot x + a_4 \cdot xy + a_5 \cdot y^2 + a_6 \cdot x^2.$$

Dann gilt $f \circ \kappa_B(g) = f(\kappa_B(g)) = f(a_1, \ldots, a_6)$ für alle $f \in K[x_1, \ldots, x_6]$. Ein K-Morphismus aus $K[X]$ operiert also auf den Koordinaten aller Polynomfunktionen aus $\mathscr{P}_{\leq 2}(L^2, L)$, womit der Begriff *Koordinatenring* greifbar wird. ◁

Auch der Begriff eines K-Morphismus lässt sich verallgemeinern (vgl. [Kra85], AI.2.1, S. 239).

Definition 3.3.4. (Morphismus)
Seien X, Y affine Varietäten. Eine Abbildung $\varphi : X \to Y$ heißt ein **Morphismus**, falls $f \circ \varphi$ für alle $f \in K[Y]$ ein Element von $K[X]$ ist.

Sind X, Y affine K-Varietäten, so fällt diese Definition mit Definition 3.2.1 zusammen. Auch die Koordinatenabbildung $\varphi^* : K[Y] \to K[X]$ ist analog zu verstehen.

Bemerkung 3.3.5. Außerdem lassen sich sämtliche Begriffe und Ergebnisse über K-Morphismen sinngemäß auch auf affine Varietäten und deren Morphismen übertragen (siehe [Kra85]).

Dazu betrachten wir folgendes Beispiel.

Beispiel 3.3.6. Sei L algebraisch abgeschlossen. Wir betrachten als Teilmenge von $\mathrm{Mat}_n(L)$ die spezielle lineare Gruppe $\mathrm{SL}_n(L) := \{\mathcal{A} \in \mathrm{Mat}_n(L) : \det(\mathcal{A}) = 1\}$. Sei I das Ideal $\langle \det -1 \rangle$ in $K[\mathrm{Mat}_n(L)]$. Dann gilt $\mathrm{SL}_n(L) = \mathcal{Z}_{\mathrm{Mat}_n(L)}(I)$, d.h. $\mathrm{SL}_n(L)$ ist eine affine Untervarietät von $\mathrm{Mat}_n(L)$. Da das Ideal I zudem ein Radikalideal ist, folgt aus dem Hilbertschen Nullstellensatz $\mathcal{I}_{\mathrm{Mat}_n(L)}(\mathrm{SL}_n(L)) = \langle \det -1 \rangle$ und damit erhalten wir für den Koordinatenring von $\mathrm{SL}_n(L)$ die Darstellung

$$K[\mathrm{SL}_n(L)] \cong K[x_{11}, \ldots, x_{nn}]/\langle \det -1 \rangle.$$

◁

Affine K-Varietäten $V \subseteq \mathbb{A}_L^n$ sind stets abgeschlossen bzgl. der Zariski-Topologie. Wir werden nun sehen, dass Zariski-offene Mengen auch affine Varietäten sein können. Dazu definieren wir folgende Mengen (vgl. [Kra85], AI.1.7, S. 235).

Definition 3.3.7. (Spezielle offene Mengen)
Sei X eine affine Varietät und sei $f \in K[X]$. Die Zariski-offene Menge

$$X_f := X \setminus \mathcal{Z}_X(\langle f \rangle) = \{x \in X : f(x) \neq 0\}$$

heißt **spezielle offene Mengen**.

Die speziellen offenen Teilmengen X_f einer affinen Varietät X bilden eine Basis der Zariski-Topologie von X (vgl. [Kra85], AI.1.7). Ist $f \in K[X]$, so ist die Funktion $\frac{1}{f}$ auf X_f definiert. Wir bezeichnen mit $K[X_f]$ die K-Algebra aller Funktionen, die erzeugt werden durch die Funktion $\frac{1}{f}$ und den Einschränkungen $h|_{X_f}$ aller Funktionen $h \in K[X]$. Wie uns der nächste Satz zeigt, sind die speziellen offenen Mengen affine Varietäten (vgl. [Kra85], AI.1.7).

Satz 3.3.8. (Spezielle offene Mengen als affine Varietäten)
Sei L algebraisch abgeschlossen, X eine affine Varietät und $f \in K[X]$. Dann ist X_f zusammen mit $K[X_f]$ eine affine Varietät und es gilt $K[X_f] \cong K[X][z]/\langle f \cdot z - 1 \rangle$.

Ein bekanntes Beispiel für eine spezielle offene Menge ist die allgemeine lineare Gruppe, die Menge aller invertierbaren $n \times n$-Matrizen.

Beispiel 3.3.9. Sei $X = \mathrm{Mat}_n(L)$. Dann ist $X_{\det} = \mathrm{Mat}_n(L) \setminus \{\mathcal{A} \in \mathrm{Mat}_n(L) : \det(\mathcal{A}) \neq 0\}$ die allgemeine lineare Gruppe $\mathrm{GL}_n(L)$, d.h. die Menge aller invertierbaren Matrizen. Somit ist $\mathrm{GL}_n(L)$ eine spezielle offene Menge von $\mathrm{Mat}_n(L)$ und mit dem letzten Satz eine affine Varietät mit Koordinatenring $K[\mathrm{GL}_n(L)] = K[x_{11}, \ldots, x_{nn}, \frac{1}{\det}] \cong K[x_{11}, \ldots, x_{nn}, z]/\langle z \cdot \det -1 \rangle$. ◁

KAPITEL 4

Lineare algebraische Gruppen

Ludwig Marcuse[12]

Wissenschaft steht im Dienste eines Ideals oder im Dienste einer herrschenden Gruppe.

Das Zitat des deutsch-amerikanischen Philosophen und Schriftstellers Ludwig Marcuse (1894–1971) kann natürlich auch mit den Augen eines Mathematikers gelesen werden: Besonders die Invariantentheorie, die ab Kapitel 6 behandelt werden wird, steht im Dienste von Gruppen. Wie wir bereits in der Einleitung gesehen haben, bilden Gruppen das entscheidende Element in der Invariantentheorie.

Von besonderem Interesse und großer Bedeutung für diese Arbeit und die Invariantentheorie ist eine ganz spezielle Klasse von Gruppen, die sogenannten **linearen algebraischen Gruppen**, denen dieses Kapitel gewidmet ist. Wir werden hier zunächst den Begriff einer linear algebraischen Gruppe definieren und erste Eigenschaften betrachten. Im zweiten Abschnitt wollen wir Gruppenoperationen behandeln, wobei wir speziell Operationen von linear algebraischen Gruppen im Fokus haben. Ebenfalls um Gruppenoperationen dreht sich der dritte Abschnitt, in dem wir lineare Darstellungen von Gruppen untersuchen. Mit Hilfe von linearen Darstellungen lassen sich schließlich spezielle linear algebraische Gruppen definieren, die sich durch eine schöne Eigenschaft auszeichnen werden, die sogenannten linear reduktiven Gruppen. Zuletzt werden wir uns der Lie-Algebra linear algebraischer Gruppen zuwenden, da diese uns auch im Kapitel über Invariantentheorie begegnen wird, wenn wir uns mit dem Reynolds-Operator beschäftigen werden.

Wir wollen hier allerdings die einzelnen Abschnitt möglichst kurz und prägnant darstellen. Für eine vertiefte und ausführliche Behandlung der einzelnen Themen sei auf die Bücher [Fis08] von Gerd Fischer, [Art93] von Michael Artin, [Spr80] von Tonny Albert Springer, [Bor91] von Armand Borel, [Hum81] von James E. Humphreys, [Ber07] von Rolf Berndt, [GW10] von Roe Goodman und Nolan R. Wallach, [Kra85] von Hanspeter Kraft oder auf den Anhang des Buches [DK02] von Harm Derksen und Gregor Kemper verwiesen.

[12]Bildquelle: http://hpd.de/node/3762 vom 29.12.2013.

4.1 Definition und erste Eigenschaften

In diesem Abschnitt wollen wir in erster Linie den Darstellungen aus [Spr80] und [Kra85] folgen. Weitere Quellen zur Theorie linear algebraischer Gruppen finden sich beispielsweise in den bereits erwähnten Büchern [Bor91], [Hum81] und [GW10] sowie in Ansätzen auch in [DK02]. Bevor wir auf die Definition eingehen, wollen wir den Begriff einer Gruppe wiederholen, allerdings auf eine leicht unterschiedliche Weise, als dies in den meisten Standardlehrbüchern der Fall ist. Sei dazu im Folgenden K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$.

Eine Menge G wird zusammen mit einer Verknüpfung $*$ zu einer Gruppe, falls es ein eindeutig bestimmtes Element $e \in G$ und Abbildungen $\pi : G \times G \to G$ sowie $\iota : G \to G$ gibt, die für alle $a, b, c \in G$ folgende Eigenschaften erfüllen:

(i) $\pi(a, e) = \pi(e, a) = a$, (neutrales Element)

(ii) $\pi(a, \iota(a)) = \pi(\iota(a), a) = e$, (inverse Elemente)

(iii) $\pi(a, \pi(b, c)) = \pi(\pi(a, b), c)$. (Assoziativität)

Die Abbildungen π und ι legen also die Gruppenstruktur von G fest. Wie es aus Lehrbüchern bekannt ist, schreibt man dann auch $a*b$ und a^{-1} anstatt $\pi(a, b)$ und $\iota(a)$. Mit den angegebenen Abbildungen π und ι ist es möglich, in folgendem Sinne den Begriff einer „linearen algebraischen Gruppe" einzuführen (vgl. [Spr80], S. 31).

Definition 4.1.1. (Lineare algebraische Gruppe)
Eine **lineare algebraische Gruppe** ist eine affine Varietät G, die gleichzeitig die Struktur einer Gruppe hat, so dass $\pi : G \times G \to G$ und $\iota : G \to G$ Morphismen von affinen Varietäten sind.

Es ist durchaus auch gebräuchlich, nur kurz von **algebraischer Gruppe** zu reden, häufig ist dabei aber die Gruppe in einem etwas allgemeineren Kontext zu verstehen (vgl. z.B. [Spr80]). Auf die Unterschiede zwischen algebraischer und linear algebraischer Gruppe soll hier jedoch nicht näher eingegangen werden, da es für unsere Zwecke nicht von Belang ist. Sollte dennoch kurz nur von einer „algebraischen Gruppe" die Rede sein, haben wir im weiteren stets eine lineare algebraische Gruppe im Sinn.

Beispiel 4.1.2. (Allgemeine lineare Gruppe)
Die allgemeine lineare Gruppe $\mathrm{GL}_n(K) = \{\mathcal{A} \in \mathrm{Mat}_n(K) : \det(\mathcal{A}) \neq 0\}$ mit der Matrixmultiplikation als Verknüpfung zugleich eine affine Varietät. Mit den beiden Morphismen $\pi : \mathrm{GL}_n(K) \times \mathrm{GL}_n(K) \to \mathrm{GL}_n(K)$ und $\iota : \mathrm{GL}_n(K) \to \mathrm{GL}_n(K)$ definiert durch $\pi(\mathcal{A}, \mathcal{B}) = \mathcal{A} \cdot \mathcal{B}$ und $\iota(\mathcal{A}) = \mathcal{A}^{-1}$ ist $\mathrm{GL}_n(K)$ also eine lineare algebraische Gruppe. ◁

Wie in [DK02], A.1, oder [Spr80], 2.1, dargestellt, lassen sich die linearen algebraischen Gruppen auch mit Hilfe der zu den Morphismen $\pi : G \times G \to G$ und $\iota : G \to G$ gehörenden Koordinatenabbildungen $\pi^* : K[G] \to K[G] \otimes_K K[G]$ und $\iota^* : K[G] \to K[G]$ beschreiben. Das neutrale Element $e \in G$ korrespondiert mit der zur konstanten Abbildung $a \mapsto e$ gehörenden Koordinatenabbildung $\varepsilon : K[G] \to K[G]$. Da diese Abbildung konstant ist, gilt

$$\varepsilon(f) = f \circ \varepsilon = f(e),$$

womit insbesondere $\varepsilon(K[G]) \subseteq K$, also $\varepsilon \in K[G]^*$ gilt. Nun wollen wir weitere Beispiele für lineare algebraische Gruppen angeben (vgl. [DK02], A.1, [Spr80], 2.1.3 und [Kra85], II.1.1).

Beispiel 4.1.3.

a) Sei $G = \mathbb{A}^1_K$ mit der gewöhnlichen Addition als Verknüpfung. Definiere den Morphismus π durch $\pi(x, y) = x + y$ und ι durch $\iota(x) = -x$. Dann ist G eine lineare algebraische Gruppe mit neutralem Element $e = 0$. Der Koordinantenring ist isomorph zu $K[t]$. Die Koordinatenabbildung $\pi^* : K[t] \to K[t] \otimes_K K[t]$ ist definiert durch $\pi^*(t) = t \otimes_K 1 + 1 \otimes_K t$ und $\iota^* : K[t] \to K[t]$ durch $\iota^*(t) = -t$. Diese Gruppe wird die **additive Gruppe** genannt und mit $\mathrm{Add}(K)$ bezeichnet.

b) Sei $G = \mathbb{A}^1_K \setminus \{0\}$ mit der gewöhnlichen Multiplikation als Verknüpfung. Dann sind die Morphismen π und ι definiert durch $\pi(x, y) = x \cdot y$ und $\iota(x) = x^{-1}$. Der Koordinantenring $K[G]$ ist isomorph zu $K[t, \frac{1}{t}]$ und die zu π und ι gehörenden Koordinatenabbildungen $\pi^* : K[t, \frac{1}{t}] \to K[t, \frac{1}{t}] \otimes_K K[t, \frac{1}{t}]$ und $\iota^* : K[t, \frac{1}{t}] \to K[t, \frac{1}{t}]$ sind definiert durch

$$\pi^*(t) = t \otimes_K t \quad \text{und} \quad \iota(t) = t^{-1}.$$

Weiter ist $\varepsilon : K[t, \frac{1}{t}] \to K$ definiert durch $\varepsilon(t) = 1$. Diese Gruppe heißt die **multiplikative Gruppe**, die mit $\mathrm{Mult}(K)$ bezeichnet wird. ◁

Naheliegend ist die Tatsache, dass jede Zariski-abgeschlossene Untergruppe H einer linearen algebraischen Gruppe G ebenfalls linear algebraisch ist (vgl. [Spr80], 2.1.3(4)). Dazu sind lediglich die Einschränkungen der Morphismen π und ι auf H zu betrachten. Insbesondere die Zarisiki-abgeschlossenen Untergruppen von $\mathrm{GL}_n(K)$ liegen dabei in unserem Interesse. Sie werden häufig auch nur kurz als **Matrixgruppen** bezeichnet. Die endlichen Untergruppen von $\mathrm{GL}_n(K)$, also die endlichen Matrixgruppen, sind somit linear algebraisch, aber auch andere bekannte Beispiele (vgl. [Kra85], II.1.1, [Hum81], S. 52 und [Spr80], 2.1.3).

Beispiel 4.1.4. (Abgeschlossene Untergruppen von $\mathrm{GL}_n(K)$)

a) **Spezielle lineare Gruppe:** $\mathrm{SL}_2(K) = \{\mathcal{A} \in \mathrm{GL}_n(K) : \ \det(\mathcal{A}) = 1\}$

b) Gruppe der **invertierbaren oberen Dreiecksmatrizen:**

$$T_n(K) = \{(a_{i,j}) \in \mathrm{GL}_n(K) : a_{ij} = 0 \text{ für } i > j\}$$

c) **Orthogonale Gruppe:** $\mathrm{O}_n(K) := \{\mathcal{A} \in \mathrm{GL}_n(K) : \ \mathcal{A} \cdot \mathcal{A}^{\mathrm{tr}} = \mathcal{I}_n\}$

d) **Spezielle orthogonale Gruppe:**

$$\mathrm{SO}_n(K) := \mathrm{SL}_n(K) \cap \mathrm{O}_n(K) = \{\mathcal{A} \in \mathrm{GL}_n(K) : \ \mathcal{A} \cdot \mathcal{A}^{\mathrm{tr}} = \mathcal{I}_n \text{ und } \det(\mathcal{A}) = 1\}$$

e) Die **unitäre Gruppe**, d.h. die Gruppe der **oberen Dreiecksmatrizen** mit Einsen auf der Diagonale: $U_n(K) := \{(a_{i,j}) \in \mathrm{GL}_n(K) : \ a_{i,j} = 0 \text{ für } i > j \text{ und } a_{i,i} = 1\}$.

f) Die **spezielle unitäre Gruppe**: $\mathrm{SU}_n(K) := U_n(K) \cap \mathrm{SL}_n(K)$.

g) Die Gruppe der **invertierbaren Diagonalmatrizen**:

$$D_n(K) := \{(a_{i,j}) \in \mathrm{GL}_n(K) : \ a_{i,j} \neq 0 \text{ für } i = j \text{ und } a_{i,j} = 0 \text{ für } i \neq j\}.$$

Eine linear algebraische Gruppe T, die isomorph ist zu $D_n(K)$ für ein $n \in \mathbb{N}_+$, heißt ein n-dimensionaler **Torus**. Die multiplikative Gruppe $\mathrm{Mult}(K)$ ist ein 1-dimensionaler Torus. ◁

In [Kra85] werden lineare algebraische Gruppen sogar als abgeschlossene Untergruppen von $\mathrm{GL}_n(K)$ definiert und eingeführt, was durchaus seine Berechtigung hat, wie wir später sehen werden. Zunächst wollen wir erste Eigenschaften linearer algebraischer Gruppen angeben. Dabei bezeichnen wir mit G^0 die Zusammenhangskomponente von G, die das neutrale Element $e \in G$ enthält. Laut [Spr80], 2.2.1, S. 36 ist G^0 für eine lineare algebraische Gruppe eindeutig bestimmt. Damit können wir folgenden Satz festhalten (vgl. [Kra85], II.1.2, Satz 1 und [Bor91], I.1.2).

Satz 4.1.5. (Eigenschaften linearer algebraischer Gruppen)
Sei $(G, *, e)$ eine lineare algebraische Gruppe.

a) Die irreduziblen Komponenten von G sind die Zusammenhangskomponenten.

b) G^0 ist ein offener und abgeschlossener Normalteiler von G.

c) Jede abgeschlossene Untergruppe von G mit endlichem Index enthält G^0, insbesondere ist also der Index von G^0 endlich.

*d) Das **Zentrum** $Z(G) = \{a \in G : \ a*x = x*a$ für alle $x \in G\}$ von G ist ein abgeschlossener Normalteiler von G.*

Aus dem Satz folgt sofort, dass für lineare algebraische Gruppen zusammenhängend und irreduzibel äquivalente Begriffe sind (vgl. [Kra85], S. 55). In der Literatur spricht man aber dennoch bevorzugt in diesem Kontext von *zusammenhängenden* linearen algebraischen Gruppen, da in einem späteren Abschnitt über Darstellungen von Gruppen (siehe Abschnitt 4.3) der Begriff „irreduzibel" gebräuchlicher ist (vgl. [Spr80], S. 37).

Definition 4.1.6. (Zusammenhängende lineare algebraische Gruppe)
Eine lineare algebraische Gruppe G heißt **zusammenhängend**, wenn $G = G^0$ gilt.

So sind beispielsweise $\mathrm{Add}(K)$ und $\mathrm{Mult}(K)$ zusammenhängend, ebenso wie $\mathrm{GL}_n(K)$ und $\mathrm{SL}_n(K)$ [Hum81], S. 53). Ein Beispiel einer nicht zusammenhängenden Gruppe ist die orthogonale Gruppe (vgl. [Kra85], S. 56 f.).

Beispiel 4.1.7. Betrachten wir die orthogonale Gruppe $\mathrm{O}_n(K)$. Wegen $\det(\mathrm{O}_n(K)) = \{\pm 1\}$ ist $\mathrm{O}_n(K)$ nicht zusammenhängend. Es gilt $\mathrm{O}_n(K)^0 = \mathrm{SO}_n(K)$ und $\mathrm{SO}_n(K)$ ist zusammenhängend. Für die Zentren gilt:

$$Z(\mathrm{O}_n(K)) = \{\pm \mathcal{I}_n\} \quad \text{und} \quad Z(\mathrm{SO}_n(K)) = \begin{cases} \{\pm \mathcal{I}_n\}, & \text{für } n > 2 \text{ gerade}, \\ \{\mathcal{I}_n\}, & \text{für } n \text{ ungerade}. \end{cases}$$

◁

Nach diesen ersten Eigenschaften wollen wir nun bestimmte Abbildungen zwischen linearen algebraischen Gruppen betrachten, deren Namen wir zunächst angeben wollen [Hum81], S. 51).

Definition 4.1.8. (Regulärer Gruppenhomomorphismus)
Seien G und H zwei lineare algebraische Gruppen.

a) Eine Abbildung $\Phi : G \to H$ heißt ein **regulärer Gruppenhomomorphismus**, wenn Φ gleichzeitig ein Homomorphismus von Gruppen und ein Morphismus von affinen Varietäten ist.

b) Ein regulärer Gruppenhomomorphismus $\Phi : G \to H$ heißt ein **regulärer Isomorphismus**, wenn Φ bijektiv und Φ^{-1} ebenfalls ein regulärer Gruppenhomomorphismus ist. Im Falle von $G = H$ heißt ein regulärer Isomorphismus Φ ein **regulärer Automorphismus**.

Ist aus dem Zusammenhang klar, dass Φ eine Abbildung linearer algebraischer Gruppen ist, so sprechen wir kurz auch nur von einem Isomorphismus bzw. von einem Automorphismus. Ein bekanntes Beispiel eines Automorphismus ist die Konjugation linearer algebraischer Gruppen.

Beispiel 4.1.9.

a) Die Determinante $\det : \mathrm{GL}_n(K) \to \mathrm{Mult}(K)$ ist ein surjektiver regulärer Gruppenhomomorphismus.

b) Die Gruppe $D_n(K)$ der invertierbaren Diagonalmatrizen ist eine Untergruppe der Gruppe der invertierbaren oberen Dreiecksmatrizen $T_n(K)$. Sei $\Phi : T_n(K) \to D_n(K)$ die kanonische Abbildung von $T_n(K)$ in $D_n(K)$, d.h. die Abbildung, die die Diagonaleinträge fest lässt und alle anderen Einträge auf 0 abbildet. Dann ist Φ ein surjektiver regulärer Gruppenhomomorphismus mit $\mathrm{Ker}(\Phi) = U_n(K)$. ◁

Im letzten Beispiel ist der Kern von Φ eine abgeschlossene Untergruppe von $\mathrm{GL}_n(K)$, also eine lineare algebraische Gruppe, ebenso sind in beiden Fällen die Bilder der regulären Gruppenhomomorphismen lineare algebraische Gruppen. Dieses Ergebnis ist kein Zufall, denn Kern und Bild regulärer Gruppenhomomorphismen sind stets abgeschlossen, der Kern ist sogar ein Normalteiler (vgl. [Spr80], 2.2.4, S. 38). Somit sind Kern und Bild von regulären Gruppenhomomorphismen also selbst linear algebraische Gruppen. Nun wollen wir zu der Begründung kommen, warum beispielsweise in [Kra85] als lineare algebraische Gruppen nur abgeschlossene Untergruppen von $\mathrm{GL}_n(K)$ betrachtet werden (vgl. [Bor91], I.1.10 und [Kra85], S. 236).

Satz 4.1.10. *Jede lineare algebraische Gruppe G ist isomorph zu einer (Zarisiki-) abgeschlossenen Untergruppe H der allgemeinen linearen Gruppe $\mathrm{GL}_n(K)$ (für ein $n \in \mathbb{N}_+$), d.h. für jede lineare algebraische Gruppe G gibt es eine abgeschlossene Untergruppe H von $\mathrm{GL}_n(K)$ (für ein $n \in \mathbb{N}_+$) und einen regulären Isomorphismus $\Phi : G \to H$.*

Dieser Satz liefert in gewisser Weise auch eine Rechtfertigung für das Adjektiv *linear*. Da lineare algebraische Gruppen die Struktur einer affinen Varietät haben, wäre es intuitiv vielleicht naheliegender, nicht von *linearen* algebraischen Gruppen, sondern von *affinen* algebraischen Gruppen zu sprechen. Der letzte Satz zeigte uns nun aber, dass jede derartige Gruppe isomorph zu einer Untergruppe der allgemeinen *linearen* Gruppe ist, womit das Adjektiv „linear" seine Berechtigung und Erklärung erfährt (vgl. [Spr80]).

Beispiel 4.1.11.

a) Die multiplikative Gruppe $\mathrm{Mult}(K)$ ist auf natürliche Weise isomorph zu $\mathrm{GL}_1(K)$.

b) Die additive Gruppe $\mathrm{Add}(K)$ ist isomorph zur Untergruppe

$$H = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} : \ a \in K \right\}$$

von $\mathrm{GL}_2(K)$. Der Isomorphismus $\Phi : \mathrm{Add}(K) \to H$ ist dabei gegeben durch $a \mapsto \left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right)$.

c) Ein Automorphismus φ von K^n heißt eine **affine Transformation**, wenn es eine Matrix $\mathcal{A} \in \mathrm{GL}_n(K)$ und einen Vektor $v \in K^n$ gibt mit $\varphi(x) = \mathcal{A} \cdot x + v$ für alle $x \in K^n$. Die Menge der affinen Transformationen bildet mit der Komposition als Verknüpfung eine Gruppe mit neutralem Element id_{K^n}, wobei sich die Identität id_{K^n} schreiben lässt durch $\mathrm{id}_{K^n}(x) = \mathcal{I}_n \cdot x + 0$ für alle $x \in K^n$. Seien $\varphi : K^n \to K^n$ und $\psi : K^n \to K^n$ definiert durch $\varphi(x) = \mathcal{A} \cdot x + v$ und $\psi(x) = \mathcal{B} \cdot x + w$ mit $\mathcal{A}, \mathcal{B} \in \mathrm{GL}_n(K)$ und $v, w \in K^n$ zwei affine Transformationen. Dann ist die affine Transformation $\varphi \circ \psi : K^n \to K^n$ definiert

durch $\varphi \circ \psi(x) = \mathcal{A} \cdot (\mathcal{B} \cdot x + w) + v = \mathcal{AB} \cdot x + (\mathcal{A} \cdot w + v)$. Das zu φ inverse Element ist die affine Transformation $\varphi^{-1} : K^n \to K^n$ mit $\varphi^{-1}(x) = \mathcal{A}^{-1} \cdot x + \mathcal{A}^{-1} \cdot (-v)$. Durch die Zuordnung

$$\varphi \mapsto \begin{pmatrix} \mathcal{A} & v \\ 0 & 1 \end{pmatrix}$$

ist ein injektiver Gruppenhomomorphismus gegeben, durch den sich die Gruppe der affinen Transformationen in die Gruppe $\mathrm{GL}_{n+1}(K)$ einbetten lässt (vgl. [Küh11], Lemma 5.3, S. 37). Somit ist die Gruppe der affinen Transformationen auf natürliche Weise isomorph zum Bild dieses injektiven Gruppenhomomorphismus, also zur abgeschlossenen Untergruppe

$$\left\{ \begin{pmatrix} \mathcal{A} & v \\ 0 & 1 \end{pmatrix} : \ \mathcal{A} \in \mathrm{GL}_n(K) \text{ und } v \in K^n \right\}$$

von $\mathrm{GL}_{n+1}(K)$. Diese Matrixgruppe heißt die **n-dimensionale affine Gruppe** und wird mit $\mathrm{AGL}_n(K)$ bezeichnet (vgl. [Küh11], Def. 5.2, S. 36). Da die Elemente von $\mathrm{AGL}_n(K)$ eindeutig durch die Paare $(\mathcal{A}, v) \in \mathrm{GL}_n(K) \times K^n$ identifiziert werden können, verwenden wir auch diese Paare um die Elemente von $\mathrm{AGL}_n(K)$ zu charakterisieren. ◁

Wir haben bereits bemerkt, dass jede endliche Untergruppe von $\mathrm{GL}_n(K)$ eine lineare algebraische Gruppe ist. Mit dem letzten Satz folgt nun, dass jede beliebige endliche Gruppe linear algebraisch ist. Damit ist also die Theorie der endlichen Gruppen Teil der Theorie der linearen algebraischen Gruppen. Wir wollen an dieser Stelle nun auf die Beschreibung linear algebraischer Gruppen eingehen, für die wir sowohl ihre Struktur einer affinen Varietät als auch den letzten Satz ausnutzen können.

Bemerkung 4.1.12. (Algorithmische Beschreibung linear algebraischer Gruppen)
Da eine linear algebraische Gruppe G insbesondere eine affine Varietät ist, lässt sie sich in $\mathbb{A}^\ell_K$ für ein $\ell \in \mathbb{N}_+$ einbetten, d.h. es gibt eine abgeschlossene Einbettung $\iota : G \hookrightarrow \mathbb{A}^\ell_K$. Dank dieser Einbettung lässt sich G als Zarisiki-abgeschlossene Teilmenge von $\mathbb{A}^\ell_K$ betrachten. Wegen $K[\mathbb{A}^\ell_K] \cong K[z_1, \ldots, z_\ell]$ können wir die zugehörige Koordinatenabbildung $\iota^* : K[\mathbb{A}^\ell_K] \to K[G]$ als surjektiven Ringhomomorphismus $\iota^* : K[z_1, \ldots, z_\ell] \twoheadrightarrow K[G]$ auffassen. Laut Satz 3.2.7 gilt

$$K[G] \cong K[z_1, \ldots, z_\ell]/\mathcal{I}(G),$$

wobei das Verschwindungsideal $\mathcal{I}(G)$ genau der Kern von ι^* ist. Da wir meistens Matrixgruppen $G \subseteq \mathrm{GL}_m(K)$ für ein $m \in \mathbb{N}_+$ verwenden werden, werden wir in diesen Fällen G in $\mathrm{Mat}_m(K)$ einbetten. Wir verwenden dann eine intuitiv naheliegendere Doppelindizierung der Unbestimmten zur Darstellung von $K[G]$, d.h. wir schreiben $K[G] \cong K[z_{1,1}, \ldots, z_{m,m}]/\mathcal{I}(G)$.

Dies wollen wir kurz an typischen Beispielen demonstrieren.

Beispiel 4.1.13.

a) Zunächst betrachten wir als Beispiel (vgl. in Ansätzen [DK02], Beispiel 4.1.7, S. 143) die spezielle lineare Gruppe $G := \mathrm{SL}_2(K) = \{\mathcal{A} \in \mathrm{Mat}_2(K) : \ \det(\mathcal{A}) = 1\}$, die auf natürliche Weise durch $\iota : G \hookrightarrow \mathrm{Mat}_2(K)$ in $\mathrm{Mat}_2(K)$ eingebettet werden kann. Somit ist die Koordinatenabbildung der Einbettung ι der surjektive Ringhomomorphismus $\iota^* : K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}] \twoheadrightarrow K[G]$. Mit dem Verschwindungsideal von G, also mit $\mathcal{I}(\mathrm{SL}_2(K)) = \langle \det((z_{i,j})_{1 \le i,j \le 2}) - 1 \rangle$, gilt $K[G] \cong K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$.

b) Die orthogonale Gruppe $G := \mathrm{O}_2(K)$ lässt sich ebenfalls in $\mathrm{Mat}_2(K)$ einbetten. Wegen

$$\begin{pmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & z_{2,2} \end{pmatrix} \cdot \begin{pmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \end{pmatrix} = \begin{pmatrix} z_{1,1}^2 + z_{1,2}^2 & z_{1,1}z_{2,1} + z_{1,2}z_{2,2} \\ z_{1,1}z_{2,1} + z_{1,2}z_{2,2} & z_{2,1}^2 + z_{2,2}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

wird das Verschwindungsideal von G von den Polynomen $z_{1,1}^2 + z_{1,2}^2 - 1$, $z_{1,1}z_{2,1} + z_{1,2}z_{2,2}$ und $z_{2,1}^2 + z_{2,2}^2 - 1$ erzeugt. ◁

Zum Abschluss dieses Abschnitts über linear algebraische Gruppen wollen wir noch kurz deren Produkte beleuchten. Mittels $\Phi : \mathrm{GL}_n(K) \times \mathrm{GL}_m(K) \to \mathrm{GL}_{n+m}(K)$ definiert durch

$$(\mathcal{A}, \mathcal{B}) \mapsto \begin{pmatrix} \mathcal{A} & 0 \\ 0 & \mathcal{B} \end{pmatrix}$$

lässt sich $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ in $\mathrm{GL}_{n+m}(K)$ (mit $n, m \in \mathbb{N}_+$) einbetten. Das direkte Produkt $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ ist mit der üblichen „komponentenweisen" Verknüpfung eine Gruppe (vgl. [Fis08], S. 40) und wie wir wissen, ist dieses Produkt auch eine affine Varietät (vgl. Satz 3.1.6). Da $\mathrm{Im}(\Phi)$ eine abgeschlossene Untergruppe von $\mathrm{GL}_{n+m}(K)$ ist, ist $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ isomorph zu einer abgeschlossenen Untergruppe einer allgemeinen linearen Gruppe und damit eine lineare algebraische Gruppe. Also ist sofort klar, dass das direkte Produkt zweier linear algebraischer Gruppen G und H erneut linear algebraisch ist (vgl. [Hum81], S. 52). Das lässt sich induktiv natürlich auf das Produkt endlich vieler Gruppen ausdehnen. Durch Produktbildung sehen wir nun, dass weitere bekannte Gruppen linear algebraisch sind (vgl. [Hum81], S. 52).

Beispiel 4.1.14. (Produkte linear algebraischer Gruppen)

a) Das n-fache Produkt der multiplikativen Gruppe $\mathrm{Mult}(K)$ ist eine linear algebraische Gruppe. Sie ist isomorph zur Gruppe $D_n(K)$ der invertierbaren Diagonalmatrizen und damit ein n-dimensionaler Torus. Mit anderen Worten: $(K^*)^n$ mit der gewöhnlichen Multiplikation auf K als Verknüpfung ist eine linear algebraische Gruppe.

b) Das n-fache Produkt der additiven Gruppe $\mathrm{Add}(K)$ ist eine lineare algebraische Gruppe. Die Zuordnung

$$v \mapsto \begin{pmatrix} \mathcal{I}_n & v \\ 0 & 1 \end{pmatrix}$$

liefert uns eine Einbettung von $(\mathrm{Add}(K))^n$ in $\mathrm{GL}_{n+1}(K)$, genauer sogar in $U_{n+1}(K)$. Offensichtlich ist das Bild von $(\mathrm{Add}(K))^n$ unter dieser Einbettung,

$$\left\{ \begin{pmatrix} \mathcal{I}_n & v \\ 0 & 1 \end{pmatrix} \in U_{n+1}(K) : \ v \in K^n \right\},$$

eine Untergruppe der affinen Gruppe $\mathrm{AGL}_n(K)$, die wir mit $\mathrm{Trans}_n(K)$ bezeichnen. Um die Beziehung zwischen $\mathrm{AGL}_n(K)$ und $\mathrm{Trans}_n(K)$ deutlich zu machen, werden wir die Elemente von $\mathrm{Trans}_n(K)$ auch als Paar $(\mathcal{I}_n, v)$ mit $v \in K^n$ schreiben. Wie sich leicht nachweisen lässt, ist $\mathrm{Trans}_n(K)$ sogar ein Normalteiler von $\mathrm{AGL}_n(K)$.

Offensichtlich ist $(\mathrm{Add}(K))^n$ somit isomorph zur Gruppe aller Abbildungen T_v, wobei $T_v : K^n \to K^n$ für $v \in K^n$ definiert ist durch $T_v(x) = x + v = \mathcal{I}_n \cdot x + v$, also zur Menge aller **Translationen** in K^n. Deshalb wird die Gruppe $\mathrm{Trans}_n(K)$ auch als **Translationsgruppe** bezeichnet. Die Translationsgruppe ist mit anderen Worten also isomorph zu K^n versehen mit der üblichen Vektoraddition (vgl. [Küh11], Folgerung 5.5, S. 38). ◁

Da direkte Produkte bekanntlich nur Spezialfälle von semidirekten Produkten sind, stellt sich an dieser Stelle sofort die Frage, ob bereits semidirekte Produkte linear algebraischer Gruppen linear algebraisch sind. Wir stellen die Antwort auf diese Frage jedoch etwas zurück, da wir zur besseren Beantwortung dieser Frage noch weitere Begriffe benötigen, wie z.B. Gruppenoperationen, denen der übernächste Abschnitt gewidmet ist.

4.2 Gruppenoperationen

Der Begriff, auf dem die gesamte Invariantentheorie basiert und der erst den Anstoß zu dieser Theorie gegeben hat, ist die Gruppenoperation, also die Operation einer Gruppe auf einer Menge. Nach wie vor sei K ein nicht-endlicher Körper der Charakteristik 0.

Definition 4.2.1. (Gruppenoperation)
Sei M eine nicht-leere Menge und $(G, *, e)$ eine Gruppe. Eine **Operation** von G auf M ist eine Abbildung $\tau : G \times M \to M$ mit den Eigenschaften

(i) $\tau(a * b, x) = \tau(a, \tau(b, x))$,

(ii) $\tau(e, x) = x$

für alle $a, b \in G$ und $x \in M$. Man nennt M dann auch eine **(Links-) G-Menge**.

Genau genommen erhalten wir auf diese Weise eine Linksoperation von G auf M; Rechtsoperationen sind analog definiert, werden für unsere Zwecke aber nicht näher von Bedeutung sein. Operiert G auf einem K-Vektorraum V, so spricht man von einer **linearen Operation**. Wir wollen zuletzt noch den Begriff einer rationalen Gruppenoperation angeben (vgl. [DK02], Definition A.1.7, S. 238 sowie [Dre04], Definition 2.5, S. 7).

Definition 4.2.2. (Rationale Gruppenoperation)
Sei G eine lineare algebraische Gruppe und V ein K-Vektorraum. Eine lineare Operation $\tau : G \times V \to V$ von G auf V heißt **rational**, falls es eine Abbildung $\tau^\star : V \to V \otimes K[G]$ mit

$$\tau(a, v) = \sum_{i=1}^{l} v_i f_i(a)$$

für alle $a \in G$ und alle $v \in V$ mit $\tau^\star(v) = \sum_{i=1}^{l} v_i \otimes f_i$.

Durch Festhalten eines Elements $a \in G$ ergibt sich aus einer Gruppenoperation $\tau : G \times M \to M$ von G auf M eine Abbildung $\tau_a : M \to M$, die stets bijektiv ist, wie in folgender Bemerkung festgehalten ist (vgl. [Fis08], Bemerkung 4.1), d.h. τ_a ist für alle $a \in G$ ein Element der symmetrischen Gruppe $\mathrm{Aut}(M)$. Wir schreiben im Folgenden auch $\tau_a(x)$ oder kurz $a(x)$ an Stelle von $\tau(a, x)$.

Bemerkung 4.2.3. Sei G eine Gruppe und M eine nicht-leere Menge.

a) Jede Gruppenoperation $\tau : G \times M \to M$ induziert durch $a \mapsto \tau_a$ einen Gruppenhomomorphismus $\rho : G \to \mathrm{Aut}(M)$.

b) Jeder Gruppenhomomorphismus $\rho : G \to \mathrm{Aut}(M)$ induziert durch $(a, x) \mapsto \rho_a(x)$ eine Gruppenoperation $\tau : G \times M \to M$.

Eine Operation von G auf M lässt sich gemäß dieser Bemerkung also auch als Gruppenhomomorphismus $\rho : G \to \mathrm{Aut}(M)$ erklären, was wir später in Abschnitt 4.3 ausführlich behandeln werden. Dabei ist das Bild $\rho(G)$ von ρ stets eine Untergruppe von $\mathrm{Aut}(M)$ (vgl. [Fis08], S. 71). Ist der Homomorphismus ρ injektiv, so nennt man die Operation **effektiv**. Durch Übergang zu $\overline{\rho} : G/\mathrm{Ker}(\rho) \to \mathrm{Aut}(M)$ können wir jede Gruppenoperation zu einer effektiven Operation machen. Gibt es zu allen $x, y \in M$ ein $a \in G$ mit $a(x) = y$, so wird die Operation von G auf M **transitiv** genannt. Ist a sogar eindeutig bestimmt, so heißt die Operation **einfach transitiv** (vgl. [Fis08], S. 71).

Beispiel 4.2.4.

a) Durch $\tau : \mathrm{GL}_n(K) \times K^n \to K^n$, definiert durch $(\mathcal{A}, x) \mapsto \mathcal{A} \cdot x$, operiert $\mathrm{GL}_n(K)$ auf K^n. Wegen $\mathcal{A} \cdot 0 = 0$ für alle $\mathcal{A} \in \mathrm{GL}_n(K)$ ist diese Operation nicht transitiv. Ebenso operiert $\mathrm{SL}_n(K)$ nicht transitiv auf K^n. Schließt man allerdings 0 aus, so operiert $\mathrm{GL}_n(K)$ einfach transitiv auf $K^n \setminus \{0\}$. Analog verhält es sich mit der Operation von $\mathrm{SL}_n(K)$.

b) Die affine Gruppe $\mathrm{AGL}_n(K)$ operiert mittels $(\mathcal{T}, x) \mapsto \mathcal{T} \cdot \binom{x}{1}$ transitiv auf K^n, aber nicht einfach transitiv. ◁

Insbesondere die Operationen linearer algebraischer Gruppen auf affinen K-Varietäten werden im weiteren Verlauf von Interesse sein; diese werden als **regulär** bezeichnet (vgl. [Kra85], S. 64).

Definition 4.2.5. (Reguläre Gruppenoperation)
Eine Operation $\tau : G \times M \to M$ einer linearen algebraischen Gruppe G auf einer affinen Varietät M heißt **regulär**, falls τ ein Morphismus ist. Operiert G auf M regulär, so nennt man die G-Menge M auch **(affine) G-Varietät**.

Der durch τ induzierte Gruppenhomomorphismus $G \to \mathrm{Aut}(M)$ ist dann regulär. Wir wollen nun die Operation einer Gruppe $(G, *, e_G)$ auf einer weiteren Gruppe $(H, \cdot, e_H)$ betrachten, d.h. es gibt eine Abbildung $\tau : G \times H \to H$ mit den in Definition 4.2.1 geforderten Eigenschaften. Damit kommen wir nun zu einer Frage zurück, die wir im letzten Abschnitt unbeantwortet ließen. Dort wurden bereits direkte Produkte von linearen algebraischen Gruppen untersucht, mit dem Ergebnis, dass auch diese wieder lineare algebraische Gruppen sind. Nun können wir analog semidirekte Produkte untersuchen. Operiert G regulär auf einer linearen algebraischen Gruppe H, so ist das (äußere) semidirekte Produkt $H \rtimes_\Phi G$ ebenfalls eine lineare algebraische Gruppe (vgl. [Hum81], S. 61). Sei nun umgekehrt eine lineare algebraische Gruppe G' gegeben, seien G und H abgeschlossene Untergruppen von G', wobei H ein Normalteiler von G' sei, und sei $H \rtimes G$ wie oben das (innere) semidirekte Produkt von G und H. Dann ist die obige Abbildung $H \rtimes G \to G'$ ein regulärer Isomorphismus und somit ist auch $H \rtimes G$ eine lineare algebraische Gruppe.

Beispiel 4.2.6. (Semidirekte Produkte von linear algebraischen Gruppen)

a) Setze

$$G := \left\{ \begin{pmatrix} \mathcal{A} & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_{n+1}(K) : \ \mathcal{A} \in \mathrm{GL}_n(K) \right\}.$$

Offensichtlich ist G dann isomorph zu $\mathrm{GL}_n(K)$. Die Gruppen G und $\mathrm{Trans}_n(K)$ sind Untergruppen der affinen Gruppe $\mathrm{AGL}_n(K)$, wobei $\mathrm{Trans}_n(K)$ ein Normalteiler von $\mathrm{AGL}_n(K)$ ist (vgl. Beispiel 4.1.11). Auch hier gilt $G \cap \mathrm{Trans}_n(K) = \{\mathcal{I}_n\}$ und

$$\mathrm{AGL}_n(K) = \mathrm{Trans}_n(K) \cdot G = \{\mathcal{B} \cdot \mathcal{A} : \ \mathcal{B} \in \mathrm{Trans}_n(K), \ \mathcal{A} \in G\}.$$

Die Gruppe G operiere durch Konjugation auf $\mathrm{Trans}_n(K)$, d.h. $\Phi : G \to \mathrm{Aut}(\mathrm{Trans}_n(K))$ ist definiert durch $\Phi_\mathcal{T}(\mathcal{X}) = \mathcal{T} \cdot \mathcal{X} \cdot \mathcal{T}^{-1}$. Dann ist $\mathrm{AGL}_n(K)$ isomorph zum semidirekten Produkt $\mathrm{Trans}_n(K) \rtimes G$.

b) Setze

$$H := \left\{ \begin{pmatrix} \mathcal{A} & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_{n+1}(K) : \ \mathcal{A} \in \mathrm{O}_n(K). \right\}$$

Dann ist das semidirekte Produkt $\mathrm{Trans}_n(K) \rtimes H$ offensichtlich isomorph zur Untergruppe

$$\left\{ \begin{pmatrix} \mathcal{A} & v \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_{n+1}(K) : \ \mathcal{A} \in \mathrm{O}_n(K), \ v \in K^n \right\}$$

von $\mathrm{AGL}_n(K)$. Sie heißt die **Bewegungsgruppe**, **Gruppe der Isometrien** oder **Euklidische Gruppe** und wird mit $\mathrm{Iso}_n(K)$ bezeichnet (vgl. [Küh11], Definition 5.10, S. 41). Auch hier werden wir die Elemente der Euklidischen Gruppe mit den Paaren $(\mathcal{A}, v)$ für $\mathcal{A} \in \mathrm{O}_n(K)$ und $v \in K^n$ identifizieren. Die Untergruppe

$$\left\{ \begin{pmatrix} \mathcal{A} & v \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_{n+1}(K) : \ \mathcal{A} \in \mathrm{SO}_n(K), v \in K^n \right\}$$

von $\mathrm{AGL}_n(K)$ heißt die **eigentliche euklidische Gruppe** und wird mit $\mathrm{Iso}_n^+(K)$ bezeichnet (vgl. [Küh11], Definition 5.10, S. 41). ◁

Nach der Behandlung des semidirekten Produkts werden wir nun einige Begriffe aus der Algebra kurz aufgreifen, die im Zusammenhang mit Gruppenoperationen erwähnenswert sind, und deren spezielle Eigenschaften im Falle von linearen algebraischen Gruppen betrachten. Lässt man beispielsweise jedes Gruppenelement auf einem festen Element einer Menge M operieren, so erhält man eine bestimmte Teilmenge von M, die sogenannte Bahn (vgl. [Fis08], S. 72).

Definition 4.2.7. (Bahn)
Sei M eine nicht-leere Menge und G eine Gruppe, die auf M operiert. Für jedes $x \in M$ heißt die Teilmenge $G(x) := \{a(x) : \ a \in G\}$ von M die **Bahn** oder der **Orbit** von x (unter der Operation von G).

Die Bahn enthält alle Elemente von M, die durch Operationen der Gruppe aus x hervorgehen. Aus dieser Definition erklärt sich sofort eine Äquivalenzrelation auf der Menge M, was sich leicht nachrechnen lässt (vgl. [Fis08], S. 72).

Bemerkung 4.2.8. (Bahnenraum)
Durch

$$x \sim_G y \quad :\Longleftrightarrow \quad y \in G(x)$$

ist eine Äquivalenzrelation auf M erklärt. Die Äquivalenzklassen bzgl. $\sim_G$ sind gerade die Bahnen. Die Menge $M/\sim_G$ der Äquivalenzklassen bzgl. $\sim_G$ heißt der **Bahnenraum** und wird mit M/G bezeichnet.

Mit anderen Worten können zwei Elemente von M als äquivalent angesehen werden, wenn sie sich in derselben Bahn befinden. Damit bilden die Bahnen eine Partition der Menge M. Weiter gilt bereits folgende schwächere Charakterisierung der Äquivalenz zweier Elemente.

Bemerkung 4.2.9. Sei M eine nicht-leere Menge und seien $x, y \in M$. Dann gilt:

$$x \sim_G y \quad \Longleftrightarrow \quad G(x) \cap G(y) \neq \emptyset.$$

Ein Begriff, der besonders im Zusammenhang mit Bahnen auftaucht, ist der Begriff einer G-stabilen Teilmenge (vgl. [Kra85], S. 64).

Definition 4.2.10. (G-stabil)
Sei M eine nicht-leere Menge und G eine Gruppe, die auf M operiert. Eine Teilmenge $N \subseteq M$ heißt **G-stabil**, falls $a(N) \subseteq N$ gilt für alle $a \in G$.

Eine G-stabile Menge verhält sich also invariant unter der Operation von G, was uns später erneut begegnen wird. Wie man auch an dem letzten Beispiel sehen kann, ist jede Bahn unter einer Gruppenoperation eine G-stabile Menge. Nach der Betrachtung typischer Begriffe, die in Zusammenhang mit Gruppenoperationen stehen, werden wir uns zum Abschluss dieses Abschnitts für einen kurzen Moment Abbildungen zwischen G-Mengen zuwenden und dabei insbesondere den Begriff einer G-äquivarianten Abbildung behandeln (vgl. (in speziellerer Form) [Kra85], S. 65), der uns an späterer Stelle erneut begegnen wird.

Definition 4.2.11. (G-äquivariant)
Sei G eine Gruppe und seien M sowie M' zwei Mengen, auf denen G operiert. Eine Abbildung $\varphi : M \to M'$ heißt **G-äquivariant**, wenn $a(\varphi(x)) = \varphi(a(x))$ für alle $a \in G$ und alle $x \in M$ gilt.

Eine Abbildung zwischen zwei G-Mengen heißt mit anderen Worten also G-äquivariant, wenn sie mit den Operationen von G auf M und M' verträglich ist (vgl. [Art93], S. 372). Sind $\tau : G \times M \to M$ und $\tau' : G \times M' \to M'$ die Operationen von G auf M bzw. M', so ist eine Abbildung $\varphi : M \to M'$ genau dann G-äquivariant, wenn für alle $a \in G$ und alle $x \in M$ gilt

$$\tau'(a, \varphi(x)) = \varphi(\tau(a, x)).$$

Hält man ein $a \in G$ fest, so erhalten wir bekanntlich aus τ und τ' Abbildungen $\tau_a : M \to M$ und $\tau'_a : M' \to M'$. Somit ist eine Abbildung $\varphi : M \to M'$ genau dann G-äquivariant, wenn für alle $a \in G$ das folgende Diagramm kommutiert:

$$\begin{array}{ccc} M & \xrightarrow{\tau_a} & M \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ M' & \xrightarrow[\tau'_a]{} & M' \end{array}$$

Beispiel 4.2.12. Sei $M = M' = \{(x, y) \in \mathbb{R}^2 : \ x^2 + y^2 = 1\}$ der Einheitskreis in $\mathbb{R}^2$, auf dem $G = \mathrm{SO}_2(\mathbb{R})$ durch Matrixmultiplikation operiere. Sei $\varphi : M \to M$ definiert durch $\varphi(x, y) = (-x, -y)$. Sei $\mathcal{A}(\vartheta) \in G$ von der Form $\left(\begin{smallmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{smallmatrix}\right)$. Dann gilt für alle $(x, y) \in M$:

$$\begin{aligned} \mathcal{A}(\vartheta) \cdot \varphi \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix} \cdot \begin{pmatrix} -x \\ -y \end{pmatrix} = \begin{pmatrix} -x\cos(\vartheta) + y\sin(\vartheta) \\ -x\sin(\vartheta) - y\cos(\vartheta) \end{pmatrix} \\ &= -\begin{pmatrix} x\cos(\vartheta) - y\sin(\vartheta) \\ x\sin(\vartheta) + y\cos(\vartheta) \end{pmatrix} = \varphi\left(\mathcal{A}(\vartheta) \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) \end{aligned}$$

Somit ist φ eine G-äquivariante Abbildung. ◁

4.3 Lineare Darstellungen

Wie wir aus Bemerkung 4.2.3 wissen, lässt sich die Operation einer Gruppe G auf einer Menge M auch als Gruppenhomomorphismus $G \to \mathrm{Aut}(M)$ erklären. Operiert die Gruppe auf einem K-Vektorraum, so nennt man diesen Gruppenhomomorphismus eine lineare Darstellung. Einige wichtige Eigenschaften linearer Darstellungen sollen hier nun kurz beleuchtet werden. Wir folgen dabei in erster Linie den Büchern [Ber07] von Rolf Berndt und [Kra85] von Hanspeter Kraft. Manches der hier präsentierten Theorie über lineare Darstellungen von Gruppen findet sich auch in dem bekannten Buch [Art93] von Michael Artin. Wie zuvor sei K ein nicht-endlicher Körper der Charakteristik 0. Zu Beginn ist es angebracht, den Begriff einer „linearen Darstellung" einzuführen und zu definieren (vgl. [Ber07], S.7 und [DK02], S. 39).

Definition 4.3.1. (Lineare Darstellung)
Sei G eine lineare algebraische Gruppe und V ein K-Vektorraum. Eine **lineare Darstellung** von G in V ist ein Gruppenhomomorphismus $\rho : G \to \mathrm{Aut}_K(V)$. Wir bezeichnen eine lineare Darstellung von G in V kurz mit $(\rho, V)_G$. Die Dimension des K-Vektorraums V wird auch als **Dimension** der linearen Darstellung $(\rho, V)_G$ bezeichnet. Wir schreiben dann $\dim_G(\rho, V)$ für die Dimension von $(\rho, V)_G$.

Um an vielen Stellen die Notation einfacher gestalten zu können, werden wir analog zu den Gruppenoperationen häufig auch ρ_a anstatt $\rho(a)$ schreiben. Mit anderen Worten ist eine Abbildung $\rho : G \to \mathrm{Aut}_K(V)$ genau dann eine lineare Darstellung einer Gruppe $(G, *, e)$ in einem K-Vektorraum V, wenn für alle $a, b \in G$ gilt:

$$\rho(a * b) = \rho(a) \circ \rho(b). \qquad (*)$$

Bevor wir mit den Eigenschaften linearer Darstellungen beginnen, wollen wir ein einfaches Beispiel einer linearen Darstellung betrachten (vgl. [Ber07], S. 10).

Beispiel 4.3.2. Sei $G = \mathrm{SO}_2(\mathbb{R})$. Bekanntlich lässt sich jedes Element von G in der Form

$$\mathcal{A}(\vartheta) := \begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

mit $\vartheta \in \mathbb{R}$ schreiben. Für $k \in \mathbb{Z}$ sei $\rho^{(k)} : G \to \mathrm{GL}_1(\mathbb{C})$ definiert durch $\rho^{(k)}(\mathcal{A}(\vartheta)) = \exp(ik\vartheta)$ und seien $\mathcal{A}(\vartheta_1), \mathcal{A}(\vartheta_2) \in G$. Dann gilt:

$$\begin{aligned} \rho^{(k)}(\mathcal{A}(\vartheta_1) \cdot \mathcal{A}(\vartheta_2)) &= \rho^{(k)}(\mathcal{A}(\vartheta_1 + \vartheta_2)) = \exp(ik(\vartheta_1 + \vartheta_2)) = \exp(ik\vartheta_1) \cdot \exp(ik\vartheta_2) \\ &= \rho^{(k)}(\mathcal{A}(\vartheta_1)) \cdot \rho^{(k)}(\mathcal{A}(\vartheta_2)) \end{aligned}$$

Somit ist $\rho^{(k)}$ für alle $k \in \mathbb{Z}$ eine 1-dimensionale lineare Darstellung von G in $\mathbb{C}$. ◁

Ist V ein endlich-dimensionaler Vektorraum der Dimension n, so erhalten wir bekanntlich durch Wahl einer Basis B von V einen Isomorphismus $\mathcal{M}_B^B : \mathrm{Aut}_K(V) \to \mathrm{GL}_n(K)$, der jedem Automorphismus $f \in \mathrm{Aut}_K(V)$ die eindeutig bestimmte Darstellungsmatrix bzgl. dem Basenpaar B, B zuordnet (vgl. [Fis05], S. 139).

Bemerkung 4.3.3. Damit induziert durch Wahl einer Basis B eine lineare Darstellung ρ von G in V einen Gruppenhomomorphismus $\tilde{\rho} : G \to \mathrm{GL}_n(K)$ durch $\tilde{\rho} = \mathcal{M}_B^B \circ \rho$. Gemäß obiger Definition ist $\tilde{\rho}$ ebenfalls eine lineare Darstellung der Gruppe G, nun in dem K-Vektorraum K^n. Diese lineare Darstellung wird als **Matrix-Darstellung** bezeichnet (vgl. [Ber07], S. 8). Gleichung $(*)$ lässt sich mit Hilfe der Darstellungsmatrizen auch wie folgt ausdrücken:

$$\mathcal{M}_B^B(\rho_{a*b}) = \mathcal{M}_B^B(\rho_a) \cdot \mathcal{M}_B^B(\rho_b).$$

Wir wollen nun als weitere Beispiele spezielle lineare Darstellungen betrachten (vgl. [Ber07], S. 8).

Beispiel 4.3.4. (Besondere Darstellungen)
Sei G eine lineare algebraische Gruppe und V ein K-Vektorraum.

a) Der durch $a \mapsto \mathrm{id}_V$ definierte Gruppenhomomorphismus $G \to \mathrm{Aut}_K(V)$ ist eine lineare Darstellung von G in V. Sie heißt die **triviale Darstellung** von G in V und wird mit $(\mathrm{id}, V)_G$ bezeichnet.

b) Ist G eine Untergruppe von $\mathrm{GL}_n(K)$ und $V = K^n$, so ist durch $\mathcal{A} \mapsto \mathcal{A}$ eine lineare Darstellung $G \to \mathrm{GL}_n(K)$ von G in V definiert. Sie heißt die **natürliche Darstellung** von G in K^n und wird mit $(\mathrm{nat}, K^n)_G$ bezeichnet. ◁

Nun wollen wir als weiteres Beispiel eine spezielle Darstellung näher betrachten, die später eine zentrale Rolle einnehmen wird (vgl. [Ber07] S. 12 sowie [Kra85], S. 72). Dazu verwenden wir die Notation des folgenden Lemmas.

Lemma 4.3.5. *Sei X eine affine Varietät, auf der eine lineare algebraische Gruppe $(G, *, e)$ durch Linksoperation operiert, und sei $V = K[X]$. Dann ist für alle $f \in V$ und alle $a \in G$ die Funktion $x \mapsto f(a^{-1}(x))$ ein Element von V. Diese Funktion bezeichnen wir mit f^a. Zudem gilt für alle $a, b \in G$ und alle $f, f_1, \ldots, f_n \in V$:*

(i) $f^{a*b} = (f^b)^a$,

(ii) $(f_1 + \ldots + f_n)^a = f_1^a + \ldots + f_n^a$,

(iii) $(f_1 \cdots f_n)^a = f_1^a \cdots f_n^a$.

Beweis: Zunächst sei daran erinnert, dass $(a * b)(x) = a(b(x))$ für alle $a, b \in G$ und $x \in X$ gilt (siehe Definition 4.2.1). Außerdem gilt $(a*b)^{-1} = b^{-1} * a^{-1}$ für alle $a, b \in G$. Seien nun $a, b \in G$ und $f, f_1, \ldots, f_n \in V$.

(i) Für alle $x \in X$ gilt:

$$\begin{aligned} f^{a*b}(x) &= f((a*b)^{-1}(x)) = f((b^{-1} * a^{-1})(x)) = f(b^{-1}(a^{-1}(x))) = f^b(a^{-1}(x)) \\ &= (f^b)^a(x), \end{aligned}$$

also $f^{a*b} = (f^b)^a$.

(ii) Für alle $x \in X$ gilt:

$$\begin{aligned} (f_1 + \ldots + f_n)^a(x) &= (f_1 + \ldots + f_n)(a^{-1}(x)) = f_1(a^{-1}(x)) + \ldots + f_n(a^{-1}(x)) \\ &= f_1^a(x) + \ldots + f_n^a(x), \end{aligned}$$

also $(f_1 + \ldots + f_n)^a = f_1^a + \ldots + f_n^a$.

(iii) Für alle $x \in X$ gilt:

$$(f_1 \cdots f_n)^a(x) = (f_1 \cdots f_n)(a^{-1}(x)) = f_1(a^{-1}(x)) \cdots f_n(a^{-1}(x)) = f_1^a(x) \cdots f_n^a(x),$$

also $(f_1 \cdots f_n)^a = f_1^a \cdots f_n^a$.

□

Nun zu dem angekündigten Beispiel, der sogenannten **regulären Darstellung** von G in $K[X]$ (vgl. [Kra85], 2.4, S. 72).

Definition 4.3.6. (Reguläre Darstellung)
Sei G eine lineare algebraische Gruppe, X eine G-Varietät und $V = K[X]$. Die durch $a \mapsto (f \mapsto f^a)$ definierte Abbildung $\rho : G \to \mathrm{Aut}_K(V)$ ist eine lineare Darstellung von G in V. Sie heißt die **reguläre Darstellung** von G in $K[X]$.

Wir wollen nun kurz betrachten, wie sich aus linearen Darstellungen einer Gruppe neue lineare Darstellungen konstruieren lassen (vgl. [Ber07], S. 14 ff. und [Kra85], S. 67).

Bemerkung 4.3.7. (Konstruktion linearer Darstellungen)
Sei G eine Gruppe, seien V sowie V' zwei K-Vektorräume und seien $\rho : G \to \mathrm{Aut}_K(V)$ sowie $\rho' : G \to \mathrm{Aut}_K(V')$ lineare Darstellungen von G in V bzw. V'.

a) Sei $\rho \oplus \rho' : G \to \mathrm{Aut}_K(V \oplus V')$ definiert durch $(\rho \oplus \rho')_a(v \oplus v') = \rho_a(v) \oplus \rho'_a(v')$ für alle $v \oplus v' \in V \oplus V' = \{(v, v') : \ v \in V \text{ und } v' \in V'\}$. Dann ist $\rho \oplus \rho'$ eine lineare Darstellung von G in $V \oplus V'$. Sie heißt die **direkte Summe** von ρ und ρ'.

   Betrachten wir speziell den endlichen Fall, also $\dim_K(V) = n$ und $\dim_K(V') = m$. Sei B eine Basis von V, B' eine Basis von V' und $C := B \oplus B'$ eine Basis von $V \oplus V'$. Dann gilt für alle $a \in G$:

$$\mathcal{M}_C^C((\rho \oplus \rho')_a) = \begin{pmatrix} \mathcal{M}_B^B(\rho_a) & 0 \\ 0 & \mathcal{M}_{B'}^{B'}(\rho'_a) \end{pmatrix} \in \mathrm{GL}_{n+m}(K).$$

b) Wie üblich wird mit V^* der Dualraum von V bezeichnet, also die Menge aller Linearformen von V nach K. Sei $\rho^* : G \to \mathrm{Aut}_K(V^*)$ definiert durch $a \mapsto \left(\rho_a^{-1}\right)^*$, wobei $\left(\rho_a^{-1}\right)^* : V^* \to V^*$ die durch $f \mapsto f \circ \rho_a^{-1}$ definierte, zu $\rho_a^{-1} : V \to V$ duale Abbildung sei (vgl. [Fis05], S. 334). Folglich gilt also für alle $f \in V^*$ und $v \in V$:

$$(\rho_a^*(f))(v) = \left(\rho_a^{-1}\right)^*(f)(v) = f(\rho_a^{-1}(v)).$$

   Dann ist ρ^* eine lineare Darstellung von G in V^*. Sie heißt die **kontragradiente Darstellung** von ρ. Sei V endlich dimensional mit $\dim_K(V) = n$, sei B eine Basis von V und B^* die zugehörige Dualbasis von V^*. Dann gilt für alle $a \in G$ (vgl. [Fis05], S. 334):

$$\mathcal{M}_{B^*}^{B^*}(\rho_a^*) = \left(\mathcal{M}_B^B\left(\rho_a^{-1}\right)\right)^{\mathrm{tr}}.$$

   Auf die kontragradiente Darstellung werden wir später noch zurückkommen. ◁

Wie wir aus Bemerkung 4.2.3 wissen, induziert jeder Gruppenhomomorphismus eine Gruppenoperation, d.h. auch jede lineare Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ induziert eine Operation von G auf V. Man spricht daher in diesem Zusammenhang auch davon, dass G auf V **linear operiert** (vgl. [Kra85], S. 66). Umgekehrt folgt aus derselben Bemerkung, dass jede Gruppenoperation $\tau : G \times V \to V$ auf einem K-Vektorraum V via $a \mapsto \tau_a$ auch eine lineare Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ induziert. Ein K-Vektorraum, auf dem eine Gruppe G linear operiert, wird auch als G-**Modul** bezeichnet (vgl. [Kra85], S. 66). Analog zu linearen Abbildungen von K-Vektorräumen lassen sich auch Abbildungen zwischen G-Moduln betrachten, die dann auch entsprechend als Homomorphismen bezeichnet werden (vgl. [Kra85], S. 69).

Definition 4.3.8. (G-Homomorphismus)
Sei G eine Gruppe, seien V, V' zwei K-Vektorräume und seien $(\rho, V)_G$ sowie $(\rho', V')_G$ lineare Darstellungen von G in V bzw. V'.

a) Eine K-lineare Abbildung $\varphi : V \to V'$ heißt ein G-**Homomorphismus**, falls für alle $a \in G$ gilt:

$$\varphi \circ \rho_a = \rho'_a \circ \varphi.$$

   Die Menge aller G-Homomorphismen auf V und V' wird mit $\mathrm{Hom}_G(V, V')$ bezeichnet. Im Falle von $V = V'$ heißt ein G-Homomorphismus φ auch ein G-**Endomorphismus**. Die Menge aller G-Endomorphismen wird entsprechend mit $\mathrm{End}_G(V)$ bezeichnet.

b) Ein bijektiver G-Homomorphismus $\varphi : V \to V'$ heißt ein G-**Isomorphismus** und im Falle von $V = V'$ ein G-**Automorphismus**. Die Menge aller G-Automorphismen wird mit $\mathrm{Aut}_G(V)$ bezeichnet.

Eine K-lineare Abbildung $\varphi : V \to V'$ ist also genau dann ein G-Homomorphismus, wenn φ mit der durch ρ gegebenen Operation verträglich ist. Da in gewisser Weise ein G-Homomorphismus also die beiden linearen Darstellungen ρ und ρ' „verflechtet", werden sie auch **Verflechtungsoperatoren** genannt (vgl. [Ber07], S. 9). Die Mengen $\mathrm{Hom}_G(V, V')$, $\mathrm{End}_G(V)$ und $\mathrm{Aut}_G(V)$ sind jeweils K-Vektorräume. Offensichtlich bilden die G-Homomorphismen eine Teilmenge der Menge aller K-Homomorphismen $\mathrm{Hom}_K(V, V')$. Die Dimension $\dim_K(\mathrm{Hom}_G(V, V'))$ des K-Vektorraums $\mathrm{Hom}_G(V, V')$ wird auch als **Multiplizität** bezeichnet (vgl. [Ber07], S. 9). Im Falle von $\dim_K(\mathrm{Hom}_G(V, V')) = 0$ spricht man von **disjunkten** Darstellungen, da es in diesem Fall keinen echten G-Homomorphismus gibt, der die beiden Darstellungen „verflechtet". Betrachten wir G-Homomorphismen genauer, so sieht man, dass sie uns bereits wohl bekannt sind: Ein G-Homomorphismus auf V, V' ist nichts anderes als eine lineare Abbildung, die zusätzlich G-äquivariant ist (siehe Definition 4.2.11 und vgl. [Kra85], S. 69). Mit Hilfe von G-Homomorphismen lässt sich die Äquivalenz von zwei Darstellungen im folgenden Sinne definieren (vgl. [Kra85], S. 66).

Definition 4.3.9. (Äquivalente Darstellungen)
Sei G eine Gruppe und seien V, V' zwei K-Vektorräume. Zwei lineare Darstellungen $(\rho, V)_G$ und $(\rho', V')_G$ heißen **äquivalent**, wenn es einen G-Isomorphismus $\varphi : V \to V'$ gibt. Die G-Moduln V und V' werden dann auch als G**-isomorph** bezeichnet.

Im endlichen Fall bedeutet das also, dass zwei lineare Darstellungen $(\rho, V)_G$ und $(\rho', V')_G$ genau dann äquivalent sind, wenn es eine invertierbare Matrix $\mathcal{T} \in \mathrm{GL}_n(K)$ gibt mit

$$\mathcal{T} \cdot \mathcal{M}(\rho_a) = \mathcal{M}(\rho'_a) \cdot \mathcal{T} \quad \text{bzw.} \quad \mathcal{M}(\rho_a) = \mathcal{T}^{-1} \cdot \mathcal{M}(\rho'_a) \cdot \mathcal{T}$$

für alle $a \in G$, wobei n die Dimension von V bzw. V' ist. Mit anderen Worten, ρ und ρ' sind genau dann äquivalent, wenn die beiden Darstellungsmatrizen $\mathcal{M}(\rho_a)$ und $\mathcal{M}(\rho'_a)$ für alle $a \in G$ ähnlich sind.

Den Begriff einer G-stabilen Teilmenge haben wir bereits kennengelernt (siehe Definition 4.2.10), d.h. eine Teilmenge, die durch Operationen der Gruppe G nicht „verlassen" wird. Eine lineare Darstellung $(\rho, V)_G$ induziert bekanntlich eine lineare Operation der Gruppe auf V. Die G-stabilen Untervektorräume von V werden in diesem Zusammenhang in der Literatur jedoch meist als ρ-invariant bezeichnet, was wir hier übernehmen wollen (vgl. [Ber07], S. 8).

Definition 4.3.10. (ρ-invariant)
Sei G eine Gruppe, V ein K-Vektorraum und $(\rho, V)_G$ eine lineare Darstellung von G in V. Ein K-Untervektorraum U von V heißt ρ**-invariant**, wenn U für alle $a \in G$ ein ρ_a-invarianter Untervektorraum von V ist, d.h. wenn $\rho_a(U) \subseteq U$ für alle $a \in G$ gilt.

Kern und Bild eines G-Homomorphismus sind jeweils invariante Untervektorräume bzgl. der jeweiligen Darstellung (vgl. [Art93], 9.5, S. 372). Ist $(\rho, V)_G$ eine lineare Darstellung von G in V und U ein ρ-invarianter K-Untervektorraum von V, so gilt $\rho_a|_U \in \mathrm{Aut}_K(U)$ für alle $a \in G$. Somit erhalten wir aus ρ einen Gruppenhomomorphismus $\overline{\rho} : G \to \mathrm{Aut}_K(U)$ mit $\overline{\rho}_a = \rho_a|_U$. Mit anderen Worten, wir erhalten eine lineare Darstellung von G in U. Diese Darstellung wird als lineare Unterdarstellung von $(\rho, V)_G$ bezeichnet.

Bemerkung 4.3.11. Sei $(\rho, V)_G$ eine lineare Darstellung von G in V und sei U ein ρ-invarianter Untervektorraum von V. Dann bezeichnen wir den durch $a \mapsto \rho_a|_U$ definierten Gruppenhomomorphismus $\overline{\rho} : G \to \mathrm{Aut}_K(U)$ auch kurz mit $\rho|_U$.

Mit der abkürzenden Schreibweise der letzten Bemerkung können wir eine lineare Unterdarstellung wie folgt definieren (vgl. [Ber07], S. 8).

Definition 4.3.12. (Lineare Unterdarstellung)
Sei G eine Gruppe, V ein K-Vektorraum und $(\rho, V)_G$ eine lineare Darstellung von G in V. Eine lineare Darstellung $(\overline{\rho}, U)_G$ heißt eine **lineare Unterdarstellung** von $(\rho, V)_G$, wenn U ein ρ-invarianter Untervektorraum von V ist und $\overline{\rho} : G \to \mathrm{Aut}_K(U)$ definiert ist durch $\overline{\rho} = \rho|_U$.

Mit diesen Begriffen können wir nun die Irreduzibilität linearer Darstellungen definieren (vgl. [Kra85], S. 68).

Definition 4.3.13. (Irreduzible Darstellung)
Sei G eine Gruppe und V ein K-Vektorraum.

a) Eine lineare Darstellung $(\rho, V)_G$ heißt **irreduzibel**, falls $\{0\}$ und $V \neq \{0\}$ die einzigen ρ-invarianten Untervektorräume von V sind.

b) Eine lineare Darstellung $(\rho, V)_G$ heißt **vollständig reduzibel**, wenn es irreduzible Unterdarstellungen $(\rho^{(1)}, U_1)_G, \ldots, (\rho^{(r)}, U_r)_G$ von $(\rho, V)_G$ gibt mit $V = U_1 \oplus \ldots \oplus U_r$. Wir schreiben dann kurz $(\rho, V)_G = (\rho^{(1)}, U_1)_G \oplus \ldots \oplus (\rho^{(r)}, U_r)_G$.

Eine irreduzible Darstellung $(\rho, V)_G$ besitzt also keine echten und nicht-trivialen Unterdarstellungen. Um dies in die Sprache der G-Moduln zu übersetzen, nennt man einen irreduziblen G-Modul **einfach** und einen vollständig reduziblen G-Modul **halbeinfach** (vgl. [Kra85], S. 68). Für eine reduzible Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ gibt es also einen echten ρ-invarianten Untervektorraum U von V. Das hat aber natürlich nicht zwangsläufig zur Folge, dass es einen komplementären ρ-invarianten Untervektorraum W von V geben muss. Wäre dies der Fall, so wäre ρ vollständig reduzibel. Eine vollständig reduzible Darstellung ρ ist also die direkte Summe ihrer linearen Unterdarstellungen, wie uns der folgende Satz bestätigt (vgl. [Kra85], S. 69).

Satz 4.3.14. (Charakterisierung vollständig reduzibler Darstellungen)
Sei G eine Gruppe, V ein endlich-dimensionaler K-Vektorraum und $(\rho, V)_G$ eine lineare Darstellung. Dann sind folgende Aussagen äquivalent:

(i) $(\rho, V)_G$ *ist vollständig reduzibel.*

(ii) Es gibt irreduzible lineare Unterdarstellungen $(\rho^{(1)}, U_1)_G, \ldots, (\rho^{(r)}, U_r)_G$ *von* $(\rho, V)_G$ *mit* $V = U_1 \oplus \ldots \oplus U_r$.

(iii) Zu jeder linearen Unterdarstellung $(\rho^{(1)}, U_1)_G$ *von* $(\rho, V)_G$ *gibt es eine lineare Unterdarstellung* $(\rho^{(2)}, U_2)$ *von* $(\rho, V)_G$ *mit* $V = U_1 \oplus U_2$.

Eine lineare Darstellung $(\rho, V)_G$ von G in V, für die es eine lineare Unterdarstellung $(\rho^{(1)}, U_1)_G$ von V gibt, die eine lineare komplementäre Unterdarstellung $(\rho^{(2)}, U_2)_G$ impliziert, d.h. mit $V = U_1 \oplus U_2$, heißt auch **zerlegbar** (vgl. [Ber07], S. 16). Eine zerlegbare Darstellung ist im Allgemeinen jedoch nicht vollständig reduzibel. Wie der Satz es aussagt, ist dies erst dann der Fall, wenn *jede* lineare Unterdarstellung $(\rho', U)_G$ eine komplementäre lineare Unterdarstellung nach sich zieht.

Beispiel 4.3.15.

a) Für alle $n \in \mathbb{N}_+$ sind die natürlichen Darstellungen von $\mathrm{GL}_n(K), \mathrm{SL}_n(K)$ und $\mathrm{O}_n(K)$ in K^n irreduzibel, für $n > 2$ auch die lineare Darstellung von $\mathrm{SO}_n(K)$ in K^n (vgl. [Kra85], S. 59).

b) Sei $G = \mathrm{Add}(K)$, $\rho : G \to \mathrm{Aut}_K(K^2)$ definiert durch $\rho(x) = \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ und sei $U_1 := \langle (1,0) \rangle_K$. Dann ist U_1 ein ρ-invarianter Untervektorraum von K^2, d.h. $(\rho, K^2)_G$ ist reduzibel, aber es gibt keinen ρ-invarianten Untervektorraum U_2 mit $K^2 = U_1 \oplus U_2$.

Angenommen, es gibt einen derartigen Untervektorraum. Dann ist U_2 aus Dimensionsgründen von der Form $U_2 = \langle (u,v) \rangle_K$ mit $(u,v) \in K^2$ und es gilt

$$\rho(x) \cdot \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + xv \\ v \end{pmatrix} = \lambda(x) \cdot \begin{pmatrix} u \\ v \end{pmatrix}$$

mit $\lambda(x) \in K$ für alle $x \in K$. Insbesondere gilt also $u + xv = \lambda(x)u$ und $v = \lambda(x)v$. Da U_2 komplementär zu U_1 ist, gilt $v \neq 0$. Wegen $v \neq 0$ folgt $\lambda(x) = 1$ und damit $xv = 0$ für alle $x \in K$, was aber nur für $v = 0$ möglich ist und damit einen Widerspruch bedeutet.

Aus dem letzten Satz folgt nun, dass $(\rho, K^2)_G$ nicht vollständig reduzibel ist. Es lässt sich sogar weiter zeigen, dass $(\rho, K^2)_G$ auch nicht zerlegbar ist (vgl. [Ber07], S. 17). ◁

Man sieht an Teil c) des Beispiels, dass „reduzibel" natürlich nicht „vollständig reduzibel" impliziert, aber es gilt die Umkehrung. In manchen Fällen lässt es sich sofort aussagen, ob eine lineare Darstellung vollständig reduzibel ist. Ein berühmter Satz über lineare Darstellungen endlicher Gruppen, der sich daraus folgern lässt, ist der nun folgende von Heinrich Maschke (1853–1908), den der deutsche Mathematiker 1899 bewiesen hatte (vgl. [Kra85], S. 68 oder [Art93], S. 361).

Satz 4.3.16. (Maschke)
Sei G eine endliche Gruppe und V ein K-Vektorraum. Dann ist jede lineare Darstellung $(\rho, V)_G$ vollständig reduzibel.

Eine interessante Aussage über irreduzible Darstellungen ergibt sich aus dem *Lemma von* Schur, das nach dem in Weißrussland geborenen Mathematiker Issai Schur (1875–1941) benannt ist und auf seine Arbeit zurückgeht. Bevor wir auf die Bedeutung des Lemmas eingehen, wollen wir dieses zunächst angeben (vgl. [Art93], S. 373).

Satz 4.3.17. (Lemma von Schur)
Sei G eine Gruppe, seien V, V' zwei K-Vektorräume, seien $(\rho, V)_G$ sowie $(\rho', V')_G$ zwei irreduzible lineare Darstellungen von G in V bzw. V' und sei $\varphi : V \to V'$ ein G-Homomorphismus.

a) Der G-Homomorphismus φ ist entweder bijektiv oder die Nullabbildung.

b) Gilt $V = V'$ und $\rho = \rho'$, so ist φ eine Homothetie, d.h. es gilt $\varphi = \lambda \cdot \mathrm{id}_V$ mit $\lambda \in K$.

Mit anderen Worten sagt uns Teil a) des Satzes, dass zwei irreduzible lineare Darstellungen einer Gruppe entweder äquivalent, falls φ bijektiv ist, oder disjunkt sind, falls $\varphi = 0$ gilt. Der zweite Teil lässt sich äquivalent so ausdrücken, dass die Menge $\mathrm{End}_G(V)$ der G-Endomorphismen isomorph ist zu K (vgl. [Kra85], S. 70).

Wir wollen nun zurück kommen zur regulären Darstellung (siehe Definition 4.3.6) und uns mit dieser noch etwas genauer auseinandersetzen, und zwar besonders im Falle einer Operation einer linearen algebraischen Gruppe G auf einer affinen Varietät X. Die reguläre Darstellung $\rho : G \to \mathrm{Aut}_K(K[X])$ ist bekanntlich (siehe Definition 4.3.6) definiert durch $a \mapsto (f \mapsto f^a)$ wobei $f^a \in K[X]$ gegeben ist durch $f^a(x) = f(a^{-1}(x))$. Nun führen wir Begriffe ein, die auch im nächsten Abschnitt sowie in späteren Kapiteln von großer Bedeutung sein werden (vgl. [Kra85], S. 72 unter Verwendung von [Dre04], S. 7 sowie [DK02], S. 238).

Definition 4.3.18. (Rationale und lokal endliche Darstellung)
Sei G eine lineare algebraische Gruppe und V ein K-Vektorraum.

a) Eine lineare Darstellung $(\rho, V)_G$ von G in V heißt **lokal endlich**, wenn für alle $v \in V$ der K-Untervektorraum $\langle G(v)\rangle_K$ von V endlich-dimensional ist.

b) Eine lineare Darstellung $(\rho, V)_G$ von G in V heißt **rational** oder eine **rationale Darstellung**, falls G auf jedem endlich-dimensionalen G-stabilen K-Untervektorraum U (als affine K-Varietät betrachtet) von V regulär operiert.

Insbesondere jede lineare Darstellung einer linearen algebraischen Gruppe G, die auf einem *endlich-dimensionalen* K-Vektorraum V regulär operiert, ist somit rational (vgl. [DK02], S. 39). Wie aus dem nächsten Satz hervorgeht, ist jede rationale Darstellung auch lokal endlich (vgl. [DK02], Lemma A.1.8, S. 238).

Satz 4.3.19. *Sei G eine lineare algebraische Gruppe und $(\rho, V)_G$ eine rationale Darstellung von G in einem K-Vektorraum V. Dann ist für jedes $v \in V$ der K-Untervektorraum $U := \langle G(v)\rangle_K$ endlich-dimensional. Die Unterdarstellung $(\rho|_U, U)_G$ ist dann ebenfalls rational.*

Wie uns der folgende Satz zeigt, ist die Darstellung einer linearen algebraischen Gruppe G in $K[X]$ sowohl lokal endlich als auch rational (vgl. [Kra85], S. 72).

Satz 4.3.20. *Sei G eine lineare algebraische Gruppe und X eine affine G-Varietät. Dann ist die reguläre Darstellung $(\rho, K[X])_G$ von G in $K[X]$ rational (und damit auch lokal endlich).*

Die nächste Folgerung erweist sich häufig als äußerst nützlich. Sie erlaubt es, eine Gruppenoperation einer linearen algebraischen Gruppe auf einer K-Varietät zurückzuführen auf eine rationale Darstellung von G in einem endlich-dimensionalen K-Vektorraum (vgl. [DK02], Lemma A.1.9, S. 239, und [Kra85], S. 74).

Korollar 4.3.21. *Sei G eine lineare algebraische Gruppe und sei X eine G-Varietät. Dann gibt es eine rationale Darstellung $(\rho, V)_G$ und eine G-äquivariante abgeschlossene Einbettung von X in einen G-stabilen endlich-dimensionalen K-Untervektorraum von V.*

Wie man an dem Beweis des letzten Korollars sehen kann, ist die rationale Darstellung, von der hier die Rede ist, die reguläre Darstellung von G in $K[X]$. Dann stellt sich natürlich sofort die Frage, welche Unterdarstellungen der regulären Darstellung von G in $K[X]$ überhaupt vorkommen (vgl. [Kra85], S. 73). Dazu werden wir nun einen speziellen Fall der regulären Darstellung betrachten: Die lineare algebraische Gruppe $(G, *, e)$ soll nun auf sich selbst durch Linksmultiplikation operieren. Dann sagt uns der folgende Satz, dass jede irreduzible Darstellung eine Unterdarstellung der regulären Darstellung von G in $K[G]$ ist (vgl. [Kra85], S. 73).

Satz 4.3.22. *Sei G eine lineare algebraische Gruppe und sei $(\rho, V)_G$ eine lineare Darstellung von G in einem K-Vektorraum V derart, dass V^* G-zyklisch ist, d.h. es gibt ein $f \in V^*$ mit $V^* = \langle G(f)\rangle_K$. Dann ist $(\rho, V)_G$ eine Unterdarstellung der regulären Darstellung von G in $K[G]$.*

Eine unmittelbare Folgerung aus diesem Satz wird im nächsten Abschnitt von Bedeutung sein und später eine Charakterisierung sogenannter linear reduktiver Gruppen erlauben (vgl. [Kra85], S. 73).

Korollar 4.3.23. *Sei G eine lineare algebraische Gruppe und sei $K[G]^n := K[G]\times\ldots\times K[G]$ mit $n \in \mathbb{N}_+$. Dann ist jede lineare Darstellung $(\rho, V)_G$ der Dimension $m \leq n$ eine Unterdarstellung der regulären Darstellung von G in $K[G]^n$.*

Wir wollen an dieser Stelle noch folgende Aussage über die Menge der Fixpunkte der Operation von G auf $K[G]$ beweisen, die wir später benötigen werden.

Lemma 4.3.24. *Sei G eine linear algebraische Gruppe. Dann sind $K[G]^G$ und K isomorph.*

Beweis: Sei $e \in G$ das neutrale Element. Wir betrachten die Abbildung $\Psi : K[G]^G \to K$, definiert durch $f \mapsto f(e)$, und zeigen, dass sie bijektiv ist. Seien zunächst $f, g \in K[G]^G$ mit $\Psi(f) = \Psi(g)$, also mit $f(e) = g(e)$. Dann gilt

$$f(a) = f(a * e) = f^{a^{-1}}(e) = f(e) = g(e) = g^{a^{-1}}(e) = g(a * e) = g(a),$$

für alle $a \in G$, also $f = g$. Sei nun $\lambda \in K$ und sei $f \in K[G]$ die konstante Abbildung $a \mapsto \lambda$. Sei $a \in G$. Dann gilt $f^a(b) = f(a^{-1} * b) = \lambda = f(b)$ für alle $b \in G$, d.h. es gilt $f^a = f$ und damit $f \in K[G]^G$. Offensichtlich ist f das Urbild von λ. □

Nun zurück zur Definition einer rationalen Darstellung. Wie uns das nächste Lemma zeigt, lassen sich rationale Darstellungen in endlich-dimensionalen K-Vektorräumen in folgendem Sinne charakterisieren (vgl. [Kra85], 2.3, Lemma 1, S. 66).

Lemma 4.3.25. *Sei G eine linear algebraische Gruppe und V ein n-dimensionaler K-Vektorraum. Eine lineare Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ ist genau dann rational, wenn es für eine (und damit jede) Basis B von V reguläre Funktionen $q_{i,j} \in K[G]$, $1 \leq i, j \leq n$, gibt so, dass für alle $a \in G$ gilt:*

$$\mathcal{M}_B^B(\rho_a) = (q_{i,j}(a))_{1\leq i,j\leq n}.$$

Wir wollen nun eine rationale Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ einer linear algebraischen Gruppe G in einem n-dimensionalen K-Vektorraum V näher betrachten und algorithmisch fassen. Durch Wahl einer beliebigen Basis B von V erhalten wir aus der Darstellung von G in V die Matrix-Darstellung $\tilde{\rho} : G \to \mathrm{GL}_n(K)$ von G in K^n, die definiert ist durch $\tilde{\rho} = \mathcal{M}_B^B \circ \rho$. Laut dem letzten Lemma gibt es reguläre Funktionen $q_{i,j} \in K[G]$ mit $\tilde{\rho}(a) = \mathcal{M}_B^B(\rho_a) = (q_{i,j}(a))_{1\leq i,j\leq n}$ für alle $a \in G$. Wegen $K[G] \cong K[z_1, \ldots, z_\ell]/\mathcal{I}(G)$ für ein $\ell \in \mathbb{N}_+$ (siehe Bemerkung 4.1.12) können wir $q_{i,j} \in K[z_1, \ldots, z_\ell]/\mathcal{I}(G)$ wählen. Auf diese Weise lässt sich die durch ρ induzierte Operation von G auf V allgemein durch eine Matrix $(q_{i,j})_{1\leq i,j\leq n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_\ell]/\mathcal{I}(G))$ beschreiben. Dieser Matrix wollen wir einen Namen geben.

Definition 4.3.26. (Darstellungsmatrix der Gruppenoperation)
Sei $G \hookrightarrow \mathbb{A}_K^\ell$ eine linear algebraische Gruppe, sei $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V und sei B eine Basis von V. Die oben konstruierte Matrix $(q_{i,j})_{1\leq i,j\leq n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l]/\mathcal{I}(G))$ heißt die **Darstellungsmatrix der (durch ρ induzierten) Gruppenoperation bzgl. der Basis** B und wird mit $\mathcal{M}_B^B(\rho)$ bezeichnet.

Die Bezeichnung als „Darstellungsmatrix" ist durchaus gerechtfertigt, stellt sie doch gewissermaßen die Darstellungsmatrix der linearen Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ bzgl. einer Basis B von V dar. Zu der Beschreibung der Gruppe mittels ihrem Verschwindungsideal und der Beschreibung der Gruppenoperation durch ihre Darstellungsmatrix wollen wir nun ein Beispiel angeben.

Beispiel 4.3.27. Wir untersuchen die orthogonale Gruppe $G := \mathrm{O}_2(K)$, für die bekanntlich $K[G] \cong K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$ gilt, wobei laut Beispiel 4.1.13 das Verschwindungsideal $\mathcal{I}(G)$ von $z_{1,1}^2 + z_{1,2}^2 - 1,\ z_{1,1}z_{2,1} + z_{1,2}z_{2,2},\ z_{2,1}^2 + z_{2,2}^2 - 1$ erzeugt wird. Als K-Vektorraum V betrachten wir $(K^2)^3$, die Menge aller (inklusive der degenerierten) Dreiecke in K^2. Wegen $(K^2)^3 \cong \mathrm{Mat}_{3,2}(K)$ können wir jedes Dreieck als Matrix

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{pmatrix} \in \mathrm{Mat}_{3,2}(K)$$

auffassen, wobei die i-te Zeile dem i-ten Punkt des Dreiecks entspricht. Die Gruppe G operiert dann durch Matrixmultiplikation auf V, d.h. die Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ ist definiert durch $\rho_{\mathcal{A}}(\mathcal{D}) = \mathcal{D} \cdot \mathcal{A}$. Als Basis B wählen wir die Matrizen $\mathcal{I}_{i,j}$, $1 \le i \le 3$ und $1 \le j \le 2$, die an der (i,j)-ten Stelle eine 1 haben und sonst nur Nullen. Dann gilt mit $\mathcal{A} = \left(\begin{smallmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & z_{2,2} \end{smallmatrix}\right)$:

$$\begin{aligned}
\rho_{\mathcal{A}}(\mathcal{I}_{1,1}) &= \mathcal{I}_{1,1} \cdot \mathcal{A} = z_{1,1}\mathcal{I}_{1,1} + z_{1,2}\mathcal{I}_{1,2} \\
\rho_{\mathcal{A}}(\mathcal{I}_{1,2}) &= \mathcal{I}_{1,2} \cdot \mathcal{A} = z_{2,1}\mathcal{I}_{1,1} + z_{1,2}\mathcal{I}_{1,2} \\
\rho_{\mathcal{A}}(\mathcal{I}_{2,1}) &= \mathcal{I}_{2,1} \cdot \mathcal{A} = z_{1,1}\mathcal{I}_{2,1} + z_{1,2}\mathcal{I}_{2,2} \\
\rho_{\mathcal{A}}(\mathcal{I}_{2,2}) &= \mathcal{I}_{2,2} \cdot \mathcal{A} = z_{2,1}\mathcal{I}_{2,1} + z_{2,2}\mathcal{I}_{2,2} \\
\rho_{\mathcal{A}}(\mathcal{I}_{3,1}) &= \mathcal{I}_{3,1} \cdot \mathcal{A} = z_{1,1}\mathcal{I}_{3,1} + z_{1,2}\mathcal{I}_{3,2} \\
\rho_{\mathcal{A}}(\mathcal{I}_{3,2}) &= \mathcal{I}_{3,2} \cdot \mathcal{A} = z_{2,1}\mathcal{I}_{3,1} + z_{2,2}\mathcal{I}_{3,2}
\end{aligned}$$

Wir wählen erneut die Termordnung DegRevLex und erhalten damit folgende Darstellungsmatrix in $K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$:

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} z_{1,1} & z_{2,1} & 0 & 0 & 0 & 0 \\ z_{1,2} & z_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & z_{1,1} & z_{2,1} & 0 & 0 \\ 0 & 0 & z_{1,2} & z_{2,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & z_{1,1} & z_{2,1} \\ 0 & 0 & 0 & 0 & z_{1,2} & z_{2,2} \end{pmatrix}.$$

◁

Weitere Beispiele von Darstellungsmatrizen von Gruppenoperationen werden wir in Kapitel 11 sehen. Wir wollen an dieser Stelle zwei Darstellungen einer linear algebraischen Gruppe G in zwei isomorphen n-dimensionale K-Vektorräume V und W betrachten. Seien also $(\rho, V)_G$ und $(\tau, W)_G$ rationale Darstellungen von G in V bzw. W. Ein Zusammenhang zwischen den beiden Darstellungen lässt sich aber nur herstellen, wenn die Vektorräume V und W nicht nur isomorph, sondern G-isomorph sind, also $(\rho, V)_G$ und $(\tau, V)_G$ äquivalente Darstellungen sind (vgl. Definition 4.3.9). Mit anderen Worten, muss es einen Isomorphismus geben, der mit den Gruppenoperationen verträglich ist, also einen G-Isomorphismus. Wir wählen nun eine Basis B von V sowie eine Basis C von W. Dann lässt sich die Eigenschaft, dass $\varphi : V \to W$ eine G-äquivariante Abbildung ist, d.h. dass $\varphi \circ \rho_a = \tau_a \circ \varphi$ für alle $a \in G$ gilt, äquivalent durch die Darstellungsmatrizen ausdrücken. Somit gilt für alle $a \in G$

$$\mathcal{M}_C^B(\varphi) \cdot \mathcal{M}_B^B(\rho_a) = \mathcal{M}_C^C(\tau_a) \cdot \mathcal{M}_C^B(\varphi)$$

Dieser Zusammenhang lässt sich analog auf die Darstellungsmatrizen der Gruppenoperationen übertragen. Dann gilt in $\mathrm{Mat}_n(K[z_1, \ldots, z_\ell]/\mathcal{I}(G))$:

$$\mathcal{M}_C^C(\tau) = \mathcal{M}_C^B(\varphi) \cdot \mathcal{M}_B^B(\rho) \cdot \left(\mathcal{M}_C^B(\varphi)\right)^{-1}.$$

Eine besondere Situation entsteht natürlich dann, wenn die Darstellungen $(\rho, V)_G$ und $(\tau, W)_G$ identisch sind, also wenn $\rho = \tau$ und $V = W$ gilt, sowie φ die Identität ist, also $\varphi = \mathrm{id}_V$ gilt. Dann ist $\mathcal{M}_C^B(\varphi)$ die Basistransformationsmatrix von B nach C und $\left(\mathcal{M}_B^C(\varphi)\right)^{-1}$ gerade die Basistransformationsmatrix von C nach B. In diesem Fall erhalten wir also auch allgemein eine Basistransformationsformel für die Darstellungsmatrizen der Gruppenoperation.

Bemerkung 4.3.28. (Transformationsformel für Darstellungsmatrizen der Gruppenoperation) Sei $(\rho, V)_G$ eine rationale Darstellung von G in V und seien B, C Basen von V. Dann gilt für alle $a \in G$ die aus der linearen Algebra bekannte Transformationsformel (vgl. [Fis05], 2.6.5, S. 158 f.)

$$\mathcal{M}_C^C(\rho_a) = \mathcal{M}_C^B(\mathrm{id}_V) \cdot \mathcal{M}_B^B(\rho_a) \cdot \mathcal{M}_B^C(\mathrm{id}_V)$$

sowie allgemein in $\mathrm{Mat}_n(K[z_1, \ldots, z_\ell]/\mathcal{I}(G))$:

$$\mathcal{M}_C^C(\rho) = \mathcal{M}_C^B(\mathrm{id}_V) \cdot \mathcal{M}_B^B(\rho) \cdot \mathcal{M}_B^C(\mathrm{id}_V).$$

4.4 Lineare reduktive Gruppen

Eine Klasse von Gruppen, genauer eine Teilmenge der linearen algebraischen Gruppen, spielt in der Invariantentheorie eine zentrale Rolle, wie wir ab Kapitel 6 sehen werden. Wir werden in unserer Darstellung in erster Linie dem Buch [Kra85] von Hanspeter Kraft folgen, da diese Definition sich nahtlos an die letzten Abschnitte anfügt. Weitere Resultate über linear reduktive Gruppen finden sich beispielsweise in [DK02] und teilweise auch in [Bor91], [Hum81] sowie [Spr80]. Wir werden zunächst kurz auf Variationen des Begriffs „Reduktivität", wie sie in verschiedenen Büchern zu finden sind, eingehen und deren Zusammenhänge angeben.

Sei wie bisher K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$ und sei G eine linear algebraische Gruppe. Mit $\mathcal{B}(G)$ wird die Menge der maximal zusammenhängenden auflösbaren Untergruppen von G bezeichnet. Dann ist das **Radikal** $\mathrm{Rad}(G)$ von G die Zusammenhangskomponente der Eins des Durchschnitts aller Gruppen in $\mathcal{B}(G)$, d.h. konkret gilt (vgl. [DK02], Definition A.3.5, S. 244)

$$\mathrm{Rad}(G) = \left(\bigcap_{B \in \mathcal{B}(G)} B \right)^0.$$

Eine linear algebraische Gruppe G heißt dann **reduktiv** oder auch **gruppentheoretisch reduktiv**, wenn $\mathrm{Rad}(G)$ ein Torus ist (vgl. [DK02], Definition A.3.6, S. 244). Ist $\mathrm{Rad}(G)$ die triviale Gruppe, so wird G als **halbeinfach** bezeichnet. Typische Beispiele reduktiver Gruppen sind $\mathrm{GL}_n(K)$, $\mathrm{SL}_n(K)$, $\mathrm{O}_n(K)$, $\mathrm{SO}_n(K)$, endliche Gruppen, Tori und halbeinfache Gruppen (vgl. [DK02], S. 50). Häufig taucht in diesem Zusammenhang auch der Begriff „geometrisch reduktiv" auf. Eine linear algebraische Gruppe G heißt dabei **geometrisch reduktiv**, wenn es für jede rationale Darstellung $(\rho, V)_G$ und jeden Fixpunkt $v \in V^G$ mit $v \neq 0$ ein homogenes Polynom $f \in K[V]^G$ gibt mit $\deg(f) > 0$ und $f(v) \neq 0$, wobei die Operation von G auf $K[V]$ durch die reguläre Darstellung von G in $K[V]$ induziert wird (vgl. [DK02], Definition 2.2.14, S. 50). In [NM63] haben Masayoshi Nagata und Takehiko Miyata bewiesen, dass jede geometrisch reduktive Gruppe auch gruppentheoretisch reduktiv ist. Dass die Umkehrung ebenfalls korrekt ist, wurde von William J. Haboush 1975 bewiesen (vgl. [Hab75]). Der für uns entscheide Begriff von „Reduktivität" ist der einer *linear* reduktiven Gruppe, die wir wie in [Kra85], S. 89, definieren wollen.

Definition 4.4.1. (Linear reduktive Gruppe)
Eine lineare algebraische Gruppe G heißt **linear reduktiv**, wenn jede rationale Darstellung $(\rho, V)_G$ von G in einem K-Vektorraum V vollständig reduzibel ist.

In der Sprache der G-Moduln übersetzt, ist eine linear algebraische Gruppe G linear reduktive, wenn jeder G-Modul V halbeinfach ist. Um diesen Begriff der „Reduktivität" im Zusammenhang mit den beiden oben genannten einordnen zu können, betrachten wir das folgende Theorem, das ebenfalls von Nagata und Miyata bewiesen wurde (vgl. [NM63]).

Theorem 4.4.2. *Ist* $\mathrm{char}(K) = 0$*, so ist eine linear algebraische Gruppe G genau dann linear reduktiv, wenn sie (gruppentheoretisch) reduktiv, und damit auch geometrisch reduktiv, ist.*

In Charakteristik 0 sind somit alle drei Variationen von Reduktivität äquivalent. In positiver Charakteristik sind (gruppentheoretisch) reduktiv und geometrisch reduktiv nach wie vor äquivalente Begriffe, linear reduktiv ist jedoch ein stärkerer Begriff. Gilt $\mathrm{char}(K) = p > 0$, so ist eine linear algebraische Gruppe G nur dann linear reduktiv, wenn G^0 ein Torus ist und $\#(G/G^0)$ nicht durch p teilbar ist (vgl. [Nag61]). Somit sind in positiver Charakteristik neben Tori unter anderem nur diejenigen endlichen Gruppen linear reduktiv, deren Ordnung nicht durch die Charakteristik geteilt wird. Der Fall $\mathrm{char}(K) = 0$ ist genau der für die späteren Anwendungen interessante Fall. Da, wie bereits erwähnt, in diesem Fall alle drei Variationen äquivalent sind, werden wir hier auf deren Unterscheidung auch nicht näher eingehen und im Folgenden nur über linear reduktive Gruppen sprechen. Dazu gelte nun stets wieder wie bisher $\mathrm{char}(K) = 0$. Beispiele von linear reduktiven Gruppen sind uns laut der Definition bereits bekannt: Gemäß dem Satz von Maschke (siehe Satz 4.3.16) sind alle endlichen Gruppen linear reduktiv. Darüber hinaus sind viele weitere bekannte, nicht-endliche und interessante Gruppen linear reduktiv.

Beispiel 4.4.3.

a) In Beispiel 4.3.15 haben wir eine rationale Darstellung der additiven Gruppe $\mathrm{Add}(K)$ in K^2 kennengelernt, die nicht vollständig reduzibel ist. Somit ist $\mathrm{Add}(K)$ nicht linear reduktiv.

b) Die Gruppen $\mathrm{GL}_n(K)$, $\mathrm{SL}_n(K)$, $\mathrm{O}_n(K)$ und $\mathrm{SO}_n(K)$ sowie alle endlichen Gruppen sind linear reduktiv (vgl. [Kra85], S. 89). ◁

Wie wir später noch sehen werden, zeichnet linear reduktive Gruppen eine wunderbare Eigenschaft aus, die deren Bedeutung für die Invariantentheorie erklärt: Die Menge der Invarianten unter der Operation einer linear reduktiven Gruppe ist stets endlich erzeugt. Dieser berühmte Satz wurde von David Hilbert im Jahre 1890 veröffentlicht und wird uns später beschäftigen. Direkte Produkte linear reduktiver Gruppen sind ebenfalls linear reduktiv, was für die linear reduktiven Gruppen aus Beispiel 4.4.3 in [Kra85], S. 288, und im Allgemeinen in [Kra11], S. 76, erwähnt ist.

Bemerkung 4.4.4. Seien G und H linear algebraische Gruppen. Das Produkt $G \times H$ ist genau dann linear reduktiv, wenn G und H linear reduktiv sind.

Somit folgt mit Beispiel 4.4.3 induktiv aus dieser Bemerkung, dass die additive Gruppe $(K^n, +)$ mit der Vektoraddition als Verknüpfung nicht linear reduktiv ist. Wir wollen nun verschiedene Charakterisierungen linear reduktiver Gruppen angeben. Das folgende Theorem beinhaltet unter anderem als eine Charakterisierung (Punkt (ii) des Theorems) die Definition einer linear reduktiven Gruppe, wie sie Derksen und Kemper in [DK02] verwenden (vgl. [DK02], Definition 2.2.1 und Theorem 2.2.5, S. 45 f.).

Theorem 4.4.5. (Erste Charakterisierungen linear reduktiver Gruppen)
Sei G eine lineare algebraische Gruppe. Dann sind die folgenden Aussagen äquivalent:

(i) *G ist linear reduktiv.*

(ii) *Für jede rationale Darstellung $(\rho, V)_G$ und jeden Fixpunkt $v \in V^G$ mit $v \neq 0$ gibt es eine invariante Linearform $f \in (V^*)^G$ mit $f(v) \neq 0$.*

(iii) *Für jede rationale Darstellung $(\rho, V)_G$ gibt es eine eindeutig bestimmte Unterdarstellung $(\rho', U)_G$ mit $V = V^G \oplus U$. In diesem Fall gilt $(U^*)^G = \{0\}$.*

Eine weitere äquivalente Charakterisierung verwendet den sogenannten Reynolds-Operator, den wir in Abschnitt 7.1 studieren werden. Um weitere Charakterisierungen sowie erste Eigenschaften linear reduktiver Gruppen zu erhalten, betrachten wir nun die Operation durch Linksmultiplikation einer linear algebraischen Gruppe $(G, *, e)$ auf sich selbst, d.h. die Operation $G \times G \to G$ ist gegeben durch $(a, x) \mapsto a * x$. Sei $(\rho, K[G])_G$ die reguläre Darstellung von G in $K[G]$, d.h. für alle $a \in G$ ist $\rho_a \in \mathrm{Aut}_K(K[G])$ definiert durch $\rho_a(f) = f^a$, wobei $f^a : G \to K$ gegeben ist durch $f^a(x) = f(a^{-1}(x))$. Damit erhalten wir folgende Charakterisierungen (vgl. [Kra85], 3.5, Satz 1, S. 107, und Lemma 1, S. 108).

Theorem 4.4.6. *Sei G eine lineare algebraische Gruppe. Dann sind folgende Aussagen äquivalent:*

(i) *G ist linear reduktiv.*

(ii) *Die reguläre Darstellung von G in $K[G]$ ist vollständig reduzibel.*

(iii) *Für jeden surjektiven G-Homomorphismus $\varphi : V \to W$ rationaler Darstellungen $(\rho, V)_G$ und $(\rho', W)_G$ gilt $\varphi(V^G) = W^G$.*

Die Definition einer linear reduktiven Gruppe machte die Entscheidung, ob eine lineare algebraische Gruppe nun linear reduktiv ist oder nicht, doch recht schwer, da jede rationale Darstellung auf vollständige Reduzibilität zu untersuchen war. Mit diesem Theorem genügt es, alleine die reguläre Darstellung der Gruppe in ihrem Koordinantenring zu betrachten. Der nächste Satz beinhaltet nützliche weitere Kriterien zur Untersuchung der linearen Reduktivität von linearen algebraischen Gruppen (vgl. [Kra85], S. 108).

Satz 4.4.7.

a) *Jeder Normalteiler einer linear reduktiven Gruppe ist linear reduktiv.*

b) *Ist G eine linear reduktive Gruppe und $\varphi : G \to H$ ein Gruppenhomomorphismus mit einer weiteren Gruppe H, so ist $\varphi(G)$ linear reduktiv.*

c) *Ist H Normalteiler einer linearen algebraischen Gruppe G und sind H sowie G/H linear reduktiv, so ist auch G linear reduktiv.*

Damit ist es uns nun ein Leichtes, für die affine, und damit auch für die euklidische, Gruppe nachzuweisen, dass diese nicht linear reduktiv ist.

Beispiel 4.4.8. Die affine Gruppe $\mathrm{AGL}_n(K)$ ist nicht linear reduktiv. Denn angenommen, $\mathrm{AGL}_n(K)$ wäre linear reduktiv. Dann wäre laut dem letzten Satz jeder Normalteiler von $\mathrm{AGL}_n(K)$ ebenfalls linear reduktiv. Die Translationsgruppe $\mathrm{Trans}_n(K)$ ist ein Normalteiler von $\mathrm{AGL}_n(K)$ und isomorph zur additiven Gruppe $(\mathrm{Add}(K))^n$. Diese wiederum ist aber nicht linear reduktiv. ◁

Eine weitere unmittelbare Folgerung ergibt sich aus diesem Satz und zieht die bereits bekannte Zusammenhangskomponente des neutralen Elements heran (vgl. [Kra85], S. 109).

Korollar 4.4.9. *Eine lineare algebraische Gruppe G ist genau dann linear reduktiv, wenn G^0 linear reduktiv ist.*

Teil II

Invariantentheorie

KAPITEL 5

SAGBI-Basen

Johann Wolfgang von Goethe[13]

Jedes Existierende ist ein Analogon alles Existierenden; daher erscheint uns das Dasein immer zu gleicher Zeit gesondert und verknüpft.

Nachdem sich Ende der 1980er Jahre sich die Gröbner-Basen Theorie längst etabliert hatte, begannen sich verschiedene Mathematiker die Frage zu stellen, ob sich auch für Unteralgebren von Polynomringen eine analoge Theorie entwickeln lässt. In etwa gleichzeitig beschäftigten sich unabhängig voneinander Deepak Kapur und Klaus Madlener sowie Lorenzo Robbiano und Moss Sweedler mit dieser Frage. Erste Ergebnisse auf diesem Gebiet wurden von Kapur und Madlener im Jahre 1989 ([KM89]) sowie von Robbiano und Sweedler im Jahre 1990 ([RS90]) veröffentlicht. Letztere führten schließlich die Bezeichnung *SAGBI* ein, die sich mittlerweile in der Mathematik durchgesetzt hat. Das Akronym „SAGBI" steht dabei für

„**S**ubalgebra **A**nalogs to **G**röbner **B**ases for **I**deals",

womit der Name sofort Programm ist. Natürlich stellt sich bereits an dieser Stelle die Frage, wie weit sich diese Analogie zur Welt der Gröbner-Basen treiben lässt? Wir werden in diesem Kapitel unter anderem auch sehen, dass sich weite Teile tatsächlich analog übertragen lassen, aber dass es durchaus auch Unterschiede gibt.

Jede K-Unteralgebra S eines Polynomrings $P = K[x_1, \ldots, x_n]$ über einem Körper K versehen mit einer Termordnung σ besitzt ein **Erzeugendensystem** $G \subseteq P \setminus \{0\}$ als K-Algebra, d.h. es gibt für alle Elemente $f \in S$ endlich viele Polynome $g_1, \ldots, g_s \in G$ und weitere Unbestimmte $y_1, \ldots, y_s$ so, dass $f = h(g_1, \ldots, g_s)$ gilt für ein Polynom $h \in K[y_1, \ldots, y_s]$. Für eine nicht-leere Teilmenge $G \subseteq P \setminus \{0\}$ schreiben wir auch kurz $K[G]$ anstatt $K[g : \ g \in G]$ für die von G erzeugte K-Unteralgebra von P. Zu bedenken ist allerdings, dass das Erzeugendensystem G nicht notwendigerweise endlich sein muss. Besitzt S jedoch sogar ein endliches Erzeugendensystem, so heißt S **endlich erzeugt**. Man beachte, dass aber auch in diesem Fall nicht jedes Algebra-Erzeugendensystem von S endlich sein muss! Beispielsweise erzeugt S natürlich sich

[13] Bildquelle: http://www.literaturwelt.com/autoren/goethe.html vom 21.03.2013.

selbst, ist aber nicht endlich. Wie man an dem einfachen Beispiel $S = K[x, xy^i | \ i \in \mathbb{N}_+]$ als K-Unteralgebra von $P = K[x, y]$ sieht, ist nicht jede K-Unteralgebra endlich erzeugt, denn diese K-Unteralgebra besitzt kein endliches Erzeugendensystem. Wir werden uns hier allerdings ausschließlich mit endlich erzeugten K-Unteralgebren beschäftigen. Unser Interesse gilt dabei ganz speziellen Erzeugendensystemen von K-Unteralgebren, die **σ-SAGBI-Basen** genannt werden. Was Gröbner-Basen für Ideale sind, werden σ-SAGBI-Basen für K-Unteralgebren sein. Wir werden im ersten Abschnitt sehen, dass anders als in der Welt der Gröbner-Basen, in der jedes endlich erzeugte Ideal auch eine endliche Gröbner-Basis besitzt, in der Welt der Unteralgebren nicht jede endlich erzeugte K-Unteralgebra eine endliche SAGBI-Basis besitzt. Insbesondere an dieser Stelle stößt die Analogie also an ihre Grenzen. SAGBI-Basen tauchen seit ihrer Entstehung in verschiedensten Anwendungen auf. Geometrisch interpretiert liefern sie z.B. eine „Deformation" einer parametrisierten Varietät in eine torische Varietät (vgl. [Stu96]). Karin Gatermann hat SAGBI-Basen beispielsweise für dynamische Systeme verwendet (vgl. [Gat00] oder [Gat03]). Ein großes Anwendungsfeld ist allerdings das Gebiet der Invariantentheorie. Besonders zu erwähnen sind hier die Bücher von Harm Derksen und Gregor Kemper (vgl. [DK02]) sowie Bernd Sturmfels (vgl. [Stu08]). Weitere Anwendungen von SAGBI-Basen finden sich beispielsweise in [Göb98], [Göb99], [Göb01], [Nor02], [RS98], [ST99] und [TT02].

In dieser Arbeit wollen wir den Fokus auf die algorithmische Betrachtung von SAGBI-Basen legen, insbesondere werden wir uns dabei der Berechnung von SAGBI-Basen in verschiedenen Situationen widmen. Die hier betrachteten Algorithmen wurden allesamt in der dem Computeralgebrasystem CoCoA (siehe [RAB15]) eigenen Skriptsprache CoCoAL implementiert. Sie sind als Paket `sagbi.cpkg` (siehe Anhang C) Teil des Open-Source-Computeralgebrasystems ApCoCoA (siehe [KAT13]), einer Erweiterung von CoCoA. Der Großteil der in diesem Paket implementierten Funktionalität ist bisher in keinem Computeralgebrasystem verfügbar. Zum derzeitigen Stand sind lediglich im Computeralgebrasystem Singular Funktionen zu SAGBI-Basen zu finden, und das in sehr geringem Umfang (siehe [DGPS15], sagbi.lib). Neben der Umsetzung in einem Software-Paket ist es uns aber auch gelungen, die Theorie der SAGBI-Basen auszubauen. Nachdem wir im ersten Abschnitt den Begriff einer SAGBI-Basis angeben und deren Existenz behandeln, führen wir im zweiten Abschnitt eine Unteralgebra-Version eines Divisionsalgorithmus sowie eine Unteralgebra-Version eines normalen Restes ein. Beides sucht man in der Literatur vergebens und fehlt so beispielsweise auch in [KR05]. Dieser Divisionsalgorithmus bildet den Kern für anschließende Berechnungen von SAGBI-Basen. Im dritten Abschnitt bauen wir auf dem Divisionsalgorithmus auf und gehen zunächst auf den Begriff einer SAGBI-Normalform ein. Dazu greifen wir die in [KR05], Tutorial 96 gemachten Vorschläge auf, beweisen sie und setzen sie algorithmisch um. Dabei konnte auch ein Fehler in [KR05] korrigiert werden sowie weitere Anwendungen der SAGBI-Normalform aufgezeigt werden. Ebenfalls im dritten Abschnitt behandeln wir den Begriff einer reduzierten SAGBI-Basis, der erstmals in [RS90] erwähnt wurde. Wir werden hier ebenfalls einen Algorithmus zur Berechnung reduzierter SAGBI-Basen vorstellen. Der letzte Abschnitt steht ganz im Zeichen von Algorithmen. So werden wir zunächst die SAGBI-Prozedur vorstellen, wie sie in [KR05] zu finden ist, und auf die Implementation dieser Prozedur eingehen. Anschließend betrachten wir den homogenen Fall. Dazu gehen wir auf eine Prozedur zur Berechnung homogener SAGBI-Basen ein, die in [RS90] nur gegen Ende kurz erwähnt wird und in [KR05], Tutorial 96 ohne Beweis aufgegriffen wird. Die Korrektheit werden hier beweisen. Außerdem ergibt sich aus dieser Prozedur ein Algorithmus zur Berechnung von Grad-beschränkten SAGBI-Basen, dessen Korrektheit und Endlichkeit wir ebenfalls beweisen werden. Darüberhinaus zeigen wir weitere Anwendungen und Charakterisierungen von Grad-beschränkten SAGBI-Basen.

5.1 Definition und Existenz von SAGBI-Basen

Sei K ein Körper, $P = K[x_1, \ldots, x_n]$ ein Polynomring über K in den Unbestimmten $x_1, \ldots, x_n$ und sei σ eine Termordnung auf $\mathbb{T}^n$. Sei weiter S stets eine *endlich erzeugte* K-Unteralgebra von P, d.h. es gebe endlich viele Polynome $f_1, \ldots, f_s \in P \setminus \{0\}$ mit $S = K[f_1, \ldots, f_s]$. Ist P' der Polynomring $K[y_1, \ldots, y_s]$ über K in weiteren Unbestimmten $y_1, \ldots, y_s$, so ist S das Bild von P' unter dem K-Algebra-Homomorphismus $\lambda : P' \to P$, definiert durch $\lambda(y_i) = f_i$ für alle $i \in \{1, \ldots, s\}$ (vgl. [KR00], Korollar 1.1.14). Mit $J := \mathrm{Ker}(\lambda)$ ist $\overline{\lambda} : K[y_1, \ldots, y_s]/J \to S$ ein K-Algebra-Isomorphismus. Wir sagen, $\overline{\lambda}$ sei eine **Darstellung** der K-Unteralgebra S durch Erzeuger und Relationen. Das Ideal J ist das Ideal der algebraischen Relationen von $f_1, \ldots, f_s$ (vgl. Definition 2.2.6). Zunächst wollen wir den Begriff einer σ-SAGBI-Basis angeben (vgl. [KR05], Definition 6.6.2, S. 479).

Definition 5.1.1. (σ-SAGBI-Basis)
Eine nicht-leere Menge $B \subseteq S \setminus \{0\}$ heißt eine **σ-SAGBI-Basis** von S, falls gilt:

$$K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}] = K[\mathrm{LT}_\sigma(g) : \ g \in B]$$

Man erkennt an der Definition einer σ-SAGBI-Basis sofort, dass jede K-Unteralgebra eine σ-SAGBI-Basis besitzt, indem man $B = S \setminus \{0\}$ setzt. Natürlich ist diese σ-SAGBI-Basis wegen ihrer Nichtendlichkeit insbesondere für Anwendungen wenig interessant. Deshalb wird das Hauptaugenmerk in diesem Kapitel auf endlichen σ-SAGBI-Basen liegen. Wir wollen dazu den Fragen nachgehen, wann solche existieren und wie man diese ggf. berechnen kann. Zuvor wollen wir aber ein erstes Beispiel einer σ-SAGBI-Basis betrachten, das an der einen oder anderen Stelle im weiteren Verlauf wieder auftauchen wird.

Beispiel 5.1.2. Sei $P = \mathbb{Q}[x_1, x_2, x_3]$ und sei $\sigma = \mathrm{DegLex}$. Sei $G = \{f_1, f_2, f_3\}$ mit den Polynomen $f_1 := x_1^2 - x_2x_3$, $f_2 := x_1x_2 + x_3^2$ sowie $f_3 := x_2^2 - x_3^2$ und sei S die von G erzeugte $\mathbb{Q}$-Unteralgebra von P. Setze $f_4 := f_2^2 - f_1f_3 = x_1^2x_3^2 + 2x_1x_2x_3^2 + x_2^3x_3 - x_2x_3^3 + x_3^4$. Dann gilt $f_4 \in S$ und wie man leicht erkennen kann, ist G keine σ-SAGBI-Basis von S, da $\mathrm{LT}_\sigma(f_4)$ kein Element von $K[\mathrm{LT}_\sigma(g) : \ g \in G]$ ist. Aber $B := G \cup \{f_4\}$ ist eine σ-SAGBI-Basis von S. ◁

Dieses Beispiel bestätigt auch genau das, was man ohnehin vermuten würde: Ein endliches Erzeugendensystem einer K-Unteralgebra muss nicht automatisch eine σ-SAGBI-Basis sein. Darüberhinaus ist aus der Definition der σ-SAGBI-Basis auch nicht sofort ersichtlich, dass eine σ-SAGBI-Basis die K-Unteralgebra S auch stets als K-Algebra erzeugt. Allein schon im wörtlichen Sinne der Bezeichnung „Basis von S" wäre diese Eigenschaft natürlich äußerst wünschenswert. In [KR05], Satz 6.6.3 wird genau das nachgewiesen.

Satz 5.1.3. *Jede σ-SAGBI-Basis von S ist ein K-Algebra-Erzeugendensystem von S.*

Bevor wir uns weiter mit allgemeinen Unteralgebren beschäftigen, wollen wir zunächst einen einfachen Spezialfall betrachten: K-Unteralgebren, die von Termen erzeugt werden. In Definition 5.1.1 tauchten derartige K-Unteralgebren auch bereits auf, die analog zu monomialen Idealen als **monomiale K-Unteralgebren** bezeichnet werden und auch ähnlich interessante Eigenschaften wie monomiale Ideale aufweisen (vgl. [KR05], Satz 6.6.4, S. 480).

Satz 5.1.4. (Eigenschaften monomialer Unteralgebren)
Sei $T \subseteq P$ eine monomiale K-Unteralgebra von P und sei $G \subseteq \mathbb{T}^n$ ein nicht-leeres Algebra-Erzeugendensystem von T.

a) *Sei $f \in T \setminus \{0\}$. Dann gibt es für jeden Term $t \in \mathrm{Supp}(f)$ Terme $t_1, \ldots, t_s \in G$ und Exponenten $a_1, \ldots, a_s \in \mathbb{N}$ mit $t = t_1^{a_1} \cdots t_s^{a_s}$, d.h. insbesondere gilt $t \in T$.*

b) *Sei $\mathfrak{G}$ die Menge aller Algebra-Erzeugendensysteme von S bestehend aus Termen. Dann gibt es in $\mathfrak{G}$ ein eindeutig bestimmtes und bzgl. Inklusion minimales Erzeugendensystem von T als K-Algebra. Dieses Erzeugendensystem wird das **minimale monomiale Algebra-Erzeugendensystem** von T genannt.*

c) *Ist T eine endlich erzeugte K-Algebra, so ist auch das minimale monomiale Algebra-Erzeugendensystem von T endlich.*

Aus dem Satz folgt natürlich sofort, dass jedes Erzeugendensystem $G \subseteq \mathbb{T}^n$ einer monomialen K-Unteralgebra T stets eine σ-SAGBI-Basis von T ist. Denn Teil a) des Satzes liefert uns

$$K[\mathrm{LT}_\sigma(f) : \ f \in T \setminus \{0\}] = K[\mathrm{LT}_\sigma(g) : \ g \in G].$$

Außerdem besitzt damit eine monomiale K-Unteralgebra eine eindeutig bestimmte, bzgl. Inklusion minimale σ-SAGBI-Basis. Ist T endlich erzeugt, so besitzt T also auch eine endliche σ-SAGBI-Basis.

Betrachten wir nun wieder eine beliebige, endlich erzeugte K-Unteralgebra S von P. Dann ist insbesondere $K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}]$ eine monomiale K-Unteralgebra von P. Für diese spezielle monomiale K-Unteralgebra von P ergeben sich aus dem letzten Satz sofort die folgenden Eigenschaften (vgl. [KR05], Korollar 6.6.5, S. 481).

Korollar 5.1.5. *Sei $S \subseteq P$ eine endlich erzeugte K-Unteralgebra von P.*

a) *Für jeden Term $t \in K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}]$ gibt es ein Polynom $f \in S \setminus \{0\}$ mit $t = \mathrm{LT}_\sigma(f)$.*

b) *Es gibt ein eindeutig bestimmtes und bzgl. Inklusion minimales Erzeugendensystem von $K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}]$ als K-Algebra, das aus Leittermen von Polynomen aus $S \setminus \{0\}$ besteht.*

c) *Ein Erzeugendensystem $G \subseteq P \setminus \{0\}$ von S ist genau dann eine σ-SAGBI-Basis von S, wenn das multiplikative Monoid $\{\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}\}$ von der Menge $\{\mathrm{LT}_\sigma(f) : \ f \in G\}$ erzeugt wird.*

Nachdem wir bereits wissen, dass es stets eine σ-SAGBI-Basis von S gibt, wollen wir nun der Frage nachgehen, wann eine endlich erzeugte K-Unteralgebra von P eine *endliche* σ-SAGBI-Basis besitzt. Dass im Allgemeinen eine K-Unteralgebra S selbst dann nicht immer eine endliche σ-SAGBI-Basis besitzt, wenn sie endlich erzeugt ist, zeigen die folgenden Beispiele, in denen sogar die verwendete Termordnung σ keine Rolle spielt (vgl. [KR05], Beispiel 6.6.7 und 6.6.8, S. 482 f.).

Beispiel 5.1.6. Sei $P = \mathbb{Q}[x_1, x_2]$.

a) Die endlich erzeugte $\mathbb{Q}$-Unteralgebra $S = \mathbb{Q}[x_1 + x_2,\ x_1x_2,\ x_1x_2^2]$ von P besitzt für keine Termordnung σ eine endliche σ-SAGBI-Basis.

b) Auch die endlich erzeugte $\mathbb{Q}$-Unteralgebra $S = \mathbb{Q}[x_1 - x_2,\ x_1x_2 - x_2^2,\ x_1x_2^2]$ von P besitzt für keine Termordnung σ eine endliche σ-SAGBI-Basis. ◁

Allerdings gibt es auch Fälle, in denen eine K-Unteralgebra S bzgl. einer Termordnung eine endliche SAGBI-Basis besitzt, aber bzgl. einer anderen Termordnung nicht (vgl. [RS90], Beispiel 4.11). In folgenden Fällen ist die Situation besonders einfach und die Frage nach der Existenz einer endlichen σ-SAGBI-Basis stets positiv zu beantworten:

- Im univariaten Polynomring $K[x]$ ist jede K-Unteralgebra S von $K[x]$ endlich erzeugt und sie besitzt zudem stets eine endliche σ-SAGBI-Basis, unabhängig von der Wahl der Termordnung σ (vgl. [KR05], Satz 6.6.9, S. 484).
- Laut Satz 5.1.4 besitzt eine endlich erzeugte monomiale K-Unteralgebra stets ein endliches minimales monomiales Erzeugendensystem. Dieses Erzeugendensystem ist eine endliche σ-SAGBI-Basis, und zwar unabhängig von der Wahl der Termordnung.

Wie sieht die Situation nun aber im allgemeinen Fall aus? Aus Korollar 5.1.5 können wir unmittelbar die folgende Aussage über die Existenz einer endlichen σ-SAGBI-Basis festhalten (vgl. [KR05], Korollar 6.6.5, S. 481).

Korollar 5.1.7. *Eine endlich erzeugte K-Unteralgebra S von P besitzt genau dann eine endliche σ-SAGBI-Basis, wenn das minimale monomiale Algebra-Erzeugendensystem der monomialen K-Unteralgebra $K[\mathrm{LT}_\sigma(f) : f \in S \setminus \{0\}]$ endlich ist.*

Laut Satz 5.1.4 ist dieses minimale monomiale Erzeugendensystem endlich, falls die monomiale K-Unteralgebra endlich erzeugt ist. Der nächste Satz liefert uns ein notwendiges Kriterium zur Überprüfung, ob die monomiale K-Unteralgebra $K[\mathrm{LT}_\sigma(f) : f \in S \setminus \{0\}]$ endlich erzeugt ist (vgl. [KR05], Satz 6.6.13).

Satz 5.1.8. *Sei S eine endlich erzeugte K-Unteralgebra von P. Gibt es für jede Unbestimmte x_i ein $a_i > 0$ mit $x_i^{a_i} \in K[\mathrm{LT}_\sigma(f) : f \in S \setminus \{0\}]$, so ist $K[\mathrm{LT}_\sigma(f) : f \in S \setminus \{0\}]$ eine endlich erzeugte K-Algebra.*

Konkret bedeutet das, dass S eine endliche σ-SAGBI-Basis besitzt, falls es für alle $i \in \{1, \ldots, n\}$ ein Polynom $f_i \in S$ gibt mit $\mathrm{LT}_\sigma(f_i) \in K[x_i]$. Auch auf andere Weise lässt sich mit Hilfe der Leitterme der Erzeuger einer K-Unteralgebra S auf die Existenz einer endlichen σ-SAGBI-Basis schließen (vgl. [KR05], Satz 6.6.11).

Satz 5.1.9. *Sei S eine endlich erzeugte K-Unteralgebra von P, d.h. es gibt endlich viele Polynome $f_1, \ldots, f_s \in P \setminus \{0\}$ mit $S = K[f_1, \ldots, f_s]$. Sind die Leitterme $\mathrm{LT}_\sigma(f_1), \ldots, \mathrm{LT}_\sigma(f_s)$ algebraisch unabhängig, so ist $\{f_1, \ldots, f_s\}$ eine σ-SAGBI-Basis von S.*

Die Umkehrung dieses Satzes gilt im Allgemeinen jedoch nicht, wie uns das bereits bekannte Beispiel 5.1.2 zeigt.

Beispiel 5.1.10. Sei $P = \mathbb{Q}[x_1, x_2, x_3]$ und sei $\sigma = \mathrm{DegLex}$. Sei $G = \{f_1, f_2, f_3\}$ mit den Polynomen $f_1 := x_1^2 - x_2x_3$, $f_2 := x_1x_2 + x_3^2$ und $f_3 := x_2^2 - x_3^2$. Setze

$$f_4 := f_2^2 - f_1f_3 = x_1^2x_3^2 + 2x_1x_2x_3^2 + x_2^3x_3 - x_2x_3^3 + x_3^4.$$

Dann ist $B := \{f_1, f_2, f_3, f_4\}$ bekanntlich eine σ-SAGBI-Basis von $S := K[G]$, aber die Leitterme $\mathrm{LT}_\sigma(f_1) = x_1^2$, $\mathrm{LT}_\sigma(f_2) = x_1x_2$, $\mathrm{LT}_\sigma(f_3) = x_2^2$ und $\mathrm{LT}_\sigma(f_4) = x_1^2x_3^2$ sind nicht algebraisch unabhängig. Betrachte dazu das Polynom $h = y_2^2 - y_1y_3 \in K[y_1, y_2, y_3, y_4]$. Dann gilt $h \neq 0$, aber $h(\mathrm{LT}_\sigma(f_1), \mathrm{LT}_\sigma(f_2), \mathrm{LT}_\sigma(f_3), \mathrm{LT}_\sigma(f_4)) = 0$. ◁

5.2 Der Unteralgebra-Divisionsalgorithmus und Ersetzungsregeln

Um die späteren Berechnungen von SAGBI-Basen effizient umsetzen zu können, benötigen wir in Analogie zur Theorie der Gröbner-Basen einen Divisionsalgorithmus. In der Literatur, insbesondere in [KR05], ist ein derartiger Algorithmus nicht unmittelbar zu finden. Diese Lücke

wollen wir nun schließen. Sei dazu wie bisher K ein Körper, $P = K[x_1, \ldots, x_n]$ und σ eine Termordnung auf $\mathbb{T}^n$. Der Divisionsalgorithmus der „Gröbner-Basen-Welt" berechnet für gegebene Polynome $f, g_1, \ldots, g_s \in P$ Polynome $q_1, \ldots, q_s \in P$ und $\tilde{f} \in P$ mit $f = q_1 g_1 + \ldots + q_s g_s + \tilde{f}$. Nun suchen wir nach einer analogen Darstellung, d.h. bei gegebenen Polynomen $f, g_1, \ldots, g_s \in P$ sind Polynome $h \in K[y_1, \ldots, y_s]$ und $\tilde{f} \in P$ zu bestimmen mit

$$f = h(g_1, \ldots, g_s) + \tilde{f}.$$

Für einen derartigen Unteralgebra-Divisionsalgorithmus wird es von zentraler Bedeutung sein, festzustellen, ob sich ein gegebener Term $t \in \mathbb{T}^n$ als Potenzprodukt von anderen Termen $t_1, \ldots, t_s \in \mathbb{T}^n$ schreiben lässt, d.h. ob es natürliche Zahlen $c_1, \ldots, c_s \in \mathbb{N}$ gibt mit

$$t = t_1^{c_1} \cdots t_s^{c_s} \tag{5.2.1}$$

Ist dies der Fall, so benötigen wir einen effizienten Algorithmus, der geeignete Exponenten $c_1, \ldots, c_s \in \mathbb{N}$ berechnet. Dazu schreiben wir für $j \in \{1, \ldots, s\}$ den Term t_j in der Form $t_j = x_1^{a_{1,j}} x_2^{a_{2,j}} \cdots x_n^{a_{n,j}}$ mit $a_{1,j}, \ldots, a_{n,j} \in \mathbb{N}$ und den Term t in der Form $t = x_1^{b_1} \cdots x_n^{b_n}$ mit $b_1, \ldots, b_n \in \mathbb{N}$. Sei weiter $\mathcal{A} \in \mathrm{Mat}_{n,s}(\mathbb{N})$ die Matrix $\mathcal{A} = (a_{i,j})$, d.h. die j-te Spalte von $\mathcal{A}$ stimmt mit dem Logarithmus $\log(t_j) \in \mathbb{N}^n$ von t_j überein. Somit gibt es also genau dann eine Darstellung von t in der Form aus Gleichung (5.2.1), wenn das lineare diophantische Gleichungssystem

$$\mathcal{A} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \tag{5.2.2}$$

eine Lösung in $\mathbb{N}^s$ besitzt. Setze $b := (b_1, \ldots, b_n)$ und $\mathcal{A}' := (\mathcal{A} | -b) \in \mathrm{Mat}_{n,s+1}(\mathbb{Z})$. Dann erhalten wir aus dem Gleichungssystem aus Gleichung (5.2.2) ein homogenes lineares diophantisches Gleichungssystem

$$\mathcal{A}' \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_s \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

das es zunächst zu lösen gilt. Sei $\mathcal{L}(\mathcal{A}') \subseteq \mathbb{Z}^{s+1}$ die Lösungsmenge des zu $\mathcal{A}'$ gehörenden, homogenen linearen diophantischen Gleichungssystems und sei $\mathcal{L}_+(\mathcal{A}') = \mathcal{L}(\mathcal{A}') \cap \mathbb{N}^{s+1}$ die Menge der nicht-negativen Lösungen dieses Systems. Kennt man die Lösungsmenge $\mathcal{L}_+(\mathcal{A}')$, so lässt sich die Lösungsmenge von Gleichung (5.2.2) unmittelbar ablesen, denn jedes Element von $\mathcal{L}_+(\mathcal{A}')$ der Form $(c_1, \ldots, c_s, 1)$ löst Gleichung (5.2.2). Die Lösungsmenge $\mathcal{L}_+(\mathcal{A}')$ wird erzeugt von der endlichen Hilbertbasis (siehe Abschnitt 2.3), die sich effizient berechnen lässt (siehe Algorithmus 2.3). Dazu wollen wir uns nun ein Beispiel betrachten.

Beispiel 5.2.1. Sei $P = \mathbb{Q}[x, y, z]$ und $\sigma = \mathrm{DegLex}$. Weiter seien $t_1 = x^2 y$, $t_2 = yz^2$ und $t_3 = xyz$ gegeben. Wir wollen nun den Term $t = x^4 y^3 z^2$ in den Termen t_1, t_2, t_3 darstellen. Zu lösen ist also das lineare diophantische Gleichungssystem

$$\begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$$

Wir setzen also $\mathcal{A}' := \left(\begin{smallmatrix} 2 & 0 & 1 & -4 \\ 1 & 1 & 1 & -3 \\ 0 & 2 & 1 & -2 \end{smallmatrix}\right)$ und erhalten mit Algorithmus 2.3 folgende Hilbertbasis von $\mathcal{L}_+(\mathcal{A}')$:

$$H = \{(1, 0, 2, 1), (2, 1, 0, 1)\},$$

d.h. alle Lösungen in $\mathcal{L}_+(\mathcal{A})$ lassen sich durch $\mathbb{N}$-Linearkombinationen aus diesen beiden Tupeln erzeugen. Da wir an einer Darstellung von t interessiert sind, sind nur die Tupel von Relevanz, deren letzten Komponente 1 ist. Wie man leicht sieht, sind die beiden Tupel aus H die einzigen Elemente in $\mathcal{L}_+(\mathcal{A})$ mit dieser Eigenschaft. Somit erhalten wir die beiden Darstellungen $t = t_1 t_3^2$ und $t = t_1^2 t_2$. ◁

Klar ist auch, dass die Hilbertbasis genau dann leer ist, wenn das Gleichungssystem keine Lösung in $\mathbb{N}^{s+1}$ besitzt. Wie man an dem Beispiel aber auch sieht, gibt es im Allgemeinen mehrere Lösungstupel, aber stets endlich viele.

Lemma 5.2.2. *Seien $t_1, \ldots, t_s \in \mathbb{T}^n$ für $s \in \mathbb{N}_+$. Dann gibt es für alle $t \in \mathbb{T}^n$ nur endlich viele Tupel $(c_1, \ldots, c_s, 1) \in \mathbb{N}^{s+1}$ mit $t = t_1^{c_1} \cdots t_s^{c_s}$. Dies sind genau diejenigen Elemente der Hilbertbasis der Form $(c_1, \ldots, c_s, 1)$.*

Beweis: Sei wie oben $\mathcal{A}' \in \mathrm{Mat}_{n,s+1}(\mathbb{Z})$ die aus den Termen $t, t_1, \ldots, t_s$ resultierende Matrix des homogenen linearen diophantischen Gleichungssystems, d.h. in der letzten Spalte von $\mathcal{A}'$ sind alle Einträge ungleich 0 negativ und die restlichen Einträge der Matrix sind nicht-negative ganze Zahlen. Laut [KR05], Theorem 6.1.17, S. 361 ist die Hilbertbasis $H \subseteq \mathbb{N}^{s+1}$ von $\mathcal{L}_+(\mathcal{A}')$ stets endlich. Sei H nun nicht leer. Dann ist die letzte Komponente jedes Elements von H ungleich 0. Denn angenommen, es gibt ein Tupel $(a_1, \ldots, a_s, 0)$ mit $(a_1, \ldots, a_s) \neq 0$ in H. Dann gilt $t_1^{a_1} \cdots t_s^{a_s} \neq 1$ wegen $t_1, \ldots, t_s \neq 1$ und damit $t_1^{a_1} \cdots t_s^{a_s} \neq t^0$ für alle $t \in \mathbb{T}^n$. Somit kann es in $\mathcal{L}_+(\mathcal{A}')$ nur endlich viele Tupel der Form $(c_1, \ldots, c_s, 1)$ geben, nämlich genau die, die in H enthalten sind. Denn durch jede nicht-triviale $\mathbb{N}$-Linearkombination mit Elementen aus H entstehen Tupel, deren letzte Komponente ungleich 1 ist. □

Damit lässt sich unter Verwendung von Algorithmus 2.3 ein Algorithmus zur Bestimmung der endlichen Lösungsmenge von Gleichung (5.2.2) festhalten.

Algorithmus 5.1 : Berechnung aller Termdarstellungen

Input : $t = x_1^{b_1} \cdots x_n^{b_n}$ mit $b := (b_1, \ldots, b_n) \in \mathbb{N}^n$
Input : $t_1, \ldots, t_s \in \mathbb{T}^n$ der Form $t_j = x_1^{a_{1,j}} x_2^{a_{2,j}} \cdots x_n^{a_{n,j}}$ mit $a_{1,j}, \ldots, a_{n,j} \in \mathbb{N}$ für alle j.
Result : $C = \{(c_1, \ldots, c_s) \in \mathbb{N}^s : \ t = t_1^{c_1} \cdots t_s^{c_s}\}$

1 $\mathcal{A} := (a_{i,j}) \in \mathrm{Mat}_{n,s}(\mathbb{N})$ und $\mathcal{A}' := (\mathcal{A} | -b) \in \mathrm{Mat}_{n,s+1}(\mathbb{Z})$;
2 Berechne mit Algorithmus 2.3 die Hilbertbasis $H \subseteq \mathbb{N}^{s+1}$ von $\mathcal{L}_+(\mathcal{A}')$;
3 $C := \{(c_1, \ldots, c_s) : \ (c_1, \ldots, c_s, 1) \in H\}$;
4 **return** C;

Dieser Algorithmus ist Teil des ApCoCoA-Paketes `sagbi.cpkg`, er ist implementiert in der Hilfsfunktion `SB.TermReprToric` (siehe Anhang C, Seite 308). Diese Hilfsfunktion bildet einen Teil der Metafunktion `SB.TermRepr` (siehe Anhang C, Seite 304). Dort wird ggf. eine Darstellung ausgewählt und zurückgegeben, worauf wir gleich zurück kommen werden. Zuvor beweisen wir die Korrektheit und Endlichkeit des Algorithmus.

Satz 5.2.3. (Berechnung aller Termdarstellungen)
Seien $t_1, \ldots, t_s \in \mathbb{T}^n$ mit $s \in \mathbb{N}_+$ echte Terme und sei $t \in \mathbb{T}^n$ ein beliebiger Term. Dann berechnet Algorithmus 2.3 die endliche Menge C aller Tupel $(c_1, \ldots, c_s) \in \mathbb{N}^s$ mit $t = t_1^{c_1} \cdots t_s^{c_s}$. Insbesondere ist C genau dann leer, wenn es keine derartige Darstellung gibt.

Beweis: Die Endlichkeit von Algorithmus 5.1 folgt sofort aus der Endlichkeit von Algorithmus 2.3. Ist die Hilbertbasis H leer, gibt es keine nicht-negativen Lösungen des zugehörigen

homogenen linearen diophantischen Gleichungssystems und dementsprechend auch keine Termdarstellung. Sei nun $H \neq \emptyset$. Laut Lemma 5.2.2 gibt es nur endliche viele Elemente in C. Ist C leer, gibt es folglich keine Termdarstellung. Umgekehrt ist C offensichtlich leer, wenn es keine Termdarstellung gibt. □

Besitzt ein Term $t \in \mathbb{T}^n$ also eine Darstellung in Termen $t_1, \ldots, t_s \in \mathbb{T}^n$, d.h. ist C nicht leer, so gibt es im Allgemeinen mehr als ein Element in der stets endlichen Menge C, wie auch am letzten Beispiel zu sehen war. Allerdings benötigen wir – im Falle der Existenz – nur *eine* Darstellung von t in $t_1, \ldots, t_s$. Wie oben bereits erwähnt, benötigen wir also eine Auswahlstrategie.

Bemerkung 5.2.4. (Auswahlstrategien)
Sei $C \neq \emptyset$. Dann lassen sich die Elemente in $C \subseteq \mathbb{N}^s$ lexikographisch ordnen, d.h. für zwei Tupel $(u_1, \ldots, u_s), (v_1, \ldots, v_s) \in \mathbb{N}^s$ betrachten wir die durch

$$(u_1, \ldots, u_s) \succeq_{\text{Lex}} (v_1, \ldots, v_s) \quad :\Longleftrightarrow \\ u_1 \geq v_1 \text{ oder } (u_1 = v_1 \text{ und } (u_2, \ldots, u_s) \succeq_{\text{Lex}} (v_2, \ldots, v_s))$$

rekursiv definierte Ordnung auf $\mathbb{N}^s$. Da C stets endlich ist, können wir also Minimum oder Maximum bestimmen und als Wahl für eine Termdarstellung festlegen. Wir werden im Folgenden stets das Maximum wählen. Natürlich sind durchaus auch andere Auswahlstrategien denkbar, entscheidend für den späteren Divisionsalgorithmus ist v.a., dass stets ein und dieselbe Strategie konsequent verfolgt wird. Wir werden die „Maximumstrategie" verfolgen.

Durch die Festlegung einer Auswahlstrategie ist es uns gelungen, eine Darstellung eines Terms $t \in \mathbb{T}^n$ in Termen $t_1, \ldots, t_s \in \mathbb{T}^n$ zu berechnen, und das in gewissem Sinne auf eindeutige Weise. Allerdings ist das nicht die einzige Art und Weise, eine derartige Darstellung zu bestimmen. Wie wir aus Abschnitt 2.3 wissen, beruht die Berechnung der Hilbertbasis auf der Berechnung torischer Ideale, genauer in dieser Situation auf der Berechnung des torischen Ideals $I(\mathcal{A}')$ bzw. $I(t_1, \ldots, t_s, t')$, wobei t' der erweiterter Term $x_1^{-b_1} \cdots x_n^{-b_n} \in \mathbb{E}^n$ ist (siehe Abschnitt 2.3). Für Terme $t_1, \ldots, t_s \in \mathbb{T}^n$ ist das torische Ideal $I(t_1, \ldots, t_s)$ genau das Relationenideal $\text{Rel}(t_1, \ldots, t_s)$, das sich somit ebenfalls effizient berechnen lässt (siehe Algorithmus 2.2). Das Relationenideal bietet eine alternative Methode zur Bestimmung einer Termdarstellung.

Bemerkung 5.2.5. (Alternative Berechnungsmethode für Termdarstellungen)
Sei $\Phi : K[y_1, \ldots, y_s] \to P$ der durch $y_i \mapsto t_i$ für alle $i \in \{1, \ldots, s\}$ definierte K-Algebra-Homomorphismus. Mit dem Diagonalideal $\triangle_\Phi = \langle y_1 - t_1, y_2 - t_2, \ldots, y_s - t_s \rangle$ im Polynomring $Q := K[x_1, \ldots, x_n, y_1, \ldots, y_n]$ gilt $\text{Rel}(t_1, \ldots, t_s) = \text{Ker}(\Phi) = \triangle_\Phi \cap K[y_1, \ldots, y_s]$. Sei nun weiter $S = K[t_1, \ldots, t_s]$ und σ eine Eliminationsordnung für $\{x_1, \ldots, x_n\}$, so gilt $t \in S$ genau dann, wenn $\text{NF}_{\triangle_\Phi}(t) \in K[y_1, \ldots, y_s]$ gilt (siehe Satz 2.2.8). Damit ist die Frage nach der Existenz einer Darstellung von t in $t_1, \ldots, t_s$ beantwortet. Ist $t \in S$, so ist $\tilde{t} := \text{NF}_{\triangle_\Phi}(t)$ ein Term in $K[y_1, \ldots, y_s]$ mit $t = \tilde{t}(t_1, \ldots, t_s)$ (siehe Satz 2.2.8) und wir erhalten eine explizite Darstellung von t in $t_1, \ldots, t_s$.

Diese Alternative ist mehr oder weniger direkt in allgemeinerer Form in CoCoA verfügbar. Auf unsere Bedürfnisse zugeschnitten, ist sie ebenfalls Teil des ApCoCoA-Pakets `sagbi.cpkg` als Hilfsfunktion `SB.TermReprAlghom` (siehe Anhang C, Seite 307). Diese Hilfsfunktion ist analog zu `SB.TermReprToric` durch die Metafunktion `SB.TermRepr` aufrufbar. Darüber hinaus gibt es eine dritte „Variante" der Berechnung einer Termdarstellung im Paket `sagbi.cpkg`, die ebenfalls über die Metafunktion aufrufbar ist: In der Funktion `SB.TermReprDio` wurde

die Methode aus der letzten Bemerkung „zu Fuß" umgesetzt, d.h. ohne die direkten CoCoA-Befehle zu verwenden (siehe dazu [KR00], Tutorial 36, S. 207). Nun aber wollen wir endlich den Unteralgebra-Divisionsalgorithmus präsentieren.

Algorithmus 5.2 : Unteralgebra-Divisionsalgorithmus

Input : $f \in P \setminus \{0\}$
Input : Polynome $g_1, \ldots, g_s \in P \setminus \{0\}$ mit $s \in \mathbb{N}_+$.
Result : $\tilde{f} \in P$ und $h \in K[y_1, \ldots, y_s]$ mit $f = h(g_1, \ldots, g_s) + \tilde{f}$.

```
1  f̃ := f und h := 0 in K[y_1,...,y_s];
2  T := Supp(f);
3  while T ≠ ∅ do
4      t := max_σ T;
5      T := T \ {t};
6      Bestimme mit Algorithmus 5.1 C = {(c_1,...,c_s) ∈ N^s : t = LT_σ(g_1)^{c_1} ··· LT_σ(g_s)^{c_s}};
7      if C ≠ ∅ then
8          (c_1,...,c_s) := max_{≥Lex} C;
9          t̃ := y_1^{c_1} ··· y_s^{c_s} ∈ K[y_1,...,y_s];
10         λ_1 := Koeffizient von t in f̃;
11         λ_2 := Koeffizient von t in t̃(g_1,...,g_s);
12         λ := λ_1/λ_2;
13         f̃ := f̃ − λ · t̃(g_1,...,g_s);
14         h := h + λ · t̃;
15         T := Supp(f̃);
16 return (h, f̃);
```

Die Korrektheit und Endlichkeit von Algorithmus 5.2 folgt aus dem nächsten Theorem.

Theorem 5.2.6. (Der Unteralgebra-Divisionsalgorithmus)
Seien $f, g_1, \ldots, g_s \in P \setminus \{0\}$ für $s \in \mathbb{N}_+$. Dann berechnet Algorithmus 5.2 in endlich vielen Schritten ein Polynom $h \in K[y_1, \ldots, y_s]$ und ein Polynom $\tilde{f} \in P$ mit

$$f = h(g_1, \ldots, g_s) + \tilde{f}$$

so, dass die folgenden Bedingungen erfüllt sind:

(i) $\mathrm{Supp}(\tilde{f}) \cap K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)] = \emptyset$.

(ii) Für alle $t \in \mathrm{Supp}(h)$ gilt $\mathrm{LT}_\sigma(t(g_1, \ldots, g_s)) \leq_\sigma \mathrm{LT}_\sigma(f)$.

Beweis: Zu Beginn des Algorithmus gilt $f = h(g_1, \ldots, g_s) + \tilde{f}$ wegen $\tilde{f} = f$ und $h = 0$. Weiter gilt in jedem Schleifendurchlauf und damit zu jedem Zeitpunkt des Algorithmus

$$f = h(g_1, \ldots, g_s) + \lambda\tilde{t}(g_1, \ldots, g_s) + \tilde{f} - \lambda\tilde{t}(g_1, \ldots, g_s) = h(g_1, \ldots, g_s) + \tilde{f}.$$

Wegen $\mathrm{LT}_\sigma(\tilde{t}(g_1, \ldots, g_s)) = \tilde{t}(\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)) = \mathrm{LT}_\sigma(\tilde{f})$ wird in Schritt 13 der Leitterm $\mathrm{LT}_\sigma(\tilde{f})$ durch Terme ersetzt, die bzgl. σ kleiner sind. Da zudem in Zeile 5 aus $T = \mathrm{Supp}(\tilde{f})$ jeweils der bzgl. σ größte Term entfernt wird, muss T nach endlich vielen Schritten leer sein. Somit terminiert die Schleife und damit der Algorithmus. Am Ende des Algorithmus gilt also $f = h(g_1, \ldots, g_s) + \tilde{f}$ mit Polynomen $h \in K[y_1, \ldots, y_s]$ und $\tilde{f} \in P$.

Nun zu den beiden Bedingungen. Solange C nicht leer ist, wird jeder Term t in $\mathrm{Supp}(\tilde{f})$ mit $t \in K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)]$ durch bzgl. σ kleinere Terme ersetzt. Somit ist die Bedingung $\mathrm{Supp}(\tilde{f}) \cap K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)] = \emptyset$ erfüllt, sobald $T = \emptyset$ erreicht ist.

Zu Beginn der ersten Schleife wird $t = \mathrm{LT}_\sigma(f)$ betrachtet und für diesen Term C berechnet. Ist C nicht leer, so wird in Zeile 13 der Leitterm von f durch Terme ersetzt, die kleiner sind als $\mathrm{LT}_\sigma(f)$. Somit wird zu $\mathrm{Supp}(h)$ ein Term hinzugefügt, der $\mathrm{LT}_\sigma(\tilde{t}(g_1, \ldots, g_s)) \leq_\sigma \mathrm{LT}_\sigma(f)$ erfüllt. Ist $C = \emptyset$, bleibt zunächst auch $\mathrm{Supp}(h)$ leer. Am Anfang der zweiten Schleife ist der betrachtete Term t also in jedem Fall kleiner als $\mathrm{LT}_\sigma(f)$. Existiert für diesen Term eine Darstellung in den Leittermen von $g_1, \ldots, g_s$, wird zu $\mathrm{Supp}(h)$ erneut ein Term $\tilde{t}$ hinzugefügt, für den $\mathrm{LT}_\sigma(\tilde{t}(g_1, \ldots, g_s)) \leq_\sigma \mathrm{LT}_\sigma(f)$ erfüllt ist. Das setzt sich nun induktiv so fort, sodass am Ende des Algorithmus auch die zweite Bedingung für alle Terme $t \in \mathrm{Supp}(h)$ erfüllt ist. □

Der Unteralgebra-Divisionsalgorithmus ist in der CoCoA-Funktion `SB.SubalgebraDivAlg` des ApCoCoA-Pakets `sagbi.cpkg` implementiert (siehe Anhang C, Seite 303). Zum Ablauf dieses Algorithmus wollen wir nun ein Beispiel angeben, an dem auch deutlich wird, wie das resultierende Polynom $\tilde{f}$ bei Berechnung durch den Algorithmus von der Reihenfolge der Polynome $g_1, \ldots, g_s$ abhängt.

Beispiel 5.2.7. Sei $P = \mathbb{Q}[x]$ und sei $f = x^8$. Zunächst seien $g_1 = x^3 - x$, $g_2 = x^4$ und $g_3 = x^5 - 1$. Wir beginnen in Schritt 6 des Algorithmus mit $t = \mathrm{LT}_\sigma(f) = x^8$. Algorithmus 5.1 liefert die Menge $C = \{(1, 0, 1), (0, 2, 0)\}$. Somit wird in Schritt 8 das Tupel $(1, 0, 1)$ gewählt und wir erhalten

$$\tilde{f} = x^8 - (x^3 - x)(x^5 - 1) = x^6 + x^3 - x$$

sowie $h = y_1 y_3$. Im nächsten Schleifendurchlauf liefert Algorithmus 5.1 für $t = x^6$ die Menge $C = \{(2, 0, 0)\}$. Damit erhalten wir $\tilde{f} = 2x^4 + x^3 - x^2 - x$ und $h = y_1 y_3 + y_1^2$. Im dritten Schleifendurchlauf folgt $C = \{(0, 1, 0)\}$ für $t = x^4$. Somit folgt $\tilde{f} = x^3 - x^2 - x$ und $h = y_1 y_3 + y_1^2 + 2y_2$. Im vierten Schleifendurchlauf erhalten wir $C = \{(1, 0, 0)\}$ für $t = x^3$. Damit folgt $\tilde{f} = -x^2$ und $h = y_1 y_3 + y_1^2 + 2y_2 + y_1$. Nun ist im fünften Schleifendurchlauf $C = \emptyset$ und dann auch $T = \emptyset$, d.h. wir erhalten die Darstellung

$$f = h(g_1, g_2, g_3) - x^2 = (g_1 g_3 + g_1^2 + 2g_2 + g_1) - x^2.$$

Nun ändern wir die Reihenfolge. Sei $g_1 = x^4$, $g_2 = x^3 - x$ und $g_3 = x^5 - 1$. Dann erhalten wir im ersten Schleifendurchlauf $C = \{(2, 0, 0), (0, 1, 1)\}$, d.h. wir wählen in Schritt 8 das Tupel $(2, 0, 0)$. Damit folgt $\tilde{f} = 0$ sowie $h = y_1^2$ und der Algorithmus terminiert bereits. ◁

Dem von der Reihenfolge des Tupels $(g_1, \ldots, g_s)$ abhängigen Polynom $\tilde{f}$, das Algorithmus 5.2 berechnet, wollen wir in Analogie zur Theorie der Gröbner-Basen noch einen passenden Namen geben.

Definition 5.2.8. (Normaler Unteralgebra-Rest)
Seien $f, g_1, \ldots, g_s \in P \setminus \{0\}$ und sei $\mathcal{G} = (g_1, \ldots, g_s) \in P^s$ für $s \in \mathbb{N}_+$. Seien $h \in K[y_1, \ldots, y_s]$ und $\tilde{f} \in P$ mit $f = h(g_1, \ldots, g_s) + \tilde{f}$ die Ergebnisse von Algorithmus 5.2 bzgl. $(f, g_1, \ldots, g_s)$. Dann heißt das Polynom $\tilde{f} \in P$ der **normale Unteralgebra-Rest** von f bzgl. $\mathcal{G}$, der mit $\mathrm{NRS}_{\sigma,\mathcal{G}}(f)$ oder kurz mit $\mathrm{NRS}_{\mathcal{G}}(f)$ bezeichnet wird. Für $f = 0$ setzen wir $\mathrm{NRS}_{\mathcal{G}}(f) = 0$.

Für den normalen Unteralgebra-Rest ist auch eine CoCoA-Funktion im Paket `sagbi.cpkg` enthalten. Mit `SB.NRS` kann man sich den normalen Unteralgebra-Rest eines Polynoms f bzgl. eines Tupels $(g_1, \ldots, g_s)$ ausgeben lassen. Der normale Unteralgebra-Rest ermöglicht als erste einfache Anwendung sofort einen notwendigen Unteralgebra-Mitgliedschaftstest.

Korollar 5.2.9. (Notwendiger Unteralgebra-Mitgliedschaftstest)
Sei $f \in P$, seien $g_1, \ldots, g_s \in P \setminus \{0\}$ und sei $\mathcal{G} = (g_1, \ldots, g_s)$. Ist $\mathrm{NRS}_{\mathcal{G}}(f) = 0$, so gilt $f \in K[g_1, \ldots, g_s]$.

Beweis: Für $f = 0$ ist die Behauptung klar. Sei $f \neq 0$. Laut Theorem 5.2.6 berechnet Algorithmus 5.2 ein Polynom $h \in K[y_1, \ldots, y_s]$ und ein Polynom $\tilde{f} \in P$ mit $f = h(g_1, \ldots, g_s) + \tilde{f}$. Gilt also $\tilde{f} = \mathrm{NRS}_{\mathcal{G}}(f) = 0$, so folgt $f \in K[g_1, \ldots, g_s]$. □

Die Umkehrung gilt im Allgemeinen jedoch nicht, wie bereits an Beispiel 5.2.7 zu sehen war. Der entscheidende Schritt zur Berechnung des normalen Unteralgebra-Rests $\mathrm{NRS}_{\mathcal{G}}(f)$ mit Algorithmus 5.2 steckt in Zeile 13. Nachdem ein passender Term $\tilde{t} \in K[y_1, \ldots, y_s]$ gefunden wurde, wird an dieser Stelle ein Term aus $\tilde{f}$ durch bzgl. σ kleinere Terme ersetzt. Man kann also davon reden, dass $\tilde{f}$ *reduziert* wird. Dieser Vorgang des Reduzierens lässt sich analog zur Welt der Gröbner-Basen durch Ersetzungsregeln ausdrücken (vgl. [KR05], Definition 6.6.16, S. 487).

Definition 5.2.10. (Unteralgebra-Ersetzungsregel)
Sei σ eine Termordnung auf $\mathbb{T}^n$ und sei $G \subseteq P \setminus \{0\}$ eine nicht-leere Menge von Polynomen.

a) Seien $f, \tilde{f} \in P$. Gibt es eine Konstante $c \in K$, Polynome $g_1, \ldots, g_s \in G$ und einen Term $t \in K[y_1, \ldots, y_s]$ mit
$$\tilde{f} = f - c \cdot t(g_1, \ldots, g_s)$$
und $t(\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)) \notin \mathrm{Supp}(\tilde{f})$, so wird f in einem Schritt **Unteralgebra-reduziert** bzgl. $g_1, \ldots, g_s$ zu $\tilde{f}$. Dieser Vorgang des Reduzierens heißt ein **Unteralgebra-Reduktionsschritt** und wird mit $f \xrightarrow{g_1,\ldots,g_s}_{\mathrm{SS}} \tilde{f}$ oder kurz mit $f \xrightarrow{G}_{\mathrm{SS}} \tilde{f}$ bezeichnet.

b) Der transitive Abschluss der Relation $\xrightarrow{G}_{\mathrm{SS}}$ heißt die durch G definierte **Unteralgebra-Ersetzungsregel** und wird mit $\xrightarrow{G}_{\mathrm{S}}$ bezeichnet.

c) Sei $f \in P$. Gibt es kein $\tilde{f} \in P \setminus \{f\}$ und keine Polynome $g_1, \ldots, g_s \in G$ mit $f \xrightarrow{g_1,\ldots,g_s}_{\mathrm{SS}} \tilde{f}$, so heißt f **irreduzibel** bzgl. $\xrightarrow{G}_{\mathrm{S}}$, ansonsten **reduzibel**.

d) Die durch $\xrightarrow{G}_{\mathrm{S}}$ definierte Äquivalenzrelation wird mit $\xleftrightarrow{G}_{\mathrm{S}}$ bezeichnet.

Die Buchstaben „SS" bzw. „S" in der Notation von $\xrightarrow{G}_{\mathrm{SS}}$ bzw. $\xrightarrow{G}_{\mathrm{S}}$ stehen dabei für „**S**ubalgebra reduction **S**tep" bzw. „**S**ubalgebra rewrite relation". Somit hat der Buchstabe „S" also keinen Bezug zu einer speziellen K-Unteralgebra S, wie man leicht hätte vermuten können. Wie man an der Definition sofort sehen kann, lässt sich mit $c = 0$ jedes Polynom $f \in P$ trivialerweise zu sich selbst reduzieren. Außerdem ist ein konstantes Polynom immer zu 0 reduzierbar.

Bemerkung 5.2.11. Sei σ eine Termordnung und sei $G \subseteq P \setminus \{0\}$. Ein Polynom $f \in P$ ist genau dann reduzibel bzgl. der Unteralgebra-Ersetzungsregel $\xrightarrow{G}_{\mathrm{S}}$, wenn es Polynome $g_1, \ldots, g_s \in G$ und einen Term $t \in K[y_1, \ldots, y_s]$ gibt mit $t(\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)) \in \mathrm{Supp}(f)$, oder anders ausgedrückt, genau dann, wenn $\mathrm{Supp}(f) \cap K[\mathrm{LT}_\sigma(G)] \neq \emptyset$ gilt. Umgekehrt ist somit f genau dann irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$, wenn $\mathrm{NRS}_{\mathcal{G}}(f) = f$ für alle $\mathcal{G} = (g_1, \ldots, g_s) \in G^s$ gilt.

Die Unteralgebra-Ersetzungsregel besitzt darüber hinaus einige der Ersetzungsregel in der Gröbner Basen Theorie entsprechende Eigenschaften (vgl. [KR05], Satz 6.6.17, S. 488). So verhält sich die durch G definierte Unteralgebra-Ersetzungsregel wie die Ersetzungsregel für Gröbner-Basen und ist ebenfalls im Allgemeinen nicht konfluent (vgl. [KR05], Beispiel 6.6.19). Dies ist genau dann der Fall, wenn G eine σ-SAGBI-Basis von S bildet.

Beispiel 5.2.12. (Gegenbeispiel zur Konfluenz)
Sei $P = K[x]$, $\sigma = \mathrm{Deg}$ und seien $g_1 := x^3 - x$, $g_2 := x^4$ und $g_3 := x^5 - 1$. Sei $G := \{g_1, g_2, g_3\}$

und $S := K[G]$. Wegen $x^2 = g_1^2 + g_1 g_3 - g_2^2 + g_1 + 2g_2$ gilt $x^2 \in S$, weshalb G keine σ-SAGBI-Basis von S ist, da $\mathrm{Supp}(x^2) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$ gilt. Weiter gilt $x^8 \xrightarrow{-g_2^2}_{\mathrm{SS}} 0$ und

$$x^8 \xrightarrow{-g_1 g_3}_{\mathrm{SS}} x^8 - (x^3 - x)(x^5 - 1) = x^6 + x^3 - x \xrightarrow{-g_1^2}_{\mathrm{SS}} 2x^4 + x^3 - x^2 - x$$
$$\xrightarrow{-2g_2}_{\mathrm{SS}} x^3 - x^2 - x \xrightarrow{-g_1}_{\mathrm{SS}} -x^2.$$

Somit gibt es kein eindeutiges bzgl. $\xrightarrow{G}_{\mathrm{S}}$ irreduzibles Polynom $\tilde{f}$ zu $f = x^8$. Wie wir aus Beispiel 5.2.7 wissen gilt $\mathrm{NRS}_{\mathcal{G}}(f) = -x^2$ für $\mathcal{G} = (g_1, g_2, g_3)$ und $\mathrm{NRS}_{\mathcal{G}}(f) = 0$ für $\mathcal{G} = (g_2, g_1, g_3)$. Indem man in der obigen Reduktionskette die durchgeführten Reduktionen „aufsammelt" erhalten wir eine Darstellungen gemäß Theorem 5.2.6. So gilt $f = g_2^2 + 0$ für $\mathcal{G} = (g_2, g_1, g_3)$ und $f = g_1 g_3 + g_1^2 + 2g_2 + g_1 - x^2$ für $\mathcal{G} = (g_1, g_2, g_3)$. ◁

Weiter ist es ebenfalls möglich, in Analogie zur Theorie der Gröbner Basen ein fundamentales SAGBI-Diagramm anzugeben (vgl. [KR05], Satz 6.6.22, S. 490). Außerdem lassen sich nun neben den bereits bekannten noch weitere Charakterisierungen von SAGBI-Basen formulieren. Einen Teil davon wollen wir hier abschließend noch angeben (vgl. [KR05], Theorem 6.6.25, S. 492 f.).

Theorem 5.2.13. (Weitere Charakterisierungen von SAGBI-Basen)
Sei S eine endlich erzeugte K-Unteralgebra von P und sei $G \subseteq P \setminus \{0\}$ eine nicht-leere (aber nicht notwendigerweise endliche) Menge von Polynomen mit $S = K[G]$. Sei σ eine Termordnung auf $\mathbb{T}^n$ und sei $\xrightarrow{G}_S$ die durch G definierte Unteralgebra Ersetzungsregel. Dann sind die folgenden Aussagen äquivalent:

(i) *G ist eine σ-SAGBI-Basis von S.*

(ii) *Für jedes Polynom $f \in S \setminus \{0\}$ gibt es $g_1, \ldots, g_s \in G$ und ein Polynom $h \in K[y_1, \ldots, y_s]$ mit $f = h(g_1, \ldots, g_s)$ und $\mathrm{LT}_\sigma(f) \geq_\sigma \mathrm{LT}_\sigma(t(g_1, \ldots, g_s))$ für alle Terme $t \in \mathrm{Supp}(h)$.*

(iii) *Für jedes Polynom $f \in S \setminus \{0\}$ gibt es $g_1, \ldots, g_s \in G$ und ein Polynom $h \in K[y_1, \ldots, y_s]$ mit $f = h(g_1, \ldots, g_s)$ und $\mathrm{LT}_\sigma(f) = \max_\sigma\{\mathrm{LT}_\sigma(t(g_1, \ldots, g_s)) \mid t \in \mathrm{Supp}(h)\}$.*

(iv) *Für alle $f \in P$ gilt $f \xrightarrow{G}_S 0$ genau dann, wenn $f \in S$ ist.*

(v) *Ist $f \in S$ bzgl. $\xrightarrow{G}_S$ irreduzibel, so gilt $f = 0$.*

(vi) *Für jedes Polynom $f \in P$ gibt es ein eindeutig bestimmtes Polynom $\tilde{f} \in P$ mit $f \xrightarrow{G}_S \tilde{f}$, das bzgl. $\xrightarrow{G}_S$ irreduzibel ist.*

(vii) *Die Unteralgebra Ersetzungsregel $\xrightarrow{G}_S$ ist konfluent.*

5.3 Normalform und reduzierte SAGBI-Basis

In diesem Abschnitt soll die Analogie zur Gröbner Basen Theorie noch etwas ausgebaut werden. Die folgenden Ausführungen sind v.a. motiviert durch Tutorial 96 in [KR05], S. 500 ff., das in diesem und den folgenden Abschnitten vollständig ausgearbeitet wurde. Insbesondere die reduzierte SAGBI-Basis findet sich auch bereits in [RS90]. Außerdem konnten wir noch über Tutorial 96 hinausgehende Folgerungen beweisen. Dadurch wird insgesamt die Theorie der SAGBI-Basen erweitert. Es stellte sich zudem heraus, dass Tutorial 96 mehrere kleine Fehler beinhaltete, die hier ebenfalls korrigiert wurden.

Wie wohl bekannt ist, ist die Normalform eines Polynoms bzgl. eines Ideals genau der normale Rest dieses Polynoms bei Division durch eine Gröbner Basis des Ideals. Dieser normale Rest ist zudem unabhängig von der Wahl der Gröbner Basis (vgl. [KR00], Satz 2.4.7). Wir wollen nun für Unteralgebren ein Analogon zu den Normalformen für Gröbner-Basen angeben. Sei dazu wie bisher K ein Körper, $P = K[x_1, \ldots, x_n]$ ein Polynomring in den Unbestimmten $x_1, \ldots, x_n$ über K und sei σ eine Termordnung auf $\mathbb{T}^n$. Zunächst können wir folgenden Satz beweisen.

Satz 5.3.1. *Sei S eine endlich erzeugte K-Unteralgebra von P und sei $f \in P$.*

a) Ist $G \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S, so gibt es ein eindeutig bestimmtes Polynom $f_G \in P$ mit $f - f_G \in S$ und $\mathrm{Supp}(f_G) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$.

b) Das eindeutig bestimmte Polynom f_G aus Teil a) ist unabhängig von der Wahl der σ-SAGBI-Basis.

Beweis:

a) Da G eine σ-SAGBI-Basis von S ist, folgt aus Theorem 5.2.13, dass es ein eindeutig bestimmtes, bzgl. $\xrightarrow{G}_{\mathrm{S}}$ irreduzibles Polynom $f_G \in P$ gibt mit $f \xrightarrow{G}_{\mathrm{S}} f_G$. Aus [KR05], Satz 6.6.17, S. 488 folgt $f - f_G \in S$ und aus Bemerkung 5.2.11 $\mathrm{Supp}(f_G) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$.

b) Seien G und H σ-SAGBI-Basen von S. Dann gibt es nach Teil a) für G und H jeweils eindeutig bestimmte Polynome $f_G, f_H \in P$ mit $f - f_G \in S$ und $\mathrm{Supp}(f_G) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$ sowie $f - f_H \in S$ und $\mathrm{Supp}(f_H) \cap K[\mathrm{LT}_\sigma(H)] = \emptyset$. Aus $f - f_G \in S$ und $f - f_H \in S$ folgt $(f - f_G) - (f - f_H) \in S$, also $f_H - f_G \in S$. Da sowohl G als auch H eine σ-SAGBI-Basis von S ist, gilt

$$K[\mathrm{LT}_\sigma(G)] = K[\mathrm{LT}_\sigma(f) |\ f \in S \setminus \{0\}] = K[\mathrm{LT}_\sigma(H)].$$

Wegen $\mathrm{Supp}(f_G) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$ und $\mathrm{Supp}(f_H) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$ hat auch $\mathrm{Supp}(f_H - f_G)$ mit $K[\mathrm{LT}_\sigma(G)]$ bzw. $K[\mathrm{LT}_\sigma(H)]$ keine gemeinsamen Elemente, d.h. $f_H - f_G$ ist irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$. Gemäß Theorem 5.2.13 gilt dann $f_H - f_G = 0$, also $f_H = f_G$.

□

Diesem eindeutig bestimmten und von der Wahl der σ-SAGBI-Basis unabhängigen Polynom f_G werden wir nun einen analogen Namen geben.

Definition 5.3.2. (SAGBI-Normalform)
Sei S eine endlich erzeugte K-Unteralgebra von P und sei $f \in P$. Das eindeutig bestimmte und von der Wahl einer σ-SAGBI-Basis unabhängige Polynom aus obigem Satz heißt die **SAGBI-Normalform** oder kurz **Normalform** von f bzgl. S und wird mit $\mathrm{NF}_{\sigma,S}(f)$ oder kurz $\mathrm{NF}_S(f)$ bezeichnet, falls die Termordnung σ aus dem Zusammenhang klar ist.

Anders als bei der Unteralgebra-Ersetzungsregel steht hier „S" im Index von NF_S auch für die jeweils betrachtete K-Unteralgebra S. Besitzt S eine endliche σ-SAGBI-Basis, so ist die Berechnung der Normalform mit Algorithmus 5.2 möglich und im ApCoCoA-Paket `sagbi.cpkg` verfügbar (siehe Anhang C, `SB.NFS`, Seite 302).

Korollar 5.3.3. (Berechnung der SAGBI-Normalform)
Sei S eine endlich erzeugte K-Unteralgebra von P, sei $G = \{g_1, \ldots, g_s\} \subseteq P$ eine endliche σ-SAGBI-Basis von S und sei $\mathcal{G} = (g_1, \ldots, g_s)$. Dann gilt $\mathrm{NF}_S(f) = \mathrm{NRS}_{\mathcal{G}}(f)$ für alle $f \in P$. Insbesondere ist $\mathrm{NRS}_{\mathcal{G}}(f)$ unabhängig von der Reihenfolge des Tupels $\mathcal{G}$.

Beweis: Laut Theorem 5.2.6 gibt es ein $h \in K[y_1, \ldots, y_s]$ mit $f - \mathrm{NRS}_{\mathcal{G}}(f) = h(g_1, \ldots, g_s)$, d.h. es gilt $f - \mathrm{NRS}_{\mathcal{G}}(f) \in S$. Ebenfalls aus Theorem 5.2.6 folgt $\mathrm{Supp}(\mathrm{NRS}_{\mathcal{G}}(f)) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$. Da G eine σ-SAGBI-Basis von S ist, ist $\mathrm{NRS}_{\mathcal{G}}(f)$ das eindeutig bestimmte Polynom f_G aus Satz 5.3.1. Mit anderen Worten gilt $\mathrm{NF}_S(f) = \mathrm{NRS}_{\mathcal{G}}(f)$. □

Die Normalform bzgl. einer Unteralgebra S erfüllt ähnlich der Normalform in der Gröbner Basen Theorie (siehe [KR00], Korollar 2.4.9) verschiedene Rechenregeln (vgl. [KR05], Tutorial 96), die im nachfolgenden Satz bewiesen werden. Die in [KR05], Tutorial 96, angegebene Regel $\mathrm{NF}_S(f_1 f_2) = \mathrm{NF}_S(\mathrm{NF}_S(f_1)\mathrm{NF}_S(f_2))$ gilt allerdings im Allgemeinen nicht, wie das folgende Beispiel zeigt.

Beispiel 5.3.4. Sei $P = \mathbb{Q}[x, y]$, $\sigma = \mathrm{DegLex}$ und seien $g_1 = x + y$ sowie $g_2 = xy$. Sei weiter $G := \{g_1, g_2\}$ sowie $\mathcal{G} := (g_1, g_2)$ und $S := K[G]$. Wie man leicht einsieht, ist G eine σ-SAGBI-Basis von S. Seien $f_1 = x^3 + x^2 y$ und $f_2 = x^2 - xy$. Dann gilt

$$f_1 \xrightarrow{-g_1^3}_{\mathrm{SS}} -2x^2y - 3xy^2 - y^3 \xrightarrow{+2g_1g_2}_{\mathrm{SS}} -xy^2 - y^3$$

und $-xy^2 - y^3$ ist irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$. Somit gilt $\mathrm{NF}_S(f_1) = \mathrm{NRS}_{\mathcal{G}}(f_1) = -xy^2 - y^3$. Analog erhalten wir mit

$$f_2 \xrightarrow{-g_1^2}_{\mathrm{SS}} -3xy - y^2 \xrightarrow{+3g_2}_{\mathrm{SS}} -y^2$$

ein irreduzibles Polynom und es folgt $\mathrm{NF}_S(f_2) = \mathrm{NRS}_{\mathcal{G}}(f_2) = -y^2$. Somit folgt für das Produkt der Normalformen $\mathrm{NF}_S(f_1) \cdot \mathrm{NF}_S(f_2) = xy^4 + y^5$. Wegen $\{xy^4, y^5\} \cap K[\mathrm{LT}_\sigma(g_1), \mathrm{LT}_\sigma(g_2)] = \emptyset$ ist dieses Polynom irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$, d.h. es gilt $\mathrm{NF}_S(\mathrm{NF}_S(f_1) \cdot \mathrm{NF}_S(f_2)) = xy^4 + y^5$. Weiter gilt

$$\begin{aligned} f_1 f_2 = x^5 - x^3y^2 &\xrightarrow{-g_1^5}_{\mathrm{SS}} -5x^4y - 11x^3y^2 - 10x^2y^3 - 5xy^4 - y^5 \\ &\xrightarrow{+5g_1^3g_2}_{\mathrm{SS}} 4x^3y^2 + 5x^2y^3 - y^5 \xrightarrow{-4g_1g_2^2}_{\mathrm{SS}} x^2y^3 - y^5 \end{aligned}$$

Das Polynom $x^2y^3 - y^5$ ist irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$, also gilt $\mathrm{NF}_S(f_1 f_2) = x^2y^3 - y^5$. Damit folgt $\mathrm{NF}_S(f_1 f_2) \neq \mathrm{NF}_S(\mathrm{NF}_S(f_1)\mathrm{NF}_S(f_2))$. ◁

Der folgende Satz beinhaltet neben den beiden in [KR05] erwähnten, korrekten Regeln in Teil c) eine korrigierte Regel für die Normalform des Produkts zweier Polynome.

Satz 5.3.5. (Rechenregeln für die SAGBI-Normalform)
Sei S eine endlich erzeugte K-Unteralgebra von P.

a) Für alle $f \in P$ gilt $\mathrm{NF}_S(\mathrm{NF}_S(f)) = \mathrm{NF}_S(f)$.

b) Für alle $f_1, f_2 \in P$ gilt $\mathrm{NF}_S(f_1 - f_2) = \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2)$.

c) Für alle $f_1, f_2 \in P$ gilt $\mathrm{NF}_S(f_1 f_2) = \mathrm{NF}_S(f_1 \cdot \mathrm{NF}_S(f_2) + f_2 \cdot \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_1) \cdot \mathrm{NF}_S(f_2))$.

Beweis: Sei $G \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S.

a) Sei $f \in P$. Dann sind $\mathrm{NF}_S(f)$ und $\mathrm{NF}_S(\mathrm{NF}_S(f))$ irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$. Somit ist auch $\mathrm{NF}_S(f) - \mathrm{NF}_S(\mathrm{NF}_S(f))$ irreduzibel bzgl. $\xrightarrow{G}_{\mathrm{S}}$. Aus der Definition der Normalform folgt $\mathrm{NF}_S(f) - \mathrm{NF}_S(\mathrm{NF}_S(f)) \in S$ und damit $\mathrm{NF}_S(f) - \mathrm{NF}_S(\mathrm{NF}_S(f)) = 0$ aus Theorem 5.2.13. Folglich gilt $\mathrm{NF}_S(f) = \mathrm{NF}_S(\mathrm{NF}_S(f))$.

b) Seien $f_1, f_2 \in P$. Dann gilt $f_1 - \mathrm{NF}_S(f_1) \in S$ und $f_2 - \mathrm{NF}_S(f_2) \in S$. Wegen

$$(f_1 - f_2) - (\mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2)) = (f_1 - \mathrm{NF}_S(f_1)) - (f_2 - \mathrm{NF}_S(f_2))$$

ist $(f_1-f_2)-(\mathrm{NF}_S(f_1)-\mathrm{NF}_S(f_2)) \in S$. Außerdem gilt $(f_1-f_2)-\mathrm{NF}_S(f_1-f_2) \in S$. Damit folgt $\mathrm{NF}_S(f_1 - f_2) - (\mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2)) \in S$. Da $\mathrm{NF}_S(f_1 - f_2) - (\mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2))$ irreduzibel ist bzgl. $\xrightarrow{G}_{\mathrm{s}}$, folgt die Behauptung erneut aus Theorem 5.2.13.

c) Seien $f_1, f_2 \in P$. Dann gilt $f_1 - \mathrm{NF}_S(f_1) \in S$ und $f_2 - \mathrm{NF}_S(f_2) \in S$. Weiter ist

$$f_1f_2-(f_1 \cdot \mathrm{NF}_S(f_2) + f_2 \cdot \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_1)\mathrm{NF}_S(f_2)) = (f_1-\mathrm{NF}_S(f_1))\cdot(f_2-\mathrm{NF}_S(f_2))$$

ein Element von S. Aus b) folgt

$$\mathrm{NF}_S(f_1f_2) - \mathrm{NF}_S\left(f_1 \cdot \mathrm{NF}_S(f_2) + f_2 \cdot \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_1)\mathrm{NF}_S(f_2)\right) \in S.$$

Da dieses Polynom irreduzibel ist bzgl. $\xrightarrow{G}_{\mathrm{s}}$, gilt gemäß Theorem 5.2.13:

$$\mathrm{NF}_S(f_1f_2) = \mathrm{NF}_S\left(f_1 \cdot \mathrm{NF}_S(f_2) + f_2 \cdot \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_1)\mathrm{NF}_S(f_2)\right).$$

□

Wir wollen diese Regeln, insbesondere die neue Regel für die Normalform des Produkts zweier Polynome, nun an einem Beispiel anwenden. Dazu betrachten wir das bereits bekannte Beispiel 5.3.4.

Beispiel 5.3.6. Sei $P = \mathbb{Q}[x, y]$, $\sigma = \mathrm{DegLex}$ und seien $g_1 = x + y$ sowie $g_2 = xy$. Sei weiter $G := \{g_1, g_2\}$ und $S := K[G]$. Wie man leicht einsieht, ist G eine σ-SAGBI-Basis von S. Seien $f_1 = x^3 + x^2y$ und $f_2 = x^2 - xy$. Zunächst gilt mit $f_1 - f_2 = x^3 + x^2y - x^2 + xy$:

$$\begin{aligned} f_1 - f_2 &\xrightarrow{-g_1^3}_{\mathrm{ss}} -2x^2y - 3xy^2 - y^3 - x^2 + xy \xrightarrow{+2g_1g_2}_{\mathrm{ss}} -xy^2 - y^3 - x^2 + xy \\ &\xrightarrow{+g_1^2}_{\mathrm{ss}} -xy^2 - y^3 + 3xy + y^2 \xrightarrow{-3g_2}_{\mathrm{ss}} -xy^2 - y^3 + y^2, \end{aligned}$$

also $\mathrm{NF}_S(f_1-f_2) = -xy^2-y^3+y^2 = \mathrm{NF}_S(f_1)-\mathrm{NF}_S(f_2)$. Aus Beispiel 5.3.4 ist bereits bekannt, dass $\mathrm{NF}_S(f_1f_2) = x^2y^3 - y^5$ gilt. Weiter gilt:

$$\begin{aligned} h &:= f_1 \cdot \mathrm{NF}_S(f_2) + f_2 \cdot \mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_1) \cdot \mathrm{NF}_S(f_2) \\ &= (x^3 + x^2y) \cdot (-y^2) + (x^2 - xy) \cdot (-xy^2 - y^3) - (-y^2) \cdot (-xy^2 - y^3) \\ &= -x^3y^2 - x^2y^3 - x^3y^2 - x^2y^3 + x^2y^3 + xy^4 - xy^4 - y^5 \\ &= -2x^3y^2 - x^2y^3 - y^5 \end{aligned}$$

und $h \xrightarrow{+2g_1g_2^2}_{\mathrm{ss}} x^2y^3 - y^5$. Somit gilt $\mathrm{NF}_S(f_1f_2) = \mathrm{NF}_S(h)$. ◁

Mit Hilfe der Normalform ist es nun möglich, in Analogie zur Gröbner Basen Theorie (vgl. Satz 2.2.1) auch für Unteralgebren einen Unteralgebra-Mitgliedschaftstest anzugeben, im Gegensatz zu Korollar 5.2.9 ist dieser nun auch hinreichend.

Satz 5.3.7. (Unteralgebra Mitgliedschaftstest)
Seien $S = K[g_1, \ldots, g_s]$ und $T = K[h_1, \ldots, h_t]$ endlich erzeugte K-Unteralgebren von P mit Polynomen $g_1, \ldots, g_s, h_1, \ldots, h_t \in P \setminus \{0\}$.

a) Für alle $f_1, f_2 \in P$ gilt genau dann $f_1 - f_2 \in S$, wenn $\mathrm{NF}_S(f_1) = \mathrm{NF}_S(f_2)$ gilt. Insbesondere gilt für ein Polynom $f \in P$ genau dann $f \in S$, wenn $\mathrm{NF}_S(f) = 0$ gilt.

b) Genau dann gilt $T \subseteq S$, wenn $\mathrm{NF}_S(h_j) = 0$ für alle $j = 1, \ldots, t$ gilt.

c) Genau dann gilt $T = S$, wenn $\mathrm{NF}_T(g_i) = 0$ für alle $i \in \{1, \ldots, s\}$ und $\mathrm{NF}_S(h_j) = 0$ für alle $j \in \{1, \ldots, t\}$ gilt.

Beweis:

a) Sei $G \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S. Seien $f_1, f_2 \in P$ und sei zunächst $f_1 - f_2 \in S$. Wegen $(f_1 - f_2) - \mathrm{NF}_S(f_1 - f_2) \in S$ folgt $\mathrm{NF}_S(f_1 - f_2) \in S$. Mit Satz 5.3.5 b) folgt $\mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2) \in S$. Da $\mathrm{NF}_S(f_1) - \mathrm{NF}_S(f_2)$ irreduzibel ist bzgl. $\xrightarrow{G}_{\mathrm{s}}$, gilt somit $\mathrm{NF}_S(f_1) = \mathrm{NF}_S(f_2)$ gemäß Theorem 5.2.13.

Gelte nun $\mathrm{NF}_S(f_1) = \mathrm{NF}_S(f_2)$. Dann gilt:

$$f_1 - f_2 = (f_1 - \mathrm{NF}_S(f_1)) - (f_2 - \mathrm{NF}_S(f_1)) = (f_1 - \mathrm{NF}_S(f_1)) - (f_2 - \mathrm{NF}_S(f_2)).$$

Wegen $f_1 - \mathrm{NF}_S(f_1),\ f_2 - \mathrm{NF}_S(f_2) \in S$ folgt $f_1 - f_2 \in S$. Die Zusatzbehauptung folgt sofort mit $f_1 := f$ und $f_2 := 0$.

b) Ist T eine K-Unteralgebra von S, so gilt insbesondere $h_1, \ldots, h_t \in S$. Aus a) folgt damit $\mathrm{NF}_S(h_j) = 0$ für alle $j = 1, \ldots, t$.

Gelte nun $\mathrm{NF}_S(h_j) = 0$ für alle $j = 1, \ldots, t$. Erneut aus a) folgt $h_1, \ldots, h_t \in S$. Damit ist klar, dass jedes Polynom $f \in T = K[h_1, \ldots, h_t]$ ein Polynom aus S ist, d.h. T ist eine K-Unteralgebra von S.

c) Dass sowohl $\mathrm{NF}_T(g_i) = 0$ für alle $i \in \{1, \ldots, s\}$ als auch $\mathrm{NF}_S(h_j) = 0$ für alle $j \in \{1, \ldots, t\}$ gilt, folgt mit a) sofort aus $T = S$.

Umgekehrt folgt aus $\mathrm{NF}_S(h_j) = 0$ für alle $j \in \{1, \ldots, t\}$ unter Verwendung von Teil b), dass T eine K-Unteralgebra von S ist, und aus $\mathrm{NF}_T(g_i) = 0$ für alle $i \in \{1, \ldots, s\}$, dass S eine K-Unteralgebra von T ist. Somit gilt $T = S$.

□

Besitzt die K-Unteralgebra S eine endliche SAGBI-Basis, so ist dieser Mitgliedschaftstest auch effizient umsetzbar. Er bildet die CoCoA-Funktion `SB.IsInSubalg` des Pakets `sagbi.cpkg` (siehe Anhang C, Seite 299). Mathematisch festhalten wollen wir das im folgenden Korollar.

Korollar 5.3.8. *Sei $S \subseteq P$ eine endlich erzeugte K-Unteralgebra, sei $\mathcal{G} = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S und sei $\mathcal{G} = (g_1, \ldots, g_s)$. Für alle $f \in P$ gilt genau dann $f \in S$, wenn $\mathrm{NF}_S = \mathrm{NRS}_{\mathcal{G}}(f) = 0$ gilt.*

Beweis: Folgt sofort aus dem letzten Satz. □

Außerdem ist es mit Hilfe des Unteralgebra-Divisionsalgorithmus möglich eine explizite Darstellung zu erhalten, falls ein Polynom Element von S ist. Die Funktion `SB.SubalgRepr` liefert genau diese Darstellung (siehe Anhang C, Seite 304).

Bemerkung 5.3.9. (Explizite Darstellung)
Sei $S \subseteq P$ eine endlich erzeugte K-Unteralgebra, sei $G = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S und sei $f \in P \setminus \{0\}$. Laut Theorem 5.2.6 berechnet Algorithmus 5.2 für $f \neq 0$ ein Polynom $h \in K[y_1, \ldots, y_s]$ mit $f = h(g_1, \ldots, g_s) + \mathrm{NF}_S(f)$. Gilt also $f \in S$, folgt $\mathrm{NF}_S(f) = 0$ und wir erhalten mit h eine explizite Darstellung von f in den Polynomen $g_1, \ldots, g_s$.

Mit Hilfe der Normalform lässt sich noch ein weiterer Satz über die Identität von Unteralgebren beweisen, der uns später von Nutzen sein wird. Eine ähnliche Aussage findet sich auch in [Ric89] (vgl. [Ric89], Satz 1).

Satz 5.3.10. *Sei T eine beliebige K-Unteralgebra von P und S eine endlich erzeugte K-Unteralgebra von P mit $S \subseteq T$. Gilt $K[\mathrm{LT}_\sigma(f) : \ f \in T \setminus \{0\}] \subseteq K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}]$, so folgt $S = T$.*

Beweis: Angenommen, es gelte $T \setminus S \neq \emptyset$. Dann gilt insbesondere $0 \notin T \setminus S$. Weiter ist auch die Menge $\{\mathrm{LT}_\sigma(f) : \ f \in T \setminus S\}$ nicht leer und besitzt ein bzgl. σ minimales Element. Sei $f_0 \in T \setminus S$ so, dass $\mathrm{LT}_\sigma(f_0)$ dieser minimale Term ist. Da $K[\mathrm{LT}_\sigma(f) : f \in T \setminus \{0\}]$ Teilmenge von $K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}]$ ist, gibt es ein $\tilde{f} \in S \setminus \{0\}$ mit $\mathrm{LT}_\sigma(f_0) = \mathrm{LT}_\sigma(\tilde{f})$ und weiter gibt es ein $c \in K$ mit $\mathrm{LM}_\sigma(f_0) = \mathrm{LM}_\sigma(c \cdot \tilde{f})$.

Gilt $f_0 = c \cdot \tilde{f}$, so folgt $f_0 \in S$ im Widerspruch zur Wahl von f_0. Sei also $f_0 \neq c \cdot \tilde{f}$. Dann gilt $f_0 - c \cdot \tilde{f} \in T \setminus S$. Denn wäre $f_0 - c \cdot \tilde{f} \in S$, so würde $0 = \mathrm{NF}_S(f_0 - c \cdot \tilde{f}) = \mathrm{NF}_S(f_0) - \mathrm{NF}_S(c \cdot \tilde{f})$ gelten. Wegen $\tilde{f} \in S$ würde dann auch $\mathrm{NF}_S(f_0) = 0$, also $f_0 \in S$, folgen. Somit ist $\mathrm{LT}_\sigma(f_0 - c \cdot \tilde{f})$ ein Element der Menge $\{\mathrm{LT}_\sigma(f) : \ f \in T \setminus S\}$. Nun gilt aber $\mathrm{LT}_\sigma(f_0 - c \cdot h) <_\sigma \mathrm{LT}_\sigma(f_0)$ im Widerspruch zur Minimalität von $\mathrm{LT}_\sigma(f_0)$. Damit folgt $T = S$. □

Nach dem Unteralgebra-Mitgliedschaftstest sowie der Identität von Unteralgebren unter bestimmten Voraussetzungen wollen wir analog zur Welt der Gröbner Basen auch hier der Frage nachgehen, ob eine endlich erzeugte K-Unteralgebra von P eine eindeutig bestimmte σ-SAGBI-Basis besitzt. Diese σ-SAGBI-Basis nimmt dann die Rolle einer reduzierten Gröbner Basis in der Welt der Unteralgebren ein. Wir legen nun den Begriff einer reduzierten σ-SAGBI-Basis fest, der erstmals von Lorenzo Robbiano und Moss Sweedler 1998 eingeführt wurde (vgl. [RS98]).

Definition 5.3.11. (Reduzierte σ-SAGBI-Basis)
Sei S eine endlich erzeugte K-Unteralgebra von P. Eine σ-SAGBI-Basis $G \subseteq P \setminus \{0\}$ von S heißt **reduzierte σ-SAGBI-Basis** von S, falls gilt:

(i) Für alle $g \in G$ gilt $\mathrm{LC}_\sigma(g) = 1$.

(ii) Die Menge $\mathrm{LT}_\sigma(G)$ der Leitterme aus G bildet das minimale monomiale Erzeugendensystem von $K[\mathrm{LT}_\sigma(G)] = K[\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}])$ als K-Algebra.

(iii) Für alle $g \in G$ gilt $\mathrm{Supp}(g - \mathrm{LT}_\sigma(g)) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$.

Es bleibt natürlich die Frage zu klären, ob so eine reduzierte σ-SAGBI-Basis überhaupt existiert. Das folgende Theorem wird uns dies nicht nur bestätigen, sondern auch zeigen, dass eine reduzierte σ-SAGBI-Basis eindeutig bestimmt ist. Ein Beweis dieses Theorems findet sich erstmals in [RS98], aber wir werden hier einen neuen Beweis liefern.

Theorem 5.3.12. (Existenz und Eindeutigkeit der reduzierten σ-SAGBI-Basis)
Jede endlich erzeugte K-Unteralgebra S von P besitzt eine eindeutig bestimmte reduzierte σ-SAGBI-Basis.

Beweis: Sei $G \subseteq P \setminus \{0\}$ eine σ-SAGBI-Basis von S und ohne Einschränkung seien alle Elemente von G bereits normiert. Wir zeigen zuerst die Existenz einer reduzierten σ-SAGBI-Basis. Aus Korollar 5.1.5 folgt, dass es ein eindeutig bestimmtes minimales Algebra-Erzeugendensystem von $K[\mathrm{LT}_\sigma(f) | \ f \in S \setminus \{0\}]$ gibt, das aus Leittermen von Polynomen aus $S \setminus \{0\}$ besteht.

Da G eine σ-SAGBI-Basis von S ist, wird das Monoid $\{\mathrm{LT}_\sigma(f)|\ f \in S \setminus \{0\}\}$ von der Menge $\mathrm{LT}_\sigma(G) = \{\mathrm{LT}_\sigma(g)|\ g \in G\}$ erzeugt (vgl. Korollar 5.1.5). Somit ist das minimale monomiale Algebra-Erzeugendensystem eine Teilmenge von $\mathrm{LT}_\sigma(G)$. Sei $G' \subseteq G$ diejenige Teilmenge von G so, dass $\mathrm{LT}_\sigma(G')$ das minimale monomiale Algebra-Erzeugendensystem der K-Unteralgebra $K[\mathrm{LT}_\sigma(f) \mid f \in S \setminus \{0\}]$ ist. Dann folgt erneut aus Korollar 5.1.5, dass auch G' eine σ-SAGBI-Basis von S ist. Die σ-SAGBI-Basis G' erfüllt bereits die Axiome (i) und (ii) der obigen Definition.

Da die Elemente von G' normiert sind, können wir jedes $g \in G'$ in der Form $g = \mathrm{LT}_\sigma(g) + \hat{g}$ mit $\hat{g} \in P$ schreiben. Setze $H := \{\mathrm{LT}_\sigma(g) + \mathrm{NF}_S(\hat{g})|\ g \in G'\}$. Wir zeigen nun, dass H eine reduzierte σ-SAGBI-Basis von S ist. Dazu ist zunächst zu zeigen, dass H eine σ-SAGBI-Basis von S ist. Sei $h \in H$. Insbesondere gilt $h \neq 0$. Dann gibt es ein $g \in G'$ mit $h = \mathrm{LT}_\sigma(g) + \mathrm{NF}_S(\hat{g})$. Wegen $\hat{g} - \mathrm{NF}_S(\hat{g}) \in S$ und $g \in S$ gilt

$$h = \mathrm{LT}_\sigma(g) + \mathrm{NF}_S(\hat{g}) = g - \hat{g} + \mathrm{NF}_S(\hat{g}) = g - (\hat{g} - \mathrm{NF}_S(\hat{g})) \in S.$$

Somit erhalten wir $H \subseteq S \setminus \{0\}$. Wegen $\mathrm{LT}_\sigma(h) = \mathrm{LT}_\sigma(g)$ folgt aus Korollar 5.1.5, dass auch H eine σ-SAGBI-Basis von S ist. Außerdem sind offensichtlich die Bedingungen (i) und (ii) der Definition einer reduzierten σ-SAGBI-Basis bereits erfüllt. Gemäß Definition 5.3.2 gilt $\mathrm{Supp}(\mathrm{NF}_S(\hat{g})) \cap K[\mathrm{LT}_\sigma(H)] = \emptyset$. Wegen $\mathrm{NF}_S(\hat{g}) = h - \mathrm{LT}_\sigma(h)$ folgt also $\mathrm{Supp}(h - \mathrm{LT}_\sigma(h)) \cap K[\mathrm{LT}_\sigma(H)] = \emptyset$. Somit erfüllt H auch Bedingung (iii) der Definition einer reduzierten σ-SAGBI-Basis, womit die Existenz einer reduzierten σ-SAGBI-Basis gezeigt ist.

Es bleibt die Eindeutigkeit zu zeigen. Seien G und H reduzierte σ-SAGBI-Basen von S. Da G und H insbesondere σ-SAGBI-Basen von S sind, gilt per Definition:

$$K[\mathrm{LT}_\sigma(G)] = K[\mathrm{LT}_\sigma(f) \mid f \in S \setminus \{0\}] = K[\mathrm{LT}_\sigma(H)].$$

Gemäß Korollar 5.1.5 ist das minimale monomiale Algebra-Erzeugendensystem von $K[\mathrm{LT}_\sigma(G)]$ bzw. $K[\mathrm{LT}_\sigma(H)]$ eindeutig bestimmt. Da G und H reduzierte σ-SAGBI-Basen sind, folgt somit $\mathrm{LT}_\sigma(G) = \mathrm{LT}_\sigma(H)$.

Sei nun $g \in G$. Dann gibt es ein $h \in H$ mit $\mathrm{LT}_\sigma(g) = \mathrm{LT}_\sigma(h)$. Insbesondere gilt $g, h \in S$, womit $g - h \in S$ folgt. Aufgrund der Eigenschaften einer reduzierten σ-SAGBI-Basis erhalten wir $\mathrm{Supp}(g - \mathrm{LT}_\sigma(g)) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$ und $\mathrm{Supp}(h - \mathrm{LT}_\sigma(h)) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$. Wegen $g - h = (g - \mathrm{LT}_\sigma(g)) - (h - \mathrm{LT}_\sigma(h))$ gilt $\mathrm{Supp}(g - h) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$, d.h. $g - h$ ist irreduzibel bzgl. $\xrightarrow{G}_S$. Laut Theorem 5.2.13 folgt $g - h = 0$, also $g = h$ und damit $g \in H$. Analog zeigt man $h \in G$ für alle $h \in H$. Somit gilt $G = H$, womit auch die Eindeutigkeit bewiesen ist. □

Da S bekanntlich genau dann eine endliche σ-SAGBI-Basis besitzt, wenn das minimale monomiale Algebra-Erzeugendensystem von $K[\mathrm{LT}_\sigma(f)|\ f \in S \setminus \{0\}]$ endlich ist, ist die reduzierte σ-SAGBI-Basis von S genau dann endlich, wenn S eine endliche σ-SAGBI-Basis besitzt. Im ApCoCoA-Paket `sagbi.cpkg` ist eine Funktion `SB.IsReducedSagbi` enthalten, die für eine gegebene σ-SAGBI-Basis überprüft, ob sie eine reduzierte σ-SAGBI-Basis ist (siehe Anhang C, Seite 300). Wir wollen nun ein Beispiel einer reduzierten σ-SAGBI-Basis angeben, bevor wir uns der Berechnung dieser widmen werden.

Beispiel 5.3.13. Sei $P = \mathbb{Q}[x_1, x_2]$ und $\sigma = \mathrm{DegLex}$. Sei $G = \{g_1, \ldots, g_4\}$ mit den Polynomen $g_1 = x_1^2 - x_2^2$, $g_2 = x_1^2 x_2$, $g_3 = x_1^2 x_2^2 - x_2^4$ und $g_4 = x_1^2 x_2^4$ sowie $S = \mathbb{Q}[G]$. Sei weiter $g_5 = x_2^6$ und $B := G \cup \{g_5\}$ sowie $\mathcal{B} := (g_1, \ldots, g_5)$. Dann ist B eine reduzierte σ-SAGBI-Basis von S. ◁

Besitzt eine endlich erzeugte K-Unteralgebra S von P eine endliche σ-SAGBI-Basis G, so berechnet der folgende Algorithmus eine endliche reduzierte σ-SAGBI-Basis aus der „normalen" endlichen σ-SAGBI-Basis.

Algorithmus 5.3 : Berechnung einer reduzierten SAGBI-Basis

Input : $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$: σ-SAGBI-Basis von S.

Result : Eine reduzierte σ-SAGBI-Basis G von S.

```
1  G := ∅, G' := ∅, L := (LT_σ(g_1), ..., LT_σ(g_s)) ∈ P^s;
2  for i = 1 → s do
3      g_i := 1/LC_σ(g_i) · g_i;
4  if s = 1 then
5      return {g_1};
6  for i = 1 → s do
7      H := (LT_σ(g_1), ..., LT_σ(g_{i-1}), LT_σ(g_{i+1}), ..., LT_σ(g_s)) ∈ P^{s-1};
8      t := NRS_H(LT_σ(g_i));
9      if t ≠ 0 then
10         G' := G' ∪ {g_i};
11 foreach g' ∈ G' do
12     h := NF_S(g' − LT_σ(g'));
13     g := LT_σ(g') + h;
14     G := G ∪ {g};
15 return G;
```

Auch dieser Algorithmus ist Teil des ApCoCoA-Pakets `sagbi.cpkg`. Die zugehörige Funktion ist `SB.ReducedSagbi` (siehe Anhang C, Seite 303). Folgendes Beispiel soll nun zunächst den Ablauf von Algorithmus 5.3 demonstrieren, bevor wir dessen Korrektheit beweisen werden. Das Ergebnis, das uns dieser Algorithmus liefert, ist dieselbe reduzierte σ-SAGBI-Basis, die uns aus Beispiel 5.3.13 bereits bekannt ist und von der wir bereits nachgewiesen haben, dass sie tatsächlich eine reduzierte σ-SAGBI-Basis ist.

Beispiel 5.3.14. Sei $P = \mathbb{Q}[x_1, x_2]$ und $\sigma = \text{DegLex}$. Sei $S = \mathbb{Q}[g_1, \ldots, g_4]$ mit den Polynomen $g_1 = x_1^2 - x_2^2$, $g_2 = x_1^2 x_2$, $g_3 = x_1^2 x_2^2 - x_2^4$ und $g_4 = x_1^2 x_2^4$. Seien weiter $g_5 = x_2^6$ und $g_6 = x_1^2 x_2^6 - x_2^8$. Dann ist bekanntlich $\{g_1, \ldots, g_6\}$ eine σ-SAGBI-Basis von S.

Um den Ablauf des Algorithmus besser verdeutlichen zu können, ersetzen wir den Erzeuger g_3 durch $g_3' := g_3 + g_2$. Dann ist natürlich auch $\{g_1, g_2, g_3', g_4, g_5, g_6\}$ eine σ-SAGBI-Basis von S. Mit dieser σ-SAGBI-Basis starten wir nun den Algorithmus, wobei wir zur Vereinfachung wieder g_3 anstatt g_3' schreiben.

Zu Beginn gilt $G = \emptyset$, $G' = \emptyset$ und $\mathcal{L} = (x_1^2,\ x_1^2 x_2,\ x_1^2 x_2^2,\ x_1^2 x_2^4,\ x_2^6,\ x_1^2 x_2^6)$. Wir setzen zur Abkürzung $\mathcal{H}_i$ für alle $i \in \{1, \ldots, 6\}$ für das Tupel, das aus $\mathcal{L}$ durch Streichen der i-ten Komponente entsteht. Dann gilt:

$$\begin{aligned} &\text{NRS}_{\mathcal{H}_1}(\text{LT}_\sigma(g_1)) = x_1^2, && \text{NRS}_{\mathcal{H}_2}(\text{LT}_\sigma(g_2)) = x_1^2 x_2, && \text{NRS}_{\mathcal{H}_3}(\text{LT}_\sigma(g_3)) = x_1^2 x_2^2 \\ &\text{NRS}_{\mathcal{H}_4}(\text{LT}_\sigma(g_4)) = x_1^2 x_2^4, && \text{NRS}_{\mathcal{H}_5}(\text{LT}_\sigma(g_5)) = x_2^6, && \text{NRS}_{\mathcal{H}_6}(\text{LT}_\sigma(g_6)) = 0. \end{aligned}$$

Somit ist $\text{LT}_\sigma(g_6)$ der einzige Leitterm, der sich als Produkt der übrigen Leitterme darstellen lässt. Nach Ende der ersten Schleife gilt also $G' = \{g_1, \ldots, g_5\}$. Da g_2, g_4 und g_5 Terme sind, folgt sofort $g_2, g_4, g_5 \in G$. Für g_1 gilt $h = \text{NF}_S(g_1 - \text{LT}_\sigma(g_1)) = \text{NF}_S(-x_2^2) = -x_2^2$ und es wird $\text{LT}_\sigma(g_1) + h = g_1$ in G eingefügt. Weiter gilt $h = \text{NF}_S(g_3 - \text{LT}_\sigma(g_3)) = \text{NF}_S(-x_2^4 + x_1^2 x_2) = -x_2^4$, d.h. $\text{LT}_\sigma(g_3) + h = x_1^2 x_2 - x_2^4$ wird in G eingefügt. Somit erhalten wir

$$G = \{x_1^2 - x_2^2,\ x_1^2 x_2,\ x_1^2 x_2 - x_2^4,\ x_1^2 x_2^4,\ x_2^6\}$$

als reduzierte σ-SAGBI-Basis von S, wie wir in Beispiel 5.3.13 ebenfalls bewiesen hatten. ◁

Nachdem wir uns mit dem Ablauf des Algorithmus vertraut gemacht haben, wollen wir nun dessen Korrektheit und Endlichkeit beweisen.

Satz 5.3.15. *Sei σ eine Termordnung auf $\mathbb{T}^n$ und sei $\{g_1, \ldots, g_s\} \subseteq P \backslash \{0\}$ eine σ-SAGBI-Basis einer K-Unteralgebra S von P. Dann berechnet Algorithmus 5.3 in endlich vielen Schritten eine reduzierte σ-SAGBI-Basis G von S.*

Beweis: Dass Algorithmus 5.3 nach endlich vielen Schritten terminiert, ist aufgrund der Endlichkeit sämtlicher Mengen und der Tatsache, dass nur for-Schleifen vorhanden sind, sofort klar.

Nach der ersten for-Schleife sind alle Polynome der σ-SAGBI-Basis normiert. Besteht die σ-SAGBI-Basis nur aus einem Polynom, sind wir damit offensichtlich bereits fertig. Andernfalls betrachten wir nun die Schleife zwischen den Zeilen 6 bis 10. Wie wir wissen, ist eine Menge von Termen immer eine σ-SAGBI-Basis derjenigen K-Unteralgebra, die diese Terme erzeugen. Somit ist der Term $t \in \mathbb{T}^n$ eindeutig bestimmt. Gilt $t = 0$, so ist $\mathrm{LT}_\sigma(g_i)$ für ein $i \in \{1, \ldots, s\}$ ein Produkt der übrigen Leitterme in $\mathcal{L}$. Nach Ende dieser Schleife bilden die Leitterme der Elemente in G' also ein minimales monomiales Algebra-Erzeugendensystem von $K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)]$. Die Menge G' ist somit offensichtlich ebenfalls eine σ-SAGBI-Basis von S wegen

$$K[\mathrm{LT}_\sigma(f) |\ f \in S \setminus \{0\}] = K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)] = K[\mathrm{LT}_\sigma(G')].$$

Die Polynome in G' erfüllen damit bereits die ersten beiden Bedingungen aus Definition 5.3.11.

Nun zur dritten Schleife in den Zeilen 11 bis 14. Sei $g' \in G'$ und sei $h = \mathrm{NF}_S(g' - \mathrm{LT}_\sigma(g'))$. Aus der Definition der Normalform folgt sofort $\mathrm{Supp}(h) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$. Setze $g := \mathrm{LT}_\sigma(g') + h$. Wegen $\mathrm{LT}_\sigma(g) = \mathrm{LT}_\sigma(g')$ gilt $h = g - \mathrm{LT}_\sigma(g)$ und es folgt $\mathrm{Supp}(g - \mathrm{LT}_\sigma(g)) \cap K[\mathrm{LT}_\sigma(G)] = \emptyset$. Somit liefert die zweite Schleife eine Menge G, die neben den ersten beiden Bedingungen aus Definition 5.3.11 auch die dritte Bedingung erfüllt. Natürlich ist auch G nach wie vor eine σ-SAGBI-Basis von S. Insgesamt ist die Menge G nach Ende der zweiten Schleife folglich eine reduzierte σ-SAGBI-Basis von S. □

5.4 Berechnung von SAGBI-Basen

In diesem Abschnitt wollen wir nun auf die Berechnung von SAGBI-Basen eingehen. Da, wie wir wissen, endliche SAGBI-Basen nicht immer existieren, kann es leider auch keinen Algorithmus zur Berechnung von σ-SAGBI-Basen geben. Allerdings lässt sich eine sogenannte **enumerierte Prozedur** angeben, die im Falle der Existenz einer endlichen SAGBI-Basis diese auch berechnet. Diese Prozedur wollen wir zu Beginn dieses Abschnitts beleuchten. Eine Implementation dieser Prozedur war Teil dieser Arbeit und ist im Paket `sagbi.cpkg` des Computer-Algebra-Systems ApCoCoA vorhanden (siehe Anhang C).

Anschließend werden wir homogene und Grad-beschränkte SAGBI-Basen betrachten. Beide werden am Ende von [RS90] nur sehr kurz erwähnt. In [KR05] als Tutorial bzw. in [RS90] sind zwar Verfahren zur Berechnung von homogenen bzw. Grad-beschränkten SAGBI-Basen angegeben, wir wollen hier aber zeigen, wie sich diese Verfahren – im Falle homogener SAGBI-Basen durch eine Prozedur, im Falle von Grad-beschränkten SAGBI-Basen durch einen Algorithmus – im Detail umsetzen lassen. Die Korrektheit beider Verfahren werden wir ebenfalls beweisen, fehlt sie doch in [KR05] gänzlich und ist in [RS90] nur zum Teil vorhanden. Außerdem werden wir die Analogie zur Theorie der Gröbner-Basen ausbauen, indem wir unter anderem Charakterisierungen von Grad-beschränkten SAGBI-Basen angeben und beweisen.

5.4.1 Die SAGBI-Prozedur

Sei wie bisher K ein Körper, $P = K[x_1, \ldots, x_n]$ der Polynomring über K in den Unbestimmten $x_1, \ldots, x_n$ und sei σ eine Termordnung auf $\mathbb{T}^n$. Sei außerdem S eine endlich erzeugte K-Unteralgebra von P. Wir schreiben anstatt $K[y_1, \ldots, y_s]$ im Folgenden erneut auch nur kurz P'. Auch in diesem Abschnitt werden wir versuchen, eine Analogie zur Gröbner-Basen-Theorie aufzubauen, und deshalb zunächst nach einem Analogon zum Buchberger-Kriterium suchen. Dazu müssen wir zuerst angeben, wie in der Welt der Unteralgebren die bekannten S-Polynome aussehen könnten. Da wir gleich sehen werden, dass diese Elemente eines torischen Ideals sind, werden sie zur Abgrenzung **T-Polynome** genannt (vgl. [KR05], Definition 6.6.26, S. 494).

Definition 5.4.1. (T-Polynome)
Sei $\mathcal{G} = (g_1, \ldots, g_s)$ ein Tupel normierter Polynome aus $P \setminus \{0\}$ und sei $b \in \mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G})) \subseteq P'$ ein echtes Binom. Dann heißt das Polynom $b(g_1, \ldots, g_s) \in P$ das **T-Polynom** von b.

Nun lässt sich ein zum Buchberger-Kriterium (vgl. [KR00], Korollar 2.5.3) analoges Kriterium angeben, das als SAGBI-Basis Kriterium bezeichnet wird (vgl. [KR05], Satz 6.6.28, S. 495). Mit diesem Kriterium erhalten wir ein effizientes Werkzeug, um zu untersuchen, ob eine Menge $G \subseteq P \setminus \{0\}$ normierter Polynome eine σ-SAGBI-Basis der von G erzeugten K-Unteralgebra ist.

Satz 5.4.2. (SAGBI-Basis Kriterium)
Sei $S \subseteq P$ eine endlich erzeugte K-Unteralgebra und sei $G \subseteq P \setminus \{0\}$ eine (nicht notwendigerweise endliche) Menge normierter Polynome mit $S = K[G]$. Dann sind die folgenden Aussagen äquivalent:

(i) Die Menge G ist eine σ-SAGBI-Basis von S.

(ii) Für jedes Tupel $\mathcal{G} = (g_1, \ldots, g_s) \in G^s$ gibt es eine Menge $B \subseteq K[y_1, \ldots, y_s]$ von echten Binomen, die $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ erzeugt und die Eigenschaft besitzt, dass $b(g_1, \ldots, g_s) \xrightarrow{G}_S 0$ gilt für alle $b \in B$.

Insbesondere lässt sich daraus auf einfache Weise ein effizienter Test für endliche SAGBI-Basen ableiten. Dieser Test ist in der CoCoA-Funktion `SB.IsSagbi` des ApCoCoA-Pakets `sagbi.cpkg` implementiert (siehe Anhang C, Seite 300).

Algorithmus 5.4 : Endlicher SAGBI-Test

Input : Eine Menge $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ normierter, nicht konstanter Polynome.
Result : TRUE, falls $g_1, \ldots, g_s$ eine σ-SAGBI-Basis bilden.

```
1 G := (g_1, ..., g_s);
2 Berechne mit Algorithmus 2.2 ein Erzeugendensystem B ⊆ K[y_1, ..., y_s] von Rel(LT_σ(G));
3 foreach b ∈ B do
4     if NRS_G(b(g_1, ..., g_s)) ≠ 0 then
5         return FALSE;
6 return TRUE;
```

Die Korrektheit des Algorithmus folgt aus dem folgenden Korollar.

Korollar 5.4.3. (SAGBI-Test)
Sei $S \subseteq P$ eine endlich erzeugte K-Unteralgebra und sei $G \subseteq P \setminus \{0\}$ eine endliche Menge

normierter Polynome mit $S = K[G]$. Dann überprüft Algorithmus 5.4, ob G eine endliche σ-SAGBI-Basis von S ist.

Beweis: Sei $G = \{g_1, \ldots, g_s\}$ und $\mathcal{G} = (g_1, \ldots, g_s)$. Ohne Einschränkung können wir $g_1, \ldots, g_s$ als nicht konstant annehmen. Zunächst zeigen wir, dass es genügt, nur das Tupel $\mathcal{G}$ im SAGBI-Kriterium zu betrachten. Sei $\mathcal{H} = (g_{i_1}, \ldots, g_{i_r})$ mit $1 \leq i_1 < \ldots < i_r \leq s$ für $1 \leq r \leq s$ ein beliebiges Tupel mit Einträgen aus G und sei $B' \subseteq K[y_{i_1}, \ldots, y_{i_r}]$ ein Erzeugendensystem von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{H}))$. Dann ist jede Relation $b' \in B'$ auch in $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ enthalten. Ist $\mathcal{H}'$ eine Permutation von $\mathcal{H}$, so bleiben die Relationen an sich unberührt, nur die Unbestimmten y_i werden umbenannt. Somit ist es ausreichend, nur das Tupel $\mathcal{G}$ zu betrachten.

Damit folgt die Behauptung sofort aus dem SAGBI-Basis Kriterium. Gilt insbesondere explizit $B = \emptyset$, so ist $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ das Nullideal, d.h. die Leitterme $\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)$ sind algebraisch unabhängig. Dann ist G laut Satz 5.1.9 eine endliche σ-SAGBI-Basis von S. □

Mit Hilfe des SAGBI-Basis Kriteriums ist es nun möglich, eine Prozedur anzugeben, die im Falle der Existenz einer endlichen σ-SAGBI-Basis nach endlichen vielen Schritten terminiert. Diese Prozedur ist in der Funktion `SB.Sagbi` des ApCoCoA-Pakets `sagbi.cpkg` implementiert (siehe Anhang C, Seite 303).

Prozedur SAGBI

Input : $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$: Endliche Menge normierter Polynome mit $S = K[g_1, \ldots, g_s]$.

Result : Eine endliche σ-SAGBI-Basis G von S, falls diese existiert.

1 $s' := s$, $G := \{g_1, \ldots, g_s\}$, $\mathcal{G} := (g_1, \ldots, g_s)$;
2 Berechne mit Algorithmus 2.2 ein Erzeugendensystem $B \subseteq K[y_1, \ldots, y_{s'}]$ von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$;
3 **while** $B \neq \emptyset$ **do**
4 $B' := \emptyset$;
5 **foreach** $b \in B$ **do**
6 $b' := \mathrm{NRS}_{\mathcal{G}}(b(g_1, \ldots, g_{s'}))$;
7 **if** $b' \neq 0$ **then**
8 $B' := B' \cup \left\{ \frac{1}{\mathrm{LC}_\sigma(b')} \cdot b' \right\}$;
9 **if** $B' = \emptyset$ **then**
10 **return** G;
11 $t := \#B'$, $s' := s' + t$;
  `// Die Elemente von` B' `werden mit` $g_{s'-t+1}, \ldots, g_{s'}$ `bezeichnet`
12 $G := G \cup \{g_{s'-t+1}, \ldots, g_{s'}\}$, $\mathcal{G} := \mathcal{G} \oplus (g_{s'-t+1}, \ldots, g_{s'})$;
13 Berechne mit Algorithmus 2.2 ein Erzeugendensystem $B \subseteq K[y_1, \ldots, y_{s'-t}, y_{s'-t+1}, \ldots, y_{s'}]$ von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$;
14 Berechne mit Algorithmus 5.5 die Teilmenge $\widehat{B} \subseteq B$, deren Binome mindestens eine Unbestimmte aus $\{y_{s'-t+1}, \ldots, y_{s'}\}$ enthalten;
15 $B := \widehat{B}$;
16 **return** G;

Die Korrektheit dieser Prozedur und im Falle der Existenz einer endlichen σ-SAGBI-Basis auch die Endlichkeit folgen aus folgendem Theorem (vgl. [KR05], Theorem 6.6.29).

Theorem 5.4.4. (Die SAGBI-Prozedur)
Sei σ eine Termordnung auf $\mathbb{T}^n$ und seien $g_1, \ldots, g_s \in P \setminus \{0\}$ normierte Polynome, die die K-Unteralgebra S erzeugen, also mit $S = K[g_1, \ldots, g_s]$. Die Prozedur SAGBI enumeriert eine σ-SAGBI-Basis G von S und endet genau dann nach endlich vielen Schritten, wenn S eine endliche σ-SAGBI-Basis besitzt.

Wie man sehen kann, ist es zur Umsetzung des Algorithmus notwendig, Erzeugendensysteme von Relationenidealen berechnen zu können. Dazu haben wir bereits verschiedene Möglichkeiten kennengelernt: Wie in der Prozedur angegeben mit Algorithmus 2.2 oder alternativ auf folgende Weise.

Bemerkung 5.4.5. (Berechnung von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$)
Seien $t_1, \ldots, t_s \in \mathbb{T}^n$ mit $\mathrm{LT}_\sigma(\mathcal{G}) = (t_1, \ldots, t_s)$. Sei $\Phi : K[y_1, \ldots, y_s] \to P$ definiert durch $y_i \mapsto t_i$ für alle $i \in \{1, \ldots, s\}$. Dann gilt $\mathrm{Rel}(t_1, \ldots, t_s) = \mathrm{Ker}(\Phi)$ und der Kern lässt sich mittels Elimination berechnen, d.h. mit dem Diagonalideal $\triangle_\Phi = \langle y_1 - t_1, \ldots, y_s - t_s \rangle$ gilt (vgl. Korollar 2.2.7)

$$\mathrm{Rel}(t_1, \ldots, t_s) = \triangle_\Phi \cap K[y_1, \ldots, y_s].$$

Außerdem stellt sich bei Betrachtung der Prozedur die Frage, wie sich in Schritt 14 eine Teilmenge $\widehat{B} \subseteq B$ berechnen lässt, deren Binome mindestens eine Unbestimmte aus $\{y_{s'-t+1}, \ldots, y_{s'}\}$ enthalten. Auf dem Papier ist das natürlich sofort zu sehen. Auch dazu können wir aber einen einfachen Algorithmus angeben, der in der Hilfsfunktion `SB.ReplaceB` des ApCoCoA-Pakets `sagbi.cpkg` umgesetzt ist (siehe Anhang C, Seite 307).

Algorithmus 5.5 : Hilfs-Algorithmus für die SAGBI-Prozedur

Input : Endliche Menge $B \subseteq K[y_1, \ldots, y_r, y_{r+1}, \ldots, y_s]$ echter Binome mit $s > r$.
Input : Unbestimmte $y_{r+1}, \ldots, y_s$.
Result : Teilmenge $\widehat{B} \subseteq B$, deren Binome mindestens eine Unbestimmte aus $\{y_{r+1}, \ldots, y_s\}$ enthalten.

```
1 B̂ := ∅, n := s − r;
2 foreach b ∈ B do
3 |   {t1, t2} := Supp(b);
4 |   t := t1 · t2;
5 |   û := Projektion von log(t) auf die letzten n Komponenten;
6 |   if û ≠ 0 then
7 |   |   B̂ := B̂ ∪ {b};
8 return B̂;
```

Betrachtet man sich nun die SAGBI-Prozedur etwas genauer, wäre es natürlich auch denkbar, in Analogie zum Buchberger-Algorithmus auch hier immer dann ein Element in G einzufügen, wenn $b' \neq 0$ gilt, und anschließend ein neues Erzeugendensystem B zu berechnen. Wie in [KR05] bereits erwähnt ist, führt eine derartige Modifikation ggf. zu Problemen. So ist es möglich, dass diese Prozedur nicht mehr terminiert, obwohl eine endliche σ-SAGBI-Basis existiert. Im nachfolgenden und abschließenden Beispiel wäre dies bei entsprechender Modifikation der Prozedur auch der Fall (vgl. [KR05], Beispiel 6.6.30, S. 497, allerdings hier mit kleineren Fehlern).

Beispiel 5.4.6. Sei $P = \mathbb{Q}[x_1, x_2]$ versehen mit der Termordnung $\sigma = \mathrm{DegLex}$ und seien $g_1 = x_1^2 x_2,\ g_2 = x_1^2 - x_2^2,\ g_3 = x_1^2 x_2^2 - x_2^4,\ g_4 = x_1^2 x_2^4$. Sei $S \subseteq P$ die von $\{g_1, g_2, g_3, g_4\}$ erzeugte $\mathbb{Q}$-Unteralgebra von P.

Vor Beginn der while-Schleife gilt $s' := 4$, $G := \{g_1, g_2, g_3, g_4\}$ und $\mathcal{G} = (g_1, g_2, g_3, g_4)$. Die Berechnung von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ liefert das Erzeugendensystem $B = \{y_2y_3 - y_1^2,\ y_3^2 - y_2y_4\}$ im Polynomring $K[y_1, y_2, y_3, y_4]$. Somit ist mindestens ein Schleifendurchlauf notwendig. Im Folgenden werden die einzelnen Durchgänge näher betrachtet:

(1) Setze $b_1 := y_2y_3 - y_1^2$ und $b_2 := y_3^2 - y_2y_4$. Dann gilt

$$\mathrm{NRS}_\mathcal{G}(b_1(g_1, g_2, g_3, g_4)) = \mathrm{NRS}_\mathcal{G}(g_2g_3 - g_1^2) = \mathrm{NRS}_\mathcal{G}(-2x_1^2x_2^4 + x_2^6) = x_2^6$$
$$\mathrm{NRS}_\mathcal{G}(b_2(g_1, g_2, g_3, g_4)) = \mathrm{NRS}_\mathcal{G}(g_3^2 - g_2g_4) = \mathrm{NRS}_\mathcal{G}(-x_1^2x_2^6 + x_2^8) = -x_1^2x_2^6 + x_2^8$$

Somit gilt $B' = \{x_2^6,\ x_1^2x_2^6 - x_2^8\}$, also $t = 2$ und $s' = 6$ in Zeile 11. Setze $g_5 := x_2^6$ sowie $g_6 := x_1^2x_2^6 - x_2^8$ und füge beide Polynome in G sowie in $\mathcal{G}$ ein. Wir berechnen dann ein Erzeugendensystem B von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ in $K[y_1, \ldots, y_6]$ und erhalten:

$$B = \{-y_1^2 + y_2y_3,\ -y_3^2 + y_2y_4,\ -y_2y_5 + y_6,\ -y_3y_4 + y_2y_6,\ -y_4^2 + y_3y_6\}.$$

Mit Algorithmus 5.5 berechnen wir die Teilmenge $\widehat{B}$ von B, die aus den Binomen besteht, die die Unbestimmten y_5 und y_6 enthalten. Hier gilt:

$$B = \{-y_2y_5 + y_6,\ -y_3y_4 + y_2y_6,\ -y_4^2 + y_3y_6\}.$$

(2) Da $B \neq \emptyset$ gilt, ist ein zweiter Durchlauf nötig. Seien $b_1 := -y_2y_5 + y_6$, $b_2 := -y_3y_4 + y_2y_6$ und $b_3 := -y_4^2 + y_3y_6$. Dann gilt $\mathrm{NRS}_\mathcal{G}(b_1(g_1, \ldots, g_6)) = 0$, $\mathrm{NRS}_\mathcal{G}(b_2(g_1, \ldots, g_6)) = 0$ und $\mathrm{NRS}_\mathcal{G}(b_3(g_1, \ldots, g_6)) = 0$. Somit folgt $B' = \emptyset$, d.h.

$$G = \{x_1^2x_2,\ x_1^2 - x_2^2,\ x_1^2x_2^2 - x_2^4,\ x_1^2x_2^4,\ x_2^6,\ x_1^2x_2^6 - x_2^8\}$$

ist eine σ-SAGBI-Basis von S. ◁

5.4.2 Die homogene SAGBI-Prozedur

Nach der Berechnung von σ-SAGBI-Basen und reduzierten σ-SAGBI-Basen wollen wir uns einer Besonderheit zuwenden, nämlich Unteralgebren, die von homogenen Polynomen erzeugt werden. Daraus lassen sich, falls sie existieren, homogene σ-SAGBI-Basen berechnen. Außerdem wird sich zeigen, dass im homogenen Fall ein effizienter Unteralgebra-Mitgliedschaftstest möglich ist, und zwar sogar dann, wenn keine endliche σ-SAGBI-Basis existiert. Die nachfolgenden Ausführungen sind ebenfalls durch Tutorial 96 aus [KR05] motiviert und wurden hier vollständig ausgearbeitet und bewiesen. Auch wir wollen Homogenität nicht im Allgemeinen untersuchen, sondern nur den standardgraduierten Fall betrachten. Sei also im Folgenden der Polynomring P stets mit der Standardgraduierung versehen. Zunächst werden wir sehen, dass K-Unteralgebren, die von homogenen Polynomen erzeugt werden, ebenfalls standardgraduiert sind (vgl. [KR05], Satz 6.6.6, S. 481).

Satz 5.4.7. *Seien $f_1, \ldots, f_s \in P\backslash\{0\}$ homogene Polynome, die eine K-Unteralgebra S erzeugen.*

a) Setze $S_d := P_d \cap S$ für alle $d \in \mathbb{N}$. Dann gilt $S = \bigoplus_{d\in\mathbb{N}} S_d$, d.h. S ist eine standardgraduierte K-Unteralgebra von P.

b) Setze $d_i := \deg(f_i)$ für alle $i \in \{1, \ldots, s\}$ und sei der Polynomring $K[y_1, \ldots, y_s]$ mit der durch $\mathcal{U} = \begin{pmatrix} d_1 & \cdots & d_s \end{pmatrix} \in \mathrm{Mat}_{1,s}(\mathbb{N})$ gegebenen $\mathbb{N}$-Graduierung versehen. Dann ist der surjektive K-Algebra-Homomorphismus $\lambda : K[y_1, \ldots, y_s] \to S$, definiert durch $\lambda(y_i) = f_i$ für alle $i \in \{1, \ldots, s\}$, homogen. Für $\mathcal{G} = (f_1, \ldots, f_s)$ induziert dieser Homomorphismus einen Isomorphismus $\overline{\lambda} : K[y_1, \ldots, y_s]/\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G})) \to S$ von graduierten K-Algebren.

Sind $g_1, \ldots, g_s \in P \setminus \{0\}$ homogene Polynome und $h \in K[y_1, \ldots, y_s]$ homogen bzgl. der durch $\mathcal{U}$ gegebenen Graduierung, so ist nach dem Satz also $\lambda(h) = h(g_1, \ldots, g_s)$ homogen bzgl. der Standardgraduierung. Ist umgekehrt ein Polynom $f \in S$ homogen bzgl. der Standardgraduierung, so gibt es ein bzgl. $\mathcal{U}$ homogenes Polynom $h \in K[y_1, \ldots, y_s]$ mit $\lambda(h) = f$. Zur Unterscheidung der Graduierungen schreiben wir $\deg_{\mathcal{U}}(h)$ für alle $h \in K[y_1, \ldots, y_s]$ bzgl. der durch $\mathcal{U}$ gegebenen $\mathbb{N}$-Graduierung. In dieser Situation gilt also $\deg_{\mathcal{U}}(h) = \deg(h(g_1, \ldots, g_s))$ für alle $h \in K[y_1, \ldots, y_s]$. Außerdem besteht ein Zusammenhang zu der durch $(\sigma, \mathcal{G})$ induzierten Graduierung (vgl. [KR05], Definition 6.6.21, S. 489).

Definition 5.4.8. (Induzierte Graduierung, σ-Grad)
Sei $\mathcal{G} = (g_1, \ldots, g_s)$ ein Tupel homogener Polynome aus $P \setminus \{0\}$ und sei der Polynomring $K[y_1, \ldots, y_s]$ mit der durch $\deg(1) := 1$ und $\deg(y_i) := \mathrm{LT}_\sigma(g_i)$ für alle $i \in \{1, \ldots, s\}$ gegebenen $\mathbb{T}^n$-Graduierung versehen. Diese Graduierung heißt die (durch $(\sigma, \mathcal{G})$) **induzierte Graduierung**. Wir schreiben $\deg_{\sigma,\mathcal{G}}(1)$ und $\deg_{\sigma,\mathcal{G}}(y_i)$ anstatt $\deg(1)$ und $\deg(y_i)$. Für ein Polynom $h \in K[y_1, \ldots, y_s]$ heißt dann $\deg_{\sigma,\mathcal{G}}(h) := \max_\sigma\{\deg_{\sigma,\mathcal{G}}(t) : \ t \in \mathrm{Supp}(h)\}$ der **σ-Grad** von h.

Gemäß der Definition der induzierten Graduierung gilt also $\deg_{\sigma,\mathcal{G}}(t) = \mathrm{LT}_\sigma(t(g_1, \ldots, g_s)) \in \mathbb{T}^n$ für jeden Term $t \in \mathbb{T}(y_1, \ldots, y_s)$. Damit erhalten wir nun folgenden Zusammenhang zwischen der durch $\mathcal{U}$ gegebenen $\mathbb{N}$-Graduierung und der durch $(\sigma, \mathcal{G})$ induzierten $\mathbb{T}^n$-Graduierung.

Lemma 5.4.9. *Seien $g_1, \ldots, g_s \in P \setminus \{0\}$ homogene Polynome, sei $\mathcal{G} = (g_1, \ldots, g_s)$ und sei $h \in K[y_1, \ldots, y_s]$ homogen bzgl. der durch $(\sigma, \mathcal{G})$ induzierten $\mathbb{T}^n$-Graduierung. Dann ist h auch homogen bzgl. der durch $\mathcal{U}$ gegebenen $\mathbb{N}$-Graduierung.*

Beweis: Sei $\tilde{t} = \deg_{\sigma,\mathcal{G}}(h)$. Da h homogen bzgl. der durch $(\sigma, \mathcal{G})$ induzierten Graduierung ist, gilt $\deg_{\sigma,\mathcal{G}}(t) = \tilde{t}$ für alle $t \in \mathrm{Supp}(h)$. Wegen $\deg_{\sigma,\mathcal{G}}(t) = \mathrm{LT}_\sigma(t(g_1, \ldots, g_s))$ für alle $t \in \mathrm{Supp}(h)$ gilt $\deg_{\mathcal{U}}(t) = \deg(\tilde{t})$ für alle $t \in \mathrm{Supp}(h)$ und es folgt die Behauptung. □

Sei im Folgenden für ein gegebenes Tupel $\mathcal{G} = (g_1, \ldots, g_s)$ homogener Polynome aus $P \setminus \{0\}$ der Polynomring $K[y_1, \ldots, y_s]$ stets neben der durch $(\sigma, \mathcal{G})$ induzierten Graduierung auch mit der durch $\mathcal{U} := \begin{pmatrix} \deg(g_1) & \cdots & \deg(g_s) \end{pmatrix} \in \mathrm{Mat}_{1,s}(\mathbb{N})$ gegebenen $\mathbb{N}$-Graduierung versehen.

Bemerkung 5.4.10. Betrachten wir nun ein homogenes Polynom $f \in P$ vom Grad d und sei $G \subseteq P \setminus \{0\}$ eine Menge normierter homogener Polynome. Falls f reduzibel ist bzgl. $\xrightarrow{G}_{\mathrm{S}}$, gibt es homogene Polynome $g_1, \ldots, g_s \in G$ und einen Term $t \in K[y_1, \ldots, y_s]$ mit $t(\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)) \in \mathrm{Supp}(f)$. Setze $\mathcal{G} = (g_1, \ldots, g_s)$. Dann ist $\mathrm{NRS}_{\mathcal{G}}(f)$ ebenfalls homogen vom Grad d.

Diese Tatsache werden wir nun verwenden, um zu zeigen, dass die Prozedur SAGBI für eine von homogenen Polynomen erzeugten K-Unteralgebra eine homogene σ-SAGBI-Basis berechnet (vgl. [RS90], dort ohne Beweis).

Satz 5.4.11. (Berechnung homogener SAGBI-Basen mit der SAGBI-Prozedur)
Seien $g_1, \ldots, g_s \in P \setminus \{0\}$ homogene, normierte Polynome, die eine K-Unteralgebra S erzeugen. Dann berechnet die Prozedur SAGBI eine homogene σ-SAGBI-Basis von S.

Beweis: Zu Beginn der Prozedur SAGBI bestehen sowohl G, als auch $\mathcal{G}$ aus homogenen Polynomen. Alle im Laufe der Prozedur berechneten Erzeugendensysteme B von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ enthalten bzgl. der durch $(\sigma, \mathcal{G})$ induzierten Graduierung homogene Binome. Da die Erzeuger

von S homogen sind, sind die Erzeuger von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ auch homogen bzgl. der durch $\mathcal{U}$ gegebenen Graduierung. Aus Satz 5.4.7 folgt, dass $b(g_1, \ldots, g_{s'})$ für alle $b \in B$ homogen in P ist. Wegen der obigen Bemerkung ist $b' = \mathrm{NRS}_{\mathcal{G}}(b(g_1, \ldots, g_{s'}))$ ebenfalls ein bzgl. der Standardgraduierung homogenes Polynom vom Grad $\deg(b(g_1, \ldots, g_{s'}))$, falls $b' \neq 0$ gilt. Somit werden in Zeile 8 der Prozedur in B' und damit in Zeile 12 auch in G bzw. $\mathcal{G}$ nur bzgl. der Standardgraduierung homogene Polynome eingefügt, d.h. die Prozedur SAGBI berechnet eine homogene σ-SAGBI-Basis. Analog zu Theorem 5.4.4 ist diese homogene σ-SAGBI-Basis von S genau dann endlich, wenn die Prozedur SAGBI terminiert. □

Insbesondere geht aus diesem Satz hervor, dass eine von endlich vielen homogenen Polynomen erzeugte K-Unteralgebra S stets eine homogene σ-SAGBI-Basis besitzt. Wir wollen nun eine homogene Version der Prozedur SAGBI betrachten. In [KR05] wird eine solche enumerierte Prozedur angegeben (vgl. [KR05], Tutorial 96).

Prozedur HomSagbi

Input : $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$: Homogene, normierte Polynome mit $S = K[g_1, \ldots, g_s]$.
Result : Eine homogene σ-SAGBI-Basis G von S.

1 $s' := 0$, $B := \emptyset$, $W := \{g_1, \ldots, g_s\}$, $G := \emptyset$, $\mathcal{G} := \emptyset$;
2 **while** $B \neq \emptyset$ *oder* $W \neq \emptyset$ **do**
3 $d := \min\{\{\deg(h) : h \in B\} \cup \{\deg(f) : f \in W\}\}$;
4 Berechne $B_d = B \cap P_d$ und $W_d = W \cap P_d$ sowie $B := B \setminus B_d$ und $W := W \setminus W_d$;
5 $B' := \emptyset$ und $W' := \emptyset$;
6 **foreach** $b \in B_d$ **do**
7 $b' := \mathrm{NRS}_{\mathcal{G}}(b(g_1, \ldots, g_{s'}))$;
8 **if** $b' \neq 0$ **then**
9 $B' := B' \cup \left\{\frac{1}{\mathrm{LC}_\sigma(b')} b'\right\}$;
10 $t := \#B'$ und $s' := s' + t$;
11 **if** $t \neq 0$ **then**
  `// Elemente von` B' `werden mit` $g_{s'-t+1}, \ldots, g_{s'}$ `bezeichnet`
12 $G := G \cup \{g_{s'-t+1}, \ldots, g_{s'}\}$ und $\mathcal{G} := \mathcal{G} \oplus (g_{s'-t+1}, \ldots, g_{s'})$;
13 **foreach** $g \in W_d$ **do**
14 $g' := g$;
15 **if** $\mathcal{G} \neq \emptyset$ **then**
16 $g' := \mathrm{NRS}_{\mathcal{G}}(g)$;
17 **if** $g' \neq 0$ **then**
18 $W' := W' \cup \left\{\frac{1}{\mathrm{LC}_\sigma(g')} g'\right\}$;
19 $t := \#W'$ und $s' := s' + t$;
20 **if** $t \neq 0$ **then**
  `// Elemente von` W' `werden mit` $g_{s'-t+1}, \ldots, g_{s'}$ `bezeichnet`
21 $G := G \cup \{g_{s'-t+1}, \ldots, g_{s'}\}$ und $\mathcal{G} := \mathcal{G} \oplus (g_{s'-t+1}, \ldots, g_{s'})$;
22 **if** $B' \neq \emptyset$ *oder* $W' \neq \emptyset$ **then**
23 Berechne mit Algorithmus 2.2 ein $B \subseteq K[y_1, \ldots, y_{s'}]$ mit $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G})) = \langle B \rangle$;
24 $B := B_{>d}$;
25 **return** G;

Diese Prozedur ist ebenfalls Teil des ApCoCoA-Pakets `sagbi.cpkg` und in der CoCoA-Funktion `SB.HomSagbi` implementiert (siehe Anhang C, Seite 299). Den Ablauf dieser Prozedur werden wir zunächst an einem wohlbekannten Beispiel demonstrieren, bevor wir deren Korrektheit zeigen. Die in Beispiel 5.4.6 betrachtete Unteralgebra S wird von homogenen Polynomen erzeugt und besitzt eine endliche σ-SAGBI-Basis. Diese SAGBI-Basis ist sogar homogen, was wir in Satz 5.4.11 allgemein bewiesen hatten. Nun wollen wir die Prozedur HomSagbi auf dieses Beispiel anwenden.

Beispiel 5.4.12. Sei $P = \mathbb{Q}[x_1, x_2]$, sei $\sigma = \text{DegLex}$ und seien $f_1, f_2, f_3, f_4 \in P$ mit

$$f_1 = x_1^2x_2,\ f_2 = x_1^2 - x_2^2,\ f_3 = x_1^2x_2^2 - x_2^4,\ f_4 = x_1^2x_2^4.$$

Sei $S \subseteq P$ die von den homogenen Polynomen f_1, f_2, f_3, f_4 erzeugte $\mathbb{Q}$-Unteralgebra von P.

Zu Beginn der Prozedur HomSagbi gilt also $s' = 0$, $B = \emptyset$, $W = \{f_1, \ldots, f_4\}$, $G = \emptyset$ und $\mathcal{G} = \emptyset$. Somit wird die while-Schleife mindestens ein Mal ausgeführt. Wir erhalten folgende Schleifendurchläufe:

(1) Es gilt $d := \deg(f_2) = 2$, $W_2 = \{x_1^2 - x_2^2\}$ und $W := W \setminus W_2 = \{x_1^2x_2, x_1^2x_2^2 - x_2^4, x_1^2x_2^4\}$. Wegen $B = \emptyset$, also auch $B_2 = \emptyset$, muss nur W_2 abgearbeitet werden. Aufgrund von $\mathcal{G} = \emptyset$ kann $g_1 := x_1^2 - x_2^2$ sofort in W' und anschließend in G bzw. $\mathcal{G}$ eingefügt werden. Als Erzeugendensystem B von $\text{Rel}(\text{LT}_\sigma(\mathcal{G}))$ erhalten wir $B = \emptyset$.

(2) Nun gilt $d := \deg(f_1) = \deg(x_1^2x_2) = 3$, also folgt $W_3 = \{x_1^2x_2\}$ und damit erhalten wir $W := W \setminus W_3 = \{x_1^2x_2^2 - x_2^4, x_1^2x_2^4\}$. Es ist erneut nur W_3 abzuarbeiten. Mit dem Tupel $\mathcal{G} = (x_1^2 - x_2^2)$ folgt $\text{NRS}_\mathcal{G}(x_1^2x_2) = x_1^2x_2$, d.h. $g_2 := x_1^2x_2$ wird in W' und anschließend in G sowie $\mathcal{G}$ eingefügt. Somit gilt $G = \{x_1^2 - x_2^2, x_1^2x_2\}$ und $\mathcal{G} = (x_1^2 - x_2^2, x_1^2x_2)$. Erneut gilt $\text{Rel}(\text{LT}_\sigma(\mathcal{G})) = 0$, also $B = \emptyset$.

(3) Im dritten Durchgang erhalten wir $d := \deg(x_1^2x_2^2 - x_2^4) = 4$. Damit gilt $W_4 = \{x_1^2x_2^2 - x_2^4\}$ und $W = \{x_1^2x_2^4\}$. Wegen $B_4 = \emptyset$ ist wiederum nur W_4 zu betrachten. Wir erhalten $\text{NRS}_\mathcal{G}(x_1^2x_2^2 - x_2^4) = x_1^2x_2^2 - x_2^4$, d.h. $g_3 := x_1^2x_2^2 - x_2^4$ wird zu G bzw. $\mathcal{G}$ hinzugefügt. Insgesamt erhalten wir also

$$G = \{x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4\} \quad \text{bzw.} \quad \mathcal{G} = (x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4).$$

Die Berechnung eines Erzeugendensystems von $\text{Rel}(\text{LT}_\sigma(\mathcal{G}))$ liefert nun $B = \{y_1y_3 - y_2^2\}$. Es gilt $\deg_\mathcal{U}(y_1y_3 - y_2^2) = \deg(g_1g_3 - g_2^2) = \deg(-2x_1^2x_2^4 + x_2^6) = 6$. Somit gilt $B_{>4} = \{y_1y_3 - y_2^2\}$ und schließlich $B := B_{>4}$ für den nächsten Durchlauf.

(4) Wegen $\deg(x_1^2x_2^4) = \deg_\mathcal{U}(y_1y_3 - y_2^2) = 6$ gilt $d = 6$. Somit folgt $W_6 = \{x_1^2x_2^4\}$ sowie $B_6 = \{y_1y_3 - y_2^2\}$. Wir erhalten in Zeile 4 also $B = \emptyset$ und $W = \emptyset$. Für $b := y_1y_3 - y_2^2$ gilt $b' := \text{NRS}_\mathcal{G}(b(g_1, g_2, g_3)) = -2x_1^2x_2^4 + x_2^6$. Dann wird $g_4 := \frac{1}{\text{LC}_\sigma(b')}b' = x_1^2x_2^4 - \frac{1}{2}x_2^6$ in G bzw. $\mathcal{G}$ eingefügt, d.h. es gilt $G = \{x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4, x_1^2x_2^4 - \frac{1}{2}x_2^6\}$. Anschließend wird $x_1^2x_2^4 \in W_6$ betrachtet. Es gilt $\text{NRS}_\mathcal{G}(x_1^2x_2^4) = \frac{1}{2}x_2^6$. Somit wird $g_5 := x_2^6$ in G bzw. $\mathcal{G}$ eingefügt. Wir erhalten also

$$G = \{x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4, x_1^2x_2^4 - \tfrac{1}{2}x_2^6, x_2^6\}$$

bzw. $\mathcal{G} = (x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4, x_1^2x_2^4 - \frac{1}{2}x_2^6, x_2^6)$. Die Berechnung von $\text{Rel}(\text{LT}_\sigma(\mathcal{G}))$ liefert $B = \{y_1y_3 - y_2^2,\ y_3^2 - y_1y_4,\ y_1^2y_5 - y_3y_4,\ y_2^2y_5 - y_4^2\}$. Setze $b_1 := y_3^2 - y_1y_4$, $b_2 := y_1^2y_5 - y_3y_4$ und $b_3 := y_2^2y_5 - y_4^2$. Dann gilt:

$$\begin{aligned}
\text{NRS}_\mathcal{G}(b_1(g_1, \ldots, g_5)) &= -\tfrac{1}{2}x_1^2x_2^6 + \tfrac{1}{2}x_2^8 \\
\text{NRS}_\mathcal{G}(b_2(g_1, \ldots, g_5)) &= -\tfrac{1}{2}x_1^2x_2^8 + \tfrac{1}{2}x_2^{10} \\
\text{NRS}_\mathcal{G}(b_3(g_1, \ldots, g_5)) &= x_1^2x_2^{10} - \tfrac{1}{4}x_2^{12}
\end{aligned}$$

also $B_{>6} = \{b_1, b_2, b_3\}$. Wir ersetzen B durch $B_{>6}$ und gehen über zum nächsten Schleifendurchlauf.

(5) Ab sofort gilt stets $W = \emptyset$ und damit auch $W_d = \emptyset$ bzw. $W' = \emptyset$. Aus den obigen Berechnungen folgt sofort $\deg_{\mathcal{U}}(b_1) = 8$, $\deg_{\mathcal{U}}(b_2) = 10$ und $\deg_{\mathcal{U}}(b_3) = 12$. Somit gilt $d := 8$, also $B_8 = \{b_1\}$ und $B := \{b_2, b_3\}$. Es gilt $\mathrm{NRS}_{\mathcal{G}}(b_1(g_1, \ldots, g_5)) = 0$. Somit bleibt B' leer und in diesem Durchgang ist nichts weiter zu tun.

(6) Nun gilt $d := 10$ wegen $\deg_{\mathcal{U}}(b_2) = 10$ und $\deg_{\mathcal{U}}(b_3) = 12$. Es gilt also $B_{10} = \{b_2\}$ und $B := \{b_3\}$. Wegen $\mathrm{NRS}_{\mathcal{G}}(b_2(g_1, \ldots, g_5)) = 0$ gilt erneut $B' = \emptyset$.

(7) Es bleibt nur noch b_3 übrig, d.h. es gilt $d := \deg_{\mathcal{U}}(b_3) = 12$. Somit folgt $B_{12} = \{b_3\}$ und $B := \emptyset$. Erneut gilt $B' = \emptyset$ wegen $\mathrm{NRS}_{\mathcal{G}}(b_3(g_1, \ldots, g_5)) = 0$. Somit gilt $B = \emptyset$ und $W = \emptyset$, weshalb die Prozedur nach diesem Schleifendurchlauf endet.

Die Prozedur gibt entsprechend das Tupel $\mathcal{G} = (x_1^2 - x_2^2, x_1^2x_2, x_1^2x_2^2 - x_2^4, x_1^2x_2^4 - \frac{1}{2}x_2^6, x_2^6)$ zurück. Dieses Ergebnis unterscheidet sich zwar leicht von dem Ergebnis der Prozedur SAGBI, allerdings sieht man sofort, dass beide dieselbe $\mathbb{Q}$-Unteralgebra von P erzeugen. Ebenfalls sieht man hier sofort, dass die Polynome in $\mathcal{G}$ homogen sind. ◁

Nachdem wir den Ablauf der Prozedur HomSagbi an einem Beispiel demonstriert hatten, wollen wir uns nun von der Korrektheit dieser Prozedur überzeugen, indem wir das folgende Theorem beweisen werden.

Theorem 5.4.13. (Die homogene SAGBI-Basis Prozedur)
Sei σ eine Termordnung auf $\mathbb{T}^n$ und seien $g_1, \ldots, g_s \in P \setminus \{0\}$ homogene, normierte Polynome, die die K-Unteralgebra S erzeugen. Die Prozedur HomSagbi enumeriert eine homogene σ-SAGBI-Basis G von S und endet genau dann dann nach endlich vielen Schritten, wenn S eine endliche homogene σ-SAGBI-Basis besitzt. In diesem Fall liefert die Prozedur HomSagbi ein nach Grad aufsteigend geordnetes Tupel $\mathcal{G}$, dessen Polynome eine homogene σ-SAGBI-Basis G von S bilden.

Beweis: Für $n \geq 1$ bezeichnen wir mit $s'(n), B(n), B'(n), W(n), W'(n), G(n), \mathcal{G}(n)$ und $d(n)$ die Größen $s, B, B', W, W', G, \mathcal{G}$ und d der Prozedur HomSagbi nach Ende des n-ten Schleifendurchlaufs. Da W endlich ist und in Zeile 4 mit jedem Schleifendurchlauf verkleinert wird, ist zunächst klar, dass $W = \emptyset$ nach endlich vielen Schritten erreicht wird. Da die while-Schleife genau dann endet, wenn sowohl $B = \emptyset$ als auch $W = \emptyset$ gilt, terminiert die Prozedur frühestens dann, wenn alle Elemente von W abgearbeitet sind. Sei also ohne Einschränkung n der erste Zeitpunkt mit $W(n) = \emptyset$ und damit auch $W'(n) = \emptyset$. Damit folgt sofort $W(k) = W'(k) = \emptyset$ und $G(k) \neq \emptyset$ sowie $\mathcal{G}(k) \neq \emptyset$ für alle $k \geq n$.

Wir nehmen zunächst an, dass die Prozedur nicht terminiert. Dann folgt sofort $B(k) \neq \emptyset$ für alle $k \geq n$. Das bedeutet, dass in die Menge G laufend Polynome mit streng wachsendem Grad eingefügt werden. Somit ist G eine nicht-endliche Menge, genauer gilt $G = \bigcup_{k \in \mathbb{N}_+} G(k)$ wegen $G(k) \subseteq G(k+1)$ für alle $k \in \mathbb{N}_+$. Es bleibt zu zeigen, dass G eine homogene σ-SAGBI-Basis von S ist. Sei $\{g_1, \ldots, g_{s'}\} \subseteq G$ und $\mathcal{G} = (g_1, \ldots, g_{s'})$. Wegen der Endlichkeit dieser Teilmenge gibt es ein minimales $n' \in \mathbb{N}_+$ mit $\{g_1, \ldots, g_{s'}\} \subseteq G(n')$. Die Menge $B(n')$ enthält alle Erzeuger von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}(n')))$ mit Grad größer $d(n')$. Sei $\tilde{B}(n')$ eine Menge echter Binome, die das Ideal $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}(n')))$ erzeugen. Dann ist $\tilde{B}(n')$ die disjunkte Vereinigung von $B(n')$ und $B(n')_{\leq d(n')}$. Für alle $b \in B(n')_{\leq d(n')}$ gilt $\mathrm{NRS}_{\mathcal{G}(n')}(b(g_1, \ldots, g_{s'})) = 0$, da diese Polynome in vorherigen Durchgängen entweder bereits zu 0 reduziert oder in G eingefügt wurden, womit sie dann spätestens im n'-ten Durchgang zu 0 reduziert werden.

Sei nun $b \in B(n')$. Gilt $\deg_{\mathcal{U}}(b) = d(n'+1)$, so wird $b(g_1, \ldots, g_{s'})$ im $(n'+1)$-ten Durchlauf reduziert, d.h. es gilt $b' := \mathrm{NRS}_{\mathcal{G}(n')}(b(g_1, \ldots, g_{s'}))$. Entweder es gilt $b' = 0$ oder $\frac{1}{\mathrm{LC}_\sigma(b')} b'$ wird zu $G(n')$ hinzugefügt. Dann folgt $\mathrm{NRS}_{\mathcal{G}(n'+1)}(b(g_1, \ldots, g_{s'})) = 0$. Für $\deg_{\mathcal{U}}(b) > d(n'+1)$ erfolgt eine analoge Behandlung in späteren Schleifendurchläufen. Somit gibt es für jedes Tupel $\mathcal{G} = (g_1, \ldots, g_{s'})$ von Elementen aus G ein Erzeugendensystem B von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}))$ so, dass $\mathrm{NRS}_{\mathcal{G}}(b(g_1, \ldots, g_{s'})) = 0$ gilt für alle $b \in B$. Gemäß dem SAGBI-Kriterium, Satz 5.4.2, ist somit G eine σ-SAGBI-Basis von S. Da sämtliche Polynome, die im Laufe der Prozedur in G eingefügt werden, homogen sind, bildet G eine homogene σ-SAGBI-Basis.

Falls die Prozedur nach endlich vielen Schritten terminiert, gibt es ein $n' \geq n$ mit $B(n') = \emptyset$ und $B(n'-1) \neq \emptyset$. Natürlich ist die Menge $G(n')$ endlich und nicht-leer. Sei $G(n') = \{g_1, \ldots, g_{s'(n')}\}$ und $\mathcal{G}(n') = (g_1, \ldots, g_{s'(n')})$. Die Elemente von $B(n'-1)$ sind homogen vom Grad $d(n')$. Gilt $\mathrm{NRS}_{\mathcal{G}(n'-1)}(b(g_1, \ldots, g_{s'(n'-1)})) = 0$ für jedes $b \in B(n'-1)_{d(n')}$, so folgt $B'(n') = \emptyset$. Mit derselben Argumentation wie oben folgt aus dem SAGBI-Kriterium, Satz 5.4.2, dass die Menge $G(n') := G(n'-1)$ eine σ-SAGBI-Basis von S ist. Andererseits gilt $B'(n') \neq \emptyset$. Dann wird zunächst in Zeile 23 ein Erzeugendensystem $\tilde{B}(n')$ von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}(n')))$ berechnet. Wegen $B(n') = \emptyset$ muss entweder $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}(n'))) = 0$ oder $\tilde{B}(n') \neq \emptyset$, aber $\tilde{B}(n')_{>d(n')} = \emptyset$ gelten.

Im ersten Fall folgt, dass die Leitterme der Polynome aus $G(n')$ algebraisch unabhängig sind. Laut Satz 5.1.9 ist $G(n')$ dann eine σ-SAGBI-Basis von S. Gilt $\tilde{B}(n')_{>d(n')} = \emptyset$, gibt es keine weiteren Binome abzuarbeiten und es folgt erneut aus dem SAGBI-Kriterium, dass $G(n')$ eine σ-SAGBI-Basis von S bildet. Da die Prozedur die Mengen B bzw. W aufsteigend nach dem Grad der Elemente abarbeitet und entsprechend auch aufsteigend in G bzw. $\mathcal{G}$ einfügt, ist $\mathcal{G}(n')$ ein nach Grad aufsteigend sortiertes Tupel homogener Polynome. Nach dem Ende der while-Schleife gibt die Prozedur somit $\mathcal{G} := \mathcal{G}(n')$ als SAGBI-Basis zurück.

Besitze nun S eine endliche homogene σ-SAGBI-Basis. Dann ist zu zeigen, dass die Prozedur terminiert. Nach wie vor gilt $W(n) = \emptyset$. Es bleibt also zu zeigen, dass es ein $n' \geq n$ gibt mit $B(n') = \emptyset$. Da $G = \bigcup_{k \in \mathbb{N}_+} G(k)$ eine σ-SAGBI-Basis von S ist, erzeugt gemäß Theorem 5.2.13 die Menge $\{\mathrm{LT}_\sigma(g) |\ g \in G\}$ die monomiale K-Unteralgebra $K[\mathrm{LT}_\sigma(f) |\ f \in S \setminus \{0\}]$. Diese K-Unteralgebra besitzt ein minimales monomiales Erzeugendensystem, das in $\{\mathrm{LT}_\sigma(g) |\ g \in G\}$ enthalten ist. Weil S eine endliche σ-SAGBI-Basis besitzt, ist laut Korollar 5.1.7 das minimale monomiale Erzeugendensystem ebenfalls endlich. Somit gibt es ein $n' \geq n$ so, dass das endliche minimale monomiale Erzeugendensystem in $\{\mathrm{LT}_\sigma(g) |\ g \in G(n')\}$ enthalten ist, d.h. $G(n')$ ist eine endliche homogene σ-SAGBI-Basis von S. Es folgt sofort, dass $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G}(n')))$ das Nullideal ist, also $B(n') = \emptyset$ gilt und die Schleife terminiert. □

5.4.3 Grad-beschränkte SAGBI-Basen

Die Prozedur HomSagbi geht aufsteigend Grad für Grad vor und die Elemente des resultierenden Tupels $\mathcal{G}$ sind entsprechend auch anhand des Grads aufsteigend sortiert. Stoppt man also die Prozedur HomSagbi nach Erreichen eines bestimmten Grades d_0, so enthält die Menge G bzw. das Tupel $\mathcal{G}$ zu diesem Zeitpunkt Polynome mit maximalem Grad d_0, die in einer σ-SAGBI-Basis von S enthalten sind. Diese Teilmenge einer σ-SAGBI-Basis wird als d_0-Grad-beschränkte σ-SAGBI-Basis bezeichnet. Wie im letzten Abschnitt sei der Polynomring $P = K[x_1, \ldots, x_n]$ über einem Körper K stets mit der Standardgraduierung versehen.

Definition 5.4.14. (Grad-beschränkte σ-SAGBI-Basis)
Sei S eine endlich erzeugte K-Unteralgebra von P und sei $G \subseteq P \setminus \{0\}$ eine homogene σ-SAGBI-Basis von S. Für $d \in \mathbb{N}_+$ heißt die Menge $G_{\leq d} := \{g \in G :\ \deg(g) \leq d\}$ bzw. das Tupel $\mathcal{G}_{\leq d}$,

das die nach Grad aufsteigend sortierten Elemente von $G_{\leq d}$ enthält, eine **d-Grad-beschränkte σ-SAGBI-Basis** von S.

Die deutsche Mathematikerin Karin GATERMANN (1961-2005) verwendete beispielsweise Grad-beschränkte SAGBI-Basen zur Lyapunov-Schmidt-Reduktion in dynamischen Systemen (vgl. [Gat03]). Die Berechnung von d-Grad-beschränkten σ-SAGBI-Basen geht allein aus der Definition hervor. Es ist einzig die Prozedur HomSagbi um den Grad d als zusätzliches Abbruchkriterium zu erweitern. Wegen des zusätzlichen Abbruchkriteriums ist die Terminierung der while-Schleife sicher gestellt, weshalb wir einen Algorithmus zur Berechnung von d-Grad-beschränkten σ-SAGBI-Basen erhalten (siehe Algorithmus 5.6).

Algorithmus 5.6 : Berechnung einer d-Grad-beschränkten SAGBI-Basis

Input : $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$: Homogene, normierte Polynome mit $S = K[g_1, \ldots, g_s]$.
Input : d_0: Maximaler Grad mit $d_0 \geq \min\{\deg(g_i) : \ i = 1, \ldots, s\}$.
Result : Eine d_0-Grad-beschränkte σ-SAGBI-Basis $\mathcal{G}_{\leq d_0}$ von S.

1 $s' := 0$, $B := \emptyset$, $W := \{g_1, \ldots, g_s\}$, $G := \emptyset$, $\mathcal{G} := \emptyset$;
2 **while** $B \neq \emptyset$ *oder* $W \neq \emptyset$ **do**
3 $d := \min(\{\deg(h) : \ h \in B\} \cup \{\deg(f) : \ f \in W\})$;
4 **if** $d > d_0$ **then**
5 break;
6 $B' := \emptyset$ und $W' := \emptyset$;
7 **foreach** $b \in B_d$ **do**
8 $b' := \mathrm{NRS}_{\mathcal{G}}(b(g_1, \ldots, g_{s'}))$;
9 **if** $b' \neq 0$ **then**
10 $B' := B' \cup \left\{\frac{1}{\mathrm{LC}_\sigma(b')} \cdot b'\right\}$;
11 $t := \#B'$ und $s' := s' + t$;
12 **if** $t \neq 0$ **then**
 `// Elemente von` B' `werden mit` $g_{s'-t+1}, \ldots, g_{s'}$ `bezeichnet`
13 $G := G \cup \{g_{s'-t+1}, \ldots, g_{s'}\}$ und $\mathcal{G} := \mathcal{G} \oplus (g_{s'-t+1}, \ldots, g_{s'})$;
14 **foreach** $g \in W_d$ **do**
15 $g' := g$;
16 **if** $\mathcal{G} \neq \emptyset$ **then**
17 $g' := \mathrm{NRS}_{\mathcal{G}}(g)$;
18 **if** $g' \neq 0$ **then**
19 $W' := W' \cup \left\{\frac{1}{\mathrm{LC}_\sigma(g')} \cdot g'\right\}$;
20 $t := \#W'$ und $s' := s' + t$;
21 **if** $t \neq 0$ **then**
 `// Elemente von` W' `werden mit` $g_{s'-t+1}, \ldots, g_{s'}$ `bezeichnet`
22 $G := G \cup \{g_{s'-t+1}, \ldots, g_{s'}\}$ und $\mathcal{G} := \mathcal{G} \oplus (g_{s'-t+1}, \ldots, g_{s'})$;
23 **if** $B' \neq \emptyset$ *oder* $W' \neq \emptyset$ **then**
24 Berechne mit Algorithmus 2.2 ein $B \subseteq K[y_1, \ldots, y_{s'}]$ mit $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{G})) = \langle B \rangle$;
25 $B := B_{>d}$;
26 **return** $\mathcal{G}$;

Dieser Algorithmus ist in der CoCoA-Funktion `SB.TruncSagbi` des Pakets `sagbi.cpkg` im-

plementiert (siehe Anhang C, Seite 305). Die Korrektheit und Endlichkeit dieses Algorithmus folgt aus dem nachstehenden Korollar.

Korollar 5.4.15. (d-Grad-beschränkter SAGBI-Algorithmus)
Sei σ eine Termordnung auf $\mathbb{T}^n$ und seien $g_1, \ldots, g_s \in P \setminus \{0\}$ homogene, normierte Polynome, die die K-Unteralgebra S erzeugen. Sei $d_0 \in \mathbb{N}_+$ mit $d_0 \geq \min\{\deg(g_1), \ldots, \deg(g_s)\}$. Dann berechnet Algorithmus 5.6 eine d_0-Grad-beschränkte σ-SAGBI-Basis $G_{\leq d_0}$ von S und gibt ein nach Grad aufsteigend sortiertes Tupel $\mathcal{G}_{\leq d_0}$ zurück, das die Elemente von $G_{\leq d_0}$ beinhaltet.

Beweis: Dass durch die zusätzliche Abbruchbedingung die Prozedur HomSagbi in jedem Fall terminiert, ist klar. Außerdem wurde in Theorem 5.4.13 bereits bewiesen, dass die Prozedur HomSagbi irgendwann eine σ-SAGBI-Basis von S berechnet. Die Elemente der Menge G bzw. des Tupels $\mathcal{G}$ sind zu einem bestimmten Zeitpunkt nach Grad aufsteigend sortiert. Bricht man also diese Prozedur im Grad d_0 ab, ist das bis dahin berechnete Tupel $\mathcal{G}$ eine d_0-Grad-beschränkte σ-SAGBI-Basis. □

Analog zu d-Grad-beschränkten σ-Gröbner Basen (siehe [KR05], Kapitel 4.5.B) lassen sich auch d-Grad-beschränkte σ-SAGBI-Basen auf verschiedene Arten charakterisieren.

Satz 5.4.16. (Charakterisierungen d-Grad-beschränkter SAGBI-Basen)
Sei $G = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine Menge normierter und homogener Polynome, die die K-Unteralgebra S von P erzeugen. Sei $d \in \mathbb{N}_+$ und S' die von den Polynomen aus $G_{\leq d}$ erzeugte K-Unteralgebra. Dann sind folgende Aussagen äquivalent:

(i) $G_{\leq d}$ bzw. $\mathcal{G}_{\leq d}$ ist eine d-Grad-beschränkte σ-SAGBI-Basis von S.

(ii) Für alle $f \in S'$ mit $f \neq 0$ ist $\mathrm{LT}_\sigma(f)$ ein Element des von $\{\mathrm{LT}_\sigma(g) : \ g \in G_{\leq d}\}$ erzeugten Monoids.

(iii) Für jedes Tupel $\mathcal{H} = (h_1, \ldots, h_k)$ von Polynomen aus $G_{\leq d}$ gibt es ein Erzeugendensystem $B \subseteq K[y_1, \ldots, y_k]$ von $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{H}))$ so, dass $b(h_1, \ldots, h_k) \xrightarrow{G_{\leq d}}_s 0$ gilt für alle $b \in B_{\leq d}$.

Beweis: Sei $\mathcal{G} = (g_1, \ldots, g_s)$ bzgl. des Grades aufsteigend geordnet. Dann gibt es ein $s' \leq s$ mit $\mathcal{G}_{\leq d} = (g_1, \ldots, g_{s'})$ und es gilt weiter $S' = K[g_1, \ldots, g_{s'}]$.

$(i) \Rightarrow (ii)$**:** Da $G_{\leq d}$ eine d-Grad-beschränkte σ-SAGBI-Basis von S ist, gibt es eine homogene σ-SAGBI-Basis $\tilde{G}$ von S mit $\tilde{G}_{\leq d} = G_{\leq d}$. Aus Theorem 5.2.13 folgt, dass das multiplikative Monoid $\{\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}\}$ von $\{\mathrm{LT}_\sigma(g) : \ g \in \tilde{G}\}$ erzeugt wird. Ist also $f \in S'$, so folgt sofort, dass $\mathrm{LT}_\sigma(f)$ ein Produkt von Leittermen von Polynomen aus $G_{\leq d}$ ist.

$(ii) \Rightarrow (i)$**:** Aus der Voraussetzung folgt, dass $\mathrm{LT}_\sigma(f)$ für jedes $f \in S'$ mit $f \neq 0$ auch ein Element des von $\{\mathrm{LT}_\sigma(g) : \ g \in G\}$ erzeugten Monoids ist. Durch Hinzufügen von ggf. unendlich vielen Termen $t_1, t_2, \ldots$ mit $\deg(t_i) > d$ für alle $i \in \mathbb{N}_+$ kann $\{\mathrm{LT}_\sigma(g) : \ g \in G\}$ zu einem Erzeugendensystem von $\{\mathrm{LT}_\sigma(f) : \ f \in S \setminus \{0\}\}$ ergänzt werden. Laut [KR00], Satz 1.5.6, gibt es homogene Polynome $h_1, h_2, \ldots \in P$ mit $\mathrm{LT}_\sigma(h_i) = t_i$ für alle $i \in \mathbb{N}_+$. Gemäß Theorem 5.2.13 ist $\tilde{G} := \{g_1, \ldots, g_s, h_1, h_2, \ldots\}$ dann eine homogene σ-SAGBI-Basis von S mit $\tilde{G}_{\leq d} = G_{\leq d}$.

$(i) \Rightarrow (iii)$**:** Sei $\tilde{G}$ eine homogene σ-SAGBI-Basis der K-Unteralgebra S mit $\tilde{G}_{\leq d} = G_{\leq d}$ und sei $\mathcal{H} = (h_1, \ldots, h_k)$ ein Tupel von Polynomen aus $G_{\leq d}$. Laut dem SAGBI-Basis Kriterium, Satz 5.4.2, gibt es eine Menge $B \subseteq K[y_1, \ldots, y_k]$ echter Binome, die das Ideal $\mathrm{Rel}(\mathrm{LT}_\sigma(\mathcal{H}))$ erzeugen, und mit der Eigenschaft, dass $b(h_1, \ldots, h_k) \xrightarrow{\tilde{G}}_{\mathrm{S}} 0$ gilt für alle $b \in B$. Für alle $b \in B_{\leq d}$ ist $b(h_1, \ldots, h_k) \in P$ ein homogenes Polynom vom Grad $\leq d$. Somit wird $b(h_1, \ldots, h_k)$ nur mit Polynomen aus $G_{\leq d}$ zu 0 reduziert.

(iii) $\Rightarrow$ *(i)*: Ohne Einschränkung können wir das Tupel $\mathcal{G}$ durch das Teiltupel $\widehat{\mathcal{G}}$ von $\mathcal{G}$ ersetzen, das aus allen Polynomen besteht, deren Leitterme kein Produkt von anderen Leittermen aus $\mathrm{LT}_\sigma(G)$ sind. Nun wenden wir Algorithmus 5.6 auf das ersetzte Tupel $\mathcal{V} := \mathcal{G}$ und $d_0 := d$ an. Aus der Voraussetzung folgt sofort, dass bis zum Abbruch der while-Schleife bzw. bis zum Erreichen von Grad d_0 alle Polynome zu 0 reduziert werden. Somit ist B' stets leer, was bedeutet, dass in die d_0-Grad-beschränkte σ-SAGBI-Basis, die der Algorithmus berechnet, nur die Polynome aus $\mathcal{V}_{\leq d_0}$ eingefügt werden. Somit ist $\mathcal{V}_{\leq d}$ eine d-Grad-beschränkte SAGBI-Basis und wegen $\mathcal{G}_{\leq d} = \mathcal{V}_{\leq d}$ ist auch $\mathcal{G}_{\leq d}$ eine d-Grad-beschränkte SAGBI-Basis.

□

Aufbauend auf diesem Satz lässt sich analog zum SAGBI-Test ein einfacher Test angeben, der überprüft, ob eine Menge eine d-Grad-beschränkte SAGBI-Basis ist. Dieser Test ist in der CoCoA-Funktion `SB.IsTruncSagbi` zu finden (siehe Anhang C, Seite 301).

Algorithmus 5.7 : d-Grad-beschränkte SAGBI-Basis-Test

Input : Eine Menge $\{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ normierter, homogener und nicht konstanter Polynome.
Input : $d \in \mathbb{N}_+$.
Result : TRUE, falls $g_1, \ldots, g_s$ eine d-Grad-beschränkte σ-SAGBI-Basis bilden.

```
1 G = (g_1, ..., g_s);
2 Berechne mit Algorithmus 2.2 ein Erzeugendensystem B ⊆ K[y_1, ..., y_s] von Rel(LT_σ(G));
3 B := B_{≤d};
4 foreach b ∈ B do
5     if NRS_G(b(g_1, ..., g_s) ≠ 0 then
6         return FALSE;
7 return TRUE;
```

In Theorem 5.2.13 haben wir gesehen, dass die Unteralgebra-Ersetzungsregel $\xrightarrow{G}_{\mathrm{S}}$ genau dann konfluent ist, wenn G eine σ-SAGBI-Basis von S ist. Wir erhalten also bei einer Unteralgebra-Reduktion nur dann ein eindeutiges Polynom, wenn wir mit einer σ-SAGBI-Basis reduzieren (vgl. dazu auch Beispiel 5.2.12). Will man das in die Praxis umsetzen, stößt man unter Umständen unweigerlich auf ein Problem: Sollte S keine endliche σ-SAGBI-Basis besitzen, ist die SAGBI-Normalform so nicht berechenbar. Im homogenen Fall bieten die d-Grad-beschränkten σ-SAGBI-Basen allerdings einen Ausweg, wie uns der nächste Satz zeigen wird.

Satz 5.4.17. *Sei S eine von homogenen Polynomen endlich erzeugte K-Unteralgebra von P. Für $d \in \mathbb{N}_+$ sei $G_{\leq d} = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine d-Grad-beschränkte σ-SAGBI-Basis von S und sei $\mathcal{G}_{\leq d} = (g_1, \ldots, g_s)$ aufsteigend geordnet. Dann gilt $\mathrm{NF}_S(f) = \mathrm{NRS}_{\mathcal{G}_{\leq d}}(f)$ für alle homogenen Polynome $f \in P$ vom Grad $\leq d$.*

Beweis: Sei $H \subseteq P \setminus \{0\}$ eine homogene σ-SAGBI-Basis von S mit $H_{\leq d} = G_{\leq d}$. Da f homogen vom Grad $\leq d$ ist, gilt entweder $\mathrm{NF}_S(f) = 0$ oder $\mathrm{NF}_S(f)$ ist auch homogen vom Grad $\leq d$. Weiter ist $\mathrm{NF}_S(f)$ das bzgl. $\xrightarrow{H}_{\mathrm{S}}$ eindeutig bestimmte Polynom mit $f \xrightarrow{H}_{\mathrm{S}} \mathrm{NF}_S(f)$. Zur Reduktion von f kommen also nur Polynome vom Grad $\leq d$ in Frage. Somit gilt schließlich $\mathrm{NF}_S(f) = \mathrm{NRS}_{\mathcal{G}_{\leq d}}(f)$. □

Aus diesem Satz folgt nun sofort ein einfacher und effizienter Unteralgebra-Mitgliedschaftstest im homogenen Fall. Wir hatten bisher bereits einen Unteralgebra-Mitgliedschaftstest mit Hilfe der Normalform kennengelernt. Dazu ist allerdings eine σ-SAGBI-Basis notwendig, und um diesen Test auch praktisch umsetzen zu können, sogar eine endliche σ-SAGBI-Basis. Im homogenen Fall spielt die Existenz einer endlichen σ-SAGBI-Basis keine Rolle. Zu einem gegebenen homogenen Polynom vom Grad $d \in \mathbb{N}_+$ ist lediglich eine d-Grad-beschränkte σ-SAGBI-Basis zu berechnen und dann folgendes Korollar anzuwenden, das sofort aus dem letzten Satz folgt.

Korollar 5.4.18. (Homogener Unteralgebra-Mitgliedschaftstest)
Sei S eine von homogenen Polynomen endlich erzeugte K-Unteralgebra von P. Für $d \in \mathbb{N}_+$ sei $G_{\leq d} = \{g_1, \ldots, g_s\} \subseteq P \setminus \{0\}$ eine d-Grad-beschränkte σ-SAGBI-Basis von S und sei $\mathcal{G}_{\leq d}$ das zugehörige, aufsteigend geordnete Tupel. Genau dann ist ein homogenes Polynom $f \in P$ vom Grad $\leq d$ Element von S, wenn $\mathrm{NRS}_{\mathcal{G}_{\leq d}}(f) = 0$ *gilt.*

Auch dieser Mitgliedschaftstest ist mit der CoCoA-Funktion `SB.HomIsInSubalg` das ApCoCoA-Pakets `sagbi.cpkg` zu finden (siehe Anhang C, Seite 298). Analog zu Bemerkung 5.3.9 lässt sich dann sofort mit Hilfe des Unteralgebra-Divisionsalgorithmus eine explizite Darstellung eines homogenen Polynoms $f \in S$ angeben. Diese explizite Darstellung kann mit der CoCoA-Funktion `SB.HomSubalgRepr` berechnet werden (siehe Anhang C, Seite 299). Zum Abschluss dieses Kapitels wollen wir den homogenen Unteralgebra-Mitgliedschaftstest an Beispielen anwenden.

Beispiel 5.4.19. Sei stets $P = \mathbb{Q}[x_1, x_2]$ und $\sigma = \mathrm{DegLex}$. Seien $g_1 := x_1 - x_2$, $g_2 := x_1x_2 - x_2^2$ sowie $g_3 := x_1x_2^2$. Dann gilt zunächst

$$g_1g_2^2 - g_1^2g_3 + g_2g_3 = x_1x_2^4 - x_2^5.$$

Somit ist das Polynom $f := x_1x_2^4 - x_2^5$ ein Element der $\mathbb{Q}$-Unteralgebra $S := \mathbb{Q}[g_1, g_2, g_3]$. Dies lässt sich auch mit Hilfe d-Grad-beschränkter σ-SAGBI-Basen zeigen. Die Menge $G_{\leq 5} = \{g_1, g_2, g_3, x_1x_2^3 - x_2^4, x_1x_2^4 - x_2^5\}$ ist eine 5-Grad-beschränkte σ-SAGBI-Basis von S. Wegen $f \in G_{\leq 5}$ folgt sofort $f \in S$.

Wir betrachten nun das Polynom $f := g_1^3g_3 - 2g_2^3 = x_1^4x_2^2 - 5x_1^3x_2^3 + 9x_1^2x_2^4 - 7x_1x_2^5 + 2x_2^6 \in S$. Eine 6-Grad-beschränkte σ-SAGBI-Basis von S lautet wie folgt:

$$G_{\leq 6} = \{g_1, g_2, g_3, x_1x_2^3 - x_2^4, x_1x_2^4 - x_2^5, x_1x_2^5 - \tfrac{1}{2}x_2^6\}$$

In diesem Fall ist nicht mehr so leicht anhand von $G_{\leq 6}$ zu sehen, dass $f \in S$ gilt. Wegen $\mathrm{NRS}_{\mathcal{G}_{\leq 6}}(f) = 0$ erhalten wir auch auf diese Weise das korrekte Ergebnis $f \in S$. ◁

KAPITEL 6

Die Theorie der Invarianten

Paul Albert GORDAN[14]

Das ist keine Mathematik,
das ist Theologie!

Der berühmte Ausspruch „Das ist keine Mathematik, das ist Theologie!", der dem 1837 in Breslau (Polen, damals Deutsches Reich) geborenen deutschen Mathematiker Paul Albert GORDAN (1837–1912) zugeschrieben wird, war dessen angebliche Reaktion auf den Beweis des berühmten Endlichkeitssatzes von David HILBERT (1862–1943) (siehe Theorem 6.2.5). Da GORDAN die Beweistechniken von HILBERT in späteren Veröffentlichungen mehrmals aufgegriffen hat, lässt sich vermuten, dass der als „König der Invarianten" bekannte GORDAN mit diesem Spruch wohl seine Bewunderung ob des Beweises von HILBERT zum Ausdruck bringen wollte.

6.1 Historische Entwicklung der Invariantentheorie

Die Geschichte hinter diesem Zitat und damit verbunden die Geschichte der Mathematischen Disziplin der Invariantentheorie soll zu Beginn dieses Kapitels kurz beleuchtet werden. Dazu betrachten wir zunächst alle Polynome der Form

$$a \cdot x^2 + 2b \cdot xy + c \cdot y^2 \qquad (*)$$

im Polynomring $\mathbb{C}[x, y]$, die sogenannten **binären quadratischen Formen** oder kurz **Binärformen**, deren Menge wir mit V_2 abkürzen. Ersetzen wir nun x durch $x + y$, so erhalten wir das Polynom

$$a \cdot x^2 + 2(a + b) \cdot xy + (a + 2b + c) \cdot y^2.$$

Mit $a' := a$, $b' := a + b$ und $c' := a + 2b + c$ liegt also erneut ein Polynom der Form $(*)$ vor. Die Koeffizienten a, b, c und a', b', c' erfüllen dabei folgenden Zusammenhang:

$$a'c' - b'^2 = a(a + 2b + c) - (a + b)^2 = a^2 + 2ab + ac - a^2 - 2ab - b^2 = ac - b^2,$$

[14] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Gordan.html vom 14.08.2013.

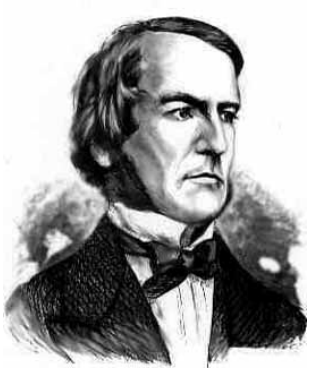
d.h. die sogenannte **Diskriminante**, oder auch **Determinante** genannt, der Binärform bleibt unverändert, sie verhält sich also **invariant** (lat. *invariare*: sich nicht verändern) unter dieser Transformation. Dies lässt sich noch etwas allgemeiner formulieren: Sei für $\lambda \in \mathbb{C}$ die Abbildung $T_\lambda^* : V_2 \to V_2$ definiert durch

$$T_\lambda^*(p) = p(x + \lambda \cdot y, y),$$

dann ist die Diskriminante von p und $T_\lambda^*(p)$ stets gleich.

Joseph-Louis Lagrange[15]

Diese Beobachtung machte der französische Mathematiker Joseph-Louis Lagrange (1736–1813) im Jahre 1773 (vgl. [Neu07]). Man könnte dies als die heimliche Geburtsstunde der Invariantentheorie bezeichnen oder, wie es der deutsche Mathematiker Friedrich Wilhem Franz Meyer (1856–1934) in [Mey98] nannte, einen „Keim" der Invariantentheorie. Von nun an war das Interesse an sich invariant verhaltenden Größen geweckt. Johann Carl Friedrich Gauss (1777–1855) nahm die Beobachtungen von Lagrange auf und verallgemeinerte diese weiter, indem er allgemeine lineare Substitutionen der Variablen von binären und auch ternären quadratischen Formen, also homogener Polynome in drei Unbestimmten vom Grad zwei, behandelte. Die Ergebnisse veröffentlichte Gauss in seinem 1801 erschienenen Buch *Disquisitiones Arithmeticae* (siehe [Gau01]). Weitere „Keime" der Invariantentheorie finden sich im Produktsatz (vgl. [Gan86], S. 27) von Jacques Philippe Marie Binet (1786–1856) und Augustin-Louis Cauchy (1789–1857) über die Determinante des Produkts zweier Matrizen, bei orthogonalen Transformationen von quadratischen Formen in Summen von Quadraten oder in der zu dieser Zeit enstandenen projektiven Geometrie (vgl. [Mey98] 1.).

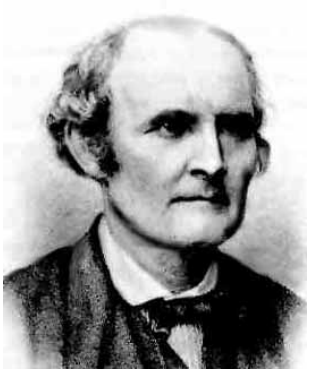


George Boole[16]

Das eigentliche Geburtsjahr der Invariantentheorie datiert Meyer 1891 in seinem *Bericht über den gegenwärtigen Stand der Invariantentheorie* für den Jahresbericht der 1890 gegründeten Deutschen Mathematiker Vereinigung (kurz DMV) (siehe [Mey92]) auf das Jahr 1841. In diesem Jahr veröffentlichte der britische Mathematiker George Boole (1815–1864) eine Arbeit über die Theorie linearer Transformationen (siehe [Boo41]), die als das Fundament der Invariantentheorie angesehen wird (vgl. [Fis66] und [Wol08]). Neben Boole prägten viele der bekanntesten Mathematiker dieser Zeit die ersten Jahrzehnte der Invariantentheorie: Unter anderem sind hier Ludwig Otto Hesse (1811–1874), James Joseph Sylvester (1814–1897), Siegfried Heinrich Aronhold (1819–1884), Arthur Cayley (1821–1895), Charles Hermite (1822–1901), Ferdinand Gotthold Max Eisenstein (1823–1852), Rudolf Friedrich Alfred Clebsch (1833–1872), Gordan, Marius Sophus Lie (1842–1899), George Salmon (1819–1904) und Felix Christian Klein (1849–1925) zu nennen. Einer der ersten, der die Ideen von Boole aufnahm, war Cayley. Er begann 1843, die Invarianten homogener Polynome vom Grad n unter linearen Transformationen zu berechnen (vgl. [Cay45]). 1856 veröffentlichte Cayley eine Arbeit über die Invarianten in zwei Unbestimmten: Er behauptete, dass es unterhalb einer bestimmten Gradschranke endlich viele unabhängige Invarianten gibt und darüber unendlich viele (vgl. [Fis66]); eine Behauptung, die sich später jedoch als falsch herausstellen sollte (siehe auch [Cri86] und

[15] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Lagrange.html vom 14.08.2013
[16] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Boole.html vom 14.08.2013

[Cri88]). Darüber hinaus entwickelte Cayley den sogenannten „Hyperdeterminantenkalkül", der es ermöglichte, beliebig viele Invarianten zu erzeugen (vgl. [Kra85], S. 1). Neben Cayley sind in den ersten Jahren der Invariantentheorie besonders Sylvester und Salmon aktiv, die in der Invariantentheorie den Schlüssel zu einer modernen Algebra sahen (vgl. [Fis66] und [Sal66]).



Arthur Cayley[17]

In Deutschland griffen in erster Linie Aronhold und Clebsch die Ideen von Cayley auf und entwickelten symbolische Methoden zur Erzeugung und Berechnung von Invarianten homogener Polynome (vgl. [Fis66]). Ab dem Jahre 1860 führte Clebsch seinen Schüler Paul Gordan in die Theorie der Invarianten ein. Beide lieferten zahlreiche Beiträge zur Theorie der Invarianten - häufig auch in gemeinsamer Arbeit, u.a. zeigte Gordan 1868, dass Cayley's Behauptung mit den Gradschranken falsch war. Er konnte beweisen, dass es für homogene Polynome in zwei Unbestimmten vom Grad n immer nur endlich viele unabhängige Invarianten gibt, d.h. es gibt ein endliches Erzeugendensystem für die Menge dieser Invarianten; ein Ergebnis, das auch seinen Namen trägt - der Endlichkeitssatz von Gordan (vgl. [Gor85]). Die Frage, ob ein endliches Erzeugendensystem für eine Menge spezieller Invarianten existiert, ist bis in die heutige Zeit hinein stets die drängendste und bedeutendste Frage der Invariantentheorie, mit der auch wir uns im Laufe dieser Arbeit vordergründig beschäftigen werden. Gordan konnte nun zeigen, dass für lineare Koordinantentransformationen auf homogenen Polynomen in zwei Unbestimmten dies der Fall ist. Mit seinem konstruktiven Beweis lieferte er zugleich eine Methode zur Bestimmung der Elemente dieses endlichen Erzeugendensystems, die wiederum auf Techniken von Aronhold und Clebsch basiert. Allerdings ist dieser äußerst komplizierte Beweis durch aufwendige Rechnungen und kombinatorische Methoden geprägt (vgl. [Kra85], S. 1).

David Hilbert[18]

Gordan's Beweis glich einer Art „Initialzündung"; immer mehr Mathematiker seiner Zeit begannen, sich mit der Theorie von Invarianten zu beschäftigen, auch der um 1900 vielleicht bedeutendste Mathematiker, der in Königsberg (damals Ostpreußen) geborene David Hilbert (1862–1943). Die Invariantentheorie näherte sich damit langsam ihrer Blütezeit. Gordan selbst konnte die Existenz von endlichen Erzeugendensystemen in vielen Spezialfällen zeigen und entwickelte zahlreiche Methoden (vgl. [Gor85]) auf dem Weg zu seinem großen Ziel: Den Beweis eines endlichen Erzeugendensystems der Invarianten linearer Transformationen auf homogenen Polynomen in beliebig vielen Unbestimmten (vgl. [Fis66]). Allerdings blieb Gordan das Erreichen dieses Ziels verwehrt. Er scheiterte an diesem Beweis, obwohl sein wissenschaftliches Wirken und Leben von der Invariantentheorie geprägt war, was ihm bereits zu Lebzeiten den Beinamen „König der Invarianten" einbrachte. Hilbert schaffte 1888 das, was Gordan verwehrt blieb: Der Beweis der Endlichkeit im allgemeinen Fall. Sein Beweis war dabei rein abstrakter Natur, d.h. Hilbert bewies nur die Existenz eines endlichen Erzeugendensystems, ohne eine Methode anzugeben, wie man eine solche finden könne. Er reichte seine Arbeit zur Veröffentlichung in den *Mathematischen Annalen* bei Klein ein, der deren aktueller Herausgeber war.

[17] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Cayley.html vom 14.08.2013
[18] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Hilbert.html vom 14.08.2013

Felix Klein[19]

Klein bat daraufhin seinen Freund Gordan um eine Einschätzung von Hilbert's Beweis, der eine Veröffentlichung in den Mathematischen Annalen mit der Begründung ablehnte, dass der Beweis zu abstrakt und daher für die Mathematischen Annalen ungeeignet sei (vgl. [OR13]). Den Beweis von Hilbert soll Gordan mit dem eingangs erwähnten Zitat „Das ist keine Mathematik, das ist Theologie!" kommentiert haben (vgl. [Fis66]). Hilbert, der durch Zufall von der Ablehnung Gordan's Wind bekam, schrieb mit Nachdruck an Klein, dass er nichts ändern werde. Klein, der sich seinem Freund Gordan durchaus verpflichtet fühlte, erkannte jedoch die Bedeutung von Hilbert's Arbeit und veröffentlichte diese (vgl. [Hil90]) trotz aller Bedenken 1890 in den Mathematischen Annalen (vgl. [OR13]). Außerdem lieferte Hilbert in den folgenden Jahren noch weitere Ergebnisse, mit denen er auch die Kritik von Gordan kontern konnte. So gab er in einer 1893 veröffentlichten Arbeit Methoden an, wie sich Invarianten unter den Operationen der speziellen und allgemeinen linearen Gruppe finden lassen (vgl. [Hil93]).

Mit den beiden Arbeiten von Hilbert (siehe [Hil90] und [Hil93]) hat die Invariantentheorie einen Höhepunkt erreicht. Und gleichzeitig auch ihr vorläufiges Ende? Charles S. Fisher untersuchte 1966 in seinem Aufsatz *The Death of a Mathematical Theory: a Study in the Sociology of Knowledge* (vgl. [Fis66]) diese Frage. Tatsächlich schienen mit dem Endlichkeitssatz von Hilbert die wesentlichen Ziele der Invariantentheorie erreicht zu sein und das Interesse an Invarianten nahm zunehmend ab: Waren es 1890 noch 42 Veröffentlichungen im *Jahrbuch über die Fortschritte der Mathematik*, so ging deren Zahl auf 5 Veröffentlichungen im Jahre 1940 zurück (vgl. [Fis66]). Hilbert widmete in dem Glauben, dass der deutsche Mathematiker Ludwig Maurer (1859–1927) sein Ergebnis sogar noch auf Operationen beliebiger Gruppen verallgemeinern konnte, das 14. seiner berühmten 23 Probleme, die er 1900 beim Internationalen Mathematiker- Kongress in Paris vorgestellt hatte, sogar einer Verallgemeinerung der Frage nach einem endlichen Invariantensystem:

Problem 14: *Nachweis der Endlichkeit gewisser voller Funktionensysteme.*

Die Arbeit von Maurer erwies sich allerdings als falsch, womit die allgemeine Fragestellung auch für die Invariantentheorie nicht geklärt war (vgl. [Kra85], S. 50). Als im Jahre 1959 der japanische Mathematiker Masayoshi Nagata (1927–2008) ein Gegenbeispiel zu Hilberts 14. Problem fand, war auch die allgemeine Endlichkeitsfrage der Invariantentheorie negativ beantwortet, da Nagata's Beispiel auch diesen Spezialfall widerlegte.

Die Invariantentheorie legte nun einen „Dornröschenschlaf" ein, der mehrere Jahrzehnte andauern sollte. Gleichzeitig geben die beiden Arbeiten von Hilbert einen entscheidenden Anstoß zur Entwicklung neuer Mathematischer Disziplinen wie der Kommutativen Algebra und Algebraischen Geometrie (vgl. [DK02], S. 1). So wurden in [Hil90] und [Hil93] bedeutende Sätze wie der Nullstellensatz, der Basissatz und der Syzygiensatz sowie Hilbertreihen eingeführt. Die Resultate, die in den folgenden Jahren in der Kommutativen Algebra sowie der Algebraischen Geometrie erzielt werden, werden später aber auch wieder ihren Beitrag zur Invariantentheorie liefern. Einen weiteren großen Einfluss übt die Invariantentheorie auf die Gruppentheorie aus, wie das bekannte Werk *The Classical Groups. Their Invariants and Representations* (siehe [Wey46]) von Hermann Weyl (1885–1955) zeigt, das den Stand der Invariantentheorie um 1940

[19] Bildquelle: http://www-history.mcs.st-andrews.ac.uk/Biographies/Klein.html vom 14.08.2013

beinhaltet. Aus einer geometrischen Sichtweise heraus entstehen dann ab 1960 wieder erste bedeutende Werke wie *Geometric invariant theory* (siehe [MFK65], 1. Auflage aus dem Jahr 1965) von David Mumford, John Fogarty und Frances C. Kirwan oder *Geometrische Methoden der Invariantentheorie* (siehe [Kra85]) von Hanspeter Kraft, sodass von einem wirklichen Tod der Invariantentheorie nur schwer zu sprechen ist. Besonders deutlich wird der neuerliche Aufschwung der Invariantentheorie in den viel zitierten ersten Zeilen aus einer Arbeit aus dem Jahr 1984 von Joseph Kung und Gian-Carlo Rota (vgl. [KR84]), die auch wir hier aufgreifen wollen.

> „Like the Arabian phoenix rising out of the ashes, the theory of invariants, pronounced dead at the turn of the century, is once again at the forefront of mathematics. During its long eclipse, the language of modern algebra was developed, a sharp tool now at last being applied to the very purpose for which it was invented."

Mit Einführung der Gröbner-Basen 1965 entsteht ein neuer Zweig der Kommutativen Algebra, Computational Commutative Algebra, der wiederum auch der Invariantentheorie in den folgenden Jahrzehnten einen neuen Schub gibt. Insbesondere lassen immer leistungsfähigere Computer Berechnungen möglich werden, die zu Zeiten von Gordan oder Hilbert noch undenkbar schienen (vgl. [DK02], S. 2). Einen Meilenstein auf dem Weg hin zur modernen Invariantentheorie, zur *Computational Invariant Theory*, stellt sicher das Buch [Stu08] von Bernd Sturmfels (geb. 1962) dar, das in erster Auflage bereits 1993 erschienen ist. Durch die Entwicklung neuer Algorithmen wird Invariantentheorie auch zunehmend für Anwendungen interessanter. Eine Reihe von möglichen Anwendungsgebieten listet das 2002 erschienene Buch [DK02] mit dem passenden Titel *Computational Invariant Theory* von Harm Derksen und Gregor Kemper auf (vgl. [DK02], Kapitel 5). Ein weiteres modernes Buch zur Invariantentheorie, das hier Erwähnung finden sollte, ist [Neu07] von Mara D. Neusel, das 2007 erschienen ist. Allerdings behandelt dieses Buch nur die Invariantentheorie endlicher Gruppen, die hier nicht im Fokus liegen wird.

6.2 Der Invariantenring

Bereits in den historischen Einführungen zu Beginn des Kapitels wurde angedeutet, welche Inhalte die Invariantentheorie ausmachen und welche Probleme dabei im Vordergrund stehen. Wir wollen nun zunächst das zentrale Objekt, den Invariantenring, definieren und erste Eigenschaften der Menge der Invarianten untersuchen. Dazu sei im Folgenden stets K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$. Weiter sei G eine linear algebraische Gruppe (vgl. hierzu Abschnitt 4.1), die auf einer affinen K-Varietät X (durch Linksoperation) regulär operiert (vgl. Definition 4.2.5), d.h. X ist eine affine G-Varietät. Die Operation von G auf X lässt sich auf den Koordinatenring $K[X]$ wie folgt fortführen: Durch $a \mapsto (f \mapsto f^a)$ ist eine rationale Darstellung $\rho : G \to \mathrm{Aut}_K(K[X])$ von G in $K[X]$ gegeben, die sogenannte **reguläre Darstellung** von G in $K[X]$ (vgl. Definition 4.3.6 und Satz 4.3.20). Dabei ist f^a das durch $f^a(x) = f(a^{-1}(x))$ für alle $x \in X$ definierte Element von $K[X]$. Die Objekte der Begierde sind nun zunächst die Fixpunkte der durch die reguläre Darstellung induzierten Operation von G auf $K[X]$. Die Fixpunkte selbst und auch die Menge der Fixpunkte werden in der Invariantentheorie allerdings mit einem neuen Namen versehen (vgl. [DK02], S. 39).

Definition 6.2.1. (Invariantenring)
Sei G eine lineare algebraische Gruppe und X eine affine G-Varietät. Die Menge

$$\{f \in K[X] : \ f^a = f \text{ für alle } a \in G\}$$

der Fixpunkte der Operation von G auf $K[X]$ heißt der **Invariantenring** der Operation von G auf $K[X]$ und wird mit $K[X]^G$ bezeichnet. Die Elemente von $K[X]^G$ werden als **invariant** bezeichnet und auch **Invarianten** genannt.

Der Invariantenring ist - wie es der Name vermuten lässt - nicht nur ein kommutativer Ring, sondern hat die algebraische Struktur einer K-Algebra, genauer einer K-Unteralgebra von $K[X]$, d.h. für alle invarianten Elemente $f, g \in K[X]^G$ und alle $\lambda \in K$ gilt $f+g \in K[X]^G$, $f \cdot g \in K[X]^G$ und $\lambda f \in K[X]^G$. Dabei sind Addition und Multiplikation punktweise zu verstehen. Die Elemente des Invariantenrings lassen sich auf anschauliche Weise über die G-Bahnen charakterisieren (vgl. [Kra85], II.3.1., Bemerkung 4, S. 94).

Lemma 6.2.2. (Charakterisierung von Invarianten durch Bahnen)
Sei X eine affine G-Varietät. Eine reguläre Funktion $f \in K[X]$ ist genau dann invariant, wenn f auf den Bahnen konstant ist, d.h. wenn $f(a(x)) = f(x)$ gilt für alle $a \in G$ und alle $x \in X$.

Wir werden in dieser Arbeit und im Folgenden überwiegend den Fall $X = V$ betrachten, wobei V ein endlich-dimensionaler K-Vektorraum ist. In diesem speziellen Fall ist der Koordinatenring $K[V]$ isomorph zum Polynomring $K[x_1, \ldots, x_n]$ in n Unbestimmten, wobei $n = \dim_K(V)$ gilt (vgl. Satz 3.2.7). Wie wir in Kapitel 5 bereits anhand des Beispiels $K[xy^d : \ d \geq 1] \subseteq K[x, y]$ gesehen haben, ist jedoch nicht jede K-Unteralgebra einer K-Algebra endlich erzeugt. Womit wir sofort an der zentralen Fragestellung der Invariantentheorie angelangt sind. Neben der Frage der Existenz eines endlichen Erzeugendensystems beschäftigt sich die Invariantentheorie besonders mit folgenden sogenannten fundamentalen Problemen:

1. Ist der Invariantenring $K[V]^G$ endlich erzeugt? Im Falle einer positiven Antwort wäre es natürlich wünschenswert, ein K-Algebra-Erzeugendensystem von $K[V]^G$ auch explizit angeben zu können. Die Elemente eines solchen Erzeugendensystems werden auch **fundamentale Invarianten** genannt.
2. In dem Fall, dass der Invariantenring endlich erzeugt ist, stellt sich weiter die Frage: Lassen sich die algebraischen Relationen zwischen den fundamentalen Invarianten $f_1, \ldots, f_s$ beschreiben? Eine Beschreibung erhalten wir, sobald wir das Relationenideal $\mathrm{Rel}(f_1, \ldots, f_s)$ (vgl. Definition 2.2.6) berechnet haben.
3. Wie kann man eine beliebige Invariante $g \in K[V]^G$ als Polynom in den fundamentalen Invarianten $f_1, \ldots, f_s$ darstellen? Dies wird insbesondere dann effektiv umsetzbar sein, wenn $K[V]^G$ eine endliche SAGBI-Basis besitzt.

Wir werden uns hier insbesondere mit dem ersten Problem befassen. Wie uns allerdings bekannt ist, existiert eine endliche SAGBI-Basis einer K-Unteralgebra selbst dann nicht immer, wenn die K-Unteralgebra endlich erzeugt ist. In den historischen Ausführungen zu Beginn dieses Kapitels haben wir bereits gesehen, dass die Frage nach einem endlichen Erzeugendensystem eines Invariantentrings stets eine der größten Fragen der Invariantentheorie war.

Offensichtlich ist der Invariantenring $K[V]^G$ isomorph zum Durchschnitt des Unterkörpers $K(x_1, \ldots, x_n)^G$ von $K(x_1, \ldots, x_n)$ mit dem Polynomring $K[x_1, \ldots, x_n]$. Wie bereits erwähnt, widmete David Hilbert eines seiner berühmten 23 Probleme einer Verallgemeinerung der Endlichkeitsfrage von Invariantenringen. In seinem 14. Problem warf er die Frage auf, ob der Durchschnitt eines beliebigen Unterkörpers L von $K(x_1, \ldots, x_n)$ mit $K[x_1, \ldots, x_n]$ stets endlich erzeugt sei. Im Jahre 1959 hatte Masayoshi Nagata diese Frage durch ein Gegenbeispiel (vgl. [DK02], Beispiel 2.1.4, S. 43) negativ beantworten können (vgl. [Nag59]). Wir wollen nun erste Beispiele angeben für endlich erzeugte Invariantenringe, die auch historisch von Bedeutung sind (vgl. [DK02], S. 40 f.).

Beispiel 6.2.3. (Symmetrische Polynome)
Die symmetrische Gruppe S_n operiert auf $V = K^n$ durch Permutation der Koordinaten, d.h. für eine Permutation σ gilt $\sigma(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. Die Menge der Invarianten der Operation von S_n auf $K[x_1, \ldots, x_n]$ sind gerade die symmetrischen Polynome. Ein klassisches Resultat der Invariantentheorie besagt, dass der Invariantenring $K[V]^{S_n}$ endlich erzeugt wird von den sogenannten **elementarsymmetrischen Polynomen**:

$$s_r = \sum_{1 \le i_1 < i_2 < \ldots < i_r \le n} x_{i_1} x_{i_2} \cdots x_{i_r}, \quad 1 \le r \le n.$$

◁

Ein weiteres klassisches Beispiel ist uns bereits begegnet, die Invarianten der Binärformen unter der Operation der speziellen linearen Gruppe.

Beispiel 6.2.4. (Binärformen)
Sei K algebraisch abgeschlossen mit $\mathrm{char}(K) = 0$. Für $d \ge 2$ sei

$$V_d := \{a_0 x^d + a_1 x^{d-1} y + \ldots + a_d y^d : \ a_0, \ldots, a_d \in K\}$$

der K-Vektorraum aller homogenen Polynome vom Grad d in den Unbestimmten x und y. Derartige Polynome werden auch als **Binärformen** bezeichnet. Die Gruppe $G := \mathrm{SL}_2(K)$ operiert auf V_d durch $\mathcal{A}(g)(x, y) = g(\alpha x + \gamma y, \ \beta x + \delta y)$, wobei $\mathcal{A} = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right) \in \mathrm{SL}_2(K)$ gilt. Wie oben angegeben setzt sich die Operation von G via $f \mapsto f^{\mathcal{A}}$ auf den Koordinatenring $K[V_d]$ fort, wobei $f^{\mathcal{A}} \in K[V_d]$ definiert ist durch $f^{\mathcal{A}}(g) = f(\mathcal{A}^{-1}(g))$. Da V_d isomorph ist zu K^{d+1} können wir den Koordinatenring $K[V_d]$ mit $K[a_0, \ldots, a_d]$ identifizieren. Da K algebraisch abgeschlossen ist, lässt sich jede Binärform $g_d = a_0 x^d + a_1 x^{d-1} y + \ldots + a_d y^d$ faktorisieren, d.h. es gibt für alle $i \in \{1, \ldots, d\}$ Koeffizienten $\alpha_i, \beta_i \in K$ mit $g_d = \prod_{i=1}^{d} (\alpha_i x + \beta_i y)$. Nun definieren wir die sogenannte **Diskriminante** von $g_d \in V_d$ durch

$$\Delta(g_d) := \prod_{1 \le i < j \le d} (\alpha_i \beta_j - \beta_i \alpha_j)^2$$

Die Diskriminante $\Delta(g_d)$ ist eine Invariante, also ein Element des Invariantentrings $K[V_d]^G$. Für $d = 2$ wollen wir dies explizit nachweisen: Sei also $g_2 = a_0 x^2 + a_1 xy + a_2 y^2$. Dann gibt es $\alpha_1, \alpha_2, \beta_1, \beta_2 \in K$ mit

$$g_2 = (\alpha_1 x + \beta_1 y)(\alpha_2 x + \beta_2 y) = \alpha_1 \alpha_2 x^2 + (\alpha_1 \beta_2 + \beta_1 \alpha_2) xy + \beta_1 \beta_2 y^2.$$

Durch Koeffizientenvergleich folgt also $a_0 = \alpha_1 \alpha_2$, $a_1 = \alpha_1 \beta_2 + \beta_1 \alpha_2$ und $a_2 = \beta_1 \beta_2$. Für die Diskriminante von g_2 erhalten wir somit

$$\Delta(g_2) = (\alpha_1 \beta_2 + \beta_1 \alpha_2)^2 - 4 \alpha_1 \alpha_2 \beta_1 \beta_2 = a_1^2 - 4 a_0 a_2,$$

die wohl bekannte Diskriminante eines quadratischen Polynoms, wie sie bereits in Schulbüchern zu finden ist. Sei nun $\mathcal{A} \in G$ mit $\mathcal{A}^{-1} = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$. Dann gilt:

$$\begin{aligned}
\mathcal{A}^{-1}(g_2) &= a_0(\alpha x + \gamma y)^2 + a_1(\alpha x + \gamma y)(\beta x + \delta y) + a_2(\beta x + \delta y)^2 \\
&= a_0(\alpha^2 x^2 + 2\alpha\gamma xy + \gamma^2 y^2) + a_1(\alpha\beta x^2 + (\alpha\delta + \beta\gamma) xy + \gamma\delta y^2) \\
&\quad + a_2(\beta^2 x^2 + 2\beta\delta xy + \delta^2 y^2) \\
&= \underbrace{(a_0\alpha^2 + a_1\alpha\beta + a_2\beta^2)}_{=:\tilde{a}_0} x^2 + \underbrace{(2a_0\alpha\gamma + a_1(\alpha\delta + \beta\gamma) + 2a_2\beta\delta)}_{=:\tilde{a}_1} xy \\
&\quad + \underbrace{(a_0\gamma^2 + a_1\gamma\delta + a_2\delta^2)}_{=:\tilde{a}_2} y^2
\end{aligned}$$

und es folgt

$$\begin{aligned}\tilde{a}_1 - 4\tilde{a}_0\tilde{a}_2 &= (2a_0\alpha\gamma + a_1(\alpha\delta+\beta\gamma) + 2a_2\beta\delta)^2 \\ &\quad - 4(a_0\alpha^2 + a_1\alpha\beta + a_2\beta^2)(a_0\gamma^2 + a_1\gamma\delta + a_2\delta^2) \\ &= 8a_0a_2\alpha\beta\gamma\delta + a_1^2\alpha^2\delta^2 - 2a_1^2\alpha\beta\gamma\delta + a_1^2\beta^2\gamma^2 - 4a_0a_2\alpha^2\delta^2 - 4a_0a_2\beta^2\gamma^2 \\ &= a_1^2(\alpha^2\delta^2 - 2\alpha\beta\gamma\delta + \beta^2\gamma^2) - 4a_0a_2(\alpha^2\delta^2 - 2\alpha\beta\gamma\delta + \beta^2\gamma^2) \\ &= (a_1^2 - 4a_0a_2)(\alpha\delta - \beta\gamma)^2 = (a_1^2 - 4a_0a_2)\det(\mathcal{A}^{-1})^2 = a_1 - 4a_0a_2.\end{aligned}$$

Somit gilt $\Delta(g_2) \in K[V_2]^G$. Es lässt sich sogar zeigen, dass $\Delta(g_2)$ den Invariantenring $K[V_2]^G$ erzeugt, d.h. dass $K[V_2]^G = K[\Delta(g_2)]$ gilt. Paul Gordan hat 1868 bewiesen, dass der Invariantenring $K[V_d]^{\mathrm{SL}_2(K)}$ für alle $d \geq 2$ stets endlich erzeugt ist (vgl. [Gor68]). Erzeugendensysteme sind aber nur in wenigen Fällen explizit bekannt. Für $d = 3$ gilt analog zu $d = 2$, dass $K[V_3]^G$ von der Diskriminante

$$\Delta(g_3) = a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3$$

erzeugt wird. Der Invariantenring $K[V_4]^G$ wird von den beiden Polynomen

$$f_2 = a_0a_4 - \frac{1}{4}a_1a_3 + \frac{1}{12}a_2^2 \quad \text{und} \quad f_3 = \det\begin{pmatrix} a_0 & \frac{a_1}{4} & \frac{a_2}{6} \\ \frac{a_1}{4} & \frac{a_2}{6} & \frac{a_3}{4} \\ \frac{a_2}{6} & \frac{a_3}{4} & a_4 \end{pmatrix}$$

erzeugt. Wie bereits erwähnt, ist die Diskriminante $\Delta(g_4)$ auch eine Invariante von $K[V_4]^G$, genauer gilt $\Delta(g_4) = 2^8(f_2^3 - 27f_3^2)$. Für $d \in \{5, 6, 8\}$ sind ebenfalls Erzeugendensysteme bekannt, die allerdings nicht mehr so einfach aufzuschreiben sind (vgl. hierzu [Spr70]).

◁

Wie wir bereits wissen, existiert nicht in jedem Fall ein endliches Erzeugendensystem des Invariantenrings. Es wäre nun natürlich wünschenswert, eine Klasse von Gruppen zu kennen, für die die Existenz eines endlichen Erzeugendensystems garantiert werden kann. An dieser Stelle kommen die uns bereits bekannten linear reduktiven Gruppen ins Spiel (siehe Abschnitt 4.4). Paul Gordan hatte die Existenz eines endlichen Erzeugendensystems für eine spezielle linear reduktive Gruppe, nämlich für die spezielle lineare Gruppe $\mathrm{SL}_2(K)$, in dem speziellen Fall der Operation auf die Binärformen 1868 bewiesen (vgl. Beispiel 6.2.4). Im Jahre 1890 lieferte David Hilbert einen Beweis (vgl. [Hil90]) für alle linear reduktiven Gruppen, die auf endlichdimensionalen Vektorräumen operieren (vgl. [DK02], Theorem 2.2.10, S. 49).

Theorem 6.2.5. (Hilberts Endlichkeitssatz)
Sei G eine linear reduktive Gruppe und $(\rho, V)_G$ eine rationale Darstellung von G in einem endlich-dimensionalen K-Vektorraum V. Dann ist der Invariantenring $K[V]^G$ eine endlich erzeugte K-Unteralgebra des Koordinatenrings $K[V]$.

Aufgrund der großen Bedeutung dieses Theorems werden wir den Beweis explizit angeben, allerdings erst zu einem späteren Zeitpunkt (siehe Seite 142), nachdem wir den sogenannten Reynolds-Operator eingeführt haben (vgl. Abschnitt 7.1) und sobald wir uns mit der Berechnung eines endlichen Algebra-Erzeugendensystems beschäftigen werden. Für den Moment nehmen wir dieses Resultat auch ohne Beweis als gegeben hin.

Dank Korollar 4.3.21 können wir aus diesem Theorem sofort folgern, dass der Invariantenring der regulären Operation einer linear reduktiven Gruppe G auf einer affinen G-Varietät X ebenfalls endlich erzeugt ist (vgl. [DK02], Korollar 2.2.11, S. 49). Was Hilbert für linear reduktive

Gruppen gezeigt hat, wurde 1963 von Masayoshi NAGATA auch für geometrisch reduktive Gruppen bewiesen (vgl. [NM63]), d.h. operiert eine geometrisch reduktive Gruppe regulär auf einer affinen G-Varietät X, so ist $K[X]^G$ ebenfalls endlich erzeugt. Hierbei gilt sogar die Umkehrung: Ist $K[X]^G$ endlich erzeugt, so ist G geometrisch reduktiv. Dieser Satz wurde von Vladimir L. POPOV 1979 bewiesen (vgl. [Pop79]). Da endliche Gruppen, wie wir in Abschnitt 4.4 gesehen haben, insbesondere reduktiv sind, folgt aus dem Theorem von NAGATA sofort die Existenz eines endlichen Erzeugendensystems auch für endliche Gruppen. Dieses Ergebnis ist allerdings schon sehr viel länger bekannt. Emmy NOETHER hat die Endlichkeitsfrage für endliche Gruppen 1916 für den nicht-modularen Fall (vgl. [Noe16]) bewiesen. Dieses Resultat konnte sie selbst 1926 schließlich auf beliebige Körper verallgemeinern (vgl. [Noe26]). Wir wollen uns nun noch etwas ausführlicher mit der Struktur des Invariantenrings beschäftigen.

Der Koordinatenring $K[V]$ ist – wie bereits erwähnt – eine zu $P = K[x_1, \ldots, x_n]$ isomorphe K-Algebra. Sei der Polynomring $P = K[x_1, \ldots, x_n]$ nun stets mit der Standardgraduierung versehen. Dann ist auch $K[V]$ standardgraduiert, d.h. es gilt $K[V] = \bigoplus_{d \geq 0} K[V]_d$. Bei der Operation von G auf $f \in K[V]$ bleibt der Grad von f erhalten. Somit erbt der Invariantenring $K[V]^G$ die Standardgraduierung vom Polynomring $K[V]$, was uns der folgende Satz zeigen wird.

Satz 6.2.6. *Sei G eine linear algebraische Gruppe, die auf einem endlich-dimensionalen K-Vektorraum V regulär operiert. Dann ist der Invariantenring $K[V]^G$ eine standardgraduierte K-Unteralgebra von $K[V]$.*

Beweis: Sei $n := \dim_K(V)$, $\rho : G \to \mathrm{Aut}_K(V)$ eine lineare Darstellung von G in V und sei B eine Basis von V. Es genügt zu zeigen, dass $K[x_1, \ldots, x_n]^G$ eine standardgraduierte K-Unteralgebra von $K[x_1, \ldots, x_n]$ ist. Wegen $V \cong K^n$ folgt die Behauptung.

Sei dazu $f \in K[x_1, \ldots, x_n]^G$ mit $d := \deg(f)$ und sei $a \in G$. Dann operiert G auf f durch die Darstellungsmatrix $\mathcal{M}_B^B(\rho_{a^{-1}}) := (a_{i,j})_{1 \leq i,j \leq n} \in \mathrm{GL}_n(K)$.

Sei $t \in \mathrm{Supp}(f)$ von der Form $t = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ mit $\alpha_1, \ldots, \alpha_n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} t^a(x_1, \ldots, x_n) &= t\left(\mathcal{M}_B^B(\rho_{a^{-1}}) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}}\right) \\ &= t(a_{1,1}x_1 + \ldots + a_{1,n}x_n,\ \ldots,\ a_{n,1}x_1 + \ldots + a_{n,n}x_n) \\ &= (a_{1,1}x_1 + \ldots + a_{1,n}x_n)^{\alpha_1} \cdots (a_{n,1}x_1 + \ldots + a_{n,n}x_n)^{\alpha_n}, \end{aligned}$$

d.h. t^a ist ein homogenes Polynom vom Grad $\deg(t)$. Somit bleibt der Grad von f durch die Operation von G erhalten. Seien nun $f_0, \ldots, f_d \in K[x_1, \ldots, x_n]$ die eindeutig bestimmten homogenen Komponenten von f. Dann gilt $f = f_0 + f_1 + \ldots + f_d$ und $f^a = f_0^a + f_1^a + \ldots + f_d^a$. Wegen $f = f^a$ und der Eindeutigkeit der Darstellung von f bzw. f^a in homogene Komponenten folgt $f_i = f_i^a$ für alle $i \in \{1, \ldots, d\}$. Somit ist $K[x_1, \ldots, x_n]^G$ standardgraduiert. □

Aus diesem Satz folgt nun also mit anderen Worten, dass die Operation von G auf $K[V]$ den Grad erhält und die Graduierung von $K[V]$ auf $K[V]^G$ vererbt wird (vgl. auch [DK02], S. 40). Ist insbesondere $f \in K[V]^G$, so sind auch die homogenen Komponenten von f Invarianten von $K[V]$, also Elemente des Invariantenrings $K[V]^G$.

6.3 Die Invarianten der speziellen und allgemeinen linearen Gruppe

Die linear reduktiven Gruppen, die uns Dank Hilberts Endlichkeitssatz stets die Existenz eines endlichen Erzeugendensystems des Invariantenrings garantieren, werden genau diejenigen

Gruppen sein, die in dieser Arbeit überwiegend betrachtet werden. Insbesondere zwei, auch historisch sehr bedeutsame, Gruppen wollen wir uns nun genauer ansehen: Die allgemeine lineare Gruppe $\mathrm{GL}_n := \mathrm{GL}_n(K)$ und die spezielle lineare Gruppe $\mathrm{SL}_n := \mathrm{SL}_n(K)$. Eine Anwendung der folgenden Resultate, deren Beweise z.B. in [PV94] nachzulesen sind, haben wir in Kapitel 1 bereits kennengelernt. Hier sei K wie bisher ein nicht-endlicher Körper mit $\mathrm{char}(K) = 0$ und V ein n-dimensionaler K-Vektorraum. Um die ohnehin schon etwas komplizierten Sachverhalte einfacher aufschreiben zu können, betrachten wir nur $V := K^n$. Dadurch sparen wir uns die ständig notwendigen Wechsel zwischen Vektoren und Koordinatenvektoren. Die Resultate lassen sich dann natürlich ohne Weiteres auf beliebige n-dimensionale K-Vektorräume übertragen.

Sei $(\rho, V)_{\mathrm{GL}_n}$ mit $\rho : \mathrm{GL}_n \to \mathrm{Aut}_K(V)$, definiert durch $\mathcal{A} \mapsto (v \mapsto \mathcal{A}\cdot v)$, eine lineare Darstellung von GL_n in V, d.h. GL_n operiert durch Linksmultiplikation auf V. Wir betrachten nun den Dualraum V^* von V und die kontragradiente Darstellung $(\rho^*, V^*)_{\mathrm{GL}_n}$ von GL_n in V^* (siehe Bemerkung 4.3.7), d.h. $\rho^* : \mathrm{GL}_n \to \mathrm{Aut}_K(V^*)$ ist definiert durch $\mathcal{A} \mapsto (\psi \mapsto \psi \circ \rho_{\mathcal{A}}^{-1})$, womit die Operation von GL_n auf dem Dualraum V^* erklärt ist. Weiter benötigen wir die sogenannte **duale Paarung** von V und V^*. Dabei handelt es sich um die durch $(v, \psi) \mapsto \psi(v)$ definierte Abbildung $\langle \cdot, \cdot \rangle : V \times V^* \to K$. In der allgemeinen Situation wird nun der Vektorraum $V^r \oplus (V^*)^s$ mit $r, s \in \mathbb{N}$ untersucht. Aus Bemerkung 4.3.7 folgt, dass $\tilde{\rho} : \mathrm{GL}_n \to \mathrm{Aut}_K(V^r \oplus (V^*)^s)$, definiert durch

$$\mathcal{A} \quad \mapsto \quad ((v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) \mapsto (\rho_{\mathcal{A}}(v_1), \ldots, \rho_{\mathcal{A}}(v_r), \rho^*_{\mathcal{A}}(\psi_1), \ldots, \rho^*_{\mathcal{A}}(\psi_s))),$$

eine lineare Darstellung von GL_n in $V^r \oplus (V^*)^s$ ist. Dann operiert GL_n auf dem Koordinatenring $K[V^r \oplus (V^*)^s]$ für alle $\delta \in K[V^r \oplus (V^*)^s]$, $\mathcal{A} \in \mathrm{GL}_n$ und $(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) \in V^r \oplus (V^*)^s$ durch

$$\delta^{\mathcal{A}}(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) = \delta(\tilde{\rho}_{\mathcal{A}^{-1}}(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s)).$$

Für alle $i \in \{1, \ldots, r\}$ und $j \in \{1, \ldots, s\}$ ist die Abbildung $\delta_{i,j} : V^r \otimes (V^*)^s \to K$, definiert durch

$$(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) \mapsto \langle v_i, \psi_j \rangle,$$

eine Invariante von $K[V^r \oplus (V^*)^s]^{\mathrm{GL}_n}$ (vgl. [DK02], Abschnitt 4.4, S. 162). Diese Invarianten $\delta_{i,j}$ erzeugen den Invariantenring, was als das **erste fundamentale Theorem für** GL_n bekannt ist (vgl. [DK02], Theorem 4.4.1, S. 162).

Theorem 6.3.1. (Erstes fundamentales Theorem für GL_n)
Seien $r, s \in \mathbb{N}$. Der Invariantenring $K[V^r \oplus (V^*)^s]^{\mathrm{GL}_n}$ wird erzeugt von den Invarianten $\delta_{i,j}$ mit $i \in \{1, \ldots, r\}$ und $j \in \{1, \ldots, s\}$.

Wenn es ein erstes fundamentales Theorem gibt, muss es natürlich auch ein zweites geben. Das zweite fundamentale Theorem gibt die Relationen zwischen den Invarianten an (vgl. [DK02], Theorem 4.4.3, S. 162).

Theorem 6.3.2. (Zweites fundamentales Theorem für GL_n)
Seien $r, s \in \mathbb{N}$. Das Ideal der algebraischen Relationen der Invarianten $\delta_{i,j}$ mit $i \in \{1, \ldots, r\}$ und $j \in \{1, \ldots, s\}$ wird erzeugt von

$$\det \begin{pmatrix} \delta_{i_1,j_1} & \delta_{i_1,j_2} & \ldots & \delta_{i_1,j_{n+1}} \\ \delta_{i_2,j_1} & \delta_{i_2,j_2} & \ldots & \delta_{i_2,j_{n+1}} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{i_{n+1},j_1} & \delta_{i_{n+1},j_2} & \ldots & \delta_{i_{n+1},j_{n+1}} \end{pmatrix},$$

wobei $1 \leq i_1 < i_2 < \ldots < i_{n+1} \leq r$ und $1 \leq j_1 < j_2 < \ldots < j_{n+1} \leq s$ gilt.

Wir wollen uns nun den Invarianten der speziellen linearen Gruppe SL_n zuwenden. Auch hier werden wir für $r, s \in \mathbb{N}$ den K-Vektorraum $V^r \oplus (V^*)^s$ betrachten. Natürlich erklärt sich die Operation von SL_n auf V, auf V^* sowie auf $V^r \oplus (V^*)^s$ analog zur entsprechenden Operation von GL_n und wir übernehmen hier der Einfachheit halber die jeweiligen linearen Darstellungen. Weiter ist ebenfalls klar, dass die Invarianten $\delta_{i,j}$ bzgl. GL_n auch Invarianten bzgl. SL_n sind. Auf folgende Weise lassen sich weitere Invarianten bzgl. SL_n erzeugen: Seien $r, s \geq n \geq 2$ und seien $i_1, \ldots, i_n \in \{1, \ldots, r\}$ mit $i_1 < \ldots < i_n$ sowie $j_1, \ldots, j_n \in \{1, \ldots, s\}$ mit $j_1 < \ldots < j_n$. Dann ist die Abbildung $V^r \oplus (V^*)^s \to K$ mit

$$(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) \mapsto \det(v_{i_1}\ v_{i_2}\ \ldots\ v_{i_n})$$

eine Invariante von $K[V^r \oplus (V^*)^s]^{\mathrm{SL}_n}$. Diese wird kurz mit $[i_1 i_2 \ldots i_n]$ bezeichnet. Weiter ist die Abbildung $V^r \oplus (V^*)^s \mapsto K$ mit

$$(v_1, \ldots, v_r, \psi_1, \ldots, \psi_s) \mapsto \det(\psi_{j_1}\ \psi_{j_2}\ \ldots\ \psi_{j_n})$$

eine Invariante von $K[V^r \oplus (V^*)^s]^{\mathrm{SL}_n}$. Diese wird mit $|j_1 j_2 \ldots j_n|$ bezeichnet (vgl. jeweils [DK02], Abschnitt 4.4, S. 163). Diese Invarianten erzeugen zusammen mit den bereits bekannten Invarianten bzgl. GL_n den Invariantenring $K[V^r \oplus (V^*)^s]^{\mathrm{SL}_n}$, was analog zu oben als das **erste fundamentale Theorem für** SL_n bekannt ist (vgl. [DK02], Theorem 4.4.4, S. 163).

Theorem 6.3.3. (Erstes fundamentales Theorem für SL_n)
Seien $r, s \in \mathbb{N}$. Der Invariantenring $K[V^r \oplus (V^*)^s]^{\mathrm{SL}_n}$ wird erzeugt von allen Invarianten $\delta_{i,j}$ mit $i \in \{1, \ldots, r\},\ j \in \{1, \ldots, s\}$, allen Invarianten $[i_1 \ldots i_n]$ mit $1 \leq i_1 < \ldots < i_n \leq r$ und allen Invarianten $|j_1 \ldots j_n|$ mit $1 \leq j_1 < \ldots < j_n \leq s$.

Das zweite fundamentale Theorem für SL_n ist nicht ganz einfach anzugeben. So verzichten Kemper und Derksen in [DK02] auch darauf, was wir ebenfalls tun werden. Zu finden ist dies erneut in [PV94]. Den interessanten Spezialfall, der in [DK02] auch angegeben ist, wollen wir aber an dieser Stelle noch erwähnen. Sei dazu $s = 0$, was bedeutet, dass $K[V^r]^{\mathrm{SL}_n}$ dann nur von den Invarianten $[i_1 \ldots i_n]$ mit $1 \leq i_1 < \ldots < i_n \leq r$ erzeugt wird (siehe auch Kapitel 1). In diesem Fall lautet das zweite fundamentale Theorem wie folgt (vgl. [DK02], Theorem 4.4.5, S. 163).

Theorem 6.3.4. (Zweites fundamentales Theorem für SL_n)
Sei $r \in \mathbb{N}$. Das Ideal der algebraischen Relationen der Invarianten $[i_1 \ldots i_n]$ in $K[V^r]^{\mathrm{SL}_n}$ mit $1 \leq i_1 < \ldots < i_r \leq r$ wird erzeugt von

$$\sum_{k=1}^{n+1} (-1)^{k-1} [i_1 \ldots i_{n-1} j_k][j_1 \ldots j_{k-1} j_{k+1} \ldots j_{n+1}],$$

wobei $i_1, \ldots, i_{n-1}, j_1, \ldots, j_{n+1} \in \{1, \ldots, r\}$ gilt mit $i_1 < \ldots < i_{n-1}$ und $j_1 < \ldots < j_{n+1}$.

Diese Relationen sind auch als **Grassman-Plücker-Relationen** bekannt.

6.4 Die Vektorinvarianten

Wir wollen diesen Abschnitt mit einem sehr anschaulichen Beispiel beginnen und ein Dreieck $\triangle = (x_1, x_2, x_3)$ mit den Eckpunkten $x_1, x_2, x_3 \in \mathbb{A}^2_{\mathbb{R}}$ in der euklidischen Ebene $\mathbb{A}^2_{\mathbb{R}}$ betrachten. Nach Wahl eines Ursprungs $\mathcal{O}$ und eines kartesischen Koordinatensystems, weisen wir den

Ecken x_1, x_2 und x_3 dieses Dreiecks die Koordinaten $(x_{1,1}, x_{1,2})$, $(x_{2,1}, x_{2,2})$ und $(x_{3,1}, x_{3,2})$ zu. Nun drehen wir dieses Dreieck um einen bestimmten Winkel entgegen dem Uhrzeigersinn um den Ursprung $\mathcal{O}$ und erhalten ein neues Dreieck mit den Ecken x_1', x_2' und x_3'. Natürlich ist es leicht einzusehen, welche geometrischen Größen von dieser Drehung unberührt bleiben: Die Fläche des Dreiecks, die Abstände der einzelnen Punkte und die Winkel des Dreiecks. Diese geometrischen Größen, die wir so selbstverständlich als invariant unter der Operation der speziellen orthogonalen Gruppe ansehen, lassen sich alle erzeugen durch die Determinanten jeweils zweier Punkte und durch die Skalarprodukte der Punkte untereinander. Wir erhalten also neun Polynome im Koordinatenring $P := \mathbb{R}[x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2}, x_{3,1}, x_{3,2}]$ der Menge $(\mathbb{R}^2)^3$ aller (auch der degenerierten) Dreiecke in der euklidischen Ebene, die uns auf natürliche Weise als *fundamental* erscheinen. Dies sind:

$$f_1 = \det \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = x_{1,1}x_{2,2} - x_{1,2}x_{2,1}$$
$$f_2 = \det \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{3,1} & x_{3,2} \end{pmatrix} = x_{1,1}x_{3,2} - x_{1,2}x_{3,1}$$
$$f_3 = \det \begin{pmatrix} x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{pmatrix} = x_{2,1}x_{3,2} - x_{2,2}x_{3,1}$$

und

$$\begin{aligned}
f_4 &= \langle (x_{1,1}, x_{1,2}), (x_{1,1}, x_{1,2}) \rangle = x_{1,1}^2 + x_{1,2}^2 \\
f_5 &= \langle (x_{1,1}, x_{1,2}), (x_{2,1}, x_{2,2}) \rangle = x_{1,1}x_{2,1} + x_{1,2}x_{2,2} \\
f_6 &= \langle (x_{1,1}, x_{1,2}), (x_{3,1}, x_{3,2}) \rangle = x_{1,1}x_{3,1} + x_{1,2}x_{3,2} \\
f_7 &= \langle (x_{2,1}, x_{2,2}), (x_{2,1}, x_{2,2}) \rangle = x_{2,1}^2 + x_{2,2}^2 \\
f_8 &= \langle (x_{2,1}, x_{2,2}), (x_{3,1}, x_{3,2}) \rangle = x_{2,1}x_{3,1} + x_{2,2}x_{3,2} \\
f_9 &= \langle (x_{3,1}, x_{3,2}), (x_{3,1}, x_{3,2}) \rangle = x_{3,1}^2 + x_{3,2}^2
\end{aligned}$$

Die Invarianz dieser Polynome lässt sich nicht nur geometrisch interpretieren, sondern auch beweisen, was wir in Kürze in allgemeiner Form machen werden. Allerdings wirft diese Beobachtung sofort eine entscheidende Frage auf:

> *Bilden diese neun Polynome ein Algebra-Erzeugendensystem des Invariantenrings $P^{\mathrm{SO}_2(\mathbb{R})}$ als $\mathbb{R}$-Unteralgebra von P?*

Anders formuliert: Sind diese neun Polynome die fundamentalen Invarianten von $P^{\mathrm{SO}_2(\mathbb{R})}$ bzw. lässt sich jedes invariante Polynom $f \in P^{\mathrm{SO}_2(\mathbb{R})}$ als Polynom in diesen neun Polynomen ausdrücken? Es wird eines der Ziele dieses Abschnittes sein, auf diese Frage eine Antwort zu liefern. Darüber hinaus werden wir auch noch weitere Beispiele für Invariantenringe behandeln, die alle zu einer speziellen Klasse von Invariantenringen gehören, zu den sogenannten **Vektorinvarianten**.

In der allgemeinen Situation werden also $m \in \mathbb{N}_+$ Punkte in K^n betrachtet. Diese Punkte werden üblicherweise in einer $m \times n$-Matrix angeordnet, sodass wir die Gruppe G auf dem K-Vektorraum $V = \mathrm{Mat}_{m,n}(K)$ operieren lassen. Als Körper verwenden wir nach wie vor nicht-endliche Körper mit Charakteristik $\mathrm{char}(K) = 0$. Den Koordinatenring $K[V]$ werden wir entsprechend der Anordnung der m Punkte in einer Matrix mit einem Polynomring identifizieren, dessen Unbestimmte diese Anordnung widerspiegeln. So wie es in dem Eingangsbeispiel

bereits angedeutet wurde, identifizieren wir den Koordinatenring $K[V]$ mit dem Polynomring $P := K[x_{1,1}, \ldots, x_{1,n}, \ \ldots, \ x_{m,1}, \ldots, x_{m,n}]$.

In diesem Zusammenhang werden die Invarianten aus P^G auch als **Vektorinvarianten** bezeichnet. Richman hat 1989 in seinem Artikel [Ric89] insbesondere die Vektorinvarianten der (speziellen) orthogonalen Gruppe untersucht, bevor 1995 John P. Dalbec die Arbeit von Richman aufgegriffen hat und die Vektorinvarianten der Euklidischen Gruppe untersucht hat (vgl. [Dal95]). Wir werden hier nun auch zuerst die Vektorinvarianten der speziellen orthogonalen Gruppe angeben. Aufbauend auf diesem Ergebnis lassen sich dann leicht die Vektorinvarianten der orthogonalen Gruppe herleiten. Später werden wir zudem Algorithmen kennenlernen, mit denen sich insbesondere die Vektorinvarianten der (speziellen) orthogonalen Gruppe für konkrete Parameter n, m berechnen lassen.

6.4.1 Die Vektorinvarianten der (speziellen) orthogonalen Gruppe

Gerade die Beweise Richmans über die Vektorinvarianten der speziellen orthogonalen Gruppe SO_n erfordern einiges an Vorarbeit, inklusive zahlreicher neuer Begriffe und sehr technischen Beweisen. Auf das alles wollen wir an dieser Stelle aber großteils nicht eingehen. Die Theorie der Vektorinvarianten bildet aber auch einen Teil dieser Arbeit und wie in Kapitel 1 bereits zu sehen war, sind sie besonders für die Bildverarbeitung von Interesse. Im Rahmen der Erarbeitung dieser Theorie wurde insbesondere der Artikel [Ric89] von Richman über die Vektorinvarianten der speziellen orthogonalen Gruppe neu aufgerollt, anschaulich aufbereitet und z.T. auf andere Weisen bewiesen. Allerdings ist die komplette Theorie dieses Artikels an dieser Stelle zu umfangreich und führt uns zu weit weg von dem Ziel, Invarianten zu berechnen. Dennoch wollen wir die Beweise hinter den folgenden Resultaten dem Leser nicht vorenthalten, sodass wir sie in Anhang A präsentieren. Interessant an diesem Artikel ist auch eine völlig andere Sichtweise auf Polynomringe.

Wir werden im Weiteren die $m \cdot n$ Unbestimmten $x_{1,1}, \ldots, x_{m,n}$ aus P passend zu ihrer Indizierung in einer $m \times n$-Matrix $(x_{i,j})$ anordnen, die wir kurz mit $\mathcal{X}$ bezeichnen. Den i-ten Zeilenvektor von $\mathcal{X}$ schreiben wir entsprechend als x_i, d.h. für alle $i \in \{1, \ldots, m\}$ gilt $x_i = (x_{i,1}, \ldots, x_{i,n})$. Wir bitten also zu beachten, dass in diesem Abschnitt entgegen der Gewohnheit x_i keine Unbestimmte bezeichnet, sondern einen Vektor von Unbestimmten. Die Gruppenoperation der orthogonalen Gruppe $\mathrm{O}_n := \mathrm{O}_n(K)$, und damit auch der speziellen orthogonalen Gruppe $\mathrm{SO}_n := \mathrm{SO}_n(K)$ als Untergruppe von O_n, auf V wird induziert durch die lineare Darstellung $\rho : \mathrm{O}_n \to \mathrm{Aut}_K(V)$ definiert durch $\rho_{\mathcal{A}}(\mathcal{B}) = \mathcal{B} \cdot \mathcal{A}$. Für ein $\mathcal{A} \in \mathrm{O}_n$ lässt sich diese Operation also durch das Matrizenprodukt $\mathcal{X} \cdot \mathcal{A}$ beschreiben. Die i-te Zeile von $\mathcal{X} \cdot \mathcal{A}$ bezeichnen wir dann kurz mit $x_i^{\mathcal{A}}$ und mit $x_{i,j}^{\mathcal{A}}$ den (i,j)-ten Eintrag von $\mathcal{X} \cdot \mathcal{A}$, d.h. $x_i^{\mathcal{A}}$ ist der Koordinatenvektor, der durch die Operation von $\mathcal{A}$ aus x_i hervorgeht. Somit operiert O_n, und damit SO_n, auf einem Polynom $f \in P$ des Koordinatenrings durch $f^{\mathcal{A}} := f(x_{1,1}^{\mathcal{A}}, \ldots, x_{m,n}^{\mathcal{A}})$, also durch Auswerten der Polynome des Koordinatenrings an den „neuen" Koordinaten $x_{i,j}^{\mathcal{A}}$.

Betrachtet man nun die neun Polynome des Eingangsbeispiels genauer, so stellt man fest, dass sie entweder 2-Minoren der Matrix $\mathcal{X}$ oder Elemente der symmetrischen 3×3- Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ sind. Das ist natürlich kein Zufall: Die n-Minoren von $\mathcal{X}$, die Einträge der Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ sowie alle k-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ für $k \in \{1, \ldots, n\}$ sind invariant unter der Operation von SO_n (vgl. Lemma A.2.1). Damit folgt ebenfalls unmittelbar die Invarianz der neun Polynome des Eingangsbeispiels. Die Idee von Richman, ein Erzeugendensystem für den Invariantenring P^{SO_n} nachzuweisen, sieht nun vor, zunächst den Fall $n = 2$ zu untersuchen. Der allgemeine Fall lässt sich dann auf diesen Basisfall zurückführen. Diesem Aufbau folgend werden wir als erstes

Resultat ebenfalls ein Erzeugendensystem des Invariantenrings P^{SO_2} angeben. Der Beweis des Theorems sowie die umfangreiche Theorie, die für diesen Beweis notwendig ist, finden sich im Anhang wieder (siehe Theorem A.2.8)

Theorem 6.4.1. (Erzeugendensystem von P^{SO_2})
Sei K ein nicht-endlicher Körper mit $\mathrm{char}(K) \neq 2$ *und sei $P = K[x_{1,1}, x_{1,2}, \ldots, x_{m,1}, x_{m,2}]$. Dann wird der Invariantenring P^{SO_2} als K-Unteralgebra von P erzeugt von den 2-Minoren von $\mathcal{X}$ und von den Einträgen der Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$, d.h.* $\mathrm{Min}(2, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}})$ *ist ein K-Algebra-Erzeugendensystem von P^{SO_2}.*

Man beachte, dass für $m = 1$ die Menge $\mathrm{Min}(2, \mathcal{X})$ der 2-Minoren von $\mathcal{X}$ nur die Null enthält und somit nichts zum Erzeugendensystem beiträgt. Weiter lässt es sich zeigen, dass dieses Erzeugendensystem eine Lex-SAGBI-Basis von P^{SO_2} ist (vgl. Korollar A.2.9). Die folgende Abbildung wird für uns nun die „Brücke" schlagen zwischen dem speziellen Fall $n = 2$ und dem allgemeinen Fall $n \geq 2$. Dazu sei im Folgenden stets $n \geq 2$ und für $k, \ell \in \{1, \ldots, n\}$ mit $k < \ell$ sei der K-Algebra-Homomorphismus

$$\Phi_{k,\ell} : K[x_{1,1}, \ldots, x_{1,n}, \ldots, x_{m,1}, \ldots, x_{m,n}] \to K[x_{1,1}, x_{1,2}, \ldots, x_{m,1}, x_{m,2}]$$

definiert durch

$$x_{i,j} \mapsto \begin{cases} x_{i,1}, & j = k, \\ x_{i,2}, & j = \ell, \\ 1, & \text{sonst.} \end{cases}$$

Für $n = 2$ gilt stets $k = 1$ und $\ell = 2$, d.h. $\Phi_{k,\ell}$ ist in diesem Fall die Identität. Mit Hilfe dieses Homomorphismus zeigt Richman in [Ric89], dass es für jedes invariante Polynom $f \in K[x_{1,1}, \ldots, x_{m,n}]^{\mathrm{SO}_n}$ ungleich Null und jedes $k \in \{1, \ldots, n-1\}$ auch ein invariantes Polynom $f_k \in K[x_{1,1}, \ldots, x_{m,2}]^{\mathrm{SO}_2}$ mit $f_k \neq 0$ und $\mathrm{LT}_{\mathrm{Lex}}(f_k) = \Phi_{k,k+1}(\mathrm{LT}_{\mathrm{Lex}}(f))$ gibt (vgl. Satz A.3.2). Durch weitere Folgerungen, die auf diesem Resultat aufbauen und deren Beweise sehr umfangreich und technisch sind, gelingt es Richman, folgendes Theorem zu beweisen. Der Beweis und die dazu vorgreifende Theorie können ebenfalls im Anhang nachgelesen werden (vgl. insbesondere Theorem A.3.7).

Theorem 6.4.2. (Lex-SAGBI-Basis von P^{SO_n})
Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $P = K[x_{1,1}, \ldots, x_{m,n}]$. Sei weiter

$$S = \mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}).$$

Dann ist S eine Lex*-SAGBI-Basis von P^{SO_n}.*

Für $n = 2$ stimmt dieses Theorem mit dem bereits bekannten Theorem 6.4.1 überein. Es ist außerdem sehr leicht einzusehen, dass es durchaus kleinere Erzeugendensysteme für P^{SO_n} gibt. Denn offensichtlich ist für $k > 1$ jeder k-Minor von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ ein Polynom in den Einträgen von $X\mathcal{X}^{\mathrm{tr}}$, also in den 1-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$. Somit können wir als Erzeugendensystem von P^{SO_n} bereits die Menge $\mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$ festhalten.

Da die spezielle orthogonale Gruppe eine Untergruppe der orthogonalen Gruppe O_n ist, stellt sich natürlich sofort die Frage, ob sich nicht Teile des Erzeugendensystems von P^{SO_n} auf die orthogonale Gruppe übertragen lassen. In der Tat ist es so, dass am Beweis von Lemma A.2.1 leicht zu erkennen ist, dass die Elemente von $\mathrm{Min}(n, \mathcal{X})$, also die n-Minoren von $\mathcal{X}$, unter der orthogonalen Gruppe nicht mehr invariant sind, die Elemente der Menge

$$\mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$$

jedoch nach wie vor. Durch eine leichte Adaption des Beweises von RICHMAN für das Erzeugendensystem von P^{SO_n} lässt sich zeigen, dass die k-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ den Invariantenring der orthogonalen Gruppe als Lex-SAGBI-Basis erzeugen (vgl. [Dal95], S. 99).

Korollar 6.4.3. (Lex-SAGBI-Basis von P^{O_n})
Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $P = K[x_{1,1}, \ldots, x_{m,n}]$. Sei weiter

$$S = \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}).$$

Dann ist S eine Lex-*SAGBI-Basis von P^{O_n}.*

Auch hier können wir als unmittelbare Folgerung ein kleineres Erzeugendensystem angeben.

Korollar 6.4.4. *Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $P = K[x_{1,1}, \ldots, x_{m,n}]$. Dann wird P^{O_n} als K-Algebra erzeugt von den paarweise verschiedenen Einträgen der Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$.*

6.4.2 Die Vektorinvarianten der Euklidischen Gruppe

Wie wir in Beispiel 4.2.6 gesehen haben, ist die Euklidische Gruppe $\mathrm{Iso}_n := \mathrm{Iso}_n(K)$ isomorph zum semidirekten Produkt der Translationsgruppe $\mathrm{Trans}_n(K)$ und der orthogonalen Gruppe $\mathrm{O}_n(K)$. DALBEC hat in seinem Artikel [Dal95] eine Verbindung hergestellt zwischen den Invarianten der orthogonalen Gruppe und den Invarianten der Euklidischen Gruppe. Dies werden wir hier nun auch tun.

Dazu betrachten wir zu unseren $m \in \mathbb{N}_+$ Punkten in K^n für $n \geq 2$ einen zusätzlichen ausgewählten Punkt. Man denke dabei beispielsweise an die Festlegung eines Ursprungs in der Euklidischen Ebene $\mathbb{A}^2_{\mathbb{R}}$. Wir bezeichnen hier nun die Koordinaten der $m+1$ Punkte mit $y_{1,1}, \ldots, y_{m+1,n}$, um die folgenden Resultate besser von den vorherigen abgrenzen zu können. Der Koordinatenring dieser $m+1$ Punkte im K^n ist dann isomorph zum Polynomring in den Unbestimmten $y_{1,1}, \ldots, y_{m+1,n}$, den wir mit R bezeichnen. Die Unbestimmten lassen sich entsprechend ihrer Indizierung auch hier in einer $(m+1) \times n$-Matrix $\mathcal{Y} := (y_{i,j})$ anordnen. Die i-te Zeile von $\mathcal{Y}$ wird entsprechend mit y_i notiert, d.h. auch hier enthält y_i die Koordinaten des i-ten Punktes. Für ein Element $(\mathcal{A}, v) \in \mathrm{Iso}_n(K)$ der Euklidischen Gruppe setzen wir hier analog $y_i^{(\mathcal{A},v)} := y_i \cdot \mathcal{A} + v$. Mit $y_i^{(\mathcal{A},v)}$ wird der Koordinatenvektor des Punkts bezeichnet, der durch die Operation von $(\mathcal{A}, v)$ aus dem Koordinatenvektor y_i hervorgeht. Die einzelnen Koordinaten von $y_i^{(\mathcal{A},v)}$ werden auch hier mit $y_{i,j}^{(\mathcal{A},v)}$ bezeichnet. Damit operiert die Euklidische Gruppe Iso_n auf einem Polynom $f \in P$ des Koordinatenrings durch

$$f^{(\mathcal{A},v)} := f\left(y_{1,1}^{(\mathcal{A},v)}, \ldots, y_{m+1,n}^{(\mathcal{A},v)}\right),$$

also ebenfalls durch Auswerten an den „neuen" Koordinaten $y_{1,1}^{(\mathcal{A},v)}, \ldots, y_{m+1,n}^{(\mathcal{A},v)}$. Wenn wir nun also unseren ausgewählten Punkt y_{m+1} und die übrigen m Punkte $y_1, \ldots, y_m$ betrachten und uns zunächst die Frage stellen, welche anschaulichen Größen bei einer reinen Translation invariant bleiben, so ist diese Frage nicht allzu schwer zu beantworten: Es sind gerade die Differenzen der einzelnen Koordinaten zu den entsprechenden Koordinaten des ausgewählten, $(m+1)$-ten Punkts, d.h. für alle $j \in \{1, \ldots, n\}$ bleiben

$$y_{1,j} - y_{m+1,j}, \ldots, y_{m,j} - y_{m+1,j}$$

invariant. Hier ist natürlich der Einwand berechtigt, dass dies doch lange nicht alle geometrischen Größen sind, die invariant bleiben. So bleiben z.B. auch die Differenzen der Koordinaten

zwischen zwei beliebigen Punkten invariant. Wie man aber auch sofort sieht, lassen sich diese Größen aus den obigen erzeugen.

Und hier sind wir auch am entscheidenden ersten Punkt angelangt. Wir wollen also beantworten, wie ein Erzeugendensystem der Invarianten der Translationsgruppe $\mathrm{Trans}_n := \mathrm{Trans}_n(K)$ aussieht. Dalbec hat in seinem Artikel [Dal95] bewiesen, dass die Polynome, die aus den obigen Differenzen hervorgehen, eine Lex-SAGBI-Basis des Invariantenrings der Translationsgruppe $\mathrm{Trans}_n(K)$ bilden und diesen somit als K-Algebra erzeugen (vgl. [Dal95], Lemma 2.1).

Satz 6.4.5. (Lex-SAGBI-Basis von R^{Trans_n})
Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $R = K[y_{1,1}, \ldots, y_{m+1,n}]$. Die Menge

$$S := \{y_{1,j} - y_{m+1,j},\ \ldots,\ y_{m,j} - y_{m+1,j} :\ j \in \{1, \ldots, n\}\}$$

ist eine Lex-*SAGBI-Basis des Invariantenrings R^{Trans_n}.*

Wir werden nun zeigen, wie Dalbec eine „Brücke" schlägt zwischen den Invarianten der Translationsgruppe und den Invarianten der orthogonalen Gruppe. Wieder anschaulich gesprochen geht er über von Punkten zu Ortsvektoren zwischen den Punkten. Dazu betrachten wir den K-Algebra-Homomorphismus $\Phi : P \to R$, definiert durch $x_{i,j} \mapsto y_{i,j} - y_{m+1,j}$. Dieser Homomorphismus hat folgende Eigenschaften (vgl. zum Teil [Dal95], Lemma 2.2).

Lemma 6.4.6. *Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und seien $P = K[x_{1,1}, \ldots, x_{m,n}]$ sowie $R = K[y_{1,1}, \ldots, y_{m+1,n}]$. Der K-Algebra-Homomorphismus $\Phi : P \to R$ mit $x_{i,j} \mapsto y_{i,j} - y_{m+1,j}$ ist injektiv und es gilt $\Phi(P^{\mathrm{O}_n}) = R^{\mathrm{Iso}_n}$, d.h. die Invariantenringe P^{O_n} und R^{Iso_n} sind isomorph.*

Dieser Homomorphismus Φ lässt sich nun verwenden, um ein Erzeugendensystem des Invariantenrings bzgl. der Operation der Euklidischen Gruppe anzugeben. Überlegt man auch hier rein anschaulich in $\mathbb{R}^n$, ist es nicht allzu schwer festzustellen, dass die Abstände zwischen zwei Punkten von Bewegungen, also von Operationen der Euklidischen Gruppe, unberührt bleiben. Mit anderen Worten, für alle $i, k \in \{1, \ldots, m+1\}$ sind die Größen

$$d_{i,k} := (y_{i,1} - y_{k,1})^2 + (y_{i,2} - y_{k,2})^2 + \ldots + (y_{i,n} - y_{k,n})^2 = \sum_{j=1}^{n} (y_{i,j} - y_{k,j})^2$$

Invarianten der Euklidischen Gruppe. Diese anschaulichen Überlegen werden wir später natürlich noch beweisen. Man beachte, dass die bei den Invarianten der Translationsgruppe betrachteten Differenzen der einzelnen Koordinaten natürlich nicht mehr invariant sind. Wir betrachten nun weiter die Einträge der symmetrischen Matrix $m \times m$-Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ und bezeichnen für alle $i, k \in \{1, \ldots, m\}$ den (i,k)-ten Eintrag von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ mit $c_{i,k}$, d.h. es gilt:

$$c_{i,k} = x_i \cdot x_k^{\mathrm{tr}} = (x_{i,1}, \ldots, x_{i,n}) \cdot \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,n} \end{pmatrix} = \sum_{j=1}^{n} x_{i,j} x_{k,j}.$$

Daraus ergeben sich folgende Zusammenhänge zu den Differenzen $d_{i,k}$, die im Beweis des folgenden Theorems von Bedeutung sind (vgl. [Dal95], S. 102, dort ohne Beweis).

Lemma 6.4.7. *Sei $\Phi : P \to R$ definiert durch $x_{i,j} \mapsto y_{i,j} - y_{m+1,j}$. Dann gilt für alle $i, k \in \{1, \ldots, m\}$:*

(i) $2\Phi(c_{i,k}) = d_{i,m+1} - d_{i,k} + d_{k,m+1}$,

(ii) $\Phi(c_{i,i} - 2c_{i,k} + c_{k,k}) = d_{i,k}$,

(iii) $\Phi(c_{i,i}) = d_{i,m+1}$.

Beweis: Seien $i, k \in \{1, \ldots, m\}$. Dann gilt:

$$\begin{aligned} d_{i,m+1} - d_{i,k} + d_{k,m+1} &= \sum_{j=1}^{n}(y_{i,j} - y_{m+1,j})^2 - \sum_{j=1}^{n}(y_{i,j} - y_{k,j})^2 + \sum_{j=1}^{n}(y_{k,j} - y_{m+1,j})^2 \\ &= \sum_{j=1}^{n} -2y_{i,j}y_{m+1,j} + 2y_{m+1,j}^2 + 2y_{i,j}y_{k,j} - 2y_{k,j}y_{m+1,j} \\ &= 2\sum_{j=1}^{n}(y_{i,j} - y_{m+1,j})(y_{k,j} - y_{m+1,j}) = 2\sum_{j=1}^{n}\Phi(x_{i,j}x_{k,j}) = 2\Phi(c_{i,k}) \end{aligned}$$

Weiter gilt sofort $\Phi(c_{i,i}) = \sum_{j=1}^{n}\Phi(x_{i,j}^2) = \sum_{j=1}^{n}(y_{i,j} - y_{m+1,j})^2 = d_{i,m+1}$. Mit diesen beiden Gleichungen folgt schließlich auch die zweite:

$$\begin{aligned} \Phi(c_{i,i} - 2c_{i,k} + c_{k,k}) &= \Phi(c_{i,i}) - 2\Phi(c_{i,k}) + \Phi(c_{k,k}) \\ &= d_{i,m+1} - (d_{i,m+1} - d_{i,k} + d_{k,m+1}) + d_{k,m+1} = d_{i,k}. \end{aligned}$$

□

Wie aus dem nächsten Theorem hervorgeht, dessen Beweis wir ausnahmsweise explizit angeben wollen, erzeugen die Polynome $d_{i,k}$ auch den Invariantenring der Euklidischen Gruppe (vgl. [Dal95], Theorem 2.1).

Theorem 6.4.8. (Erzeugendensystem von R^{Iso_n})
Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $R = K[y_{1,1}, \ldots, y_{m+1,n}]$. Die Menge

$$S := \{d_{i,k} : \ 1 \leq i < k \leq m+1\} \subseteq R$$

ist ein Algebra-Erzeugendensystem des Invariantenrings R^{Iso_n} der Euklidischen Gruppe.

Beweis: Sei $h \in R^{\mathrm{Iso}_n}$. Laut Lemma 6.4.6 gibt es genau ein $f \in P^{\mathrm{O}_n}$ mit $\Phi(f) = h$. Aus Korollar 6.4.4 folgt $f \in K[c_{i,k} : \ 1 \leq i \leq k \leq m]$. Gemäß Lemma 6.4.7 ist $h = \Phi(f)$ ein Polynom in den quadrierten Abständen $d_{i,k}$, also ein Element der K-Unteralgebra $K[d_{i,k} : \ 1 \leq i < k \leq m+1]$. Damit folgt $R^{\mathrm{Iso}_n} \subseteq K[S]$. Es bleibt zu zeigen, dass die Elemente von S tatsächlich invariant unter den Operationen von Iso_n sind. Seien $i, k \in \{1, \ldots, m+1\}$ mit $i < k$. Aus Lemma 6.4.7 folgt, dass $f_{i,k} := c_{i,i} - 2c_{i,k} + c_{k,k} \in P$ das Urbild von $d_{i,k} \in R$ unter Φ ist. Wegen $P^{\mathrm{O}_n} = K[c_{i,k} : \ 1 \leq i \leq k \leq m]$ gilt $f_{i,k} \in P^{\mathrm{O}_n}$. Laut Lemma 6.4.6 gilt dann $d_{i,k} = \Phi(f_{i,k}) \in R^{\mathrm{Iso}_n}$. □

Dalbec gibt auch eine Lex-SAGBI-Basis von R^{Iso_n} an. Dazu verwendet er die sogenannten **Cayley-Menger Bideterminanten** aus dem Artikel [Hav91] von Timothy F. Havel, die benannt sind nach dem englischen Mathematiker Arthur Cayley (1821–1895) und dem österreichischen Mathematiker Karl Menger (1902–1985). Dazu betrachtet man die Notation betreffend analog zu den k-Minoren die Matrizen

$$D(i_1, \ldots, i_\ell | k_1, \ldots, k_\ell) := 2\left(-\frac{1}{2}\right)^{\ell} \begin{pmatrix} 0 & 1 & 1 & \ldots & 1 \\ 1 & d_{i_1,k_1} & d_{i_1,k_2} & \ldots & d_{i_1,k_\ell} \\ 1 & d_{i_2,k_1} & d_{i_2,k_2} & \ldots & d_{i_2,k_\ell} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & d_{i_\ell,k_1} & d_{i_\ell,k_2} & \ldots & d_{i_\ell,k_\ell} \end{pmatrix}$$

mit jeweils paarweise verschiedenen $i_1, \ldots, i_\ell, k_1, \ldots, k_\ell \in \{1, \ldots, m+1\}$ und $\ell \geq 2$. Die Determinante solcher Matrizen heißt eine $(\ell+1)\times(\ell+1)$-**Cayley-Menger Bideterminante**, die Determinante der speziellen Matrix $D(i_1, \ldots, i_\ell | i_1, \ldots, i_\ell)$ heißt die **Cayley-Menger Determinante**. Zwischen den Bideterminanten und den ℓ-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ besteht folgender Zusammenhang, d.h. für $\ell \geq 1$ und alle paarweise verschiedenen $i_1, \ldots, i_\ell \in \{1, \ldots, m+1\}$ sowie $k_1, \ldots, k_\ell \in \{1, \ldots, m+1\}$ gilt (vgl. [Dal95], S. 102):

$$\Phi(\det(\mathcal{X}\mathcal{X}^{\mathrm{tr}}(i_1, \ldots, i_\ell | k_1, \ldots, k_\ell))) = \det(D(i_1, \ldots, i_\ell, m+1 | k_1, \ldots, k_\ell, m+1)). \tag{6.4.1}$$

Damit lässt sich für R^{Iso_n} folgende Lex-SAGBI-Basis nachweisen (vgl. [Dal95], Theorem 2.2, S. 103).

Theorem 6.4.9. (Lex-SAGBI-Basis von R^{Iso_n})
Sei $n \in \mathbb{N}$ mit $n \geq 2$, sei $m \in \mathbb{N}_+$ und $R = K[y_{1,1}, \ldots, y_{m+1,n}]$. Die $(\ell+1)\times(\ell+1)$ Cayley-Menger *Bideterminanten mit $2 \leq \ell \leq n+1$ bilden eine* Lex-*SAGBI-Basis des Invariantenrings R^{Iso_n} der Euklidischen Gruppe.*

Hier ein Beispiel anzugeben, ist leider schier unmöglich, denn selbst für kleine Werte von m und n, d.h. für wenige Punkte in nicht besonders hoch dimensionalen Räumen, enthält die SAGBI-Basis zu viele Elemente mit viel zu „großen" Polynomen, um sie irgendwie sinnvoll darstellen zu können. So enthält die SAGBI-Basis von R^{Iso_n} beispielsweise für $m = 4$ und $n = 3$ bereits 230 Elemente.

6.5 Homogene Parametersysteme

Bevor wir uns mit homogenen Parametersystemen beschäftigen können, müssen wir einen anderen Begriff betrachten, der später bei der Berechnung von Erzeugendensystemen auch eine große Rolle spielen wird: Der Hilbertsche Nullstellenkegel. Dazu sei im Folgenden G stets eine linear reduktive Gruppe und $(\rho, V)_G$ eine rationale Darstellung von G in einem K-Vektorraum V. Dann können wir den Hilbertschen Nullstellenkegel wie folgt definieren (vgl. [DK02], Definition 2.4.1, S. 60).

Definition 6.5.1. (Hilbertscher Nullstellenkegel)
Die Nullstellenmenge der Menge aller homogenen Invarianten positiven Grades,

$$\{v \in V : \ f(v) = 0 \text{ für alle } f \in K[V]_+^G\},$$

heißt der **Hilbertsche Nullstellenkegel** bzgl. der Operation von G und wird mit $\mathcal{N}_{V,G}$ bezeichnet oder kurz mit $\mathcal{N}_V$, falls die Gruppe aus dem Zusammenhang hervorgeht.

Der Hilbertsche Nullstellenkegel lässt sich anschaulich auf folgende Weise interpretieren (vgl. [DK02], Lemma 2.4.2, S. 60).

Satz 6.5.2. (Interpretation des Hilbertschen Nullstellenkegels)
Der Hilbertsche Nullstellenkegel $\mathcal{N}_V$ ist die Menge aller Vektoren $v \in V$ mit $0 \in \overline{G(v)}$.

Für endliche Gruppen besteht der Hilbertsche Nullstellenkegel somit nur aus dem Nullvektor (vgl. [DK02], S. 60). Es lässt sich weiter stets für alle $v \in \mathcal{N}_{V,G}$ eine multiplikative Untergruppe $H \subseteq G$ finden so, dass v in $\overline{H(v)}$ enthalten ist. Dies ist auch als **Hilbert-Mumford-Kriterium**

bekannt (vgl. [Hil90] und [MFK94]). Es ermöglicht in vielen Situation eine Entscheidung darüber, welche Bahnen im Hilbertschen Nullstellenkegel liegen (vgl. [DK02], S. 60). Wir führen nun noch den folgenden Satz an, der uns weitere Informationen über die algebraische Struktur des Invariantenrings liefert (Teil a) vgl. [DK02], Lemma 2.4.5, S. 60).

Satz 6.5.3. *Seien $f_1, \ldots, f_r \in K[V]^G$ homogene Invarianten.*

a) Gilt $\mathcal{N}_{V,G} = \mathcal{Z}_K(\langle f_1, \ldots, f_r\rangle)$, so ist $K[V]^G$ ein endlich erzeugter $K[f_1, \ldots, f_r]$-Modul.

b) Sind $f_1, \ldots, f_r \in K[V]^G_+$ und ist $K[V]^G$ ein endlich erzeugter $K[f_1, \ldots, f_r]$-Modul, so gilt $\mathcal{N}_{V,G} = \mathcal{Z}_K(\langle f_1, \ldots, f_r\rangle)$.

Beweis: Wir beweisen hier nur Teil b) des Satzes. Da $K[V]^G$ als Modul endlich erzeugt ist über $F := K[f_1, \ldots, f_r]$, gibt es endlich viele $g_1, \ldots, g_s \in K[V]^G$ mit

$$K[V]^G = Fg_1 + \ldots + Fg_s. \qquad (*)$$

Die Inklusion $\mathcal{N}_{V,G} \subseteq \mathcal{Z}_K(\langle f_1, \ldots, f_r\rangle)$ ist wegen $f_1, \ldots, f_r \in K[V]^G_+$ sofort klar. Sie nun umgekehrt $v \in \mathcal{Z}_K(\langle f_1, \ldots, f_r\rangle)$ und sei $f \in K[V]^G_+$. Wegen $(*)$ gibt es Polynome $h_1, \ldots, h_s \in F$ mit $f = h_1g_1 + \ldots + h_sg_s$. Wegen $v \in \mathcal{Z}_K(\langle f_1, \ldots, f_r\rangle)$ folgt dann $h_1(v) = \ldots = h_s(v) = 0$ und damit auch $f(v) = 0$, d.h. es gilt $v \in \mathcal{N}_{V,G}$. □

Damit wollen wir nun betrachten, was ein homogenes Parametersystem ist und welche Eigenschaften dieses hat. Zunächst werden ein homogenes Parametersystem in folgendem Sinne definieren (vgl. [DK02], Definition 2.4.6, S. 61).

Definition 6.5.4. (Homogenes Parametersystem)
Sei $R = \bigoplus_{d=0}^{\infty} R_d$ eine positiv graduierte K-Algebra mit $R_0 = K$. Eine Menge $\{f_1, \ldots, f_r\} \subseteq R$ homogener Elemente heißt ein **homogenes Parametersystem von** R, wenn gilt:

(i) $f_1, \ldots, f_r$ sind algebraisch unabhängig,

(ii) R ist ein endlich-erzeugter $K[f_1, \ldots, f_r]$-Modul.

Es stellt sich natürlich sofort die Frage, ob ein Invariantenring stets ein homogenes Parametersystem besitzt, der wir nun nachgehen wollen. Dazu betrachten wir zunächst das **Noethersche Normalisierungslemma** (vgl. [DK02], Lemma 2.4.7, S. 61).

Theorem 6.5.5. (Noethersches Normalisierungslemma)
Sei K ein unendlicher Körper und sei $R = \bigoplus_{d=0}^{\infty} R_d$ eine positiv graduierte K-Algebra mit $R_0 = K$. Sei $d \in \mathbb{N}$, seien $f_1, \ldots, f_r \in R_d$ und sei R ein endlich-erzeugter $K[f_1, \ldots, f_r]$-Modul. Dann gibt es $g_1, \ldots, g_s \in R_d$, die K-Linearkombinationen der $f_1, \ldots, f_r$ sind und ein homogenes Parametersystem von R bilden.

Aus diesem Theorem folgt nun sofort, dass jeder endlich erzeugte, positiv graduierte Ring ein homogenes Parametersystem besitzt, was wir in folgendem Korollar festhalten (vgl. [DK02], Korollar 2.4.8, S. 61).

Korollar 6.5.6. *Sei $R = \bigoplus_{d=0}^{\infty} R_d$ eine endlich erzeugte, positiv graduierte K-Algebra mit $R_0 = K$. Dann besitzt R ein homogenes Parametersystem.*

Somit besitzt der Invariantenring $K[V]^G$ als standardgraduierter Ring für linear reduktive Gruppen G stets ein homogenes Parametersystem. In diesem Fall ist $K[V]^G$ ein endlich erzeugter $K[f_1, \ldots, f_r]$-Modul. Die Elemente dieses Systems erhalten in diesem Fall einen neuen Namen (vgl. [DK02], S. 61).

Definition 6.5.7. (Primäre und sekundäre Invarianten)
Die Elemente $f_1, \ldots, f_r \in K[V]^G$ eines homogenen Parametersystems $\{f_1, \ldots, f_r\} \subseteq K[V]^G$ von $K[V]^G$ heißen **primäre Invarianten** und Invarianten $g_1, \ldots, g_s \in K[V]^G$ mit

$$K[V]^G = Fg_1 + \ldots + Fg_s$$

für $F = K[f_1, \ldots, f_r]$ heißen **sekundäre Invarianten**.

Wir haben nun unter anderem gesehen, dass $K[V]^G$ ein endlich erzeugter F-Modul ist mit $F = K[f_1, \ldots, f_r]$, wenn $f_1, \ldots, f_r \in K[V]^G$ primäre Invarianten sind. Die Frage, die wir im Folgenden noch kurz beantworten wollen, ist folgende: Wann ist $K[V]^G$ ein *freier* Modul über F? Der Schlüssel zur Beantwortung dieser Frage ist die sogenannte **Cohen-Macaulay-Eigenschaft** (vgl. [DK02], Definition 2.5.1, S. 62).

Definition 6.5.8. (Cohen-Macaulay-Eigenschaft)
Sei R ein Noetherscher Ring und M ein endlich erzeugter R-Modul.

a) Ein Element $f \in R$ heißt **regulär**, wenn für alle Elemente $m \in M$ mit $f \cdot m = 0$ stets $m = 0$ folgt. Eine Folge $f_1, \ldots, f_k \in R$ heißt M**-regulär**, wenn $M/\langle f_1, \ldots, f_k\rangle M \neq 0$ gilt und wenn für alle $i \in \{1, \ldots, k\}$ die Multiplikation mit f_i eine injektive Abbildung auf $M/\langle f_1, \ldots, f_{i-1}\rangle M$ induziert.

b) Sei $I \subseteq R$ ein Ideal mit $IM \neq M$. Dann heißt die maximale Länge k einer M-regulären Folge $f_1, \ldots, f_k \in I$ die **Tiefe von** I **auf** M. Sie wird mit $\operatorname{depth}(I, M)$ bezeichnet.

c) Ist R ein lokaler Ring mit maximalem Ideal $\mathfrak{m}$, so heißt M **Cohen-Macaulay**, wenn $\operatorname{depth}(\mathfrak{m}, M) = \operatorname{Kdim}(R/\operatorname{Ann}(M))$ gilt. Ist R kein lokaler Ring, so heißt M **Cohen-Macaulay**, wenn für alle maximalen Ideale $\mathfrak{m} \subseteq R$ mit $M_\mathfrak{m} \neq 0$ der Modul $M_\mathfrak{m}$ Cohen-Macaulay ist als $R_\mathfrak{m}$-Modul.

d) R heißt **Cohen-Macaulay**, wenn R als R-Modul im Sinne von c) Cohen-Macaulay ist.

Beispielsweise hat jeder Polynomring die Cohen-Macaulay-Eigenschaft (vgl. [DK02], Lemma 2.5.2, S. 63). Wie ein bedeutendes Resultat von Melvin Hochster und Joel L. Roberts aus dem Jahr 1974 besagt, hat auch der Invariantenring $K[V]^G$ die Cohen-Macaulay-Eigenschaft, wenn G eine linear reduktive Gruppe ist (vgl. im Original [HR74] oder [DK02], Theorem 2.5.5, S. 64).

Theorem 6.5.9. (Hochster und Roberts)
Sei G eine linear reduktive Gruppe und $(\rho, V)_G$ eine rationale Darstellung von G in einem endlich dimensionalen K-Vektorraum V. Dann ist der Invariantenring $K[V]^G$ ein Cohen-Macaulay-Ring.

Warum die Cohen-Macaulay-Eigenschaft so interessant ist, wird uns der nächste Satz zeigen, der eine von vielen möglichen Charakterisierungen der Cohen-Macaulay-Eigenschaft enthält (vgl. [DK02], Satz 2.5.3, S. 63).

Satz 6.5.10. *Sei R eine Noethersche, graduierte K-Algebra mit $R_0 = K$. Dann sind die folgenden Aussagen äquivalent:*

(i) R ist ein Cohen-Macaulay-Ring.

(ii) Ist $\{f_1, \ldots, f_r\} \subseteq R$ ein homogenes Parametersystem, dann ist R ein freier $K[f_1, \ldots, f_r]$-Modul.

Wenn G also eine linear reduktive Gruppe und $(\rho, V)_G$ eine rationale Darstellung von G in V ist, so ist $K[V]^G$ wegen des Theorems von HOCHSTER und ROBERTS ein Cohen-Macaulay-Ring. Wegen Korollar 6.5.6 besitzt $K[V]^G$ ein homogenes Parametersystem $\{f_1, \ldots, f_r\} \subseteq K[V]^G$. Somit folgt aus dem Satz, dass der Invariantenring $K[V]^G$ ein endlich-erzeugter freier F-Modul ist, wobei $F = K[f_1, \ldots, f_r]$ ist. Die Zerlegung

$$K[V]^G = Fg_1 \oplus \ldots \oplus Fg_s$$

mit homogenen Invarianten $g_1, \ldots, g_s \in K[V]^G$ geht zurück auf den japanischen Mathematiker Heisuke HIRONAKA (geb. 1931) und wird deshalb oft auch die **HIRONAKA-Zerlegung** von $K[V]^G$ genannt. Weitere Informationen zu diesen Themen finden sich in geringem Maße in [DK02] und in besonderem Maße in dem Buch [BH93] von Winfried BRUNS und Jürgen HERZOG.

6.6 Die Hilbert-Reihe von Invariantenringen

Eine für die Invariantentheorie bedeutsame „Informationsquelle", die Informationen, wie z.B. der Dimension des Invariantenrings oder Gradschranken für fundamentale Invarianten, liefert, ist die sogenannte **Hilbert-Reihe**, mit der wir uns hier in kompakter Form auseinandersetzen wollen. Für einen vertieften Einstieg in die Thematik sei besonders auf [KR05], Kapitel 5, verwiesen. Das Interessante an der Hilbert-Reihe von Invariantenringen ist vor allem, dass diese in vielen Situationen bereits berechnet werden kann, obwohl man gar kein Erzeugendensystem des Rings kennt. Außerdem liefert die Hilbert-Reihe Kenntnisse über die Struktur des Invariantenrings, wie z.B. die Anzahl fundamentaler Invarianten in einem bestimmten Grad. Die Kenntnis der Hilbert-Reihe des Invariantenrings ermöglicht wiederum die Berechnung fundamentaler Invarianten, wie wir später sehen werden (siehe Abschnitt 7.2.3).

Zunächst wollen wir die **Hilbert-Funktion** definieren. Sofern nicht anders angegeben, sei dazu im Folgenden wie bisher K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$ und sei der Polynomring $P = K[x_1, \ldots, x_n]$ durch $W = (w_1, \ldots, w_n) \in \mathbb{Z}^n$ graduiert, wobei die Einträge von W entweder alle positiv oder alle negativ sind, d.h. mit anderen Worten gilt insbesondere $\deg(x_i) = w_i$ für alle $i \in \{1, \ldots, n\}$. Für einen endlich erzeugten, graduierten P-Modul $M = \bigoplus_{d \in \mathbb{Z}} M_d$ gilt dann zunächst $P_0 = K$ und $\dim_K(M_d) < \infty$ für alle $d \in \mathbb{Z}$ (vgl. [KR05], Satz 4.1.19, S. 24). Damit ist folgende Definition gerechtfertigt (vgl. [KR05], Definition 5.8.8, S. 325).

Definition 6.6.1. (Hilbert-Funktion)
Sei M ein endlich erzeugter, graduierter P-Modul. Dann heißt die wohldefinierte Abbildung $\mathrm{HF}_M : \mathbb{Z} \to \mathbb{Z}$, definiert durch $d \mapsto \dim_K(M_d)$, die **Hilbert-Funktion** von M.

Die Hilbert-Funktion $\mathrm{HF}_M : \mathbb{Z} \to \mathbb{Z}$ ist eine sogenannte **ganzzahlige Laurent-Funktion**, d.h. es gibt ein $i_0 \in \mathbb{Z}$ mit $\mathrm{HF}_M(i) = 0$ für alle $i < i_0$. Die Zahl $i_0 \in \mathbb{Z}$ heißt auch der **Anfangsgrad** von M und wird mit $\alpha(M)$ bezeichnet (vgl. [KR05], Satz 5.1.14, S. 186). Diese Eigenschaft spiegelt sich auch in folgendem klassischen Beispiel wider (siehe auch [KR05], Satz 5.1.13, S. 186).

Beispiel 6.6.2. Der Polynomring $P = K[x_1, \ldots, x_n]$ sei standardgraduiert, d.h. es gilt also $\deg(x_i) = 1$ für alle $i \in \{1, \ldots, n\}$. Wir setzen $P_d := \{0\}$ für $d < 0$. Dadurch wird $P = \bigoplus_{d \in \mathbb{Z}} P_d$ zu einer $\mathbb{Z}$-graduierten K-Algebra. Im Sinne des obigen Satzes gilt $\alpha(P) = 0$, d.h. 0 ist der Anfangsgrad von P. Weiter gilt $\mathrm{HF}_P(i) = \binom{n+i-1}{n-1}$ für jedes $i \in \mathbb{Z}$. Sei G nun eine

linear algebraische Gruppe. Da der Invariantenring P^G ebenfalls standardgraduiert ist, folgt offensichtlich

$$\mathrm{HF}_{P^G}(i) \leq \binom{n+i-1}{n-1}$$

für alle $i \in \mathbb{Z}$. ◁

Wie sich Funktionswerte der Hilbert-Funktion konkret berechnen lassen, soll hier nicht vertieft werden. Dazu sei ebenfalls auf [KR05], Kapitel 5, verwiesen, insbesondere auf Korollar 5.1.19, S. 188, für den standardgraduierten Fall und Satz 5.8.10, S. 326, für den allgemeinen Fall. Ist M ein endlich erzeugter P-Modul und $i_0 \in \mathbb{Z}$ der Anfangsgrad von M, so gibt es weiter ein eindeutig bestimmtes Polynom $p \in \mathbb{Q}[t]$ mit $p(i) \in \mathbb{Z}$ für alle $i \in \mathbb{Z}$ und $\mathrm{HF}_M(i) = p(i)$ für alle $i \geq i_0$ (vgl. [KR05], Theorem 5.1.21, S. 190 unter Verwendung von [KR05], Satz 5.8.9, S. 326). Dieses Polynom p heißt das **Hilbert-Polynom** von M und wird mit HP_M bezeichnet (vgl. [KR05], Definition 5.4.11, S. 239). Viele weitere Eigenschaften von Hilbert-Funktionen finden sich ebenfalls in [KR05], Kapitel 5, aber wir wollen es an dieser Stelle bei den vorgestellten Eigenschaften belassen und uns nun der Hilbert-Reihe zuwenden. Mit anderen Worten haben wir bisher gesehen, dass der Hauptteil der zur Hilbert-Funktion HF_M assoziierten Laurentreihe $\sum_{n=-\infty}^{\infty} \mathrm{HF}_M(n) \cdot z^n$ im Entwicklungspunkt 0 endlich ist. Daher lässt sich die Hilbert-Reihe wie folgt definieren (vgl. [DK02], Definition 5.8.11, S. 326 f.).

Definition 6.6.3. (Hilbert-Reihe)
Sei M ein endlich erzeugter, graduierter P-Modul. Die formale Laurentreihe $\sum\limits_{d=\alpha(M)}^{\infty} \mathrm{HF}_M(d) \cdot z^d$ heißt die **Hilbert-Reihe** von M und wird mit $\mathrm{HS}_M(z)$ bezeichnet.

Als erstes Beispiel wollen wir auch hier die Hilbert-Reihe eines standardgraduierten Polynomrings betrachten (vgl. [KR05], Satz 5.2.14, S. 202).

Beispiel 6.6.4. Sei P standardgraduiert, d.h. es gilt $\deg(x_i) = 1$ für alle $i \in \{1, \ldots, n\}$. Bekanntlich ist dann $\mathrm{HF}_P : \mathbb{Z} \to \mathbb{Z}$ definiert durch

$$\mathrm{HF}_P(i) = \begin{cases} 0, & i < 0 \\ \binom{n+i-1}{n-1}, & i \geq 0 \end{cases}$$

Damit folgt $\mathrm{HS}_P(z) = \sum_{d=0}^{\infty} \binom{n+d-1}{n-1} \cdot z^d$. Dies ist gerade die Potenzreihenentwicklung der Funktion $z \mapsto \frac{1}{(1-z)^n}$, d.h. es gilt $\mathrm{HS}_P(z) = \frac{1}{(1-z)^n}$. ◁

Wie das für Polynomringe aussieht, die mit einer anderen Graduierung versehen sind, werden wir gleich beantworten. Zunächst lässt sich folgende Formel für die Hilbert-Reihe des Polynomrings P festhalten (vgl. [KR05], Theorem 5.8.15, S. 328).

Theorem 6.6.5. (Hilbert-Reihe von Polynomringen)
Sei $P = K[x_1, \ldots, x_n]$ durch $W = (w_1, \ldots, w_n) \in \mathbb{Z}^n$ graduiert, wobei die Einträge von W entweder alle positiv oder alle negativ sind. Dann gilt:

$$\mathrm{HS}_P(z) = \frac{1}{(1 - z^{w_1}) \cdots (1 - z^{w_n})}.$$

Aus diesem Theorem lässt sich ebenfalls eine explizite Formel für freie P-Moduln angeben (vgl. [KR05], Korollar 5.8.16, S. 329).

Korollar 6.6.6. *Sei $P = K[x_1, \ldots, x_n]$ durch $W = (w_1, \ldots, w_n) \in \mathbb{Z}^n$ graduiert, wobei die Einträge von W entweder alle positiv oder alle negativ sind. Sei weiter $F = \bigoplus_{i=1}^{r} P(-\delta_i)$ mit $\delta_i \in \mathbb{Z}$ für alle $i \in \{1, \ldots, r\}$ ein graduierter, freier P-Modul. Dann gilt:*

$$\mathrm{HS}_F(z) = \frac{\sum_{i=1}^{r} z^{\delta_i}}{(1 - z^{w_1}) \cdots (1 - z^{w_n})}.$$

Damit ist es sofort möglich, die Hilbert-Reihe des Invariantenrings anzugeben, wenn wir die primären und sekundären Invarianten kennen. Sei P standardgraduiert, d.h. es gilt $\deg(x_i) = 1$ für alle $i \in \{1, \ldots, n\}$, und sei G eine linear reduktive Gruppe. Laut Satz 6.5.10 gibt es ein homogenes Parametersystem $f_1, \ldots, f_r$ so, dass P^G ein endlich erzeugter, freier $K[f_1, \ldots, f_r]$-Modul ist, d.h. mit $F := K[f_1, \ldots, f_r]$ gibt es homogene Polynome $g_1, \ldots, g_s \in P^G$ mit $P^G = Fg_1 \oplus \ldots \oplus Fg_s$. Sei $R = K[y_1, \ldots, y_r]$ mit Unbestimmten $y_1, \ldots, y_r$ graduiert durch $W = (w_1, \ldots, w_r)$, wobei $w_i = \deg(f_i)$ für alle $i \in \{1, \ldots, r\}$ gilt, und sei $\Phi : R \to P$ definiert durch $y_i \mapsto f_i$. Dann ist F das Bild von Φ. Da $f_1, \ldots, f_r$ als Elemente eines homogenen Parametersystems algebraisch unabhängig sind, gilt $R \cong K[f_1, \ldots, f_r]$. Setze $\delta_j := \deg(g_j)$ für alle $j \in \{1, \ldots, s\}$. Dann ist P^G isomorph zu dem graduierten, freien R-Modul $S = \bigoplus_{j=1}^{s} R(-\delta_j)$. Gemäß dem letzten Korollar gilt somit

$$\mathrm{HS}_{P^G}(z) = \mathrm{HS}_S(z) = \frac{\sum_{j=1}^{s} z^{\delta_j}}{(1 - z^{w_1}) \cdots (1 - z^{w_r})}.$$

Die bisherigen Resultate wollen wir nun rekapitulieren lassen und uns besonders damit auseinandersetzen, welche Kenntnis sie uns über die Hilbert-Reihe des Invariantenrings liefern. Seien $f_1, \ldots, f_r \in P$ die fundamentalen Invarianten von P^G, d.h. es gilt $P^G = K[f_1, \ldots, f_r]$. Weiter versehen wir den Polynomring $R := K[y_1, \ldots, y_r]$ mit Unbestimmten $y_1, \ldots, y_r$ mit der durch $W = (w_1, \ldots, w_r)$ gegebenen $\mathbb{Z}^r$-Graduierung, wobei $w_i = \deg(f_i)$ gilt für alle $i \in \{1, \ldots, r\}$. Dann ist P^G das Bild des K-Algebra-Homomorphismus $\Phi : R \to P$ definiert durch $y_i \mapsto f_i$ für alle $i \in \{1, \ldots, r\}$ und der Kern von Φ ist ein homogenes Ideal. Nun betrachten wir die homogene exakte Sequenz $0 \to \mathrm{Ker}(\Phi) \to R \to R/\mathrm{Ker}(\Phi) \to 0$ und erhalten (vgl. [KR05], Satz 5.8.13, S. 327)

$$\mathrm{HS}_{R/\mathrm{Ker}(\Phi)}(z) = \mathrm{HS}_R(z) - \mathrm{HS}_{\mathrm{Ker}(\Phi)}(z).$$

Kennen wir also die Hilbert-Reihe $\mathrm{HS}_{R/\mathrm{Ker}(\Phi)}(z)$ von $R/\mathrm{Ker}(\Phi)$, so kennen wir wegen des Isomorphismus $P^G \cong R/\mathrm{Ker}(\Phi)$ auch die Hilbert-Reihe des Invariantenrings P^G. Aus Theorem 6.6.5 ist die Hilbert-Reihe von $R = K[y_1, \ldots, y_r]$ bereits bekannt. Somit wäre auf diesem Weg nur noch die Hilbert-Reihe des homogenen Ideals $\mathrm{Ker}(\Phi) \subseteq R$ zu bestimmen.

Die spezielle Bestimmung der Hilbert-Reihe von P/I für ein homogenes Ideal $I \subseteq P$ wollen wir nun im Folgenden zum Thema machen. Zunächst lässt sich das folgende Theorem über die Gestalt der Hilbert-Reihe festhalten (vgl. [KR05], Korollar 5.8.19, S. 330).

Theorem 6.6.7. (Gestalt der Hilbert-Reihe)
Sei $P = K[x_1, \ldots, x_n]$ durch $W = (w_1, \ldots, w_n) \in \mathbb{Z}^n$ graduiert, sei M ein endlich erzeugter, graduierter P-Modul und sei $k := \alpha(M)$ der Anfangsgrad von M. Dann ist die Hilbert-Reihe von M von der Form

$$\mathrm{HS}_M(z) = \frac{z^k \cdot p}{(1 - z^{w_1}) \cdots (1 - z^{w_n})},$$

wobei $p \in \mathbb{Z}[z]$ ein Polynom ist mit $p(0) = \mathrm{HF}_M(k) > 0$.

Das Polynom p aus diesem Theorem bekommt ebenfalls einen passenden Namen, es heißt der **Hilbert-Zähler** von M und wird mit HN_M bezeichnet (vgl. [KR05], Definition 5.2.21, S. 204).

Die Bezeichnung HN kommt dabei von dem Englischen Begriff „Hilbert Numerator". Für den für uns besonders interessanten Fall P/I ergibt sich aus dem letzten Theorem folgende Darstellung der Hilbert-Reihe (vgl. [KR05], Theorem 5.8.18, S. 329).

Korollar 6.6.8. *Sei $0 \neq I \subseteq P$ ein echtes homogenes Ideal. Dann ist die Hilbert-Reihe von P/I von der Form*

$$\mathrm{HS}_{P/I}(z) = \frac{\mathrm{HN}_{P/I}(z)}{(1-z^{w_1})\cdots(1-z^{w_n})}.$$

Dabei gilt $\mathrm{HN}_{P/I}(0) = 1$.

Damit ist die Hilbert-Reihe von P/I durch den Hilbert-Zähler bestimmt, d.h. die Berechnung der Hilbert-Reihe lässt sich auf die Berechnung des Hilbert-Zählers reduzieren. Analog zur Hilbert-Funktion kann man auch die Bestimmung der Hilbert-Reihe eines homogenen Ideals I auf die Bestimmung der Hilbert-Reihe des zugehörigen Leittermideals bzgl. einer Termordnung zurückführen (vgl. [KR05], Satz 5.8.13, S. 327 f. unter Berücksichtigung von Theorem 5.2.18, S. 204), was als **Macaulay's Theorem für Hilbert-Reihen** bekannt ist. Wie wir bereits wissen, gilt $\mathrm{HS}_{P/I}(z) = \mathrm{HS}_P(z) - \mathrm{HS}_I(z)$. Somit folgt aus diesem Theorem sofort $\mathrm{HS}_{P/I}(z) = \mathrm{HS}_{P/\mathrm{LT}_\sigma(I)}(z)$, d.h. dieses Theorem gestattet es, die Berechnung der Hilbert-Reihe bzw. des Hilbert-Zählers auf monomiale Ideale zurückzuführen. Jedes monomiale Ideal besitzt ein eindeutig bestimmtes minimales Erzeugendensystem bestehend aus Termen. Mit diesem Erzeugendensystem als Eingabe ist es möglich, einen rekursiven Algorithmus für die Berechnung des Hilbert-Zählers von P/I für ein echtes monomiales Ideal I anzugeben (vgl. [KR05], Theorem 5.8.18, S. 329). Für die Umsetzung dieses Algorithmus gibt es eine Vielzahl verschiedener Ansätze und Strategien, auf die wir aber nicht näher eingehen werden. Diese können in [KR05], Abschnitt 5.3, S. 214–224 (unter Berücksichtigung von [KR05], Bemerkung 5.8.20, S. 330), nachgelesen werden.

Für eine Berechnung der Hilbert-Reihe des Invariantenrings $K[V]^G$ ist bisher stets die Kenntnis eines Erzeugendensystem von $K[V]^G$ notwendig und vorausgesetzt. Allerdings lässt sich die Hilbert-Reihe in vielen Situation schon im Vorhinein berechnen, ohne ein Erzeugendensystem zu kennen. Darauf wollen wir hier aber nur kurz eingehen. Eine Situation, in der eine Berechnung der Hilbert-Reihe im Voraus möglich ist, liegt bei endlichen Gruppen vor. Diese Aussage liefert die sogenannte **Molien-Formel**, benannt nach dem deutsch-baltischen Mathematiker Theodor Molien (1861–1941). Der Einfachheit halber geben wir dieses Theorem nur für Körper der Charakteristik Null an; eine Verallgemeinerung auf beliebige Körper K mit $\mathrm{char}(K) \nmid \#G$ ist aber möglich (vgl. [DK02], Theorem 3.2.2, S. 77).

Theorem 6.6.9. (Molien-Formel)
Sei G eine endliche Gruppe und sei $(\rho, V)_G$ eine rationale Darstellung von G in einem endlichdimensionalen K-Vektorraum V. Dann gilt:

$$\mathrm{HS}_{P^G}(z) = \frac{1}{\#G}\sum_{a\in G}\frac{1}{\det(1-z\cdot\rho_a)}.$$

Die Hilbert-Reihe des Invariantenrings lässt sich also im Falle von endlichen Gruppen als eine Art Durchschnitt über alle Elemente der Gruppe berechnen. Diese Idee der Durchschnittsbildung kann auch auf beliebige linear reduktive Gruppen erweitert werden (vgl. [DK02], Abschnitt 4.6.1, S. 180–196), wenngleich man hier natürlich nicht alle Gruppenelemente in analoger Form betrachten kann. Wir wollen nun stark vereinfacht die Möglichkeiten kurz erläutern. Sei dazu nun zur Vereinfachung $K = \mathbb{C}$. Jede linear reduktive Gruppe G enthält eine maximal kompakte Untergruppe C, auf der es ein Haarsches Maß μ gibt, mit dem sich stetige Funktionen

$f : C \to \mathbb{C}$ in folgendem Sinne „über die Gruppe C integrieren lassen" (vgl. [Kra85], AII.3, S. 285): Es gibt ein lineares Funktional $f \mapsto \int_C f(a)d\mu$ mit $a \in C$ auf der Menge aller stetigen Funktionen $f : C \to \mathbb{C}$, das folgende Eigenschaften erfüllt:

(i) $\int_C d\mu = 1$ (Normierung)

(ii) Für alle $b \in C$ gilt $\int_C f(ab)d\mu = \int_C f(a)d\mu = \int_C f(ba)d\mu$. (Rechts-/Linksinvarianz)

Mit diesem Maß lässt sich die Molien-Formel verallgemeinern zu (vgl. [DK02], S. 180)

$$\mathrm{HS}_{\mathbb{C}[V]^G}(z) = \int_C \frac{1}{\det(1 - z \cdot \rho_a)} d\mu, \tag{6.6.1}$$

wobei $(\rho, V)_G$ eine rationale Darstellung von G in einem endlich-dimensionalen $\mathbb{C}$-Vektorraum V ist. Die Hilbert-Reihe $\mathrm{HS}_{\mathbb{C}[V]^G}(z)$ konvergiert für $|z| < 1$, denn aus Korollar 6.6.8 folgt, dass die Hilbert-Reihe eine rationale Funktion ist, die nur Pole in $z = 1$ besitzt. Ebenso lässt sich zeigen, dass das Integral in Gleichung (6.6.1) für $|z| < 1$ definiert ist.

Die Formel in Gleichung (6.6.1) kann nun noch weiter vereinfacht werden. Dazu nehmen wir G als zusammenhängend an, betrachten einen maximalen Torus T von G und eine maximal kompakte Untergruppe D von T mit $D \subseteq C$. Dann lässt sich T mit $(\mathbb{C}^*)^r$ und D mit $(S^1)^r$ für ein $r \in \mathbb{N}_+$ identifizieren, wobei $S^1 := \{z \in \mathbb{C} : \ |z| = 1\}$ der Einheitskreis in der komplexen Ebene ist. Auch für D lässt sich ein Haarsches Maß ν mit den obigen Eigenschaften wählen. Dann gibt es eine Gewichtsfunktion $\varphi : D \to \mathbb{R}$ mit der Eigenschaft, dass für alle stetigen Funktionen $f : C \to \mathbb{C}$ gilt (vgl. [DK02], S. 181):

$$\int_C f(a)d\mu = \int_D \varphi(a)f(a)d\nu.$$

Damit vereinfacht sich Gleichung (6.6.1) zu

$$\mathrm{HS}_{\mathbb{C}[V]^G}(z) = \int_D \frac{\varphi(a)}{\det(1 - z \cdot \rho_a)} d\nu.$$

In vielen Fällen, wie z.B. eines r-dimensionalen Torus, lassen sich nun explizite Formeln für die Hilbert-Reihe angeben. Dazu sei allerdings nur auf [DK02], Abschnitt 4.6, S. 180–196, verwiesen. Bei der Berechnung der dabei auftretenden Integrale werden Methoden der Funktionentheorie herangezogen, wie z.B. der bekannte **Residuensatz** von Augustin-Louis Cauchy (siehe z.B. [FB06], Theorem 6.3, S. 164). Wir wollen nun noch in sehr kompakter Form ein Beispiel betrachten, das wir allerdings nicht vollständig ausführen wollen, sondern nur skizzieren wollen (vgl. [DK02], Beispiel 4.6.1, S. 182, und Beispiel 4.6.15, S. 190).

Beispiel 6.6.10. Sei $G = \mathrm{SL}_2(\mathbb{C})$ und $C = \mathrm{SU}_2(\mathbb{C})$. Sei weiter

$$T := \left\{ \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} : \ z \in \mathbb{C}^* \right\} \quad \text{und} \quad D := \left\{ \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} : \ z \in S^1 \right\}.$$

Dann ist C eine maximal kompakte Untergruppe von G, T ein maximaler Torus von G und D eine maximal kompakte Untergruppe von T. Sei V_k der $\mathbb{C}$-Vektorraum der Binärform vom Grad k. Dann operiert die Gruppe D auf V_k durch die $(k+1) \times (k+1)$-Matrix

$$\begin{pmatrix} z^k & 0 & 0 & \dots & 0 \\ 0 & z^{k-2} & 0 & & 0 \\ 0 & 0 & z^{k-4} & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & z^{-k} \end{pmatrix}.$$

Durch Wahl geeigneter Maße (vgl. [DK02], Beispiel 4.6.1, S. 182 f.) folgt

$$\mathrm{HS}_{\mathbb{C}[V_k]^{\mathrm{SL}_2}}(t) = \frac{1}{\pi}\int_0^{2\pi} \frac{\sin^2(u)}{(1-e^{iku}t)(1-e^{i(k-2)u}t)\cdots(1-e^{-iku}t)}du$$

und durch Umformen in ein Kurvenintegral erhalten wir

$$\mathrm{HS}_{\mathbb{C}[V_k]^{\mathrm{SL}_2}}(t) = \frac{1}{2\pi i}\int_{S^1} \frac{1-z^2}{z(1-z^kt)(1-z^{k-2}t)\cdots(1-z^{-k}t)}dz,$$

wobei $|t| < 1$ gelte. Setze nun $f_{k,t}(z) := \frac{1-z^2}{z(1-z^kt)(1-z^{k-2}t)\cdots(1-z^{-k}t)}$. Die Menge der Pole der komplexen Funktion $f_{k,t}$ in der Einheitskreisscheibe $B_1(0)$ ist

$$P := \left\{\zeta_{k-2j}^{\ell} t^{\frac{1}{k-2j}} : \ 0 \le j < \tfrac{k}{2},\ 0 \le \ell < k-2j\right\},$$

wobei ζ_p eine primitive p-te Einheitswurzel ist. Dann folgt $\mathrm{HS}_{\mathbb{C}[V_k]^{\mathrm{SL}_2}}(t) = \sum_{x\in P}\mathrm{Res}(f_{k,t},x)$ mit dem Residuensatz. Dabei ist $\mathrm{Res}(f_{k,t},x)$ das Residuum der Funktion $f_{k,t}$ an der Stelle $x \in P$. Betrachten wir noch den konkreten Fall $k = 2$. Dann gilt zunächst

$$f_{2,t}(z) = \frac{1-z^2}{z(1-z^2t)(1-t)(1-z^{-2}t)}.$$

Die einzigen Pole von $f_{2,t}$ in $B_1(0)$ sind $z_1 := \zeta_2^0\sqrt{t} = \sqrt{t}$ und $z_2 := \zeta_2^1\sqrt{t} = -\sqrt{t}$. Für die Residuen an den Stellen z_1 bzw. z_2 folgt:

$$\begin{aligned}\mathrm{Res}(f_{2,t},z_1) &= \lim_{z\to\sqrt{t}} \frac{(1-z^2)(z-\sqrt{t})}{z(1-z^2t)(1-t)z^{-2}(z+\sqrt{t})(z-\sqrt{t})} = \frac{1}{2(1-t^2)}\\ \mathrm{Res}(f_{2,t},z_2) &= \frac{1}{2(1-t^2)}.\end{aligned}$$

Somit gilt schließlich $\mathrm{HS}_{\mathbb{C}[V_2]^{\mathrm{SL}_2}}(t) = \frac{1}{1-t^2}$. ◁

KAPITEL 7

Algorithmische Invariantentheorie

Theodor Fontane[20]

Wer rechnet, ist immer in Gefahr, sich zu verrechnen. Die dumme Kuh trifft immer das richtige Gras.

Auf die „dumme Kuh", die Theodor Fontane hier anspricht, wollen wir natürlich nicht vertrauen. Auch ist es wenig erstrebenswert, Invarianten „per Hand" zu berechnen. Was wir also benötigen, sind Algorithmen zur Berechnung fundamentaler Invarianten. Wie wir im letzten Kapitel durch Hilberts Endlichkeitssatz bereits gesehen haben, ist der Invariantenring für linear reduktive Gruppe stets endlich erzeugt. Und genau für diese Klasse von Gruppen lieferte Harm Derksen einen effizienten Algorithmus, den wir unter anderem in diesem Kapitel vorstellen wollen. Wir stellen uns aber nicht mit dem bloßen Algorithmus zufrieden, sondern wir wollen diesen verstehen und beweisen. Eine Implementierung des Derksen-Algorithmus ist im Computeralgebrasystem Magma zu finden. Um den Algorithmus von Derksen vollends erfassen zu können, ist es notwendig, sich mit dem sogenannten **Reynolds-Operator** zu beschäftigen, einer K-linearen Abbildung, die nur für linear reduktive Gruppen existiert. Dieser Abbildung ist der erste Abschnitt dieses Kapitels gewidmet. Darüber hinaus werden wir hier als Alternative zum Derksen-Algorithmus einen naiven Algorithmus näher beleuchten, der zwar weniger effizient ist, aber dennoch Beachtung erfahren soll. Dieser Algorithmus ist in etwas anderer Form, als wir in vorstellen werden, Teil des Open-source-Computeralgebrasystem CoCoA und kommt ohne Reynolds-Operator aus.

7.1 Der Reynolds-Operator

Wir wollen in diesem Abschnitt ein „Schlüsselwerkzeug" zur effizienten Berechnung von Erzeugendensystemen von Invariantenringen betrachten, den sogenannten **Reynolds-Operator**. Dieser Operator geht zurück auf den irischen Mathematiker Osborne Reynolds (1842–1912) und wird hier wie folgt definiert. Dabei weichen wir leicht von der Darstellung in [DK02], Definition 2.2.2, S. 45 ab. Wie bisher sei K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$.

[20] Bildquelle: http://de.wikipedia.org/wiki/Theodor_Fontane vom 06.04.2015.

Definition 7.1.1. (Reynolds-Operator)
Sei G eine linear algebraische Gruppe, die auf einer affinen G-Varietät X regulär operiert. Eine K-lineare Abbildung $\mathrm{Rey} : K[X] \to K[X]$ mit

(i) $\mathrm{Rey} \circ \mathrm{Rey} = \mathrm{Rey}$,

(ii) $\mathrm{Rey}(f) = f$ für alle $f \in K[X]^G$,

(iii) $\mathrm{Rey}(f^a) = \mathrm{Rey}(f)$ für alle $f \in K[X]$ und alle $a \in G$

heißt ein **Reynolds-Operator** bzgl. G (in $K[X]$).

Ein Reynolds-Operator ist zunächst also eine idempotente, lineare Abbildung, d.h. eine Projektion. Somit gilt sofort $K[X] = \mathrm{Ker}(\mathrm{Rey}) \oplus \mathrm{Im}(\mathrm{Rey})$. Ist $(\rho, K[X])_G$ die reguläre Darstellung von G in $K[X]$, so lässt sich die dritte Eigenschaft so ausdrücken, dass ein Reynolds-Operator zusätzlich eine ρ-invariante Abbildung ist. In [DK02] wird ein Reynolds-Operator sofort als Projektion auf $K[X]^G$ definiert. Für das bessere Verständnis wollen wir hier explizit nachweisen, dass das Bild eines Reynolds-Operators tatsächlich der Invariantenring ist.

Satz 7.1.2. (Bilder von Reynolds-Operatoren)
Sei G eine linear algebraische Gruppe, die auf einer affinen G-Varietät X regulär operiert, und sei $\mathrm{Rey} : K[X] \to K[X]$ *ein Reynolds-Operator bzgl. G in $K[X]$. Dann gilt* $\mathrm{Im}(\mathrm{Rey}) = K[X]^G$, *d.h.* Rey *ist eine Projektion von $K[X]$ auf den Invariantenring $K[X]^G$.*

Beweis: Zunächst zeigen wir, dass genau dann $f \in \mathrm{Im}(\mathrm{Rey})$ gilt, wenn $\mathrm{Rey}(f) = f$ erfüllt ist, wobei $f \in \mathrm{Im}(\mathrm{Rey})$ trivialerweise aus $\mathrm{Rey}(f) = f$ folgt. Sei nun $f \in \mathrm{Im}(\mathrm{Rey})$. Dann gibt es ein $g \in K[X]$ mit $\mathrm{Rey}(g) = f$ und es folgt $\mathrm{Rey}(f) = \mathrm{Rey}(\mathrm{Rey}(g)) = \mathrm{Rey}(g) = f$. Damit ist die Inklusion $K[X]^G \subseteq \mathrm{Im}(\mathrm{Rey})$ per Definition klar.

Für den Beweis der Inklusion $\mathrm{Im}(\mathrm{Rey}) \subseteq K[X]^G$ sei $f \in \mathrm{Im}(\mathrm{Rey})$ und o.B.d.A. $f \neq 0$. Dann gibt es ein $g \in K[X]$ mit $\mathrm{Rey}(g) = f$. Sei weiter $a \in G$. Dann gilt:

$$\mathrm{Rey}(f^a - f) = \mathrm{Rey}(f^a) - \mathrm{Rey}(f) = \mathrm{Rey}(f) - \mathrm{Rey}(f) = 0,$$

d.h. $f^a - f \in \mathrm{Ker}(\mathrm{Rey})$. Weiter gilt $f^a \in \mathrm{Im}(\mathrm{Rey})$. Denn sonst wäre $f^a \in \mathrm{Ker}(\mathrm{Rey})$, d.h. es würde $\mathrm{Rey}(f^a) = 0$ gelten und damit auch $\mathrm{Rey}(f) = 0$. Wegen $f \in \mathrm{Im}(\mathrm{Rey})$ und, da $\mathrm{Im}(\mathrm{Rey})$ und $\mathrm{Ker}(\mathrm{Rey})$ nur die Null gemeinsam haben, würde dies $f = 0$ bedeuten im Widerspruch zur Wahl von f.

Somit gilt insgesamt $f^a - f \in \mathrm{Im}(\mathrm{Rey})$ und $f^a - f \in \mathrm{Ker}(\mathrm{Rey})$, also folgt $f^a - f = 0$, d.h. es gilt $f^a = f$ und damit $f \in K[X]^G$. □

Ein Reynolds-Operator berechnet also genau den Invariantenring der Operation von G auf X. Als erstes Beispiel eines Reynolds-Operators wollen wir speziell einen Reynolds-Operator bzgl. endlichen Gruppen betrachten (vgl. [DK02], Beispiel 2.2.3, S. 45).

Beispiel 7.1.3. (Reynolds-Operator bei endlichen Gruppen)
Sei $(G, *, e)$ eine endliche Gruppe mit $\mathrm{char}(K) \nmid \#G$ und X eine affine G-Varietät. Sei weiter $\Phi : K[X] \to K[X]$ definiert durch

$$f \mapsto \frac{1}{\#G} \cdot \sum_{a \in G} f^a.$$

Dass Φ eine wohldefinierte K-lineare Abbildung ist, ist klar. Weiter gilt unter Verwendung der Rechenregeln aus Lemma 4.3.5:

(i) Sei $f \in K[X]$. Dann gilt:

$$\begin{aligned}\Phi(\Phi(f)) &= \Phi\left(\frac{1}{\#G}\cdot\sum_{a\in G} f^a\right) = \frac{1}{\#G}\cdot\sum_{b\in G}\left(\frac{1}{\#G}\cdot\sum_{a\in G} f^a\right)^b \\ &= \frac{1}{\#G}\cdot\sum_{b\in G}\frac{1}{\#G}\cdot\sum_{a\in G}(f^a)^b = \frac{1}{\#G}\cdot\sum_{b\in G}\frac{1}{\#G}\cdot\sum_{a\in G} f^{b*a} \\ &= \frac{1}{\#G}\cdot\sum_{b\in G}\Phi(f) = \frac{1}{\#G}\cdot\#G\cdot\Phi(f) = \Phi(f)\end{aligned}$$

(ii) Sei $f \in K[X]^G$. Dann gilt $f^a = f$ für alle $a \in G$ und es folgt:

$$\Phi(f) = \frac{1}{\#G}\cdot\sum_{a\in G} f^a = \frac{1}{\#G}\cdot\sum_{a\in G} f = \frac{1}{\#G}\cdot\#G\cdot f = f$$

(iii) Sei $f \in K[X]$ und sei $a \in G$. Dann gilt:

$$\Phi(f^a) = \frac{1}{\#G}\cdot\sum_{b\in G}(f^a)^b = \frac{1}{\#G}\cdot\sum_{b\in G} f^{b*a} = \Phi(f)$$

Somit ist die Abbildung Φ ein Reynolds-Operator bzgl. G in $K[X]$. ◁

Dieser Reynolds-Operator bzgl. einer endlichen Gruppe, bildet somit für jedes Polynom eine Art „Durchschnitt" über die ganze Gruppe hinweg. Allein durch diese Durchschnittsbildung lassen sich in diesem Fall also Invarianten und damit letztendlich ein Erzeugendensystem das Invariantentrings berechnen. Wie das im Detail funktioniert, werden wir im allgemeinen Rahmen für beliebige linear reduktive Gruppe später betrachten. Die Invariantentheorie endlicher Gruppen ist nicht explizit Gegenstand dieser Arbeit, bildet jedoch einen bedeutenden Zweig der Invariantentheorie. Wir verweisen an dieser Stelle nur auf die bereits erwähnten Bücher [DK02] oder [Neu07]. Ungeklärt ist bisher auch, ob es stets einen Reynolds-Operator gibt. Wie auch das letzte Beispiel zeigt, hängt ein Reynolds-Operator bzgl. einer Gruppe natürlich wesentlich von der Gruppe ab. Für die uns bereits bekannten linear reduktiven Gruppen gibt es stets einen Reynolds-Operator und dieser ist zudem sogar eindeutig bestimmt, was aus dem nächsten Theorem folgt (vgl. [DK02], Theorem 2.2.5, S. 46). Da der Beweis des Theorems konstruktiv ist und wir bei der Definition eines Reynolds-Operators von der Darstellung in [DK02] leicht abweichen, wollen wir das Theorem explizit beweisen.

Theorem 7.1.4. (Existenz und Eindeutigkeit von Reynolds-Operatoren)
Sei G eine linear algebraische Gruppe. Es gibt genau dann für jede affine G-Varietät X einen eindeutig bestimmten Reynolds-Operator $\mathrm{Rey} : K[X] \to K[X]$ *bzgl. G in $K[X]$, wenn G linear reduktiv ist.*

Beweis: Wir zeigen zunächst unter Verwendung von Theorem 4.4.5, dass G linear reduktiv ist, wenn es für jede affine G-Varietät einen eindeutig bestimmten Reynolds-Operator bzgl. G gibt. Sei dazu V ein endlich-dimensionaler K-Vektorraum, $(\rho, V)_G$ eine rationale Darstellung von G in V und sei $v \in V^G \setminus \{0\}$. Der K-Vektorraum V ist enthalten im Bidualraum zu V, d.h. es gilt mit anderen Worten $V \subseteq K[V^*]$. Laut Voraussetzung gibt es einen eindeutig bestimmten Reynolds-Operator $\mathrm{Rey} : K[V^*] \twoheadrightarrow K[V^*]^G$ bzgl. G in $K[V^*]$. Sei weiter $\pi : K[V^*]^G \to K$ eine Projektion auf K mit $\pi(v) \neq 0$. Dann ist $f := \pi \circ \mathrm{Rey}|_V$ ein Element von $(V^*)^G$ und es gilt

$$f(v) = (\pi \circ \mathrm{Rey}|_V)(v) = \pi(\mathrm{Rey}|_V(v)) = \pi(v) \neq 0.$$

Gemäß Theorem 4.4.5 ist G somit linear reduktiv.

Sei G nun linear reduktiv und X eine affine G-Varietät. Laut Satz 4.3.20 ist die reguläre Darstellung $(\rho, K[X])_G$ von G in $K[X]$ rational und lokal endlich. Sei $V \subseteq K[X]$ ein beliebiger endlich-dimensionaler, G-stabiler K-Vektorraum. Dann ist auch $(\rho|_V, V)_G$ eine rationale Darstellung. Da G linear reduktiv ist, folgt aus Theorem 4.4.5, dass es einen eindeutig bestimmten K-Untervektorraum U von V gibt mit $V = V^G \oplus U$ und $(U^*)^G = \{0\}$. Für jeden endlich-dimensionalen, G-stabilen K-Vektorraum $V \subseteq K[X]$ sei $R_V : V \to V^G$ die Projektion von V auf V^G längs U.

Sei nun $V' \subseteq K[X]$ ein weiterer endlich-dimensionaler, G-stabiler K-Vektorraum mit $V \subseteq V'$. Dann gilt $R_{V'}|_U = 0$, denn sonst wäre $(U^*)^G \neq \{0\}$. Da R_V die Projektion auf V^G längs U ist, gilt insbesondere $f - R_V(f) \in U$ für alle $f \in V$. Somit folgt wegen $V^G \subseteq (V')^G$

$$0 = R_{V'}(f - R_V(f)) = R_{V'}(f) - R_{V'}(R_V(f)) = R_{V'}(f) - R_V(f),$$

d.h. $R_{V'}|_V = R_V$. Laut Satz 4.3.19 ist für jedes $f \in K[X]$ der K-Vektorraum $V_f := \langle G(f) \rangle_K$ endlich-dimensional und offensichtlich G-stabil. Außerdem ist V_f der bzgl. Inklusion kleinste endlich-dimensionale und G-stabile K-Untervektorraum von $K[X]$, der f enthält. Auch für V_f gibt es einen eindeutig bestimmten K-Untervektorraum $U_f \subseteq V_f$ mit $V_f = V_f^G \oplus U_f$.

Sei nun $\mathcal{R} : K[X] \to K[X]$ definiert durch $\mathcal{R}(f) = R_{V_f}(f)$. Diese Abbildung ist zunächst wohldefiniert, da für jeden endlich-dimensionalen, G-stabilen K-Untervektorraum V' von $K[X]$, der ebenfalls f enthält, $R_{V'}(f) = R_{V_f}(f)$ folgt. Nun sind die Axiome aus Definition 7.1.1 nachzuweisen:

(i) Sei $f \in K[X]$. Setze $g := \mathcal{R}(f) = R_{V_f}(f) \in V_f^G$. Dann gilt $V_g = V_f$ und $R_{V_f}(g) = g$. Somit folgt

$$\mathcal{R}(\mathcal{R}(f)) = \mathcal{R}(R_{V_f}(f)) = R_{V_f}(g) = g = \mathcal{R}(f).$$

(ii) Sei $f \in K[X]^G$. Dann gilt $f \in V_f^G$ und es folgt $\mathcal{R}(f) = R_{V_f}(f) = f$.

(iii) Sei $f \in K[X]$ und $a \in G$. Ist $f \in V_f^G$, so gilt $f^a = f$ und es folgt $\mathcal{R}(f^a) = \mathcal{R}(f)$. Ist $f \notin V_f^G$, also $f \in U_f$, dann gilt auch $f^a \in U_f$, denn wäre $f^a \in V_f^G$, so müsste auch f Element von V_f^G sein. Somit gilt $f^a - f \in U_f$ und es folgt

$$0 = R_{V_f}(f^a - f) = \mathcal{R}(f^a - f) = \mathcal{R}(f^a) - \mathcal{R}(f),$$

also $\mathcal{R}(f^a) = \mathcal{R}(f)$.

Es bleibt die Eindeutigkeit zu zeigen. Sei $\tilde{\mathcal{R}} : K[X] \to K[X]$ ein weiterer Reynolds-Operator bzgl. G in $K[X]$ und sei $f \in K[X]$. Dann gilt analog zu oben $V_f = V_f^G \oplus U_f$ mit einem endlich-dimensionalen K-Untervektorraum U_f von V_f. Es gilt $f - \mathcal{R}(f) \in U_f$ und $f - \tilde{\mathcal{R}}(f) \in U_f$, also $\tilde{\mathcal{R}}(f) - \mathcal{R}(f) \in U_f$. Somit gilt wegen $\tilde{\mathcal{R}}(f) \in K[X]^G$:

$$0 = \mathcal{R}(\tilde{\mathcal{R}}(f) - \mathcal{R}(f)) = \mathcal{R}(\tilde{\mathcal{R}}(f)) - \mathcal{R}(\mathcal{R}(f)) = \tilde{\mathcal{R}}(f) - \mathcal{R}(f),$$

also $\tilde{\mathcal{R}}(f) = \mathcal{R}(f)$ und damit folgt die Eindeutigkeit. □

Dieses Theorem liefert uns somit neben der eindeutigen Existenz eines Reynolds-Operators bzgl. linear reduktiver Gruppen auch eine weitere Charakterisierung eben dieser: Die linear reduktiven Gruppen sind genau diejenigen linear algebraischen Gruppen, für die es einen eindeutig bestimmten Reynolds-Operator gibt. Aus diesem Grund können wir in diesem Fall von nun an

auch von *dem* Reynolds-Operator bzgl. einer linear reduktiven Gruppe G in $K[X]$ sprechen, den wir mit $\mathrm{Rey}_G : K[X] \twoheadrightarrow K[X]^G$ bezeichnen. Da sich die linear reduktiven Gruppen außerdem als die Gruppen erwiesen, die stets ein endliches Erzeugendensystem für den Invariantenring garantieren, lässt sich leicht erkennen, dass der Reynolds-Operator eine zentrale Rolle in der Berechnung eines solchen Erzeugendensystems einnehmen wird.

Der Reynolds-Operator bzgl. endlichen Gruppen ist uns bereits bekannt. Dieser lässt sich wie in Beispiel 7.1.3 angegeben konstruieren. Der Beweis des Theorems zeigt uns, wie für beliebige linear reduktive Gruppen der Reynolds-Operator konstruiert werden kann: Betrachte dazu für $f \in K[X]$ den K-Vektorraum $V_f := \langle G(f) \rangle_K$. Dieser ist endlich-dimensional und G-stabil. Ist G linear reduktiv, so gibt es ein G-stabiles Komplement $U_f \subseteq K[X]$ mit $V_f = V_f^G \oplus U_f$. Dann ist das Bild von $f \in K[X]$ unter Rey_G die Projektion von f auf $V_f^G \subseteq K[X]^G$ längs U_f. Eine effiziente Umsetzung dieser Idee werden wir später betrachten. Bevor wir uns allerdings weiter mit der Berechnung von Reynolds-Operatoren beschäftigen, werden wir unmittelbare Folgerungen aus dem letzten Theorem angeben (vgl. [DK02], Korollar 2.2.7, S. 48).

Korollar 7.1.5. *Sei G eine linear reduktive Gruppe, sei X eine affine G-Varietät und sei $\mathrm{Rey}_G : K[X] \to K[X]$ der eindeutig bestimmte Reynolds-Operator bzgl. G in $K[X]$.*

a) Ist $V \subseteq K[X]$ ein G-stabiler K-Untervektorraum von $K[X]$, so gilt $\mathrm{Rey}_G(V) = V^G$.

b) Der Reynolds-Operator ist $K[X]^G$-linear, d.h. insbesondere gilt $\mathrm{Rey}_G(g \cdot f) = g \cdot \mathrm{Rey}_G(f)$ für alle $g \in K[X]^G$ und alle $f \in K[X]$.

Nun wollen wir uns langsam einem Algorithmus zur Berechnung des Reynolds-Operators nähern. Dazu betrachten wir als ersten Schritt ein weiteres Beispiel eines Reynolds-Operators (vgl. [DK02], Beispiel 2.2.4, S. 46).

Beispiel 7.1.6. (Reynolds-Operator der multiplikativen Gruppe)
Sei $G = \mathrm{Mult}(K)$ die multiplikative Gruppe. Laut Beispiel 4.1.3 ist der Koordinatenring von G isomorph zum Laurent-Polynomring $K[t, t^{-1}]$. Sei nun X eine affine G-Varietät. Die durch die Gruppenoperation $\tau : G \times X \to X$ induzierte Koordinatenabbildung $\tau^* : K[X] \to K[G] \otimes K[X]$ ist definiert durch $\tau^*(f)(a, x) = f(a(x))$. Da $K[G] \otimes K[X]$ isomorph ist zu $K[X][t, t^{-1}]$ besitzt $\tau^*(f)$ für alle $f \in K[X]$ eine Darstellung der Form

$$\tau^*(f) = \sum_{i \in I_f} f_i t^i$$

mit endlich vielen $f_i \in K[X]$, d.h. $I_f \subseteq \mathbb{Z}$ ist endlich. O.B.d.A. können wir stets $0 \in I_f$ annehmen, denn andernfalls fügen wir 0 zu I_f hinzu und setzen $f_0 := 0$. Dann ist die durch

$$\mathrm{Rey}_G(f) = f_0$$

definierte Abbildung $\mathrm{Rey}_G : K[X] \to K[X]$ der Reynolds-Operator bzgl. G in $K[X]$.

Dieses Beispiel lässt sich kanonisch auf das n-fache kartesische Produkt der multiplikativen Gruppe erweitern, d.h. wir sehen an dem folgenden Beispiel, wie sich der Reynolds-Operator bzgl. eines n-dimensionalen Torus berechnen lässt (vgl. [DK02], Beispiel 4.5.2, S. 166).

Beispiel 7.1.7. (Reynolds-Operator von Tori)
Sei G ein n-dimensionaler Torus und sei X eine affine G-Varietät. Der Fall $n = 1$ wurde bereits in Beispiel 7.1.6 behandelt. Sei also $n \geq 2$.

Der Koordinatenring $K[G]$ ist isomorph zum Laurent-Polynomring $K[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}]$. Sei $\tau : G \times X \to X$ die Operation von G auf X und $\tau^* : K[X] \to K[G] \otimes K[X]$ die zugehörige Koordinatenabbildung. Wegen $K[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}] \cong K[X][x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}]$ ist analog zu Beispiel 7.1.6 für $f \in K[X]$ das Bild $\tau^*(f)$ ein Laurent-Polynom in $x_1, \ldots, x_n$ mit Koeffizienten in $K[X]$. Sei $f_0 \in K[X]$ der Koeffizient in $\tau^*(f)$ bei $x_1^0, \ldots, x_n^0$. Dann ist die Abbildung $\mathrm{Rey}_G : K[X] \to K[X]$ definiert durch

$$\mathrm{Rey}_G(f) = f_0$$

der Reynolds-Operator bzgl. G in $K[X]$. ◁

Um den nächsten Schritt zur effektiven Berechnung des Reynolds machen zu können, werden wir kurz einige Ergebnisse, die uns bisher begegnet sind, rekapitulieren lassen. Aus Satz 4.1.5 ist uns bekannt, dass die Zusammenhangskomponente G^0 einer linear algebraischen Gruppe G ein abgeschlossener Normalteiler von G mit endlichem Index ist, d.h. die Faktorgruppe G/G^0 ist endlich. Die Kommutatorgruppe $\mathrm{Kom}(G^0)$ von G^0 ist zusammenhängend, halbeinfach (vgl. [DK02], S. 167) und ein abgeschlossener Normalteiler von G^0. Die Faktorgruppe $G^0/\mathrm{Kom}(G^0)$ ist ein Torus (vgl. [DK02], S. 167). Für Faktorgruppen werden wir zunächst den relativen Reynolds-Operator angeben (vgl. [DK02], S. 75 f.).

Bemerkung 7.1.8. (Der relative Reynolds-Operator)
Sei G eine linear reduktive Gruppe, X eine affine G-Varietät und H eine Untergruppe von G. Dann ist $\mathrm{Rey}_{G/H} : K[X]^H \to K[X]^G$ der Reynolds-Operator bzgl. G/H. Er heißt auch der **relative Reynolds-Operator** bzgl. G/H in $K[X]$.

Speziell für endliche Gruppen lässt sich der relative Reynolds-Operator leicht angeben (vgl. [DK02], S. 75 f.).

Beispiel 7.1.9. Ist G eine endliche Gruppe und H derart, dass $\mathrm{ind}(G : H)$ nicht durch die Charakteristik von K geteilt wird, so ist $\mathrm{Rey}_{G/H}$ definiert durch

$$\mathrm{Rey}_{G/H}(f) = \frac{1}{\mathrm{ind}(G : H)} \sum_{a \in G/H} f^a.$$

Aufbauend auf der letzten Bemerkung betrachten wir folgendes Lemma (vgl. [DK02], Lemma 4.5.3, S. 167).

Lemma 7.1.10. *Sei G eine linear reduktive Gruppe, sei X eine affine G-Varietät und sei $N \subseteq G$ ein Normalteiler von G. Sei $\mathrm{Rey}_G : K[X] \to K[X]^G$ der Reynolds-Operator bzgl. G in $K[X]$, sei weiter $\mathrm{Rey}_N : K[X] \to K[X]^N$ der Reynolds-Operator bzgl. N und $\mathrm{Rey}_{G/N} : K[X]^N \to K[X]^G$ der relative Reynolds-Operator bzgl. G/N. Dann gilt:*

$$\mathrm{Rey}_G = \mathrm{Rey}_{G/N} \circ \mathrm{Rey}_N.$$

Durch zweimalige Anwendung dieses Lemmas ist es möglich, die Berechnung des Reynolds-Operators wie folgt zu vereinfachen.

Bemerkung 7.1.11. Für eine linear reduktive Gruppe G betrachten wir den Normalteiler G^0 von G und erhalten zunächst $\mathrm{Rey}_G = \mathrm{Rey}_{G/G^0} \circ \mathrm{Rey}_{G^0}$. Da G/G^0 eine endliche Gruppe ist, lässt sich Rey_{G/G^0} wie in Beispiel 7.1.3 bestimmen. Somit reduziert sich die Aufgabe auf die

Bestimmung von Rey_{G^0}. Dazu betrachtet man den Normalteiler $\mathrm{Kom}(G^0)$ von G^0 und wendet erneut das Lemma an. Dadurch folgt nun

$$\mathrm{Rey}_{G^0} = \mathrm{Rey}_{G^0/\mathrm{Kom}(G^0)} \circ \mathrm{Rey}_{\mathrm{Kom}(G^0)}.$$

Da $G^0/\mathrm{Kom}(G^0)$ ein Torus ist, wissen wir auch, wie sich $\mathrm{Rey}_{G^0/\mathrm{Kom}(G^0)}$ berechnen lässt (vgl. Beispiel 7.1.7). Somit bleibt nur die Berechnung von $\mathrm{Rey}_{\mathrm{Kom}(G^0)}$ offen, d.h. die Berechnung des Reynolds-Operators bzgl. einer beliebigen linear reduktiven Gruppe G lässt sich auf die Berechnung des Reynolds-Operators bzgl. der halbeinfachen und zusammenhängenden Untergruppe $\mathrm{Kom}(G^0)$ reduzieren.

Es genügt also, sich auf die Berechnung des Reynolds-Operators bzgl. zusammenhängender und halbeinfacher Gruppen G zu beschränken. Die Berechnung von $\mathrm{Rey}_G(f)$ für ein $f \in K[V]$ für derartige Gruppen wollen wir hier nur kurz umschreiben.

Bemerkung 7.1.12. Da G als lineare algebraische Varietät insbesondere eine affine Varietät ist, können wir den Reynolds-Operator bzgl. G in $K[G]$ betrachten, den wir mit $\mathcal{R}_G : K[G] \twoheadrightarrow K[G]^G$ bezeichnen. Wegen der Isomorphie $K[G]^G \cong K$ kann $\mathcal{R}_G$ auch als Element des Dualraums $K[G]^*$ angesehen werden. Durch die Operation $\bullet$ von $\mathcal{R}_G$ auf $K[V]$ lässt sich ein Reynolds-Operator bzgl. G in $K[V]$ definieren. Zur Berechnung von $\mathcal{R}_G \bullet f$ wird der sogenannte **Casimir-Operator** $c \in K[G]^*$ verwendet. Dieser operiert ebenfalls durch $\bullet$ auf f. Dieser wird nun iterativ solange auf f angewandt, bis die resultierenden Polynome linear abhängig sind, d.h. mit $f_0 := f$ werden folgende Berechnung durchgeführt:

$$\begin{aligned} f_1 &= c \bullet f_0 \\ f_2 &= c \bullet f_1 \\ &\vdots \\ f_\ell &= c \bullet f_{\ell-1} \end{aligned}$$

Aus diesem Ergebnis lässt sich dann sofort $\mathcal{R}_G \bullet f$ ablesen.

Für unsere Zwecke soll diese Ausführung reichen. Wir werden den Reynolds-Operator im algorithmischen Sinne also als „Black-Box" betrachten. Für weitere Informationen sei auf [DK02], Abschnitt 4.5, S. 166–180, verwiesen.

7.2 Berechnung fundamentaler Invarianten

In diesem Abschnitt wollen wir uns der Berechnung von Erzeugendensystemen von Invariantenringen linear reduktiver Gruppen zuwenden. Dank des Hilbertschen Endlichkeitssatzes ist die Existenz eines endlichen Erzeugendensystems in diesem Fall stets garantiert, wie wir bereits wissen.

Falls nicht anders angegeben sei K ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$ und sei G in diesem Abschnitt stets eine linear reduktive Gruppe, die sich für ein $l \in \mathbb{N}_+$ in $\mathbb{A}^l_K$ einbetten lässt und deren Koordinatenring $K[G]$ isomorph ist zum Restklassenring $K[z_1, \ldots, z_l]/\mathcal{I}(G)$, wobei $\mathcal{I}(G) \subseteq K[z_1, \ldots, z_l]$ das Verschwindungsideal der linear reduktiven Gruppe G ist (vgl. Bemerkung 4.1.12). Um die folgenden Resultate einfacher darstellen zu können und um eine klare Linie durchziehen zu können, werden wir die Aussagen stets über dem

Polynomring $P := K[x_1, \ldots, x_n]$ formulieren. Dieses Vorgehen ist aus folgendem Grund gerechtfertigt: Sei $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V. Dann ist der Koordinatenring $K[V]$ isomorph zum Polynomring $K[x_1, \ldots, x_n]$, d.h. alle folgenden Resultate über $K[x_1, \ldots, x_n]$ übertragen sich kanonisch auf den Koordinatenring $K[V]$, insbesondere gilt $K[V]^G \cong K[x_1, \ldots, x_n]^G$. Gesucht werden nun also die fundamentalen Invarianten von P^G, d.h. Polynome $f_1, \ldots, f_r \in P$ mit $P^G = K[f_1, \ldots, f_r]$. Sei dazu $\Phi : K[y_1, \ldots, y_r] \to P$ mit Unbestimmten $y_1, \ldots, y_r$ der durch $y_i \mapsto f_i$ definierte K-Algebra-Homomorphismus. Die von $f_1, \ldots, f_r$ erzeugte K-Unteralgebra von P ist dann das Bild von Φ.

7.2.1 Ein Invarianzkriterium

Bevor wir uns verschiedenen Methoden der Berechnung fundamentaler Invarianten zuwenden, wollen wir ein effizientes Kriterium angeben, mit dem wir entscheiden können, ob ein gegebenes Polynom $f \in K[x_1, \ldots, x_n]$ invariant ist oder nicht. Sind die fundamentalen Invarianten von P^G bereits bekannt, ist die Frage einfach zu beantworten und uns auch bereits bekannt. Dann ist schließlich nur zu untersuchen, ob f ein Element der K-Unteralgebra $K[f_1, \ldots, f_r]$ von P ist. Ist σ eine Eliminationsordnung für $\{x_1, \ldots, x_n\}$ in $Q := K[x_1, \ldots, x_n, y_1, \ldots, y_r]$, so folgt aus Satz 2.2.8:

$$f \in P^G \quad \Longleftrightarrow \quad \mathrm{NF}_{\sigma, \triangle_\Phi}(f) \in K[y_1, \ldots, y_r].$$

Dabei ist $\triangle_\Phi$ das Diagonalideal von Φ (vgl. Definition 2.2.4). In diesem Fall erhalten wir außerdem sofort eine explizite Darstellung von f in den fundamentalen Invarianten, denn mit $h := \mathrm{NF}_{\sigma, \triangle_\Phi}(f) \in K[y_1, \ldots, y_r]$ gilt $f = h(f_1, \ldots, f_r)$. Bilden $f_1, \ldots, f_r$ sogar eine σ-SAGBI-Basis von P^G für eine Termordnung σ, so können wir die Mitgliedschaft zu P^G mit Hilfe der SAGBI-Normalform (vgl. Definition 5.3.2) feststellen. In diesem Fall gilt (vgl. Satz 5.3.7)

$$f \in P^G \quad \Longleftrightarrow \quad \mathrm{NF}_{P^G}(f) = 0.$$

Beide Kriterien haben natürlich einen unübersehbaren Haken: Wir müssen die fundamentalen Invarianten kennen! Daher wollen wir nun noch ein Kriterium angeben, das ohne Kenntnis der fundamentalen Invarianten auskommt und für beliebige linear algebraische Gruppen G gültig ist. Durch Wahl einer Basis B von V lässt sich die durch die rationale Darstellung $(\rho, V)_G$ implizierte Gruppenoperation durch ihre Darstellungsmatrix beschreiben. Sei also im Folgenden B stets eine Basis von V und $\mathcal{M}_B^B(\rho) \in \mathrm{Mat}_n(K[z_1, \ldots, z_l]/\mathcal{I}(G))$ die Darstellungsmatrix der Gruppenoperation. Die Beschreibung der Gruppe durch ihr Verschwindungsideal und der Gruppenoperation durch ihre Darstellungsmatrix gibt uns nun die Möglichkeit, einen effizienten Test anzugeben für die Invarianz eines gegebenen Polynoms $f \in K[x_1, \ldots, x_n]$.

Satz 7.2.1. (Invarianzkriterium)
Sei $Q := K[z_1, \ldots, z_l, x_1, \ldots, x_n]$ und σ eine Termordnung im Polynomring Q. Ein Polynom $f \in P$ ist genau dann invariant, also ein Element von P^G, wenn gilt:

$$\mathrm{NF}_{\sigma, \mathcal{I}(G)Q}\left(f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})\right) = 0.$$

Dabei ist $f \in P$ als Polynom in Q zu betrachten.

Beweis: Es gilt genau dann $f \in P^G$, wenn $f^a = f$, also wenn $f - f(\mathcal{M}_B^B(\rho_a) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}}) = 0$ für alle $a \in G$ gilt. Da $\mathcal{M}_B^B(\rho)$ nur modulo $\mathcal{I}(G)$ eindeutig bestimmt ist, ist f also mit anderen

Worten genau dann invariant, wenn $f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})$ ein Element von $\mathcal{I}(G)Q$ ist. Laut Satz 2.2.1 ist dies genau dann der Fall, wenn gilt:

$$\mathrm{NF}_{\sigma,\mathcal{I}(G)Q}\left(f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})\right) = 0$$

□

Aufbauend auf diesem Invarianzkriterium wollen wir den durch

$$f \mapsto f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})$$

definierten K-Vektorraum-Homomorphismus $\Psi : P \to Q/\mathcal{I}(G)Q$ betrachten. Der letzte Satz impliziert sofort, dass alle Invarianten im Kern von Ψ enthalten sind, d.h. es gilt $P^G \subseteq \mathrm{Ker}(\Psi)$. Die Umkehrung gilt im allgemeinen jedoch nicht. Wir werden später auf diese Abbildung zurückkommen und zeigen, wie man sie zur Berechnung von Erzeugendensystemen verwenden kann. Zuletzt wollen wir dieses Kriterium noch als Algorithmus formulieren.

Algorithmus 7.1 : Invarianzkriterium.

Input : Erzeuger $h_1, \ldots, h_t \in K[z_1, \ldots, z_l]$ von $\mathcal{I}(G)$.
Input : Darstellungsmatrix $\mathcal{M}_B^B(\rho) = (q_{i,j})_{1 \leq i,j \leq n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l])$ der Gruppenoperation.
Input : $f \in P$.
Result : TRUE, falls $f \in P^G$, und FALSE sonst.

```
1 Q := K[z_1, ..., z_l, x_1, ..., x_n], Termordnung σ;
2 h := NF_{σ,I(G)Q}(f − f(M_B^B(ρ) · (x_1, ..., x_n)^tr));
3 return h = 0;
```

7.2.2 Der Derksen-Algorithmus

Wie auch bereits zu Beginn von Kapitel 6 ausgeführt wurde, dominiert die Endlichkeitsfrage die Invariantentheorie. Leider wissen wir, dass ein endliches Erzeugendensystem des Invariantenrings nicht immer garantiert werden kann. Dank Hilberts Endlichkeitssatz (siehe Theorem 6.2.5) gibt es für eine linear reduktive Gruppe und eine rationale Darstellung der Gruppe stets ein endliches Erzeugendensystem. Der Nachweis, dass eine Menge von Polynomen ein Erzeugendensystem des Invariantenrings bildet, kann - wie wir in Abschnitt 6.4 gesehen haben - unter Umständen recht aufwendig sein. Im Jahre 1999 hat Harm Derksen einen Algorithmus vorgestellt (vgl. [Der99] und [DK02], Abschnitt 4.1), der mit Hilfe von Gröbner-Basen und unter Verwendung des Reynolds-Operators im Falle von linear reduktiven Gruppen ein endliches Erzeugendensystem des Invariantenrings berechnet. Diesen Algorithmus und dessen Anwendung auf verschiedene interessante Beispiele wollen wir nun vorstellen. Eine Implementierung dieses Algorithmus findet sich seit der Version 2.14 im Computer-Algebra System Magma wieder (siehe [Mg14]).

Sei dazu im Folgenden G stets eine linear reduktive Gruppe über einem nicht-endlichen Körper K der Charakteristik $\mathrm{char}(K) = 0$. Wie in Abschnitt 6.2 versprochen, wollen wir nun den Beweis des Endlichkeitssatzes nachholen. Dieser ist besonders deshalb interessant für uns, weil die Ideen des Beweises den Schlüssel für den späteren Algorithmus liefern werden. Wir werden den Beweis im Spezialfall $K[x_1, \ldots, x_n]^G$ angeben, eine Verallgemeinerung auf $K[V]$ folgt aber unmittelbar. Der Polynomring $P := K[x_1, \ldots, x_n]$ sei dabei stets mit der Standardgraduierung versehen.

Theorem (Hilberts Endlichkeitssatz, Theorem 6.2.5)
Sei G eine linear reduktive Gruppe. Dann ist der Invariantenring $K[x_1, \ldots, x_n]^G$ eine endlich erzeugte K-Unteralgebra von $K[x_1, \ldots, x_n]$.

Beweis: Sei σ eine Termordnung auf der Menge der Terme $\mathbb{T}^n$ von $P = K[x_1, \ldots, x_n]$. Sei $I \subseteq P$ das Ideal, das von allen homogenen Invarianten positiven Grades erzeugt wird, d.h. es gilt $I = \langle P_+^G \rangle$. Dann gibt es endlich viele homogene Invarianten $f_1, \ldots, f_r \in P^G$, die I erzeugen. Nun wird gezeigt, dass diese Polynome den Invariantenring als K-Algebra erzeugen, d.h. das $P^G = K[f_1, \ldots, f_r]$ gilt, wobei die Inklusion $K[f_1, \ldots, f_r] \subseteq P^G$ offensichtlich richtig ist.

Da P^G eine standardgraduierte K-Unteralgebra von P ist, genügt es, ein homogenes Polynom $h \in P^G$ vom Grad d zu betrachten. Mittels Induktion nach d wird nun $h \in K[f_1, \ldots, f_r]$ bewiesen.

$d = 0$: Dann gilt $h \in K$ und wegen $K \subseteq K[f_1, \ldots, f_r]$ folgt die Behauptung.

$d > 0$: Wegen $h \in I$ und $\deg(h) > 0$ gibt es $g_1, \ldots, g_r \in P$ mit $h = \sum_{i=1}^r g_i f_i$. Ohne Einschränkung können $g_1, \ldots, g_r$ als homogen vom Grad $d_i := \deg(g_i) = d - \deg(f_i)$ für $i \in \{1, \ldots, r\}$ angenommen werden. Sei nun $\mathrm{Rey}_G : P \twoheadrightarrow P^G$ der Reynolds-Operator bzgl. G in P. Laut Korollar 7.1.5 gilt $\mathrm{Rey}_G(P_{d'}) = P_{d'}^G$ für alle $d' \in \mathbb{N}_+$. Ebenfalls aus Korollar 7.1.5 folgt

$$h = \mathrm{Rey}_G(h) = \mathrm{Rey}_G\left(\sum_{i=1}^r g_i f_i\right) = \sum_{i=1}^r \mathrm{Rey}_G(g_i f_i) = \sum_{i=1}^r \mathrm{Rey}_G(g_i) f_i.$$

Sei $i \in \{1, \ldots, r\}$. Da $\mathrm{Rey}_G(g_i)$ homogen vom Grad $< d$ ist, folgt laut Induktionsvoraussetzung also $\mathrm{Rey}_G(g_i) \in K[f_1, \ldots, f_r]$ und somit $h \in K[f_1, \ldots, f_r]$.

□

Insbesondere geht aus dem Beweis also hervor, dass ein homogenes Erzeugendensystem des homogenen Ideals $I = \langle P_+^G \rangle$ den Invariantenring als K-Algebra erzeugt. Dieses von allen homogenen Invarianten positiven Grades erzeugte Ideal, das im Beweis von Hilberts Endlichkeitssatz auftaucht, wird auch das **Hilbert-Ideal** genannt (vgl. [KK12]).

Definition 7.2.2. (Hilbert-Ideal)
Sei G eine linear reduktive Gruppe. Das von allen homogenen Invarianten positiven Grades erzeugte Ideal von P heißt das **Hilbert-Ideal** und wird mit HI_G bezeichnet.

Mit anderen Worten gilt also $\mathrm{HI}_G := \langle P_+^G \rangle$. Das Hilbert-Ideal werden wir nun genauer betrachten. Die fundamentalen Invarianten von P^G bilden ein Erzeugendensystem von HI_G, aber dieses ist natürlich nicht das einzige Erzeugendensystem des Hilbert-Ideals. Obwohl es von allen homogenen Invarianten positiven Grades erzeugt wird, besitzt es ein Erzeugendensystem mit homogenen, aber nicht notwendigerweise invarianten Polynomen (siehe Kapitel 2). Wie aus dem nächsten Satz hervorgeht, liefert uns der Reynolds-Operator schließlich aus einem beliebigen Erzeugendensystem des Hilbert-Ideals HI_G ein Erzeugendensystem des Invariantenrings (vgl. [DK02], Satz 4.1.1, S. 139).

Satz 7.2.3. (Hilbert-Ideal und Invariantenring)
Sei G eine linear reduktive Gruppe und sei $\mathrm{Rey}_G : P \twoheadrightarrow P^G$ der Reynolds-Operator bzgl. G in P. Seien $f_1, \ldots, f_r \in P$ homogene (aber nicht notwendigerweise invariante) Erzeuger von HI_G. Dann erzeugen auch $\mathrm{Rey}_G(f_1), \ldots, \mathrm{Rey}_G(f_r)$ das Hilbert-Ideal HI und es gilt:

$$P^G = K[\mathrm{Rey}_G(f_1), \ldots, \mathrm{Rey}_G(f_r)].$$

Aus den Eigenschaften des Reynolds-Operators (siehe Abschnitt 7.1) folgt sofort, dass auch die Bilder $\mathrm{Rey}_G(f_1), \ldots, \mathrm{Rey}_G(f_r)$ homogene Polynome sind, nun eben homogene Invarianten. Über dieses Erzeugendensystem des Invariantenrings lässt sich noch etwas mehr aussagen (vgl. [DK02], Bemerkung 4.1.2, S. 140).

Korollar 7.2.4. *In der Situation des letzten Satzes sei $\{f_1, \ldots, f_r\}$ ein minimales homogenes Erzeugendensystem von HI_G. Dann ist $\{\mathrm{Rey}_G(f_1), \ldots, \mathrm{Rey}_G(f_r)\}$ ein minimales Erzeugendensystem des Invariantenrings P^G.*

Aus dem Hilbertschen Endlichkeitssatz wissen wir nun, dass für linear reduktive Gruppen der Invariantenring stets eine endlich erzeugte K-Unteralgebra ist. Es drängt sich zwangsläufig eine Frage auf: Besitzt der Invariantenring in diesem Fall auch stets eine endliche SAGBI-Basis?

Bemerkung 7.2.5. Wie uns aus Kapitel 5 bestens bekannt ist, gibt es durchaus endlich erzeugte K-Unteralgebren, die keine endliche SAGBI-Basis besitzen (siehe Beispiel 5.1.6). Da in allen in dieser Arbeit betrachteten Beispielen stets auch eine endliche SAGBI-Basis existiert, liegt die Vermutung nahe, dass der Invariantenring einer linear reduktiven Gruppe stets eine endliche SAGBI-Basis besitzt. In diesem Fall wäre es möglich, aus einem Erzeugendensystem des Invariantenrings mit Hilfe von Prozedur SAGBI (siehe Seite 94) eine endliche SAGBI-Basis zu berechnen. Leider muss die Frage aber verneint werden: Es gibt Invariantenringe linear reduktiver Gruppen, die unabhängig von der Wahl einer Termordnung σ keine endliche σ-SAGBI-Basis besitzen. So hat Manfred Göbel im Jahre 2000 bewiesen, dass der Invariantenring $\mathbb{C}[x_1, \ldots, x_n]^{A_n}$ der alternierenden Gruppe A_n bzgl. jeder Termordnung σ keine endliche σ-SAGBI-Basis besitzt (vgl. [Göb00]). Allerdings ist es natürlich stets möglich, d-Grad-beschränkte SAGBI-Basen zu berechnen.

Damit kommen wir wieder zum eigentlichen Thema zurück, der Berechnung eines Erzeugendensystems. Gemäß Satz 7.2.3 ist es also möglich, aus einem beliebigen homogenen Erzeugendensystem des Hilbert-Ideals HI mit Hilfe des Reynolds-Operators ein Erzeugendensystem des Invariantenrings P^G zu erhalten. Somit kann das Problem der Bestimmung eines Erzeugendensystems des Invariantenrings auf die Bestimmung eines Erzeugendensystems von HI_G zurückgeführt werden, der wir uns nun widmen wollen. Wir betrachten dazu den K-Morphismus $\varphi : G \times V \to V \times V$ mit $\varphi(a, v) = (v, a(v))$, wobei wie zu Beginn $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V sei. Die zu φ gehörige Koordinatenabbildung ist der K-Algebra-Homomorphismus $\varphi^* : K[V \times V] \to K[G \times V]$. Der Koordinatenring $K[V \times V]$ ist dann isomorph zum Polynomring $K[x_1, \ldots, x_n, y_1, \ldots, y_n]$ und die Nullstellenmenge des Hilbert-Ideals $\mathrm{HI}_G = \langle P_+^G \rangle$,

$$\mathcal{Z}_K(\mathrm{HI}_G) = \{x \in K^n : \ f(x) = 0 \text{ für alle } f \in P_+^G\},$$

ist isomorph zum Hilbertschen Nullstellenkegel (siehe Definition 6.5.1)

$$\mathcal{N}_V = \{v \in V : \ f(v) = 0 \text{ für alle } f \in K[V]_+^G\},$$

d.h. mit anderen Worten gilt $\mathcal{Z}_K(\mathrm{HI}_G) \cong \mathcal{N}_V$. Nun betrachten wir den Zariski-Abschluss des Bildes des K-Morphismus φ und erhalten folgendes Lemma (vgl. [DK02], S. 140).

Lemma 7.2.6. *In der obigen Situation gilt $\overline{\mathrm{Im}(\varphi)} \cap (V \times \{0\}) = \mathcal{N}_V \times \{0\}$.*

Sei $Q := K[x_1, \ldots, x_n, y_1, \ldots, y_n] \cong K[V \times V]$ und sei $J \subseteq Q$ das zu $\mathcal{I}(\overline{\mathrm{Im}(\varphi)}) \subseteq K[V \times V]$ isomorphe Ideal. Dann folgt $\mathcal{Z}(J + \langle y_1, \ldots, y_n \rangle) = \mathcal{Z}(\mathrm{HI}_G Q + \langle y_1, \ldots, y_n \rangle)$ aus Lemma 7.2.6,

wobei $\mathrm{HI}_G Q$ das Hilbert-Ideal $\mathrm{HI}_G \subseteq P$ in Q bezeichnet. Der nachfolgende Satz zeigt uns, dass die beiden Ideale in $Q/\langle y_1, \ldots, y_n \rangle$ für linear reduktive Gruppen identisch sind (vgl. [DK02], Theorem 4.1.3, S. 140).

Satz 7.2.7. *Sei G eine linear reduktive Gruppe und sei $\varphi : G \times V \to V \times V$ der durch $\varphi(a, v) = (v, a(v))$ definierte K-Morphismus. Sei $J \subseteq Q$ das zum Ideal $\mathcal{I}(\overline{\mathrm{Im}(\varphi)}) \subseteq K[V \times V]$ isomorphe Ideal. Dann gilt:*

$$J + \langle y_1, \ldots, y_n \rangle = \mathrm{HI}Q + \langle y_1, \ldots, y_n \rangle.$$

Somit erhalten wir aus einem Erzeugendensystem des zum Verschwindungsideal $\mathcal{I}(\overline{\mathrm{Im}(\varphi)})$ isomorphen Ideals $J \subseteq Q$ sofort ein Erzeugendensystem des Hilbert-Ideals HI: Sind $f_1, \ldots, f_r$ Polynome in Q mit $J = \langle f_1, \ldots, f_r \rangle$, so gilt

$$\mathrm{HI} = \langle f_1(x_1, \ldots, x_n, 0, \ldots, 0), \ldots, f_r(x_1, \ldots, x_n, 0, \ldots, 0) \rangle \subseteq P.$$

Mit Satz 7.2.3 erhalten wir dann ein Erzeugendensystem des Invariantenrings P^G durch Anwendung des Reynolds-Operators. Bleibt also nur noch die Frage zu klären, wie man $\mathcal{I}(\overline{\mathrm{Im}(\varphi)})$ bzw. das Ideal J berechnen kann? Die Antwort auf diese Frage finden wir in Kapitel 3, genauer in Satz 3.2.19: Da V isomorph ist zu $\mathbb{A}^n_K$ folgt aus Satz 3.2.19, dass sich ein zu $\mathcal{I}(\overline{\mathrm{Im}(\varphi)})$ isomorphes Ideal J in Q mittels Elimination berechnen lässt (vgl. hierzu Satz 2.2.5). Dank der Beschreibung der Gruppe G und ihrer Operation auf V durch ihr Verschwindungsideal $\mathcal{I}(G) \subseteq K[z_1, \ldots, z_l]$ bzw. der Darstellungsmatrix $\mathcal{M}^B_B(\rho) \in \mathrm{Mat}_n(K[z_1, \ldots, z_l]/\mathcal{I}(G))$, können wir den Algorithmus zur Berechnung eines Erzeugendensystems des Invariantenrings damit formulieren (vgl. [DK02], Algorithmus 4.1.9, S. 146).

Algorithmus 7.2 : Derksen-Algorithmus zur Berechnung eines Erzeugendensystems von P^G

Input : Erzeuger $h_1, \ldots, h_t \in K[z_1, \ldots, z_l]$ von $\mathcal{I}(G)$.
Input : Darstellungsmatrix $\mathcal{M}^B_B(\rho) = (q_{i,j})_{1 \le i,j \le n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l])$ der Gruppenoperation.
Result : $S \subseteq P$ mit $P^G = K[S]$.

1 $Q := K[z_1, \ldots, z_l, x_1, \ldots, x_n, y_1, \ldots, y_n]$, σ eine Eliminationsordnung für $\{z_1, \ldots, z_l\}$;
2 **for** $i := 1$ **to** n **do**
3 $\quad g_i := y_i - \sum_{j=1}^n q_{i,j} x_j$;
4 $I := \langle h_1, \ldots, h_t, g_1, \ldots, g_n \rangle$;
5 Berechne eine σ-Gröbner-Basis G von I in Q;
6 $\widehat{G} := G \cap K[x_1, \ldots, x_n, y_1, \ldots, y_n]$;
7 $H := \{f(x_1, \ldots, x_n, 0, \ldots, 0) : \ f \in \widehat{G}\}$;
8 $S := \{\mathrm{Rey}_G(f) : \ f \in H\}$;
9 **return** S;

In Anlehnung an den Erfinder des Algorithmus, Harm Derksen, wird das von $\widehat{G}$ erzeugte Ideal in Algorithmus 7.2 mittlerweile auch das **Derksen-Ideal** genannt. In allgemeinerer Form liefert es auch die Basis für die Berechnung von Erzeugendensystemen von Invariantenringen nichtreduktiver Gruppen (vgl. [KK12]). Was in [DK02], Abschnitt 4.1, S. 139–146, nicht als Aussage fixiert ist, wird nun als Theorem festgehalten und anschließend bewiesen: Die Endlichkeit und Korrektheit von Algorithmus 7.2.

Theorem 7.2.8. (Korrektheit und Endlichkeit des DERKSEN-Algorithmus)
Sei G eine linear reduktive Gruppe, $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V und B eine Basis von V. Sei $P = K[x_1, \ldots, x_n]$ der zum Koordinatenring $K[V]$ isomorphe Polynomring. Dann berechnet Algorithmus 7.2 in endlich vielen Schritten ein endliches Algebra-Erzeugendensystem des Invariantenrings P^G.

Beweis: Der n-dimensionale K-Vektorraum V ist isomorph zu K^n. Deshalb zeigen wir ohne Einschränkung die Behauptung für $V = K^n$. Zunächst folgt aus dem Hilbertschen Endlichkeitssatz (siehe Theorem 6.2.5), dass die Existenz eines endlichen Algebra-Erzeugendensystems garantiert ist. Die Endlichkeit des Algorithmus ist offensichtlich. Es bleibt also nur die Korrektheit zu zeigen.

Dazu betrachten wir den durch $\varphi(a, v) = (v, a(v))$ definierten K-Morphismus $\varphi : G \times V \to V \times V$ mit seiner zugehörigen Koordinatenabbildung $\varphi^* : K[V \times V] \to K[G \times V]$. Aus Satz 3.2.19 folgt $\mathcal{I}(\overline{\mathrm{Im}(\varphi)}) = (\varphi^*)^{-1}(0)$ und somit $\mathcal{I}(\overline{\mathrm{Im}(\varphi)}) = \triangle_{\varphi^*} \cap K[V \times V]$, wobei $\triangle_{\varphi^*}$ das Diagonalideal von φ^* ist. Das Diagonalideal $\triangle_{\varphi^*}$ ist gerade das Verschwindungsideal des Graphen

$$\Gamma_\varphi = \{(a, v, v, a(v)) : \ a \in G, \ v \in V\} \subseteq (G \times V) \times (V \times V)$$

von φ. Somit folgt $\mathcal{I}(\overline{\mathrm{Im}(\varphi)}) = \mathcal{I}(\Gamma_\varphi) \cap K[V \times V]$.

Wie man sieht, ist in Γ_φ eine Komponente doppelt vorhanden. Deshalb genügt es, zur Bestimmung des Verschwindungsideals $\mathcal{I}(\Gamma_\varphi)$ die Teilmenge $\Gamma := \{(a, v, a(v)) : \ a \in G, \ v \in V\}$ von $G \times V \times V$ zu betrachten. Dank der Einbettung $\iota : G \hookrightarrow \mathbb{A}^l_K$ für ein $l \in \mathbb{N}_+$ können wir Γ auch als Teilmenge von $\mathbb{A}^l_K \times V \times V$ ansehen. Wegen $K[V \times V] \cong K[x_1, \ldots, x_n, y_1, \ldots, y_n]$ und $K[\mathbb{A}^l_K \times V \times V] \cong K[z_1, \ldots, z_l, x_1, \ldots, x_n, y_1, \ldots, y_n]$ können wir $\mathcal{I}(\Gamma)$ als Ideal in dem Polynomring $Q := K[z_1, \ldots, z_l, x_1, \ldots, x_n, y_1, \ldots, y_n]$ auffassen. Dank der Einbettung ι ist G eine affine Untervarietät von $\mathbb{A}^\ell_K$ mit Verschwindungsideal $\mathcal{I}(G)$, das erzeugt wird von den Polynomen $h_1, \ldots, h_t \in K[z_1, \ldots, z_\ell]$. Sei $\mathcal{M}^B_B(\rho) = (q_{i,j})_{1 \le i,j \le n}$ mit $q_{i,j} \in K[z_1, \ldots, z_\ell]/\mathcal{I}(G)$ für alle $i, j \in \{1, \ldots, n\}$ die Darstellungsmatrix der Gruppenoperation von G auf V bzgl. B. Dann besteht folgender Zusammenhang zwischen den Koordinaten bzw. den Unbestimmten $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathcal{M}^B_B(\rho) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

der sich in den Polynomen $g_i := y_i - \sum_{j=1}^n q_{i,j} x_j \in Q$ mit $i \in \{1, \ldots, n\}$ widerspiegelt. Somit gilt $\mathcal{I}(\Gamma) = \langle h_1, \ldots, h_t, g_1, \ldots, g_n \rangle$. In den Schritten (1)-(5) von Algorithmus 7.2 wird also eine σ-Gröbner-Basis von $\mathcal{I}(\Gamma)$ in Q bestimmt, wobei σ eine Eliminationsordnung für $\{z_1, \ldots, z_l\}$ ist. Wegen $\mathcal{I}(\Gamma) = \mathcal{I}(\Gamma_\varphi)$ folgt aus Theorem 2.2.3, dass $\widehat{G}$ eine $\widehat{\sigma}$-Gröbner-Basis von $\mathcal{I}(\Gamma_\varphi) \cap K[x_1, \ldots, x_n, y_1, \ldots, y_n]$ ist und damit eine $\widehat{\sigma}$-Gröbner-Basis von $\mathcal{I}(\overline{\mathrm{Im}(\varphi)})$ als Ideal in $K[x_1, \ldots, x_n, y_1, \ldots, y_n]$.

Aus Satz 7.2.7 folgt, dass $H := \{f(x_1, \ldots, x_n, 0, \ldots, 0) : \ f \in \widehat{G}\}$ ein Erzeugendensystem des Hilbert-Ideals $\mathrm{HI} = \langle P^G_+ \rangle$ ist. Gemäß Satz 7.2.3 ist $S := \{\mathrm{Rey}_G(f) : \ f \in H\}$ ein Algebra-Erzeugendensystem von P^G, d.h. es gilt $P^G = K[S]$. □

Zur Demonstration dieses Algorithmus wollen wir nun an dieser Stelle ein Beispiel betrachten, von dem wir das Ergebnis bereits kennen. Denn es handelt sich hier nur um Vektorinvarianten der orthogonalen Gruppe (siehe Korollar 6.4.4). Weitere Beispiele folgen in den späteren Anwendungen. Die Anwendung des Reynolds-Operators wurde hier nicht per Hand durchgeführt, sondern es wurden die Ergebnisse aus MAGMA (siehe [Mg14]) verwendet.

Beispiel 7.2.9. Wir betrachten die orthogonale Gruppe $G := \mathrm{O}_2(K)$, deren Koordinatenring $K[G]$ isomorph ist zu $K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$, wobei $\mathcal{I}(G)$ von den Polynomen

$$\begin{aligned} h_1 &:= z_{1,1}^2 + z_{1,2}^2 - 1, \\ h_2 &:= z_{1,1}z_{2,1} + z_{1,2}z_{2,2}, \\ h_3 &:= z_{2,1}^2 + z_{2,2}^2 - 1 \end{aligned}$$

erzeugt wird (vgl. Beispiel 4.3.27). Wir lassen G wie in Beispiel 4.3.27 auf dem K-Vektorraum $\mathrm{Mat}_{3,2}(K)$ operieren. Sei $(\rho, \mathrm{Mat}_{3,2}(K))_G$ die durch die Operation von G induzierte rationale Darstellung. Bzgl. der kanonischen Basis $B = (\mathcal{I}_{i,j} : \ 1 \leq i \leq 3, \ 1 \leq j \leq 2)$ von $\mathrm{Mat}_{3,2}(K)$ ist $\mathcal{M}_B^B(\rho)$ bekanntlich die Matrix

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} z_{1,1} & z_{2,1} & 0 & 0 & 0 & 0 \\ z_{1,2} & z_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & z_{1,1} & z_{2,1} & 0 & 0 \\ 0 & 0 & z_{1,2} & z_{2,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & z_{1,1} & z_{2,1} \\ 0 & 0 & 0 & 0 & z_{1,2} & z_{2,2} \end{pmatrix}.$$

Wir berechnen in $Q := K[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}, x_1, \ldots, x_6, y_1, \ldots, y_6]$ zunächst die restlichen Erzeuger des Derksen-Ideals I. Es gilt:

$$\begin{aligned} g_1 &:= y_1 - z_{1,1}x_1 - z_{2,1}x_2 \\ g_2 &:= y_2 - z_{1,2}x_1 - z_{2,2}x_2 \\ g_3 &:= y_3 - z_{1,1}x_3 - z_{2,1}x_4 \\ g_4 &:= y_4 - z_{1,2}x_3 - z_{2,2}x_4 \\ g_5 &:= y_5 - z_{1,1}x_5 - z_{2,1}x_5 \\ g_6 &:= y_6 - z_{1,2}x_5 - z_{2,2}x_6 \end{aligned}$$

Die σ-Gröbner-Basis G von $I := \langle h_1, h_2, h_3, g_1, \ldots, g_6 \rangle$ mit $\sigma = \mathrm{Elim}(z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2})$ besteht aus $\#G = 105$ Elementen, weshalb wir an dieser Stelle auf eine Auflistung derer verzichten wollen. Die $\widehat{\sigma}$-Gröbner-Basis von $I \cap K[x_1, \ldots, x_6, y_1, \ldots, y_6]$ besteht immer noch aus $\#\widehat{G} = 30$ Elementen, weshalb wir auch sie hier nicht explizit angeben.

Für die Menge H erhalten wir schließlich:

$$\begin{aligned} H := \{&x_1^2 + x_2^2,\ x_1x_3 + x_2x_4,\ x_3^2 + x_4^2,\ x_1x_5 + x_2x_6,\ x_3x_5 + x_4x_6,\ x_5^2 + x_6^2, \\ &x_1x_2x_4 - x_2^2x_3,\ x_1x_2x_6 - x_2^2x_5,\ x_1x_4^2 - x_2x_3x_4,\ x_1x_4x_6 - x_2x_4x_5, \\ &x_3x_4x_6 - x_4^2x_5,\ x_1x_4x_6 - x_2x_3x_6,\ x_1x_6^2 - x_2x_5x_6,\ x_3x_6^2 - x_4x_5x_6\} \end{aligned}$$

Dieses Erzeugendensystem des Hilbert-Ideals HI_G lässt sich noch minimieren. Das minimale Erzeugendensystem von HI_G lautet:

$$H' := \{x_1^2 + x_2^2,\ x_1x_3 + x_2x_4,\ x_3^2 + x_4^2,\ x_1x_5 + x_2x_6,\ x_3x_5 + x_4x_6,\ x_5^2 + x_6^2\}.$$

Diese Polynome sind bereits alle invariant, sodass wir $S = H'$ durch Anwendung des Reynolds-Operators als Ergebnis erhalten. Eine σ-SAGBI-Basis des Invariantenrings lautet wie folgt:

$$\begin{aligned} &\{x_1^2 + x_2^2, x_1x_3 + x_2x_4, x_3^2 + x_4^2, x_1x_5 + x_2x_6, \\ &x_3x_5 + x_4x_6, x_5^2 + x_6^2, x_2^2x_5^2 - 2x_1x_2x_5x_6 + x_1^2x_6^2, x_4^2x_5^2 - 2x_3x_4x_5x_6 + x_3^2x_6^2, \\ &x_2x_4x_5^2 - x_2x_3x_5x_6 - x_1x_4x_5x_6 + x_1x_3x_6^2, x_2^2x_3x_5 - x_1x_2x_4x_5 - x_1x_2x_3x_6 + x_1^2x_4x_6, \\ &x_2x_3x_4x_5 - x_1x_4^2x_5 - x_2x_3^2x_6 + x_1x_3x_4x_6, x_2^2x_3^2 - 2x_1x_2x_3x_4 + x_1^2x_4^2\} \end{aligned}$$

◁

Der Derksen-Algorithmus (siehe Algorithmus 7.2) beinhaltet als wesentlichen Schritt die Berechnung des Hilbert-Ideals HI_G. Um spätere Algorithmen besser darstellen zu können, werden wir diese Berechnung noch als eigenständigen Algorithmus angeben. Die Endlichkeit und Korrektheit folgt natürlich unmittelbar aus dem Beweis des letzten Satzes.

Algorithmus 7.3 : Berechnung des Hilbert-Ideals

Input : Erzeuger $h_1, \ldots, h_t \in K[z_1, \ldots, z_l]$ von $\mathcal{I}(G)$.
Input : Darstellungsmatrix $\mathcal{M}_B^B(\rho) = (q_{i,j})_{1 \leq i,j \leq n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l])$ der Gruppenoperation.
Result : $H \subseteq P$ mit $\mathrm{HI}_G = \langle H \rangle$.

1 $Q := K[z_1, \ldots, z_l, x_1, \ldots, x_n, y_1, \ldots, y_n]$, $\sigma := \mathrm{Elim}(\{z_1, \ldots, z_l\})$;
2 **for** $i := 1$ **to** n **do**
3 $\quad g_i := y_i - \sum_{j=1}^{n} q_{i,j} x_j$;
4 $I := \langle h_1, \ldots, h_t, g_1, \ldots, g_n \rangle$;
5 Berechne eine σ-Gröbner-Basis G von I in Q;
6 $\widehat{G} := G \cap K[x_1, \ldots, x_n, y_1, \ldots, y_n]$;
7 $H := \{f(x_1, \ldots, x_n, 0, \ldots, 0) : \ f \in \widehat{G}\}$;
8 **return** H;

7.2.3 Die Hilbert-Reihen-Methode

Natürlich ist der Derksen-Algorithmus 7.2 nicht die einzige Möglichkeit ein Erzeugendensystem des Invariantenrings einer linear reduktiven Gruppe zu berechnen. Auch wenn dieser Algorithmus sehr einfach wirkt, liegt die größte Schwierigkeit bei seiner Umsetzung in der Implementierung des Reynolds-Operators. Wie wir in Abschnitt 7.1 angedeutet haben, ist diese nicht ganz trivial. Hat man allerdings den Reynolds-Operator zur Verfügung, ist die Berechnung eines Erzeugendensystems mit obigem Algorithmus sehr effizient machbar. Wir wollen hier aber noch andere Möglichkeiten angeben, die fundamentalen Invarianten zu berechnen, die ohne Reynolds-Operator auskommen. Dazu werden wir Methoden der Linearen Algebra verwenden. Dies geschieht auf Kosten geringerer Effizienz, was für unsere Anwendungen aber keine gravierende Einschränkung darstellt. Eine erste Möglichkeit bietet sich durch die Hilbert-Reihe (vgl. Abschnitt 6.6) des Invariantenrings P^G. Wir wollen nun also zusätzlich voraussetzen, dass uns die Hilbert-Reihe HS_{P^G} bereits bekannt ist. Dann können wir sofort ein einfaches Kriterium angeben, um zu überprüfen, ob bestimmte Polynome ein Erzeugendensystem des Invariantenrings bilden (vgl. [DK02], S. 69).

Lemma 7.2.10. (Hilbert-Reihen-Kriterium)
Sei G eine linear algebraische Gruppe und sei $S := \{f_1, \ldots, f_r\} \subseteq P^G$ eine Menge homogener Invarianten. Dann gilt:

$$P^G = K[S] \quad \Longleftrightarrow \quad \mathrm{HS}_{P^G}(z) = \mathrm{HS}_{K[S]}(z)$$

Aufbauend auf diesem Kriterium lässt sich ein weiterer Algorithmus zur Berechnung fundamentaler Invarianten für linear reduktive Gruppen wie folgt formulieren (vgl. [DK02], Algorithmus 2.6.1, S. 69). Dieses Verfahren funktioniert allerdings auch für nicht linear reduktive Gruppen. In diesem Fall erhalten wir eine enumerierte Prozedur, die genau dann terminiert, wenn ein

endliches Erzeugendensystem existiert.

Prozedur InvHilbert

Input : Hilbert-Reihe $\mathrm{HS}_{P^G}(z)$ einer linear algebraischen Gruppe.

Result : Ein endliches Erzeugendensystem $S \subseteq P$ mit $P^G = K[S]$, falls dieses existiert.

1 $S := \emptyset$;
2 Berechne homogene Invarianten $f_1, \ldots, f_r \in P^G$;
3 $S := S \cup \{f_1, \ldots, f_r\}$;
4 **while** $\mathrm{HS}_{P^G}(z) \neq \mathrm{HS}_{K[S]}(z)$ **do**
5 $\quad \psi(z) := \mathrm{HS}_{P^G}(z) - \mathrm{HS}_{K[S]}(z)$;
6 $\quad$ Schreibe $\psi(z)$ in der Form $\psi(z) = a_d z^d +$ (Monome höheren Grades) mit $a_d \in \mathbb{N}_+$;
7 $\quad$ Bestimme $g_1, \ldots, g_{a_d} \in P_d^G$ so, dass P_d^G erzeugt wird von $K[S]_d$ und $g_1, \ldots, g_{a_d}$;
8 $\quad S := S \cup \{g_1, \ldots, g_{a_d}\}$;
9 **return** S;

Für linear reduktive Gruppen ist die Terminierung der Prozedur also sichergestellt, d.h. in diesem Fall wird die Prozedur zum Algorithmus. Die Korrektheit folgt aus folgendem Satz (vgl. in Ansätzen [DK02], S. 69 f.).

Satz 7.2.11. *Sei G eine linear algebraische Gruppe und $\mathrm{HS}_{P^G}(z)$ die Hilbert-Reihe des Invariantenrings P^G. Die enumerierte Prozedur InvHilbert terminiert genau dann, wenn P^G ein endliches Algebra-Erzeugendensystem besitzt. In diesem Fall berechnet Prozedur InvHilbert ein minimales Algebra-Erzeugendensystem von P^G.*

Beweis: Seien $f_1, \ldots, f_r \in P^G$ homogene Invarianten. Die Darstellung in Schritt (6) der Differenz $\psi(z) = \mathrm{HS}_{P^G} - \mathrm{HS}_{K[S]}$ in der Form $\psi(z) = a_d z^d +$ (Monome höheren Grades) bedeutet, dass $K[f_1, \ldots, f_r]$ bereits alle Invarianten vom Grad $< d$ beinhaltet und genau a_d invariante Erzeuger vom Grad d fehlen. Diese a_d Erzeuger $g_1, \ldots, g_{a_d}$ werden im nächsten Schritt berechnet.

Sei $d_z(\psi(z))$ die kleinste natürliche Zahl d, für die der Term z^d in $\psi(z)$ vorkommt. Dann wird $d_z(\psi(z))$ in jedem Schleifendurchlauf erhöht. Sei $D \in \mathbb{N}$ eine obere Schranke für $\beta(P^G)$, d.h. P^G wird von Invarianten vom Grad $\leq D$ erzeugt. Nach D Durchläufen gilt entweder

$$\psi(z) = a_d z^d + (\text{Monome höheren Grades})$$

mit $d > D$ oder $\mathrm{HS}_{P^G}(z) = \mathrm{HS}_{K[S]}(z)$. In letzterem Fall folgt aus dem letzten Lemma sofort $P^G = K[S]$, d.h. P^G ist endlich erzeugt mit Erzeugendensystem S. Die Minimalität von S folgt unmittelbar aus der Konstruktion der Polynome g_i. Im anderen Fall besitzt P^G kein endliches Algebra-Erzeugendensystem und der Algorithmus terminiert nicht. □

7.2.4 Die Lineare-Algebra-Methode

Befasst man sich etwas intensiver mit der Prozedur InvHilbert, so wird man sofort auf zwei „Probleme" stoßen:

(1) Wie berechnet man Invarianten $f_1, \ldots, f_r \in P^G$?

(2) Wie bestimmt man die fehlenden Erzeuger $g_1, \ldots, g_{a_d} \in P_d^G$, die P_d^G zusammen mit $K[S]_d$ erzeugen?

Eine mögliche Antwort auf die erste Frage kennen wir bereits: Das Hilbert-Ideal liefert hier unter Umständen eine einfache Möglichkeit. Die zweite Frage lässt sich mit Methoden der Linearen Algebra lösen. Wir werden beide Lösungsansätze mit der Idee von Prozedur InvHilbert kombinieren und einen weiteren Algorithmus angeben, der die Hilbert-Reihe von P^G nicht erfordert und auch keine Polynome $f_1, \ldots, f_r$ vorweg berechnen muss. Zunächst wollen wir hier an verschiedene vorausgegangene Resultate erinnern. Allen voran zu nennen ist die Eigenschaft, dass P^G eine standardgraduierte K-Algebra ist. Auch von Bedeutung ist die Tatsache, dass der Reynolds-Operator Rey_G den Grad erhält. Weiter haben wir in Algorithmus 7.2 gesehen, dass der Schlüssel zu einem Algebra-Erzeugendensystem des Invariantenrings im Hilbert-Ideal $\mathrm{HI}_G = \langle P_+^G \rangle$ liegt. Wie dieses berechnet werden kann, ist ebenfalls Algorithmus 7.2 oder gesondert Algorithmus 7.3 zu entnehmen. Sind $f_1, \ldots, f_r \in P$ homogene Erzeuger von HI_G, so wissen wir aus Satz 7.2.3, dass P^G als K-Algebra von den Polynomen $\mathrm{Rey}_G(f_1), \ldots, \mathrm{Rey}_G(f_r)$ erzeugt wird. Ist ein Polynom f_i für ein $i \in \{1, \ldots, r\}$ bereits invariant, würde uns das Hilbert-Ideal den Ansatzpunkt für den obigen Algorithmus liefern. Aber wir werden nun sehen, dass selbst dann, wenn dies nicht der Fall ist, ein Erzeugendensystem aus dem Hilbert-Ideal ohne Verwendung des Reynolds-Operators berechnet werden kann.

Was wir dazu benötigen, ist ein Algorithmus, der uns zu vorgegeben Polynomen eine ganz bestimmte Menge von Invarianten berechnet. Tobias Kamke und Gregor Kemper geben in [KK12] nachfolgenden Algorithmus an, allerdings in einer etwas allgemeineren Version, die für uns aber nicht relevant ist, weshalb wir uns auf folgende speziellere Version beschränken.

Algorithmus 7.4 : Berechnung einer Menge von Invarianten.

Input : $\mathcal{I}(G)$.
Input : Darstellungsmatrix $\mathcal{M}_B^B(\rho) = (q_{i,j})_{1 \le i,j \le n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l])$ der Gruppenoperation.
Input : Eine Menge $S = \{h_1, \ldots, h_r\} \subseteq P$ von Polynomen.
Result : Eine K-Basis C von $P^G \cap \langle S \rangle_K$.

1 Setze $Q := K[z_1, \ldots, z_\ell, x_1, \ldots, x_n]$ und wähle Termordnung σ in Q;
2 $T := \emptyset$;
3 **for** $i = 1$ **to** r **do**
4 $\quad f_i := \mathrm{NF}_{\sigma, \mathcal{I}(G)Q}(h_i - h_i(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}}))$;
5 $\quad T := T \cup \{f_i\}$;
6 Berechne eine K-Basis D des K-Vektorraums $\{(\alpha_1, \ldots, \alpha_r) \in K^r : \sum_{i=1}^r \alpha_i f_i = 0\}$;
7 Berechne eine K-Basis D' des K-Vektorraums $\{(\alpha_1, \ldots, \alpha_r) \in K^r : \sum_{i=1}^r \alpha_i h_i = 0\}$;
8 Wähle eine Menge $D'' \subseteq D$ so, dass $D' \cup D''$ linear unabhängig ist und $\#(D' \cup D'') = \#D$ gilt;
9 $C := \{\sum_{i=1}^r \alpha_i h_i : (\alpha_1, \ldots, \alpha_r) \in D''\}$;
10 **return** C;

Die Endlichkeit des Algorithmus ist hier ohnehin kein kritischer Punkt, die Korrektheit folgt aus dem nächsten Satz.

Satz 7.2.12. *Sei G eine linear algebraische Gruppe, $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V, sei B eine Basis von V und $S \subseteq P$ eine endliche Menge von Polynomen. Dann berechnet Algorithmus 7.4 in endlich vielen Schritten eine K-Basis des K-Vektorraums $P^G \cap \langle S \rangle_K$.*

Beweis: Die Endlichkeit von Algorithmus 7.4 ist klar. Sei $S := \{h_1, \ldots, h_r\} \subseteq P$. Wir betrachten das Polynom $f = \sum_{i=1}^r \alpha_i h_i$ mit $(\alpha_1, \ldots, \alpha_r) \in D$, wobei D eine K-Basis des K-

Vektorraums $\{(\alpha_1, \ldots, \alpha_r) \in K^r : \sum_{i=1}^r \alpha_i f_i = 0\}$ ist. Dann folgt sofort $f \in \langle S \rangle_K$. Weiter gilt

$$\begin{aligned}
&\mathrm{NF}_{\mathcal{I}(G)Q}(f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})) \\
&\qquad = \mathrm{NF}_{\mathcal{I}(G)Q}\left(\sum_{i=1}^r \alpha_i h_i - \sum_{i=1}^r \alpha_i h_i(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})\right) \\
&\qquad\qquad = \sum_{i=1}^r \alpha_i \mathrm{NF}_{\mathcal{I}(G)Q}(h_i - h_i(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})) = \sum_{i=1}^r \alpha_i f_i = 0, \quad (*)
\end{aligned}$$

also folgt auch $f \in P^G$ aus Satz 7.2.1 und damit $f \in P^G \cap \langle S \rangle_K$. Somit ist die Menge

$$C' := \left\{ \sum_{i=1}^r \alpha_i h_i : \ (\alpha_1, \ldots, \alpha_r) \in D \right\}$$

eine Teilmenge von $P^G \cap \langle S \rangle_K$.

Sei nun $f \in P^G \cap \langle S \rangle_K$. Dann gibt es $(\alpha_1, \ldots, \alpha_r) \in K^r$ mit $f = \sum_{i=1}^r \alpha_i h_i$ wegen $f \in \langle S \rangle_K$. Wegen $f \in P^G$ gilt $\mathrm{NF}_{\mathcal{I}(G)Q}(f - f(\mathcal{M}_B^B(\rho) \cdot (x_1, \ldots, x_n)^{\mathrm{tr}})) = 0$ und mit $(*)$ folgt schließlich $(\alpha_1, \ldots, \alpha_r) \in D$. Somit ist C' ein Erzeugendensystem von $P^G \cap \langle S \rangle_K$. Durch Auswahl der Teilmenge D'' von D mit $\sum_{i=1}^r \alpha_i h_i \neq 0$ erhalten wir eine Basis C von $P^G \cap \langle S \rangle_K$ aus C'. □

Die Situation in Algorithmus 7.4 ist zwar recht allgemein gehalten, wenngleich sich in [KK12] eine noch allgemeinere Version finden lässt, aber eine spezielle Menge S ist dennoch besonders interessant und wird deshalb in folgendem Korollar festgehalten.

Korollar 7.2.13. *Sei G eine linear algebraische Gruppe, $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V, sei B eine Basis von V und $S \subseteq P$ die Menge aller Terme vom Grad $d \in \mathbb{N}_+$. Dann berechnet Algorithmus 7.4 eine K-Basis von P_d^G, des K-Vektorraums aller Invarianten vom Grad d.*

In diesem Fall sind die Schritte (7) und (8) von Algorithmus 7.4 überflüssig. Denn da S eine K-Basis von P_d ist, folgt

$$\left\{ (\alpha_1, \ldots, \alpha_r) \in K^r : \ \sum_{i=1}^r \alpha_i h_i = 0 \right\} = \{0\}$$

und damit $D' = \emptyset$. Dies gilt natürlich in analoger Weise für alle linear unabhängigen Mengen S. Wir wollen uns nun damit beschäftigen, wie sich Algorithmus 7.4 verwenden lässt, um ein Erzeugendensystem des Invariantenrings zu berechnen. Dazu bauen wir auf dem Hilbert-Ideal auf. Sei H eine Menge homogener (aber nicht notwendigerweise invarianter) Erzeuger von $\mathrm{HI}_G = \langle P_+^G \rangle$ und sei $D = \{\deg(f) : \ f \in H\} \subseteq \mathbb{N}$. Ohne Einschränkung sei H ein minimales Erzeugendensystem von HI_G. Weiter sei $H_d = \{f \in H : \ \deg(f) = d\}$ für ein $d \in D$ die Teilmenge von H bestehend aus allen Polynomen vom Grad d in H. Da der Reynolds-Operator den Grad erhält, folgt aus Satz 7.2.3, dass es maximal $\#H_d$ fundamentale Invarianten vom Grad d geben kann. Damit lassen sich aus einem Erzeugendensystem des Hilbert-Ideals ähnliche Informationen ablesen wie aus der Hilbert-Reihe. Wir können dies nun ausnutzen, unser Wissen aus Algorithmus 7.2, Algorithmus 7.1 und Algorithmus 7.4 einfließen lassen und erhalten einen

einfachen Algorithmus, der uns ein Erzeugendensystem des Invariantenrings berechnet.

Algorithmus 7.5 : Lineare-Algebra-Methode zur Berechnung eines Erzeugendensystems des Invariantenrings.

Input : $\mathcal{I}(G)$ einer linear reduktiven Gruppe G.
Input : Darstellungsmatrix $\mathcal{M}_B^B(\rho) = (q_{i,j})_{1 \le i,j \le n} \in \mathrm{Mat}_n(K[z_1, \ldots, z_l])$ der Gruppenoperation.
Result : $S \subseteq P$ mit $P^G = K[S]$.

1 Berechne mit Algorithmus 7.3 minimales Erzeugendensystem H des Hilbert-Ideals HI_G;
2 $S := H \cap P^G$; /* Verwende Algorithmus 7.1 */
3 **if** $S \neq H$ **then**
4 $D := \{\deg(f) : \ f \in S \setminus H\}$;
5 **while** $D \neq \emptyset$ **do**
6 $d := \min D$;
7 $D := D \setminus \{d\}$;
8 $S_{\le d} := \{g_1, \ldots, g_s\}$;
9 Berechne K-Basis C' von $\mathrm{HI}_d / \langle g_1^{x_1} \cdots g_s^{x_s} : \ x_1 \deg(g_1) + \ldots + x_s \deg(g_s) = d \rangle$;
10 Berechne mit Algorithmus 7.4 eine K-Basis C von $P^G \cap \langle C' \rangle_K$;
11 $S := S \cup C$;
12 **return** S;

Ein sehr ähnlicher Algorithmus ist im Computer-Algebra-System CoCoA (siehe [RAB15]) in dem Paket `invariants.cpkg` von Alesandro Del Padrone bereits implementiert (vgl. [Del02]), wenngleich diese Implementierung etwas von Algorithmus 7.5 abweicht. Zudem mussten die Implementierung der aktuellen CoCoA-Syntax angepasst und kleine Fehler korrigiert werden. Die Korrektheit und Endlichkeit von Algorithmus 7.5 folgt aus dem nächsten Theorem.

Theorem 7.2.14. *Sei G eine linear reduktive Gruppe, $(\rho, V)_G$ eine rationale Darstellung von G in einem n-dimensionalen K-Vektorraum V und B eine Basis von V. Sei $P = K[x_1, \ldots, x_n]$ der zum Koordinatenring $K[V]$ isomorphe Polynomring. Dann berechnet Algorithmus 7.5 in endlich vielen Schritten ein endliches Algebra-Erzeugendensystem des Invariantenrings P^G.*

Beweis: Für die Endlichkeit des Algorithmus ist einzig die while-Schleife zwischen den Schritten (5) und (11) kritisch. Wegen $S \neq H$ ist $D = \{\deg(f) : \ f \in S \setminus H\}$ in Schritt (4) nicht leer, d.h. die Schleife wird mindestens ein Mal durchlaufen. Da D endlich ist und in Schritt (7) laufend verkleinert wird, terminiert die while-Schleife und damit der Algorithmus nach endlich vielen Schritten.

Wir betrachten zuerst den Fall, dass nach Schritt (2) $S = H$ gilt. Das heißt mit anderen Worten, alle Erzeuger des Hilbert-Ideals HI_G sind invariant. Laut Satz 7.2.3 ist das Bild von H unter dem Reynolds-Operator, $\mathrm{Rey}_G(H) = \{\mathrm{Rey}_G(h) : \ h \in H\}$, ein Algebra-Erzeugendensystem von P^G. Wegen $\mathrm{Rey}_G(f) = f$ für alle $f \in P^G$ folgt somit $\mathrm{Rey}_G(H) = H$, d.h. S ist ein Algebra-Erzeugendensystem von P^G.

Sei nun $S \neq H$ nach Schritt (2). Dann beinhaltet S bereits alle fundamentalen Invarianten vom Grad $< \min(D)$. Da der Reynolds-Operator den Grad erhält, wissen wir aus H sofort, wie viele fundamentale Invarianten eines bestimmten Grades es maximal geben kann. Nun werden aufsteigend nach dem Grad die fundamentalen Invarianten höheren Grades gesucht. Sei nun also $d = \min(D)$. Dann gilt $S_{\le d} = \{h \in S : \ \deg(h) \le d\}$ und seien $g_1, \ldots, g_s$ die paarweise verschiedenen Polynome in $S_{\le d}$. Dann erhalten wir in Schritt (9) eine K-Basis des Vektorraums

aller Polynome im Hilbert-Ideal vom Grad d modulo allen Polynomen, die sich als Produkte aus $g_1, \ldots, g_s$ bilden lassen. In

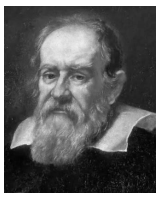
$$(\mathrm{HI}_G)_d / \langle g_1^{x_1} \cdots g_s^{x_s} : \; x_1 \deg(g_1) + \ldots + x_s \deg(g_s) = d \rangle$$

sind auch die fundamentalen Invarianten vom Grad d enthalten, falls es welche gibt. Da aber die Polynome der Basis C' nicht notwendigerweise invariant sind, wird in Schritt (10) eine K-Basis von $P^G \cap \langle C' \rangle_K$ berechnet. Dadurch erhalten wir genau die fundamentalen Invarianten vom Grad d. Diese werden in Schritt (11) S hinzugefügt und dann wird die Prozedur mit dem nächst höheren Grad wiederholt, solange bis D leer ist. Auf diese Weise enthält schließlich S alle fundamentalen Invarianten des Invariantenrings P^G. □

Die Anwendung dieses Algorithmus auf das Beispiel aus Beispiel 7.2.9 bestätigt zunächst die Korrektheit des Algorithmus. Wir werden später noch weitere Anwendungen dieses Algorithmus kennenlernen.

KAPITEL 8

Geometrische Invariantentheorie



Wer die Geometrie begreift, vermag in dieser Welt alles zu verstehen.

Galileo Galilei[21]

Das Zitat, das dem großen italienischen Philosophen, Mathematiker, Physiker und Astronomen Galileo Galilei zugeschrieben wird, ist auch für uns als Aufforderung zu verstehen. Sowohl in der Einleitung als auch in Kapitel 6 wurde bereits deutlich, dass auch in der Invariantentheorie sehr viel Geometrie steckt und viele Anwendungen der Invariantentheorie sich gerade diese geometrischen Eigenschaften zunutze machen. Einen Teil dieser geometrischen Eigenschaften der Invariantentheorie wollen wir in diesem Kapitel näher beleuchten. Wir werden in späteren Anwendungen auch genau auf diese Inhalte zurückgreifen. In unserem Interesse steht dabei besonders die sogenannte Trennungseigenschaft von Invarianten. Die geometrische Invariantentheorie ist zwar auch Teil des Buches [DK02], steht dort allerdings nicht Vordergrund. Wie es der Titel bereits erwarten lässt, enthält das Buch *Geometrische Methoden der Invariantentheorie* von Hanspeter Kraft (siehe [Kra85]) weite Teile der nachfolgenden Eigenschaften.

8.1 Der algebraische Quotient

Will man über die Trennungseigenschaft von Invarianten reden, kommt man besonders bei linear reduktiven Gruppen kaum an dem sogenannten „algebraischen Quotienten" vorbei. Dazu sei im Folgenden K stets ein nicht-endlicher Körper der Charakteristik $\mathrm{char}(K) = 0$, auch wenn die meisten Resultate ihre Gültigkeit auch für beliebige Körper behalten. Weiter sei G eine linear reduktive Gruppe über K und X eine affine Varietät, auf der G regulär operiert. Wie bisher liegt stets die Zariski-Topologie zugrunde, sofern keine andere Topologie angegeben ist. Wir wollen zunächst den Begriff eines algebraischen Quotienten festlegen, der - wie sich herausstellen wird - auf gewisse Weise den Bahnenraum X/G bestmöglich approximiert.

Da G linear reduktiv ist, folgt zunächst aus dem Endlichkeitssatz von Hilbert (vgl. Theorem 6.2.5), dass der Invariantenring $K[X]^G$ eine endlich erzeugte K-Unteralgebra von $K[X]$

[21] Bildquelle: http://de.wikipedia.org/wiki/Galileo_Galilei vom 06.04.2015.

ist. Somit gibt es eine affine Varietät Y und einen Morphismus $\pi : X \to Y$, der die regulären Funktionen auf Y mit den G-invarianten Funktionen auf X identifiziert (vgl. [Kra85], II.3., S. 89). Mit anderen Worten induziert die zu π gehörige Koordinatenabbildung $\pi^* : K[Y] \to K[X]$ einen Isomorphismus $K[Y] \cong K[X]^G$. Ein derartiger Morphismus $\pi : X \to Y$ heißt ein algebraischer Quotient (vgl. [Kra85], II.3.2, S. 95).

Definition 8.1.1. (Algebraischer Quotient)
Sei X eine affine G-Varietät. Ein Morphismus $\pi : X \to Y$ mit einer affinen Varietät Y heißt ein **(algebraischer) Quotient** von X bzgl. G, wenn die zugehörige Koordinatenabbildung $\pi^* : K[Y] \to K[X]$ einen Isomorphismus $K[Y] \cong K[X]^G$ induziert.

Im Folgenden sei nun Y stets eine affine Varietät, sofern nichts anderes angegeben wird. Da, wie oben bereits dargelegt, der HILBERTsche Endlichkeitssatz die Existenz eines algebraischen Quotienten garantiert, lässt sich auf folgende Weise stets ein algebraischer Quotient „konstruieren", womit wir ein für uns bedeutendes Beispiel eines algebraischen Quotienten formulieren können.

Lemma 8.1.2. (Konstruktion algebraischer Quotienten)
Seien $f_1, \ldots, f_r \in K[X]$ fundamentale Invarianten des Invariantenrings $K[X]^G$ und sei der Morphismus $\pi : X \to \mathbb{A}^r_K$ definiert durch

$$\pi(x) = (f_1(x), \ldots, f_r(x)).$$

Setze $Y := \overline{\pi(X)} \subseteq \mathbb{A}^r_K$. Dann ist Y eine affine Varietät und $\pi : X \to Y$ ein algebraischer Quotient von X bzgl. G.

Beweis: Es gilt $K[Y] \cong K[y_1, \ldots, y_r]/\mathcal{I}(Y)$ und $K[X]^G = K[f_1, \ldots, f_r]$. Wir betrachten nun die Abbildung $\varphi : K[y_1, \ldots, y_r] \to K[f_1, \ldots, f_r]$ mit $y_i \mapsto f_i$. Diese Abbildung ist surjektiv und weiter gilt

$$\begin{aligned}\mathrm{Ker}(\varphi) &= \{p \in K[y_1, \ldots, y_r] : \ p(f_1, \ldots, f_r) = 0\} \\ &= \{p \in K[y_1, \ldots, y_r] : \ p(f_1, \ldots, f_r)(y) = 0 \text{ für alle } y \in Y\} = \mathcal{I}(Y).\end{aligned}$$

Somit ist $\overline{\varphi} : K[y_1, \ldots, y_r]/\mathcal{I}(Y) \to K[f_1, \ldots, f_r]$ ein Isomorphismus und es folgt schließlich $K[Y] \cong K[X]^G$, d.h. $\pi : X \to Y$ ist ein algebraischer Quotient. □

Nachdem die Existenz algebraischer Quotienten also in dem hier betrachteten Fall linear reduktiver Gruppen stets gegeben ist, wollen wir uns nun den Eigenschaften algebraischer Quotienten zuwenden. Erste elementare Eigenschaften algebraischer Quotienten ergeben sich dabei sofort (für Surjektivität vgl. [DK02], Lemma 2.3.2, S. 52).

Satz 8.1.3. (Elementare Eigenschaften algebraischer Quotienten)
Sei X eine affine G-Varietät und $\pi : X \to Y$ ein algebraischer Quotient von X bzgl. G. Dann ist π eine surjektive und G-invariante Abbildung, d.h. insbesondere gilt $\pi(a(x)) = \pi(x)$ für alle $a \in G$ und alle $x \in X$.

Beweis: Wir zeigen zuerst die G-Invarianz von π. Angenommen, es gibt ein $a \in G$ und ein $x \in X$ mit $\pi(a(x)) \neq \pi(x)$. Dann gibt es ein $h \in K[Y]$ mit $h(\pi(a(x))) \neq h(\pi(x))$. Mit anderen Worten, es gilt $\pi^*(h)(a(x)) \neq \pi^*(h)(x)$ im Widerspruch dazu, dass $\pi^*(h)$ ein Element des Invariantenrings $K[X]^G$ ist.

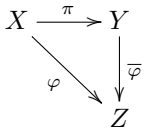
Sei $y \in Y$ und $\mathfrak{m}_y \subseteq K[X]^G$ das Verschwindungsideal von y, also das zu y korrespondierende maximale Ideal. Angenommen, es gilt $\pi^{-1}(y) = \emptyset$. Dann ist $\pi^{-1}(y)$ die Nullstellenmenge von $\mathfrak{m}_y \cdot K[X]$. Somit gilt $1 \in \mathfrak{m}_y \cdot K[X]$ und es gibt endlich viele $f_i \in \mathfrak{m}_y$ und $g_i \in K[X]$ mit $1 = \sum_{i=1}^{r} g_i f_i$. Mit dem Reynolds-Operator bzgl. G in $K[X]$ folgt

$$1 = \mathrm{Rey}_G(1) = \sum_{i=1}^{r} \mathrm{Rey}_G(g_i f_i) = \sum_{i=1}^{r} \mathrm{Rey}_G(g_i) \cdot \mathrm{Rey}_G(f_i) = \sum_{i=1}^{r} \mathrm{Rey}_G(g_i) \cdot f_i,$$

wobei $\mathrm{Rey}_G(g_i) \in K[X]^G$ gilt. Folglich gilt $1 \in \mathfrak{m}_y$, was aber einen Widerspruch bedeutet. Damit gilt $\pi^{-1}(y) \neq \emptyset$, d.h. π ist surjektiv. □

Mit anderen Worten sagt uns die G-Invarianz eines algebraischen Quotienten, dass er entlang der Bahnen konstant ist. Nach den ersten elementaren Eigenschaften algebraischer Quotienten wollen wir die universelle Eigenschaft algebraischer Quotienten angeben (vgl. [Kra85], II.3.2, S. 95).

Satz 8.1.4. (Universelle Eigenschaft algebraischer Quotienten)
Sei X eine affine G-Varietät und sei $\pi : X \twoheadrightarrow Y$ ein algebraischer Quotient von X bzgl. G. Dann gibt es für jeden G-invarianten Morphismus $\varphi : X \to Z$ (mit einer affinen Varietät Z) einen eindeutig bestimmten Morphismus $\overline{\varphi} : Y \to Z$ mit $\varphi = \overline{\varphi} \circ \pi$.



Beweis: Da der Morphismus $\varphi : X \to Z$ G-invariant ist, also konstant entlang den Bahnen, folgt $\varphi^*(K[Z]) \subseteq K[X]^G$. Außerdem induziert die Koordinatenabbildung $\pi^* : K[Y] \to K[X]$ einen Isomorphismus $K[Y] \xrightarrow{\sim} K[X]^G$. Somit kommutiert das nachfolgende Diagramm:

$$\begin{array}{ccc} K[X] & \xleftarrow{\pi^*} & K[Y] \\ {\scriptstyle\varphi^*}\uparrow & \nearrow & \uparrow{\scriptstyle\sim} \\ K[Z] & \xrightarrow{\psi^*} & K[X]^G \end{array}$$

Die Komposition von ψ^* und dem Isomorphismus $K[Y] \xrightarrow{\sim} K[X]^G$ liefert einen eindeutig bestimmten K-Algebra-Homomorphismus $\mu^* : K[Z] \to K[Y]$ mit $\pi^* \circ \mu^* = \varphi^*$. Der zu μ^* gehörende Morphismus $\mu : Y \to Z$ ist laut Satz 3.2.14 ebenfalls eindeutig bestimmt. Mit $\overline{\varphi} := \mu$ folgt die Behauptung. □

Wie aus bisherigen Erkenntnissen bereits leicht ersichtlich ist, gibt es zu einer affinen G-Varietät X viele algebraische Quotienten. Aus der universellen Eigenschaft lässt sich nun aber unmittelbar folgern, dass ein algebraischer Quotient damit bis auf eindeutige bestimmte Isomorphie eindeutig festgelegt ist (vgl. [Kra85], II.3.2, S. 96), was es erlaubt, in gewissem Sinne von *dem* algebraischen Quotienten von X bzgl. G zu sprechen. Unter allen algebraischen Quotienten können wir also einen auswählen, für den die zu X und G resultierende affine Varietät Y mit dem Invariantenring $K[X]^G$ korrespondiert, d.h. für die sogar die Gleichheit $K[Y] = K[X]^G$ gilt, nicht nur Isomorphie (vgl. [DK02], 2.3.1, S. 51). Diese affine Varietät wird mit $X/\!\!/G$ bezeichnet und den zugehörigen algebraischen Quotienten von X bzgl. G notieren wir als

$$\pi_X : X \twoheadrightarrow X/\!\!/G.$$

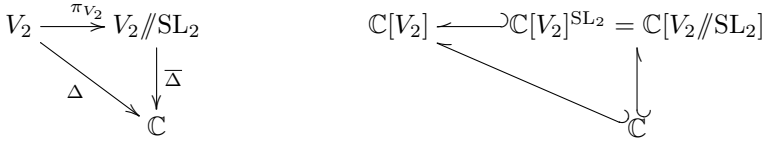
Manchmal wird auch die Varietät $X/\!\!/G$ selbst als **(algebraischer) Quotient** bezeichnet und der Morphismus $\pi_X : X \twoheadrightarrow X/\!\!/G$ die **Quotientenabbildung** genannt. Auf die Begründung für

die Notation $X/\!\!/G$ kommen wir später zurück. Ist nun $\pi : X \to Y$ ein algebraischer Quotient, so sind wegen $K[Y] \cong K[X]^G$ und $K[X/\!\!/G] = K[X^G]$ also $K[Y]$ und $K[X/\!\!/G]$ isomorph. Somit sind auch die Varietäten Y und $X/\!\!/G$ isomorph. Wir werden nun *den* algebraischen Quotienten $\pi_X : X \twoheadrightarrow X/\!\!/G$ von X bzgl. G bzw. die affine Varietät $X/\!\!/G$ weiter untersuchen. Dank der universellen Eigenschaft und der gerade angesprochenen Isomorphie übertragen sich die Eigenschaften kanonisch auf jeden algebraischen Quotienten $\pi : X \to Y$ bzw. die entsprechende Varietät Y. Wir wollen an dieser Stelle ein erstes Beispiel eines algebraischen Quotienten angeben (vgl. auch [Kra85], II.3.3, Beispiel 1, S. 102).

Beispiel 8.1.5. Sei V_2 der $\mathbb{C}$-Vektorraum der Binärformen vom Grad 2 und als Gruppe wählen wir die spezielle lineare Gruppe $\mathrm{SL}_2 := \mathrm{SL}_2(\mathbb{C})$. Aus Beispiel 6.2.4 wissen wir bereits, dass der Invariantenring $\mathbb{C}[V_2]^{\mathrm{SL}_2}$ von der Diskriminante $\Delta : V_2 \to \mathbb{C}$, definiert durch

$$\Delta(a_0x^2 + a_1xy + a_2y^2) = a_1^2 - 4a_0a_2,$$

erzeugt wird. Sei $\varphi : V_2 \to \mathbb{C}$ definiert durch $\varphi(f) = \Delta(f)$. Offensichtlich ist Δ und damit φ surjektiv. Gemäß Lemma 8.1.2 ist damit Δ selbst ein algebraischer Quotient von V_2 bzgl. SL_2. Ist $\pi_{V_2} : V_2 \to V_2/\!\!/\mathrm{SL}_2$ *der* algebraische Quotient, so erhalten wir die kommutativen Diagramme



und schließlich die Isomorphie von $V_2/\!\!/\mathrm{SL}_2$ und $\mathbb{C}$. ◁

Die erste bedeutende Eigenschaft, die wir nun betrachten wollen, ist die sogenannte **G-Abgeschlossenheit**, die mit Hilfe der Surjektivität des algebraischen Quotienten beweisen werden kann (vgl. [Kra85], II.3.2, S. 96, oder [DK02], Korollar 2.3.4, S. 52).

Satz 8.1.6. (G-Abgeschlossenheit)
Sei X eine affine G-Varietät, sei $U \subseteq X$ eine abgeschlossene, G-stabile Teilmenge von X und sei $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G. Dann ist $\pi_X(U) \subseteq X/\!\!/G$ abgeschlossen und der Morphismus $\tilde{\pi} := \pi_X|_U : U \twoheadrightarrow \pi_X(U)$ ein algebraischer Quotient.

Aus diesem Satz lässt sich unmittelbar eine Eigenschaft der Varietät $X/\!\!/G$ folgern. Genauer sagt uns das folgende Korollar, welche Topologie $X/\!\!/G$ trägt (vgl. [Kra85], II.3.2, S. 96). Da wir hier zum ersten Mal mit einer Quotiententopologie zu tun haben, wollen wir diese Aussage explizit beweisen.

Korollar 8.1.7. *Sei X eine affine G-Varietät und sei $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G. Die Topologie von $X/\!\!/G$ ist die durch π_X induzierte Quotiententopologie.*

Beweis: Zu zeigen ist, dass eine Teilmenge $U \subseteq X/\!\!/G$ genau dann offen ist, wenn ihre Urbildmenge $\pi_X^{-1}(U)$ offen ist in X.

Sei also $U \subseteq X/\!\!/G$ und zunächst offen. Dann ist $A := (X/\!\!/G) \setminus U$ abgeschlossen und es folgt sofort, dass $\pi_X^{-1}(A)$ abgeschlossen ist. Wegen $\pi_X^{-1}(U) = X \setminus \pi_X^{-1}(A)$ ist $\pi_X^{-1}(U)$ offen in X. Sei nun $\pi_X^{-1}(U)$ offen und $A := X \setminus \pi_X^{-1}(U)$. Da A abgeschlossen und G-stabil ist, folgt wegen der G-Abgeschlossenheit, dass $\pi_X(A)$ abgeschlossen ist. Wegen $U = (X/\!\!/G) \setminus \pi_X(A)$ ist U offen in $X/\!\!/G$. □

An dieser Stellen wollen wir kurz noch bei den topologischen Eigenschaften des algebraischen Quotienten bleiben, indem wir folgenden Satz betrachten (vgl. [Kra85], II.3.3, Satz 1, S. 100). Damit lassen sich entscheidende topologische Eigenschaften von X auf die affine Varietät $X/\!\!/G$ übertragen.

Satz 8.1.8. *Sei X eine affine G-Varietät und $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G.*

a) Ist X irreduzibel, dann ist auch $X/\!\!/G$ irreduzibel.

b) Ist X normal, dann ist auch $X/\!\!/G$ normal.

8.2 Die Trennungseigenschaft

Eine weitere wichtige, für uns sicherlich *die* wichtigste Eigenschaft algebraischer Quotienten ist die sogenannte **Trennungseigenschaft**. D.h. wir untersuchen, inwiefern ein algebraischer Quotient in der Lage ist, die Bahnen zu trennen. Ein erstens Mal mit Bahnen in Berührung gekommen sind wir zu Beginn dieses Kapitels, als wir gesehen haben, dass algebraische Quotienten entlang der Bahnen konstant sind (siehe Satz 8.1.3). Daraus lässt sich sofort ein notwendiges Kriterium für die Trennung von Bahnen folgern.

Korollar 8.2.1. (Notwendiges Kriterium zur Trennung von Bahnen)
Sei X eine affine G-Varietät, sei $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G und seien $x, x' \in X$ mit $\pi_X(x) \neq \pi_X(x')$. Dann gilt $G(x) \cap G(x') = \emptyset$, insbesondere also $G(x) \neq G(x')$.

Beweis: Angenommen, es gilt $G(x) \cap G(x') \neq \emptyset$. Dann gibt es ein $y \in X$ mit $y \in G(x)$ und $y \in G(x')$, d.h. es gibt ein $a \in G$ mit $y = a(x)$ und ein $a' \in G$ mit $y = a'(x)$. Somit folgt unmittelbar $\pi(a(x)) = \pi(a'(x))$ und mit Satz 8.1.3 sofort $\pi(x) = \pi(x')$ im Widerspruch zur Voraussetzung. □

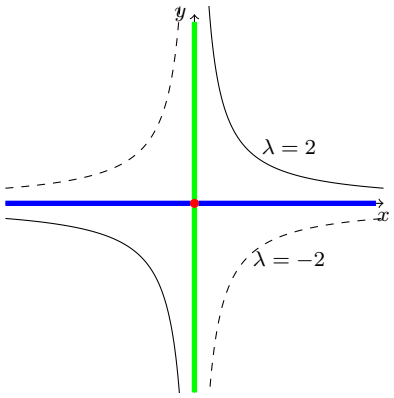
Mit anderen Worten können zwei Punkte, die verschiedene Werte in $X/\!\!/G$ liefern, niemals derselben Bahn angehören. Die Umkehrung ist im Allgemein jedoch falsch, wie uns die folgenden Beispiele zeigen (vgl. [DK02], Beispiel 2.3.1, S. 52 und [DK02], Beispiel 2.3.7, S. 53 bzw. [Kra85], S. 130). D.h. zwei Punkte können durchaus in verschiedenen Bahnen liegen, obwohl sie dieselben Werte unter π_X liefern.

Beispiel 8.2.2. Sei $G = \mathrm{Mult}(K) = K^* \cong \mathrm{GL}_1(K)$ die multiplikative Gruppe und $X = \mathbb{A}^2_K$.

a) Zunächst operiere G durch $a(x, y) = (a \cdot x, a \cdot y)$ für alle $a \in G$ und $(x, y) \in X$ auf X. Dann gilt $K[X]^G = K$, d.h. $X/\!\!/G$ ist ein einziger Punkt. Sei z dieser Punkt. Für $(x, y) \in X \setminus \{0\}$ enthält $G(x, y)$ alle skalaren Vielfachen $a \cdot (x, y)$ mit $a \neq 0$; anschaulich und etwas salopp formuliert also eine Ursprungsgerade ohne dem Ursprung selbst. Für $(x, y) = 0$ gilt $G(0) = \{0\}$. Somit gibt es unendlich viele G-Bahnen. Da aber alle Punkte auf den einzigen Punkt in $X/\!\!/G$ abgebildet werden, trennt π_X hier nicht die Bahnen.

b) Nun operiere G auf X durch $a(x, y) = (a \cdot x, a^{-1} \cdot y)$ für alle $a \in G$ und $(x, y) \in X$.

Dann gilt $K[X]^G = K[xy]$ und $X/\!\!/G \cong K$. Somit ist die Quotientenabbildung $\pi_X : X \to X/\!\!/G \cong K$ definiert durch $(x, y) \mapsto xy$. Für ein $\lambda \in K$ mit $\lambda \neq 0$ ist die Faser $\pi_X^{-1}(\lambda) = \{(x, y) \in X : \ xy = \lambda\}$ selbst eine einzige abgeschlossene Bahn. Die Nullfaser $\pi_X^{-1}(0)$ besteht aus den drei disjunkten Bahnen $\{(0, 0)\}$, $\{(x, 0) : \ x \in K, \ x \neq 0\}$ und $\{(0, y) : \ y \in K, \ y \neq 0\}$. Damit trennt auch hier π_X nicht die Bahnen. Dass für $\lambda \neq 0$ beide Hyperbeläste zu einer Bahn gehören folgt aus der folgenden Tatsache: Sei $\lambda \in K$ mit $\lambda \neq 0$ und $(x, y) \in \pi_V^{-1}(\lambda)$. Für alle $a \in G$ gilt dann



$$a(x, y) = (a \cdot x) \cdot (a^{-1} \cdot y) = (a \cdot a^{-1}) \cdot xy = xy$$

und es folgt $a(x, y) \in \pi_V^{-1}(\lambda)$. ◁

Die Beispiele zeigen, dass ein algebraischer Quotient im Allgemeinen nicht in der Lage ist, die Bahnen selbst zu trennen. Wie wir später sehen werden, kann es aber algebraische Quotienten geben, die dies durchaus vermögen. Algebraische Quotienten besitzen jedoch immer eine Trennungseigenschaft in folgendem Sinne (vgl. [Kra85], II.3.2, S. 96).

Satz 8.2.3. (Trennungseigenschaft)
Sei X eine affine G-Varietät, $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G und sei $(A_i)_{i \in I}$ eine Familie abgeschlossener, G-stabiler Teilmengen von X. Dann gilt:

$$\pi_X \left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} \pi_X(A_i).$$

Insbesondere sind also die Bilder zweier disjunkter, G-stabiler und abgeschlossener Teilmengen von X disjunkt.

Aus der Trennungseigenschaft lassen sich unmittelbar eine Reihe von Eigenschaften herleiten, die wir im Folgenden betrachten wollen. So geht zunächst unter anderem aus dem Beweis zu Korollar 8.1.7 hervor, dass eine Teilmenge von $X/\!\!/G$ genau dann abgeschlossen ist, wenn ihre Urbildmenge unter π_X abgeschlossen in X ist. Somit sind insbesondere die Fasern

$$\pi_X^{-1}(y) = \{x \in X : \ \pi_X(x) = y\}$$

als Urbildmengen der abgeschlossenen einelementigen Mengen $\{y\} \subseteq X/\!\!/G$ abgeschlossen in X. Zudem ist eine Faser $\pi_X^{-1}(y)$ stets eine G-stabile Menge.

Lemma 8.2.4. (Elementare Eigenschaften von Fasern)
Sei X eine affine G-Varietät und sei $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G.

a) Für jedes $y \in X/\!\!/G$ ist die Faser $\pi_X^{-1}(y)$ eine G-stabile Menge.

b) Für jedes $x \in X$ gibt es genau ein $y \in X/\!\!/G$ mit $G(x) \subseteq \pi_X^{-1}(y)$.

Beweis:

a) Sei $y \in X/\!\!/G$ und $a \in G$. Laut Satz 8.1.3 ist π_X eine G-invariante Abbildung, d.h. für alle $x \in X$ gilt $\pi_X(a(x)) = \pi_X(x)$. Somit gilt $a(x) \in \pi_X^{-1}(y)$ für alle $x \in \pi_X^{-1}(y)$, also ist $\pi_X^{-1}(y)$ eine G-stabile Menge.

b) Sei $x \in X$. Da π_X surjektiv ist, gibt es genau ein $y \in X/\!\!/G$ mit $x \in \pi_X^{-1}(y)$. Da $\pi_X^{-1}(y)$ G-stabil ist, folgt sofort $G(x) \subseteq \pi_X^{-1}(y)$.

□

Das Verhältnis zwischen Bahnen und Fasern wollen wir noch etwas genauer beleuchten. Gemäß diesem Lemma ist jede Bahn in genau einer Faser enthalten. Diese kann jedoch durchaus mehrere Bahnen enthalten, wie in Beispiel 8.2.2 schön zu sehen ist. Aus dem nächsten Korollar folgt nun, dass jede Faser stets genau eine *abgeschlossene* Bahn enthält (vgl. [DK02], Korollar 2.3.6, S. 52).

Korollar 8.2.5. *Sei X eine affine G-Varietät und $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G. Für jedes $y \in X/\!\!/G$ enthält die Faser $\pi_X^{-1}(y)$ genau eine abgeschlossene Bahn. Diese Bahn ist enthalten im Zariski-Abschluss aller (anderen) Bahnen in $\pi^{-1}(y)$.*

Mit anderen Worten parametrisiert der algebraische Quotient $X/\!\!/G$ die abgeschlossenen Bahnen in X (vgl. [Kra85], II.3.2, Bemerkung 1, S. 96). Dazu wollen wir auf die bereits bekannten Beispiele aus Beispiel 8.2.2 zurückgreifen.

Beispiel 8.2.6. Sei $G = \mathrm{Mult}(K) = K^* \cong \mathrm{GL}_1(K)$ die multiplikative Gruppe und $X = \mathbb{A}^2_K$.

a) In der Situation von Beispiel 8.2.2 a) besteht $X/\!\!/G$ aus einem einzigen Punkt z. Die eindeutig bestimmte abgeschlossene Bahn in der einzigen Faser $\pi_X^{-1}(z)$ ist die „Nullbahn" $G(0) = \{0\}$. Diese ist im Zariski-Abschluss aller anderen Bahnen enthalten.

b) In der Situation von Beispiel 8.2.2 b) ist für jedes $\lambda \in K$ mit $\lambda \neq 0$ die Faser

$$\pi_X^{-1}(\lambda) = \{(x, y) \in X : \ xy = \lambda\}$$

selbst eine einzige abgeschlossene Bahn. Die Nullfaser $\pi_X^{-1}(0)$ besteht aus den drei disjunkten Bahnen $\{(0,0)\}$, $\{(x,0) : \ x \in K, \ x \neq 0\}$, $\{(0,y) : \ y \in K, \ y \neq 0\}$. Nur die erste Bahn ist als einzelner Punkt abgeschlossen und diese ist enthalten im Abschluss der beiden anderen Bahnen. ◁

Aus dem letzten Korollar bzw. der G-Abgeschlossenheit (siehe Satz 8.1.6) ergeben sich noch eine Reihe weiterer Folgerungen, die in folgendem Korollar zusammengefasst sind (vgl. [Kra85], II.3.3, Satz 3, S. 101).

Korollar 8.2.7. *Sei X eine affine G-Varietät, $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G und sei $x \in X$ sowie $y := \pi_X(x)$.*

a) Der Abschluss $\overline{G(x)}$ der G-Bahn von x enthält genau eine abgeschlossene Bahn.

b) Ist die G-Bahn $G(x)$ abgeschlossen, so gilt

$$\pi_X^{-1}(y) = \{z \in X : \ x \in \overline{G(z)}\}$$

und $G(x)$ ist die einzige abgeschlossene G-Bahn in der Faser $\pi_X^{-1}(y)$.

c) Jede G-stabile abgeschlossene Teilmenge von X, die $G(x)$ als einzige abgeschlossene G-Bahn enthält, ist in der Faser $\pi_X^{-1}(y)$ enthalten.

Die Aussage b) des Korollars ist uns aus einem anderen Zusammenhang bereits bekannt. Der Hilbertsche Nullstellenkegel lässt sich genau auf diese Weise darstellen. Wir erhalten mit diesem Korollar nun eine andere Sichtweise auf den Hilbertschen Nullstellenkegel, nämlich als Faser der 0 unter einem algebraischen Quotienten (vgl. [Kra85], II.3.3, S. 102).

Beispiel 8.2.8. (Nullfaser)
Laut Satz 6.5.2 lässt sich der Hilbertsche Nullstellenkegel $\mathcal{N}_V$ für eine rationale Darstellung $(\rho, V)_G$ von G in einem K-Vektorraum V als die Menge $\mathcal{N}_V = \{v \in V : \ 0 \in \overline{G(v)}\}$ auffassen. Ist $\pi_V : V \twoheadrightarrow V/\!\!/G$ der algebraische Quotient von V bzgl. G, so gilt

$$\mathcal{N}_V = \{v \in V : \ 0 \in \overline{G(v)}\} = \pi_V^{-1}(\pi_V(0)).$$

Man nennt in diesem Zusammenhang den Nullstellenkegel von V auch die **Nullfaser** des algebraischen Quotienten π_V. ◁

Mit Hilfe von Korollar 8.2.5 können wir nun die Trennungseigenschaft aus Satz 8.2.3 präzisieren.

Korollar 8.2.9. *Sei X eine affine G-Varietät und $\pi_X : X \twoheadrightarrow X/\!\!/G$ der algebraische Quotient von X bzgl. G. Dann gilt für alle $x, x' \in X$:*

$$\pi_X(x) = \pi_X(x') \qquad \Longleftrightarrow \qquad \overline{G(x)} \cap \overline{G(x')} \neq \emptyset$$

Beweis: Seien $x, x' \in X$ und zunächst $y \in X/\!\!/G$ mit $y = \pi_X(x) = \pi_X(x')$, d.h. es gilt $x, x' \in \pi_X^{-1}(y)$. Gemäß dem letzten Korollar gibt es genau eine abgeschlossene Bahn in $\pi_X^{-1}(y)$. Diese ist sowohl in $\overline{G(x)}$, als auch in $\overline{G(x')}$ enthalten. Somit folgt $\overline{G(x)} \cap \overline{G(x')} \neq \emptyset$.

Sei nun $B := \overline{G(x)} \cap \overline{G(x')} \neq \emptyset$, sei $z \in B$ und $y := \pi_X(z)$. Aus der Trennungseigenschaft folgt $\pi_X(B) = \pi_X(\overline{G(x)}) \cap \pi_X(\overline{G(x')}) \neq \emptyset$. Da π_X eine G-invariante Abbildung ist, sind $\pi_X(\overline{G(x)})$ und $\pi_X(\overline{G(x')})$ beide einelementig. Das einzige Element in $\pi_X(\overline{G(x)}) \cap \pi_X(\overline{G(x')})$ ist also y. Somit folgt $x, x' \in \pi_X^{-1}(y)$ bzw. $\pi_X(x) = \pi_X(x')$. □

Allerdings ist hier Vorsicht geboten! Denn $\pi_X(x) = \pi_X(x')$ impliziert nicht, dass x und x' ein und derselben Bahn angehören. An Beispiel 8.2.2 ist das sehr schön zu erkennen. An dieser Stelle sind wir nun aber in der Lage, die Notation „$X/\!\!/G$" zu erklären. Diese ergibt sich aus der Tatsache, dass $X/\!\!/G$ isomorph ist zur Menge von Äquivalenzklassen bzgl. der folgenden Äquivalenzrelation.

Lemma 8.2.10. *Sei X eine affine G-Varietät. Dann ist durch*

$$x \sim x' \quad :\Longleftrightarrow \quad \overline{G(x)} \cap \overline{G(x')} \neq \emptyset$$

eine Äquivalenzrelation auf X gegeben. Die Äquivalenzklassen bzgl. $\sim$ entsprechen eineindeutig den Elementen von $X/\!\!/G$.

Beweis: Reflexivität und Symmetrie sind klar. Seien $x_1, x_2, x_3 \in X$ mit $x_1 \sim x_2$ und $x_2 \sim x_3$. Dann gilt $\overline{G(x_1)} \cap \overline{G(x_2)} \neq \emptyset$ und $\overline{G(x_2)} \cap \overline{G(x_3)} \neq \emptyset$. Mit dem letzten Korollar folgt nun sofort $\pi_X(x_1) = \pi_X(x_2)$ und $\pi_X(x_2) = \pi_X(x_3)$. Somit erhalten wir $\pi_X(x_1) = \pi_X(x_3)$, also erneut mit dem letzten Korollar $\overline{G(x_1)} \cap \overline{G(x_3)} \neq \emptyset$ und damit $x_1 \sim x_3$.

Sei M die Menge der Äquivalenzklassen bzgl. $\sim$. Wir betrachten die Abbildung $\varphi : M \to X/\!\!/G$, definiert durch $\alpha \mapsto \pi_X(x)$, wobei $x \in \alpha$ ein Repräsentant von α ist, und zeigen, dass φ wohldefiniert und bijektiv ist. Sei also $\alpha \in M$ und seien $x, x' \in \alpha$, d.h. es gilt $x \sim x'$. Mit dem letzten Korollar folgt $\pi_X(x) = \pi_X(x')$ und damit ist φ unabhängig von der Wahl des Repräsentanten, also wohldefiniert.

Seien nun $\alpha, \beta \in M$ mit $\varphi(\alpha) = \varphi(\beta)$. Sei $x \in \alpha$ und $y \in \beta$. Dann folgt $\pi_X(x) = \pi_X(y)$, also $x \sim y$ und damit $\alpha = \beta$, d.h. φ ist injektiv. Sei weiter $z \in X/\!\!/G$. Da $\pi_X : X \to X/\!\!/G$ surjektiv ist, ist die Faser $\pi_X^{-1}(z)$ nicht leer. Seien weiter $x \in \pi_X^{-1}(z)$ und $\alpha = [x]_\sim$. Dann gilt $\varphi(\alpha) = z$, d.h. φ ist surjektiv. □

Dass die Elemente von $X/\!\!/G$ genau den Äquivalenzklassen bzgl. $\sim$ entsprechen, erklärt auch die Notation dieser Menge, die sich aber dennoch in Sachen Notation vom Bahnenraum X/G unterscheiden muss, weshalb die Menge mit Doppelstrich notiert wird. Ist jede G-Bahn allerdings abgeschlossen, d.h. gilt $\overline{G(x)} = G(x)$ für alle $x \in X$, so stimmt die Äquivalenzrelation aus Bemerkung 4.2.8 mit dieser hier überein (vgl. auch Bemerkung 4.2.9). Und genau hier können wir weiter anknüpfen. Denn sind alle Bahnen abgeschlossen, so enthält jede Faser genau eine Bahn, d.h. die Bahnen stimmen genau mit den Fasern überein. In diesem Fall nennt man einen algebraischen Quotienten auch **geometrisch** (vgl. [Kra85], II.3.2, S. 96).

Definition 8.2.11. (Geometrischer Quotient)
Sei X eine affine G-Varietät. Der algebraische Quotient $\pi_X : X \twoheadrightarrow X/\!\!/G$ heißt **geometrisch** oder der **geometrische Quotient** von X bzgl. G, wenn jede Faser von π_X eine G-Bahn ist.

Ist $\pi_X : X \twoheadrightarrow X/\!\!/G$ ein geometrischer Quotient, so entsprechen die G-Bahnen in X also eineindeutig den Punkten in $X/\!\!/G$, d.h. in diesem Fall gilt $X/G = X/\!\!/G$. Für endliche Gruppen existiert stets ein geometrischer Quotient (vgl. [Kra85], II.3.6, Satz 1, S. 111).

Satz 8.2.12. (Existenz geometrischer Quotienten bei endlichen Gruppen)
Sei G endlich und X eine affine G-Varietät. Dann ist der Quotient $\pi_X : X \twoheadrightarrow X/\!\!/G$ geometrisch und π_X ein endlicher Morphismus.

Für nicht-endliche Gruppen ist die Existenz eines geometrischen Quotienten im Allgemeinen nicht gegeben, wie auch die Beispiele in Beispiel 8.2.2 bzw. Beispiel 8.2.6 zeigen. Allerdings „enthält" Teil b) von Beispiel 8.2.2 einen geometrischen Quotienten, was wir an dieser Stelle als Beispiel angeben wollen.

Beispiel 8.2.13. Sei wie in Beispiel 8.2.2 $G = \mathrm{Mult}(K)$ die multiplikative Gruppe, die auf $X = \mathbb{A}^2_K$ durch $a(x, y) = (a \cdot x, a^{-1} \cdot y)$ für alle $a \in G$ und alle $(x, y) \in X$ operiere. Mit Ausnahme der Nullfaser $\pi_X^{-1}(0)$ sind alle Fasern G-Bahnen, d.h. $\pi_X : X \to X/\!\!/G$ ist kein geometrischer Quotient. Allerdings ist der Quotient $\pi : X \setminus \pi_X^{-1}(0) \to K \setminus \{0\}$ geometrisch. ◁

Geometrische Quotienten lassen sich auch auf folgende Weise charakterisieren.

Satz 8.2.14. (Charakterisierung geometrischer Quotienten durch Bahnen)
Sei X eine affine G-Varietät. Der algebraische Quotient $\pi_X : X \twoheadrightarrow X/\!\!/G$ ist genau dann geometrisch, wenn jede G-Bahn abgeschlossen ist.

Beweis: Sei zunächst π_X geometrisch, d.h. jede Faser von π_X ist eine G-Bahn. Sei $x \in X$. Gemäß Lemma 8.2.4 gibt es dann genau ein $y \in X/\!\!/G$ mit $G(x) = \pi_X^{-1}(y)$, also ist $G(x)$ insbesondere abgeschlossen.

Sei nun jede Bahn abgeschlossen und sei $y \in X/\!\!/G$. Laut Korollar 8.2.5 enthält $\pi_X^{-1}(y)$ genau eine abgeschlossene Bahn und diese ist im Zariski-Abschluss aller anderen Bahnen in $\pi_X^{-1}(y)$ enthalten. Da aber alle Bahnen abgeschlossen sind, enthält $\pi_X^{-1}(y)$ genau eine G-Bahn. Da $\pi_X^{-1}(y)$ eine G-stabile Menge (vgl. Lemma 8.2.4) ist, ist diese G-Bahn mit der Faser identisch. Somit ist jede Faser eine G-Bahn, π_X also geometrisch. □

In Korollar 8.2.1 hatten wir gesehen, dass $\pi_X(x) = \pi_X(x')$ aus $G(x) \cap G(x') \neq \emptyset$ und damit aus $G(x) = G(x')$ folgt und dass die Umkehrung im Allgemeinen falsch ist. Ist der algebraische Quotient geometrisch, gilt auch die Umkehrung, und wir erhalten folgenden Satz.

Satz 8.2.15. (Trennungseigenschaft geometrischer Quotienten)
Sei $\pi_X : X \to X/\!\!/G$ der algebraische Quotient von X bzgl. G und seien $x, x' \in X$. Ist π_X geometrisch, so gilt:

$$\pi_X(x) = \pi_X(x') \qquad \Longleftrightarrow \qquad G(x) = G(x')$$

Beweis: Die Rückrichtung folgt allgemein aus Korollar 8.2.1. Sei also π_X geometrisch und gelte $\pi_X(x) = \pi_X(x') =: y \in X/\!\!/G$. Da nun alle Bahnen abgeschlossen sind, folgt aus Korollar 8.2.9 sofort $G(x) \cap G(x') \neq \emptyset$ und damit auch $G(x) = G(x')$. □

Ist $\pi_X : X \to X/\!\!/G$ ein geometrischer Quotient und X zusammenhängend, so gilt

$$\dim(\pi_X^{-1}(\pi_X(x)) = \dim(G(x))$$

für alle $x \in X$ (vgl. [DK02], S. 53) und alle G-Bahnen in X haben dieselbe Dimension (vgl. [Kra85], II.3.2, Bem. 2, S. 97). Hier gilt auch die andere Richtung, d.h. wenn alle G-Bahnen von X dieselbe Dimension haben, liegt keine Bahn im Abschluss einer anderen Bahn, womit π_X die Bahnen trennt (vgl. [DK02], S. 54). Mit anderen Worten ist der algebraische Quotient π_X genau dann geometrisch.

Zum Ende dieses Abschnitts wollen wir noch eine spezielle Situation kurz beleuchten. Wir betrachten im Folgenden als affine G-Varietät einen endlich-dimensionalen K-Vektorraum V und eine rationale Darstellung von G in V. Zunächst geben wir eine Aussage über die Struktur von $V/\!\!/G$ in einem Spezialfall an (vgl. [Kra85], II.3.3, Satz 4, S. 103).

Satz 8.2.16. *Sei V ein endlich-dimensionaler K-Vektorraum, auf dem G linear operiert. Gibt es eine Bahn mit Kodimension ≤ 1, so ist $V/\!\!/G$ entweder ein Punkt oder isomorph zu K.*

Will man den algebraischen Quotienten $\pi_V : V \twoheadrightarrow V/\!\!/G$ genauer verstehen und studieren, ist es unerlässlich, die Bahnen der Gruppe G in V zu untersuchen, was aber im Allgemeinen nicht einfach ist. Ein lohnenswerter Ansatzpunkt ist die Nullfaser. Denn in [Kra85], II.4.2, S. 129 heißt es: „Es wird sich zeigen, daß die Nullfaser in gewissem Sinne die 'schlechteste' aller Fasern ist, oder umgekehrt, daß alle 'guten' Eigenschaften der Nullfaser auch allen anderen Fasern zukommen." Aus diesem Grund wollen wir nun die Nullfaser genauer beleuchten. Zunächst können wir aus der Anzahl der Bahnen in der Nullfaser auf die Anzahl der Bahnen in allen Fasern schließen (vgl. [Kra85], II.4.2, S. 129).

Satz 8.2.17. *Sei $\pi_V : V \twoheadrightarrow V/\!\!/G$ der algebraische Quotient von V in G. Enthält die Nullfaser $\mathcal{N}_V = \pi_V^{-1}(\pi_V(0))$ nur endlich viele Bahnen, so gilt dies auch für jede Faser von π_V. Zudem haben alle irreduziblen Komponenten aller Fasern von π_V dieselbe Dimension.*

Wir wollen es nun an dieser Stelle dabei belassen und verweisen für ein vertieftes Studium der Nullfaser auf [Kra85], insbesondere auf die Abschnitte II.4.2, II.4.3 und III.2. Bislang haben sind wir stets davon ausgegangen, dass ein endliches Erzeugendensystem des Invariantenrings existiert und bekannt ist. Wir wollen nun noch kurz ansprechen, ob es selbst im Falle der Nichtexistenz eines endlichen Erzeugendensystems eine Menge von Invarianten gibt, die ähnliche Trennungseigenschaften haben wie die Erzeuger. Dazu werden wir zunächst etwas allgemeiner den Begriff einer separierenden Menge angeben (nach [Kem09], Definition 1.1).

Definition 8.2.18. (R-separierend)
Sei X eine affine Varietät und $R \subseteq K[X]$. Eine Teilmenge $S \subseteq R$ heißt **R-separierend**, wenn für alle $x, y \in X$ gilt:

$$\exists f \in R : \ f(x) \neq f(y) \quad \Longrightarrow \quad \exists g \in S : \ g(x) \neq g(y).$$

Eine R-separierende Teilmenge S von R ist also eine Menge, die die Punkte von X genauso gut trennt, wie R selbst. Betrachten wir speziell die K-Unteralgebra $K[X]^G$ als Menge R, so nennt man die Elemente einer $K[X]^G$-separierenden Menge $S \subseteq K[X]^G$ entsprechend auch **separierende Invarianten**. Klar ist, dass eine Menge $S \subseteq K[X]^G$ genau dann $K[X]^G$-separierend ist, wenn die von S erzeugte K-Unteralgebra $K[S]$ von $K[X]^G$ ebenfalls $K[X]^G$-separierend ist (vgl. [DK02], S. 55). Damit ist insbesondere jedes Erzeugendensystem von $K[X]^G$ eine $K[X]^G$-separierende Menge (vgl. [Kem03]). Wie bereits an verschiedenen Stellen angedeutet und verwendet wurde, lässt sich mit dem Invariantenring $K[X]^G$ durch

$$x \sim_{K[X]^G} y \quad :\Longleftrightarrow \quad \forall f \in K[X]^G : \ f(x) = f(y)$$

eine Äquivalenzrelation auf X definieren. Dafür reicht es bereits, nur die Erzeuger von $K[X]^G$ zu betrachten. Eine Teilmenge $S \subseteq K[X]^G$ ist nun genau dann $K[X]^G$-separierend, wenn die durch S induzierte Äquivalenzrelation

$$x \sim_S y \quad :\Longleftrightarrow \quad \forall f \in S : \ f(x) = f(y)$$

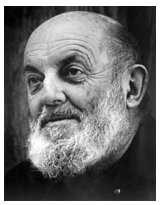
auf X mit die obigen Äquivalenzrelation $\sim_{K[X]^G}$ übereinstimmt (vgl. [Kem09]). Für die Berechnung separierender Invarianten gibt es einen effizienten Algorithmus, auf den wir aber nicht näher eingehen wollen. Dazu sei nur auf [Kem03] verwiesen.

Teil III

Bildverarbeitung

KAPITEL 9

Von Bildern zu Polynomen



Ansel Adams[22]

Ein Foto wird meistens nur angeschaut – selten schaut man in es hinein.

Das Zitat eines der bedeutendsten US-amerikanischen Fotografen, Ansel Adams (1902–1984), klingt beinahe wie eine Aufforderung für uns. Aus Sicht eines Mathematikers lässt es sich jedoch gar nicht vermeiden, *in* ein Bild hinein zu schauen. Genau das wollen wir in diesem Kapitel tun; wir wollen hier die grundlegenden Objekte für die späteren Anwendungen erarbeiten. Dazu werden wir uns mit dem Entstehungsprozess von Bildern auseinandersetzen, um verstehen zu können, was unter sogenannten lokalen Bildmerkmalen zu verstehen ist.

9.1 Der Bildentstehungsprozess

Um mit Bildern Mathematik betreiben zu können, ist zunächst etwas Vorarbeit zu leisten. Wir werden uns in diesem Abschnitt deshalb mit dem Bildentstehungsprozess in vereinfachter Form auseinandersetzen. Dazu betrachten wir zu Beginn das für unsere Zwecke ausreichende, einfache Modell einer Lochkamera und werden uns anschließend mit der mathematischen Modellierung von Bildern beschäftigen, wobei wir ausschließlich Grauwertbilder, und damit keine Farbbilder, modellieren werden. Außerdem nehmen wir hier stets an, dass Bilder durch Aufnahme mit einer Kamera entstanden sind, und schließen andere Möglichkeiten der Bildentstehung, wie z.B. durch Röntgen oder Scannen, aus.

9.1.1 Das Lochkameramodell

Die wesentlichen Elemente einer modernen, handelsüblichen Kamera sind Objektiv und Bildsensor (vgl. [Han10]). Der Bildsensor, der sich in der sogenannten **Bildebene** befindet, sammelt dabei die einfallenden Photonen und wandelt die Energie in elektrische Spannung um. Die heute

[22] Bildquelle: http://www.anseladams.com/ vom 12.09.2014.

gängigen Aufnahmesensoren sind entweder sogenannte **CCD-Sensoren** (CCD = Charge Coupled Device) oder **CMOS-Sensoren** (CMOS = Complementary Metal Oxide Semiconductor). Beide Sensoren bestehen aus lichtempfindlichen Elementen, auch **Pixel**[23] genannt, die in einer Matrix oder in einer Zeile angeordnet sind, wobei die Anordnung in einer Matrix die häufigste Form ist. Es handelt sich in diesem Fall bei einem Bildsensor also genauer um eine matrixartig angeordnete Ansammlung vieler Einzelsensoren, weshalb man in diesem Fall auch von einem **Sensorarray** spricht. Auch wenn in der Darstellung die Pixel als quadratisch skizziert sind, muss dies in der Praxis nicht immer der Fall sein; sie sind auch nicht notwendigerweise rechteckig.

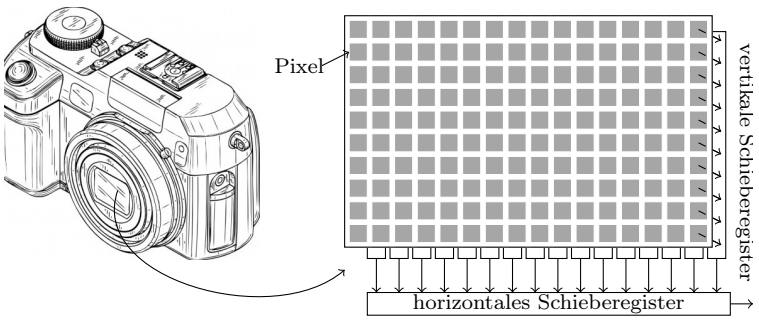


Abbildung 9.1: Schematische Darstellung eines CCD-Sensorchips

Die einzelnen Pixel sammeln nun für einen gewissen Belichtungszeitraum die einfallenden Photonen und wandeln deren Energie in Ladung um. Die Ladungen der einzelnen Pixel werden durch vertikale und horizontale Schieberegister in einen Bereich der Kamera transportiert, wo die Ladung in Spannung umgewandelt wird. Nähere Informationen zu Bildsensoren findet man in [Han10], Abschnitt 2.3, [FP03], Abschnitt 1.4.1, oder [Jäh05], Abschnitt 1.7. Wie aus der erzeugten Spannung schließlich der Grauwert eines Pixels wird, werden wir später betrachten.

Das Objektiv besteht aus einem ganzen System von Linsen, wir nehmen vereinfacht jedoch nur eine Linse an, die wir zudem als radial symmetrisch und infinitesimal „dünn" voraussetzen wollen (vgl. [Han10], S. 11 oder [BB05], S. 8). Die Realität spiegelt diese einfache Annahme natürlich nicht wider, aber als Modell ist dies für uns ausreichend. Eine Darstellung eines etwas realistischeren optischen Systems findet man z.B. in [Jäh05], S. 205 ff., oder in [FP03], worauf wir nicht näher eingehen wollen. Die Rotationsachse der Linse wird auch **optische Achse** genannt. Wir betrachten hier nur die sogenannte Optik erster Ordnung (vgl. [Han10], S. 14), d.h. wir nehmen an, dass alle Lichtstrahlen nahezu parallel zur optischen Achse auf die Linse treffen. Die Ebene, die vertikal durch die Linse verläuft, nennt man die **Hauptebene** und den Schnittpunkt der optischen Achse mit der Hauptebene das **optische Zentrum** oder das **Kamerazentrum** (vgl. [HZ06], S. 154). Häufig wird das optische Zentrum auch als Hauptpunkt bezeichnet (vgl. [Jäh05], S. 206), wir aber werden mit diesem Begriff später einen anderen Punkt verbinden. Das aufzunehmende Objekt soll vereinfacht ebenfalls als planar angenommen werden. Die Ebene, die durch das Objekt geht und parallel zur Bildebene verläuft, wird die **Objektebene** genannt.

Die Größe $f > 0$ ist die **Brennweite** (engl.: *focal length*) der Linse. Die Ebenen, die parallel zur Hauptebene im Abstand f verlaufen, nennt man auch die **Fokalebenen**. Vereinfacht ausgedrückt, erscheint ein Bild scharf, mit anderen Worten *fokussiert*, wenn die Fokalebene bis auf eine gewisse Toleranz mit der Bildebene zusammenfällt. Der Strahl, der unter der Annahme

[23] Das Kunstwort „Pixel" ist eine Neologismus, der entstanden ist aus der Kombination der Wörter „Picture" und „Element".

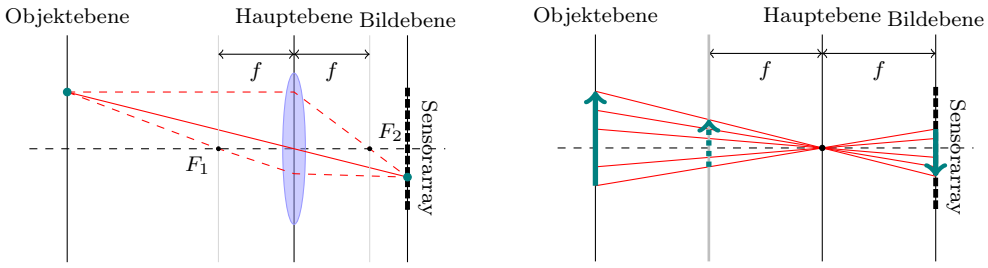


Abbildung 9.2: Vereinfachte Darstellung der Optik erster Ordnung unter der Annahme einer „dünnen" Linse und des Lochkameramodells.

einer infinitesimal „dünnen" Linse ungebrochen durch das optische Zentrum verläuft, wird auch der **Mittelpunktstrahl** genannt. Betrachtet man nur die Abbildungseigenschaft dieses Mittelpunktstrahls, so spricht man von dem **Lochkameramodell**, mit dem sich die Eigenschaften realer Objektive approximieren lassen (vgl. [Han10], Abschnitt 2.2.4).Durch die Zentralprojektion am optischen Zentrum wird ein Objekt im Lochkameramodell in der Bildebene auf dem Kopf stehend abgebildet. Dasselbe Bild lässt sich allerdings auch in einer Ebene zwischen Objekt- und Hauptebene abgreifen, die ebenfalls als **Bildebene** bezeichnet wird (vgl. [HZ06], S. 153), was zur folgenden Festlegung eines kartesischen **Kamerakoordinatensystems** führt (vgl. [Han10], Abschnitt 2.4):

- Das optische Zentrum bildet den Ursprung des Koordinatensystems.
- Die x-Achse verläuft parallel zur x-Achse des Sensorarrays in der Hauptebene, die y-Achse verläuft orthogonal zur x-Achse ebenfalls in der Hauptebene.
- Die z-Achse steht senkrecht auf der Hauptebene, d.h. die z-Achse fällt mit der optischen Achse zusammen.

Die Bildebene ist dann die Ebene $\{(x, y, z) \in \mathbb{R}^3 : \ z = f\}$, jedoch wird üblicherweise die Bildebene mit der Ebene

$$E := \{(x, y, z) \in \mathbb{R}^3 : \ z = 1\}$$

identifiziert, was es ermöglicht, Sätze der projektiven Geometrie anzuwenden (vgl. [Han10], S. 23). Mit anderen Worten entspricht eine Längeneinheit auf der z-Achse der Brennweite f. Da die Maßstäbe auf allen Achsen gleich bleiben sollen, entspricht also auch auf der x- und y-Achse eine Längeneinheit der Brennweite f.

Nun sind Objekte bzw. Punkte auf den Objekten üblicherweise nicht im Kamerakoordinatensystem gegeben, sondern in einem „übergeordneten", sogenannten **Weltkoordinatensystem**. Ein Punkt $P \in \mathbb{R}^3$ im Weltkoordinatensystem kann durch eine (eigentliche) Bewegung $(\mathcal{A}, v) \in \mathrm{Iso}_3^+(\mathbb{R})$ stets in einen Punkt des Kamerakoordinatensystems abgebildet werden. Mit anderen Worten gibt es stets eine Bewegung $(\mathcal{A}, v) \in \mathrm{Iso}_3^+(\mathbb{R})$, die den Koordinatensystemwechsel vom Weltkoordinatensystem ins Kamerakoordinatensystem durchführt. Diese Bewegung $(\mathcal{A}, v)$ bzw. die spezielle orthogonale Matrix $\mathcal{A} \in \mathrm{SO}_3(\mathbb{R})$ zusammen mit dem Translationsvektor $v \in \mathbb{R}^3$ bezeichnet man auch als **extrinsische Kameraparameter** (vgl. [FP03], S. 30). Sei $\epsilon : \mathbb{R}^3 \to \mathbb{R}^3$ definiert durch

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \mathcal{A} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + v.$$

Dann ist ϵ die Abbildung, die den Wechsel vom Welt- in das Kamerakoordinatensystem vollzieht. Wie man sieht, ist sie durch die extrinsischen Kameraparameter festgelegt. Indem wir die

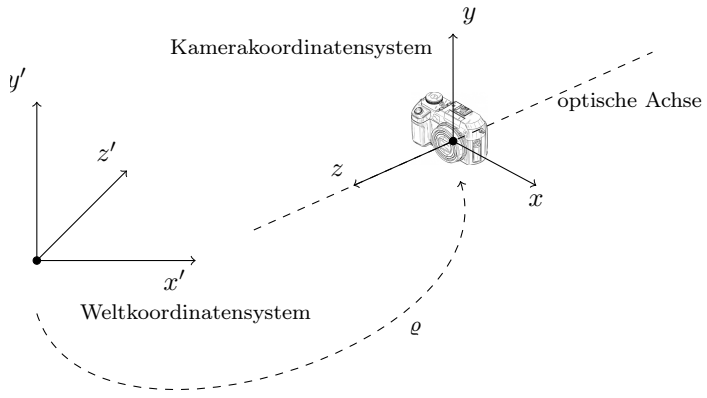


Abbildung 9.3: Übergang vom Welt- ins Kamerakoordinatensystem

Vektoren $(x, y, z) \in \mathbb{R}^3$ mit homogenen Koordinaten $(x, y, z, 1) \in \mathbb{R}^4$ identifizieren, können wir ϵ auch durch die 4×4-Matrix $\mathcal{T}_\epsilon := \left(\begin{smallmatrix} \mathcal{A} & v \\ 0 & 1 \end{smallmatrix}\right)$ repräsentieren. Wir wollen von nun an stets davon ausgehen, dass Objektpunkte im Kamerakoordinatensystem vorliegen. Im Bildaufnahmeprozess wird die optische Achse auch gerne **Hauptachse** genannt. Der Schnittpunkt $H = (0, 0, 1)$ der Hauptachse mit der Bildebene wird der **Hauptpunkt** genannt und jeder Strahl vom Ursprung zu einem Objektpunkt P heißt **Hauptstrahl** (vgl. [HZ06], S. 154).

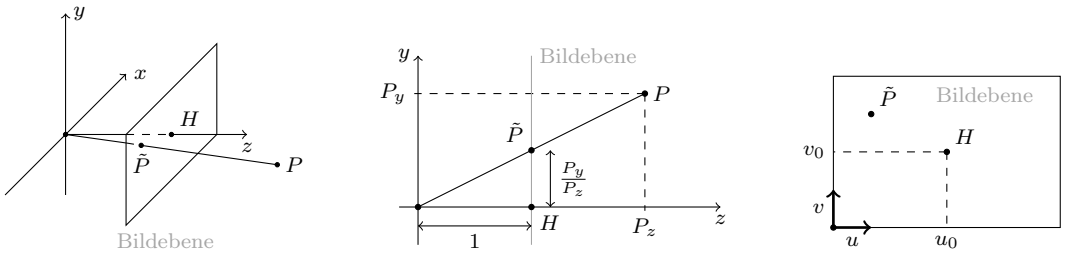


Abbildung 9.4: Schematische Darstellung der Transformation vom Kamera- ins Bildkoordinatensystem.

Ein Punkt $P \in \mathbb{R}^3$ im dreidimensionalen Kamerakoordinatensystem, der nicht in der x-y-Ebene liegt, wird durch die Abbildung $\mathrm{pr} : \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3 : \ z = 0\} \to \mathbb{R}^3$, definiert durch

$$(x, y, z) \mapsto \left(\frac{x}{z}, \frac{y}{z}, 1\right),$$

auf den Punkt $\tilde{P}$ in der Bildebene abgebildet. Dank der bijektiven Abbildung $\pi : E \to \mathbb{R}^2$, definiert durch $\pi(x_1, x_2, x_3) = (x_1, x_2)$, können wir $\tilde{P}$ auch als zweidimensionalen Punkt identifizieren. Dieser zweidimensionale Punkt wird nun im sogenannten **Bildkoordinatensystem** dargestellt. Als Ursprung des zweidimensionalen Bildkoordinatensystems wäre natürlich der Hauptpunkt H mit den Koordinaten $(0, 0, 1)$ im Kamerakoordinatensystem die kanonische Wahl, da er durch π auf $(0, 0)$ abgebildet wird. In der Praxis wird der Ursprung des Bildkoordinatensystems allerdings mit einer Ecke des Bildes identifiziert. In der nebenstehenden Abbildung ist dies die Ecke links unten, häufig wird aber auch der Ursprung mit der Ecke links oben identifiziert (vgl. [BB05], S. 12), da in der Praxis die Pixel eines Bildes entsprechend indiziert sind. Alle Punkte müssen im Bildaufnahmeprozess also nun noch um einen festen Vektor $(u_0, v_0) \in \mathbb{R}^2$ verschoben werden. Die Achsen des Bildkoordinatensystems sind parallel zu den Achsen des Sensorarrays. Wir wollen hier davon ausgehen, dass die Achsen des Sensorarrays orthogonal sind, was aber in der Praxis nicht immer der Fall sein muss. Modellieren

lässt sich das über einen Parameter γ, der durch $\gamma = \frac{f \cot(\theta)}{d_v}$ mit dem Winkel θ zwischen den Koordinatenachsen zusammenhängt (vgl. [FP03], S. 29). Für $\gamma = 0$ sind die Koordinatenachsen orthogonal. Als Längeneinheiten in der Bildebene werden üblicherweise die Ausmaße der photosensitiven Elemente, also der Einzelsensoren des Sensorarrays, verwendet. Diese sind im Allgemeinen jedoch nicht quadratisch. Sind d_u und d_v die Länge und die Breite eines photosensitiven Elements, dann entspricht eine Längeneinheit im Kamerakoordinatensystem $\alpha_u := \frac{f}{d_u}$ bzw. $\alpha_v := \frac{f}{d_v}$ Längeneinheiten im Bildkoordinatensystem (vgl. [Han10], S. 24 bzw. [FP03], S. 29). Die Abbildung $\iota : E \to \mathbb{R}^2$, definiert durch

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \mapsto \mathcal{K} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \quad \text{mit} \quad \mathcal{K} = \begin{pmatrix} \alpha_u & \gamma \\ 0 & \alpha_v \end{pmatrix},$$

heißt die **Kamerakoordinatensystemtransformation** im Lochkameramodell (vgl. [Han10], S. 23f). Die Parameter $f, \gamma, \alpha_u, \alpha_v, u_0, v_0$ werden auch die **intrinsischen Kameraparameter** genannt (vgl. [FP03], S. 29 f.). Damit lässt sich das Lochkameramodell mathematisch wie folgt definieren (vgl. [Han10], Definition 2.2.6, S. 24).

Definition 9.1.1. (Lochkameramodell)
Seien $(\mathcal{A}, v) \in \mathrm{Iso}_3(\mathbb{R})$ die extrinsischen Kameraparameter und $f, \gamma, \alpha_u, \alpha_v, u_0, v_0 \in \mathbb{R}$ die intrinsischen Kameraparameter einer Kamera. Sei $\epsilon : \mathbb{R}^3 \to \mathbb{R}^3$ die durch $(\mathcal{A}, v)$ definierte Bewegung, die den Koordinatenwechsel vom Welt- ins Kamerakoordinatensystem vollzieht, und sei $\iota : E \to \mathbb{R}^2$ die durch die intrinsischen Kameraparameter definierte Kamerakoordinatensystemtransformation. Weiter sei $D := \mathbb{R}^3 \setminus \epsilon^{-1}(\{(x, y, z) \in \mathbb{R}^3 : \ z = 0\}$. Dann heißt die Abbildung $\kappa : D \to \mathbb{R}^2$ mit

$$\kappa = \iota \circ \mathrm{pr} \circ \epsilon$$

die **Kameraabbildung bzgl. des Lochkameramodells**.

Die Kameraabbildung κ bildet also in der vereinfachten Darstellung des Lochkameramodells einen von einem Punkt der realen Welt reflektierten Lichtstrahl auf einen Punkt in der Bildebene ab, d.h. diese Abbildung gibt an, wo ein Lichtstrahl, der von einem Objektpunkt reflektiert wird, auf die Bildebene trifft. Obwohl das Lochkameramodell die Realität nicht exakt modelliert, wollen wir es dabei belassen, da dieses Modell für unsere Zwecke ausreichend ist. Wir wollen außerdem auch darauf verzichten, sogenannte **Verzeichnungen** zu modellieren oder auf die **Schärfentiefe** einzugehen. Weitere Informationen zu Kameramodellen finden sich in [Han10], Abschnitt 2.2, in [HZ06], Kapitel 6 oder in [Jäh05], Kapitel 7.

9.1.2 Mathematische Modellierung von Bildern

Wir wissen nun, wie der von einem Punkt der realen Welt reflektierte Lichtstrahl im Lochkameramodell auf die Bildebene und damit auf das Sensorarray trifft. Allerdings ist damit noch nicht klar, wie ein Grauwertbild entsteht und wie sich Bilder mathematisch modellieren lassen. Um diese Modellierung zu leisten, muss man sich zwangsläufig mit der Modellierung von Sensoren auseinandersetzen, was wir hier in prägnanter Form tun wollen. Dabei werden wir uns nicht vordergründig mit den technischen Details auseinandersetzen, sondern uns vorrangig für das mathematische Modell interessieren. Grundlegende Informationen zur technischen Arbeitsweise von Sensoren findet man z.B. in [Han10] oder [Jäh05].

Das Sensorarray der Kamera befindet sich bekanntlich in der Bildebene $E \subseteq \mathbb{R}^3$. Wir werden hier vereinfachend annehmen, dass die Einzelsensoren nahtlos aneinander grenzen und orthogonal sind. In der Realität ist wie in Abbildung 9.1 ein kleiner Offset zwischen den einzelnen

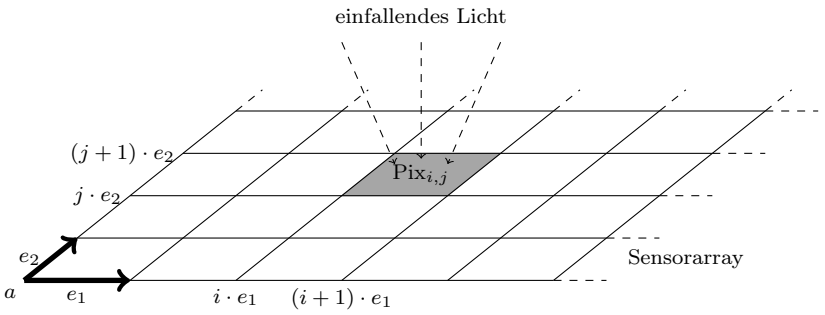


Abbildung 9.5: Schematische Darstellung der mathematischen Modellierung von Pixeln.

photosensitiven Elementen zu erwarten, den wir jedoch vernachlässigen wollen. Wie in Abbildung 9.5 dargestellt, können wir die Bildebene im Kamerakoordinatensystem mit einem Vektor $a \in \mathbb{R}^3$ wie folgt modellieren:

$$E := a + \langle\{e_1, e_2\}\rangle_{\mathbb{R}},$$

wobei $e_1, e_2 \in \mathbb{R}^3$ ein Orthogonalsystem bilden. Die Längen $\|e_1\|$ sowie $\|e_2\|$ definieren die Ausmaße eines photosensitiven Elements, d.h. $\|e_1\|$ und $\|e_2\|$ entsprechen den Größen d_u bzw. d_v aus dem letzten Abschnitt. Sie werden auch **Pixelbreite** und **Pixelhöhe** genannt. Den von e_1 und e_2 aufgespannten $\mathbb{R}$-Untervektorraum $\langle\{e_1, e_2\}\rangle_{\mathbb{R}}$ von $\mathbb{R}^3$ bezeichnen wir mit U. Wir wollen hier also insbesondere annehmen, dass alle Einzelsensoren gleiche Ausmaße haben. Da zudem jeder Einzelsensor dieselbe Energieform, nämlich die Lichtleistung, misst, spricht man bei einem derartigen Sensorarray auch von einem **homogenen Multisensorensystem** (vgl. [Don09], S. 6). Sei durch $r, s \in \mathbb{N}_+$ das Ausmaß des Sensorarrays festgelegt, d.h. es gibt in einer Zeile des Sensorarrays genau r und in einer Spalte des Sensorarrays genau s Einzelsensoren. Die einzelnen Sensoren lassen sich also durch Paare aus $\mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ identifizieren. Die Menge der Sensoren werden wir im Folgenden mit S bezeichnen, d.h. es gilt stets $S = \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$. Damit können wir wie folgt ein Pixel definieren.

Definition 9.1.2. (Pixel)
Sei $a \in \mathbb{R}^3$, sei (e_1, e_2) ein Orthogonalsystem in $\mathbb{R}^3$, sei $E = a + \langle\{e_1, e_2\}\rangle_{\mathbb{R}}$ die Bildebene und sei durch $r, s \in \mathbb{N}$ das Ausmaß des Sensorarrays gegeben. Für $(i, j) \in S$ heißt die kompakte Teilmenge

$$\mathrm{Pix}_{i,j} := \{a + \beta_1 e_1 + \beta_2 e_2 : \ \beta_1 \in [i, i+1],\ \beta_2 \in [j, j+1]\}$$

von E das **(Hardware-)Pixel** mit **Pixelindex** (i, j).

Ein Pixel entspricht also in dieser vereinfachten Modellierung dem lichtempfindlichen Teil des zugehörigen Sensors. Die kompakte Teilmenge

$$B := \{a + \beta_1 e_1 + \beta_2 e_2 : \ \beta_1 \in [0, r],\ \beta_2 \in [0, s]\} = \bigcup_{(i,j)\in S} \mathrm{Pix}_{i,j}$$

von E wird auch als **Bildrechteck** bezeichnet (vgl. [Don09], S. 14). Sei nun $\varrho : \mathbb{R}^2 \to E$ definiert durch

$$(x_1, x_2) \mapsto a + x_1 \cdot e_1 + x_2 \cdot e_2. \tag{9.1.1}$$

Dann ist ϱ ein Isomorphismus, der die Bildebene E mit der reellen Ebene $\mathbb{R}^2$ identifiziert, d.h. insbesondere wird die x_1-Achse auf die Gerade $a + \mathbb{R} \cdot e_1$ und die x_2-Achse auf die Gerade

$a + \mathbb{R} \cdot e_2$ abgebildet, wobei wir $\mathbb{R}^2$ wie üblich als affine Ebene mit kartesischem Koordinatensystem und Ursprung im Nullpunkt verstehen wollen. Für einen Sensor $(i,j) \in S$ bezeichnen wir die Urbildmenge des Pixels $\text{Pix}_{i,j}$ unter ϱ mit $A_{i,j} \subseteq \mathbb{R}^2$. Dabei gilt offensichtlich

$$A_{i,j} = [i, i+1] \times [j, j+1]$$

und die Vereinigung über alle $(i,j) \in S$ liefert genau die Urbildmenge des Bildrechtecks, d.h. es gilt $B = \varrho\left(\bigcup_{(i,j)\in S} A_{i,j}\right)$. Den aktiven Oberflächenanteil eines Sensors bezeichnet man allgemein auch als die **Apertur** eines Sensors. Dank dem Isomorphismus ϱ identifizieren wir im Folgenden die Apertur eines Sensors (i,j) mit der Urbildmenge $A_{i,j} \subseteq \mathbb{R}^2$ und führen ohne Einschränkung alle weiteren Modellierungen in $\mathbb{R}^2$ anstatt auf der Bildebene E durch.

Wir wollen hier nun zunächst einen einzelnen Sensor modellieren. Da alle Sensoren als baugleich vorausgesetzt sind, bezeichnen wir die Apertur dieses einen Sensors im Folgenden nur kurz mit A. Außerdem werden alle Einzelsensoren gleichzeitig für eine gewisse Zeitdauer $\tau > 0$ belichtet. Man spricht deshalb in diesem Fall von einem **zeitsynchronen homogenen Multisensorensystem** (vgl. [Don09], S. 7). Das von einem Objekt reflektierte Licht trifft nun wie im letzten Abschnitt beschrieben auf die photosensitiven Elemente eines Sensors und setzt hier Ladung frei, was man auch als resultierende elektronisch-physikalische Sekundargröße bezeichnet (vgl. [Pis02], S. 12). Sei $f(x,t) \in \mathbb{R}$ die zu einem Zeitpunkt t im Punkt $x \in A$ auftretende Lichtleistungsdichte, also hier die Lichtleistung pro Flächeneinheit, und $\ell(f)$ die Sekundargröße, also die Ladung oder Spannung (vgl. [Don09], S. 1). Die Leistungsdichte $f(x,t)$ wird dabei im Punkt $x \in A$ in einem bestimmten kleinen Zeitintervall $[t_0, t_0+\tau] \subseteq \mathbb{R}$ für einen Anfangszeitpunkt $t_0 \in \mathbb{R}$ gemessen, dem sogenannten **Akquisitions-** oder **Zeitintegrationsintervall** (vgl. [Pis02], S. 13). Dann ergibt sich die Sekundargröße $\ell(f)$ als Integral über diesem Intervall und der Apertur (vgl. [Pis02], S. 13 sowie [Don09], S. 2 und [FP03], Abschnitt 1.4.2, S. 17 f.):

$$\ell(f) = \int_{t_0}^{t_0+\tau} \int_A f(x,t) d\lambda_E(x) dt.$$

Da der Isomorphismus ϱ Σ_2-Σ_3-messbar ist, ist $\lambda_E := \lambda^2 \circ \varrho^{-1}$ das Bildmaß des Lebesgue-Borel-Maßes λ^2 bzgl. ϱ (vgl. [For11], § 6, S. 76), also das Lebesgue-Maß auf E. Ohne Einschränkung verwenden wir im Folgenden das Lebesgue-Maß λ^2 anstatt diesem Maß.

Nun ist es aber weiter so, dass die Sensorfläche nicht zeitlich und räumlich homogen reagieren kann. Um diesen Effekt zu modellieren, erfolgt eine Gewichtung mit einer sensorspezifischen Dichtefunktion (vgl. [Don09], S. 2). Die Menge $A \times [t_0, t_0+\tau] \subseteq \mathbb{R}^3$ ist Lebesgue-messbar und die Einschränkung $\lambda^3|_{\Sigma_3(A\times[t_0,t_0+\tau])}$ ist ein Maß, das wir kurz mit $\lambda_{A\times[t_0,t_0+\tau]}$ bezeichnen. Weiter bezeichnen wir wie üblich mit $\mathscr{L}^2(A \times [t_0, t_0+\tau])$ die Menge aller 2-fach $\lambda_{A\times[t_0,t_0+\tau]}$-integrierbaren Funktionen $f : A \times [t_0, t_0+\tau] \to \mathbb{R}$ und mit $L^2(A \times [t_0, t_0+\tau])$ die Menge der Äquivalenzklassen aller 2-fach $\lambda_{A\times[t_0,t_0+\tau]}$-fast-überall gleichen reellwertigen Funktionen auf $A \times [t_0, t_0+\tau]$, die wir stets mit einem beliebigen Element der Äquivalenzklasse repräsentieren (vgl. [For11], S. 49 und S. 62). Sei die nicht-negative $\lambda_{A\times[t_0,t_0+\tau]}$-messbare Funktion $k : A \times [t_0, t_0+\tau] \to \overline{\mathbb{R}}_+$ die sensorspezifische Gewichtsfunktion. Dann erhalten wir

$$\ell(f) = \int_{t_0}^{t_0+\tau} \int_A f(x,t) k(x,t) d\lambda^2(x) dt$$

und können die wesentlichen Begriffe in folgendem Sinne definieren (vgl. [Pis02], Definition 2.1.1, S. 13).

Definition 9.1.3. (Sensormaß, Sensorinputfunktion, Sensorfunktional)
Sei $A \subseteq \mathbb{R}^2$ die Apertur und $[t_0, t_0 + \tau] \subseteq \mathbb{R}$ das Akquisitionsintervall eines Sensors. Sei weiter $k : A \times [t_0, t_0 + \tau] \to \overline{\mathbb{R}}_+$ die nicht-negative $\lambda_{A\times[t_0,t_0+\tau]}$-messbare sensorspezifische Gewichtsfunktion. Dann ist

$$\mu := k \cdot \lambda_{A\times[t_0,t_0+\tau]}$$

ein Maß mit Dichte k (vgl. [For11], §4, Satz 11, S. 52); es heißt das **Sensormaß** des Sensors. Die Elemente in $L^2(A \times [t_0, t_0 + \tau], \mu)$ heißen **Sensorinputfunktionen** des Sensors und die Abbildung $\ell : L^2(A \times [t_0, t_0 + \tau]), \mu) \to \mathbb{R}$, definiert durch

$$f \mapsto \ell(f) := \int f d\mu,$$

heißt das **Sensorfunktional**.

Von einem Grauwert sind wir damit aber immer noch ein gutes Stück entfernt. Die Werte der resultierenden Sekundargröße $\ell(f)$ sind beliebige positive reelle Zahlen, weshalb man auch vom **Analogsignal** spricht (vgl. [Don09], S. 2). Bei der Sensormodellierung müssen auch zufällige Phänomene berücksichtigt werden, die sich durch eine reellwertige quadratintegrierbare Zufallsvariable auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ modellieren lassen, den wir hier nicht näher spezifizieren wollen. Diese Zufallsvariable ist zudem abhängig vom Analogsignal, weshalb wir die Zufallsvariable als Abbildung $X : \mathbb{R} \times \Omega \to \mathbb{R}$ formulieren, und hat Erwartungswert 0 (vgl. [Don09], S. 3). Somit kann die Ausgabe des Sensors durch

$$\ell(f) + X(\ell(f), \omega)$$

für ein $\omega \in \Omega$ modelliert werden (vgl. [Don09], S. 3). Dieses Analogsignal ist für den Menschen oder auch den Computer nicht immer unmittelbar zur Weiterverarbeitung geeignet. So muss z.B. die Bewertung der Lichtintensität auf den Menschen angepasst werden (vgl. [Don09]). Nähere Informationen zur quantitativen Visualisierung findet man z.B. in [Jäh05], Kapitel 6. Mathematisch lässt sich das durch Nachschalten einer Umskalierungsfunktion $[0, \infty[\to \mathbb{R}$ modellieren. Diese Umskalierung führt aber in der Weiterverarbeitung unter Umständen zu Problemen (vgl. [Don09], S. 9 f.). Wir werden hier nicht nur aus diesen Gründen ganz auf deren Modellierung verzichten.

Wir wollen nun wieder alle $r \cdot s$ Sensoren des Sensorarrays ins Blickfeld rücken. Jeder Sensor $(i, j) \in S$ besitzt eine Apertur $A_{i,j}$, ein Sensormaß $\mu_{i,j}$, ein Sensorfunktional $\ell_{i,j}$ sowie eine sensorspezifische Zufallsvariable $X_{i,j}$. Das Akquisitionsintervall ist hier für alle Sensoren wegen der gleichzeitigen Belichtung gleich. Wir erhalten somit eine endliche Familie von Analogsignalen

$$\left(\ell_{i,j}(f) + X_{i,j}(\ell_{i,j}(f), \omega)\right)_{(i,j)\in S},$$

die man auch als **Analogbild** bezeichnet (vgl. [Don09], S. 7). Sei $\mathscr{L}^2(\mathbb{R}^2 \times \mathbb{R}, \mu_{i,j})$ die Menge aller Funktionen $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, deren Einschränkung $f|_{A_{i,j}\times[t_0,t_0+\tau]}$ zweifach $\mu_{i,j}$-integrierbar ist. Dann beinhaltet der Durchschnitt

$$\mathscr{L}^2\left((\mu_{i,j})_{(i,j)\in S}\right) := \bigcap_{(i,j)\in S} \mathscr{L}^2(\mathbb{R}^2 \times \mathbb{R}, \mu_{i,j})$$

alle Funktionen $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ mit der Eigenschaft, dass für alle $(i, j) \in S$ die Einschränkung $f|_{A_{i,j}\times[t_0,t_0+\tau]}$ zweifach $\mu_{i,j}$-integrierbar ist. Dieser Durchschnitt ist ein $\mathbb{R}$-Vektorraum (vgl. [Pis02], S. 13). Ohne näher darauf einzugehen, sei $L^2\left((\mu_{i,j})_{(i,j)\in S}\right)$ die Menge der Äquivalenzklassen aller fast überall gleichen Funktionen aus $\mathscr{L}^2\left((\mu_{i,j})_{(i,j)\in S}\right)$ (vgl. [Pis02], S. 13). Dies ist die Menge der in Frage kommenden Sensorinputfunktion für das Sensorarray.

Zum Grauwert wird ein kontinuierliches Analogsignal erst durch einen sogenannten **Framegrabber** (vgl. [Han10], Abschnitt 2.4.2, S. 39), dessen wesentlicher Bestandteil ein Analog-Digital-Wandler ist. Dieser erzeugt mit einer modifizierten Potentiometerschaltung diskrete Messwerte, die sich digital speichern lassen. Dieser gesamte Vorgang wird auch **Quantisierung** genannt. Eine Einführung in die Arbeitsweise eines Analog-Digital-Wandlers findet man z.B. in [Jäh05], Kapitel 9. Die mathematische Definition eines **Quantisierers** lautet wie folgt (vgl. [GL00]).

Definition 9.1.4. (k-Quantisierer)
Sei $P : \mathcal{B}(\mathbb{R}) \to [0,1]$ ein Wahrscheinlichkeitsmaß. Für $k \in \mathbb{N}_+$ heißt eine Funktion $q : \mathbb{R} \to \mathbb{R}$ ein **k-Quantisierer** oder kurz **Quantisierer**, falls q Borel-messbar ist und $\#q(\mathbb{R}) \leq k$ gilt. Die endliche Menge $q(\mathbb{R})$ heißt dann ein **Codebuch** von q.

Für eine Borelmenge $B \in \mathcal{B}(\mathbb{R})$ ist dabei $P(B)$ die Wahrscheinlichkeit, dass ein Grauwert in B liegt. In der Praxis sind $k = 2^e$ für $e \in \mathbb{N}_+$ verschiedene Grauwerte üblich. Für die meisten Anwendungen sind bereits 8 Bit völlig ausreichend, d.h. $2^8 = 256$ verschiedene Grauwerte, da das menschliche Auge ohnehin nicht mehr Graustufen zu differenzieren vermag. Nur in speziellen Anwendungen der Medizin oder der Astronomie können auch bis zu 2^{16} verschiedene Graustufen notwendig und sinnvoll sein (vgl. [BB05], Abschnitt 2.2.6, S. 12 f.). Es wird versucht, einen Quantisierer möglichst optimal zu wählen, d.h. man versucht q so zu wählen, dass der **Quantisierungsfehler** $e(q) := \left(\int_{\mathbb{R}} |x - q(x)|^2 dP(x)\right)^{1/2}$ von q möglichst klein wird. Ein k-Quantisierer q heißt dabei k-optimal, wenn $e(q) = \inf\{e(q') : \ q' \ k\text{-Quantisierer}\}$ gilt. Damit ist gewährleistet, dass das Analogsignal $\ell(f) + X(\ell(f), \omega)$ nur wenig vom sogenannten **Digitalsignal** $q(\ell(f) + X(\ell(f), \omega))$ abweicht. Für unsere Zwecke ist es völlig ausreichend, unter einem Quantisierer eine Funktion zu verstehen, die jedem Analogsignal einen Grauwert zuordnet. Für weiter führende Informationen zu Quantisierern und Quantisierung sei auf das Buch [GL00] von Siegfried Graf und Harald Luschgy verwiesen. Ein einfaches Beispiel eines Quantisierers ist die sogenannte **uniforme Quantisierung**.

Beispiel 9.1.5. (Uniforme Quantisierung)
Sei $k \in \mathbb{N}$ die Anzahl erlaubter Graustufen und sei $I := [a,b] \subseteq \mathbb{R}$ mit $a < b$ ein reelles Intervall, das die Elemente des Analogbilds umfasst. Setze $t := \frac{1}{2k}(b-a)$ und $a_i := a + (2i+1) \cdot t$ für alle $i \in \mathbb{Z}_{0,k-1}$. Dann ist $\mathbb{G} = \{a_0, \ldots, a_{k-1}\}$ die Menge der Grauwerte und $q : \mathbb{R} \to \mathbb{G}$, definiert durch

$$q(x) = \begin{cases} a_0, & x \in \ ]-\infty, a_0 + t[\ , \\ a_i, & x \in \ [a_i - t, a_i + t[\text{ für } i \in \{1, \ldots, k-2\}, \\ a_{k-1}, & x \in \ [a_{k-1} - t, \infty[\end{cases}$$

eine Quantisierungsfunktion. Sei heißt ein **uniformer k-Quantisierer**. ◁

Somit ist es nun möglich, formal zu definieren, was mathematisch unter einem digitalen Grauwertbild zu verstehen ist (vgl. [Pis02], Definition 2.1.3, S. 14).

Definition 9.1.6. (Digitales Bild)
Sei $f \in L^2\left((\mu_{i,j})_{(i,j)\in S}\right)$, sei $(\ell_{i,j}(f) + X_{i,j}(\ell_{i,j}(f), \omega))_{(i,j)\in S}$ das Analogbild des Sensorarrays und sei $q : \mathbb{R} \to \mathbb{R}$ eine Quantisierungsfunktion. Dann heißt die endliche Familie reeller Zahlen

$$(q(\ell_{i,j}(f) + X_{i,j}(\ell_{i,j}(f), \omega)))_{(i,j)\in S}$$

ein **digitales Bild**.

Das Sensormaß, das Sensorfunktional, die sensorspezifische Zufallsvariable, all das sind Größen, die fest einem Sensor zugeordnet sind. Was also ein Bild ausmacht, ist die Sensorinputfunktion f. Das, was wir dann am Bildschirm als Bild sehen, ist nur eine approximative Interpretation, die durch die Quantisierung des Sensoroutputs entsteht. Es scheint somit nur logisch zu sein, sich mit der Sensorinputfunktion genauer auseinanderzusetzen, was wir im folgenden Abschnitt tun wollen.

9.2 Lokale Bildmerkmale

Da durch die Quantisierung das Analogbild nur approximiert wird, reicht die Genauigkeit, die ein digitales Bild bietet, für viele Anwendungen leider nicht aus. Hier wäre es äußerst hilfreich, die Sensorinputfunktion zu kennen. Diese erlaubt eine viel korrektere Betrachtungsweise. Leider ist eine Rekonstruktion der Sensorinputfunktion ohne Zusatzannahmen nicht möglich, allerdings ist häufig bereits eine partielle Rekonstruktion ausreichend, in unserem Fall der sogenannten **zeitintegrierten Sensorinputfunktion** (vgl. [Don09], S. 23). Diese auf dem ganzen Bild zu rekonstruieren, wäre aber viel zu komplex und zu aufwendig, sodass in der Praxis die Sensorinputfunktion nur auf lokalen Bereichen rekonstruiert wird (vgl. [Pis02], S. 15). Diese lokale Rekonstruktion der zeitintegrierten Sensorinputfunktion wird uns dann zu sogenannten **lokalen Bildmerkmalen** führen.

9.2.1 Die zeitintegrierte Sensorinputfunktion

Gegeben ist im Folgenden für eine unbekannte Sensorinputfunktion $f \in L^2((\mu_{i,j})_{(i,j)\in S})$ auf einem Sensorarray der Dimension $r \times s$, also mit $S = \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$, Sensormaßen $\mu_{i,j}$, Sensorfunktionalen $\ell_{i,j}$, Zufallsvariablen $X_{i,j}$ und einer Quantisierungsfunktion q nur das digitale Bild

$$(q(\ell_{i,j}(f) + X_{i,j}(\ell_{i,j}(f), \omega)))_{(i,j)\in S}\,.$$

Mit anderen Worten, gegeben ist einzig eine endliche Familie reeller Zahlen. Für alle Sensoren $(i,j) \in S$ wird im Folgenden die sensorspezifische Dichtefunktion $k_{i,j}$, die auch als Faltungskern bezeichnet wird, als konstant 1 auf $A_{i,j} \times [t_0, t_0 + \tau]$ angenommen und außerhalb mit 0 fortgesetzt, d.h. wir nehmen $k_{i,j} = \mathbb{1}_{A_{i,j}\times[t_0,t_0+\tau]}$ für alle $(i,j) \in S$ an und können sie damit im Folgenden vernachlässigen (vgl. [Pis02], S. 14). In den meisten Anwendungen - und so werden auch wir das hier annehmen - kann jedes Sensormaß $\mu_{i,j}$ als **separierbar** vorausgesetzt werden (vgl. [Pis02], S. 14), d.h. es gibt ein Maß $\alpha_{i,j}$ auf der Apertur $A_{i,j}$ und ein Maß $\nu_{i,j}$ auf dem Akquisitionsintervall $[t_0, t_0 + \tau]$ so, dass $\mu_{i,j}$ das Produktmaß (vgl. [Els11], Kapitel V, Satz und Definition 1.5, S. 167) der Maße $\alpha_{i,j}$ und $\nu_{i,j}$ ist, mit anderen Worten, es gilt

$$\mu_{i,j} = \alpha_{i,j} \otimes \nu_{i,j}.$$

Es lassen sich also sozusagen Zeit und Raum trennen. Da wir die Pixel und damit die Apertur als orthogonal vorausgesetzt haben - genauer verwenden wir $A_{i,j} = [i, i+1] \times [j, j+1]$, werden wir vereinfacht für die Maße $\alpha_{i,j}$ und ebenso für $\mu_{i,j}$ ausschließlich die Lebesgue-Borel-Maße λ^2 bzw. λ verwenden. Das Tripel $(\alpha_{i,j}, \nu_{i,j}, X_{i,j})$ heißt auch die **Sensorcharakteristik** des Sensors $(i,j) \in S$ und das Quadrupel $\big((\alpha_{i,j}, \nu_{i,j}, X_{i,j})_{(i,j)\in S},\ q,\ a,\ \{e_1, e_2\}\big)$ die **Sensorcharakteristik** des Sensorarrays (vgl. [Don09], Definition 1.8.1, S. 14). Wir wollen für die Rekonstruktion zusätzlich noch folgende Nebenbedingungen für die Sensormodellierung festlegen (vgl. auch [Fuc00], S. 162):

(1) Da wir in einem Bild davon ausgehen können, dass alle Sensoren zugleich belichtet werden, weil wir also mit anderen Worten ein zeitsynchrones homogenes Multisensorensystem betrachten, können wir annehmen, dass die Maße $\nu_{i,j}$ alle identisch sind (vgl. [Pis02], S. 15). Dieses eine Maß bezeichnen wir kurz nur mit ν.

(2) Das Sensorarray wird als **pixelperiodisch** (vgl. [Don09], Definition 1.8.3, S. 15) angenommen. Das bedeutet insbesondere, dass die Messwerte jedes Sensors auf dieselbe Weise entstehen, d.h. jeder Sensor arbeitet nicht nur zeit-, sondern auch ortsinvariant. Ist also für alle $(i,j) \in S$ die Abbildung $T_{(i,j)} : \mathbb{R}^2 \to \mathbb{R}^2$ die Translation um den Vektor $(i,j) \in \mathbb{R}^2$, so ist für alle $(i,j) \in S$ das Maß $\alpha_{i,j}$ das Bildmaß (vgl. [For11], § 6, S. 76) von $\alpha_{0,0}$ bzgl. $T_{(i,j)}$, d.h. es gilt
$$\alpha_{i,j} = \alpha_{0,0} \circ T_{(i,j)}^{-1}. \tag{9.2.1}$$

(3) Der Sensor verursacht keine Skalierung, d.h. für alle $(i,j) \in S$ gilt $\ell_{i,j}(\mathbb{1}) = 1$. Insbesondere heißt das, dass das Sensorarray **1-normiert** (vgl. [Don09], Definition 2.4.14, S. 57) ist, d.h. für alle $(i,j) \in S$ gilt $\alpha_{i,j}(\mathbb{1}) = 1$.

Da das Sensormaß separierbar ist, folgt aus dem Satz von Fubini (vgl. z.B. [Els11], Kapitel V, Satz 2.1, S. 175), dass die Funktion $t \mapsto f(x,t)$ für $\alpha_{i,j}$-fast alle $x \in A_{i,j}$ ν-integrierbar und die Funktion $x \mapsto f(x,t)$ für ν-fast alle $t \in [t_0, t_0+\tau]$ $\alpha_{i,j}$-integrierbar ist, und unter anderem gilt für alle $(i,j) \in S$ und alle $f \in L^2((\mu_{i,j})_{(i,j)\in S})$:
$$\int\limits_{A_{i,j}\times[t_0,t_0+\tau]} f(x,t)d\mu_{i,j}(x,t) = \int\limits_{A_{i,j}} \left(\int\limits_{[t_0,t_0+\tau]} f(x,t)d\nu(t) \right) d\alpha_{i,j}(x).$$

Dies führt uns zur Definition der zeitintegrierten Sensorinputfunktion (vgl. [Pis02], S. 15 oder [Don09], S. 24).

Definition 9.2.1. (Zeitintegrierte Sensorinputfunktion)
Sei $\big((\alpha_{i,j}, \nu, X_{i,j})_{(i,j)\in S},\ q,\ a,\ \{e_1, e_2\}\big)$ die Sensorcharakteristik eines zeitsynchronen homogenen Sensorarrays, sei $A_{i,j} \subseteq \mathbb{R}^2$ für alle $(i,j) \in S$ die Apertur des Sensors (i,j) sowie $\mu_{i,j} = \alpha_{i,j} \otimes \nu$ dessen Sensormaß. Sei $[t_0, t_0+\tau] \subseteq \mathbb{R}$ mit $t_0, \tau \in \mathbb{R}_+$ das gemeinsame Akquisitionsintervall der Sensoren und sei $f \in L^2((\mu_{i,j})_{(i,j)\in S})$ eine Sensorinputfunktion. Die Abbildung $\zeta_f : \bigcup_{(i,j)\in S} A_{i,j} \to \mathbb{R}$, definiert durch
$$x \mapsto \int\limits_{[t_0,t_0+\tau]} f(x,t)d\nu(t),$$
heißt die **zeitintegrierte Sensorinputfunktion** von f.

Es gilt somit der folgende offensichtliche Zusammenhang zwischen der zeitintegrierte Sensorinputfunktion, der Sensorinputfunktion und dem Sensorfunktional: Für alle $(i,j) \in S$ gilt
$$\ell_{i,j}(f) = \int\limits_{A_{i,j}} \zeta_f(x)d\alpha_{i,j}(x).$$

Für sogenannte ortssynchrone homogene Multisensorensysteme kann man analog die aperturintegrierte Sensorinputfunktion betrachten und deren lokale Rekonstruktion betreiben, was wir aber nicht tun wollen (vgl. [Don09], Abschnitt 2.2, S. 24 ff.). Somit wird also das, was wir als „Bild" wahrnehmen, d.h. eine Ansammlung von Pixeln unterschiedlicher Graufärbung, insbesondere durch die zeitintegrierte Sensorinputfunktion bestimmt. Es ist damit durchaus berechtigt,

ein Bild durch seine zeitintegrierte Sensorinputfunktion zu identifizieren. Die Kenntnis dieser Funktion lässt natürlich wesentlich genauere Untersuchungen eines Bildes zu, als ein digitales Grauwertbild. Deshalb wollen wir im folgenden Abschnitt die Frage klären, wie sich die zeitintegrierte Sensorinputfunktion aus einer Ansammlung von Grauwerten rekonstruieren lässt, was aus Performancegründen aber nur auf kleinen Teilbereichen eines digitalen Bildes geschieht.

Wir wollen im Folgenden auf die Betrachtung der zufälligen Rauscheffekte, die durch die Zufallsvariablen $X_{i,j}$ modelliert wurden, der Einfachheit halber verzichten. Gegeben ist somit also einzig ein digitales Bild in der Form

$$(q(\ell_{i,j}(f)))_{(i,j)\in S}$$

für eine unbekannte Sensorinputfunktion $f \in L^2((\mu_{i,j})_{(i,j)\in S})$ auf einem Sensorarray der Dimension $r \times s$, das wir nur kurz mit der Indexmenge $S := \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ identifizieren. Wir vereinfachen die Notation etwas, indem wir für jeden Pixelindex $(i,j) \in S$ die reelle Zahl $q(\ell_{i,j}(f))$ eines digitalen Bildes kurz mit $\mathrm{gv}_{i,j}$ bezeichnen. Wie in den meisten Anwendungen üblich und ausreichend, wollen wir $\mathrm{gv}_{i,j} \in \mathbb{Z}_{0,255}$ annehmen, und $\mathrm{gv}_{i,j}$ entsprechend als **Grauwert** des Pixels (i,j) bezeichnen.

Definition 9.2.2. (Grauwertbild)
Die Abbildung $\mathrm{gv} : S \to \mathbb{Z}_{0,255}$ mit $(i,j) \mapsto \mathrm{gv}_{i,j}$, die also jedem Pixel bzw. Sensor ihren Grauwert zuordnet, wird folgerichtig als **Grauwertbild** bezeichnet.

Die zeitintegrierte Sensorinputfunktion würde sich grundsätzlich auch auf dem gesamten Bild, rekonstruieren lassen, allerdings wäre die Komplexität für dieses Unterfangen viel zu hoch, sodass wir uns auf eine lokale Rekonstruktion beschränken, wie es auch in der Praxis, wie z.B. in [Pis02], der Fall ist. Allerdings ist die Darstellung in [Pis02] auf Seite 43 nicht ganz korrekt, da sich die Sensorinputfunktion $f \in \mathrm{Abb}(\mathbb{R}^2 \times \mathbb{R}, \mathbb{R})$ offensichtlich nicht durch Polynome in $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ rekonstruieren lässt. Vielmehr war wohl auch hier die zeitintegrierte Sensorinputfunktion im Sinne des Autors. Zunächst wollen wir darauf eingehen, auf welchen lokalen Bildbereichen bzw. „Pixelfenstern" wir die zeitintegrierte Sensorinputfunktion rekonstruieren können und wollen. Offensichtlich sinnvoll scheint es, zusammenhängende Pixelmengen zu verwenden. Zusammenhang allein wird jedoch nicht ausreichen. Die erste Eigenschaft, die deshalb an lokale Pixelfenster gestellt wird, ist die **diskrete Konvexität** (vgl. [Don09], Definition 2.4.1, S. 49).

Definition 9.2.3. (diskret konvex)
Eine endliche Teilmenge $M \subseteq \mathbb{Z} \times \mathbb{Z}$ heißt **diskret konvex**, wenn es eine konvexe (also insbesondere zusammenhängende) Menge $K \subseteq \mathbb{R}^2$ gibt mit $M = K \cap (\mathbb{Z} \times \mathbb{Z})$.

Dazu wollen wir Beispiele verschiedener diskret konvexer Menge angeben, die im weiteren Verlauf von Bedeutung sein werden.

Beispiel 9.2.4.

a) Für natürliche Zahlen $k, \ell \in \mathbb{N}$ ist die Menge

$$M_{k,\ell} := \{(i,j) \in \mathbb{N} \times \mathbb{N} : \ i \leq k \text{ und } j \leq \ell\}$$

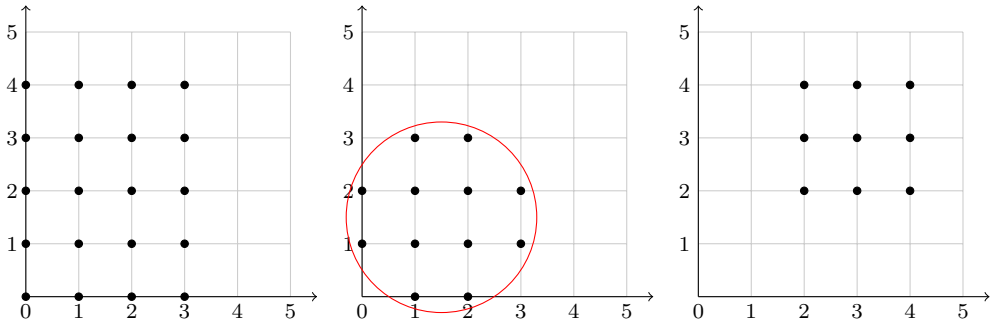
wegen $M_{k,\ell} = (\mathbb{Z} \times \mathbb{Z}) \cap ([0,k] \times [0,\ell])$ eine diskret konvexe Menge.

b) Die annähernd kreisförmige Menge $M_{3,3} \setminus \{0,3\}^2$ (siehe Abbildung 9.6) ist diskret konvex. Denn wählt man für $K \subseteq \mathbb{R}^2$ eine offene Kreisscheibe $B_r((\frac{3}{2}, \frac{3}{2}))$ mit einem Radius r im Intervall $\left]\frac{1}{2}\sqrt{10}, \frac{3}{2}\sqrt{2}\right[\subseteq \mathbb{R}$, so gilt $M_{3,3} \setminus \{0,3\}^2 = (\mathbb{Z} \times \mathbb{Z}) \cap K$.

c) Sei $(i,j) \in \mathbb{Z} \times \mathbb{Z}$ und $k \in \mathbb{N}$. Dann ist

$$G^k_{(i,j)} := \mathbb{Z}_{i-k,i+k} \times \mathbb{Z}_{j-k,j+k}$$

wegen $G^k_{(i,j)} = (\mathbb{Z} \times \mathbb{Z}) \cap ([i-k, i+k] \times [j-k, j+k])$ eine diskret konvexe Menge. Sie heißt ein quadratisches **Gitter der Länge** $2k+1$ **mit Zentrum im Punkt** (i,j). ◁



(a) $M_{3,4}$ (b) $M_{3,3} \setminus \{0,3\}^2$ (c) $G^1_{(3,3)}$

Abbildung 9.6: Darstellungen diskret konvexer Mengen.

Wie man auch an diesen Beispielen bereits erahnen kann, bleibt die Eigenschaft der diskreten Konvexität unter einer Verschiebung um einen ganzzahligen Vektor erhalten.

Lemma 9.2.5. *Sei $M \subseteq \mathbb{Z} \times \mathbb{Z}$ eine diskret konvexe Menge. Dann ist für alle $(i,j) \in \mathbb{Z} \times \mathbb{Z}$ auch das translatorische Bild $T_{(i,j)}(M)$ von M diskret konvex.*

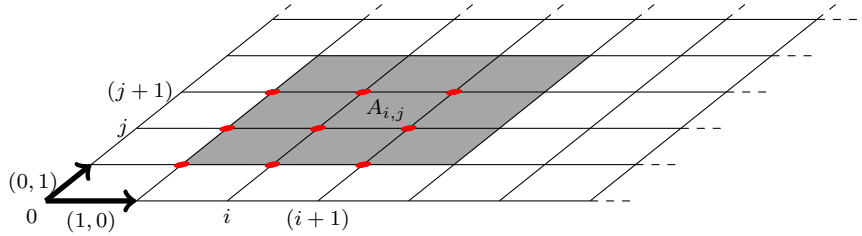
Beweis: Sei $(i,j) \in \mathbb{Z} \times \mathbb{Z}$. Da M diskret konvex ist, gibt es eine konvexe Menge $K \subseteq \mathbb{R}^2$ mit $M = (\mathbb{Z} \times \mathbb{Z}) \cap K$. Offensichtlich gilt $T_{(i,j)}(M) = (\mathbb{Z} \times \mathbb{Z}) \cap T_{(i,j)}(K)$ und $T_{(i,j)}(K) \subseteq \mathbb{R}^2$ ist erneut konvex. Somit ist auch $T_{(i,j)}(M)$ diskret konvex. □

Dieser Sachverhalt wird gegen Ende des Abschnitts noch von großer Bedeutung sein. An dieser Stelle bedeutet das insbesondere, dass für das obige Beispiel eines Gitters der Länge 3 z.B. gilt $G^1_{(3,3)} = T_{(3,3)}(G^1_{(0,0)})$. Das Konzept der diskret konvexen Mengen lässt sich unmittelbar auf das Sensorarray, also auf die Aperturebene, übertragen. Man spricht in diesem Fall vom zugehörigen **Lokalisierungsfenster** (vgl. [Don09], Definition 2.4.1, S. 49).

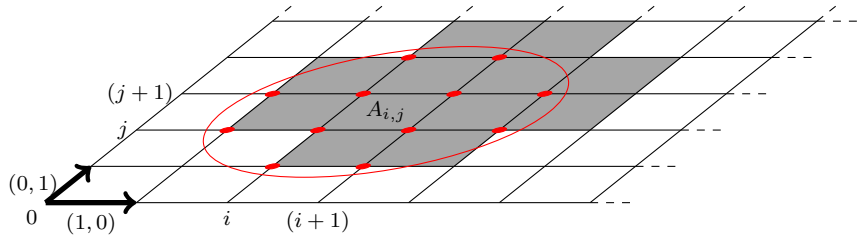
Definition 9.2.6. (Lokalisierungsfenster)
Sei das Sensorarray kurz durch $S = \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ repräsentiert und sei $M \subseteq S$ eine diskret konvexe Menge. Dann heißt die Teilmenge $\mathrm{Loc}_M := \bigcup_{(k,\ell) \in M} A_{k,\ell}$ von $\mathbb{R}^2$ das zu M gehörige **Lokalisierungsfenster**.

Abbildung 9.7 stellt anhand zweier Beispiele den Zusammenhang zwischen diskret konvexen Mengen und den zugehörigen Lokalisierungsfenstern graphisch dar. Somit ist implizit klar, dass sich Randpixel nicht als Zentren solcher Lokalisierungsfenster eignen, hier müsste man andersartige Lokalisierungsfenster wählen. Wir wollen auf die Betrachtung von Randpixel jedoch verzichten, da sie auch für die folgenden Anwendungen nur wenig Bedeutung haben. D.h. insbesondere wollen wir voraussetzen, dass für die verwendeten diskret konvexen Mengen M stets $M \subseteq S$ gilt. Wie man an den Abbildungen auch sofort erkennen kann, ist Loc_M genau

die Urbildmenge der Menge $\bigcup_{(k,\ell)\in M} \mathrm{Pix}_{k,\ell} \subseteq \mathbb{R}^3$ unter der affin-linearen bijektiven Abbildung $\varrho : \mathbb{R}^2 \to E$ aus Gleichung (9.1.1). Das Ziel ist es nun, die zeitintegrierte Sensorinputfunktion auf $\mathrm{Loc}_M \neq \emptyset$, also die Einschränkung $\zeta_f|_{\mathrm{Loc}_M} \in \mathrm{Abb}(\mathrm{Loc}_M, \mathbb{R})$ zu rekonstruieren, d.h. durch eine geeignete Funktion zu approximieren.



(a) $G^1_{(i,j)}$



(b) $(1,1) + M_{3,3} \setminus \{0,3\}^2$

Abbildung 9.7: Darstellung verschiedener Lokalisierungsfensters.

Gesucht ist dazu zunächst ein endlich-dimensionaler reeller Untervektorraum des Vektorraums $\mathrm{Abb}(\mathrm{Loc}_M, \mathbb{R})$ als sogenannter **Rekonstruktionsraum** zu einer nicht-leeren, diskret-konvexen Menge $M \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$. Die Antwort findet man im Vektorraum $\mathscr{P}_{\leq n}(\mathrm{Loc}_M, \mathbb{R})$ der reellen Polynomfunktionen vom Grad $\leq n$ auf Loc_M, der dafür optimal geeignet ist; dazu später mehr. Dabei spielt es keine Rolle, ob man Polynomfunktionen aus $\mathscr{P}_{\leq n}(\mathrm{Loc}_M, \mathbb{R})$, also auf Loc_M, oder aus $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, also auf $\mathbb{R}^2$, verwendet (vgl. [Don09], S. 50), d.h. insbesondere ist jede Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ $\alpha_{i,j}$-integrierbar für alle $(i,j) \in S$. Die Gradschranke n ist hierbei in Abhängigkeit vom Lokalisierungsfenster zu wählen. So soll n derart gewählt werden, dass die Dimension von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ die Kardinalität von M nicht übersteigt, d.h. n ist so zu wählen, dass

$$\dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})) = \binom{n+2}{2} \leq \#M \tag{9.2.2}$$

gilt (vgl. [Don09], S. 50 oder [Pis02], S. 43). Dadurch werden Rauscheffekte unterdrückt (vgl. [Pis02], S. 43). Eine zu große Wahl von M bei geringer Dimension ist allerdings auch nachteilig, da in diesem Fall Detailinformationen verloren gehen, was nicht gewünscht ist (vgl. [Pis02], S. 43). Somit ist es optimal, wenn die Differenz $\#M - \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ größer oder gleich Null und möglichst klein ist. Dies ist z.B. sehr gut erfüllt für $n = 2$ und Gitter $G^1_{(i,j)}$ der Länge 3 oder für $n = 3$ und $M = (i,j) + M_{3,3} \setminus \{0,3\}^2$. Zur Rekonstruktion mit Polynomfunktionen ist die Eigenschaft der diskreten Konvexität alleine jedoch nicht ausreichend. An die diskret konvexen Mengen ist noch eine andere Anforderung zu stellen, die diskret konvexe Mengen der Form aus Abbildung 9.7 bereits erfüllen, jedoch eine diskret konvexe Menge im Allgemeinen nicht erfüllt: Sie müssen zulässig für Polynomapproximationen bis zu einem bestimmten Grad n sein, was man kurz auch als **n-zulässig** bezeichnet (vgl. [Don09], Definition 2.4.5).

Definition 9.2.7. (n-zulässig)
Eine diskret konvexe Teilmenge $M \subseteq \mathbb{Z} \times \mathbb{Z}$ heißt **zulässig** für Polynomapproximation vom Grad $\leq n$ oder kurz **n-zulässig**, wenn für alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ gilt:

$$(\forall (i,j) \in M : \ p(i,j) = 0) \quad \Longrightarrow \quad p = 0.$$

Eine nicht-leere, diskret konvexe Menge ist mit anderen Worten genau dann n-zulässig, wenn das Nullpolynom das einzige Polynom ist, das an allen Punkten von M verschwindet. Außerdem erfüllt eine nicht-leere, n-zulässige Menge M Gleichung (9.2.2). Die leere Menge ist demnach zulässig für Polynomapproximation in jedem Grad, aber natürlich eher wenig interessant. Wie bereits erwähnt, sind die betrachteten diskret konvexen Mengen aus Abbildung 9.6 zulässig für Polynomapproximation bis zu einem bestimmten Grad. Bis zu welchem Grad geht aus folgenden Beispielen hervor (vgl. [Don09], Beispiel 2.4.6, S. 53).

Beispiel 9.2.8.

a) Seien $m, k \in \mathbb{N}_+$. Dann ist die diskret konvexe Menge $M_{m,k}$ zulässig für Polynomapproximationen vom Grad $\leq n \leq \min\{m, k\}$. Somit ist insbesondere $M_{2,2}$ zulässig für Polynomapproximationen vom Grad ≤ 2.

b) Die diskret konvexe Menge $M_{3,3} \setminus \{0,3\}^2$ ist 3-zulässig.

c) Die diskret konvexe Menge $M_{2,2} \setminus \{0,2\}^2$ ist 1-zulässig, jedoch nicht 2-zulässig. ◁

Die Eigenschaft der n-Zulässigkeit bleibt auch hier unter ganzzahliger Verschiebung erhalten (vgl. [Don09], Beispiel 2.4.6, S. 53).

Lemma 9.2.9. *Sei $M \subseteq \mathbb{Z} \times \mathbb{Z}$ n-zulässig. Dann ist für alle $(i,j) \in \mathbb{Z} \times \mathbb{Z}$ auch das translatorische Bild $T_{(i,j)}(M)$ eine n-zulässige Menge.*

Damit folgt aus den letzten Beispielen, dass insbesondere das im Nullpunkt zentrierte quadratische Gitter $G^1_{(0,0)}$ wegen $G^1_{(0,0)} = T_{(-1,-1)}(M_{2,2})$ zulässig für Polynomapproximationen vom Grad ≤ 2 ist. Betrachten wir nun eine n-zulässige Menge $M \subseteq S := \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ der Kardinalität $m \in \mathbb{N}_+$. Die Einschränkung $\mathrm{gv}|_M$ des Grauwertbildes $\mathrm{gv} : S \to \mathbb{R}$ auf die Menge $M \subseteq S$ können wir auch als endliche reelle Folge $(\mathrm{gv}_{i,j})_{(i,j)\in M} \in \mathrm{Abb}(M, \mathbb{R})$ notieren. Indem wir die Elemente von M mittels der lexikographischen Ordnung

$$(i,j) \preceq (k,l) \quad :\Longleftrightarrow \quad i \leq k \text{ oder } [i = k \text{ und } j \leq l]$$

aufsteigend ordnen, können wir die endliche reelle Folge $(\mathrm{gv}_{i,j})_{(i,j)\in M} \in \mathrm{Abb}(M, \mathbb{R})$ auch als Vektor $(\mathrm{gv}_{i_1,j_1}, \ldots, \mathrm{gv}_{i_m,j_m}) \in \mathbb{R}^m$ auffassen, den wir den zu M gehörigen **Grauwertvektor** nennen. Wir werden im Folgenden das Bild $\mathrm{gv}(M)$ von M unter gv stets als Vektor in $\mathbb{R}^m$ auffassen. An dieser Stelle erinnern wir daran, dass $\mathrm{gv}_{i,j} = \ell_{i,j}(f)$ und

$$\ell_{i,j}(f) = \int_{A_{i,j}} \zeta_f(x) d\alpha_{i,j}(x)$$

für alle $(i,j) \in S$ gilt. Sei nun für alle Pixel (i,j) einer diskret konvexen, n-zulässigen Menge $M \subseteq S$ die Linearform $\psi_{i,j} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ definiert durch

$$\psi_{i,j}(p) = \int\limits_{A_{i,j}} p(x) d\alpha_{i,j}(x) = \int_i^{i+1} \int_j^{j+1} p(x,y) dx dy.$$

Wie man hier sehr schön sehen kann, ist nun also eine Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu suchen, die die zeitintegrierte Sensorinputfunktion ζ_f auf einer n-zulässigen Menge $M \subseteq S$ möglichst gut approximiert. Es sei hier auch kurz erwähnt, dass die laufend zu berechnenden mehrdimensionalen Integrale auch möglichst effektiv und numerisch stabil berechnet werden sollen. Hier ist es möglich, die Quadraturformeln der Gauß-Quadratur für eindimensionale Integrale auf höher-dimensionale Integrale zu übertragen. Diese **mehrdimensionalen Quadraturformeln** werden in [Pis02], Abschnitt 2.5, S. 36 ff., erläutert und sollen hier nicht näher beleuchtet werden.

Sei weiter $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathrm{Abb}(M, \mathbb{R})$ der durch $\Psi_M(p) = (\psi_{i,j}(p))_{(i,j)\in M}$ definierte Vektorraum-Homomorphismus. Dann können wir für jede Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ analog mit Hilfe der obigen lexikographischen Ordnung auch $\Psi_M(p)$ als Vektor in $\mathbb{R}^m$ und damit die Abbildung Ψ_M als $\mathbb{R}$-Vektorraum-Homomorphismus $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ auffassen. Wir suchen also eine Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ so, dass $\Psi_M(p)$ und $\mathrm{gv}(M)$ möglichst gut übereinstimmen, d.h. eine Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, für die

$$\|\Psi_M(p) - \mathrm{gv}(M)\| = \sum_{(i,j)\in M} \left(\psi_{i,j}(p) - g_{i,j}\right)^2$$

minimal wird (vgl. [Don09], S. 50 und [Pis02], S. 43). Mit anderen Worten suchen wir eine Lösung des euklidischen, linearen Ausgleichsproblems

$$\|\Psi_M(p) - \mathrm{gv}(M)\| = \min\left\{\|\Psi_M(q) - \mathrm{gv}(M)\| :\ q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})\right\}. \tag{9.2.3}$$

Bevor wir allerdings auf die Lösbarkeit sowie ggf. die Lösung dieses linearen Ausgleichsproblems zu sprechen kommen, wollen wir noch auf die versprochenen Gründe eingehen, warum sich gerade Polynomfunktionen für die Rekonstruktion besonders gut eignen.

Bemerkung 9.2.10. (Gründe für die Verwendung von Polynomfunktionen)
Donner und Pisinger nennen in [Don09] bzw. [Pis02] folgende Gründe für die Verwendung von Polynomfunktionen zur Rekonstruktion der zeitintegrierten Sensorinputfunktion:

(1) Sensorinputfunktionen glätten im Allgemeinen auch zunächst erwartete Intensitätssprünge. Daher ist es nicht unrealistisch anzunehmen, dass Sensorinputfunktionen sehr hoch differenzierbar sind. Wie aus dem Taylorschen Satz hervorgeht, approximieren Polynome ausreichend glatte Funktionen besonders gut.

(2) Pisinger kommt in [Pis02] durch Tests an großen Stichproben zu dem Ergebnis, dass die Approximation mit Polynomen der optimalen Karhunen-Loève-Approximation recht nahe kommt. Jedoch kommt die Polynomapproximation ohne Auswertung statistischer Daten und der Berechnung einer Autokovarianzmatrix aus und ist daher der Karhunen-Loève-Approximation vorzuziehen (vgl. [Pis02], S. 68).

(3) Grundsätzlich wäre auch eine Rekonstruktion mit trigonometrischen Polynomen denkbar. Auch diese wurde von Pisinger in [Pis02] untersucht, mit dem Ergebnis, dass Polynome zur Approximation besser geeignet sind als trigonometrische Polynome, da sie bessere Ergebnisse liefern und die Performance steigern.

(4) Wie es auch hier das Ziel sein wird, werden lokale Rekonstruktionen von zeitintegrierten Sensorinputfunktionen zur Bestimmung von lokalen Bildmerkmalen verwendet. Diese Bildmerkmale sollten nach Möglichkeit invariant unter Translationen, Rotationen und Skalierungen sein. Auch diese Eigenschaften sollte der verwendete Rekonstruktionsraum

haben, d.h. der Rekonstruktionsraum V sollte ein translations-, rotations- und skalierungsinvarianter Untervektorraum endlicher Dimension des Vektorraums der stetigen Funktionen $\mathscr{C}(\mathbb{R}^2, \mathbb{R})$ auf $\mathbb{R}^2$ sein, was für endlich-dimensionale $\mathbb{R}$-Vektorräume nur im Fall $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ für ein $n \in \mathbb{N}_+$ gegeben ist.

Im weiteren Verlauf wollen wir die Lösbarkeit des euklidischen, linearen Ausgleichsproblems aus Gleichung (9.2.3) untersuchen und im Falle der Lösbarkeit die eindeutig bestimmte Lösung möglichst effizient angeben.

9.2.2 Rekonstruktion der zeitintegrierten Sensorinputfunktion

Es gelten auch in diesem Abschnitt zunächst dieselben Voraussetzungen wie im letzten Abschnitt, d.h. insbesondere sei $M \subseteq S$ eine n-zulässige Menge der Kardinalität $m \in \mathbb{N}_+$, sei stets $\mathrm{gv}(M) \in \mathbb{R}^m$ der zugehörige Grauwertvektor, sei $\psi_{i,j} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ für jedes Pixel $(i,j) \in M$ diejenige Linearform, die über die Apertur $A_{i,j}$ des Sensors (i,j) integriert, und sei $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ die durch

$$\Psi_M(p) = (\psi_{i_1,j_1}(p), \ldots, \psi_{i_m,j_m}(p))$$

definierte lineare Abbildung, wobei M nach wie vor lexikographisch geordnet sein soll. Das euklidische, lineare Ausgleichsproblem aus Gleichung (9.2.3) wollen wir zunächst etwas umformulieren. Sei dazu in diesem Abschnitt stets $\nu := \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$. Wir wählen nun die kanonische Termbasis $B := (p_1, \ldots, p_\nu)$ von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, die bzgl. einer gradkompatiblen Termordnung aufsteigend geordnet ist. Wir bestimmen nun anschließend die Darstellungsmatrix $\mathcal{M}_E^B(\Psi_M) \in \mathrm{Mat}_{m,\nu}(\mathbb{R})$ der linearen Abbildung Ψ_M bzgl. des Basenpaares B, E, wobei E die kanonische Basis von $\mathbb{R}^m$ sei. Das euklidische, lineare Ausgleichsproblem

$$\left\|\mathcal{M}_E^B(\Psi_M) \cdot c - \mathrm{gv}(M)\right\| = \min\left\{\left\|\mathcal{M}_E^B(\Psi_M) \cdot x - \mathrm{gv}(M)\right\| : \ x \in \mathbb{R}^\nu\right\} \tag{9.2.4}$$

ist zu dem aus Gleichung (9.2.3) äquivalent, d.h. genau dann wenn Gleichung (9.2.4) lösbar ist, ist auch Gleichung (9.2.3) lösbar. Ein Lösungsvektor $c \in \mathbb{R}^\nu$ von Gleichung (9.2.4) ist dabei genau der Koordinatenvektor bzgl. B einer Lösung $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ des linearen Ausgleichsproblems aus Gleichung (9.2.3).

Die Untersuchung der Lösbarkeit eines euklidischen, linearen Ausgleichsproblems wie in Gleichung (9.2.4) stellt sicherlich kein Problem dar. Ist ein derartiges Ausgleichsproblem lösbar, kommen aufbauend auf der Darstellungsmatrix $\mathcal{M}_E^B(\Psi_M)$ viele bekannte Verfahren in Frage, wie z.B.:

- durch Lösen der sogenannten **Normalengleichung**, die aber in vielen praktischen Anwendungen schlecht konditioniert ist,
- mittels **Singulärwertzerlegung** und Verwendung der Pseudoinversen von $\mathcal{M}_E^B(\Psi_M)$ oder
- mit Hilfe der $\mathcal{QR}$**-Zerlegung** von $\mathcal{M}_E^B(\Psi_M)$ (siehe z.B. [SK09]).

Wie Fuchs in [Fuc00] ausführt, ist für den auch hier vorliegenden Fall, dass diskrete Messdaten - in unserem Fall die endlich vielen Grauwerte der n-zulässigen Menge M - durch Polynome approximiert werden sollen, eine Orthogonalentwicklung vorteilhafter (vgl. [Fuc00], Abschnitt 2.4, S. 36). Diese verwendet beispielsweise auch Pisinger in [Pis02]. Wir wollen diese deshalb auch für unsere Zwecke verwenden und das Lösungsverfahren mittels Orthogonalentwicklung im

Folgenden kurz vorstellen. Bevor wir jedoch über Orthogonalität von Polynomfunktionen reden, wollen wir kurz angeben, wie man aufbauend auf Gleichung (9.2.5) mit Orthogonalentwicklung eine Lösung bestimmen kann. Dies spiegelt auch das weitere Vorgehen wider.

Bemerkung 9.2.11. (Naive Orthogonalentwicklung)
Wir wollen hier die lineare Abbildung $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ als injektiv voraussetzen. Dann ist $\mathrm{Im}(\Psi_M) \subseteq \mathbb{R}^m$ ein $\mathbb{R}$-Untervektorraum der Dimension $\nu = \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ von $\mathbb{R}^m$. Sei $B := (p_1, \ldots, p_\nu)$ wiederum die Termbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ in einer gradkompatiblen Termordnung. Dann ist $(\Psi_M(p_1), \ldots, \Psi_M(p_\nu))$ eine Basis von $\mathrm{Im}(\Psi_M) \subseteq \mathbb{R}^m$. Aus dieser Basis kann man z.B. mit dem Schmidtschen Orthogonalisierungsverfahren eine Orthogonalbasis $(b_1, \ldots, b_\nu)$ von $\mathrm{Im}(\Psi_M)$ bzgl. des Standardskalarprodukts in $\mathbb{R}^m$ bestimmen. Dann ist

$$x^{(M)} := \sum_{k=1}^{\nu} \frac{\langle \mathrm{gv}(M), b_k \rangle}{\langle b_k, b_k \rangle} b_k$$

die Orthogonalprojektion des Grauwertvektors $\mathrm{gv}(M) = (g_{i_1,j_1}, \ldots, g_{i_m,j_m}) \in \mathbb{R}^m$ auf $\mathrm{Im}(\Psi_M)$, d.h. die eindeutig bestimmte Lösung des euklidischen, linearen Ausgleichsproblems

$$\|x^{(M)} - \mathrm{gv}(M)\| = \min\{\|x - \mathrm{gv}(M)\| : \ x \in \mathrm{Im}(\Psi_M)\}.$$

Durch Lösen des Gleichungssystems $\mathcal{M}_E^B(\Psi_M) \cdot c = x^{(M)}$ erhalten wir den Koordinatenvektor einer Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, die das Ausgleichsproblem 9.2.3 löst. Diese Lösung p ist dann ebenfalls eindeutig bestimmt.

Die entscheidende Voraussetzung bei dieser Lösungsmethode ist die Injektivität der linearen Abbildung Ψ_M. Somit wird es für das weitere Vorgehen entscheidend sein, die Frage nach der Injektivität von Ψ_M zu klären. Ist diese Voraussetzung gegeben, so zeigt die letzte Bemerkung, dass die Lösung des Rekonstruktionsproblems aus Gleichung (9.2.3) grundsätzlich kein Problem darstellt, das uns größere Schwierigkeiten bereitet. Allerdings lässt in diesem Lösungsansatz die Effizienz und auch die numerische Stabilität zu wünschen übrig, man denke besonders an die bekannte numerische Instabilität des Schmidtschen Orthogonalisierungsverfahrens. Deshalb wollen wir hier auch einen effizienten und numerisch stabilen Weg vorstellen, der in [Haa00], [Fuc00] und [Pis02] zu finden ist und in der Praxis viele Anwendungen findet. Dazu benötigen wir allerdings ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, wobei uns der folgende Satz entscheidend weiterhelfen wird (folgt mit [Fis05], Kapitel 6).

Satz 9.2.12. *Seien $\varphi_1, \ldots, \varphi_m \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ Linearformen und sei $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ die durch $\Phi(p) = (\varphi_1(p), \ldots, \varphi_m(p))$ definierte lineare Abbildung. Dann sind folgende Aussagen äquivalent:*

(i) Φ *ist injektiv.*

(ii) $\{\varphi_1, \ldots, \varphi_m\}$ *ist ein Erzeugendensystem des Dualraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$.*

(iii) Die Abbildung $\langle \cdot, \cdot \rangle_\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \times \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}$ mit $\langle p, q \rangle_\Phi = \langle \Phi(p), \Phi(q) \rangle$ ist ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.

Ist eine der äquivalenten Aussagen erfüllt, gilt insbesondere $\dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})) \leq m$.

Wenn wir ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ durch eine geeignete Abbildung Φ gegeben haben, können wir natürlich auch eine Orthogonal- bzw. Orthonormalbasis bzgl. dieses Skalarprodukts bestimmen, z.B. mit dem Orthonormalisierungsverfahren nach Gram-Schmidt. Mit einer Orthonormalbasis ist es dann wiederum ein Leichtes, eine Funktion $p_0 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu bestimmen,

die zu einem gegebenen Vektor $y \in \mathbb{R}^m$ das lineare Ausgleichsproblem

$$\|\Phi(p_0) - y\| = \min\{\|\Phi(p) - y\| : \ p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})\} \tag{9.2.5}$$

löst. Wie diese Lösung p_0 in diesem Fall aussieht, zeigt uns der nächste Satz, der die Orthogonalprojektion auf $\mathrm{Im}(\Phi)$ analog zu Bemerkung 9.2.11 bereits beinhaltet (vgl. zur Orthogonalprojektion im Allgemeinen [Fis05], Kapitel 5, und hier im Speziellen [Pis02], Satz 3.1.1, S. 44).

Satz 9.2.13. (Orthogonalprojektion)
Seien $\varphi_1, \ldots, \varphi_m \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ Linearformen so, dass $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$, definiert durch $\Phi(p) = (\varphi_1(p), \ldots, \varphi_m(p))$, ein Skalarprodukt

$$\langle \cdot, \cdot \rangle_\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \times \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}, \quad (p, q) \mapsto \langle \Phi(p), \Phi(q) \rangle$$

auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ induziert, und sei $(q_1, \ldots, q_\nu)$ eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_\Phi$. Dann ist für alle $y \in \mathbb{R}^m$ die Polynomfunktion

$$p_0 := \sum_{i=1}^{\nu} \langle y, \Phi(q_i) \rangle \cdot q_i$$

die eindeutig bestimmte Lösung des linearen Ausgleichsproblems

$$\|\Phi(p_0) - y\| = \min\{\|\Phi(p) - y\| : \ p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})\}.$$

Ist $(q_1, \ldots, q_\nu)$ „nur" eine Orthogonalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_\Phi$, so erhält man mit den reellen Koeffizienten

$$\frac{\langle y, \Phi(q_1) \rangle}{\langle q_1, q_1 \rangle_\Phi}, \ldots, \frac{\langle y, \Phi(q_\nu) \rangle}{\langle q_\nu, q_\nu \rangle_\Phi}$$

eine analoge Darstellung. Die reellen Koeffizienten der Orthogonalentwicklung von p_0 bzgl. einer Orthonormalbasis werden entsprechend als **Orthonormalkoeffizienten** von p_0 bzgl. der Orthonormalbasis $(q_1, \ldots, q_\nu)$ bezeichnet. Analog sprechen wir auch von **Orthogonalkoeffizienten**, falls $(q_1, \ldots, q_\nu)$ eine Orthogonalbasis ist.

Damit bleibt jetzt „nur" noch die Frage zu klären, wann Ψ_M injektiv ist, bzw. wann Ψ_M ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ induziert. Dazu knüpfen wir an den letzten Satz an und betrachten die Frage allgemeiner. Ob eine lineare Abbildung $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ der Form des letzten Satzes ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ induziert, hängt offensichtlich maßgeblich von den Linearformen φ_i ab. Ein Beispiel für Linearformen, die eine injektive lineare Abbildung bzw. ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ induzieren, liefern uns die sogenannten **Auswertungsfunktionale** $\mathrm{eval}_x : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}$ mit $\mathrm{eval}_x(p) = p(x)$ für alle $x \in \mathbb{R}^2$, die eine Polynomfunktion p an einer bestimmten Stelle x „auswerten". Um zu untersuchen, für welche Menge von Punkten $x \in \mathbb{R}^2$ die Auswertungsfunktionale ein Erzeugendensystem von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ bilden und damit ein Skalarprodukt induzieren, wird der diskrete Begriff einer n-zulässigen Menge erweitert auf Teilmengen von $\mathbb{R}^2$ (vgl. [Don09], Definition 2.3.9, S. 40).

Definition 9.2.14. (n-Eindeutigkeitsmenge)
Eine Menge $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ von $m \geq \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ Punkten heißt eine **n-Eindeutigkeitsmenge** bzgl. $\mathbb{R}^2$, falls für alle Polynomfunktionen $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ gilt:

$$(\forall i \in \{1, \ldots, m\} : \ p(x_i) = 0) \quad \Longrightarrow \quad p = 0.$$

Für univariate Polynome ist also jede Menge mit mindestens $n+1$ Punkten eine n-Eindeutigkeitsmenge, denn aus dem Fundamentalsatz der Algebra ist bekannt, dass ein univariates Polynom vom Grad n höchstens n Nullstellen besitzen kann. Für multivariate Polynome ist die Sachlage jedoch schwieriger, so verschwindet z.B. die Polynomfunktion $p(x,y) = y$ auf jeder Teilmenge M von $\mathbb{R} \times \{0\}$, obwohl sie nicht die Nullfunktion ist.

Beispiel 9.2.15. Jede diskret konvexe, n-zulässige Menge $M \subseteq \mathbb{Z} \times \mathbb{Z}$, deren Kardinalität mindestens so groß wie die Dimension von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ ist, ist eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$. ◁

Wie sich leicht zeigen lässt, bleibt die Eigenschaft einer n-Eindeutigkeitsmenge unter bestimmten Abbildungen erhalten, sodass insbesondere auf einfache Weise aus n-zulässigen Mengen n-Eindeutigkeitsmengen durch Translation konstruiert werden können (vgl. [Don09], Folgerung 2.3.14, S. 43).

Lemma 9.2.16. *Ist $M \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$ und $u \in \mathbb{R}^2$. Dann ist auch $T_u(M)$ eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$.*

Die n-Eindeutigkeitsmengen lassen sich nicht nur aus bekannten Beispielen konstruieren, sondern auch auf verschiedene Weisen charakterisieren. Eine für uns bedeutende Charakterisierung stellt einen Zusammenhang her zwischen den Auswertungsfunktionalen und einer n-Eindeutigkeitsmenge (vgl. [Pis02], Satz 3.2.1, S. 46).

Satz 9.2.17. *Seien $x_1, \ldots, x_\nu \in \mathbb{R}^2$ paarweise verschiedene Punkte, wobei ν die Dimension von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bezeichne. Dann sind folgende Aussagen äquivalent:*

(i) $\{x_1, \ldots, x_\nu\}$ ist eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$.

(ii) Die Auswertungsfunktionale $\mathrm{eval}_{x_1}, \ldots, \mathrm{eval}_{x_\nu} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ an den Stellen $x_1, \ldots, x_\nu$ bilden eine Basis des Dualraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$.

Aus dem letzten Satz und Satz 9.2.12 folgt sofort, dass sich mit den Auswertungsfunktionalen ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ definieren lässt. Dieses wird uns auch weiter begleiten, weshalb wir auf einer endlichen n-Eindeutigkeitsmenge $X = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ die durch

$$p \mapsto (\mathrm{eval}_{x_1}(p),\ \ldots,\ \mathrm{eval}_{x_m}(p)) = (p(x_1),\ \ldots,\ p(x_m))$$

definierte lineare Abbildung $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ im Folgenden stets mit Eval_X bezeichnen. Somit erhalten wir auf folgende Weise ein erstes Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Dieses Korollar ist eine unmittelbare Folgerung aus dem letzten Satz und Satz 9.2.12 (siehe dazu auch [Pis02], Satz 3.2.6, S. 47).

Korollar 9.2.18. *Sei $m \geq \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ und sei $X := \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$. Dann ist die lineare Abbildung $\mathrm{Eval}_X : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ injektiv und die bilineare Abbildung $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X} : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \times \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}$, definiert durch*

$$\langle p, q \rangle_{\mathrm{Eval}_X} = \langle \mathrm{Eval}_X(p), \mathrm{Eval}_X(q) \rangle = \sum_{i=1}^{m} p(x_i) q(x_i),$$

ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.

Auf einer speziellen n-Eindeutigkeitsmenge, nämlich auf einer n-zulässigen, diskret konvexen Menge (vgl. Beispiel 9.2.15), können wir also ebenfalls ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ angeben, welches wir hier explizit in folgendem Beispiel festhalten wollen.

Beispiel 9.2.19. Sei $M \subseteq \mathbb{Z} \times \mathbb{Z}$ eine diskret konvexe und n-zulässige Menge von Punkten mit $\#M \geq \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$. Dann ist die Abbildung $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_M} : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \times \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}$, definiert durch

$$\langle p, q \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} p(i,j) \cdot q(i,j),$$

ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. ◁

Haben wir ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ gegeben, ist es natürlich möglich, auch orthogonale oder sogar orthonormale Polynome $(p_1, \ldots, p_\nu)$ zu berechnen. Eine naive Möglichkeit besteht darin, aufbauend auf der Termbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ das Schmidtsche Orthonormalisierungsverfahren zu verwenden. Allerdings kann die Anwendung dieses Verfahrens numerisch instabil sein. Dennoch wollen wir im folgenden Beispiel das Gram-Schmidt-Verfahren verwenden, um erste Beispiele für Orthogonalbasen von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ kennenzulernen (vgl. in Ansätzen [Pis02], Beispiel 3.3.8, S. 58).

Beispiel 9.2.20. Wir betrachten nachfolgend die beiden hier abgebildeten Eindeutigkeitsmengen.

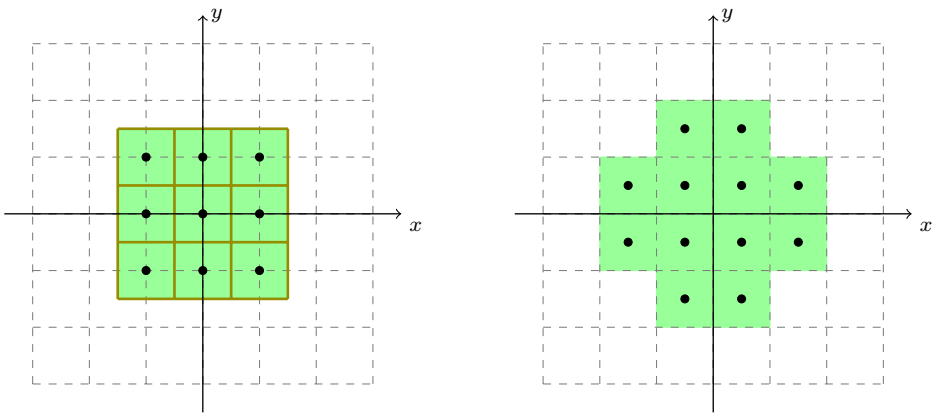


Abbildung 9.8: Darstellungen von n-Eindeutigsmengen.

a) Sei $M = \mathbb{Z}_{-1,1} \times \mathbb{Z}_{-1,1}$, also $M = G^1_{(0,0)}$ mit $\#M = 9$, die in Abbildung 9.8 links dargestellt ist. Sei $(p_{0,0}, p_{0,1}, p_{1,0}, p_{0,2}, p_{1,1}, p_{2,0})$ die Termbasis des $\mathbb{R}$-Vektorraums $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$, wobei für alle $i, j \in \mathbb{Z}_{0,2}$ die Funktion $p_{i,j} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ definiert ist durch $p_{i,j}(x, y) = x^i y^j$. Setze $q_1 := p_{0,0}$. Dann gilt

$$\langle q_1, q_1 \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} q_1(i,j) q_1(i,j) = \sum_{(i,j) \in M} 1 = 9$$

$$\langle q_1, p_{0,1} \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} q_1(i,j) p_{0,1}(i,j) = \sum_{(i,j) \in M} j = -3 + 3 = 0$$

und es folgt $q_2 := p_{0,1}$. Weiter gilt

$$\langle q_2, q_2 \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} p^2_{0,1}(i,j) = \sum_{(i,j) \in M} j^2 = 3 + 3 = 6$$

$$\langle q_1, p_{1,0} \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} p_{0,0}(i,j) p_{1,0}(i,j) = \sum_{(i,j) \in M} i = -3 + 3 = 0$$

$$\langle q_2, p_{1,0} \rangle_{\mathrm{Eval}_M} = \sum_{(i,j) \in M} p_{0,1}(i,j) p_{1,0}(i,j) = \sum_{(i,j) \in M} i \cdot j = 2 - 2 = 0$$

und es folgt

$$q_3 := p_{1,0} - \frac{\langle q_1, p_{1,0}\rangle_{\mathrm{Eval}_M}}{\langle q_1, q_1\rangle_{\mathrm{Eval}_M}} q_1 - \frac{\langle q_2, p_{1,0}\rangle_{\mathrm{Eval}_M}}{\langle q_2, q_2\rangle_{\mathrm{Eval}_M}} q_2 = p_{1,0}.$$

Nun gilt:

$$\begin{aligned}
\langle q_3, q_3\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{1,0}^2(i,j) = \sum_{(i,j)\in M} i^2 = 3+3 = 6\\
\langle q_1, p_{0,2}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,0}(i,j)p_{0,2}(i,j) = \sum_{(i,j)\in M} j^2 = 3+3 = 6\\
\langle q_2, p_{0,2}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,1}(i,j)p_{0,2}(i,j) = \sum_{(i,j)\in M} j^3 = -3+3 = 0\\
\langle q_3, p_{0,2}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{1,0}(i,j)p_{0,2}(i,j) = \sum_{(i,j)\in M} ij^2 = -2+2 = 0
\end{aligned}$$

und es folgt

$$\begin{aligned}
q_4 &:= p_{0,2} - \frac{\langle q_1, p_{0,2}\rangle_{\mathrm{Eval}_M}}{\langle q_1, q_1\rangle_{\mathrm{Eval}_M}} q_1 - \frac{\langle q_2, p_{0,2}\rangle_{\mathrm{Eval}_M}}{\langle q_2, q_2\rangle_{\mathrm{Eval}_M}} q_2 - \frac{\langle q_3, p_{0,2}\rangle_{\mathrm{Eval}_M}}{\langle q_3, q_3\rangle_{\mathrm{Eval}_M}} q_3\\
&= p_{0,2} - \frac{6}{9} p_{0,0} = p_{0,2} - \frac{2}{3}.
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
\langle q_4, q_4\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} \left(p_{0,2}(i,j) - \frac{2}{3}\right)^2 = \sum_{(i,j)\in M} j^4 - \sum_{(i,j)\in M} \frac{4}{3} j^2 + \sum_{(i,j)\in M} \frac{4}{9}\\
&= 6 - \frac{4}{3}\cdot 6 + \frac{4}{9}\cdot 9 = 2\\
\langle q_1, p_{1,1}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,0}(i,j)p_{1,1}(i,j) = \sum_{(i,j)\in M} ij = -2+2 = 0\\
\langle q_2, p_{1,1}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,1}(i,j)p_{1,1}(i,j) = \sum_{(i,j)\in M} ij^2 = -2+2 = 0\\
\langle q_3, p_{1,1}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{1,0}(i,j)p_{1,1}(i,j) = \sum_{(i,j)\in M} i^2 j = -2+2 = 0\\
\langle q_4, p_{1,1}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} \left(p_{0,2} - \frac{2}{3}\right)(i,j)p_{1,1}(i,j) = \sum_{i,j\in M} \left(ij^3 - \frac{2}{3}ij\right) = \frac{2}{3} - \frac{2}{3} = 0
\end{aligned}$$

und es folgt $q_5 := p_{1,1}$. Schließlich gilt

$$\begin{aligned}
\langle q_5, q_5\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{1,1}^2(i,j) = \sum_{(i,j)\in M} i^2 j^2 = 4\\
\langle q_1, p_{2,0}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,0}(i,j)p_{2,0}(i,j) = \sum_{(i,j)\in M} i^2 = 3+3 = 6\\
\langle q_2, p_{2,0}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{0,1}(i,j)p_{2,0}(i,j) = \sum_{(i,j)\in M} i^2 j = -2+2 = 0\\
\langle q_3, p_{2,0}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} p_{1,0}(i,j)p_{2,0}(i,j) = \sum_{(i,j)\in M} i^3 = -3+3 = 0\\
\langle q_4, p_{2,0}\rangle_{\mathrm{Eval}_M} &= \sum_{(i,j)\in M} \left(p_{0,2} - \frac{2}{3}\right)(i,j)p_{2,0}(i,j) = \sum_{i,j\in M} \left(i^2 j^2 - \frac{2}{3}i^2\right) = -\frac{4}{3} + \frac{4}{3} = 0
\end{aligned}$$

und es folgt zuletzt

$$q_6 := p_{2,0} - \frac{\langle q_1, p_{2,0}\rangle_{\mathrm{Eval}_M}}{\langle q_1, q_1\rangle_{\mathrm{Eval}_M}} q_1 = p_{2,0} - \frac{2}{3}.$$

Zusammengefasst bilden die folgenden sechs Polynomfunktionen eine Orthogonalbasis des $\mathbb{R}$-Vektorraums $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_M}$:

$$\begin{aligned} q_1(x,y) &= 1 & q_2(x,y) &= y & q_3(x,y) &= x \\ q_4(x,y) &= y^2 - \frac{2}{3} & q_5(x,y) &= xy & q_6(x,y) &= x^2 - \frac{2}{3} \end{aligned}$$

Auf analoge Weise erhält man mit dem Schmidtschen Orthonormalisierungsverfahren folgende Orthonormalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_M}$:

$$\begin{aligned} p_1(x,y) &= \frac{1}{3} & p_2(x,y) &= \frac{1}{\sqrt{6}} y & p_3(x,y) &= \frac{1}{\sqrt{6}} x \\ p_4(x,y) &= \frac{1}{\sqrt{2}} \left(y^2 - \frac{2}{3} \right) & p_5(x,y) &= \frac{1}{2} xy & p_6(x,y) &= \frac{1}{\sqrt{2}} \left(x^2 - \frac{2}{3} \right) \end{aligned}$$

b) Sei nun M die in Abbildung 9.8 rechts dargestellte Menge, also

$$\begin{aligned} M = \{ &\left(-\tfrac{1}{2}, -\tfrac{3}{2}\right), \left(\tfrac{1}{2}, -\tfrac{3}{2}\right), \left(-\tfrac{3}{2}, -\tfrac{1}{2}\right), \left(-\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{3}{2}, -\tfrac{1}{2}\right), \\ &\left(-\tfrac{3}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{3}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{3}{2}\right), \left(\tfrac{1}{2}, \tfrac{3}{2}\right) \} \end{aligned}$$

Auf dieser Menge mit 12 Elementen betrachten wir den $\mathbb{R}$-Vektorraum $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ mit Termbasis $(p_{0,0}, p_{0,1}, p_{1,0}, p_{0,2}, p_{1,1}, p_{2,0}, p_{0,3}, p_{1,2}, p_{2,1}, p_{3,0})$. Dann erhalten wir auf analoge Weise durch Anwendung des Schmidtschen Orthogonalisierungsverfahrens folgende Orthogonalbasis von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_M}$:

$$\begin{aligned} q_1(x,y) &= 1 & q_2(x,y) &= y \\ q_3(x,y) &= x & q_4(x,y) &= y^2 - \frac{11}{12} \\ q_5(x,y) &= xy & q_6(x,y) &= x^2 + \frac{1}{2} y^2 - \frac{11}{8} \\ q_7(x,y) &= y^3 - \frac{83}{44} y & q_8(x,y) &= xy^2 - \frac{19}{44} x \\ q_9(x,y) &= x^2 y + \frac{1}{2} y^3 - \frac{11}{8} y & q_{10}(x,y) &= x^3 + \frac{9}{10} xy^2 - \frac{91}{40} x \end{aligned}$$

Die folgenden Polynome bilden eine Orthonormalbasis von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_M}$:

$$\begin{aligned} q_1(x,y) &= \frac{1}{\sqrt{12}} & q_2(x,y) &= \frac{1}{\sqrt{11}} y \\ q_3(x,y) &= \frac{1}{\sqrt{11}} x & q_4(x,y) &= \sqrt{\frac{3}{32}} \left(y^2 - \frac{11}{12} \right) \\ q_5(x,y) &= \frac{2}{\sqrt{19}} xy & q_6(x,y) &= \frac{1}{\sqrt{8}} \left(x^2 + \frac{1}{2} y^2 - \frac{11}{8} \right) \\ q_7(x,y) &= \sqrt{\frac{11}{72}} \left(y^3 - \frac{83}{44} y \right) & q_8(x,y) &= \sqrt{\frac{11}{40}} \left(xy^2 - \frac{19}{44} x \right) \\ q_9(x,y) &= \frac{1}{\sqrt{2}} \left(x^2 y + \frac{1}{2} y^3 - \frac{11}{8} y \right) & q_{10}(x,y) &= \sqrt{\frac{5}{18}} \left(x^3 + \frac{9}{10} xy^2 - \frac{91}{40} x \right) \end{aligned}$$

◁

Wie man schnell erkennen kann, ist das Schmidtsche Orthogonalisierungsverfahren zusätzlich zu der fragwürdigen numerischen Stabilität auch recht aufwendig. Deshalb stellt PISINGER in [Pis02] ein effizientes und numerisch stabiles Rekursionsverfahren vor (vgl. [Pis02], S. 52 ff.), das ähnlich der 3-Term-Rekursion für univariate Polynomfunktionen funktioniert (vgl. [Fuc00]), aber auf das wir hier nicht näher eingehen wollen. Nachdem wir nun basierend auf den Auswertungsfunktionalen erste Skalarprodukte auf dem reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ kennen, wollen wir die Abbildung

$$\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m \quad \text{mit} \quad \Psi_M(p) = (\psi_{i_1,j_1}(p), \ldots, \psi_{i_m,j_m}(p))$$

auf einer n-zulässigen Menge M wieder ins Blickfeld rücken und zur Frage, wann Ψ_M ein Skalarprodukt induziert, zurückkehren. Dazu betrachten wir zunächst folgende Abbildung auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.

Lemma 9.2.21. *Sei $z \in \mathbb{R}^2$ und sei $T_z^* : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit $T_z^*(p) = p \circ T_z$.*

a) Die wohldefinierte Abbildung T_z^* ist linear und bijektiv. Insbesondere ist T_{-z}^* die Umkehrabbildung von T_z^*.

b) Für alle $w \in \mathbb{R}^2$ gilt $T_z^* \circ T_w^* = T_{z+w}^*$.

Beweis:

a) Da $p \circ T_z$ für alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ ebenfalls eine Polynomfunktion auf $\mathbb{R}^2$ vom Grad $\leq n$ ist, ist die Abbildung T_z^* wohldefiniert. Weiter gilt für alle $p, q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und alle $x \in \mathbb{R}^2$

$$\begin{aligned} T_z^*(p+q)(x) &= (p+q) \circ T_z(x) = (p+q)(x+z) = p(x+z) + q(x+z) \\ &= p \circ T_z(x) + q \circ T_z(x) = (p \circ T_z + q \circ T_z)(x) = (T_z^*(p) + T_z^*(q))(x) \end{aligned}$$

und für alle $\lambda \in \mathbb{R}$, $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und $x \in \mathbb{R}^2$:

$$T_z^*(\lambda p)(x) = (\lambda p) \circ T_z(x) = (\lambda p)(x+z) = \lambda p(x+z) \lambda (p \circ T_z(x)) = \lambda T_z^*(p)(x)$$

Somit ist T_z^* linear. Schließlich gilt

$$T_z^* \circ T_{-z}^*(p) = T_z^*(p \circ T_{-z}) = p \circ T_{-z} \circ T_z = p \circ T_z^{-1} \circ T_z = p$$

und analog $T_{-z}^* \circ T_z^*(p) = p$ für alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, d.h. T_z^* ist bijektiv mit Umkehrabbildung T_{-z}^*.

b) Sei $w \in \mathbb{R}^2$. Dann gilt für alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$:

$$T_z^* \circ T_w^*(p) = T_z^*(T_w^*(p)) = T_z^*(p \circ T_w) = p \circ T_w \circ T_z = p \circ T_{w+z} = T_{w+z}^*(p)$$

□

Diese lineare und bijektive Abbildung wird der Translationsoperator auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ genannt (vgl. [Pis02], Definition 3.3.1, S. 55).

Definition 9.2.22. (Translationsoperator)
Für $z \in \mathbb{R}^2$ heißt die lineare Funktion $T_z^* : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit $T_z^*(p) = p \circ T_z$ der **Translationsoperator** auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.

Mit diesem Translationsoperator können wir uns nun die Linearformen $\psi_{i,j}$, die Ψ_M definieren, genauer betrachten. Wir haben das Sensorarray als pixelperiodisch und homogen vorausgesetzt, d.h. für alle Sensoren $(i,j) \in S$ gilt $\alpha_{i,j} = \alpha_{0,0} \circ T^{-1}_{(i,j)}$ und alle Sensoren messen die gleiche Energieform. Somit folgt für alle $(i,j) \in S$ und alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$

$$\begin{aligned}\psi_{i,j}(p) &= \int_{A_{i,j}} p(x) d\alpha_{i,j}(x) = \int_{A_{i,j}} p(x) d\alpha_{0,0} \circ T^{-1}_{(i,j)}(x) \\ &= \int_{A_{i,j}} p(x) d\alpha_{0,0}(x-(i,j)) \overset{\tilde{x}:=x-(i,j)}{=} \int_{A_{0,0}} p(\tilde{x}+(i,j)) d\alpha_{0,0}(\tilde{x}) = \psi_{0,0} \circ T^*_{(i,j)}(p)\end{aligned}$$

Das bedeutet, dass sich die Linearform $\psi_{i,j}$ für alle $(i,j) \in S$ mit Hilfe des Translationsoperators auf die Linearform $\psi_{0,0}$ zurückführen lässt. Wir bezeichnen die Linearform $\psi_{0,0}$ im Folgenden nur kurz mit ψ. Dann gilt also für alle $(i,j) \in S$

$$\psi_{i,j} = \psi \circ T^*_{(i,j)}. \tag{9.2.6}$$

Die Linearform ψ ist dabei gegeben durch

$$\psi(p) = \int_{A_{0,0}} p(x) d\alpha_{0,0}(x) = \int_0^1 \int_0^1 p(x,y) dx dy;$$

sie ist eine sogenannte **strikt positive Linearform**, d.h. für alle $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit positivem Grad gilt $\psi(p) > 0$ (siehe dazu [Pis02], Abschnitt 2.4, S. 31 ff., insbesondere Satz 2.5.5, S. 38). Ist nun X eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$, so ist $\{\psi \circ T^*_x : x \in X\}$ ein Erzeugendensystem des Dualraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ (vgl. auch [Don09], Satz 2.4.11, S. 55). Das ist eine unmittelbare Folgerung aus dem nächsten Satz, der diese Tatsache in allgemeiner Form präsentiert und der wiederum auf Satz 9.2.17 basiert.

Satz 9.2.23. *Sei $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge bzgl. $\mathbb{R}^2$ und sei die Linearform $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ strikt positiv. Dann bilden $\varphi \circ T^*_{x_1}, \ldots, \varphi \circ T^*_{x_m}$ ein Erzeugendensystem von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$.*

Zusammen mit den Überlegungen vor diesem Satz folgt unter Verwendung von Satz 9.2.12 nun sofort, dass die Abbildung Ψ_M auf einer n-zulässigen Menge $M \subseteq \mathbb{Z} \times \mathbb{Z}$ ein Skalarprodukt induziert, womit insbesondere auch die Frage nach der Injektivität von Ψ_M beantwortet ist. Wir wollen dieses Ergebnis an dieser Stelle etwas allgemeiner in dem folgenden Korollar festhalten. In ähnlicher Form ist diese Aussage in [Pis02], Satz 3.3.3, S. 56 zu finden.

Korollar 9.2.24. *Sei $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge, sei $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ eine strikt positive Linearform und sei $\varphi_i = \varphi \circ T^*_{x_i}$ für alle $i \in \{1, \ldots, m\}$. Dann ist die lineare Abbildung $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ mit $\Phi(p) = (\varphi_1(p), \ldots, \varphi_m(p))$ injektiv und induziert damit durch*

$$\langle p, q \rangle_\Phi = \langle \Phi(p), \Phi(q) \rangle = \sum_{i=1}^m \varphi_i(p) \varphi_i(q),$$

ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.

Damit sind wir nun in der Lage, das lineare Ausgleichsproblem aus Gleichung (9.2.3) auf einer n-zulässigen Menge M mit $m \geq \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ Punkten durch Orthogonalentwicklung zu lösen. Denn in diesem Fall ist die lineare Abbildung $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ injektiv und sie induziert ein Skalarprodukt $\langle \cdot, \cdot \rangle_{\Psi_M}$ auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Die eindeutig bestimmte Lösung $p^{(M)} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ des linearen Ausgleichsproblems aus Gleichung (9.2.3) bezeichnen wir als

das **lokale Rekonstruktionspolynom** bzgl. M. Dadurch lassen sich lokale Bildbereiche eindeutig durch Polynomfunktionen beschreiben. Ist $B \subseteq \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des durch Ψ_M induzierten Skalarprodukts, so werden die Orthonormalkoeffizienten von $p^{(M)}$ bzgl. der Basis B als lokale Bildmerkmale bezeichnet (vgl. [Pis02], S. 72).

Definition 9.2.25. (Lokale Bildmerkmale)
Sei $\mathrm{gv} : S \to \mathbb{Z}_{0,255}$ ein digitales Grauwertbild, sei $M \subseteq S$ eine n-zulässige Menge und sei B eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\Psi_M}$. Der Koordinatenvektor des Rekonstruktionspolynoms $p^{(M)}$ bzgl. der Orthonormalbasis B heißt ein **lokales Bildmerkmal** oder **lokaler Merkmalsvektor** von gv auf M.

Analog ist es auch möglich, eine Orthogonalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu verwenden und die Orthogonalkoeffizienten als lokale Bildmerkmale zu verwenden. Ist M insbesondere eine in einem Pixel $(i_0, j_0) \in S$ zentrierte n-zulässige Menge, so wird der Merkmalsvektor $(\lambda_1, \ldots, \lambda_\nu) \in \mathbb{R}^\nu$ häufig um die Pixelkoordinaten erweitert zu einem Vektor $(\lambda_1, \ldots, \lambda_\nu, i_0, j_0) \in \mathbb{R}^{\nu+2}$. Damit wird der lokale Merkmalsvektor auch von der Notation her im Bild verortet und zudem werden die lokalen Merkmalsvektoren dadurch eindeutig. Dieser Vektor wird entsprechend auch als **erweiterter Merkmalsvektor** bezeichnet (vgl. [Pis02], Definition 4.1.8, S. 76). Im folgenden Kapitel werden lokale Bildmerkmale genauer analysiert, wobei besonders die Frage nach invarianten Größen der Koordinaten eines lokalen Bildmerkmals im Vordergrund stehen wird. Nebenbei werden wir allerdings noch sehen, dass sich lokale Bildmerkmale äußerst effektiv berechnen lassen. Das bisherige Ergebnis kann uns bzgl. der Effizienz noch nicht ganz zufrieden stellen. Schließlich ist für jede n-zulässige Menge $M \subseteq S$ auf diese Weise zuerst eine Orthonormalbasis bzgl. $\langle \cdot, \cdot \rangle_{\Psi_M}$ zu bestimmen, um die eindeutig bestimmte gesuchte Polynomfunktion $p^{(M)} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ angeben zu können, die das lineare Ausgleichsproblem aus Gleichung (9.2.3) löst. Im nächsten Abschnitt werden wir sehen, dass eine Berechnung von Orthonormalbasen für jeden lokalen Bildbereich M nicht notwendig ist. Zum Abschluss geben wir noch eine erste Interpretation lokaler Bildmerkmale an.

Lemma 9.2.26. *Sei $M \subseteq S$ eine n-zulässige Menge der Kardinalität m, sei $B = (b_1, \ldots, b_\nu)$ eine Orthogonalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\Psi_M}$ mit $b_1 = 1$ und sei $(a_1, \ldots, a_\nu) \in \mathbb{R}^\nu$ das zugehörige lokale Bildmerkmal. Dann gilt*

$$a_1 = \frac{1}{m} \sum_{(i,j) \in M} \mathrm{gv}_{i,j},$$

d.h. a_1 ist das arithmetische Mittel der Grauwerte aus M.

Beweis: Mit dem zu M gehörigen Grauwertvektor $\mathrm{gv}(M) \in \mathbb{R}^m$ gilt laut Satz 9.2.13

$$a_1 = \frac{\langle \mathrm{gv}(M), \Psi_M(b_1) \rangle}{\langle \Psi_M(b_1), \Psi_M(b_1) \rangle}.$$

Für alle $(i,j) \in M$ gilt $\psi_{i,j}(b_1) = 1$ und damit $\Psi_M(b_1) = (1, \ldots, 1) \in \mathbb{R}^m$. Somit folgt $\langle \Psi_M(b_1), \Psi_M(b_1) \rangle = m$ und $\langle \mathrm{gv}(M), \Psi_M(b_1) \rangle = \sum_{(i,j) \in M} \mathrm{gv}_{i,j}$, also schließlich die Behauptung. □

Dass das arithmetische Mittel der Grauwerte aus M mit dem Koeffizienten der konstanten Basisfunktion zusammenhängt, ist nicht allzu verwunderlich. Schließlich minimiert das arithmetische Mittel $\overline{\mathrm{gv}}$ der Grauwerte auf M die Summe

$$\sum_{(i,j) \in M} (c - \mathrm{gv}_{i,j})^2$$

und auch das Rekonstruktionspolynom $p^{(M)}$ minimiert eine Summe. Dabei handelt es sich um ein klassisches Problem der deskriptiven Statistik. Der Grauwertvektor $\mathrm{gv}(M)$ kann als **Merkmal** $\mathrm{gv} : M \to \mathbb{R}$ auf der Menge M von Individuen angesehen werden, das jedem Pixel in M seinen Grauwert $\mathrm{gv}_{i,j} \in \mathbb{Z}_{0,255}$ zuordnet. In dieser Sichtweise ist

$$\mathrm{Str}_2(\mathrm{gv}, c) := \sum_{(i,j)\in M} (c - \mathrm{gv}_{i,j})^2$$

die quadratische Streuung des Merkmals gv um c, die für $c = \overline{\mathrm{gv}}$ minimal wird, d.h. die quadratische Funktion $\mathrm{Str}_2(\mathrm{gv}, \cdot) : \mathbb{R} \to \mathbb{R}$ besitzt im arithmetischen Mittel $\overline{\mathrm{gv}}$ ein globales Minimum. Wählt man eine Orthonormalbasis bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\Psi_M}$, so sieht die Situation etwas anders aus. Dann stimmt a_1 nur bis auf einen Faktor mit dem arithmetischen Mittel überein.

Korollar 9.2.27. *Sei $M \subseteq S$ eine n-zulässige Menge der Kardinalität m, sei $B = (b_1, \ldots, b_\nu)$ eine Orthogonalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\Psi_M}$, wobei b_1 die konstante Basisfunktion sein soll, und sei $(a_1, \ldots, a_\nu) \in \mathbb{R}^\nu$ das zugehörige lokale Bildmerkmal. Dann gilt*

$$a_1 = \sqrt{m} \cdot \left(\frac{1}{m} \sum_{(i,j)\in M} \mathrm{gv}_{i,j} \right).$$

Beweis: Wegen $\langle \Psi_M(1), \Psi_M(1) \rangle = m$ gilt $b_1 = \frac{1}{\sqrt{m}}$, also gilt $\Psi_M(b_1) = \frac{1}{\sqrt{m}} \cdot (1, \ldots, 1) \in \mathbb{R}^m$. Somit folgt

$$a_1 = \langle \mathrm{gv}(M), \Psi_M(b_1) \rangle = \frac{1}{\sqrt{m}} \cdot \sum_{(i,j)\in M} \mathrm{gv}_{i,j} = \sqrt{m} \cdot \left(\frac{1}{m} \sum_{(i,j)\in M} \mathrm{gv}_{i,j} \right).$$

□

9.2.3 Effiziente Berechnung lokaler Bildmerkmale

In diesem Abschnitt werden wir sehen, dass sich lokale Bildmerkmale äußerst effektiv und einfach berechnen lassen. Wir betrachten also wie bisher eine n-zulässige Menge $M \subseteq S$ mit $m \geq \nu$ Elementen, wobei $\nu \in \mathbb{N}_+$ die Dimension $\dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ bezeichne. Bisher ist es notwendig, für jeden lokalen Bildbereich $M \subseteq S$ zunächst eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des durch Ψ_M induzierten Skalarprodukts zu bestimmen; ein Aufwand, der unnötig ist, wie wir bald sehen werden. Denn wir wollen nun zunächst darauf eingehen, wie man das lineare Ausgleichsproblem Gleichung (9.2.3) effektiv lösen kann, ohne für jedes $M \subseteq S$ zuvor eine Orthonormalbasis berechnen zu müssen. Laut Korollar 9.2.24 ist die lineare Abbildung $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$, definiert durch

$$p \mapsto (\varphi_i(p), \ldots, \varphi_m(p))$$

injektiv und induziert ein Skalarprodukt auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, falls für alle $i \in \{1, \ldots, m\}$ die Linearform $\varphi_i \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ gegeben ist durch $\varphi_i = \varphi \circ T^*_{x_i}$ für eine strikt positive Linearform $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ und eine n-Eindeutigkeitsmenge $X = \{x_1, \ldots, x_m\}$. Der nächste Satz stellt nun einen Zusammenhang her zwischen dem durch Auswertungsfunktionale induzierten Skalarprodukt (siehe Korollar 9.2.18) und dem durch Φ induzierten Skalarprodukt bzw. zwischen den jeweiligen Orthogonal- bzw. Orthonormalbasen (vgl. [Pis02], Satz 3.3.6, S. 58). Insbesondere zeigt der Satz, dass die Bilder von Eval_X und Φ identisch sind (vgl. [Pis02], Satz 3.3.5, S. 57).

Satz 9.2.28. *Sei* $\nu := \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ *und sei* $X := \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ *für* $m \geq \nu$ *eine* n*-Eindeutigkeitsmenge bzgl.* $\mathbb{R}^2$*. Sei* $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ *eine strikt positive Linearform und sei* $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ *definiert durch* $\Phi(p) = (\varphi_1(p), \ldots, \varphi_m(p))$*, wobei* $\varphi_i = \varphi \circ T^*_{x_i}$ *für alle* $i \in \{1, \ldots, m\}$ *gilt.*

a) Dann gilt $\mathrm{Im}(\mathrm{Eval}_X) = \mathrm{Im}(\Phi)$*, d.h. insbesondere ist* $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathrm{Im}(\mathrm{Eval}_X)$ *bijektiv.*

b) Ist $(q_1, \ldots, q_\nu)$ *eine Orthogonal- bzw. Orthonormalbasis von* $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ *bzgl.* $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X}$*, so ist* $(p_1, \ldots, p_\nu)$ *mit*

$$p_i := \Phi^{-1}\left(\mathrm{Eval}_X(q_i)\right)$$

für alle $i \in \{1, \ldots, \nu\}$ *eine Orthogonal- bzw. Orthonormalbasis von* $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ *bzgl. des durch* Φ *induzierten Skalarprodukts* $\langle \cdot, \cdot \rangle_\Phi$*.*

Da Orthogonal- bzw. Orthonormalbasen von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X}$ auf n-Eindeutigkeitsmengen $X \subseteq \mathbb{R}^2$ besonders einfach und Dank eines Rekursionsverfahrens auch numerisch stabil berechnet werden können, eröffnet dieser Satz eine Möglichkeit, auf dieser gegebenen Orthogonal- bzw. Orthonormalbasis eine Orthogonal- bzw. Orthonormalbasis bzgl. ein durch eine strikt positive Linearform φ induziertes Skalarprodukt zu berechnen. Ist wie in der Situation des letzten Satzes $(q_1, \ldots, q_\nu)$ eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X}$, so lässt sich die Orthonormalbasis $(p_1, \ldots, p_\nu)$ von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_\Phi$ numerisch stabil (ohne Verwendung des Schmidtschen Orthogonalisierungsverfahrens) wie folgt bestimmen (angelehnt an [Pis02], Bemerkung 3.3.7, S. 58):

(1) Wähle eine beliebige Basis $B := (b_1, \ldots, b_\nu)$ von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, z.B. die Termbasis.

(2) Bilde die Darstellungsmatrix $\mathcal{M}_E^B(\Phi) = (\varphi_i(b_j))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq \nu}} \in \mathrm{Mat}_{m,\nu}(\mathbb{R})$ von Φ, wobei E die kanonische Basis im $\mathbb{R}^m$ sei. Da X eine n-Eindeutigkeitsmenge ist, gilt $m \geq \nu$.

(3) Betrachte für alle $i \in \{1, \ldots, \nu\}$ das lineare Gleichungssystem

$$\mathcal{M}_E^B(\Phi) \cdot \begin{pmatrix} \lambda_{i,1} \\ \vdots \\ \lambda_{i,\nu} \end{pmatrix} = \begin{pmatrix} q_i(x_1) \\ \vdots \\ q_i(x_m) \end{pmatrix}$$

Wegen $\mathrm{Rang}(\mathcal{M}_E^B(\Phi)) = \dim_{\mathbb{R}}(\mathrm{Im}(\Phi)) = \dim_{\mathbb{R}}(\mathrm{Im}(\mathrm{Eval}_X)) = \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})) = \nu$ ist dieses Gleichungssystem eindeutig lösbar. Mit dem Lösungsvektor $(\lambda_{i,1}, \ldots, \lambda_{i,\nu}) \in \mathbb{R}^\nu$ folgt

$$p_i = \lambda_{i,1} b_1 + \ldots + \lambda_{i,\nu} b_\nu.$$

Es ist also möglich, aufbauend auf einer Orthogonal- bzw. Orthonormalbasis bzgl. $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X}$, die einfach und numerisch stabil berechnet werden kann, mit Hilfe einfacher linearer Algebra und numerisch stabiler mehrdimensionaler Integration eine Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_\Phi$ zu berechnen. Dazu wollen wir zwei konkrete Beispiele angeben, auf denen wir später aufbauen werden (vgl. auch [Pis02], Beispiel 3.3.8, S. 59 f.). In diesen Beispielen verwenden wir im Nullpunkt zentrierte n-Eindeutigkeitsmengen X (siehe auch Abbildung 9.9). Wir werden später verstehen, wofür wir genau diese benötigen, die Dank der Zentrierung im Nullpunkt verhältnismäßig einfach aufgebaut sind.

Beispiel 9.2.29. Sei die Linearform $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ definiert durch

$$\varphi(p) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} p(x, y) dx dy.$$

Dann ist φ strikt positiv.

a) Sei $X = \mathbb{Z}_{-1,1} \times \mathbb{Z}_{-1,1} = G^1_{(0,0)}$ und $n = 2$. Sei weiter für alle $(i,j) \in X$ die Linearformen $\varphi_{i,j} \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ gegeben durch $\varphi_{i,j} = \varphi \circ T^*_{i,j}$, also definiert durch

$$\varphi_{i,j}(p) = \int_{j-\frac{1}{2}}^{j+\frac{1}{2}} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} p(x,y)dxdy.$$

Dann sind die Eckpunkte der quadratischen Integrationsbereiche der Linearformen $\varphi_{i,j}$ genau die Schwerpunkte der quadratischen Gitterkästchen (siehe auch Abbildung 9.9). Sei weiter $\Phi : \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \to \mathrm{Abb}(X, \mathbb{R})$ die durch $\Phi(p) = (\varphi_{i,j}(p))_{(i,j)\in X}$ definierte und zu dem im Nullpunkt zentrierten Gitter X gehörige lineare Abbildung. Dann induziert Φ laut Korollar 9.2.24 ein Skalarprodukt $\langle \cdot, \cdot \rangle_\Phi : \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \times \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}$ durch

$$\langle p, q \rangle_\Phi = \sum_{(i,j)\in M} \varphi_{i,j}(p)\varphi_{i,j}(q).$$

Bzgl. dieses Skalarprodukts bilden die Polynomfunktionen

$$\begin{aligned}
&p_{0,0}(x,y) = 1 && p_{0,1}(x,y) = y && p_{1,0}(x,y) = x \\
&p_{0,2}(x,y) = y^2 - \frac{3}{4} && p_{1,1}(x,y) = xy && p_{2,0}(x,y) = x^2 - \frac{3}{4}
\end{aligned}$$

eine Orthogonalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ und die Funktionen

$$\begin{aligned}
&\tilde{p}_{0,0}(x,y) = \frac{1}{3} && \tilde{p}_{0,1}(x,y) = \frac{1}{\sqrt{6}}y && \tilde{p}_{1,0}(x,y) = \frac{1}{\sqrt{6}}x \\
&\tilde{p}_{0,2}(x,y) = \frac{1}{\sqrt{2}}\left(y^2 - \frac{3}{4}\right) && \tilde{p}_{1,1}(x,y) = \frac{1}{2}xy && \tilde{p}_{2,0}(x,y) = \frac{1}{\sqrt{2}}\left(x^2 - \frac{3}{4}\right)
\end{aligned}$$

eine Orthonormalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$.

b) Sei nun $n = 3$ und

$$\begin{aligned}
X = \{&\left(-\tfrac{1}{2}, -\tfrac{3}{2}\right), \left(\tfrac{1}{2}, -\tfrac{3}{2}\right), \left(-\tfrac{3}{2}, -\tfrac{1}{2}\right), \left(-\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{3}{2}, -\tfrac{1}{2}\right), \\
&\left(-\tfrac{3}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{3}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{3}{2}\right), \left(\tfrac{1}{2}, \tfrac{3}{2}\right)\}
\end{aligned}$$

Dann ist X eine 3-Eindeutigkeitsmenge, die in Abbildung 9.9 unten dargestellt ist. Auch hier sind für alle $(i,j) \in X$ die Linearformen $\varphi_{i,j} \in \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ durch

$$\varphi_{i,j}(p) = \int_{j-\frac{1}{2}}^{j+\frac{1}{2}} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} p(x,y)dxdy$$

gegeben. Nun stimmen die Integrationsbereiche genau mit den Gitterkästchen überein. Auf analoge Weise induziert $\Phi : \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R}) \to \mathrm{Abb}(X, \mathbb{R})$ mit $\Phi(p) = (\varphi_{i,j}(p))_{(i,j)\in X}$ ein Skalarprodukt auf $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$. Bzgl. dieses Skalarprodukts bilden die folgenden Polynome eine Orthonormalbasis von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$:

$$\begin{aligned}
&p_1(x,y) = \tfrac{1}{6}\sqrt{3} && p_2(x,y) = \tfrac{1}{\sqrt{11}}y \\
&p_3(x,y) = \tfrac{1}{\sqrt{11}}x && p_4(x,y) = \tfrac{1}{4(1+\sqrt{3})}(y^2 + (2+\sqrt{3})x^2 - 3 - \sqrt{3}) \\
&p_5(x,y) = \tfrac{2}{\sqrt{19}}xy && p_6(x,y) = \tfrac{1}{4(1+\sqrt{3})}(x^2 + (2+\sqrt{3})y^2 - 3 - \sqrt{3}) \\
&p_7(x,y) = \tfrac{\sqrt{10}}{6}\left(y^3 - \frac{13}{5}y + \frac{9}{10}x^2y\right) && p_8(x,y) = \tfrac{\sqrt{110}}{20}\left(y^2x - \frac{17}{33}x\right) \\
&p_9(x,y) = \tfrac{\sqrt{110}}{20}\left(x^2y - \frac{17}{33}y\right) && p_{10}(x,y) = \tfrac{\sqrt{10}}{6}\left(x^3 - \frac{13}{5}x + \frac{9}{10}xy^2\right)
\end{aligned}$$

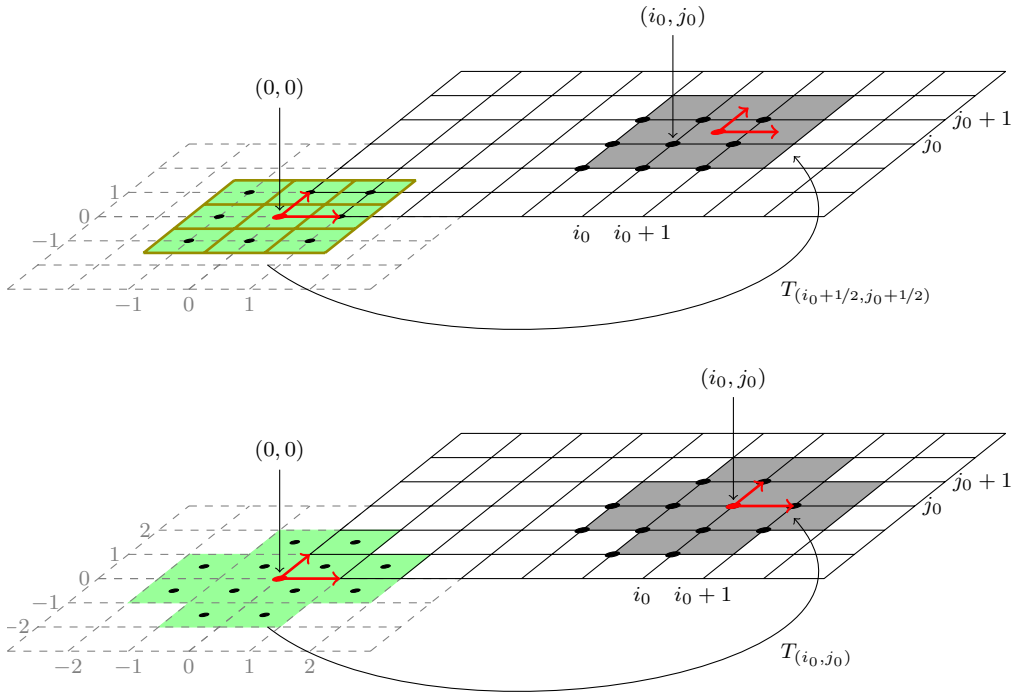


Abbildung 9.9: Darstellung der Rückführung eines Lokalisierungsfensters auf ein im Nullpunkt zentriertes Fenster (grün).

◁

Für die Berechnung der Orthogonalentwicklung aus Satz 9.2.13 ist es insbesondere notwendig, die Bilder $\Phi(p_i)$ für alle Basispolynome p_i zu berechnen. Dank des letzten Satzes ist dies wegen $\Phi(p_i) = \mathrm{Eval}_X(p_i)$ durch einfach Polynomauswertung möglich.

Beispiel 9.2.30. In der Situation von Beispiel 9.2.29 a) liefert für alle i, j die Linearform $\varphi_{i,j}$ das Integral einer Funktion $p \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ über dem Rechteck $[i - \frac{1}{2}, i + \frac{1}{2}] \times [j - \frac{1}{2}, j + \frac{1}{2}]$ (siehe auch Abbildung 9.9). Sei $k, \ell \in \mathbb{N}$ mit $k + \ell \leq 2$. Sei weiter $p_{k,\ell}$ ein Element einer Orthogonalbasis bzgl. $\langle \cdot, \cdot \rangle_\Phi$ und $q_{k,\ell}$ ein Element einer Orthogonalbasis bzgl. $\langle \cdot, \cdot \rangle_{\mathrm{Eval}_X}$ (siehe Beispiel 9.2.20). Dann gilt:

$$\varphi_{i,j}(p_{k,\ell}) = q_{k,\ell}(i, j),$$

d.h. konkret gilt beispielsweise mit $p_{0,2}(x, y) = y^2 - \frac{3}{4}$ und $q_{0,2}(x, y) = y^2 - \frac{2}{3}$:

$$\varphi_{-1,1}(p_{0,2}) = \int_{\frac{1}{2}}^{\frac{3}{2}} \int_{-\frac{3}{2}}^{-\frac{1}{2}} (y^2 - \tfrac{3}{4}) dx dy = q_{0,2}(-1, 1) = 1 - \tfrac{2}{3} = \tfrac{1}{3}.$$

◁

Wir haben nun also insbesondere Beispiele gesehen von Skalarprodukten und orthogonalen bzw. orthonormalen Polynomen auf im Nullpunkt zentrierten Gittern. Allerdings benötigen wir Skalarprodukte und orthogonale bzw. orthonormale Polynome auf Lokalisierungsfenstern, die zu n-zulässigen Mengen $M \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ gehören, die also sozusagen „im Bild" liegen. In der Tat ist es so, dass die Kenntnis einer Orthogonal- oder Orthonormalbasis auf Nullpunkt zentrierten Mengen völlig ausreicht, um beliebige lokale Fenster, die dieselbe Struktur haben wie

diejenigen, die Nullpunkt-zentriert sind, innerhalb des Bildes zu betrachten. Diese lassen sich einfach verschieben, wobei die Verschiebung sich in der Orthogonal- bzw. Orthonormalbasis widerspiegelt. Abbildung 9.9 stellt diesen Sachverhalt anschaulich dar. Zunächst ist klar, dass wir aus einer injektiven Abbildung $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ durch den bijektiven Translationsoperator T_u^* wieder eine injektive Abbildung $\Phi \circ T_u^*$ erhalten, die erneut ein Skalarprodukt induziert. Sind also $\gamma_1, \ldots, \gamma_m \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ die Linearformen mit $\Gamma(p) = (\gamma_1(p), \ldots, \gamma_m(p))$, so gilt für alle $i \in \{1, \ldots, m\}$:

$$\gamma_i = \varphi_i \circ T_u^* = \varphi \circ T_{x_i}^* \circ T_u^* = \varphi \circ T_{x_i+u}^*,$$

d.h. auch Γ beruht auf der strikt positiven Linearform φ. Die Punkte $x_1 + u, \ldots, x_m + u$ bilden ebenfalls eine n-Eindeutigkeitsmenge derselben Struktur (vgl. auch Lemma 9.2.16). Der nächste Satz stellt uns den einfachen Zusammenhang zwischen den jeweiligen Orthogonal- bzw. Orthonormalbasen dar.

Satz 9.2.31. (Orthogonal- bzw. Orthonormalbasen bei Translationen)
Sei $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge und sei $\varphi \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})^*$ eine strikt positive Linearform. Sei $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ definiert durch $\Phi(p) = (\varphi_1(p), \ldots, \varphi_m(p))$ mit Linearformen $\varphi_i = \varphi \circ T_{x_i}^*$ für alle $i \in \{1, \ldots, m\}$ und sei $(q_1, \ldots, q_\nu)$ eine Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_\Phi$. Dann ist $(p_1, \ldots, p_\nu)$ mit

$$p_i := T_{-u}^*(q_i) = q_i \circ T_{-u} \quad \text{für alle } i \in \{1, \ldots, \nu\}$$

für alle $u \in \mathbb{R}^2$ eine Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des durch die lineare Abbildung $\Gamma := \Phi \circ T_u^*$ induzierten Skalarprodukts $\langle \cdot, \cdot \rangle_\Gamma$.

Beweis: Sei $u \in \mathbb{R}^2$. Wegen der Bijektivität des Translationsoperators (vgl. Lemma 9.2.21) ist klar, dass $(p_1, \ldots, p_\nu)$ eine Basis des $\mathbb{R}$-Vektorraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ ist. Weiter gilt für alle $i, j \in \{1, \ldots, \nu\}$

$$\begin{aligned}
\langle p_i, p_j \rangle_\Gamma &= \langle \Gamma(p_i), \Gamma(p_j) \rangle = \langle \Phi(T_u^*(p_i)), \Phi(T_u^*(p_j)) \rangle = \sum_{k=1}^{m} \varphi_k(T_u^*(p_i)) \cdot \varphi_k(T_u^*(p_j)) \\
&= \sum_{k=1}^{m} \varphi_k(T_u^*(T_{-u}^*(q_i))) \cdot \varphi_k(T_u^*(T_{-u}^*(q_j))) = \sum_{k=1}^{m} \varphi_k(q_i) \cdot \varphi_k(q_j) = \langle \Phi(q_i), \Phi(q_j) \rangle \\
&= \langle q_i, q_j \rangle_\Phi
\end{aligned}$$

Da $(q_1, \ldots, q_\nu)$ eine Orthogonal- bzw. Orthonormalbasis bzgl. $\langle \cdot, \cdot \rangle_\Phi$ ist, folgt somit, dass auch $(p_1, \ldots, p_\nu)$ eine Orthogonal- bzw. Orthonormalbasis bzgl. $\langle \cdot, \cdot \rangle_\Gamma$ ist. □

Dieser Satz liefert uns nun den Schlüssel dazu, wie man auf beliebigen Lokalisierungsfenstern ohne großen Aufwand eine Orthogonal- bzw. Orthonormalbasis angeben kann, und zwar alleine durch Rückführung auf eine bekannte Basis auf beispielsweise einem im Nullpunkt zentrierten Lokalisierungsfenster (siehe auch Abbildung 9.9). Im nächsten Beispiel soll dies demonstriert werden.

Beispiel 9.2.32. Sei $M_0 := G_{(0,0)}^1$. Aus Beispiel 9.2.29 ist uns eine Orthogonal- bzw. Orthonormalbasis auf M_0 bekannt bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_\Phi$ (vgl. Beispiel 9.2.29), wobei $\Phi : \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^9$ definiert ist durch $\Phi(p) = (\varphi_{i,j}(p))_{(i,j) \in M_0}$ für die strikt positive Linearform $\varphi \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})^*$ gegeben durch

$$\varphi(p) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} p(x, y) dx dy$$

und $\varphi_{i,j} = \varphi \circ T^*_{(i,j)}$ für alle $(i,j) \in M_0$. Die zugehörige Orthogonalbasis lautet wie folgt:

$$q_1(x,y) = 1 \qquad q_2(x,y) = y \qquad q_3(x,y) = x$$
$$q_4(x,y) = y^2 - \frac{3}{4} \qquad q_5(x,y) = xy \qquad q_6(x,y) = x^2 - \frac{3}{4}$$

Sei nun $(i_0, j_0) \in \mathbb{Z} \times \mathbb{Z}$ so, dass $M := G^1_{(i_0,j_0)} \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ gilt. In Abbildung 9.9 sind oben links durch die Punkte die Elemente von M_0 sowie oben rechts die Elemente von M gekennzeichnet. Setze $Y := [-\frac{3}{2}, \frac{3}{2}] \times [-\frac{3}{2}, \frac{3}{2}]$, d.h. Y ist der grün schraffierte Bereich in Abbildung 9.9. Dann gilt $M = T_{(i_0,j_0)}(M_0)$ und $\mathrm{Loc}_M = T_u(Y)$ mit $u := (i_0 + \frac{1}{2}, j_0 + \frac{1}{2})$. Weiter ist die Linearform $\psi \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})^*$ bekanntlich definiert durch

$$\psi(p) = \int_0^1 \int_0^1 p(x,y)dxdy$$

und mit $\psi_{i,j} = \psi \circ T^*_{(i,j)}$ für alle $(i,j) \in M$ ist $\Psi_M : \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^9$ definiert durch $\Psi_M(p) = (\psi_{i,j}(p))_{(i,j)\in M}$. Es gilt $\Psi_M = \Phi \circ T^*_u$ und der letzte Satz zeigt uns, dass wir auf folgende Weise aus der Orthogonalbasis $(q_1, \ldots, q_6)$ bzgl. $\langle \cdot, \cdot \rangle_\Phi$ folgende Orthogonalbasis bzgl. des Skalarprodukts $\langle \cdot, \cdot \rangle_{\Psi_M}$ erhalten:

$$\begin{aligned}
p_1(x,y) &= q_1 \circ T_{-u}(x,y) = 1 \\
p_2(x,y) &= q_2 \circ T_{-u}(x,y) = q_2((x,y) - u) = y - (j_0 + \tfrac{1}{2}) \\
p_3(x,y) &= q_3 \circ T_{-u}(x,y) = q_3((x,y) - u) = x - (i_0 + \tfrac{1}{2}) \\
p_4(x,y) &= q_4 \circ T_{-u}(x,y) = q_4((x,y) - u) = \big(y - (j_0 + \tfrac{1}{2})\big)^2 - \tfrac{3}{4} \\
&= y^2 - 2(j_0 + \tfrac{1}{2})y + (j_0 + \tfrac{1}{2})^2 - \tfrac{3}{4} \\
p_5(x,y) &= q_5 \circ T_{-u}(x,y) = q_5((x,y) - u) = \big(x - (i_0 + \tfrac{1}{2})\big) \cdot \big(y - (j_0 + \tfrac{1}{2}\big) \\
&= xy - (j_0 + \tfrac{1}{2})x - (i_0 + \tfrac{1}{2})y + (i_0 + \tfrac{1}{2})(j_0 + \tfrac{1}{2}) \\
p_6(x,y) &= q_6 \circ T_{-u}(x,y) = q_6((x,y) - u) = \big(x - (i_0 + \tfrac{1}{2})\big)^2 - \tfrac{3}{4} \\
&= x^2 - 2(i_0 + \tfrac{1}{2})x + (i_0 + \tfrac{1}{2})^2 - \tfrac{3}{4}
\end{aligned}$$

Analog lässt sich eine Orthonormalbasis auf M bestimmen. ◁

Da sich also Orthonormalbasen für n-zulässige, diskret konvexe Mengen $M \subseteq S$ auf einfache Weise finden lassen, ist es auch kein Problem das Rekonstruktionspolynom $p^{(M)} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und damit das zu M gehörige lokale Bildmerkmal zu bestimmen. Der letzte Satz führt uns auch sofort zur Translationsinvarianz lokaler Bildmerkmale, was eine unmittelbare Folgerung aus dem nächsten Satz ist.

Satz 9.2.33. (Translationsinvarianz)
Sei $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^2$ eine n-Eindeutigkeitsmenge und sei $\Phi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ wie in Satz 9.2.31. Sei $B := (q_1, \ldots, q_\nu)$ eine Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. $\langle \cdot, \cdot \rangle_\Phi$ und sei $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Dann sind für alle $u \in \mathbb{R}^2$ die Koordinatenvektoren von p (bgzl. B) und von $T^*_{-u}(p)$ (bzgl. C) identisch, wobei $C := (p_1, \ldots, p_\nu)$ mit $p_i := T^*_{-u}(q_i)$ für alle $i \in \{1, \ldots, \nu\}$ eine Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. des durch $\Gamma := \Phi \circ T^*_u$ induzierten Skalarprodukts $\langle \cdot, \cdot \rangle_\Gamma$ ist.

Beweis: Sei $u \in \mathbb{R}^2$ und sei $(a_1, \ldots, a_\nu) \in \mathbb{R}^\nu$ der Koordinatenvektor von p bzgl. der Orthonormalbasis $B = (q_1, \ldots, q_\nu)$. Dann gilt:

$$\begin{aligned}
T^*_{-u}(p) &= T^*_{-u}(a_1 q_1 + \ldots + a_\nu q_\nu) = a_1 T^*_{-u}(q_1) + \ldots + a_\nu T^*_{-u}(q_\nu) \\
&= a_1 p_1 + \ldots + a_\nu p_\nu,
\end{aligned}$$

d.h. die Polynomfunktion $T_u^*(p)$ hat bzgl. C ebenfalls den Koordinatenvektor $(a_1, \ldots, a_\nu)$. $\square$

Betrachten wir abschließend speziell eine n-zulässige, diskret konvexe Menge $M \subseteq S$ der Kardinalität m und die darauf aufbauende lineare Abbildung $\Psi_M : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$.

Bemerkung 9.2.34. Das Rekonstruktionspolynom $p^{(M)} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ der zeitintegrierte Sensorinputfunktion auf M hat denselben Koordinatenvektor bzgl. einer Orthogonal- bzw. Orthonormalbasis auf M wie das translatorische Bild $T^*_{-u}(p^{(M)})$ bzgl. einer Orthogonal- bzw. Orthonormalbasis auf $T_u(M)$. Somit sind lokale Bildmerkmale per Konstruktion invariant unter Translationen. Es ist damit also möglich, lokale Bildmerkmale stets bzgl. einer einzigen festen Orthogonal- bzw. Orthonormalbasis, die zu einer im Nullpunkt zentrierten n-Eindeutigkeitsmenge gehört, zu berechnen. Die Berechnung der Koeffizienten selbst ist effizient durch Auswertung geeigneter Polynomfunktionen machbar (vgl. Beispiel 9.2.30).

KAPITEL 10

Invarianten von Polynomfunktionen

Jane Austen[24]

Wieso gönnen wir uns den Genuss nicht sofort? Wie oft wird die Freude durch Vorbereitungen verdorben, durch törichte Vorbereitungen!

Leider ist es zum jetzigen Zeitpunkt noch nicht ganz möglich, dass wir uns dem „Genuss" hingeben, mit dem finalen Algorithmus des letzten Kapitels zu arbeiten. Dieser Algorithmus (siehe Algorithmus 11.1) benötigt als Eingabe fundamentale Invarianten ganz spezieller Invariantenringe, denen wir uns in diesem Kapitel widmen wollen. Somit ist noch etwas Vorbereitung zu leisten, die in diesem Fall aber alles andere als „töricht" ist, sondern essentiell für Algorithmus 11.1. In diesem Algorithmus spielen die lokalen Bildmerkmale, die wir im letzten Kapitel kennengelernt haben, die zentrale Rolle. Dabei handelt es sich um Koordinaten von Polynomfunktionen bzgl. spezieller Orthonormalbasen. Und damit ist das weitere Vorgehen und das Ziel dieses Kapitels schnell klar: Wir benötigen fundamentale Invarianten von Polynomfunktionen unter linearen algebraischen Gruppen. Dazu beschränken wir uns auf die Betrachtung linear reduktiver Untergruppen G der allgemeinen linearen Gruppe $\mathrm{GL}_2(\mathbb{R})$ und geben explizit fundamentale Invarianten für die spezielle orthogonale Gruppe $\mathrm{SO}_2(\mathbb{R})$ und die orthogonale Gruppe $\mathrm{O}_2(\mathbb{R})$ an, da diese für die Praxis die größte Bedeutung besitzen. Dabei belassen wir es allerdings bei Polynomfunktionen vom Grad ≤ 3, da diese in den Beispielen des nächsten Kapitels Verwendung finden.

Die allgemeine lineare Gruppe $\mathrm{GL}_2(\mathbb{R})$, und damit jede Untergruppen G, operiert auf dem reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ durch die Abbildung $\tau : G \times \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, die definiert ist durch $(\mathcal{A}, p) \mapsto p \circ T_{\mathcal{A}}$, wobei $T_{\mathcal{A}} : \mathbb{R}^2 \to \mathbb{R}^2$ die durch $\mathcal{A} \in G$ induzierte lineare Abbildung $x \mapsto \mathcal{A} \cdot x$ bezeichnet. Diese Gruppenoperation induziert bekanntlich einen Gruppenhomomorphismus $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$, definiert durch

$$\mathcal{A} \mapsto (p \mapsto p \circ T_{\mathcal{A}}),$$

die **rationale Darstellung** von G in $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ (siehe Definition 4.3.18). Auch hier schreiben wir zur besseren Lesbarkeit $\rho_{\mathcal{A}}$ für den Automorphismus $\rho(\mathcal{A}) \in \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$. Nun müssen wir uns mit der Operation von G auf den Koordinatenring $\mathbb{R}[V]$ von $V = \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$

[24] Bildquelle: https://de.wikipedia.org/wiki/Jane_Austen vom 14.09.2015.

beschäftigen. Die Operation der Gruppe G lässt sich auf den Koordinatenring $\mathbb{R}[V]$ fortsetzen, genauer ist durch $\mathcal{A} \mapsto (f \mapsto f^{\mathcal{A}})$ eine rationale Darstellung $\tilde{\rho} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathbb{R}[V])$ von G in $\mathbb{R}[V]$ gegeben. Dabei ist bekanntlich $f^{\mathcal{A}} \in \mathbb{R}[V]$ für alle $f \in \mathbb{R}[V]$ die durch $f^{\mathcal{A}}(p) = f(p \circ T_{\mathcal{A}^{-1}})$ definierte reguläre Funktion auf V. Wir betrachten die Situation nun etwas konkreter, indem wir den Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zunächst in diesem Kapitel stets mit der Standardtermbasis

$$B := (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0}, \ldots, b_{n,0})$$

versehen, wobei für alle $i, j \in \mathbb{N}$ mit $i + j \leq n$ die Basisfunktion $b_{i,j} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ definiert ist durch $b_{i,j}(x, y) = x^i y^j$. Wir verzichten der Einfachheit halber also auf die Betrachtung anhand einer Orthonormalbasis. Offensichtlich stellt dies aber keine Einschränkung dar, denn lokale Bildmerkmale lassen sich problemlos in Koordinatenvektoren bzgl. der Standardtermbasis umrechnen. Sei weiter $\kappa_B : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^\nu$ mit $\nu = \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ die Koordinatenfunktion auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. der Basis B. Die Koordinaten $\kappa_B(p) \in \mathbb{R}^\nu$ bezeichnen wir analog zur Indizierung der Basis B mit $(a_{0,0}, a_{0,1}, a_{1,0}, \ldots, a_{n,0}) \in \mathbb{R}^\nu$, d.h. $a_{i,j} \in \mathbb{R}$ ist die Koordinate von p bzgl. der Basisfunktion $b_{i,j}$. Dann ist der Koordinatenring $\mathbb{R}[V]$ isomorph zum Polynomring

$$P := \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}, a_{0,2}, a_{1,1}, a_{2,0}, \ldots, a_{n,0}]$$

und die rationale Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ induziert eine zu $\tilde{\rho}$ äquivalente Darstellung $\widehat{\rho} : G \to \mathrm{Aut}_{\mathbb{R}}(P)$, die ebenfalls durch $\mathcal{A} \mapsto (f \mapsto f^{\mathcal{A}})$ definiert ist. Nun ist aber für $f \in P$ das Polynom $f^{\mathcal{A}} \in P$ gegeben durch $f^{\mathcal{A}} = f(\kappa_B(p \circ T_{\mathcal{A}^{-1}}))$. Um nun alle invarianten Größen unter G angeben zu können, genügt es, die fundamentalen Invarianten des Invariantenrings (siehe auch Definition 6.2.1)

$$P^G = \{f \in P : \ f^{\mathcal{A}} = f \text{ für alle } \mathcal{A} \in G\}$$

zu berechnen. Die Existenz eines endlichen Algebra-Erzeugendensystem von P^G folgt in diesem Fall sofort aus dem Hilbertschen Endlichkeitssatz (siehe Theorem 6.2.5), der die Existenz eines endlichen Erzeugendensystem für linear reduktive Gruppen (vgl. Definition 4.4.1) garantiert. Für die hier betrachteten Untergruppen G, d.h. für linear reduktive Gruppen, haben wir in Kapitel 7 zwei Verfahren zur Berechnung fundamentaler Invarianten kennengelernt. Wir werden im Folgenden beide anwenden und damit deren Funktionalität demonstrieren. Da sich beide Algorithmen auf dieselben Eingabeparameter stützen, wollen diese nun im ersten Abschnitt näher beleuchten.

10.1 Vorbereitungen

Zur Berechnung fundamentaler Invarianten von Invariantenringen linear reduktiver Gruppen haben wir in Abschnitt 7.2 zwei Varianten kennengelernt: Die eine nutzt geschickt einfache Methoden der Linearen Algebra, die andere, wesentlich effizientere Variante stützt sich auf den Reynolds-Operator $\mathrm{Rey}_G : P \twoheadrightarrow P^G$, dessen Existenz nur für linear reduktive Gruppen gegeben ist (vgl. Theorem 7.1.4). Das als Derksen-Algorithmus bekannte Verfahren (siehe Algorithmus 7.2) und die sogenannte Lineare-Algebra-Methode wollen wir anschließend im Detail anwenden und Schritt für Schritt nachvollziehen, wobei wir die Berechnungen natürlich auf Computeralgebrasysteme stützen. Es gibt zwar eine Vielzahl von Computeralgebrasystemen, die Funktionen zur Berechnung von Invarianten unter *endlichen* Gruppen enthalten, allerdings gibt es derzeit nur wenige Computeralgebrasysteme, die auch Funktionen zur Berechnung von Erzeugendensystemen von Invariantenringen *unendlicher* Gruppen bereitstellen. Zwei davon sind:

- **Magma:** In Magma (siehe [Mg14]) findet sich eine Implementierung des DERKSEN-Algorithmus wieder. Einfache Berechnungen können auch in einem Online-Rechner[25] durchgeführt werden.
- **CoCoA:** Im CoCoA-Paket `invariants.cpkg`, das nicht zum Standard-Reportoire des Open-source-Computeralgebrasystems CoCoA zählt, ist mehr oder weniger direkt die in Abschnitt 7.2.4 beschriebene naive Lineare-Algebra-Methode umgesetzt. Da ein paar kleinere Fehler in dieser Implementierung enthalten waren, musste dieses Paket vor der Anwendung geringfügig angepasst werden. CoCoA (siehe [RAB15]) bzw. die Erweiterung ApCoCoA (siehe [KAT13]) sind als Open-source-Software jeweils kostenlos verfügbar.

Der DERKSEN-Algorithmus (siehe Algorithmus 7.2) ist im Grunde sehr einfach aufgebaut (siehe auch Abschnitt 7.2.2): Er stützt sich nur auf die Berechnung von Gröbner-Basen und den Reynolds-Operator, dessen Umsetzung für nicht-endliche Gruppen allerdings nicht so einfach zu machen ist (siehe dazu im Detail Abschnitt 7.1). Bis zur Berechnung des Hilbert-Ideals (siehe Definition 7.2.2) sind beide Methoden identisch. Dementsprechend stützen sich auch beide Verfahren auf dieselben Eingabeparameter:

- Das Verschwindungsideal $\mathcal{I}(G)$ der Gruppe G in einem geeigneten Polynomring.
- Die Beschreibung der Operation der Gruppe G auf $\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R})$ durch die Darstellungsmatrix $\mathcal{M}_B^B(\rho)$ der linearen Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R}))$ von G in $\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R})$ bzgl. einer Basis B.

Zunächst wollen wir für die Angabe des Verschwindungsideals einen geeigneten Polynomring angeben. Die hier betrachteten Untergruppen G sind als Zariski-abgeschlossene Untergruppen von $\mathrm{GL}_2(\mathbb{R})$ lineare algebraische Gruppen (vgl. Beispiel 4.1.4), d.h. affine Varietäten, die gleichzeitig die Struktur einer Gruppe besitzen (vgl. Definition 4.1.1). Wir betrachten G zunächst weiter als affine Varietät. Laut Bemerkung 4.1.12 lässt sich G in $\mathbb{A}^4_{\mathbb{R}}$ einbetten, d.h. genauer gibt es eine abgeschlossene Einbettung $\iota : G \hookrightarrow \mathbb{A}^4_{\mathbb{R}}$. Somit können wir G als Zariski-abgeschlossene Teilmenge von $\mathbb{A}^4_{\mathbb{R}}$ betrachten. Die zu ι gehörige Koordinatenabbildung $\iota^* : \mathbb{R}[\mathbb{A}^4_{\mathbb{R}}] \to \mathbb{R}[G]$ können wir Dank $\mathbb{R}[\mathbb{A}^4_{\mathbb{R}}] \cong \mathbb{R}[z_1,z_2,z_3,z_4]$ auch als surjektiven Ringhomomorphismus $\iota^* : \mathbb{R}[z_1,z_2,z_3,z_4] \to \mathbb{R}[G]$ auffassen. Wegen der Doppelindizierung der Einträge der Elemente von G als Matrixgruppe ist allerdings eine Bezeichnung der Unbestimmten durch Doppelindizes naheliegender. Deshalb betrachten wir den Polynomring $\mathbb{R}[z_{1,1},z_{1,2},z_{2,1},z_{2,2}]$, den wir mit der Termordnung $\sigma := \mathrm{DegRevLex}$ versehen. Dieser Polynomring ist isomorph zum Koordinatenring des $\mathbb{R}$-Vektorraums $\mathrm{Mat}_2(\mathbb{R})$ bzgl. der kanonischen Basis

$$\mathcal{E}_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{E}_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{E}_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{E}_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Somit ist der Koordinatenring $\mathbb{R}[G]$ der Gruppe G isomorph zu $\mathbb{R}[z_{1,1},z_{1,2},z_{2,1},z_{2,2}]/\mathcal{I}(G)$, wobei $\mathcal{I}(G) \subseteq \mathbb{R}[z_{1,1},z_{1,2},z_{2,1},z_{2,2}]$ das Verschwindungsideal von G ist (vgl. Satz 3.2.7).

Neben der Beschreibung der Gruppe durch ihr Verschwindungsideal spielt für die Berechnung der Invarianten unter G die algorithmische Beschreibung der Gruppenoperation durch die Darstellungsmatrix $\mathcal{M}_B^B(\rho) \in \mathrm{Mat}_\nu(\mathbb{R}[z_{1,1},z_{1,2},z_{2,1},z_{2,2}]/\mathcal{I}(G))$ von ρ bzgl. B eine zentrale Rolle (vgl. Definition 4.3.26), wobei $\nu = \dim_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R})) = \binom{n+2}{2}$ gelte. Bekanntlich ist der reelle Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R})$ eine direkte Summe

$$\mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R}) = \bigoplus_{k=0}^{n} \mathscr{P}_k(\mathbb{R}^2,\mathbb{R})$$

[25] Siehe http://magma.maths.usyd.edu.au/calc/.

und $B_k := (b_{i,j} \in B : \ i + j = k)$ für alle $k \in \{0, \ldots, n\}$ eine Basis des $\mathbb{R}$-Untervektorraums $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$ von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Nun betrachten wir für alle $k \in \{0, \ldots, n\}$ die rationale Unterdarstellung $\rho^{(k)} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_k(\mathbb{R}^2, \mathbb{R}))$ von G in $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$. Dann ist ρ die direkte Summe dieser linearen Unterdarstellungen, d.h. es gilt $\rho = \bigoplus_{k=0}^{n} \rho^{(k)}$ (vgl. Bemerkung 4.3.7). Laut Bemerkung 4.3.7 hat für alle $\mathcal{A} \in G$ die Darstellungsmatrix

$$\mathcal{M}_B^B(\rho_{\mathcal{A}}) = \begin{pmatrix} \mathcal{M}_{B_0}^{B_0}\left(\rho_{\mathcal{A}}^{(0)}\right) & & \cdots & 0 \\ & \mathcal{M}_{B_1}^{B_1}\left(\rho_{\mathcal{A}}^{(1)}\right) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \mathcal{M}_{B_n}^{B_n}\left(\rho_{\mathcal{A}}^{(n)}\right) \end{pmatrix} \in \mathrm{GL}_\nu(\mathbb{R})$$

des Automorphismus $\rho_{\mathcal{A}} \in \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ als Blockmatrizen auf der Diagonale genau die Darstellungsmatrizen $\mathcal{M}_{B_k}^{B_k}\left(\rho_{\mathcal{A}}^{(k)}\right)$ der Automorphismen $\rho_{\mathcal{A}}^{(k)} \in \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_k(\mathbb{R}^2, \mathbb{R}))$. Mit dieser Darstellungsmatrix lässt sich die Operation eines Gruppenelements $\mathcal{A} \in G$ auf eine Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ durch $\mathcal{M}_B^B(\rho_{\mathcal{A}}) \cdot \kappa_B(p)$ auf Koordinatenebene ausdrücken, d.h. mit anderen Worten, es gilt $\kappa_B(p \circ T_{\mathcal{A}}) = \mathcal{M}_B^B(\rho_{\mathcal{A}}) \cdot \kappa_B(p)$. Die Struktur der Darstellungsmatrix $\mathcal{M}_B^B(\rho_{\mathcal{A}})$ des Automorphismus $\rho_{\mathcal{A}}$ setzt sich auf natürliche Weise auf die Darstellungsmatrix von ρ fort. Analog gilt also

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} \mathcal{M}_{B_0}^{B_0}\left(\rho^{(0)}\right) & & \cdots & 0 \\ & \mathcal{M}_{B_1}^{B_1}\left(\rho^{(1)}\right) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \mathcal{M}_{B_n}^{B_n}\left(\rho^{(n)}\right) \end{pmatrix},$$

wobei $\mathcal{M}_{B_k}^{B_k}\left(\rho^{(k)}\right) \in \mathrm{Mat}_{\nu'}(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$ mit $\nu' := \dim_{\mathbb{R}}(\mathscr{P}_k(\mathbb{R}^2, \mathbb{R}))$ für alle k die Darstellungsmatrix der linearen Unterdarstellung $\rho^{(k)}$ ist. Damit können wir nun im Folgenden konkrete Gruppen untersuchen.

Wir werden die Einzelschritte beider Algorithmen nachfolgend angeben, um das Vorgehen nachvollziehen zu können. Damit ein Computeralgebrasystem genutzt werden kann, werden außerdem sämtliche Berechnungen, die mit Computeralgebrasystemen durchgeführt wurden, über einem berechenbaren Körper, hier über $\mathbb{Q}$ durchgeführt. Die oben eingeführte Doppelindizierung der Unbestimmten in P hat immense Vorteile, wenn es um die Zuordnung der Koordinaten zur jeweiligen Basisfunktion oder um die spätere Interpretation der Ergebnisse geht, aber die nachfolgenden Berechnungen sind im Polynomring $\tilde{P} := \mathbb{Q}[x_1, \ldots, x_\nu]$ leichter zu beschreiben und nachzuvollziehen, weshalb wir im Folgenden diesen anstatt P für Berechnungen mit Computeralgebrasystemen verwenden wollen. Dabei steht die Unbestimmte x_i für die Koordinate bzgl. der i-ten Basisfunktion in B in der angegebenen Ordnung. Den Polynomring $\tilde{P}$ versehen wir mit der Termordnung $\sigma := \mathrm{DegRevLex}$. Zudem betrachten wir nun die betrachteten Gruppen G über $\mathbb{Q}$. Die fundamentalen Invarianten über $\mathbb{Q}$ bilden dann aber insgesamt auch fundamentale Invarianten über $\mathbb{R}$ (vgl. dazu [KR00], Lemma 2.4.16, S. 116), was diese Vorgehensweise rechtfertigt.

10.2 Invarianten der speziellen orthogonalen Gruppe

Als erste Untergruppe der allgemeinen linearen Gruppe $\mathrm{GL}_2(\mathbb{R})$ werden wir die spezielle orthogonale Gruppe $G := \mathrm{SO}_2(\mathbb{R})$ als Matrixgruppe

$$\mathrm{SO}_2(\mathbb{R}) = \{(z_{i,j})_{1 \leq i,j \leq 2} \in \mathrm{GL}_2(\mathbb{R}) : \ (z_{i,j}) \cdot (z_{i,j})^{\mathrm{tr}} = \mathcal{I}_2 \text{ und } \det(z_{i,j}) = 1\}$$

betrachten. Das Verschwindungsideal der Gruppe in $\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ ergibt sich aus dem Mengenprädikat der obigen Beschreibung von $\mathrm{SO}_2(\mathbb{R})$. Mit $(z_{i,j})_{1\leq i,j\leq 2}$ gilt zunächst

$$\begin{pmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & z_{2,2} \end{pmatrix} \cdot \begin{pmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \end{pmatrix} = \begin{pmatrix} z_{1,1}^2 + z_{1,2}^2 & z_{1,1}z_{2,1} + z_{1,2}z_{2,2} \\ z_{1,1}z_{2,1} + z_{1,2}z_{2,2} & z_{2,1}^2 + z_{2,2}^2 \end{pmatrix}$$

und

$$\det \begin{pmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & z_{2,2} \end{pmatrix} = z_{1,1}z_{2,2} - z_{1,2}z_{2,1}$$

Wegen $(z_{i,j})_{1\leq i,j\leq 2} \cdot (z_{i,j})_{1\leq i,j\leq 2}^{\mathrm{tr}} = \mathcal{I}_2$ bzw. wegen $(z_{i,j})_{1\leq i,j\leq 2} \cdot (z_{i,j})_{1\leq i,j\leq 2}^{\mathrm{tr}} - \mathcal{I}_2 = 0$ und wegen $\det((z_{i,j})) = 1$ wird das Verschwindungsideal $\mathcal{I}(G) \subseteq \mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ somit von den folgenden vier Polynomen erzeugt:

$$\begin{aligned} g_1 &:= z_{1,1}^2 + z_{1,2}^2 - 1, & g_2 &:= z_{1,1}z_{2,1} + z_{1,2}z_{2,2}, \\ g_3 &:= z_{2,1}^2 + z_{2,2}^2 - 1, & g_4 &:= z_{1,1}z_{2,2} - z_{1,2}z_{2,1} - 1 \end{aligned}$$

Für die Darstellungsmatrix der rationalen Darstellung $\rho : \mathrm{SO}_2(\mathbb{R}) \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ müssen wir nun konkreter werden. Da der Fall $n = 0$ offensichtlich eher uninteressant ist, wollen wir hier für $n \in \{1, 2, 3\}$ die Invarianten explizit angeben. Die Invarianten für $n \geq 4$ lassen sich aber auf analoge Weise berechnen, allerdings klar zu Lasten der Lesbarkeit. Allgemeine Invarianten, d.h. von n abhängige „Formeln" für Invarianten wie z.B. im Falle der Vektorinvarianten (siehe Abschnitt 6.4) lassen sich leider nicht erkennen.

Der Fall $n = 1$

Zunächst bestimmen wir die Darstellungsmatrix von $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B = (b_{0,0}, b_{0,1}, b_{1,0})$ von $\mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R})$. Mit $\mathcal{Z} := (z_{i,j})_{1\leq i,j\leq 2}$ gilt:

$$\begin{aligned} \rho_{\mathcal{Z}}(b_{0,0})(x, y) &= b_{0,0} \circ T_{\mathcal{Z}}(x, y) = b_{0,0}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) = 1 = 1 \cdot b_{0,0}(x, y) \\ \rho_{\mathcal{Z}}(b_{0,1})(x, y) &= b_{0,1} \circ T_{\mathcal{Z}}(x, y) = b_{0,1}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\ &= z_{2,1}x + z_{2,2}y = (z_{2,2}b_{0,1} + z_{2,1}b_{1,0})(x, y) \\ \rho_{\mathcal{Z}}(b_{1,0})(x, y) &= b_{1,0} \circ T_{\mathcal{Z}}(x, y) = b_{1,0}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\ &= z_{1,1}x + z_{1,2}y = (z_{1,2}b_{0,1} + z_{1,1}b_{1,0})(x, y) \end{aligned}$$

Wegen $\dim_{\mathbb{R}}(\mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R})) = 3$ erhalten wir zunächst folgende 3×3-Matrix, deren Einträge Polynome in $\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ sind:

$$\mathcal{Q}_1 := (q_{i,j})_{1\leq i,j\leq 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & z_{2,2} & z_{1,2} \\ 0 & z_{2,1} & z_{1,1} \end{pmatrix}.$$

Indem wir für jeden Eintrag der Matrix $\mathcal{Q}_1$, also für jedes Polynom $q_{i,j} \in \mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ die Normalform $\mathrm{NF}_{\sigma,\mathcal{I}(G)}(q_{i,j})$ bzgl. des Verschwindungsideals $\mathcal{I}(G)$ berechnen, erhalten wir die eindeutig bestimmte Darstellungsmatrix der Gruppenoperation bzgl. der Basis B:

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & z_{2,2} & -z_{2,1} \\ 0 & z_{2,1} & z_{2,2} \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$$

Nun rechnen wir über $\mathbb{Q}$. Für $1 \leq i, j \leq 3$ sei $\tilde{q}_{i,j} \in \mathbb{Q}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$ der (i, j)-te Eintrag der Matrix $\mathcal{M}_B^B(\rho)$. Weiter betrachten wir den Polynomring

$$Q := \mathbb{Q}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}, x_1, x_2, x_3, y_1, y_2, y_3],$$

den wir mit der Eliminationsordnung $\tau := \mathrm{Elim}(\{z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}\})$ versehen. Die Polynome g_1, g_2, g_3, g_4 und $\tilde{q}_{i,j}$ sind nun als Polynome in Q anzusehen. Anschließend bestimmen wir im Polynomring Q für alle $i \in \{1,2,3\}$ das Polynom $h_i := y_i - \sum_{j=1}^{3} \tilde{q}_{i,j} x_j$, d.h. wir erhalten hier die Polynome

$$h_1 = y_1 - x_1, \qquad h_2 = y_2 - z_{2,2}x_2 + z_{2,1}x_3, \qquad h_3 = y_3 - z_{2,1}x_2 - z_{2,2}x_3.$$

Das Derksen-Ideal $I \subseteq Q$ wird von g_1, g_2, g_3, g_4 und h_1, h_2, h_3 erzeugt. Im nächsten Schritt wird eine τ-Gröbner-Basis H von I berechnet. Diese besteht aus folgenden neun Polynomen:

$$\begin{aligned} H = \{&-x_1 + y_1,\ -z_{2,1}x_2 - z_{2,2}x_3 + y_3,\ -z_{2,2}x_2 + z_{2,1}x_3 + y_2,\ z_{2,1}^2 + z_{2,2}^2 - 1,\\ &z_{2,1}y_2 - z_{2,2}y_3 + x_3,\ -z_{2,2}y_2 - z_{2,1}y_3 + x_2,\ x_2^2 + x_3^2 - y_2^2 - y_3^2,\\ &- z_{1,2} - z_{2,1}, -z_{1,1} + z_{2,2}\} \end{aligned}$$

Darauf aufbauend erhalten wir die $\widehat{\tau}$-Gröbner-Basis $\widehat{H} = H \cap \mathbb{Q}[x_1, x_2, x_3, y_1, y_2, y_3]$ des Eliminationsideals $I \cap \mathbb{Q}[x_1, x_2, x_3, y_1, y_2, y_3]$, die nur noch zwei Polynome enthält. Es gilt:

$$\widehat{H} = \{-x_1 + y_1,\ x_2^2 + x_3^2 - y_2^2 - y_3^2\}.$$

Indem wir in den Polynomen aus $\widehat{H}$ sämtliche Unbestimmten y durch 0 ersetzen, erhalten wir ein Erzeugendensystem des Hilbert-Ideals $\mathrm{HI}_G \subseteq \mathbb{Q}[x_1, x_2, x_3]$:

$$\mathrm{HI}_G = \langle f(x_1, x_2, x_3, 0, 0, 0) :\ f \in \widehat{H} \rangle = \langle -x_1, x_2^2 + x_3^2 \rangle,$$

wobei wir $f(x_1, x_2, x_3, 0, 0, 0)$ für $f \in \widehat{H}$ als Polynom in $\mathbb{Q}[x_1, x_2, x_3]$ betrachten. Wie man sofort sieht, wird HI_G minimal erzeugt von diesen beiden Polynomen. Von hier an gehen die beiden Algorithmen unterschiedlich vor:

(1) **Derksen-Algorithmus:** Laut Satz 7.2.3 und Korollar 7.2.4 erhalten wir mit dem Reynolds-Operator ein minimales Erzeugendensystem von $\mathbb{Q}[x_1, x_2, x_3]^G$. Wie man sich leicht überzeugt, sind beide Erzeuger von HI_G bereits invariant (siehe Satz 7.2.1). Somit erhalten wir die folgenden Erzeuger des Invariantenrings:

$$f_1 := \mathrm{Rey}_G(x_1) = x_1, \qquad f_2 := \mathrm{Rey}_G(x_2^2 + x_3^2) = x_2^2 + x_3^2$$

(2) Im ersten auf die Berechnung der Hilbert-Ideals folgenden Schritt sucht die **Lineare-Algebra-Methode** nach Invarianten unter den Erzeugern des Hilbert-Ideals. Da hier beide Erzeuger schon invariant sind, ist der Algorithmus an dieser Stelle bereits sofort zu Ende. Wir werden die Funktionsweise dieser Methode deshalb erst im nächsten Fall genauer untersuchen können.

Wie in Kapitel 6 erwähnt, ist neben der Bestimmung eines Erzeugendensystems des Invariantenrings die Frage nach den Relationen dieser Erzeuger von zentraler Bedeutung für die Invariantentheorie. Zu bestimmen ist dazu lediglich das Relationenideal

$$\mathrm{Rel}(f_1, f_2) = \{h \in \mathbb{Q}[y_1, y_2] :\ h(f_1, f_2) = 0\}.$$

Allerdings ist hier offensichtlich zu sehen, dass dieses Ideal das Nullideal ist, was nichts anderes bedeutet, als dass die beiden Erzeuger algebraisch unabhängig sind. Wir übertragen das Ergebnis der Berechnungen nun auf P mit der oben angebenen Notation der Unbestimmten durch Doppelindizes. Wie bereits erwähnt, erzeugen zunächst die Polynome f_1 und f_2 auch den Invariantenring $\mathbb{R}[x_1, x_2, x_3]^G$. Nun ist weiter lediglich der durch

$$x_1 \mapsto a_{0,0},\ x_2 \mapsto a_{0,1},\ x_3 \mapsto a_{1,0}$$

definierte $\mathbb{R}$-Algebren-Homomorphismus $\Phi : \mathbb{R}[x_1, x_2, x_3] \to P$ zu betrachten. Dann erzeugen für $i \in \{1, 2\}$ die Polynome $\Phi(f_i)$ den Invariantenring P^G, was wir in folgendem Satz festhalten wollen, dessen Beweis sich unmittelbar zum einen aus den vorausgegangenen Überlegungen und zum anderen aus der Korrektheit beider Algorithmen ergibt.

Satz 10.2.1. (Rotationsinvarianten für $n = 1$)
Sei $V = \mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R})$ mit Basis $B = (b_{0,0}, b_{0,1}, b_{1,0})$ und sei $G = \mathrm{SO}_2(\mathbb{R})$. Dann wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring $\mathbb{R}[V]$ isomorphen Polynomrings $P = \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}]$ minimal von den Polynomen $f_1 = a_{0,0}$ und $f_2 = a_{0,1}^2 + a_{1,0}^2$ erzeugt. Diese Erzeuger sind zudem algebraisch unabhängig.

Beweis: Folgt aus der Korrektheit von Algorithmus 7.2 aus Theorem 7.2.8 und von Algorithmus 7.5 aus Theorem 7.2.14. □

Der Fall $n = 2$

Um im Fall $n = 2$ die Darstellungsmatrix der rationalen Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0})$ angeben zu können, genügt es wegen der Blockstruktur von $\mathcal{M}_B^B(\rho)$ und der Kenntnis der Darstellungsmatrix für $n = 1$, die Darstellungsmatrix der linearen Unterdarstellung $\rho^{(2)} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_2(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B_2 = (b_{0,2}, b_{1,1}, b_{2,0})$ von $\mathscr{P}_2(\mathbb{R}^2, \mathbb{R})$ zu bestimmen. Mit $\mathcal{Z} := (z_{i,j})_{1 \leq i,j \leq 2}$ gilt:

$$\begin{aligned}
\rho_{\mathcal{Z}}(b_{0,2})(x, y) &= b_{0,2} \circ T_{\mathcal{Z}}(x, y) = b_{0,2}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{2,1}x + z_{2,2}y)^2 = z_{2,1}^2x^2 + 2z_{2,1}z_{2,2}xy + z_{2,2}^2y^2 \\
&= (z_{2,2}^2 b_{0,2} + 2z_{2,1}z_{2,2}b_{1,1} + z_{2,1}^2 b_{2,0})(x, y) \\
\rho_{\mathcal{Z}}(b_{1,1})(x, y) &= b_{1,1} \circ T_{\mathcal{Z}}(x, y) = b_{1,1}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}x + z_{1,2}y) \cdot (z_{2,1}x + z_{2,2}y) \\
&= z_{1,1}z_{2,1}x^2 + (z_{1,1}z_{2,2} + z_{1,2}z_{2,1})xy + z_{1,2}z_{2,2}y^2 \\
&= (z_{1,2}z_{2,2}b_{0,2} + (z_{1,1}z_{2,2} + z_{1,2}z_{2,1})b_{1,1} + z_{1,1}z_{2,1}b_{2,0})(x, y) \\
\rho_{\mathcal{Z}}(b_{2,0})(x, y) &= b_{2,0} \circ T_{\mathcal{Z}}(x, y) = b_{2,0}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}x + z_{1,2}y)^2 = z_{1,1}^2x^2 + 2z_{1,1}z_{1,2}xy + z_{1,2}^2y^2 \\
&= (z_{1,2}^2 b_{0,2} + 2z_{1,1}z_{1,2}b_{1,1} + z_{1,1}^2 b_{2,0})(x, y)
\end{aligned}$$

Damit erhalten wir analog zunächst eine Matrix

$$\mathcal{Q}_2 := (q_{i,j})_{1 \leq i,j \leq 3} = \begin{pmatrix} z_{2,2}^2 & z_{1,2}z_{2,2} & z_{1,2}^2 \\ 2z_{2,1}z_{2,2} & z_{1,1}z_{2,2} + z_{1,2}z_{2,1} & 2z_{1,1}z_{1,2} \\ z_{2,1}^2 & z_{1,1}z_{2,1} & z_{1,1}^2 \end{pmatrix}$$

deren Einträge Polynome in $\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ sind. Durch Berechnung der Normalform der einzelnen Einträge bzgl. $\mathcal{I}(G)$ erhalten wir die gesuchte Darstellungsmatrix von $\rho^{(2)}$:

$$\mathcal{M}_{B_2}^{B_2}\left(\rho^{(2)}\right) = \begin{pmatrix} z_{2,2}^2 & -z_{2,1}z_{2,2} & -z_{2,2}^2 + 1 \\ 2z_{2,1}z_{2,2} & 2z_{2,2}^2 - 1 & -2z_{2,1}z_{2,2} \\ -z_{2,2}^2 + 1 & z_{2,1}z_{2,2} & z_{2,2}^2 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)).$$

Somit lautet die Darstellungsmatrix von $\mathcal{M}_B^B(\rho)$ wie folgt:

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & -z_{2,1} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & -z_{2,1}z_{2,2} & -z_{2,2}^2+1 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & 2z_{2,2}^2-1 & -2z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2+1 & z_{2,1}z_{2,2} & z_{2,2}^2 \end{pmatrix}.$$

Ab jetzt rechnen wir erneut wieder über $\mathbb{Q}$. Für $1 \leq i,j \leq 6$ bezeichnen wir nun den (i,j)-ten Eintrag von $\mathcal{M}_B^B(\rho)$ mit $\tilde{q}_{i,j} \in \mathbb{Q}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G)$. Wir betrachten analog zum Fall $n = 1$ den Polynomring $Q := \mathbb{Q}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}, x_1, \ldots, x_6, y_1, \ldots, y_6]$, den wir mit der Eliminationsordnung $\tau := \mathrm{Elim}(\{z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}\})$ versehen. Aufgrund der Blockstruktur der Darstellungsmatrix $\mathcal{M}_B^B(\rho)$ genügt es, im Polynomring Q für alle $i \in \{4, \ldots, 6\}$ das Polynom $h_i := y_i - \sum_{j=1}^{6} \tilde{q}_{i,j} x_j$ zu bestimmen, die Polynome h_1, h_2, h_3 von Fall $n = 1$ können wir übernehmen. Wir erhalten also im vorliegenden Fall die folgenden zusätzlichen Polynome:

$$\begin{aligned} h_4 &:= y_4 - z_{2,2}^2 x_4 + z_{2,1}z_{2,2}x_5 + z_{2,2}^2 x_6 - x_6 \\ h_5 &:= y_5 - 2z_{2,1}z_{2,2}x_4 - 2z_{2,2}^2 x_5 + x_5 + 2z_{2,1}z_{2,2}x_6 \\ h_6 &:= y_6 + z_{2,2}^2 x_4 - x_4 - z_{2,1}z_{2,2}x_5 - z_{2,2}^2 x_6 \end{aligned}$$

Somit wird im Fall $n = 2$ das Derksen-Ideal $I \subseteq Q$ von den Polynomen $g_1, \ldots, g_4$ des Verschwindungsideals $\mathcal{I}(G)$, die dabei als Polynome in Q betrachtet werden, und von den Polynomen $h_1, \ldots, h_6 \in Q$ erzeugt. Auf die Angabe der τ-Gröbner-Basis H von I verzichten wir hier allerdings, da diese aus 25 Polynomen besteht. Die $\widehat{\tau}$-Gröbner-Basis $\widehat{H}$ des Eliminationsideals $I \cap \mathbb{Q}[x_1, \ldots, x_6, y_1, \ldots, y_6]$ besteht nur noch aus 10 Elementen, im Einzelnen gilt:

$$\begin{aligned} \widehat{H} = \{ & -x_4 - x_6 + y_4 + y_6,\ -x_1 + y_1,\ x_2^2 + x_3^2 - y_2^2 - y_3^2, \\ & -\tfrac{1}{2}x_5^2 - 2x_6^2 + 2x_6y_4 + 1/2y_5^2 + 2x_6y_6 - 2y_4y_6, \\ & -\tfrac{1}{2}x_6y_2^2 - \tfrac{1}{2}x_5y_2y_3 + \tfrac{1}{2}x_6y_3^2 + \tfrac{1}{2}x_3^2y_4 - \tfrac{1}{2}y_3^2y_4 + \tfrac{1}{2}x_2x_3y_5 - \tfrac{1}{2}x_3^2y_6 + \tfrac{1}{2}y_2^2y_6, \\ & \tfrac{1}{4}x_5y_2^2 - x_6y_2y_3 - \tfrac{1}{4}x_5y_3^2 - \tfrac{1}{2}x_2x_3y_4 + \tfrac{1}{2}y_2y_3y_4 + \tfrac{1}{2}x_3^2y_5 - \tfrac{1}{4}y_2^2y_5 - \tfrac{1}{4}y_3^2y_5 \\ & \quad + \tfrac{1}{2}x_2x_3y_6 + \tfrac{1}{2}y_2y_3y_6, \\ & -\tfrac{1}{2}x_2x_5y_2 - x_3x_6y_2 - \tfrac{1}{2}x_3x_5y_3 + x_2x_6y_3 + x_3y_2y_4 - x_2y_3y_4 + \tfrac{1}{2}x_2y_2y_5 + \tfrac{1}{2}x_3y_3y_5, \\ & -\tfrac{1}{2}x_3x_5y_2 + x_2x_6y_2 + \tfrac{1}{2}x_2x_5y_3 + x_3x_6y_3 - \tfrac{1}{2}x_3y_2y_5 + \tfrac{1}{2}x_2y_3y_5 - x_2y_2y_6 - x_3y_3y_6, \\ & -x_2x_3x_5 - 2x_3^2x_6 + x_6y_2^2 + x_6y_3^2 + x_3^2y_4 - y_3^2y_4 + y_2y_3y_5 + x_3^2y_6 - y_2^2y_6, \\ & -x_3^2x_5 + 2x_2x_3x_6 + \tfrac{1}{2}x_5y_2^2 + \tfrac{1}{2}x_5y_3^2 - x_2x_3y_4 + y_2y_3y_4 - \tfrac{1}{2}y_2^2y_5 \\ & \quad + \tfrac{1}{2}y_3^2y_5 - x_2x_3y_6 - y_2y_3y_6\} \end{aligned}$$

Aufbauend auf dem letzten Erzeugendensystem erhalten wir auf einfache Weise ein homogenes Erzeugendensystem des Hilbert-Ideals $\mathrm{HI}_G \subseteq \mathbb{Q}[x_1, \ldots, x_6]$. Es gilt:

$$\begin{aligned} \mathrm{HI}_G &= \langle f(x_1, \ldots, x_6, 0, \ldots, 0) :\ f \in \widehat{H} \rangle \\ &= \langle -x_4 - x_6,\ -x_1,\ x_2^2 + x_3^2,\ -\tfrac{1}{2}x_5^2 - 2x_6^2,\ -x_2x_3x_5 - 2x_3^2x_6,\ -x_3^2x_5 + 2x_2x_3x_6 \rangle, \end{aligned}$$

wobei wir auch hier $f(x_1, \ldots, x_6, 0, \ldots, 0)$ für $f \in \widehat{H}$ als Polynom in $\mathbb{Q}[x_1, \ldots, x_6]$ betrachten. Dies ist zudem ein minimales Erzeugendensystem von HI_G. Was zu erwarten war, sehen wir auch hier: Die beiden Erzeuger des Hilbert-Ideals aus Fall $n = 1$ sind auch hier enthalten. Nun zur Berechnung der fundamentalen Invarianten:

(1) **Derksen-Algorithmus:** Laut Satz 7.2.3 ist nun ebenfalls nur noch der Reynolds-Operator auf die Erzeuger des Hilbert-Ideals anzuwenden und wir erhalten ein minimales Erzeugendensystem von $\mathbb{Q}[x_1,\dots,x_6]^G$ (vgl. Korollar 7.2.4). Da der Reynolds-Operator Grad-erhaltend ist (folgt aus Korollar 7.1.5), folgt somit bereits, dass wir sechs fundamentale Invarianten erhalten. Dies sind gerade die Bilder der Erzeuger des Hilbert-Ideals, wobei uns $f_1 = x_1$ und $f_2 = x_2^2 + x_3^2$ bereits bekannt sind. Auch das Polynom $x_4 + x_6$ ist bereits invariant, sodass wir als dritte offensichtliche Invariante $f_3 := x_4 + x_6$ erhalten. Für die restlichen drei Polynome erhalten wir:

$$\begin{aligned} f_4 &:= \mathrm{Rey}_G(-\tfrac{1}{2}x_5^2 - 2x_6^2) = x_4x_6 - \tfrac{1}{4}x_5^2,\\ f_5 &:= \mathrm{Rey}_G(-x_2x_3x_5 - 2x_3^2x_6) = x_2^2x_6 - x_2x_3x_5 + x_3^2x_4,\\ f_6 &:= \mathrm{Rey}_G(-x_3^2x_5 + 2x_2x_3x_6) = x_2^2x_5 - 2x_2x_3x_4 + 2x_2x_3x_6 - x_3^2x_5 \end{aligned}$$

Somit liefert der Derksen-Algorithmus als Algebra-Erzeugendensystem des Invariantenrings $\mathbb{Q}[x_1,\dots,x_6]^G$ die Invarianten $f_1,\dots,f_6$. Wie sich auch hierauf aufbauend leicht nachrechnen lässt, ist dies ein minimales Erzeugendensystem, da sich keines der sechs Polynome als Polynom in den restlichen fünf Polynomen schreiben lässt.

(2) **Lineare-Algebra-Methode:** Die Lineare-Algebra-Methode berechnet fundamentale Invarianten ohne Verwendung des Reynolds-Operators. Dennoch wollen wir diesen zu Beginn kurz heranziehen, da er das weitere Vorgehen offenlegt. Wir wissen, dass laut Satz 7.2.3 $\mathrm{Rey}_G(H)$ ein Algebra-Erzeugendensystem von $\mathbb{Q}[x_1,\dots,x_6]^G$ ist. Da der Reynolds-Operator den Grad erhält, können wir noch weitere Informationen aus H ablesen. Wir benötigen zwei fundamentale Invarianten vom Grad 1, zwei vom Grad 2 und zwei vom Grad 3. Da der Invariantenring eine standardgraduierte $\mathbb{Q}$-Unteralgebra ist, kann man zur Berechnung fundamentaler Invarianten gradweise vorgehen. Zunächst lässt sich mit dem Invarianz-Kriterium aus Algorithmus 7.1 leicht feststellen, dass die folgenden Polynome bereits invariant sind:

$$r_1 := -x_1, \qquad r_2 := -x_4 - x_6, \qquad r_3 := x_2^2 + x_3^2.$$

Umgekehrt enthält H drei Erzeuger des Hilbert-Ideals, die nicht invariant sind: Ein Polynom vom Grad 2 und zwei Polynome vom Grad 3. Offensichtlich fehlt zunächst noch eine fundamentale Invariante vom Grad 2. Um die fehlende Invariante vom Grad 2 zu erhalten, ist zunächst eine $\mathbb{Q}$-Basis C' des $\mathbb{Q}$-Vektorraums

$$(\mathrm{HI}_G)_2/\langle r_1^{c_1}r_2^{c_2}r_3^{c_3} :\ c_1\deg(r_1) + c_2\deg(r_2) + c_3\deg(r_3) = 2\rangle$$

zu bestimmen, wobei $(\mathrm{HI}_G)_2 = \mathrm{HI}_G \cap \mathbb{Q}[x_1,\dots,x_6]_2$ gilt. Wir erhalten

$$\begin{aligned} C' = \{&-\tfrac{1}{2}x_5^2 - 2x_6^2,\ 2x_4^2 - x_4x_6 - 3x_6^2,\ -x_4x_5 - x_5x_6,\\ &- x_3x_4 - x_3x_6,\ -x_2x_4 - x_2x_6,\ -x_1x_6,\ -x_1x_5,\ -x_1x_3,\ -x_1x_2\} \end{aligned}$$

Darauf aufbauend kann mit Algorithmus 7.4 eine $\mathbb{Q}$-Basis C des mit C' erzeugten $\mathbb{Q}$-Vektorraums $\mathbb{Q}[x_1,\dots,x_6]^G \cap \langle C'\rangle_{\mathbb{Q}}$ berechnet werden. In diesem Fall erhalten wir als Ergebnis die Basis $C = \{-4x_4^2 - \frac{5}{2}x_5^2 + 2x_4x_6 - 4x_6^2\}$ und damit die fehlende fundamentale Invariante vom Grad 2. Auf analoge Weise erhalten wir im Grad 3 die fundamentalen Invarianten

$$-2x_2x_3x_4 + x_2^2x_5 + 2x_2x_3x_6 \quad \text{und} \quad -x_2^2x_4 + x_3^2x_4 - 2x_2x_3x_5 + x_2^2x_6 - x_3^2x_6.$$

Somit liefert zusammengefasst die Lineare-Algebra-Methode zunächst folgende fundamentalen Invarianten:

$$\begin{aligned}
\tilde{f}_1 &= -x_1, & \tilde{f}_4 &= -4x_4^2 - \tfrac{5}{2}x_5^2 + 2x_4x_6 - 4x_6^2,\\
\tilde{f}_2 &= x_2^2 + x_3^2, & \tilde{f}_5 &= -2x_2x_3x_4 + x_2^2x_5 + 2x_2x_3x_6,\\
\tilde{f}_3 &= -x_4 - x_6, & \tilde{f}_6 &= -x_2^2x_4 + x_3^2x_4 - 2x_2x_3x_5 + x_2^2x_6 - x_3^2x_6
\end{aligned}$$

Damit stellt sich natürlich sofort die Frage, ob beide Ergebnisse identisch sind, d.h. ob diese Polynome dieselbe $\mathbb{Q}$-Unteralgebra von $\mathbb{Q}[x_1,\ldots,x_6]$ erzeugen wie $f_1,\ldots,f_6$ zuvor. Das lässt sich mit den Methoden in Kapitel 2 leicht überprüfen, wir werden hier aber einen anderen Weg gehen, der die Zusammenhänge der beiden Erzeugendensysteme deutlicher macht. Sei dazu σ = DegRevLex. Wir betrachten nun die durch $\tilde{f}_1,\ldots,\tilde{f}_6$ erzeugte $\mathbb{Q}$-Unteralgebra S von $\mathbb{Q}[x_1,\ldots,x_6]$ und berechnen eine σ-SAGBI-Basis von S. Mit Hilfe der SAGBI-Prozedur erhalten wir folgende σ-SAGBI-Basis von S:

$$\begin{aligned}
\{&x_1,\ x_4+x_6,\ x_2^2+x_3^2,\ x_4^2+\tfrac{5}{8}x_5^2-\tfrac{1}{2}x_4x_6+x_6^2,\ x_2x_3x_4-\tfrac{1}{2}x_2^2x_5+\tfrac{1}{2}x_3^2x_5-x_2x_3x_6,\\
&x_2^2x_4-x_3^2x_4+2x_2x_3x_5-x_2^2x_6+x_3^2x_6,\ x_5^2-4x_4x_6,\ x_3^2x_4-x_2x_3x_5+x_2^2x_6\}
\end{aligned}$$

Da wir nun sogar insgesamt acht Polynome in der σ-SAGBI-Basis wiederfinden, können wir natürlich noch nicht beurteilen, ob $S = \mathbb{Q}[f_1,\ldots,f_6]^G$ gilt. Auf dieser σ-SAGBI-Basis aufbauend wollen wir nun mit Algorithmus 5.3 die reduzierte σ-SAGBI-Basis von S berechnen. Wir erhalten schließlich folgende reduzierte σ-SAGBI-Basis von S:

$$\begin{aligned}
\{&x_1,\ x_4+x_6,\ x_2^2+x_3^2,\ x_2x_3x_4-\tfrac{1}{2}x_2^2x_5+\tfrac{1}{2}x_3^2x_5-x_2x_3x_6,\\
&x_5^2-4x_4x_6,\ x_3^2x_4-x_2x_3x_5+x_2^2x_6\}
\end{aligned}$$

Die Elemente dieser reduzierten σ-SAGBI-Basis von S stimmen also bis auf skalare Vielfache genau mit den Polynomen $f_1,\ldots,f_6$ überein, die wir als Ergebnis des Derksen-Algorithmus erhalten hatten. Somit stimmen S und $\mathbb{Q}[f_1,\ldots,f_6]$, also auch S und $\mathbb{Q}[x_1,\ldots,x_6]^G$, tatsächlich überein.

Auch hier wollen wir die Relationen der fundamentalen Invarianten untersuchen. Seien dazu $y_1,\ldots,y_6$ weitere Unbestimmte. Dann gilt für das Relationenideal von $f_1,\ldots,f_6$ im Polynomring $\mathbb{Q}[y_1,\ldots,y_6]$:

$$\begin{aligned}
\mathrm{Rel}(f_1,\ldots,f_6) &= \{h\in\mathbb{Q}[y_1,\ldots,y_6]:\ h(f_1,\ldots,f_6)=0\}\\
&= \langle 2y_2^2y_4 - 2y_2y_3y_5 + 2y_5^2 + \tfrac{1}{2}y_6^2\rangle
\end{aligned}$$

Somit sind die fundamentalen Invarianten $f_1,\ldots,f_6$ algebraisch abhängig. Wir übertragen das Ergebnis der Berechnungen nun wieder auf P mit der oben angebenen Notation der Unbestimmten durch Doppelindizes. Wie bereits erwähnt, erzeugen zunächst die Polynome $f_1,\ldots,f_6$ auch den Invariantenring $\mathbb{R}[x_1,\ldots,x_6]^G$. Nun ist weiter lediglich der durch

$$x_1\mapsto a_{0,0},\ x_2\mapsto a_{0,1},\ x_3\mapsto a_{1,0},\ x_4\mapsto a_{0,2},\ x_5\mapsto a_{1,1},\ x_6\mapsto a_{2,0}$$

definierte $\mathbb{R}$-Algebren-Homomorphismus $\Phi:\mathbb{R}[x_1,\ldots,x_6]\to P$ zu betrachten. Dann erzeugen für $i\in\{1,\ldots,6\}$ die Polynome $\Phi(f_i)$ den Invariantenring P^G, was wir in folgendem Satz festhalten wollen, dessen Beweis sich unmittelbar zum einen aus den vorausgegangenen Überlegungen und zum anderen aus der Korrektheit von Algorithmus 7.2 ergibt.

Satz 10.2.2. (Rotationsinvarianten für $n=2$)
Sei $V=\mathscr{P}_{\leq 2}(\mathbb{R}^2,\mathbb{R})$ mit Basis $B=(b_{0,0},b_{0,1},b_{1,0},b_{0,2},b_{1,1},b_{2,0})$ und sei $G=\mathrm{SO}_2(\mathbb{R})$. Dann

wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring $\mathbb{R}[V]$ isomorphen Polynomrings $P = \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}, a_{0,2}, a_{1,1}, a_{2,0}]$ minimal durch folgende Polynome erzeugt:

$$\begin{aligned} f_1 &= a_{0,0} & f_4 &= a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2 \\ f_2 &= a_{0,1}^2 + a_{1,0}^2 & f_5 &= a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2} \\ f_3 &= a_{0,2} + a_{2,0} & f_6 &= a_{0,1}^2 a_{1,1} - 2a_{0,1}a_{1,0}a_{0,2} + 2a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1} \end{aligned}$$

Diese Erzeuger sind algebraisch abhängig und erfüllen folgenden Zusammenhang:

$$f_2^2 f_4 - f_2 f_3 f_5 + f_5^2 + \tfrac{1}{4} f_6^2 = 0.$$

Beweis: Folgt aus Theorem 7.2.8 und Theorem 7.2.14. □

Wir wollen uns an dieser Stelle exemplarisch explizit von der Invarianz eines dieser Polynome überzeugen, auch um den Begriff „Invarianz" besser „fassen" zu können; die restlichen Polynome könnte man analog behandeln.

Beispiel 10.2.3. Wie wohl bekannt ist, gibt es für jedes Element $\mathcal{A} \in \mathrm{SO}_2(\mathbb{R})$ ein $\varphi \in \mathbb{R}$ mit $\mathcal{A} = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$. Wir schreiben zur Abkürzung $c := \cos(\varphi)$ und $s := \sin(\varphi)$. Betrachten wir nun eine Polynomfunktion $p \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$, so operiert $\mathcal{A}$ wie folgt auf p:

$$\begin{aligned} p \circ T_{\mathcal{A}}(x,y) &= p(xc - ys, xs + yc) \\ &= a_{0,0} + a_{0,1}(xs + yc) + a_{1,0}(xc - ys) \\ &\quad + a_{0,2}(xs + yc)^2 + a_{1,1}(xc - ys)(xs + yc) + a_{2,0}(xc - ys)^2 \\ &= a_{0,0} + \underbrace{(a_{0,1}c - a_{1,0}s)}_{=:\tilde{a}_{0,1}} y + \underbrace{(a_{0,1}s + a_{1,0}c)}_{=:\tilde{a}_{1,0}} x \\ &\quad + \underbrace{(a_{0,2}c^2 - a_{1,1}cs + a_{2,0}s^2)}_{=:\tilde{a}_{0,2}} y^2 + \underbrace{(2a_{0,2}cs + (c^2 - s^2)a_{1,1} - 2a_{2,0}cs)}_{=:\tilde{a}_{1,1}} xy \\ &\quad + \underbrace{(a_{0,2}s^2 + a_{1,1}cs + a_{2,0}c^2)}_{=:\tilde{a}_{2,0}} x^2 \end{aligned}$$

Dies erhalten wir analog mit Hilfe der Darstellungsmatrix von $\rho_{\mathcal{A}}$:

$$\begin{aligned} \mathcal{M}_B^B(\rho_{\mathcal{A}}) \cdot \kappa_B(p) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & c & -s & 0 & 0 & 0 \\ 0 & s & c & 0 & 0 & 0 \\ 0 & 0 & 0 & c^2 & -s \cdot c & -c^2 + 1 \\ 0 & 0 & 0 & 2s \cdot c & 2c^2 - 1 & -2s \cdot c \\ 0 & 0 & 0 & -c^2 + 1 & s \cdot c & c^2 \end{pmatrix} \cdot \begin{pmatrix} a_{0,0} \\ a_{0,1} \\ a_{1,0} \\ a_{0,2} \\ a_{1,1} \\ a_{2,0} \end{pmatrix} \\ &= \begin{pmatrix} a_{0,0} \\ a_{0,1}c - a_{1,0}s \\ a_{0,1}s + a_{1,0}c \\ a_{0,2}c^2 - a_{1,1}cs + a_{2,0}(-c^2 + 1) \\ 2a_{0,2}cs + a_{1,1}(2c^2 - 1) - a_{2,0}cs \\ a_{0,2}(-c^2 + 1) + a_{1,1}cs + a_{2,0}c^2 \end{pmatrix} = \begin{pmatrix} a_{0,0} \\ a_{0,1}c - a_{1,0}s \\ a_{0,1}s + a_{1,0}c \\ a_{0,2}c^2 - a_{1,1}cs + a_{2,0}s^2 \\ 2a_{0,2}cs + a_{1,1}(c^2 - s^2) - a_{2,0}cs \\ a_{0,2}s^2 + a_{1,1}cs + a_{2,0}c^2 \end{pmatrix} \end{aligned}$$

Wir werden nun exemplarisch f_5 betrachten und dessen Invarianz nachweisen. Es gilt zunächst:

$$\begin{aligned} f_5^{\mathcal{A}} &= f_5(a_{0,0}, \tilde{a}_{0,1}, \tilde{a}_{1,0}, \tilde{a}_{0,2}, \tilde{a}_{1,1}, \tilde{a}_{2,0}) = \tilde{a}_{0,1}^2 \tilde{a}_{2,0} - \tilde{a}_{0,1}\tilde{a}_{1,0}\tilde{a}_{1,1} + \tilde{a}_{1,0}^2 \tilde{a}_{0,2} \\ &= a_{0,2}a_{1,0}^2 c^4 - a_{0,1}a_{1,0}a_{1,1}c^4 + a_{0,1}^2 a_{2,0}c^4 + 2a_{0,2}a_{1,0}^2 c^2 s^2 - 2a_{0,1}a_{1,0}a_{1,1}c^2 s^2 \\ &\quad + 2a_{0,1}^2 a_{2,0}c^2 s^2 + a_{0,2}a_{1,0}^2 s^4 - a_{0,1}a_{1,0}a_{1,1}s^4 + a_{0,1}^2 a_{2,0}s^4 \end{aligned}$$

Wegen $\cos(\varphi)^2 + \sin(\varphi)^2 = 1$ lässt sich dieses Polynom noch vereinfachen. Dazu betrachten wir im Polynomring $\mathbb{Q}[a_{0,1}, a_{1,0}, a_{0,2}, a_{1,1}, a_{2,0}, c, s]$, versehen mit der Termordnung DegRevLex, das Ideal $I := \langle c^2 + s^2 - 1 \rangle$. Dann folgt:

$$\mathrm{NF}_I(\tilde{f}_5) = a_{0,2}a_{1,0}^2 - a_{0,1}a_{1,0}a_{1,1} + a_{0,1}^2 a_{2,0} = f_5,$$

was die Invarianz von f_5 zeigt. Analog folgt dies für die restlichen fundamentalen Invarianten. ◁

Der Fall $n = 3$

Erneut benötigen wir zunächst die Darstellungsmatrix von $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R}))$, nun bzgl. der Basis

$$B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0}, b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0}).$$

Da uns auch hier die Darstellungsmatrix für $n = 2$ bereits bekannt ist, genügt es erneut, nur die Darstellungsmatrix der linearen Unterdarstellung $\rho^{(3)} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_3(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B_3 = (b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0})$ zu bestimmen. Mit $\mathcal{Z} = (z_{i,j})_{1 \leq i,j \leq 2}$ gilt zunächst:

$$\begin{aligned}
\rho_{\mathcal{Z}}(b_{0,3})(x,y) &= b_{0,3} \circ T_{\mathcal{Z}}(x,y) = b_{0,3}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{2,1}x + z_{2,2}y)^3 = z_{2,1}^3x^3 + 3z_{2,1}z_{2,2}^2xy^2 + 3z_{2,1}^2z_{2,2}x^2y + z_{2,2}^3y^3 \\
&= (z_{2,2}^3 b_{0,3} + 3z_{2,1}z_{2,2}^2 b_{1,2} + 3z_{2,1}^2 z_{2,2} b_{2,1} + z_{2,1}^3 b_{3,0})(x,y) \\
\rho_{\mathcal{Z}}(b_{1,2})(x,y) &= b_{1,2} \circ T_{\mathcal{Z}}(x,y) = b_{1,2}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}x + z_{1,2}y) \cdot (z_{2,1}x + z_{2,2}y)^2 \\
&= (z_{1,1}x + z_{1,2}y) \cdot (z_{2,1}^2x^2 + 2z_{2,1}z_{2,2}xy + z_{2,2}^2y^2) \\
&= z_{1,1}z_{2,1}^2x^3 + 2z_{1,1}z_{2,1}z_{2,2}x^2y + z_{1,1}z_{2,2}^2xy^2 \\
&\quad + z_{1,2}z_{2,1}^2x^2y + 2z_{1,2}z_{2,1}z_{2,2}xy^2 + z_{1,2}z_{2,2}^2y^3 \\
&= (z_{1,2}z_{2,2}^2 b_{0,3} + (z_{1,1}z_{2,2}^2 + 2z_{1,2}z_{2,1}z_{2,2})b_{1,2} + \\
&\quad (2z_{1,1}z_{2,1}z_{2,2} + z_{1,2}z_{2,1}^2)b_{2,1} + z_{1,1}z_{2,1}^2 b_{3,0})(x,y) \\
\rho_{\mathcal{Z}}(b_{2,1})(x,y) &= b_{2,1} \circ T_{\mathcal{Z}}(x,y) = b_{2,1}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}x + z_{1,2}y)^2 \cdot (z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}^2x^2 + 2z_{1,1}z_{1,2}xy + z_{1,2}^2y^2) \cdot (z_{2,1}x + z_{2,2}y) \\
&= z_{1,1}^2z_{2,1}x^3 + 2z_{1,1}z_{1,2}z_{2,1}x^2y + z_{1,2}^2z_{2,1}xy^2 \\
&\quad + z_{1,1}^2z_{2,2}x^2y + 2z_{1,1}z_{1,2}z_{2,2}xy^2 + z_{1,2}^2z_{2,2}y^3 \\
&= (z_{1,2}^2z_{2,2}b_{0,3} + (z_{1,2}^2z_{2,1} + 2z_{1,1}z_{1,2}z_{2,2})b_{1,2} + \\
&\quad (2z_{1,1}z_{1,2}z_{2,1} + z_{1,1}^2z_{2,2})b_{2,1} + z_{1,1}^2z_{2,1}b_{3,0})(x,y) \\
\rho_{\mathcal{Z}}(b_{3,0})(x,y) &= b_{3,0} \circ T_{\mathcal{Z}}(x,y) = b_{3,0}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y) \\
&= (z_{1,1}x + z_{1,2}y)^3 = z_{1,1}^3x^3 + 3z_{1,1}z_{1,2}^2xy^2 + 3z_{1,1}^2z_{1,2}x^2y + z_{1,2}^3y^3 \\
&= (z_{1,2}^3 b_{0,3} + 3z_{1,1}z_{1,2}^2 b_{1,2} + 3z_{1,1}^2 z_{1,2} b_{2,1} + z_{1,1}^3 b_{3,0})(x,y)
\end{aligned}$$

Damit erhalten wir erneut zuerst folgende Matrix $\mathcal{Q}_3$, deren Einträge Polynome im Polynomring $\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$ sind:

$$\mathcal{Q}_3 = \begin{pmatrix}
z_{2,2}^3 & z_{1,2}z_{2,2}^2 & z_{1,2}^2z_{2,2} & z_{1,2}^3 \\
3z_{2,1}z_{2,2}^2 & z_{1,1}z_{2,2}^2 + 2z_{1,2}z_{2,1}z_{2,2} & z_{1,2}^2z_{2,1} + 2z_{1,1}z_{1,2}z_{2,2} & 3z_{1,1}z_{1,2}^2 \\
3z_{2,1}^2z_{2,2} & 2z_{1,1}z_{2,1}z_{2,2} + z_{1,2}z_{2,1}^2 & 2z_{1,1}z_{1,2}z_{2,1} + z_{1,1}^2z_{2,2} & 3z_{1,1}^2z_{1,2} \\
z_{2,1}^3 & z_{1,1}z_{2,1}^2 & z_{1,1}^2z_{2,1} & z_{1,1}^3
\end{pmatrix}$$

Durch Berechnung der Normalform bzgl. $\mathcal{I}(G)$ für jedes Polynom der Matrix $\mathcal{Q}_3$ erhalten wir die Darstellungsmatrix von $\rho^{(3)}$ bzgl. B_3. Diese lautet wie folgt:

$$\mathcal{M}_{B_3}^{B_3}\left(\rho^{(3)}\right) = \begin{pmatrix} z_{2,2}^3 & -z_{2,1}z_{2,2}^2 & -z_{2,2}^3 + z_{2,2} & z_{2,1}z_{2,2}^2 - z_{2,1} \\ 3z_{2,1}z_{2,2}^2 & 3z_{2,2}^3 - 2z_{2,2} & -3z_{2,1}z_{2,2}^2 + z_{2,1} & -3z_{2,2}^3 + 3z_{2,2} \\ -3z_{2,2}^3 + 3z_{2,2} & 3z_{2,1}z_{2,2}^2 - z_{2,1} & 3z_{2,2}^3 - 2z_{2,2} & -3z_{2,1}z_{2,2}^2 \\ -z_{2,1}z_{2,2}^2 + z_{2,1} & -z_{2,2}^3 + z_{2,2} & z_{2,1}z_{2,2}^2 & z_{2,2}^3 \end{pmatrix}$$

Auf die Angabe der gesamten Darstellungsmatrix $\mathcal{M}_B^B(\rho) \in \mathrm{Mat}_{10}(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$ wollen wir hier verzichten. In abgekürzter Form sieht die Matrix wie folgt aus:

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} \mathcal{M}_{B_0}^{B_0}(\rho^{(0)}) & 0 & 0 & 0 \\ 0 & \mathcal{M}_{B_1}^{B_1}(\rho^{(1)}) & 0 & 0 \\ 0 & 0 & \mathcal{M}_{B_2}^{B_2}(\rho^{(2)}) & 0 \\ 0 & 0 & 0 & \mathcal{M}_{B_3}^{B_3}(\rho^{(3)}) \end{pmatrix}$$

Analog zu den vorherigen Fällen rechnen wir auch von nun an wieder über $\mathbb{Q}$. Ebenfalls bezeichnen wir die Einträge der 10×10-Matrix $\mathcal{M}_B^B(\rho)$ wieder mit $\tilde{q}_{i,j}$ und betrachten im Folgenden den Polynomring $Q := \mathbb{Q}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}, x_1, \ldots, x_{10}, y_1, \ldots, y_{10}]$ versehen mit der Eliminationsordnung $\tau := \mathrm{Elim}(\{z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}\})$. Auch hier ist es wieder ausreichend, die Polynome $h_i := y_i - \sum_{j=1}^{10} \tilde{q}_{i,j}x_j$ für $i \in \{7, \ldots, 10\}$ zu bestimmen. Diese lauten:

$$\begin{aligned} h_7 &= y_7 - z_{2,2}^3x_7 + z_{2,1}z_{2,2}^2x_8 + z_{2,2}^3x_9 - z_{2,2}x_9 - z_{2,1}z_{2,2}^2x_{10} + z_{2,1}x_{10}, \\ h_8 &= y_8 - 3z_{2,1}z_{2,2}^2x_7 - 3z_{2,2}^2x_8 + 2z_{2,2}x_8 + 3z_{2,1}z_{2,2}^2x_9 - z_{2,1}x_9 + 3z_{2,2}^3x_{10} - 3z_{2,2}x_{10}, \\ h_9 &= y_9 + 3z_{2,2}^3x_7 - 3z_{2,2}x_7 - 3z_{2,1}z_{2,2}^2x_8 + z_{2,1}x_8 - 3z_{2,2}^3x_9 + 2z_{2,2}x_9 + 3z_{2,1}z_{2,2}^2x_{10}, \\ h_{10} &= y_{10} + z_{2,1}z_{2,2}^2x_7 - z_{2,1}x_7 + z_{2,2}^3x_8 - z_{2,2}x_8 - z_{2,1}z_{2,2}^2x_9 - z_{2,2}^3x_{10} \end{aligned}$$

Wie wir bereits im letzten Abschnitt gesehen haben, beruhen beide Methoden auf einem Erzeugendensystem des Hilbert-Ideals. Die einzelnen Schritte lassen wir der Übersichtlichkeit wegen hier weg. Die τ-Gröbner-Basis des Hilbert-Ideals umfasst nun bereits 32 Polynome, die wir hier verständlicherweise nicht angeben wollen. Somit ist klar, dass der Derksen-Algorithmus ebenfalls 32 fundamentale Invarianten liefert, die in Anhang B nachzulesen sind. Dieselben Ergebnisse lassen sich auch mit der Linearen-Algebra-Methode erzielen, was wir aber nicht durchführen wollen.

Die Ergebnisse dieser Berechnung wollen wir nun ebenfalls auf den Polynomring P übertragen. Die Polynome $f_1, \ldots, f_{32}$ aus Anhang B erzeugen den Invariantenring $\mathbb{R}[x_1, \ldots, x_{10}]^G$. Sei nun $\Phi : \mathbb{R}[x_1, \ldots, x_{10}] \to P$ definiert durch

x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$\Phi(x_i)$	$a_{0,0}$	$a_{0,1}$	$a_{1,0}$	$a_{0,2}$	$a_{1,1}$	$a_{2,0}$	$a_{0,3}$	$a_{1,2}$	$a_{2,1}$	$a_{3,0}$

Dann erhalten wir Polynome $\Phi(f_1), \ldots, \Phi(f_{32})$, die den Invariantenring P^G erzeugen und die wir wieder mit $f_1, \ldots, f_{32}$ bezeichnen.

Satz 10.2.4. (Rotationsinvarianten für $n = 3$)
Sei $V = \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ mit Basis $B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0}, b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0})$ und sei $G = \mathrm{SO}_2(\mathbb{R})$. Dann wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring isomorphen Polynomrings P minimal erzeugt von den Bildern der 32 Polynome aus Abschnitt B.1 unter Φ.

Beweis: Folgt aus der Korrektheit von Algorithmus 7.2 (siehe Theorem 7.2.8) und der Korrektheit von Algorithmus 7.5 (siehe Theorem 7.2.14). □

An dieser Stelle können wir bereits eine Beobachtung machen, die so auch zu erwarten war: Die fundamentalen Invarianten bzgl. des reellen Vektorraums $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ sind auch fundamentale Invarianten bzgl. $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$. Somit lassen sich aus dem Erzeugendensystem bzgl. $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ die fundamentalen Invarianten bzgl. $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$ bzw. $\mathscr{P}_{\leq k}(\mathbb{R}^2, \mathbb{R})$ für $k \in \{0, 1, 2, 3\}$ ablesen. Es ist umgekehrt aber nicht möglich, aus den Erzeugendensystemen bzgl. $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$ mit $k \leq n$ ein Erzeugendensystem für $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu „kombinieren". Zu Ersterem betrachten wir noch ein Beispiel.

Beispiel 10.2.5. Sei $V = \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ und $G = \mathrm{SO}_2(\mathbb{R})$. Analog zum vorherigen Abschnitt sei $B = (b_{0,0}, \ldots, b_{3,0})$ die kanonische Basis des zehn-dimensionalen $\mathbb{R}$-Vektorraums $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$. Dann erhalten wir für $k \in \{0, \ldots, 3\}$ folgende (geordneten) Basen für die $\mathbb{R}$-Untervektorräume $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$:

k	0	1	2	3
Basis	$(b_{0,0})$	$(b_{0,1}, b_{1,0})$	$(b_{0,2}, b_{1,1}, b_{2,0})$	$(b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0})$

Für $k \in \{0, 1, 2, 3\}$ sei $\iota_k : \mathscr{P}_k(\mathbb{R}^2, \mathbb{R}) \to \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ die kanonische Einbettung. Beispielsweise ist ι_2 für $k = 2$ definiert durch

$$p \mapsto 0 \cdot b_{0,0} + 0 \cdot b_{0,1} + 0 \cdot b_{1,0} + p + 0 \cdot b_{0,3} + 0 \cdot b_{1,2} + 0 \cdot b_{2,1} + 0 \cdot b_{3,0}.$$

und die zugehörige Koordinatenabbildung $\iota_2^* : \mathbb{R}[x_1, \ldots, x_{10}] \to \mathbb{R}[x_4, x_5, x_6]$ ist dann definiert durch

$$x_i \mapsto \begin{cases} x_i, & i \in \{4, 5, 6\}, \\ 0, & \text{sonst.} \end{cases}$$

Damit erhalten wir aus den fundamentalen Invarianten $f_1, \ldots, f_{32}$ von $\mathbb{R}[\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})]^G$ fundamentalen Invarianten von $\mathbb{R}[\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})]^G$ durch Anwendung der jeweiligen Koordinatenabbildung ι_k^*:

k	fundamentale Invarianten
0	$f_1 = x_1$
1	$f_3 = x_2^2 + x_3^2$
2	$f_2 = x_4 + x_6$
	$f_4 = x_4 x_6 - \frac{1}{4} x_5^2$
3	$f_5 = x_7 x_9 - \frac{1}{3} x_8^2 + x_8 x_{10} - \frac{1}{3} x_9^2$,
	$f_6 = x_7^2 + x_7 x_9 + x_8 x_{10} + x_{10}^2$
	$f_{23} = x_7^3 x_{10} - \frac{1}{3} x_7^2 x_8 x_9 + \frac{2}{27} x_7 x_8^3 + \frac{1}{3} x_7 x_8^2 x_{10} - \frac{2}{9} x_7 x_8 x_9^2$ $-\frac{1}{3} x_7 x_9^2 x_{10} - x_7 x_{10}^3 + \frac{1}{27} x_8^3 x_9 + \frac{2}{9} x_8^2 x_9 x_{10} - \frac{1}{27} x_8 x_9^3 + \frac{1}{3} x_8 x_9 x_{10}^2 - \frac{2}{27} x_9^3 x_{10}$
	$f_{24} = x_7^2 x_{10}^2 - \frac{2}{3} x_7 x_8 x_9 x_{10} + \frac{4}{27} x_7 x_9^3 + \frac{4}{27} x_8^3 x_{10} - \frac{1}{27} x_8^2 x_9^2$

Auf analoge Weise können wir die Einbettung $\iota_{\leq 2} : \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \hookrightarrow \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ betrachten. Die zugehörige Koordinatenabbildung $\iota_{\leq 2}^* : \mathbb{R}[x_1, \ldots, x_{10}] \to \mathbb{R}[x_1, \ldots, x_6]$ ist also definiert durch

$$x_i \mapsto \begin{cases} x_i, & i \in \{1, \ldots, 6\}, \\ 0, & \text{sonst.} \end{cases}$$

Dann liefert $\iota_{\leq 2}^*$ die bekannten fundamentalen Invarianten f_1, f_2, f_3, f_4, f_9 und f_{12} des Invariantenrings $\mathbb{R}[x_1, \ldots, x_6]^G$ bzgl. der kanonischen Basis. ◁

10.3 Invarianten der orthogonalen Gruppe

Für die Praxis am bedeutendsten sind zwar Invarianten unter der Operation der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$. Allerdings wollen wir es nicht bei Rotationen belassen, sondern auch Drehspiegelungen, d.h. die orthogonale Gruppe $G = \mathrm{O}_2(\mathbb{R})$ betrachten. Auf diese Weise erhalten wir weniger fundamentale Invarianten, die in vielen Anwendungen für eine Korrespondenzfindung bereits ausreichen können. Bekanntlich hat G als Matrixgruppe die Form

$$\mathrm{O}_2(\mathbb{R}) = \{(z_{i,j})_{1\leq i,j\leq 2} : \ (z_{i,j}) \cdot (z_{i,j})^{\mathrm{tr}} = \mathcal{I}_2\},$$

d.h. um den Bezug zur bisher behandelten speziellen orthogonalen Gruppe herzustellen, gilt also $\mathrm{SO}_2(\mathbb{R}) = \mathrm{O}_2(\mathbb{R}) \cap \mathrm{SL}_2(\mathbb{R})$. Analog zum vorherigen Abschnitt lässt sich das Verschwindungsideal der orthogonalen Gruppe sofort an der Mengenbeschreibung ablesen. Für $G = \mathrm{O}_2(\mathbb{R})$ gilt also

$$\mathcal{I}(G) = \langle z_{1,1}^2 + z_{1,2}^2 - 1,\ z_{1,1}z_{2,1} + z_{1,2}z_{2,2},\ z_{2,1}^2 + z_{2,2}^2 - 1\rangle$$

in $\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$. Besonders viel ist ansonsten in den vorausgegangenen Abschnitten nicht zu ändern. Die Gruppe operiert natürlich auf exakt dieselbe Art und Weise auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, aber die Darstellungsmatrix $\mathcal{M}_B^B(\rho)$ der rationalen Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ von G in $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bzgl. der Standardtermbasis B ändert sich mit dem veränderten Verschwindungsideal. Wir werden nun analog zum vorherigen Abschnitt die fundamentalen Invarianten von Polynomfunktionen für $n \in \{1, 2, 3\}$ angeben, allerdings in einer verkürzten Darstellung ohne Angabe der Einzelschritte der betrachteten Algorithmen, da die wesentlichen Schritte identisch zum vorherigen Abschnitt sind.

Der Fall $n = 1$

Zur Bestimmung der Darstellungsmatrix der rationalen Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B = (b_{0,0}, b_{0,1}, b_{1,0})$ können wir zunächst völlig analog zu Abschnitt 10.2 vorgehen und erhalten dieselbe Matrix $\mathcal{Q}_1 = (q_{i,j})_{1\leq i,j\leq 3} \in \mathrm{Mat}_3(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}])$ wie in Abschnitt 10.2 (siehe Seite 205). Nun ist jedoch für jeden Eintrag von $\mathcal{Q}_1$ die Normalform bzgl. des neuen Verschwindungsideals zu bestimmen und bzgl. dieses Ideals bleibt die Matrix $\mathcal{Q}_1$ unverändert, d.h. es gilt

$$\mathcal{M}_B^B(\rho) = \mathcal{Q}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & z_{2,2} & z_{1,2} \\ 0 & z_{2,1} & z_{1,1} \end{pmatrix}.$$

Beide Algorithmen liefern darauf aufbauend das wenig überraschende Ergebnis, das sich vom Ergebnis aus Abschnitt 10.2 nicht unterscheidet. Die fundamentalen Invarianten sind $f_1 = x_1$ und $f_2 = x_2^2 + x_3^2$. Übertragen auf den Polynomring P können wir dieses Ergebnis in folgendem Satz festhalten.

Satz 10.3.1. (Invarianten der orthogonalen Gruppe für $n = 1$)
Sei $V = \mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R})$ mit Basis $B = (b_{0,0}, b_{0,1}, b_{1,0})$ und sei $G = \mathrm{O}_2(\mathbb{R})$. Dann wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring $\mathbb{R}[V]$ isomorphen Polynomrings $P = \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}]$ minimal von den Polynomen $f_1 = a_{0,0}$ und $f_2 = a_{0,1}^2 + a_{1,0}^2$ erzeugt. Diese Erzeuger sind zudem algebraisch unabhängig.

Beweis: Folgt aus der Korrektheit von Algorithmus 7.2 aus Theorem 7.2.8 und von Algorithmus 7.5 aus Theorem 7.2.14. □

Der Fall $n = 2$

Auch hier gehen wir zur Bestimmung der Darstellungsmatrix von $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0})$ analog zu Abschnitt 10.2 vor und bestimmen zunächst die Darstellungsmatrix der linearen Unterdarstellung $\rho^{(2)} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_2(\mathbb{R}^2, \mathbb{R}))$ bzgl. der Basis $B_2 = (b_{0,2}, b_{1,1}, b_{2,0})$. Dabei erhalten wir als erstes Zwischenergebnis eine zu Abschnitt 10.2 (siehe Seite 207) identische Matrix $\mathcal{Q}_2 \in \mathrm{Mat}_3(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}])$. Durch Berechnung der Normalformen bzgl. $\mathcal{I}(G)$ erhalten wir die Darstellungsmatrix $\mathcal{M}_{B_2}^{B_2}(\rho^{(2)})$ und somit die Darstellungsmatrix von ρ bzgl. B in $\mathrm{Mat}_6(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$:

$$\mathcal{M}_B^B(\rho) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & z_{1,2} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{1,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & z_{1,2}z_{2,2} & -z_{2,2}^2 + 1 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & z_{1,2}z_{2,1} + z_{1,1}z_{2,2} & -2z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2 + 1 & -z_{1,2}z_{2,2} & z_{2,2}^2 \end{pmatrix}$$

Beide Algorithmen liefern die folgenden fundamentalen Invarianten als minimales Erzeugendensystem von $\mathbb{Q}[x_1, \ldots, x_6]^G$:

$$f_1 = x_1 \qquad f_2 = x_2^2 + x_3^2 \qquad f_3 = x_4 + x_6$$
$$f_4 = x_4x_6 - \frac{1}{4}x_5^2 \qquad f_5 = x_2^2x_6 - x_2x_3x_5 + x_3^2x_4$$

Im Vergleich zur speziellen orthogonalen Gruppe (siehe Abschnitt 10.2) stimmen also bis auf eine alle Invarianten überein, nur die fundamentale Invariante

$$f_6 = x_2^2x_5 - 2x_2x_3x_4 + 2x_2x_3x_6 - x_3^2x_5$$

ist „verloren gegangen", was nichts anderes bedeutet, als dass dieses Polynom nicht mehr invariant unter der Operation der orthogonalen Gruppe ist. Davon wollen wir uns explizit überzeugen.

Beispiel 10.3.2. Wie wir wissen, operiert die Gruppe G auf $\mathbb{Q}[V]$ durch $(\mathcal{A}^{-1}, f) \mapsto f^{\mathcal{A}^{-1}}$, wobei $f^{\mathcal{A}^{-1}} \in \mathbb{Q}[V]$ definiert ist durch $f^{\mathcal{A}^{-1}}(p) = f(p \circ T_{\mathcal{A}})$. Diese Darstellung lässt sich nahtlos auf die Koordinaten übertragen. Sei $v \in \mathbb{Q}^6$ der Koordinatenvektor von $p \in \mathscr{P}_{\leq 2}(\mathbb{Q}^2, \mathbb{Q})$ bzgl. der Basis B. Dann operiert G auf V durch die Darstellungsmatrix des Automorphismus ρ, d.h. für $f \in \mathbb{Q}[x_1, \ldots, x_6]$ ist $f^{\mathcal{A}^{-1}} \in \mathbb{Q}[x_1, \ldots, x_6]$ definiert durch

$$f^{\mathcal{A}^{-1}}(v) = f(\mathcal{M}_B^B(\rho_{\mathcal{A}}) \cdot v).$$

Dies lässt sich analog auf allgemeine Weise mit der Darstellungsmatrix der linearen Darstellung $\rho : G \to \mathrm{Aut}_{\mathbb{Q}}(\mathscr{P}_{\leq 2}(\mathbb{Q}^2, \mathbb{Q}))$ beschreiben. Zunächst gilt

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} := \mathcal{M}_B^B(\rho) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} x_1 \\ z_{2,2}x_2 + z_{1,2}x_3 \\ z_{2,1}x_2 + z_{1,1}x_3 \\ z_{2,2}^2x_4 + z_{1,2}z_{2,2}x_5 - z_{2,2}^2x_6 + x_6 \\ 2z_{2,1}z_{2,2}x_4 + (z_{1,2}z_{2,1} + z_{1,1}z_{2,2})x_5 - 2z_{2,1}z_{2,2}x_6 \\ -z_{2,2}^2x_4 + x_4 - z_{1,2}z_{2,2}x_5 + z_{2,2}^2x_6 \end{pmatrix}$$

Nun betrachten wir das Polynom $\tilde{f}_6 = y_2^2y_5 - 2y_2y_3y_4 + 2y_2y_3y_6 - y_3^2y_5$. Indem wir wie in der obigen Gleichung angegeben die Unbestimmten y_i ersetzen, erhalten wir ein Polynom $\tilde{f}_6$ im Polynomring $Q := \mathbb{Q}[x_1, \ldots, x_6, z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]$, den wir mit der Termordnung $\sigma = \mathrm{DegRevLex}$

versehen. Sei $I := \mathcal{I}(\mathrm{O}_2(\mathbb{Q}))$ und $J := \mathcal{I}(\mathrm{SO}_2(\mathbb{Q}))$. Nun ist f_6 genau dann invariant, wenn $\mathrm{NF}_{\sigma,IQ}(\tilde{f}_6) = f_6$ gilt. Allerdings gilt:

$$\begin{aligned}\mathrm{NF}_{\sigma,IQ}(\tilde{f}_6) = {} & 2x_2x_3x_4z_{1,2}z_{2,1} - x_2^2x_5z_{1,2}z_{2,1} + x_3^2x_5z_{1,2}z_{2,1} - 2x_2x_3x_6z_{1,2}z_{2,1} \\ & - 2x_2x_3x_4z_{1,1}z_{2,2} + x_2^2x_5z_{1,1}z_{2,2} - x_3^2x_5z_{1,1}z_{2,2} + 2x_2x_3x_6z_{1,1}z_{2,2}\end{aligned}$$

und wie wir bereits wussten $\mathrm{NF}_{\sigma,JQ}(\tilde{f}_6) = f_6$. Dies zeigt, dass in der Tat f_6 invariant unter $\mathrm{SO}_2(\mathbb{Q})$ ist, jedoch nicht unter $\mathrm{O}_2(\mathbb{Q})$. ◁

Auch hier wollen wir die Relationen unter den fundamentalen Invarianten $f_1, \ldots, f_5$ untersuchen. Wie sich leicht berechnen lässt, ist

$$\mathrm{Rel}(f_1, \ldots, f_5) = \{h \in \mathbb{Q}[y_1, \ldots, y_5] : \ h(f_1, \ldots, f_5) = 0\}$$

das Nullideal, d.h. die Erzeuger $f_1, \ldots, f_5$ sind algebraisch unabhängig. Wir übertragen nun erneut das Ergebnis der Berechnungen auf P mit der oben angebenen Notation der Unbestimmten durch Doppelindizes. Dies liefert folgenden Satz.

Satz 10.3.3. (Invarianten der orthogonalen Gruppe für $n = 2$)
Sei $V = \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ mit Basis $B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0})$ und sei $G = \mathrm{O}_2(\mathbb{R})$. Dann wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring $\mathbb{R}[V]$ isomorphen Polynomrings $P = \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}, a_{0,2}, a_{1,1}, a_{2,0}]$ durch folgende Polynome erzeugt:

$$\begin{aligned} &f_1 = a_{0,0} && f_2 = a_{0,1}^2 + a_{1,0}^2 && f_3 = a_{0,2} + a_{2,0} \\ &f_4 = a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2 && f_5 = a_{0,1}^2a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2a_{0,2} &&\end{aligned}$$

Diese fundamentalen Invarianten sind zudem algebraisch unabhängig.

Beweis: Folgt aus der Korrektheit von Algorithmus 7.2 aus Theorem 7.2.8 und von Algorithmus 7.5 aus Theorem 7.2.14). Die algebraische Unabhängigkeit der Erzeuger lässt sich leicht nachrechnen. □

Es ist hier ebenso wie zuvor bei der speziellen orthogonalen Gruppe zu beobachten, dass sich leider ähnlich wie für Vektorinvarianten kein „Muster" oder gar eine von n abhängige „Formel" erkennen lässt, wie die fundamentalen Invarianten aufgebaut sind. Jedoch sind bei Wahl der kanonischen Basis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ in der oben verwendeten Ordnung für $n \geq 2$ stets folgende Polynome unter den fundamentalen Invarianten von $\mathbb{R}[\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})]^G$ sowohl für $G = \mathrm{O}_2(\mathbb{R})$ als auch $G = \mathrm{SO}_2(\mathbb{R})$ zu finden:

$$\begin{aligned} &f_1 = a_{0,0}, && f_2 = a_{0,2} + a_{2,0}, \\ &f_3 = a_{0,1}^2 + a_{1,0}^2, && f_4 = a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2 \end{aligned}$$

Die Bedeutung und die Invarianz dieser Polynome lässt sich auf anschauliche Weise interpretieren. Aus dieser anschaulichen Interpretation heraus kann Pisinger in [Pis02] ein Kriterium dafür angeben, wann zwei Polynomfunktionen $p, q \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ in einer Bahn unter der Operation der speziellen orthogonalen Gruppe sind (siehe Beispiel 11.1.5), allerdings nur unter speziellen Voraussetzungen.

Bemerkung 10.3.4. (Interpretation bestimmter fundamentaler Invarianten)
Sei $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und $G = \mathrm{O}_2(\mathbb{R})$ oder $G = \mathrm{SO}_2(\mathbb{R})$. Die nachfolgenden Überlegungen sind für beide Gruppen korrekt. Offensichtlich erscheint, dass $p(0,0) = a_{0,0}$ unter der Operation

von G invariant ist, was letztendlich daran liegt, dass die Untervektorräume $\mathscr{P}_k(\mathbb{R}^2, \mathbb{R})$ für alle $k \in \{0, \ldots, n\}$ G-stabil sind. Beschränken wir uns kurz auf $n = 2$, so ist $p \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ definiert durch

$$p(x, y) = a_{0,0} + a_{0,1}y + a_{1,0}x + a_{0,2}y^2 + a_{1,1}xy + a_{2,0}x^2.$$

Bekanntlich ist p als Polynomfunktion beliebig oft stetig partiell differenzierbar. Wir erhalten somit folgende partiellen Ableitungen:

$$\frac{\partial p}{\partial x}(x, y) = a_{1,0} + a_{1,1}y + 2a_{2,0}x, \qquad \frac{\partial p}{\partial y}(x, y) = a_{0,1} + 2a_{0,2}y + a_{1,1}x,$$
$$\frac{\partial^2 p}{\partial x^2}(x, y) = 2a_{2,0}, \qquad \frac{\partial^2 p}{\partial y^2}(x, y) = 2a_{0,2}, \qquad \frac{\partial^2 p}{\partial x \partial y}(x, y) = a_{1,1}$$

Allgemein folgt somit für die ersten partiellen Ableitungen von $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ im Nullpunkt:

$$p(0, 0) = a_{0,0}, \qquad \frac{\partial p}{\partial x}(0, 0) = a_{1,0}, \qquad \frac{\partial p}{\partial y}(0, 0) = a_{0,1},$$
$$\frac{\partial^2 p}{\partial x^2}(0, 0) = 2a_{2,0}, \qquad \frac{\partial^2 p}{\partial y^2}(0, 0) = 2a_{0,2}, \qquad \frac{\partial^2 p}{\partial x \partial y}(0, 0) = a_{1,1}.$$

Somit ist der Gradient von p im Nullpunkt durch

$$\operatorname{grad} p(0, 0) = \left(\frac{\partial p}{\partial x}(0, 0), \frac{\partial p}{\partial y}(0, 0) \right) = (a_{1,0}, a_{0,1})$$

gegeben und die Hesse-Matrix von p im Nullpunkt durch

$$\mathcal{H}_p(0, 0) = \begin{pmatrix} \frac{\partial^2 p}{\partial x^2}(0, 0) & \frac{\partial^2 p}{\partial x \partial y}(0, 0) \\ \frac{\partial^2 p}{\partial x \partial y}(0, 0) & \frac{\partial^2 p}{\partial y^2}(0, 0) \end{pmatrix} = \begin{pmatrix} 2a_{2,0} & a_{1,1} \\ a_{1,1} & 2a_{0,2} \end{pmatrix}.$$

Damit bekommen die Invarianten f_2, f_3 und f_4 eine sehr anschauliche Interpretation: Das Polynom f_3 korrespondiert genau mit dem Quadrat der Länge des Gradienten im Nullpunkt:

$$\|\operatorname{grad} p(0, 0)\|^2 = a_{0,1}^2 + a_{1,0}^2.$$

Diese Länge ist ein Maß für die Steilheit des Anstiegs im Nullpunkt. Es ist mit der Anschauung durchaus im Einklang, dass diese Länge durch die Operation von G invariant bleiben sollte. Weiter stimmt f_2 genau mit der Hälfte der Spur der Hesse-Matrix von p im Nullpunkt überein:

$$\frac{1}{2} \cdot \operatorname{Spur}(\mathcal{H}_p(0, 0)) = \frac{1}{2} \cdot (2a_{0,2} + 2a_{2,0}) = a_{0,2} + a_{2,0}$$

und f_4 entspricht genau einem Viertel der Determinante dieser Hesse-Matrix:

$$\frac{1}{4} \cdot \det(\mathcal{H}_p(0, 0)) = \frac{1}{4} \cdot (4a_{0,2}a_{2,0} - a_{1,1}^2) = a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2.$$

Die Hesse-Matrix $\mathcal{H}_p(0, 0) \in \operatorname{Mat}_2(\mathbb{R})$ ist eine symmetrische reelle Matrix und als solche diagonalisierbar. Damit sind die reellen Eigenwerte $\lambda_1, \lambda_2 \in \mathbb{R}$ von $\mathcal{H}_p(0, 0)$ eindeutig festgelegt. Das charakteristische Polynom von $\mathcal{H}_p(0, 0)$ lautet:

$$\chi_{\mathcal{H}_p(0,0)}(t) = t^2 - \operatorname{Spur}(\mathcal{H}_p(0, 0)) \cdot t + \det(\mathcal{H}_p(0, 0)).$$

Gemäß dem Satz von Vieta gilt $\operatorname{Spur}(\mathcal{H}_p(0, 0)) = \lambda_1 + \lambda_2$ und $\det(\mathcal{H}_p(0, 0)) = \lambda_1 \cdot \lambda_2$. Da Spur und Determinante der Hesse-Matrix invariant sind, sind folglich auch die Eigenwerte der

Hesse-Matrix im Nullpunkt invariant unter der Operation der orthogonalen, und damit auch der speziellen orthogonalen Gruppe. Weiter gilt

$$\frac{\mathrm{Spur}(\mathcal{H}_p(0,0))^2}{\det(\mathcal{H}_p(0,0)} = \frac{(\lambda_1+\lambda_2)^2}{\lambda_1\cdot\lambda_2} = \frac{\lambda_1^2+2\lambda_1\lambda_2+\lambda_2^2}{\lambda_1\lambda_2} = \frac{\lambda_1}{\lambda_2}+\frac{\lambda_2}{\lambda_1}+2,$$

d.h. die rationale Invariante $\frac{(a_{0,2}+a_{2,0})^2}{a_{0,2}a_{2,0}-\frac{1}{4}a_{1,1}}$ korrespondiert auf diese Weise mit den Eigenwertverhältnissen. Außerdem gilt

$$\mathrm{Spur}(\mathcal{H}_p(0,0))^2 - 4\det(\mathcal{H}_p(0,0)) = (\lambda_1+\lambda_2)^2 - 4\lambda_1\lambda_2 = (\lambda_1-\lambda_2)^2,$$

d.h. die Invariante

$$\begin{aligned}\mathrm{Spur}(\mathcal{H}_p(0,0))^2 - 4\det(\mathcal{H}_p(0,0)) &= 4(a_{0,2}+a_{2,0})^2 - 16(a_{0,2}a_{2,0}-\tfrac{1}{4}a_{1,1}^2)\\ &= 4\cdot\left[(a_{0,2}+a_{2,0})^2 - 4(a_{0,2}a_{2,0}-\tfrac{1}{4}a_{1,1}^2)\right]\\ &= 4\cdot\left[(a_{0,2}-a_{2,0})^2 + a_{1,1}^2\right]\end{aligned}$$

korrespondiert mit der Differenz der Eigenwerte. Aus diesen anschaulichen Überlegungen heraus konstruiert Pisinger in [Pis02] ein Distanzmaß (siehe Beispiel 11.1.5), deren Bestandteile wir quasi „nebenbei" erhalten haben.

Kehrt man die Sichtweise um, so haben wir in den letzten Abschnitten quasi „nebenbei" mit Methoden der Invariantentheorie die wohl bekannte Aussage bewiesen, dass für eine Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2,\mathbb{R})$ mit $n \geq 2$ der konstante Term $p(0,0)$, die Länge des Gradienten $\mathrm{grad}\, p(0,0)$ im Nullpunkt und die Spur sowie Determinante der Hesse-Matrix $\mathcal{H}_p(0,0)$ im Nullpunkt invariant unter der Operation der orthogonalen Gruppe $\mathrm{O}_2(\mathbb{R})$, und damit insbesondere unter der Operation der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$, sind.

Der Fall $n = 3$

Wir wählen auch in diesem Fall ein analoges Vorgehen und bestimmen zunächst die Darstellungsmatrix der linearen Unterdarstellung $\rho^{(3)} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_3(\mathbb{R}^2,\mathbb{R}))$, der linearen Darstellung von G in $\mathscr{P}_3(\mathbb{R}^2,\mathbb{R})$ bzgl. der Basis $B_3 = (b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0})$. Erneut erhalten wir eine Darstellungsmatrix $\mathcal{Q}_3 \in \mathrm{Mat}_4(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}])$, die mit der Matrix $\mathcal{Q}_3$ aus Abschnitt 10.2 übereinstimmt. Die Darstellungsmatrix $\mathcal{M}_{B_3}^{B_3}(\rho^{(3)})$ von $\rho^{(3)}$ lautet dann wie folgt:

$$\mathcal{M}_{B_3}^{B_3}(\rho^{(3)}) = \begin{pmatrix} z_{2,2}^3 & z_{1,2}z_{2,2}^2 & -z_{2,2}^3+z_{2,2} & -z_{1,2}z_{2,2}^2+z_{1,2}\\ 3z_{2,1}z_{2,2}^2 & 3z_{1,1}z_{2,2}^2-2z_{1,1} & -3z_{2,1}z_{2,2}^2+z_{2,1} & -3z_{1,1}z_{2,2}^2+3z_{1,1}\\ -3z_{2,2}^3+3z_{2,2} & -3z_{1,2}z_{2,2}^2+z_{1,2} & 3z_{2,2}^3-2z_{2,2} & 3z_{1,2}z_{2,2}^2\\ -z_{2,1}z_{2,2}^2+z_{2,1} & -z_{1,1}z_{2,2}^2+z_{1,1} & z_{2,1}z_{2,2}^2 & z_{1,1}z_{2,2}^2 \end{pmatrix}$$

Mit der daraus resultierenden Darstellungsmatrix $\mathcal{M}_B^B(\rho)$ erhalten wir 19 fundamentale Invarianten, die in Anhang B nachzulesen sind. Auch dieses Ergebnis übertragen wir auf P mit der oben angebenen Notation der Unbestimmten durch Doppelindizes. Dies liefert uns erneut einen Satz, dessen Beweis sich abermals unmittelbar zum einen aus den vorausgegangenen Überlegungen und zum anderen aus der Korrektheit von Algorithmus 7.2 ergibt.

Satz 10.3.5. (Invarianten der orthogonalen Gruppe für $n = 3$)
Sei $V = \mathscr{P}_{\leq 3}(\mathbb{R}^2,\mathbb{R})$ mit Basis $B = (b_{0,0}, b_{0,1}, b_{1,0}, b_{0,2}, b_{1,1}, b_{2,0}, b_{0,3}, b_{1,2}, b_{2,1}, b_{3,0})$ und sei $G = \mathrm{O}_2(\mathbb{R})$. Dann wird der Invariantenring P^G als $\mathbb{R}$-Unteralgebra des zum Koordinatenring $\mathbb{R}[V]$ isomorphen Polynomrings P erzeugt von den Polynomen $f_1, \ldots, f_{19}$ aus Abschnitt B.2.

Beweis: Folgt aus der Korrektheit von Algorithmus 7.2 aus Theorem 7.2.8 und von Algorithmus 7.5 aus Theorem 7.2.14. □

10.4 Invarianten lokaler Bildmerkmale

Was wir in den letzten beiden Abschnitten berechnet haben, sind fundamentale Invarianten bzgl. der Standardtermbasis. Diese ist aber nicht orthogonal bzw. orthonormal bzgl. des Skalarproduktes, das wir auf den lokalen Pixelfenstern verwenden wollen. Wir benötigen also fundamentale Invarianten bzgl. einer geeigneten Orthogonal- bzw. Orthonormalbasis. Sei nun B wie zuvor die Standardtermbasis des reellen Vektorraums $V := \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit Dimension $\nu = \binom{n+2}{2}$ und sei $C = (c_{0,0}, c_{0,1}, c_{1,0}, \ldots, c_{n,0})$ eine geeignete Orthogonal- bzw. Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ (siehe dazu Abschnitt 9.2.2). Sei $\kappa_B : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^\nu$ die Koordinatenfunktion bzgl. B und analog $\kappa_C : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^\nu$ die Koordinatenfunktion bzgl. C. Dann lassen sich die Ergebnisse aus den vorigen Abschnitten durch Basiswechsel auf lokale Bildmerkmale anwenden. Dazu gibt es grundsätzlich zwei verschiedene Möglichkeiten:

(1) Mit der Basistransformationsmatrix $\mathcal{M}_B^C(\mathrm{id}_V) \in \mathrm{Mat}_\nu(\mathbb{R})$ kann jedes lokale Bildmerkmal $\kappa_C(p) \in \mathbb{R}^\nu$ in einen Koordinatenvektor $\kappa_B(p) \in \mathbb{R}^\nu$ bzgl. der Standardtermbasis umgerechnet werden. Dann lassen sich die Invarianten und Ergebnisse aus den letzten Abschnitten anwenden. Für die Theorie und die Darstellung in dieser Arbeit, ist diese Vorgehensweise zu vertreten, für die Praxis ist sie offensichtlich wenig empfehlenswert, da sehr viele Koordinatenwechsel nötig wären.

(2) Indem man die Darstellungsmatrix $\mathcal{M}_C^C(\rho) \in \mathrm{Mat}_\nu(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$ einer linearen Darstellung $\rho : G \to \mathrm{Aut}_\mathbb{R}(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}))$ von G in $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bestimmt, lassen sich die Invarianten von Polynomfunktionen bzgl. der Orthogonalbasis C auf analoge Weise wie oben bestimmen. Mit diesen Invarianten ist eine Umrechnung der lokalen Bildmerkmale in Koordinatenvektoren bzgl. B nicht mehr nötig.

Eine dritte, vermeintlich naheliegende Möglichkeit lässt sich leider nicht im Allgemeinen anwenden. Durch den Basiswechsel werden die Koordinaten bzgl. der einen Basis in die Koordinaten bzgl. der anderen Basis transformiert. Leider machen die Invarianten diese Transformation nur bedingt mit, wie das folgende Beispiel zeigt. Dennoch ist es unter bestimmten Voraussetzungen möglich, auf diese Weise die Invarianten bzgl. B in Invarianten bzgl. C „umzurechnen". Dazu werden wir später ein Beispiel betrachten.

Beispiel 10.4.1. Wir betrachten den Vektorraum $V = \mathscr{P}_{\leq 1}(\mathbb{R}^2, \mathbb{R})$ und die spezielle orthogonale Gruppe $G = \mathrm{SO}_2(\mathbb{R})$. Bzgl. der Standardtermbasis B wird der Invariantenring P^G mit $P = \mathbb{R}[a_{0,0}, a_{0,1}, a_{1,0}]$ von den folgenden Polynomen erzeugt (vgl. Satz 10.2.1):

$$f_1 = a_{0,0} \qquad\qquad f_2 = a_{0,1}^2 + a_{1,0}^2.$$

Nun betrachten wir die Basis $C = (c_{0,0}, c_{0,1}, c_{1,0})$ mit $c_{0,0}(x, y) = 1$, $c_{0,1}(x, y) = y + 1$ und $c_{1,0}(x, y) = x + y + 1$. Mit der Darstellungsmatrix

$$\mathcal{M}_C^C(\rho) = \begin{pmatrix} 1 & -z_{2,2} + 1 & z_{2,1} - z_{2,2} - 1 \\ 0 & -z_{2,1} + z_{2,2} & -2z_{2,1} \\ 0 & z_{2,1} & z_{2,1} + z_{2,2} \end{pmatrix}$$

erhalten wir eine einzige fundamentale Invariante $g = a_{0,1}^2 + 2a_{0,1}a_{1,0} + 2a_{1,0}^2$ für den Invariantenring P^G. Die beiden Basistransformationsmatrizen lauten wie folgt:

$$\mathcal{M}_C^B(\mathrm{id}_V) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad \mathcal{M}_B^C(\mathrm{id}_V) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Seien nun $\Phi : P \to P$ und $\Psi : P \to P$ mittels $\mathcal{M}_C^B(\mathrm{id}_V)$ bzw. $\mathcal{M}_B^C(\mathrm{id}_V)$ definiert durch

$$a_{0,0} \mapsto a_{0,0} - a_{0,1}, \qquad a_{0,1} \mapsto a_{0,1} - a_{1,0}, \qquad a_{1,0} \mapsto a_{1,0}$$

bzw.

$$a_{0,0} \mapsto a_{0,0} + a_{0,1} + a_{1,0}, \qquad a_{0,1} \mapsto a_{0,1} + a_{1,0}, \qquad a_{1,0} \mapsto a_{1,0}$$

Wie man leicht nachrechnen kann, sind Φ und Ψ beide bijektiv und es gilt $\Phi^{-1} = \Psi$ (siehe vgl. [KR00], Satz 3.6.12, S. 232). Weiter gilt:

$$\begin{aligned}\Phi(g) &= (a_{0,1} - a_{1,0})^2 + 2(a_{0,1} - a_{1,0})a_{1,0} + 2a_{1,0}^2 = a_{0,1}^2 + a_{1,0}^2 = f_2,\\ \Psi(f_2) &= (a_{0,1} + a_{1,0})^2 + a_{1,0}^2 = a_{0,1}^2 + 2a_{0,1}a_{1,0} + 2a_{1,0}^2 = g\end{aligned}$$

aber $\Psi(f_1) = a_{0,0} + a_{0,1} + a_{1,0}$ ist nicht invariant. ◁

Wie dieses Beispiel zeigt, bildet im Allgemeinen ein Algebra-Automorphismus die Invarianten nicht aufeinander ab, weshalb es im Allgemeinen leider auch nicht möglich ist, die Invarianten der Polynomfunktionen bzgl. B einfach in Invarianten bzgl. einer Orthogonalbasis „umzurechnen". Wir werden nun Beispiele von Invarianten lokaler Bildmerkmale angeben.

Beispiel 10.4.2. (Invarianten lokaler Bildmerkmale bzgl. $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$)
Wir betrachten den reellen Vektorraum $V := \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$. Sei $B = (b_{i,j} : \ i + j \leq 2)$ die Standardtermbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ und sei $C := (c_{i,j} : \ i, j \in \mathbb{N},\ i + j \leq 2)$ mit

$$\begin{aligned} c_{0,0}(x,y) &= 1 & c_{0,1}(x,y) &= y & c_{1,0}(x,y) &= x\\ c_{0,2}(x,y) &= y^2 - \tfrac{3}{4} & c_{1,1}(x,y) &= xy & c_{2,0}(x,y) &= x^2 - \tfrac{3}{4}\end{aligned}$$

die Orthogonalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ aus Beispiel 9.2.29, also bzgl. des in Beispiel 9.2.29 angegebenen Skalarprodukts. Die Basistransformationsmatrizen lauten somit wie folgt:

$$\mathcal{M}_C^B(\mathrm{id}_V) = \begin{pmatrix} 1 & 0 & 0 & \frac{3}{4} & 0 & \frac{3}{4} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad \mathcal{M}_B^C(\mathrm{id}_V) = \begin{pmatrix} 1 & 0 & 0 & -\frac{3}{4} & 0 & -\frac{3}{4} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

a) Sei zunächst $G = \mathrm{SO}_2(\mathbb{R})$ und $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}))$ die rationale Darstellung von G in $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$. Wir bestimmen hier nun die Darstellungsmatrix von ρ bzgl. C auf zwei verschiedene Arten: Zunächst ganz klassisch und anschließend unter Verwendung der Basistransformationsformel (siehe Bemerkung 4.3.28). Da nur zwei Basiselemente von C von denen aus B abweichen, reicht es, diese zu betrachten. Mit $\mathcal{Z} = (z_{i,j})_{1 \leq i,j \leq 2}$ gilt:

$$\begin{aligned}\rho_{\mathcal{Z}}(p_{0,2})(x,y) &= p_{0,2}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y)\\ &= (z_{2,1}x + z_{2,2}y)^2 - \tfrac{3}{4} = z_{2,1}^2x^2 + 2z_{2,1}z_{2,2}xy + z_{2,2}^2y^2 - \tfrac{3}{4}\\ &= \left(\tfrac{3}{4}(z_{2,1}^2 + z_{2,2}^2 - 1)p_{0,0} + z_{2,2}^2p_{0,2} + 2z_{2,1}z_{2,2}p_{1,1} + z_{2,1}^2p_{2,0}\right)(x,y),\\ \rho_{\mathcal{Z}}(p_{2,0})(x,y) &= p_{2,0}(z_{1,1}x + z_{1,2}y, z_{2,1}x + z_{2,2}y)\\ &= (z_{1,1}x + z_{1,2}y)^2 - \tfrac{3}{4} = z_{1,1}^2x^2 + 2z_{1,1}z_{1,2}xy + z_{1,2}^2y^2 - \tfrac{3}{4}\\ &= \left(\tfrac{3}{4}(z_{1,1}^2 + z_{1,2}^2 - 1)p_{0,0} + z_{1,2}^2p_{0,2} + 2z_{1,1}z_{1,2}p_{1,1}(x,y) + z_{1,1}^2p_{2,0}\right)(x,y)\end{aligned}$$

Somit erhalten wir zunächst folgende Matrix in $\mathrm{Mat}_6(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}])$:

$$\mathcal{Q} = \begin{pmatrix} 1 & 0 & 0 & \frac{3}{4}(z_{2,1}^2 + z_{2,2}^2 - 1) & 0 & \frac{3}{4}(z_{1,1}^2 + z_{1,2}^2 - 1) \\ 0 & z_{2,2} & z_{1,2} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{1,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & z_{1,2}z_{2,2} & z_{1,2}^2 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & z_{1,1}z_{2,2} + z_{1,2}z_{2,1} & 2z_{1,1}z_{1,2} \\ 0 & 0 & 0 & z_{2,1}^2 & z_{1,1}z_{2,1} & z_{1,1}^2 \end{pmatrix}$$

Durch Wahl der Termordnung DegRevLex und der Berechnung der Normalform jedes Eintrags der Matrix bzgl. dem Verschwindungsideal $\mathcal{I}(G)$ erhalten wir die Darstellungsmatrix der Gruppenoperation bzgl. der Basis C in $\mathrm{Mat}_6(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$:

$$\mathcal{M}_C^C(\rho) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & -z_{2,1} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & -z_{2,1}z_{2,2} & -z_{2,2}^2 + 1 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & 2z_{2,2}^2 - 1 & -2z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2 + 1 & z_{2,1}z_{2,2} & z_{2,2}^2 \end{pmatrix}.$$

Mit der Basistransformationsformel aus Bemerkung 4.3.28 folgt analog aus der Darstellungsmatrix $\mathcal{M}_B^B(\rho)$:

$$\begin{aligned} \mathcal{M}_C^C(\rho) &= \mathcal{M}_C^B(\mathrm{id}_V) \cdot \mathcal{M}_B^B(\rho) \cdot \mathcal{M}_B^C(\mathrm{id}_V) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & -z_{2,1} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & -z_{2,1}z_{2,2} & -z_{2,2}^2 + 1 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & 2z_{2,2}^2 - 1 & -2z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2 + 1 & z_{2,1}z_{2,2} & z_{2,2}^2 \end{pmatrix} \end{aligned}$$

Es lässt sich in jedem Fall festhalten, dass $\mathcal{M}_C^C(\rho) = \mathcal{M}_B^B(\rho)$ gilt. Somit ist klar, dass die Invarianten $f_1, \ldots, f_6$ aus Satz 10.2.2 bzgl. B auch Invarianten bzgl. C sind.

b) Sei nun $G = \mathrm{O}_2(\mathbb{R})$ und analog $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}))$ die rationale Darstellung von G in $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$. Es ist nicht schwer zu sehen, dass wir zunächst dieselbe Matrix $\mathcal{Q}$ wie in a) erhalten. Modulo dem Verschwindungsideal $\mathcal{I}(G)$ gilt

$$\mathcal{M}_C^C(\rho) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & z_{1,2} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{1,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & z_{1,2}z_{2,2} & -z_{2,2}^2 + 1 \\ 0 & 0 & 0 & 2z_{2,1}z_{2,2} & z_{1,2}z_{2,1} + z_{1,1}z_{2,2} & -2z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2 + 1 & -z_{1,2}z_{2,2} & z_{2,2}^2 \end{pmatrix},$$

d.h. auch hier stimmen $\mathcal{M}_B^B(\rho)$ und $\mathcal{M}_C^C(\rho)$ überein. Somit sind für die orthogonale Gruppe die Invarianten $f_1, \ldots, f_5$ aus Satz 10.3.3 bzgl. B auch Invarianten bzgl. der Orthogonalbasis C.

Bei Betrachtung der obigen Orthogonalbasis C ändern sich also die fundamentalen Invarianten nicht. Wir erhalten somit auch bzgl. dieser Orthogonalbasis die fundamentalen Invarianten

$$\begin{aligned} f_1 &= a_{0,0}, & f_4 &= a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2, \\ f_2 &= a_{0,1}^2 + a_{1,0}^2, & f_5 &= a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2} \\ f_3 &= a_{0,2} + a_{2,0}, & f_6 &= a_{0,1}^2 a_{1,1} - 2a_{0,1}a_{1,0}a_{0,2} + 2a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1} \end{aligned}$$

bzw. im Falle von $\mathrm{O}_2(\mathbb{R})$ nur die ersten fünf. ◁

Das Ergebnis aus dem letzten Beispiel lässt sich natürlich nicht verallgemeinern. Wie wir im nächsten Beispiel sehen werden, ändern sich die fundamentalen Invarianten bei Verwendung einer bekannten Orthonormalbasis durchaus.

Beispiel 10.4.3. (Invarianten lokaler Bildmerkmale bzgl. $\mathscr{P}_{\leq 2}(\mathbb{R}^2,\mathbb{R})$)
Auch hier betrachten wir den $\mathbb{R}$-Vektorraum $V := \mathscr{P}_{\leq 2}(\mathbb{R}^2,\mathbb{R})$. Analog zum letzten Beispiel sei $B = (b_{i,j} : i+j \leq 2)$ die Standardtermbasis von V und weiter sei $C = (c_{i,j} : i+j \leq 2)$ mit

$$\begin{aligned} &c_{0,0}(x,y) = \tfrac{1}{3} && c_{0,1}(x,y) = \tfrac{1}{\sqrt{6}}y && c_{1,0}(x,y) = \tfrac{1}{\sqrt{6}}x,\\ &c_{0,2}(x,y) = \tfrac{1}{\sqrt{2}}(y^2 - \tfrac{3}{4}) && c_{1,1}(x,y) = \tfrac{1}{2}xy && c_{2,0}(x,y) = \tfrac{1}{\sqrt{2}}(x^2 - \tfrac{3}{4}) \end{aligned}$$

die Orthonormalbasis von V aus Beispiel 9.2.29. Somit erhalten wir folgende Basistransformationsmatrizen:

$$\mathcal{M}_C^B(\mathrm{id}_V) = \begin{pmatrix} 3 & 0 & 0 & \frac{9}{4} & 0 & \frac{9}{4} \\ 0 & \sqrt{6} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{pmatrix} \text{ und } \mathcal{M}_B^C(\mathrm{id}_V) = \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{-3}{4\sqrt{2}} & 0 & \frac{-3}{4\sqrt{2}} \\ 0 & \frac{1}{\sqrt{6}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{6}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Sei nun $G = \mathrm{SO}_2(\mathbb{R})$ und $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(V)$ die rationale Darstellung von G in V. Mit der Basistransformationsformel erhalten wir aus der bekannten Darstellungsmatrix $\mathcal{M}_B^B(\rho)$ in $\mathrm{Mat}_6(\mathbb{R}[z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}]/\mathcal{I}(G))$ die Darstellungsmatrix bzgl. C:

$$\begin{aligned} \mathcal{M}_C^C(\rho) &= \mathcal{M}_C^B(\mathrm{id}_V) \cdot \mathcal{M}_B^B(\rho) \cdot \mathcal{M}_B^C(\mathrm{id}_V) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & z_{2,2} & -z_{2,1} & 0 & 0 & 0 \\ 0 & z_{2,1} & z_{2,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{2,2}^2 & -\frac{1}{\sqrt{2}}z_{2,1}z_{2,2} & -z_{2,2}^2 + 1 \\ 0 & 0 & 0 & 2\sqrt{2}z_{2,1}z_{2,2} & 2z_{2,2}^2 - 1 & -2\sqrt{2}z_{2,1}z_{2,2} \\ 0 & 0 & 0 & -z_{2,2}^2 + 1 & \frac{1}{\sqrt{2}}z_{2,1}z_{2,2} & z_{2,2}^2 \end{pmatrix} \end{aligned}$$

Somit ist klar, dass sich manche fundamentalen Invarianten ändern werden. Ebenfalls schnell ersichtlich ist, dass die ersten beiden fundamentalen Invarianten

$$f_1 = a_{0,0} \quad \text{und} \quad f_2 = a_{0,1}^2 + a_{1,0}^2$$

unverändert bleiben. In diesem Fall ist es möglich, die Invarianten bzgl. C aus den Invarianten bzgl. B zu berechnen, ohne erneut beispielsweise Algorithmus 7.2 zu bemühen. Sei $\Phi : P \to P$ definiert durch

$$a_{0,0} \mapsto a_{0,0},\ a_{0,1} \mapsto \tfrac{1}{\sqrt{6}}a_{0,1},\ a_{1,0} \mapsto \tfrac{1}{\sqrt{6}},\ a_{0,2} \mapsto \tfrac{1}{\sqrt{2}}a_{0,2},\ a_{1,1} \mapsto \tfrac{1}{2}a_{1,1},\ a_{2,0} \mapsto \tfrac{1}{\sqrt{2}}a_{2,0}.$$

Dann folgt für die letzten vier Invarianten bzgl. B:

$$\begin{aligned} \Phi(f_3) &= \tfrac{1}{\sqrt{2}} \cdot (a_{0,2} + a_{2,0}),\\ \Phi(f_4) &= \tfrac{1}{2}a_{0,2}a_{2,0} - \tfrac{1}{16}a_{1,1}^2,\\ \Phi(f_5) &= \tfrac{1}{6\sqrt{2}}a_{0,1}^2a_{2,0} - \tfrac{1}{12}a_{0,1}a_{1,0}a_{1,1} + \tfrac{1}{6\sqrt{2}}a_{1,0}^2a_{0,2},\\ \Phi(f_6) &= \tfrac{1}{12}a_{0,1}^2a_{1,1} - \tfrac{1}{3\sqrt{2}}a_{0,1}a_{1,0}a_{0,2} + \tfrac{2}{3\sqrt{2}}a_{0,1}a_{1,0}a_{2,0} - \tfrac{1}{12}a_{1,0}^2a_{1,1}. \end{aligned}$$

Wir normieren die berechneten Polynome und erhalten zusammen mit f_1 und f_2 folgende weiteren fundamentalen Invarianten bzgl. der Orthonormalbasis C:

$$
\begin{aligned}
f_3 &= a_{0,2} + a_{2,0}, \\
f_4 &= a_{0,2}a_{2,0} - \tfrac{1}{8}a_{1,1}^2, \\
f_5 &= a_{0,1}^2 a_{2,0} - \tfrac{1}{2}\sqrt{2}a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}, \\
f_6 &= a_{0,1}^2 a_{1,1} - 2\sqrt{2}a_{0,1}a_{1,0}a_{0,2} + 2\sqrt{2}a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1}
\end{aligned}
$$

◁

KAPITEL 11

Korrespondenzfindung lokaler Bildmerkmale

Jimmy Carter[26]

Die Theorie ist eine Vermutung mit Hochschulbildung.

Das Zitat des 39. Präsidenten der Vereinigten Staaten von Amerika (1977–1981) bringt eines treffend auf den Punkt: Auch die schönste Theorie bleibt irgendwie so lange fruchtlos und wertlos, bis sich eine Anwendungsmöglichkeit für sie auftut. Dieses „Problem" stellt sich für die Invariantentheorie in keinster Weise! Wie wir bereits in der Einleitung dargelegt haben, ist gerade die Bildverarbeitung bzw. das Rechnersehen ein sehr fruchtbarer Boden für Anwendungen der Invariantentheorie. Und es gibt noch zahlreiche weitere Anwendungsfelder der Invariantentheorie. Wir wollen in diesem Kapitel, insbesondere im ersten Abschnitt, nun eine neue Anwendungsmöglichkeit der Invariantentheorie auf die lokalen Bildmerkmale des letzten Kapitels vorstellen und anschließend exemplarisch anwenden.

Zu Beginn dieser Arbeit sind wir in der Einleitung auf die verschiedenen Transformationen eingegangen, die lokal auf Bildern operieren, die sogenannten geometrischen Transformationen, d.h. folgende Menge von Abbildungen:

$$\mathrm{AGL}_2(\mathbb{R}) = \{T : \mathbb{R}^2 \to \mathbb{R}^2 : \text{ Es gibt } \mathcal{A} \in \mathrm{GL}_2(\mathbb{R}), v \in \mathbb{R}^2 \text{ mit } T(x) = \mathcal{A} \cdot x + v\}.$$

Im letzten Kapitel haben wir nun gesehen, wie die Operation dieser Gruppe auf ein Grauwertbild $\mathrm{gv} : S \to \mathbb{Z}_{0,255}$ mit $S \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ für $r, s \in \mathbb{N}_+$ im Detail vonstatten geht. Hinter jedem Grauwert $\mathrm{gv}_{i,j}$ eines Pixels $(i, j) \in S$ steckt – in vereinfachter Form – der quantisierte Wert des Analogsignals $\ell_{i,j}(f) \in \mathbb{R}$. Die Gruppe $\mathrm{AGL}_2(\mathbb{R})$ der geometrischen Transformationen operiert also auf der Sensorinputfunktion $f : \bigcup_{(i,j)\in S} A_{i,j} \times [t_0, t_0 + \tau] \to \mathbb{R}$ durch $f^T(x, t) = f(T(x), t)$.

[26] Bildquelle: http://de.wikipedia.org/wiki/Jimmy_Carter vom 14.06.2014.

Dadurch erhalten wir transformierte analoge Ausgangssignale

$$\ell_{i,j}(f^T) = \int_{A_{i,j}\times[t_0,t_0+\tau]} f(T(x),t)d\mu_{i,j}(x,t) = \int_{A_{i,j}} \left(\int_{[t_0,t_0+\tau]} f(T(x),t)d\nu(t) \right) d\alpha_{i,j}(x)$$
$$= \int_{A_{i,j}} \zeta_f(T(x))d\alpha_{i,j}(x),$$

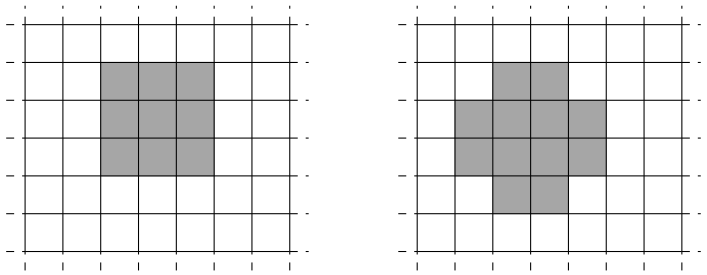
die zu transformierten quantisierten Grauwerten $q(\ell_{i,j}(f^T)) \in \mathbb{Z}_{0,255}$ führen. Somit induziert die Operation auf der Sensorinputfunktion eine Operation auf der zeitintegrierten Sensorinputfunktion $\zeta_f : \bigcup_{(i,j)\in S} A_{i,j} \to \mathbb{R}$ durch $\zeta_f^T := \zeta_f \circ T$ für alle $T \in \mathrm{AGL}_2(\mathbb{R})$. Der letzte Abschnitt des letzten Kapitels hat gezeigt, dass wir für diskret konvexe, n-zulässige Mengen $M \subseteq S$ auf dem zugehörigen Lokalisierungsfenster $\mathrm{Loc}_M \subseteq \mathbb{R}^2$ die unbekannte zeitintegrierte Sensorinputfunktion ζ_f durch eine Polynomfunktion $p^{(M)} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ approximieren können, d.h. wir können die zeitintegrierte Sensorinputfunktion auf geeigneten Bildbereichen lokal rekonstruieren. Die Koordinatenvektoren dieser Polynomfunktionen $p^{(M)}$ bzgl. geeigneter Orthogonal- oder Orthonormalbasen von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ sind die lokalen Bildmerkmale. Somit ist es ausreichend, die Wirkung der Operation der geometrischen Transformationen auf den endlich-dimensionalen, reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu untersuchen. Zunächst klar und wohl bekannt ist die Tatsache, dass für alle Transformationen $T \in \mathrm{AGL}_2(\mathbb{R})$ der reelle Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ ein T-invarianter Untervektorraum des Vektorraums $\mathscr{C}(\mathbb{R}^2, \mathbb{R})$ der stetigen Funktionen ist, d.h. es gilt $T(\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})) \subseteq \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ für alle $T \in \mathrm{AGL}_2(\mathbb{R})$ (vgl. u.a. [Fuc00]).

Im Folgenden werden wir nun zunächst ein Verfahren zur Ermittlung korrespondierender Bildmerkmale vorstellen, das sich auf Methoden der Invariantentheorie stützt. Dazu benötigen wir Invarianten von Polynomfunktionen unter bestimmten geometrischen Transformationen. Insbesondere wird dabei die Frage im Vordergrund stehen, wie es mit Hilfe fundamentaler Invarianten möglich ist, die Korrespondenz von Polynomfunktionen $p, q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ unter bestimmten Gruppen festzustellen. Konkret werden wir dabei behandeln, wann zwei Polynome p, q in einer Bahn unter der Operation der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$ liegen, d.h. wann es eine Rotation $R \in \mathrm{SO}_2(\mathbb{R})$ gibt mit $p = q \circ R$. Diese Ergebnisse lassen sich dann unmittelbar auf lokale Bildmerkmale übertragen. Wir können also beispielsweise beantworten, wann ein lokaler Bildausschnitt näherungsweise eine gedrehte Version eines anderen lokalen Bildausschnitts ist. Diese Fragen werden für bestimmte Lokalisierungsfenster im Falle der speziellen orthogonalen Gruppe z.B. in [Pis02] schon beantwortet, allerdings ist eine Antwort auf diese Frage mit Invarianten sehr viel eleganter und einfacher zu geben. Zudem ist die hier vorgestellte Methode nicht auf spezielle Lokalisierungsfenster oder spezielle Gruppen beschränkt, wie z.B. in [Pis02], sondern auf beliebige zulässige Lokalisierungsfenster und verschiedene Gruppen anwendbar. Im zweiten Abschnitt werden wir das Verfahren im dritten Abschnitt an Beispielen demonstrieren, wobei wir auf die Invarianten von Polynomfunktionen aus dem letzten Kapitel zurückgreifen. Im dritten Abschnitt geben wir noch einen kleinen Ausblick, insbesondere darauf, wie sich skalierungsinvariante Größen erzeugen lassen und wie es möglich ist, eine höhere Robustheit gegenüber photometrischen Einflüssen zu erzielen.

11.1 Das Verfahren zur Korrespondenzfindung lokaler Bildmerkmale

Wie in der Einleitung bereits ausgeführt wurde, ist das Ziel dieser Arbeit, ein einfaches Verfahren anzugeben, mit dessen Hilfe sich die Korrespondenz zweier lokaler Bildausschnitte feststellen lässt. Dazu ist zunächst zu klären, wann zwei lokale Pixelfenster als „korrespondierend"

bezeichnet werden können. Wir betrachten im Folgenden also zwei diskret-konvexe, n-zulässige und formgleiche Mengen $M_1, M_2 \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ der Kardinalität $m \geq n$ in unterschiedlichen Bildern und deren zugehörige Lokalisierungsfenster $\mathrm{Loc}_1, \mathrm{Loc}_2 \subseteq \mathbb{R}^2$ sowie die jeweiligen Grauwertvektoren $\mathrm{gv}(M_1), \mathrm{gv}(M_2) \in \mathbb{Z}_{0,255}^m \subseteq \mathbb{R}^m$. Typische Beispiele von geeigneten diskret-konvexen Mengen sind in Abbildung 11.1 dargestellt. Klar ist, dass die beiden Bildausschnitte genau dann exakt übereinstimmen, wenn die Grauwertvektoren identisch sind. Das wäre genau dann der Fall, wenn die beiden Bilder sich nur durch eine Translation um eine bestimmte Anzahl an Pixeln unterscheiden; eine Situation, die in der Praxis so nicht vorzufinden ist. Man muss also zwangsläufig übergehen zur Frage „Wann sind die Bildausschnitte näherungsweise gleich?". Diese Frage lässt sich beantworten, indem man für $\varepsilon > 0$ die Bildausschnitte als „identisch" ansieht, wenn $\|\mathrm{gv}(M_1) - \mathrm{gv}(M_2)\| < \varepsilon$ ist, wobei $\|\cdot\|$ die Euklidische Norm in $\mathbb{R}^m$ bezeichnet. Auch das ist natürlich noch viel zu „grob". Besser wäre es, wenn man die Bildausschnitte im Subpixelbereich anhand ihrer Sensorinputfunktion vergleichen könnte. Wie wir wissen, ist es leider nicht möglich, die Sensorinputfunktion allein aus den Grauwerten zu rekonstruieren, es ist aber sehr wohl möglich, die Sensorinputfunktion partiell zu rekonstruieren. Genauer können wir die zeitintegrierte Sensorinputfunktion durch Polynomfunktionen rekonstruieren, womit wir bei den lokalen Bildmerkmalen angelangt sind. Dies sind bekanntlich die Koordinatenvektoren der approximierenden Polynomfunktionen bzgl. einer geeigneten Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.



(a) 3×3-Fenster. (b) „Kreisförmige" Fenster.

Abbildung 11.1: Die gängigsten Beispiele für Lokalisierungsfenster.

Die Situation lässt sich also wie folgt modellieren. Sie $M_0 \subseteq \mathbb{Z} \times \mathbb{Z}$ eine diskret-konvexe und n-zulässige Menge der Kardinalität m, die dieselbe Form hat wie M_1 bzw. M_2 und im Nullpunkt zentriert ist. Weiter sei $\Psi : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^m$ eine geeignete lineare Abbildung, die ein Skalarprodukt $\langle \cdot, \cdot \rangle_\Psi$ auf M_0 induziert (siehe Abschnitt 9.2). Dann lässt sich bzgl. dieses Skalarprodukts eine Orthonormalbasis $B = (b_1, \ldots, b_\nu)$ von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ bestimmen. Nun ist es möglich mittels Orthogonalentwicklung die eindeutig bestimmten Approximationspolynome $p_1, p_2 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu berechnen, die $\mathrm{gv}(M_1)$ bzw. $\mathrm{gv}(M_2)$ bestmöglich approximieren, d.h. für die

$$\|\Psi(p_1) - \mathrm{gv}(M_1)\| = \min\{\|\Psi(q) - \mathrm{gv}(M_1)\| : \ q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})\}$$

bzw.

$$\|\Psi(p_2) - \mathrm{gv}(M_2)\| = \min\{\|\Psi(q) - \mathrm{gv}(M_2)\| : \ q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})\}$$

gilt (siehe Abschnitt 9.2.2). Sei $\kappa_B : \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \to \mathbb{R}^\nu$ die durch

$$\alpha_1 \cdot b_1 + \ldots + \alpha_\nu \cdot b_\nu \quad \mapsto \quad (\alpha_1, \ldots, \alpha_\nu)$$

definierte Koordinatenfunktion bzgl. der Orthonormalbasis B. Dann sind die Koordinatenvektoren $v_1 := \kappa_B(p_1) \in \mathbb{R}^\nu$ und $v_2 := \kappa_B(p_2) \in \mathbb{R}^\nu$ von p_1 und p_2 bzgl. B die zu Loc_1 bzw.

Loc_2 gehörenden lokalen Bildmerkmale. Damit lassen sich die beiden Bildausschnitte im Subpixelbereich zunächst anhand von p_1 und p_2 vergleichen. Für ein $\varepsilon > 0$ stimmen die beiden Bildausschnitte näherungsweise überein, falls $\|p_1 - p_2\|_\Psi < \varepsilon$ gilt. Dabei bezeichnet $\|\cdot\|_\Psi$ die durch $\langle\cdot,\cdot\rangle_\Psi$ induzierte Norm auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Dies lässt sich nahtlos auf die Koordinatenvektoren übertragen.

Bemerkung 11.1.1. Sei $B = (b_1, \ldots, b_\nu)$ eine Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und seien $v_1 = (\alpha_1, \ldots, \alpha_\nu)$ und $v_2 = (\beta_1, \ldots, \beta_\nu)$ Koordinatenvektoren zweier Polynomfunktionen $p_1, p_2 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Dann gilt:

$$\begin{aligned}
\|p_1 - p_2\|_\Psi^2 &= \|(\alpha_1 - \beta_1) \cdot b_1 + \ldots + (\alpha_\nu - \beta_\nu) \cdot b_\nu\|_\Psi^2 \\
&= \langle (\alpha_1 - \beta_1) \cdot b_1 + \ldots + (\alpha_\nu - \beta_\nu) \cdot b_\nu, (\alpha_1 - \beta_1) \cdot b_1 + \ldots + (\alpha_\nu - \beta_\nu) \cdot b_\nu \rangle_\Psi \\
&= \langle (\alpha_1 - \beta_1) \cdot b_1, (\alpha_1 - \beta_1) \cdot b_1 \rangle + \ldots + \langle (\alpha_\nu - \beta_\nu) \cdot b_\nu, (\alpha_\nu - \beta_\nu) \cdot b_\nu \rangle \\
&= (\alpha_1 - \beta_1)^2 + \ldots + (\alpha_\nu - \beta_\nu)^2 = \|v_1 - v_2\|^2
\end{aligned}$$

Somit lassen sich die beiden lokalen Bildausschnitte durch ihre lokalen Bildmerkmale miteinander vergleichen.

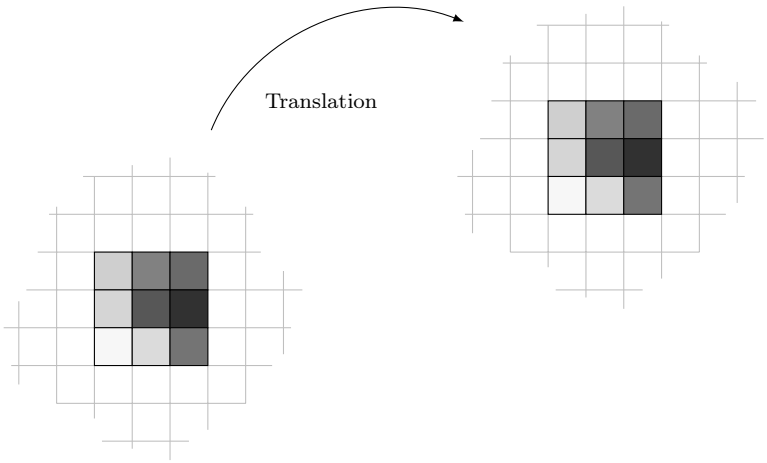


Abbildung 11.2: Darstellung der Wirkung einer Translation auf lokale Bildmerkmale.

Allerdings ist ein direkter Vergleich lokaler Bildmerkmale relativ uninteressant, da die näherungsweise Korrespondenz der beiden Bildausschnitte nur dann festgestellt werden kann, wenn auf die Bilder mehr oder weniger nur reine Translationen wirken. Hier kommen nun die geometrischen Transformationen ins Spiel. Wie zu Beginn des Kapitels dargelegt wurde, induziert die Operation der Gruppe $\mathrm{AGL}_2(\mathbb{R})$ auf den Bildern eine Operation auf den Approximationspolynomen, also auf den reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Die Operation dieser Gruppe auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ induziert wiederum eine Operation auf den Koordinatenring. Wir wollen also nicht die Approximationspolynome anhand ihrer Koordinatenvektoren bzgl. einer Orthonormalbasis auf Gleichheit testen, sondern wir wollen wissen, ob es eine Transformation $T \in \mathrm{AGL}_2(\mathbb{R})$ gibt so, dass p_1 und $p_2 \circ T$ näherungsweise übereinstimmen. Mit anderen Worten, wir sind daran interessiert zu entscheiden, ob sich zwei Polynomfunktionen ungefähr in derselben Bahn befinden. Wann dies nicht ungefähr, sondern exakt der Fall ist, liegt zunächst auf der Hand.

Bemerkung 11.1.2. Sei G eine Untergruppe von $\mathrm{AGL}_2(\mathbb{R})$ und seien $p_1, p_2 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Genau dann gilt $p_2 \in G(p_1)$, wenn es eine Transformation $T \in G$ gibt mit $p_1 = p_2 \circ T$.

Das führt uns zur Definition einer approximativen Bahn oder, wie wir sie nennen wollen, eine δ-Bahn, abhängig von einer Schranke $\delta > 0$.

Definition 11.1.3. (δ-Bahn)
Sei G eine Untergruppe von $\mathrm{AGL}_2(\mathbb{R})$ und sei $\delta > 0$. Dann bezeichnen wir die Teilmenge

$$G_\delta(p) := \{q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) : \text{ Es gibt ein } T \in G \text{ mit } \|p - q \circ T\|_\Psi < \delta\}$$

von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ als die δ**-Bahn** von p.

Damit können wir die Frage „Gibt es ein $T \in G \subseteq \mathrm{AGL}_2(\mathbb{R})$ mit $p_1 \approx p_2 \circ T$?" exakter ausdrücken durch die Frage „Gilt $p_2 \in G_\delta(p_1)$ für ein $\delta > 0$?". Wegen Bemerkung 11.1.1 lässt sich die Zugehörigkeit einer Polynomfunktion $q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zur δ-Bahn einer Polynomfunktion p auch mit Hilfe der Koordinatenvektoren feststellen, d.h. äquivalent gilt:

$$G_\delta(p) = \{q \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) : \text{ Es gib ein } T \in G \text{ mit } \|\kappa_B(p) - \kappa_B(q \circ T)\| < \delta\}.$$

Einen einfachen Spezialfall einer Untergruppe G haben wir bereits angesprochen: Operieren nur reine Translationen auf den Bildern, können wir bereits jetzt entscheiden, wann diese ungefähr in einer Bahn sind. Da lokale Bildmerkmale translationsinvariant sind, lässt sich für $\delta > 0$ mittels $\|v_1 - v_2\| < \delta$ feststellen, ob sich zwei Bildausschnitte näherungsweise nur durch eine Translation voneinander unterscheiden.

Lemma 11.1.4. (Klassifikation von approximativen Bahnen bei Translationen)
Sei $G := \mathrm{Trans}_2(\mathbb{R})$*, seien* $p_1, p_2 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ *und sei* $\delta > 0$*. Genau dann gilt* $p_2 \in G_\delta(p_1)$*, wenn* $\|\kappa_B(p_1) - \kappa_B(p_2)\| < \delta$ *ist.*

Beweis: Ist $p_2 \in G_\delta(p_1)$, so gibt es ein $T \in G$ mit $\|p_1 - p_2 \circ T\|_\Psi < \delta$. Wegen $\kappa_B(p_2 \circ T) = \kappa_B(p_2)$ folgt aus Bemerkung 11.1.1 die Behauptung. Gilt umgekehrt $\|\kappa_B(p_1) - \kappa_B(p_2)\| < \delta$, so gilt $\|p_1 - p_2 \circ T\| < \delta$ für alle $T \in G$, also $p_2 \in G_\delta(p_1)$. □

Diese Situation und die entsprechende Klassifikation ist in vielen Anwendungen auch zu finden. Sie bildet die typische Ausgangslage des sogenannten optischen Flusses. Es gibt in der Theorie des optischen Flusses sogar Methoden, die mehr oder weniger direkt die einzelnen Grauwerte der Bildausschnitte miteinander vergleichen. Anstatt die Grauwerte zu vergleichen, ist es natürlich wesentlich genauer, lokale Bildmerkmale miteinander zu vergleichen, da dadurch eine Zuordnung im Subpixelbereich möglich ist. In [Haa00] wird beispielsweise neben anderen Ansätzen auch dieser Ansatz zur Verfolgung von Fahrbahnmarkierungen in Fahrerassistenzsystemen untersucht. Offensichtlich reicht dieses Vergleichsmaß nicht mehr aus, sobald man Anwendungen betrachtet, bei denen insbesondere Rotationen im Spiel sind, wie beispielsweise bei der Erkennung eines Objekts an einem Fließband.

Da lokale Bildmerkmale translationsinvariant sind, genügt es, sich im Weiteren mit abgeschlossenen Untergruppen G von $\mathrm{GL}_2(\mathbb{R})$, also mit linearen algebraischen Gruppen, zu beschäftigen. Wegen der Translationsinvarianz ist dann auch der Fall $\mathrm{Trans}_2(\mathbb{R}) \rtimes G$ abgedeckt, d.h. sind p_1 und p_2 näherungsweise in einer G-Bahn, so sind sie auch näherungsweise in einer Bahn bzgl. der Gruppe $\mathrm{Trans}_2(\mathbb{R}) \rtimes G$. Ganz besonders interessant sind dabei die linear reduktiven Gruppen (siehe Abschnitt 4.4). Die Elemente von $\mathrm{GL}_2(\mathbb{R})$ wollen wir aber nach wie vor in der Regel als Transformationen $\mathbb{R}^2 \to \mathbb{R}^2$ betrachten. Die für Anwendungen bedeutendste linear reduktive Gruppe ist sicherlich die spezielle orthogonale Gruppe $\mathrm{SO}_2(\mathbb{R})$. Nun benötigen wir ein von der betrachteten Gruppe G abhängiges, invariantes Distanzmaß d im Koordinatenraum $\mathbb{R}^\nu$, das die

Zugehörigkeit zu einer Bahn anhand der Koordinatenvektoren „messen" kann. Dabei soll offensichtlich $d(v,w) = 0$ genau dann gelten, wenn sich die Polynomfunktionen $\kappa_B^{-1}(v)$ und $\kappa_B^{-1}(w)$ in einer Bahn befinden. Wir suchen also nach einer geeigneten **Pseudometrik** bzw. **Halbmetrik**. Bei dieser Suche kommt uns die Invariantentheorie zu Hilfe. Bevor wir uns damit beschäftigen, wollen wir uns an dieser Stelle im folgenden Beispiel ansehen, welches Distanzmaß es bereits gibt. Pisinger hat in [Pis02] mit vergleichsweise großem Aufwand für einen Beispielfall ein Distanzmaß hergeleitet, das sich in der Praxis sehr gut bewährt hat (siehe [Pis02], Abschnitt 4.2, S. 77–94).

Beispiel 11.1.5. Pisinger betrachtet die spezielle orthogonale Gruppe $G = \mathrm{SO}_2(\mathbb{R})$ und den „einfachsten" Fall von 3×3-Lokalisierungsfenstern. Er benötigt zur Charakterisierung ganz spezielle Orthonormalbasen von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$: Orthonormalbasen, die „stabil" sind unter Koordinatenvertauschung. Beispiel 9.2.29 enthält eine derartige Basis, in der sich die Symmetrie der 3×3-Lokalisierungsfenster widerspiegelt. Sei B eine entsprechende Orthonormalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ und seien $v_1 = (\alpha_1, \ldots, \alpha_6) \in \mathbb{R}^6$ sowie $v_2 = (\beta_1, \ldots, \beta_6) \in \mathbb{R}^6$ die Koordinatenvektoren zweier Polynomfunktionen $p_1, p_2 \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ bzgl. der Basis B. Sei $\mathcal{R}_\varphi \in \mathrm{SO}_2(\mathbb{R})$ eine Drehmatrix um einen Winkel $\varphi \in \mathbb{R}$ und $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ die durch $\mathcal{R}_\varphi$ induzierte lineare Abbildung. Pisinger zeigt (siehe [Pis02], Satz 4.2.11, S. 87 ff.), dass genau dann $p_1 = p_2 \circ R_\varphi$ gilt, also genau dann $p_2 \in G(p_1)$ ist, wenn folgende Bedingungen erfüllt sind:

(i) $\alpha_1 = \beta_1$,

(ii) $\binom{\alpha_2}{\alpha_3} = \mathcal{R}_\varphi \cdot \binom{\beta_2}{\beta_3}$,

(iii) $\mathcal{H}_{p_1}(0) = \mathcal{R}_\varphi^{\mathrm{tr}} \cdot \mathcal{H}_{p_2}(0) \cdot \mathcal{R}_\varphi$, wobei $\mathcal{H}_{p_1}(0)$ bzw. $\mathcal{H}_{p_2}(0)$ die Hesse-Matrix von p_1 bzw. p_2 im Nullpunkt bezeichnet.

Die letzten beiden Bedingungen lassen sich äquivalent etwas einfacher auf folgende Weise überprüfen (siehe [Pis02], Bemerkung 4.2.12, S. 90 f.):

(1) Teste, ob $\alpha_2^2 + \alpha_3^2 = \beta_2^2 + \beta_3^2$ gilt. In diesem Fall gibt es einen Winkel $\varphi_1 \in [0, 2\pi[$ mit $\binom{\alpha_2}{\alpha_3} = \mathcal{R}_{\varphi_1} \cdot \binom{\beta_2}{\beta_3}$.

(2) Teste, ob $\mathrm{Spur}(\mathcal{H}_{p_1}(0)) = \mathrm{Spur}(\mathcal{H}_{p_2}(0))$ und $\det(\mathcal{H}_{p_1}(0)) = \det(\mathcal{H}_{p_2}(0))$ gilt. In diesem Fall gibt es ein $\varphi_2 \in [0, 2\pi[$ mit $\mathcal{H}_{p_1}(0) = \mathcal{R}_{\varphi_2}^{\mathrm{tr}} \cdot \mathcal{H}_{p_2}(0) \cdot \mathcal{R}_{\varphi_2}$.

(3) Teste, ob $\varphi_1 = \varphi_2$ gilt.

Aufbauend auf dieser anschaulichen Charakterisierung konstruiert Pisinger das folgende translations- und rotationsinvariante Distanzmaß $d : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}$ (siehe [Pis02], 4.2.15).

$$\begin{aligned} d(v_1, v_2) = \nu_1 \cdot \left| \sqrt{\alpha_2^2 + \alpha_3^2} - \sqrt{\beta_2^2 + \beta_3^2} \right| + \nu_2 \cdot \left| \frac{\mathrm{Spur}(\mathcal{H}_{p_1}(0))^2}{\det(\mathcal{H}_{p_1}(0))} - \frac{\mathrm{Spur}(\mathcal{H}_{p_2}(0))^2}{\det(\mathcal{H}_{p_2}(0))} \right| \\ + \nu_3 \cdot \min\{\mathrm{Spur}(\mathcal{H}_{p_1}(0))^2 - 4\det(\mathcal{H}_{p_1}(0)), \mathrm{Spur}(\mathcal{H}_{p_2}(0))^2 - 4\det(\mathcal{H}_{p_2}(0))\} \\ \cdot \left| \psi - \sphericalangle\left(\begin{pmatrix} \alpha_2 \\ \alpha_3 \end{pmatrix}, \begin{pmatrix} \beta_2 \\ \beta_3 \end{pmatrix} \right) \right| \end{aligned}$$

Dabei sind $\nu_1, \nu_2, \nu_3 \in \mathbb{R}$ experimentell ermittelte Gewichte. Pisinger konnte mit $\nu_1 = 0.3$, $\nu_2 = 0.25$ und $\nu_3 = 0.2$ gute Ergebnisse erzielen. Die einzelnen Teile dieses Distanzmaßes lassen sich alle anschaulich interpretieren. So ist der erste Summand im Wesentlichen die Differenz der Gradienten $\operatorname{grad} p_1(0)$ und $\operatorname{grad} p_2(0)$ am Nullpunkt, der zweite Summand beschreibt die Differenz der Eigenwertverhältnisse der beiden Hesse-Matrizen. Die Differenz im dritten Summanden ist die Winkeldifferenz zwischen dem Drehwinkel ψ der Eigenvektorsysteme der beiden Hesse-Matrizen und dem Drehwinkel der Gradienten. Diese Differenz wird mit dem Minimum

der quadrierten Differenzen der Eigenwerte multipliziert, da die Winkeldifferenz nur von Wert ist, wenn die Eigenwerte verschieden sind. PISINGER sieht die lokalen Bildmerkmale v_1, v_2 nun aufbauend auf diesem Distanzmaß als korrespondierend an, falls für ein $\varepsilon > 0$ der Abstand $d(v_1, v_2) < \varepsilon$ ist. Wie man sehen kann, lässt PISINGER α_1 und β_1 außen vor. Das begründet er damit, dass α_1 bzw. β_1 nur Vielfache des arithmetischen Mittels der Grauwerte aus dem betrachteten Pixelfenster sind (siehe Lemma 9.2.26 und Korollar 9.2.27) und somit stark durch photometrische Einflüsse beeinflusst werden. ◁

Wie an diesem Beispiel schnell zu sehen ist, steckt in der „manuellen" Konstruktion aufbauend auf der Anschauung sehr viel Arbeit. So ist es für größere Lokalisierungsfenster oder auch für andere Gruppen nahezu unmöglich, auf ähnlichem Weg ein Distanzmaß anzugeben. Außerdem ist diese Vorgehensweise auf bestimmte Orthonormalbasen beschränkt. Dank der Invariantentheorie bereiten diese Schwierigkeiten und Einschränkungen keine großen Probleme mehr. Wie wir zum Teil bereits in Bemerkung 10.3.4 gesehen haben, hängen die einzelnen Bestandteile des Distanzmaßes aus dem letzten Beispiel eng mit den fundamentalen Invarianten für $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ und $\mathrm{SO}_2(\mathbb{R})$ zusammen.

Beispiel 11.1.6. Sei $C = (c_{i,j} : \ i + j \leq 2)$ die Orthonormalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ aus Beispiel 10.4.3 und sei $p \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ mit Koordinatenvektor $(\alpha_{0,0}, \ldots, \alpha_{2,0}) \in \mathbb{R}^6$, also definiert durch

$$\begin{aligned} p(x,y) = \alpha_{0,0} \cdot \tfrac{1}{3} + \alpha_{0,1} \cdot \tfrac{1}{\sqrt{6}} y + \alpha_{1,0} \cdot \tfrac{1}{\sqrt{6}} x \\ + \alpha_{0,2} \cdot \tfrac{1}{\sqrt{2}} (y^2 - \tfrac{3}{4}) + \alpha_{1,1} \cdot \tfrac{1}{2} xy + \alpha_{2,0} \cdot \tfrac{1}{\sqrt{2}} \cdot (x^2 - \tfrac{3}{4}). \end{aligned}$$

Dann folgt:

$$\operatorname{grad} p(0,0) = \frac{1}{\sqrt{6}} (\alpha_{1,0}, \alpha_{0,1}), \qquad \mathcal{H}_p(0,0) = \begin{pmatrix} \sqrt{2}\alpha_{2,0} & \frac{1}{2}\alpha_{1,1} \\ \frac{1}{2}\alpha_{1,1} & \sqrt{2}\alpha_{0,2} \end{pmatrix}$$

also weiter

$$\operatorname{Spur}(\mathcal{H}_p(0,0)) = 2\sqrt{2} \cdot (\alpha_{0,2} + \alpha_{2,0}), \qquad \det(\mathcal{H}_p(0,0)) = 2 \cdot (\alpha_{0,2}\alpha_{2,0} - \tfrac{1}{8}\alpha_{1,1}).$$

Laut Beispiel 10.4.3 sind bzgl. der Orthonormalbasis C folgende Polynome fundamentale Invarianten:

$$\begin{aligned} f_1 &= a_{0,0}, & f_4 &= a_{0,2}a_{2,0} - \tfrac{1}{8}a_{1,1}^2, \\ f_2 &= a_{0,1}^2 + a_{1,0}^2, & f_5 &= a_{0,1}^2 a_{2,0} - \tfrac{1}{2}\sqrt{2} a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}, \\ f_3 &= a_{0,2} + a_{2,0}, & f_6 &= a_{0,1}^2 a_{1,1} - 2\sqrt{2} a_{0,1}a_{1,0}a_{0,2} + 2\sqrt{2} a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1} \end{aligned}$$

Man sieht bis auf die Winkel, die im Distanzmaß aus dem letzten Beispiel vorkommen, sofort, dass sie sich nur aus den fundamentalen Invarianten zusammensetzen. Auch die beiden Winkel lassen sich über diese fundamentalen Invarianten bestimmen, was aber nur mit großem Aufwand nachgerechnet werden kann, worauf wir an dieser Stelle verzichten wollen. Betrachtet man allein nur die beiden Beträge

$$\left| \sqrt{f_2(\alpha_{0,0}, \ldots, \alpha_{2,0})} - \sqrt{f_2(\beta_{0,0}, \ldots, \beta_{2,0})} \right| \text{ und } \left| \frac{f_3(\alpha_{0,0}, \ldots, \alpha_{2,0})^2}{f_4(\alpha_{0,0}, \ldots, \alpha_{2,0})} - \frac{f_3(\beta_{0,0}, \ldots, \beta_{2,0})^2}{f_4(\beta_{0,0}, \ldots, \beta_{2,0})} \right|,$$

die mit den ersten beiden Teilen des Distanzmaßes korrespondieren, so lässt sich das weitere Vorgehen bereits erkennen. ◁

Man kann leicht erkennen, dass hinter einer anschaulichen Herangehensweise bereits für $n = 2$ ein großer Aufwand steckt. Wie sieht die Situation nun für größere Pixelfenster, also größere Parameter n aus?

Bemerkung 11.1.7. In [Pis02] verwendet PISINGER für die Klassifikation von Polynomfunktionen aus $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ neben den in Beispiel 11.1.5 genannten Kriterien nur das zuätzliche „rotationsinvariante Maß" (siehe [Pis02], S. 101):

$$a_{0,3}^2 + \tfrac{1}{3}(a_{1,2}^2 + a_{2,1}^2) + a_{3,0}^2$$

In Abschnitt B.1 sind unter anderem die beiden fundamentalen Invarianten

$$\begin{aligned} f_5 &= a_{0,3}a_{2,1} - \tfrac{1}{3}a_{1,2}^2 + a_{1,2}a_{3,0} - \tfrac{1}{3}a_{2,1}^2, \\ f_6 &= a_{0,3}^2 + a_{0,3}a_{2,1} + a_{1,2}a_{3,0} + a_{3,0}^2 \end{aligned}$$

für $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ zu finden. Es gilt

$$f_6 - f_5 = a_{0,3}^2 + \tfrac{1}{3}(a_{1,2}^2 + a_{2,1}^2) + a_{3,0}^2,$$

d.h. es handelt sich tatsächlich um eine Invariante. Allerdings ist offensichtlich diese eine zusätzliche Invariante zur Klassifikation der Bahnen nicht ausreichend.

Nachdem wir nun die Situation aus der bekannten Sicht präsentiert haben, wollen sie nun aus Sicht der Invariantentheorie beleuchten, allerdings allgemeiner. Dazu betrachten wir nun eine linear reduktive Gruppe G, die sich bekanntlich dadurch auszeichnet, dass der Invariantenring $\mathbb{R}[x_1, \ldots, x_\nu]^G$ stets endlich erzeugt ist (siehe Theorem 6.2.5), d.h. es gibt endliche viele Polynome $f_1, \ldots, f_r \in \mathbb{R}[x_1, \ldots, x_\nu]$ mit $\mathbb{R}[x_1, \ldots, x_\nu]^G = \mathbb{R}[f_1, \ldots, f_r]$. Mit $V := \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R}) \cong \mathbb{R}^\nu$ existiert damit stets der algebraische Quotient $\pi : V \to V /\!\!/ G$ (siehe Abschnitt 8.1), definiert durch

$$p \mapsto (f_1(\kappa_B(p)), \ldots, f_r(\kappa_B(p))).$$

Dieser bringt genau die entscheide Eigenschaft mit, denn es gilt genau dann $\pi(p_1) = \pi(p_2)$, wenn $\overline{G(p_1)} \cap \overline{G(p_2)} \neq \emptyset$ gilt. Mit anderen Worten, die Polynomfunktion p_2 ist genau dann im Abschluss der Bahn $G(p_1)$ enthalten, wenn $\pi(p_1) = \pi(p_2)$ gilt. Ist der algebraische Quotient geometrisch, trennt er nicht nur die abgeschlossenen Bahnen, sondern die Bahnen selbst, d.h. in diesem Fall gilt $p_2 \in G(p_1)$ genau dann, wenn $\pi(p_1) = \pi(p_2)$ gilt. Dann ist $V /\!\!/ G$ identisch mit dem Bahnenraum V/G.

Bemerkung 11.1.8. Für einen geometrischen Quotienten $\pi : V \to V/G$ lässt sich die Zugehörigkeit zu der Bahn einer Polynomfunktion $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ wie folgt charakterisieren:

$$q \in G(p) \quad \Longleftrightarrow \quad \pi(p) = \pi(q).$$

Die Frage, ob der algebraische Quotient nun geometrisch ist oder nicht, ist im Allgemeinen nicht einfach zu beantworten. Nun arbeiten wir aber über den reellen Zahlen und mit Bildern, womit wir eine exakte Übereinstimmung von $\pi(p_1)$ und $\pi(p_2)$ in der Praxis ohnehin nicht zu erwarten haben. Somit ist es über den reellen Zahlen recht „bedeutungslos", ob der algebraische Quotient geometrisch ist oder nicht und ob wir die Bahn oder den Abschluss der Bahn betrachten. Denn der algebraische Quotient „trennt" die Bahnen in einem für unsere Zwecke ausreichenden Maße. Um nun eine approximative Variante von Bemerkung 11.1.8 zu erhalten, ist allerdings eine Metrik auf $V /\!\!/ G \subseteq \mathbb{A}_{\mathbb{R}}^r$ notwendig. Die Wahl einer derartigen geeigneten Metrik ist leider nicht im Allgemeinen zu beantworten. Sie hängt maßgeblich von der Anwendungssituation ab und muss experimentell in der jeweiligen Situation ermittelt werden, wie auch in Beispiel 11.1.5 zu sehen ist.

Bemerkung 11.1.9. (Wahl der Metrik $\tilde{d}$)
Es lässt sich leider nicht endgültig klären, was eine gute Wahl für $\tilde{d}$ ist, geschweige denn, was die „richtige" Wahl wäre. Es ist durchaus denkbar, dass in manchen praktischen Anwendungen auch sehr phantasievolle Metriken gut funktionieren können. Wenn man z.B. das Distanzmaß in [Pis02] betrachtet, ist diese Halbmetrik doch weit von „Standardmetriken" entfernt, dennoch hat sich dieses Maß in der Praxis bewährt. Problematisch an den naheliegenden Metriken, wie z.B. die Euklidische Metrik, die Maximum-Metrik oder die Manhattan-Metrik, ist sicherlich die Tatsache, dass manche Teile im Vektor $(f_1(v), \ldots, f_r(v))$ durch den Grad der Polynome $f_1, \ldots, f_r$ stärker gewichtet werden als andere, vielleicht manchmal zu stark. Um dies etwas auszugleichen, erscheint eine gewichtete Manhattan-Metrik, wie sie ja auch zum Teil in Beispiel 11.1.5 vorkommt, keine schlechte Wahl zu sein. Aber abschließend klären lässt sich diese Frage in dieser Arbeit nicht.

Haben wir allerdings eine geeignete Metrik auf $\mathbb{A}^r_{\mathbb{R}}$ gewählt, so erhalten wir sofort eine Halbmetrik auf dem Koordinatenraum $\mathbb{R}^\nu$.

Lemma 11.1.10. *Sei $\pi : V \to V/\!\!/G \subseteq \mathbb{A}^r_{\mathbb{R}}$ der algebraische Quotient von G in $V = \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ und sei $\tilde{d} : \mathbb{A}^r_{\mathbb{R}} \times \mathbb{A}^r_{\mathbb{R}} \to \mathbb{R}$ eine Metrik. Dann ist $d_G : \mathbb{R}^\nu \times \mathbb{R}^\nu \to \mathbb{R}$, definiert durch*

$$d_G(v, w) = \tilde{d}(\pi(\kappa_B^{-1}(v)), \pi(\kappa_B^{-1}(w)))$$

*eine Halbmetrik auf $\mathbb{R}^\nu$. Wir bezeichnen sie als **die durch G auf V induzierte Halbmetrik**.*

Beweis: Seien $u, v, w \in \mathbb{R}^\nu$ und seien $p = \kappa_B^{-1}(u)$, $q = \kappa_B^{-1}(v)$ und $r = \kappa_B^{-1}(w)$ die zugehörigen Polynomfunktionen. Dann gilt $d_G(v, v) = \tilde{d}(\pi(q), \pi(q)) = 0$. Weiter gilt

$$d_G(v, w) = \tilde{d}(\pi(q), \pi(r)) = \tilde{d}(\pi(r), \pi(q)) = d_G(w, v)$$

und schließlich

$$d_G(u, v) + d_G(v, w) = \tilde{d}(\pi(p), \pi(q)) + \tilde{d}(\pi(q), \pi(r)) \geq \tilde{d}(\pi(p), \pi(r)) = d_G(u, w).$$

□

Der Unterschied zu einer Metrik liegt also darin, dass zwei verschiedene Koordinatenvektoren $v \neq w$ Abstand $d_G(v, w) = 0$ haben können. Ist der algebraische Quotient π geometrisch, so gilt $d_G(v, w) = 0$ genau dann, wenn die Polynomfunktionen $\kappa_B^{-1}(v)$ und $\kappa_B^{-1}(w)$ in einer Bahn sind. Ist der Abstand $d_G(v, w)$ klein, so ist auch der Abstand der zugehörigen Polynomfunktionen klein, wie aus dem nächsten Lemma hervorgeht.

Lemma 11.1.11. *Seien $p_1, p_2 \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit Koordinatenvektoren $v_1 = \kappa_B(p_1) \in \mathbb{R}^\nu$ und $v_2 = \kappa_B(p_2) \in \mathbb{R}^\nu$. Weiter sei $d : \mathbb{R}^\nu \times \mathbb{R}^\nu \to \mathbb{R}$ eine Halbmetrik auf $\mathbb{R}^\nu$. Gilt $d_G(v_1, v_2) < \varepsilon$ für ein $\varepsilon > 0$, so gibt es ein $\delta > 0$ mit $p_2 \in G_\delta(p_1)$.*

Beweis: Zunächst gilt $p_2 \circ T \in G(p_2)$ und damit $\pi(p_2) = \pi(p_2 \circ T)$ für alle $T \in G$. Sei $T \in G$ und sei $U_\varepsilon(\pi(p_1))$ eine ε-Umgebung von $\pi(p_1)$. Dann folgt $\pi(p_2 \circ T) \in U_\varepsilon(\pi(p_1))$. Da π stetig ist, ist $U_\delta(p_1) := \pi^{-1}(U_\varepsilon(\pi(p_1)))$ eine Umgebung von p_1 für ein $\delta > 0$. Wegen $p_2 \circ T \in U_\delta(p_1)$ gilt $\|p_1 - p_2 \circ T\|_\Psi < \delta$ und damit $p_2 \in G_\delta(p_1)$. □

Damit ist es gerechtfertigt in folgendem Sinne von „korrespondierenden Bildmerkmalen" zu reden.

Definition 11.1.12. (Korrespondierende lokale Bildmerkmale)
Sei G eine linear reduktive Untergruppe von $\mathrm{GL}_2(\mathbb{R})$, die linear auf $V := \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ operiert, sei $d_G : \mathbb{R}^\nu \times \mathbb{R}^\nu \to \mathbb{R}$ die durch G induzierte Halbmetrik auf V und sei $\varepsilon > 0$. Zwei lokale Bildmerkmale $v, w \in \mathbb{R}^\nu$ werden als **ε-korrespondierend** bezeichnet, wenn $d_G(v, w) < \varepsilon$ gilt.

Gilt für $\varepsilon > 0$ also $d_G(v_1, v_2) < \varepsilon$, so betrachten wir die Polynomfunktionen $p_1 = \kappa_B^{-1}(v_1)$ und $p_2 = \kappa_B^{-1}(v_2)$ als näherungsweise in einer Bahn. Damit ist es nun zusammenfassend möglich, einen einfachen Algorithmus zur Ermittlung korrespondierender lokaler Bildmerkmale zu formulieren. Um die Korrespondenzpaare in den Bildern verorten zu können, betrachten wir die um Pixelkoordinaten erweiterten Bildmerkmale in $\mathbb{R}^\nu \times \mathbb{R}^2$.

Algorithmus 11.1 : Algorithmus zur Bestimmung von Korrespondenzpaaren lokaler Bildmerkmale

Input : Ein Erzeugendensystem $\{f_1, \ldots, f_r\} \subseteq \mathbb{R}[x_1, \ldots, x_\nu]$ des Invariantenrings $\mathbb{R}[x_1, \ldots, x_\nu]^G$ einer linear reduktiven Untergruppe G von $\mathrm{GL}_2(\mathbb{R})$.
Input : Suchbild gv_1 und Musterbild gv_2.
Input : Eine Metrik $\tilde{d}$ auf $\mathbb{A}^r_{\mathbb{R}}$ zur Konstruktion von $d_G : \mathbb{R}^\nu \times \mathbb{R}^\nu \to \mathbb{R}$.
Input : $\varepsilon > 0$.
Result : Eine Menge K mit Paaren erweiterter korrespondierender Bildmerkmale.

```
1 K := ∅;
2 Bestimme eine Menge G_1 ⊆ ℝ^ν × ℝ^2 markanter erweiterter Bildmerkmale in gv_1;
3 Bestimme eine Menge G_2 ⊆ ℝ^ν × ℝ^2 aller erweiterten Bildmerkmale in gv_2;
4 foreach (v, x) ∈ G_1 do
5     M := {(w, y) ∈ G_2 : d_G(v, w) < ε};
6     foreach (w, y) ∈ M do
7         K := K ∪ {((v, x), (w, y))}
8 return K;
```

Die Klassifikation „markanter" Bildmerkmale kann mit Hilfe eines Ähnlichkeitsmaßes durchgeführt werden, wie es beispielsweise auch in [Pis02] der Fall ist (siehe auch Kapitel 1). Wie man schnell erkennen kann, ist es durchaus möglich, dass es zu einem Merkmal im Musterbild mehrere (vermeintlich) korrespondierende Merkmale im Suchbild gibt, unter denen sich nicht zwangsläufig ein korrektes finden muss. Es ist also im Allgemeinen mit Fehlern zu rechnen. Eine Beurteilung, ob eine Zuordnung richtig ist oder nicht, kann nur nach Augenschein erfolgen. Die Idee, lokale Bildmerkmale zur Korrespondenzfindung in Stereobildern einzusetzen, ist nicht neu (vgl. z.B. [Pis02]), jedoch bietet dieses Idee in Verbindung mit Methoden der Invariantentheorie ganz neue Möglichkeiten. So betrachtet Pisinger in [Pis02] lediglich einen Spezialfall (siehe Beispiel 11.1.5). Die in [Pis02] bewiesene Charakterisierung von Bahnen – auch wenn das dort nicht so genannt wird – unter der Operation der speziellen orthogonalen Gruppe bekommen wir beinahe eins zu eins sozusagen „geschenkt", wie wir später noch sehen werden. Eine anschauliche Vorgehensweise ist für diese Situation auch noch gut machbar, allerdings nahezu unmöglich, wenn man größere Pixelfenster und dementsprechend höherdimensionale Polynomfunktionen betrachten will. Diese Einschränkung hat das hier vorgestellte Verfahren nicht.

Bemerkung 11.1.13. (Vorteile des Verfahrens)
Das hier vorgestellte Verfahren weist im Vergleich zu herkömmlichen Methoden folgende Vorteile auf:

- **Erweiterbarkeit auf andere Gruppen:** Auch wenn die spezielle orthogonale Gruppe $\mathrm{SO}_2(\mathbb{R})$ und die orthogonale Gruppe $\mathrm{O}_2(\mathbb{R})$ sicherlich die höchste Relevanz für die Praxis

haben, ist unser Verfahren nicht auf diese Gruppen limitiert. Es lässt sich ohne zusätzlichen Aufwand auf weitere linear reduktive Untergruppen G von $\mathrm{GL}_2(\mathbb{R})$ erweitern, falls eine Anwendung dies erfordert. Die herkömmlichen Methoden können das nicht „automatisch", man müsste ggf. erst neue Vergleichskriterien entwickeln, was mit hohem Aufwand verbunden sein kann. Somit lassen sich also auf einfache Weise Invarianten lokaler Bildmerkmale, und damit Entscheidungsregeln für die Korrespondenz, für das semidirekte Produkt $\mathrm{Trans}_2(\mathbb{R}) \rtimes G$ bestimmen.

- **Erweiterbarkeit auf beliebige zulässige Pixelfenster:** Für beliebige n-zulässige, diskret-konvexe Mengen M ist es notwendig, sich zuerst ein Vergleichskriterium für die näherungsweise Übereinstimmung von p und $q \circ T$ zu erarbeiten. Das ist für kleine Mengen M noch halbwegs gut machbar und wurde auch im Falle von 3×3-Pixelfenstern bereits gemacht (siehe [Pis02]), allerdings gestaltet sich das zunehmend schwieriger, wenn die Mengen größer werden. Unser Verfahren hat mit der Größe der verwendeten Pixelfenster keine Probleme. Ebenso wäre es möglich, nicht-symmetrische Lokalisierungsfenster zu verwenden. Es sind lediglich die fundamentalen Invarianten des Polynomvektorraums $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ zu berechnen. Das ist ein immenser Vorteil, da eine anschauliche Herangehensweise bei der Entwicklung eines Vergleichskriteriums, wie wir schon gesehen haben, eigentlich nur für 3×3-Pixelfenster funktionieren wird. Wir sind nun grundsätzlich in der Lage, beliebige n-zulässige Pixelfenster zu bedienen, wobei natürlich aus anderen Gründen Grenzen bestehen.
- **Keine zusätzlichen Anforderungen an die Orthonormalbasis:** In den bisherigen Verfahren werden nur symmetrische Pixelfenster eingesetzt. Die daraus resultierenden Orthonormalbasen müssen für das Verfahren in [Pis02] diese Symmetrie widerspiegeln. Diese Einschränkung besteht mit unserem Verfahren nicht, da einfach nur die fundamentalen Invarianten bzgl. der gewählten Basis zu bestimmen sind.
- **Effiziente Umsetzung:** In den herkömmlichen Verfahren müssen für ein Vergleichskriterium in der Regel zunächst Sekundargrößen aus den lokalen Bildmerkmalen, wie z.B. bestimmte Drehwinkel, berechnet werden (siehe Beispiel 11.1.5). Die Berechnung solcher Sekundargrößen ist in unserem Verfahren nicht notwendig. Aufbauend auf den lokalen Bildmerkmalen sind lediglich Polynome auszuwerten. Somit lässt sich dieses sehr einfache Verfahren äußerst effizient umsetzen, auch Hardware nah und damit sehr schnell. Zudem ist es gut möglich, dass in der Praxis bereits mit weniger fundamentalen Invarianten gute Ergebnisse erzielt werden können, sodass vielleicht nicht das ganze Erzeugendensystem verwendet werden muss. Das würde die Effizienz natürlich erneut steigern.

Nach der Vorstellung des Verfahrens zur Korrespondenzfindung lokaler Bildmerkmale wollen wir es im weiteren Verlauf auch an Beispielen demonstrieren. Dabei werden wir die beiden gängigsten Pixelfenster verwenden, die in Abbildung 11.1 abgebildet sind. Es handelt sich bei den betrachteten Beispielen aber lediglich um eine Art „proof of concept". Da wir in dieser Arbeit keine konkrete Anwendung mit Praxisbezug verfolgen, können also Fragen nach der Praxistauglichkeit des Verfahrens oder nach eventuellen Problemen nicht beantwortet werden, was aber auch nicht Ziel dieser Arbeit sein soll! Nichtsdestotrotz benötigen wir als Eingabe für den Algorithmus fundamentale Invarianten für Polynomfunktionen für die betrachteten Pixelfenster: Im Falle von 3×3-Pixelfenster ist der Vektorraum $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ zu betrachten, im Falle von annähernd kreisförmigen Pixelfenstern der Vektorraum $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$. Beide wurden im letzten Kapitel bereits behandelt.

11.2 Demonstrationsbeispiele

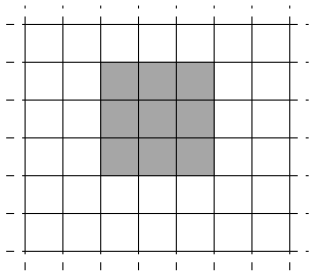
Damit wollen wir den oben beschriebenen Algorithmus an Beispielen demonstrieren. Wir wollen also feststellen, ob zwei lokale Bildausschnitte korrespondieren, d.h. ob sie näherungsweise Abbilder ein und desselben Oberflächenausschnitts eines Objekts sind. Dazu werden wir Algorithmus 11.1 in mehreren Beispielen demonstrieren, wobei wir allerdings leicht von Algorithmus 11.1 abweichen, indem wir uns manuell ein Pixelfenster im Musterbild suchen, das wir im zweiten Bild, also im Suchbild, wiederfinden wollen. Wie in Anwendungen üblich betrachten wir zur Suche nicht das gesamte Suchbild, sondern wir schränken das Suchbild zur Korrespondenzsuche manuell auf einen kleineren Suchbereich ein. In praktischen Anwendungen werden in der Regel Suchfenster der Dimension 40×40 Pixel verwendet (siehe z.B. [Pis02]). Wir werden in den folgenden Beispielen Suchfenster der Dimension 80×80 Pixel verwenden, was die Problemstellung grundsätzlich eher erschwert. Da wir wissen, dass das zu suchende Pixelfenster in transformierter Form im Suchbild zu finden ist, werden wir den „nächsten Nachbarn" als korrespondierend ansehen, d.h. das Pixelfenster im Suchbild, dessen lokales Bildmerkmal dem gesuchten am nächsten kommt. Wir betrachten dabei die beiden klassischen Arten n-zulässiger, diskret-konvexer Mengen aus Abbildung 11.1 als lokale Pixelfenster.

Bemerkung 11.2.1. (Wahl der Bilder)
Da wir in dieser Arbeit keine konkrete Anwendung verfolgen, sondern ein Konzept vorstellen und hier nun demonstrieren wollen, liegt uns natürlich auch kein Datenmaterial mit konkretem Praxisbezug vor. Außerdem ist es auch gar nicht das Ziel dieser Arbeit, nur eine spezielle Anwendung zu bedienen. Als Beispielbilder wählen wir deshalb typische Bilder aus Datensätzen der University of Oxford[27], die im Zusammenhang mit Stereobildern mit breiter Basis häufig referenziert werden. Derartige Bilder können durchaus als „schwieriger" eingestuft werden, da die Situation in einer bestimmten Anwendung der Praxis meist kontrollierter ist, wenn man z.B. an die Situation der Erkennung eines Werkstücks an einem Fließband denkt.

Im ersten Unterabschnitt werden wir nun 3×3-Pixelfenster verwenden. Wir werden dabei insbesondere sehen, dass wir durch unser Verfahren ein sehr ähnliches Distanzmaß wie in [Pis02] erhalten. Und dieses Distanzmaß ist bereits erfolgreich in der Praxis im Einsatz.

11.2.1 Korrespondenzfindung auf 3×3-Pixelfenstern

Auf 3×3-Pixelfenstern ist der reelle Vektorraum $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ ein geeigneter Rekonstruktionsraum für die zeitintegrierte Sensorinputfunktion. Dann ist $C' \subseteq \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ mit den Elementen



$$c'_{0,0}(x,y) = 1 \qquad c'_{0,1}(x,y) = y \qquad c'_{1,0}(x,y) = x,$$
$$c'_{0,2}(x,y) = y^2 - \tfrac{3}{4} \qquad c'_{1,1}(x,y) = xy \qquad c'_{2,0}(x,y) = x^2 - \tfrac{3}{4}$$

eine geeignete Orthogonalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ und C mit den Elementen

$$c_{0,0}(x,y) = \tfrac{1}{3} \qquad c_{0,1}(x,y) = \tfrac{1}{\sqrt{6}}y \qquad c_{1,0}(x,y) = \tfrac{1}{\sqrt{6}}x,$$
$$c_{0,2}(x,y) = \tfrac{1}{\sqrt{2}}(y^2 - \tfrac{3}{4}) \qquad c_{1,1}(x,y) = \tfrac{1}{2}xy \qquad c_{2,0}(x,y) = \tfrac{1}{\sqrt{2}}(x^2 - \tfrac{3}{4})$$

[27] siehe http://www.robots.ox.ac.uk/~vgg/data/

eine geeignete Orthonormalbasis von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ auf 3×3-Pixelfenstern. Laut Beispiel 10.4.2 sind bzgl. der Basis C' die folgenden Polynome fundamentale Invarianten des Invariantenrings $\mathbb{R}[a_{0,0}, \ldots, a_{2,0}]^G$ mit $G = \mathrm{SO}_2(\mathbb{R})$:

$$\begin{aligned}
f'_1 &= a_{0,0} & f'_2 &= a_{0,1}^2 + a_{1,0}^2, \\
f'_3 &= a_{0,2} + a_{2,0} & f'_4 &= a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2, \\
f'_5 &= a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2} & f'_6 &= a_{0,1}^2 a_{1,1} - 2a_{0,1}a_{1,0}a_{0,2} + 2a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1}
\end{aligned}$$

Dabei bilden die ersten fünf Polynome ein Erzeugendensystem bzgl. der orthogonalen Gruppe $\mathrm{O}_2(\mathbb{R})$. Bzgl. der Orthonormalbasis C sind folgende Polynome fundamentale Invarianten für die spezielle orthogonale Gruppe:

$$\begin{aligned}
f_1 &= a_{0,0}, & f_4 &= a_{0,2}a_{2,0} - \tfrac{1}{8}a_{1,1}^2, \\
f_2 &= a_{0,1}^2 + a_{1,0}^2, & f_5 &= a_{0,1}^2 a_{2,0} - \tfrac{1}{2}\sqrt{2}a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}, \\
f_3 &= a_{0,2} + a_{2,0}, & f_6 &= a_{0,1}^2 a_{1,1} - 2\sqrt{2}a_{0,1}a_{1,0}a_{0,2} + 2\sqrt{2}a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1},
\end{aligned}$$

wobei auch hier die ersten fünf Polynome die fundamentalen Invarianten bzgl. der orthogonalen Gruppe bilden. In Bemerkung 10.3.4 haben wir die vier einfachen fundamentalen Invarianten $f'_1, \ldots, f'_4$ anschaulich interpretieren können. So anschaulich diese Polynome auch waren, sie alleine reichen jedoch nicht, um die Bahnen zu trennen, wie das folgende einfache Beispiel zeigt.

Beispiel 11.2.2. Wir betrachten die Gruppe $G = \mathrm{O}_2(\mathbb{R})$ und die folgenden beiden Polynomfunktionen in $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$:

$$\begin{aligned}
p(x, y) &= 0 \cdot 1 + 1 \cdot y + 1 \cdot x + 1 \cdot y^2 + 2 \cdot xy + 1 \cdot x^2, \\
\tilde{p}(x, y) &= 0 \cdot 1 + \sqrt{2} \cdot y + 0 \cdot x + 0 \cdot y^2 + 0 \cdot xy + 2 \cdot x^2
\end{aligned}$$

Zur besseren Lesbarkeit setzen wir:

$$\begin{aligned}
&a_{0,0} := 0, \quad a_{0,1} := 1, \quad a_{1,0} := 1, \quad a_{0,2} := 1, \quad a_{1,1} := 2, \quad a_{2,0} := 1, \\
&\tilde{a}_{0,0} := 0, \quad \tilde{a}_{0,1} := \sqrt{2}, \quad \tilde{a}_{1,0} := 0, \quad \tilde{a}_{0,2} := 0, \quad \tilde{a}_{1,1} := 0, \quad \tilde{a}_{2,0} := 2.
\end{aligned}$$

Dann folgt mit den fundamentalen Invarianten aus Satz 10.3.3:

$$\begin{aligned}
f_1(a_{0,0}, \ldots, a_{2,0}) &= a_{0,0} = 0, \\
f_2(a_{0,0}, \ldots, a_{2,0}) &= a_{0,1}^2 + a_{1,0}^2 = 2, \\
f_3(a_{0,0}, \ldots, a_{2,0}) &= a_{0,2} + a_{2,0} = 2, \\
f_4(a_{0,0}, \ldots, a_{2,0}) &= a_{0,2}a_{2,0} - \tfrac{1}{4}a_{1,1}^2 = 0, \\
f_5(a_{0,0}, \ldots, a_{2,0}) &= a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2} = 0
\end{aligned}$$

und

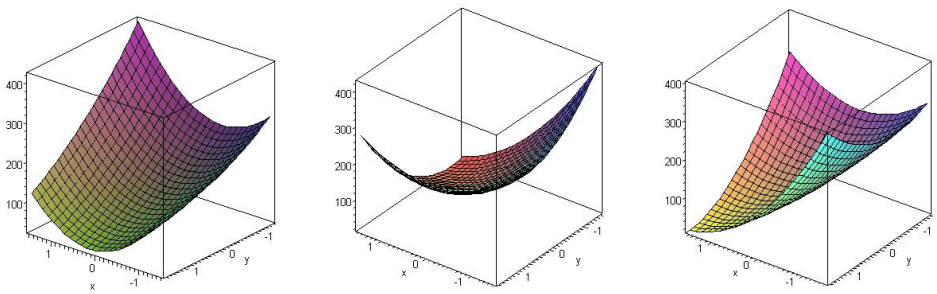
$$\begin{aligned}
f_1(\tilde{a}_{0,0}, \ldots, \tilde{a}_{2,0}) &= \tilde{a}_{0,0} = 0, \\
f_2(\tilde{a}_{0,0}, \ldots, \tilde{a}_{2,0}) &= \tilde{a}_{0,1}^2 + \tilde{a}_{1,0}^2 = 2, \\
f_3(\tilde{a}_{0,0}, \ldots, \tilde{a}_{2,0}) &= \tilde{a}_{0,2} + \tilde{a}_{2,0} = 2, \\
f_4(\tilde{a}_{0,0}, \ldots, \tilde{a}_{2,0}) &= \tilde{a}_{0,2}\tilde{a}_{2,0} - \tfrac{1}{4}\tilde{a}_{1,1}^2 = 0, \\
f_5(\tilde{a}_{0,0}, \ldots, \tilde{a}_{2,0}) &= \tilde{a}_{0,1}^2 \tilde{a}_{2,0} - \tilde{a}_{0,1}\tilde{a}_{1,0}\tilde{a}_{1,1} + \tilde{a}_{1,0}^2 \tilde{a}_{0,2} = 4
\end{aligned}$$

Somit ist klar, dass die Invarianten $f_1, \ldots, f_4$ nicht ausreichen, um die Polynomfunktionen p und $\tilde{p}$ zu trennen. ◁

Zur Trennung der Bahnen sind also zunächst sicher alle fundamentalen Invarianten nötig. Bevor wir zu Bildern übergehen, wollen wir noch halbwegs „exakte" Beispiele betrachten.

Beispiel 11.2.3. Gegeben sind in $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ drei Polynomfunktionen p_1, p_2, p_3, deren Graphen in Abbildung 11.3 dargestellt sind und die wie folgt definiert sind:

$$p_1(x,y) = 52.16717x^2 - 20.24995xy + 13.49949y^2 + 12.87347x - 68.80928y + 103.30555$$
$$p_2(x,y) = 29.37445x^2 + 43.09750xy + 36.29221y^2 - 67.28353x + 19.32278y + 103.30555$$
$$p_3(x,y) = 37.13983x^2 - 42.79095xy + 28.52683y^2 - 62.77151x - 30.98679y + 103.30555$$



(a) $p_1(x,y)$ (b) $p_2(x,y)$ (c) $p_3(x,y)$

Abbildung 11.3: Graphen verschiedener Polynomfunktionen. Die ersten beiden Polynomfunktionen (von links nach rechts) unterscheiden sich durch eine Drehung, die erste und die dritte durch eine Spiegelung.

Seien nun $f_1, \ldots, f_6$ bzw. $f_1, \ldots, f_5$ die fundamentalen Invarianten bzgl. $\mathrm{SO}_2(\mathbb{R})$ bzw. $\mathrm{O}_2(\mathbb{R})$ und sei $v_j \in \mathbb{R}^6$ der Koordinatenvektor von p_j bzgl. der Standardtermbasis B. Da die Invarianten bzgl. B und bzgl. C identisch sind, können wir hier in diesem Fall ohne Einschränkung auch die leicht ablesbaren Koeffizienten verwenden. Klar ist, dass $f_1(v_1) = f_1(v_2) = f_1(v_3)$ gilt. Weiter gilt:

$$f_2(v_1) \approx 4900.44324 \qquad f_3(v_1) \approx 65.66666 \qquad f_4(v_1) \approx 601.71507$$
$$f_5(v_1) \approx 231296.31366 \qquad f_6(v_1) \approx -161026.59512$$

und

$$f_2(v_2) \approx 4900.44323 \qquad f_3(v_2) \approx 65.66666 \qquad f_4(v_2) \approx 601.71508$$
$$f_5(v_2) \approx 231296.30084 \qquad f_6(v_2) \approx -161026.61351$$

sowie

$$f_2(v_3) \approx 4900.44362 \qquad f_3(v_3) \approx 65.66666 \qquad f_4(v_3) \approx 601.71526$$
$$f_5(v_3) \approx 231296.30858 \qquad f_6(v_2) \approx 161026.58943$$

Da v_1 und v_2 näherungsweise an allen Invarianten $f_1, \ldots, f_6$ übereinstimmen, liegen p_1 und p_2 näherungsweise in einer Bahn bzgl. der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$. Die Polynomfunktion p_3 liegt sicher nicht in dieser Bahn, da die Funktionswerte von f_6 für v_3 stark von den Funktionswerten von f_6 der anderen beiden Koordinatenvektoren abweicht. Allerdings liegen alle drei Polynomfunktionen in einer Bahn bzgl. der orthogonalen Gruppe $\mathrm{O}_2(\mathbb{R})$, da alle drei an den fundamentalen Invarianten $f_1, \ldots, f_5$ der orthogonalen Gruppe näherungsweise übereinstimmen. Polynomfunktion p_2 entstand aus p_1 durch eine Drehung um 90° und p_3 entstand aus p_1 durch die Transformation anhand der orthogonalen Matrix $\frac{1}{5}\left(\begin{smallmatrix} -3 & 4 \\ 4 & 3 \end{smallmatrix}\right)$. ◁

Bei der praktischen Arbeit mit Bildern werden die Ungenauigkeiten natürlich sehr viel größer sein. Womit wir nun Beispiele mit Bildern betrachten wollen. Dazu verwenden wir für die folgenden Beispiele die spezielle orthogonale Gruppe $G := \mathrm{SO}_2(\mathbb{R})$ und die Orthonormalbasis C von $V := \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R}) \cong \mathbb{R}^6$. Der algebraische Quotient $\pi : V \to V/\!\!/G \subseteq \mathbb{R}^5$ ist definiert durch

$$p \mapsto (f_1(\kappa_C(p)), \ldots, f_6(\kappa_C(p))).$$

Als Metrik $\tilde{d} : \mathbb{R}^5 \times \mathbb{R}^5 \to \mathbb{R}$ werden wir die Manhattan-Metrik verwenden, die zumindest in den betrachteten Beispielen die richtigen Ergebnisse lieferte. Das von Pisinger in [Pis02] konstruierte Distanzmaß kommt einer ebenfalls einer Manhattan-Metrik sehr nahe, wenngleich einer mit Gewichten versehen Manhattan-Metrik (siehe Beispiel 11.1.5). Wie schon zu Beginn des Kapitels erläutert, kann die Frage nach einer geeigneten Metrik nicht abschließend geklärt werden. Sie muss experimentell abhängig von der konkreten Anwendung ermittelt werden. Zusammengefasst betrachten wir die Halbmetrik $d_G : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}$ auf dem Koordinatenraum $\mathbb{R}^6$, die mit $v = (v_1, \ldots, v_6)$ und $w = (w_1, \ldots, w_6)$ definiert ist durch

$$\begin{aligned}
d_G(v,w) &= \tilde{d}(\pi(\kappa_C^{-1}(v)), \pi(\kappa_C^{-1}(w))) = |f_1(v) - f_1(w)| + \ldots + |f_6(v) - f_6(w)| \\
&= |v_1 - w_1| + |(v_2^2 + v_3^2) - (w_2^2 + w_3^2)| + |(v_4 + v_6) - (w_4 + w_6)| \\
&\quad + |(v_4 v_6 - \tfrac{1}{8} v_5^2) - (w_4 w_6 - \tfrac{1}{8} w_5^2)| \\
&\quad + |(v_2^2 v_6 - \tfrac{1}{2}\sqrt{2} v_2 v_3 v_5 + v_3^2 v_4) - (w_2^2 w_6 - \tfrac{1}{2}\sqrt{2} w_2 w_3 w_5 + w_3^2 w_4)| \\
&\quad + |(v_2^2 v_5 - 2\sqrt{2} v_2 v_3 v_4 + 2\sqrt{2} v_2 v_3 v_6 - v_3^2 v_5) \\
&\qquad - (w_2^2 w_5 - 2\sqrt{2} w_2 w_3 w_4 + 2\sqrt{2} w_2 w_3 w_6 - w_3^2 w_5)|
\end{aligned}$$

Die Parallelen zum Distanzmaß aus Beispiel 11.1.5 sind also klar zu erkennen. Gut möglich, dass in bestimmten Praxissituationen ein auf ähnliche Weise wie in Beispiel 11.1.5 aus den Invarianten konstruiertes Distanzmaß gute Resultate erzielen kann.

Beispiel 11.2.4. Wir betrachten als erstes Beispiel die nachfolgenden Bilder einer Abteikirche, die wir zu Beginn in Kapitel 1 schon gesehen haben.

Abbildung 11.4: Dargestellt ist die Kirche aus zwei verschiedenen Blickrichtungen, wobei wir jeweils die obere Hälfte des Turms heranzoomen.

Wir wollen nun die Ecke des Sims des linken Turms (rot umrandeter 3×3-Fenster) im rechten Bild wiederfinden. Den Grauwertvektor dieser diskret-konvexen Menge $M_1 \subseteq \mathbb{Z} \times \mathbb{Z}$ lautet:

$$\mathrm{gv}(M_1) = (207, 129, 106, 213, 87, 49, 247, 219, 116)$$

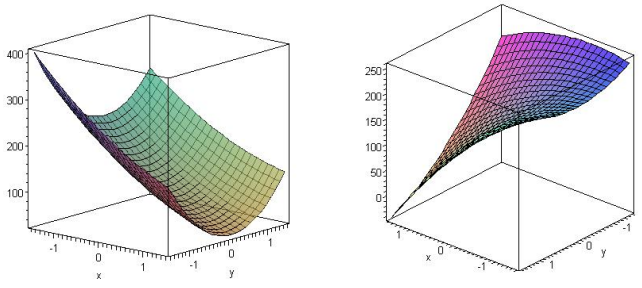
und das zugehörige lokale Bildmerkmal bzgl. der Basis C ist

$$v_1 = (457.666, -57.154, -161.666, 76.838, 15, 16.027).$$

Somit erhalten wir auf M_1 das folgende Rekonstruktionspolynom:

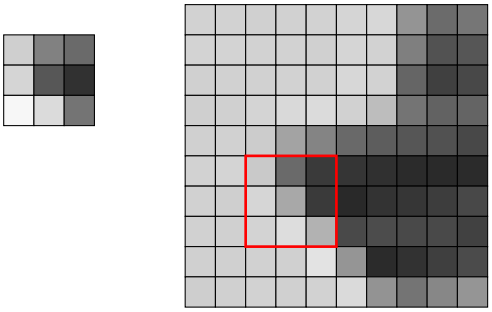
$$p_1(x,y) = \tfrac{3719}{36} - \tfrac{70}{3}y - 66x + \tfrac{163}{3}y^2 + \tfrac{15}{2}xy + \tfrac{34}{3}x^2,$$

dessen Graph untenstehend abgebildet ist.



(a) $p_1(x,y)$ (b) $p_2(x,y)$

Zur Suche schränken wir das Suchbild auf ein Fenster der Dimension 80×80 Pixel ein, wir stellen hier aber nur das rot markierte 10×10-Fenster dar. Wir zoomen die Situation also etwas weiter heran, d.h. wir betrachten das folgende Lokalisierungsfenster im Musterbild (links) und den abgebildeten Teil des Suchbildes (rechts).



Wir berechnen nun für für jedes innere Pixel des Suchbildes die Orthonormalkoeffizienten auf 3×3-Lokalisierungsfenstern bzgl. der Orthonormalbasis C. Anschließend werden die Bilder aller lokaler Bildmerkmale unter π bestimmt. Für p_1 erhalten wir:

$$\pi(p_1) = (457.666, 29402.666, 92.866, 1203.430, 1962614.783, -1932320.000).$$

Mit der Manhattan-Metrik erhalten wir das rot umrandete Lokalisierungsfenster M_2 im Suchbild mit Grauwertvektor $\mathrm{gv}(M_2) = (201, 106, 58, 214, 168, 58, 211, 221, 178)$ und zugehörigem Bildmerkmal

$$v_2 = (471.666, -100.020, -135.538, 22.391, -55, -16.499).$$

Ein durchaus passendes Ergebnis, wobei hier augenscheinlich Interpretationsspielraum vorhanden ist. Das zugehörige Rekonstruktionspolynom lautet (näherungsweise) wie folgt:

$$p_2(x,y) = 154.097 - 40.833y - 55.333x + 15.833y^2 - 27.5xy - 11.666x^2.$$

Als Funktionswert unter π erhalten wir

$$\pi(p_2) = (471.666, 28374.833, 5.892, -747.569, 773520.999, -1031075.833).$$

Wie man sieht, sind die beiden Auswertungen $\pi(p_1)$ und $\pi(p_2)$ nicht so nah, wie man sich das vielleicht vorstellen würde, aber das Ergebnis kann durchaus als korrekt bezeichnet werden. ◁

Dieses erste Beispiel ist vergleichsweise harmlos und so war es durchaus zu erwarten, dass in diesem Fall ein gutes Ergebnis erzielt werden kann. Denn in diesen Bildern wirkt v.a. eine starke Translation, aber nur eine verhältnismäßig geringe Rotation und ebenfalls eine geringe Skalierung. Dies ist im folgenden Beispiel anders. Hier wird sehr stark rotiert und auch etwas mehr skaliert. So waren auch die Ergebnisse des Verfahrens nach [Pis02] angewandt auf diese Bilder nicht besonders gut (siehe [Sta07]). Diese Bilder sind besonders aufgrund ihrer sehr heterogenen Struktur als äußerst schwer einzustufen, dennoch wollen wir unser Glück versuchen.

Beispiel 11.2.5. Als zweites Beispiel wollen wir die beiden nachfolgenden Bilder[28] betrachten. Wir wollen uns auch hier exemplarisch nur einen Teil beider Bilder näher ansehen. Dazu wählen wir den kleinen Turm im Hintergrund des Schiffes, genauer die Spitze dieses Turms.

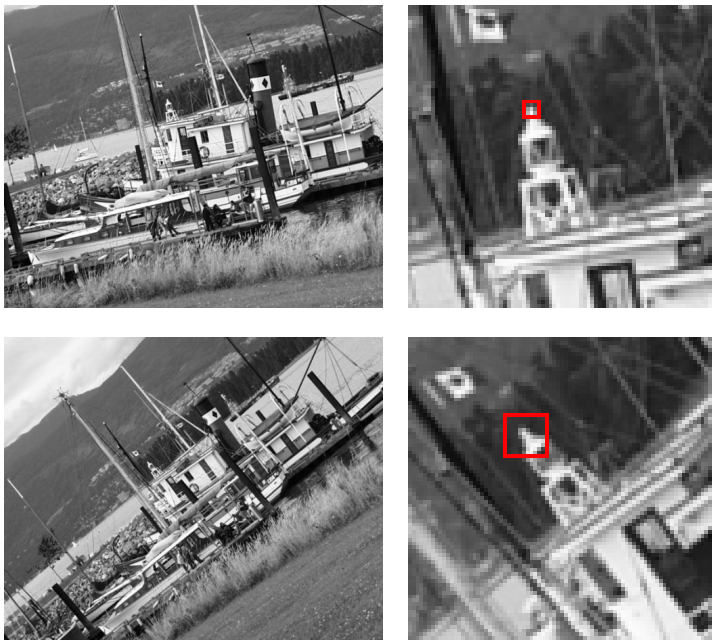
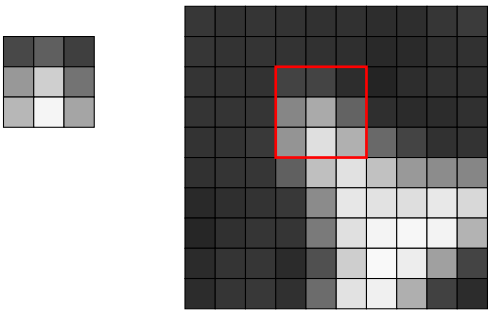


Abbildung 11.5: Die geometrische Transformation, die auf diese beiden Bilder wirkt, ist sehr viel ausgeprägter als in den Bildern aus Teil a).

Auch hier schränken wir die Suche nach dem 3×3-Muster des Musterbildes, das die Turmspitze beinhaltet, im Suchbild auf ein Suchfenster der Dimension 80×80 Pixel ein. Detailliert betrachten werden in diesem Suchfenster den rot umrandeten 10×10-Bereich, der nachfolgend dargestellt ist.



[28] Bildquelle: http://www.robots.ox.ac.uk/~vgg/data/data-aff.html vom 01.08.2015

Das Musterbild auf $M_1 \subseteq \mathbb{Z} \times \mathbb{Z}$ hat folgenden Grauwertvektor:

$$\mathrm{gv}(M_1) = (72, 95, 63, 152, 207, 115, 183, 245, 165)$$

mit zugehörigem lokalen Bildmerkmal

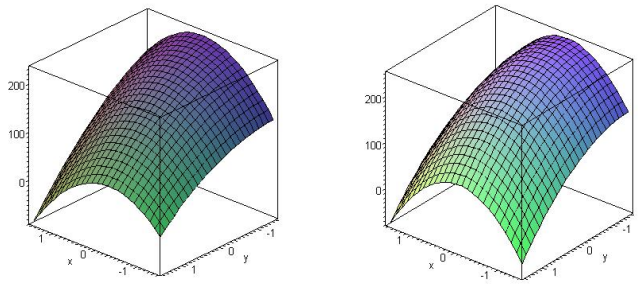
$$v_1 = (432.333, -148.194, -26.127, -29.462, 4.5, -81.081).$$

Das zugehörige Rekonstruktionspolynom $p_1 \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ ist definiert durch

$$p_1(x, y) = \tfrac{3007}{18} - \tfrac{367}{6}y - \tfrac{29}{6}x - \tfrac{79}{6}y^2 - \tfrac{25}{2}xy - \tfrac{253}{6}x^2.$$

Der Graph von p_1 ist nachfolgend zu sehen. Das Bild von p_1 unter π lautet (näherungsweise):

$$\pi(p_1) = (432.333, 22644.166, -110.544, 2386.357, -1813106.954, -469557.250).$$



(a) $p_1(x, y)$ (b) $p_2(x, y)$

Nun verfahren wir wie zuvor: Wir berechnen für alle inneren Pixel des Suchfensters die Orthonormalkoeffizienten auf 3×3-Lokalisierungsfenstern bzgl. der Orthonormalbasis C. Anschließend berechnen wir die Bilder der zu den Koeffizientenvektoren gehörenden Polynomfunktionen unter π. Wir suchen erneut mit der Manhattan-Metrik als Distanzmaß den nächsten Nachbarn zu dem Auswertungsvektor des Musters. Das Ergebnis ist durchaus überzeugend. Als „korrespondierend" wird das rot umrandete 3×3-Fenster eingestuft, das tatsächlich die Spitze des Turms zeigt. Der zu $M_2 \subseteq \mathbb{Z} \times \mathbb{Z}$ gehörende Grauwertvektor ist

$$\mathrm{gv}(M_2) = (67, 68, 45, 134, 170, 99, 148, 223, 176)$$

mit lokalem Bildmerkmal $v_2 = (376.666, -149.827, -11.839, -18.620, -25, -59.632)$. Der Graph des entsprechenden Rekonstruktionspolynoms

$$p_2(x, y) = \tfrac{14597}{72} - \tfrac{121}{2}y - \tfrac{32}{2}x - \tfrac{125}{6}y^2 + \tfrac{9}{4}xy - \tfrac{172}{3}x^2$$

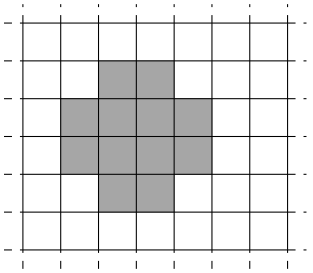
ist oben bereits abgebildet. Das Bild von p_2 unter π lautet (näherungsweise):

$$\pi(p_2) = (376.666, 22588.333, -78.253, 1032.263, -1309896.888, -763464.666).$$

◁

Das waren natürlich nur sehr kleine Beispiele, die zum einen der Veranschaulichung dienen sollten und zum anderen ein Gefühl von dem Potential vermitteln sollten, das die Invariantentheorie nicht nur in Bezug auf die Bildverarbeitung im Allgemeinen, sondern ganz speziell in Bezug auf Anwendungen hat, die sich auf lokale Bildmerkmale stützen. In der Praxis würde man z.B. auch nicht ein Merkmal des Musterbildes willkürlich auswählen und mit allen Merkmalen des Suchbildes vergleichen, sondern zuvor besonders markante Merkmale im Musterbild ermitteln, die zum Vergleich herangezogen werden würden (siehe auch Kapitel 1). Selbstverständlich müssen die Invarianten auch ihre Praxistauglichkeit in einer konkreten Anwendungssituation erst unter Beweis stellen. Dies ist jedoch nicht Ziel und Teil dieser Arbeit.

11.2.2 Korrespondenzfindung auf annähernd kreisförmigen Pixelfenstern



Als zweite Beispielklasse von Lokalisierungsfenstern werden wir annähernd kreisförmige Pixelfenster betrachten. Hat man es in Anwendungen mit Rotationen zu tun, sollten möglichst „kreisförmige" Pixelfenster verwendet werden. Diese Eigenschaft erfüllen die hier betrachteten Pixelfenster mit insgesamt 12 Pixeln in bestmöglicher Weise. Derartige Pixelfenster bieten also im Vergleich zu 3×3-Pixelfenstern insbesondere bei stark rotierten Bildern eine höhere Genauigkeit. Allerdings ist hier der Polynomvektorraum $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ als Rekonstruktionsraum nicht mehr geeignet, da in diesem Fall zu viele Detailinformationen verloren gehen würden. Anders ausgedrückt, die Kardinalität der Pixelfenster und die Dimension von $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ sind zu unterschiedlich. Einen geeigneten Rekonstruktionsraum für die zeitintegrierte Sensorinputfunktion finden wir folglich in dem zehndimensionalen $\mathbb{R}$-Vektorraum $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$. Eine geeignete Orthonormalbasis C von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ lautet wie folgt (siehe Beispiel 9.2.29):

$$
\begin{aligned}
c_{0,0}(x,y) &= \tfrac{1}{6}\sqrt{3} & c_{0,1}(x,y) &= \tfrac{1}{\sqrt{11}}y \\
c_{1,0}(x,y) &= \tfrac{1}{\sqrt{11}}x & c_{0,2}(x,y) &= \tfrac{1}{4(1+\sqrt{3})}(y^2 + (2+\sqrt{3})x^2 - 3 - \sqrt{3}) \\
c_{1,1}(x,y) &= \tfrac{2}{\sqrt{19}}xy & c_{2,0}(x,y) &= \tfrac{1}{4(1+\sqrt{3})}(x^2 + (2+\sqrt{3})y^2 - 3 - \sqrt{3}) \\
c_{0,3}(x,y) &= \tfrac{\sqrt{10}}{6}\left(y^3 - \frac{13}{5}y + \frac{9}{10}x^2y\right) & c_{1,2}(x,y) &= \tfrac{\sqrt{110}}{20}\left(y^2x - \frac{17}{33}x\right) \\
c_{2,1}(x,y) &= \tfrac{\sqrt{110}}{20}\left(x^2y - \frac{17}{33}y\right) & c_{3,0}(x,y) &= \tfrac{\sqrt{10}}{6}\left(x^3 - \frac{13}{5}x + \frac{9}{10}xy^2\right)
\end{aligned}
$$

Wir werden hier nun die orthogonale Gruppe $G = \mathrm{O}_2(\mathbb{R})$ verwenden. In diesem Fall gibt es bzgl. der Standardtermbasis $B = (b_{i,j} : \ i + j \leq 3)$ insgesamt 19 fundamentale Invarianten des Invariantenrings $\mathbb{R}[a_{0,0}, \ldots, a_{3,0}]^G$, die in Abschnitt B.2 zu finden sind. Aus Gründen der Übersichtlichkeit verzichten wir auf die Berechnung fundamentaler Invarianten bzgl. der Basis C. Stattdessen werden wir für die folgenden Beispiele die lokalen Bildmerkmale, also die Koordinatenvektoren bzgl. der Orthonormalbasis C, jeweils in Koordinatenvektoren bzgl. der Standardtermbasis B umrechnen und die Invarianten $f_1, \ldots, f_{19}$ aus Abschnitt B.2 zur Konstruktion eines Distanzmaßes verwenden, jedoch hier als Polynom im Polynomring $\mathbb{R}[a_{0,0}, \ldots, a_{3,0}]$. Mit $V := \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ ist der algebraische Quotient $\pi : V \to V /\!\!/ G \subseteq \mathbb{R}^{19}$ definiert durch

$$p \mapsto (f_1(\kappa_B(p)), \ldots, f_{19}(\kappa_B(p))).$$

Als Metrik $\tilde{d} : \mathbb{R}^{19} \times \mathbb{R}^{19} \to \mathbb{R}$ werden wir ebenfalls die Manhattan-Metrik verwenden. Damit erhalten wir eine Halbmetrik $d_G : \mathbb{R}^{10} \times \mathbb{R}^{10} \to \mathbb{R}$, die definiert ist durch

$$d_G(v,w) = \tilde{d}(\pi(\kappa_B^{-1}(v)), \pi(\kappa_B^{-1}(w))) = |f_1(v) - f_1(w)| + \ldots + |f_{19}(v) - f_{19}(w)|$$

und für die genau dann $d_G(v,w) = 0$ gilt, wenn die zugehörigen Polynomfunktionen $\kappa_B^{-1}(v)$ und $\kappa_B^{-1}(w)$ in einer Bahn sind.

Beispiel 11.2.6. Als erstes Beispiel werden wir auch hier verhältnismäßig schwierige Bilder betrachten, nämlich die bereits oben verwendeten Bilder eines Schiffes[29]. Wir wollen uns auch hier exemplarisch nur einen Teil beider Bilder näher ansehen. Wir betrachten hier den Kamin des Schiffes etwas genauer.

[29] Bildquelle: http://www.robots.ox.ac.uk/~vgg/data/data-aff.html vom 01.08.2015

Abbildung 11.6: Beispiel zur Korrespondenzfindung lokaler Bildmerkmale mit kreisförmigen Pixelfenstern.

Wir schränken die Suche nach dem rot markierten kreisrunden Muster, das eine Ecke der Raute auf dem Kamin des Schiffes beinhaltet, im Suchbild (unten) auf das rechts unten abgebildete 80×80-Bild ein. Das rot markierte Muster auf $M_1 \subseteq \mathbb{Z} \times \mathbb{Z}$ hat folgenden Grauwertvektor (zeilenweise von oben nach unten):

$$\mathrm{gv}(M_1) = (208, 203, 203, 136, 129, 195, 165, 61, 56, 166, 15, 27)$$

mit zugehörigem lokalen Bildmerkmal

$$v_1 = (451.487, 199.298, -3.919, -41.845, -9.176, 122.683, 9.913, 4.910, -28.637, 4.585).$$

Das entsprechende Rekonstruktionspolynom $p_1 \in \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ ist (näherungsweise) definiert durch

$$\begin{aligned} p_1(x, y) = 95.329 + 54.242y - 8.791x - 3.064y^2 - 4.210xy + 38.067x^2 \\ + 5.225y^3 + 4.750xy^2 - 10.315x^2y + 2.416x^3. \end{aligned}$$

Den Bildvektor $\pi(p_1) \in \mathbb{R}^{19}$ wollen wir hier aus verständlichen Gründen aber nicht präsentieren. Wir berechnen nun im Suchbild für jedes innere Pixel die lokalen Bildmerkmale bzgl. der Basis C auf kreisförmigen Pixelfenstern und suchen nach dem Merkmal v_2 bzw. der Polynomfunktion $p_2 \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$, dessen Auswertung $\pi(p_2)$ dem Vektor $\pi(p_1)$ am nächsten kommt. Wir erhalten in Bildern das nachfolgende Ergebnis.

Wie wir sehen, liefert der Algorithmus das falsche Muster, was aber nicht ungewöhnlich ist. Denn das gewählte Muster im Musterbild war sicherlich zu wenig markant. Selbst mit bloßem Auge lässt sich zwischen oberer und unterer Ecke nur schwer unterscheiden. Außerdem sind die beiden Bilder bzgl. der Skalierung recht unterschiedlich. Demnach ist es nicht verwunderlich, dass anstatt der oberen Ecke der Raute die untere Ecke erkannt wurde. Das oben rot markierte Muster $M_2 \subseteq \mathbb{Z} \times \mathbb{Z}$ besitzt folgenden Grauwertvektor:

$$\mathrm{gv}(M_2) = (18, 43, 21, 14, 34, 181, 109, 52, 36, 167, 177, 132).$$

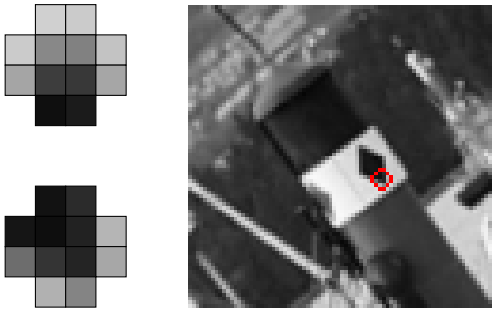
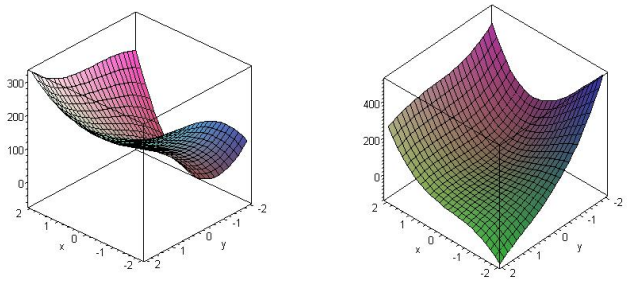


Abbildung 11.7: Links oben ist das zu suchende Muster dargestellt. Dieses wurde im Suchbild (rechts) falsch wieder gefunden. Links unten ist das gefundene Muster abgebildet.

Das zugehörige lokale Bildmerkmal lautet wie folgt:

$$v_2 = (284.056, -129.348, 96.182, 25.719, 63.318, 120.915, -16.416, -40.903, -12.020, 32.571).$$

Dementsprechend erhalten wir ein Rekonstruktionspolynom $p_2 \in \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$, das (näherungs-



(a) $p_1(x, y)$ (b) $p_2(x, y)$

weise) wie folgt definiert ist:

$$\begin{aligned} p_2(x, y) = 18.505 - 13.256y - 4.583x + 19.847y^2 + 29.052xy + 43.646x^2 \\ - 8.652y^3 - 6.000xy^2 - 14.090x^2y + 17.166x^3 \end{aligned}$$

Die Graphen der beiden Polynomfunktionen sind oben stehend abgebildet. ◁

Wir haben mit dem letzten Beispiel bewusst einen Fall gewählt, bei dem der Algorithmus an seine Grenzen stößt, aber auch die gängigen Verfahren würden in dieser Situation ihre Probleme haben und ebenfalls mit hoher Wahrscheinlichkeit scheitern.

Bemerkung 11.2.7. (Probleme bei Skalierungen)
Mit dem oben gewählten Distanzmaß erhält man natürlich keine Skalierungsinvarianz. Dieses Maß ist auch nicht besonders robust gegenüber Skalierungen. Somit war auch aus diesem Grund zu erwarten, dass der Algorithmus bei den obigen Bildern falsch liegt. Abhilfe könnten hier aus Invarianten konstruierte Maße sein, die auf rationalen Invarianten aufbauen, wie es zum Teil in Beispiel 11.1.5 der Fall ist. Einen theoretischen Ansatz dazu stellen wir im nächsten Abschnitt vor.

Zum Abschluss wollen wir noch ein letztes Beispiel präsentieren.

Beispiel 11.2.8. Wir betrachten Bilder einer Sequenz, die wir zum Teil schon kennen, d.h. wir verwenden das nachfolgend abgebildete Paar von Such- und Musterbild, jeweils in der Dimension 80×80 Pixel.

Abbildung 11.8: Das rot markierte Muster im Musterbild (links) soll im Suchbild (rechts) gefunden werden.

Das zu suchende Muster $M_1 \subseteq \mathbb{Z} \times \mathbb{Z}$ ist im Musterbild rot markiert. Es besitzt folgenden Grauwertvektor:

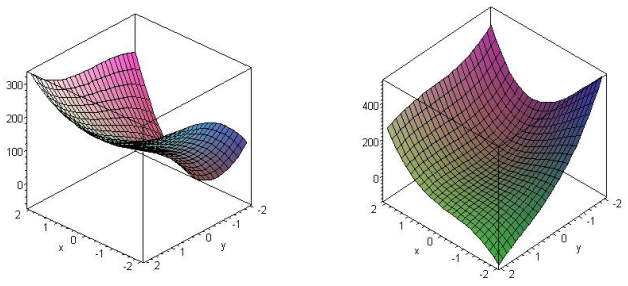
$$\mathrm{gv}(M_1) = (202, 214, 208, 198, 138, 93, 211, 197, 46, 12, 218, 188)$$

mit zugehörigem lokalen Bildmerkmal

$$v_1 = (555.699, 30.301, -176.534, 110.431, 53.797, -19.445, -52.553, 46.385, -5.303, 50.438).$$

Somit ist das Rekonstruktionspolynom $p_1 \in \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ (näherungsweise) definiert durch

$$\begin{aligned} p_1(x, y) = {} & 121.018 + 82.584y - 134.875x + 35.933y^2 + 24.684xy + 3.464x^2 \\ & - 27.698y^3 + 48.250xy^2 - 27.709x^2y + 26.583x^3 \end{aligned}$$



(a) $p_1(x, y)$ (b) $p_2(x, y)$

Algorithmus 11.1 stuft das nachfolgend abgebildete lokale Bildmerkmal $M_2 \subseteq \mathbb{Z} \times \mathbb{Z}$ mit Grauwertvektor

$$\mathrm{gv}(M_2) = (204, 191, 196, 193, 125, 108, 197, 206, 86, 31, 205, 107)$$

und (approximativem) Merkmalsvektor

$$v_2 = (533.760, 52.915, -159.951, 55.521, 62.056, -27.577, -14.924, -7.627, 17.677, 49.015)$$

als korrespondierend ein.

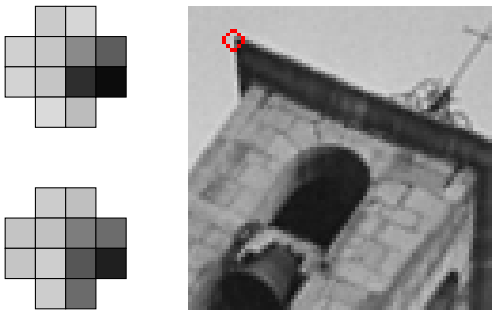


Abbildung 11.9: Links oben ist das zu suchende Muster dargestellt. Dieses wurde im Suchbild (rechts) augenscheinlich korrekt gefunden. Links unten ist das gefundene Muster abgebildet.

Das zu v_2 gehörende Rekonstruktionspolynom $p_2 \in \mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ ist also (näherungsweise) definiert durch

$$\begin{aligned} p_2(x,y) = 141.982 + 31.629y - 113.333x + 16.437y^2 + 28.473xy - 4.337x^2 \\ - 7.865y^3 + 19.250xy^2 + 2.191x^2y + 25.833x^3 \end{aligned}$$

Die Graphen der beiden Rekonstruktionspolynome sind oben bereits abgebildet. ◁

11.3 Ausblick

Natürlich muss das hier vorgestellte Verfahren seine Praxistauglichkeit erst nachweisen. Dazu müsste man die Invarianten lokaler Bildmerkmale in einem konkreten Praxisbezug einsetzen und der Tauglichkeit mit den gängigen Verfahren vergleichen und bewerten. Das war allerdings explizit nicht Ziel dieser Arbeit und bleibt damit offen. Die ersten zarten Versuche, die als „proof of concept" unternommen wurden, lassen aber durchaus positive Erwartungen zu. Neben den bisher untersuchten Translationen und Rotationen spielen natürlich auch Skalierungen eine große Rolle. Wir werden im ersten Unterabschnitt nun unter anderem einen Vorschlag unterbreiten, wie Invarianten erzeugt werden können, die zusätzlich skalierungsinvariant sind.

Darüber hinaus sind geometrische Transformationen bei weitem nicht die einzigen Probleme, mit denen man bei der Erkennung und Lokalisation von Objekten in Bildern zu tun hat. So ist es möglich und in vielen Anwendungen wahrscheinlich, dass man zusätzlich mit unterschiedlichen Kontraststufen, unterschiedlichen Kompressionsstufen oder unterschiedlichen Helligkeitsstufen zu kämpfen hat. Man denke an das Beispiel der Verfolgung von Fahrbahnmarkierungen in Fahrerassistenzsystemen: Hier wäre es fatal, wenn die Verfolgung an unterschiedlichen Lichtverhältnissen scheitern würde. Aber genau mit diesen hat man sicherlich zu tun, z.B. wenn man sich am Anfang oder am Ende eines Tunnels befindet. Auch unter diesen „Störaspekten" müssen sich die Invarianten erst beweisen. Eine mögliche Methode, von Helligkeitsschwankungen weitgehend unabhängig zu sein, wollen wir im zweiten Unterabschnitt noch andeuten.

11.3.1 Invarianten unter Skalierungen

Invarianz unter Rotationen oder unter der Operation der orthogonalen Gruppe ist nicht alles, was für praxisnahe Anwendungen von Bedeutung ist. Von zentralem Interesse für die Praxis ist die zusätzliche Invarianz unter Skalierungen, genauer unter simultanen Skalierungen in x- und y-Richtung. Wir sind hier also insgesamt an Invarianten unter der Operation einer der folgenden Gruppen interessiert:

$$\{\alpha \cdot \mathcal{A} :\ \alpha \in \mathbb{R} \setminus \{0\} \text{ und } \mathcal{A} \in \mathrm{SO}_2(\mathbb{R})\} \quad \text{oder} \quad \{\alpha \cdot \mathcal{A} :\ \alpha \in \mathbb{R} \setminus \{0\} \text{ und } \mathcal{A} \in \mathrm{O}_2(\mathbb{R})\}$$

Dazu betrachten wir zunächst die Operation der Gruppe $G = \{\alpha \cdot \mathcal{I}_2 :\ \alpha \in \mathbb{R} \setminus \{0\}\}$. Diese Gruppe operiert wie die Gruppen zuvor durch $(\mathcal{A}, p) \mapsto p \circ T_{\mathcal{A}}$ auf dem reellen Vektorraum $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, den wir mit der Standardtermbasis $B = (b_{i,j})$ betrachten wollen. Sei $\mathcal{A} \in G$, d.h. es gibt ein $\alpha \in \mathbb{R} \setminus \{0\}$ mit $\mathcal{A} = \alpha \cdot \mathcal{I}_2$. Für die einzelnen Basisfunktionen $b_{i,j} \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit $i + j \leq n$ folgt dann

$$b_{i,j} \circ T_{\mathcal{A}}(x, y) = b_{i,j}(\alpha \cdot x, \alpha \cdot y) = \alpha^{i+j} \cdot b_{i,j}(x, y),$$

insbesondere hat die Operation der Gruppe G also erwartungsgemäß keine Auswirkung auf die Basisfunktion $b_{0,0}$. Ist nun $p \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ mit Koordinaten $a_{i,j}$ bzgl. der Standardtermbasis B gegeben, dann ist für alle $i, j \in \mathbb{N}$ mit $i + j \leq n$ die Koordinate der mit $\mathcal{A}$ transformierten Polynomfunktion $p \circ T_{\mathcal{A}}$ bzgl. $b_{i,j}$ gegeben durch

$$\tilde{a}_{i,j} := \alpha^{i+j} \cdot a_{i,j}.$$

So wie die Basisfunktionen und Koordinaten hier gewählt sind, korrespondiert der Grad des Skalierungsfaktors also in eindeutiger Weise mit den Indizes der Unbestimmten des zum Koordinatenring von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$ isomorphen Polynomrings $P = \mathbb{R}[a_{i,j}]$. Diesem Grad wollen wir einen Namen und eine Bezeichnung geben.

Definition 11.3.1. (Skalierungsgrad von Unbestimmten)
Für alle $i, j \in \mathbb{N}$ mit $i, j \leq n$ heißt die natürliche Zahl $\mathrm{sdeg}(a_{i,j}) := i + j$ der **Skalierungsgrad** der Unbestimmten $a_{i,j}$.

Dieses Konzept lässt sich nahtlos auf Terme und Polynome in P übertragen. Wie wir zu Beginn von Kapitel 10 gesehen haben, operiert auch diese Gruppe G durch $\mathcal{A} \mapsto (f \mapsto f^{\mathcal{A}})$ auf P, wobei $f^{\mathcal{A}} = f(\kappa_B(p \circ T_{\mathcal{A}^{-1}}))$ gilt. Dazu wollen wir ein konkretes Beispiel betrachten.

Beispiel 11.3.2. Sei $n = 2$ und $f = a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}$, d.h. f ist eine fundamentale Invariante der orthogonalen und der speziellen orthogonalen Gruppe. Dann gilt für alle Parameter $\alpha \in \mathbb{R} \setminus \{0\}$:

$$p \circ T_{\alpha\mathcal{I}_2}(x, y) = p(\alpha x, \alpha y) = a_{0,0} + \alpha a_{0,1} y + \alpha a_{1,0} x + \alpha^2 a_{0,2} y^2 + \alpha^2 a_{1,1} xy + \alpha^2 a_{2,0} x^2$$

und es folgt

$$\begin{aligned} f(\kappa_B(p \circ T_{\alpha\mathcal{I}_2})) &= f(a_{0,0}, \alpha a_{0,1}, \alpha a_{1,0}, \alpha^2 a_{0,2}, \alpha^2 a_{1,1}, \alpha^2 a_{2,0}) \\ &= (\alpha a_{0,1})^2 \cdot \alpha^2 a_{2,0} - \alpha a_{0,1} \cdot \alpha a_{1,0} \cdot \alpha^2 a_{1,1} + (\alpha a_{1,0})^2 \cdot \alpha^2 a_{0,2} \\ &= \alpha^4 \cdot (a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}) = \alpha^4 \cdot f \end{aligned}$$

◁

Wie man auch an diesem Beispiel sehen kann, ist es sinnvoll, die folgenden Begriffe, die durch die Operation der Gruppe G auf P induziert werden, zu definieren.

Definition 11.3.3. (Skalierungsgrad)
Sei $t = a_{0,0}^{e_{0,0}} a_{0,1}^{e_{0,1}} a_{1,0}^{e_{1,0}} \cdots a_{n,0}^{e_{n,0}}$ mit $e_{0,0}, e_{0,1}, e_{1,0}, \ldots, e_{n,0} \in \mathbb{N}$ ein Term in $\mathbb{T}^\nu$ und sei $f \in P$ mit $f \neq 0$. Dann heißt die natürliche Zahl

$$\operatorname{sdeg}(t) := \sum_{i+j \leq n} e_{i,j} \cdot \operatorname{sdeg}(a_{i,j})$$

der **Skalierungsgrad** des Terms t und die natürliche Zahl $\max\{\operatorname{sdeg}(t) : \ t \in \operatorname{Supp}(f)\}$ heißt der **Skalierungsgrad** des Polynoms f, der mit $\operatorname{sdeg}(f)$ bezeichnet wird. Ein Polynom $f \in P \setminus \{0\}$ mit $\operatorname{sdeg}(t) = \operatorname{sdeg}(f)$ für alle $t \in \operatorname{Supp}(f)$ bezeichnen wir als **skalierungshomogen**.

Auch dazu wollen wir zunächst ein Beispiel betrachten.

Beispiel 11.3.4. Sei $n = 2$ und $f = a_{0,1}^2 a_{2,0} - a_{0,1} a_{1,0} a_{1,1} + a_{1,0}^2 a_{0,2}$ wie im letzten Beispiel. Für die einzelnen Terme gilt:

$$\begin{aligned}
\operatorname{sdeg}(a_{0,1}^2 a_{2,0}) &= 2 \cdot \operatorname{sdeg}(a_{0,1}) + \operatorname{sdeg}(a_{2,0}) = 2 \cdot (0+1) + (2+0) = 4, \\
\operatorname{sdeg}(a_{0,1} a_{1,0} a_{1,1}) &= \operatorname{sdeg}(a_{0,1}) + \operatorname{sdeg}(a_{1,0}) + \operatorname{sdeg}(a_{1,1}) = 1 + 1 + 2 = 4, \\
\operatorname{sdeg}(a_{1,0}^2 a_{0,2}) &= 2 \cdot \operatorname{sdeg}(a_{1,0}) + \operatorname{sdeg}(a_{0,2}) = 2 + 2 = 4.
\end{aligned}$$

Somit ist f nicht nur homogen bzgl. des Grades der Terme, sondern auch bzgl. des Skalierungsgrades, also skalierungshomogen. Folglich gilt $\operatorname{sdeg}(f) = 4$. ◁

Wie aus Beispiel 11.3.2 deutlich wird, ist ein Polynom $f \in P$ offensichtlich genau dann skalierungsinvariant, wenn $\operatorname{sdeg}(f) = 0$ gilt. Und das wiederum ist aber nur für Potenzen von $a_{0,0}$ der Fall. Somit wird die Suche nach skalierungsinvarianten Polynomen wenig erfolgreich sein. Abhilfe schaffen hier ähnlich wie in Beispiel 1.2.1 **rationale Invarianten**, d.h. wir betrachten den Qotientenkörper

$$\operatorname{Quot}(P) = \left\{ \frac{f}{g} : \ f, g \in P \text{ und } g \neq 0 \right\}$$

des Polynomrings P. Auf folgende Weise lassen sich unter den hier betrachteten Voraussetzungen skalierungsinvariante rationale Funktionen erzeugen.

Satz 11.3.5. (Erzeugung skalierungsinvarianter rationaler Funktionen)
Seien $g, f_1, \ldots, f_\ell \in P \setminus \{0\}$ skalierungshomogene Polynome mit $\operatorname{sdeg}(f_i) \mid \operatorname{sdeg}(g)$ für alle $i \in \{1, \ldots, \ell\}$ gilt. Setze $k_i := \frac{\operatorname{sdeg}(g)}{\operatorname{sdeg}(f_i)}$ für alle $i \in \{1, \ldots, \ell\}$ Dann sind die rationalen Funktionen

$$\frac{f_1^{k_1}}{g}, \ldots, \frac{f_\ell^{k_\ell}}{g} \in \operatorname{Quot}(P)$$

skalierungsinvariant.

Beweis: Sei $\mathcal{A} \in G$, d.h. von der Form $\mathcal{A} = \alpha \cdot \mathcal{I}_2$ mit $\alpha \in \mathbb{R} \setminus \{0\}$, und sei $i \in \{1, \ldots, \ell\}$. Da f_i und g skalierungshomogen sind, gilt $f_i^{\mathcal{A}} = \alpha^{\operatorname{sdeg}(f_i)} \cdot f_i$ und $g^{\mathcal{A}} = \alpha^{\operatorname{sdeg}(g)} \cdot g$. Damit folgt

$$\left(f_i^{k_i}\right)^{\mathcal{A}} = \alpha^{k_i \cdot \operatorname{sdeg}(f_i)} \cdot f_i^{k_i} = \alpha^{\operatorname{sdeg}(g)} \cdot f_i^{k_i},$$

also

$$\frac{\left(f_i^{k_i}\right)^{\mathcal{A}}}{g^{\mathcal{A}}} = \frac{\alpha^{\operatorname{sdeg}(g)} \cdot f_i^{k_i}}{\alpha^{\operatorname{sdeg}(g)} \cdot g} = \frac{f_i^{k_i}}{g}$$

und damit die Behauptung. □

Mit diesem Satz ist es nun möglich, neben dem Polynom $a_{0,0}$ weitere rotations- und skalierungsinvariante Größen bzw. Invarianten der orthogonalen Gruppe, die zusätzlich skalierungsinvariant sind, zu erzeugen, was wir in folgendem Beispiel demonstrieren wollen.

Beispiel 11.3.6. Sei $n = 2$, d.h. wir betrachten den Vektorraum $\mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$.

a) Sei zunächst $G = \mathrm{SO}_2(\mathbb{R})$. Die fundamentalen Invarianten aus Satz 10.2.2 sind nicht nur homogen bzgl. des Grades, sondern auch skalierungshomogen. Genauer gilt im Einzelnen:

$$\begin{aligned} &\mathrm{sdeg}(f_1) = 0, && \mathrm{sdeg}(f_2) = 2, && \mathrm{sdeg}(f_3) = 2, \\ &\mathrm{sdeg}(f_4) = 4, && \mathrm{sdeg}(f_5) = 4, && \mathrm{sdeg}(f_6) = 4 \end{aligned}$$

Wir legen als Polynom g im Sinne des letzten Satzes $g := f_4 = a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2$ fest. Dann sind mit f_2, f_3, f_5, f_6 die Voraussetzungen des letzten Satzes erfüllt. Somit sind die rationalen Funktionen

$$\begin{aligned} r_1 &= \frac{f_2^2}{g} = \frac{(a_{0,1}^2 + a_{1,0}^2)^2}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2}, \\ r_2 &= \frac{f_3^2}{g} = \frac{(a_{0,2} + a_{2,0})^2}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2}, \\ r_3 &= \frac{f_5}{g} = \frac{a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2}, \\ r_4 &= \frac{f_6}{g} = \frac{a_{0,1}^2 a_{1,1} - 2a_{0,1}a_{1,0}a_{0,2} + 2a_{0,1}a_{1,0}a_{2,0} - a_{1,0}^2 a_{1,1}}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2} \end{aligned}$$

skalierungs- und rotationsinvariant.

b) Sei nun $G = \mathrm{O}_2(\mathbb{R})$. Laut Satz 10.3.3 sind die in a) erwähnten Polynome $f_1, \ldots, f_5$ die fundamentalen Invarianten der orthogonalen Gruppe. Somit sind die rationalen Funktionen

$$\begin{aligned} r_1 &= \frac{f_2^2}{g} = \frac{(a_{0,1}^2 + a_{1,0}^2)^2}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2}, \\ r_2 &= \frac{f_3^2}{g} = \frac{(a_{0,2} + a_{2,0})^2}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2}, \\ r_3 &= \frac{f_5}{g} = \frac{a_{0,1}^2 a_{2,0} - a_{0,1}a_{1,0}a_{1,1} + a_{1,0}^2 a_{0,2}}{a_{0,2}a_{2,0} - \frac{1}{4}a_{1,1}^2} \end{aligned}$$

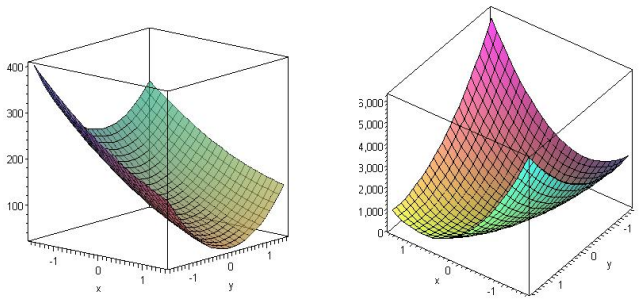
aus a) nicht nur skalierungsinvariant, sondern zudem invariant unter der Operation der orthogonalen Gruppe. ◁

Von der Invarianz dieser rationalen Funktionen wollen wir uns nun an einem konkreten Beispiel vergewissern.

Beispiel 11.3.7. Wir betrachten die Polynomfunktion $p \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ mit

$$p(x, y) = \tfrac{3719}{36} - \tfrac{70}{3}y - 66x + \tfrac{163}{3}y^2 + \tfrac{15}{2}xy + \tfrac{34}{3}x^2.$$

Dabei handelt es sich um das Rekonstruktionspolynom zu dem in Abbildung 1.9 dargestellten Lokalisierungsfenster, dessen Graph auch in nachfolgender Abbildung zu finden ist. Auf dieser Polynomfunktion lassen wir nun die Matrix $\mathcal{A} = \left(\begin{smallmatrix} 3 & 4 \\ 4 & -3 \end{smallmatrix}\right)$ operieren. Wegen $\det(\mathcal{A}) = -25$



(a) $p(x, y)$ (b) $q(x, y)$

handelt sich bei $\mathcal{A}$ also um ein Element der Gruppe $G = \{\alpha \cdot \mathcal{B} : \ \alpha \in \mathbb{R} \setminus \{0\}$ und $\mathcal{B} \in \mathrm{O}_2(\mathbb{R})\}$. Die resultierende Polynomfunktion $q \in \mathscr{P}_{\leq 2}(\mathbb{R}^2, \mathbb{R})$ ist definiert durch

$$q(x, y) = p \circ T_{\mathcal{A}}(x, y) = \tfrac{3719}{36} - 194y - \tfrac{874}{3}x + \tfrac{1741}{3}y^2 - \tfrac{1959}{2}xy + \tfrac{3184}{3}x^2.$$

Dass die konstanten Terme übereinstimmen, war zu erwarten. Nun zu den drei rationalen Invarianten bzgl. $\mathrm{O}_2(\mathbb{R})$ aus dem letzten Beispiel. Damit gilt:

$$\begin{aligned} r_1(\kappa_B(p)) &\approx 39909.83217 & r_1(\kappa_B(q)) &\approx 39909.83217 \\ r_2(\kappa_B(p)) &\approx 7.16636 & r_2(\kappa_B(q)) &\approx 7.16636 \\ r_3(\kappa_B(p)) &\approx 384.39504 & r_3(\kappa_B(q)) &\approx 384.39504 \end{aligned}$$

was die Invarianz der rationalen Funktionen r_1, r_2, r_3 zeigt. ◁

Diese Art der Konstruktion skalierungsinvarianter Größen ist sicherlich ein interessantes theoretisches Konzept, allerdings erwiesen sie sich in ersten Versuchen als nicht besonders gut. Der Grund dürfte relativ einfach sein: Die rationalen Funktionen besitzen zwar einen Skalierungsgrad 0, aber haben im Allgemeinen unterschiedlichen Grad im Zähler und Nenner. Dadurch werden die Ergebnisse verfälscht. Möglich erscheint es aber, Teile davon zu übernehmen. So tauchen in dem in der Praxis erprobten Distanzmaß aus [Pis02] teilweise ähnliche rationale Konzepte auf (siehe Beispiel 11.1.5).

11.3.2 Photometrische Invarianz

Intensitätsschwankungen oder unterschiedliche Kontraststufen lassen sich bekanntlich auf vielfältige Weise durch eine Vorverarbeitung von Such- und Musterbild abmildern. Diese sogenannte **Photometrische Normierung** wird durch eine affin-lineare Grauwerttransformation $t : \mathbb{Z}_{0,255} \to \mathbb{Z}$, definiert durch

$$t(i) = \lfloor a \cdot i + b \rfloor,$$

modelliert, wobei der Parameter $a \in \mathbb{R}$ den Kontrast steuert und der Parameter $b \in \mathbb{R}$ die Helligkeit des Bildes (siehe z.B. [GWE04]). Zur Umsetzung dieser Transformation gibt es verschiedene Ansätze:

- **Kontrastnormierung** (siehe [GWE04], S. 68 f.),
- **Histogrammebnung** (siehe [GWE04], S. 81 ff.)
- **Standardisierung** durch arithmetisches Mittel und Standardabweichung: Dieser Ansatz schneidet in Vergleichstests am besten ab (siehe z.B. [SP05]).

Anstatt die Standardisierung global zu betreiben, könnte man sie auch lokal auf das jeweilige Lokalisierungsfenster anwenden. Sei also $M \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ eine n-zulässige, diskret konvexe Menge mit Kardinalität $m \in \mathbb{N}_+$ und $\mathrm{gv} : \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1} \to \mathbb{Z}_{0,255}$ die Grauwertfunktion des Bildes. Dann ist $\mathrm{gv}(M) \in \mathbb{Z}_{0,255}^m$ der zu M gehörige Grauwertvektor. Wir können uns also erneut in der deskriptiven Statistik bedienen und $\mathrm{gv}|_M$ als Merkmal $X : M \to \mathbb{R}$ auffassen, das jedem Individuum, also jedem Pixel in M, seinen Grauwert zuordnet, d.h. X ist definiert durch $(i,j) \mapsto \mathrm{gv}_{i,j}$. Mit dem arithmetischen Mittel und der Standardabweichung,

$$\overline{x} := \frac{1}{\#M} \sum_{(i,j)\in M} \mathrm{gv}_{i,j} \quad \text{und} \quad \sigma_X := \sqrt{\frac{1}{\#M} \sum_{(i,j)\in M} (\mathrm{gv}_{i,j} - \overline{x})^2},$$

heißt das Merkmal $X^* : M \to \mathbb{R}$ definiert durch

$$X^*(i,j) = \frac{X(i,j) - \overline{x}}{\sigma_X},$$

die **Standardisierung** von X, sie hat arithmetisches Mittel 0 und Standardabweichung 1. Wir können also durch eine derartige Transformation einen **standardisierten Grauwertvektor** $\mathrm{gv}^*(M) \in \mathbb{R}^m$ aus $\mathrm{gv}(M)$ berechnen. Nimmt man diesen Vektor anstatt des Grauwertvektors zur Rekonstruktion der zeitintegrierten Sensorinputfunktion, so erhält man ebenfalls eine Polynomfunktion $p^* \in \mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$, deren Koeffizienten bzgl. einer Orthogonalbasis auch entsprechend benannt werden.

Definition 11.3.8. (Standardisierte lokale Bildmerkmale)
Sei $\mathrm{gv} : \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1} \to \mathbb{Z}_{0,255}$ ein digitales Grauwertbild, $M \subseteq \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$ eine n-zulässige Menge und sei B eine geeignete Orthonormalbasis von $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$. Dann heißt der Koordinatenvektor des auf $\mathrm{gv}^*(M) \in \mathbb{R}^m$ basierenden Rekonstruktionspolynoms bzgl. B ein **standardisiertes lokales Bildmerkmal**.

Damit ist sofort klar, dass die Koordinate, die zur konstanten Basisfunktion in B gehört, bei standardisierten lokalen Bildmerkmalen stets 0 sein muss, da diese Koordinate nur ein Vielfaches des arithmetischen Mittels ist (siehe Korollar 9.2.27). Zudem sind diese lokalen Bildmerkmale invariant unter affin-linearen Grauwerttransformationen.

Satz 11.3.9. *Sei $S := \mathbb{Z}_{0,r-1} \times \mathbb{Z}_{0,s-1}$, $\mathrm{gv} : S \to \mathbb{Z}_{0,255}$ ein digitales Grauwertbild und $M \subseteq S$ eine n-zulässige Menge mit Grauwertvektor $\mathrm{gv}(M) \in \mathbb{Z}_{0,255}^m$. Sei $t : \mathbb{Z}_{0,255} \to \mathbb{R}$ definiert durch $t(g) = a \cdot g + b$ mit $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ sowie $Y : M \to \mathbb{R}$ das durch $Y = t \circ \mathrm{gv}$ definierte Merkmal. Dann gilt:*

$$Y^*(M) = \frac{a}{|a|} \cdot \mathrm{gv}^*(M)$$

Beweis: Sei $(i,j) \in M$. Sei weiter $\overline{x}$ das arithmetische Mittel von $\mathrm{gv}|_M$ und σ die Standardabweichung. Dann gilt:

$$\overline{y} = a \cdot \overline{x} + b \quad \text{und} \quad \sigma_Y = |a| \cdot \sigma$$

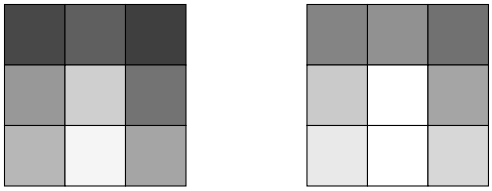
Damit folgt:

$$Y^*(i,j) = \frac{t(\mathrm{gv}_{i,j}) - \overline{y}}{\sigma_Y} = \frac{a \cdot \mathrm{gv}_{i,j} + b - a \cdot \overline{x} - b}{|a| \cdot \sigma} = \frac{a}{|a|} \cdot \frac{\mathrm{gv}_{i,j} - \overline{x}}{\sigma} = \frac{a}{|a|} \cdot \mathrm{gv}^*(i,j)$$

□

Aus diesem Satz lässt sich sofort folgern, dass standardisierte lokale Bildmerkmale weitgehend invariant unter photometrischen Einflüssen sind. Leider nur *weitgehend*, denn eine 100%-ige Invarianz lässt sich nicht erzielen. Zum einen wird der durch t transformierte Grauwert auf die nächst kleinere ganze Zahl abgerundet, was wir hier vernachlässigt haben, und zum anderen müssen diese ganzen Zahlen in die Menge $\mathbb{Z}_{0,255}$ abgebildet werden, also dorthin „verschoben" werden, was dazu führt, das manche Werte „abgeschnitten" werden müssen, weil sie über diesen Bereich hinauslaufen. Dadurch wird das Ergebnis etwas verfälscht, sodass von keiner totalen Invarianz gesprochen werden kann. Dazu wollen wir abschließend ein Beispiel betrachten.

Beispiel 11.3.10. Wir untersuchen die beiden nachfolgend dargestellten 3×3-Lokalisierungsfenster Loc_1 und Loc_2.



Dabei gilt für die Grauwerte $\mathrm{gv}'_{i,j}$ des Lokalisierungsfensters Loc_2 (rechts):

$$\mathrm{gv}'_{i,j} = \min\{\mathrm{gv}_{i,j} + 50, 255\},$$

wobei $\mathrm{gv}_{i,j}$ die Grauwerte des linken Lokalisierungsfensters Loc_1 bezeichnen. Es handelt sich dabei um eine recht deutliche Aufhellung. Die standardisierten „Grauwerte" sind nachfolgend abgebildet.

−1.23	−0.83	−1.38
0.13	1.07	−0.49
0.66	1.71	0.35

−1.16	−0.90	−1.53
0.23	1.27	−0.51
0.84	1.27	0.48

Die lokalen Bildmerkmale für Loc_1 und Loc_2 lauten wie folgt:

$$\begin{aligned} v_1 &= \left(\tfrac{1297}{9},\ -\tfrac{121}{2},\ -\tfrac{32}{3},\ -\tfrac{125}{6},\ \tfrac{9}{4},\ -\tfrac{172}{3}\right) \\ &= (144.111,\ -60.5,\ -10.666,\ -20.833,\ 2.25,\ -57.333), \\ v_2 &= \left(\tfrac{1715}{9},\ -\tfrac{313}{6},\ -\tfrac{37}{3},\ -\tfrac{151}{6},\ -\tfrac{1}{4},\ -\tfrac{125}{3}\right) \\ &= (190.555,\ -52.166,\ -12.333,\ -25.166,\ -0.25,\ -41.666) \end{aligned}$$

Dies zeigt besonders die zu erwartende starke Abweichung in der ersten Komponente, da das arithmetische Mittel der Grauwerte von der Transformation stark beeinflusst werden. Die standardisierten lokalen Bildmerkmale

$$\begin{aligned} v_1^* &= (0,\ -1.027,\ -0.181,\ -0.353,\ 0.038,\ -0.974) \\ v_2^* &= (0,\ -1.031,\ -0.243,\ -0.497,\ -0.004,\ -0.824) \end{aligned}$$

stimmen erwartungsgemäß in der ersten Komponente überein und weichen auch sonst weniger voneinander ab, d.h. diese starke Aufhellung macht sich auf die standardisierten lokalen Bildmerkmale weniger stark bemerkbar. ◁

ANHANG A

Vektorinvarianten der speziellen orthogonalen Gruppe SO_n

Für den Beweis des Erzeugendensystems des Invariantenrings der Operation der speziellen orthogonalen Gruppe auf einer endlichen Menge von Punkten (siehe Abschnitt 6.4), wie in David Richman 1989 in [Ric89] präsentiert hatte, sind mehrere grundlegende Begriffe und Resultate von Nöten, die im Allgemeinen nicht als bekannt vorauszusetzen sein dürften. Da mit Ausnahme von Abschnitt 6.4 diese Begriffe keine Rolle spielen, werden wir sie nur hier im Anhang erwähnen und angeben.

Dazu seien im Folgenden stets $m, n \in \mathbb{N}$ mit $n \geq 2$ und $m \geq 1$. Sei K ein nicht-endlicher Körper mit Charakteristik $\mathrm{char}(K) \neq 2$ und sei $P = K[x_{i,j} : \ i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}]$ der Polynomring über K in den $m \cdot n$ Unbestimmten $x_{1,1}, \ldots, x_{m,n}$. Wir schreiben P auch kurz in der Form $K[x_{i,j}]$, wenn m und n aus dem Zusammenhang ersichtlich sind. Der Polynomring P sei in diesem Kapitel außerdem stets versehen mit der lexikographischen Termordnung Lex, d.h. es gilt:

$$x_{1,1} >_{\mathrm{Lex}} x_{1,2} >_{\mathrm{Lex}} \ldots >_{\mathrm{Lex}} x_{1,n} >_{\mathrm{Lex}} x_{2,1} >_{\mathrm{Lex}} \ldots >_{\mathrm{Lex}} x_{m,n}.$$

Die Unbestimmten lassen sich anhand ihrer Indizierung auf kanonische Weise in einer $m \times n$-Matrix $(x_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ darstellen, die wir kurz mit $\mathcal{X}$ bezeichnen. Die i-te Zeile von $\mathcal{X}$, d.h. das n-Tupel $(x_{i,1}, \ldots, x_{i,n})$, bezeichnen wir abkürzend nur mit x_i und bitten gleichzeitig zu beachten, dass x_i in diesem Kapitel *keine* Unbestimmte bezeichnet.

A.1 Grundlegende Begriffe

Um die entscheidenden Beweise angeben zu können, sind verschiedene Begriffe und Strukturen notwendig sowie Eigenschaften dieser. In diesem Abschnitt sollen diese kurz erläutert werden.

A.1.1 Spezielle Determinanten

In [Ric89] spielen Determinanten und Minoren der Matrix $\mathcal{X}$ oder aus $\mathcal{X}$ entstehender Matrizen eine große Rolle. Aus diesem Grund werden wir uns hier zunächst mit speziellen Determinanten

und entsprechenden Resultaten auseinandersetzen. Dazu werden wir zunächst den Begriff eines k-Minors festlegen (vgl. [Fis05], 3.3.6, S. 206 f.) und die Notation dabei zwar [Ric89] annähern, ansonsten aber neu festlegen.

Definition A.1.1. (k-reihige Teilmatrix, k-Minor)
Sei R ein kommutativer Ring mit Eins und sei $\mathcal{M} = (m_{i,j}) \in \mathrm{Mat}_{m,n}(R)$ mit $m, n \in \mathbb{N}_+$. Sei $k \in \mathbb{N}_+$ mit $k \leq \min\{m, n\}$ und seien $i_1, \ldots, i_k \in \{1, \ldots, m\}$ mit $i_1 < \ldots < i_k$ sowie $j_1, \ldots, j_k \in \{1, \ldots, n\}$ mit $j_1 < \ldots < j_k$. Dann heißt die quadratische $k \times k$-Matrix

$$\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k) := (m_{i,j})_{\substack{i_1 \leq i \leq i_k \} \\ j_1 \leq j \leq j_k \}}} \in \mathrm{Mat}_k(R)$$

eine **k-reihige Teilmatrix** von $\mathcal{M}$ und ihre Determinante ein **k-reihiger Minor** oder kurz ein **k-Minor**. Die Menge aller k-Minoren der Matrix $\mathcal{M}$ bezeichnen wir mit $\mathrm{Min}(k, \mathcal{M})$.

Die Matrix $\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k)$ ensteht somit aus $\mathcal{M}$ durch Streichen der Zeilen mit den Nummern $\{1, \ldots, m\} \setminus \{i_1, \ldots, i_k\}$ sowie der Spalten mit den Nummern $\{1, \ldots, n\} \setminus \{j_1, \ldots, j_k\}$. Als erste Eigenschaften von k-Minoren wollen wir folgenden Satz betrachten, der weiteren Verlauf eine zentrale Rolle spielt (vgl. [Fis05], 3.3.6, S. 207).

Satz A.1.2. *Sei $\mathcal{M} \in \mathrm{Mat}_{m,n}(R)$ und $r \in \mathbb{N}_+$ mit $r \leq \min\{m, n\}$. Dann sind folgende Aussagen äquivalent:*

(i) $r = \mathrm{Rang}(\mathcal{M})$,

(ii) Es gibt eine r-reihige Teilmatrix $\mathcal{M}'$ von $\mathcal{M}$ mit $\det(\mathcal{M}') \neq 0$, und für $k > r$ ist jeder k-Minor gleich Null.

Wir wollen zudem die Minoren der symmetrischen Matrix $\mathcal{M} := \mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}} \in \mathrm{Mat}_m(P)$ genauer betrachten, da auch sie von großer Bedeutung sein werden. Wie man sich leicht überlegt, lassen sich k-reihige Teilmatrizen von $\mathcal{M}$ auf folgende Weise darstellen: Ist $k \leq m$ und sind $i_1, \ldots, i_k, j_1, \ldots, j_k \in \{1, \ldots, m\}$ Indizes mit $i_1 < \ldots < i_k$ sowie $j_1 < \ldots < j_k$, so gilt

$$\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k) = \mathcal{A} \cdot \mathcal{B}^{\mathrm{tr}}$$

mit den Matrizen

$$\mathcal{A} := \begin{pmatrix} x_{i_1,1} & \ldots & x_{i_1,n} \\ \vdots & & \vdots \\ x_{i_k,1} & \ldots & x_{i_k,n} \end{pmatrix} \quad \text{und} \quad \mathcal{B} := \begin{pmatrix} x_{j_1,1} & \ldots & x_{j_1,n} \\ \vdots & & \vdots \\ x_{j_k,1} & \ldots & x_{j_k,n} \end{pmatrix}.$$

Da alle Einträge von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ homogene Polynome vom Grad 2 sind, ist jeder k-Minor von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ ein homogenes Polynom vom Grad $2k$. Für dieses Polynom gelten folgende Eigenschaften.

Satz A.1.3. *Sei $\mathcal{X}$ die $m \times n$-Matrix der Unbestimmten, $\mathcal{M} = \mathcal{X}\mathcal{X}^{\mathrm{tr}} \in \mathrm{Mat}_m(P)$, sei $k \in \mathbb{N}_+$ mit $k \leq m$ und seien $i_1, \ldots, i_k, j_1, \ldots, j_k \in \{1, \ldots, m\}$ mit $i_1 < \ldots < i_k$ sowie $j_1 < \ldots < j_k$.*

a) Ist $k > n$, so gilt $\det(\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k)) = 0$.

b) Ist $k \leq n$, so gilt $\mathrm{LT}_{\mathrm{Lex}}(\det(\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k))) = x_{i_1,1} \cdots x_{i_k,k} x_{j_1,1} \cdots x_{j_k,k}$

Beweis: a) Seien $\mathcal{A}, \mathcal{B} \in \mathrm{Mat}_{k,n}(P)$ diejenigen Matrizen mit $\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k) = \mathcal{A} \cdot \mathcal{B}^{\mathrm{tr}}$. Für eine bestimmte Menge $S \subseteq \{1, \ldots, n\}$ von Spaltennummern, sei $\mathcal{A}_S$ bzw. $\mathcal{B}_S$ die Teilmatrix von $\mathcal{A}$ bzw. $\mathcal{B}$, die entsteht, indem man genau die Spalten auswählt, die S

angibt. Für $S \subseteq \{1, \ldots, n\}$ mit $\#S = k$ ist $\mathcal{A}_S$ bzw. $\mathcal{B}_S$ also eine k-reihige Teilmatrix von $\mathcal{X}$. Dann folgt aus dem Satz von Binet-Cauchy (vgl. [Gan86], S. 27):

$$\det(\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k)) = \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ \#S = k}} \det(\mathcal{A}_S) \cdot \det(\mathcal{B}_S).$$

Für $k > n$ gibt es keine derartige Menge S. Somit folgt die Behauptung.

b) Sei $S = \{1, \ldots, k\}$ und $\mathcal{A}_S$ wie in a), d.h. $\mathcal{A}_S$ ist die k-reihige Teilmatrix von $\mathcal{X}$ der Form

$$\mathcal{A}_S = \begin{pmatrix} x_{i_1,1} & \ldots & x_{i_1,k} \\ \vdots & \vdots & \\ x_{i_k,1} & \ldots & x_{i_k,k} \end{pmatrix}$$

Dann gilt $\mathrm{LT}_{\mathrm{Lex}}(\det(\mathcal{A}_S)) = x_{i_1,1} \cdots x_{i_k,k}$ und für jede weitere Teilmenge $S' \subseteq \{1, \ldots, n\}$ mit $\#S' = k$ folgt $\mathrm{LT}_{\mathrm{Lex}}(\det(\mathcal{A}_S)) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(\det(\mathcal{A}_{S'}))$.

Setze $f := \det(\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k))$. Dann gilt erneut mit dem Satz von Binet-Cauchy:

$$\begin{aligned} \mathrm{LT}_{\mathrm{Lex}}(f) &= \mathrm{LT}_{\mathrm{Lex}}\left(\sum_{\substack{S \subseteq \{1,\ldots,n\} \\ \#S = k}} \det(\mathcal{A}_S) \cdot \det(\mathcal{B}_S) \right) \\ &= \mathrm{LT}_{\mathrm{Lex}}(\det(\mathcal{A}_{\{1,\ldots,k\}}) \cdot \det(\mathcal{B}_{\{1,\ldots,k\}})) = x_{i_1,1} \cdots x_{i_k,k} \cdot x_{j_1,1} \cdots x_{j_k,k} \end{aligned}$$

□

Aus dem Beweis dieses Satzes geht insbesondere hervor, dass sich für $k \leq \min\{m, n\}$ jeder k-Minor von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ als Produkt von k-Minoren von $\mathcal{X}$ schreiben lässt.

A.1.2 $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupel

Im folgenden Abschnitt werden wir uns mit sogenannten $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupeln beschäftigen. Sei dazu $k \in \mathbb{N}_+$ und zunächst $\mathcal{H} \in \mathrm{Mat}_{k,n}(P)$ eine beliebige $k \times n$-Matrix mit Einträgen aus P. Weiter seien $h_1, \ldots, h_k \in P^n$ die Zeilenvektoren der Matrix $\mathcal{H}$. Der zentrale Punkt wird nun sein, diese Vektoren zu ordnen. Dazu würden wir also eine Ordnung $\preceq$ auf P^n benötigen. Hätten wir eine Ordnung $\leq$ auf P gegeben, könnten wir zwei Vektoren $(f_1, \ldots, f_n), (g_1, \ldots, g_n) \in P^n$ beispielsweise lexikographisch ordnen, d.h. induktiv würde gelten:

$$(f_1, \ldots, f_n) \preceq (g_1, \ldots, g_n) \quad \Longleftrightarrow \quad f_1 \leq g_1 \text{ oder } (f_1 = g_1 \text{ und } (f_2, \ldots, f_n) \preceq (g_2, \ldots, g_n)).$$

Allerdings ist es etwas heikel, eine geeignete Ordnung $\leq$ auf P anzugeben. Auch Richman schweigt sich in seiner Arbeit [Ric89] darüber aus. Für uns stellt das aber kein großes Problem dar, da wir die Begriffe der $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupel in allgemeiner Form nicht benötigen werden. Für unsere Zwecke ist nur folgende Matrix $\mathcal{H}$ von Interesse:

$$\mathcal{H} = \begin{pmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,n} \\ x_{2,1} & x_{2,2} & \ldots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \ldots & x_{m,n} \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \in \mathrm{Mat}_{m+n,n}(P)$$

Deren Zeilenvektoren $h_1, \ldots, h_{m+n} \in P^n$ lassen sich durch

$$h_i \preceq h_j \quad \Longleftrightarrow \quad \mathrm{LT}_{\mathrm{PosLex}}(h_i) \leq_{\mathrm{PosLex}} \mathrm{LT}_{\mathrm{PosLex}}(h_j)$$

für alle $i, j \in \{1, \ldots, m+n\}$ total ordnen, d.h. die Zeilen der Matrix $\mathcal{H}$ sind von oben nach unten absteigend geordnet. Wir werden im Folgenden für die Matrix $\mathcal{H}$ abkürzend auch $\left(\begin{smallmatrix} \mathcal{X} \\ \mathcal{I}_n \end{smallmatrix}\right)$ schreiben, wobei $\mathcal{I}_n \in \mathrm{Mat}_n(P)$ die $n \times n$ Einheitsmatrix bezeichnet. Aufbauend auf dieser Matrix $\mathcal{H}$ wollen wir nun die Begriffe $\mathcal{H}$-Matrix und $\mathcal{H}$-Tupel definieren. Da die Anzahl der Spalten von $\mathcal{H}$-Matrizen fest vorgegeben sein wird, bezeichnen wir generell eine beliebige $l \times n$-Matrix $\mathcal{M} = (m_{i,j})$ und ein beliebiges Tupel $W = (w_1, \ldots, w_l)$ als **Matrix bzw. Tupel der Länge** l. Für $l = 0$ setzen wir $\mathcal{M} := \emptyset$ bzw. $W := \emptyset$ und bezeichnen $\mathcal{M}$ bzw. W in diesem Zusammenhang auch als **leere Matrix** bzw. **leeres Tupel**. Außerdem werden wir Tupel und Vektoren, deren Komponenten in einem bestimmten Sinne absteigend geordnet sind, als **monoton absteigend** bzw. **streng monoton absteigend** bezeichnen (vgl. [Ric89], S. 51).

Definition A.1.4. ($\mathcal{H}$-Matrix)
Für $l \in \mathbb{N}$ heißt eine $l \times n$-Matrix $\mathcal{M}$, deren Einträge *Zeilenvektoren* der Matrix $\mathcal{H}$ sind, eine **$\mathcal{H}$-Matrix** der Länge l. Mit $\mathrm{Mat}_{l,n}(\mathcal{H})$ wird die Menge aller $\mathcal{H}$-Matrizen der Länge l und mit $\mathrm{Mat}(\mathcal{H})$ die Menge aller $\mathcal{H}$-Matrizen bezeichnet.

Bislang sind die Anforderungen an eine $\mathcal{H}$-Matrix sehr gering. Wir werden nun den stärkeren Begriff der Standard $\mathcal{H}$-Matrix einführen. Würde man, wie es in [Ric89] ansatzweise gemacht wurde, generell beliebige Matrizen für die Matrix $\mathcal{H}$ zulassen, so wäre für diese Definition die Tatsache, dass die Zeilenvektoren von $\mathcal{H}$ paarweise verschieden sind und bzgl. $\preceq$ absteigend von oben nach unten geordnet sind, von entscheidender Bedeutung. Allerdings betrachten wir nur Matrizen $\mathcal{H}$ der Form $\left(\begin{smallmatrix} \mathcal{X} \\ \mathcal{I}_n \end{smallmatrix}\right)$ und diese erfüllen bereits diese Anforderungen (vgl. [Ric89], S. 52).

Definition A.1.5. (Standard $\mathcal{H}$-Matrix)
Sei $l \in \mathbb{N}$. Eine $\mathcal{H}$-Matrix $\mathcal{M} = (m_{i,j})_{\substack{1 \leq i \leq l \\ 1 \leq j \leq n}}$ heißt **standard** oder eine **Standard $\mathcal{H}$-Matrix** der Länge l, falls entweder $l = 0$ gilt oder für $l > 0$ die folgenden Eigenschaften erfüllt sind:

(i) Die Einträge jeder Zeile von $\mathcal{M}$ sind streng monoton absteigend geordnet, d.h. für alle $i \in \{1, \ldots, l\}$ gilt:

$$m_{i,1} \succ m_{i,2} \succ \ldots \succ m_{i,n}.$$

(ii) Die Einträge jeder Spalte von $\mathcal{M}$ sind monoton absteigend geordnet, d.h. es gilt

$$m_{1,j} \succeq m_{2,j} \succeq \ldots \succeq m_{l,j}$$

für alle $j \in \{1, \ldots, n\}$.

Gemäß dieser Definition ist also insbesondere die leere Matrix $\emptyset$ eine Standard $\mathcal{H}$-Matrix. Die Einträge $m_{i,j}$ einer $\mathcal{H}$-Matrix $\mathcal{M} = (m_{i,j})$ der Länge $l > 0$ sind also Zeilenvektoren der Matrix $\mathcal{H}$, d.h. es gilt $m_{i,j} \in \{h_1, \ldots, h_{m+n}\}$. Betrachten wir nun für $i \in \{1, \ldots, l\}$ die i-te Zeile $(m_{i,1}, \ldots, m_{i,n})$ von $\mathcal{M}$, so lässt sich das transponierte Zeilentupel $(m_{i,1}, \ldots, m_{i,n})^{\mathrm{tr}}$ als $n \times n$-Matrix mit Einträgen aus $\mathbb{T}^{mn} \cup \{0, 1\}$ auffassen. Diese Matrix werden wir für $i \in \{1, \ldots, l\}$ kurz mit $\mathcal{M}_i$ bezeichnen und deren Determinante mit $|\mathcal{M}_i|$. Wir schreiben manchmal auch $|m_{i,1}\ m_{i,2}\ \ldots\ m_{i,n}|$ anstatt $|\mathcal{M}_i|$. Offensichtlich ist $|\mathcal{M}_i|$ für alle $i \in \{1, \ldots, l\}$ ein Polynom aus P. Von besonderer Bedeutung ist das Produkt der l Zeilendeterminanten, die wir aus einer $\mathcal{H}$-Matrix $\mathcal{M}$ der Länge l erhalten (vgl. [Ric89], S. 52).

Definition A.1.6. (Wert einer $\mathcal{H}$-Matrix)
Sei $\mathcal{M}$ eine $\mathcal{H}$-Matrix der Länge $l \in \mathbb{N}$. Das Produkt $|\mathcal{M}_1| \cdots |\mathcal{M}_l|$ der l Zeilendeterminanten von $\mathcal{M}$ heißt der **Wert** der $\mathcal{H}$-Matrix $\mathcal{M}$ und wird mit $|\mathcal{M}|$ bezeichnet. Der Wert einer Standard $\mathcal{H}$-Matrix $\mathcal{M}$ heißt auch **Standard Produkt**.

Die Werte von $\mathcal{H}$-Matrizen sind also ganz spezielle Polynome in P. Dabei kann ein bestimmtes Polynom $f \in P$ der Wert von vielen verschiedenen $\mathcal{H}$-Matrizen sein. So gilt beispielsweise $|\emptyset| = 1$, aber auch $|\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}| = 1$. Außerdem ist sofort aus der Definition eines Wertes einer $\mathcal{H}$-Matrix $\mathcal{M}$ klar, dass $|\mathcal{M}| = 0$ gilt, falls eine Zeile von $\mathcal{M}$ zwei oder mehr identische Einträge besitzt. Da die Zeilenvektoren von $\mathcal{H}$ paarweise verschieden sind, ist dies auch die einzige Möglichkeit, das Nullpolynom als Wert einer $\mathcal{H}$-Matrix darzustellen, d.h. es gilt auch die Umkehrung. Stellt man weiter an eine $\mathcal{H}$-Matrix $\mathcal{M}$ die Anforderung, dass jede Zeile von $\mathcal{M}$ paarweise verschiedene Einträge, also paarweise verschiedene Zeilen von $\mathcal{H}$, enthält, so folgt sofort $|\mathcal{M}| \neq 0$. Außerdem ist jeder der Faktoren $|\mathcal{M}_i|$ in $|\mathcal{M}|$ dann ein n-Minor der Matrix $\mathcal{H}$ und folglich $|\mathcal{M}|$ ein Produkt von n-Minoren von $\mathcal{H}$.

Wir wollen nun noch den Spezialfall $n = 2$ näher betrachten, da dieser später von zentraler Bedeutung sein wird. Sei also im Folgenden stets $n = 2$, d.h. wir betrachten den Polynomring $P = K[x_{1,1}, x_{1,2}, \ldots, x_{m,1}, x_{m,2}]$ und $\mathcal{H} = \left(\begin{smallmatrix}\mathcal{X}\\ \mathcal{I}_2\end{smallmatrix}\right)$. Wie üblich bezeichne $\{e_1, e_2\}$ mit $e_1 := (1,0)$ und $e_2 := (0,1)$ die kanonische Basis des freien Moduls P^2 und wie in [Ric89] bezeichnen wir mit $e_1(\mathcal{M})$ bzw. $e_2(\mathcal{M})$ die Anzahl der e_1- bzw. e_2-Zeilen in einer $\mathcal{H}$-Matrix $\mathcal{M}$. Eine Standard $\mathcal{H}$-Matrix $\mathcal{M}$ mit $e_1(\mathcal{M}) = 0$ kann wegen der Ordnung auf den Zeilen und Spalten der Matrix $\mathcal{M}$ mögliche e_2-Zeilenvektoren höchstens in der zweiten Spalte von $\mathcal{M}$ als Eintrag enthalten. Wird dies berücksichtigt, so ist es möglich, in dieser Situation gewisse Information über $\mathcal{M}$ allein aus dem Leitterm $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ abzulesen. Dies hat Richman in [Ric89] im Beweis zu Proposition 6 nur kurz erwähnt und soll hier ausführlicher behandelt werden.

Lemma A.1.7. *Sei $\mathcal{M} \in \mathrm{Mat}_{l,n}(\mathcal{H})$ eine Standard $\mathcal{H}$-Matrix der Länge $l > 0$ mit $e_1(\mathcal{M}) = 0$. Dann entspricht für alle $i \in \{1, \ldots, m\}$ und alle $j \in \{1, 2\}$ der Exponent der Unbestimmten $x_{i,j}$ im Leitterm $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ genau der Anzahl an Zeilen der Form $(x_{i,1}, x_{i,2})$ in der j-ten Spalte von $\mathcal{M}$.*

Beweis: Seien $i, k \in \{1, \ldots, m\}$. Dann gilt zunächst:

$$\begin{aligned} |x_i\ x_k| &= \det\begin{pmatrix} x_{i,1} & x_{i,2} \\ x_{k,1} & x_{k,2} \end{pmatrix} = x_{i,1}x_{k,2} - x_{i,2}x_{k,1} \\ |x_i\ (0,\ 1)| &= \det\begin{pmatrix} x_{i,1} & x_{i,2} \\ 0 & 1 \end{pmatrix} = x_{i,1} \end{aligned}$$

Wegen $e_1(\mathcal{M}) = 0$ und wegen der Eigenschaft, dass eventuelle e_2-Zeilen nur in der zweiten Spalte von $\mathcal{M}$ vorkommen können, ist der Wert von $\mathcal{M}$ ein Produkt von Polynomen dieser Art. Aus der Ordnung der Zeilen der Matrix $\mathcal{H}$ folgt:

$$\mathrm{LT}_{\mathrm{Lex}}(|x_i\ x_k|) = x_{i,1}x_{k,2} \qquad \mathrm{LT}_{\mathrm{Lex}}(|x_i\ (0,\ 1)|) = x_{i,1},$$

d.h. $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ ist ein Produkt von Termen dieser Art. Somit korrespondiert die Anzahl der Unbestimmten $x_{i,1}$ in $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ eindeutig mit der Anzahl der Zeile x_i in der ersten Spalte von $\mathcal{M}$. Analog korrespondiert die Anzahl der Unbestimmten $x_{k,2}$ in $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ eindeutig mit der Anzahl der Zeile x_k in der zweiten Spalte von $\mathcal{M}$. Diese Anzahlen sind an den Exponenten der Unbestimmten im Leitterm ablesbar. □

Daraus ergibt sich sofort folgende Formel für den Grad des Polynoms $|\mathcal{M}|$ einer Standard $\mathcal{H}$-Matrix, die keine e_1-Einträge enthält. In diesem Fall gilt:

$$\deg(|\mathcal{M}|) = 2l - e_2(\mathcal{M}). \tag{A.1.1}$$

Wir wollen hier außerdem noch anmerken, dass $\mathcal{M} = \emptyset$ die einzige Standard $\mathcal{H}$-Matrix ist mit $e_1(\mathcal{M}) = 0$ und $|\mathcal{M}| = 1$. Denn angenommen, es gäbe eine Standard $\mathcal{H}$-Matrix $\mathcal{M} \neq \emptyset$ mit $e_1(\mathcal{M}) = 0$ und $|\mathcal{M}| = 1$. Dann würde $\mathcal{M}$ mindestens eine Zeile der Form $(x,\ y)$ oder $(x,\ e_2)$ enthalten im Widerspruch zu $|\mathcal{M}| = 1$. Mit Hilfe dieser Hilfsaussagen lässt sich der folgende Satz leicht beweisen (vgl. [Ric89], Prop. 6, S. 53).

Satz A.1.8. *Sei $\mathcal{M} \in \mathrm{Mat}_{l,n}(\mathcal{H})$ eine Standard $\mathcal{H}$-Matrix der Länge $l > 0$ mit $e_1(\mathcal{M}) = 0$. Dann ist die Matrix $\mathcal{M}$ eindeutig durch den Leitterm von $|\mathcal{M}|$ bestimmt.*

Beweis: Sei $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) = x_{1,1}^{\alpha_{1,1}} x_{1,2}^{\alpha_{1,2}} \cdots x_{m,1}^{\alpha_{m,1}} x_{m,2}^{\alpha_{m,2}}$ mit $\alpha_{i,1}, \alpha_{i,2} \in \mathbb{N}$ für alle $i \in \{1, \ldots, m\}$ der Leitterm von $|\mathcal{M}|$. Wegen $e_1(\mathcal{M}) = 0$ gilt $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) \neq 1$. Laut Lemma A.1.7 gibt für jedes $i \in \{1, \ldots, m\}$ der Exponent $\alpha_{i,1}$ an, wie oft der Zeilenvektor x_i in der ersten Spalte von $\mathcal{M}$ vorkommt, und der Exponent $\alpha_{i,2}$, wie oft die Zeile x_i in der zweiten Spalte von $\mathcal{M}$ vorkommt. Zudem lässt sich die Anzahl der Zeilenvektoren e_2 in der zweiten Spalte von $\mathcal{M}$ mit Hilfe von Gleichung (A.1.1) berechnen.

Da $\mathcal{M}$ eine Standard $\mathcal{H}$-Matrix ist, enthält die erste Spalte von $\mathcal{M}$ nur Zeilenvektoren aus $\mathcal{X}$, die bzgl. $\preceq$ absteigend von oben nach unten geordnet sind. Damit ist die erste Spalte von $\mathcal{M}$ eindeutig festgelegt. Analog ist die zweite Spalte von $\mathcal{M}$ und damit die Matrix $\mathcal{M}$ selbst eindeutig bestimmt. □

Eine unmittelbare Folgerung aus diesem Satz ist in dem abschließenden Korollar festgehalten.

Korollar A.1.9. *Zwei Standard $\mathcal{H}$-Matrizen $\mathcal{M}$ und $\mathcal{L}$ derselben Länge mit $e_1(\mathcal{M}) = e_1(\mathcal{L}) = 0$ sind genau dann identisch, wenn die Leitterme ihrer Werte übereinstimmen.*

Nun werden wir in Analogie zu $\mathcal{H}$-Matrizen den Begriff des $\mathcal{H}$-Tupels angeben. Dabei handelt es sich um Tupel gerader Länge, die als Komponenten ebenfalls Zeilenvektoren der Matrix $\mathcal{H}$ beinhalten. Diese Tupel verwendet Richman zwar in seiner Arbeit [Ric89], benennt diese jedoch nicht und geht auch nur äußerst kurz auf Eigenschaften solcher Tupel ein. Ohne ausufern zu wollen, soll dieser Struktur dennoch etwas mehr Bedeutung eingeräumt werden, da sie eine wesentliche Komponente eines SO_2-Tableaus bildet, wie wir im nächsten Abschnitt sehen werden. Im Zuge dessen werden wir an dieser Stelle den Begriff eines $\mathcal{H}$-Tupels einführen und definieren.

Definition A.1.10. ($\mathcal{H}$-Tupel)
Ein Tupel W der Länge $2u$ mit $u \in \mathbb{N}$, dessen Einträge Zeilenvektoren der Matrix $\mathcal{H}$ sind, heißt ein **$\mathcal{H}$-Tupel** (der Länge $2u$).

Dabei ist für $u = 0$ das Tupel W das leere Tupel. Ist $W = (w_1, \ldots, w_{2u})$ für $u \in \mathbb{N}_+$ ein nichtleeres $\mathcal{H}$-Tupel, so ist es möglich, die Komponenten des Tupels paarweise zu gruppieren. Da jede Komponente von W ein Zeilenvektor von $\mathcal{H}$ ist, können wir diesen Zeilenvektor auch als einzeilige Matrix betrachten. Somit ist es möglich, für alle $i \in \{1, \ldots, u\}$ durch Anwendung des gewöhnlichen Produkts zweier Matrizen Polynome $w_{2i-1} \cdot w_{2i}^{\mathrm{tr}} \in P$ zu bilden. Wir bezeichnen ein Produkt $w_{2i-1} \cdot w_{2i}^{\mathrm{tr}}$ auch mit $\langle w_{2i-1}, w_{2i} \rangle$. Das Produkt über alle diese Polynome von W bezeichnen wir als Wert von W.

Definition A.1.11. (Wert eines $\mathcal{H}$-Tupels)
Sei $u \in \mathbb{N}$ und $W = (w_1, \ldots, w_{2u})$ ein $\mathcal{H}$-Tupel der Länge $2u$. Das Produkt $\prod_{i=1}^{u} \langle w_{2i-1}, w_{2i} \rangle$ heißt der **Wert** des $\mathcal{H}$-Tupels W. Dieser wird mit $\langle W \rangle$ bezeichnet.

Wir schreiben auch $\langle w_1, \ldots, w_{2u} \rangle$ anstatt $\langle W \rangle$ bzw. $\langle (w_1, \ldots, w_{2u}) \rangle$. Man beachte auch, dass $\mathcal{H}$-Tupel der Länge 0 damit per Definition Wert 1 haben, d.h. im Speziellen gilt also $\langle \emptyset \rangle := 1$. Außerdem ist der Wert eines $\mathcal{H}$-Tupels per Definition ein normiertes Polynom. Analog zu den $\mathcal{H}$-Matrizen lassen sich bestimmte Polynome auch hier als Wert verschiedener $\mathcal{H}$-Tupel schreiben. So gilt beispielsweise $\langle (e_1, e_2) \rangle = 0$, aber natürlich auch $\langle (e_1, e_2, e_1, e_2) \rangle = 0$, oder $\langle \emptyset \rangle = 1$ und $\langle (e_1, e_1) \rangle = 1$. Wir wollen nun auch hier bestimmte $\mathcal{H}$-Tupel W betrachten, die uns zum einen garantieren, dass stets $\langle W \rangle \neq 0$ gilt, und es uns zum anderen ermöglichen, analog zu $\mathcal{H}$-Matrizen ein $\mathcal{H}$-Tupel allein aus dessen Leitterm zu rekonstruieren.

Definition A.1.12. (Standard $\mathcal{H}$-Tupel)
Sei $u \in \mathbb{N}$. Ein $\mathcal{H}$-Tupel $W = (w_1, \ldots, w_{2u})$ heißt **Standard $\mathcal{H}$-Tupel** der Länge $2u$, falls entweder $u = 0$ gilt oder für $u > 0$ die folgenden Eigenschaften erfüllt sind:

(i) Alle Komponenten von W sind Zeilenvektoren der Matrix $\mathcal{X}$.

(ii) Die Komponenten W sind monoton absteigend geordnet, d.h. es gilt $w_1 \succeq w_2 \succeq \ldots \succeq w_{2u}$.

Dadurch dass Standard $\mathcal{H}$-Tupel W der Länge $2u$ nur Zeilenvektoren aus $\mathcal{X}$ als Einträge beinhalten, gilt nun stets $\langle W \rangle \neq 0$. Genauer ist der Wert $\langle W \rangle$ ein homogenes Polynom vom Grad $2u$ und, falls $u > 0$ gilt, ist jeder der Faktoren in $\langle W \rangle$ ein homogenes Polynom vom Grad 2. Wir erhalten als Werte von Standard $\mathcal{H}$-Matrizen nur spezielle homogene Polynome. Wie man leicht sieht, ist der Wert eines Standard $\mathcal{H}$-Tupels ein Produkt von Einträgen der symmetrischen $m \times m$-Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}} = (x_i \cdot x_j^{\mathrm{tr}})_{1 \leq i,j \leq m}$. Diese Polynome werden im weiteren Verlauf noch von zentraler Bedeutung sein. Ein Standard $\mathcal{H}$-Tupel W der Länge $2u > 0$ hat somit die Form $W = (x_{c(1)}, x_{c(2)}, \ldots, x_{c(2u)})$ mit $c(1), \ldots, c_{2u} \in \{1, \ldots, m\}$ und $c(1) \leq \ldots \leq c_{2u}$. Damit ist sofort klar, wie der Leitterm des Wertes von W aussieht:

$$\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle) = x_{c(1),1} x_{c(2),1} \cdots x_{c(2u),1}. \tag{A.1.2}$$

Außerdem ist nun sofort klar, dass $W = \emptyset$ das einzige Standard $\mathcal{H}$-Tupel ist mit $\langle W \rangle = 1$. Aufbauend auf der Form des Leitterms lässt sich folgender Satz beweisen.

Satz A.1.13. *Sei W ein Standard $\mathcal{H}$-Tupel der Länge $2u$ mit $u \in \mathbb{N}$. Dann ist W eindeutig durch $\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle)$ bestimmt.*

Beweis: Aus Gleichung (A.1.2) folgt zunächst, dass der Leitterm von $\langle W \rangle$ von der Form

$$\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle) = x_{1,1}^{\alpha_{1,1}} x_{2,1}^{\alpha_{2,1}} \cdots x_{m,1}^{\alpha_{m,1}}$$

mit $\alpha_{1,1}, \ldots, \alpha_{m,1} \in \mathbb{N}$ ist. Gilt $\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle) = 1$, also $\langle W \rangle = 1$, so folgt $W = \emptyset$.

Sei nun also $\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle) \neq 1$. Dann folgt aus Gleichung (A.1.2), dass für alle $i = 1, \ldots, m$ der Exponent $\alpha_{i,1}$ genau angibt, wie oft die i-te Zeile von $\mathcal{X}$ als Komponente in W vorkommt. Durch die Ordnung auf W ist W eindeutig bestimmt. □

Eine unmittelbare Folgerung aus diesem Satz stellt das folgende Korollar dar.

Korollar A.1.14. *Zwei Standard $\mathcal{H}$-Tupel W und V sind genau dann identisch, wenn die Leitterme $\mathrm{LT}_{\mathrm{Lex}}(\langle W \rangle)$ und $\mathrm{LT}_{\mathrm{Lex}}(\langle V \rangle)$ übereinstimmen.*

Zum Abschluss dieses Abschnitts werden wir einige Rechenregeln für $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupel im Fall $n = 2$ angeben, die sich alle unmittelbar nachrechnen lassen, worauf wir hier verzichten wollen. Zum Teil werden diese Aussagen auch einen Zusammenhang herstellen zwischen $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupeln. Für den Rest dieses Abschnitts sei deshalb stets $n = 2$.

Lemma A.1.15. *Seien a, b, c, d beliebige Zeilenvektoren von $\mathcal{X}$. Dann gilt:*

(i) $\begin{vmatrix} a & e_1 \\ b & e_1 \end{vmatrix} = \langle (a, b) \rangle - \begin{vmatrix} a & e_2 \\ b & e_2 \end{vmatrix}$.

(ii) $|c\ e_2|\langle (a, b) \rangle = \begin{vmatrix} a & c \\ b & e_1 \end{vmatrix} + |a\ e_2|\langle (b, c) \rangle$

(iii) $|c\ e_1|\langle (a, b) \rangle = -\begin{vmatrix} a & c \\ b & e_2 \end{vmatrix} + |a\ e_1|\langle (b, c) \rangle$

(iv) $\begin{vmatrix} a & e_1 \\ d & e_2 \end{vmatrix} \langle (b, c) \rangle = -\begin{vmatrix} a & e_2 \\ b & d \\ c & e_2 \end{vmatrix} + |b\ d|\langle (a, c) \rangle + \begin{vmatrix} a & e_1 \\ b & e_2 \end{vmatrix} \langle (c, d) \rangle$

Kommen in einer $\mathcal{H}$-Matrix $\mathcal{M}$ zwei (oder mehr) e_1-Einträge in der zweiten Spalte vor, so können wir also die zugehörigen zwei Faktoren im Wert von $\mathcal{M}$ durch eine Differenz von Werten ersetzen, in denen die e_1-Einträge in gewissem Sinne eliminiert wurden. Unter Anwendung dieser Formeln ergeben sich weitere Zusammenhänge, die sich ebenfalls leicht nachrechnen lassen.

Lemma A.1.16. *Seien a, b, c, d beliebige Zeilenvektoren von $\mathcal{H}$. Dann gilt:*

(i) $\begin{vmatrix} a & d \\ b & c \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} - \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

(ii) $\langle (a, c, b, d) \rangle = \langle (a, b, c, d) \rangle - \begin{vmatrix} a & d \\ b & c \end{vmatrix}$

(iii) $|c\ d|\langle (a, b) \rangle = -|a\ c|\langle (b, d) \rangle + |a\ d|\langle (b, c) \rangle$

A.1.3 SO_2-Tableaus

Wir werden in diesem Abschnitt $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupel zusammen als ein Paar betrachten, das SO_2-Tableau genannt wird. Erfüllen SO_2-Tableaus bestimmte zusätzliche Eigenschaften, spricht man von sogenannten Standard SO_2-Tableaus. Diese bilden den Schlüssel für den Beweis eines Erzeugendensystem des Invariantenrings P^{SO_2} (siehe Abschnitt A.2). Wir werden zudem sehen, dass man mit SO_2- bzw. Standard SO_2-Tableaus im Grunde nur spezielle Erzeugendensysteme des Polynomrings P erhält, d.h. wir werden insbesondere zeigen, dass sich jedes Polynom in P durch Standard SO_2-Tableaus darstellen lässt.

Definition und grundlegende Eigenschaften von SO_2-Tableaus

Sei im Folgenden nun stets $n = 2$. Weiter sei $m \in \mathbb{N}_+$ und $\mathcal{H} = \binom{\mathcal{X}}{\mathcal{I}_2}$. Dann lassen sich SO_2-Tableaus wie folgt definieren (vgl. [Ric89], S. 58).

Definition A.1.17. (SO_2-Tableau)
Sei $\mathcal{M} \in \mathrm{Mat}_{l,2}(\mathcal{H})$ eine $\mathcal{H}$-Matrix der Länge $l \in \mathbb{N}$ und W ein $\mathcal{H}$-Tupel der Länge $2u$ mit $u \in \mathbb{N}$, dessen Einträge Zeilen von $\mathcal{X}$ sind. Dann heißt das Paar $(\mathcal{M}, W)$ ein **SO_2-Tableau**. Die Menge aller SO_2-Tableaus bezeichnen wir mit Tab.

Um die Paare $(\mathcal{M}, W)$ besser als ein Objekt betrachten zu können, werden wir diese auch in einer alternativen Form darstellen. Wir lesen die Matrix $\mathcal{M}$ dazu zeilenweise aus und schreiben die Einträge der Reihe nach in ein Tupel. Ist also $\mathcal{M}$ eine $\mathcal{H}$-Matrix der Länge l, so erhalten wir auf diese Weise ein Tupel der Länge $2l$. Anschließend konkatenieren wir dieses Tupel und das Tupel W der Länge $2u$. Somit lässt sich das Paar $(\mathcal{M}, W)$ als Tupel der Länge $2l + 2u$, deren Einträge Zeilenvektoren der Matrix $\mathcal{H}$ sind, darstellen. Dieses Tupel bezeichnen wir mit $\gamma_{(\mathcal{M},W)}$. RICHMAN erwähnt in seiner Arbeit diese Tupel nur in einem Beweis. Wir werden deren Eigenschaften und Bedeutung später noch etwas ausführlicher beleuchten. Dieser Prozess der Darstellung eines SO_2-Tableaus als ein Tupel ist jedoch nicht eindeutig umkehrbar, d.h. im Allgemeinen gibt es zu einem Tupel γ viele verschiedene SO_2-Tableaus $(\mathcal{M}, W)$ mit $\gamma = \gamma_{(\mathcal{M},W)}$. Somit ist eine Rekonstruktion von $\mathcal{M}$ und W aus $\gamma_{(\mathcal{M},W)}$ im Allgemeinen nicht möglich. Haben zwei Tupel $\gamma_{(\mathcal{M},W)}$ und $\gamma_{(\mathcal{L},V)}$ dieselbe Länge, so induziert die Ordnung $\preceq$ auf den Zeilen von $\mathcal{H}$ eine lexikographische Ordnung auf diesen Tupeln, die wir mit $\preceq_{\text{Lex}}$ bezeichnen. Genauer gilt also für zwei Tupel $\gamma_{(\mathcal{M},W)} = (a_1, \ldots, a_k)$ und $\gamma_{(\mathcal{L},V)} = (b_1, \ldots, b_k)$ derselben Länge k:

$$\gamma_{(\mathcal{M},W)} \preceq_{\text{Lex}} \gamma_{(\mathcal{L},V)} \iff a_1 \prec b_1 \text{ oder } (a_1 = b_1 \text{ und } (a_2, \ldots, a_k) \preceq_{\text{Lex}} (b_2, \ldots, b_k)).$$

Wir werden nun analog zu $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupeln auch den SO_2-Tableaus einen Wert zuweisen. RICHMAN verwendet dazu das Produkt der Werte der $\mathcal{H}$-Matrix und des $\mathcal{H}$-Tupels (vgl. [Ric89], S. 58).

Definition A.1.18. (Wert eines SO_2-Tableaus)
Sei $(\mathcal{M}, W)$ ein SO_2-Tableau. Das Polynom $|\mathcal{M}|\langle W\rangle$ heißt der **Wert** des SO_2-Tableaus $(\mathcal{M}, W)$. Ist $T \subseteq \text{Tab}$ eine Menge von SO_2-Tableaus, so bezeichnen wir die Menge der Werte der SO_2-Tableaus aus T mit $\text{Val}(T)$, d.h. es gilt $\text{Val}(T) = \{|\mathcal{M}|\langle W\rangle : \ (\mathcal{M}, W) \in T\}$.

Die Menge $\text{Val}(\text{Tab})$ der Werte aller SO_2-Tableaus ist eine Menge normierter und von Null verschiedener Polynome. Leicht einzusehen ist natürlich, dass es verschiedenste Möglichkeiten gibt, Paare aus $\mathcal{H}$-Matrizen und $\mathcal{H}$-Tupeln zu bilden, d.h. es gibt zu einem bestimmten Polynom $f \in P$ unter Umständen viele verschiedene SO_2-Tableaus, die den Wert f besitzen. Gilt beispielsweise $f = |\mathcal{M}|\langle W\rangle$ für ein SO_2-Tableau $(\mathcal{M}, W)$, so lassen sich durch Hinzufügen von $(e_1,\ e_2)$-Zeilen zu $\mathcal{M}$ beliebige neue SO_2-Tableaus erzeugen, die allesamt denselben Wert haben. Es drängt sich an dieser Stelle natürlich zwangsläufig die Frage auf, ob denn jedes Polynom $f \in P$ der Wert eines SO_2-Tableaus ist. Darauf werden wir später eingehen.

Zunächst wollen wir aber auch den Begriff des SO_2-Tableaus standardisieren. Wie man vermuten könnte, geschieht dies dadurch, dass wir für das Paar $(\mathcal{M}, W)$ einfach Standard $\mathcal{H}$-Matrizen und Standard $\mathcal{H}$-Tupel verwenden. Dies werden wir auch tun, allerdings mit zusätzlichen Anforderungen (vgl. [Ric89], S. 58).

Definition A.1.19. (Standard SO_2-Tableau)
Ein SO_2-Tableau $(\mathcal{M}, W)$ heißt **standard** oder ein **Standard SO_2-Tableau**, falls gilt:

(i) $\mathcal{M}$ ist eine Standard $\mathcal{H}$-Matrix mit $e_1(\mathcal{M}) \leq 1$ und jeder Eintrag der ersten Spalte von $\mathcal{M}$ ist ein Zeilenvektor von $\mathcal{X}$.

(ii) W ist ein Standard $\mathcal{H}$-Tupel und jede Komponente von W ist bzgl. $\preceq$ kleiner oder gleich jedem Eintrag in der ersten Spalte von $\mathcal{M}$.

Wir bezeichnen die Menge aller Standard SO_2-Tableaus mit STab.

Wegen $\text{STab} \subseteq \text{Tab}$ ist $\text{Val}(\text{STab})$ die Menge der Werte aller Standard SO_2-Tableaus. Da die $\mathcal{H}$-Matrix $\mathcal{M}$ eines Standard SO_2-Tableaus $(\mathcal{M}, W)$ standard ist, sind insbesondere die Einträge in

jeder Zeile von $\mathcal{M}$ verschieden. Somit folgt sofort $|\mathcal{M}|\langle W\rangle \neq 0$ für alle Standard SO_2-Tableaus $(\mathcal{M}, W)$. Außerdem werden uns die Tatsachen, dass $\mathcal{M}$ in der ersten Spalte nur Zeilenvektoren der Matrix $\mathcal{X}$ enthält und dass jeder Eintrag von W kleiner oder gleich jedem Eintrag in der ersten Spalte von $\mathcal{M}$ ist, später dabei helfen, in bestimmten Situationen das Standard SO_2-Tableau $(\mathcal{M}, W)$ allein aus dem Leitterm von $|\mathcal{M}|\langle W\rangle$ zu rekonstruieren. Zunächst wollen wir aber überlegen, was wir über die Struktur einer Standard $\mathcal{H}$-Matrix eines Standard SO_2-Tableaus aussagen können. Die Struktur ist aufgrund der Ordnung auf den Zeilen und Spalten der Matrix bereits per Definition weitgehend festgelegt. Sind x, y Zeilenvektoren von $\mathcal{X}$ mit $x \succ y$, dann enthält $\mathcal{M}$ nur Zeilen der Form $(x,\ y)$, $(x,\ e_1)$, $(x,\ e_2)$ oder $(e_1,\ e_2)$. Durch die zusätzlichen Einschränkungen aus der Definition eines Standard SO_2-Tableaus lässt sich eine $\mathcal{H}$-Matrix eines Standard SO_2-Tableaus noch präziser angeben. Zeilen der Form $(e_1,\ e_2)$ sind somit nicht mehr möglich. Außerdem enthält $\mathcal{M}$ den Zeilenvektor e_1 maximal ein Mal. Die Ordnung auf den Zeilen und Spalten von $\mathcal{M}$ legt die Struktur letztendlich im Sinne der folgenden Bemerkung fest.

Bemerkung A.1.20. (Struktur von Standard SO_2-Tableaus)

a) Eine $\mathcal{H}$-Matrix $\mathcal{M}$ ist genau dann Teil eines Standard SO_2-Tableaus, wenn es ein $s \in \mathbb{N}$ gibt so, dass $\mathcal{M}$ die folgende Form hat:

$$\begin{pmatrix} x_{a(1)} & x_{b(1)} \\ \vdots & \vdots \\ x_{a(s)} & x_{b(s)} \\ x_{a(s+e_1(\mathcal{M}))} & e_1 \\ x_{a(s+e_1(\mathcal{M})+1)} & e_2 \\ \vdots & \vdots \\ x_{a(s+e_1(\mathcal{M})+e_2(\mathcal{M}))} & e_2 \end{pmatrix}.$$

Entsprechend der Definition eines Standard SO_2-Tableaus erfüllt $\mathcal{M}$ die nachfolgenden Eigenschaften:

(i) Für alle $i \in \{1, \ldots, s\}$ gilt $x_{a(i)} \succ x_{b(i)}$.

(ii) $x_{a(1)} \succeq \ldots \succeq x_{a(s)} \succeq \ldots \succeq x_{a(s+e_1(\mathcal{M})+e_2(\mathcal{M}))}$.

(iii) $x_{b(1)} \succeq \ldots \succeq x_{b(s)}$.

b) Sei $W = (w_1, \ldots, w_{2u})$ mit $w_1 \succeq w_2 \succeq \ldots \succeq w_{2u}$ für $u \in \mathbb{N}_+$ ein Standard $\mathcal{H}$-Tupel und $\mathcal{M}$ von der Form aus a). Dann ist W aufgrund der Ordnung auf der ersten Spalte von $\mathcal{M}$ genau dann Teil eines Standard SO_2-Tableaus, wenn $\mathcal{M}$ entweder leer ist oder gilt:

$$x_{a(s+e_1(\mathcal{M})+e_2(\mathcal{M}))} \succeq w_1.$$

Anders als Standard $\mathcal{H}$-Matrizen oder Standard $\mathcal{H}$-Tupel lassen sich Standard SO_2-Tableaus ohne zusätzliches Wissen im Allgemeinen nicht eindeutig aus dem Leitterm ihres Wertes rekonstruieren. Dennoch können wir gewisse Informationen aus dem Leitterm ablesen.

Lemma A.1.21. *Sei $(\mathcal{M}, W)$ ein Standard SO_2-Tableau mit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) \neq 1$ und seien $t_1 \in \mathbb{T}(x_{1,1}, \ldots, x_{m,1})$ sowie $t_2 \in \mathbb{T}(x_{1,2}, \ldots, x_{m,2})$ Terme mit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = t_1 \cdot t_2$. Dann gibt es einen eindeutig bestimmten Term $t \in \mathbb{T}^{mn}$ mit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) = t \cdot t_2$ und $t_1 = t \cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$.*

Beweis: Laut Gleichung (A.1.2) ist der Leitterm von $\langle W\rangle$ von der Form $\mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle) = x_{c(1),1}\cdots x_{c(2u),1}$ mit $c(1),\ldots,c(2u) \in \{1,\ldots,m\}$. Somit ist $\mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$ ein Teiler von t_1, d.h. es gibt einen eindeutig bestimmten Term $t \in \mathbb{T}^{n\cdot m}$ mit $t_1 = t \cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$.

Wegen $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)\cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle) = t_1\cdot t_2$ folgt $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)\cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle) = t\cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)\cdot t_2$ und somit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) = t\cdot t_2$. □

Somit lässt sich aus dem Leitterm $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = t_1\cdot t_2$ ein Teil der Struktur von $\mathcal{M}$ ablesen. Da der Term t_2 nur $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$, aber nicht $\mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$ teilt, lässt sich gemäß Lemma A.1.7 die zweite Spalte von $\mathcal{M}$ bis auf eventuelle e_1- und e_2-Zeilen eindeutig bestimmen. Der eindeutig bestimmte Term $t \in \mathbb{T}^{m\cdot n}$ aus Lemma A.1.21 lässt sich aber im Allgemeinen allein aus der Kenntnis des Leitterms $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle)$ nicht berechnen. Möglich wird dies allerdings, wenn man $\mathcal{H}$-Matrizen betrachtet, die keine e_1-Zeilen enthalten. Bevor wir auf die Problematik eines e_1-Eintrags eingehen, zeigen wir, wie sich der Term t aus $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle)$ im Falle von $e_1(\mathcal{M}) = 0$ berechnen lässt.

Korollar A.1.22. *Sei $(\mathcal{M}, W)$ ein Standard SO_2-Tableau mit $e_1(\mathcal{M}) = 0$ und einer Zerlegung $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = t_1\cdot t_2$ des Leitterms als Produkt von Termen $t_1 \in \mathbb{T}(x_{11}, x_{21},\ldots,x_{m1})$ sowie $t_2 \in \mathbb{T}(x_{12}, x_{22},\ldots,x_{m2})$ vom Grad d_1 bzw. d_2. Dann gilt $d_1 \geq d_2 + e_2(\mathcal{M})$ und der eindeutig bestimmte Term $t \in \mathbb{T}^{n\cdot m}$ aus Lemma A.1.21 ist wie folgt festgelegt:*

(1) Gilt $d_1 = 0$ oder $d_2 + e_2(\mathcal{M}) = 0$, so gilt $t = 1$.

(2) Gilt $d_1 > 0$ und $d_2 + e_2(\mathcal{M}) > 0$, so schreibe $t_1 = x_{a(1),1}\cdots x_{a(d_2+e_2(\mathcal{M})),1}\cdots x_{a(d_1)1}$ mit $a(1),\ldots,a(d_1) \in \{1,\ldots,m\}$ und $a(1) \leq \ldots \leq a(d_1)$. Dann ist $x_{a(1),1}\cdots x_{a(d_2+e_2(\mathcal{M})),1}$ der gesuchte Term t.

Wir wollen nun noch die Frage beantworten, warum es so entscheidend ist, dass $e_1(\mathcal{M}) = 0$ gilt. Angenommen, es würde $e_1(\mathcal{M}) = 1$ gelten. Dann würde der Eintrag e_1 in der zweiten Spalte von $\mathcal{M}$ stehen, d.h. $\mathcal{M}$ enthält eine Zeile der Form $(x_i,\ e_1)$ mit $i \in \{1,\ldots,m\}$. Wegen

$$|x_i\ e_1| = \det\begin{pmatrix} x_{i1} & x_{i2} \\ 1 & 0 \end{pmatrix} = -x_{i2}$$

würde somit ein Faktor x_{i2} in $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ der Determinante dieser Zeile entsprechen. Das wiederum würde bedeuten, dass der Grad d_2 nicht mehr mit der Anzahl der Zeilenvektoren von $\mathcal{X}$ in der zweiten Spalte von $\mathcal{M}$ korrespondiert. Die Zahl $e_1(\mathcal{M}) = 1$ verrät uns zwar, dass ein Faktor x_{i2} zu einer Zeile der Form $(x_i,\ e_1)$ gehört, allerdings nicht, welche das ist. Somit ist eine eindeutige Berechnung von t nicht mehr möglich. Unter der zusätzlichen Voraussetzung $e_1(\mathcal{M}) = 0$ haben wir damit die Möglichkeit, aus dem Leitterm des Wertes eines Standard SO_2-Tableaus dieses eindeutig zu rekonstruieren, sofern wir die Anzahl der e_2-Einträge kennen (nach [Ric89], Proposition 12, S. 60).

Satz A.1.23. *Sei $(\mathcal{M}, W)$ ein Standard SO_2-Tableau mit $e_1(\mathcal{M}) = 0$. Dann ist $(\mathcal{M}, W)$ eindeutig durch den Leitterm $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle)$ und die Zahl $e_2(\mathcal{M})$ bestimmt.*

Beweis: Seien $t_1 \in \mathbb{T}(x_{1,1},\ldots,x_{m,1})$ und $t_2 \in \mathbb{T}(x_{1,2},\ldots,x_{m,2})$ Terme mit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = t_1\cdot t_2$. Mit Hilfe von Korollar A.1.22 und unter Verwendung von $e_2(\mathcal{M})$ erhalten wir den eindeutig bestimmten Term $t \in \mathbb{T}^{n\cdot m}$ mit $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) = t\cdot t_2$ und $t_1 = t\cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$.

Somit lassen sich $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ und $\mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$ aus $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle)$ bestimmen. Mit Hilfe von Satz A.1.8 können wir die Standard $\mathcal{H}$-Matrix $\mathcal{M}$ aus $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|)$ rekonstruieren. Analog liefert Satz A.1.13 das Standard $\mathcal{H}$-Tupel W aus $\mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle)$. □

Das nachfolgende Korollar stellt ein unmittelbares Resultat aus dem letzten Satz dar und lässt uns Standard SO_2-Tableaus anhand ihrer Leitterme vergleichen.

Korollar A.1.24. *Zwei Standard* SO_2*-Tableaus* $(\mathcal{M}, W), (\mathcal{L}, V)$ *mit* $e_1(\mathcal{M}) = e_1(\mathcal{L}) = 0$ *und* $e_2(\mathcal{M}) = e_2(\mathcal{L})$ *sind genau dann identisch, wenn* $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{L}|\langle V\rangle)$ *gilt.*

Spezielle Erzeugendensysteme des Polynomrings

Wir wollen nun zu einer bereits aufgeworfenen Frage zurückkehren: Können wir jedes Polynom mit Werten von SO_2-Tableaus darstellen? Der nächste Satz liefert uns die einfache Antwort auf diese Frage (vgl. Bemerkung in Beweis zu [Ric89], Proposition 13, S. 61).

Satz A.1.25. *Die Werte der* $\mathcal{H}$*-Matrizen bilden ein Erzeugendensystem des Polynomrings* P *als* K*-Vektorraum. Damit erzeugen insbesondere auch die Werte der* SO_2*-Tableaus* P *als* K*-Vektorraum.*

Beweis: Wegen $|e_1\ e_2| = 1$ und $|e_1\ x_i| = x_{i,2}$ sowie $|x_i\ e_2| = x_{i,1}$ für alle $i \in \{1, \ldots, m\}$ gibt es für alle Terme $t \in \mathbb{T}^{mn}$ eine $\mathcal{H}$-Matrix $\mathcal{M}$ mit $t = |\mathcal{M}|$. Gemäß dem Basissatz von Macaulay (vgl. [KR00], Theorem 1.5.7) ist $\mathbb{T}^{mn}$ eine K-Vektorraumbasis von P. Somit erzeugen die Werte der $\mathcal{H}$-Matrizen den Polynomring P als K-Vektorraum.

Da $(\mathcal{M}, \emptyset)$ für alle $\mathcal{H}$-Matrizen $\mathcal{M}$ ein SO_2-Tableau ist und zudem $\langle\emptyset\rangle = 1$ gilt, folgt die Zusatzbehauptung sofort. □

Damit drängt sich natürlich sofort die Frage auf, ob sich jedes Polynom auch mit Werten von Standard SO_2-Tableaus darstellen lässt. In Standard SO_2-Tableaus kommen Zeilen der Form $(e_1,\ x)$ nicht vor. Somit kann der naive Zugang, jeden Term durch einen geeigneten Wert zu ersetzen, nicht funktionieren. Allerdings lässt sich zeigen, dass der Wert jedes SO_2-Tableaus durch Werte von Standard SO_2-Tableaus darstellbar ist. Zum Beweis des entsprechenden Satzes ist jedoch einiges an Vorarbeit zu leisten, die hier im Gegensatz zu [Ric89] ausführlicher und strukturierter aufbereitet ist. Wir betrachten zunächst eine ganz spezielle Menge von SO_2-Tableaus.

Sei $\mathrm{Tab}' \subseteq \mathrm{Tab}$ die Menge der SO_2-Tableaus $(\mathcal{L}, V)$, deren $\mathcal{H}$-Matrix $\mathcal{L}$ die folgenden Eigenschaften erfüllt:

(i) Die Einträge jeder Zeile von $\mathcal{L}$ sind streng monoton absteigend geordnet.

(ii) Es gilt $e_1(\mathcal{L}) \leq 1$, d.h. $\mathcal{L}$ besitzt höchstens einen e_1 Eintrag.

(iii) Die Zeile $(e_1\ e_2)$ kommt in $\mathcal{L}$ nicht vor.

Wie man sofort sieht, erfüllen die $\mathcal{H}$-Matrizen der Elemente aus Tab' bis auf die Ordnung innerhalb der Spalten alle Eigenschaften einer $\mathcal{H}$-Matrix eines Standard SO_2-Tableaus. Außerdem haben die Elemente der Menge Tab' noch weitere Eigenschaften, die wir für den Beweis des Satzes benötigen.

Lemma A.1.26. (Eigenschaften von Tab')
Sei $\mathrm{Tab}' \subseteq \mathrm{Tab}$ *wie oben. Dann gilt:*

(i) Für alle $\mathcal{H}$*-Tupel* V *ist das* SO_2*-Tableau* $(\emptyset, V)$ *ein Element von* Tab'*. Insbesondere gilt also* $(\emptyset, \emptyset) \in \mathrm{Tab}'$*.*

(ii) Jedes Standard SO_2*-Tableau ist ein Element von* Tab'*, d.h. es gilt* $\mathrm{STab} \subseteq \mathrm{Tab}' \subseteq \mathrm{Tab}$*.*

(iii) Für alle $(\mathcal{L}_1, V_1), \ldots, (\mathcal{L}_k, V_k) \in \mathrm{Tab}'$ *gibt es ein* $(\mathcal{L}, V) \in \mathrm{Tab}'$ *mit*

$$|\mathcal{L}|\langle V\rangle = |\mathcal{L}_1|\langle V_1\rangle \cdots |\mathcal{L}_k|\langle V_k\rangle.$$

Beweis: (i) Klar.

(ii) Jede $\mathcal{H}$-Matrix $\mathcal{M}$ eines Standard SO_2-Tableaus $(\mathcal{M}, W)$ erfüllt per Definition die Anforderungen der Elemente aus Tab'. Somit gilt $(\mathcal{M}, W) \in \mathrm{Tab}'$ für alle $(\mathcal{M}, W) \in \mathrm{STab}$.

(iii) Seien $(\mathcal{L}_1, V_1), \ldots, (\mathcal{L}_k, V_k) \in \mathrm{Tab}'$. Setze $\mathcal{L} := (\mathcal{L}_1, \ldots, \mathcal{L}_k)^{\mathrm{tr}}$ und sei V die Konkatenation der Tupel $V_1, \ldots, V_k$. Da $\mathcal{L}$ die Anforderungen an die Elemente aus Tab' erfüllt, gilt $(\mathcal{L}, V) \in \mathrm{Tab}'$ und es folgt $|\mathcal{L}|\langle V\rangle = |\mathcal{L}_1| \cdots |\mathcal{L}_k| \cdot \langle V_1\rangle \cdots \langle V_k\rangle = |\mathcal{L}_1|\langle V_1\rangle \cdots |\mathcal{L}_k|\langle V_k\rangle$.

□

Wir wollen nun als ersten Schritt zeigen, dass sich der Wert jedes SO_2-Tableaus mit Werten der Elemente aus Tab' schreiben lässt (erwähnt im Beweis zu [Ric89], Proposition 11).

Lemma A.1.27. *Sei* $\mathrm{Tab}' \subseteq \mathrm{Tab}$ *wie oben und sei* $G_{\mathrm{Tab}'}$ *die von* $\mathrm{Val}(\mathrm{Tab}')$ *erzeugte additive Gruppe. Dann gilt* $|\mathcal{M}|\langle W\rangle \in G_{\mathrm{Tab}'}$ *für alle* SO_2*-Tableaus* $(\mathcal{M}, W)$*.*

Beweis: Sei $(\mathcal{M}, W)$ ein SO_2-Tableau. Gilt $(\mathcal{M}, W) \in \mathrm{Tab}'$, so ist nichts weiter zu zeigen. Sei also im Folgenden $(\mathcal{M}, W) \notin \mathrm{Tab}'$, d.h. mindestens eine der Bedingungen an die Elemente aus Tab' ist nicht erfüllt. Laut Lemma A.1.26 reicht es, nur den Wert $|\mathcal{M}|$ von $\mathcal{M}$ zu betrachten, d.h. zu zeigen, dass es SO_2-Tableaus $(\mathcal{L}_1, V_1), \ldots, (\mathcal{L}_k, V_k) \in \mathrm{Tab}'$ und $\alpha_1, \ldots, \alpha_k \in \{-1, 1\}$ gibt mit

$$|\mathcal{M}| = \alpha_1|\mathcal{L}_1|\langle V_1\rangle + \ldots + \alpha_k|\mathcal{L}_k|\langle V_k\rangle.$$

Denn wegen $(\emptyset, W) \in \mathrm{Tab}'$ folgt die Behauptung dann mit Lemma A.1.26 aus

$$|\mathcal{M}|\langle W\rangle = \alpha_1|\mathcal{L}_1|\langle V_1\rangle \cdot |\emptyset|\langle W\rangle + \ldots + \alpha_k|\mathcal{L}_k|\langle V_k\rangle \cdot |\emptyset|\langle W\rangle.$$

Sei o.B.d.A. $|\mathcal{M}| \neq 0$ und $\mathcal{M} \neq \emptyset$, d.h. $\mathcal{M}$ ist eine $\mathcal{H}$-Matrix der Länge $l > 0$, also von der Form

$$\mathcal{M} = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix},$$

wobei $a_1, \ldots, a_l$ und $b_1, \ldots, b_l$ Zeilenvektoren von $\mathcal{H}$ sind. Wegen $|\mathcal{M}| \neq 0$ gilt $a_i \neq b_i$ für alle $i \in \{1, \ldots, l\}$.

Wir nehmen zunächst an, dass genau eine der drei Bedingungen an die Elemente aus Tab' nicht erfüllt ist, die wir nun im Einzelnen betrachten wollen.

(i) Es gibt mindestens eine Zeile in $\mathcal{M}$, deren Einträge nicht streng monoton absteigend geordnet sind. Wegen $|a_i\ b_i| = -|b_i\ a_i|$ für alle $i \in \{1, \ldots, l\}$ gibt es eine $\mathcal{H}$-Matrix $\mathcal{M}'$ der Länge $l' = l$ mit $|\mathcal{M}| = \pm|\mathcal{M}'|$, in der die Einträge jeder Zeile bzgl. $\prec$ absteigend geordnet sind. Aus $(\mathcal{M}', \emptyset) \in \mathrm{Tab}'$ folgt $|\mathcal{M}| \in G_{\mathrm{Tab}'}$.

(ii) Es gilt $e_1(\mathcal{M}) \geq 2$. Da die Einträge der einzelnen Zeilen absteigend geordnet sind und (e_1, e_2)-Zeilen in $\mathcal{M}$ ebenfalls nicht vorkommen, befinden sich die Einträge e_1 in der zweiten Spalte von $\mathcal{M}$. Ebenso befinden sich eventuelle Einträge e_2 in der zweiten Spalte, d.h. die erste Spalte von $\mathcal{M}$ enthält nur Zeilenvektoren aus $\mathcal{X}$. Seien x, y beliebige Zeilenvektoren von $\mathcal{X}$. Dann enthält $\mathcal{M}$ also nur Zeilen der Form (x, y), (x, e_2) oder (x, e_1).

Folglich gibt es eine $\mathcal{H}$-Matrix $\tilde{\mathcal{M}}$ mit $|\tilde{\mathcal{M}}| = |\mathcal{M}|$, deren Zeilen so geordnet sind, dass zuerst alle Zeilen der Form $(x,\ y)$, dann alle Zeilen der Form $(x,\ e_2)$ und schließlich alle Zeilen der Form $(x,\ e_1)$ kommen.

Jeweils ein Paar von Faktoren der Form $|x\ e_1|\cdot|y\ e_1|$ in $|\tilde{\mathcal{M}}|$ lässt sich laut Lemma A.1.15(i) durch die Differenz $\langle(x,y)\rangle - |x\ e_2|\cdot|y\ e_2|$ ersetzen. Für $\mathcal{L} := \left(\begin{smallmatrix} x & e_2 \\ y & e_2 \end{smallmatrix}\right)$ und $V := (x,y)$ gilt $(\emptyset, V), (\mathcal{L}, \emptyset) \in \mathrm{Tab}'$. Somit ist jedes Produkt $|x\ e_1|\cdot|y\ e_1|$ wegen

$$|x\ e_1|\cdot|y\ e_1| = |\emptyset|\langle V\rangle - |\mathcal{L}|\langle\emptyset\rangle$$

ein Element von $G_{\mathrm{Tab}'}$. Es gibt $k := \left\lfloor \frac{e_1(\mathcal{M})}{2} \right\rfloor$ derartige Paare in $|\tilde{\mathcal{M}}|$, d.h. ersetzen wir alle Paare wie angegeben, so bleibt höchstens ein Faktor der Form $|x\ e_1|$ übrig.

Sei $\mathcal{L}$ die Teilmatrix von $\tilde{\mathcal{M}}$, die alle Zeilen der Form $(x,\ y)$ und ggf. eine Zeile der Form $(x,\ e_1)$ enthält. Für die k Paare gibt es also $\mathcal{H}$-Matrizen $\mathcal{L}_1, \ldots, \mathcal{L}_k$ und $\mathcal{H}$-Tupel $V_1, \ldots, V_k$ mit $(\emptyset, V_1), (\mathcal{L}_1, \emptyset), \ldots, (\emptyset, V_k), (\mathcal{L}_k, \emptyset) \in \mathrm{Tab}'$ so, dass gilt:

$$\begin{aligned} |\mathcal{M}| &= |\tilde{\mathcal{M}}| = |\mathcal{L}|\cdot(|\emptyset|\langle V_1\rangle - |\mathcal{L}_1|\langle\emptyset\rangle)\cdot\ldots\cdot(|\emptyset|\langle V_k\rangle - |\mathcal{L}_k|\langle\emptyset\rangle) \\ &= |\mathcal{L}|\langle\emptyset\rangle\cdot(|\emptyset|\langle V_1\rangle - |\mathcal{L}_1|\langle\emptyset\rangle)\cdot\ldots\cdot(|\emptyset|\langle V_k\rangle - |\mathcal{L}_k|\langle\emptyset\rangle) \end{aligned}$$

Wegen $(\mathcal{L}, \emptyset) \in \mathrm{Tab}'$ folgt $|\mathcal{M}| \in G_{\mathrm{Tab}'}$ durch Ausmultiplizieren unter Verwendung von Lemma A.1.26.

(iii) Es gibt mindestens eine $(e_1,\ e_2)$-Zeile in $\mathcal{M}$. Wegen $|e_1\ e_2| = 1$ gibt es eine $\mathcal{H}$-Matrix $\mathcal{M}'$ der Länge $l' < l$ mit $|\mathcal{M}| = |\mathcal{M}'|$, die keine Zeilen der Form (e_1, e_2) enthält. Wegen $\mathcal{M}' \in \mathrm{Tab}'$ folgt $|\mathcal{M}| \in G_{\mathrm{Tab}'}$.

Durch wiederholte Anwendung dieser drei Fälle folgt auch im allgemeinen Fall $|\mathcal{M}| \in G_{\mathrm{Tab}'}$ und damit $|\mathcal{M}|\langle W\rangle \in G_{\mathrm{Tab}'}$. □

Somit ist es also möglich, den Wert eines beliebigen SO_2-Tableaus $(\mathcal{M}, W)$ durch Werte von Elementen der Menge Tab' darzustellen. Als unmittelbare Folgerung ergibt sich folgendes Korollar.

Korollar A.1.28. *Die Werte der SO_2-Tableaus aus Tab' erzeugen den Polynomring P als K-Vektorraum.*

Dies verwendet Richman als ersten Schritt, um zu zeigen, dass sich $|\mathcal{M}|\langle W\rangle$ auch durch Werte von Standard SO_2-Tableaus darstellen lässt. Es ist natürlich leicht ersichtlich, dass wir diesem Ziel durch den letzten Satz tatsächlich ein Stück näher gekommen sind, da bekanntlich die $\mathcal{H}$-Matrizen der Elemente aus Tab' den Anforderungen einer $\mathcal{H}$-Matrix eines Standard SO_2-Tableaus schon beinahe gerecht werden. Dies liefert darüber hinaus sofort die Idee, wie Richman beweist, dass sich der Wert jedes SO_2-Tableaus durch Werte von Standard SO_2-Tableaus darstellen lässt. Zunächst können wir uns aufgrund des obigen Satzes auf die Elemente von Tab' beschränken. Da Tab' auch die Standard SO_2-Tableaus enthält, ist die Behauptung in dem Falle, dass $(\mathcal{M}, W) \in \mathrm{Tab}'$ standard ist, sofort klar. Ist ein Element $(\mathcal{M}, W) \in \mathrm{Tab}'$ nicht standard, so bedeutet das, dass mindestens einer der folgenden Fälle vorliegt:

(1) $\mathcal{M}$ ist nicht standard. Wegen $(\mathcal{M}, W) \in \mathrm{Tab}'$ ist das genau dann der Fall, wenn die Einträge der Spalten von $\mathcal{M}$ nicht absteigend geordnet sind.

(2) W ist nicht standard, d.h. die Komponenten von W sind nicht monoton absteigend geordnet.

(3) Es gibt einen Eintrag a in der ersten Spalte von $\mathcal{M}$ und eine Komponente w von W mit $w \prec a$.

Wir werden nun jeden der Fälle einzeln betrachten, womit der Beweis des eigentlichen Satzes unmittelbar folgt. Zum Beweis dieser Fälle schränkt RICHMAN die Menge Tab' noch weiter ein. Wir werden dem auch hier folgen, allerdings diese Theorie ausführlicher darstellen und zudem kleine Fehler in [Ric89] korrigieren. Dazu bezeichnen wir mit $\widehat{\mathcal{M}}$ diejenige $\mathcal{H}$-Matrix, die aus einer $\mathcal{H}$-Matrix $\mathcal{M}$ entsteht, indem man jeden e_1-Eintrag durch e_2 ersetzt. Wir werden nun auf die anfangs eingeführte Darstellung eines SO_2-Tableaus $(\mathcal{M}, W)$ als Tupel $\gamma_{(\mathcal{M},W)}$ zurückkommen und folgende Teilmengen von Tab' in Abhängigkeit von $(\mathcal{M}, W)$ betrachten: Für $(\mathcal{M}, W) \in \mathrm{Tab}'$ sei $\Gamma_{(\mathcal{M},W)} \subseteq \mathrm{Tab}'$ die Menge aller $(\mathcal{L}, V) \in \mathrm{Tab}'$, für die gilt:

(i) Es gibt ein Teiltupel von $\gamma_{(\widehat{\mathcal{M}},W)}$, das bis auf eine Permutation der Einträge mit $\gamma_{(\widehat{\mathcal{L}},V)}$ übereinstimmt.

(ii) Es gilt eine der folgenden Bedingungen:

(1) $\#\gamma_{(\mathcal{L},V)} = \#\gamma_{(\mathcal{M},W)}$ und $\gamma_{(\mathcal{L},V)} \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$,

(2) $\gamma_{(\mathcal{L},V)}$ ist eine Permutation von $\gamma_{(\mathcal{M},W)}$ und $\#V < \#W$,

(3) $\#\gamma_{(\mathcal{L},V)} < \#\gamma_{(\mathcal{M},W)}$.

RICHMAN schreibt als Bedingung (2) unter anderem $\gamma_{(\mathcal{L},V)}$ „equals" $\gamma_{(\mathcal{M},W)}$. Wie ein späterer Beweis zeigen wird, kann er damit allerdings entweder keine komponentenweise Übereinstimmung beider Tupel gemeint haben, oder es handelt sich tatsächlich um einen Fehler. Die Menge $\Gamma_{(\mathcal{M},W)}$ weist eine Reihe von Eigenschaften auf, die RICHMAN in [Ric89] im Beweis zu Proposition 11 erwähnt, aber nicht weiter darauf eingeht.

Lemma A.1.29. (Eigenschaften von $\Gamma_{(\mathcal{M},W)}$)
Sei $\mathrm{Tab}' \subseteq \mathrm{Tab}$ *wie oben und* $(\mathcal{M}, W) \in \mathrm{Tab}'$. *Dann gilt:*

(i) Die Menge $\Gamma_{(\mathcal{M},W)}$ *ist endlich.*

(ii) Es gilt $(\mathcal{M}, W) \notin \Gamma_{(\mathcal{M},W)}$. *Damit folgt insbesondere* $\Gamma_{(\emptyset,\emptyset)} = \emptyset$.

(iii) Für jedes $(\mathcal{L}, V) \in \Gamma_{(\mathcal{M},W)}$ *ist* $\Gamma_{(\mathcal{L},V)}$ *eine echte Teilmenge von* $\Gamma_{(\mathcal{M},W)}$.

(iv) Ist $(\mathcal{M}, W)$ *nicht standard, so ist* $\Gamma_{(\mathcal{M},W)}$ *nicht-leer und es gibt ein von* $(\emptyset, \emptyset)$ *verschiedenes Standard* SO_2*-Tableau in* $\Gamma_{(\mathcal{M},W)}$.

Beweis: (i) Da $\#\gamma_{(\widehat{\mathcal{M}},W)}$ endlich ist und es nur endlich viele Permutationen der Einträge von $\gamma_{(\widehat{\mathcal{M}},W)}$ gibt, ist aufgrund von Eigenschaft (i) der Definition die Menge $\Gamma_{(\mathcal{M},W)}$ endlich.

(ii) Folgt sofort, da $(\mathcal{M}, W)$ in $\Gamma_{(\mathcal{M},W)}$ keine der Eigenschaften aus (ii) der Definition erfüllt.

(iii) Sei $(\mathcal{L}, V) \in \Gamma_{(\mathcal{M},W)}$. Aus (ii) folgt $(\mathcal{L}, V) \notin \Gamma_{(\mathcal{L},V)}$. Somit gilt $\Gamma_{(\mathcal{L},V)} \subsetneqq \Gamma_{(\mathcal{M},W)}$.

(iv) Ist $\mathcal{M} = \emptyset$, so muss lediglich W absteigend geordnet werden und es folgt die Behauptung. Sei nun $\mathcal{M} \neq \emptyset$, also eine $\mathcal{H}$-Matrix der Länge $l > 0$, und sei W ohne Einschränkung absteigend geordnet. Sei $\gamma_{(\mathcal{M},W)} = (a_1, b_1, a_2, b_2, \ldots, a_l, b_l, w_1, \ldots, w_{2u})$ mit $u \in \mathbb{N}$. Wegen $(\mathcal{M}, W) \in \mathrm{Tab}'$ gilt $a_i \succ b_i$ für alle $i \in \{1, \ldots, l\}$. Somit gibt es stets ein Teiltupel γ von $\gamma_{(\mathcal{M},W)}$, das die Bedingungen der Definition erfüllt und ein Standard SO_2-Tupel repräsentiert.

□

Bekanntlich erfüllen die $\mathcal{H}$-Matrizen der Elemente aus Tab' bereits alle Anforderungen an $\mathcal{H}$-Matrizen eines Standard SO_2-Tableaus mit Ausnahme der Ordnung auf den Spalten. Ohne Auswirkungen auf den Wert einer $\mathcal{H}$-Matrix können wir jedoch die Zeilen so anordnen, dass die Einträge der ersten Spalte der $\mathcal{H}$-Matrix absteigend geordnet sind. Dabei werden Zeilen mit identischer erster Komponente weiter anhand der zweiten Komponente geordnet. Stimmen die Zeilen auch in dieser überein, werden die identischen Zeilen beliebig angeordnet. Wir werden nun zeigen, dass sich der Wert eines nicht-standard SO_2-Tableaus $(\mathcal{M}, W) \in \mathrm{Tab}'$ mit Werten von Elementen aus $\Gamma_{(\mathcal{M},W)}$ darstellen lässt (enthalten im Beweis zu [Ric89], Proposition 11, S. 59).

Lemma A.1.30. *Sei* $\mathrm{Tab}' \subseteq \mathrm{Tab}$ *wie oben, sei* $(\mathcal{M}, W) \in \mathrm{Tab}'$ *nicht standard und sei* $G_{\Gamma_{(\mathcal{M},W)}}$ *die von den Werten der Elemente aus* $\Gamma_{(\mathcal{M},W)}$ *erzeugte additive Gruppe. Dann gilt* $|\mathcal{M}|\langle W\rangle \in G_{\Gamma_{(\mathcal{M},W)}}$.

Beweis: Gilt $\mathcal{M} = \emptyset$, so ist W ein nicht-leeres $\mathcal{H}$-Tupel, dessen Komponenten nicht monoton absteigend geordnet sind. Durch Ordnen der Komponenten von W erhalten wir ein monoton absteigend geordnetes $\mathcal{H}$-Tupel W' mit $(\emptyset, W') \in G_{\Gamma_{(\mathcal{M},W)}}$. Somit folgt die Behauptung.

Sei also $\mathcal{M} \neq \emptyset$. Ohne Einschränkung seien die Zeilen von $\mathcal{M}$ so angeordnet, dass die Einträge der ersten Spalte von $\mathcal{M}$ monoton absteigend geordnet sind. Sei $\mathcal{M}$ also von der Form

$$\mathcal{M} = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix},$$

wobei $a_1, \ldots, a_l, b_1, \ldots, b_l$ Zeilenvektoren von $\mathcal{H}$ sind mit $a_1 \succeq \ldots \succeq a_l$, und sei W von der Form $W = (w_1, \ldots, w_{2u})$ mit Zeilenvektoren $w_1, \ldots, w_{2u}$ von $\mathcal{X}$.

Da $(\mathcal{M}, W)$ nicht standard ist, liegt mindestens einer der oben erwähnten Fälle vor. Wir nehmen zunächst an, dass genau einer dieser Fälle vorliegt und betrachten diese im Einzelnen.

(1) $\mathcal{M}$ ist keine Standard $\mathcal{H}$-Matrix, d.h. es gibt mindestens ein Paar $(i, j) \in \{1, \ldots, l\}^2$ mit $a_i \succeq a_j$ und $b_i \prec b_j$. Gemäß Lemma A.1.16 (i) gilt:

$$\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = \begin{vmatrix} a_i & b_j \\ a_j & b_i \end{vmatrix} - \begin{vmatrix} a_i & a_j \\ b_j & b_i \end{vmatrix}. \qquad (*)$$

Setze

$$\begin{aligned} \lambda_1 &:= (a_1, b_1, \ldots, a_i, b_j, \ldots, a_j, b_i, \ldots, a_l, b_l, w_1, \ldots, w_{2u}) \\ \lambda_2 &:= (a_1, b_1, \ldots, a_i, a_j, \ldots, b_j, b_i, \ldots, a_l, b_l, w_1, \ldots, w_{2u}). \end{aligned}$$

Offensichtlich sind λ_1 und λ_2 Permutationen von

$$\gamma_{(\mathcal{M},W)} = (a_1, b_1, \ldots, a_i, b_i, \ldots, a_j, b_j, \ldots, a_l, b_l, w_1, \ldots, w_{2u})$$

Wegen $a_i \succeq a_j \succ b_j \succ b_i$ gilt $\lambda_1 \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$ und $\lambda_2 \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$. Somit gibt es SO_2-Tableaus $(\mathcal{L}_1, W), (\mathcal{L}_2, W) \in \Gamma_{(\mathcal{M},W)}$ mit $\lambda_1 = \gamma_{(\mathcal{L}_1,W)}$ und $\lambda_2 = \gamma_{(\mathcal{L}_2,W)}$ sowie

$$|\mathcal{M}|\langle W\rangle = |\mathcal{L}_1|\langle W\rangle - |\mathcal{L}_2|\langle W\rangle,$$

womit $|\mathcal{M}|\langle W\rangle \in G_{\Gamma_{(\mathcal{M},W)}}$ folgt.

Ein Fall muss noch gesondert betrachtet werden. Es ist möglich, dass $b_i = e_2$ und $b_j = e_1$ gilt. In diesem Fall erhalten wir auf der rechten Seite von Gleichung $(*)$ eine $(e_1\ e_2)$-Zeile.

Dann wäre diese Matrix aber kein Element von Tab′. Ohne Auswirkung auf den Wert streichen wir in diesem Fall diese Zeile, womit schließlich $\#\lambda_2 < \#\gamma_{(\mathcal{M},W)}$ gilt. Somit gibt es auch hier ein $\mathcal{L}_2$ mit $(\mathcal{L}_2, W) \in \Gamma_{(\mathcal{M},W)}$ und die Behauptung folgt erneut.

(2) $\mathcal{M}$ ist eine Standard $\mathcal{H}$-Matrix, aber W ist kein Standard $\mathcal{H}$-Tupel, d.h. es gilt insbesondere $W \neq \emptyset$ und W ist nicht absteigend geordnet. Da W nicht absteigend geordnet ist, gibt es mindestens ein Paar in W, das aufsteigend geordnet ist. Sei W in der Form

$$W = (w_1, w_2, \ldots, w_{2i-1}, w_{2i}, \ldots, w_{2u-1}, w_{2u}).$$

1. Fall: Für ein $i \in \{1, \ldots, u\}$ gilt $w_{2i-1} \prec w_{2i}$. Setze

$$V := (w_1, w_2, \ldots, w_{2i}, w_{2i-1}, \ldots, w_{2u-1}, w_{2u}).$$

Da $\gamma_{(\mathcal{M},V)}$ eine Permutation von $\gamma_{(\mathcal{M},W)}$ ist und $\gamma_{(\mathcal{M},V)} \succ_{\text{Lex}} \gamma_{(\mathcal{M},W)}$ gilt, folgt sofort $(\mathcal{M}, V) \in \Gamma_{(\mathcal{M},W)}$. Wegen $\langle (w_{2i-1}, w_{i2}) \rangle = \langle (w_{i2}, w_{2i-1}) \rangle$ gilt $\langle V \rangle = \langle W \rangle$, also $|\mathcal{M}|\langle W \rangle = |\mathcal{M}|\langle V \rangle$, und es folgt die Behauptung.

2. Fall: Es gibt ein $i \in \{2, \ldots, u\}$ mit $(w_{2i-3}, w_{2i-2}) \prec_{\text{Lex}} (w_{2i-1}, w_{2i})$, $w_{2i-3} \succeq w_{2i-2}$ sowie $w_{2i-1} \succeq w_{2i}$. Setze

$$V := (w_1, w_2, \ldots, w_{2i-1}, w_{2i}, w_{2i-3}, w_{2i-2}, \ldots, w_{2u-1}, w_{2u}).$$

Offensichtlich ist das Tupel $\gamma_{(\mathcal{M},V)}$ des Paares $(\mathcal{M}, V)$ erneut eine Permutation von $\gamma_{(\mathcal{M},W)}$. Wegen $(w_{2i-1}, w_{2i}) \succ_{\text{Lex}} (w_{2i-3}, w_{2i-2})$ gilt $\gamma_{(\mathcal{M},V)} \succ_{\text{Lex}} \gamma_{(\mathcal{M},W)}$, also $(\mathcal{M}, V) \in \Gamma_{(\mathcal{M},W)}$. Weiter gilt:

$$\begin{aligned} \langle (w_{2i-3}, w_{2i-2}, w_{2i-1}, w_{2i}) \rangle &= \langle (w_{2i-3}, w_{2i-2}) \rangle \cdot \langle (w_{2i-1}, w_{2i}) \rangle \\ &= \langle (w_{2i-1}, w_{2i}, w_{2i-3}, w_{2i-2}) \rangle. \end{aligned}$$

Somit folgt $\langle V \rangle = \langle W \rangle$, also $|\mathcal{M}|\langle W \rangle = |\mathcal{M}|\langle V \rangle$, und damit die Behauptung.

3. Fall: Es gibt ein $i \in \{1, \ldots, u-1\}$ mit $w_{2i} \prec w_{2i+1}$. Wir betrachten das Teiltupel

$$(w_{2i-1}, w_{2i}, w_{2i+1}, w_{2i+2})$$

von W. Laut Lemma A.1.16 (ii) gilt für dessen Wert:

$$\langle (w_{2i-1}, w_{2i}, w_{2i+1}, w_{2i+2}) \rangle = \langle (w_{2i-1}, w_{2i+1}, w_{2i}, w_{2i+2}) \rangle - \begin{vmatrix} w_{2i-1} & w_{2i+2} \\ w_{2i} & w_{2i+1} \end{vmatrix}$$

Sei $\tilde{W} = (w_1, w_2, \ldots, w_{2i-2}, w_{2i+3}, \ldots, w_{2u})$ und setze

$$V := (w_1, w_2, \ldots, w_{2i-1}, w_{2i+1}, w_{2i}, w_{2i+2}, \ldots, w_{2u})$$

sowie

$$\mathcal{L} = \begin{pmatrix} & \mathcal{M} & \\ w_{2i-1} & & w_{2i+2} \\ & w_{2i} & w_{2i+1} \end{pmatrix}.$$

Dann sind $\gamma_{(\mathcal{M},V)}$ und $\gamma_{(\mathcal{L},\tilde{W})}$ Permutationen von $\gamma_{(\mathcal{M},W)}$. Wegen $\gamma_{(\mathcal{M},V)} \succ \gamma_{(\mathcal{M},W)}$ und $\#\tilde{W} < \#W$ gilt $(\mathcal{M}, V) \in \Gamma_{(\mathcal{M},W)}$ und $(\mathcal{L}, \tilde{W}) \in \Gamma_{(\mathcal{M},W)}$. Somit gilt:

$$\begin{aligned} |\mathcal{M}|\langle W \rangle &= |\mathcal{M}|\langle \tilde{W} \rangle \cdot \langle (w_{2i-1}, w_{2i}, w_{2i+1}, w_{2i+2}) \rangle \\ &= |\mathcal{M}|\langle \tilde{W} \rangle \langle (w_{2i-1}, w_{2i}, w_{2i+1}, w_{2i+2}) \rangle - |\mathcal{M}| \cdot \begin{vmatrix} w_{2i-1} & w_{2i+2} \\ w_{2i} & w_{2i+1} \end{vmatrix} \cdot \langle \tilde{W} \rangle \\ &= |\mathcal{M}|\langle V \rangle - |\mathcal{L}|\langle \tilde{W} \rangle \end{aligned}$$

und es folgt die Behauptung.

(3) Sowohl $\mathcal{M}$, als auch W sind standard und beide nicht leer. Da $(\mathcal{M}, W)$ aber nicht standard ist, bedeutet das, dass es ein Element in W gibt, das größer ist als ein Element in der ersten Spalte von $\mathcal{M}$, d.h. es gibt ein $i \in \{1, \ldots, l\}$ und ohne Einschränkung ein $j \in \{1, \ldots, u\}$ mit $a_i \prec w_{2j}$. Denn wegen $w_{2j-1} \succeq w_{2j}$ folgt dann sofort $a_i \prec w_{2j-1}$.

Gelte zunächst $e_1(\mathcal{M}) = 0$. Dann ist die i-te Zeile von $\mathcal{M}$ entweder von der Form $(a_i,\ b_i)$, wobei b_i ein Zeilenvektor von $\mathcal{X}$ ist, oder von der Form $(a_i,\ e_2)$, falls $e_2(\mathcal{M}) > 0$ gilt. Im ersten Fall kann man die Faktoren $|a_i\ b_i|\langle(w_{2j-1}, w_{2j})\rangle$ gemäß Lemma A.1.16 (iii) wie folgt ersetzen:

$$|a_i\ b_i|\langle(w_{2j-1}, w_{2j})\rangle = -|w_{2j-1}\ a_i|\langle(w_{2j}, b_i)\rangle + |w_{2j-1}\ b_i|\langle(w_{2j}, a_i)\rangle.$$

Setze

$$\mathcal{L}_1 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & a_i \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}, \quad \mathcal{L}_2 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & b_i \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}$$

und

$$V_1 := (w_1, w_2, \ldots, w_{2j}, b_i, \ldots, w_{2u}), \quad V_2 := (w_1, w_2, \ldots, w_{2j}, a_i, \ldots, w_{2u}).$$

Wegen $w_{2j-1} \succ a_i$ gilt $\gamma_{(\mathcal{L}_1,V_1)} \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$ und $\gamma_{(\mathcal{L}_2,V_2)} \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$, also sind die Paare $(\mathcal{L}_1, V_1)$ und $(\mathcal{L}_2, V_2)$ Elemente von $\Gamma_{(\mathcal{M},W)}$. Somit folgt

$$|\mathcal{M}|\langle W\rangle = -|\mathcal{L}_1|\langle V_1\rangle + |\mathcal{L}_2|\langle V_2\rangle$$

und damit die Behauptung.

Ist die i-te Zeile von $\mathcal{M}$ von der Form $(a_i,\ e_2)$, so gilt laut Lemma A.1.15 (ii):

$$|a_i\ e_2|\langle(w_{2j-1}, w_{2j})\rangle = \begin{vmatrix} w_{2j-1} & a_i \\ w_{2j} & e_1 \end{vmatrix} + |w_{2j-1}\ e_2|\langle(w_{2j}, a_i)\rangle. \qquad (**)$$

Setze

$$\mathcal{L}_1 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & a_i \\ w_{2j} & e_1 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}, \quad \mathcal{L}_2 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & e_2 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}$$

und

$$V_1 := (w_1, \ldots, w_{2j-2}, w_{2j+1}, \ldots, w_{2u}), \quad V_2 := (w_1, w_2, \ldots, w_{2j}, a_i, \ldots, w_{2u}).$$

Wegen $w_{2j-1} \succ a_i$ gilt $\gamma_{(\mathcal{L}_2,V_2)} \succ_{\mathrm{Lex}} \gamma_{(\mathcal{M},W)}$, also $(\mathcal{L}_2, V_2) \in \Gamma_{(\mathcal{M},W)}$, und wegen $\#V_1 < \#W$ gilt $(\mathcal{L}_1, V_1) \in \Gamma_{(\mathcal{M},W)}$. Es folgt:

$$|\mathcal{M}|\langle W\rangle = |\mathcal{L}_1|\langle V_1\rangle + |\mathcal{L}_2|\langle V_2\rangle.$$

Gelte nun $e_1(\mathcal{M}) = 1$. Dann kann die i-te Zeile von $\mathcal{M}$ auch von der Form $(a_i,\ e_1)$ sein. Laut Lemma A.1.15 (iii) gilt

$$|a_i\ e_1|\langle(w_{2j-1}, w_{2j})\rangle = -\begin{vmatrix} w_{2j-1} & a_i \\ w_{2j} & e_2 \end{vmatrix} + |w_{2j-1}\ e_1|\langle(w_{2j}, a_i)\rangle.$$

Setze

$$\mathcal{L}_1 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & a_i \\ w_{2j} & e_2 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}, \quad \mathcal{L}_2 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j-1} & e_1 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}$$

und

$$V_1 := (w_1, \ldots, w_{2j-2}, w_{2j+1}, \ldots, w_{2u}), \quad V_2 := (w_1, w_2, \ldots, w_{2j}, a_i, \ldots, w_{2u}).$$

Dann gilt analog zu oben $(\mathcal{L}_1, V_1), (\mathcal{L}_2, V_2) \in \Gamma_{(\mathcal{M},W)}$ und

$$|\mathcal{M}|\langle W\rangle = -|\mathcal{L}_1|\langle V_1\rangle + |\mathcal{L}_2|\langle V_2\rangle.$$

Der Fall $(a_i,\ e_2)$ muss nun noch eigens betrachtet werden, da sich Gleichung $(*)$ hier nicht anwenden lässt. Würden wir dies tun, so würden wir eine Matrix $\mathcal{L}_1$ mit $e_1(\mathcal{L}_1) = 2$ erhalten und damit kein Element aus Tab$'$, also folglich auch kein Element aus $\Gamma_{(\mathcal{M},W)}$. Wir betrachten stattdessen zusätzlich die Zeile von $\mathcal{M}$ mit dem e_1-Eintrag. Wir nehmen an, diese Zeile sei für $k \in \{1, \ldots, l\}$ die k-te Zeile von $\mathcal{M}$. Laut Lemma A.1.15 (iv) gilt für den Faktor $|a_k\ e_1| \cdot |a_i\ e_2|$ in $|\mathcal{M}|$:

$$\begin{vmatrix} a_k & e_1 \\ a_i & e_2 \end{vmatrix} \langle (w_{2j-1}, w_{2j})\rangle = -\begin{vmatrix} a_k & e_2 \\ w_{2j-1} & a_i \\ w_{2j} & e_2 \end{vmatrix} + |w_{2j-1}\ a_i|\langle (a_k, w_{2j})\rangle + \begin{vmatrix} a_k & e_1 \\ w_{2j-1} & e_2 \end{vmatrix} \langle (w_{2j}, a_i)\rangle.$$

Setze

$$\mathcal{L}_1 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_k & e_1 \\ \vdots & \vdots \\ w_{2j-1} & a_i \\ w_{2j} & e_2 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}, \quad \mathcal{L}_2 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ w_{2j} & a_i \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}, \quad \mathcal{L}_3 := \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_k & e_1 \\ \vdots & \vdots \\ w_{2j-1} & e_2 \\ \vdots & \vdots \\ a_l & b_l \end{pmatrix}$$

und $V_1 := (w_1, w_2, \ldots, w_{2j-2}, w_{2j+1}, \ldots, w_{2u})$, $V_2 := (w_1, w_2, \ldots, a_k, w_{2j}, \ldots, w_{2u})$ sowie $V_3 := (w_1, \ldots, w_{2j}, a_i, \ldots, w_{2u})$. Dann folgt sofort

$$|\mathcal{M}|\langle W\rangle = -|\mathcal{L}_1|\langle V_1\rangle + |\mathcal{L}_2|\langle V_2\rangle + |\mathcal{L}_3|\langle V_3\rangle.$$

Wegen $\#V_1 < \#W$ gilt $(\mathcal{L}_1, V_1) \in \Gamma_{(\mathcal{M},W)}$ und wegen $a_i \prec w_{2j}$ bzw. $a_i \prec w_{2j-1}$ gilt $(\mathcal{L}_2, V_2) \in \Gamma_{(\mathcal{M},W)}$ bzw. $(\mathcal{L}_3, V_3) \in \Gamma_{(\mathcal{M},W)}$. Somit folgt auch hier die Behauptung.

Durch ggf. wiederholte Anwendung dieser Fälle folgt auch allgemein $|\mathcal{M}|\langle W\rangle \in G_{\Gamma_{(\mathcal{M},W)}}$. □

Mit diesen Vorbereitungen lässt sich zum Abschluss dieses Abschnitts der ursprünglich in [Ric89] recht lange und etwas undurchsichtige Beweis, dass sich der Wert eines beliebigen SO_2-Tableaus durch Werte von Standard SO_2-Tableaus darstellen lässt, kürzer und strukturierter fassen (vgl. [Ric89], Proposition 11).

Satz A.1.31. *Sei G_{STab} die von den Werten der Standard SO_2-Tableaus erzeugte additive Gruppe. Dann ist für alle SO_2-Tableaus $(\mathcal{M}, W)$ der Wert $|\mathcal{M}|\langle W\rangle$ ein Element dieser Gruppe.*

Beweis: Sei $(\mathcal{M}, W)$ ein SO_2-Tableau und sei $\mathrm{Tab}' \subseteq \mathrm{Tab}$ wie oben. Gemäß Lemma A.1.27 ist der Wert $|\mathcal{M}|\langle W\rangle$ ein Element der von den Werten der Elemente aus Tab' erzeugten additiven Gruppe. Somit reicht es, im Folgenden nur die Elemente aus Tab' zu betrachten.

Sei also $(\mathcal{M}, W) \in \mathrm{Tab}'$. Bekanntlich gilt $\mathrm{STab} \subseteq \mathrm{Tab}'$ (vgl. Lemma A.1.26). Ist also $(\mathcal{M}, W)$ bereits ein Standard SO_2-Tableau, so ist nichts weiter zu zeigen. Sei deshalb $(\mathcal{M}, W)$ nicht standard. Laut Lemma A.1.30 ist $|\mathcal{M}|\langle W\rangle$ ein Element der von den Werten der Elemente aus $\Gamma_{(\mathcal{M},W)}$ erzeugten additiven Gruppe, d.h. es gibt $(\mathcal{L}_1, V_1), \ldots, (\mathcal{L}_k, V_k) \in \Gamma_{(\mathcal{M},W)}$ und $\alpha_1, \ldots, \alpha_k \in \{-1, 1\}$ mit $|\mathcal{M}|\langle W\rangle = \alpha_1|\mathcal{L}_1|\langle V_1\rangle + \ldots + \alpha_k|\mathcal{L}_k|\langle V_k\rangle$.

Ist für $i \in \{1, \ldots, k\}$ ein $(\mathcal{L}_i, V_i)$ kein Standard SO_2-Tableau, so wenden wir erneut Lemma A.1.30 an und stellen $|\mathcal{L}_i|\langle V_i\rangle$ mit Werten von Elementen aus $\Gamma_{(\mathcal{L}_i,V_i)}$ dar. Aus Lemma A.1.29 wissen wir, dass $\Gamma_{(\mathcal{M},W)}$ endlich ist, dass $\Gamma_{(\mathcal{L}_i,V_i)} \subsetneqq \Gamma_{(\mathcal{M},W)}$ gilt und dass $\Gamma_{(\mathcal{M},W)}$ Standard SO_2-Tableaus enthält. Durch die rekursive Anwendung von Lemma A.1.30 erhalten wir also nach endlich vielen Schritten eine Darstellung von $|\mathcal{M}|\langle W\rangle$ durch Werte von Standard SO_2-Tableaus. □

Bekanntlich lässt sich jedes Polynom $f \in P$ als Linearkombination von Werten von SO_2-Tableaus darstellen. Da im vorigen Satz gezeigt wurde, dass jeder dieser Werte ein Element der von Werten von Standard SO_2-Tableaus erzeugten additiven Gruppe ist, ist es somit auch möglich, f mit Werten von Standard SO_2-Tableaus darzustellen.

Satz A.1.32. *Die Werte der Standard SO_2-Tableaus erzeugen den Polynomring P als K-Vektorraum.*

Beweis: Laut Satz A.1.25 erzeugen die Werte der SO_2-Tableaus P als K-Vektorraum. Da sich laut dem letzten Satz der Wert eines SO_2-Tableaus mit Werten von Standard SO_2-Tableaus schreiben lässt, bilden die Werte der Standard SO_2-Tableaus ebenfalls ein Erzeugendensystem des K-Vektorraums P. □

Wir wollen zum Abschluss dieses Abschnitts an einem Beispiel demonstrieren, wie man ein beliebiges Polynom durch Werte von Standard SO_2-Tableaus repräsentieren kann.

Beispiel A.1.33. Wir betrachten den Polynomring $P = K[x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2}, x_{3,1}, x_{3,2}]$ und das Polynom $f = x_{1,1}x_{2,1}x_{3,1}x_{3,2} + x_{1,1}x_{2,2}x_{3,2}^2$. Mit $\mathcal{M} = \left(\begin{smallmatrix} x_1 & e_2 \\ x_3 & e_1 \end{smallmatrix}\right)$ und $W = (x_3,\ x_2)$ gilt $f = -|\mathcal{M}|\langle W\rangle$, wobei hier ohne Einschränkung die Zeilen von $\mathcal{M}$ bereits anhand der ersten Spalte geordnet wurden. Wie man sofort sieht, ist $(\mathcal{M}, W)$ kein Standard SO_2-Tableau, da sowohl die Ordnung auf der zweiten Spalte von $\mathcal{M}$, als auch auf W nicht den Anforderungen entspricht. Allerdings ist $(\mathcal{M}, W)$ ein Element von Tab'. Wir wollen nun $|\mathcal{M}|\langle W\rangle$ mit Werten von Standard SO_2-Tableaus darstellen.

Dazu wenden wir zunächst Schritt (1) des Beweises von Lemma A.1.30 an und erhalten

$$\begin{vmatrix} x_1 & e_2 \\ x_3 & e_1 \end{vmatrix} = \begin{vmatrix} x_1 & e_1 \\ x_3 & e_2 \end{vmatrix} - \begin{vmatrix} x_1 & x_3 \\ e_1 & e_2 \end{vmatrix}.$$

Setze $\mathcal{M}_1^{(1)} := \begin{pmatrix} x_1 & e_1 \\ x_3 & e_2 \end{pmatrix}$ und $\mathcal{M}_2^{(1)} := \begin{pmatrix} x_1 & x_3 \end{pmatrix}$ sowie $W_1^{(1)} := W$ und $W_2^{(1)} := W$. Dann gilt:

$$|\mathcal{M}|\langle W\rangle = |\mathcal{M}_1^{(1)}|\langle W_1^{(1)}\rangle - |\mathcal{M}_2^{(1)}|\langle W_2^{(1)}\rangle.$$

Beide Matrizen sind Standard $\mathcal{H}$-Matrizen, jedoch ist weder $(\mathcal{M}_1^{(1)}, W_1^{(1)})$, noch $(\mathcal{M}_2^{(1)}, W_2^{(1)})$ ein Standard SO_2-Tableau.

Die fehlende Ordnung auf $W_1^{(1)}$ bzw. $W_2^{(1)}$ ist natürlich leicht zu korrigieren. Dazu setzen wir $W_1^{(2)} := (x_2,\ x_3)$ bzw. $W_2^{(2)} := (x_2,\ x_3)$ sowie $\mathcal{M}_1^{(2)} := \mathcal{M}_1^{(1)}$ und $\mathcal{M}_2^{(2)} := \mathcal{M}_2^{(1)}$. Dann gilt unverändert $|\mathcal{M}|\langle W\rangle = |\mathcal{M}_1^{(2)}|\langle W_1^{(2)}\rangle - |\mathcal{M}_2^{(2)}|\langle W_2^{(2)}\rangle$. Nun ist $(\mathcal{M}_2^{(2)}, W_2^{(2)})$ ein Standard SO_2-Tableau, $(\mathcal{M}_1^{(2)}, W_1^{(2)})$ jedoch nicht.

Da sowohl $\mathcal{M}_1^{(2)}$, als auch $W_1^{(2)}$ bereits standard sind, wenden wir Schritt (3) des Beweises von Lemma A.1.30 auf $(\mathcal{M}_1^{(2)}, W_1^{(2)})$ an. Demnach gilt:

$$|\mathcal{M}_1^{(2)}|\langle W_1^{(2)}\rangle \quad = \quad \begin{vmatrix} x_1 & e_1 \\ x_3 & e_2 \end{vmatrix}\langle (x_2, x_3)\rangle = -\begin{vmatrix} x_1 & e_2 \\ x_2 & x_3 \\ x_3 & e_2 \end{vmatrix} + |x_2\ x_3|\langle (x_1, x_3)\rangle + \begin{vmatrix} x_1 & e_1 \\ x_2 & e_2 \end{vmatrix}\langle (x_3, x_3)\rangle$$

Setze $\mathcal{M}_1^{(3)} := \left(\begin{smallmatrix} x_1 & e_2 \\ x_2 & x_3 \\ x_3 & e_2 \end{smallmatrix}\right)$, $W_1^{(3)} := \emptyset$, $\mathcal{M}_2^{(3)} := \left(\begin{smallmatrix} x_2 & x_3 \end{smallmatrix}\right)$, $W_2^{(3)} := (x_1, x_3)$ und $\mathcal{M}_3^{(3)} := \left(\begin{smallmatrix} x_1 & e_1 \\ x_2 & e_2 \end{smallmatrix}\right)$, $W_3^{(3)} := (x_3, x_3)$. Dann gilt also

$$|\mathcal{M}_1^{(2)}|\langle W_1^{(2)}\rangle = -|\mathcal{M}_1^{(3)}|\langle W_1^{(3)}\rangle + |\mathcal{M}_2^{(3)}|\langle W_2^{(3)}\rangle + |\mathcal{M}_3^{(3)}|\langle W_3^{(3)}\rangle.$$

Dabei ist nur $(\mathcal{M}_3^{(3)}, W_3^{(3)})$ ein Standard SO_2-Tableau.

Betrachten wir weiter zuerst $(\mathcal{M}_1^{(3)}, W_1^{(3)})$. Wir wenden hier Schritt (1) des Beweises von Lemma A.1.30 auf die ersten beiden Zeilen von $\mathcal{M}_1^{(3)}$ an und erhalten:

$$\begin{vmatrix} x_1 & e_2 \\ x_2 & x_3 \end{vmatrix} = \begin{vmatrix} x_1 & x_3 \\ x_2 & e_2 \end{vmatrix} - \begin{vmatrix} x_1 & x_2 \\ x_3 & e_2 \end{vmatrix}$$

Setze $\mathcal{M}_1^{(4)} := \left(\begin{smallmatrix} x_1 & x_3 \\ x_2 & e_2 \\ x_3 & e_2 \end{smallmatrix}\right)$, $W_1^{(4)} := W_1^{(3)} = \emptyset$ und $\mathcal{M}_2^{(4)} := \left(\begin{smallmatrix} x_1 & x_2 \\ x_3 & e_2 \\ x_3 & e_2 \end{smallmatrix}\right)$, $W_2^{(4)} := W_1^{(3)} = \emptyset$. Dann gilt

$$|\mathcal{M}_1^{(3)}|\langle W_1^{(3)}\rangle = |\mathcal{M}_1^{(4)}|\langle W_1^{(4)}\rangle - |\mathcal{M}_2^{(4)}|\langle W_2^{(4)}\rangle$$

und sowohl $(\mathcal{M}_1^{(4)}, W_1^{(4)})$, als auch $(\mathcal{M}_2^{(4)}, W_2^{(4)})$ ist ein Standard SO_2-Tableau.

Betrachten wir nun weiter $(\mathcal{M}_2^{(3)}, W_2^{(3)})$. Hier lässt sich Schritt (3) des Beweises von Lemma A.1.30 anwenden. Demnach gilt:

$$\begin{aligned} |\mathcal{M}_2^{(3)}|\langle W_2^{(3)}\rangle &= |x_2\ x_3|\langle (x_1, x_3)\rangle \\ &= -|x_1\ x_2|\langle (x_3, x_3)\rangle + |x_1\ x_3|\langle (x_3, x_2)\rangle. \end{aligned}$$

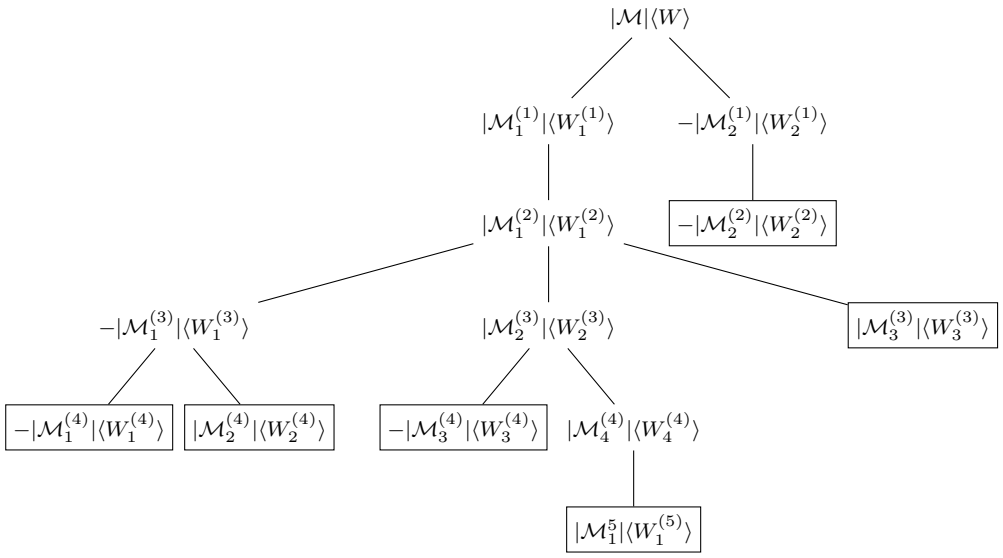
Setze $\mathcal{M}_3^{(4)} := \begin{pmatrix} x_1 & x_2 \end{pmatrix}$, $W_3^{(4)} := (x_3, x_3)$ und $\mathcal{M}_4^{(4)} := \begin{pmatrix} x_1 & x_3 \end{pmatrix}$, $W_4^{(4)} := (x_3, x_2)$. Dann gilt:

$$|\mathcal{M}_2^{(3)}|\langle W_2^{(3)}\rangle = -|\mathcal{M}_3^{(4)}|\langle W_3^{(4)}\rangle + |\mathcal{M}_4^{(4)}|\langle W_4^{(4)}\rangle.$$

Mit $\mathcal{M}_1^{(5)} := \mathcal{M}_4^{(4)}$ und $W_1^{(5)} := (x_2, x_3)$ erhalten wir eine Darstellung von $|\mathcal{M}_2^{(3)}|\langle W_2^{(3)}\rangle$ durch Standard SO_2-Tableaus, d.h. es gilt:

$$|\mathcal{M}_2^{(3)}|\langle W_2^{(3)}\rangle = -|\mathcal{M}_3^{(4)}|\langle W_3^{(4)}\rangle + |\mathcal{M}_1^{(5)}|\langle W_1^{(5)}\rangle.$$

Wir fassen zur besseren Übersicht die einzelnen Ersetzungsschritte in folgendem Baumdiagramm zusammen, wobei die entsprechenden Koeffizienten mitberücksichtigt wurden:



Durch Aufsummieren der Blätter dieses Baumes erhalten wir eine Darstellung von $|\mathcal{M}|\langle W\rangle$ durch Werte von Standard SO_2-Tableaus. Dabei gilt $|\mathcal{M}_2^{(2)}|\langle W_2^{(2)}\rangle = |\mathcal{M}_1^{(5)}|\langle W_1^{(5)}\rangle$. Somit folgt schließlich:

$$\begin{aligned}
|\mathcal{M}|\langle W\rangle &= -|\mathcal{M}_1^{(4)}|\langle W_1^{(4)}\rangle + |\mathcal{M}_2^{(4)}|\langle W_2^{(4)}\rangle - |\mathcal{M}_3^{(4)}|\langle W_3^{(4)}\rangle + |\mathcal{M}_1^{(5)}|\langle W_1^{(5)}\rangle \\
&\quad -|\mathcal{M}_2^{(2)}|\langle W_2^{(2)}\rangle + |\mathcal{M}_3^{(3)}|\langle W_3^{(3)}\rangle \\
&= -|\mathcal{M}_1^{(4)}|\langle W_1^{(4)}\rangle + |\mathcal{M}_2^{(4)}|\langle W_2^{(4)}\rangle - |\mathcal{M}_3^{(4)}|\langle W_3^{(4)}\rangle + |\mathcal{M}_3^{(3)}|\langle W_3^{(3)}\rangle \\
&= -\begin{vmatrix} x_1 & x_3 \\ x_2 & e_2 \\ x_3 & e_2 \end{vmatrix}\langle\emptyset\rangle + \begin{vmatrix} x_1 & x_2 \\ x_3 & e_2 \\ x_3 & e_2 \end{vmatrix}\langle\emptyset\rangle - |x_1\ x_2|\langle(x_3,x_3)\rangle + \begin{vmatrix} x_1 & e_1 \\ x_2 & e_2 \end{vmatrix}\langle(x_3,x_3)\rangle \\
&= -(x_{11}x_{32} - x_{12}x_{31})x_{21}x_{31} + (x_{11}x_{22} - x_{12}x_{21})x_{31}^2 \\
&\quad -(x_{11}x_{22} - x_{12}x_{21})(x_{31}^2 + x_{32}^2) - x_{12}x_{21}(x_{31}^2 + x_{32}^2) \\
&= -x_{11}x_{21}x_{31}x_{32} + x_{12}x_{21}x_{31}^2 + x_{11}x_{22}x_{31}^2 - x_{12}x_{21}x_{31}^2 \\
&\quad -x_{11}x_{22}x_{31}^2 - x_{11}x_{22}x_{32}^2 + x_{12}x_{21}x_{31}^2 + x_{12}x_{21}x_{32}^2 - x_{12}x_{21}x_{31}^2 - x_{12}x_{21}x_{32}^2 \\
&= -x_{11}x_{21}x_{31}x_{32} - x_{11}x_{22}x_{32}^2 \\
&= -f
\end{aligned}$$

◁

A.2 Die Vektorinvarianten von SO_2

Nach wie vor seien die $m{\cdot}n$ Unbestimmten $x_{1,1},\ldots,x_{m,n}$ aus P in einer $m\times n$-Matrix $\mathcal{X} := (x_{i,j})$ angeordnet. Die Gruppenoperation von $\mathrm{SO}_n := \mathrm{SO}_n(K)$ auf V wird induziert durch die lineare Darstellung $\rho : \mathrm{SO}_n \to \mathrm{Aut}_K(V)$ definiert durch $\rho_{\mathcal{A}}(\mathcal{B}) = \mathcal{B}\cdot\mathcal{A}$. Für ein $\mathcal{A}\in\mathrm{SO}_n$ lässt sich diese Operation also durch $\mathcal{X}\cdot\mathcal{A}$ beschreiben. Die i-te Zeile von $\mathcal{X}\cdot\mathcal{A}$ bezeichnen wir dann kurz mit $x_i^{\mathcal{A}}$ und mit $x_{i,j}^{\mathcal{A}}$ den (i,j)-ten Eintrag von $\mathcal{X}\cdot\mathcal{A}$. Somit operiert SO_n auf $f\in P$ durch $f^{\mathcal{A}} := f(x_{1,1}^{\mathcal{A}},\ldots,x_{i,j}^{\mathcal{A}})$, also durch Auswerten der Polynome des Koordinatenrings an den Stellen $x_{i,j}^{\mathcal{A}}$. Zunächst werden wir nachweisen, dass die n-Minoren von $\mathcal{X}$, die Einträge der Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ sowie alle k-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ für $k\in\{1,\ldots,n\}$ invariant sind unter der Operation von SO_n.

Lemma A.2.1. *Sei $G = \mathrm{SO}_n(K)$ und sei $m \in \mathbb{N}_+$. Dann sind die n-Minoren von $\mathcal{X}$ und die k-Minoren der symmetrischen $m \times m$-Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$ für $k \in \{1, \ldots, n\}$ Elemente von P^G.*

Beweis: Da für $n > m$ sowie für $k > m$ die n-Minoren von $\mathcal{X}$ sowie die k-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ alle Null sind (vgl. Satz A.1.2), können wir ohne Einschränkung $m \geq n$ annehmen.

Sei $\mathcal{G} \in \mathrm{SO}_n$. Für $i_1, \ldots, i_n \in \{1, \ldots, m\}$ sei $\mathcal{F}$ die $n \times n$-Teilmatrix $\mathcal{X}(i_1, \ldots, i_n | 1, \ldots, n)$ von $\mathcal{X}$ (Notation siehe Anhang A). Dann ist $f := \det(\mathcal{F})$ ein n-Minor von $\mathcal{X}$ und es gilt

$$f^{\mathcal{G}} = (\det(\mathcal{F}))^{\mathcal{G}} = \det(\mathcal{F}^{\mathcal{G}}) = \det(\mathcal{F} \cdot \mathcal{G}) = \det(\mathcal{F}) \cdot \det(\mathcal{G}) = \det(\mathcal{F}) = f.$$

Sei $k \in \{1, \ldots, n\}$ und seien $i_1, \ldots, i_k, j_1, \ldots, j_k \in \{1, \ldots, m\}$. Sei nun $\mathcal{M} := \mathcal{X}\mathcal{X}^{\mathrm{tr}}$ und sei $\mathcal{F}$ die $k \times k$-Teilmatrix $\mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k)$ von $\mathcal{M}$. Dann ist $f := \det(\mathcal{F})$ ein k-Minor von $\mathcal{M}$. Gemäß Abschnitt A.1.1 gibt es $k \times n$-Teilmatrizen $\mathcal{A}$ und $\mathcal{B}$ von $\mathcal{X}$ mit

$$\mathcal{F} = \mathcal{M}(i_1, \ldots, i_k | j_1, \ldots, j_k) = \mathcal{A} \cdot \mathcal{B}^{\mathrm{tr}}.$$

Dann sind $\mathcal{A}^{\mathcal{G}} := \mathcal{A} \cdot \mathcal{G}$ und $\mathcal{B}^{\mathcal{G}} := \mathcal{B} \cdot \mathcal{G}$ jeweils $k \times n$-Teilmatrizen von $\mathcal{X}^{\mathcal{G}} = \mathcal{X} \cdot \mathcal{G}$. Somit folgt:

$$\begin{aligned} f^{\mathcal{G}} &= \det(\mathcal{F})^{\mathcal{G}} = \det(\mathcal{A} \cdot \mathcal{B}^{\mathrm{tr}})^{\mathcal{G}} = \det(\mathcal{A})^{\mathcal{G}} \cdot \det(\mathcal{B}^{\mathrm{tr}})^{\mathcal{G}} = \det(\mathcal{A})^{\mathcal{G}} \cdot \det(\mathcal{B})^{\mathcal{G}} \\ &= \det(\mathcal{A} \cdot \mathcal{G}) \cdot \det(\mathcal{B} \cdot \mathcal{G}) = \det(\mathcal{A}) \cdot \det(\mathcal{G}) \cdot \det(\mathcal{B}) \cdot \det(\mathcal{G}) = \det(\mathcal{A}) \cdot \det(\mathcal{B}) \\ &= \det(\mathcal{A}) \cdot \det(\mathcal{B}^{\mathrm{tr}}) = \det(\mathcal{A} \cdot \mathcal{B}^{\mathrm{tr}}) = \det(\mathcal{F}) = f, \end{aligned}$$

also insgesamt die Behauptung. □

Der erste entscheidende Schritt hin zu einem Erzeugendensystem ist der Nachweis, dass die Werte von Standard SO_2-Tableaus, in denen die Einträge $e_1 = (1, 0)$ und $e_2 = (0, 1)$ nicht vorkommen, gerade die invarianten Polynome sind. Die Anzahl der Einträge e_1 wird dabei mit $e_1(\mathcal{M})$ und die der Einträge e_2 mit $e_2(\mathcal{M})$ bezeichnet. Um diese erste Etappe erreichen zu können, werden wir sehr viel detaillierter vorgehen als RICHMAN und einiges exakter aufbereiten sowie kleinere Fehler und Ungenauigkeiten in [Ric89] korrigieren. Wir werden uns zuerst eine spezielle Art der Darstellung für die Elemente der Gruppe $\mathrm{SO}_2(K)$ betrachten, die für den weiteren Verlauf eine große Rolle spielen wird.

Bemerkung A.2.2. Für alle $t \in K$ mit $t^2 + 1 \neq 0$ ist die 2×2-Matrix

$$\mathcal{A}(t) := \frac{1}{t^2 + 1} \begin{pmatrix} 1 - t^2 & -2t \\ 2t & 1 - t^2 \end{pmatrix}$$

ein Element von $\mathrm{SO}_2(K)$. Umgekehrt gibt es für jede Matrix $\mathcal{A} \in \mathrm{SO}_2(K) \setminus \{\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\}$ ein $t \in K$ mit $t^2 + 1 \neq 0$ und $\mathcal{A} = \mathcal{A}(t)$.

Falls wir im Folgenden eine Element von $\mathrm{SO}_2(K) \setminus \{\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\}$ in der Form aus Bemerkung A.2.2 betrachten, schreiben wir zur Verdeutlichung $\mathcal{A}(t)$ anstatt $\mathcal{A}$. Wie man leicht sieht, lässt sich für die Matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ keine derartige Darstellung finden, obwohl sie ein Element von $\mathrm{SO}_2(K)$ ist. Denn dazu müsste man ein $t \in K$ bestimmen mit $1 - t^2 = -1$ und $2t = 0$. Allerdings lässt sich die Matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ als Produkt von zwei Matrizen $\mathcal{A}, \mathcal{B} \in \mathrm{SO}_2(K) \setminus \{\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\}$ schreiben. Dank der Darstellung der Elemente der Gruppe SO_2 in der Form aus Bemerkung A.2.2 können wir die spezielle orthogonale Gruppe SO_2 auch wie folgt charakterisieren:

$$\mathrm{SO}_2(K) = \left\{ \frac{1}{t^2 + 1} \begin{pmatrix} 1 - t^2 & -2t \\ 2t & 1 - t^2 \end{pmatrix} : \ t \in K, \ t^2 + 1 \neq 0 \right\} \cup \left\{ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

Wir führen nun eine zusätzliche Unbestimmte z ein, womit wir die Elemente der Gruppe $\mathrm{SO}_2(K)$ allgemein durch die Matrix

$$\mathcal{A}(z) := \frac{1}{z^2+1}\begin{pmatrix} 1-z^2 & -2z \\ 2z & 1-z^2 \end{pmatrix}$$

beschreiben können. Die Matrix $\mathcal{A}(z)$ ist eine 2×2-Matrix mit Einträgen aus dem Funktionenkörper

$$K(z) = \left\{ \frac{f}{g} : \ f, g \in K[z], g \neq 0 \right\},$$

d.h. es gilt $\mathcal{A}(z) \in \mathrm{Mat}_2(K(z))$. Durch Auswertung der rationalen Funktionen $\frac{1-z^2}{z^2+1}$ und $\frac{2z}{z^2+1}$ an allen Punkte $t \in K$ mit $t^2+1 \neq 0$ erhalten wir mit Ausnahme der Matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ jede Matrix von $\mathrm{SO}_2(K)$. Da wir, wie bereits erwähnt, die Matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ als Produkt zweier Elemente von $\mathrm{SO}_2(K) \setminus \{\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\}$ schreiben können, reicht es nur die Auswertungen der Matrix $\mathcal{A}(z)$ zu betrachten. Mit Hilfe der Matrix $\mathcal{A}(z)$ können wir die Operation der Gruppe SO_2 auf P allgemein beschreiben. Sei wie üblich $x_i = (x_{i,1}, x_{i,2})$ für $i \in \{1, \ldots, m\}$ die i-te Zeile von $\mathcal{X}$. Dann gilt:

$$\begin{aligned} x_i^{\mathcal{A}(z)} &= (x_{i,1}, x_{i,2}) \cdot \mathcal{A}(z) = (x_{i,1}, x_{i,2}) \cdot \begin{pmatrix} \frac{1-z^2}{z^2+1} & -\frac{2z}{z^2+1} \\ \frac{2z}{z^2+1} & \frac{1-z^2}{z^2+1} \end{pmatrix} \\ &= \left(\frac{1-z^2}{z^2+1} x_{i,1} + \frac{2z}{z^2+1} x_{i,2}, \ -\frac{2z}{z^2+1} x_{i,1} + \frac{1-z^2}{z^2+1} x_{i,2} \right) \end{aligned}$$

im Einzelnen also

$$x_{i,1}^{\mathcal{A}(z)} = \frac{1-z^2}{z^2+1} x_{i,1} + \frac{2z}{z^2+1} x_{i,2}, \qquad x_{i,2}^{\mathcal{A}(z)} = -\frac{2z}{z^2+1} x_{i,1} + \frac{1-z^2}{z^2+1} x_{i,2}. \tag{A.2.1}$$

Wir erhalten dadurch Polynome aus $P' := K(z)[x_{1,1}, x_{1,2}, \ldots, x_{m,1}, x_{m,2}]$ und folglich einen K-Algebra-Homomorphismus $\Psi : P \to P'$ definiert durch $x_{i,j} \mapsto x_{i,j}^{\mathcal{A}(z)}$.

Wir geben nun einige nützliche Rechenregeln an, die die Operation von $\mathrm{SO}_2(K)$ auf spezielle Polynome beschreiben. Bei den betrachteten Polynomen handelt es sich um Werte spezieller SO_2-Tableaus, die wir im weiteren Verlauf benötigen werden. Da sich diese Regeln unmittelbar nachrechnen lassen, verzichten wir an dieser Stelle auf einen Nachweis.

Lemma A.2.3. *Sei $\mathcal{A}(z)$ wie oben, seien $i, j \in \{1, \ldots, m\}$ sowie e_1, e_2 die Standardbasisvektoren von P^2. Dann gilt:*

(i) $|x_i \ x_j|^{\mathcal{A}(z)} = |x_i \ x_j|$.

(ii) $\langle (x_i, x_j) \rangle^{\mathcal{A}(z)} = \langle (x_i, x_j) \rangle$.

(iii) $|x_i \ e_1|^{\mathcal{A}(z)} = \frac{2z}{1+z^2}|x_i \ e_2| + \frac{1-z^2}{1+z^2}|x_i \ e_1|$.

(iv) $|x_i \ e_2|^{\mathcal{A}(z)} = \frac{1-z^2}{1+z^2}|x_i \ e_2| - \frac{2z}{1+z^2}|x_i \ e_1|$.

Mit diesen Rechenregeln werden wir nun die Operation von $\mathcal{A}(z)$ auf den Wert eines beliebigen Standard SO_2-Tableaus $(\mathcal{M}, W)$ (siehe Definition A.1.19) untersuchen. Teil (ii) des vorausgegangen Lemmas sagt uns, dass wir uns dabei auf den Wert von $\mathcal{M}$ beschränken können. Für den Wert dieser Matrix kennen wir im Falle von $e_1(\mathcal{M}) = e_2(\mathcal{M}) = 0$ auch sofort die Antwort; in diesem Fall gilt $|\mathcal{M}|^{\mathcal{A}(z)} = |\mathcal{M}|$. Einen weiteren Sonderfall deckt das folgende Lemma ab. Diese Aussage findet sich in aller Kürze und ohne bewiesen zu werden im Beweis zu Proposition 13 in [Ric89].

Lemma A.2.4. *Sei $\mathcal{M}$ eine Standard $\mathcal{H}$-Matrix der Länge $l \in \mathbb{N}$ mit $\varepsilon_2 := e_2(\mathcal{M}) = l$, d.h. für $l > 0$ ist $\mathcal{M}$ von der Form $\begin{pmatrix} x_{a(1)} & e_2 \\ \vdots & \vdots \\ x_{a(l)} & e_2 \end{pmatrix}$ mit $a(1), \ldots, a(l) \in \{1, \ldots, m\}$. Dann gilt:*

$$|\mathcal{M}|^{\mathcal{A}(z)} = \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} \cdot |\mathcal{M}| + \sum_{k \geq 0} \sum_{j=0}^{\varepsilon_2} \frac{z^k}{(1+z^2)^j} h_{j,k}$$

mit Polynomen $h_{j,k} \in P$. Für $h_{j,k} \neq 0$ gilt insbesondere $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(h_{j,k})$.

Beweis: Gilt $l = 0$, also $\varepsilon_2 = 0$, so ist die Behauptung klar. Sei daher im Folgenden $l > 0$ bzw. $\varepsilon_2 > 0$.

Laut Lemma A.2.3 (iv) gilt für alle $i \in \{1, \ldots, l\}$:

$$|x_{a(i)}\ e_2|^{\mathcal{A}(z)} = \frac{1-z^2}{1+z^2}|x_{a(i)}\ e_2| - \frac{2z}{1+z^2}|x_{a(i)}\ e_1| = \frac{1}{1+z^2} \cdot \left((1-z^2)x_{a(i),1} + 2z x_{a(i),2}\right).$$

Damit folgt zunächst:

$$|\mathcal{M}|^{\mathcal{A}(z)} \quad = \quad \prod_{i=1}^{\varepsilon_2} |x_{a(i)}\ e_2|^{\mathcal{A}(z)} = \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} \cdot \prod_{i=1}^{\varepsilon_2} \left((1-z^2)x_{a(i),1} + 2z x_{a(i),2}\right)$$

Wir betrachten nun die l-Tupel $\mathcal{I}_1 := (x_{a(1),1}, \ldots, x_{a(l),1})$ sowie $\mathcal{I}_2 := (x_{a(1),2}, \ldots, x_{a(l),2})$, die alle in $\mathcal{M}$ auftretenden Unbestimmten der ersten sowie zweiten Spalte von $\mathcal{X}$ beinhalten, und bilden für alle $i \in \{0, \ldots, \varepsilon_2\}$ wie folgt Mengen T_i von Termen: Es gilt $t \in T_i$ genau dann, wenn es eine i-elementige Teilmenge $\{k_1, \ldots, k_i\}$ von $\{1, \ldots, l\}$ gibt mit

$$t = \mathcal{I}_1(k_1) \cdots \mathcal{I}_1(k_i) \cdot \mathcal{I}_2(k_{i+1}) \cdots \mathcal{I}_2(k_l).$$

Dabei sind $k_{i+1}, \ldots, k_l$ die zu $k_1, \ldots, k_i$ komplementären Indizes in $\{1, \ldots, l\}$ und $\mathcal{I}_1(j)$ bzw. $\mathcal{I}_2(j)$ ist der j-te Eintrag des l-Tupels $\mathcal{I}_1$ bzw. $\mathcal{I}_2$. Dann folgt:

$$|\mathcal{M}|^{\mathcal{A}(z)} \quad = \quad \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} \cdot \sum_{i=0}^{\varepsilon_2} (1-z^2)^i (2z)^{\varepsilon_2 - i} \cdot \sum_{t \in T_i} t$$

Insbesondere gilt $T_{\varepsilon_2} = \{x_{a(1),1} \cdots x_{a(l),1}\}$ und damit $\sum_{t \in T_{\varepsilon_2}} t = |\mathcal{M}|$. Setze $h_i := \sum_{t \in T_i} t$ für alle $i \in \{0, \ldots, \varepsilon_2\}$ und es folgt schließlich:

$$\begin{aligned} |\mathcal{M}|^{\mathcal{A}(z)} &= \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} \cdot \sum_{i=0}^{\varepsilon_2} (1-z^2)^i (2z)^{\varepsilon_2 - i} \cdot h_i \\ &= \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\mathcal{M}| + \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} \sum_{i=0}^{\varepsilon_2 - 1} (1-z^2)^i (2z)^{\varepsilon_2 - i} \cdot h_i \end{aligned}$$

Setze $h := \sum_{i=0}^{\varepsilon_2 - 1} (1-z^2)^i (2z)^{\varepsilon_2 - i} \cdot h_i$. Dann ist h ein Polynom in $K[x_{1,1}, \ldots, x_{m,2}, z]$. Durch Anwendung des Divisions-Algorithmus (vgl. [KR00], Theorem 1.6.4) erhalten wir Polynome $q_0, \ldots, q_{\varepsilon_2}$ aus $K[x_{1,1}, \ldots, x_{m,2}, z]$ mit $h = \sum_{j=0}^{\varepsilon_2} (1+z^2)^j q_j$. Sei $a_j \in \mathbb{N}$ für alle $j \in \{0, \ldots, \varepsilon_2\}$ der höchste Exponent der Unbestimmten z in q_j. Erneut mit dem Divisions-Algorithmus folgt, dass es Polynome $h_{j,0}, \ldots, h_{j,a_j} \in P$ gibt mit

$$q_j = z^0 h_{j,0} + \ldots + z^{a_j} h_{j,a_j} = \sum_{k=0}^{a_j} z^k h_{j,k}.$$

Wir schreiben kurz nur $q_j = \sum_{k\geq 0} z^k h_{j,k}$. Damit erhalten wir die gewünschte Darstellung:

$$\begin{aligned}
|\mathcal{M}|^{\mathcal{A}(z)} &= \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\mathcal{M}| + \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} h \\
&= \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\mathcal{M}| + \left(\frac{1}{1+z^2}\right)^{\varepsilon_2} \sum_{j=0}^{\varepsilon_2} (1+z^2)^j q_j \\
&= \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\mathcal{M}| + \sum_{j=0}^{\varepsilon_2} \frac{1}{(1+z^2)^j} \sum_{k\geq 0} z^k h_{j,k} \\
&= \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\mathcal{M}| + \sum_{k\geq 0} \sum_{j=0}^{\varepsilon_2} \frac{z^k}{(1+z^2)^j} h_{j,k}
\end{aligned}$$

Die Zusatzbehauptung folgt sofort aus der Konstruktion der Polynome $h_{j,k}$ und dem Divisions-Algorithmus. □

Wir werden nun aufbauend auf diesem Spezialfall eine ähnliche Darstellung für den allgemeinen Fall beweisen, also für beliebige Standard $\mathcal{H}$-Matrizen von Standard SO_2-Tableaus. Damit erhalten wir auch automatisch eine analoge Darstellung für den Wert eines Standard SO_2-Tableaus, wie sie Richman in [Ric89] erwähnt hat. Dabei werden wir die Struktur von Standard $\mathcal{H}$-Matrizen von Standard SO_2-Tableaus ausnutzen. Da für derartige $\mathcal{H}$-Matrizen $e_1(\mathcal{M}) \leq 1$ gilt, sind diese allgemein von der Form

$$\mathcal{M} = \begin{pmatrix} x_{a(1)} & x_{b(1)} \\ \vdots & \vdots \\ x_{a(s)} & x_{b(s)} \\ x_{a(s+\varepsilon_1)} & e_1 \\ x_{a(s+\varepsilon_1+1)} & e_2 \\ \vdots & \vdots \\ x_{a(s+\varepsilon_1+\varepsilon_2)} & e_2 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} & =: \mathcal{T} \\ \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} & =: \mathcal{B} \end{matrix}$$

mit $\varepsilon_1 := e_1(\mathcal{M})$ und $\varepsilon_2 := e_2(\mathcal{M})$. Aufgrund der Ordnung in den Zeilen und Spalten gilt:

(i) $a(i) \leq a(j)$ für alle i, j mit $i < j$,

(ii) $a(i) < b(i)$ für alle i.

Wir können also jede Standard $\mathcal{H}$-Matrix eines Standard SO_2-Tableaus in drei Teile aufteilen. Damit gilt insbesondere $|\mathcal{M}| = |\mathcal{T}| \cdot |x_{a(s+\varepsilon_1)}\ e_1| \cdot |\mathcal{B}|$, wobei natürlich $\mathcal{T} = \emptyset$, $\mathcal{B} = \emptyset$ oder $\varepsilon_1 = 0$ möglich ist, was jeweils einen Wert 1 der entsprechenden Teilmatrix zur Folge hat. So gilt beispielsweise $\mathcal{T} = \emptyset$ im Falle von $m = 1$, d.h. wenn nur ein Punkt im K^n betrachtet wird. Diese Form verwenden wir nun, um eine zum vorherigen Lemma analoge Aussage zu beweisen.

Lemma A.2.5. *Sei $(\mathcal{M}, W)$ ein Standard SO_2-Tableau mit $\varepsilon_1 := e_1(\mathcal{M})$ und $\varepsilon_2 := e_2(\mathcal{M})$. Weiter sei $\widehat{\mathcal{M}}$ diejenige $\mathcal{H}$-Matrix, die aus $\mathcal{M}$ entsteht, indem jeder e_1-Eintrag durch e_2 ersetzt wird. Dann gilt:*

$$(|\mathcal{M}|\langle W\rangle)^{\mathcal{A}(z)} = \left(\frac{2z}{1+z^2}\right)^{\varepsilon_1} \left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2} |\widehat{\mathcal{M}}|\langle W\rangle + \sum_{k\geq 0} \sum_{j=0}^{\varepsilon_1+\varepsilon_2} \frac{z^k}{(1+z^2)^j} h_{j,k}$$

mit Polynomen $h_{j,k} \in P$. Für $h_{j,k} \neq 0$ gilt insbesondere $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(h_{j,k})$.

Beweis: Aus Lemma A.2.3 (iii) folgt sofort $\langle W\rangle^{\mathcal{A}(z)} = \langle W\rangle$. Somit reicht es, nur die Standard $\mathcal{H}$-Matrix $\mathcal{M}$ zu betrachten und zu zeigen, dass

$$|\mathcal{M}|^{\mathcal{A}(z)} = \left(\frac{2z}{1+z^2}\right)^{\varepsilon_1}\left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2}|\widehat{\mathcal{M}}| + \sum_{k\geq 0}\sum_{j=0}^{\varepsilon_1+\varepsilon_2}\frac{z^k}{(1+z^2)^j}\tilde{h}_{j,k}$$

gilt mit Polynomen $\tilde{h}_{j,k} \in P$. Mit Polynomen $h_{j,k} := \tilde{h}_{j,k}\cdot\langle W\rangle$ folgt dann die eigentliche Behauptung.

Seien also $\mathcal{T}$ und $\mathcal{B}$ wie oben. Für $\varepsilon_1 = \varepsilon_2 = 0$ ist die Behauptung klar und es gilt insbesondere $|\mathcal{M}|^{\mathcal{A}(z)} = |\mathcal{M}|$, was sofort aus Lemma A.2.3 folgt. Ist $\varepsilon_1 = 0$ und $\varepsilon_2 > 0$, so folgt wegen $|\mathcal{T}| \in P$ die Behauptung sofort mit Lemma A.2.4 aus

$$|\mathcal{M}|^{\mathcal{A}(z)} = |\mathcal{T}|^{\mathcal{A}(z)}\cdot|\mathcal{B}|^{\mathcal{A}(z)} = |\mathcal{T}|\cdot|\mathcal{B}|^{\mathcal{A}(z)}.$$

Gelte nun $\varepsilon_1 = 1$ und $\varepsilon_2 = 0$. Mit Lemma A.2.3 (iii) folgt dann:

$$\begin{aligned}
|\mathcal{M}|^{\mathcal{A}(z)} &= |\mathcal{T}|^{\mathcal{A}(z)}\cdot|x_{a(s+\varepsilon_1)}\ e_1|^{\mathcal{A}(z)} = |\mathcal{T}|\cdot|x_{a(s+\varepsilon_1)}\ e_1|^{\mathcal{A}(z)}\\
&= \frac{2z}{1+z^2}\underbrace{|\mathcal{T}|\cdot|x_{a(s+\varepsilon_1)}\ e_2|}_{=|\widehat{\mathcal{M}}|} + \frac{1-z^2}{1+z^2}|\mathcal{T}|\cdot|x_{a(s+\varepsilon_1)}\ e_1|\\
&= \frac{2z}{1+z^2}|\widehat{\mathcal{M}}| + \frac{z^2}{1+z^2}|\mathcal{T}|x_{a(s+\varepsilon_1)2} - \frac{1}{1+z^2}|\mathcal{T}|x_{a(s+\varepsilon_1)2}\\
&= \frac{2z}{1+z^2}|\widehat{\mathcal{M}}| + \sum_{k=0}^{2}\sum_{j=0}^{1}\frac{z^k}{(1+z^2)^j}\tilde{h}_{j,k}
\end{aligned}$$

mit $\tilde{h}_{1,0} := -|\mathcal{T}|x_{a(s+\varepsilon_1)2}$ und $\tilde{h}_{1,2} := |\mathcal{T}|x_{a(s+\varepsilon_1)2}$ sowie $\tilde{h}_{0,0} = \tilde{h}_{0,1} = \tilde{h}_{11} = \tilde{h}_{0,2} = 0$.

Es bleibt noch der allgemeine Fall zu zeigen. Sei also $\varepsilon_1 = 1$ und $\varepsilon_2 > 0$. Wir wenden Lemma A.2.4 auf $\mathcal{B}$ an und Lemma A.2.3 auf $|x_{a(s+\varepsilon_1)}\ e_1|$. Dann folgt:

$$\begin{aligned}
|\mathcal{M}|^{\mathcal{A}(z)} &= |\mathcal{T}|^{\mathcal{A}(z)}\cdot|x_{a(s+\varepsilon_1)}\ e_1|^{\mathcal{A}(z)}\cdot|\mathcal{B}|^{\mathcal{A}(z)}\\
&= |\mathcal{T}|\cdot|x_{a(s+\varepsilon_1)}\ e_1|^{\mathcal{A}(z)}\cdot\left(\left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2}|\mathcal{B}| + \sum_{k\geq 0}\sum_{j=0}^{\varepsilon_2}\frac{z^k}{(1+z^2)^j}\cdot h'_{j,k}\right)\\
&= |\mathcal{T}|\cdot\left(\frac{2z}{1+z^2}|x_{a(s+\varepsilon_1)}\ e_2| + \frac{1-z^2}{1+z^2}|x_{a(s+\varepsilon_1)}\ e_1|\right)\\
&\quad\cdot\left(\left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2}|\mathcal{B}| + \sum_{k\geq 0}\sum_{j=0}^{\varepsilon_2}\frac{z^k}{(1+z^2)^j}\cdot h'_{j,k}\right)\\
&= \left(\frac{2z}{1+z^2}\right)\left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2}\underbrace{|\mathcal{T}|\cdot|x_{a(s+\varepsilon_1)}\ e_2|\cdot|\mathcal{B}|}_{=|\widehat{\mathcal{M}}|} + \sum_{k\geq 0}\sum_{j=0}^{\varepsilon_1+\varepsilon_2}\frac{z^k}{(1+z^2)^j}\tilde{h}_{j,k}\\
&= \left(\frac{2z}{1+z^2}\right)^{\varepsilon_1}\left(\frac{1-z^2}{1+z^2}\right)^{\varepsilon_2}|\widehat{\mathcal{M}}| + \sum_{k\geq 0}\sum_{j=0}^{\varepsilon_1+\varepsilon_2}\frac{z^k}{(1+z^2)^j}\tilde{h}_{j,k}
\end{aligned}$$

mit Polynomen $\tilde{h}_{j,k} \in P$. Mit Lemma A.2.4 und wegen $x_{a(s+\varepsilon_1),1} >_{\mathrm{Lex}} x_{a(s+\varepsilon_2),2}$ folgt aus der Konstruktion der Polynome $\tilde{h}_{j,k}$ sofort $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(\tilde{h}_{j,k})$ für alle j,k mit $\tilde{h}_{j,k} \neq 0$. Wegen $\langle W\rangle \neq 0$ folgt somit $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(h_{j,k})$ für alle j,k aus $h_{j,k} = \tilde{h}_{j,k}\cdot\langle W\rangle$. □

Wir haben in den Beweisen der letzten beiden Lemmata insbesondere gesehen, dass es im Falle von $\varepsilon_1 + \varepsilon_2 > 0$ mindestens ein Polynom $h_{j,k}$ mit $h_{j,k} \neq 0$ gibt. Wie wir aus Satz A.1.32 bereits wissen, erzeugen die Werte der Standard SO_2-Tableaus den Polynomring P als K-Vektorraum. Somit ist klar, dass sich auch jedes invariante Polynom mit Werten von Standard SO_2-Tableaus schreiben lässt. RICHMAN zeigt nun sogar, dass sich die invarianten Polynome als Linearkombination von Werten derjeniger Standard SO_2-Tableaus darstellen lassen, deren $\mathcal{H}$-Matrizen weder e_1-, noch e_2-Einträge beinhalten (vgl. [Ric89], Proposition 13). Dies ist ganz besonders deshalb interessant, weil derartige Standard SO_2-Tableaus genau diejenigen sind, deren Werte invariante Polynome sind, wie aus dem nächsten Satz folgt, den wir nun zuerst zeigen wollen. Dieser Satz taucht in dieser Form nicht in [Ric89] auf, nur die Aussage des Satzes wird im Beweis von Proposition 13 erwähnt und verwendet.

Satz A.2.6. *Sei $(\mathcal{M}, W)$ ein Standard SO_2-Tableau. Der Wert $|\mathcal{M}|\langle W\rangle$ ist genau dann ein Element von P^{SO_2}, wenn $e_1(\mathcal{M}) = e_2(\mathcal{M}) = 0$ gilt.*

Beweis: Setze $\varepsilon_1 := e_1(\mathcal{M})$ und $\varepsilon_2 := e_2(\mathcal{M})$ sowie $P' := K[x_{11}, \ldots, x_{m2}, z]$. Gilt $\varepsilon_1 = \varepsilon_2 = 0$, so gilt entweder $\mathcal{M} = \emptyset$ oder $|\mathcal{M}|$ ist ein Produkt von 2-Minoren von $\mathcal{X}$. Außerdem ist bekanntlich entweder $W = \emptyset$ oder $\langle W\rangle$ ein Produkt von Einträgen der symmetrischen Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$. Laut Lemma A.2.1 sind sowohl die 2-Minoren von $\mathcal{X}$, als auch die Einträge von $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$ invariant unter den Operationen von SO_2. Somit ist $|\mathcal{M}|\langle W\rangle$ stets ein Produkt von invarianten Polynomen und es gilt $|\mathcal{M}|\langle W\rangle \in P^{\mathrm{SO}_2}$.

Sei nun $|\mathcal{M}|\langle W\rangle \in P^{\mathrm{SO}_2}$, d.h. $|\mathcal{M}|\langle W\rangle$ ist ein invariantes Polynom oder mit anderen Worten, es gilt $(|\mathcal{M}|\langle W\rangle)^{\mathcal{A}(z)} = |\mathcal{M}|\langle W\rangle$. Nun nehmen wir an, es gilt $\varepsilon_1 + \varepsilon_2 > 0$. Aus Lemma A.2.5 folgt mit $g := \sum_{k\geq 0} \sum_{j=0}^{\varepsilon_1+\varepsilon_2} z^k (1+z^2)^{\varepsilon_1+\varepsilon_2-j} h_{j,k} \in K[x_{1,1}, \ldots, x_{m,2}, z]$:

$$(|\mathcal{M}|\langle W\rangle)^{\mathcal{A}(z)} = \left(\frac{1}{1+z^2}\right)^{\varepsilon_1+\varepsilon_2} \cdot \left((2z)^{\varepsilon_1}(1-z^2)^{\varepsilon_2} \cdot |\widehat{\mathcal{M}}|\langle W\rangle + g\right)$$

Setze $f := (2z)^{\varepsilon_1}(1-z^2)^{\varepsilon_2} \cdot |\widehat{\mathcal{M}}|\langle W\rangle + g$. Division mit Rest des ersten Summanden durch $(1+z^2)$ liefert zunächst

$$(2z)^{\varepsilon_1}(1-z^2)^{\varepsilon_2} \cdot |\widehat{\mathcal{M}}|\langle W\rangle = q_{\varepsilon_2}(2z)^{\varepsilon_1}|\widehat{\mathcal{M}}|\langle W\rangle \cdot (1+z^2) + 2^{\varepsilon_2}(2z)^{\varepsilon_1}|\widehat{\mathcal{M}}|\langle W\rangle$$

und damit $\mathrm{NR}_{1+z^2}(f) = 2^{\varepsilon_1+\varepsilon_2} z^{\varepsilon_1} |\widehat{\mathcal{M}}|\langle W\rangle + \mathrm{NR}_{1+z^2}(g)$. Weiter ist $\mathrm{NR}_{1+z^2}(g)$ von der Form $\mathrm{NR}_{1+z^2}(g) = s \cdot z + t$ für eindeutig bestimmte Polynome $s, t \in P$. Somit gilt:

$$\mathrm{NR}_{1+z^2}(f) = 2^{\varepsilon_1+\varepsilon_2} z^{\varepsilon_1} |\widehat{\mathcal{M}}|\langle W\rangle + s \cdot z + t.$$

Wegen $\mathrm{LT}_{\mathrm{Lex}}(h_{j,k}) <_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle)$ für alle j, k mit $h_{j,k} \neq 0$ (vgl. Lemma A.2.5) gilt

$$\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s) \quad \text{bzw.} \quad \mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t) \qquad (*)$$

falls $s \neq 0$ bzw. $t \neq 0$ gilt. Wegen $\varepsilon_1 + \varepsilon_2 > 0$ und $(|\mathcal{M}|\langle W\rangle)^{\mathcal{A}} = |\mathcal{M}|\langle W\rangle$ gilt $f \in \langle 1+z^2\rangle \subseteq P'$, also $\mathrm{NR}_{1+z^2}(f) = 0$ und damit $2^{\varepsilon_1+\varepsilon_2} z^{\varepsilon_1} |\widehat{\mathcal{M}}|\langle W\rangle + s \cdot z + t = 0$. Wir unterscheiden deshalb folgende Fälle:

1. Fall: $\varepsilon_1 = 0$. Dann gilt $2^{\varepsilon_2}|\widehat{\mathcal{M}}|\langle W\rangle + t = 0$. Insbesondere gilt also $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle) = \mathrm{LT}_{\mathrm{Lex}}(t)$ im Widerspruch zu $(*)$.

2. Fall: $\varepsilon_1 = 1$. Dann gilt $2^{\varepsilon_1+\varepsilon_2}|\widehat{\mathcal{M}}|\langle W\rangle + s = 0$. Analog zum ersten Fall folgt im Widerspruch zu $(*)$ auch hier $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}|\langle W\rangle) = \mathrm{LT}_{\mathrm{Lex}}(s)$.

Somit folgt insgesamt $\varepsilon_1 + \varepsilon_2 = 0$, also die Behauptung. □

Aus diesem Satz folgt unmittelbar auch, dass die Standard $\mathcal{H}$-Matrizen $\mathcal{M}$ mit $e_1(\mathcal{M}) = 0$ und $e_2(\mathcal{M}) = 0$ genau die $\mathcal{H}$-Matrizen sind, deren Wert invariant unter SO_2 sind. Denn mit $W = \emptyset$ ist $(\mathcal{M}, W)$ ein Standard SO_2-Tableau. Dies schließt natürlich auch $\mathcal{M} = \emptyset$ mit ein. Nun können wir uns dem bedeutenderen Satz widmen. RICHMAN greift dazu nun die Idee des Beweises des letzten Satzes auf und zeigt unter Verwendung von Lemma A.2.5, dass sich jedes invariante Polynom $f \in P^{\mathrm{SO}_2}$ als Linearkombination von Werten von Standard SO_2-Tableaus darstellen lässt, deren Standard $\mathcal{H}$-Matrizen weder e_1- noch e_2-Einträge beinhalten (vgl. [Ric89], Proposition 13).

Satz A.2.7. *Der Invariantenring P^{SO_2} wird als K-Vektorraum von den Werten der Standard SO_2-Tableaus erzeugt, deren $\mathcal{H}$-Matrizen keine e_1- und e_2-Einträge beinhalten.*

Beweis: Für $f = 0$ ist die Behauptung klar. Sei also $f \in P^{\mathrm{SO}_2} \setminus \{0\}$, d.h. insbesondere gilt $f^{\mathcal{A}(z)} = f$, und setze $P' := K[x_{1,1}, \ldots, x_{m,2}, z]$. Aus Satz A.1.32 folgt, dass es paarweise verschiedene Standard SO_2-Tableaus $(\mathcal{M}_1, W_1), \ldots, (\mathcal{M}_k, W_k)$ und Koeffizienten $\alpha_1, \ldots, \alpha_k \in K$ gibt mit

$$f = \alpha_1 |\mathcal{M}_1|\langle W_1\rangle + \ldots + \alpha_k |\mathcal{M}_k|\langle W_k\rangle. \qquad (*)$$

Somit bleibt $e_1(\mathcal{M}_i) = e_2(\mathcal{M}_i) = 0$ zu zeigen für alle $i \in \{1, \ldots, k\}$. Setze $\varepsilon_{i,1} := e_1(\mathcal{M}_i)$ bzw. $\varepsilon_{i,2} := e_2(\mathcal{M}_i)$ für alle $i \in \{1, \ldots, k\}$ und $\varepsilon := \max\{\varepsilon_{1,1} + \varepsilon_{1,2}, \ldots, \varepsilon_{k,1} + \varepsilon_{k,2}\}$. Dann reicht es, $\varepsilon = 0$ zu zeigen.

Angenommen, es gilt $\varepsilon > 0$, d.h. es gibt mindestens eine Standard $\mathcal{H}$-Matrix in der Darstellung $(*)$ von f, die mindestens einen e_1- oder e_2-Eintrag in der zweiten Spalte enthält. Da alle SO_2-Tableaus in der Darstellung $(*)$ von f Standard SO_2-Tableaus sind, gilt insbesondere $\varepsilon_{i,1} \leq 1$ für alle $i \in \{1, \ldots, k\}$. Wir betrachten nun folgende disjunkte Indexmengen:

$$I_0 := \{i \in \{1, \ldots, k\} : \ \varepsilon_{i,1} = 0, \ \varepsilon_{i,2} = \varepsilon\} \text{ und } I_1 := \{i \in \{1, \ldots, k\} : \ \varepsilon_{i,1} = 1 \ \varepsilon_{i,2} = \varepsilon - 1\},$$

die zusammen die Menge derjenigen SO_2-Tableaus in der Darstellung von f identifizieren, deren $\mathcal{H}$-Matrix genau ε Einträge der Form e_1 bzw. e_2 enthält. Wir können damit f wie folgt darstellen:

$$f = \sum_{i \in I_0} \alpha_i |\mathcal{M}_i|\langle W_i\rangle + \sum_{i \in I_1} \alpha_i |\mathcal{M}_i|\langle W_i\rangle + \sum_{i \notin I_0 \cup I_1} \alpha_i |\mathcal{M}_i|\langle W_i\rangle. \qquad (**)$$

Wegen $\varepsilon > 0$ gilt $I_0 \cup I_1 \neq \emptyset$, d.h. es gibt insbesondere mindestens ein $i \in \{1, \ldots, k\}$ mit $\mathcal{M}_i \neq \emptyset$. Aus Lemma A.2.5 erhalten wir für alle $i \in I_0$:

$$(1 + z^2)^{\varepsilon} \left(|\mathcal{M}_i|\langle W_i\rangle\right)^{\mathcal{A}(z)} = (1 - z^2)^{\varepsilon} |\widehat{\mathcal{M}_i}|\langle W_i\rangle + \sum_{l \geq 0} \sum_{j=0}^{\varepsilon} z^l (1 + z^2)^{\varepsilon - j} h_{j,l}^{(i)}$$

und für alle $i \in I_1$:

$$(1 + z^2)^{\varepsilon} \left(|\mathcal{M}_i|\langle W_i\rangle\right)^{\mathcal{A}(z)} = 2z(1 - z^2)^{\varepsilon - 1} |\widehat{\mathcal{M}_i}|\langle W_i\rangle + \sum_{l \geq 0} \sum_{j=0}^{\varepsilon} z^l (1 + z^2)^{\varepsilon - j} h_{j,l}^{(i)}$$

mit Polynomen $h_{j,l}^{(i)} \in P$. Dabei gilt $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}_i}|\langle W_i\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(h_{j,l}^{(i)})$ für alle Indizes i, j, l mit $h_{j,l}^{(i)} \neq 0$. Setze $m_i := (1 + z^2)^{\varepsilon} (|\mathcal{M}_i|\langle W_i\rangle)^{\mathcal{A}(z)}$ für alle $i \in \{1, \ldots, k\}$.

Analog zum Beweis von Satz A.2.6 folgt $\mathrm{NR}_{1+z^2}(m_i) = 2^{\varepsilon} |\widehat{\mathcal{M}_i}|\langle W_i\rangle + s_i \cdot z + t_i$ für alle $i \in I_0$ bzw. $\mathrm{NR}_{1+z^2}(m_j) = 2^{\varepsilon} z |\widehat{\mathcal{M}_j}|\langle W_j\rangle + s_j \cdot z + t_j$ für alle $j \in I_1$, wobei s_i bzw. s_j und t_i bzw. t_j jeweils Polynome aus P sind. Ist $s_i \neq 0$ bzw. $t_i \neq 0$ für $i \in I_0$, so folgt

$$\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}_i}|\langle W_i\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s_i) \quad \text{bzw.} \quad \mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}_i}|\langle W_i\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t_i)$$

und analog

$$\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_j|\langle W_j\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s_j) \quad \text{bzw.} \quad \mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_j|\langle W_j\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t_j),$$

falls $s_j \neq 0$ bzw. $t_j \neq 0$ gilt für $j \in I_1$. Für alle $i \notin I_0 \cup I_1$ gilt $\mathrm{NR}_{1+z^2}(m_i) = 0$ wegen $\varepsilon_{i,1} + \varepsilon_{i,2} < \varepsilon$. Somit folgt insgesamt aus $(**)$:

$$\begin{aligned}
\mathrm{NR}_{1+z^2}\left((1+z^2)^{\varepsilon} f^{\mathcal{A}(z)}\right) &= \sum_{i\in I_0} \mathrm{NR}_{1+z^2}(\alpha_i m_i) + \sum_{j\in I_1} \mathrm{NR}_{1+z^2}(\alpha_j m_j) \\
&= \sum_{i\in I_0} \alpha_i \left(2^{\varepsilon}|\widehat{\mathcal{M}}_i|\langle W_i\rangle + s_i z + t_i\right) + \sum_{j\in I_1} \alpha_j \left(2^{\varepsilon}|\widehat{\mathcal{M}}_j|\langle W_j\rangle \cdot z + s_j z + t_j\right) \\
&= 2^{\varepsilon} \sum_{i\in I_0} \alpha_i |\widehat{\mathcal{M}}_i|\langle W_i\rangle + z \cdot 2^{\varepsilon} \sum_{j\in I_1} \alpha_j |\widehat{\mathcal{M}}_j|\langle W_j\rangle + z \cdot \sum_{l\in I_0\cup I_1} \alpha_l s_l + \sum_{l\in I_0\cup I_1} \alpha_l t_l
\end{aligned}$$

Setze $s := \sum_{l\in I_0\cup I_1} \alpha_l s_l$ und $t := \sum_{l\in I_0\cup I_1} \alpha_l t_l$. Dann sind s, t Polynome in P. Gilt $s \neq 0$ bzw. $t \neq 0$, so gibt es ein $l \in I_0 \cup I_1$ mit

$$\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}_l}|\langle W_l\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s) \quad \text{bzw.} \quad \mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}_l}|\langle W_l\rangle) >_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t). \qquad (\dagger)$$

Wegen $f^{\mathcal{A}(z)} = f$ gilt $f^{\mathcal{A}(z)} \in P$ und wegen $\varepsilon > 0$ folgt $\mathrm{NR}_{1+z^2}\left((1+z^2)^{\varepsilon} f^{\mathcal{A}(z)}\right) = 0$, d.h. es gilt:

$$2^{\varepsilon} \sum_{i\in I_0} \alpha_i \cdot |\widehat{\mathcal{M}}_i|\langle W_i\rangle + t = 0 \quad \text{und} \quad 2^{\varepsilon} \sum_{j\in I_1} \alpha_j \cdot |\widehat{\mathcal{M}}_j|\langle W_j\rangle + s = 0.$$

Da die Standard SO_2-Tableaus $(\mathcal{M}_1, W_1), \ldots, (\mathcal{M}_k, W_k)$ paarweise verschieden sind, sind auch die Standard SO_2-Tableaus $(\widehat{\mathcal{M}}_i, W_i)$ für alle $i \in I_0$ bzw. $(\widehat{\mathcal{M}}_j, W_j)$ für alle $j \in I_1$ jeweils paarweise verschieden. Aus Korollar A.1.24 folgt, dass die Leitterme der Werte der Standard SO_2-Tableaus $(\widehat{\mathcal{M}}_i, W_i)$ für alle $i \in I_0$ bzw. $(\widehat{\mathcal{M}}_j, W_j)$ für alle $j \in I_1$ ebenfalls paarweise verschieden sind.

Wegen $I_0 \cup I_1 \neq \emptyset$ unterscheiden wir folgende Fälle: Gilt $I_0 \neq \emptyset$ und $I_1 = \emptyset$, so gilt $s = 0$ und $t \neq 0$. Aus $2^{\varepsilon} \sum_{i\in I_0} \alpha_i |\widehat{\mathcal{M}}_i|\langle W_i\rangle + t = 0$ folgt $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_i|\langle W_i\rangle) \leq_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t)$ für alle $i \in I_0 \cup I_1$ im Widerspruch zu $(\dagger)$. Gilt $I_1 \neq \emptyset$ und $I_0 = \emptyset$, dann folgt analog $t = 0$ und $s \neq 0$ sowie $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_i|\langle W_i\rangle) \leq_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s)$ erneut im Widerspruch zu $(\dagger)$. Sei nun $I_0 \neq \emptyset$ und $I_1 \neq \emptyset$. Dann gilt $s \neq 0$ und $t \neq 0$. Somit folgt analog zu oben $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_i|\langle W_i\rangle) \leq_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(t)$ für alle $i \in I_0$ und $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_i|\langle W_i\rangle) \leq_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s)$ für alle $i \in I_1$. Sei o.B.d.A. $\mathrm{LT}_{\mathrm{Lex}}(t) \leq \mathrm{LT}_{\mathrm{Lex}}(s)$. Dann gilt $\mathrm{LT}_{\mathrm{Lex}}(|\widehat{\mathcal{M}}_i|\langle W_i\rangle) \leq_{\mathrm{Lex}} \mathrm{LT}_{\mathrm{Lex}}(s)$ für alle $i \in I_0 \cup I_1$ wiederum im Widerspruch zu $(\dagger)$. Somit erhalten wir $\varepsilon = 0$, womit die Behauptung bewiesen ist. □

Mit diesem Satz ist der Beweis des abschließenden Theorems, das ein Erzeugendensystem von P^{SO_2} als K-Unteralgebra von P angibt, sofort klar. Die Einträge in jeder Zeile einer Standard $\mathcal{H}$-Matrix sind verschieden. Zudem enthält eine nicht-leere Standard $\mathcal{H}$-Matrix $\mathcal{M}$ mit $e_1(\mathcal{M}) = 0$ und $e_2(\mathcal{M}) = 0$ nur Einträge aus $\mathcal{X}$, d.h. $|\mathcal{M}|$ ist ein Produkt von 2-Minoren von $\mathcal{X}$. Der Wert eines nicht-leeren Standard $\mathcal{H}$-Tupels W eines Standard SO_2-Tableaus $(\mathcal{M}, W)$ ist bekanntlich ein Produkt von Einträgen der Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$.

Theorem A.2.8. (Erzeugendensystem von P^{SO_2})
Sei K ein nicht-endlicher Körper mit $\mathrm{char}(K) \neq 2$ *und sei $P = K[x_{1,1}, x_{1,2}, \ldots, x_{m,1}, x_{m,2}]$. Dann wird der Invariantenring P^{SO_2} als K-Unteralgebra von P erzeugt von den 2-Minoren von $\mathcal{X}$ und von den Einträgen der Matrix $\mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}}$.*

Beweis: Setze $G := \mathrm{Min}(2, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$. Die Inklusion $K[G] \subseteq P^{\mathrm{SO}_2}$ folgt sofort aus Lemma A.2.1.

Sei $f \in P^{\mathrm{SO}_2} \setminus \{0\}$. Nach Satz A.2.7 gibt es Standard SO_2-Tableaus $(\mathcal{M}_1, W_1), \ldots, (\mathcal{M}_k, W_k)$ mit $e_1(\mathcal{M}_i) = e_2(\mathcal{M}_i) = 0$ für alle $i \in \{1, \ldots, k\}$ und Koeffizienten $\alpha_1, \ldots, \alpha_k \in K$ mit

$$f = \alpha_1 |\mathcal{M}_1| \langle W_1 \rangle + \ldots + \alpha_k |\mathcal{M}_k| \langle W_k \rangle.$$

Da für alle $i \in \{1, \ldots, k\}$ mit $\mathcal{M}_i \neq \emptyset$ der Wert $|\mathcal{M}_i|$ ein Produkt von Elementen von $\mathrm{Min}(2, \mathcal{X})$ und mit $W_i \neq \emptyset$ der Wert $\langle W_i \rangle$ ein Produkt von Elementen aus $\mathrm{Min}(1, \mathcal{X} \cdot \mathcal{X}^{\mathrm{tr}})$ ist, folgt $f \in K[G]$. □

Man beachte, dass für $m = 1$ die Menge $\mathrm{Min}(2, \mathcal{X})$ der 2-Minoren von $\mathcal{X}$ nur die Null enthält und somit nichts zum Erzeugendensystem beiträgt. Wir wollen zum Abschluss dieses Abschnitts noch der Frage nachgehen, ob die Menge der 2-Minoren von $\mathcal{X}$ zusammen mit den Einträgen der symmetrischen Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ den Invariantenring P^{SO_2} nicht nur erzeugt, sondern sogar eine Lex-SAGBI-Basis von P^{SO_2} bildet.

Korollar A.2.9. *Die Menge* $G := \mathrm{Min}(2, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$ *ist eine* Lex-*SAGBI-Basis von* P^{SO_2}.

Beweis: Gemäß Theorem 5.2.13 reicht es zu zeigen, dass die Menge $\{\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G\}$ das multiplikative Monoid $\{\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in P^{\mathrm{SO}_2} \setminus \{0\}\}$ erzeugt.

Sei $f \in P^{\mathrm{SO}_2} \setminus \{0\}$. Gemäß Satz A.2.7 gibt es paarweise verschiedene Standard SO_2-Tableaus $(\mathcal{M}_1, W_1), \ldots, (\mathcal{M}_k, W_k)$ mit $e_1(\mathcal{M}_i) = e_2(\mathcal{M}_i) = 0$ für alle $i \in \{1, \ldots, k\}$ und Koeffizienten $\alpha_1, \ldots, \alpha_k \in K$ mit $f = \alpha_1 |\mathcal{M}_1| \langle W_1 \rangle + \ldots + \alpha_k |\mathcal{M}_k| \langle W_k \rangle$.

Aus Korollar A.1.24 folgt, dass auch die Leitterme $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_1| \langle W_1 \rangle), \ldots, \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k| \langle W_k \rangle)$ der Werte dieser Standard SO_2-Tableaus paarweise verschieden sind. Somit gibt es genau ein Standard SO_2-Tableau, also genau ein $i \in \{1, \ldots, k\}$ mit $\mathrm{LT}_{\mathrm{Lex}}(f) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_i| \langle W_i \rangle)$. Bekanntlich ist $|\mathcal{M}_i|$ für $\mathcal{M}_i \neq \emptyset$ wegen $e_1(\mathcal{M}_i) = e_2(\mathcal{M}_i) = 0$ ein Produkt von 2-Minoren von $\mathcal{X}$ und $\langle W_i \rangle$ für $W_i \neq \emptyset$ ein Produkt von Einträgen der symmetrischen Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$, d.h. es gibt $g_1, \ldots, g_l \in G$ und $a_1, \ldots, a_l \in \mathbb{N}$ mit $|\mathcal{M}_i| \langle W_i \rangle = g_1^{a_1} \cdots g_l^{a_l}$. Damit folgt

$$\mathrm{LT}_{\mathrm{Lex}}(f) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_i| \langle W_i \rangle) = \mathrm{LT}_{\mathrm{Lex}}(g_1)^{a_1} \cdots \mathrm{LT}_{\mathrm{Lex}}(g_l)^{a_l}.$$

□

A.3 Verallgemeinerung auf SO_n

Nachdem wir im letzten Abschnitt ein Erzeugendensystem für die Invarianten der Gruppe SO_2 erhalten haben, wollen wir nun die Invarianten der Gruppe SO_n für beliebige $n > 2$ betrachten. Um die Korrektheit eines Erzeugendensystems von P^{SO_n} zu beweisen, führt Richman den allgemeinen Fall auf den Fall $n = 2$ zurück und verwendet zudem die Theorie der SAGBI-Basen. Sei also $n \in \mathbb{N}$ mit $n > 2$ nun beliebig. Einige Informationen über Invarianten im allgemeinen Fall haben wir bereits erhalten. So wurde in Lemma A.2.1 bewiesen, dass die n-Minoren der Matrix $\mathcal{X}$ sowie die k-Minoren der symmetrischen Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ für $k \in \{1, \ldots, n\}$ invariant sind. Wie wir in Lemma A.2.1 gezeigt haben, sind die Einträge von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ invariant, denn diese sind nichts anderes als 1-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$. Außerdem ist natürlich jeder k-Minor von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ ein Element der von den 1-Minoren von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ erzeugten Unteralgebra $K[\mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})]$ von P.

Wir wollen nun beginnen, die „Brücke" zwischen dem allgemeinen Fall mit beliebigem $n > 2$ und dem bereits bekannten Fall $n = 2$ zu bauen. Nach wie vor sei P der Polynomring in $n \cdot m$ Unbestimmten mit $n \geq 2$ und $m \in \mathbb{N}_+$. Um in manchen Situationen besser erkennen zu können, welches n bzw. welcher Polynomring genau betrachtet wird, aber nicht alle Unbestimmten ständig angeben zu müssen, werden wir P an entsprechenden Stellen auch wie folgt in Abhängigkeit von n schreiben: Für $n \geq 2$ bezeichnen wir mit $P(n)$ den Polynomring $K[x_{i1}, \ldots, x_{in} : \ i \in \{1, \ldots, m\}]$ in den $m \cdot n$ Unbestimmten $x_{11}, \ldots, x_{mn}$. Unter Verwendung dieser Notation betrachten wir im Folgenden für $k, l \in \{1, \ldots, n\}$ mit $k < l$ den K-Algebra-Homomorphismus $\Phi_{k,l} : P(n) \to P(2)$, der definiert ist durch

$$x_{i,j} \mapsto \begin{cases} x_{i,1}, & j = k, \\ x_{i,2}, & j = l, \\ 1, & \text{sonst.} \end{cases}$$

Für $m = 4$ und $n = 3$ sowie $t = x_{1,1}^2 x_{3,1} x_{1,2}^3 x_{2,2}^4 x_{3,3}^5 x_{4,3}$ gilt also beispielsweise:

$$\Phi_{2,3}(t) = 1^2 \cdot 1 \cdot x_{1,1}^3 \cdot x_{2,1}^4 x_{3,2}^5 x_{4,2} = x_{1,1}^3 x_{2,1}^4 x_{3,2}^5 x_{4,2}.$$

Für die folgenden Aussagen werden wir nun Terme $t \in \mathbb{T}^{mn}$ auf besondere Art und Weise darstellen. Um dies besser notieren zu können, betrachten wir für $k \in \{1, \ldots, n\}$ den K-Algebra-Homomorphismus $\varphi_k : P(n) \to P(n)$, definiert durch

$$\varphi_k(x_{i,j}) = \begin{cases} x_{i,j}, & j = k, \\ 1, & j \neq k. \end{cases}$$

Dann ist für jeden Term $t \in \mathbb{T}^{mn}$ und jedes $k \in \{1, \ldots, n\}$ das Bild $\varphi_k(t)$ ein Term aus $\mathbb{T}(x_{1k}, \ldots, x_{mk})$, also ein Term, der nur Unbestimmte der k-ten Spalte von $\mathcal{X}$ enthält. Setzen wir $t_k := \varphi_k(t)$ für $k \in \{1, \ldots, n\}$, so lässt sich t schreiben als Produkt $t = t_1 \cdots t_n$ von entsprechenden Termen. So gilt z.B. für den Term $t = x_{1,1}^2 x_{3,1} x_{1,2}^3 x_{2,2}^4 x_{3,3}^5 x_{4,3}$

$$\varphi_1(t) = x_{1,1}^2 x_{3,1}, \quad \varphi_2(t) = x_{1,2}^3 x_{2,2}^4, \quad \varphi_3(t) = x_{3,3}^5 x_{4,3}$$

und wir erhalten die Zerlegung $t = t_1 \cdot t_2 \cdot t_3$ mit $t_1 = x_{1,1}^2 x_{3,1}$, $t_2 = x_{1,2}^3 x_{2,2}^4$ und $t_3 = x_{3,3}^5 x_{4,3}$. Das folgende Lemma liefert uns erste Eigenschaften des K-Algebra-Homomorphismus $\Phi_{k,l}$, insbesondere den zweiten Teil des Lemmas werden wir im weiteren Verlauf benötigen.

Lemma A.3.1. *Seien $k, l \in \{1, \ldots, n\}$ mit $k < l$. Sei weiter $t \in \mathbb{T}^{mn}$.*

a) Es gilt $\deg(\Phi_{k,l}(t)) = \deg(\varphi_k(t)) + \deg(\varphi_l(t))$.

b) Für jedes Standard SO_2*-Tableau $(\mathcal{M}, W)$ mit $e_1(\mathcal{M}) = e_2(\mathcal{M}) = 0$ und der Eigenschaft $\Phi_{k,l}(t) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W \rangle)$ gilt:*

$$\deg(|\mathcal{M}|) = 2\deg(\varphi_l(t)) \quad \textit{und} \quad \deg(\langle W \rangle) = \deg(\varphi_k(t)) - \deg(\varphi_l(t)).$$

Beweis: Schreibe t in der Form $t = t_1 \cdots t_n$ mit $t_j := \varphi_j(t)$ für alle $j \in \{1, \ldots, n\}$. Dann gilt:

$$\Phi_{k,l}(t) = \Phi_{k,l}(t_k t_l) = \Phi_{k,l}(t_k) \cdot \Phi_{k,l}(t_l).$$

a) Aus $\deg(\Phi_{k,l}(t_k)) = \deg(\varphi_k(t))$ bzw. $\deg(\Phi_{k,l}(t_l)) = \deg(\varphi_l(t))$ folgt die Behauptung.

b) Es gilt $\Phi_{k,l}(t_k) \in \mathbb{T}(x_{11}, \ldots, x_{m1})$ und $\Phi_{k,l}(t_l) \in \mathbb{T}(x_{12}, \ldots, x_{m2})$. Mit

$$\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = \Phi_{k,l}(t) \quad \text{bzw.} \quad \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|\langle W\rangle) = \Phi_{k,l}(t_k) \cdot \Phi_{k,l}(t_l)$$

folgt aus Lemma A.1.21, dass es einen eindeutig bestimmten Term $\tilde{t} \in \mathbb{T}(x_{11}, \ldots, x_{m1})$ gibt mit

$$\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}|) = \tilde{t} \cdot \Phi_{k,l}(t_l) \quad \text{und} \quad \tilde{t} \cdot \mathrm{LT}_{\mathrm{Lex}}(\langle W\rangle) = \Phi_{k,l}(t_k).$$

Wegen $e_1(\mathcal{M}) = e_2(\mathcal{M}) = 0$ gilt $\deg(\tilde{t}) = \deg(\Phi_{k,l}(t_l)) = \deg(t_l)$. Somit folgt sofort $\deg(|\mathcal{M}|) = 2\deg(t_l)$ und $\deg(\langle W\rangle) = \deg(t_k) - \deg(t_l)$.

□

Wir werden nun den speziellen K-Algebra-Homomorphismus $\Phi_{k,k+1}$ für ein $k \in \{1, \ldots, n-1\}$ näher betrachten. Dieser Homomorphismus wird den Schlüssel dazu bilden, eine Lex-SAGBI-Basis, und damit ein Erzeugendensystem, von P^{SO_n} angeben zu können. Zunächst ist es damit möglich, die besagte „Brücke" zwischen $P(n)^{\mathrm{SO}_n}$ und $P(2)^{\mathrm{SO}_2}$ zu schlagen, indem wir zeigen, dass das Bild des Leitterms jedes invarianten Polynoms $f \in P(n)^{\mathrm{SO}_n}$ mit $f \neq 0$ unter $\Phi_{k,k+1}$ der Leitterm eines Polynoms $f_k \in P(2)^{\mathrm{SO}_2}$ mit $f_k \neq 0$ ist (nach [Ric89], Proposition 2).

Satz A.3.2. *Sei $f \in P(n)^{\mathrm{SO}_n}$ mit $f \neq 0$. Dann gibt es für alle $k \in \{1, \ldots, n-1\}$ ein Polynom $f_k \in P(2)^{\mathrm{SO}_2}$ mit $f_k \neq 0$ und $\mathrm{LT}_{\mathrm{Lex}}(f_k) = \Phi_{k,k+1}(\mathrm{LT}_{\mathrm{Lex}}(f))$.*

Beweis: Die Behauptung wird mit Induktion nach n bewiesen. Für $n = 2$ ist $\Phi_{1,2}$ die Identität und damit die Aussage trivial.

Sei nun $n > 2$. Wegen $f \neq 0$ ist $\mathrm{Supp}(f) \neq \emptyset$. Wir schreiben nun jeden Term $t \in \mathrm{Supp}(f)$ in der Form $t = t_1 \cdots t_n$ mit $t_j = \varphi_j(t)$ für alle $j \in \{1, \ldots, n\}$. Für einen Term $t = t_1 \cdots t_n \in \mathrm{Supp}(f)$ setzen wir $\tilde{t} := t_1 \cdots t_{n-1}$. Dann gilt $f = \sum_{t \in \mathrm{Supp}(f)} a_t \tilde{t} t_n$. Da für alle $\mathcal{A}_{n-1} \in \mathrm{SO}_{n-1}$ die Matrix $\mathcal{A}_n := \left(\begin{smallmatrix} \mathcal{A}_{n-1} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ ein Element von SO_n ist, gilt:

$$f^{\mathcal{A}_n} = \left(\sum_{t \in \mathrm{Supp}(f)} a_t \cdot \tilde{t} t_n \right)^{\mathcal{A}_n} = \sum_{t \in \mathrm{Supp}(f)} a_t \cdot (\tilde{t} t_n)^{\mathcal{A}_n} = \sum_{t \in \mathrm{Supp}(f)} a_t \cdot \tilde{t}^{\mathcal{A}_{n-1}} t_n$$

Die Menge $\mathbb{T}(x_{1n}, \ldots, x_{mn})$ der Terme in den Unbestimmten $x_{1n}, \ldots, x_{mn}$ bildet eine Basis des freien $P(n-1)$-Moduls $P(n)$. Damit folgt $\tilde{t}^{\mathcal{A}_{n-1}} = \tilde{t}$ für alle $t \in \mathrm{Supp}(f)$ mit $t = \tilde{t} t_n$ und alle $\mathcal{A}_{n-1} \in \mathrm{SO}_{n-1}$ aus $f^{\mathcal{A}_n} = f$, d.h. es gilt also $\tilde{t} \in P(n-1)^{\mathrm{SO}_{n-1}}$. Schreiben wir also insbesondere den Leitterm von f in der Form $\mathrm{LT}_{\mathrm{Lex}}(f) = \tilde{t} \cdot t_n$, so gilt auch hier $\tilde{t} \in P(n-1)^{\mathrm{SO}_{n-1}}$.

Für alle $k \in \{1, \ldots, n-2\}$ gilt $\Phi_{k,k+1}(\tilde{t}) = \Phi_{k,k+1}(\tilde{t} t_n)$. Somit gibt es laut Induktionsannahme für alle $k \in \{1, \ldots, n-2\}$ ein Polynom $f_k \in P(2)^{\mathrm{SO}_2} \setminus \{0\}$ mit

$$\mathrm{LT}_{\mathrm{Lex}}(f_k) = \Phi_{k(k+1)}(\mathrm{LT}_{\mathrm{Lex}}(\tilde{t})) = \Phi_{k(k+1)}(\tilde{t} \cdot t_n) = \Phi_{k(k+1)}(\mathrm{LT}_{\mathrm{Lex}}(f))$$

Setze nun $\tilde{t} := t_2 \cdots t_n$ für alle $t \in \mathrm{Supp}(f)$ mit $t = t_1 \cdots t_n$. Da analog für alle $\mathcal{A}_{n-1} \in \mathrm{SO}_{n-1}$ auch $\mathcal{A}_n := \left(\begin{smallmatrix} 1 & 0 \\ 0 & \mathcal{A}_{n-1} \end{smallmatrix}\right) \in \mathrm{SO}_n$ gilt, folgt hier:

$$f^{\mathcal{A}_n} = \left(\sum_{t \in \mathrm{Supp}(f)} a_t \cdot t_1 \tilde{t} \right)^{\mathcal{A}_n} = \sum_{t \in \mathrm{Supp}(f)} a_t \cdot (t_1 \tilde{t})^{\mathcal{A}_n} = \sum_{t \in \mathrm{Supp}(f)} a_t \cdot t_1 \tilde{t}^{\mathcal{A}_{n-1}}$$

Setzen wir $R := K[x_{12}, \ldots, x_{1n}, \ldots, x_{m2}, \ldots, x_{mn}]$, so folgt auf analoge Weise $\tilde{t} \in R^{\mathrm{SO}_{n-1}}$ für alle $t \in \mathrm{Supp}(f)$ mit $t = t_1\tilde{t}$. Inbesondere gilt also $\tilde{t} \in R^{\mathrm{SO}_{n-1}}$ für $\mathrm{LT}_{\mathrm{Lex}}(f) = t_1\tilde{t}$. Laut Induktionsannahme gibt es nun für alle $k \in \{2, \ldots, n-1\}$ ein Polynom $f_k \in P(2)^{\mathrm{SO}_2}$ mit

$$\mathrm{LT}_{\mathrm{Lex}}(f_k) = \Phi_{k(k+1)}(\mathrm{LT}_{\mathrm{Lex}}(\tilde{t})) = \Phi_{k(k+1)}(t_1\tilde{t}) = \Phi_{k(k+1)}(\mathrm{LT}_{\mathrm{Lex}}(f)).$$

Damit folgt die Behauptung für alle $k \in \{1, \ldots, n-1\}$. □

Unter Verwendung verschiedener Eigenschaften von Standard SO_2-Tableaus (vgl. die entsprechenden Sätze in Abschnitt A.1.3) lässt sich aus diesem Satz sofort folgern, dass es für alle invarianten Polynome $f \in P(n)^{\mathrm{SO}_n}$ auch ein Standard SO_2-Tableau gibt, dessen Standard $\mathcal{H}$-Matrix keine e_1- und e_2-Einträge enthält und das dieselben Eigenschaften besitzt wie das Polynom f_k des Satzes. Dieses Standard SO_2-Tableau ist sogar eindeutig bestimmt, wie aus dem folgenden Korollar hervorgeht.

Korollar A.3.3. *Sei $f \in P(n)^{\mathrm{SO}_n}$ mit $f \neq 0$. Dann gibt es für alle $k \in \{1, \ldots, n-1\}$ genau ein Standard SO_2-Tableau $(\mathcal{M}_k, W_k)$ mit $e_1(\mathcal{M}_k) = e_2(\mathcal{M}_k) = 0$ und $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle) = \Phi_{k,k+1}(\mathrm{LT}_{\mathrm{Lex}}(f))$.*

Beweis: Aus dem letzten Satz folgt, dass es für alle $k \in \{1, \ldots, n-1\}$ ein Polynom $f_k \in P(2)^{\mathrm{SO}_2}$ gibt mit $\mathrm{LT}_{\mathrm{Lex}}(f_k) = \Phi_{k,k+1}(\mathrm{LT}_{\mathrm{Lex}}(f))$. Gemäß Satz A.2.7 gibt es paarweise verschiedene Standard SO_2-Tableaus $(\mathcal{M}_{k_1}, W_{k_1}), \ldots, (\mathcal{M}_{k_l}, W_{k_l})$ mit $e_1(\mathcal{M}_{k_j}) = 0$ und $e_2(\mathcal{M}_{k_j}) = 0$ für alle $j \in \{1, \ldots, l\}$ sowie Koeffizienten $\alpha_{k_1}, \ldots, \alpha_{k_l} \in K$ so, dass sich dieses Polynom f_k als Linearkombination

$$f_k = \alpha_{k_1}|\mathcal{M}_{k_1}|\langle W_{k_1}\rangle + \ldots + \alpha_{k_l}|\mathcal{M}_{k_l}|\langle W_{k_l}\rangle.$$

schreiben lässt. Somit gibt es ein $j \in \{1, \ldots, l\}$ mit $\mathrm{LT}_{\mathrm{Lex}}(f_k) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_{k_j}|\langle W_{k_j}\rangle)$. Da die Leitterme der paarweise verschiedenen Standard SO_2-Tableaus laut Korollar A.1.24 ebenfalls paarweise verschieden sind, ist j eindeutig bestimmt. Mit $(\mathcal{M}_k, W_k) := (\mathcal{M}_{k_j}, W_{k_j})$ folgt die Existenz.

Seien nun $(\mathcal{M}_k, W_k), (\mathcal{M}'_k, W'_k)$ zwei Standard SO_2-Tableaus mit $e_1(\mathcal{M}_k) = e_2(\mathcal{M}_k) = 0$, $e_1(\mathcal{M}'_k) = e_2(\mathcal{M}'_k) = 0$ und $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle) = \Phi_{k,k+1}(\mathrm{LT}_{\mathrm{Lex}}(f)) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}'_k|\langle W'_k\rangle)$. Dann gilt also $\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}'_k|\langle W'_k\rangle)$, und es folgt $(\mathcal{M}_k, W_k) = (\mathcal{M}'_k, W'_k)$ aus Korollar A.1.24. □

Somit erfüllt insbesondere der Leitterm $\mathrm{LT}_{\mathrm{Lex}}(f)$ die Anforderungen aus Lemma A.3.1, d.h. die Aussagen über den Grad aus diesem Lemma sind auf das eindeutig bestimmte Standard SO_2-Tableau anwendbar, was wir in einem späteren Beweis auch tun werden. Wir betrachten zuvor allerdings die Menge aller Terme, die die Eigenschaft des letzten Korollars erfüllen, und zeigen, dass diese Terme von ganz bestimmten Leittermen erzeugt werden (nach [Ric89], Proposition 14).

Lemma A.3.4. *Sei $G := \mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$ und sei $t \in \mathbb{T}^{mn}$ ein Term mit der Eigenschaft, dass es für alle $k \in \{1, \ldots, n-1\}$ genau ein Standard SO_2-Tableau $(\mathcal{M}_k, W_k)$ gibt mit $e_1(\mathcal{M}) = e_2(\mathcal{M}) = 0$ und $\Phi_{k,k+1}(t) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle)$. Dann ist t ein Element des von $\{\mathrm{LT}_{\mathrm{Lex}}(g) : g \in G\}$ erzeugten multiplikativen Monoids.*

Beweis: Setze $G' := \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$. Seien $g_1, \ldots, g_s$ die Elemente von $\mathrm{Min}(n, \mathcal{X})$ und seien $g_{s+1}, \ldots, g_l$ die Elemente von G'.

Die Behauptung wird per Induktion nach dem Grad d von t bewiesen, genauer zeigen wir, dass es $a_1, \ldots, a_l \in \mathbb{N}$ gibt mit $t = \mathrm{LT}_{\mathrm{Lex}}(g_1)^{a_1} \cdots \mathrm{LT}_{\mathrm{Lex}}(g_l)^{a_l}$. Für $d = 0$ gilt $t = 1$ und mit $a_1 = \ldots = a_l = 0$ folgt die Behauptung.

Sei nun $d > 0$. Wir schreiben wieder $t = t_1 \cdots t_n$ mit $t_j := \varphi_j(t)$ für alle $j \in \{1, \ldots, n\}$. Da $(\mathcal{M}_k, W_k)$ für alle $k \in \{1, \ldots, n-1\}$ ein Standard SO_2-Tableau ist, ist $\deg(\langle W_k \rangle)$ durch 2 teilbar, d.h. mit Lemma A.3.1 ist auch $\deg(t_k) - \deg(t_{k+1})$ für alle $k \in \{1, \ldots, n-1\}$ durch 2 teilbar. Ebenso folgt aus diesem Lemma, dass $\deg(|\mathcal{M}_k|)$ durch 2 teilbar ist. Sei $m_k \in \mathbb{N}$ die Länge der Standard $\mathcal{X}$-Matrix $\mathcal{M}_k$, also die Anzahl der Zeilen der Matrix $\mathcal{M}_k$. Da $|\mathcal{M}_k|$ ein Produkt von 2-Minoren von $\mathcal{X}$ ist, hat jeder der Faktoren in $|\mathcal{M}_k|$ Grad 2. Somit ist die Anzahl der Faktoren durch 2 teilbar und damit ebenso die Länge m_k, da die Anzahl der Faktoren in $|\mathcal{M}_k|$ genau der Anzahl an Zeilen von $\mathcal{M}_k$ entspricht.

Wegen $\deg(\langle W_k \rangle) \geq 0$ folgt $\deg(t_k) \geq \deg(t_{k+1})$ für alle $k \in \{1, \ldots, n-1\}$, also

$$\deg(t_1) \geq \deg(t_2) \geq \ldots \geq \deg(t_n).$$

Setze $u := \max\{j \in \{1, \ldots, n\} : \deg(t_j) > 0\}$. Wegen $d > 0$ existiert dieses Maximum. Damit folgt $\deg(t_k) > 0$ für alle $k \leq u$ und $t_k = 1$ für alle $k > u$, d.h. es gilt $t = t_1 \cdots t_u$. Für alle $k \in \{1, \ldots, u\}$ setze $a(k) := \min\{i \in \{1, \ldots, m\} : x_{i,k} \mid t_k\}$. Dann ist $x_{a(k),1}$ für alle $k \in \{1, \ldots, u\}$ ein Teiler des Terms $\Phi_{k,k+1}(t) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k \rangle)$.

Mit Lemma A.3.1 gilt $\deg(|\mathcal{M}_k|) > 0$ für alle $k < u$ wegen $\deg(t_{k+1}) > 0$, d.h. insbesondere ist die Matrix $\mathcal{M}_k$ in diesem Fall nicht leer. Genauer ist unter Berücksichtigung der bisherigen Ergebnisse die erste Zeile von $\mathcal{M}_k$ für alle $k \in \{1, \ldots, u-1\}$ von der Form $(x_{a(k)}, x_{a(k+1)})$. Wir unterscheiden nun anhand von u zwei Fälle.

Sei zunächst $u = n$, d.h. t enthält aus jeder Spalte von $\mathcal{X}$ mindestens eine Unbestimmte. Sei $\tilde{t}_k$ für alle $k \in \{1, \ldots, n-1\}$ der Term mit $t_k = x_{a(k)k} \cdot \tilde{t}_k$. Dann gilt:

$$t = t_1 \cdots t_n = x_{a(1)1}\tilde{t}_1 \cdots x_{a(n)n}\tilde{t}_n = x_{a(1)1} \cdots x_{a(n)n} \cdot \tilde{t}_1 \cdots \tilde{t}_n.$$

Setze $\tilde{t} := \tilde{t}_1 \cdots \tilde{t}_n$. Dann gilt $\deg(\tilde{t}) < d$. Sei $\tilde{\mathcal{M}}_k$ die Matrix, die aus $\mathcal{M}_k$ entsteht, indem man die erste Zeile streicht. Dann ist $(\tilde{\mathcal{M}}_k, W_k)$ für alle $k \in \{1, \ldots, n-1\}$ erneut ein Standard SO_2-Tableau und es gilt:

$$\begin{aligned}\Phi_{k,k+1}(t) &= \Phi_{k,k+1}(x_{a(1),1} \cdots x_{a(n),n} \cdot \tilde{t}) = \Phi_{k,k+1}(x_{a(1),1} \cdots x_{a(n),n}) \cdot \Phi_{k,k+1}(\tilde{t}) \\ &= x_{a(k),1}x_{a(k+1),2} \cdot \Phi_{k,k+1}(\tilde{t})\end{aligned}$$

sowie

$$\begin{aligned}\mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k \rangle) &= \mathrm{LT}_{\mathrm{Lex}}(|x_{a(k)}\ x_{a(k+1)}|) \cdot \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k \rangle) \\ &= x_{a(k),1}x_{a(k+1),2} \cdot \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k \rangle)\end{aligned}$$

Es folgt $\Phi_{k,k+1}(\tilde{t}) = \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k \rangle)$. Somit gibt es laut Induktionsannahme $c_1, \ldots, c_l \in \mathbb{N}$ mit $\tilde{t} = \mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_l^{c_l})$. Da weiter $(x_{a(k)}, x_{a(k+1)})$ die erste Zeile der Standard $\mathcal{H}$-Matrix $\mathcal{M}_k$ ist, gilt $a(k) < a(k+1)$ für alle $k \in \{1, \ldots, n-1\}$, d.h. es gilt $a(1) < a(2) < \ldots < a(n)$. Damit ist der Term $x_{a(1),1} \cdots x_{a(n),n}$ der Leitterm eines n-Minors von $\mathcal{X}$, d.h. es gibt also ein $j \in \{1, \ldots, s\}$ mit $\mathrm{LT}_{\mathrm{Lex}}(g_j) = x_{a(1),1} \cdots x_{a(n),n}$. Dann gilt:

$$t = x_{a(1),1} \cdots x_{a(n),n} \cdot \tilde{t} = \mathrm{LT}_{\mathrm{Lex}}(g_j) \cdot \mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_l^{c_l}) = \mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_j^{c_j+1} \cdots g_l^{c_l})$$

Setze $a_i := c_i$ für alle $i \in \{1, \ldots, l\}$ mit $i \neq j$ und $a_j := c_j + 1$. Dann folgt die Behauptung.

Sei nun $u < n$. Da die Länge m_k der Standard $\mathcal{H}$-Matrix $\mathcal{M}_k$ für alle $k \in \{1, \ldots, n\}$ durch 2 teilbar ist und für alle $k < u$ zudem $\mathcal{M}_k$ nicht leer ist, gilt $m_k \geq 2$ für alle $k < u$. Damit besteht $|\mathcal{M}_k|$ für alle $k < u$ aus mindestens zwei Faktoren mit jeweils Grad 2, d.h. es gilt $\deg(|\mathcal{M}_k|) \geq 4$ und folglich $\deg(t_{k+1}) \geq 2$ für alle $k < u$. Wegen $\deg(t_k) \geq \deg(t_{k+1})$ gilt schließlich $\deg(t_k) \geq 2$ für alle $k \leq u$.

Setze $b(k) := \min\{i \in \{1, \ldots, m\} : \ x_{a(k),k}x_{i,k} \mid t_k\}$ für alle $k \in \{1, \ldots, u\}$. Dieses Minimum existiert wegen $\deg(t_k) \geq 2$ für alle $k \leq u$. Sei $\tilde{t}_k$ der Term mit $t_k = x_{a(k),k}x_{b(k),k} \cdot \tilde{t}_k$ und $\tilde{t}$ der Term mit $t = x_{a(1),1}x_{b(1),1} \cdots x_{a(u),u}x_{b(u),u} \cdot \tilde{t}$. Dann gilt für alle $k \in \{1, \ldots, u-1\}$:

$$\begin{aligned}\Phi_{k,k+1}(t) &= \Phi_{k,k+1}(t_1 \cdots t_u) = \Phi_{k,k+1}(x_{a(1),1}x_{b(1),1} \cdots x_{a(u),u}x_{b(u),u}) \cdot \Phi_{k,k+1}(\tilde{t}) \\ &= x_{a(k),1}x_{b(k),1}x_{a(k+1),2}x_{b(k+1),2} \cdot \Phi_{k,k+1}(\tilde{t}),\end{aligned}$$

d.h. $x_{a(k),1}x_{b(k),1}x_{a(k+1),2}x_{b(k+1),2}$ teilt den Term $\Phi_{k,k+1}(t) = \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle)$. Somit ist für alle $k \in \{1, \ldots, u-1\}$ die zweite Zeile der Matrix $\mathcal{M}_k$ von der Form $(x_{b(k)}, x_{b(k+1)})$.

Für $k = u$ gilt nun $\deg(|\mathcal{M}_u|) = 2\deg(t_{u+1}) = 0$, also $|\mathcal{M}_u| = 1$ und damit $\mathcal{M}_u = \emptyset$, und $\deg(\langle W_u\rangle) = \deg(t_u) \geq 2$, d.h. W_u besitzt mindestens zwei Einträge und ist damit von der Form $W_u = (x_{a(u)}, x_{b(u)}, \ldots)$.

Für $k \in \{1, \ldots, u-1\}$ sei $\tilde{\mathcal{M}}_k$ analog zu oben die Matrix, die durch Streichen der ersten beiden Zeilen entsteht, und sei $\tilde{W}_k$ das Tupel, das entsteht, indem man die ersten beiden Einträge streicht. Dann gilt für alle $k < u$:

$$\begin{aligned}\Phi_{k,k+1}(t) &= \mathrm{LT}_{\mathrm{Lex}}(|\mathcal{M}_k|\langle W_k\rangle) = \mathrm{LT}_{\mathrm{Lex}}(|x_{a(k)}, x_{a(k+1)}| \cdot |x_{b(k)}, x_{b(k+1)}|) \cdot \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k\rangle) \\ &= x_{a(k),1}x_{a(k+1),2}x_{b(k),1}x_{b(k+1),2} \cdot \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k\rangle),\end{aligned}$$

also $\Phi_{k,k+1}(\tilde{t}) = \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k\rangle)$. Für $k = u$ (und nach wie vor $u < n$) gilt

$$\Phi_{u,u+1}(t) = \mathrm{LT}_{\mathrm{Lex}}(|\emptyset|\langle W_u\rangle) = x_{a(u),1}x_{b(u),1} \cdot \mathrm{LT}_{\mathrm{Lex}}(\langle \tilde{W}_u\rangle)$$

sowie

$$\begin{aligned}\Phi_{u,u+1}(t) &= \Phi_{u,u+1}(t_u) = \Phi_{u,u+1}(x_{a(u),u}x_{b(u),u}) \cdot \Phi_{u,u+1}(\tilde{t}_u) \\ &= x_{a(u),1}x_{b(u),1} \cdot \Phi_{u,u+1}(\tilde{t}).\end{aligned}$$

Somit folgt $\Phi_{u,u+1}(\tilde{t}) = \mathrm{LT}_{\mathrm{Lex}}(\langle \tilde{W}_u\rangle)$. Zusammenfassend gilt also für alle $k \leq u < n$:

$$\Phi_{k,k+1}(\tilde{t}) = \begin{cases} \mathrm{LT}_{\mathrm{Lex}}(|\tilde{\mathcal{M}}_k|\langle W_k\rangle), & \text{für } k < u, \\ \mathrm{LT}_{\mathrm{Lex}}(\langle \tilde{W}_k\rangle), & \text{für } k = u, \\ 1, & \text{für } k > u. \end{cases}$$

Damit erfüllt auch $\tilde{t}$ die Voraussetzungen. Wegen $\deg(\tilde{t}) < \deg(t)$ gibt es deshalb laut Induktionsannahme $c_1, \ldots, c_l \in \mathbb{N}$ mit $\mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_l^{c_l}) = \tilde{t}$. Da die Matrizen $\mathcal{M}_k$ für alle $k \in \{1, \ldots, n-1\}$ standard sind, folgt $a(k) < a(k+1)$ sowie $b(k) < b(k+1)$ für alle $k < u$, also $a(1) < \ldots < a(u)$ sowie $b(1) < \ldots < b(u)$. Somit ist der Term $x_{a(1),1}x_{b(1),1} \cdots x_{a(u),u}x_{b(u),u}$ der Leitterm eines u-Minors von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$. Sei $g_j \in G'$ mit $j \in \{s+1, \ldots, l\}$ der u-Minor von $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ mit $\mathrm{LT}_{\mathrm{Lex}}(g_j) = x_{a(1),1}x_{b(1),1} \cdots x_{a(u),u}x_{b(u),u}$. Dann gilt:

$$\begin{aligned}t &= x_{a(1),1}x_{b(1),1} \cdots x_{a(u),u}x_{b(u),u} \cdot \tilde{t} = \mathrm{LT}_{\mathrm{Lex}}(g_j) \cdot \mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_j^{c_j} \cdots g_l^{c_l}) \\ &= \mathrm{LT}_{\mathrm{Lex}}(g_1^{c_1} \cdots g_j^{c_j+1} \cdots g_l^{c_l}),\end{aligned}$$

womit die Behauptung auch für $u < n$ bewiesen ist. □

Damit lässt sich unmittelbar folgern, dass der Leitterm eines unter SO_n invarianten Polynoms ebenfalls ein Produkt von n-Minoren von $\mathcal{X}$ sowie von k-Minoren der symmetrischen Matrix $\mathcal{X}\mathcal{X}^{\mathrm{tr}}$ für $k \in \{1, \ldots, n-1\}$ ist.

Korollar A.3.5. *Sei* $G := \mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$. *Dann ist* $\mathrm{LT}_{\mathrm{Lex}}(f)$ *für alle* $f \in P^{\mathrm{SO}_n}$ *mit* $f \neq 0$ *ein Element des von* $\{\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G\}$ *erzeugten multiplikativen Monoids.*

Beweis: Sei $f \in P^{\mathrm{SO}_n}$ mit $f \neq 0$. Laut Korollar A.3.3 erfüllt $\mathrm{LT}_{\mathrm{Lex}}(f)$ die Voraussetzungen von Lemma A.3.4. Aus diesem Lemma folgt sofort, dass der Leitterm $\mathrm{LT}_{\mathrm{Lex}}(f)$ ein Element des von $\{\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G\}$ erzeugten Monoids ist. □

Eine weitere Folgerung lässt etwas mehr auf die von G erzeugte Algebra blicken und zeigt, dass G diese K-Algebra als Lex-SAGBI-Basis erzeugt.

Korollar A.3.6. *Sei* $G := \mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$. *Dann ist* G *eine* Lex-*SAGBI-Basis der von* G *erzeugten* K*-Unteralgebra von* P.

Beweis: Sei $S := K[G]$. Wegen Lemma A.2.1 sowie Lemma A.2.1 sind die Elemente von G Elemente von P^{SO_n}. Somit gilt $S \subseteq P^{\mathrm{SO}_n}$. Aus dem letzten Korollar folgt also insbesondere, dass für alle $f \in S \setminus \{0\}$ der Leitterm $\mathrm{LT}_{\mathrm{Lex}}(f)$ ein Element des von $\{\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G\}$ erzeugten multiplikativen Monoids ist, d.h. das Monoid $\{\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in S \setminus \{0\}\}$ wird erzeugt von $\{\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G\}$. Gemäß Theorem 5.2.13 ist G eine Lex-SAGBI-Basis von S. □

In Theorem A.2.8 wurde bereits ein Erzeugendensystem von P^{SO_n} für den Fall $n = 2$ angegeben und bewiesen. Im anschließenden Korollar A.2.9 wurde zudem bereits gezeigt, dass es sich bei diesem Erzeugendensystem um eine Lex-SAGBI-Basis von P^{SO_2} handelt. Mit Hilfe des letzten Lemmas und der Folgerungen daraus ist es nun möglich, auch eine Lex-SAGBI-Basis von P^{SO_n} anzugeben (nach [Ric89], Proposition 14).

Theorem A.3.7. (Lex-SAGBI-Basis von P^{SO_n})
Sei $n \in \mathbb{N}$ *mit* $n \geq 2$, *sei* $m \in \mathbb{N}_+$ *und* $P = K[x_{1,1}, \ldots, x_{m,n}]$. *Sei weiter* G *die Menge der Polynome in* $\mathrm{Min}(n, \mathcal{X}) \cup \mathrm{Min}(1, \mathcal{X}\mathcal{X}^{\mathrm{tr}}) \cup \ldots \cup \mathrm{Min}(n-1, \mathcal{X}\mathcal{X}^{\mathrm{tr}})$. *Dann ist* G *eine* Lex-*SAGBI-Basis von* P^{SO_n}.

Beweis: Sei S die von G erzeugte K-Unteralgebra von P. In Lemma A.2.1 und Lemma A.2.1 wurde bereits gezeigt, dass alle Elemente von G invariant unter den Operationen der SO_n sind, d.h. es gilt $G \subseteq P^{\mathrm{SO}_n}$ und damit $S \subseteq P^{\mathrm{SO}_n}$. Ohne Einschränkung können wir alle Nullen aus G streichen. Diese existieren insbesondere im Fall $m = 1$. Aus Korollar A.3.5 folgt

$$K[\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in P^{\mathrm{SO}_n} \setminus \{0\}] \subseteq K[\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G]$$

und aus Korollar A.3.6 wegen $K[\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in S \setminus \{0\}] = K[\mathrm{LT}_{\mathrm{Lex}}(g) : \ g \in G]$ somit

$$K[\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in P^{\mathrm{SO}_n} \setminus \{0\}] \subseteq K[\mathrm{LT}_{\mathrm{Lex}}(f) : \ f \in S \setminus \{0\}]$$

Damit folgt schließlich $P^{\mathrm{SO}_n} = S$ aus Satz 5.3.10. Erneut aus Korollar A.3.6 folgt, dass G auch eine Lex-SAGBI-Basis von P^{SO_n} ist. □

ANHANG B

Invarianten von Polynomfunktionen ausgewählter Gruppen

B.1 Spezielle orthogonale Gruppe

Folgende Polynome sind fundamentale Invarianten des Invariantenrings $\mathbb{R}[x_1, \ldots, x_{10}]^{\mathrm{SO}_2(\mathbb{R})}$. Betrachtet wird die Operation der speziellen orthogonalen Gruppe $\mathrm{SO}_2(\mathbb{R})$ auf dem reellen Vektorraum $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ der Polynomfunktionen vom Grad ≤ 3 in zwei Unbestimmten. Zugrunde liegt die Standardtermbasis B von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ in folgender Ordnung:

$$B = (1, y, x, y^2, xy, x^2, y^3, y^2x, yx^2, x^3).$$

Die fundamentalen Invarianten werden aufsteigend nach ihrem Grad angegeben:

Grad $d = 1$: Hier erhalten wir die beiden Invarianten

$$\begin{aligned} f_1 &= x_1, \\ f_2 &= x_4 + x_6 \end{aligned}$$

Grad $d = 2$: Im Grad $d = 2$ liegen sechs fundamentale Invarianten vor:

$$\begin{aligned} f_3 &= x_2^2 + x_3^2, \\ f_4 &= x_4x_6 - \tfrac{1}{4}x_5^2, \\ f_5 &= x_7x_9 - \tfrac{1}{3}x_8^2 + x_8x_{10} - \tfrac{1}{3}x_9^2, \\ f_6 &= x_7^2 + x_7x_9 + x_8x_{10} + x_{10}^2, \\ f_7 &= x_2x_8 + 3x_2x_{10} - 3x_3x_7 - x_3x_9, \\ f_8 &= x_2x_7 + \tfrac{1}{3}x_2x_9 + \tfrac{1}{3}x_3x_8 + x_3x_{10}, \end{aligned}$$

Grad $d = 3$: Weiter gibt es zehn fundamentale Invarianten vom Grad 3:

$$
\begin{aligned}
f_9 &= x_2^2x_5 - 2x_2x_3x_4 + 2x_2x_3x_6 - x_3^2x_5,\\
f_{10} &= x_2x_4x_9 - x_2x_5x_8 + 3x_2x_6x_7 + 3x_3x_4x_{10} - x_3x_5x_9 + x_3x_6x_8,\\
f_{11} &= x_2x_4x_{10} - \tfrac{1}{3}x_2x_5x_9 + \tfrac{1}{3}x_2x_6x_8 - \tfrac{1}{3}x_3x_4x_9 + \tfrac{1}{3}x_3x_5x_8 - x_3x_6x_7,\\
f_{12} &= x_2^2x_6 - x_2x_3x_5 + x_3^2x_4,\\
f_{13} &= x_2x_5x_7 + \tfrac{1}{3}x_2x_5x_9 + \tfrac{2}{3}x_2x_6x_8 + 2x_2x_6x_{10} - 2x_3x_4x_7 - \tfrac{2}{3}x_3x_4x_9\\
&\quad - \tfrac{1}{3}x_3x_5x_8 - x_3x_5x_{10},\\
f_{14} &= x_2x_5x_8 + 3x_2x_5x_{10} - 6x_2x_6x_7 - 2x_2x_6x_9 - 2x_3x_4x_8 - 6x_3x_4x_{10}\\
&\quad + 3x_3x_5x_7 + x_3x_5x_9,\\
f_{15} &= x_4x_7x_8 + \tfrac{2}{3}x_4x_8x_9 + x_4x_9x_{10} - \tfrac{3}{2}x_5x_7^2 - \tfrac{1}{6}x_5x_8^2 + \tfrac{1}{6}x_5x_9^2 + \tfrac{3}{2}x_5x_{10}^2\\
&\quad - x_6x_7x_8 - \tfrac{2}{3}x_6x_8x_9 - x_6x_9x_{10},\\
f_{16} &= x_4x_7x_{10} - \tfrac{1}{9}x_4x_8x_9 - \tfrac{1}{3}x_5x_7x_9 + \tfrac{1}{9}x_5x_8^2 + \tfrac{1}{3}x_5x_8x_{10} - \tfrac{1}{9}x_5x_9^2\\
&\quad - x_6x_7x_{10} + \tfrac{1}{9}x_6x_8x_9,\\
f_{17} &= x_4x_8^2 + 2x_4x_9^2 + 9x_4x_{10}^2 - 3x_5x_7x_8 - 2x_5x_8x_9 - 3x_5x_9x_{10} + 9x_6x_7^2 + 2x_6x_8^2 + x_6x_9^2,\\
f_{18} &= x_4x_8x_{10} - \tfrac{1}{3}x_4x_9^2 - \tfrac{3}{2}x_5x_7x_{10} + \tfrac{1}{6}x_5x_8x_9 + x_6x_7x_9 - \tfrac{1}{3}x_6x_8^2,
\end{aligned}
$$

Grad $d = 4$: Außerdem erhalten wir zwölf fundamentale Invarianten vom Grad 4:

$$
\begin{aligned}
f_{19} &= x_2x_4x_5x_9 - 6x_2x_4x_6x_8 - 12x_2x_4x_6x_{10} + \tfrac{1}{2}x_2x_5^2x_8 + \tfrac{9}{2}x_2x_5^2x_{10} + 3x_2x_5x_6x_7\\
&\quad - 2x_2x_5x_6x_9 + 2x_2x_6^2x_8 - 2x_3x_4^2x_9 + 2x_3x_4x_5x_8 - 3x_3x_4x_5x_{10} + 12x_3x_4x_6x_7\\
&\quad + 6x_3x_4x_6x_9 - \tfrac{9}{2}x_3x_5^2x_7 - \tfrac{1}{2}x_3x_5^2x_9 - x_3x_5x_6x_8,\\
f_{20} &= x_2x_4x_5x_{10} - \tfrac{2}{3}x_2x_4x_6x_9 - \tfrac{1}{3}x_2x_5^2x_9 + x_2x_5x_6x_8 - 2x_2x_6^2x_7 - 2x_3x_4^2x_{10}\\
&\quad + x_3x_4x_5x_9 - \tfrac{2}{3}x_3x_4x_6x_8 - \tfrac{1}{3}x_3x_5^2x_8 + x_3x_5x_6x_7,\\
f_{21} &= x_4^2x_8x_9 + 3x_4^2x_9x_{10} - \tfrac{3}{2}x_4x_5x_7x_9 - x_4x_5x_8^2 - \tfrac{3}{2}x_4x_5x_8x_{10} - \tfrac{1}{2}x_4x_5x_9^2\\
&\quad + \tfrac{9}{2}x_4x_5x_{10}^2 + 3x_4x_6x_7x_8 - 3x_4x_6x_9x_{10} + \tfrac{3}{2}x_5^2x_7x_8 - \tfrac{3}{2}x_5^2x_9x_{10} - \tfrac{9}{2}x_5x_6x_7^2\\
&\quad + \tfrac{3}{2}x_5x_6x_7x_9 + \tfrac{1}{2}x_5x_6x_8^2 + \tfrac{3}{2}x_5x_6x_8x_{10} + x_5x_6x_9^2 - 3x_6^2x_7x_8 - x_6^2x_8x_9,\\
f_{22} &= x_4^2x_9^2 + 9x_4^2x_{10}^2 - 2x_4x_5x_8x_9 - 6x_4x_5x_9x_{10} + 2x_4x_6x_8^2 + 2x_4x_6x_9^2 + \tfrac{3}{2}x_5^2x_7x_9\\
&\quad + \tfrac{1}{2}x_5^2x_8^2 + \tfrac{3}{2}x_5^2x_8x_{10} + \tfrac{1}{2}x_5^2x_9^2 - 6x_5x_6x_7x_8 - 2x_5x_6x_8x_9 + 9x_6^2x_7^2 + x_6^2x_8^2,\\
f_{23} &= x_7^3x_{10} - \tfrac{1}{3}x_7^2x_8x_9 + \tfrac{2}{27}x_7x_8^3 + \tfrac{1}{3}x_7x_8^2x_{10} - \tfrac{2}{9}x_7x_8x_9^2\\
&\quad - \tfrac{1}{3}x_7x_9^2x_{10} - x_7x_{10}^3 + \tfrac{1}{27}x_8^3x_9 + \tfrac{2}{9}x_8^2x_9x_{10} - \tfrac{1}{27}x_8x_9^3 + \tfrac{1}{3}x_8x_9x_{10}^2 - \tfrac{2}{27}x_9^3x_{10},\\
f_{24} &= x_7^2x_{10}^2 - \tfrac{2}{3}x_7x_8x_9x_{10} + \tfrac{4}{27}x_7x_9^3 + \tfrac{4}{27}x_8^3x_{10} - \tfrac{1}{27}x_8^2x_9^2,\\
f_{25} &= x_2^2x_7x_8 + \tfrac{3}{2}x_2^2x_7x_{10} + \tfrac{1}{2}x_2^2x_8x_9 + x_2^2x_9x_{10} - 3x_2x_3x_7^2 - x_2x_3x_7x_9\\
&\quad + x_2x_3x_8x_{10} + 3x_2x_3x_{10}^2 - x_3^2x_7x_8 - \tfrac{3}{2}x_3^2x_7x_{10} - \tfrac{1}{2}x_3^2x_8x_9 - x_3^2x_9x_{10},\\
f_{26} &= x_2x_7^2x_9 - \tfrac{1}{3}x_2x_7x_8^2 + \tfrac{1}{2}x_2x_7x_8x_{10} + \tfrac{1}{3}x_2x_7x_9^2 + \tfrac{3}{2}x_2x_7x_{10}^2\\
&\quad - \tfrac{1}{6}x_2x_8^2x_9 - \tfrac{1}{6}x_2x_8x_9x_{10} + \tfrac{3}{2}x_3x_7^2x_{10} - \tfrac{1}{6}x_3x_7x_8x_9 + \tfrac{1}{2}x_3x_7x_9x_{10}\\
&\quad + \tfrac{1}{3}x_3x_8^2x_{10} - \tfrac{1}{6}x_3x_8x_9^2 + x_3x_8x_{10}^2 - \tfrac{1}{3}x_3x_9^2x_{10},\\
f_{27} &= x_2x_7^2x_8 + \tfrac{9}{2}x_2x_7^2x_{10} + \tfrac{5}{6}x_2x_7x_8x_9 + \tfrac{7}{2}x_2x_7x_9x_{10} + \tfrac{4}{3}x_2x_8^2x_{10} - \tfrac{1}{6}x_2x_8x_9^2\\
&\quad + 5x_2x_8x_{10}^2 - \tfrac{1}{3}x_2x_9^2x_{10} + 3x_2x_{10}^3 - 3x_3x_7^3 - 5x_3x_7^2x_9 + \tfrac{1}{3}x_3x_7x_8^2\\
&\quad - \tfrac{7}{2}x_3x_7x_8x_{10} - \tfrac{4}{3}x_3x_7x_9^2 - \tfrac{9}{2}x_3x_7x_{10}^2 + \tfrac{1}{6}x_3x_8^2x_9 - \tfrac{5}{6}x_3x_8x_9x_{10} - x_3x_9x_{10}^2,
\end{aligned}
$$

$$
\begin{aligned}
f_{28} &= x_2^2x_8^2 + 9x_2^2x_8x_{10} - x_2^2x_9^2 + 9x_2^2x_{10}^2 - 6x_2x_3x_7x_8 - 27x_2x_3x_7x_{10} - x_2x_3x_8x_9 \\
&\quad - 6x_2x_3x_9x_{10} + 9x_3^2x_7^2 + 9x_3^2x_7x_9 - x_3^2x_8^2 + x_3^2x_9^2, \\
f_{29} &= x_2^3x_{10} - x_2^2x_3x_9 + x_2x_3^2x_8 - x_3^3x_7, \\
f_{30} &= x_2^3x_9 - 2x_2^2x_3x_8 + 3x_2^2x_3x_{10} + 3x_2x_3^2x_7 - 2x_2x_3^2x_9 + x_3^3x_8,
\end{aligned}
$$

Grad $d = 5$: Schließlich gibt es noch zwei fundamentale Invarianten vom Grad 5:

$$
\begin{aligned}
f_{31} &= x_4^3x_9x_{10} - x_4^2x_5x_8x_{10} - \tfrac{1}{2}x_4^2x_5x_9^2 + \tfrac{3}{2}x_4^2x_5x_{10}^2 + 3x_4^2x_6x_7x_{10} + \tfrac{1}{3}x_4^2x_6x_8x_9 \\
&\quad - x_4^2x_6x_9x_{10} + \tfrac{2}{3}x_4x_5^2x_8x_9 - x_4x_5^2x_9x_{10} - 2x_4x_5x_6x_7x_9 - \tfrac{1}{3}x_4x_5x_6x_8^2 \\
&\quad + 2x_4x_5x_6x_8x_{10} + \tfrac{1}{3}x_4x_5x_6x_9^2 + x_4x_6^2x_7x_8 - 3x_4x_6^2x_7x_{10} - \tfrac{1}{3}x_4x_6^2x_8x_9 \\
&\quad - \tfrac{1}{6}x_5^3x_8^2 + \tfrac{1}{6}x_5^3x_9^2 + x_5^2x_6x_7x_8 - \tfrac{2}{3}x_5^2x_6x_8x_9 - \tfrac{3}{2}x_5x_6^2x_7^2 + x_5x_6^2x_7x_9 \\
&\quad + \tfrac{1}{2}x_5x_6^2x_8^2 - x_6^3x_7x_8, \\
f_{32} &= x_4^3x_{10}^2 - x_4^2x_5x_9x_{10} + \tfrac{1}{3}x_4^2x_6x_9^2 + \tfrac{1}{2}x_4x_5^2x_8x_{10} + \tfrac{1}{6}x_4x_5^2x_9^2 - \tfrac{2}{3}x_4x_5x_6x_8x_9 \\
&\quad + \tfrac{1}{3}x_4x_6^2x_8^2 - \tfrac{1}{4}x_5^3x_7x_{10} - \tfrac{1}{12}x_5^3x_8x_9 + \tfrac{1}{2}x_5^2x_6x_7x_9 + \tfrac{1}{6}x_5^2x_6x_8^2 \\
&\quad - x_5x_6^2x_7x_8 + x_6^3x_7^2
\end{aligned}
$$

B.2 Orthogonale Gruppe

Folgende Polynome sind fundamentale Invarianten des Invariantenrings $\mathbb{R}[x_1, \ldots, x_{10}]^{\mathrm{O}_2(\mathbb{R})}$. Betrachtet wird die Operation der orthogonalen Gruppe $\mathrm{O}_2(\mathbb{R})$ auf dem reellen Vektorraum $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ der Polynomfunktionen vom Grad ≤ 3 in zwei Unbestimmten. Zugrunde liegt die Standardtermbasis B von $\mathscr{P}_{\leq 3}(\mathbb{R}^2, \mathbb{R})$ in folgender Ordnung:

$$B = (1, y, x, y^2, xy, x^2, y^3, y^2x, yx^2, x^3).$$

Die fundamentalen Invarianten werden aufsteigend nach ihrem Grad angegeben:

Grad $d = 1$: $f_1 = x_1$ und $f_2 = x_4 + x_6$.

Grad $d = 2$:

$$
\begin{aligned}
f_3 &= x_4x_6 - \tfrac{1}{4}x_5^2, \\
f_4 &= x_7x_9 - \tfrac{1}{3}x_8^2 + x_8x_{10} - \tfrac{1}{3}x_9^2, \\
f_5 &= x_7^2 + x_7x_9 + x_8x_{10} + x_{10}^2, \\
f_6 &= x_2x_7 + \tfrac{1}{3}x_2x_9 + \tfrac{1}{3}x_3x_8 + x_3x_{10}, \\
f_7 &= x_2^2 + x_3^2,
\end{aligned}
$$

Grad $d = 3$:

$$
\begin{aligned}
f_8 &= x_2x_4x_9 - x_2x_5x_8 + 3x_2x_6x_7 + 3x_3x_4x_{10} - x_3x_5x_9 + x_3x_6x_8, \\
f_9 &= x_2^2x_6 - x_2x_3x_5 + x_3^2x_4, \\
f_{10} &= x_2x_5x_8 + 3x_2x_5x_{10} - 6x_2x_6x_7 - 2x_2x_6x_9 - 2x_3x_4x_8 - 6x_3x_4x_{10} + 3x_3x_5x_7 + x_3x_5x_9, \\
f_{11} &= x_4x_8^2 + 2x_4x_9^2 + 9x_4x_{10}^2 - 3x_5x_7x_8 - 2x_5x_8x_9 - 3x_5x_9x_{10} + 9x_6x_7^2 + 2x_6x_8^2 + x_6x_9^2, \\
f_{12} &= x_4x_8x_{10} - \tfrac{1}{3}x_4x_9^2 - \tfrac{3}{2}x_5x_7x_{10} + \tfrac{1}{6}x_5x_8x_9 + x_6x_7x_9 - \tfrac{1}{3}x_6x_8^2,
\end{aligned}
$$

Grad $d = 4$:

$$
\begin{aligned}
f_{13} &= x_2x_4x_5x_{10} - \tfrac{2}{3}x_2x_4x_6x_9 - \tfrac{1}{3}x_2x_5^2x_9 + x_2x_5x_6x_8 - 2x_2x_6^2x_7 - 2x_3x_4^2x_{10} + x_3x_4x_5x_9 \\
&\quad - \tfrac{2}{3}x_3x_4x_6x_8 - \tfrac{1}{3}x_3x_5^2x_8 + x_3x_5x_6x_7, \\
f_{14} &= x_4^2x_9^2 + 9x_4^2x_{10}^2 - 2x_4x_5x_8x_9 - 6x_4x_5x_9x_{10} + 2x_4x_6x_8^2 + 2x_4x_6x_9^2 + \tfrac{3}{2}x_5^2x_7x_9 \\
&\quad + \tfrac{1}{2}x_5^2x_8^2 + \tfrac{3}{2}x_5^2x_8x_{10} + \tfrac{1}{2}x_5^2x_9^2 - 6x_5x_6x_7x_8 - 2x_5x_6x_8x_9 + 9x_6^2x_7^2 + x_6^2x_8^2, \\
f_{15} &= x_2x_7^2x_9 - \tfrac{1}{3}x_2x_7x_8^2 + \tfrac{2}{3}x_2x_7x_9^2 + 3x_2x_7x_{10}^2 - \tfrac{2}{9}x_2x_8^2x_9 - \tfrac{2}{3}x_2x_8x_9x_{10} + \tfrac{1}{9}x_2x_9^3 \\
&\quad + 3x_3x_7^2x_{10} - \tfrac{2}{3}x_3x_7x_8x_9 + \tfrac{1}{9}x_3x_8^3 + \tfrac{2}{3}x_3x_8^2x_{10} - \tfrac{2}{9}x_3x_8x_9^2 + x_3x_8x_{10}^2 - \tfrac{1}{3}x_3x_9^2x_{10}, \\
f_{16} &= x_7^2x_{10}^2 - \tfrac{2}{3}x_7x_8x_9x_{10} + \tfrac{4}{27}x_7x_9^3 + \tfrac{4}{27}x_8^3x_{10} - \tfrac{1}{27}x_8^2x_9^2, \\
f_{17} &= x_2^2x_8^2 + 9x_2^2x_8x_{10} - x_2^2x_9^2 + 9x_2^2x_{10}^2 - 6x_2x_3x_7x_8 - 27x_2x_3x_7x_{10} - x_2x_3x_8x_9 \\
&\quad - 6x_2x_3x_9x_{10} + 9x_3^2x_7^2 + 9x_3^2x_7x_9 - x_3^2x_8^2 + x_3^2x_9^2, \\
f_{18} &= x_2^3x_9 - 2x_2^2x_3x_8 + 3x_2^2x_3x_{10} + 3x_2x_3^2x_7 - 2x_2x_3^2x_9 + x_3^3x_8,
\end{aligned}
$$

Grad $d = 5$:

$$
\begin{aligned}
f_{19} &= x_4^3x_{10}^2 - x_4^2x_5x_9x_{10} + \tfrac{1}{3}x_4^2x_6x_9^2 + \tfrac{1}{2}x_4x_5^2x_8x_{10} + \tfrac{1}{6}x_4x_5^2x_9^2 - \tfrac{2}{3}x_4x_5x_6x_8x_9 + \tfrac{1}{3}x_4x_6^2x_8^2 \\
&\quad - \tfrac{1}{4}x_5^3x_7x_{10} - \tfrac{1}{12}x_5^3x_8x_9 + \tfrac{1}{2}x_5^2x_6x_7x_9 + \tfrac{1}{6}x_5^2x_6x_8^2 - x_5x_6^2x_7x_8 + x_6^3x_7^2
\end{aligned}
$$

ANHANG C

Das ApCoCoA-Paket `sagbi.cpkg`

Im Zuge dieser Arbeit sind zahlreiche Algorithmen und Funktionen für SAGBI-Basen (siehe Kapitel 5) im Computeralgebrasystem ApCoCoA[30] in der CoCoA eigenen Skriptsprache CoCoAL implementiert worden. Sie bilden das ApCoCoA-Paket `sagbi.cpkg`. In diesem Anhang werden die Funktionen dieses Pakets kurz erläutert und deren Funktionsweise anhand von Beispielen illustriert. Das Paket `sagbi.cpkg` ist in der hier präsentierten Form nicht Teil der aktuellsten Version 1.9.1 von ApCoCoA (Stand: 31. August 2016), das dort enthaltene Paket ist als veraltet anzusehen. Folgende Informationen sind für die Arbeit mit dem Paket von Bedeutung:

Paketpfad: `$apcocoa/sagbi`

Globaler Alias: `SB`

Das Paket `sagbi.cpkg` enthält die nachfolgend aufgelisteten Funktionen in alphabetischer Reihenfolge, unterteilt in Haupt- und Hilfsfunktionen:

[30] http:\www.apcocoa.org

C.1 Hauptfunktionen des Pakets

Die folgenden CoCoA-Funktionen bilden (in alphabetischer Reihenfolge) den funktionalen Kern des Pakets `sagbi.cpkg`.

```
SB.EvalSubalgPoly(H:POLY,PolyList:LIST of POLY):POLY
```

Beschreibung:
Wertet ein Polynom `H` in $K[y_1, \ldots, y_s]$ an den Polynomen $g_1, \ldots, g_s \in K[x_1, \ldots, x_n]$ der Liste `PolyList` aus und gibt das Polynom $H(g_1, \ldots, g_s)$ in $K[x_1, \ldots, x_n]$ zurück.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
SARing::=QQ[y[1..3]];
G:=[x-y, xy-y^2, xy^2];
F:=SARing::(y[1]^3-y[1]y[2]+y[3]^2);

SB.EvalSubalgPoly(F,G);
G[1]^3-G[1]*G[2]+G[3]^2;

x^2y^4 + x^3 - 4x^2y + 5xy^2 - 2y^3
------------------------------
x^2y^4 + x^3 - 4x^2y + 5xy^2 - 2y^3
------------------------------
```

```
SB.HomIsInSubalg(F:POLY,Gens:LIST of POLY):BOOL
```

Beschreibung:
Diese Funktion bildet den Unteralgebra-Mitgliedschaftstest im homogenen Fall (bzgl. der Standardgraduierung) aus Korollar 5.4.18. Die Eingabeparameter `F` und `Gens` müssen entsprechend jeweils homogen sein. Die Funktion erwartet mit `Gens` als Eingabe eine d-Grad-beschränkte SAGBI-Basis in der zugrunde liegenden Termordnung für $d = \deg(F)$.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x-y, xy-y^2, xy^2];
F:=xy^4-y^5;

SB.HomIsInSubalg(F,G);

H:=SB.TruncSagbi(G,Deg(F));
H;

SB.HomIsInSubalg(F,H);

ERROR: SB.HomIsInSubalg: ARGV[2] is not a Deg(F)-truncated SAGBI-Basis!
CONTEXT: Error("SB.HomIsInSubalg: ARGV[2] is not a Deg(F)-truncated SAGBI-Basis!")
------------------------------
[x - y, xy - y^2, xy^2, xy^3 - y^4, xy^4 - y^5]
------------------------------
True
------------------------------
```

```
SB.HomSagbi(Gens:LIST of POLY):LIST of POLY
```

Beschreibung:
Implementation der homogenen SAGBI-Prozedur (siehe Prozedur HomSagbi, Seite 98). Diese enummerierte Prozedur berechnet im homogenen Fall (bzgl. der Standardgraduierung) eine homogene SAGBI-Basis der von `Gens` erzeugten K-Unteralgebra S bzgl. der betrachteten Termordnung. Die Prozedur terminiert genau dann, wenn S eine endliche homogene SAGBI-Basis besitzt. Die Funktion erwartet eine Liste homogener Polynome als Eingabe.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2y, x^2-y^2, x^2y^2-y^4, x^2y^4];

SB.HomSagbi(G);

[x^2 - y^2, x^2y, x^2y^2 - y^4, x^2y^4 - 1/2y^6, y^6]
-------------------------------
```

```
SB.HomSubalgRepr(F:POLY, Gens:LIST of POLY):POLY or NULL
```

Beschreibung:
Diese Funktion versucht eine explizite Darstellung des homogenen Polynoms `F` in den homogenen Polynomen $g_1, \ldots, g_s \in K[x_1, \ldots, x_n]$ der Liste `Gens` zu berechnen. Ist `F` ein Polynom in $K[g_1, \ldots, g_s]$, wird ein Polynom $h \in K[y_1, \ldots, y_s]$ ausgegeben mit $F = h(g_1, \ldots, g_s)$, andernfalls gibt die Funktion `NULL` zurück. Die Funktion erwartet als Eingabe eine $\deg(F)$-Gradbeschränkte SAGBI-Basis `Gens`.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x-y, xy-y^2, xy^2];
F:=xy^4-y^5;

H:=SB.TruncSagbi(G,Deg(F));
H;

SB.HomSubalgRepr(F,H);

[x - y, xy - y^2, xy^2, xy^3 - y^4, xy^4 - y^5]
-------------------------------
SubalgebraRing :: y[5]
-------------------------------
```

```
SB.IsInSubalg(Poly:POLY, Gens:LIST of POLY):BOOL
```

Beschreibung:
Diese Funktion bildet den Unteralgebra-Mitgliedschaftstest aus Korollar 5.3.8, d.h. sie überprüft, ob das Polynom `Poly` ein Element der von `Gens` erzeugten K-Unteralgebra S ist. Dazu erwartet die Funktion mit `Gens` eine endliche SAGBI-Basis von S.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2y, x^2-y^2, x^2y^2-y^4, x^2y^4];
H:=SB.Sagbi(G);

F1:=G[1]^3*G[4]-2*G[2]^3+G[3]^3;
F2:=G[1]^3*G[4]-2*G[2]^3+G[3]^3+x;

H;
SB.IsInSubalg(F1,H);
SB.IsInSubalg(F2,H);

[x^2y, x^2 - y^2, x^2y^2 - y^4, x^2y^4, y^6, x^2y^6 - y^8]
-------------------------------
True
-------------------------------
False
-------------------------------
```

`SB.IsReducedSagbi(G:LIST of POLY):BOOL`

Beschreibung:
Die Funktion überprüft, ob `G` eine reduzierte SAGBI-Basis bzgl. der gegebenen Termordnung von der von `G` erzeugten K-Unteralgebra ist (siehe Definition 5.3.11).

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2-y^2, x^2y, x^2y^2-y^4, x^2y^4];

H:=SB.Sagbi(G);
L:=SB.ReducedSagbi(H);

SB.IsReducedSagbi(H);
SB.IsReducedSagbi(L);

False
-------------------------------
True
-------------------------------
```

`SB.IsSagbi(G:LIST of POLY):BOOL`

Beschreibung:
Die Funktion überprüft, ob `G` eine SAGBI-Basis bzgl. der gegebenen Termordnung von der von `G` erzeugten K-Unteralgebra ist (siehe Satz 5.3.7). Dazu wird Algorithmus 5.4 implementiert.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2y, x^2-y^2, x^2y^2-y^4, x^2y^4];

SB.IsSagbi(G);
```

```
H:=SB.Sagbi(G);
SB.IsSagbi(H);

False
-------------------------------
True
-------------------------------
```

```
SB.IsSagbiOf(GensS:LIST of POLY, Basis:LIST of POLY):BOOL
```

Beschreibung:
Die Funktion überprüft, ob `Basis` eine SAGBI-Basis bzgl. der gegebenen Termordnung von der von `GensS` erzeugten K-Unteralgebra ist. Als Eingabe erwartet die Funktion also unter anderem mit `Basis` eine SAGBI-Basis der von `Basis` erzeugten K-Unteralgebra bzgl. der gegebenen Termordnung.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2y, x^2-y^2, x^2y^2-y^4, x^2y^4];

H:=SB.Sagbi(G);
H;
SB.IsSagbiOf(G,H);

[x^2y, x^2 - y^2, x^2y^2 - y^4, x^2y^4, y^6, x^2y^6 - y^8]
-------------------------------
True
-------------------------------
```

```
SB.IsTruncSagbi(Basis:LIST of POLY, D:INT):BOOL
```

Beschreibung:
Die Funktion überprüft, ob `Basis` eine `D`-Grad-beschränkte SAGBI-Basis bzgl. der gegebenen Termordnung ist (siehe Algorithmus 5.7).

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x-y, xy-y^2, xy^2];

H:=SB.TruncSagbi(G,6);
H;
SB.IsTruncSagbi(H,6);
SB.IsTruncSagbi(H,4);
SB.IsTruncSagbi(H,8);

[x - y, xy - y^2, xy^2, xy^3 - y^4, xy^4 - y^5, xy^5 - 1/2y^6]
-------------------------------
True
-------------------------------
True
-------------------------------
False
-------------------------------
```

```
SB.NFS(F:POLY, Polys:LIST of POLY):POLY
```

Beschreibung:
Diese Funktion berechnet die SAGBI-Normalform (siehe Definition 5.3.2) von `F` bzgl. einer endlichen SAGBI-Basis. Somit erwartet die Funktion als Eingabe mit der Liste `Polys` eine endliche SAGBI-Basis $\{g_1, \ldots, g_s\}$ bzgl. der gegebenen Termordnung. Die Funktion nutzt zur Berechnung die Funktion `SB.NRS` (siehe Korollar 5.3.3).

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x+y, xy];
F1:=G[1]^3*G[2]-5*G[2]^3;
F2:=x^3+x^2y;

SB.NFS(F1,G);
SB.NFS(F2,G);

0
-------------------------------
-xy^2 - y^3
-------------------------------
```

```
SB.NRS(F:POLY, Polys:LIST of POLY):POLY
```

Beschreibung:
Berechnet den normalen Unteralgebra-Rest von `F` bei Unteralgebra-Division durch `Polys` (siehe Definition 5.2.8). Dazu wird die Funktion `SB.SubalgebraDivAlg` aufgerufen und dessen zweiter Rückgabeparameter hier ausgegeben.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x+y, xy];
F1:=G[1]^3*G[2]-5*G[2]^3;
F2:=x^3+x^2y;

SB.NRS(F1,G);
SB.SubalgebraDivAlg(F1,G);
SB.NRS(F2,G);
SB.SubalgebraDivAlg(F2,G);

0
-------------------------------
[SubalgebraRing :: y[1]^3y[2] - 5y[2]^3, 0]
-------------------------------
-xy^2 - y^3
-------------------------------
[SubalgebraRing :: y[1]^3 - 2y[1]y[2], -xy^2 - y^3]
-------------------------------
```

```
SB.ReducedSagbi(Basis:LIST of POLY):LIST of POLY
```

Beschreibung:
Diese Funktion berechnet die reduzierte SAGBI-Basis der von `Basis` erzeugten K-Unteralgebra bzgl. der gegebenen Termordnung (siehe Algorithmus 5.3). Die Funktion erwartet mit `Basis` als Eingabe eine endliche SAGBI-Basis bzgl. der gegebenen Termordnung.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2-y^2, x^2y, x^2y^2-y^4, x^2y^4];

H:=SB.Sagbi(G);
H;
SB.ReducedSagbi(H);

[x^2 - y^2, x^2y, x^2y^2 - y^4, x^2y^4, x^2y^6 - y^8, y^6]
------------------------------
[x^2 - y^2, x^2y, x^2y^2 - y^4, x^2y^4, y^6]
------------------------------
-- Done.
------------------------------
```

```
SB.Sagbi(GensS:LIST of POLY):LIST of POLY
```

Beschreibung:
Implementation der Prozedur SAGBI. Die Prozedur berechnet eine endliche SAGBI-Basis der von `GensS` erzeugten K-Unteralgebra S bzgl. der gegebenen Termordnung und terminiert genau dann, wenn S eine endliche SAGBI-Basis besitzt.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2-y^2, x^2y, x^2y^2-y^4, x^2y^4];

SB.Sagbi(G);

[x^2 - y^2, x^2y, x^2y^2 - y^4, x^2y^4, y^6, x^2y^6 - y^8]
------------------------------
```

```
SB.SubalgebraDivAlg(F:POLY, Polys:LIST of POLY):LIST of POLY
```

Beschreibung:
Implementation des Unteralgebra-Divisionsalgorithmus (siehe Algorithmus 5.2). Der Algorithmus berechnet für das Polynom `F` und das Tupel von Polynomen $(g_1, \ldots, g_s)$ der Liste `Polys` ein Polynom $h \in K[y_1, \ldots, y_s]$ und ein Polynom $\tilde{f} \in K[x_1, \ldots, x_n]$ mit $F = h(g_1, \ldots, g_s) + \tilde{f}$, die folgende Eigenschaften erfüllen:

(i) $\mathrm{Supp}(\tilde{f}) \cap K[\mathrm{LT}_\sigma(g_1), \ldots, \mathrm{LT}_\sigma(g_s)] = \emptyset$.

(ii) Für alle $t \in \mathrm{Supp}(h)$ gilt $\mathrm{LT}_\sigma(t(g_1, \ldots, g_s)) \leq_\sigma \mathrm{LT}_\sigma(f)$.

Beispiel:

```
Use P::=QQ[x];
G:=[x^3-x, x^4, x^5-1];
F1:=x^8;
F2:=x^2;

SB.SubalgebraDivAlg(F1,G);
SB.SubalgebraDivAlg(F2,G);

[SubalgebraRing :: y[2]^2, 0]
------------------------------
[SubalgebraRing :: 0, x^2]
------------------------------
```

```
SB.SubalgRepr(F:POLY, Basis:LIST of POLY):POLY
```

Beschreibung:
Diese Funktion berechnet eine Darstellung von `F` in den Polynomen $g_1, \ldots, g_s$ aus der Liste `Basis`, falls `F` Element der von `Basis` erzeugten K-Unteralgebra ist (siehe Bemerkung 5.3.9), d.h. in diesem Fall wird ein Polynom $h \in K[y_1, \ldots, y_s]$ berechnet mit $F = h(g_1, \ldots, g_s)$. Ansonsten wird `NULL` zurückgegeben. Die Funktion erwartet eine SAGBI-Basis bzgl. der gegebenen Termordnung als Eingabe.

Beispiel:

```
Use P::=QQ[x];
G:=[x^3-x, x^4, x^5-1];
F:=x^8;

SB.SubalgRepr(F,G);
H:=SB.Sagbi(G);
H;

SB.SubalgRepr(F,H);

ERROR: SB.IsSagbiOf: ARGV[2] is not a SAGBI-Basis!
CONTEXT: Error("SB.IsSagbiOf: ARGV[2] is not a SAGBI-Basis!")
------------------------------
[x^3 - x, x^4, x^5 - 1, x^2, x]
------------------------------
SubalgebraRing :: y[2]^2
------------------------------
```

```
SB.TermRepr(Term:POLY, TermList:LIST of POLY):LIST of INT
SB.TermRepr(Term:POLY, TermList:LIST of POLY, ReprType:INT):LIST
of INT
```

Beschreibung:
Die Funktion berechnet eine Darstellung eines Terms $t \in \mathbb{T}^n$ in Termen $t_1, \ldots, t_s \in \mathbb{T}^n$, d.h. ein Tupel $(a_1, \ldots, a_s) \in \mathbb{N}^s$ mit $t = t_1^{a_1} \cdots t_s^{a_s}$, falls eine derartige Darstellung existiert. Andernfalls wird die leere Liste `[]` ausgegeben. Zur Berechnung stehen drei Alternativen zur Auswahl, die über den optionalen Parameter `ReprType` angesteuert werden können:

- `ReprType=0`: (Default) Die Darstellung wird mittels torischer Ideale berechnet (siehe Algorithmus 5.1 und `SB.TermReprToric`). In dieser Variante bildet diese Funktion die Implementation von Algorithmus 5.1.
- `ReprType=1`: Diese Variante nutzt einen K-Algebra-Homomorphismus (siehe Satz 2.2.8 und `SB.TermReprAlghom`).
- `ReprType=2`: Die zugrunde liegende lineare diophantische Gleichung wird mit naiven und weniger effizienten Methoden gelöst (siehe `SB.TermReprDio`).

Beispiel:

```
Use P::=QQ[x,y], DegLex;
TermList:=[xy^2, x^3y, y^2];
T:=x^5y^5;

SB.TermRepr(T,TermList);
SB.TermRepr(T,TermList,1);
SB.TermRepr(T,TermList,2);

[2, 1, 0]
-------------------------------
[2, 1, 0]
-------------------------------
[2, 1, 0]
-------------------------------
```

`SB.TruncSagbi(Gens:LIST of POLY, Deg:INT):LIST of POLY`

Beschreibung:
Implementation von Algorithmus 5.6. Dieser Algorithmus berechnet eine `Deg`-Grad-beschränkte SAGBI-Basis von der von `Gens` erzeugten K-Unteralgebra bzgl. der gegebenen Termordnung.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x-y, xy-y^2, xy^2];

SB.TruncSagbi(G,4);
SB.TruncSagbi(G,6);
SB.TruncSagbi(G,8);

[x - y, xy - y^2, xy^2, xy^3 - y^4]
-------------------------------
[x - y, xy - y^2, xy^2, xy^3 - y^4, xy^4 - y^5, xy^5 - 1/2y^6]
-------------------------------
[x - y, xy - y^2, xy^2, xy^3 - y^4, xy^4 - y^5, xy^5 - 1/2y^6,
  xy^6 - y^7, xy^7 - y^8]
-------------------------------
```

C.2 Hilfsfunktionen des Pakets

Im Folgenden sind die weniger bedeutenden Hilfsfunktionen alphabetisch aufgelistet.

```
SB.CompareLex(L1:LIST of INT, L2:LIST of INT):BOOL
```

Beschreibung:
Die Funktion vergleicht zwei Tupel ganzer Zahlen gleicher Länge lexikographisch. Sie gibt TRUE zurück, wenn L1 lexikographisch größer ist als L2.

Beispiel:

```
L1:=[1,0,0];
L2:=[1,2,0];

SB.CompareLex(L1,L2);

False
-------------------------------
L1:=[1,0,0];
L2:=[1,-2,0];

SB.CompareLex(L1,L2);

True
-------------------------------
```

```
SB.EvalPolyList(G:LIST of POLY, L:LIST of INT):POLY
```

Beschreibung:
Diese Funktion bildet das Potenzprodukt mit den Polynomen $\{g_1, \ldots, g_s\}$ der Liste G und den Exponenten $\{a_1, \ldots, a_s\}$ der Liste L, d.h. es wird das Polynom $g_1^{a_1} \cdots g_s^{a_s}$ berechnet. Die beiden Listen müssen gleiche Länge haben. Zudem wird erwartet, dass L nur nicht-negative ganze Zahlen enthält.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2-y, x+y, y^2];
L:=[1,2,1];

SB.EvalPolyList(G,L);
G[1]^1*G[2]^2*G[3]^1;

x^4y^2 + 2x^3y^3 + x^2y^4 - x^2y^3 - 2xy^4 - y^5
-------------------------------
x^4y^2 + 2x^3y^3 + x^2y^4 - x^2y^3 - 2xy^4 - y^5
-------------------------------
```

```
SB.RelationIdeal(Terms:LIST of POLY):LIST of POLY
```

Beschreibung:
Implementation von Algorithmus 2.2. Dieser Algorithmus berechnet das Relationenideal eines Tupels $(t_1, \ldots, t_s)$ von Termen der Liste Terms mit Hilfe torischer Ideale (siehe Definition 2.3.1), d.h. als Eingabe erwartet die Funktion eine Liste $t_1, \ldots, t_s \in \mathbb{T}^n \setminus \{1\}$ von Termen. Die Funktion gibt eine Liste von echten Binomen in $K[y_1, \ldots, y_s]$ zurück.

Beispiel:

```
Use P::=QQ[x,y], DegLex;
G:=[x^2-y, x+y, y^2];

SB.RelationIdeal(G);

[SARing :: -y[2]^2 + y[1]]
------------------------------
```

`SB.ReplaceB(B:LIST of POLY, IndetsInvolved:LIST of POLY):LIST of POLY`

Beschreibung:
Diese Funktion ist eine Implementation von Algorithmus 5.5 und stellt eine Hilfsfunktion für Prozedur SAGBI dar. Sie wählt aus einer Liste von Binomen `B` diejenigen Binome aus, die mindestens eine Unbestimmte aus der Liste `IndetsInvolved` enthalten, und gibt diese Auswahl zurück. Erwartet wird als Eingabe als insbesondere eine Liste `IndetsInvolved` von Unbestimmten des aktuellen Polynomrings.

Beispiel:

```
Use P::=QQ[x,y,z], DegLex;
B:=[x^2y-y, y^2z, xz^2-y, 2x+y];

SB.ReplaceB(B,[z]);
SB.ReplaceB(B,[x]);

[y^2z, xz^2 - y]
------------------------------
[y^2z, xz^2 - y]
------------------------------
```

`SB.TermReprAlghom(Term:POLY, TermList:LIST of POLY):LIST of INT`

Beschreibung:
Diese Funktion berechnet eine Darstellung eines Terms $t \in \mathbb{T}^n$ in den Termen $t_1, \ldots, t_s \in \mathbb{T}^n$ der Liste `TermList` mit Hilfe eines K-Algebra-Homomorphismus (siehe Satz 2.2.8). Dazu wird die CoCoA-Funktion `SubalgebraRepr` verwendet. Berechnet wird ein Tupel $(c_1, \ldots, c_s) \in \mathbb{N}^s$ mit $t = t_1^{c_1} \cdots t_s^{c_s}$. Im Gegensatz zu `SB.TermReprToric` berechnet diese Funktion ein beliebiges Tupel unter allen möglichen.

Beispiel:

```
Use P::=QQ[x,y,z], DegLex;
TermList:=[x^2y, yz^2, xyz];
Term:=x^4y^3z^2;

SB.TermReprAlghom(Term, TermList);

[1, 0, 2]
------------------------------
```

```
SB.TermReprDio(Term:POLY, TermList:LIST of POLY):LIST of INT
```

Beschreibung:
Diese Funktion berechnet eine Darstellung eines Terms $t \in \mathbb{T}^n$ in den Termen $t_1, \ldots, t_s \in \mathbb{T}^n$ der Liste `TermList` mit naiven Lösungsmethoden für lineare diophantische Gleichungen (siehe [KR00], Tutorial 36, S. 207). Berechnet wird ein Tupel $(c_1, \ldots, c_s) \in \mathbb{N}^s$ mit $t = t_1^{c_1} \cdots t_s^{c_s}$. Im Gegensatz zu `SB.TermReprToric` berechnet diese Funktion ein beliebiges Tupel unter allen möglichen.

Beispiel:

```
Use P::=QQ[x,y,z], DegLex;
TermList:=[x^2y, yz^2, xyz];
Term:=x^4y^3z^2;

SB.TermReprDio(Term, TermList);

[1, 0, 2]
-------------------------------
```

```
SB.TermReprToric(Term:POLY, TermList:LIST of POLY):LIST of
(LIST of INT)
```

Beschreibung:
Diese Funktion bildet die Implementation von Algorithmus 5.1. Sie berechnet alle Darstellungen eines Terms $t \in \mathbb{T}^n$ in den Termen $t_1, \ldots, t_s \in \mathbb{T}^n$ der Liste `TermList` berechnet, d.h. berechnet wird die Menge $C := \{(c_1, \ldots, c_s) \in \mathbb{N}^s : \ t = t_1^{c_1} \cdots t_s^{c_s}\}$. Dazu wird mit der CoCoA-Funktion `HilbertBasis` die Hilbertbasis des zugehörigen linearen diophantischen Gleichungssystems berechnet (siehe Algorithmus 2.3) und daraus die passenden Tupel ausgewählt. Diese Funktion wird standardmäßig bei Aufruf von `TermRepr` aufgerufen und dort dann das lexikographisch größte Element der Liste C ausgewählt.

Beispiel:

```
Use P::=QQ[x,y,z], DegLex;
TermList:=[x^2y, yz^2, xyz];
Term:=x^4y^3z^2;

SB.TermReprToric(Term, TermList);

[[2, 1, 0], [1, 0, 2]]
-------------------------------
```

Symbolverzeichnis

Grundlegende Mengen

$\mathfrak{P}(X)$	Potenzmenge einer Menge X.	
$\mathbb{N}$	Menge der natürlichen Zahlen (einschließlich 0).	
$\mathbb{N}_+$	Menge der positiven natürlichen Zahlen.	
$\mathbb{Z}$	Menge der ganzen Zahlen.	
$\mathbb{Z}_{m,n}$	Menge der ganzen Zahlen von $m \in \mathbb{Z}$ bis einschließlich $n \in \mathbb{Z}$.	
$\mathbb{Q}$	Menge der rationalen Zahlen.	
$\mathbb{R}$	Menge der reellen Zahlen.	
$\mathbb{C}$	Menge der komplexen Zahlen.	
$\overline{K}$	Algebraischer Abschluss eines Körpers K.	
K^*	$K \setminus \{0\}$ für eine Körper K.	
V^*	Der Dualraum eines K-Vektorraums V.	
$\mathbb{A}^n_K$	n-dimensionaler affiner Raum über einem Körper K.	
$\mathrm{Abb}(X, Y)$	Menge aller Abbildungen von X nach Y.	

Spezielle Mengen

$\mathcal{L}(\mathcal{A})$	Lösungsmenge eines linearen diophantischen Gleichungssystems mit Koeffizientenmatrix $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$.	
$\mathcal{L}_+(\mathcal{A})$	$\mathcal{L}(\mathcal{A}) \cap \mathbb{N}^n$.	
$\mathbb{E}^m$	Menge der erweiterten Terme im Laurent-Polynomring $K[y_1, \ldots, y_m, y_1^{-1}, \ldots, y_m^{-1}]$.	28
$\mathbb{T}^n$	Menge der Terme im Polynomring $K[x_1, \ldots, x_n]$.	
$\mathrm{Min}(k, \mathcal{M})$	Menge aller k-Minoren einer Matrix $\mathcal{M}$.	256
$\mathrm{Mat}_{l,n}(\mathcal{H})$	Menge aller $\mathcal{H}$-Matrizen bzgl. einer $m \times n$-Matrix $\mathcal{H}$.	258
Tab	Menge aller SO_2-Tableaus.	262
$\mathrm{Val}(T)$	Menge aller Werte der SO_2-Tableaus einer Menge $T \subseteq$ Tab.	263
STab	Menge aller Standard SO_2-Tableaus.	263
M^G	Menge aller Fixpunkte der Gruppenoperation von G auf M.	56
M/G	Bahnenraum	56
$\mathcal{N}_{V,G}$	Der Hilbertsche Nullstellenkegel bzgl. der Operation einer Gruppe G in einem K-Vektorraum V.	124
$G(x)$	Die Bahn eines Elements $x \in M$ bei Gruppenoperation von G auf M.	56
$\widehat{\mathbb{T}}$	Menge von Termen in den Unbestimmten $\{x_1, \ldots, x_n\} \setminus L$ für eine Menge $L \subseteq \{x_1, \ldots, x_n\}$	
$\mathrm{Ker}(f)$	Der Kern einer Abbildung f.	
$X /\!\!/ G$	Der algebraische Quotient.	155
$\mathcal{Z}_L(I)$	Nullstellenmenge eines Ideals $I \subseteq K[x_1, \ldots, x_n]$ mit einem Erweiterungskörper $L \supseteq K$.	34
$\mathcal{Z}_V(I)$	Nullstellenmenge eines Ideals $I \subseteq K[V]$ auf einer affinen K-Varietät V.	40
$\overline{B}$	Abschluss einer Menge B in einem topologischen Raum.	

Γ_φ	Graph eines K-Morphismus φ.	38
$\mathrm{Pix}_{i,j}$	Pixel mit Pixelindex (i,j).	172

Abbildungen

φ^*	Koordinatenabbildung eines K-Morphismus φ.	41
HF_M	Die Hilbert-Funktion eines Moduls M.	127
HS_M	Die Hilbert-Reihe eines Moduls M.	128
HP_M	Das Hilbert-Polynom eines Moduls M.	128
HN_M	Der Hilbert-Zähler eines Moduls M.	
Rey_G	Der Reynolds-Operator bzgl. G.	134
$(\rho, V)_G$	Eine lineare Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ einer Gruppe G in einem K-Vektorraum V.	58
f^a	Die durch $x \mapsto f(a^{-1}(x))$ für ein Element a einer Gruppe G und $f \in K[X]$ definierte Funktion aus $K[X]$.	59
T_u	Translation um einen Vektor u in $\mathbb{R}^n$.	
T_u^*	Für $u \in \mathbb{R}^2$ der durch $T_u^*(p) = p \circ T_u$ definierte Translationsoperator auf $\mathscr{P}_{\leq n}(\mathbb{R}^2, \mathbb{R})$.	190
gv	Grauwertfunktion oder Grauwertbild.	178

Matrizen und Tupel

$\mathcal{M}(i_1, \ldots, i_k \vert j_1, \ldots, j_k)$	k-reihige Teilmatrix einer $m \times n$-Matrix $\mathcal{M}$.	256
$\vert\mathcal{M}\vert$	Wert einer $\mathcal{H}$-Matrix $\mathcal{M}$.	259
$e_i(\mathcal{M})$	Für $i \in \{1,2\}$ Anzahl an e_i-Zeilen in einer $\mathcal{H}$-Matrix $\mathcal{M}$.	259
$\langle W \rangle$	Wert eines $\mathcal{H}$-Tupels W.	261
$\mathcal{M}_C^B(f)$	Darstellungsmatrix einer linearen Abbildung $f : V \to W$, wobei B eine Basis von V und C eine Basis von W ist.	
$\mathcal{M}_B^B(\rho)$	Darstellungsmatrix einer linearen Darstellung $\rho : G \to \mathrm{Aut}_K(V)$ von G in einem endlich-dimensionalen K-Vektorraum V bzgl. einer Basis B von V.	65

Gruppen

$\mathrm{Aut}(M)$	Menge aller bijektiven Abbildungen auf M.	
G^0	Zusammenhangskompente einer linearen algebraischen Gruppe $(G, *, e)$, die e enthält.	50
$Z(G)$	Zentrum einer Gruppe.	50
$\mathrm{Sta}_G(N)$	Stabilisator von $N \subseteq M$ bei Gruppenoperation von G auf M.	56
$\mathrm{Nor}_G(H)$	Normalisator einer Untergruppe H von G.	
$\mathrm{Add}(K)$	Die additive Gruppe $(K, +, 0)$.	49
$\mathrm{Mult}(K)$	Die multiplikative Gruppe $(K^*, \cdot, 1)$.	49
$\mathrm{ind}(G : H)$	Der Index einer Untergruppe H von G.	

Matrixgruppen

$\mathrm{GL}_n(K)$	Die n-dimensionale allgemeine lineare Gruppe.	48
$\mathrm{SL}_n(K)$	Die n-dimensionale spezielle lineare Gruppe.	49
$T_n(K)$	Die Gruppe der n-dimensionalen invertierbaren oberen Dreiecksmatrizen.	49
$\mathrm{O}_n(K)$	Die n-dimensionale orthogonale Gruppe.	49
$\mathrm{SO}_n(K)$	Die n-dimensionale spezielle orthogonale Gruppe.	49
$U_n(K)$	Die n-dimensionale unitäre Gruppe.	49
$\mathrm{SU}_n(K)$	Die n-dimensionale spezielle unitäre Gruppe.	49
$D_n(K)$	Die Gruppe der n-dimensionalen invertierbaren Diagonalmatrizen.	49
$P_n(K)$	Die Gruppe der n-dimensionalen Permutationsmatrizen.	49
$\mathrm{AGL}_n(K)$	Die n-dimensionale affine Gruppe.	52

$\mathrm{Trans}_n(K)$	Die n-dimensionale Translationsgruppe.	53
$\mathrm{Iso}_n(\mathbb{R})$	Die n-dimensionale Euklidische Gruppe.	56
$\mathrm{Iso}_n^+(\mathbb{R})$	Die eigentliche n-dimensionale Euklidische Gruppe.	56

Polynome

$\deg(f)$	Grad eines Polynoms f.	22
$\mathrm{LT}_\sigma(f)$	Leitterm eines Polynoms $f \neq 0$ bzgl. einer Termordnung σ.	23
$\mathrm{LM}_\sigma(f)$	Leitmonom eines Polynoms $f \neq 0$ bzgl. einer Termordnung σ.	23
$\mathrm{LC}_\sigma(f)$	Leitkoeffizient eines Polynoms $f \neq 0$ bzgl. einer Termordnung σ.	23
eval	Auswertungsfunktional, das jedem Polynom $f \in R[x_1, \ldots, x_n]$ seine Polynomfunktion $R^n \to R$ zuweist.	
eval_x	Auswertungsfunktional, das ein Polynom $f \in R[x_1, \ldots, x_n]$ an der Stelle x auswertet.	
$\mathrm{NF}_{\sigma,I}(f)$	Normalform von f bzgl. einer Termordnung σ und einem Ideal I.	
$\mathrm{NRS}_{\sigma,\mathcal{G}}(f)$	Normaler Unteralgebra-Rest bzgl. einer Termordnung σ und einem Tupel $\mathcal{G} = (g_1, \ldots, g_s)$.	82
$\mathrm{NF}_{\sigma,S}(f)$	SAGBI-Normalform von f bzgl. einer Unteralgebra $S \subseteq P$ und einer Termordnung σ.	85

Termordnungen

Lex	Lexikographische Termordnung auf $\mathbb{T}^n$.	23
DegLex	Graduiert-lexikographische Termordnung.	23
DegRevLex	Umgekehrt graduiert-lexikographische Termordnung.	23
$\mathrm{Elim}(L)$	Eliminationsordnung für eine Menge $L \subseteq \{x_1, \ldots, x_n\}$.	25
$\widehat{\sigma}$	Einschränkung einer Termordnung auf $\widehat{\mathbb{T}}$.	

Ideale

$\mathrm{Id}(R)$	Menge aller Ideale eines (kommutativen) Rings R.	
$\mathrm{Rad}(R)$	Menge aller Radikalideale von R.	
$\mathrm{LT}_\sigma\{I\}$	Monoideal von Termen aus I.	
$\mathrm{LT}_\sigma(I)$	Leittermideal des Ideals I.	
$\langle M \rangle$	Das von einer Menge M erzeugte Ideal.	
$\sqrt{I}$	Radikal eines Ideals I.	
$I :_P J$	Quotientenideal von I durch J im Polynomring P.	
$\mathrm{Ann}_P(J)$	Annulator von J in einem Polynomring P.	
fI	Das von $\{f \cdot g \mid g \in I\}$ für $f \in K[y_1, \ldots, y_m]$ und $I \subseteq K[x_1, \ldots, x_n]$ erzeugte Ideal in $K[x_1, \ldots, x_n, y_1, \ldots, y_m]$.	
IQ	Das von den Polynomen eines Ideals $I \subseteq K[x_1, \ldots, x_n]$ im Polynomring $Q = K[x_1, \ldots, x_n, y_1, \ldots, y_m]$ erzeugte Ideal.	
$I :_P J^\infty$	Saturierung von I durch J in einem Polynomring P.	
$\triangle_\varphi$	Diagonalideal eines K-Algebra-Homomorphismus $\varphi : P' \to P$.	26
$\mathrm{Rel}(\mathcal{G})$	Das Relationenideal eines Tupels $\mathcal{G} \in P^m$.	27
$I(\mathcal{A})$	Zur Matrix $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ gehörige torische Ideal.	28
$I(t_1, \ldots, t_n)$	Zu Termen $t_1, \ldots, t_n \in \mathbb{T}^n$ gehörige torische Ideal.	28
HI_G	Das Hilbert-Ideal.	142
$\mathcal{I}(V)$	Verschwindungsideal in $K[x_1, \ldots, x_n]$ einer affinen K-Varietät $V \subseteq \mathbb{A}_L^n$.	36
$\mathcal{I}_V(W)$	Verschwindungsideal in $K[V]$ einer Teilmenge W von V.	40

Ringe

$K[x_1, \ldots, x_n]$	Polynomring über einem Körper K in den Unbestimmten $x_1, \ldots, x_n$.
$K[x_1, \ldots, x_n]_d$	$\{f \in K[x_1, \ldots, x_n] : \deg(t) = d \text{ für alle } t \in \mathrm{Supp}(f)\}$ mit $d \in \mathbb{N}$

$\mathscr{P}(K^n, K)$	Menge aller Polynomfunktionen von K^n nach K.	
$\mathscr{P}_d(K^n, K)$	Menge aller Polynomfunktionen von K^n nach K vom Grad $d \in \mathbb{N}$.	
$\mathscr{P}_{\leq d}(K^n, K)$	Menge aller Polynomfunktionen von K^n nach K vom Grad $\leq d \in \mathbb{N}$.	
$\mathrm{Mat}_{m,n}(M)$	Menge aller $m \times n$-Matrizen mit Einträgen aus M.	
$K[X]$	Koordinatenring einer affinen K-Varietät X.	
$K[X]^G$	Invariantenring der Operation einer Gruppe G auf $K[X]$.	111
R_+	Die Teilmenge $\bigoplus_{d>0} R_d$ eines standardgraduierten Rings R.	

Operatoren

$\#M$	Kardinalität einer Menge M	
$\mathrm{char}(K)$	Charakteristik eines Körpers K.	
$\dim_K(V)$	Vektorraumdimension eines K-Vektorraums V.	
log	Logarithmus auf $\mathbb{T}^n$.	22
$\xrightarrow{G}$	Durch G definierte Ersetzungsregel.	
$\xleftrightarrow{G}$	Die durch $\xrightarrow{G}$ definierte Äquivalenzrelation.	
$\xrightarrow{G}_{\mathrm{SS}}$	Ein Unteralgebra-Reduktionsschritt bzgl. einer Menge G von Polynomen.	83
$\xrightarrow{G}_{\mathrm{S}}$	Unteralgebra-Ersetzungsregel bzgl. einer Menge G von Polynomen.	83
$\xleftrightarrow{G}_{\mathrm{S}}$	Die durch $\xrightarrow{G}_{\mathrm{S}}$ definierte Äquivalenzrelation.	83
$\langle \cdot, \cdot \rangle$	Standardskalarprodukt in $\mathbb{R}^n$.	
$\| \cdot \|$	Vom Standardskalarprodukt in $\mathbb{R}^n$ induzierte Norm.	
$\langle \cdot, \cdot \rangle_\Phi$	Von einer injektiven linearen Abbildung $\Phi : V \to \mathbb{R}^n$ induziertes Skalarprodukt.	
$\| \cdot \|_\Phi$	Von $\langle \cdot, \cdot \rangle_\Phi$ induzierte Norm auf einem $\mathbb{R}$-Vektorraum V	

Literaturverzeichnis

[ACG96] Aslaksen, Helmer ; Chan, Shih-Ping ; Gulliksen, Tor: Invariants of S4 and the shape of sets of vectors. In: *Applicable Algebra in Engineering, Communication and Computing* 7 (1996), Nr. 1, S. 53–57

[AM04] Adem, Alejandro ; Milgram, R. J.: Invariants and Cohomology of Gropus. In: *Cohomology of Finite Groups*. Berlin-Heidelberg : Springer Berlin Heidelberg, 2004 (Grundlehren der mathematischen Wissenschaften), S. 89–113

[Art93] Artin, Michael: *Algebra*. Basel : Birkhäuser, 1993

[BB05] Burger, Wilhelm ; Burge, Mark J.: *Digitale Bildverarbeitung*. Springer Verlag, Berlin–Heidelberg, 2005

[Ber07] Berndt, Rolf: *Representations of Linear Groups*. Vieweg Verlag, Wiesbaden, 2007

[BH93] Bruns, Winfried ; Herzog, Jürgen: *Cohen-Macaulay Rings*. Cambridge University Press, 1993

[Bis11] Bishop, Christopher M.: *Pattern Recognition and Machine Learning*. 2. Springer, 2011

[Boo41] Boole, George: Exposition of a general theory of linear transformations. In: *The Cambridge mathematical journal* 3 (1841)

[Bor91] Borel, Armand ; Ewing, J.H. (Hrsg.) ; Gehring, F.W. (Hrsg.) ; Halmos, P.R. (Hrsg.): *Linear Algebraic Groups*. 2. Springer, New York, 1991 (Graduate Texts in Mathematics)

[Bos06] Bosch, Siegfried: *Algebra*. 6. Springer Berlin Heidelberg, 2006

[BW98] Becker, Thomas ; Weispfenning, Volker: *Gröbner Bases - a computational approach to commutative algebra*. 2. Springer, New York, 1998 (Graduate texts in mathematics)

[Cay45] Cayley, Arthur: On the theory of linear transformations. In: *Cambridge Math. J* (1845)

[CH92] Campbell, Sue A. ; Holmes, Philip: Heteroclinic cycles and modulated travelling waves in a system with D4 symmetry. In: *Physica D: Nonlinear Phenomena* 59 (1992), Nr. 1-3, S. 52–78

[CLO07] Cox, David ; Little, John ; O'Shea, Donal: *Ideals, Varieties and Algortihms*. 3. Springer Verlag, 2007

[Coh93] Cohen, Henri: *Graduate Texts in Mathematics*. Bd. 138: *A Course in Computational Algebraic Number Theory*. Springer Verlag, 1993

[Cri86] Crilly, Tony: The rise of Cayley's invariant theory (1841–1862). In: *Historia Mathematica* (1986), S. 241–254

[Cri88] Crilly, Tony: The decline of Cayley's invariant theory (1863–1895). In: *Historia mathematica* 15 (1988), S. 332–347

[CT95] Collins, Michael A. ; Thompson, Keiran C.: Group theory and the global functional shapes for molecular potential energy surfaces. In: Bonchev, Danail (Hrsg.) ; Rouvray, Dennis H. (Hrsg.): *Chemical Group Theory: Techniques and Applications.* Reading : Gordon and Breach Publishers, 1995, S. 191–234

[Dal95] Dalbec, John P.: Straightening Euclidean invariants. In: *Annals of Mathematics and Artificial Intelligence* 13 (1995), Nr. 1-2, S. 97–108

[Del02] Del Padrone, Allesandro: *CoCoA package invariants.* Genua, 2002

[Der99] Derksen, Harm: Computation of invariants for reductive groups. In: *Advances in Mathematics* (1999)

[DGPS15] Decker, Wolfram ; Greuel, Gert-Martin ; Pfister, Gerhard ; Schönemann, Hans: *Singular 4.0.2 - A computer algebra system for polynomial computations.* http://www.singular.uni-kl.de/. Version: 2015

[DK02] Derksen, Harm ; Kemper, Gregor: Computational Invariant Theory. In: Gamkrelidze, R V. (Hrsg.) ; Popov, V L. (Hrsg.): *Encyclopaedia of Mathematical Sciences– Invariant Theory and Algebraic Transformation Groups*, Berlin–Heidelberg, 2002 (Encyclopaedia of Mathematical Sciences - Invariant Theory and Algebraic Transformation Groups)

[Don97] Donner, Klaus: Image interpretation based on local transform characterization. In: *Pattern Recognition and Image Analysis* 7 (1997), S. 431–447

[Don09] Donner, Klaus: *Digitale Bild- und Signalverarbeitung - Skriptum zur Vorlesung.* Passau, 2009

[Dre04] Drezet, JM: Luna's slice theorem and applications. In: *Algebraic group actions and quotients* (2004), S. 1–50

[Els11] Elstrodt, Jürgen: *Maß- und Integrationstheorie.* 7. Berlin-Heidelberg : Springer Verlag, 2011

[FB06] Freitag, Eberhard ; Busam, Rolf: *Funktionentheorie 1.* 4. Berlin-Heidelberg : Springer, 2006

[FDFH92] Foley, James ; Dam, Andries van ; Feiner, Steven K. ; Hughes, John F.: *Computer Graphics. Principles and Practice.* 2. Reading, Mass. (USA), 1992

[Fis66] Fisher, Charles S.: The death of a mathematical theory: a study in the sociology of knowledge. In: *Archive for History of exact Sciences* (1966)

[Fis05] Fischer, Gerd: *Lineare Algebra.* 15. Vieweg Verlag, Wiesbaden, 2005

[Fis08] Fischer, Gerd: *Lehrbuch der Algebra.* 1. Vieweg Verlag, Wiesbaden, 2008

[For11] Forster, Otto: *Analysis 3.* 6. Wiesbaden : Vieweg+Teubner Verlag, 2011

[FP03] Forsyth, David A. ; Ponce, Jean: *Computer Vision - A Modern Approach.* Upper Saddle River : Prentice Hall by Pearson Education, Inc., 2003

[Fuc00] Fuchs, Erich: *Schnelle Quadratmittelapproximation in gleitenden Zeitfenstern mit diskreten orthogonalen Polynomen*, Universität Passau, Dissertation, 2000

[Gan86] Gantmacher, Felix R.: *Matrizentheorie*. Springer, 1986

[Gat00] Gatermann, Karin: *Computer Algebra Methods for Equivariant Dynamical Systems*. Springer Berlin Heidelberg, 2000 (Lecture Notes in Mathematics)

[Gat03] Gatermann, Karin: Applications of SAGBI-bases in dynamics. In: *Journal of Symbolic Computation* 35 (2003), Nr. 5, S. 543–575

[Gau01] Gauss, Carl F.: *Disquisitiones arithmeticae*. 1801

[GG99] Gatermann, Karin ; Guyard, Frédéric: Gröbner Bases, Invariant Theory and Equivariant Dynamics. In: *Journal of Symbolic Computation* 28 (1999), Nr. 1-2, S. 275–302

[GK00] Geissler, Katharina ; Klüners, Jürgen: Galois Group Computation for Rational Polynomials. In: *Journal of Symbolic Computation* 30 (2000), Nr. 6, S. 653–674

[GL98] Gatermann, Karin ; Lauterbach, Reiner: Automatic classification of normal forms. In: *Nonlinear Analysis, Theory, Methods and Applications* 34 (1998), S. 157–190

[GL00] Graf, Siegfried ; Luschgy, Harald: *Foundations of Quantization for Probability Distributions*. Springer-Verlag Berlin Heidelberg, 2000 (Lecture Notes in Mathematics). – 240 S.

[Göb98] Göbel, Manfred: A constructive description of SAGBI bases for polynomial invariants of permutation groups. In: *Journal of Symbolic Computation* (1998), S. 261–272

[Göb99] Göbel, Manfred: A rewriting technique for universal polynomial invariants. In: *Information processing letters* 69 (1999), S. 271–273

[Göb00] Göbel, Manfred: Rings of polynomial invariants of the alternating group have no finite SAGBI bases with respect to any admissible order. In: *Information Processing Letters* 74 (2000), Nr. 1, S. 15–18

[Göb01] Göbel, Manfred: Visualizing Properties of Comprehensive SAGBI Bases - Two Examples. In: *Applicable Algebra in Engineering, Communication and Computing* 12 (2001), Nr. 5, S. 429–435

[Gor68] Gordan, Paul: Beweis, dass jede Covariante und Invariante einer binären Form eine ganze Funktion mit numerischen Coefficienten einer endlichen Anzahl solcher Formen ist. In: *Journal für die Reine und Angewandte Mathematik* 69 (1868)

[Gor85] Gordan, Paul ; Kerschensteiner, Georg (Hrsg.): *Vorlesungen über Invariantentheorie*. Teubner Verlag, 1885

[GW10] Goodman, Roe ; Wallach, Nolan R.: *Symmetry, Representations and Invariants*. Springer, New York, 2010 (Graduate Texts in Mathematics)

[GWE04] Gonzalez, Rafael C. ; Woods, Richard E. ; Eddins, Steven L.: *Digital Image Processing Using MATLAB*. Prentice Hall by Pearson Education, Inc., 2004

[Haa00] Haas, Jürgen: *Echtzeit-Korrespondenzprobleme in Bildsequenzen und Subpixelgenauigkeit*. Aachen, Universität Passau, Dissertation, 2000

[Hab75] Haboush, William J.: Reductive groups are geometrically reductive. In: *The Annals of Mathematics* (1975), S. 67–83

[Han10] Hanning, Tobias: *Rechnersehen.* Passau : Ralf Schuster Verlag, 2010

[Has80] Hashin, Z.: Failure Criteria for Unidirectional Fiber Composites. In: *Journal of Applied Mechanics* 47 (1980), S. 329–334

[Hav91] Havel, Timothy F.: Some examples of the use of distances as coordinates for Euclidean geometry. In: *Journal of Symbolic Computation* 11 (1991), Nr. 5-6, S. 579–593

[Hel93] Helisch, Wolfgang: *Invariantensysteme und Tensorgeneratoren bei Materialtensoren zweiter und vierter Stufe*, RWTH Aachen, Dissertation, 1993. – 159 S.

[Hil90] Hilbert, David: Ueber die Theorie der algebraischen Formen. In: *Mathematische Annalen* (1890)

[Hil93] Hilbert, David: Ueber die vollen Invariantensysteme. In: *Mathematische Annalen* (1893)

[HR74] Hochster, Melvin ; Roberts, Joel L.: Rings of invariants of reductive groups acting on regular rings are Cohen-Macaulay. In: *Advances in mathematics* 13 (1974)

[HS81] Horn, Berthold ; Schunck, Brian: Determining optical flow. In: *Artificial Intelligence* 17 (1981), Nr. 1-3, S. 185–203

[HS88] Harris, Chris ; Stephens, Mike: A combined corner and edge detector. In: *In Proc. of Fourth Alvey Vision Conference*, 1988, S. 147–151

[Hum81] Humphreys, James E. ; Gehring, F.W. (Hrsg.) ; Halmos, P.R. (Hrsg.) ; Moore, C.C. (Hrsg.): *Linear Algebraic Groups.* 2. Springer Verlag, 1981 (Graduate Texts in Mathematics)

[HZ06] Hartley, Richard ; Zisserman, Andrew: *Multiple View Geometry.* 2. Cambridge University Press, 2006

[Jäh05] Jähne, Bernd: *Digitale Bildverarbeitung.* 6. Berlin-Heidelberg : Springer Verlag, 2005

[JMS84] Jaric, Marko V. ; Michel, L. ; Sharp, R. T.: Zeros of covariant vector fields for the point groups: invariant formulation. In: *Journal Physics France* 45 (1984), S. 1–27

[KAT13] Kreuzer, Martin ; ApCoCoA-Team: *ApCoCoA 1.9.1 - Applied Computations in Computer Algebra.* http://www.apcocoa.org/. Version: 2013

[Kem03] Kemper, Gregor: Computing Invariants of reductive Groups in positive characteristic. In: *Transformation Groups* (2003)

[Kem09] Kemper, Gregor: Separating invariants. In: *Journal of Symbolic Computation* 44 (2009), Nr. 9, S. 1212–1222

[KK12] Kamke, Tobias ; Kemper, Gregor: Algorithmic Invariant Theory of Nonreductive Groups. In: *Qualitative Theory of Dynamical Systems* (2012), S. 1–25

[KM89] Kapur, Deepak ; Madlener, Klaus: A Completion Procedure for Computing a Canonical Basis for a k-Subalgebra. In: *IN COMPUTERS AND MATHEMATICS*, Springer, 1989, S. 1–11

[KM10] Karpfinger, Christian ; Meyberg, Kurt: *Algebra: Gruppen - Ringe - Körper*. 2. Spektrum Akademischer Verlag, 2010

[KR84] Kung, Joseph P. S. ; Rota, Gian-Carlo: The invariant theory of binary forms. In: *Bulletin of the American Mathematical Society* 10 (1984), Nr. 1, S. 27–86

[KR00] Kreuzer, Martin ; Robbiano, Lorenzo: *Computational Commutative Algebra*. Bd. 1. Springer Verlag, Berlin–Heidelberg–New York, 2000

[KR05] Kreuzer, Martin ; Robbiano, Lorenzo: *Computational Commutative Algebra*. Bd. 2. Springer Verlag, Berlin–Heidelberg–New York, 2005

[Kra85] Kraft, Hanspeter: *Geometrische Methoden in der Invariantentheorie*. Vieweg Verlag, Braunschweig–Wiesbaden, 1985 (Aspekte der Mathematik)

[Kra11] Kraft, Hanspeter: *Algebraic Transformation Groups - An Introduction*. http://math.unibas.ch/institut/personen/profil/profil/person/kraft/. Version: 2011

[Küh11] Kühnel, Wolfgang: *Matrizen und Lie-Gruppen*. Vieweg+Teubner, Wiesbaden, 2011

[Kun85] Kunz, Ernst: *Introduction to Commutative Algebra and Algebraic Geometry*. Birkhäuser, 1985

[Kun97] Kunz, Ernst: *Einführung in die algebraische Geometrie*. Vieweg Verlag, Wiesbaden, 1997

[LK81] Lucas, Bruce D. ; Kanade, Takeo: An Iterative Image Registration Technique with an Application to Stereo Vision. (1981), S. 674–679

[LS09] Laures, Gerd ; Szymik, Markus: *Grundkurs Topologie*. Spektrum Akademischer Verlag, 2009

[Mey92] Meyer, Friedrich Wilhelm F.: *Bericht über den gegenwärtigen Stand der Invariantentheorie*. 1892 (Jahresbericht der Deutschen Mathematiker Vereinigung)

[Mey98] Meyer, Friedrich Wilhelm F.: Invariantentheorie. In: *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*. Leipzig, 1898

[MFK65] Mumford, David ; Fogarty, John ; Kirwan, Frances C.: *Geometric invariant theory*. 1. Springer Berlin Heidelberg, 1965 (Ergebnisse der Mathematik und ihrer Grenzgebiete 34)

[MFK94] Mumford, David ; Fogarty, John ; Kirwan, Frances: *Geometric invariant theory*. 3. Springer Berlin, Heidelberg, New York, 1994 (Ergebnisse der Mathematik und ihrer Grenzgebiete 34)

[Mg14] Magma-group: *Magma Computational Algebra System*. http://magma.maths.usyd.edu.au/magma/. Version: 2014

[MZ92] Mundy, Joseph L. ; Zisserman, Andrew: *Geometric Invariance in Computer Vision*. Cambridge–Massachusetts, 1992

[Nag59] Nagata, Masayoshi: On the fourteenth problem of Hilbert. In: *American Journal of Mathematics* 81 (1959)

[Nag61] Nagata, Masayoshi: Complete reducibility of rational representations of a matrix group. In: *Journal of Mathematics of Kyoto University* 1 (1961)

[Neu07] Neusel, Mara D. ; Folland, Gerald B. (Hrsg.) ; Osgood, Brad (Hrsg.) ; Forman, Robin (Hrsg.) ; Starbird, Michael (Hrsg.): *Invariant Theory*. American Mathematical Society, 2007 (Student Mathematical Library 36)

[NM63] Nagata, Masayoshi ; Miyata, Takehiko: Note on semi-reductive groups. In: *Kyoto Journal of Mathematics* 3 (1963), S. 379–382

[Noe16] Noether, Emmy: Der Endlichkeitssatz der Invarianten endlicher Gruppen. In: *Mathematische Annalen* (1916)

[Noe26] Noether, Emmy: Der Endlichkeitssatz der Invarianten endlicher linearer Gruppen der Charakteristik p. In: *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen* (1926)

[Nor02] Nordbeck, Patrik: SAGBI Bases Under Composition. In: *Journal of Symbolic Computation* 33 (2002), Nr. 1, S. 67–76

[OR13] O'Connor, J. J. ; Robertson, E. F.: *Paul Albert Gordan*. http://www-history.mcs.st-andrews.ac.uk/Biographies/Gordan.html. Version: 2013

[Pis02] Pisinger, Georg: *Lokale Stützstrukturen zur Transformationspassung von Bildern*, Universität Passau, Dissertation, 2002

[Pop79] Popov, Vladimir L.: On Hilbert's theorem on invariants. In: *Dokl. Akad. Nauk SSSR* (1979)

[PV94] Popov, Vladimir L. ; Vinberg, Ernest B.: Invariant theory. In: Parshin, A.N. (Hrsg.) ; Shafarevich, I.R. (Hrsg.): *Algebraic Geometry IV*. Berlin-Heidelberg : Springer Verlag, 1994 (Encyclopaedia of Mathematical Sciences), Kapitel II

[Qv13] Querenburg von, Boto: *Mengentheoretische Topologie*. 3. Springer Verlag, Berlin-Heidelberg, 2013

[RAB15] Robbiano, Lorenzo ; Abbott, John ; Bigatti, Anna: *CoCoA System - Computations in Commutative Algebra*. http://cocoa.dima.unige.it/. Version: 2015

[Rei93] Reiss, Thomas H.: *Recognizing Planar Objects Using Invariant Image Features*. Springer Berlin Heidelberg, 1993 (Lecture Notes in Computer Science)

[Ric89] Richman, David R.: The Fundamental Theorems of Vector Invariants. In: *Advances in Mathematics* (1989)

[RS90] Robbiano, Lorenzo ; Sweedler, Moss: Subalgebra bases. In: *Commutative Algebra*. New York, Heidelberg, Berlin : Springer Verlag, 1990 (Lecture Notes in Mathematics), S. 61–87

[RS98] Robbiano, Lorenzo ; Sweedler, Moss: Ideal and Subalgebra Coefficients. In: *Proceedings of the American Mathematical Society*, 1998

[Sal66] Salmon, George: *Modern higher algebra*. 1866

[SK09] Schwarz, Rudolf ; Köckler, Norbert: *Numerische Mathematik*. 7. Vieweg+Teubner, Wiesbaden, 2009

[Slo77] Sloane, Neil J. A.: Error-correcting codes and invariant theory: New applications of a nineteenth-century technique. In: *The American Mathematical Monthly* 84 (1977), S. 82–107

[SP05] Santamaria, Mauricio V. ; Palacios, Roberto P.: Comparison of illumination normalization methods for face recognition. In: *Third COST 275 Workshop - Biometrics on the Internet* 275 (2005), Nr. 1, S. 1–4

[Spr70] Springer, Tonny A.: *Invariant Theory*. Springer Verlag, Berlin–Heidelberg, 1970 (Lecture Notes in Mathematics)

[Spr80] Springer, Tonny A. ; Coates, J. (Hrsg.) ; Helgason, S. (Hrsg.): *Linear Algebraic Groups*. Birkhäuser, 1980 (Progress in Mathematics)

[ST99] Stillman, Michael ; Tsai, Harrison: Using SAGBI bases to compute invariants. In: *Journal of Pure and Applied Algebra* 139 (1999), Nr. 1-3, S. 285–302

[Sta73] Stauduhar, Richard P.: The determination of Galois groups. In: *Mathematics of Computation* 27 (1973), Nr. 124, S. 981–981

[Sta79a] Stanley, Richard P.: Combinatorics and invariant theory. In: *Proceedings of Symposis in Pure Mathematics* 34 (1979), S. 345–355

[Sta79b] Stanley, Richard P.: Invariants of finite groups and their applications to combinatorics. In: *Bulletin of the American Mathematical Society* 1 (1979), Nr. 3, S. 475–512

[Sta07] Stadler, Thomas: *Korrespondenzfindung für Stereobilder mit breiter Stereobasis*, Universität Passau, Diplomarbeit, 2007. – 179 S.

[Stu96] Sturmfels, Bernd: *Gröbner Bases and Convex Polytopes*. American Mathematical Society, 1996 (University Lecture Series)

[Stu08] Sturmfels, Bernd ; Paule, Peter (Hrsg.): *Algorithms in Invariant Theory*. 2. Springer Verlag, Wien–New York, 2008 (Texts & Monographs in Symbolic Computation)

[TC92] Taubin, Gabriel ; Cooper, David B.: Object recognition based on moment (or algebraic) invariants. In: Mundy, Joseph L. (Hrsg.) ; Zisserman, Andrew (Hrsg.): *Geometric Invariance in Computer Vision*. MIT Press, 1992, S. 375–397

[Thi08] Thiéry, Nicolas M.: Algebraic invariants of graphs; a study based on computer exploration. (2008), S. 9–20

[TT02] Thiery, Nicolas ; Thomassé, Stephan: SAGBI bases of permutation groups and convex cones. In: *preprint* (2002), S. 1–4

[Wey46] Weyl, Hermann: *The Classical Groups. Their Invariants and Representations*. Princeton University Press, 1946 (Princeton mathematical series)

[Wol08] Wolfson, Paul R.: George Boole and the origins of invariant theory. In: *Historia Mathematica* 35 (2008), Nr. 1, S. 37–46

Stichwortverzeichnis

G

H

I

K

L

T

U

V

W

Z