universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

„Über die Konstruktion minimaler linearer Darstellungen von Elementen des freien Schiefkörpers (freier assoziativer Algebren)“

verfasst von / submitted by

Dipl.-Ing. Konrad Schrempf

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 405
Dissertationsgebiet lt. Studenblatt / field of study as it appears on the student record sheet:	Mathematik
Betreut von / Supervisor:	ao. Univ.-Prof. Dr. Karl Auinger

Danksagung

Zum Glück gibt es hier zu wenig Platz, um all jene aufzulisten, die meine Arbeit (direkt oder indirekt) erschwert haben. Aber es gibt auch Menschen, denen Wissenschaft wichtig ist. Dazu gehören Wolfgang Tutschke und Thomas Hirschler, die mir zum jeweils richtigen Zeitpunkt geholfen haben, dort einen Weg zu finden, wo es keinen zu geben schien. Darüber hinaus gibt es noch einige, die ich nicht nennen möchte, weil ich vermeiden will, dass sie Schwierigkeiten bekommen, weil sie Verständnis für das Wehren gegen Unrecht aufgebracht haben. Sie alle haben mich dadurch bestärkt, diesen Weg zu gehen, auf dessen Zielgeraden mich schließlich Karl Auinger unterstützt hat. Herzlichen Dank!

Diese Arbeit ist all den
motivierten und leidenschaftlichen
Mathematikern gewidmet, die nie
eine Chance bekommen, ihr
Talent zu zeigen.

Abstract: On the construction of Minimal Linear Representations of Elements in the Free Field

(of Free Associative Algebras)
aka "Calculating with Free Fractions"

Long before we learn to construct the (field of) *rational numbers* (out of the ring of *integers*) at university, we learn how to calculate with *fractions* at school. When it comes to "numbers", we are used to a *commutative* multiplication, for example $2 \cdot 3 = 6 = 3 \cdot 2$. On the other hand —even before we can write— we learn to talk (in a language) using *words*, consisting of purely *non-commuting* "letters" (or symbols), for example $xy \neq yx$ (with the concatenation as multiplication). Now, if we combine numbers (from a field) with words (from the *free monoid* of an alphabet) we get *non-commutative polynomials* which form a ring (with "natural" addition and multiplication), namely the *free associative algebra*. Adding or multiplying polynomials is easy, for example $(\frac{2}{3}xy + z) + \frac{1}{3}xy = xy + z$ or $2x(yx + 3z) = 2xyx + 6xz$. Although the integers and the non-commutative polynomials look rather different, they share many properties, for example the unique number of *irreducible* factors: $x(1 - yx) = x - xyx = (1 - xy)x$. However, the construction of the *free field* (this is the "universal" *field of fractions* of the *free associative algebra*, that is, the non-commutative counterpart to the field of *rational numbers*) is very difficult, the development of the theory took almost four decades. And until now —that is, since almost additional five decades— its practical application was hardly possible, because there were —except for special cases— no algorithms to *minimize linear representations* ("generalized non-commutative fractions"), corresponding to *cancelling* of "classical" fractions. In this work a comprehensive theory is derived that permits to calculate with elements in the free field by means of *free fractions* (that is, to *add*, *multiply* or —if non-zero— *invert* them). Additionally, algorithms are developed which facilitate an implementation in computer algebra systems, with a multitude of applications.

Remark. Appendix E is a detailed introduction (with references to the main papers) and an extended table of contents with additional information in English.

Remark. The development of a theory is hardly ever linear. Inconsistencies *between* the publications should therefore be seen as an indication of a mathematical voyage of discovery. The presentation here however is from a different perspective (looking back, putting everything into a more general context) and —at least the German version— should be sufficiently coherent to be useful.

Remark. Judging the importance of single results in a highly interwoven theory is difficult. Talking about the visibility is easier. Beside Theorem 2.5.13, Theorem 3.3.6 and Algorithm 4.5.15 some are less visible: Remark 2.3.12, Proposition 3.3.7 and Lemma 4.2.2 for the implementation, Corollary 2.5.15 and Theorem 3.4.9 from an algebraic point of view and Lemma 2.3.2 and its role in Theorem 4.5.9 for the underlying concepts.

Zusammenfassung: Über die Konstruktion minimaler linearer Darstellungen von Elementen des freien Schiefkörpers

(freier assoziativer Algebren)
alias „Das Rechnen mit freien Brüchen“

Lange bevor wir uns mit der Konstruktion (des Körpers) der *rationalen Zahlen* (aus dem Ring der *ganzen Zahlen*) an der Universität beschäftigen, lernen wir in der Schule mit *Brüchen* zu rechnen. Bei „Zahlen“ sind wir eine *kommutative* Multiplikation gewöhnt, z.B. $2 \cdot 3 = 6 = 3 \cdot 2$. Auf der anderen Seite —sogar noch vor dem Schreiben— lernen wir das Sprechen mit *Wörtern*, die aus *nicht-kommutierenden* „Buchstaben“ (oder Symbolen) bestehen, z.B. $xy \neq yx$ (mit der Aneinanderreihung als Multiplikation). Nun, wenn wir Zahlen (eines Körpers) mit Wörtern (des *freien Monoids* eines Alphabets) kombinieren, erhalten wir *nicht-kommutative Polynome*, die einen Ring (mit „natürlicher“ Addition und Multiplikation) bilden, nämlich die *freie assoziative Algebra*. Addieren oder multiplizieren zweier Polynome ist einfach, z.B. $(\frac{2}{3}xy + z) + \frac{1}{3}xy = xy + z$ oder $2x(yx + 3z) = 2xyx + 6xz$. Obwohl die ganzen Zahlen und die nicht-kommutativen Polynome ziemlich verschieden sind, teilen sie viele Eigenschaften, zum Beispiel die Eindeutigkeit der Anzahl *irreduzibler* Faktoren: $x(1 - yx) = x - xyx = (1 - xy)x$. Allerdings ist die Konstruktion des *freien Schiefkörpers* (das ist der „universelle“ *Quotientenkörper* der *freien assoziativen Algebra*, also das nicht-kommutative Pendant zum Körper der *rationalen Zahlen*) sehr schwierig, die Entwicklung der Theorie hat beinahe vier Jahrzehnte gedauert. Und bis dato —also seit weiteren knapp fünf Jahrzehnten— war dessen praktische Anwendung kaum möglich, weil es —außer für Spezialfälle— keine Algorithmen gab, um *lineare Darstellungen* („verallgemeinerte nicht-kommutative Brüche“) zu *minimieren* (was in etwa dem *Kürzen* „klassischer“ Brüche entspricht). In dieser Arbeit wird eine umfassende Theorie hergeleitet, die es erlaubt, mit Elementen des freien Schiefkörpers mittels *freier Brüche* zu rechnen (das heißt, sie zu *addieren*, *multiplizieren* oder —wenn ungleich Null— *invertieren*). Zusätzlich werden Algorithmen entwickelt, die eine Implementierung in Computer-Algebra-Systemen ermöglichen, mit einer Vielzahl an Anwendungen.

Bemerkung. Die Entwicklung einer Theorie verläuft kaum jemals linear. Die Präsentation hier jedoch ist aus einer anderen Perspektive (zurückblickend, in größeren Kontext stellend) und sollte hinreichend kohärent sein, um sie auch nutzen zu können.

Bemerkung. Die Wichtigkeit einzelner Resultate in einer hochgradig verwobenen Theorie zu beurteilen ist schwierig. Über die Sichtbarkeit zu sprechen ist einfacher. Neben Satz 2.5.13, Satz 3.3.6 und Algorithmus 4.5.15 sind ein paar weniger sichtbar: Bemerkung 2.3.12, Proposition 3.3.7 und Lemma 4.2.2 für die Implementierung, Korollar 2.5.15 und Satz 3.4.9 aus einer algebraischen Perspektive und Lemma 2.3.2 und seine Rolle in Satz 4.5.9 für die zugrundeliegenden Konzepte.

Zitierungen, Textstruktur und Orthographie

In Bezug auf die saubere wissenschaftliche Praxis *alle* externen Quellen penibel aufzulisten mache ich nur eine Ausnahme, nämlich bei Übersetzungen von kleineren Bemerkungen beziehungsweise von Fließtext und Beispielen meiner eigenen (relevanten) Arbeiten, weil das —meiner Ansicht nach— mehr stören als helfen würde. Eigene wesentliche Aussagen werden natürlich klar gekennzeichnet, damit nachvollziehbar ist, *wo* etwas zum *ersten Mal* formuliert oder bewiesen wurde, weil dort der Kontext und auch das *damalige* „Wissen“ viel klarer erkennbar ist. Trotzdem beharre ich nicht auf der alten, manchmal schwerer zu lesenden Darstellung, merke aber an, wenn etwas verändert oder verallgemeinert wurde.

Im wesentlichen halte ich mich an die mathematische Praxis, nicht vorzugreifen. Kleinere Ausnahmen sind dem Kompromiss geschuldet, den wesentlichen Teil der Theorie möglichst kurz zu halten und Ergänzungen in den Anhang zu verschieben. „Kurz“ heißt aber *nicht*, dass es keinerlei Überlappungen mehr gibt. Anhand von *nicht-trivialen* Spezialfällen lassen sich die grundlegenden Konzepte meist viel besser nachvollziehen. Diese sind dann —mit den Beispielen— in die allgemeine Theorie eingebettet.

Für das Einfügen von Gedanken verwende ich die spanische Variante von Gedankenstrichen, in der klarer hervorgeht, *was* eingefügt wurde. Zusätzliche Information kann auch in Klammern stehen, vor allem dann, wenn etwas in den Hintergrund —als Gegensatz zur Betonung— rücken soll. Begriffe in Anführungszeichen können je nach Kontext verschiedenes bedeuten: die Einführung beziehungsweise die Vorwegnahme von Definitionen, eine Abgrenzung zu einer exakten Definition im Sinne einer Verallgemeinerung oder die Verwendung im übertragenen Sinn. Dort, wo es möglich ist, werden Begriffe in Anführungszeichen durch kursive Schrift ersetzt. Um die Lesbarkeit zu verbessern, werden Teile von Eigennamen (z.B. Adjektive) nur dann mit großem Anfangsbuchstaben geschrieben, wenn es sich um einen alleinstehenden Begriff handelt oder eine exakte Abgrenzung notwendig ist. Wenn also zum Beispiel von der „freien assoziativen Algebra“ die Rede ist, sollte spätestens nach dem Vorwort klar sein, dass es sich hier um einen Fachbegriff handelt. Akronyme und Abkürzungen werden dann verwendet, wenn sie sehr häufig vorkommen, die Lesbarkeit erhöht wird und sie allgemein verständlich sind.

Was nicht hier steht

Hier gibt es *keinerlei* Einführung in die mathematischen Grundlagen. Mit *linearer Algebra* sollte man jedenfalls gut vertraut sein, eine Einführung in die *Algebra* sollte man nicht nur zur Hand haben, sondern sich darin gut zurecht finden. *Nicht-kommutative Algebra* und *Faktorisierung in Ringen* sind mathematische Spezialgebiete und als solche etwas schwerer zugänglich, das heißt, man muss sich die Theorie mit einigem Aufwand selbst erarbeiten. Das Literaturverzeichnis sollte dabei helfen, einen individuellen Weg zu finden.

Vorwort

Oder: Über die Sprache dieser Arbeit und warum sich jemand die Mühe macht, bereits (in englischer Sprache) veröffentlichte Arbeiten völlig anders zu präsentieren, wo es doch mancherorts reicht, alles zusammenzuklammern (um einen Doktor zu bekommen) und bereits ein gemeinsames Literaturverzeichnis unter den Begriff Luxus fällt.

Doch *wo* sonst wäre Raum für das „Herz der Mathematik“? [Hal80] Wie soll man sich in der Fülle von Details und Literatur zurechtfinden, *insbesondere* dann, wenn man nicht gerade Mathematiker mit Spezialgebiet „nicht-kommutative Algebra“ ist? Hier soll der Versuch unternommen werden, mehrere „Sprachen“ miteinander zu verweben. Alle einzuladen, Bleistift und Papier zur Hand zu nehmen, um die Schönheit der Mathematik zu entdecken.

Keinesfalls soll es so wie in Robert Musils „Die Verwirrungen des Zöglings Törless“ sein, als es um die *komplexen Zahlen* geht, und die Schwierigkeit sich etwas Imaginäres vorzustellen. Denn schließlich gibt es ja noch „komplexere“, nämlich die *Hamiltonschen Quaternionen*, deren Entdeckung (vor über 150 Jahren) alles andere als einfach war [vdW73]. Oft hilft es etwas zu verstehen, wenn man sieht, was *nicht* funktioniert. Und dafür sollte es irgendwo auch Platz geben.

Wer —außer den Mathematikern— denkt beim Bruchrechnen darüber nach, mit *was* man da eigentlich rechnet? Und ob man sich auch sicher sein kann, dass man jede ganze Zahl[1] als Produkt von Primzahlen (z.B. $12 = 2 \cdot 2 \cdot 3$) schreiben kann, um dann gegebenenfalls (Zähler und Nenner) zu kürzen? Ja selbst bei den Mathematikern weiß man sofort, in welchem Fach sie zuhause sind, wenn sie bei den ganzen Zahlen an einen *ZPE-Ring*[2] (ja sogar Hauptidealring) denken.

Es war eine an sich unscheinbare Bemerkung von Roland Speicher im Februar 2017

[1] Klarerweise sind hier ganze Zahlen *ohne* der Null und den Einheiten (d.h. inventierbarer Elemente) ± 1 gemeint. Und eine Faktorisierung (in Primzahlen) ist natürlich *modulo* dem Einfügen von Einheiten, z.B. $12 = 2 \cdot 2 \cdot 3 = 2(-1) \cdot (-1)2 \cdot 3$, zu verstehen

[2] Ein ZPE-Ring ist ein Ring, in dem nicht-invertierbare Elemente (mit Ausnahme der Null) „eindeutig“ in Primelemente zerlegt werden können. Zur Wiederholung: In einem Ring kann man *addieren* und *multiplizieren*, aber nicht notwendigerweise „dividieren“. Zur Verwirrung: Leider wird man später Primelemente vergeblich suchen, weil (unter anderem) im Ring der ganzen Zahlen Primelemente und irreduzible Elemente (Atome) zusammenfallen, das heißt, zwei unterschiedliche Konzepte stimmen überein. Im allgemeinen ist das nicht der Fall und im Nicht-Kommutativen gibt es „oft“ überhaupt „zu wenige“ Primelemente.

in Saarbrücken, die mich viel beschäftigt hat. Sinngemäß meinte er (in Bezug auf die nicht-kommutativen rationalen Funktionen) ob es nicht an der Zeit wäre, an die grundlegende Theorie zu glauben[3] (so, wie es ja auch bei den *rationalen Zahlen* ist) und einfach einmal damit zu rechnen? Nach und nach kann man dann tiefer und tiefer in die schier bodenlose Welt der nicht-kommutativen Algebra eintauchen.

Und in gewisser Weise ist mein Beitrag nur das Zusammenstellen von anwendbaren Rechenregeln für „nicht-kommutative Brüche“, insbesondere für die Implementierung in Computer-Algebra-Systemen (CAS) zur praktischen Verwendung. Die grundlegende Theorie gibt es zum Teil schon seit vielen Jahrzehnten, viele Fragestellungen gehen an den Anfang des zwanzigsten Jahrhunderts zurück. Streifzüge durch die Literatur zeigen eine bunte Welt der Mathematik, in der natürlich die Gefahr besteht, dass man sich verliert ...

Auch soll es niemandem so wie (der Romanfigur) Törless gehen, dem der Kopf schon nach den ersten Seiten von *Kant* schwirrte. Zwar kann ich das in Bezug auf „Die Kritik der reinen Vernunft“ nur teilweise nachvollziehen, doch an die ersten Seiten von *Cohn* kann ich mich noch sehr gut erinnern. Es waren ein paar Anläufe notwendig, um so weit im Vorwort zu kommen, dass ich etwas erahnte, das mich nicht mehr los ließ. Mit dafür verantwortlich war sicher auch, in zwei seiner Bücher, nämlich [Coh77, Coh95], auf folgendes Zitat aus Goethes Faust zu stoßen:

> O glücklich, wer noch hoffen kann,
> aus diesem Meer des Irrtums aufzutauchen!
> Was man nicht weiß, das eben brauchte man,
> Und was man weiß, kann man nicht brauchen.

In diesem Sinne soll der Übergang zwischen umgangssprachlich und mathematisch verwendeten Begriffen fließend sein. Mit Wörtern/Monomen (unser Alphabet hier besteht üblicherweise nur aus den drei Buchstaben x, y und z) beziehungsweise *Polynomen* kann man „rechnen“, z.B. $(\frac{2}{3}xy + z) + \frac{1}{3}xy = xy + z \neq yx + z$ oder $x \cdot (yx + z) = xyx + xz$. Und unser Ziel wird es sein, auch Polynome (üblicherweise mit p oder q bezeichnet) zu invertieren, z.B. $(xy - yx)^{-1} = \frac{1}{xy-yx}$. Bei diesem „einfachen“ Beispiel sollte man länger innehalten. Schon die Darstellung mit dem *Bruchstrich* ist nicht immer geeignet, weil pq^{-1} im allgemeinen eben *nicht* $q^{-1}p$ ist. Und was passiert,

[3] Hier würden bei der Verwendung des Wortes „glauben“ nicht einmal Anführungszeichen helfen. Glauben bedeutet nicht, sich blind auf irgendetwas zu verlassen. Es geht darum, sich *selbst* von der Plausibilität zu überzeugen. Auch darum geht es in dieser Arbeit. Die Details kann man dann sukzessive nachlesen. Aber das ist eben *keine* Kleinigkeit. Nur zur Orientierung: Während eine vierstündige (allgemeine) Algebravorlesung in jedem Mathematikstudium vorkommen sollte, fällt eine vertiefende Vorlesung über nicht-kommutative Algebra, das heißt, eine *Einführung* in die nicht-kommutative Algebra, meist schon unter eine Spezialisierung. Und während in letzterer die Konstruktion von Quotientenkörpern laut Ore [Ore31] Platz hat, ist es alleine schon aufgrund des Stoffumfanges unmöglich, über eine allgemeinere Einbettung zu diskutieren.

Das soll keinesfalls entmutigen! Ganz im Gegenteil. Wohl dosiert ist es herrlich, all das, was man aus dem Kommutativen kennt (z.B. Euklidscher Algorithmus, Hauptidealringe) allgemeiner zu betrachten (z.B. Schwacher Algorithmus, freie Idealringe).

Und 'mal ehrlich: Wer von den (nicht aktiv lehrenden) Mathematikern kann zum Beispiel die —im Studium gelernte— Konstruktion der reellen Zahlen jemand anderem aus dem Stegreif erklären?

wenn man für x, y und z —betrachtet als Variablen, die nicht kommutieren— Zahlen einsetzt? Für $q = xy - yx$ sollte man das unbedingt probieren! Und mit der Zeit gewöhnt man sich daran, dass es beliebig viele „Bruchstriche“ geben kann, z.B. bei $f = x^{-1}yz^{-1}$, das man noch irgendwie als „$x\backslash y/z$“ schreiben könnte (allgemeinere Elemente werden oft mit f oder g bezeichnet).

Nun, nachdem ja die Hamiltonschen Quaternionen bekanntlich einen Schiefkörper bilden, werden sich manche fragen, ob denn nicht auch das Rechnen mit rationalen Quaternionen als nicht-kommutatives Bruchrechnen bezeichnet werden kann. Ja. Diese „Art“ der Nicht-Kommutativität wird manchmal mit *schwach nicht-kommutativ*[4] als Abgrenzung zur „freien“ Nicht-Kommutativität bezeichnet. Tatsächlich würde man mit *einem einzigen* Bruchstrich auskommen, im Zähler und im Nenner würden dann ganze Quaternionen stehen. Gibt es denn irgendetwas „dazwischen“? Interessierte seien auf [Reu96a], [HS07] und [HS15] als mögliche Ausgangspunkte verwiesen.

Da selbst der Sprachgebrauch im Englischen alleine zu Verwirrung und mitunter zu falschen Aussagen führt, muss man natürlich besonders bei Übersetzungen vorsichtig sein. Aber zum einen erklärt Paul M. Cohn jedes noch so winzige Detail —so ist für ihn ein *Körper* nicht notwendigerweise *kommutativ* und „eindeutig“ kann etwas großzügiger verstanden werden, um ein Konzept sinnvoll zu verallgemeinern— und zum anderen kann man über seine ausführlichen Literaturverweise beliebig tief eintauchen (und sich total verlieren). Überblicksmäßig kann man viel in seinen Kommentaren am Ende jedes Kapitels erfahren, z.B. in [Coh85, Kapitel 7], nämlich wielange es gedauert hat (und wieviele Mathematiker daran mitgewirkt haben), um eine Theorie zu entwickeln: „Bis 1970 basierten die einzigen rein algebraischen Methoden der Einbettung von Ringen in Körpern auf Ores Methode [Ore31].“[5] Von Cohn gibt es auch eine kleinere Arbeiten in deutscher Sprache, z.B. [Coh82b].

Als Übersetzung für „free field“ verwende ich „freier Schiefkörper“ (anstatt „freier Körper“) für eine klarere Abgrenzung zum *kommutativen* (Grund-)Körper (z.B. der rationalen, reellen oder komplexen Zahlen). Damit bekäme man folgende Kette der „Inklusionen“: (kommutativer) Körper $\subsetneq$ Schiefkörper $\subsetneq$ freier Schiefkörper.

Die „Einbettung“ der ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ (als Brüche mit Nenner Eins) in die rationalen Zahlen $\mathbb{Q}$ sollte einigermaßen geläufig sein. Salopp formuliert starten wir hier mit einem *kommutativen* Körper $\mathbb{K}$ und erzeugen gemeinsam mit einem Alphabet X die *nicht-kommutativen* Polynome $\mathbb{K}\langle X\rangle$, die dann in deren *freien Schiefkörper* $\mathbb{F}$ eingebettet werden. Präziser formuliert: Für die *Algebra* der nicht-kommutativen Polynome $\mathbb{K}\langle X\rangle$, das heißt, die *freie assoziative Algebra* (über $\mathbb{K}$ und X), gibt es einen Ring-Homomorphismus in einen „universellen“ *Quotientenkörper* $\mathbb{F} = \mathbb{K}(\langle X\rangle)$. Die Theorie dazu ist alles andere als trivial.[6]

[4] Im Englischen sagt man meist „mild non-commutativity“ dazu.

[5] Die Übersetzung stammt von mir. Im folgenden werde ich nicht mehr extra darauf hinweisen. Auch bitte ich um Verständnis, nicht auf den Ursprung einzelner deutscher mathematischer Fachbegriffe (und in welchem Lehrbuch man sie findet) zu verweisen. Dort, wo es wesentlich ist, werde ich die englischen Begriffe anführen. Manchmal erwähne ich auch Fachwörter aus anderen mathematischen Bereichen um Verwechslungen oder Fehlinterpretationen zu vermeiden. Dass das in der gesamten Arbeit verwendete Fachvokabular sehr umfangreich ist, sollte nicht unterschätzt werden.

[6] Nachdem vorher bereits Oystein Ore erwähnt wurde: Seine Konstruktion sollte in allen Büchern

Damit ist immerhin schon der halbe Titel dieser Arbeit erklärt. Wenn man mit Brüchen rechnet, z.B. $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$, denkt man nicht notwendigerweise daran, dass man damit *Elemente* eines Körpers (z.B. der rationalen Zahlen) darstellt. Aber klar ist, dass es *mehrere* verschiedene Darstellungen für ein und das selbe Element gibt[7], z.B. $\frac{-1}{2} = \frac{3}{-6} = -\frac{2}{4}$. Kaum jemand würde es als *Wortproblem* bezeichnen, festzustellen, ob zwei Brüche „gleich“ sind, das heißt sie das selbe Element *repräsentieren*. Das „eigentliche“ Wortproblem hier ist einfach, weil man Buchstabe für Buchstabe vergleichen kann, z.B. um festzustellen, dass $xyz \neq xyx$ ist. Und für Polynome ergibt sich das quasi natürlich, in dem man Koeffizienten und Wörter überprüft. Im allgemeinen, nämlich für Elemente des freien Schiefkörpers $\mathbb{F}$, ist das nicht mehr ganz so einfach, weil die Darstellungen komplizierter werden. Wie man hier vorgehen kann, ist Inhalt dieser Arbeit.

Der erste Schritt ist die Einführung *linearer* Darstellungen (von Elementen des freien Schiefkörpers). Ein Bruch kann direkt als lineare Darstellung betrachtet werden. Und das ist auch ein einfacher Einstieg, nämlich die vielleicht komplizierteste Art, mit Brüchen zu rechnen. (In der Einleitung kommen wir darauf zurück.) Als zweiten Schritt kann man damit nicht nur mit (nicht-kommutativen) Polynomen[8] rechnen, man kann sie auch verwenden, um sie zu faktorisieren, z.B. $x - xyx = x(1 - yx) = (1 - xy)x$, ähnlich der Zerlegung von ganzen Zahlen in Primfaktoren. Hier sollte man dann —wenn man so weit gekommen ist— etwas innehalten, denn bis hier her ist es relativ einfach zu sehen, was „in“ der linearen Darstellung passiert. Als dritten Schritt kann man über Elemente nachdenken, die man als „Bruch“ von zwei Polynomen pq^{-1} mit „rechtem“ Nenner $q \in \mathbb{K}\langle X\rangle$ darstellen kann, nachdenken. Wobei hier zunächst q eine spezielle Form hat, nämlich der konstante Term, das ist der Koeffizient vom *leeren Wort*, darf nicht verschwinden, z.B. $q = 1 - xy$. Warum ist $q^{-1} = 1 + xy + (xy)^2 + (xy)^3 + \ldots$? (Wer das noch nie händisch gerechnet hat, sollte das unbedingt selbst nachprüfen: $qq^{-1} = \ldots$ und $q^{-1}q = \ldots$ Ein paar Terme genügen, um zu sehen, was hier gemeint ist. Gegebenenfalls kann man z.B. $x = y = \frac{1}{10}$ einsetzen.)

Als vierten Schritt kann man einen allgemeineren Nenner $q \neq 0$ betrachten. Im einfachsten Fall kann man dann auch schon „innen“ kürzen, z.B. $(x - xyx)(zx)^{-1} = (1 - xy)x \cdot x^{-1}z^{-1} = (1 - xy)z^{-1}$. Auch hier sollte man etwas innehalten, denn was in einem konkreten Beispiel so selbstverständlich aussieht, ist es nicht. Zu guter Letzt kann man dann den letzten Verallgemeinerungsschritt machen, nämlich *beliebige*

zu finden sein, in denen es um nicht-kommutative Algebra geht. Der Abstraktionsgrad ist aber zum Teil sehr unterschiedlich. Als Einstieg ist seine Originalarbeit [Ore31] besonders zu empfehlen, da er tatsächlich Brüche verwendet. Für eine „modernere“ Schreibweise könnte man [Coh03b, Abschnitt 7.1] oder [Coh06a, Abschnitt 0.7] heranziehen. Üblicherweise beginnt man mit der Konstruktion von Schiefpolynomringen [GW89, Kapitel 1] die man dann in ihren jeweiligen Schiefkörper einbetten kann.

[7] Man spricht deshalb von *Äquivalenzklassen*. Alle Darstellungen für ein (einzelnes) Element fasst man in einer Äquivalenzklasse zusammen.

[8] Nicht-Kommutativät wird im folgenden nicht extra erwähnt. Nur in wenigen Abschnitten kommen auch *kommutative* Polynome vor, dort wird dann darauf hingewiesen. Tatsächlich entspricht die Faktorisierung in *Atome* (irreduzible Elemente) der Faktorisierung in Atome im Kommutativen. Und (nicht nur) in den ganzen Zahlen sind die Atome genau die Primelemente. Diese beiden Begriffe sauber zu trennen ist am Anfang nicht so einfach, weil einem üblicherweise die Beispiele fehlen.

Elemente des freien Schiefkörpers (via linearer Darstellungen) betrachten.

In der Präsentation ist eine Trennung in so kleine Teilschritte nicht möglich. Viel mehr geht es darum, sich —vor allem in den Beispielen— bewusst zu machen, wo man gerade steht und welche Information man eigentlich braucht. An einzelnen Stellen wird explizit darauf hingewiesen. Zum Schluss hat man sozusagen eine „Black Box“, in die man in Spezialfällen hineinschauen kann. Diese Vorgehensweise ist auch die empfohlene für die Implementierung am Computer. Zwar ist die Programmierung ziemlich aufwändig und technisch, dafür kann man dann tatsächlich fast wie mit Brüchen rechnen ...

Und zum Trost für all jene, die Stunden, Tage und Wochen damit verbringen, um halbwegs nachvollziehen zu können, was im folgenden eigentlich steht, sei daran erinnert, wieviele Wochen, Monate und Jahre notwendig waren, um etwas zu finden, von dem man nicht wusste, *was* es eigentlich ist, geschweige denn, *wo* man danach suchen soll. Und das aufbauend auf eine Theorie, deren Entwicklung mehrere Jahrzehnte dauerte. In diesem Sinne wünsche ich viel Spaß beim Rechnen mit (nicht-kommutativen) „freien Brüchen“!

Inhaltsverzeichnis

Abbildungs- und Tabellenverzeichnis

Kapitel 1

Einleitung

Die Konstruktion des *universellen Quotientenkörpers* $\mathbb{F} = \mathbb{K}(\langle X \rangle)$ einer *freien assoziativen Algebra* $\mathbb{K}\langle X \rangle$ (über einem *kommutativen* Körper $\mathbb{K}$ und einem endlichen Alphabet X) ist alles andere als einfach. Neben verschiedenen „theoretischen“ Zugängen (siehe z.B. [Reu99] oder [Coh06b]) gibt es zwei „praktische“: Entweder gibt man (rationalen) Ausdrücken der Form $(xy - yx)^{-1}$ einen Sinn, indem man Matrizen einsetzt [KVV14] oder „volle“ Matrizen invertiert [Coh03b, Abschnitt 9.3]. Beide Zugänge sind sehr diffizil. Für das Beispiel, den *Kommutator* $f(x, y) = xy - yx$, findet man 2×2 Matrizen $\bar{x}$ und $\bar{y}$, sodass $f(\bar{x}, \bar{y}) = \bar{x}\bar{y} - \bar{y}\bar{x} \neq 0$ ist. (Bitte probieren vorm Weiterlesen.) Doch für ein Alphabet $X = \{w, x, y, z\}$ und $g(w, x, y, z) = wxyz - xwyz + \ldots - zyxw$ (alle Permutationen der Buchstaben mit dem „richtigen“ Vorzeichen, das heißt, $4! = 24$ Terme) reichen 2×2 Matrizen nicht mehr aus [AL50]. Daher muss man mit Matrizen *aller* Größen arbeiten. Um den Zugang über „volle“ Matrizen[1] richtig zu verstehen, muss man tiefer eintauchen [Coh95, Abschnitt 4.2], [Coh06a, Kapitel 7].

Hier begnügen wir uns daher mit einer Illustration, basierend auf [Coh03b, Abschnitt 9.3]. Die notwendigen Grundlagen folgen dann in Abschnitt 2.1. Die Einfachheit der Definition *voller Matrizen* soll aber nicht darüber hinwegtäuschen, dass dahinter etwas sehr Fundamentales steckt. Bei den rationalen Zahlen wissen wir, dass der Nenner nicht Null werden darf. Und so etwas Ähnliches brauchen wir. Betrachten wir zunächst ein *lineares* Gleichungssystem $As = v$ der Dimension $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$, das heißt, wir haben n unbekannte Komponenten $s_1, s_2, \ldots, s_n$ im Lösungsvektor s (und auch v ist ein Spaltenvektor der Größe n). Wenn A invertierbar[2] ist, können wir $s = A^{-1}v$ schreiben. Sei zunächst $n = 1$ mit $A = a \in \mathbb{Z}$ und $v \in \mathbb{Z}$ (Einträge mit ganzen Zahlen). Dann ist $s = a^{-1}v = \frac{v}{a}$ für $a \neq 0$ eine Darstellung für eine rationale Zahl $s \in \mathbb{Q}$. Nun, wie kann man „umständlich“ mit diesen Darstellungen von rationalen Zahlen rechnen? Angenommen, wir haben $s_1 = \frac{v_1}{a_1}$ und $s_2 = \frac{v_2}{a_2}$ gegeben. Um

[1] Nicht zu verwechseln mit voll (bzw. dicht) besetzten Matrizen.

[2] Es dauert ein Weile, bis man sich daran gewöhnt, immer „mitzudenken“, *über* welchem Ring man Invertierbarkeit versteht.

deren Summe $s_1 + s_2 \in \mathbb{Q}$ zu berechnen, können wir sofort das Gleichungssystem

$$\begin{bmatrix} a_1 & -a_1 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} s_1 + s_2 \\ s_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

hinschreiben und lösen. Üblicherweise sind wir an der ersten Komponente im Lösungsvektor s interessiert. In diesem Fall nennen wir $As = v$ ein *zulässiges lineares System*[3] (abgekürzt ZLS). In anderen Worten: Mit einem ZLS können wir eine rationale Zahl darstellen. Man kann sich ein ZLS auch als „verallgemeinerten Bruch" vorstellen. Wenn klar ist, was gemeint ist, wird auch nur der Begriff „System" verwendet.

Wichtig: A muss „invertierbar" sein (für $n = 1$ heißt das, dass $A \neq 0$ sein muss, für $n > 1$ müssen wir dem erst noch einen Sinn geben). Nun kann man die erste Komponente einfach über den ersten Einheits(zeilen)vektor $u = e_1^\top = [1, 0, \ldots, 0]$ extrahieren. Das heißt, dass das gewünschte Element $f = s_1 = uA^{-1}v$ ist. Man nennt dann $\pi_f = (u, A, v)$ eine *lineare Darstellung* von f. Lineare Darstellungen können noch allgemeiner formuliert werden [CR99]. Und Cohn definiert *zulässige Systeme* viel allgemeiner [Coh72].

Für den *Antikommutator* $f = xy + yx$ ist ein ZLS $\mathcal{A}_f = (u, A, v)$ der Dimension $n = 4$ gegeben durch (die Nullen sind durch Punkte ersetzt, um die Struktur hervorzuheben)

$$\begin{bmatrix} 1 & -x & -y & . \\ . & 1 & . & -y \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xy + yx \\ y \\ x \\ 1 \end{bmatrix}.$$

Sei $A = (a_{ij})$. Die Lösung kann sehr einfach (von unten nach oben) berechnet werden: $s_4 = 1$ und $s_i + a_{i,i+1}s_{i+1} + \ldots + a_{i,n}s_n = 0$ für $i = 3, 2, 1$. Für diese spezielle Form ist die Invertierbarkeit sichergestellt. Ist es möglich, $f = xy + yx$ durch ein kleineres System darzustellen? Und *wie* könnte man ein solches *minimales* ZLS gegebenenfalls konstruieren? Das sind zentrale Fragen dieser Arbeit, deren endgültige Beantwortung einige Geduld abverlangt.

Später werden wir den *Rang* eines Elementes $f \in \mathbb{F}$ über die Dimension eines *minimalen* zulässigen linearen Systems (für f) definieren. Für ein Wort/Monom, z.B. $g = xyz$, lässt sich ein System einfach hinschreiben:

$$\begin{bmatrix} 1 & -x & . & . \\ . & 1 & -y & . \\ . & . & 1 & -z \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xyz \\ yz \\ z \\ 1 \end{bmatrix}.$$

Zwar ist es hier intuitiv irgendwie klarer (als beim System für f oben), dass das minimal ist, aber das muss präzisiert werden. Das erste Ziel wird sein, „einfache" *rationale Operationen* mit diesen Darstellungen (Systemen) zu definieren, zum Beispiel

[3] Was mit „linear" gemeint ist, muss natürlich noch definiert werden. Vorerst begnügen wir uns damit, dass wir auch „Buchstaben" in die *Systemmatrix* A einsetzen können. Der englische Begriff ist „admissible linear system", abgekürzt ALS.

um Elemente zu skalieren, zu addieren oder zu multiplizieren. Das ist vergleichsweise einfach. Allerdings schnell etwas umständlich, weil die Systeme größer und größer werden. Bevor wir ein Element *invertieren* (den Kehrwert bilden) müssen wir sicher sein, dass das erlaubt ist. Wenn ein System *minimal* ist, ist auch das einfach. Ein ZLS für die *Summe* von $f_1 = 2x$ und $f_2 = 3y$ wäre

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & . & . \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 2 \\ . \\ 3 \end{bmatrix}.$$

Wie schaut der Lösungsvektor s aus? Ist dieses System für $f = f_1 + f_2 = 2x + 3y$ minimal?

Obwohl es in jedem der drei Hauptkapitel einen Abschnitt „Grundlagen“ gibt, wird es davor jeweils entweder eine informelle Einführung mit Beispielen und den wichtigsten Begriffen geben oder einen „typischen“ Abschnitt, mit dem man vorab einfacher in das Kapitel eintauchen kann.

Ein kleines Beispiel mit rationalen Zahlen erklärt sofort die Strukturierung in „Rechnen“, „Faktorisieren“ und „Minimieren“ (oder „Standardisieren“ beziehungsweise „Kürzen“):

$$\begin{aligned} \frac{2}{3} \cdot \frac{3}{4} &= \frac{6}{12} = \frac{2 \cdot 3}{2 \cdot 2 \cdot 3} = \frac{1}{2}, \\ \frac{1}{2} + \frac{3}{2} &= \frac{4}{2} = \frac{2 \cdot 2}{2} = 2. \end{aligned}$$

Irgendwann wird diese Schleife abgebrochen und man kann die Zahl anwenden. Zähler und Nenner sind *koprim* (oder *relativ prim*), das heißt, der größte gemeinsame Teiler (ggT) ist 1 oder -1. Hinweis: Das sind die einzigen ganzen Zahlen, die in $\mathbb{Z}$ selbst invertiert werden können. Invertierbare Elemente nennt man *Einheiten*.

Die Grundidee ist tatsächlich so einfach. Die Umsetzung ist es nicht. Zwar kann man relativ einfach (auf Basis von zulässigen linearen Systemen) *addieren* und *multiplizieren*, praktikabel ist das nicht. Abgesehen davon lässt sich schwer feststellen, ob ein Ausdruck nicht zwischendurch zu Null wurde und man dann eben nicht *invertieren* darf. Für einen einfachen konkreten Ausdruck $f = x^{-1}zz^{-1}yz^{-1} = x^{-1}yz^{-1}$ wird man sofort einen einfacheren (und damit ein kleineres ZLS) finden, zum Beispiel

$$\begin{bmatrix} x & y \\ . & z \end{bmatrix} s = \begin{bmatrix} . \\ 1 \end{bmatrix}.$$

Aber was macht man zum Beispiel mit

$$g = x - \left(x^{-1} + (y^{-1} - x)^{-1}\right)^{-1},$$

wenn man nicht schon weiß, dass es sich dabei um die linke Seite von Huas Identität [Ami66] handelt? (In Kapitel 2 (Rechnen), konkret im Beispiel 2.6.1, werden wir diese

Identität mit Hilfe von rationalen Operationen und einfachen Minimierungsschritten beweisen.) Spätestens in Kapitel 3 (Faktorisieren) werden wir sehen, dass wir auf *minimale* Systeme angewiesen sind. Bräuchten wir also die Faktorisierung für die Minimierung (in Kapitel 4), stünden wir vor einem Henne-Ei-Problem.

Den Schlüssel dazu kann man aber auch bereits beim normalen Bruchrechnen (siehe Beispiel oben) erahnen: Man merkt sich die Faktorisierung (von Zähler und Nenner), damit man bei der Multiplikation sofort kürzen kann und bei der Addition nur mehr den Zähler faktorisieren muss (um dann zu kürzen). In unserem Fall heißt das, dass wir eine spezielle Form eines zulässigen linearen Systems benötigen, die es uns ermöglicht „einfach“ zu minimieren. Mit der Faktorisierung lässt sich ein ZLS dann in eine Form bringen, mit der man einfacher rechnen kann. (Tatsächlich ist es etwas komplizierter, weil eine Faktorisierung alleine nicht reicht um ein ZLS in eine *Standardform* zu bringen.) In (der folgenden) Abbildung 1.1 findet man (fast) alles auf einen Blick.

Womöglich schwirrt jetzt der Kopf erst recht. Doch dafür sollte es nun möglich sein, sich vorwiegend anhand der Beispiele „durchzuschlagen“. Tatsächlich wird sich die eine oder der andere später fragen, ob das nicht alles viel einfacher ginge, weil doch eh alles schon beim Hinschauen klar ist. Spätestens dann ist der richtige Zeitpunkt gekommen, genauer die Definitionen, Lemmata, Bemerkungen, Propositionen, Sätze und —in letzter Konsequenz— auch deren Beweise zu studieren. Das alles gibt es nämlich nur deswegen, weil ich mir vor über zwei Jahren dachte, dass das eh alles so einfach ist. Erst ganz zum Schluss, nämlich mit dem Schreiben (eines Entwurfs) dieses Textes, ist mir diese verblüffende Ähnlichkeit (zum Rechnen mit den Brüchen) aufgefallen.

Und ich hoffe, dass diese kleinen, unscheinbaren und zum Teil (mathematisch) unpräzisen Bemerkungen helfen, den Zugang zu einem etwas verborgeneren Bereich der Mathematik etwas zu erleichtern ...

George M. Bergman hat etwas schön auf den Punkt gebracht [Ber78]: „Die Hauptresultate in dieser Arbeit sind trivial. Aber was trivial ist, wenn es abstrakt beschrieben ist, kann im Kontext einer komplizierten Situation, wo es gebraucht wird, alles andere als klar sein. Daher scheint es sich zu lohnen, explizite Formulierungen und Beweise dieser Resultate niederzuschreiben.“

Dieses Zitat ist hier durchaus mehrdeutig zu verstehen. Denn zum einen schildert es die Situation, wie es einem gehen kann, wenn man (scheinbar) einfache Resultate (ein Korollar aus Cohns Theorie wird uns hier —als Lemma formuliert— öfter begegnen) versucht anzuwenden. Und zum anderen sind zwar einzelne Beweise vollkommen elementar und verblüffend einfach, (z.B. der des „linearisierten“ Wort-Problems in Abschnitt 4.2), der *Kontext*, in dem es dann gebraucht (beziehungsweise implementiert) wird, ist aber nicht immer so überschaubar.

1.1 Wie man sich zurechtfinden kann

Niemand sollte sich von der zunehmenden Komplexität abschrecken lassen. In konkreten praktischen Situationen findet man die notwendigen Umformungen relativ einfach durch „Hinschauen“ (auf das *zulässige lineare System*, ZLS). Die volle Systematisierung (am Computer) hat ihren Preis. Nur sollte man sich in Erinnerung rufen, dass man es hier nicht mit einem „klassischen“ Zahlenkörper, sondern mit einem (über dem Grundkörper) *unendlichdimensionalen* nicht-kommutativen „Funktionenkörper“ zu tun hat. Zum Vergleich: Die rationalen Hamiltonschen Quaternionen sind *vierdimensional* über dem Körper der rationalen Zahlen $\mathbb{Q}$.

Kapitel 2 könnte auch „Darstellen und Rechnen, insbesondere Invertieren“ heißen. Denn neben den Grundlagen wird dort eine *minimale Inverse* entwickelt und da man für Monome und einzelne Polynome bereits *minimale* Systeme kennt, lassen sich schon nicht-triviale rationale Identitäten zeigen. Natürlich vorausgesetzt, dass man alle notwendigen Umformungs- und Minimierungsschritte auch erkennt.

Kapitel 3 könnte auch „Faktorisieren und Multiplizieren“ heißen. Für zwei (durch *minimale* Systeme gegebene) Polynome lässt sich eine *minimale Multiplikation* definieren. Der „umgekehrte“ Schritt ist dann die Faktorisierung. Einer der schwierigsten Teile ist die allgemeine Faktorisierungstheorie, weil zunächst einmal geklärt werden muss, was eigentlich ein Faktor ist (schließlich ist in einem Körper jedes Element —mit Ausnahme der Null— invertierbar). Letztlich lassen sich drei Typen der *minimalen Multiplikation* formulieren, vorausgesetzt man weiß bereits, dass es sich um „Faktoren“ handelt. Als Einstieg in dieses Kapitel sind die Abschnitte 3.2 (Polynommultiplikation) und 3.3 (Polynomfaktorisierung) geeignet.

Kapitel 4 könnte auch „Minimieren und Addieren“ heißen. Nicht ganz, denn der allgemeine Fall der Multiplikation muss auch noch gelöst werden. Der Schlüssel dazu sind zwei fundamentale Techniken: Die Linearisierung des Wortproblems und die Faktorisierungstheorie (beziehungsweise die Verfeinerung der „Pivotblöcke“ in der Diagonale der Systemmatrix als eine Art „lokale“ Faktorisierung). Für ein zulässiges lineares System eines Polynoms kann man Zeilen und Spalten prüfen und gegebenenfalls (nach entsprechender Transformation) entfernen. Im allgemeinen kann es sein, dass man nur mehrere Zeilen (oder Spalten) *gemeinsam* eliminieren kann. Man muss „lokale“ Wortprobleme lösen. Aber dazu müssen diese Probleme „klein genug“ sein. Als Einstieg in dieses Kapitel ist besonders Abschnitt 4.3 zu empfehlen.

Im Anhang A gibt es eine fragmentarische und bei weitem nicht vollständige Illustration von möglichen Anwendungen des Rechnens mit (nicht-kommutativen) „freien Brüchen“. Und im Anhang B finden sich ergänzende Bemerkungen, insbesondere für eine Implementierung in Computer-Algebra-Systemen. Möchte man eine Idee von der Definition des freien Schiefkörpers bekommen, kann man von Anhang C aus in die umfangreiche Literatur von Cohn eintauchen. Und schließlich gibt es im Anhang E noch eine englische Zusammenfassung mit einem erweiterten Inhaltsverzeichnis mit Referenzen zu den Hauptarbeiten. Wenn man mit dieser Einleitung fertig ist und alles noch einmal wiederholen möchte, bevor man mit Kapitel 2 (Rechnen) weitermacht, bietet sich also alternativ Abschnitt E.1 an.

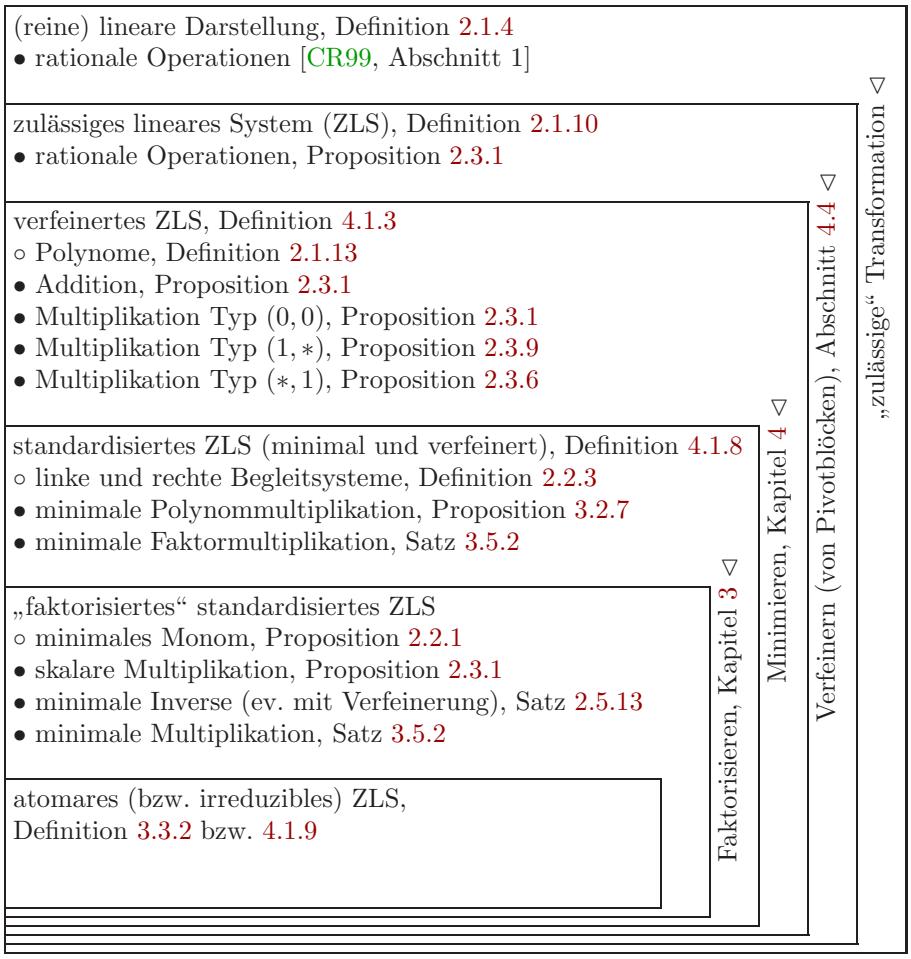


Abbildung 1.1: (Fast) alles auf einen Blick. Da jedes Element des freien Schiefkörpers $\mathbb{F}$ durch eine *reine* lineare Darstellung repräsentiert werden kann und unser Koeffizientenring $\mathbb{K}$ ein *kommutativer Körper* ist, reichen (hier) *zulässige lineare Systeme* aus. Alle Elemente, die hier in den Operationen verwendet werden sind von Null verschieden. Nachdem man sich in Abschnitt 1.2 mit zulässigen linearen Systemen vertraut gemacht hat, kann man nach Abschnitt 1.3 auch selbst „einfache“ minimale Systeme konstruieren.

In der Abbildung ist der Weg skizziert, den wir gehen wollen. Das große Ziel ist eine *Standardform*, die es uns ermöglicht mit Elementen zu rechnen und einfach (mit linearen Techniken) zu minimieren.

1.2 Freie Brüche: Eine Darstellung, zwei Gleichungssysteme

Oder: Ein Schnellstart für jene, die es nicht mehr erwarten können. Im Abschnitt 2.1 gibt es jede Menge an Begriffen (und Definitionen). Für einen Überblick werden die, die uns ständig begegnen werden, anhand eines Beispiels eingeführt. Wir betrachten ein Element f im freien Schiefkörper $\mathbb{F}$, das durch die lineare Darstellung $\mathcal{A} = (u, A, v)$ der Dimension $n = 4$ gegeben ist. Bevor wir zu der etwas geläufigeren Sichtweise der „Zeilengleichungen“ kommen, schauen wir uns die der „Spaltengleichungen“ $u = tA$ an, in der f als Linearkombination der Komponenten des Zeilenlösungsvektors $t = [t_1, t_2, \ldots, t_n]$ ausgedrückt werden kann. In dieser Kurzeinführung bedeuten unterstrichene Einträge, dass diese *statisch* sind, das heißt, nicht verändert werden dürfen. Wenn im folgenden dann (elementare) Umformungen beschrieben werden, beziehen sich diese *immer* auf die Systemmatrix A.

$$\underbrace{\begin{bmatrix}\underline{1} & \underline{0} & \underline{0} & \underline{0}\end{bmatrix}}_{u,\text{ linke Seite}} = \underbrace{\begin{bmatrix}t_1 & t_2 & t_3 & t_4\end{bmatrix}}_{t,\text{ rechte Familie}} \left.\begin{bmatrix}1-x & . & -x & -x\\ 1 & y & 1 & -2\\ . & 1 & . & -x\\ . & . & . & 1\end{bmatrix}\right\} \begin{matrix}\text{Dimension}\\ \dim(\mathcal{A}) = n\end{matrix}$$

Die Gleichungen (von links beginnend) lauten also

$$\begin{aligned}1 &= t_1(1-x) + t_2,\\ 0 &= t_2 y + t_3,\\ 0 &= -t_1 x + t_2 \quad \text{und}\\ 0 &= -t_1 x - 2t_2 - t_3 x + t_4.\end{aligned}$$

Anstatt die Lösung t sofort zu berechnen, werden wir das System so umformen, dass dies „einfacher“ wird. Zunächst aber schauen wir uns das Gleichungssystem $As = v$ genauer an:

$$\underbrace{\begin{bmatrix}1-x & . & -x & -x\\ 1 & y & 1 & -2\\ . & 1 & . & -x\\ . & . & . & 1\end{bmatrix}}_{A,\text{ Systemmatrix}} \underbrace{\begin{bmatrix}\underline{s_1}\\ s_2\\ s_3\\ s_4\end{bmatrix}}_{s,\text{ linke Familie}} = \underbrace{\begin{bmatrix}.\\ -4\\ .\\ 2\end{bmatrix}}_{v,\text{ rechte Seite}}$$

Eine Gleichung, nämlich $s_4 = 2$, lässt sich besonders einfach lösen. Hier haben wir $\kappa_1 s_1 + \kappa_2 s_2 + \kappa_3 s_3 + \kappa_4 s_4 = 1$ für $\kappa_1 = \kappa_2 = \kappa_3 = 0$ und $\kappa_4 = \frac{1}{2}$, deshalb schreiben wir

$1 \in L(\mathcal{A})$, dem linearen Span (über $\mathbb{K}$) der *linken Familie*. (Gäbe es keine solche Linearkombination, würden wir $1 \notin L(\mathcal{A})$ schreiben.) Analoges gilt für den linearen Span der *rechten Familie* $R(\mathcal{A})$. Normalerweise müssen wir zwischen dem Element f und der Darstellung $\mathcal{A}$ unterscheiden. Wenn $\mathcal{A}$ *minimal* ist (was hier der Fall ist), können wir den *Rang* von f als die *Dimension* von $\mathcal{A}$ definieren, $\operatorname{rang}(f) := \dim(\mathcal{A})$. In diesem Fall sagen wir „f ist vom Typ $(*, 1)$“ oder $1 \in L(f)$ für $1 \in L(\mathcal{A})$ beziehungsweise „f ist vom Typ $(1, *)$“ oder $1 \in R(f)$ für $1 \in R(\mathcal{A})$.

Nun werden wir diese Darstellung Schritt für Schritt so umformen, dass die Lösung beider Gleichungssysteme, das heißt die Bestimmung von s und t, einfacher wird. Denn diese beiden Familien spielen eine entscheidende Rolle bei der Charakterisierung der Minimalität einer Darstellung, siehe Proposition 2.1.8. Tatsächlich wird das Ziel aber sein, diese Lösungen gar nicht erst ausrechnen zu müssen, weil uns das im allgemeinen Fall gar nicht weiterhelfen wird. Üblicherweise schreiben wir s und t (ohne ihrer Komponenten) in „generischer“ Form. Der Blick „hinein“ (in die Darstellung) dient nur der Erklärung. Nach den folgenden Umformungen sollte man nicht auf diesen Blick vergessen und die „neuen“ Lösungen s und t ausrechnen, weil er einem auch eine Eselsbrücke für die Bezeichnungen *linke* beziehungsweise *rechte Familie* eröffnet.

Zunächst addieren wir 2-mal Zeile 4 zu Zeile 2 (für den Lösungsvektor t bedeutet das, dass wir 2-mal t_2 von t_4 subtrahieren müssen). Danach vertauschen wir die Spalten 2 und 3 (für s heißt das, s_2 und s_3 vertauschen) und subtrahieren (die neue) Spalte 2 von Spalte 1. Diese elementaren Umformungen fassen wir in der *zulässigen* Transformation (P, Q), das heißt, die erste Komponente im Lösungsvektor s ändert sich nicht, mit

$$P = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & 2 \\ . & . & 1 & . \\ . & . & . & 1 \end{bmatrix} \quad \text{und} \quad Q = \begin{bmatrix} \underline{1} & \underline{0} & \underline{0} & \underline{0} \\ . & . & 1 & . \\ -1 & 1 & . & . \\ . & . & . & 1 \end{bmatrix}$$

zusammen. Damit erhalten wir eine neue Darstellung $\mathcal{A}' = (u', A', v') = P\mathcal{A}Q$,

$$\mathcal{A}' = (uQ, PAQ, Pv) = \left(\begin{bmatrix} \underline{1} & \underline{0} & \underline{0} & \underline{0} \end{bmatrix}, \begin{bmatrix} 1 & -x & . & -x \\ . & 1 & y & . \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 2 \end{bmatrix} \right).$$

Die erste Komponente des Lösungsvektors s ist (immer noch) $f = 2x - 2xyx$. Wer damit noch nicht zufrieden ist, kann entweder Zeile 3 von Zeile 1 oder Spalte 2 von Spalte 4 subtrahieren und sich unser Element alternativ als $x(2-2yx)$ oder $(1-xy)2x$ denken.[4] Das wird uns in Kapitel 3 beschäftigen. Für Polynome findet man immer so eine Form mit n (skalaren) „Pivotblöcken“ der Größe 1×1. Im allgemeinen ist das nicht möglich, wir werden aber versuchen, möglichst „kleine“ Pivotblöcke zu bekommen.

[4] Eine noch nicht erwähnte Annahme ist, dass Skalare und Wörter/Monome miteinander kommutieren. Skalare werden üblicherweise links geschrieben.

Darum wird es in den Kapiteln 3 (Faktorisieren) und 4 (Minimieren) gehen. Das soll uns aber nicht davon abhalten, bereits früher damit zu rechnen. Die Beispiele zu Beginn sind so gestaltet, dass man sie einfach „händisch“ minimieren beziehungsweise deren Minimalität prüfen kann.

Fehlt noch eine Anmerkung zur Systemmatrix A. Wir schreiben sie immer in der kompakten Form mit (maximal) *linearen* Einträgen (nicht-kommutativer Polynome). Tatsächlich kann A auch als *lineares Matrixbüschel*[5] $A = (A_0, A_1, \ldots, A_d)$ mit Koeffizientenmatrizen $A_i \in \mathbb{K}^{n\times n}$ für ein Alphabet $X = \{x_1, \ldots, x_d\}$ interpretiert werden, geschrieben auch als $A = A_0 + A_1x_1 + \ldots + A_dx_d$. Wenn man die Elemente als *Funktionen* betrachtet, z.B. $f(x, y) = 2x - 2xyx$, und Matrizen (gleicher Größe) für die nicht-kommutativen Variablen einsetzt, muss man natürlich das Tensorprodukt verwenden, z.B. $A_1 \otimes x_1$.

1.3 Linke und rechte Minimierungsschritte

Fürs praktische Rechnen müssen wir natürlich auch immer wieder ein ZLS „verkleinern“. In konkreten Fällen schafft man auch die Minimierung. Dass da einiges im Verborgenen liegt, kann man an Kapitel 4 erahnen. Aber dazu später. Gehen wir zurück zum Beispiel $2x + 3y$ von vorhin:

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & . & . \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 2 \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x+3y \\ 2 \\ 3y \\ 3 \end{bmatrix}.$$

Zuerst versuchen wir einen „linken“ Minimierungsschritt, das heißt, eine Komponente in der linken Familie zu eliminieren. Dazu subtrahieren wir $\frac{2}{3}$-mal Zeile 4 von Zeile 2 und addieren $\frac{2}{3}$-mal Spalte 2 zu Spalte 4:

$$\begin{bmatrix} 1 & -x & -1 & -\frac{2}{3}x \\ . & 1 & 0 & 0 \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 0 \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x+3y \\ 0 \\ 3y \\ 3 \end{bmatrix}.$$

In der zweiten Zeile steht nun $s_2 = 0$. Das heißt, dass für die Lösung s_1 *kein Beitrag* von s_2 kommt. Und daher können wir sowohl die Gleichung $s_2 = 0$ als auch die Variable s_2 aus unserem Gleichungssystem löschen. Damit erhalten wir folgendes (noch

[5] Englisch „linear matrix pencil“. In der Übersetzung von Gantmachers Klassiker [Gan66] wird der Begriff „lineares Matrizenbüschel vom Typ (m, n)“ (für m Zeilen und n Spalten) verwendet. Da in unserem Fall *alle* Matrizen die gleiche Größe haben, kann man dann den Begriff „Matrizenbüschel“ für allgemeinere Büschel verwenden. Lineare Matrixbüschel spielen eine wichtige Rolle bei *nicht-linearen Eigenwertproblemen*. Aus Sicht der Matrixbüschel wäre [Ikr91] ein möglicher Einstieg. Algorithmen und weitere Literatur findet man in [MV04].

nicht minimales) ZLS für $2x + 3y$:

$$\begin{bmatrix} 1 & -1 & -\frac{2}{3}x \\ . & 1 & -y \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x+3y \\ 3y \\ 3 \end{bmatrix}.$$

Man sieht sofort, dass man nun einen „rechten“ Minimierungsschritt durchführen kann, um t_2 (in der rechten Familie) zu eliminieren. Tatsächlich braucht man weder die linke noch die rechte Familie ausrechnen, um zu „minimieren“ (ohne darüber nachzudenken, ob das Resultat auch minimal ist).

Eine Minimierung ist manchmal nur in „Blöcken“ möglich. Wir betrachten nun das ZLS $\mathcal{A} = (u, A, v)$[6] für $ff^{-1} = 1$ mit $f = xy - z$,

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & . & . \\ . & 1 & -y & . & . \\ . & . & 1 & -1 & . \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Hier können wir in A einen 3×2 Nullblock rechts oben erzeugen, indem wir (als ersten Schritt) Spalte 3 zu Spalte 4 und Zeile 4 zu Zeile 2 addieren:

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & z & . \\ . & 1 & -y & . & -1 \\ . & . & 1 & 0 & 0 \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Und (als zweiten Schritt) Spalte 2 zu Spalte 5 und Zeile 5 zu Zeile 1 addieren:

$$\mathcal{A}'' = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & 0 & 0 \\ . & 1 & -y & 0 & 0 \\ . & . & 1 & 0 & 0 \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Nun können wir den unteren 2×2 Diagonalblock invertieren (weil die Systemmatrix „invertierbar“ ist) und erhalten $t''_4 = t''_5 = 0$ (weil die entsprechenden Einträge in u Null sind). Damit bekommen wir das (nicht minimale) ZLS der Dimension 3,

$$\mathcal{A}''' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z \\ . & 1 & -x \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right).$$

[6] Dieses zulässige lineare System kann man wie folgt konstruieren: Zunächst startet man mit einem *minimalen* ZLS der Dimension 3 für das Monom xy (Proposition 2.2.1). Nachdem xy vom Typ $(1, 1)$ ist, kann man dort sofort z rechts oben in der Systemmatrix eintragen, um ein System für $xy - z$ zu bekommen. Ein System für die Inverse erhält man über Satz 2.5.13. Und mit der Multiplikation in Proposition 2.3.1 bekommt man dann ein System der Dimension 5 für ff^{-1}.

Man beachte, dass die unteren Komponenten in der rechten Seite Null sind. Wir können also sofort einen linken Block-Eliminierungs-Schritt durchführen und bekommen das minimale System $\mathcal{A}'''' = (1, [1], 1)$ für $1 \in \mathbb{F}$.

Bemerkung. Der andere Fall $f^{-1}f = 1$ ist etwas schwieriger, weil man die erste Komponente, die mit anderen „verknüpft“ ist, in der linken Familie nicht verändern darf. Der Trick ist hier, stattdessen $1 \cdot f^{-1}f$ über ein „erweitertes“ ZLS zu betrachten. Mit der Multiplikation in Abschnitt 2.3 ist dann auch das kein großes Problem mehr. Wie man mit einem erweiterten ZLS arbeitet, ist in Beispiel 4.5.6 illustriert.

Bemerkung. In manchen Fällen kann man auch einen linken und rechen Minimierungsschritt *gemeinsam* machen. In Abschnitt A.1 wird das verwendet, um den linken ggT zweier Polynome p und q über die Minimierung eines ZLS für $p^{-1}q$ zu bestimmen.

Kapitel 2

Rechnen

Eine der zentralen Teile dieses Kapitels ist die Konstruktion von *minimalen* zulässigen linearen Systemen für die Inverse in Abschnitt 2.5. Die folgende (einfache) Konstruktion findet sich (davor) in Proposition 2.3.1 (Rationale Operationen). Angenommen wir haben die Inverse eines Monoms $f = xyz$ gegeben durch das ZLS $\mathcal{A}' = (u', A', v')$,

$$\begin{bmatrix} z & -1 & . \\ . & y & -1 \\ . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} z^{-1}y^{-1}x^{-1} \\ y^{-1}x^{-1} \\ x^{-1} \end{bmatrix}.$$

Die Minimalität (laut Proposition 2.1.8) ist sofort klar, wenn man auch die $\mathbb{K}$-lineare Unabhängigkeit der rechten Familie prüft. Ein (minimales) ZLS für f ist zum Beispiel gegeben durch

$$\begin{bmatrix} . & z & -1 & . \\ . & . & y & -1 \\ -1 & . & . & x \\ . & 1 & . & . \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xyz \\ 1 \\ z \\ yz \end{bmatrix},$$

die Systemmatrix wurde links durch $-v$ und unten durch u „erweitert“. Will man die uns bereits geläufige Form (aus Proposition 2.2.1) haben, muss man die Reihenfolge der Zeilen $1, 2, 3$ und Spalten $2, 3, 4$ vertauschen und die Zeilen $1, 2, 3$ mit -1 multiplizieren. Als neues System $\mathcal{A} = (u, A, v)$ für f erhalten wir

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & . & . \\ . & 1 & -y & . \\ . & . & 1 & -z \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Hier sieht man auch sofort, dass $1 \in L(f)$ und $1 \in R(f)$ ist, das heißt, f ist *vom Typ* $(1, 1)$. Eine nochmalige Anwendung dieser „Erweiterung“ würde zwar wieder ein ZLS für f^{-1} ergeben, aber das hätte dann bereits Dimension 5. Daher wird eine

wichtige (technische) Aufgabe die Prüfung sein, *wie* man „besondere“ Formen (der Systemmatrizen) erkennt.

Um später minimieren zu können, wollen wir eine möglichst „einfache“ Struktur, das heißt, die (diagonalen) Pivotblöcke sollten so klein wie möglich sein. Für das Beispiel hier, ein ZLS für ein Monom, trifft das auch für die Standardinverse (Proposition 2.5.1) zu.

Notation. Die Menge der natürlichen Zahlen wird mit $\mathbb{N} = \{1, 2, \ldots\}$, die inklusive der Null mit $\mathbb{N}_0$ bezeichnet. Nulleinträge in Matrizen werden üblicherweise durch Punkte ersetzt, um die Struktur der anderen Einträge hervorzuheben, außer eine Null entsteht nach einer Operation dort, wo vorher ein Nicht-Null-Eintrag war. Die Identitätsmatrix wird mit I (wenn die Größe aus dem Zusammenhang klar ist) oder I_n (der Größe n) bezeichnet. Mit Σ beziehungsweise Σ_n wird die Permutationsmatrix bezeichnet, die die Reihenfolge der Zeilen/Spalten umkehrt.

2.1 Grundlagen

Sei $\mathbb{K}$ ein *kommutativer* Körper, $\overline{\mathbb{K}}$ dessen algebraischer Abschluss und $X = \{x_1, x_2, \ldots, x_d\}$ ein *endliches* (nicht-leeres) Alphabet. $\mathbb{K}\langle X\rangle$ bezeichnet die *freie assoziative Algebra* (oder *freie* $\mathbb{K}$*-Algebra*) und $\mathbb{F} = \mathbb{K}(\langle X\rangle)$ ihren *universeller Quotientenkörper* (oder „freien Schiefkörper“) [Coh95, CR99]. Ein Element in $\mathbb{K}\langle X\rangle$ heißt (nicht-kommutatives) *Polynom*. In unseren Beispielen ist das Alphabet üblicherweise $X = \{x, y, z\}$. Die Algebra der (nicht-kommutativen) *rationalen formalen Potenzreihen* einschließend, haben wir folgende Kette von echten Inklusionen:

$$\mathbb{K} \subsetneq \mathbb{K}\langle X\rangle \subsetneq \mathbb{K}^{\mathrm{rat}}\langle\langle X\rangle\rangle \subsetneq \mathbb{K}(\langle X\rangle) =: \mathbb{F}.$$

Das von X erzeugte *freie Monoid* X^* ist die Menge aller *endlichen Wörter* $x_{i_1} \cdots x_{i_n}$ mit $i_k \in \{1, 2, \ldots, d\}$. Ein Element des Alphabets heißt *Buchstabe*, eines des freien Monoids *Wort*. Die Multiplikation auf X^* ist das *Zusammenfügen* von Wörtern, das heißt, $(x_{i_1} \cdots x_{i_m}) \cdot (x_{j_1} \cdots x_{j_n}) = x_{i_1} \cdots x_{i_m} x_{j_1} \cdots x_{j_n}$, mit neutralem Element 1, dem *leeren Wort*. Die *Länge* eines Wortes $w = x_{i_1} x_{i_2} \cdots x_{i_m}$ ist m, mit $|w| = m$ oder $\ell(w) = m$ bezeichnet. Ausführliche Einleitungen finden sich unter anderem in [BR11, Kapitel 1] oder [SS78, Abschnitt I.1].

Definition 2.1.1 (Innerer Rang, volle und hohle Matrizen [Coh85], [CR99])**.** Gegeben eine Matrix $A \in \mathbb{K}\langle X\rangle^{n\times n}$, ist der *innere Rang* von A die kleinste Zahl $m \in \mathbb{N}$ sodass eine Faktorisierung $A = TU$ mit $T \in \mathbb{K}\langle X\rangle^{n\times m}$ und $U \in \mathbb{K}\langle X\rangle^{m\times n}$ existiert. Die Matrix A heißt *voll* wenn $m = n$ gilt, ansonsten *nicht-voll*. Sie heißt *hohl* wenn sie eine Null-Untermatrix der Größe $k \times l$ mit $k + l > n$ enthält.

Definition 2.1.2 (Assoziierte und stabil assoziierte Matrizen [Coh95])**.** Zwei Matrizen A und B über $\mathbb{K}\langle X\rangle$ (der gleichen Größe) heißen *assoziiert* über einem Unterring $R \subseteq \mathbb{K}\langle X\rangle$ wenn es invertierbare Matrizen P, Q über R gibt, sodass $A = PBQ$ ist. A und B (nicht notwendigerweise der gleichen Größe) heißen *stabil assoziiert* wenn

$A \oplus I_p$ und $B \oplus I_q$ assoziiert sind für Einheitsmatrizen I_p und I_q. Hier bezeichnet $C \oplus D$ die diagonale Summe $\left[\begin{smallmatrix} C & \\ & D \end{smallmatrix}\right]$.

Lemma 2.1.3 ([Coh95, Korollar 6.3.6]). *Eine nicht-volle lineare quadratische Matrix über $\mathbb{K}\langle X\rangle$ ist über $\mathbb{K}$ zu einer linearen hohlen Matrix assoziiert.*

Bemerkung. Eine *volle* (engl. *full*) Matrix kann *dünn besetzt* (engl. *sparse*) sein. Eine hohle quadratische Matrix kann nicht voll sein [Coh85, Abschnitt 3.2], illustriert anhand eines Beispiels:

$$A = \begin{bmatrix} z & . & . \\ x & . & . \\ y & -x & 1 \end{bmatrix} = \begin{bmatrix} z & 0 \\ x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ y & -x & 1 \end{bmatrix}.$$

Definition 2.1.4 (Lineare Darstellungen, Dimension, Rang [CR94], [CR99]). Sei $f \in \mathbb{F}$. Eine *lineare Darstellung* von f ist ein Tripel $\pi_f = (u, A, v)$ mit $u \in \mathbb{K}^{1\times n}$, voller Matrix $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$, das heißt, A ist invertierbar über $\mathbb{F}$, $A_\ell \in \mathbb{K}^{n\times n}$, $v \in \mathbb{K}^{n\times 1}$ und $f = uA^{-1}v$. Die *Dimension* von π_f ist $\dim(u, A, v) = n$. Eine Darstellung (u, A, v) heißt *minimal* wenn A die kleinst mögliche Dimension unter allen linearen Darstellungen von f besitzt. Die „leere“ Darstellung $\pi = (,,)$ ist die minimale für $0 \in \mathbb{F}$ mit $\dim(\pi) = 0$. Sei $f \in \mathbb{F}$ und π eine *minimale* lineare Darstellung von f. Dann ist der *Rang* von f definiert als $\operatorname{rang}(f) = \dim(\pi)$.

Notation. Wenn in einer linearen Darstellung $\pi = (u, A, v)$ der Zeilenvektor die Form $u = [\kappa, 0, \ldots, 0]$ und der Spaltenvektor die Form $v = [0, \ldots, 0, \lambda]^\top$ (für $\kappa, \lambda \in \mathbb{K}$) hat, schreiben wir auch $\pi = (\kappa, A, \lambda)$.

Bemerkung. Der Zusammenhang der Konzepte *Rang*, *Inversionshöhe* und *Tiefe* wird in [Reu96a] diskutiert, nämlich *Inversionhöhe* $\leq$ *Tiefe* $\leq$ *Rang*. Mehr zur „Tiefe“ findet sich zum Beispiel in [Coh06a, Abschnitt 7.7].

Bemerkung. Cohn und Reutenauer definieren lineare Darstellungen etwas allgemeiner, nämlich $f = c + uA^{-1}v$ mit möglicherweise von Null verschiedenem $c \in \mathbb{K}$ und nennen sie *rein* wenn $c = 0$ gilt.

Definition 2.1.5 ([CR99]). Zwei lineare Darstellungen heißen *äquivalent*, wenn sie das selbe Element repräsentieren.

Satz 2.1.6 ([CR99, Satz 1.4]). *Wenn $\pi' = (u', A', v')$ und $\pi'' = (u'', A'', v'')$ äquivalente (reine) lineare Darstellungen sind, wobei die erste minimal sei, dann ist die zweite isomorph zu einer Darstellung $\pi = (u, A, v)$ mit einer Block-Zerlegung*

$$u = \begin{bmatrix} . & u' & * \end{bmatrix}, \quad A = \begin{bmatrix} * & * & * \\ . & A' & * \\ . & . & * \end{bmatrix} \quad \text{und} \quad v = \begin{bmatrix} * \\ v' \\ . \end{bmatrix}.$$

Bemerkung. Nun, was hindert uns daran, nach Transformationen zu suchen, um sofort diese Form zu bekommen? Theoretisch nichts, für ein Alphabet mit d Buchstaben, $\dim(\pi'') = n$ und einen entsprechenden Nullblock mit k Zeilen links unten (a

priori weiß man nicht, wie groß dieser Block ist) bräuchten wir „nur“ $(d+1)k(n-k)+k$ *nichtlineare* Gleichungen mit maximal quadratischen Termen und zwei vom Grad n (damit die Invertierbarkeit der Transformationsmatrizen sichergestellt ist) in $2n^2$ *kommutativen* Unbekannten. Praktisch ist das bereits für $n = 5$ schwierig. Das Ziel ist daher, so weit wie möglich *lineare* Techniken zur Minimierung zu verwenden. Und dabei kann es helfen, eine (oder mehrere) Faktorisierung(en) eines Elementes in seine „verallgemeinerten“ Atome zu kennen. Ist die linke untere Blockstruktur hinreichend „fein“, kann man mit linearen Techniken minimieren. Direkte Verknüpfungen gibt es unter anderem zu den Abschnitten 3.3 (Polynomfaktorisierung), 3.6 (Faktorisierung), 4.2 (Wortproblem) und insbesondere 4.5 (Minimierung).

Definition 2.1.7 (Linke und rechte Familien [CR94])**.** Sei $\pi = (u, A, v)$ eine lineare Darstellung von $f \in \mathbb{F}$ mit der Dimension n. Die Familien $(s_1, s_2, \ldots, s_n) \subseteq \mathbb{F}$ mit $s_i = (A^{-1}v)_i$ und $(t_1, t_2, \ldots, t_n) \subseteq \mathbb{F}$ mit $t_j = (uA^{-1})_j$ heißen *linke Familie* beziehungsweise *rechte Familie*. $L(\pi) = \operatorname{span}\{s_1, s_2, \ldots, s_n\}$ und $R(\pi) = \operatorname{span}\{t_1, t_2, \ldots, t_n\}$ bezeichnen ihren jeweiligen linearen Span (über $\mathbb{K}$).

Proposition 2.1.8 (Minimalitätscharakterisierung [CR94, Proposition 4.7])**.** *Eine lineare Darstellung $\pi = (u, A, v)$ eines Elementes $f \in \mathbb{F}$ ist genau dann minimal, wenn beide, die linke Familie und die rechte Familie jeweils $\mathbb{K}$-linear unabhängig sind. In diesem Fall hängen $L(\pi)$ und $R(\pi)$ nur von f ab.*

Notation. ([Sch17a, Abschnitt 1]) Für zwei beliebige *minimale* lineare Darstellungen π_1 und π_2 eines Elementes $f \in \mathbb{F}$ gilt $1 \in L(\pi_1)$ genau dann, wenn $1 \in L(\pi_2)$ weil sie sonst nicht mit invertierbaren Matrizen über $\mathbb{K}$ ineinander übergeführt werden könnten. Mit $1 \in L(f)$ (bzw. $1 \in R(f)$) bezeichnen wir $1 \in L(\pi)$ (bzw. $1 \in R(\pi)$) für jede *minimale* Darstellung π von f.

Definition 2.1.9 (Typen)**.** Ein Element $f \in \mathbb{F}$ heißt *vom Typ* $(1, *)$ (bzw. $(0, *)$), wenn $1 \in R(f)$ (bzw. $1 \notin R(f)$) ist. Es heißt *vom Typ* $(*, 1)$ (bzw. $(*, 0)$), wenn $1 \in L(f)$ (bzw. $1 \notin L(f)$) ist. Beide Teiltypen können miteinander kombiniert werden.

Definition 2.1.10 (Zulässige lineare Systeme, zulässige Transformationen [Sch17b])**.** Eine lineare Darstellung $\pi = (u, A, v)$ von $f \in \mathbb{F}$ heißt *zulässiges lineares System* (ZLS) für f, auch geschrieben als $As = v$, wenn $u = e_1 = [1, 0, \ldots, 0]$ ist. Das Element f ist dann die erste Komponente des (eindeutig bestimmten) Lösungsvektors s. Gegeben eine lineare Darstellung $\pi = (u, A, v)$ der Dimension n von $f \in \mathbb{F}$ und invertierbare Matrizen $P, Q \in \mathbb{K}^{n\times n}$, ist die transformierte $P\pi Q = (uQ, PAQ, Pv)$ wieder eine lineare Darstellung (von f). Für ein ZLS π heißt die Transformation (P, Q) *zulässig* wenn die erste Zeile von Q der Einheitszeilenvektor $e_1 = [1, 0, \ldots, 0]$ ist.

Bemerkung 2.1.11. Cohn definiert *zulässige Systeme* viel allgemeiner [Coh85, Abschnitt 7.1] mit nicht notwendigerweise skalaren Einträgen in der rechten Seite von $As = v$ (der Dimension n) und schreibt dieses System als Block $B = [-v, A]$ der Größe $n \times (n + 1)$. Die ersten n Spalten von B dienen dann als Zähler, die letzten n Spalten als Nenner. Damit kann er ein Analogon zur Cramerschen Regel formulieren

[Coh82a]. Allerdings kann in diesem Fall —selbst für eine *lineare* Systemmatrix A— für reguläre Elemente die Dimension eines solchen (minimalen) Systems vom Hankel Rang [Fli74], [SS78, Abschnitt II.3] abweichen.

Definition 2.1.12. Sei $M = M_1 \otimes x_1 + \ldots + M_d \otimes x_d$ mit $M_i \in \mathbb{K}^{n\times n}$ für ein $n \in \mathbb{N}$. Ein Element in $\mathbb{F}$ heißt *regulär*, wenn es eine lineare Darstellung (u, A, v) mit $A = I - M$ gibt, das heißt, $A_0 = I$ in Definition 2.1.4, oder äquivalent, wenn A_0 regulär (invertierbar) ist.

Bemerkung. Die linke Familie $(A^{-1}v)_i$ (bzw. die rechte Familie $(uA^{-1})_j$) und der Lösungsvektor s von $As = v$ (bzw. t von $u = tA$) werden synonym verwendet.

Bemerkung. Ein zulässiges lineares System $\pi = (u, A, v)$ wird üblicherweise mit $\mathcal{A}$ bezeichnet. *Minimale* zulässige lineare Systeme für ein und das selbe Element können mit zulässigen Transformationen ineinander übergeführt werden (siehe dazu den Beweis des „linearen“ Wortproblems, Satz 4.2.3). Im allgemeinen trifft das nicht zu. Daran würde selbst eine etwas verallgemeinerte Form der Transformationsmatrizen (in Abhängigkeit eines gegebenen Elementes) nichts ändern:

$$\mathcal{A}_1 = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} 1 & -z \\ . & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ . \end{bmatrix} \right), \quad \mathcal{A}_2 = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} 1 & -y \\ . & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ . \end{bmatrix} \right).$$

Es gibt keine (invertierbaren) Matrizen P und Q, sodass $\mathcal{A}_1 = P\mathcal{A}_2Q$ ist.

Bemerkung. Die folgende Definition eines *polynomiellen* ZLS ersetzt [Sch17c, Definition 2.1], weil später ein verfeinertes ZLS allgemeiner (nicht nur für Polynome) definiert wird. Die einer *polynomiell zulässigen Transformation* ist etwas allgemeiner formuliert, weil für die Faktorisierung ohnehin eine spezielle Transformation verwendet wird.

Definition 2.1.13 (Polynomielles ZLS, polynomielle Transformation)**.** Ein ZLS $\mathcal{A} = (u, A, v)$ der Dimension n mit einer Systemmatrix $A = (a_{ij})$ für ein Polynom $0 \neq p \in \mathbb{K}\langle X\rangle$ heißt *polynomiell*, wenn

(1) $v = [0, \ldots, 0, \lambda]^\top$ für ein $\lambda \in \mathbb{K}$ und

(2) $a_{ii} = 1$ für $i = 1, 2, \ldots, n$ und $a_{ij} = 0$ für $i > j$ gilt, das heißt, A ist eine obere Dreiecksmatrix.

Ein polynomielles ZLS wird auch geschrieben als $\mathcal{A} = (1, A, \lambda)$ mit $1, \lambda \in \mathbb{K}$. Eine zulässige Transformation (P, Q) heißt *polynomiell*, wenn sie von der Form

$$(P, Q) = \left(\begin{bmatrix} 1 & \alpha_{1,2} & \ldots & \alpha_{1,n-1} & \alpha_{1,n} \\ & \ddots & \ddots & \vdots & \vdots \\ & & 1 & \alpha_{n-2,n-1} & \alpha_{n-2,n} \\ & & & 1 & \alpha_{n-1,n} \\ & & & & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ & 1 & \beta_{2,3} & \ldots & \beta_{2,n} \\ & & 1 & \ddots & \vdots \\ & & & \ddots & \beta_{n-1,n} \\ & & & & 1 \end{bmatrix} \right)$$

(2.1.14)

ist. Gilt zusätzlich $\alpha_{1,n} = \alpha_{2,n} = \ldots = \alpha_{n-1,n} = 0$ heißt (P, Q) *Polynomfaktorisierungstransformation*, siehe dazu auch Abschnitt 3.3.

2.2 Minimale Systeme

Für Monome kann man sofort ein *minimales* ZLS angeben. Darum geht es in der gleich folgenden Proposition. Mit diesen minimalen zulässigen linearen Systemen kann man bereits (im nächsten Abschnitt) rechnen. Für bestimmte Polynome ist das über eine verallgemeinerte „Begleitmatrix“ möglich, die gleich im Anschluss entwickelt wird. Die Minimalität wird in Proposition 2.2.6 gezeigt.

Proposition 2.2.1 (Minimales Monom [Sch17b, Proposition 4.1])**.** *Sei $k \in \mathbb{N}$ und $f = x_{i_1} x_{i_2} \cdots x_{i_k}$ ein Monom in $\mathbb{K}\langle X\rangle \subseteq \mathbb{F}$. Dann ist*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & \cdots & . \end{bmatrix}, \begin{bmatrix} 1 & -x_{i_1} & & & \\ & 1 & -x_{i_2} & & \\ & & \ddots & \ddots & \\ & & & 1 & -x_{i_k} \\ & & & & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \vdots \\ . \\ 1 \end{bmatrix} \right)$$

ein minimales *polynomielles ZLS der Dimension* $\dim \mathcal{A} = k + 1$.

Beweis. Für Zeilen- und Spaltenindizes $[1, x_{i_1}, x_{i_1} x_{i_2}, \ldots, x_{i_1} \cdots x_{i_k}]$ beziehungsweise $[1, x_{i_k}, x_{i_{k-1}} x_{i_k}, \ldots, x_{i_1} \cdots x_{i_k}]$ ist die Hankel-Matrix [Fli74], [SS78, Abschnitt II.3] von f

$$H(f) = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$$

mit Rang $k+1$. Die Systemmatrix von $\mathcal{A}$ ist voll und klarerweise ist $\mathcal{A}$ polynomiell. □

Bemerkung. Trivialerweise ist $\mathcal{A} = (1, [1], 1)$ ein minimales ZLS für das (multiplikative) *Einheitselement*, das *leere Wort*.

Für einen speziellen Fall, nämlich einem Alphabet mit nur einem Buchstaben, liefert die Begleitmatrix eines Polynoms $p(x)$ sofort eine *minimale* lineare Darstellung von $p \in \mathbb{K}\langle\{x\}\rangle$.

Bemerkung. Wenn p das charakteristische Polynom einer (quadratischen) Matrix $B \in \mathbb{K}^{m \times m}$ ist, dann können die Eigenwerte mit den Techniken von Abschnitt 3.3 (wenn notwendig zu $\overline{\mathbb{K}}$ übergehend) berechnet werden, illustriert in Beispiel 3.3.1.

Für eine allgemeinere Klasse von (nicht-kommutativen) Polynomen können *linke* und *rechte* Begleitsysteme definiert werden. Im allgemeinen sind *minimale* polynomielle zulässige lineare Systeme notwendig, um Begleitmatrizen zu verallgemeinern, vergleiche Definition 2.2.7.

Definition 2.2.2 (Begleitmatrix, Charakteristisches Polynom, Normalform [Gan65, Abschnitt 6.6])**.** Sei $p(x) = a_0 + a_1 x + \ldots + a_{m-1}x^{m-1} + x^m \in \mathbb{K}[x]$. Die *Begleitmatrix* $L(p)$ ist definiert als

$$L(p) = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ddots & \vdots & -a_1 \\ & \ddots & \ddots & 0 & \vdots \\ & & 1 & 0 & -a_{m-2} \\ & & & 1 & -a_{m-1} \end{bmatrix}.$$

Dann ist $p(x)$ das *charakteristische Polynom* von $L = L(p)$:

$$\det(xI - L) = \det \begin{bmatrix} x & 0 & \ldots & 0 & a_0 \\ -1 & x & \ddots & \vdots & a_1 \\ & \ddots & \ddots & 0 & \vdots \\ & & -1 & x & a_{m-2} \\ & & & -1 & x + a_{m-1} \end{bmatrix}.$$

Gegeben eine quadratische Matrix $M \in \mathbb{K}^{m\times m}$ kann die *Normalform* für M über die *Begleitmatrix* $L(M)$ ihres *charakteristischen Polynoms* $p(M) = \det(xI - M)$ definiert werden.

Bemerkung. In [Coh95, Abschnitt 8.1] wird $C(p) = xI - L(p)^\top$ auch *Begleitmatrix* genannt. Das ist gerechtfertigt in dem man $C(p)$ als *lineares Matrixbüschel* $C(p) = C_0 \otimes 1 + C_x \otimes x$ sieht. Das kann direkt für nicht-kommutative Polynome verallgemeinert werden.

Nun verlassen wir $\mathbb{K}[x] = \mathbb{K}\langle\{x\}\rangle$ und betrachten Polynome $p \in \mathbb{K}\langle X\rangle$. Es gibt zwei Fälle, in denen ein *minimales* ZLS direkt angegeben werden kann, nämlich wenn der Träger (des Polynoms) von links (in der linken Familie) oder von rechts (in der rechten Familie) mit *strikt* ansteigendem Rang „erzeugt“ werden kann. Zum Beispiel ist ein *minimales* polynomielles ZLS für $p = a_0 + a_1(x+2y) + a_2(x-z)(x+2y) + y(x-z)(x+2y)$ gegeben durch

$$\begin{bmatrix} 1 & -y - a_2 & -a_1 & -a_0 \\ . & 1 & -(x-z) & . \\ . & . & 1 & -(x+2y) \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} p \\ (x-z)(x+2y) \\ x + 2y \\ 1 \end{bmatrix}.$$

Definition 2.2.3 (Linke und rechte Begleitsysteme [Sch17c, Definition 3.2])**.** Für $i = 1, 2, \ldots, m$ sei $q_i \in \mathbb{K}\langle X\rangle$ mit $\operatorname{rang}(q_i) = 2$ und $a_i \in \mathbb{K}$. Für ein Polynom $p \in \mathbb{K}\langle X\rangle$ der Form

$$p = q_m q_{m-1} \cdots q_1 + a_{m-1} q_{m-1} \cdots q_1 + \ldots + a_2 q_2 q_1 + a_1 q_1 + a_0$$

heißt das polynomielle ZLS

$$\begin{bmatrix} 1 & -q_m - a_{m-1} & -a_{m-2} & \ldots & -a_1 & -a_0 \\ & 1 & -q_{m-1} & 0 & \ldots & 0 \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & -q_2 & 0 \\ & & & & 1 & -q_1 \\ & & & & & 1 \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{2.2.4}$$

linkes Begleitsystem. Und für ein Polynom $p \in \mathbb{K}\langle X\rangle$ der Form

$$p = a_0 + a_1 q_1 + a_2 q_1 q_2 + \ldots + a_{m-1} q_1 q_2 \cdots q_{m-1} + q_1 q_2 \cdots q_m$$

heißt das polynomielle ZLS

$$\begin{bmatrix} 1 & -q_1 & 0 & \ldots & 0 & -a_0 \\ & 1 & -q_2 & \ddots & \vdots & -a_1 \\ & & \ddots & \ddots & 0 & \vdots \\ & & & 1 & -q_{m-1} & -a_{m-2} \\ & & & & 1 & -q_m - a_{m-1} \\ & & & & & 1 \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{2.2.5}$$

rechtes Begleitsystem.

Proposition 2.2.6 ([Sch17c, Proposition 3.5])**.** *Für $i = 1, 2, \ldots, m$ sei $q_i \in \mathbb{K}\langle X\rangle$ mit $\operatorname{rang}(q_i) = 2$. Dann haben die Polynome*

$$p_l = q_m q_{m-1} \cdots q_1 + a_{m-1} q_{m-1} \cdots q_1 + \ldots + a_2 q_2 q_1 + a_1 q_1 + a_0 \quad \textit{und}$$
$$p_r = a_0 + a_1 q_1 + a_2 q_1 q_2 + \ldots + a_{m-1} q_1 q_2 \cdots q_{m-1} + q_1 q_2 \cdots q_m$$

jeweils Rang $m + 1$.

Beweis. Beide, die linke Familie $(p, q_{m-1} \cdots q_1, \ldots, q_2 q_1, q_1, 1)$ und die rechte Familie $(1, q_m + a_{m-1}, (q_m + a_{m-1})q_{m-1} + a_{m-2}, \ldots, p)$ sind für (2.2.4) $\mathbb{K}$-linear unabhängig. Damit ist das linke Begleitsystem (für p_l) minimal mit der Dimension $m+1$. Also gilt $\operatorname{rang}(p_\mathrm{l}) = m + 1$. Mit einem ähnlichen Argument erhalten wir $\operatorname{rang}(p_\mathrm{r}) = m + 1$ für das rechte Begleitsystem. □

Für ein allgemeines Polynom $p \in \mathbb{K}\langle X\rangle$ mit $\operatorname{rang}(p) = n \geq 2$ können wir jedes *minimale* polynomielle ZLS $\mathcal{A} = (1, A, \lambda)$ verwenden, um ein ZLS der Form $(1, A', 1)$

zu bekommen, in dem wir die letzte Zeile mit $1/\lambda$ und die letzte Spalte mit λ multiplizieren. Nun können wir eine (verallgemeinerte Version der) Begleitmatrix definieren. Dieses Begleitmatrizen können als Bausteine verwendet werden um Begleitmatrizen für Produkte von Polynomen zu erhalten. Das ist nur ein anderer Blick auf die *minimale Polynommultiplikation*, Proposition 3.2.7.

Bemerkung. Obwohl im allgemeinen nichts über die Minimalität von linearen Darstellungen für *kommutative* Polynome (in mehreren Variablen) gesagt werden kann, lässt sich Proposition 2.2.6 verwenden um minimale lineare Darstellungen für den kommutativen Fall zu konstruieren. Denn in diesem Fall ist der Rang das Maximum der Ränge der Monome, zum Beispiel $p = x^2y + xyz = yxx + xyz = xy(x + z)$.

Definition 2.2.7 (Begleitmatrizen [Sch17c, Definition 3.6])**.** Sei $p \in \mathbb{K}\langle X\rangle$ mit $\operatorname{rang}(p) = n \geq 2$ gegeben durch das *minimale* polynomielle ZLS $\mathcal{A} = (1, A, 1)$ und bezeichne $C(p)$ die obere rechte Teilmatrix der Größe $(n-1) \times (n-1)$. Dann heißt $C(p)$ (nicht-kommutative) *Begleitmatrix* von p.

2.3 Rationale Operationen

Proposition 2.3.1 (Rationale Operationen [CR99, Abschnitt 1])**.** *Seien $0 \neq f, g \in \mathbb{F}$ gegeben durch die zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f)$ beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g)$ und sei $0 \neq \mu \in \mathbb{K}$. Dann können zulässige lineare Systeme für die rationalen Operationen folgendermaßen konstruiert werden:*

Die skalare Multiplikation μf ist gegeben durch

$$\mu\mathcal{A}_f = \big(u_f, A_f, \mu v_f\big).$$

Die Summe $f + g$ ist gegeben durch

$$\mathcal{A}_f + \mathcal{A}_g = \left(\begin{bmatrix} u_f & . \end{bmatrix}, \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ . & A_g \end{bmatrix}, \begin{bmatrix} v_f \\ v_g \end{bmatrix}\right).$$

Das Produkt fg ist gegeben durch

$$\mathcal{A}_f \cdot \mathcal{A}_g = \left(\begin{bmatrix} u_f & . \end{bmatrix}, \begin{bmatrix} A_f & -v_f u_g \\ . & A_g \end{bmatrix}, \begin{bmatrix} . \\ v_g \end{bmatrix}\right).$$

Und die Inverse f^{-1} ist gegeben durch

$$\mathcal{A}_f^{-1} = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} -v_f & A_f \\ . & u_f \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix}\right).$$

Bemerkung. Alternativ ist die Summe $f+g$ auch gegeben durch das zulässige lineare System $\mathcal{A}_g + \mathcal{A}_f$, das damit trivialerweise zu $\mathcal{A}_f + \mathcal{A}_g$ äquivalent ist. Warum gibt es in diesem Fall (unabhängig von der Minimalität) immer eine zulässige Transformation (P, Q), sodass $\mathcal{A}_g + \mathcal{A}_f = P(\mathcal{A}_f + \mathcal{A}_g)Q$ ist?

Bemerkung. Es lässt sich einfach prüfen, dass die entsprechenden Lösungsvektoren (linken Familien) der oben definierten zulässigen linearen Systeme

$$\mu s_f, \quad \begin{bmatrix} s_f + u_f^\top g \\ s_g \end{bmatrix}, \quad \begin{bmatrix} s_f g \\ s_g \end{bmatrix} \quad \text{bzw.} \quad \begin{bmatrix} h^{-1} \\ s_h h^{-1} \end{bmatrix}$$

sind. Vergewissern sollte man sich noch, ob die jeweiligen Systemmatrizen tatsächlich *voll* sind. Die Details dazu finden sich in [Coh95, Abschnitt 4.3].

Die folgenden zwei Bemerkungen sollte man beim ersten Mal lesen überspringen. Wenn man sich am Ende des Intermezzos noch an diese Stelle erinnert, kann man hierher zurückkehren um etwas tiefer in die nicht-kommutative Algebra einzutauchen.

Bemerkung. Alternativ dazu kann man es sich auch folgendermaßen überlegen [Sch17b, Abschnitt 1]: Für die Summe und das Produkt ist die Vollheit (der Systemmatrix) aus der Tatsache klar, dass die freie assoziative Algebra —als ein *freier Idealring* (FIR)— *unbeschränkte Erzeugungsnummer* (UGN für *unbounded generating number*) hat und deshalb die diagonale Summe voller Matrizen wiederum voll ist [Coh85, Abschnitt 7.3]. Die grundlegenden Begriffe finden sich in [Coh85, Abschnitt 0.2]. Die Systemmatrix für die Inverse ist voll weil $h \neq 0$ ist und damit die Linearisierung (der „Block“) von $\mathcal{A}_h$ voll ist [CR99].

Bemerkung. Eine Diskussion, welche Bedingungen notwendig für die Einbettbarkeit eines Rings in einen Schiefkörper sind, findet man in [Coh95, Abschnitt 1.4]. Ein Ring hat *unbeschränkte Erzeugungsnummer* (UGN) genau dann, wenn jede invertierbare Matrix voll ist [Coh95, Proposition 1.4.3]. In unserem Fall, nämlich der Einbettung der freien assoziativen Algebra in den freien Schiefkörper, passt alles wunderschön zusammen. Aber das alles im Detail nachzuvollziehen ist mit einigem Aufwand verbunden. Das wird unter anderem auch im Kapitel 4 in Lams zweitem „nicht-kommutativen“ Lehrbuch [Lam99] klar, in dem einige Details von Cohn zusammengefasst sind.

Bevor wir zwei Varianten einer Konstruktion für die Multiplikation (Proposition 2.3.6 und 2.3.9) formulieren, die —unter bestimmten Voraussetzungen— zu einem kleineren System für das Produkt führt, brauchen wir Techniken um ein *minimales* System $\mathcal{A}$ in eine spezielle Form zu bringen, falls $1 \in L(\mathcal{A})$ und/oder $1 \in R(\mathcal{A})$ ist. Dazu brauchen wir zunächst ein paar technische Hilfsmittel. Die Lemmata 2.3.2 und 2.3.3 sind so formuliert, dass sie auch den trivialen Fall enthalten. Lemma 2.3.4 wird mit Hilfe von Lemma 2.3.2 bewiesen.

Lemma 2.3.2 ([Sch17c, Lemma 2.3])**.** *Sei $\mathcal{A} = (u, A, v)$ ein ZLS der Dimension $n \geq 1$ mit $\mathbb{K}$-linear unabhängiger linker Familie $s = A^{-1}v$ und $B = B_0 \otimes 1 + B_1 \otimes x_1 + \ldots + B_d \otimes x_d$ mit $B_\ell \in \mathbb{K}^{m \times n}$ sodass $Bs = 0$ ist. Dann existiert eine (eindeutige) Matrix $T \in \mathbb{K}^{m \times n}$ sodass $B = TA$ ist.*

Beweis. Für $n = 1$ gilt trivialerweise $B = 0$ und damit $T = 0$. Also sei $n \geq 2$ und wir nehmen —ohne Beschränkung der Allgemeinheit— an, dass $v = [0, \ldots, 0, 1]^\top$ und $m = 1$ ist. Nachdem A voll ist und daher über dem freien Schiefkörper invertiert

werden kann, existiert eine eindeutige Matrix T sodass $B = TA$, nämlich $T = BA^{-1}$ in $\mathbb{F}^{1\times n}$. Die letzte Spalte in T ist Null, weil $0 = Bs = BA^{-1}v = Tv$. Nun bezeichne A' die Matrix A deren letzte Zeile entfernt wurde und A'_B die Matrix, die man erhält, wenn man die letzte Zeile in A durch B ersetzt. A'_B kann nicht voll sein, da $s \in \ker A'_B$ den Widerspruch $s = (A'_B)^{-1}0 = 0$ liefert.

Daher ist, wegen Lemma 2.1.3, A'_B assoziiert über $\mathbb{K}$ zu einer linearen hohlen Matrix. Und wir behaupten, dass es nur eine einzige Möglichkeit gibt, A'_B in eine hohle Matrix zu transformieren, nämlich mit (letzter) Nullzeile. Wenn wir (mit invertierbaren Transformationen) keinen Nullblock der Größe $(n-i)\times i$ in den ersten $n-1$ Zeilen von A'_B erzeugen können, dann können wir auch keinen Nullblock der Größe $(n-i+1)\times i$ erzeugen und wir sind fertig.

Nun nehmen wir gegenteilig an, dass es invertierbare Matrizen $P' \in \mathbb{K}^{(n-1)\times(n-1)}$ und (zulässiges) $Q \in \mathbb{K}^{n\times n}$ mit $(Q^{-1}s)_1 = s_1$ gibt, sodass $P'A'Q$ einen Nullblock der Größe $(n-i)\times i$ für ein $i = 1,\dots,n-1$ enthält. Es gibt zwei Fälle. Wenn die ersten $n-i$ Einträge in der ersten Spalte nicht zu Null gemacht werden können, konstruieren wir einen rechten oberen Nullblock:

$$\hat{A} = \begin{bmatrix} A_{11} & . \\ A_{21} & A_{22} \end{bmatrix}, \quad \hat{s} = Q^{-1}s \quad \text{und} \quad \hat{v} = Pv = v$$

wobei A_{11} die Größe $(n-i)\times(n-i)$ hat. Wäre A_{11} *nicht* voll, dann wäre auch A nicht voll (die letzte Zeile ist nicht in der Transformation involviert). Daher ist dieser Pivotblock invertierbar (über $\mathbb{F}$). Somit würde $\hat{s}_1 = \hat{s}_2 = \ldots = \hat{s}_{n-i} = 0$ gelten. Im anderen Fall konstruieren wir einen oberen linken Nullblock in PAQ. Aber dann würde $\hat{s}_{i+1} = \hat{s}_{i+2} = \ldots = \hat{s}_n = 0$ gelten. Beides widerspricht der $\mathbb{K}$-linearen Unabhängigkeit der linken Familie.

Daher ist A'_B über $\mathbb{K}$ zu einer linearen hohlen Matrix mit einem $1\times n$ Nullblock assoziiert, sagen wir in der letzten Zeile (die Spalten bleiben unberührt):

$$\begin{bmatrix} I_{n-1} & . \\ T' & 1 \end{bmatrix} \begin{bmatrix} A' \\ B \end{bmatrix} I_n = \begin{bmatrix} A' \\ . \end{bmatrix}.$$

Die Matrix $T = [-T', 0] \in \mathbb{K}^{1\times n}$ erfüllt $B = TA$. □

Bemerkung. Obwohl das ZLS in Lemma 2.3.2 nicht minimal sein muss, ist die Annahme der $\mathbb{K}$-linearen Unabhängigkeit der linken Familie aus zwei Gründen wichtig. Einer betrifft „pathologische“ Situationen, vergleiche mit Beispiel 4.2.4. Ein zu einem $s_j = 0$ korrespondierender Eintrag (in der Systemmatrix), sagen wir für $j = 3$, kann beliebig sein:

$$\begin{bmatrix} 1 & -x & . \\ . & . & z \\ . & 1 & -1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}.$$

Für $B = [2, -2x, y]$ hat die Transformation T nicht-skalare Einträge: $T = [2, yz^{-1}, 0]$. Der andere Grund betrifft den Ausschluss anderer Möglichkeiten der Hohlheit ausge-

nommen der letzten Zeile. Für $B = [0, 0, 1]$ ist die Matrix

$$A'_B = \begin{bmatrix} 1 & -x & . \\ . & . & z \\ . & . & 1 \end{bmatrix}$$

hohl. Jedoch ist die Transformation, nach der wir suchen $T = [0, z^{-1}, 0]$.

Lemma 2.3.3. *Sei $\mathcal{A} = (u, A, v)$ ein ZLS der Dimension $n \geq 1$ mit $\mathbb{K}$-linear unabhängiger rechter Familie $t = uA^{-1}$ und $B = B_0 \otimes 1 + B_1 \otimes x_1 + \ldots + B_d \otimes x_d$ mit $B_\ell \in \mathbb{K}^{n\times m}$ sodass $tB = 0$ ist. Dann existiert eine (eindeutige) Matrix $U \in \mathbb{K}^{n\times m}$ sodass $B = AU$ ist.*

Bemerkung. Wenn es nicht gerade um die Zulässigkeit einer Transformation geht, ist weder zwischen der Formulierung einer Aussage über die linke und die rechte Familie noch deren Beweis ein wesentlicher Unterschied. Der Grund, Aussagen zur rechten Familie zu „wiederholen“ liegt darin, dass das Nachvollziehen in Beweisen, wo diese Resultate gebraucht werden, einfacher wird (und man nicht jedes Mal umdenken muss).

Lemma 2.3.4 (für Typ $(*, 1)$ [Sch17b, Lemma 4.18])**.** *Sei $\mathcal{A} = (u, A, v)$ ein* minimales *ZLS der Dimension $n = \dim(\mathcal{A}) \geq 2$ und $1 \in L(\mathcal{A})$. Dann existiert eine zulässige Transformation (P, Q) sodass die letzte Zeile von PAQ die Form $[0, \ldots, 0, 1]$ und Pv die Form $[0, \ldots, 0, \lambda]^\top$ für ein $\lambda \in \mathbb{K}$ hat.*

Beweis. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $v = [0, \ldots, 0, 1]^\top$ und die linke Familie $(s_1, s_2, \ldots, s_{n-1}, 1)$ ist. Andernfalls kann sie mit einer zulässigen Transformation (P°, Q°) auf diese Form gebracht werden. Nun bezeichne $\bar{A}$ den oberen linken Block der Größe $(n-1) \times (n-1)$ von A, bezeichne $\bar{s} = (s_1, \ldots, s_{n-1})$ und schreibe $As = v$ als

$$\begin{bmatrix} \bar{A} & b \\ c & d \end{bmatrix} \begin{bmatrix} \bar{s} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Nun wenden wir Lemma 2.3.2 auf $B = [c, d-1]$ an um die Matrix $T = [\bar{T}, \tau] \in \mathbb{K}^{1\times n}$ mit der Eigenschaft $B = TA$ zu erhalten. Somit haben wir die Transformation

$$(P, Q) = \left(\begin{bmatrix} I_{n-1} & . \\ -\bar{T} & -\tau \end{bmatrix} P^\circ, Q^\circ \right).$$

□

Lemma 2.3.5 (für Typ $(1, *)$ [Sch17b, Lemma 4.19])**.** *Sei $\mathcal{A} = (u, A, v)$ ein* minimales *ZLS der Dimension $n = \dim(\mathcal{A}) \geq 2$ und $1 \in R(\mathcal{A})$. Dann existiert eine zulässige Transformation (P, Q) sodass die erste Spalte von PAQ die Form $[1, 0, \ldots, 0]^\top$ und Pv die Form $[0, \ldots, 0, \lambda]^\top$ für ein $\lambda \in \mathbb{K}$ hat.*

Multipliziert man zwei Polynome auf der Ebene (minimaler) zulässiger linearer Systeme, sieht man sofort, dass man das resultierende System der Dimension n auf

$n-1$ verkleinern kann, vergleiche auch Beispiel 3.2.1. Später, in Satz 3.5.2 (Minimale Faktormultiplikation), werden wir formulieren, unter welchen Voraussetzungen die folgenden Konstruktionen zu einem *minimalen* ZLS für das Produkt führen.

Proposition 2.3.6 (Multiplikation Typ $(1,*)$ [Sch17a, Proposition 2.8])**.** *Seien $f,g \in \mathbb{F}\setminus\mathbb{K}$ gegeben durch die zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f) = (1, A_f, \lambda_f)$ der Dimension n_f der Form*

$$\mathcal{A}_f = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} a & b' & b \\ a' & B & b'' \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda_f \end{bmatrix} \right) \tag{2.3.7}$$

beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g) = (1, A_g, \lambda_g)$. Dann ist ein ZLS für fg der Dimension $n = n_f + n_g - 1$ gegeben durch

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} a & b' & \lambda_f b u_g \\ a' & B & \lambda_f b'' u_g \\ . & . & A_g \end{bmatrix}, \begin{bmatrix} . \\ . \\ v_g \end{bmatrix} \right). \tag{2.3.8}$$

Beweis. Konstruiere das ZLS $\mathcal{A}' = (u', A', v')$ der Dimension $n_f + n_g$ für das Produkt fg laut Proposition 2.3.1. Addiere λ_f-mal Spalte n_f zu Spalte (n_f+1) (in der Systemmatrix A'). Entferne Spalte n_f von A' und v' und Zeile n_f von A' und u' um das ZLS (2.3.8) der Dimension $n_f + n_g - 1$ zu erhalten. □

Proposition 2.3.9 (Multiplikation Typ $(*,1)$ [Sch17a, Proposition 2.9])**.** *Seien $f,g \in \mathbb{F}\setminus\mathbb{K}$ gegeben durch die zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f) = (1, A_f, \lambda_f)$ der Dimension n_f beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g) = (1, A_g, \lambda_g)$ der Dimension n_g der Form*

$$\mathcal{A}_g = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & c' & c \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda_g \end{bmatrix} \right). \tag{2.3.10}$$

Dann ist ein ZLS für fg der Dimension $n = n_f + n_g - 1$ gegeben durch

$$\mathcal{A} = \left(\begin{bmatrix} u_f & . & . \end{bmatrix}, \begin{bmatrix} A_f & e_{n_f}\lambda_f b' & e_{n_f}\lambda_f b \\ . & B & b'' \\ . & c' & c \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda_g \end{bmatrix} \right). \tag{2.3.11}$$

Beweis. Konstruiere das ZLS $\mathcal{A}' = (u', A', v')$ der Dimension $n_f + n_g$ für das Produkt fg laut Proposition 2.3.1. Addiere λ_f-mal Zeile (n_f+1) zu Zeile n_f (in der Systemmatrix A'). Entferne Zeile (n_f+1) von A' und v' und Spalte (n_f+1) von A' und u' um das ZLS (2.3.11) der Dimension $n_f + n_g - 1$ zu erhalten. □

Bemerkung 2.3.12 ([Sch17b, Abschnitt 4])**.** Wenn $1 \in R(\mathcal{A})$ für ein *minimales* ZLS $\mathcal{A} = (u, A, v)$ gilt, sagen wir $\dim(\mathcal{A}) = n$, dann existiert wegen Lemma 2.3.5 eine zulässige Transformation (P, Q) sodass die erste Spalte in PAQ gleich $[1, 0, \ldots, 0]^\top$

ist. Wenn die erste Spalte von $A = (a_{ij})$ nicht bereits diese Form hat, kann eine zulässige Transformation in zwei Schritten gefunden werden: Zuerst stellen wir ein lineares Gleichungssystem mit einem $(n-1)$-Tupel an Skalaren $(\mu_2, \mu_3, \ldots, \mu_n)$ auf sodass $a_{i1} + \mu_2 a_{i2} + \mu_3 a_{i3} + \ldots + \mu_n a_{in} \in \mathbb{K}$ für alle $i = 1, 2, \ldots, n$ gilt. Und danach können wir mit elementaren Zeilentransformationen (Gaußsche Elimination in der ersten Spalte) und —wenn notwendig— Permutationen die gewünschte Form der ersten Spalte erreichen. Zusammen ergeben diese Zeilen- und Spaltentransformationen eine (zulässige) Transformation (P', Q').

Für den Fall $1 \in L(\mathcal{A})$ kann man ähnlich (Zeilen und Spalten vertauscht) vorgehen, wenn die letzte Komponente der rechten Seite ungleich Null ist. Klarerweise können beide Fälle kombiniert werden, zum Beispiel für die Anwendung der Minimalen Inversen Typ $(1, 1)$, Satz 2.5.13. Diese Vorgehensweise kann auch allgemeiner für nicht-minimale Systeme funktionieren, aber in „pathologischen“ Fällen versagen, vergleiche auch mit Beispiel 4.2.4.

Bemerkung. Ist f in Proposition 2.3.6 (bzw. g in Proposition 2.3.9) durch ein *minimales* ZLS gegeben, kann es einfach in die Form (2.3.7) (bzw. (2.3.10)) gebracht werden.

Bemerkung. Die in Bemerkung 2.3.12 beschriebene Vorgehensweise sollte „behutsam“ ausgeführt werden um eine mögliche Blockstruktur nicht zu zerstören. Das heißt, dass man in solchen Fällen mit einem „kleineren“ Gleichungssystem auskommen oder teilweise ganz darauf verzichten kann. Letzteres ist in Beispiel 4.1.4 illustriert.

2.4 Disjunkte Addition

Definition 2.4.1 (Disjunkte Elemente, Primärelemente [CR99])**.** Zwei Elemente $f, g \in \mathbb{F}$ heißen *disjunkt* wenn $\operatorname{rang}(f + g) = \operatorname{rang}(f) + \operatorname{rang}(g)$ ist. Ein Element $f \in \mathbb{F}$ heißt *primär* (oder *unzerlegbar*), wenn es keine minimale lineare Darstellung $\pi_f = (u, A, v)$ mit einer zerlegbaren Systemmatrix $A = A_1 \oplus A_2 = \left[\begin{smallmatrix} A_1 & \cdot \\ \cdot & A_2 \end{smallmatrix}\right]$ von f gibt.

Satz 2.4.2 (Primärzerlegung [CR99, Satz 2.3])**.** *Jedes Element des freien Schiefkörpers kann eindeutig als Summe disjunkter Primärelemente geschrieben werden; maximal eines davon ist ein Polynom.*

Bemerkung. Die Annahme der *Minimalität* in Definition 2.4.1 kann weggelassen werden, wenn man die „zulässige“ Zerlegbarkeit der linearen Darstellung

$$\pi_f = \left(\begin{bmatrix} u_1 & u_2 \end{bmatrix}, \begin{bmatrix} A_1 & \cdot \\ \cdot & A_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right)$$

mit $u_1, u_2 \neq 0$ und $v_1, v_2 \neq 0$ betrachtet.

Für disjunkte Elemente lässt sich sofort eine minimale Addition (Proposition 2.4.3) formulieren. Zu testen, ob zwei Elemente in $\mathbb{F}$ disjunkt sind, ist schwierig, weil man

Techniken zur Minimierung linearer Darstellungen braucht. Darum wird es in Kapitel 4 gehen.

Da aber Minimalität einer linearen Darstellung zur $\mathbb{K}$-linearen Unabhängigkeit ihrer linken beziehungsweise rechten Familie äquivalent ist (Proposition 2.1.8), sind zwei Elemente disjunkt, wenn *alle* Komponenten ihrer linken beziehungsweise rechten Familien von *linear unabhängigen* Klassen des freien Schiefkörpers sind, das heißt, die Vereinigung zweier $\mathbb{K}$-linear unabhängigen Teilmengen *unterschiedlicher* Klassen ist $\mathbb{K}$-linear unabhängig, zum Beispiel

$$\mathbb{F} = \mathbb{K}\langle X\rangle \uplus \mathbb{K}^{\mathrm{rat}}\langle\langle X\rangle\rangle \setminus \mathbb{K}\langle X\rangle \uplus \mathbb{F} \setminus \mathbb{K}^{\mathrm{rat}}\langle\langle X\rangle\rangle.$$

Für $f = x + \left((1-x)^{-1} + x^{-1}\right)$ ist ein *minimales* ZLS —konstruiert laut Proposition 2.4.3— gegeben durch

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & . & . \\ . & . & 1-x & x+1 \\ . & . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} f \\ 1 \\ (1+x)^{-1} + x^{-1} \\ x^{-1} \end{bmatrix}.$$

Proposition 2.4.3 (Minimale disjunkte Addition)**.** *Seien $f, g \in \mathbb{F}$ disjunkt und gegeben durch die* minimalen *zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f)$ der Dimension n_f beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g)$ der Dimension n_g. Dann ist das ZLS*

$$\mathcal{A}_f + \mathcal{A}_g = \left(\begin{bmatrix} u_f & . \end{bmatrix}, \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ . & A_g \end{bmatrix}, \begin{bmatrix} v_f \\ v_g \end{bmatrix} \right)$$

der Dimension $n_f + n_g$ (von Proposition 2.3.1) für $f + g$ minimal.

2.5 Minimale Inverse

Bevor die wesentlichen Ergebnisse dieses Abschnitts in Satz 2.5.13 (Minimale Inverse) zusammengefasst werden, erfolgt die Herleitung in zwei Stufen: Zuerst geht es darum, die „Form“ eines zulässigen linearen Systems nach zweimaliger Anwendung $f = (f^{-1})^{-1}$ weitestgehend zu erhalten. Danach kann man mehrere Fälle unterscheiden, um ein *minimales* ZLS für die Inverse zu konstruieren.

Proposition 2.5.1 (Standardinverse [Sch17b, Proposition 4.2])**.** *Sei $0 \neq f \in \mathbb{F}$ gegeben durch das zulässige lineare System $\mathcal{A} = (u, A, v)$ der Dimension n. Dann ist ein ZLS der Dimension $n+1$ für f^{-1} gegeben durch*

$$\mathcal{A}^{-1} = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} \Sigma v & -\Sigma A \Sigma \\ . & u\Sigma \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right). \tag{2.5.2}$$

Beweis. Es ist einfach zu prüfen, dass der Lösungsvektor von $\mathcal{A}^{-1}$

$$\begin{bmatrix} f^{-1} \\ \Sigma_n s_f f^{-1} \end{bmatrix}$$

ist. Vergleiche mit Proposition 2.3.1. □

Definition 2.5.3 (Standardinverse)**.** Sei $\mathcal{A}$ ein ZLS für ein Element ungleich Null. Das ZLS (2.5.2) heißt *Standardinverse* von $\mathcal{A}$, bezeichnet mit $\mathcal{A}^{-1}$.

Lemma 2.5.4 (Inverse Typ $(1,1)$ [Sch17b, Lemma 4.5])**.** *Angenommen* $0 \neq f \in \mathbb{F}$ *habe ein* minimales *zulässiges lineares System der Dimension* n *der Form*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right).$$

Dann ist ein minimales ZLS für f^{-1} *der Dimension* $n-1$ *gegeben durch*

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} -\lambda\Sigma b'' & -\Sigma B\Sigma \\ -\lambda b & -b'\Sigma \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right)$$

mit $1 \notin R(\mathcal{A}')$ *und* $1 \notin L(\mathcal{A}')$.

Beweis. Die Standardinverse von $\mathcal{A}$ ist

$$\begin{bmatrix} \lambda & -1 & . & . \\ . & -\Sigma b'' & -\Sigma B\Sigma & . \\ . & -b & -b'\Sigma & -1 \\ . & . & . & 1 \end{bmatrix} \tilde{s} = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Addiert man Zeile 4 zu Zeile 3 und λ-mal Spalte 2 zu Spalte 1 erhält man

$$\begin{bmatrix} 0 & -1 & . & . \\ -\lambda\Sigma b'' & -\Sigma b'' & -\Sigma B\Sigma & . \\ -\lambda b & -b & -b'\Sigma & 0 \\ . & . & . & 1 \end{bmatrix} s' = \begin{bmatrix} . \\ . \\ 1 \\ 1 \end{bmatrix}.$$

Es folgt, dass $s'_2 = 0$ ist und s'_{n+1} nichts zur Lösung s'_1 beiträgt. Daher kann man sowohl die erste und die letzte Zeile als auch die zweite und die letzte Spalte entfernen. Wäre $\mathcal{A}'$ nicht minimal, würde ein ZLS $\mathcal{A}''$ der Dimension $m < n-1$ für f^{-1} existieren. Die Standardinverse $(\mathcal{A}'')^{-1}$ ergäbe ein System der Dimension $m+1 < n$ für f, was der angenommenen Minimalität von $\mathcal{A}$ widersprechen würde. Es bleibt zu zeigen, dass $1 \notin R(\mathcal{A}')$ und $1 \notin L(\mathcal{A}')$ ist. Sei $t = (t_1, t_2, \ldots, t_n)$ die rechte Familie von $\mathcal{A}$, die (wegen der Minimalität) $\mathbb{K}$-linear unabhängig ist. Dann ist die rechte Familie von $\mathcal{A}^{-1}$ gleich $(f^{-1}t_n, \ldots, f^{-1}t_2, f^{-1}t_1, f^{-1})$, die nach der ersten Zeilenoperation ist $f^{-1}(t_n, \ldots, t_2, t_1, 1 - t_1)$. Entfernt man die erste und die letzte Komponente (entsprechend der ersten und der letzten Zeile), erhält man die rechte Familie $f^{-1}(t_{n-1}, \ldots, t_2, t_1)$. Daher ist $1 \notin R(\mathcal{A}')$, weil sonst $f \in \operatorname{span}\{t_{n-1}, \ldots, t_2, t_1\}$ der $\mathbb{K}$-linearen Unabhängigkeit von t widersprechen würde. Ähnliche Argumente zeigen, dass $1 \notin L(\mathcal{A}')$ ist. □

Lemma 2.5.5 (Inverse Typ $(1,0)$ [Sch17b, Lemma 4.8])**.** *Angenommen* $0 \neq f \in \mathbb{F}$ *habe ein* minimales *zulässiges System der Dimension* n *der Form*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & c' & c \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right)$$

mit $1 \notin L(\mathcal{A})$*. Dann ist eine minimales ZLS für* f^{-1} *der Dimension* n *gegeben durch*

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -\frac{1}{\lambda}c & -\frac{1}{\lambda}c'\Sigma \\ . & -\Sigma b'' & -\Sigma B\Sigma \\ . & -b & -b'\Sigma \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right) \tag{2.5.6}$$

mit $1 \notin L(\mathcal{A}')$*.*

Beweis. Die Standardinverse von $\mathcal{A}$ ist

$$\begin{bmatrix} \lambda & -c & -c'\Sigma & . \\ . & -\Sigma b'' & -\Sigma B\Sigma & . \\ . & -b & -b'\Sigma & -1 \\ . & . & . & 1 \end{bmatrix} \tilde{s} = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}$$

und hat Dimension $n+1$. Nachdem man Zeile $n+1$ zu Zeile n addiert hat, kann man Zeile $n+1$ und Spalte $n+1$ entfernen, weil $\tilde{s}_{n+1}$ nichts zur Lösung $\tilde{s}_1 = f^{-1}$ beiträgt. Nun dividieren wir die erste Zeile durch λ und erhalten (2.5.6). Es ist noch zu zeigen, dass $\mathcal{A}'$ minimal und $1 \notin L(\mathcal{A}')$ ist. Sei $(s_1, s_2, \ldots, s_n)$ die linke Familie von $\mathcal{A}$, die (wegen der Minimalität) $\mathbb{K}$-linear unabhängig ist. Dann ist die linke Familie von $\mathcal{A}^{-1}$ gleich $(f^{-1}, s_n f^{-1}, \ldots, s_2 f^{-1}, 1)$. Man beachte, dass (zulässige) Zeilenoperationen keinen Einfluss auf die linke Familie haben. Nachdem wir den letzten Eintrag $s_1 f^{-1} = \tilde{s}_{n+1} = 1$ entfernt haben, ist die linke Familie von $\mathcal{A}'$ gleich $(1, s_n, \ldots, s_2)f^{-1}$. Laut Annahme ist $1 \notin L(\mathcal{A})$. Daher ist $1 \notin \operatorname{span}\{s_2, s_3, \ldots, s_n\}$, also ist $(1, s_n, \ldots, s_2)$ $\mathbb{K}$-linear unabhängig. Klarerweise ist $1 \notin L(\mathcal{A}')$, weil $f \notin \operatorname{span}\{1, s_n, \ldots, s_2\}$ ist. Analog sei $(t_1, t_2, \ldots, t_n)$ die rechte Familie von $\mathcal{A}$, die ebenfalls $\mathbb{K}$-linear unabhängig ist. Dann ist die rechte Familie von $\mathcal{A}^{-1}$ gleich $(f^{-1}t_n, \ldots, f^{-1}t_2, f^{-1}t_1, f^{-1})$, die nach der Zeilenoperation $(f^{-1}t_n, \ldots, f^{-1}t_2, f^{-1}t_1, f^{-1} - f^{-1}t_1)$. Nachdem wir den letzten Eintrag entfernt haben, ist die rechte Familie von $\mathcal{A}'$ gleich $f^{-1}(t_n, \ldots, t_2, t_1)f^{-1}$ und daher klarerweise $\mathbb{K}$-linear unabhängig. Damit ist $\mathcal{A}'$ laut Proposition 2.1.8 minimal. □

Lemma 2.5.7 (Inverse Typ $(0,1)$ [Sch17b, Lemma 4.11])**.** *Angenommen* $0 \neq f \in \mathbb{F}$ *habe ein* minimales *zulässiges lineares System der Dimension* n *der Form*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} a & b' & b \\ a' & B & b'' \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right) \tag{2.5.8}$$

mit $1 \notin R(\mathcal{A})$. *Dann ist ein minimales ZLS für* f^{-1} *der Dimension* n *gegeben durch*

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} -\lambda\Sigma b'' & -\Sigma B\Sigma & -\Sigma a'. \\ -\lambda b & -b'\Sigma & -a \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

mit $1 \notin R(\mathcal{A}')$.

Beweis. Die Standardinverse von $\mathcal{A}$ ist

$$\begin{bmatrix} \lambda & -1 & . & . \\ . & -\Sigma b'' & -\Sigma B\Sigma & -\Sigma a' \\ . & -b & -b'\Sigma & -a \\ . & . & . & 1 \end{bmatrix} \tilde{s} = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Nachdem man λ-mal Spalte 2 zu Spalte 1 addiert, kann man Zeile 1 und Spalte 2 entfernen, weil $\tilde{s}_2 = 0$ ist. Das Zeigen der Minimalität und $1 \notin R(\mathcal{A}')$ ist ähnlich zum Beweis von Lemma 2.5.5 (Spaltenoperationen wirken sich auf die linke Familie aus). □

Lemma 2.5.9 (Inverse Typ $(0,0)$)**.** *Sei* $\mathcal{A} = (u, A, v)$ *ein* minimales *zulässiges lineares System der Dimension* n *für* $0 \neq f \in \mathbb{F}$ *mit* $1 \notin R(\mathcal{A})$ *und* $1 \notin L(\mathcal{A})$. *Dann liefert die Standardinverse* $\mathcal{A}^{-1}$ *ein minimales ZLS der Dimension* $n+1$ *für* f^{-1}.

Beweis. Wäre $\mathcal{A}^{-1}$ nicht minimal, dann gäbe es ein System $\mathcal{A}'$ der Dimension $m < n+1$ für f^{-1}. Die Anwendung von Lemma 2.5.4 würde ein ZLS der Dimension $m-1 < n$ für f und damit einen Widerspruch zur angenommenen Minimalität von $\mathcal{A}$ liefern. □

Beispiel 2.5.10. Verwenden wir das minimale ZLS (für den Antikommutator) von der Einleitung, erhalten wir laut Lemma 2.5.4 (Inverse Typ $(1,1)$) ein *minimales* ZLS für $(xy + yx)^{-1}$:

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & -1 & . \\ y & . & -1 \\ . & y & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right).$$

Lemma 2.5.9 (Inverse Typ $(0,0)$) ergibt wieder ein minimales System für $xy + yx$.

Beispiel 2.5.11. Das Element xyz^{-1} kann mittels des *minimalen* ZLS

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & . \\ . & 1 & -y \\ . & . & z \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

mit rechter Familie $t = (1, x, xyz^{-1})$ und linker Familie $s = (xyz^{-1}, yz^{-1}, z^{-1})$ dargestellt werden. Nun kann Lemma 2.5.5 (Inverse Typ $(1,0)$) angewendet werden, um das minimale ZLS

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -z & . \\ . & y & -1 \\ . & . & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

für $zy^{-1}x^{-1}$ zu erhalten. Nachem $1 \notin L(\mathcal{A}')$ ist, kann Lemma 2.5.5 erneut angewendet werden.

Proposition 2.5.12 ([CR99, Proposition 2.1]). *Sei $f \in \mathbb{K}\langle\langle X\rangle\rangle$.*

(i) f ist eine Potenzreihe genau dann, wenn in jeder minimalen linearen Darstellung der konstante Term A_0 ihrer Systemmatrix $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$ invertierbar ist. In diesem Fall gibt es eine minimale lineare Darstellung mit $A_0 = I$ beziehungsweise $A = I - M$.

(ii) f ist ein Polynom genau dann, wenn in jeder minimalen linearen Darstellung mit einer Systemmatrix der Form $A = I - M$ die Matrix M nilpotent ist. In diesem Fall gibt es eine minimale lineare Darstellung mit einer strikten oberen Dreiecksmatrix M.

Satz 2.5.13 (Minimale Inverse [Sch17b, Satz 4.20]). *Sei $f \in \mathbb{F} \setminus \mathbb{K}$ gegeben durch das* minimale *zulässige lineare System $\mathcal{A} = (u, A, v)$ der Dimension n. Dann erhält man ein* minimales *ZLS für f^{-1} folgendermaßen:*

f vom Typ $(1,1)$ ergibt f^{-1} vom Typ $(0,0)$ mit $\dim(\mathcal{A}') = n-1$:

$$\mathcal{A}' = \left(1, \begin{bmatrix} -\lambda\Sigma b'' & -\Sigma B\Sigma \\ -\lambda b & -b'\Sigma \end{bmatrix}, 1\right) \quad \textit{für} \quad \mathcal{A} = \left(1, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & . & 1 \end{bmatrix}, \lambda\right).$$

f vom Typ $(1,0)$ ergibt f^{-1} vom Typ $(1,0)$ mit $\dim(\mathcal{A}') = n$:

$$\mathcal{A}' = \left(1, \begin{bmatrix} 1 & -\frac{1}{\lambda}c & -\frac{1}{\lambda}c'\Sigma \\ . & -\Sigma b'' & -\Sigma B\Sigma \\ . & -b & -b'\Sigma \end{bmatrix}, 1\right) \quad \textit{für} \quad \mathcal{A} = \left(1, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & c' & c \end{bmatrix}, \lambda\right).$$

f vom Typ $(0,1)$ ergibt f^{-1} vom Typ $(0,1)$ mit $\dim(\mathcal{A}') = n$:

$$\mathcal{A}' = \left(1, \begin{bmatrix} -\lambda\Sigma b'' & -\Sigma B\Sigma & -\Sigma a'. \\ -\lambda b & -b'\Sigma & -a \\ . & . & 1 \end{bmatrix}, 1\right) \quad \textit{für} \quad \mathcal{A} = \left(1, \begin{bmatrix} a & b' & b \\ a' & B & b'' \\ . & . & 1 \end{bmatrix}, \lambda\right).$$

f vom Typ $(0,0)$ ergibt f^{-1} vom Typ $(1,1)$ mit $\dim(\mathcal{A}') = n+1$:

$$\mathcal{A}' = \left(1, \begin{bmatrix} \Sigma v & -\Sigma A\Sigma \\ . & u\Sigma \end{bmatrix}, 1\right).$$

(Die Permutationsmatrix Σ dreht die Reihenfolge der Zeilen/Spalten um.)

Beweis. Siehe Lemmata 2.5.4, 2.5.5, 2.5.7 und 2.5.9. □

Korollar 2.5.14. *Sei $p \in \mathbb{K}\langle X\rangle$ mit $\operatorname{rang}(p) = n \geq 2$. Dann ist $\operatorname{rang}(p^{-1}) = n-1$.*

Beweis. Diese Aussage folgt aus Proposition 2.5.12 und der Inversen Typ $(1,1)$. □

Korollar 2.5.15. *Sei* $0 \neq f \in \mathbb{F}$. *Dann ist* $f \in \mathbb{K}$ *genau dann, wenn* $\operatorname{rang}(f) = \operatorname{rang}(f^{-1}) = 1$ *ist.*

Bemerkung. Diese einfache Konsequenz aus Satz 2.5.13 ermöglicht die Unterscheidung zwischen *trivialen* Einheiten (skalare Elemente ungleich Null) und *nicht-trivialen* Einheiten, das heißt, Elemente in $\mathbb{F} \setminus \mathbb{K}$. Darauf kommen wir (indirekt) in der Faktorisierungstheorie in Abschnitt 3.4 zurück, wo wir nur (das Einfügen von) *trivialen* Einheiten (in Faktorisierungen) erlauben, formuliert in Lemma 3.4.7.

Bemerkung. Klarerweise ist $n \geq 2$ für die Typen $(1,1)$, $(1,0)$ und $(0,1)$. Der Block B ist immer quadratisch und von der Größe $n-2$. Für $n = 2$ ist die Systemmatrix von $\mathcal{A}$

- $\left[\begin{smallmatrix} 1 & b \\ . & 1 \end{smallmatrix}\right]$ für Typ $(1,1)$,
- $\left[\begin{smallmatrix} 1 & b \\ . & c \end{smallmatrix}\right]$ für Typ $(1,0)$ und
- $\left[\begin{smallmatrix} a & b \\ . & 1 \end{smallmatrix}\right]$ für Typ $(0,1)$.

Bemerkung. Wie bereits vorher in Bemerkung 2.3.12 erwähnt, kann ein *minimales* ZLS in Abhängigkeit des Typs in die entsprechende Form gebracht werden. Die Umkehrung ist trivial. Wenn ein ZLS $\mathcal{A}$ zum Beispiel die Form von (2.5.8) hat, folgt sofort $1 \in L(\mathcal{A})$.

2.6 Rationale Identitäten

Bereits mit Hilfe der minimalen Inversen von Satz 2.5.13 kann man —ausgehend von Proposition 2.2.1— relativ systematisch rationale Identitäten zeigen. „Relativ“ bedeutet hier, dass man meist sofort sieht, welche Umformungen notwendig sind. Für die Implementierung (Programmierung) ist bereits das folgende Beispiel eine kleine Herausforderung, wenn man —abgesehen von der Inversen— nur die rationalen Konstruktionen laut Proposition 2.3.1 verwendet.

Beispiel 2.6.1 (Huas Identität [Ami66])**.** Es gilt:

$$x - \left(x^{-1} + (y^{-1} - x)^{-1}\right)^{-1} = xyx. \tag{2.6.2}$$

Beweis. Minimale lineare zulässige Systeme für y^{-1} und x sind

$$\begin{bmatrix} y \end{bmatrix} s = \begin{bmatrix} 1 \end{bmatrix} \quad \text{bzw.} \quad \begin{bmatrix} 1 & -x \\ . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 1 \end{bmatrix}.$$

Das ZLS für die Differenz $y^{-1} - x$,

$$\begin{bmatrix} y & -y & . \\ . & 1 & -x \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} 1 \\ . \\ -1 \end{bmatrix}, \quad s = \begin{bmatrix} y^{-1} - x \\ -x \\ -1 \end{bmatrix}, \quad t = \begin{bmatrix} y^{-1} & -1 & y^{-1} - x \end{bmatrix}$$

ist minimal, weil sowohl die linke Familie s als auch die rechte Familie t $\mathbb{K}$-linear unabhängig sind (Proposition 2.1.8). Klarerweise gilt $1 \in R(y^{-1} - x)$. Daher gibt es laut Lemma 2.3.5 eine zulässige Transformation

$$(P, Q) = \left(\begin{bmatrix} . & 1 & . \\ 1 & . & 1 \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . \\ 1 & 1 & . \\ . & . & 1 \end{bmatrix} \right),$$

die man einfach findet, indem man die Spalte 2 zu Spalte 1 addiert, die beiden ersten Zeilen vertauscht und dann Zeile 3 zu (der neuen) Zeile 2 addiert. Wir bekommen also

$$\begin{bmatrix} 1 & 1 & -x \\ . & -y & 1 \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ -1 \end{bmatrix}$$

und können (erneut) die Inverse vom Typ $(1, 1)$ anwenden:

$$\begin{bmatrix} 1 & y \\ -x & -1 \end{bmatrix} s = \begin{bmatrix} . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} (y^{-1} - x)^{-1} \\ -(1 - xy)^{-1} \end{bmatrix}.$$

Dieses System repräsentiert ein reguläres Element $(y^{-1} - x)^{-1} = (1 - yx)^{-1}y$, kann also in ein reguläres ZLS (Definition 2.1.12) transformiert werden, indem man die zweite Zeile mit -1 multipliziert. Danach addieren wir x^{-1} „von links“:

$$\begin{bmatrix} x & -x & . \\ . & 1 & y \\ . & x & 1 \end{bmatrix} s = \begin{bmatrix} 1 \\ . \\ -1 \end{bmatrix}, \quad s = \begin{bmatrix} x^{-1} + (y^{-1} - x)^{-1} \\ (y^{-1} - x)^{-1} \\ -(1 - xy)^{-1} \end{bmatrix}.$$

Dieses System ist minimal und —nachdem wir Zeile 3 zu Zeile 1 addiert haben (um den oberen Nicht-Null-Eintrag in der rechten Seite zu eliminieren)— wenden wir die (minimale) Inverse vom Typ $(0, 0)$ an:

$$\begin{bmatrix} -1 & -1 & -x & . \\ . & -y & -1 & . \\ . & -1 & 0 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}. \tag{2.6.3}$$

Nun multiplizieren wir Zeile 1 und die Spalten 2 und 3 mit -1 und vertauschen die Spalten 2 und 3 um folgendes System zu erhalten:

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & y & . \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} x - xyx \\ -yx \\ x \\ 1 \end{bmatrix}.$$

Der nächste Schritt wäre eine Skalierung mit -1 und der Addition von x (laut Proposition 2.3.1). Mit zwei Minimierungsschritten erreicht man wieder Minimalität. Alternativ kann die Addition eines *linearen* Terms zu einem Polynom (einem polynomiellen

ZLS) —in Abhängigkeit des Eintrags v_n der rechten Seite— direkt im rechten oberen Eintrag der Systemmatrix ausgeführt werden:

$$\begin{bmatrix} 1 & -x & -1 & x \\ . & 1 & y & . \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} -xyx \\ -yx \\ x \\ 1 \end{bmatrix}.$$

Die systematische Minimierung von polynomiellen zulässigen linearen Systemen ist in Abschnitt 4.3 beschrieben. Für eine allgemeine Minimierung ist etwas mehr an Vorbereitung notwendig. Darum geht es dann in Kapitel 4. □

Bemerkungen. Die Umformung vom ZLS (2.6.3) ist ein einfacher Fall der *Verfeinerung* eines Pivotblocks und wird im Abschnitt 4.4 im Detail diskutiert. Huas Identität kommt auch in [CR94] als Beispiel vor. Es lohnt sich, beide Vorgehensweisen miteinander zu vergleichen.

Intermezzo

Spätestens hier sollte man eine Pause einlegen und sich den Fragen widmen, die in der Zwischenzeit aufgetaucht sind. Von einem linearen Weiterlesen rate ich ausdrücklich ab, weil die Fülle an Details in den beiden folgenden Kapiteln sehr groß ist. Eine Möglichkeit, den Überblick nicht zu verlieren, wäre, mit dem Beispiel 3.3.1 (Eigenwertberechnung) zu beginnen und dann die beiden Abschnitte 3.2 (Polynommultiplikation) und 3.3 (Polynomfaktorisierung) durchzugehen, bevor man sich in Abschnitt 4.3 (Minimieren eines polynomiellen ZLS) vertieft. Das wäre sozusagen eine Minimalvariante, in der die wesentlichen Konzepte vorkommen.

Bevor man dann tiefer in die folgenden Kapitel 3 (Faktorisieren) und 4 (Minimieren) eintaucht, empfiehlt es sich, noch einmal Kapitel 2 (Rechnen), insbesondere dessen Grundlagen, zu wiederholen. Die Abschnitte 3.4 (Faktorisierungstheorie), 3.5 (Faktormultiplikation) und 3.6 (Allgemeine Faktorisierung) sollte man sich eher für den Schluss aufsparen. Diese sind weniger für die Minimierung relevant —dafür genügt der Abschnitt 4.4 (Verfeinern)— als für ein tieferes Verständnis linearer Darstellungen (von Elementen des freien Schiefkörpers).

Abgesehen vom etwas höheren Aufwand —schließlich benötigt man jetzt eine Matrix (und einen Vektor) statt eines einfachen Bruches— sollte nun die Überzeugung eingetreten sein, dass man (fast) ganz normal mit Elementen des freien Schiefkörpers rechnen kann. Der Illusion, zu *verstehen*, was denn dieser Körper nun tatsächlich wäre, sollte man sich allerdings nicht hingeben, wenn man noch nicht damit begonnen hat, in der (zitierten) Literatur zumindest etwas zu stöbern.

Da es so viele Anknüpfungspunkte zu verschiedenen Bereichen der Mathematik gibt, empfiehlt es sich, seinen eigenen Weg zu finden. Vielleicht hilft der folgende Auszug eines Artikels von C. Reutenauer dabei:

„Der Ring der rationalen Reihen (mit seinem nicht-kommutativen Produkt) ist per Definition der Ring, der die nicht-kommutativen Polynome und die Inverse jeder Reihe mit nicht-verschwindendem konstanten Term enthält. Jedoch ist er kein Schiefkörper; zum Beispiel haben die Variablen kein Inverses in diesem Ring. Aber es existieren Körper, die die nicht-kommutativen Polynome enthalten. Unter diesen gibt es einen kanonischen, und dieser hat die ‚geringst möglichen Relationen'. Er wurde von P. M. Cohn konstruiert und heißt der freie Schiefkörper. *Er ist nicht einfach zu*

beschreiben.[1] [...] Eine natürliche Frage taucht auf: Tatsächlich sind rationale Reihen im freien Schiefkörper enthalten; umgekehrt, wenn eine Reihe —im üblichen Sinn— im freien Schiefkörper enthalten ist, ist sie rational? Das wurde positiv von Fliess beantwortet [Fli70], der damit die Vorstellung klargestellt hat." [Reu08, Abschnitt 9]

Selbst, wenn man im Studium viel mit Algebra (z.B. für Zahlentheorie) zu tun hatte, braucht es seine Zeit, um sich in der nicht-kommutativen Welt[2] zurechtzufinden. Ein erster Test ist $(a+b)^2$. Ohne weitere Information —abgesehen der, dass die Addition üblicherweise kommutativ ist— sollte man das als $a^2+ab+ba+b^2$ auflösen. Tatsächlich gibt es eine ganze Reihe an Konzepten aus dem Kommutativen, die sich verallgemeinern lassen. Aber diese sind selten offensichtlich, wie man sich selbst — nur um ein paar zu nennen— in [Ore31] (Lineare Gleichungen), [Tay73] (Funktionen) oder [GGRW05] (Determinanten) überzeugen kann.

Etwas anderes, das den freien Schiefkörper eher schwer zugänglich macht, ist das Auftauchen von Phänomenen und Konzepten, die zunächst ungewohnt sind. Dazu gehört das identische Verschwinden von bestimmten Polynomen, selbst dann, wenn man Matrizen einsetzt [AL50], zum Beispiel um Ausdrücken wie $(xy-yx)^{-1}$ einen Sinn zu geben. Oder das der Inversionshöhe [Reu96a, HS07, HS15]. Schnell entstehen viele weitere Fragen und so ist es kaum möglich, sich nicht zu verlieren.

Ein neues Phänomen betrifft die Möglichkeit, dass zwei (verallgemeinerte) Atome (irreduzible Elemente) zu einem *neuen* Atom „verschmelzen" können. So ist z.B. $f=(1-xy)(1-zy)^{-1}$ irreduzibel, das heißt, dass die beiden ursprünglichen „Faktoren" nicht mehr aus einem *minimalen* ZLS für f rekonstruiert werden können. *Wie* soll oder kann man sich das vorstellen? Im Ring der ganzen Zahlen wäre „$2\cdot 3=5$" völliger Unsinn. Natürlich müsste man hier etwas präziser (mit der Sprache) sein, schließlich muss man ja zwischen Element und Darstellung unterscheiden.[3] Trotzdem darf man den Zeitaufwand nicht unterschätzen, der anfällt, wollte man das genauer untersuchen. Außerdem lässt sich nicht abschätzen, welche weiteren Überraschungen die nicht-kommutative Welt in sich birgt. Das ist auch der Grund weshalb man sich insbesondere für die (allgemeine) Faktorisierungstheorie viel Zeit nehmen sollte.

Um die allgemeine Theorie der (nicht-kommutativen) *Divisionsringe* im Detail nachvollziehen zu können, braucht man entsprechende Grundlagen. So liest man in der Einleitung des Kapitels über Divisionsringe in Lams „zweiten" Buch über nicht-kommutative Algebra auf Seite 287 [Lam99]: „Nach der Entwicklung von genug Modultheorie[4] in den letzten drei Kapiteln, ist nun die Bühne frei, um die Theorie von Divisionsringen zu studieren. Das aktuelle Kapitel ist eine allgemeine Einleitung in

[1] Die Hervorhebung stammt von mir.

[2] Eigentlich befindet man sich in der Rubrik „Assoziative Algebra".

[3] Das scheinbare Paradoxon hier löst sich auf, wenn man sich von rationalen Ausdrücken als „Darstellung" löst. Im Abschnitt 2.6 (Rationale Identitäten) ist uns bereits der Ausdruck $(y^{-1}-x)^{-1}$ untergekommen, dessen Regularität erst im äquivalenten $(1-yx)^{-1}y$ sichtbar wird. Dagegen ist die Invertierbarkeit des konstanten Teils der Systemmatrix eines minimalen ZLS sofort erkennbar.

[4] *Moduln* sind eine Verallgemeinerung von Vektorräumen, daher ist auch der (Euklidsche) dreidimensionale Raum $\mathbb{R}^3$ ein *Modul*. Viel von dem, was in Vektorräumen selbstverständlich ist, gilt in Moduln nicht automatisch.

diese Theorie, im Rahmen nicht-kommutativer Ringe.“ Grob werden dann drei Fälle unterschieden: Der „Gute“, der „Schlechte“ und der „Hässliche“[5]. In letzteren fällt die freie assoziative Algebra (für ein Alphabet mit mindestens zwei Buchstaben). Weiters schreibt Lam: „Es ist keineswegs klar, dass $\mathbb{K}\langle X\rangle$ in einen Divisionsring eingebettet werden kann.“

Nun mag zwar die Einbettung (des Rings) der nicht-kommutativen Polynome in einen Divisionsring (beziehungsweise Schiefkörper) etwas „sperriger“ zu formulieren sein, die *Eigenschaften* der freien assoziativen Algebra sind alles andere als hässlich. So hat sie zum Beispiel einen *distributiven Faktorverband* [Coh82b] bzw. [Coh85, Abschnitt 3.5], das ist der „Verband“ aller Teiler eines (vorgegebenen) Polynoms. Am schönsten sieht man das am Beispiel $x(1-yx) = x - xyx = (1-xy)x$, das uns im nächsten Kapitel in unterschiedlichen Varianten begegnen wird. Es kann im allgemeinen viele „verschiedene Wege“ geben, ein Polynom (in irreduzible Elemente) zu faktorisieren, aber *alle* „Wege“ sind gleich lang! Auch das ist etwas, das man von den ganzen Zahlen her bereits kennt: $2 \cdot 3 = 6 = 3 \cdot 2$. Richard Dedekind hat sich darüber vor über hundert Jahren den Kopf zerbrochen ...

[5] Im Original heißt es „The Ugly“, das man in diesem Zusammenhang vielleicht auch mit der „Sperrige“ oder der „Vertrackte“ übersetzen könnte.

Kapitel 3

Faktorisieren

Nachdem die gesamte Faktorisierungstheorie einem „kleinen“ Problem aus der Minimierung von linearen Darstellungen entsprungen ist, soll dieses auch als Roter Faden durch das Kapitel führen. In gewisser Weise hat sich diese Theorie verselbständigt und ist mittlerweile auch rein algebraisch interessant, ermöglicht sie doch, den freien Schiefkörper als „Ring“ zu interpretieren. Nicht in der trivialen Sicht, wonach jeder Körper ein Ring ist, sondern in der reichhaltigen „Struktur“, die man erhält, wenn man nicht-kommutative Faktorisierungstheorie und die Theorie der Einbettung von freien Idealringen[1] in ihren jeweiligen universellen Quotientenkörper miteinander verknüpft. Die Fülle an Fragestellungen ist enorm, weil man nun über verallgemeinerte Monome, Links- und Rechtsideale, mögliche Invarianten in Bezug auf die Faktorisierung, etc. nachdenken kann.

Um den Faden nicht zu verlieren, kehren wir zurück zu einem einfachen Beispiel: Angenommen, wir haben ein Element f durch das zulässige lineare System $\mathcal{A}_f$,

$$\begin{bmatrix} x & 1 & . \\ . & y & -1 \\ . & -1 & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}$$

gegeben. Mittels Proposition 2.3.6 konstruieren wir für fx das System $\mathcal{A}$,

$$\begin{bmatrix} x & 1 & . & . \\ . & y & -1 & . \\ . & -1 & x & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Ist $\mathcal{A}$ minimal? Nun wiederholen wir diesen Schritt für f gegeben durch ein anderes

[1] Einen freien Idealring (FIR) kann man sich als „verallgemeinerte“ freie assoziative Algebra vorstellen. Details findet man in [Coh85, Abschnitt 1.2].

System $\mathcal{A}'_f$ und konstruieren wiederum ein System $\mathcal{A}'$ für fx, nämlich

$$\begin{bmatrix} x & 1 & . & . \\ 1 & y & -1 & . \\ . & . & x & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix},$$

in dem man $\mathcal{A}'_f$ direkt ablesen kann. Hier sieht man sofort, dass Zeile 3 und Spalte 3 entfernt werden können, nachdem man Spalte 3 zu Spalte 4 addiert hat. Also kann $\mathcal{A}'$ und damit auch $\mathcal{A}$ nicht minimal sein. Was das mit der Faktorisierung zu tun hat, sieht man dann viel deutlicher im Beispiel 3.2.1, sobald man sich mittels der minimalen Inversen vergewissert hat, dass $f = (pq)^{-1}$ für $p = x$ und $q = 1 - yx$ ist.

Der 2×1 Nullblock (links unten) in der Systemmatrix von $\mathcal{A}_f$ wird zu einem Nullblock rechts oben in der Systemmatrix von $\mathcal{A}_f^{-1}$, der Standardinversen von $\mathcal{A}_f$,

$$\begin{bmatrix} 1 & -x & 1 & 0 \\ . & 1 & -y & 0 \\ . & . & -1 & -x \\ . & . & . & -1 \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ . \\ 1 \end{bmatrix},$$

die hier minimal ist, weil f vom Typ $(0, 0)$ ist. Und dieser Nullblock rechts oben ist der, der bei der Multiplikation $(*, 1)$ oder $(1, *)$ entsteht, vergleiche Propositionen 2.3.6 bzw. 2.3.9. Damit ergibt sich eine „natürliche“ Korrespondenz zwischen Faktorisierungen und rechter oberer Nullblock-Struktur in der Systemmatrix (vorausgesetzt die entsprechenden Einträge in der rechten Seite sind Null).

Mit anderen Worten: Man kann (nicht-triviale) Faktoren in einem Polynom finden, indem man nach „passenden“ Transformationen (eines *minimalen* ZLS) sucht. Darum geht es in Abschnitt 3.3. Faktorisiert man auf diese Art ein Polynom in zwei (nicht notwendigerweise irreduzible) Faktoren, sind „deren“ zulässige lineare Systeme *notwendigerweise* minimal. Umgekehrt, und das ist der Kern von Abschnitt 3.2, gilt das auch. So „offensichtlich“ die minimale Polynommultiplikation auch sein mag, der Beweis ist *hochgradig* nicht-trivial. (Was vermutlich daran liegt, dass außer der Minimalität keine weiteren Annahmen, wie z.B. Invertierbarkeit der Systemmatrix über den formalen Potenzreihen, notwendig sind.)

Als praktischer Zugang bietet sich Beispiel 3.3.1 an, in dem die (nicht-kommutative) Polynomfaktorisierung verwendet wird, um die Eigenwerte einer Matrix über das charakteristische Polynom auszurechnen.

Die allgemeine Faktorisierungstheorie in Abschnitt 3.4 ist sicherlich schwerer zugänglich. Zwar bekommt man durch die Polynome schnell eine Idee, wie es sein sollte, aber die „Stolpersteine“ bis hin zum Satz 3.4.9 sind zahlreich. Dafür kann man dann die freie assoziative Algebra „vergessen“ und Elemente direkt im freien Schiefkörper faktorisieren. Und auch hier gibt es wieder zwei Sichten, nämlich die der (minimalen) Multiplikation in Abschnitt 3.5 und die der (eigentlichen) Faktorisierung über

das Aufspüren von Nullblöcken in Abschnitt 3.6. Darüber hinaus ist das (multiplikative) „Verschmelzen“ zweier (verallgemeinerter) Atome zu *einem* (neuen) Atom etwas gewöhnungsbedürftig. Das dem zugrundeliegende Phänomen müsste jedenfalls genauer untersucht werden, weil das die (algorithmische) Minimierung (ohne „Verfeinerung“[2]) schwierig macht.

Dieses Kapitel schließt mit Abschnitt 3.7, in dem anhand zweier Beispiele detailliert die einzelnen Schritte für die Faktorisierung erläutert werden. In Beispiel 3.7.1 geht es um die Faktorisierung eines Polynoms und in Beispiel 3.7.6 um die (allgemeine) Faktorisierung eines regulären Elementes. Darüber hinaus gibt es Anmerkungen zur grundsätzlichen Vorgehensweise.

3.1 Grundlagen

Die folgenden Definitionen folgen hauptsächlich [BS15], sind aber für unsere Zwecke angepasst. Wir brauchen hier nicht die volle Allgemeinheit. Während es im Kommutativen eine ziemlich einheitliche Faktorisierungstheorie gibt [GHK06, Abschnitt 1.1], ist der „einfachste“ nicht-kommutative Fall, der eines „eindeutigen“ *Faktorisierungsbereiches* (UFD) wie der *freien assoziativen Algebra*, nicht so klar. Für einen allgemeinen (algebraischen) Blickwinkel empfiehlt sich die Zusammenfassung [Sme15]. Die Faktorisierung in *freien Idealringen* (FIRs) wird ausführlich in [Coh85, Kapitel 3] diskutiert. Freie Idealringe spielen eine wichtige Rolle bei der Konstruktion von freien Schiefkörpern. Mehr zur „nicht-kommutativen“ Faktorisierung findet man in [Jor89] und [BHL17] (nur um zwei zu nennen) und der dort erwähnten Literatur. Zur Abgrenzung der Faktorisierung in freien assoziativen Algebren siehe auch Abschnitt B.4.

Definition 3.1.1 (Ähnliche Rechtsideale, ähnliche Elemente [Coh85, Abschnitt 3.2])**.** Sei R ein Ring. Zwei Rechtsideale $\mathfrak{a}, \mathfrak{b} \subseteq R$ heißen *ähnlich*, geschrieben als $\mathfrak{a} \sim \mathfrak{b}$, wenn $R/\mathfrak{a} \cong R/\mathfrak{b}$ als rechte R-Moduln ist. Zwei Elemente $p, q \in R$ heißen *ähnlich*, wenn ihre Rechtsideale pR und qR ähnlich sind, das heißt, $pR \sim qR$. Siehe auch [Sme15, Abschnitt 4.1].

Definition 3.1.2 (Links-/Rechtsteilbarkeit, koprime Elemente [BS15, Abschn. 2])**.** Sei R ein Integritätsbereich und $H = R^\bullet = R\backslash\{0\}$. Ein Element $p \in H$ *teilt* $q \in H$ *von links*, geschrieben als $p \,|_\mathrm{l}\, q$, wenn $q \in pH = \{ph \mid h \in H\}$ ist. Zwei Elemente p, q heißen *links koprim* wenn für alle h mit der Eigenschaft $h \,|_\mathrm{l}\, p$ und $h \,|_\mathrm{l}\, q$ bereits $h \in H^\times = \{f \in H \mid f \text{ ist invertierbar}\}$ ist, das heißt, h ist ein Element der *Einheitengruppe*. Teilbarkeit von rechts $p \,|_\mathrm{r}\, q$ und die Notation von *rechts komprim* wird in ähnlicher Weise definiert. Zwei Elemente heißen *koprim* wenn sie links und rechts koprim sind.

Definition 3.1.3 (Atomare Integritätsbereiche, irreduzible Elemente [BS15, Abschnitt 2])**.** Sei R ein Integritätsbereich und $H = R^\bullet$. Ein Element $p \in H \setminus H^\times$,

[2] Der Abschnitt 4.4 (Verfeinern von Pivotblöcken) im folgenden Kapitel 4 (Minimieren) würde genau so gut hierher passen. Allerdings wäre das für jene, die hauptsächlich an der Minimierung interessiert sind, wohl nicht ganz so übersichtlich.

das heißt, eine nicht-null Nicht-Einheit (in R), heißt *Atom* (oder *irreduzibel*) wenn $p = q_1q_2$ mit $q_1, q_2 \in H$ impliziert, dass entweder $q_1 \in H^\times$ oder $q_2 \in H^\times$ ist. Die Menge der Atome in R wird mit $\mathbf{A}(R)$ bezeichnet. Das (kürzbare) Monoid H heißt *atomar*, wenn jede Nicht-Einheit als endliches Produkt von Atomen in H geschrieben werden kann. Der Integritätsbereich R heißt *atomar*, wenn das Monoid $R^\bullet$ atomar ist.

Definition 3.1.4 (Ähnlichkeits-eindeutige Faktorisierungsbereiche [Sme15, Definition 4.1])**.** Ein Bereich R heißt *ähnlichkeits-faktoriell* (oder ein *Ähnlichkeits-UFD*[3]), wenn R atomar ist und die folgende Bedingung erfüllt: Wenn $p_1p_2\cdots p_m = q_1q_2\cdots q_n$ für Atome (irreduzible Elemente) $p_i, q_j \in R$ gilt, dann ist $m = n$ und es existiert eine Permutation $\sigma \in \mathfrak{S}_m$ sodass p_i ähnlich zu $q_{\sigma(i)}$ für alle $i \in \{1, 2, \ldots, m\}$ ist.

Bemerkung. Ähnlichkeit zweier Elemente a, a' in einem *schwachen Bezout-Ring* R ist äquivalent zur Existenz von $b, b' \in R$ sodass $ab' = ba'$ mit ab' und ba' koprim ist, das heißt, a und b sind links koprim und b' und a' sind rechts koprim. Die freie assoziative Algebra $\mathbb{K}\langle X\rangle$ ist ein schwacher Bezout-Ring [Coh63, Proposition 5.3 und Satz 6.2].

Beispiel. Die Polynome $p = 1 - xy$ und $q = 1 - yx$ sind ähnlich, weil $px = (1 - xy)x = x - xyx = x(1 - yx) = xq$ ist. Siehe auch Beispiel 3.2.1.

Beispiel. Im *freien Monoid* X^* sind die Atome genau die Buchstaben x_i im Alphabet X.

Proposition 3.1.5 ([Coh63, Satz 6.3])**.** *Die freie assoziative Algebra $\mathbb{K}\langle X\rangle$ ist ein Ähnlichkeits-UFD.*

Bemerkung 3.1.6. In Abschnitt 3.4 wird die *Teilbarkeit* und die *Irreduzibilität* mittels der Definitionen 3.4.4 bzw. 3.4.10 auf den freien Schiefkörper übertragen. Um auch die „Ähnlichkeit“ sinnvoll verallgemeinern zu können, ist noch einiges zu tun. Als Zwischenschritt könnte man die *Ähnlichkeit* im Sinne der vorherigen Bemerkung (über schwache Bezout-Ringe) auch im freien Schiefkörper definieren, weil *Koprimheit* bereits zur Verfügung steht. Gegeben zwei Elemente $f, g \in \mathbb{F}$, wie findet man gegebenenfalls $h_1, h_2 \in \mathbb{F}$ sodass $fh_1 = h_2g$ ist, f und h_2 links koprim und h_1 und g rechts koprim sind? Ein Spezialfall davon ist das *Konjugationsproblem*: Existiert ein $0 \neq h \in \mathbb{F}$, sodass $fh = hg$ beziehungsweise $f = hgh^{-1}$ ist?

3.2 Minimale Polynommultiplikation

Beispiel 3.2.1 ([Sch17c, Beispiel 2.7])**.** Die Polynome $p = x \in \mathbb{K}\langle X\rangle$ und $q = 1 - yx \in \mathbb{K}\langle X\rangle$ können durch die minimalen polynomiellen zulässigen linearen Systeme

$$\mathcal{A}_p = \left(1, \begin{bmatrix} 1 & -x \\ . & 1 \end{bmatrix}, 1\right) \quad \text{bzw.} \quad \mathcal{A}_q = \left(1, \begin{bmatrix} 1 & y & -1 \\ . & 1 & -x \\ . & . & 1 \end{bmatrix}, 1\right)$$

[3] Wir verwenden in der Folge diesen weniger „sperrigen“ Begriff, auch wenn das Akronym vom Englischen „unique factorization domain“ kommt.

dargestellt werden. Ein polynomielles ZLS für $pq = x(1-yx)$ ist gegeben durch

$$\begin{bmatrix} 1 & -x & . & . & . \\ . & 1 & -1 & . & . \\ . & . & 1 & y & -1 \\ . & . & . & 1 & -x \\ . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} x(1-yx) \\ 1-yx \\ 1-yx \\ x \\ 1 \end{bmatrix}.$$

Addiert man Spalte 2 zu Spalte 3 (und subtrahiert s_3 von s_2) bekommt man

$$\begin{bmatrix} 1 & -x & -x & . & . \\ . & 1 & 0 & . & . \\ . & . & 1 & y & -1 \\ . & . & . & 1 & -x \\ . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} x(1-yx) \\ 0 \\ 1-yx \\ x \\ 1 \end{bmatrix},$$

also das polynomielle ZLS

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & 0 & 0 \\ . & 1 & y & -1 \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right). \tag{3.2.2}$$

Nachdem auch die rechte Familie $t = [1, x, -xy, x(1-yx)]$ $\mathbb{K}$-linear unabhängig ist, ist dieses System laut Proposition 2.1.8 (Minimalitätscharakterisierung) minimal. Man beachte den rechten oberen 1×2 Nullblock in der Systemmatrix von $\mathcal{A}$.

Um zu zeigen, dass die Konstruktion in Proposition 3.2.7 zu einer *minimalen Darstellung* (für das Produkt zweier Polynome) führt, ist einige Vorbereitung notwendig. Eines der wesentlichen Werkzeuge ist (wieder einmal) Lemma 2.1.3 [Coh95, Korollar 6.3.6]. Obwohl wir hier nur mit regulären Elementen arbeiten, muss die Invertierbarkeit der konstanten Koeffizientenmatrix A_0 (in der Systemmatrix) in Lemma 2.3.2 nicht angenommen werden. Zunächst ist nicht klar, ob es nicht auch einen viel einfacheren Beweis geben könnte. Erst später, nach der allgemeinen Faktorisierungstheorie und der etwas allgemeineren (minimalen) Multiplikation zeigt sich, dass hier viel mehr dahintersteckt, nämlich dass polynomielle Faktoren immer *verallgemeinerte* Faktoren (siehe Definition 3.4.1) sind.

Mit anderen Worten: Beim Übergang vom Ring (der freien assoziativen Algebra) zu seinem Quotientenkörper (dem freien Schiefkörper) muss man —gegeben ein Element— *fast alle* Elemente, das heißt, alle bis auf endlich viele, als Faktoren ausschließen. Dabei ist es keineswegs selbstverständlich, dass ein irreduzibles Polynom keine weiteren echten Faktoren (im freien Schiefkörper) hat. Darum geht es dann in Abschnitt 3.4.

Lemma 3.2.3 ist eine etwas verallgemeinerte Version von [Sch17c, Lemma 2.4]. Der Beweis selbst muss kaum angepasst werden. Der Beweis der (folgenden) Proposition 3.2.7 wird wesentlich einfacher über die Lemmata 3.2.5 und 3.2.6, die aus

dem Originalbeweis (der Polynommultiplikation) extrahiert wurden. Sie sind später nützlich, besonders in Lemma 3.4.7.

Bemerkung. Die Transformation im folgenden Lemma ist *nicht* notwendigerweise zulässig. Aber, außer für $n = 2$ (die über eine Permutation der letzten beiden Elemente in der linken Familie, das heißt, eine Permutation der letzten beiden Spalten in der Systemmatrix, zulässig wird) kann sie so gewählt werden, dass sie *zulässig* ist.

Lemma 3.2.3 ([Sch17a, Lemma 2.14])**.** *Sei $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ ein ZLS der Dimension $n \geq 2$ und $\mathbb{K}$-linear abhängiger linker Familie $s = A^{-1}v$. Sei $m \in \{2, 3, \ldots, n\}$ der minimale Index sodass die linke Teilfamilie $\underline{s} = (A^{-1}v)_{i=m}^{n}$ $\mathbb{K}$-linear unabhängig ist. Für $A = (a_{ij})$ sei $a_{ii} = 1$ für $1 \leq i \leq m$ und $a_{ij} = 0$ für $j < i \leq m$ (oberer $m \times m$ Dreiecksblock) und $a_{ij} = 0$ für $j \leq m < i$ (Nullblock der Größe $(n - m) \times m$ links unten). Dann existieren Matrizen $T, U \in \mathbb{K}^{1\times(n+1-m)}$ sodass*

$$U + (a_{m-1,j})_{j=m}^{n} - T(a_{ij})_{i,j=m}^{n} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix} \quad \text{und} \quad T(v_i)_{i=m}^{n} = 0 \quad \text{ist.}$$

Beweis. Laut Annahme ist die linke Teilfamilie $(s_{m-1}, s_m, \ldots, s_n)$ $\mathbb{K}$-linear abhängig. Also gibt es $\kappa_m, \ldots, \kappa_n \in \mathbb{K}$, sodass $s_{m-1} = \kappa_m s_m + \kappa_{m+1}s_{m+1} + \ldots + \kappa_n s_n$ ist. Sei $U = [\kappa_m, \kappa_{m+1}, \ldots, \kappa_n]$. Dann ist $s_{m-1} - U\underline{s} = 0$. Laut Annahme ist $v_{m-1} = 0$. Nun können wir Lemma 2.3.2 mit $B = U + [a_{m-1,m}, a_{m-1,m+1}, \ldots, a_{m-1,n}]$ und $\underline{s}$ anwenden. Daher gibt es eine Matrix $T \in \mathbb{K}^{1\times(n+1-m)}$, die folgende Gleichung erfüllt:

$$U + \begin{bmatrix} a_{m-1,m} & \ldots & a_{m-1,n} \end{bmatrix} - T \begin{bmatrix} a_{m,m} & \ldots & a_{m,n} \\ \vdots & \ddots & \vdots \\ a_{n,m} & \ldots & a_{n,n} \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix}. \tag{3.2.4}$$

Da die letzte Spalte von T Null ist, gilt auch $T(v_i)_{i=m}^{n} = 0$. □

Lemma 3.2.5 ([Sch17a, Lemma 2.15])**.** *Seien $p \in \mathbb{K}\langle X\rangle \setminus \mathbb{K}$ und $g \in \mathbb{F} \setminus \mathbb{K}$ durch die* minimalen *zulässigen linearen Systeme $A_p = (u_p, A_p, v_p)$ der Dimension n_p beziehungsweise $A_g = (u_g, A_g, v_g)$ der Dimension n_g mit $1 \in R(g)$ gegeben. Dann ist die linke Familie des ZLS $\mathcal{A} = (u, A, v)$ laut Proposition 2.3.9 für pg der Dimension $n = n_p + n_g - 1$ $\mathbb{K}$-linear unabhängig.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $v = [0, \ldots, 0, 1]^\top$, $\mathcal{A}_p$ in polynomieller Form und $[1, 0, \ldots, 0]^\top$ die erste Spalte von A_g. Seien $s_p = (s_1^p, \ldots, s_{n_p}^p)$ und $s_g = (s_1^g, \ldots, s_{n_g}^g)$ die jeweiligen linken Familien von $\mathcal{A}_p$ und $\mathcal{A}_g$. Wir müssen zeigen, dass die linke Familie

$$s = (s_1, s_2, \ldots, s_n) = (s_1^p g, \ldots, s_{n_p-1}^p g, g, s_2^g, \ldots, s_{n_g}^g).$$

von $\mathcal{A}$ $\mathbb{K}$-linear unabhängig ist. Wir nehmen gegenteilig an, dass es einen Index $1 < m \leq n_p$ gibt, sodass $(s_{m-1}, s_m, \ldots, s_n)$ $\mathbb{K}$-linear abhängig ist, während $(s_m, \ldots, s_n)$ $\mathbb{K}$-linear unabhängig ist. Dann existieren laut Lemma 3.2.3 Matrizen $T, U \in \mathbb{K}^{1\times(n-m+1)}$ als Blöcke in (invertierbaren) Matrizen $P, Q \in \mathbb{K}^{n\times n}$,

$$P = \begin{bmatrix} I_{m-2} & . & . \\ . & 1 & T \\ . & . & I_{n-m+1} \end{bmatrix} \quad \text{und} \quad Q = \begin{bmatrix} I_{m-2} & . & . \\ . & 1 & U \\ . & . & I_{n-m+1} \end{bmatrix},$$

die die Gleichung $s_{m-1} = 0$ (in Zeile $m-1$) in $P\mathcal{A}Q$ erzeugen. (Diese „potentielle“ Transformation (P, Q) ist nicht notwendigerweise zulässig. Aber das spielt hier keine Rolle, weil wir nur die lineare Unabhängig der linken Familie zeigen.) Sei $\tilde{P}$ (bzw. $\tilde{Q}$) der obere linke Teil von P (bzw. Q) der Größe $n_g \times n_g$. Dann ist $s^p_{m-1} = \alpha \in \mathbb{K}$ die Gleichung in Zeile $m-1$ in $\tilde{P}\mathcal{A}_p\tilde{Q}$ und, da $s^p_{n_p} = \lambda \in \mathbb{K}$ ist, würde das der $\mathbb{K}$-linearen Unabhängigkeit der linken Familie von $\mathcal{A}_p$ widersprechen. Also gibt es keinen solchen Index $1 < m \leq n_p$ und damit ist $s = (s_1, s_2, \ldots, s_n)$ $\mathbb{K}$-linear unabhängig. □

Bemerkung. Über die Minimalität von $\mathcal{A}$ kann nichts gesagt werden, da die rechte Familie $t = uA^{-1}$ nicht notwendigerweise $\mathbb{K}$-linear unabhängig ist. Als Beispiel nehmen wir $p = xy$ und $g = y^{-1} + z$. Ein, laut Proposition 2.3.9 konstruiertes, ZLS für pg ist

$$\begin{bmatrix} 1 & -x & . & . & . \\ . & 1 & -y & . & . \\ . & . & 1 & 1 & -z \\ . & . & . & y & 1 \\ . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix}.$$

Dessen rechte Familie ist $t = [1, x, xy, x, x + xyz]$.

Lemma 3.2.6. *Seien $f \in \mathbb{F} \setminus \mathbb{K}$ und $q \in \mathbb{K}\langle X\rangle \setminus \mathbb{K}$ gegeben durch die* minimalen *zulässigen linearen Systeme $A_f = (u_f, A_f, v_f)$ der Dimension n_f mit $1 \in L(f)$ beziehungsweise $A_q = (u_q, A_q, v_q)$ der Dimension n_q. Dann ist die rechte Familie des ZLS $\mathcal{A} = (u, A, v)$ laut Proposition 2.3.6 für fq der Dimension $n = n_f + n_q - 1$ $\mathbb{K}$-linear unabhängig.*

Proposition 3.2.7 (Minimale Polynommultiplikation [Sch17c, Proposition 2.6])**.** *Seien $p, q \in \mathbb{K}\langle X\rangle$ gegeben durch die* minimalen *polynomiellen zulässigen linearen Systeme $A_p = (1, A_p, \lambda_p)$ beziehungsweise $A_q = (1, A_q, \lambda_q)$ der Dimensionen $n_p, n_q \geq 2$. Dann ist das ZLS $\mathcal{A}$ laut Proposition 2.3.6 für pq* minimal *(und polynomiell) mit der Dimension $n = n_p + n_q - 1$.*

Beweis. Die Fälle $p \in \mathbb{K}$ oder $q \in \mathbb{K}$ sind trivial. Die linke Familie von $\mathcal{A}$ ist $\mathbb{K}$-linear unabhängig laut Lemma 3.2.5 und die rechte Familie ist $\mathbb{K}$-linear unabhängig laut Lemma 3.2.6. Proposition 2.1.8 liefert die Minimalität. Durch die Konstruktion ist $\mathcal{A}$ ein polynomielles ZLS. □

3.3 Polynomfaktorisierung

Das hier entwickelte Faktorisierungskonzept basiert auf *minimalen* polynomiellen zulässigen linearen Systemen. Ein paar Möglichkeiten, wie man solche direkt erhält, haben wir in Abschnitt 2.2 (und für die Multiplikation im vorherigen Abschnitt) kennengelernt. Wie man diese (für Polynome) im allgemeinen konstruiert, wird in Abschnitt 4.3 beschrieben.

Bemerkung. Obwohl wir hier (allgemeine) zulässige lineare Systeme verwenden, mit einer Beschränkung der Anwendung rationaler Operationen mit Ausnahme der Inversen in Proposition 2.3.1 auf ausschließlich (Systeme für) Polynome, erhält man wieder Polynome. Wenn die Inverse beschränkt wird auf (Systeme für) rationale formale Potenzreihen mit *nicht-verschwindenden* konstanten Koeffizienten, erhält man wieder rationale formale Potenzreihen. Nachdem wir nur die Multiplikation verwenden, ist die Gültigkeit der Polynomfaktorisierung in Satz 3.3.6 *unabhängig* von der Konstruktion des freien Schiefkörpers.

Obwohl es auf den ersten Blick so aussieht, als wollte man mit Kanonen auf Spatzen schießen indem man Polynome als Elemente des freien Schiefkörpers (der freien assoziativen Algebra) betrachtet, gibt es —neben der „Sichtbarkeit“ der Faktorisierung (in zulässigen linearen Systemen)— auch andere Vorteile. So kann man später den linken (oder rechten) größten gemeinsamen Teiler zweier Polynome p, q berechnen, indem man die lineare Darstellung für $p^{-1}q$ minimiert, das heißt, gemeinsame linke Faktoren eliminiert. Das wird im Anhang in Abschnitt A.1 illustriert.

Beispiel 3.3.1 ([Sch17c, Beispiel 3.7])**.** Sei

$$B = \begin{bmatrix} 6 & 1 & 3 \\ -7 & 3 & 14 \\ 1 & 0 & 1 \end{bmatrix}.$$

Dann ist $p(x) = \det(xI - B) = x^3 - 10x^2 + 31x - 30$ das charakteristische Polynom von B. Das linke Begleitsystem $\mathcal{A}_p$ von p (laut Definition 2.2.3) ist

$$\begin{bmatrix} 1 & -x+10 & -31 & 30 \\ . & 1 & -x & . \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} p(x) \\ x^2 \\ x \\ 1 \end{bmatrix}.$$

Die Anwendung der Transformation (P, Q) mit

$$P = \begin{bmatrix} 1 & . & . & . \\ & 1 & -3 & . \\ & & 1 & . \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & -5 & 6 & . \\ & 1 & . & . \\ & & 1 & . \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & -5 & 6 & . \\ & 1 & -3 & . \\ & & 1 & . \\ & & & 1 \end{bmatrix}$$

und

$$Q = \begin{bmatrix} 1 & . & . & . \\ & 1 & 5 & . \\ & & 1 & \frac{6}{5} \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & . & . & . \\ & 1 & . & . \\ & & 1 & \frac{9}{5} \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & . & . & . \\ & 1 & 5 & 9 \\ & & 1 & 3 \\ & & & 1 \end{bmatrix}$$

ergibt das zulässige lineare System $P\mathcal{A}_pQ$,

$$\begin{bmatrix} 1 & 5-x & . & . \\ . & 1 & 2-x & . \\ . & . & 1 & 3-x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} p(x) \\ (x-2)(x-3) \\ x-3 \\ 1 \end{bmatrix}.$$

Also ist die Menge der Eigenwerte von B gleich $\{2,3,5\}$. Für den Fall, dass das Polynom nicht in Linearfaktoren zerfällt, vergleiche auch mit Beispiel 3.3.4.

Definition 3.3.2 (Atomare zulässige lineare Systeme)**.** Ein *minimales* polynomielles ZLS $\mathcal{A} = (1, A, \lambda)$ der Dimension $n \geq 2$ heißt *atomar* (oder *irreduzibel*), wenn es keine Polynomfaktorisierungstransformation (P, Q) gibt, sodass PAQ einen rechten oberen Nullblock der Größe $(n-i-1) \times i$ für ein $i = 1, 2, \ldots, n-2$ hat.

Wenn wir im minimalen ZLS (3.2.2) für $pq = x(1-yx)$ aus dem Beispiel 3.2.1 Spalte 2 zu Spalte 4 addieren, erhalten wir

$$\begin{bmatrix} 1 & -x & . & -x \\ . & 1 & y & 0 \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} x(1-yx) \\ -yx \\ x \\ 1 \end{bmatrix}.$$

Subtrahieren der Zeile 3 von Zeile 1 ergibt

$$\begin{bmatrix} 1 & -x & -1 & 0 \\ . & 1 & y & 0 \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} (1-xy)x \\ -yx \\ x \\ 1 \end{bmatrix}$$

mit einem rechten oberen 2×1 Nullblock in der Systemmatrix. Wir würden das gleiche System (bezogen auf die Nulleinträge) via *minimaler Polynommultiplikation* von $1-xy$ und x erhalten. Das illustriert, dass wir die Faktoren von Polynomen finden können, indem wir nach polynomiell zulässigen Transformationen suchen, die einen entsprechenden Nullblock rechts oben in der (transformierten) Systemmatrix ergeben.

Satz 3.3.6 wird die Korrespondenz zwischen einer Faktorisierung in Atome und der Struktur der rechten oberen Nullblöcke herstellen. Also, um ein Polynom p vom Rang $n \geq 3$ in *nicht-triviale* Faktoren $p = q_1 q_2$ mit $\operatorname{rang}(q_i) = n_i \geq 2$ und $n = n_1 + n_2 - 1$ zu zerlegen, müssen wir Transformationen der Form

$$(P,Q) = \left(\begin{bmatrix} 1 & \alpha_{1,2} & \cdots & \alpha_{1,n-1} & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & 1 & \alpha_{n-2,n-1} & 0 \\ & & & 1 & 0 \\ & & & & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ & 1 & \beta_{2,3} & \cdots & \beta_{2,n} \\ & & 1 & \ddots & \vdots \\ & & & \ddots & \beta_{n-1,n} \\ & & & & 1 \end{bmatrix} \right) \tag{3.3.3}$$

mit Einträgen $\alpha_{ij}, \beta_{ij} \in \mathbb{K}$ finden, siehe Definition 2.1.13. Im allgemeinen ist das ein *nicht-lineares* Problem mit $(n-2)(n-1)$ (kommutativen) Unbekannten.

Beispiel 3.3.4 ([Sch17c, Beispiel 2.12])**.** Sei $p = x^2 - 2 \in \mathbb{K}\langle X \rangle$ gegeben durch das minimale polynomielle ZLS

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & 2 \\ . & 1 & -x \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right).$$

Wenn $\mathbb{K}$ der Körper der rationalen Zahlen $\mathbb{Q}$ ist, dann ist $\mathcal{A}$ atomar (irreduzibel). Wenn $\mathbb{K}$ der Körper der reellen Zahlen $\mathbb{R}$ ist, dann existiert die polynomiell zulässige Transformation

$$(P,Q) = \left(\begin{bmatrix} 1 & \sqrt{2} & . \\ . & 1 & . \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . \\ . & 1 & -\sqrt{2} \\ . & . & 1 \end{bmatrix} \right)$$

der Form (3.3.3) sodass $\mathcal{A}' = P\mathcal{A}Q$ gleich

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x+\sqrt{2} & 0 \\ . & 1 & -x-\sqrt{2} \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

ist, also $p = x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ in $\mathbb{R}\langle X\rangle$.

Bemerkung. Es ist einfach zu prüfen, dass $p = xy - 2$ ein Atom in $\mathbb{K}\langle X\rangle$ ist, weil beide Einträge, $\alpha_{1,2}$ in P und $\beta_{2,3}$ in Q, Null sein müssen (sonst könnte der obere rechte Eintrag in A' nicht verschwinden) und daher gibt es *keine* nicht-triviale polynomielle zulässige Transformation, das heißt, eine Transformation die die obere rechte Block-Struktur ändert.

Lemma 3.3.5 ([Sch17c, Lemma 2.13])**.** *Seien $0 \neq p, q_1, q_2 \in \mathbb{K}\langle X\rangle$ durch die* minimalen *polynomiellen zulässigen linearen Systeme $\mathcal{A} = (1, A, \lambda)$, $\mathcal{A}_1 = (1, A_1, \lambda_1)$ und $\mathcal{A}_2 = (1, A_2, \lambda_2)$ der Dimensionen $n = \operatorname{rang}(p)$ beziehungsweise $n_i = \operatorname{rang}(q_i) \geq 2$ mit $p = q_1 q_2$ gegeben. Dann existiert eine (Polynomfaktorisierungs-)Transformation (P, Q) der Form (3.3.3), sodass PAQ einen rechten oberen Nullblock der Größe $(n_1 - 1) \times (n_2 - 1) = (n - n_2) \times (n - n_1)$ hat.*

Beweis. Sei $\mathcal{A}' = (u', A', v') = (1, A', \lambda')$ das laut Proposition 3.2.7 aus $\frac{\lambda_2}{\lambda}\mathcal{A}_1$ und $\frac{\lambda}{\lambda_2}\mathcal{A}_2$ konstruierte (minimale) ZLS für $p = q_1 q_2$. Klarerweise gilt $\dim(\mathcal{A}') = n_1 + n_2 - 1 = n = \operatorname{rang}(p)$. Und laut Konstruktion hat A' einen rechten oberen Nullblock der Größe $(n_1 - 1) \times (n_2 - 1)$. Beide Systeme, $\mathcal{A}$ und $\mathcal{A}'$, repräsentieren das gleiche Element p, daher existiert laut Satz 2.1.6 eine zulässige Transformation (P, Q), sodass $P\mathcal{A}Q = \mathcal{A}'$ ist. Nachdem $\mathcal{A}'$ polynomiell ist und sich die rechte Seite $Pv = v'$ nicht ändert, ist (P, Q) eine Transformation der Form (3.3.3). □

Satz 3.3.6 (Polynomfaktorisierung [Sch17c, Satz 2.14])**.** *Sei $p \in \mathbb{K}\langle X\rangle$ durch das* minimale *polynomielle ZLS $\mathcal{A} = (1, A, \lambda)$ der Dimension $n = \operatorname{rang}(p) \geq 3$ gegeben. Dann kann p genau dann in $q_1 q_2$ mit $n_i = \operatorname{rang}(q_i) \geq 2$ faktorisiert werden, wenn es eine Polynomfaktorisierungstransformation (P, Q) gibt, sodass PAQ einen rechten oberen Nullblock der Größe $(n_1 - 1) \times (n_2 - 1)$ hat.*

Beweis. Gibt es so eine Faktorisierung, kann Lemma 3.3.5 angewendet werden. Umgekehrt, wenn es so eine (polynomiell zulässige) Transformation für einen Nullblock

der Größe $k_1 \times k_2$ gibt, erhalten wir ein ZLS $\mathcal{A}'$ in *Blockform* (p ist der erste Eintrag von $s_{\underline{1}}$)

$$\begin{bmatrix} A_{1,1} & A_{1,2} & . \\ . & 1 & A_{2,3} \\ . & . & A_{3,3} \end{bmatrix} s = \begin{bmatrix} . \\ . \\ v_{\underline{3}} \end{bmatrix}, \quad s = \begin{bmatrix} s_{\underline{1}} \\ g \\ s_{\underline{3}} \end{bmatrix}$$

mit Blöcken $A_{1,1}$ und $A_{3,3}$ der Größen $k_1 \times k_1$ beziehungsweise $k_2 \times k_2$. In diesem ZLS duplizieren wir den Eintrag s_{k_1+1} indem wir eine Zeile (und Spalte) einfügen, um folgendes ZLS der Größe $k_1 + k_2 + 2 = n + 1$ zu erhalten:

$$\begin{bmatrix} A_{1,1} & A_{1,2} & 0 & . \\ 0 & 1 & -1 & 0 \\ . & . & 1 & A_{2,3} \\ . & . & 0 & A_{3,3} \end{bmatrix} s' = \begin{bmatrix} . \\ . \\ . \\ v_{\underline{3}} \end{bmatrix}, \quad s' = \begin{bmatrix} s_{\underline{1}} \\ g \\ g \\ s_{\underline{3}} \end{bmatrix}.$$

Entsprechend der Konstruktion der Multiplikation in Proposition 2.3.1 steht $p = fg$ in der ersten Komponente von $s_{\underline{1}}$ (dem ersten Block in s') für $f, g \in \mathbb{K}\langle X \rangle$ gegeben durch die (polynomiell) zulässigen linearen Systeme

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ . & 1 \end{bmatrix} s_f = \begin{bmatrix} . \\ 1 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} 1 & A_{2,3} \\ . & A_{3,3} \end{bmatrix} s_g = \begin{bmatrix} . \\ v_{\underline{3}} \end{bmatrix}$$

der Dimension $n_1 = k_1 + 1$ beziehungsweise $n_2 = k_2 + 1$. □

Die praktische Umsetzung der Faktorisierung erfolgt dann mit einer einfachen Variante des Satzes 4.2.1 [CR99, Satz 4.1]. Die Invertierbarkeit der Transformationsmatrizen P und Q in (3.3.3) ist durch ihre Form sichergestellt. Ein Beispiel findet sich in Abschnitt 3.7. In unserem Fall ist der *kommutative* Polynomring

$$\begin{aligned} \mathbb{K}[\alpha, \beta] = \mathbb{K}[&\alpha_{1,2}, \ldots, \alpha_{1,n-1}, \alpha_{2,3}, \ldots, \alpha_{2,n-1}, \ldots, \alpha_{n-2,n-1}, \\ &\beta_{2,3}, \ldots, \beta_{2,n}, \beta_{3,4}, \ldots, \beta_{3,n}, \ldots, \beta_{n-1,n}]. \end{aligned}$$

Allerdings garantiert ein nicht-triviales Ideal *keine* Lösung über $\mathbb{K}$, obwohl Lösungen über $\overline{\mathbb{K}}$ existieren. Will man nur testen (ohne sie zu berechnen), *ob* es Lösungen (über $\mathbb{K}$) gibt, kann man das Konzept der *Resultanten* verwenden. Eine Einleitung findet man in [CLO15, Abschnitt 3.6]. Dieses Buch beinhaltet auch eine Einleitung in Gröbner-Basen und einen Überblick über Computer-Algebra-Systeme um sie zu berechnen. Zusätzlich zur Arbeit von [Buc70], kann der Überblick über Gröbner–Shirshov-Basen von [BK00] konsultiert werden.

Bemerkung. Um im allgemeinen die Multiplikation (Proposition 3.2.7) umzudrehen, um Faktoren zu finden, brauchen wir auch einen Nullblock entsprechender Größe links unten. Diese wichtige Voraussetzung ist in der Form eines polynomiellen ZLS versteckt. Siehe auch Abschnitt 3.6 (Allgemeine Faktorisierung).

Proposition 3.3.7 ([Sch17c, Proposition 2.15]). *Sei $p \in \mathbb{K}\langle X\rangle$ gegeben durch das minimale polynomielle ZLS $\mathcal{A} = (1, A, \lambda)$ der Dimension $n = \operatorname{rang}(p) \geq 3$ und sei (P, Q) wie in (3.3.3). Fixiere ein $k \in \{1, 2, \ldots, n-2\}$ und bezeichne mit I_k das Ideal von $\mathbb{K}[\alpha, \beta]$, das durch die Koeffizienten von jedem $x \in \{1\} \cup X$ in den (i, j)-Einträgen der Matrix PAQ für $1 \leq i \leq k$ und $k+2 \leq j \leq n$ erzeugt wird. Dann kann p über $\overline{\mathbb{K}}\langle X\rangle$ in $q_1 q_2$ mit $\operatorname{rang}(q_1) = k+1$ und $\operatorname{rang}(q_2) = n-k$ faktorisiert werden, genau dann, wenn das Ideal $I_k \neq \mathbb{K}[\alpha, \beta]$ ist.*

Bemerkung 3.3.8 ([Sch17c, Abschnitt 2]). Es sei ein Polynom $p \in \mathbb{K}\langle X\rangle$ durch ein minimales polynomielles ZLS der Dimension $n = \operatorname{rang}(p) \geq 2$ gegeben. Dann gibt es maximal $\phi(n) = 2^{n-2}$ (minimale) Standardsysteme (in Bezug auf die Struktur der oberen rechten Nullblöcke). Für $n = 2$ ist das klar. Für $n > 2$ kann das ZLS atomar (bezogen auf den algebraisch abgeschlossenen Grundkörper $\overline{\mathbb{K}}$) sein oder einen Nullblock der Größe $1\times(n-2)$ oder $(n-2)\times 1$ haben, also gilt $\phi(n+1) = 1+2\phi(n)-1 = 2\phi(n)$, weil das System mit der „feinsten“ oberen rechten Struktur doppelt gezählt wird.

Bemerkung. Modulo Ähnlichkeit hat jedes Element $p \in \mathbb{K}\langle X\rangle \setminus \mathbb{K}$ nur *eine* Faktorisierung in Atome. Wenn man an der Anzahl der Faktorisierungen (nicht unbedingt in Atome) modulo Permutationen (und Multiplikation der Faktoren mit Einheiten) interessiert ist, kann die Abschätzung $2^{\operatorname{rang}(p)-2}$ der Bemerkung 3.3.8 verwendet werden. Jedoch kann die Anzahl der Faktorisierungen im Sinne von [BHL17, Definition 3.1] größer sein. Als Beispiel diene das Polynom $p = (x-1)(x-2)(x-3)$, das $3! = 6$ verschiedene Faktorisierungen (modulo dem Einfügen von Einheiten) besitzt, während die Anzahl der polynomiellen zulässigen linearen Systeme mit $2^{\operatorname{rang}(p)-2} = 4$ beschränkt ist.

Sei p ein Polynom mit der Faktorisierung $p = q_1 q_2 \cdots q_m$ in Atome $q_i \in \mathbb{K}\langle X\rangle$. Da $\mathbb{K}\langle X\rangle$ ein Ähnlichkeits-UFD ist (Proposition 3.1.5), hat *jede* Faktorisierung von p in Atome m Faktoren. Daher kann man die *Länge* von p mit m definieren, geschrieben als $\ell(p) = m$. Für ein Wort $w \in X^* \subseteq \mathbb{K}\langle X\rangle$ ist die Länge $\ell(w) = |w|$. Ein Blick auf Proposition 3.2.7 (Minimale Polynommultiplikation) zeigt, dass die Länge eines Elementes $p \in \mathbb{K}\langle X\rangle^\bullet$ mit dem Rang abgeschätzt werden kann, nämlich $\ell(p) \leq \operatorname{rang}(p) - 1$. Mehr über Längenfunktionen —und Transfer-Homomorphismen im Kontext nicht-eindeutiger Faktorisierungen— (in nicht-kommutativer Umgebung) findet man in [Sme15, Abschnitt 3] oder [BS15].

3.4 Faktorisierungstheorie

In $\mathbb{F} = \mathbb{K}(\langle X\rangle)$ gilt $\mathbb{F} \setminus \{0\} = \mathbb{F}^\bullet = \mathbb{F}^\times = \{f \in \mathbb{F} \mid f \text{ ist invertierbar}\}$, das heißt, es gibt keine von Null verschiedenen Nicht-Einheiten. Daher werden wir die Elemente über ihre *minimalen* linearen Darstellungen betrachten. Wir erinnern uns, dass der Rang (Abschnitt 2.1) eines Elementes $f \in \mathbb{F}$ durch die Dimension eines minimalen ZLS von f definiert ist.

Zuerst definieren wir *Faktoren* basierend auf dem Rang, in Definition 3.4.1. Obwohl diese Definition ausreichen würde, um die klassische Teilbarkeit für Polynome zu definieren, ist sie im allgemeinen zu starr. Nachdem das alles andere als offensichtlich ist, wird das detailliert anhand eines Beispiels vor Definition 3.4.4 (Links- und Rechts-Teilbarkeit) erklärt. Danach ist einige Vorbereitung notwendig, um das Einfügen von nicht-trivialen Einheiten auszuschließen. Das ist das Wesen von Lemma 3.4.7. Zuletzt liefert der Satz 3.4.9 die Äquivalenz der „klassischen“ Teilbarkeit (in der freien assoziativen Algebra) und der neuen (im freien Schiefkörper) für Polynome.

Für eine Faktorisierung eines Polynoms $p = q_1 q_2 \cdots q_m$ in Atome q_i wollen wir auch eine Faktorisierung seiner Inversen $p^{-1} = (q_1 q_2 \cdots q_m)^{-1} = q_m^{-1} \cdots q_2^{-1} q_1^{-1}$ in Atome q_i^{-1} haben. Für zwei Polynome p, q gilt $\operatorname{rang}(p) + \operatorname{rang}(q) = \operatorname{rang}(pq) + 1$ laut Proposition 3.2.7 (Minimale Polynommultiplikation). Laut Definition 3.1.2 gilt $p \,|_1\, h$ wenn $h = pq$ für ein $q \in \mathbb{K}\langle X\rangle$ ist. Satz 2.5.13 (Minimale Inverse Typ $(1,1)$) liefert

$$\operatorname{rang}(q^{-1}) + \operatorname{rang}(p^{-1}) = \operatorname{rang}(q) - 1 + \operatorname{rang}(p) - 1 = \operatorname{rang}(q^{-1}p^{-1}),$$

oder q^{-1} „teilt“ h^{-1} von links für $h = pq$. Um das Einfügen von nicht-trivialen Einheiten von $\mathbb{F} \setminus \mathbb{K}$ zu vermeiden, müssen wir die Summe der Ränge der Faktoren beschränken:

$$\operatorname{rang}(px) + \operatorname{rang}(x^{-1}q) = \operatorname{rang}(pq) + 2.$$

Definition 3.4.1 (Linke und rechte Faktoren [Sch17a, Definition 3.1])**.** Sei $h \in \mathbb{H} = \mathbb{F}^\bullet$ gegeben. Ein Element $f \in \mathbb{H}$ heißt *linker Faktor* von h wenn

$$\begin{aligned} \operatorname{rang}(f) + \operatorname{rang}(f^{-1}h) &\leq \operatorname{rang}(h) + 1 \quad \text{und} \\ \operatorname{rang}(h^{-1}f) + \operatorname{rang}(f^{-1}) &\leq \operatorname{rang}(h^{-1}) + 1. \end{aligned}$$

Ein Element $g \in \mathbb{H}$ heißt *rechter Faktor* von h wenn

$$\begin{aligned} \operatorname{rang}(hg^{-1}) + \operatorname{rang}(g) &\leq \operatorname{rang}(h) + 1 \quad \text{und} \\ \operatorname{rang}(g^{-1}) + \operatorname{rang}(gh^{-1}) &\leq \operatorname{rang}(h^{-1}) + 1. \end{aligned}$$

Skalare und skalare Vielfache von h heißen *triviale* Faktoren. Nicht-triviale linke oder rechte Faktoren heißen *echt*. Linke und rechte Faktoren heißen auch *äußere* Faktoren um sie —wenn es notwendig ist— von (allgemeinen) Faktoren einer Faktorisierung zu unterscheiden.

Bemerkung. Unmittelbar gilt: f ist genau dann ein *linker Faktor* von h, wenn $g = f^{-1}h$ ein *rechter Faktor* von h ist. Und f ist genau dann ein *linker Faktor* von h, wenn f^{-1} ein *rechter Faktor* von h^{-1} ist.

Für zwei Polynome p und q sagt die vorhergehende Definition nichts anderes als: p (bzw. q) ist ein linker (bzw. rechter) Faktor von pq. Allerdings, im allgemeinen ist f *kein* linker Faktor von $h = fg$. Als Beispiel diene $f = (xyz)^{-1}$ und $g = x$. Es ist

$$\begin{aligned} \operatorname{rang}(f) + \operatorname{rang}(g) &= \operatorname{rang}(z^{-1}y^{-1}x^{-1}) + \operatorname{rang}(x) \\ &= 3 + 2 \\ &> \operatorname{rang}(z^{-1}y^{-1}) + 1. \end{aligned}$$

Obwohl hier einfach zu sehen ist, dass f^{-1} und g einen nicht-trivialen linken Teiler in $\mathbb{K}\langle X\rangle$ haben (im Sinne der Definition 3.1.2), kann das im allgemeinen viel diffiziler sein, illustriert in Beispiel 3.4.2. Dieses Beispiel wird auch zeigen, dass die Definition von äußeren Faktoren ziemlich restriktiv ist und nicht direkt angewandt werden kann. Später werden *linke* und *rechte Teiler* allgemeiner definiert, sodass äußere Faktoren in zumindest einer Folge „abgespalten“ werden können (siehe Definition 3.4.4). Obwohl wir später sehen werden, dass das eine Verallgemeinerung der Faktorisierung in der freien assoziativen Algebra ist, ist sie aus zwei Gründen sehr viel schwieriger anzuwenden: Zum einen muss man *alle* möglichen „Folgen“ von Faktorisierungen testen, um die Atome (modulo „Ähnlichkeit“) zu erhalten. Zum anderen muss die Invertierbarkeit der Transformationsmatrizen —um das ZLS in so einer Art zulässig zu transformieren, dass die Faktoren „extrahiert“ werden können— über eine Bedingung für nicht-verschwindende Determinanten sichergestellt werden. Letzteres könnte die praktische Anwendbarkeit auf Rang ≤ 6, ähnlich zum Test ob eine Matrix voll ist [Jan18], beschränken. Die Abschnitte 3.5 (Minimale Faktormultiplikation) und 3.6 (Allgemeine Faktorisierung) enthalten weitere Details. Experimente zeigen, dass das Testen von (Ir-)Reduzibilität von Polynomen unter der Verwendung einer polynomiellen Form praktisch für Rang ≤ 12 und in einigen Fällen bis Rang ≤ 17 funktioniert [Jan18].

Bemerkung. Selbst in „einfachen“ Fällen ist es schwierig zu entscheiden, ob ein Element irreduzibel ist oder nicht, basierend auf regulären Ausdrücken. Als Beispiel diene $f = 1 - xy$ und $g = (1 - zy)^{-1}$. Dann ist fg *irreduzibel* während gf *reduzibel* ist. Man muss deren jeweilige *minimale* lineare Darstellungen betrachten. Minimale zulässige lineare Systeme für fg und gf sind

$$\begin{bmatrix} 1 & -1 & -x \\ . & y & 1 \\ . & 1 & z \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \quad \text{bzw.} \quad \begin{bmatrix} y & 1 & . & . \\ 1 & z & -x & -1 \\ . & . & 1 & y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Das Prinzip, wie man einen (echten) linken Teiler (in einem ZLS) findet, wird im Detail im folgenden Beispiel erklärt.

Beispiel 3.4.2 ([Sch17a, Beispiel 3.2])**.** Sei $f = f_1 f_2 f_3$, mit $f_1 = (xy)^{-1}$, $f_2 = 1 - xz$ und $f_3 = (yz)^{-1}$, gegeben durch das *minimale* ZLS

$$\begin{bmatrix} y & -1 & z & 0 \\ . & x & -1 & 0 \\ 0 & 0 & z & -1 \\ 0 & 0 & . & y \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{3.4.3}$$

Dann sieht man sofort (nachdem man sich die Konstruktion eines ZLS für das Produkt laut Proposition 2.3.1 in Erinnerung ruft), dass f_3 ein rechter Faktor von f ist, indem

man s_3 dupliziert, das heißt, eine Zeile (zwischen Zeile 2 und Zeile 3) einfügt:

$$\begin{bmatrix} y & -1 & z & 0 & . \\ . & x & -1 & 0 & . \\ 0 & 0 & 1 & -1 & 0 \\ . & . & 0 & z & -1 \\ . & . & 0 & . & y \end{bmatrix} s' = \begin{bmatrix} . \\ . \\ 0 \\ . \\ 1 \end{bmatrix}, \quad s' = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_3 \\ s_4 \end{bmatrix}.$$

Wenn also ein minimales ZLS $\mathcal{A} = (u, A, v)$ für f nicht die Form von (3.4.3) hat, müssen wir eine zulässige Transformation (P, Q) finden, sodass die (transformierte) Systemmatrix PAQ einen Nullblock der Größe 2×2 links unten und einen der Größe 2×1 rechts oben hat und nur die letzte Komponente der rechten Seite Pv von Null verschieden ist, um den rechten Faktor f_3 zu erkennen. In ähnlicher Weise, wenn man in (3.4.3) Zeile 3 von Zeile 1 subtrahiert und Spalte 2 zu Spalte 4 addiert, bekommt man

$$\begin{bmatrix} y & -1 & 0 & 0 \\ . & x & -1 & x \\ 0 & 0 & z & -1 \\ 0 & 0 & . & y \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ . \\ 1 \end{bmatrix}.$$

Vergleiche mit Abbildung 3.1, $k = 2$ in Typ $(*, 1)$. Indem man t_2 dupliziert, das heißt, eine Spalte (zwischen Spalte 2 und 3) einfügt, sieht man, dass $f_1 = (xy)^{-1}$ ein linker Faktor von $f = (xy)^{-1}(1 - xz)(yz)^{-1}$ ist:

$$\begin{bmatrix} 1 & . & 0 & . & . \end{bmatrix} = \begin{bmatrix} t_1 & t_2 & t_2 & t_3 & t_4 \end{bmatrix} \begin{bmatrix} y & -1 & 0 & . & . \\ . & x & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & x \\ . & . & 0 & z & -1 \\ . & . & 0 & . & y \end{bmatrix}.$$

Allerdings ist f_1 (bzw. f_2) *kein* linker (bzw. rechter) Faktor von $f_1 f_2$ während y^{-1} ein linker Faktor von $f_1 f_2$ und x^{-1} ein linker Faktor von $x^{-1} f_2$ ist. Nun schauen wir uns dieses Phänomen genauer an. Ein *minimales* ZLS für $f' = f_1 f_2$ ist gegeben durch

$$\begin{bmatrix} y & -1 & z \\ . & x & -1 \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} y^{-1}(x^{-1} - z) \\ x^{-1} \\ 1 \end{bmatrix}.$$

Der Grund ist, dass sich der Rang nicht erhöht (wenn f_2 mit x^{-1} und y^{-1} von links multipliziert wird), weil $1 \in R(x^{-1} f_2)$ ist:

$$\begin{bmatrix} 1 & . & . \end{bmatrix} = t \begin{bmatrix} x & -x & -1 \\ . & 1 & z \\ . & . & 1 \end{bmatrix}, \quad t = \begin{bmatrix} x^{-1} & 1 & x^{-1} - z \end{bmatrix}.$$

Nachdem man Spalte 2 zu Spalte 1 addiert und die ersten beiden Zeilen vertauscht hat (das führt dazu, dass die ersten beiden Spalten in t vertauscht werden) erhält

man $[1,0,0]^\top$ in der ersten Spalte der Systemmatrix (für die Existenz dieser Transformationen siehe Lemma 2.3.5):

$$\begin{bmatrix}1 & . & .\end{bmatrix} = t \begin{bmatrix}1 & 1 & z\\ 0 & -x & -1\\ . & . & 1\end{bmatrix}, \quad t = \begin{bmatrix}1 & x^{-1} & x^{-1}-z\end{bmatrix}.$$

Nachdem man $x^{-1}f_2$ von links mit y^{-1} multipliziert, gilt $1 \notin R(f')$. Daher *erhöht* eine weitere Multiplikation mit (zum Beispiel) z^{-1} von links den Rang:

$$\begin{bmatrix}z & -1 & . & .\\ . & y & -1 & z\\ . & . & x & -1\\ . & . & . & 1\end{bmatrix} s = \begin{bmatrix}.\\ .\\ .\\ 1\end{bmatrix}, \quad s = \begin{bmatrix}z^{-1}f'\\ f'\\ x^{-1}\\ 1\end{bmatrix}.$$

Zusammenfassend gibt es im wesentlichen zwei verschiedene Faktorisierungen von $f = (xy)^{-1}(1-xz)(yz)^{-1}$ (auf der Ebene von äußeren Faktoren in Definition 3.4.1), nämlich

$$f = \Big(y^{-1}\big(x^{-1}(1-xz)\big)\Big)(yz)^{-1} = (xy)^{-1}\Big(\big((1-xz)z^{-1}\big)y^{-1}\Big).$$

Nun kommen wir zur Hauptdefinition, die die Definition 3.1.2 von linken und rechten Teilern in der freien assoziativen Algebra verallgemeinert. Um in Satz 3.4.9 zeigen zu können, dass diese Definitionen tatsächlich auf $\mathbb{K}\langle X\rangle$ äquivalent sind, ist ein wenig Vorbereitung notwendig.

Notation. Sei $m \in \mathbb{N}$ und eine Permutation $\sigma = (i_1, i_2, \ldots, i_m) \in \mathfrak{S}_m$ fixiert. Für $k \in 1, 2, \ldots, m$ sei $\underline{\sigma}_k = \min\{i_1, i_2, \ldots, i_k\}$ und $\overline{\sigma}_k = \max\{i_1, i_2, \ldots, i_k\}$. Wir schreiben $\sigma_{j,k}$ für das j-te Element in der (geordneten) Folge $\underline{\sigma}_k = \sigma_{1,k} < \sigma_{2,k} < \ldots < \sigma_{k,k} = \overline{\sigma}_k$.

Definition 3.4.4 (Links- und Rechtsteilbarkeit, koprime Elemente [Sch17a, Definition 3.4])**.** Sei $\mathbb{H} = \mathbb{F}^\bullet$. Ein Element $\mathbb{H} \ni g$ *teilt* $f \in \mathbb{H}$ *von links*, geschrieben als $g \,|_{\mathrm{l}}^{\mathbb{F}}\, f$, wenn, für ein $m' < m \in \mathbb{N}$, $f_1, f_2, \ldots, f_m \in \mathbb{H}$ und eine Permutation $\sigma = (i_1, i_2, \ldots, i_m) \in \mathfrak{S}_m$ existieren, sodass $g = f_1 f_2 \cdots f_{m'}$ und $f = g f_{m'+1} \cdots f_m$ ist und für alle $k = 2, 3, \ldots, m$ entweder

- f_{i_k} ein *linker Faktor* von $f_{\sigma_{1,k}} f_{\sigma_{2,k}} \cdots f_{\sigma_{k,k}}$ ist, wenn $i_k < \underline{\sigma}_{k-1}$ gilt oder
- f_{i_k} ein *rechter Faktor* von $f_{\sigma_{1,k}} f_{\sigma_{2,k}} \cdots f_{\sigma_{k,k}}$ ist, wenn $i_k > \overline{\sigma}_{k-1}$ gilt.

Zwei Elemente $f, g \in \mathbb{H}$ heißen *links koprim* (in $\mathbb{H}$) wenn für alle $h \in \mathbb{H}$ mit $h \,|_{\mathrm{l}}^{\mathbb{F}}\, f$ und $h \,|_{\mathrm{l}}^{\mathbb{F}}\, g$ gilt, dass $h \in \mathbb{K}^\times$ ist, das heißt, h ein Element der *trivialen Einheitengruppe* ist. Rechte Teilbarkeit $f \,|_{\mathrm{r}}^{\mathbb{F}}\, g$ und die Notation *rechts koprim* (in $\mathbb{H}$) ist in ähnlicher Weise definiert. Zwei Elemente (in $\mathbb{H}$) heißen *koprim* wenn sie links und rechts koprim sind.

Beispiel. Für $g = (xy)^{-1}$ und $f = g(1-xz)$ haben wir $m' = 2$, $m = 3$, $f_1 = y^{-1}$, $f_2 = x^{-1}$ und $f_3 = 1 - xz$ und die Permutation $\sigma = (3, 2, 1)$: $f = f_1(f_2 f_3)$, das heißt, *g teilt f von links* (in $\mathbb{F}$) weil $g = f_1 f_2$ ist, f_2 ist ein *linker Faktor* von $f_2 f_3$ und f_1 ist ein *linker Faktor* von $f_1 f_2 f_3$.

Lemma 3.4.5 (Rang-Lemma [Sch17a, Lemma 3.5]). *Sei $0 \neq p, q \in \mathbb{F}$. Dann gilt*

(i) $\operatorname{rang}(pq) = \operatorname{rang}(p) + \operatorname{rang}(q) - 1$, *wenn* $p, q \in \mathbb{K}\langle X\rangle$ *sind,*

(ii) $\operatorname{rang}(p) = \operatorname{rang}(p^{-1}) + 1$, *wenn* $p \in \mathbb{K}\langle X\rangle$ *ist,*

(iii) $\operatorname{rang}(p^{-1}q) \leq \operatorname{rang}(p^{-1}) + \operatorname{rang}(q) - 1$, *wenn* $1 \in R(q)$ *ist,*

(iv) $\operatorname{rang}(p^{-1}q) \geq \operatorname{rang}(p^{-1}) + \operatorname{rang}(q) - 1$, *wenn* p^{-1} *ein* linker Faktor *von* $p^{-1}q$ *ist,*

(v) $\operatorname{rang}(pq^{-1}) \leq \operatorname{rang}(p) + \operatorname{rang}(q^{-1}) - 1$, *wenn* $1 \in L(p)$ *ist und*

(vi) $\operatorname{rang}(pq^{-1}) \geq \operatorname{rang}(p) + \operatorname{rang}(q^{-1}) - 1$, *wenn* q^{-1} *ein* rechter Faktor *von* pq^{-1} *ist.*

Beweis. Die Rang-Identitäten (i) und (ii) sind unmittelbare Konsequenzen der minimalen Polynommultiplikation (Proposition 3.2.7) beziehungsweise der minimalen Inversen (Satz 2.5.13). Die Ungleichungen (iv) und (vi) folgen direkt aus der Definition 3.4.1. Um (iii) zu beweisen seien p^{-1} und q durch die *minimalen* zulässigen linearen Systeme $\mathcal{A}'_p = (u'_p, A'_p, v'_p)$ der Dimension n'_p beziehungsweise $\mathcal{A}_q = (u_q, A_q, v_q)$ der Dimension n_q gegeben. Die Konstruktion laut Proposition 2.3.9 ergibt ein ZLS der Dimension n' (für $p^{-1}q$), daher gilt $\operatorname{rang}(p^{-1}q) \leq n' = \operatorname{rang}(p^{-1}) + \operatorname{rang}(q) - 1$. Der Beweis von (v) ist —abgesehen von der Anwendung der Proposition 2.3.6— ähnlich dem von (iii). □

Bemerkungen. Die Rang-Abschätzungen (iii) und (v) gelten insbesondere für Polynome. Weiters, das Beispiel 3.4.2 in Erinnerung rufend, ist es notwendig aber *nicht* hinreichend, dass p und q *links koprim* sind, damit (iv) hält.

Um die Idee des folgenden Lemmas zu illustrieren, nehmen wir $p = xyz$ und ein beliebiges f mit $\operatorname{rang}(f) = 2$. Es ist einfach zu sehen, dass $\operatorname{rang}(pf^{-1}) \geq \operatorname{rang}(p) - \operatorname{rang}(f) = 2$ (mit Gleichheit für $f = yz$) ist, weil wir sonst ein ZLS der Dimension $\operatorname{rang}(pf^{-1}) + \operatorname{rang}(f^{-1}) - 1 < \operatorname{rang}(p)$ konstruieren könnten. Nun sei $f = xz$. Dann ist $\operatorname{rang}(pf^{-1}) = \operatorname{rang}(xyx^{-1}) = 3$. Jedoch, für ein weiteres Polynom q bekommen wir $\operatorname{rang}(fq) = 2 + \operatorname{rang}(q)$, daher gilt $\operatorname{rang}(pf^{-1}) + \operatorname{rang}(fq) = 3 + 2 + \operatorname{rang}(q) > 3 + \operatorname{rang}(q) = \operatorname{rang}(pq) + 1$.

Was in einem konkreten Beispiel verhältnismäßig einfach ist, nämlich zu verifizieren, dass wir keine nicht-triviale Einheit „einfügen“ können, ist im allgemeinen sehr technisch, weil wir die linke und rechte Familie im Detail untersuchen müssen. Ein subtiler Schritt wird im folgenden Beispiel diskutiert.

Beispiel 3.4.6. Sei $h = xy + z^{-1}$ (vom Typ $(1, 1)$) gegeben durch das *minimale* ZLS

$$\begin{bmatrix} 1 & x & z^{-1} & xy + z^{-1} \end{bmatrix} = \begin{bmatrix} 1 & . & . & . \end{bmatrix} \begin{bmatrix} 1 & -x & 1 & 0 \\ . & 0 & z & 1 \\ . & 1 & 0 & -y \\ . & . & . & 1 \end{bmatrix}.$$

Insbesondere ist die rechte Familie $t = (t_1, t_2, t_3, t_4)$ $\mathbb{K}$-linear unabhängig. Daher können wir t_3 nicht als Linearkombination (über $\mathbb{K}$) von h, t_1 und t_2 schreiben. Das ist trivial. Das folgende nicht: Wenn wir ein Polynom q invertieren (in Bezug auf die minimale Inverse), z.B. $q = z$, bekommen wir q^{-1} vom Typ $(0, 0)$. Was kann passieren, wenn wir ein Polynom p mit einem Element vom Typ $(1, 1)$ multiplizieren, das eine nicht-triviale Inverse eines Polynoms (additiv) „enthält“?

Lemma 3.4.7 ([Sch17a, Lemma 3.7])**.** *Sei* $p \in \mathbb{K}\langle X\rangle \setminus \mathbb{K}$ *und* $f \in \mathbb{F} \setminus \mathbb{K}$ *sodass* $p_{i_0}p_{i_0+1}\cdots p_m f^{-1} \notin \mathbb{K}^\times$ *für* alle *Faktorisierungen* $p = p_1p_2\cdots p_m$ *in Atome und* alle $i_0 \in \{1, 2, \ldots, m\}$ *gilt. Dann ist* $\operatorname{rang}(pf^{-1}) + \operatorname{rang}(f) > \operatorname{rang}(p) + 1$.

Beweis. Für ein festgehaltenes nicht-skalares Polynom p betrachten wir die Faktorisierungen $p = p_1p_2\cdots p_m$ in m Atome p_j. Um die Notation zu vereinfachen sei $p_0 = p_{m+1} = 1$. Hier bezeichnen $p_1, p_2, \ldots, p_m$ immer Atome. Sei

$$r = \min\big\{\operatorname{rang}(p_{i_0}p_{i_0+1}\cdots p_m f^{-1}) \mid p = p_1p_2\cdots p_m \text{ und } i_0 \in \{1, 2, \ldots, m\}\big\}$$

und i_0 und $p_1p_2\cdots p_m$ sodass dieses Minimum angenommen wird. Laut Annahme gilt $h = p_{i_0}p_{i_0+1}\cdots p_m f^{-1} \in \mathbb{F} \setminus \mathbb{K}$ mit $\operatorname{rang}(h) = r$. Also ist $f = h^{-1}p_{i_0}p_{i_0+1}\cdots p_m$ mit *nicht-skalarem* h. Laut Satz 2.5.13 (Minimale Inverse) gibt es vier Fälle:

- $r \geq 2$ und $\operatorname{rang}(h^{-1}) = r - 1$ für Typ $(1, 1)$,
- $r \geq 2$ und $\operatorname{rang}(h^{-1}) = r$ für Typ $(1, 0)$,
- $r \geq 2$ und $\operatorname{rang}(h^{-1}) = r$ für Typ $(0, 1)$ und
- $r \geq 1$ und $\operatorname{rang}(h^{-1}) = r + 1$ für Typ $(0, 0)$.

Nun fixieren wir eine beliebige Faktorisierung von p (in Atome q_i) und einem $1 < \ell \leq m$ und setzen $p' = q_1q_2\cdots q_{\ell-1}$ und $p'' = q_\ell q_{\ell+1}\cdots q_m$ mit den Rängen n' beziehungsweise n''. Wir müssen zeigen, dass

$$\operatorname{rang}(p'h) + \operatorname{rang}(h^{-1}p'') > \operatorname{rang}(p'p'') + 1 = \operatorname{rang}(p') + \operatorname{rang}(p'')$$

ist. Wir gehen wie folgt vor: In Abhängigkeit der vier Fälle konstruieren wir —unter Verwendung der Propositionen 2.3.6 und 2.3.9— zulässige lineare Systeme für $n'h$ beziehungsweise $h^{-1}n''$ und leiten eine obere Schranke für die Anzahl der Zeilen/Spalten her, die (wegen der $\mathbb{K}$-linearen Abhängigkeit der Einträge in der rechten beziehungsweise linken Familie) eliminiert werden können.

Wir beginnen mit der Annahme von Typ $(1, 1)$. Für $n'h$ konstruieren wir ein ZLS $\mathcal{A}'$ der Dimension $n_1 = n' + r - 1$ mit der Block-Zerlegung (als lineare Darstellung laut Satz 2.1.6)

$$\pi' = \left(\begin{bmatrix} 0 & u' & . \end{bmatrix}, \begin{bmatrix} A'_{1,1} & A'_{1,2} & A'_{1,3} \\ . & A'_{2,2} & A'_{2,3} \\ . & . & A'_{3,3} \end{bmatrix}, \begin{bmatrix} . \\ v' \\ 0 \end{bmatrix} \right).$$

Für $h^{-1}n''$ konstruieren wir $\mathcal{A}''$ der Dimension $n_2 = n''+r-2$ mit der Block-Zerlegung

$$\pi'' = \left(\begin{bmatrix} 0 & u'' & . \end{bmatrix}, \begin{bmatrix} A''_{1,1} & A''_{1,2} & A''_{1,3} \\ . & A''_{2,2} & A''_{2,3} \\ . & . & A''_{3,3} \end{bmatrix}, \begin{bmatrix} . \\ v'' \\ 0 \end{bmatrix} \right).$$

Sei k'_t (bzw. k'_s) die Größe des Blocks $A'_{1,1}$ (bzw. $A'_{3,3}$) in π' und k''_t (bzw. k''_s) die Größe des Blocks $A''_{1,1}$ (bzw. $A''_{3,3}$) in π''. Zuerst schreiben wir die linke und die rechte Familie von h^{-1} in Bezug auf die jeweilige Familie von h: Seien $(s^h_1, s^h_2, \ldots, s^h_r)$ und $(t^h_1, t^h_2, \ldots, t^h_r)$ die linke beziehungsweise rechte Familie eines minimalen ZLS für h. Dann sind $s_{h^{-1}} = (1, s^h_{r-1}, \ldots, s^h_2)h^{-1}$ und $t_{h^{-1}} = h^{-1}(t^h_{r-1}, \ldots, t^h_2, 1)$ die Familien eines minimalen ZLS für h^{-1}, konstruiert laut Satz 2.5.13 (Minimale Inverse). Wir erinnern uns, dass Zeile/Spalte n' in einem System der Dimension $n'+r$ eliminiert wurde um $\mathcal{A}'$ zu erhalten und Zeile/Spalte r in einem System der Dimension $r-1+n''$ eliminiert wurde um $\mathcal{A}''$ zu erhalten. Danach werfen wir einen genaueren Blick auf die linken Familien von $\mathcal{A}'$ und $\mathcal{A}''$. Diese sind (ohne Beschränkung der Allgemeinheit)

$$\begin{aligned} s' &= (s^{p'}_1 h, s^{p'}_2 h, \ldots, s^{p'}_{n'-1}h, s^h_1, s^h_2, \ldots, s^h_r) \quad \text{bzw.} \\ s'' &= (\underbrace{h^{-1}p'', s^h_{r-1}h^{-1}p'', \ldots, s^h_2 h^{-1}p''}_{r-1}, \underbrace{s^{p''}_2, \ldots, s^{p''}_{n''}}_{n''-1}). \end{aligned}$$

Die erste Beobachtung ist, dass $k'_s = 0$ ist, das heißt, die linke Familie von $\mathcal{A}'$ ist $\mathbb{K}$-linear unabhängig, weil $1 \in R(h)$ ist und Lemma 3.2.5 angewendet werden kann. Die ersten $r-1$ und die letzten $n''-1$ Komponenten von s'' sind $\mathbb{K}$-linear unabhängig. Maximal $r-1$ (Linearkombinationen von) Komponenten in s'' können eliminiert werden. Daher gilt $k''_s \leq r-1$. Jedoch behaupten wir, dass

$$k''_s \leq r-2$$

ist. Wir nehmen gegenteilig an, dass (der Rang von n'' groß genug ist und) $k''_s = r-1$ ist, das heißt, der Block $A''_{3,3}$ hat Dimension k''_s. Dann können *alle* zu h^{-1} korrespondierenden (Linearkombinationen von) Komponenten in s'' mittels der letzten $n''-1$ *polynomiellen* Einträge eliminiert werden. Daher ist insbesondere die erste Komponente $h^{-1}p''$ ein Polynom mit $\operatorname{rang}(h^{-1}p'') < \operatorname{rang}(p'') = n''$. Das heißt, $h = \kappa q_\ell q_{\ell+1} \cdots q_{\ell'}$ für ein ℓ' sodass $\ell < \ell' \leq m$ und $\kappa \in \mathbb{K}$ ist, was der Minimalität von $\operatorname{rang}(h)$ widerspricht. Daher ist $k''_s \leq r-2$. Als nächstes werfen wir einen genaueren Blick auf die rechten Familien

$$\begin{aligned} t' &= (\underbrace{t^{p'}_1, t^{p'}_2, \ldots, t^{p'}_{n'-1}, p't^h_1}_{n'}, p't^h_2, \ldots, p't^h_r) \quad \text{bzw.} \\ t'' &= (h^{-1}t^h_{r-1}, \ldots, h^{-1}t^h_2, h^{-1}, h^{-1}t^{p''}_2, \ldots, h^{-1}t^{p''}_{n''}) \\ &= h^{-1}(\underbrace{t^h_{r-1}, \ldots, t^h_2, 1}_{r-1}, t^{p''}_2, \ldots, t^{p''}_{n''}). \end{aligned}$$

Laut Annahme sind die ersten $r-1$ und die letzten n'' Komponenten in t'' $\mathbb{K}$-linear unabhängig. Nachdem die $t_i^{p''}$s Polynome sind, können wir maximal $k \leq r-2$ *polynomielle* (Linearkombinationen von) Komponenten in t'' eliminieren. Aber dann sind die entsprechenden k (Linearkombinationen von) Komponenten in t' $\mathbb{K}$-linear unabhängig von $(t_1^{p'}, \ldots, t_{n'-1}^{p'})$ weil sie von der Form $p'\tilde{t}_i^h$ sind (siehe Lemma 3.2.6 oder die minimale Polynommultiplikation). Laut Annahme sind die ersten n' und die letzten r Komponenten in t' $\mathbb{K}$-linear unabhängig. Also können maximal $r-1-k$ Komponenten eliminiert werden. Daher gilt $k_t' + k_t'' \leq r-1$. Jedoch behaupten wir, dass

$$k_t' + k_t'' \leq r-2$$

ist. Wir nehmen gegenteilig an, dass $k_t' + k_t'' = r-1$ ist. Dann würde $p't_r^h$ auch „verschwinden“, das heißt,

$$p'g = p'h + \sum_{j=2}^{r-1} \beta_j p' t_j^h = \sum_{i=1}^{n'-1} \alpha_i t_i^{p'} = q$$

mit $\operatorname{rang}(q) < \operatorname{rang}(p')$. Daher ist $g = \kappa q_{\ell-1}^{-1} \cdots q_{\ell'}^{-1}$ für ein ℓ' mit $1 \leq \ell' < \ell - 1$ und $\kappa \in \mathbb{K}$. Laut Annahme ist h vom Typ $(1,1)$, aber g ist vom Typ $(0,0)$ laut Satz 2.5.13. Daher ist $g = h - h_0$ für ein (nicht-verschwindendes) h_0 vom Typ $(1,1)$. Aber das würde der $\mathbb{K}$-linearen Unabhängigkeit der rechten Familie $(t_1^h, t_2^h, \ldots, t_r^h)$ widersprechen. Für eine Illustration siehe Beispiel 3.4.6. Schließlich haben wir, für h vom Typ $(1,1)$,

$$\begin{aligned}\operatorname{rang}(p'h) + \operatorname{rang}(h^{-1}p'') &= n' + r - 1 - (k_s' + k_t') + n'' + r - 2 - (k_s'' + k_t'') \\ &\geq n' + n'' + 2r - 3 - (r-2) - (r-2) \\ &= n' + n'' + 1 > \operatorname{rang}(p'p'') + 1.\end{aligned}$$

Wenn h vom Typ $(0,0)$ ist, dann ist h^{-1} vom Typ $(1,1)$ und t'' ist $\mathbb{K}$-linear unabhängig, daher können wir ähnliche Argumente verwenden. Wenn h vom Typ $(1,0)$ ist, haben die Systeme $\mathcal{A}'$ und $\mathcal{A}''$ die Dimensionen $n'+r-1$ beziehungsweise $n''+r-1$. Ihre linken Familien sind

$$\begin{aligned}s' &= (s_1^{p'}h, s_2^{p'}h, \ldots, s_{n'-1}^{p'}h, s_1^h, s_2^h, \ldots, s_r^h) \quad \text{bzw.} \\ s'' &= (\underbrace{h^{-1}p'', s_r^h h^{-1}p'', \ldots, s_2^h h^{-1}p''}_{r}, \underbrace{s_2^{p''}, \ldots, s_{n''}^{p''}}_{n''-1}).\end{aligned}$$

Mit ähnlichen Argumenten wie für den Fall h vom Typ $(1,1)$ —s' ist $\mathbb{K}$-linear unabhängig— bekommen wir $k_s'' \leq r-1$. Die rechten Familien sind

$$\begin{aligned}t' &= (\underbrace{t_1^{p'}, t_2^{p'}, \ldots, t_{n'-1}^{p'}, p't_1^h}_{n'}, \underbrace{p't_2^h, \ldots, p't_r^h}_{r-1}) \quad \text{bzw.} \\ t'' &= h^{-1}(\underbrace{t_r^h, \ldots, t_2^h, 1}_{r}, t_2^{p''}, \ldots, t_{n''}^{p''}).\end{aligned}$$

Nachdem weder h noch h^{-1} ein Polynom ist, können maximal $r-2$ Komponenten in t' oder t'' eliminiert werden, das heißt, $k'_t + k''_t \leq r-2$. Daher, für h vom Typ $(1,0)$, gilt

$$\begin{aligned}\operatorname{rang}(p'h) + \operatorname{rang}(h^{-1}p'') &= n' + r - 1 - (k'_s + k'_t) + n'' + r - 1 - (k''_s + k''_t) \\ &\geq n' + n'' + 2r - 2 - (r-1) - (r-2) \\ &= n' + n'' + 1 > \operatorname{rang}(p'p'') + 1.\end{aligned}$$

Wenn h vom Typ $(0,1)$ ist, dann ist die rechte Familie t'' (laut Lemma 3.2.6) $\mathbb{K}$-linear unabhängig. Mit ähnlichen Argumenten erhalten wir $k'_t \leq r-1$ und $k'_s + k''_s \leq 2$. Also haben wir schlussendlich gezeigt, dass $\operatorname{rang}(pf^{-1}) + \operatorname{rang}(f) > \operatorname{rang}(p) + 1$ ist. □

Lemma 3.4.8 ([Sch17a, Lemma 3.8]). *Sei $q \in H = \mathbb{K}\langle X\rangle^\bullet$ und $p \in \mathbb{H} = \mathbb{F}^\bullet$. Dann gilt: $p|_l^{\mathbb{F}} q$ impliziert $q = ph$ mit $p, h \in H$.*

Beweis. Für ein $m' < m$, seien $p = f_1 f_2 \cdots f_{m'}$ und $q = p f_{m'+1} \cdots f_m$ und $\sigma = (i_1, i_2, \ldots, i_m)$ eine Permutation sodass f_{i_k} für alle $k \in \{2, 3, \ldots, m\}$ ein äußerer Faktor von $f_{\sigma_{1,k}} f_{\sigma_{2,k}} \cdots f_{\sigma_{k,k}}$ ist. Wir müssen zeigen, dass $q_k = f_{\sigma_{1,k}} f_{\sigma_{2,k}} \cdots f_{\sigma_{k,k}} \in H$ und $f_{i_k} \in H$ für alle $k = 1, 2, \ldots, m$ ist, über Induktion nach k von m bis 2. Für $k = m$ gilt $q_m = q \in H$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass f_{i_k} ein *echter rechter* Faktor von q_k ist (für triviale Faktoren ist nichts zu zeigen), also

$$\begin{aligned}&\operatorname{rang}(q_k f_{i_k}^{-1}) + \operatorname{rang}(f_{i_k}) \leq 1 + \operatorname{rang}(q_k) \quad \text{und} \\ &\operatorname{rang}(f_{i_k}^{-1}) + \operatorname{rang}(f_{i_k} q_k^{-1}) \leq 1 + \operatorname{rang}(q_k^{-1}).\end{aligned}$$

Laut Annahme ist q_k ein Polynom und daher haben alle Faktorisierungen in Atome die gleiche Länge, sagen wir ℓ_k. Wir behaupten, dass $f_{i_k} = \kappa\, g_{\ell_0} \cdots g_{\ell_k}$ für eine Faktorisierung $q_k = g_1 g_2 \cdots g_{\ell_k}$, ein $\ell_0 \in \{1, 2, \ldots, \ell_k\}$ und $\kappa \in \mathbb{K}$ ist. Wir nehmen das Gegenteil an und wenden Lemma 3.4.7 mit $p = q_k$ und $f = f_{i_k}$ an, um den Widerspruch

$$\operatorname{rang}(q_k f_{i_k}^{-1}) + \operatorname{rang}(f_{i_k}) > 1 + \operatorname{rang}(q_k)$$

zu erhalten. Also ist $f_{i_k} \in H$ und $q_{k-1} = q_k f_{i_k}^{-1} \in H$. Insbesondere gilt $p = f_1 f_2 \cdots f'_m \in H$ und $h = f_{m'+1} f_{m'+2} \cdots f_m \in H$. □

Satz 3.4.9 (Teilbarkeitsäquivalenz [Sch17a, Satz 3.10]). *Seien $p, q \in H = \mathbb{K}\langle X\rangle^\bullet$. Dann $p|_l q$ (bzw. $p|_r q$) genau dann, wenn $p|_l^{\mathbb{F}} q$ (bzw. $p|_r^{\mathbb{F}} q$) in $\mathbb{F}$.*

Beweis. Gelte $p \mid_{\mathrm{l}} q$, das heißt, $q \in pH = \{ph \mid h \in H\}$. Wir zeigen, dass p ein linker Faktor von $q = ph$ ist. Mittels Lemma 3.4.5 (i) bekommen wir $\operatorname{rang}(q) = \operatorname{rang}(p) + \operatorname{rang}(p^{-1}q) - 1$ und laut (ii) bekommen wir $\operatorname{rang}(q^{-1}) + 1 = \operatorname{rang}(p^{-1}) + 1 + \operatorname{rang}(q^{-1}p)$, also ist p ein linker Faktor von $ph = q$ und damit ($m' = 1$ und $m = 2$ in Definition 3.4.4) $p|_{\mathrm{l}}^{\mathbb{F}} q$ (in $\mathbb{F}$). Umgekehrt müssen wir zeigen, dass $q \in pH$ ist. Aber das folgt direkt aus der Annahme $p|_{\mathrm{l}}^{\mathbb{F}} q$ und Lemma 3.4.8. □

Notation. Nachdem die linke (bzw. rechte) Teilbarkeit in $\mathbb{K}\langle X\rangle$ der in $\mathbb{F}$ entspricht, können wir die Notation vereinfachen, indem wir im folgenden $f \mid_{\mathrm{l}} g$ statt $f \mid_{\mathrm{l}}^{\mathbb{F}} g$ (bzw. $f \mid_{\mathrm{r}} g$ statt $f \mid_{\mathrm{r}}^{\mathbb{F}} g$) schreiben.

Definition 3.4.10 (Atome, Irreduzible Elemente [Sch17a, Definition 3.11])**.** Sei $\mathbb{H} = \mathbb{F}^{\bullet}$. Ein Element $f \in \mathbb{H} \setminus \mathbb{K}$, das heißt, eine nicht-triviale Einheit (in $\mathbb{F}$), heißt (verallgemeinertes) *Atom* (oder *irreduzibel*) wenn $f = g_1 g_2$ mit $g_1, g_2 \in \mathbb{H}$ und $g_1 \mid_{\mathrm{l}} f$ impliziert, dass entweder $g_1 \in \mathbb{K}^{\times}$ oder $g_2 \in \mathbb{K}^{\times}$ ist. Analog wie in Definition 3.1.3 wird die Menge der Atome in $\mathbb{F}$ mit $\mathbf{A}(\mathbb{F})$ bezeichnet.

Bemerkung. Man bemerke die zusätzliche Bedingung $g_1 \mid_{\mathrm{l}} f$ verglichen mit Definition 3.1.3. Sie ist *entscheidend*. Ohne ihr wäre $f = 1 - x = x \cdot (x^{-1} - 1)$ kein Atom. (Tatsächlich würde es überhaupt keine Atome geben.) Jedoch gilt $x \nmid_{\mathrm{l}} (1 - x)$ weil $\operatorname{rang}(x) + \operatorname{rang}\big(x^{-1}(1-x)\big) = 4 > 3 = 1 + \operatorname{rang}(1-x)$ ist.

Bemerkung. Ein reduzibles Polynom hat *ausschließlich* „polynomielle“ Teiler. Die Aussage von Lemma 3.4.8 ist stärker als die Voraussetzung für die Teilbarkeitsäquivalenz in Satz 3.4.9, in dem ein Teiler ein Polynom ist.

Proposition 3.4.11. *Ein Polynom ist ein Atom genau dann, wenn es ein verallgemeinertes Atom ist.*

Beweis. Wir müssen zeigen, dass $\mathbf{A}(\mathbb{K}\langle X\rangle) = \mathbf{A}(\mathbb{F}) \cap \mathbb{K}\langle X\rangle$ ist. Laut Lemma 3.4.8 gilt für ein Polynom p aber bereits: $g_1 \mid_{\mathrm{l}} p$ impliziert $p = g_1 g_2$ mit Polynomen g_1 und g_2. Nun sind beide Implikationen unmittelbar. □

Notation. Im folgenden verwenden wir „Atom“ als den allgemeinen Begriff und „polynomielles Atom“ um zu betonen, dass das Atom ein Element der freien assoziativen Algebra ist.

Bemerkung 3.4.12. Wie bereits in der Einleitung erwähnt (und vor Beispiel 3.4.2 illustriert), können (verallgemeinerte) Atome im freien Schiefkörper auch *multiplikativ* (aus Atomen) „erzeugt“ werden. Dieses Phänomen sollte jedenfalls genauer untersucht werden. Denn wenn man das besser versteht kann man möglicherweise das in Abschnitt 4.5 entwickelte Minimierungskonzept (zumindest teilweise) verbessern, indem man zwischen Addition und Multiplikation unterscheidet.

3.5 Minimale Faktormultiplikation

Bevor wir im folgenden Abschnitt die Korrespondenz von (linken unteren und rechten oberen) Nullblöcken in der Systemmatrix eines *minimalen* zulässigen linearen Systems und einer nicht-trivialen Faktorisierung beschreiben, befassen wir uns mit der Konstruktion eines *minimalen* ZLS $\mathcal{A} = (u, A, v) = (1, A, \lambda)$, $A = (a_{ij})$, für das Produkt von zwei (nicht-skalaren) Elementen f und g, die durch *minimale* zulässige lineare Systeme (der Dimension n_f bzw. n_g) gegeben sind.

Tatsächlich müssen wir uns hier auf den Fall beschränken, in dem f ein *linker Faktor* von fg (laut Definition 3.4.1) ist. Der allgemeine Fall kann nur algorithmisch

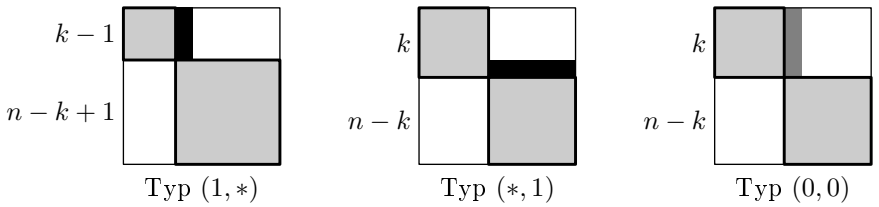


Abbildung 3.1: [Sch17a, Abbildung 1] Es gibt drei Typen der Faktorisierung eines Elementes $h = fg$ mit $\operatorname{rang}(h) = n$, $\operatorname{rang}(f) = k$ und $\operatorname{rang}(g) = n - k$ für Typ $(0, 0)$ beziehungsweise $\operatorname{rang}(g) = n - k + 1$ sonst. Diese Typen entsprechen denen der minimalen Faktormultiplikation. Für Typ $(1, *)$ und $(*, 1)$ muss die Kopplung *nicht-skalar* für *alle* Transformationen sein, die entsprechende Nullblöcke erzeugen.

gelöst werden, das heißt, gegebenenfalls müssen Blöcke eliminiert werden, siehe dazu Abschnitt 4.5. Entsprechend Satz 3.5.2 gibt es drei Fälle (siehe auch Abbildung 3.1):

Typ	Nullen links unten	„Kopplung“	Nullen rechts oben
$(1, *)$	$n_g \times (n_f - 1)$	$\exists\, 1 \leq i < n_f : a_{i,n_f} \notin \mathbb{K}$	$(n_f - 1) \times (n_g - 1)$
$(*, 1)$	$(n_g - 1) \times n_f$	$\exists\, 1 \leq j < n_g : a_{n_f,n_f+j} \notin \mathbb{K}$	$(n_f - 1) \times (n_g - 1)$
$(0, 0)$	$n_g \times n_f$	$\forall\, i = 1, \ldots, n_f : a_{i,n_f+1} \in \mathbb{K}$	$n_f \times (n_g - 1)$

Man beachte, dass die „Kopplungsbedingung“ für Typ $(1, *)$ und $(*, 1)$ für *jedes* (zulässig) transformierte System gelten muss, weil sonst diese beiden Typen einfach von Typ $(0, 0)$ „abgeleitet“ werden könnten. Siehe dazu auch Beispiel 3.6.2.

Um die Multiplikation „umzukehren“ müssen wir ein zulässiges lineares System entsprechend der Transformationen

$$(P, Q) = \left(\begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n-1} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ \alpha_{n-1,1} & \cdots & \alpha_{n-1,n-1} & 0 \\ \alpha_{n,1} & \cdots & \alpha_{n,n-1} & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \beta_{2,1} & \beta_{2,2} & \cdots & \beta_{n,2} \\ \vdots & \vdots & \cdots & \vdots \\ \beta_{n,1} & \beta_{n,2} & \cdots & \beta_{n,n} \end{bmatrix} \right) \tag{3.5.1}$$

mit Einträgen $\alpha_{ij}, \beta_{ij} \in \mathbb{K}$ umformen. Um die Invertierbarkeit zu garantieren, brauchen wir —im Unterschied zu den Transformationen (3.3.3) für die Polynomfaktorisierung— die zusätzlichen Bedingungen $\det(P) \neq 0$ und $\det(Q) \neq 0$.

Bemerkung. Die *minimale Polynommultiplikation* (Proposition 3.2.7) kann als Korollar des folgenden Satzes formuliert werden. Die Schwierigkeit des Beweises der ersteren (minimalen Multiplikation) ist in der Definition der *äußeren* Faktoren (Definition 3.4.1) versteckt. Um im allgemeinen zu testen, ob f ein linker Faktor von fg ist, benötigt man Techniken zum Minimieren von linearen Darstellungen. Dem ist Kapitel 4, im speziellen Abschnitt 4.5, gewidmet.

Satz 3.5.2 (Minimale Faktormultiplikation [Sch17a, Satz 4.2]). *Seien $f, g \in \mathbb{F} \setminus \mathbb{K}$ gegeben durch die* minimalen *zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f)$ der Dimension n_f beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g)$ der Dimension n_g. Sei $n = n_f + n_g$. Wenn f ein* linker Faktor *von fg ist, dann ist ein* minimales *ZLS für fg gegeben durch*

$$\mathcal{A} = \begin{cases} \textit{Proposition 2.3.6 mit } \dim \mathcal{A} = n-1, & \textit{wenn } 1 \in L(f), \\ \textit{Proposition 2.3.9 mit } \dim \mathcal{A} = n-1, & \textit{wenn } 1 \in R(g), \\ \textit{Proposition 2.3.1 mit } \dim \mathcal{A} = n, & \textit{wenn } 1 \notin R(g) \textit{ und } 1 \notin L(f) \textit{ ist.} \end{cases}$$

Beweis. Nachdem f ein linker Faktor von fg ist, gilt $\operatorname{rang}(f)+\operatorname{rang}(g) \leq \operatorname{rang}(fg)+1$ und damit

$$\operatorname{rang}(fg) \geq \operatorname{rang}(f) + \operatorname{rang}(g) - 1 = \dim \mathcal{A} \geq \operatorname{rang}(fg).$$

Also ist $\mathcal{A}$ minimal, wenn $1 \in R(g)$ oder $1 \in L(f)$ ist, das heißt, für die ersten beiden Fälle/Typen $(*, 1)$ und $(1, *)$. Für den letzten Fall/Typ $(0, 0)$ unterscheiden wir vier Teilfälle. Laut Satz 2.5.13 (Minimale Inverse) gilt

$$\operatorname{rang}(h^{-1}) = \begin{cases} \operatorname{rang}(h) - 1, & \text{wenn } h \text{ vom Typ } (1,1), \\ \operatorname{rang}(h), & \text{wenn } h \text{ vom Typ } (1,0) \text{ oder } (0,1), \text{ und} \\ \operatorname{rang}(h) + 1, & \text{wenn } h \text{ vom Typ } (0,0) \text{ ist.} \end{cases}$$

Nachdem f ein linker Faktor von fg ist, gilt also

$$\operatorname{rang}(g^{-1}f^{-1}) + 1 \geq \operatorname{rang}(g^{-1}) + \operatorname{rang}(f^{-1})$$

und daher —laut der minimalen Inversen angewandt auf der rechten Seite—

$$\operatorname{rang}(g^{-1}f^{-1})+1 \geq \begin{cases} \operatorname{rang}(g) + \operatorname{rang}(f), & \text{wenn } 1 \in R(f),\ 1 \in L(g), \\ \operatorname{rang}(g) + 1 + \operatorname{rang}(f), & \text{wenn } 1 \in R(f),\ 1 \notin L(g), \\ \operatorname{rang}(g) + \operatorname{rang}(f) + 1, & \text{wenn } 1 \notin R(f),\ 1 \in L(g), \text{ und} \\ \operatorname{rang}(g) + 1 + \operatorname{rang}(f) + 1, & \text{wenn } 1 \notin R(f),\ 1 \notin L(g) \text{ ist.} \end{cases}$$

Hier ist zu beachten, dass wir a priori nicht Minimalität von $\mathcal{A}$ für fg annehmen können, weil wir das zeigen müssen. Deswegen können wir nicht die minimale Inverse auf der linken Seite anwenden, weil wir nur wissen, dass (zum Beispiel) $1 \in L(g)$ laut Konstruktion $1 \in L(\mathcal{A})$ impliziert. Jedoch, unter Verwendung der minimalen Inversen, wissen wir —nachdem g vom Typ $(0, *)$ ist—, dass g^{-1} vom Typ $(1, 1)$ oder $(0, 1)$ und f^{-1} vom Typ $(1, 1)$ oder $(1, 0)$ ist. Also können wir einen der ersten beiden (bereits bewiesenen) Fälle verwenden um schließlich $\operatorname{rang}(fg) \geq \operatorname{rang}(f) + \operatorname{rang}(g)$ zu zeigen. □

Bemerkung. Sei $A = (a_{ij})$ die Systemmatrix des zulässigen linearen Systems von Satz 3.5.2 (Minimale Faktormultiplikation). Für Typ $(1, *)$ existiert ein $i \in$

$\{1, 2, \ldots, n_f - 1\}$ sodass a_{i,n_f} *nicht-skalar* ist. Für Typ $(*, 1)$ existiert ein $j \in \{n_f + 1, n_f + 2, \ldots, n\}$ sodass $a_{n_f,j}$ *nicht-skalar* ist. Und für Typ $(0, 0)$ sind die Einträge a_{i,n_f+1} *skalar* für $i \in \{1, 2, \ldots, n_f\}$. Wir beziehen uns darauf als *Kopplungsbedingungen*. Man beachte, dass es beim Typ $(1, *)$ (bzw. Typ $(*, 1)$) keine Transformation der Form (3.5.1) gibt, die die Nullblöcke respektiert und eine „skalare Kopplung“ ergibt, weil das der Minimalität von $\mathcal{A}_f$ (bzw. $\mathcal{A}_g$) widersprechen würde. Vergleiche dazu auch Beispiel 3.6.2.

3.6 Allgemeine Faktorisierung

Lemma 3.6.1 ([Sch17a, Lemma 4.3])**.** *Seien $h, f, g \in \mathbb{F}\backslash\mathbb{K}$ gegeben durch die* minimalen *zulässigen linearen Systeme $\mathcal{A} = (1, A, \lambda)$ der Dimension n, $\mathcal{A}_f = (u_f, A_f, v_f) = (1, A_f, \lambda_f)$ der Dimension n_f beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g) = (1, A_g, \lambda_g)$ der Dimension n_g sodass f ein* linker Faktor *von $h = fg$ ist.*

Typ $(1, *)$: Wenn f vom Typ $(*, 1)$ ist, dann existiert eine zulässige Transformation (P, Q) der Form (3.5.1) sodass $PAQ = (a'_{i,j})$

- *links unten einen Nullblock der Größe $n_g \times (n_f - 1)$ und*
- *rechts oben einen der Größe $(n_f - 1) \times (n_g - 1)$ hat, und*
- *es existiert ein $i \in \{1, 2, \ldots, n_f - 1\}$, sodass a'_{i,n_f}* nicht-skalar *ist.*

Typ $(*, 1)$: Wenn g vom Typ $(1, *)$ ist, dann existiert eine zulässige Transformation (P, Q) der Form (3.5.1) sodass $PAQ = (a'_{i,j})$

- *links unten einen Nullblock der Größe $(n_g - 1) \times n_f$ und*
- *rechts oben einen der Größe $(n_f - 1) \times (n_g - 1)$ hat, und*
- *es existiert ein $j \in \{n_f + 1, n_f + 2, \ldots, n\}$, sodass $a'_{n_f,j}$* nicht-skalar *ist.*

Typ $(0, 0)$: Wenn f vom Typ $(*, 0)$ und g vom Typ $(0, *)$ ist, dann existiert eine zulässige Transformation (P, Q) der Form (3.5.1) sodass $PAQ = (a'_{i,j})$

- *links unten einen Nullblock der Größe $n_g \times n_f$ und*
- *rechts oben einen der Größe $n_f \times (n_g - 1)$ hat, und*
- *$a'_{i,n_f+1} \in \mathbb{K}$ für alle $i \in \{1, 2, \ldots, n_f\}$ ist.*

Beweis. Sei $\mathcal{A}' = (u', A', v') = (1, A', \lambda')$ das jeweilige laut Satz 3.5.2 aus $\frac{\lambda_g}{\lambda}\mathcal{A}_f$ und $\frac{\lambda}{\lambda_g}\mathcal{A}_g$ konstruierte (minimale) ZLS für $h = fg$. Die Systemmatrix $A' = (a'_{ij})$ hat — laut Konstruktion— entsprechende Nullblöcke (links unten und rechts oben) und —für Typ $(0, 0)$— skalare Einträge a'_{i,n_f+1} für $i \in \{1, 2, \ldots, n_f\}$. Nachdem beide Systeme $\mathcal{A}$ und $\mathcal{A}'$ für h minimal sind, existiert laut Satz 2.1.6 eine zulässige Transformation

(P, Q) sodass $P\mathcal{A}Q = \mathcal{A}'$ ist. Die rechte Seite $Pv = v'$ ändert sich nicht, also ist (P, Q) von der Form (3.5.1). Die Kopplungsbedingungen sind wegen der Konstruktion der minimalen Multiplikation erfüllt. □

Beispiel 3.6.2. Sei $h = x^{-1}zy^{-1}x^{-1}$ gegeben durch das *minimale* zulässige lineare System

$$\begin{bmatrix} x & -z & . \\ . & y & -1 \\ . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}.$$

Die Multiplikation des Typs $(0, 1)$ für $n_f = n_g = 2$ verletzt die Kopplungsbedingung, weil dadurch ein *nicht-minimales* ZLS für g in $h = fg$ „erzeugt“ würde.

Lemma 3.6.3 (Faktorisierung Typ $(1, *)$ [Sch17a, Lemma 4.5])**.** *Sei $h = fg \in \mathbb{F} \setminus \mathbb{K}$ gegeben durch das* minimale *zulässige lineare System $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ der Dimension $n \geq 2$. Fixiere ein $1 < k \leq n$. Angenommen, A habe einen Nullblock der Größe $(n-k+1) \times (k-1)$ links unten und einen der Größe $(k-1) \times (n-k)$ rechts oben. Für eine Transformation (P, Q) bezeichne $a'_{i,j}$ die Einträge der transformierten Systemmatrix PAQ. Wenn für* jede *Transformation (P, Q) der Form (3.5.1), die diese Nullblöcke respektiert, ein $i \in \{1, 2, \ldots, k-1\}$ existiert, sodass $a'_{i,k}$* nicht-skalar *ist, dann ist f ein* linker Faktor *vom Typ $(*, 1)$ von h mit* $\operatorname{rang}(f) = k$ *und* $\operatorname{rang}(g) = n - k + 1$.

Beweis. Laut Annahme hat $\mathcal{A}$ die Blockform

$$\begin{bmatrix} A_{1,1} & A_{1,2} & . \\ . & A_{2,2} & A_{2,3} \\ . & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} s_{\underline{1}} \\ s_k \\ s_{\underline{3}} \end{bmatrix} = \begin{bmatrix} . \\ . \\ v_{\underline{3}} \end{bmatrix}$$

mit quadratischen Diagonalblöcken $A_{1,1}$, $A_{2,2}$ und $A_{3,3}$ der Größen $k-1$, 1 beziehungsweise $n-k-1$. Wir duplizieren den Eintrag s_k in der linken Familie, indem wir eine Zeile (und eine Spalte) einfügen um das folgende ZLS der Dimension $n+1$ zu erhalten:

$$\begin{bmatrix} A_{1,1} & A_{1,2} & 0 & . \\ 0 & 1 & -1 & 0 \\ . & 0 & A_{2,2} & A_{2,3} \\ . & 0 & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} s_{\underline{1}} \\ s_k \\ s_k \\ s_{\underline{3}} \end{bmatrix} = \begin{bmatrix} . \\ . \\ . \\ v_{\underline{3}} \end{bmatrix},$$

das heißt, die Konstruktion laut Proposition 2.3.6 „umdrehen“. Die Teilsysteme der Dimension k und $n-k+1$ sind *minimal* für f (wegen der Kopplungsbedingung) beziehungsweise $g = \mu s_k$, denn sonst könnte man ein ZLS für h der Dimension $n' < n$ konstruieren, was der Minimalität von $\mathcal{A}$ widerspräche. Klarerweise ist $1 \in L(f)$. Laut Konstruktion gilt $\operatorname{rang}(f) + \operatorname{rang}(g) = \operatorname{rang}(h) + 1$, also müssen wir nur zeigen, dass auch $\operatorname{rang}(g^{-1}) + \operatorname{rang}(f^{-1}) \leq \operatorname{rang}(h^{-1}) + 1$ gilt, damit f ein linker Faktor von h ist. Dafür unterscheiden wir vier Fälle (wie in der minimalen Faktormultiplikation)

und wenden die minimale Inverse an. Wenn h vom Typ $(1,1)$ ist, dann ist f vom Typ $(1,1)$ und g vom Typ $(*,1)$. Also gilt $\operatorname{rang}(g^{-1}) \leq \operatorname{rang}(g)$ und wir erhalten $\operatorname{rang}(g^{-1}) + \operatorname{rang}(f^{-1}) \leq \operatorname{rang}(g) + \operatorname{rang}(f) - 1 = \operatorname{rang}(h^{-1}) + 1$. Die anderen Fälle sind genau so einfach. □

Lemma 3.6.4 (Faktorisierung Typ $(*,1)$ [Sch17a, Lemma 4.6])**.** *Sei $h = fg \in \mathbb{F} \setminus \mathbb{K}$ gegeben durch das* minimale *zulässige lineare System $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ der Dimension $n \geq 2$. Fixiere ein $1 \leq k < n$. Angenommen, A habe einen Nullblock der Größe $(n-k)\times k$ links unten und einen der Größe $(k-1)\times(n-k)$ rechts oben. Für eine Transformation (P,Q) bezeichne a'_{ij} die Einträge der transformierten Systemmatrix PAQ. Wenn für* jede *Transformation (P,Q) der Form (3.5.1), die diese Nullblöcke respektiert, ein $j \in \{k+1, k+2, \ldots, n\}$ existiert, sodass $a'_{k,j}$* nicht-skalar *ist, dann ist f ein* linker Faktor *vom Typ $(*,1)$ von h mit* $\operatorname{rang}(f) = k$ *und* $\operatorname{rang}(g) = n - k + 1$.

Beweis. Laut Annahme hat $\mathcal{A}$ die Blockform

$$\begin{bmatrix} u_{\underline{1}} & . & . \end{bmatrix} = \begin{bmatrix} t_{\underline{1}} & t_k & t_{\underline{3}} \end{bmatrix} \begin{bmatrix} A_{1,1} & A_{1,2} & . \\ A_{2,1} & A_{2,2} & A_{2,3} \\ . & . & A_{3,3} \end{bmatrix}$$

mit quadratischen Diagonalblöcken $A_{1,1}$, $A_{2,2}$ und $A_{3,3}$ der Größen $k-1$, 1 beziehungsweise $n-k$. Wir duplizieren den Eintrag t_k in der rechten Familie, indem wir eine Spalte (und eine Zeile) einfügen, um das folgende ZLS der Dimension $n+1$ zu erhalten:

$$\begin{bmatrix} u_{\underline{1}} & . & . & . \end{bmatrix} = \begin{bmatrix} t_{\underline{1}} & t_k & t_k & t_{\underline{3}} \end{bmatrix} \begin{bmatrix} A_{1,1} & A_{1,2} & 0 & . \\ A_{2,1} & A_{2,2} & -1 & 0 \\ 0 & 0 & 1 & A_{2,3} \\ . & . & 0 & A_{3,3} \end{bmatrix},$$

das heißt, die Konstruktion laut Proposition 2.3.9 „umdrehen“. Die Teilsysteme der Dimension k und $n-k+1$ sind *minimal* für $f = \mu t_k$ beziehungsweise g (wegen der Kopplungsbedingung), denn sonst könnte man ein ZLS für h der Dimension $n' < n$ konstruieren, was der Minimalität von $\mathcal{A}$ widerspräche. Klarerweise ist $1 \in R(g)$. Ähnlich wie im Beweis von Lemma 3.6.3 zeigt man, dass f ein linker Faktor von $h = fg$ ist. □

Lemma 3.6.5 (Faktorisierung Typ $(0,0)$ [Sch17a, Lemma 4.7])**.** *Sei $h = fg \in \mathbb{F} \setminus \mathbb{K}$ gegeben durch das* minimale *zulässige lineare System $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ der Dimension $n \geq 2$. Fixiere ein $1 \leq k < n$. Angenommen, $A = (a_{ij})$ habe einen Nullblock der Größe $(n-k)\times k$ links unten und einen der Größe $k \times (n-k-1)$ rechts oben. Wenn $a_{i,k+1} \in \mathbb{K}$ für alle $i \in \{1, 2, \ldots, k\}$ gilt, dann ist f ein* linker Faktor *vom Typ $(*,0)$ von h mit* $\operatorname{rang}(f) = k$ *und g ist vom Typ $(0,*)$ mit* $\operatorname{rang}(g) = n - k$.

Beweis. Wir erhalten die Teilsysteme $\mathcal{A}_f$ (für f) und $\mathcal{A}_g$ (für g) direkt aus der Konstruktion laut Proposition 2.3.1 (Multiplikation). Wäre eines der Teilsysteme nicht minimal, würde das der angenommenen Minimalität von $\mathcal{A}$ widersprechen. So, wie

$1 \in L(f)$ oder $1 \in R(g)$ und die Anwendung der Multiplikation Typ $(*, 1)$ beziehungsweise $(1, *)$ der Minimalität widersprechen würde. Die Argumente, um zu zeigen, dass f ein linker Faktor von $h = fg$ ist, sind ähnlich zu denen im Beweis von Lemma 3.6.3. □

Satz 3.6.6 (Freie Faktorisierung [Sch17a, Satz 4.8])**.** *Sei $h \in \mathbb{F}$ mit $n = \operatorname{rang}(h) \geq 2$ gegeben durch das* minimale *zulässige lineare System $\mathcal{A} = (u, A, v)$. Dann hat h einen echten linken Faktor f mit $\operatorname{rang}(f) = k$ genau dann, wenn es eine zulässige Transformation (P, Q) der Form (3.5.1) gibt, sodass PAQ vom „Typ“ $(1, *)$, $(*, 1)$ oder $(0, 0)$ ist, wie in Abbildung 3.1.*

Beweis. Einen echten linken Faktor vom Rang k angenommen, liefert Lemma 3.6.1 die entsprechende Transformation. Umgekehrt, eine entsprechende Transformation angenommen, erhalten wir einen echten linken Faktor vom Rang k laut Lemma 3.6.3 für Typ $(1, *)$, laut Lemma 3.6.4 für Typ $(*, 1)$ und laut Lemma 3.6.5 für Typ $(0, 0)$. □

Einen Rang eines möglichen linken Faktors in $\overline{\mathbb{K}}\langle\langle X\rangle\rangle$ fixierend, kann wiederum eine Variante von Satz 4.2.1 [CR99, Satz 4.1] verwendet werden, um Nullblöcke (links unten und rechts oben in der Systemmatrix) der entsprechenden Größe (in Abhängigkeit des Typs der Faktorisierung) zu erzeugen. Hier haben wir den *kommutativen* Polynomring

$$\begin{aligned}\mathbb{K}[\alpha, \beta] = \mathbb{K}[&\alpha_{1,1}, \ldots, \alpha_{1,n-1}, \alpha_{2,1}, \ldots, \alpha_{2,n-1}, \ldots, \alpha_{n,1}, \ldots, \alpha_{n,n-1},\\ &\beta_{2,1}, \ldots, \beta_{2,n}, \beta_{3,1}, \ldots, \beta_{3,n}, \ldots, \beta_{n,1}, \ldots, \beta_{n,n}].\end{aligned}$$

Um die Invertierbarkeit der Transformationsmatrizen P und Q sicherzustellen, brauchen wir —im Gegensatz zu Proposition 3.3.7— jedenfalls eine Bedingung der Art $\det(P) = 1$ beziehungsweise $\det(Q) = 1$. Die Kopplungsbedingungen für Typ $(0, 0)$ müssen direkt implementiert werden indem man den Gleichungen, die das Ideal erzeugen, die Koeffizienten entsprechend $x \in X$ für den „Kopplungsvektor“ hinzufügt. Für Typ $(1, *)$ und $(*, 1)$ kann man zuerst auf eine „skalare“ Kopplung testen. Wenn es hier keine Lösung gibt, kann man versuchen eine entsprechende Transformation für die Nullblöcke alleine zu finden.

3.7 Beispiele Faktorisierung

Beispiel 3.7.1 (Polynomfaktorisierung [Sch17c, Abschnitt 4])**.** Gegeben seien $p = x(1 - yx)(3 - yx)$ und $q = (xy - 1)(xy - 3)x$ in $\mathbb{Q}\langle X\rangle$. Unter Verwendung von Begleitsystemen (Definition 2.2.3) für die Faktoren erhält man über die minimale Poly-

nommultiplikation (Proposition 3.2.7) das *minimale* ZLS (für p):

$$\begin{bmatrix} 1 & -x & . & . & . & . \\ . & 1 & -y & -1 & . & . \\ . & . & 1 & x & . & . \\ . & . & . & 1 & -y & -3 \\ . & . & . & . & 1 & x \\ . & . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} x(1-yx)(3-yx) \\ (1-yx)(3-yx) \\ -x(3-yx) \\ 3-yx \\ -x \\ 1 \end{bmatrix} \tag{3.7.2}$$

Klarerweise gilt $p = xyxyx - 4xyx + 3x = q$. Nun betrachten wir das folgende *minimale* ZLS (rechtes Begleitsystem) für $p = xyxyx + (3x - 4xyx)$:

$$\begin{bmatrix} 1 & -x & . & . & . & -x \\ . & 1 & -y & . & . & . \\ . & . & 1 & -x & . & \frac{4}{3}x \\ . & . & . & 1 & -y & . \\ . & . & . & . & 1 & -\frac{1}{3}x \\ . & . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ 3 \end{bmatrix}.$$

Wir versuchen, einen oberen rechten Nullblock der Größe 3×2 zu erzeugen. Das entspräche einer Faktorisierung in $p = q_1 q_2$ mit $\operatorname{rang}(q_1) = 4$ und $\operatorname{rang}(q_2) = 3$. Dafür wenden wir die (zulässige) Transformation (P, Q) direkt auf die Koeffizientenmatrizen A_0, A_x und A_y in $A = A_0 \otimes 1 + A_x \otimes x + A_y \otimes y$ an, um die Gleichungen zu bekommen. Für y erhalten wir

$$PA_yQ = P \begin{bmatrix} . & . & . & . & . & . \\ & . & -1 & . & . & . \\ & & . & . & . & . \\ & & & . & -1 & . \\ & & & & . & . \\ & & & & & . \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ & 1 & \beta_{2,3} & \beta_{2,4} & \beta_{2,5} & \beta_{2,6} \\ & & 1 & \beta_{3,4} & \beta_{3,5} & \beta_{3,6} \\ & & & 1 & \beta_{4,5} & \beta_{4,6} \\ & & & & 1 & \beta_{5,6} \\ & & & & & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} & 0 \\ & 1 & \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} & 0 \\ & & 1 & \alpha_{3,4} & \alpha_{3,5} & 0 \\ & & & 1 & \alpha_{4,5} & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} . & . & . & . & . & . \\ & . & -1 & -\beta_{3,4} & -\beta_{3,5} & -\beta_{3,6} \\ & & . & . & . & . \\ & & & . & -1 & -\beta_{5,6} \\ & & & & . & . \\ & & & & & . \end{bmatrix}$$

$$= \begin{bmatrix} . & . & -\alpha_{1,2} & -\alpha_{1,2}\beta_{3,4} & -\alpha_{1,2}\beta_{3,5} - \alpha_{1,4} & -\alpha_{1,4}\beta_{5,6} - \alpha_{1,2}\beta_{3,6} \\ & . & -1 & -\beta_{3,4} & -\beta_{3,5} - \alpha_{2,4} & -\alpha_{2,4}\beta_{5,6} - \beta_{3,6} \\ & & . & . & -\alpha_{3,4} & -\alpha_{3,4}\beta_{5,6} \\ & & & . & -1 & -\beta_{5,6} \\ & & & & . & . \\ & & & & & . \end{bmatrix}.$$

Davon nehmen wir den oberen rechten 3×2 Block. Somit haben wir die folgenden 6 Gleichungen für y:

$$\begin{bmatrix} \alpha_{1,2}\beta_{3,5} + \alpha_{1,4} & \alpha_{1,4}\beta_{5,6} + \alpha_{1,2}\beta_{3,6} \\ \beta_{3,5} + \alpha_{2,4} & \alpha_{2,4}\beta_{5,6} + \beta_{3,6} \\ \alpha_{3,4} & \alpha_{3,4}\beta_{5,6} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Mit einer ähnlichen Vorgehensweise erhalten wir die folgenden 6 Gleichungen für x:

$$\begin{bmatrix} \alpha_{1,3}\beta_{4,5} + \beta_{2,5} & \alpha_{1,3}\beta_{4,6} + \beta_{2,6} + \frac{1}{3}\alpha_{1,5} - \frac{4}{3}\alpha_{1,3} + 1 \\ \alpha_{2,3}\beta_{4,5} & \alpha_{2,3}\beta_{4,6} + \frac{1}{3}\alpha_{2,5} - \frac{4}{3}\alpha_{2,3} \\ \beta_{4,5} & \beta_{4,6} + \frac{1}{3}a_{3,5} - \frac{4}{3} \end{bmatrix} = 0^{3\times 2}.$$

Und schließlich die für 1 (ohne der Terme, die $\alpha_{3,4} = \beta_{4,5} = 0$ enthalten):

$$\begin{bmatrix} \alpha_{1,3}\beta_{3,5} + \alpha_{1,2}\beta_{2,5} + \alpha_{1,5} & \alpha_{1,5}\beta_{5,6} + \alpha_{1,4}\beta_{4,6} + \alpha_{1,3}\beta_{3,6} + \alpha_{1,2}\beta_{2,6} \\ \alpha_{2,3}\beta_{3,5} + \beta_{2,5} + \alpha_{2,5} & \alpha_{2,5}\beta_{5,6} + \alpha_{2,4}\beta_{4,6} + \alpha_{2,3}\beta_{3,6} + \beta_{2,6} \\ \beta_{3,5} + \alpha_{3,5} & \alpha_{3,5}\beta_{5,6} + \beta_{3,6} \end{bmatrix} = 0^{3\times 2}.$$

Eine Gröbner-Basis für das von diesen 18 Gleichungen erzeugte Ideal (berechnet via FriCAS [Fri18], unter Verwendung der *lexikographischen Termordnung*) ist

$$\begin{aligned} &(\alpha_{1,2} - \alpha_{1,4}\beta_{4,6}, \quad \alpha_{1,3} - \alpha_{1,5}\beta_{4,6}, \quad \alpha_{2,3} - \alpha_{2,5}\beta_{4,6}, \\ &\alpha_{2,4} + 3\beta_{4,6} - 4, \quad \alpha_{3,4}, \quad \alpha_{3,5} + 3\beta_{4,6} - 4, \\ &\beta_{2,5}, \quad \beta_{2,6} + 1, \quad \beta_{3,5} - 3\beta_{4,6} + 4, \\ &\beta_{3,6} - 3\beta_{4,6}\beta_{5,6} + 4\beta_{5,6}, \quad \beta_{4,5}, \quad \beta_{4,6}^2 - \tfrac{4}{3}\beta_{4,6} + \tfrac{1}{3}). \end{aligned}$$

Der letzte Erzeuger ist $(\beta_{4,6} - 1)(\beta_{4,6} - \frac{1}{3})$, also ist $\beta_{4,6} \in \{1, \frac{1}{3}\}$. Wir setzen unsere Berechnung mit dem Fall $\beta_{4,6} = 1$ fort. Damit bekommen wir $\alpha_{2,4} = \alpha_{3,5} = 1$, $\beta_{2,6} = \beta_{3,5} = -1$ und wählen $\alpha_{1,2} = \alpha_{1,4} = 0$, $\alpha_{1,3} = \alpha_{1,5} = 3$, $\alpha_{2,3} = \alpha_{2,5} = 0$, $\beta_{3,6} = -\beta_{5,6} = 0$. Da die Variablen $\alpha_{4,5}$, $\beta_{2,3}$, $\beta_{2,4}$ und $\beta_{3,4}$ nicht vorkommen, können wir sie auf Null setzen. Damit ist eine mögliche Transformation

$$(P, Q) = \left(\begin{bmatrix} 1 & 0 & 3 & 0 & 3 & . \\ & 1 & 0 & 1 & 0 & . \\ & & 1 & 0 & 1 & . \\ & & & 1 & 0 & . \\ & & & & 1 & . \\ & & & & & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . & . & . & . \\ & 1 & 0 & 0 & 0 & -1 \\ & & 1 & 0 & -1 & 0 \\ & & & 1 & 0 & 1 \\ & & & & 1 & -0 \\ & & & & & 1 \end{bmatrix} \right).$$

Durch deren Anwendung erhält man das zulässige lineare System $\mathcal{A}' = P\mathcal{A}Q$

$$\begin{bmatrix} 1 & -x & 3 & -3x & 0 & 0 \\ . & 1 & -y & 1 & 0 & 0 \\ . & . & 1 & -x & 0 & 0 \\ . & . & . & 1 & -y & 1 \\ . & . & . & . & 1 & -\frac{1}{3}x \\ . & . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ 3 \end{bmatrix},$$

das —so wie das ZLS (3.7.2)— rechts oben einen 3×2 Nullblock hat. Daher ist $\mathcal{A}'$ das (minimale) Produkt von

$$\begin{bmatrix} 1 & -x & 3 & -3x \\ . & 1 & -y & 1 \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xyx - x \\ yx - 1 \\ x \\ 1 \end{bmatrix} \tag{3.7.3}$$

und

$$\begin{bmatrix} 1 & -y & 1 \\ . & 1 & -\frac{1}{3}x \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} yx - 3 \\ x \\ 3 \end{bmatrix}. \tag{3.7.4}$$

Also ist $p = (xyx-x)(yx-3)$. Der erste Faktor ist *nicht* atomar, weil wir (polynomiell zulässig) rechts oben einen Nullblock entweder der Größe 1×2 (indem wir 3-mal Zeile 3 von Zeile 1 subtrahieren) oder der Größe 2×1 (indem wir Spalte 2 von Spalte 4 und 2-mal Zeile 3 von Zeile 1 subtrahieren) in der Systemmatrix von (3.7.3) erzeugen können. Andererseits genügt ein kurzer Blick auf das ZLS (3.7.4) um festzustellen, dass der rechte Faktor irreduzibel ist.

Bemerkung 3.7.5. Bevor man diesen allgemeinen Weg (mit dem nicht-linearen Gleichungssystem) beschreitet, sollte man jedenfalls mit *linearen* Methoden versuchen, einen entsprechenden Nullblock rechts oben zu erzeugen. Die Grundidee entspricht der von Lemma 4.2.2 (für das linearisierte Wortproblem). Entscheidend ist, dass sich die Zeilen- und Spaltentransformationen *nicht* „überlappen“. Für den Nullblock der Größe 3×2 im vorherigen Beispiel hieße das, entweder die Zeilen $\{4, 5\}$ und die Spalten $\{2, 3\}$ oder die Zeilen $\{5\}$ und die Spalten $\{2, 3, 4\}$ zu verwenden.

Beispiel 3.7.6 ([Sch17a, Beispiel 4.9])**.** Sei $f \in \mathbb{Q}\langle\langle X\rangle\rangle$ gegeben durch das *minimale* ZLS $\mathcal{A} = (u, A, v)$,

$$\begin{bmatrix} -1 & . & x & -1 \\ 1+x & x & -1 & . \\ y & 1 & x & -1 \\ x & . & -2 & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Nun versuchen wir einen *linken Faktor* f_1 vom Typ $(*, 0)$ mit Rang $n_1 = 2$ und einen *rechten Faktor* f_2 vom Typ $(0, *)$ mit Rang $n_2 = 2$ zu finden, das heißt, die minimale Faktormultiplikation Typ $(0, 0)$ „umzukehren“. Dazu brauchen wir eine *invertierbare* Transformation (P, Q) der Form (3.5.1) sodass $PAQ = (a'_{i,j})$ einen Nullblock der Größe 2×2 links unten und einen der Größe 2×1 rechts oben hat und $a'_{1,3}, a'_{2,3} \in \mathbb{K}$ ist. Zusätzlich zu $\det(P) = 1$ und $\det(Q) = 1$ haben wir $12 + 6 + 4$ Gleichungen. Eine Gröbner-Basis für das von diesen 24 Gleichungen erzeugte Ideal (berechnet via

FriCAS [Fri18], unter Verwendung der *lexikographischen Termordnung*) ist

$$\begin{aligned}
&(\alpha_{1,1}+\alpha_{1,2}\beta_{2,2}\beta_{3,3}\beta_{3,4}+\alpha_{1,3},\quad \alpha_{1,2}\alpha_{2,3}\alpha_{3,1}-\alpha_{1,3}\alpha_{2,2}\alpha_{3,1}-1,\\
&\alpha_{2,1}+\alpha_{2,2}\beta_{2,2}\beta_{3,3}\beta_{3,4}+\alpha_{2,3},\quad \alpha_{3,2},\quad \alpha_{3,3},\quad \alpha_{4,2},\quad \alpha_{4,3},\\
&\beta_{2,2}\beta_{3,3}^2\beta_{3,4}-\beta_{2,3},\quad \beta_{2,2}\beta_{3,3}\beta_{4,4}-\beta_{2,2}\beta_{3,4}\beta_{4,3}-1,\quad \beta_{2,3}^2,\quad \beta_{2,3}\beta_{3,4},\\
&\beta_{2,3}\beta_{4,4}-\beta_{3,3}\beta_{3,4},\quad \beta_{2,4},\quad \beta_{3,1},\quad \beta_{3,2},\quad \beta_{3,4}^2,\quad \beta_{4,1}+1,\quad \beta_{4,2}).
\end{aligned}$$

Nachdem $\beta_{3,4}=0$ ist, ist die Transformation (P,Q) von der Form

$$(P,Q)=\left(\begin{bmatrix}\alpha_{1,1}&\alpha_{1,2}&-\alpha_{1,1}&.\\ \alpha_{2,1}&\alpha_{2,2}&-\alpha_{2,1}&.\\ \alpha_{3,1}&0&0&.\\ \alpha_{4,1}&0&0&1\end{bmatrix},\begin{bmatrix}1&.&.&.\\ \beta_{2,1}&\beta_{2,2}&0&0\\ 0&0&\beta_{3,3}&0\\ -1&0&\beta_{4,3}&\beta_{4,4}\end{bmatrix}\right)$$

mit einer Lösung über $\mathbb{Q}$:

$$(P,Q)=\left(\begin{bmatrix}2&0&-2&.\\ 0&1&0&.\\ \frac{1}{2}&0&0&.\\ 0&0&0&1\end{bmatrix},\begin{bmatrix}1&.&.&.\\ 0&1&0&0\\ 0&0&1&0\\ -1&0&0&1\end{bmatrix}\right).$$

Das transformierte System $P\mathcal{A}Q$ ist

$$\begin{bmatrix}-2-2y&-2&.&0\\ 1+x&x&-1&0\\ 0&0&\frac{1}{2}x&-\frac{1}{2}\\ 0&0&-2&x\end{bmatrix}s=\begin{bmatrix}.\\.\\.\\1\end{bmatrix},$$

das heißt, $f=f_1f_2$, mit $f_1=(1-xy)^{-1}$ und $f_2=(x^2-2)^{-1}$, was man einfach sieht, wenn man die minimale Inverse auf die zwei Teilsysteme der Dimension $n_1=n_2=2$ anwendet. Beide Faktoren f_1 und f_2 sind Atome. Über $\overline{\mathbb{C}}(\langle X\rangle)$ zerfällt der zweite Faktor f_2 in $(x-\sqrt{2})^{-1}$ und $(x+\sqrt{2})^{-1}$.

Bemerkung 3.7.7. Bevor man mit „Gewalt“ versucht, eine (zulässige) Transformation (P,Q), zum Beispiel für die Multiplikation vom Typ $(0,0)$ und zwei Systemen der Dimension 2, zu finden, kann man einfach herausfinden, dass f *regulär* ist. Wenn das System nicht in eine polynomielle Form gebracht werden kann, das heißt f kein Polynom ist, könnte man prüfen, ob f^{-1} ein Polynom ist. Gegebenenfalls kann man dann die Techniken aus der Polynomfaktorisierung anwenden. Siehe dazu insbesondere Bemerkung 3.7.5.

Bemerkung. Um eine Lösung (von polynomiellen Gleichungssystemen) im allgemeinen (systematischer) zu finden, kann die Primärzerlegung von Idealen verwendet werden, siehe dazu zum Beispiel [Stu02, Kapitel 5], [CLO15, Abschnitt 4.8] und/oder [Coh03a, Abschnitt 10.8]. Für eine umfassende Diskussion empfiehlt sich [Mor05].

Bemerkung 3.7.8. Wenn die Faktorisierung an der Berechnung einer Gröbner-Basis scheitert, kann man versuchen, mehrstufig vorzugehen, das heißt, zuerst einen Nullblock links unten zu erzeugen und dann die (Suche nach) Transformationen auf „Blockfaktorisierungstransformationen“ (siehe Abbildung 4.1) einschränken. Das gilt insbesondere dann, wenn das ZLS bereits eine „passende“ Struktur (links unten) aufweist. Unter welchen Voraussetzungen man auf diese Art eine Aussage darüber bekommt, ob es nicht-triviale Faktoren gibt, wenn man keine entsprechende Transformation (bei gegebener Blockstruktur) findet, wäre noch zu klären.

Kapitel 4

Minimieren

Die Grundidee der Minimierung (einer linearen Darstellung) mit den linken und rechten Minimierungsschritten ist denkbar einfach. Wenn die Blockstruktur gröber wird und das „Hinschauen“ nicht mehr reicht, können die Zeilen- und Spaltentransformationen über ein *lineares Gleichungssystem* ermittelt werden. Das ist im wesentlichen der Inhalt von Abschnitt 4.2 (Wortproblem) dem letztlich die gesamte Theorie (hier) entsprungen ist. Meine naive Idee war, „lokale“ Wortprobleme zu lösen. Die vielen Fragen, die dabei entstanden, haben dann unter anderem zur Faktorisierungstheorie geführt ...

Aber *wann*, das heißt, unter welchen Voraussetzungen, ist ein zulässiges lineares System (konstruiert aus zwei *minimalen* laut Proposition 2.3.1) minimal? Wenn es keine linken und rechten —mit linearen Techniken durchzuführende— Minimierungsschritte mehr gibt? Reicht es aus, *eine* „feinste“ Struktur zu finden, sodass die Systemmatrix eine obere rechte Block-Dreiecksmatrix mit einer maximalen Anzahl an (quadratischen) Diagonalblöcken ist?

Für Polynome, das heißt, polynomielle zulässige lineare Systeme, lässt sich das verhältnismäßig einfach in einem Algorithmus beschreiben. Darum geht es im Abschnitt 4.3, in dem viele der bis jetzt manuell durchgeführten Schritte formalisiert werden. Kennt man „alle“ Faktorisierungen eines Polynoms, so kennt man auch alle „feinsten“ Pivotblock-Strukturen der *minimalen* zulässigen linearen Systeme ihres Inversen und kann so „einfach“ weiter rechnen, weil man weiterhin (verhältnismäßig) einfach minimieren kann.

Bereits zu Beginn des Kapitels 2 (Rechnen) haben wir uns mit den Anforderungen an die Konstruktion eines ZLS für die Inverse eines Elementes beschäftigt. Im Abschnitt 4.4 werden wir den Zusammenhang zwischen einer Faktorisierung und der Verfeinerung von Pivotblöcken im System der Inversen etwas genauer untersuchen und die Vorgehensweise für letzteres beschreiben.

Dieses „einfach“ zu präzisieren wird ein zentraler Teil von Abschnitt 4.5 sein. Zwar ist das Grundprinzip der allgemeinen Minimierung ähnlich der von polynomiellen zulässigen linearen Systemen, doch die Fragestellungen rund um die „Minimalität“

von Pivotblöcken ist wesentlich subtiler. Die zentrale Frage ist die einer *hinreichenden* Bedingung für die Minimierung mit *linearen* Techniken.

Tatsächlich lässt sich ein allgemeinerer Minimierungs-Algorithmus entwickeln, in dem dann gegebenenfalls auch nicht-lineare Gleichungssysteme gelöst werden müssen. Diese Vorgehensweise ist *unabhängig* vom Konzept der Faktorisierung im vorherigen Kapitel 3, man kann es sich aber als eine Art „lokale“ Faktorisierung vorstellen. Warum das dennoch die Faktorisierungstheorie für die Minimierung nicht überflüssig macht, wird noch (anhand eines Beispiels) zu klären sein. Eine Frage dazu sei vorab verraten: Kann man sicherstellen, dass nicht *unnötigerweise* nicht-lineare Techniken angewendet werden? Diese Frage ist deswegen so fundamental, weil das Fehlen einer Aussage über die (Nicht-)Existenz einer entsprechenden Lösung (über dem Grundkörper) eines nicht-linearen Gleichungssystems impliziert, dass man *keinerlei* Aussagen über die Minimalität des zulässigen linearen Systems treffen kann. (Die Kenntnis einer Faktorisierung *in Atome* löst dieses Problem freilich auch nicht immer. Jedenfalls sollte man versuchen, linke untere und rechte obere Nullblöcke — so, wie sie zum Beispiel bei der Multiplikation laut Proposition 2.3.1 entstehen— weitestgehend zu erhalten.)

Da es in diesem Kapitel im wesentlichen um die „Minimierung“ der Addition und der Multiplikation geht, sollen noch ein paar Gedanken aus diesen Blickwinkeln einfließen. Aus multiplikativer Sicht ist die Sinnhaftigkeit der Kenntnis einer Faktorisierung eines Elementes in seine jeweiligen (verallgemeinerten) Atome unmittelbar klar: Bei der Multiplikation kann man dann gegebenenfalls kürzen. Das kann man sich zum Beispiel auch zu Nutze machen, um den *linken ggT* zweier Polynome zu finden (Abschnitt A.1). Ganz so trivial ist das aber trotzdem nicht, weil das entsprechende Atom mit seinem Inversen nicht unbedingt „nebeneinander“ liegen muss, zum Beispiel

$$x(1-yx)\cdot x^{-1} = (1-xy)x\cdot x^{-1} = 1-xy.$$

Dazu kommt, dass zwei Atome miteinander zu einem „verschmelzen“ können (siehe Bemerkung 3.4.12) und man daher auf eine Verfeinerung von Pivotblöcken „innerhalb“ eines Atoms angewiesen sein kann. Aber auch aus additiver Sicht spielt die Faktorisierung eine entscheidende Rolle, weil man „gemeinsame“ linke und rechte Faktoren zweier Summanden nur „einmal“ im ZLS braucht. Darüber hinaus ist natürlich die Verfeinerung von Pivotblöcken (aus einer Faktorisierung) besonders wichtig. Mit einer (linearen) Technik haben wir uns bereits (indirekt) im Abschnitt der rationalen Operationen in Form der Lemmata 2.3.4 und 2.3.5 beziehungsweise der Bemerkung 2.3.12 beschäftigt. Diese Technik spielt wiederum eine Rolle für die minimale Faktormultiplikation in Abschnitt 3.5 (und die minimale Inverse, Satz 2.5.13). Aus Sicht der algorithmischen Minimierung verschwindet also die Grenze zwischen Addition und Multiplikation.

Etwas, das in diesem Zusammenhang nur am Rande eine Rolle spielt, aber aus algebraischer Sicht höchst interessant ist, ist die Frage, ob der freie Schiefkörper ein „Ähnlichkeits-UFD“ ist (analog zur Definition 3.1.4). In weiterer Folge ließen sich vielleicht verfeinerte Techniken entwickeln, um aus einer Faktorisierung (in Atome)

alle Faktorisierungen (aus Sicht der Struktur der Nullblöcke in der Systemmatrix eines ZLS) zu erhalten. Im allgemeinen werden lineare Techniken nicht ausreichen, aber vielleicht kann man die Berechnung einer Gröbner-Basis vereinfachen und insbesondere deren Lösbarkeit über dem Grundkörper (und nicht über dessen algebraischen Abschluss) „vereinfacht“ feststellen.

Zur Erinnerung: Hier operieren wir *direkt* in (der Systemmatrix) der linearen Darstellung und sind deswegen *unabhängig* von deren Regularität (das heißt, Invertierbarkeit über den formalen Potenzreihen). Und das hat natürlich seinen Preis. Die „klassischen“ Methoden zur Minimierung linearer Darstellungen für *reguläre* Elemente arbeiten im wesentlichen *indirekt*, in dem sie die linke beziehungsweise rechte Familie „berechnen“, siehe dazu Abschnitt B.3.

So ähnlich wie mit den Abschnitten 3.2 (Polynommultiplikation) und 3.3 (Polynomfaktorisierung) im letzten Kapitel verhält es sich mit den Abschnitten 4.2 (Wortproblem) und 4.3 („polynomielle“ Minimierung) hier. Sie könnten zum Teil wesentlich gekürzt werden indem die Aussagen und Resultate als Spezialfälle der allgemeineren Theorie betrachtet werden. Aber darunter würde die Klarheit leiden, die besonders in diesen Abschnitten richtig zur Geltung kommt. In diesem Sinne kann man Abschnitt 4.3 auch als Einleitung verstehen um danach mit den Grundlagen und den anderen Abschnitten weiter in die Tiefe zu gehen.

Bemerkung. In der Literatur findet man neben dem Begriff „minimieren“ (einer linearen Darstellung) auch „reduzieren“. Da der Begriff „reduziert“ in [Coh95, Abschnitt 6.3] in einem anderen Zusammenhang (für allgemeinere Systemmatrizen beziehungsweise zulässige lineare Systeme) verwendet wird, hätte dessen Verwendung statt „verfeinert“ (in Bezug auf Pivotblöcke) womöglich zur Verwirrung beigetragen. Man denke an „Die *Minimierung* eines *verfeinerten* zulässigen linearen Systems für ein *irreduzibles* Element.“ Statt des Begriffs „Blockzerlegung“ wäre auch „Blockpartitionierung“ (Cohn verwendet letzteren) möglich.

4.1 Grundlagen und eine Standardform

Definition 4.1.1 (Pivotblöcke, Blocktransformation [Sch18a, Definition 3.1])**.** Sei $\mathcal{A} = (u, A, v)$ ein zulässiges lineares System und bezeichne $A = (A_{ij})_{i,j=1}^{m}$ die Blockzerlegung (mit quadratischen Diagonalblöcken A_{ii}) mit *maximalem* m sodass $A_{ij} = 0$ für $i > j$ gilt. Die Diagonalblöcke A_{ii} heißen *Pivotblöcke* und die Anzahl m wird mit $\#_{\mathrm{pb}}(\mathcal{A}) = \#_{\mathrm{pb}}(A)$ bezeichnet. Für $i = 1, 2, \ldots, m$ heißt $\dim_i(\mathcal{A})$ die *Dimension* (oder *Größe*) des Pivotblocks A_{ii}. Für $i < 1$ oder $i > m$ sei $\dim_i(\mathcal{A}) = 0$. Eine (zulässige) Transformation (P, Q) heißt (zulässige) *Blocktransformation* (für $\mathcal{A}$), wenn $P_{ij} = Q_{ij} = 0$ für $i > j$ (für die Blockstruktur von $\mathcal{A}$) gilt.

Notation. Sei $\mathcal{A} = (u, A, v)$ ein ZLS mit m Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Dann bezeichnet $n_{i:j} = \dim_{i:j}(\mathcal{A}) = n_i + n_{i+1} + \ldots + n_j$ die Summe der Größen der Pivotblöcke A_{ii} bis A_{jj} (mit der Konvention $n_{i:j} = 0$ für $j < i$). Für ein vorgegebenes System wird die Einheitsmatrix der Größe $n_{i:j}$ mit $I_{i:j}$ bezeichnet. Wenn (P, Q) eine

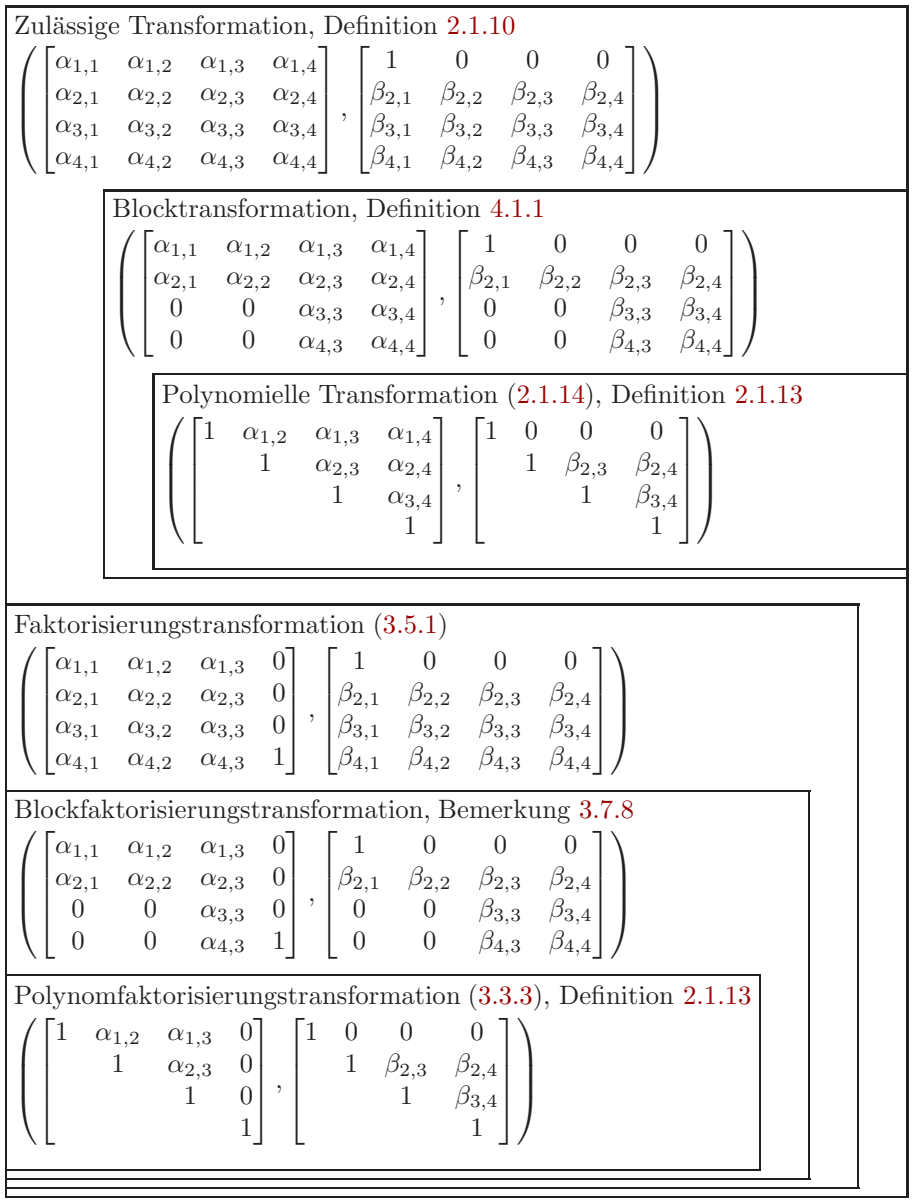


Abbildung 4.1: Die *invertierbaren* Transformationsmatrizen $P = (\alpha_{ij}) \in \mathbb{K}^{n\times n}$ und $Q = (\beta_{ij}) \in \mathbb{K}^{n\times n}$ als Paar (P, Q), angewandt auf ein (nicht notwendigerweise minimales) zulässiges lineares System $\mathcal{A} = (u, A, v)$ der Dimension n für ein Element f des freien Schiefkörpers, ergeben ein *äquivalentes* ZLS $\mathcal{A}' = P\mathcal{A}Q = (uQ, PAQ, Pv)$ für f. Ohne Beschränkung der Allgemeinheit sei hier, der Übersicht halber, $n = 4$. Gegebenenfalls stellen $\det(P) \neq 0$ und $\det(Q) \neq 0$ die Invertierbarkeit sicher.

zulässige Transformation für $\mathcal{A}$ ist, bezeichnet $(PAQ)_{ij}$ den (der Blockzerlegung von A entsprechenden) Block (i,j) der Größe $n_i \times n_j$ in PAQ und $(Pv)_{\underline{i}}$ den der Größe $n_i \times 1$ in Pv und $(uQ)_{\underline{j}}$ den der Größe $1 \times n_j$ in uQ.

Notation. Komponenten in der linken Familie $s = A^{-1}v$ werden (wie immer) mit s_i bezeichnet. Die j-te Komponente für $1 \leq j \leq \dim_i(\mathcal{A})$ des i-ten Blocks für $1 \leq i \leq \#_{\mathrm{pb}}(\mathcal{A})$ wird mit $s_{\underline{i}(j)}$ bezeichnet. Eine Teilfamilie von s bezüglich des Pivotblocks k wird mit $s_{\underline{k}}$ bezeichnet, $s_{\underline{i:j}} = (s_{\underline{i}}, s_{\underline{i+1}}, \ldots, s_{\underline{j}})$. Analog gilt das auch für die rechte Familie $t = uA^{-1}$.

Notation. Eine „Gruppierung“ von Pivotblöcken $\{i, i+1, \ldots, j\}$ der Systemmatrix wird mit $A_{i:j,i:j}$ bezeichnet. Wenn aus dem Kontext klar ist, wo so ein Blockbereich endet oder anfängt, wird auch nur $A_{i:,i:}$ beziehungsweise $A_{:j,:j}$ geschrieben, insbesondere bei einer Blockzerlegung bezüglich eines bestimmten Pivotblocks. Zum Beispiel $A_{1:,1:}$, $A_{k,k}$ und $A_{:m,:m}$.

Definition 4.1.2 (Zulässige Pivotblocktransformation [Sch18a, Definition 3.2])**.** Sei $\mathcal{A} = (u, A, v)$ ein ZLS mit $m = \#_{\mathrm{pb}}(\mathcal{A})$ Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Eine zulässige Transformation (P, Q) der Form $(I_{1:k-1} \oplus \bar{T} \oplus I_{k+1:m}, I_{1:k-1} \oplus \bar{U} \oplus I_{k+1:m})$ mit $\bar{T}, \bar{U} \in \mathbb{K}^{n_k \times n_k}$ heißt *zulässig für Pivotblock k* oder (zulässige) *k-te Pivotblocktransformation*.

Definition 4.1.3 (Verfeinerter Pivotblock, verfeinertes ZLS [Sch18a, Definition 3.3])**.** Sei $\mathcal{A} = (u, A, v)$ ein ZLS mit $m = \#_{\mathrm{pb}}(\mathcal{A})$ Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Ein Pivotblock A_{kk} (für $1 \leq k \leq m$) heißt *verfeinert*, wenn es keine zulässige Pivotblocktransformation $(P, Q)_k$ gibt, sodass $(PAQ)_{kk}$ links unten einen Nullblock der Größe $i \times (n_k - i)$ für ein $i \in \{1, 2, \ldots, n_k - 1\}$ hat. Das zulässige lineare System $\mathcal{A}$ heißt *verfeinert*, wenn alle Pivotblöcke verfeinert sind.

Bemerkungen. Zum einen ist die „Form“ eines verfeinerten ZLS nicht eindeutig. Das wird im folgenden Beispiel illustriert. Zum anderen ist ein verfeinertes ZLS nicht notwendigerweise verfeinert über dem algebraischen Abschluss des Grundkörpers $\overline{\mathbb{K}}$. Zum Beispiel sei

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} 1 & x \\ 2x & 1 \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right).$$

Addiert man $\sqrt{2}$-mal Zeile 1 zu Zeile 2 und subtrahiert man $\sqrt{2}$-mal Spalte 2 von Spalte 1, erhält man

$$\begin{bmatrix} 1 - \sqrt{2}x & x \\ 0 & 1 + \sqrt{2}x \end{bmatrix} s = \begin{bmatrix} . \\ 1 \end{bmatrix}.$$

Siehe auch Beispiel 3.3.4 (Irreduzibilität in Abhängigkeit des Grundkörpers).

Beispiel 4.1.4 ([Sch18a, Beispiel 3.4])**.** Sei $f = (y^{-1} - x)^{-1}$ gegeben durch das minimale ZLS

$$\mathcal{A}_f = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} 1 & -y \\ -x & 1 \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right).$$

Durch die Propositionen 2.2.1 (Minimales Monom) und 2.4.3 (Disjunkte Addition) erhält man für $f + 3z$ sofort das minimale und *verfeinerte* ZLS

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -y & -1 & . \\ -x & 1 & x & . \\ . & . & 1 & -z \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ 1 \\ . \\ 3 \end{bmatrix} \right).$$

Da $1 \in R(\mathcal{A})$ ist und dieses System über die Addition entstanden ist, kann sehr einfach —in einer kontrollierten Weise— eine andere *verfeinerte* Pivotblock-Struktur erzeugt werden (vergleiche dazu Bemerkung 2.3.12). Zuerst addiert man Spalte 3 zu Spalte 1,

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 0 & -y & -1 & . \\ 0 & 1 & x & . \\ 1 & . & 1 & -z \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ 1 \\ . \\ 3 \end{bmatrix} \right),$$

und dann vertauscht man die Zeilen 1 und 3:

$$\mathcal{A}'' = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & . & 1 & -z \\ . & 1 & x & . \\ . & -y & -1 & . \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ 1 \\ . \\ 3 \end{bmatrix} \right).$$

Definition 4.1.5 (Blockzerlegung eines ZLS, Blockzeilen- und -spaltentransformation [Sch18a, Definition 3.5])**.** Sei $\mathcal{A} = (u, A, v)$ ein ZLS der Dimension n mit $m = \#_{\text{pb}}(\mathcal{A}) \geq 2$ Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Für ein $1 \leq k \leq m$ ist die *Blockzerlegung* bezüglich Pivotblock k das System

$$\mathcal{A}^{[\underline{k}]} = \left(\begin{bmatrix} u_{\underline{1:}} & . & . \end{bmatrix}, \begin{bmatrix} A_{1:,1:} & A_{1:,k} & A_{1:,:m} \\ . & A_{k,k} & A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix}, \begin{bmatrix} v_{\underline{1:}} \\ v_{\underline{k}} \\ v_{\underline{:3}} \end{bmatrix} \right)$$

mit (quadratischen) Diagonalblöcken $A_{1:,1:}$, $A_{k,k}$ und $A_{:m,:m}$ der Größen $n_{1:k-1}$, n_k beziehungsweise $n_{k+1:m}$. ($\underline{k}$ wird hier verwendet um zu betonen, dass k ein Blockindex ist.) Mit $\mathcal{A}^{[-\underline{k}]}$ wird das ZLS $\mathcal{A}^{[\underline{k}]}$ ohne Blockzeile/-spalte k (der Dimension $n - n_k$) bezeichnet ($\mathcal{A}^{[-\underline{k}]}$ ist *nicht* notwendigerweise äquivalent zu $\mathcal{A}^{[\underline{k}]}$):

$$\mathcal{A}^{[-\underline{k}]} = \left(\begin{bmatrix} u_{\underline{1}} & . \end{bmatrix}, \begin{bmatrix} A_{1:,1:} & A_{1:,:m} \\ . & A_{:m,:m} \end{bmatrix}, \begin{bmatrix} v_{\underline{1}} \\ v_{\underline{3}} \end{bmatrix} \right).$$

Eine zulässige Transformation $(P, Q)_{\underline{k}} = \big(P(\bar{T}, T), Q(\bar{U}, U)\big)_{\underline{k}}$ der Form

$$(P, Q)_{\underline{k}} = \left(\begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{T} & T \\ . & . & I_{k+1:m} \end{bmatrix}, \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{U} & U \\ . & . & I_{k+1:m} \end{bmatrix} \right) \tag{4.1.6}$$

heißt k-te *Blockzeilentransformation* für $\mathcal{A}^{[\underline{k}]}$, eine $(P,Q)^{\underline{k}} = \big(P(\bar{T},T), Q(\bar{U},U)\big)^{\underline{k}}$ der Form

$$(P,Q)^{\underline{k}} = \left(\begin{bmatrix} I_{1:k-1} & T & . \\ . & \bar{T} & . \\ . & . & I_{k+1:m} \end{bmatrix}, \begin{bmatrix} I_{1:k-1} & U & . \\ . & \bar{U} & . \\ . & . & I_{k+1:m} \end{bmatrix} \right) \tag{4.1.7}$$

heißt k-te *Blockspaltentransformation* für $\mathcal{A}^{[\underline{k}]}$. Für $\bar{T} = \bar{U} = I_{n_k}$ schreiben wir auch nur $P(T)$ bzw. $Q(U)$ und nennen eine Blocktransformation $\big(P(T), Q(U)\big)$ dann auch *speziell*.

Definition 4.1.8 (Standardisiertes ZLS [Sch18a, Definition 3.8])**.** Ein *minimales* und *verfeinertes* ZLS $\mathcal{A} = (u, A, v) = (1, A, \lambda)$, das heißt, $v = [0, \ldots, 0, \lambda]$, heißt *standardisiert* (oder *Standard-ZLS*).

Mit (der allgemeinen) Definition 3.4.10 (von Atomen) und Proposition 3.4.11 kann Definition 3.1.3 (atomare zulässige lineare Systeme) verallgemeinert werden:

Definition 4.1.9 (Atomares ZLS)**.** Ein *standardisiertes* zulässiges lineares System für ein Atom (im freien Schiefkörper) heißt *atomar* (oder *irreduzibel*).

4.2 Das Wortproblem

Seien $f, g \in \mathbb{F}$ gegeben durch die linearen Darstellungen $\pi_f = (u_f, A_f, v_f)$ der Dimension n_f beziehungsweise $\pi_g = (u_g, A_g, v_g)$ der Dimension n_g. Die Matrix

$$L = \begin{bmatrix} . & u_f & u_g \\ v_f & A_f & . \\ v_g & . & -A_g \end{bmatrix}$$

ist eine *Linearisierung* von $f - g$ der Größe $n = n_f + n_g + 1$. Nun ist $f = g$ genau dann, wenn L nicht voll ist [Coh95, Abschnitt 4.5]. Bezüglich Linearisierung siehe auch Abschnitt B.1, bezüglich Wortproblem [Coh95, Abschnitt 6.6]. Ob die Matrix L voll ist (oder nicht), kann mittels des folgenden Satzes festgestellt werden. Für $P = (\alpha_{ij})$ und $Q = (\beta_{ij})$ ist der *kommutative* Polynomring

$$\begin{aligned} \mathbb{K}[\alpha,\beta] = \mathbb{K}[&\alpha_{1,1}, \ldots, \alpha_{1,n}, \alpha_{2,1}, \ldots, \alpha_{2,n}, \ldots, \alpha_{n,1}, \ldots, \alpha_{n,n}, \\ &\beta_{1,1}, \ldots, \beta_{1,n}, \beta_{2,1}, \ldots, \beta_{2,n}, \ldots, \beta_{n,1}, \ldots, \beta_{n,n}]. \end{aligned}$$

Satz 4.2.1 ([CR99, Satz 4.1])**.** *Für jedes $k \in \{1, 2, \ldots, n\}$ bezeichne I_k das Ideal von $\mathbb{K}[\alpha,\beta]$, das von den Polynomen $\det(P) - 1$, $\det(Q) - 1$ und den Koeffizienten von jedem $x \in \{1\} \cup X$ in den (i,j)-Einträgen der Matrix PLQ für $1 \le i \le k$ und $k \le j \le n$ erzeugt wird. Dann ist die lineare Matrix L genau dann voll, wenn jedes Ideal $I_r = \mathbb{K}[\alpha,\beta]$ für $r \in \{1, 2, \ldots, n\}$ ist.*

Bemerkung. In [CR99] ist an dieser Stelle ein Druckfehler, die Koeffizienten von L_0, den Koeffizienten bezüglich des *leeren Wortes*, fehlen.

Bemerkung. Eine Variante dieses Satzes (ohne Bedingung an die Determinanten der Transformationsmatrizen), nämlich Proposition 3.3.7, wird für die Faktorisierung von Polynomen über $\overline{\mathbb{K}}\langle X\rangle$ verwendet.

Bemerkung. Eine Lösung für das Wortproblem (im freien Schiefkörper) findet sich bereits in [Coh73, Coh75b].

Die praktische Anwendbarkeit dieses Satzes stößt bereits für $n \geq 5$, wo 50 oder mehr Unbekannte involviert sind, an ihre Grenzen. Hat man jedoch irgendein ZLS (oder eine lineare Darstellung) für $f - g$, zum Beispiel laut Proposition 2.3.1, dann kann man prüfen, ob es (zulässig) in ein kleineres System, zum Beispiel $A's' = 0$, transformiert werden kann. Für Polynome (mit $A = I - M$ und M einer oberen (nilpotenten) Dreiecksmatrix) kann das Zeile für Zeile (bzw. Spalte für Spalte) gemacht werden. Tatsächlich steckt diese Idee in Algorithmus 4.3.8 (im folgenden Abschnitt). Im allgemeinen können die Pivotblöcke (die quadratischen Blöcke in der Diagonalen) beliebig groß sein. Daher muss die Eliminierung (von Zeilen bzw. Spalten) *blockweise* erfolgen, indem man ein einzelnes lineares Gleichungssystem für Zeilen- *und* Spaltenoperationen aufstellt (und löst). Diese Idee steckt im folgenden Lemma.

Bemerkung. Die Existenz einer Lösung für dieses lineare Gleichungssystem ist *invariant* unter zulässigen Transformationen (der Teilsysteme). Das ist eine Schlüsselanforderung, da die Normalform [CR94] nur modulo Ähnlichkeitstransformationen eindeutig ist. Für nicht-minimale lineare Darstellungen ist die *stabile Assoziation* (Definition 2.1.2) relevant.

Lemma 4.2.2 ([Sch17b, Lemma 2.3])**.** *Seien $f, g \in \mathbb{F}$ gegeben durch die zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f)$ der Dimension n_f beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g)$ der Dimension n_g. Wenn Matrizen $T, U \in \mathbb{K}^{n_f \times n_g}$ existieren, sodass $u_f U = 0$, $TA_g - A_f U = A_f u_f^\top u_g$ und $Tv_g = v_f$ gilt, dann ist $f = g$.*

Beweis. Die Differenz $f - g$ kann durch das ZLS $As = v$ mit

$$A = \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ . & A_g \end{bmatrix}, \quad s = \begin{bmatrix} s_f - u_f^\top g \\ -s_g \end{bmatrix} \quad \text{und} \quad v = \begin{bmatrix} v_f \\ -v_g \end{bmatrix}$$

dargestellt werden. Wir definieren die (invertierbaren) Transformationsmatrizen

$$P = \begin{bmatrix} I_{n_f} & T \\ . & I_{n_g} \end{bmatrix} \quad \text{und} \quad Q = \begin{bmatrix} I_{n_f} & -U \\ . & I_{n_g} \end{bmatrix}$$

und erhalten mit $A' = PAQ$, $s' = Q^{-1}s$ und $v' = Pv$ das neue ZLS $A's' = v'$:

$$\begin{aligned}
A' &= \begin{bmatrix} I_{n_f} & T \\ . & I_{n_g} \end{bmatrix} \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ . & A_g \end{bmatrix} \begin{bmatrix} I_{n_f} & -U \\ . & I_{n_g} \end{bmatrix} \\
&= \begin{bmatrix} A_f & -A_f u_f^\top u_g + TA_g \\ . & A_g \end{bmatrix} \begin{bmatrix} I_{n_f} & -U \\ . & I_{n_g} \end{bmatrix} \\
&= \begin{bmatrix} A_f & -A_f u_f^\top u_g + TA_g - A_f U \\ . & A_g \end{bmatrix} = \begin{bmatrix} A_f & 0 \\ . & A_g \end{bmatrix}, \\
s' &= \begin{bmatrix} I_{n_f} & U \\ . & I_{n_g} \end{bmatrix} \begin{bmatrix} s_f - u_f^\top g \\ -s_g \end{bmatrix} = \begin{bmatrix} s_f - u_f^\top g - Us_g \\ -s_g \end{bmatrix}, \\
v' &= \begin{bmatrix} I_{n_f} & T \\ . & I_{n_g} \end{bmatrix} \begin{bmatrix} v_f \\ -v_g \end{bmatrix} = \begin{bmatrix} v_f - Tv_g \\ -v_g \end{bmatrix}.
\end{aligned}$$

Die Invertierbarkeit von A' über dem freien Schiefkörper impliziert $s_f - u_f^\top g - Us_g = 0$, insbesondere gilt

$$\begin{aligned}
0 &= u_f s_f - u_f u_f^\top g - u_f U s_g \\
&= f - g
\end{aligned}$$

weil $u_f U = 0$ ist. □

Sei d die Anzahl der Buchstaben im Alphabet X, $\dim(\mathcal{A}_f) = n_f$ und $\dim(\mathcal{A}_g) = n_g$. Um die Transformationsmatrizen $T, U \in \mathbb{K}^{n_f \times n_g}$ vom vorherigen Lemma zu bestimmen, brauchen wir nur ein lineares Gleichungssystem mit $(d+1)n_f(n_g+1)$ Gleichungen in $2n_f n_g$ Unbekannten lösen. Wenn es eine Lösung gibt, ist $f = g$. Weder $\mathcal{A}_f$ noch $\mathcal{A}_g$ müssen minimal sein. Computerexperimente zeigen, dass Huas Identität [Ami66]

$$x - \left(x^{-1} + (y^{-1} - x)^{-1}\right)^{-1} = xyx$$

positiv mittels Lemma 4.2.2 getestet werden kann, wenn das ZLS für die linke Seite über die rationalen Operationen von Proposition 2.3.1 konstruiert wird. Jedoch ohne der Annahme der Minimalität impliziert das Fehlen einer Lösung *nicht*, dass $f \neq g$ ist, siehe dazu das folgende Beispiel 4.2.4. Huas Identität wird in Beispiel 2.6.1 schrittweise über zulässige lineare Systeme bewiesen. Ist eine der beiden linearen Darstellungen nicht minimal, kann auch die Gaußsche Elimination (über dem freien Schiefkörper) verwendet werden, siehe dazu Abschnitt A.2.

Satz 4.2.3 („Lineares“ Wortproblem [Sch17b, Satz 2.4])**.** *Seien $f, g \in \mathbb{F}$ gegeben durch die* minimalen *zulässigen linearen Systeme $\mathcal{A}_f = (u_f, A_f, v_f)$ beziehungsweise $\mathcal{A}_g = (u_g, A_g, v_g)$ jeweils der Dimension n. Dann ist $f = g$ genau dann, wenn es Matrizen $T, U \in \mathbb{K}^{n \times n}$ gibt, sodass $u_f U = 0$, $TA_g - A_f U = A_f u_f^\top u_g$ und $Tv_g = v_f$ gilt.*

Beweis. Wenn $f = g$ ist, dann existieren —nachdem zulässige lineare Systeme (reine) lineare Darstellungen sind— laut Satz 2.1.6 invertierbare Matrizen $P, Q \in \mathbb{K}^{n\times n}$ sodass $A_f = PA_gQ$ und $v_f = Pv_g$ ist. Sei $T = P$ und $U = Q^{-1} - u_f^\top u_g$. Die zulässigen linearen Systeme sind minimal. Daher ist die linke Familie s_f $\mathbb{K}$-linear unabhängig. Nachdem die erste Komponente von s_g gleich der ersten von $s_f = Q^{-1}s_g$ ist und die linke Familie s_g ebenfalls $\mathbb{K}$-linear unabhängig ist, muss $[1, 0, \ldots, 0]$ die erste Zeile von Q^{-1} sein. Deshalb ist $u_fU = u_f(Q^{-1} - u_f^\top u_g) = 0$. Klarerweise gilt $v_f = Tv_g$ und

$$\begin{aligned} TA_g - A_fU &= PA_g - A_fQ^{-1} + A_fu_f^\top u_g \\ &= A_fu_f^\top u_g. \end{aligned}$$

Die andere Implikation folgt aus Lemma 4.2.2. □

Beispiel 4.2.4 ([Sch17b, Beispiel 2.5])**.** Seien $f = x^{-1}$ und $g = x^{-1}$ gegeben durch die zulässigen linearen Systeme

$$[x]s_f = [1] \quad \text{bzw.} \quad \begin{bmatrix} x & -z \\ . & 1 \end{bmatrix} s_g = \begin{bmatrix} 1 \\ . \end{bmatrix}.$$

Dann ist das zulässige lineare System

$$\begin{bmatrix} x & -x & . \\ . & x & -z \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} 1 \\ -1 \\ . \end{bmatrix}, \quad s = \begin{bmatrix} 0 \\ -x^{-1} \\ 0 \end{bmatrix}$$

eine lineare Darstellung von $f - g = 0$.

Obwohl es hier offensichtlich ist, dass die zweite Komponente im Lösungsvektor s_g Null ist, ist es im allgemeinen nicht klar, wie man solche „pathologischen“ linearen Darstellungen ohne Annahme der Minimalität ausschließen kann.

Bemerkung 4.2.5 ([Sch17b, Abschnitt 2])**.** Man könnte sich fragen, für welche Art der Konstruktion (rationale Operationen [CR99], Higmans Trick [Hig40, Coh85], selbstadjungierter Linearisierungstrick [And13], etc.) es hinreichende Bedingungen für die Existenz von Matrizen T, U (über $\mathbb{K}$) in Lemma 4.2.2 gibt, wenn $f = g$ ist. Leider scheint das —einige Beispiele ausgenommen— unmöglich, wie das folgende zulässige lineare System (konstruiert mittels der rationalen Operationen von Proposition 2.3.1) für $x - xyy^{-1} = 0$ zeigt (einige Nullen wurden belassen, um die Blockstruktur zu betonen):

$$\begin{bmatrix} 1 & -x & -1 & . & . & . & . & . & . \\ 0 & 1 & . & . & . & . & . & . & . \\ . & . & 1 & -x & . & . & . & . & . \\ . & . & 0 & 1 & -1 & . & . & . & . \\ . & . & . & . & 0 & 1 & -y & . & . \\ . & . & . & . & -1 & 0 & 1 & . & . \\ . & . & . & . & 0 & 1 & 0 & -1 & . \\ . & . & . & . & . & . & . & 1 & -y \\ . & . & . & . & . & . & . & 0 & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 1 \\ . \\ . \\ . \\ . \\ . \\ . \\ -1 \end{bmatrix}.$$

Hier gibt es keine Matrizen T, U und damit auch keine Transformationen P, Q (zulässig, mit Blöcken T, U), sodass PAQ einen oberen rechten Nullblock der Größe 2×7 hat und die ersten beiden Komponenten von Pv Null sind.

4.3 Minimieren eines polynomiellen ZLS

Ein genauer Blick auf den Beweis von Proposition 3.2.7 (Minimale Polynommultiplikation) bringt einen überraschend einfachen Algorithmus für die Konstruktion eines *minimalen* polynomiellen zulässigen linearen Systems zum Vorschein, vorausgesetzt es ist in dieser Form gegeben. Er kann für die Minimierung der Summe in Proposition 2.3.1 verwendet werden. „Einfach“ heißt, dass er auch für ziemlich große dünn besetzte Systeme leicht manuell umgesetzt werden kann. In Abhängigkeit der Datenstruktur ist die Implementierung (siehe dazu auch Abschnitt B.5) selbst etwas technisch. Man muss sehr vorsichtig sein, wenn die Skalare (des Grundkörpers $\mathbb{K}$) nicht exakt dargestellt werden können, insbesondere wenn lineare Gleichungssysteme (siehe weiter unten) gelöst werden müssen.

Um die zugrundeliegende Idee zu illustrieren, minimieren wir (teilweise) ein *nichtminimales* „fast“ polynomielles ZLS $\mathcal{A} = (u, A, v)$ der Dimension $n = 6$ für $p = -xy + (xy + z)$. Man beachte, dass man dazu keinerlei Wissen über die linke oder die rechte Familie benötigt. Sei

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & . & -1 & . & . \\ . & 1 & -y & . & . & . \\ . & . & 1 & . & . & . \\ . & . & . & 1 & -x & -z \\ . & . & . & . & 1 & -y \\ . & . & . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ -1 \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Zuerst führen wir einen linken Minimierungsschritt durch, das heißt, wir entfernen (wenn möglich) ein Element der $\mathbb{K}$-linear abhängigen linken Familie $s = A^{-1}v$ und konstruieren ein neues System. Wir fixieren ein $1 \leq k < n$, sagen wir $k = 3$. Wenn wir eine zulässige Transformation (P, Q) der Form

$$(P, Q) = \left(\begin{bmatrix} I_{k-1} & . & . \\ . & 1 & T \\ . & . & I_{n-k} \end{bmatrix}, \begin{bmatrix} I_{k-1} & . & . \\ . & 1 & U \\ . & . & I_{n-k} \end{bmatrix} \right) \tag{4.3.1}$$

finden, sodass Zeile k in PAQ gleich $[0, 0, 1, 0, 0, 0]$ und $(Pv)_k = 0$ ist, können wir Zeile k und Spalte k in $P\mathcal{A}Q$ eliminieren, weil $(Q^{-1}s)_k = 0$ ist. Wie können wir diese Blöcke $T, U \in \mathbb{K}^{1\times(n-k)}$ finden? Wir schreiben $\mathcal{A}$ in Blockform —Blockzeilen- und Blockspaltenindizes sind unterstrichen um sie von Komponentenindizes zu unterscheiden— bezüglich Zeile/Spalte k

$$\mathcal{A}^{[k]} = \left(\begin{bmatrix} u_{\underline{1}} & . & . \end{bmatrix}, \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ . & 1 & A_{2,3} \\ . & . & A_{3,3} \end{bmatrix}, \begin{bmatrix} v_{\underline{1}} \\ v_{\underline{2}} \\ v_{\underline{3}} \end{bmatrix} \right) \tag{4.3.2}$$

und wenden die Transformation (P, Q) an:

$$\begin{aligned} PAQ &= \begin{bmatrix} I_{k-1} & . & . \\ . & 1 & T \\ . & . & I_{n-k} \end{bmatrix} \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ . & 1 & A_{2,3} \\ . & . & A_{3,3} \end{bmatrix} \begin{bmatrix} I_{k-1} & . & . \\ . & 1 & U \\ . & . & I_{n-k} \end{bmatrix} \\ &= \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,2}U + A_{1,3} \\ . & 1 & U + A_{2,3} + TA_{3,3} \\ . & . & A_{3,3} \end{bmatrix}, \\ Pv &= \begin{bmatrix} I_{k-1} & . & . \\ . & 1 & T \\ . & . & I_{n-k} \end{bmatrix} \begin{bmatrix} v_{\underline{1}} \\ v_{\underline{2}} \\ v_{\underline{3}} \end{bmatrix} = \begin{bmatrix} v_{\underline{1}} \\ v_{\underline{2}} + Tv_{\underline{3}} \\ v_{\underline{3}} \end{bmatrix}. \end{aligned}$$

Nun können wir eine *hinreichende* Bedingung für $(Q^{-1}s)_k = 0$ ablesen, nämlich die *Existenz* von $T, U \in \mathbb{K}^{1\times n-k}$ sodass

$$U + A_{2,3} + TA_{3,3} = 0 \quad \text{und} \quad v_{\underline{2}} + Tv_{\underline{3}} = 0 \tag{4.3.3}$$

gilt. Sei d die Anzahl der Buchstaben in unserem Alphabet X. Die Blöcke $T = [\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_n]$ und $U = [\beta_{k+1}, \beta_{k+2}, \ldots, \beta_n]$ in der Transformation (P, Q) haben die Größe $1 \times (n-k)$ und die beiden Matrizen T und U sind „entkoppelt“, daher bekommen wir ein *lineares* Gleichungssystem (über $\mathbb{K}$) mit $2(n-k)$ Unbekannten (für $k > 1$) und $(d+1)(n-k)+1$ Gleichungen:

$$\begin{aligned} \begin{bmatrix} \beta_{k+1} & \beta_{k+2} & \beta_{k+3} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} + \\ + \begin{bmatrix} \alpha_{k+1} & \alpha_{k+2} & \alpha_{k+3} \end{bmatrix} \begin{bmatrix} 1 & -x & -z \\ . & 1 & -y \\ . & . & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \\ \begin{bmatrix} -1 \end{bmatrix} + \begin{bmatrix} \alpha_{k+1} & \alpha_{k+2} & \alpha_{k+3} \end{bmatrix} \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}. \end{aligned}$$

Eine Lösung ist $T = [0, 0, 1]$ und $U = [0, 0, -1]$. Wir ermitteln $\tilde{\mathcal{A}}_1 = P\mathcal{A}Q$ und entfernen Blockzeile $\underline{2}$ und -spalte $\underline{2}$, das heißt, Zeile k und Spalte k, um das neue äquivalente ZLS

$$\mathcal{A}_1 = (u, A, v) = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -1 & . & . \\ . & 1 & . & . & y \\ . & . & 1 & -x & -z \\ . & . & . & 1 & -y \\ . & . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right)$$

zu erhalten. Als nächstes führen wir einen rechten Minimierungsschritt durch, das heißt, wir entfernen (wenn möglich) ein Element der $\mathbb{K}$-linear abhängigen rechten Familie $t = uA^{-1}$ und konstruieren ein neues System. Dazu fixieren wir ein $1 < k \leq$

$n = 5$, sagen wir $k = 3$. Nun suchen wir eine Transformation (P, Q) der Form

$$(P, Q) = \left(\begin{bmatrix} I_{k-1} & T & . \\ . & 1 & . \\ . & . & I_{n-k} \end{bmatrix}, \begin{bmatrix} I_{k-1} & U & . \\ . & 1 & . \\ . & . & I_{n-k} \end{bmatrix} \right) \tag{4.3.4}$$

sodass die Spalte k in PAQ gleich $[0, 0, 1, 0, 0]^\top$ ist (für eine zulässige Transformation, das heißt, die erste Zeile von U ist Null, ist der entsprechende Eintrag u_k in u Null). Eine hinreichende Bedingung für $(tP^{-1})_k = 0$ ist die Existenz von $T, U \in \mathbb{K}^{(k-1)\times 1}$ sodass

$$A_{1,1}U + A_{1,2} + T = 0 \tag{4.3.5}$$

ist. (In Bemerkung 4.3.9 gibt es eine weniger „komprimierte“ Version dieses linearen Gleichungssystems.) Eine Lösung ist $T = [1, 0]^\top$ und $U = [0, 0]^\top$. Wir ermitteln $\tilde{\mathcal{A}}_2 = P\mathcal{A}_1Q$ und entfernen Zeile k und Spalte k um das neue (noch nicht minimale) ZLS

$$\mathcal{A}_2 = (u, A, v) = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -x & -z \\ . & 1 & . & y \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right)$$

zu erhalten. Wenn man einen linken (beziehungsweise rechten) Minimierungsschritt mit $k = 1$ (beziehungsweise $k = n$ und $v = [0, \dots, 0, \lambda]^\top$) durchführen kann, dann repräsentiert das ZLS Null und wir können sofort aufhören.

Das Folgende ist die einzige nicht-triviale Beobachtung: Wir erinnern uns, wenn Zeilen- (bzw. Spalten-) Blöcke T, U existieren, sodass (4.3.3) (bzw. (4.3.5)) eine Lösung besitzt, dann ist die linke (bzw. rechte) Familie $\mathbb{K}$-linear abhängig. Um Minimalität laut Proposition 2.1.8 zu garantieren, brauchen wir die andere Implikation, das heißt, die Existenz entsprechender Zeilen- *oder* Spaltenblöcke für nicht-minimale polynomielle zulässige lineare Systeme.

Obwohl die Argumente im Beweis von Proposition 3.2.7 (Polynommultiplikation) zu finden sind, wiederholen wir sie hier, weil sie der entscheidende Teil des Minimierungsalgorithmus sind: Sei $\mathcal{A} = (u, A, v)$ ein polynomielles ZLS der Dimension $n \geq 2$ mit linker Familie $s = (s_1, s_2, \dots, s_n)$. Wir nehmen an, dass ein $1 \leq k < n$ existiert, sodass die Teilfamilie $(s_{k+1}, s_{k+2}, \dots, s_n)$ $\mathbb{K}$-linear unabhängig ist, während $(s_k, s_{k+1}, \dots, s_n)$ $\mathbb{K}$-linear abhängig ist. Dann existieren laut Lemma 3.2.3 Matrizen $T, U \in \mathbb{K}^{1\times(n-k)}$ sodass (4.3.3) gilt. In anderen Worten: Wir müssen mit $k_s = n - 1$ für einen linken (und $k_t = 2$ für einen rechten) Minimierungsschritt starten.

Wenn wir einen Minimierungsschritt ausführen, müssen wir die jeweils andere Familie „erneut“ prüfen. Dazu betrachten wir folgendes Beispiel, das *nicht* aus zwei minimalen Systemen konstruiert wurde:

$$\mathcal{A} = (u, A, v) = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -y & x+y & . \\ . & 1 & . & . & -z \\ . & . & 1 & . & -z \\ . & . & . & 1 & -y \\ . & . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Klarerweise ist sowohl die linke Teilfamilie (s_3, s_4, s_5) von $s = A^{-1}v$ als auch die rechte (t_1, t_2, t_3) von $t = uA^{-1}$ $\mathbb{K}$-linear unabhängig. Subtrahiert man Zeile 3 von Zeile 2 und addiert man Spalte 2 zu Spalte 3 erhält man das ZLS

$$\mathcal{A}' = (u', A', v') = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -x-y & x+y & . \\ . & 1 & 0 & . & 0 \\ . & . & 1 & . & -z \\ . & . & . & 1 & -y \\ . & . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Die rechte Teilfamilie (t_1'', t_2'', t_3'') von $\mathcal{A}'' = \mathcal{A}'^{[-2]}$ ist (hier) *nicht* mehr $\mathbb{K}$-linear unabhängig, man muss also erneut einen rechten Minimierungsschritt für $k = 3$ probieren.

Definition 4.3.6 (Minimierungsgleichungen)**.** Sei $\mathcal{A} = (u, A, v)$ ein polynomielles ZLS der Dimension $n \geq 2$. Bezüglich der Blockzerlegung (4.3.2) bezeichne $\mathcal{A}^{[-k]}$ das —nicht notwendigerweise äquivalente— ZLS $\mathcal{A}^{[k]}$ ohne Zeile/Spalte k (der Dimension $n-1$):

$$\mathcal{A}^{[-k]} = \left(\begin{bmatrix} u_{\underline{1}} & . \end{bmatrix}, \begin{bmatrix} A_{1,1} & A_{1,3} \\ . & A_{3,3} \end{bmatrix}, \begin{bmatrix} v_{\underline{1}} \\ v_{\underline{3}} \end{bmatrix} \right).$$

Für $k \in \{1, 2, \ldots, n-1\}$ heißen die Gleichungen (4.3.3), das heißt, $U + A_{2,3} + TA_{3,3} = 0$ und $v_{\underline{2}} + Tv_{\underline{3}} = 0$ bezüglich der Blockzerlegung $\mathcal{A}^{[k]}$, k-te *linke Minimierungsgleichungen*. Sie werden mit $\mathcal{L}_k = \mathcal{L}_k(\mathcal{A})$ bezeichnet. Eine Lösung durch das Zeilenpaar (T, U) wird mit $\mathcal{L}_k(T, U) = 0$ bezeichnet, die induzierte Transformation als $\big(P(T), Q(U)\big)$ geschrieben. Für $k \in \{2, 3, \ldots, n\}$ heißen die Gleichungen (4.3.5), das heißt, $A_{1,1}U + A_{1,2} + T = 0$ bezüglich der Blockzerlegung $\mathcal{A}^{[k]}$, k-te *rechte Minimierungsgleichungen*. Sie werden mit $\mathcal{R}_k = \mathcal{R}_k(\mathcal{A})$ bezeichnet. Eine Lösung durch das Spaltenpaar (T, U) wird mit $\mathcal{R}_k(T, U) = 0$ bezeichnet, die induzierte Transformation als $\big(P(T), Q(U)\big)$ geschrieben.

Nun fehlt nur noch ein wichtiges Detail, nämlich die Tatsache, dass wir Lemma 3.2.3 nicht im folgenden (ersten) linken Minimierungsschritt anwenden können, weil $\big(P(T), Q(U)\big)$ nicht notwendigerweise zulässig ist. Das lässt sich aber folgendermaßen umgehen: Für $0 \neq \alpha \in \mathbb{K}$ betrachten wir das ZLS $\mathcal{A}$:

$$\begin{bmatrix} 1 & -\alpha \\ . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ \lambda \end{bmatrix}.$$

Wir transformieren die Systemmatrix A (zulässig) in der folgende Weise:

$$\underbrace{\begin{bmatrix} . & \alpha \\ 1 & . \end{bmatrix}}_{=:P} \begin{bmatrix} 1 & -\alpha \\ . & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & . \\ 1/\alpha & 1 \end{bmatrix}}_{=:Q} = \begin{bmatrix} . & \alpha \\ 1 & -\alpha \end{bmatrix} \begin{bmatrix} 1 & . \\ 1/\alpha & 1 \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ 0 & -\alpha \end{bmatrix}.$$

Mit dieser *zulässigen* Transformation (P, Q) erhalten wir das System $\mathcal{A}' = P\mathcal{A}Q$,

$$\begin{bmatrix} 1 & \alpha \\ . & -\alpha \end{bmatrix} s = \begin{bmatrix} \alpha\lambda \\ 0 \end{bmatrix}, \tag{4.3.7}$$

in dem wir die *letzte* Zeile und Spalte löschen können. Man beachte, dass für die rechte Familie kein Sonderfall zu berücksichtigen ist.

Algorithmus 4.3.8 (Minimieren eines polynomiellen ZLS)**.**
Eingabe: $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ *polynomielles ZLS (für ein Polynom p) der Dimension* $n \geq 2$.
Ausgabe: $\mathcal{A}' = (,,)$, *wenn* $p = 0$ *ist, oder ein* minimales *polynomielles ZLS* $\mathcal{A}' = (u', A', v') = (1, A', \lambda')$ *wenn* $p \neq 0$ *ist.*

$k := 2$
while $k \leq \dim(\mathcal{A})$ **do**
  $n := \dim(\mathcal{A})$
  $k' := n + 1 - k$
  Ist die linke Teilfamilie $(s_{k'}, \overbrace{s_{k'+1}, \ldots, s_n}^{\text{lin. unabhängig}})$ $\mathbb{K}$-linear abhängig?
  if $\exists\, T, U \in \mathbb{K}^{1\times(k-1)}$ zulässig : $\mathcal{L}_{k'}(\mathcal{A}) = \mathcal{L}_{k'}(T, U) = 0$ **then**
    if $k' = 1$ **then**
      return $(,,)$
    endif
    $\mathcal{A} := \big(P(T)\mathcal{A}Q(U)\big)^{[-k']}$
    if $k > \max\{2, \frac{n+1}{2}\}$ **then**
      $k := k - 1$
    endif
    continue
  endif
  if $k = 2$ **and** $s_{n-1} = \alpha s_n$ (für ein $\alpha \in \mathbb{K}$) **then**
    find zulässiges (P, Q) sodass $(Q^{-1}s)_n = 0$ und $P\mathcal{A}Q$ polynomiell ist
    $\mathcal{A} := (P\mathcal{A}Q)^{[-n]}$
    continue
  endif
  Ist die rechte Teilfamilie $(\overbrace{t_1, \ldots, t_{k-1}}^{\text{lin. unabhängig}}, t_k)$ $\mathbb{K}$-linear abhängig?
  if $\exists\, T, U \in \mathbb{K}^{(k-1)\times 1}$ zulässig : $\mathcal{R}_k(\mathcal{A}) = \mathcal{R}_k(T, U) = 0$ **then**
    $\mathcal{A} := \big(P(T)\mathcal{A}Q(U)\big)^{[-k]}$
    if $k > \max\{2, \frac{n+1}{2}\}$ **then**
      $k := k - 1$
    endif
    continue
  endif
  $k := k + 1$
done
return $P\mathcal{A}$, mit P, sodass $Pv = [0, \ldots, 0, \lambda']^\top$

Beweis. Das zulässige lineare System $\mathcal{A}$ repräsentiert $p = 0$ genau dann, wenn $s_1 = (A^{-1}v)_1 = 0$ ist. Da alle Systeme zu $\mathcal{A}$ äquivalent sind, wird dieser Fall für $k' = 1$ erkannt, weil es dann laut Lemma 3.2.3 eine *zulässige* Transformation gibt, sodass die erste linke Minimierungsgleichung erfüllt wird. Nun sei $p \neq 0$. Wir müssen zeigen, dass sowohl die linke Familie s' als auch die rechte Familie t' von $\mathcal{A}' = (u', A', v')$ $\mathbb{K}$-linear unabhängig ist. Sei $n' = \dim(\mathcal{A}')$ und für $k \in \{1, 2, \ldots, n'\}$ bezeichne $s'_{(k)} = (s'_{n'+1-k}, s'_{n'+2-k}, \ldots, s'_{n'})$ die linke und $t'_{(k)} = (t'_1, t'_2, \ldots, t'_k)$ die rechte Teilfamilie. Laut Annahme ist $\mathcal{A}$ ein polynomielles ZLS und daher ist sowohl $s'_{n'} \neq 0$ als auch $t'_1 \neq 0$, das heißt, $s'_{(1)}$ ist $\mathbb{K}$-linear unabhängig und $t'_{(1)}$ ist $\mathbb{K}$-linear unabhängig. Die Schleife beginnt mit $k = 2$. Nur, wenn sowohl $s'_{(k)}$ als auch $t'_{(k)}$ $\mathbb{K}$-linear unabhängig sind, wird k inkrementiert. Für $k = 2$ gibt es den Spezialfall (4.3.7) für die linke Familie. Ansonsten war ein linker (Lemma 3.2.3) oder rechter (Variante von Lemma 3.2.3) Minimierungsschritt erfolgreich und die Dimension des aktuellen ZLS ist echt kleiner als die des vorhergehenden ZLS. Nachdem k von unten beschränkt ist, bricht der Algorithmus in endlich vielen Schritten ab. Wir müssen uns nur noch vergewissern, dass es eine zulässige Transformation gibt, wenn ein Spaltenpaar (T, U) existiert, sodass die rechten Minimierungsgleichungen $\mathcal{R}_k(T, U) = 0$ erfüllt sind. Wäre jedoch die erste Spalte notwendig, um den ersten Eintrag in Spalte k zu eliminieren, kann stattdessen die k-te Zeile verwendet werden. Klarerweise ist $\mathcal{A}'$ polynomiell. □

Bemerkung. Bei der Implementierung kann man zusätzliche Abfragen einbauen, um die auftretenden Gleichungssysteme nicht unnötigerweise mehrfach zu lösen.

Bemerkung 4.3.9. Dieser Algorithmus kann sehr effizient implementiert werden, vorausgesetzt die Zeilen- und Spaltentransformationen werden direkt, das heißt, *ohne* Matrix-Matrix-Multiplikationen, durchgeführt. Sei d die Anzahl der Buchstaben in unserem Alphabet X. Für $\ell = 0, 1, \ldots, d$ bezeichne $A_{ij}^{(\ell)}$ die Teilmatrix entsprechend des Buchstabens x_ℓ und der (aktuellen) Blockzerlegung von $\mathcal{A}^{[k]}$. Die rechten Minimierungsgleichungen $A_{1,1}U + A_{1,2} + T = 0$ können als

$$\begin{bmatrix} I & A_{1,1}^{(0)} \\ I & A_{1,1}^{(1)} \\ \vdots & \vdots \\ I & A_{1,1}^{(d)} \end{bmatrix} \begin{bmatrix} T \\ U \end{bmatrix} = \begin{bmatrix} -A_{1,2}^{(0)} \\ -A_{1,2}^{(1)} \\ \vdots \\ -A_{1,2}^{(d)} \end{bmatrix}$$

mit $2(k-1)$ Unbekannten, $k < n$, geschrieben werden. Über die Gaußsche Elimination erhält man die Komplexität $\mathcal{O}(dn^3)$ um solch ein System zu lösen, siehe [Dem97, Abschnitt 2.3]. Das Erstellen dieses Systems und das Arbeiten mit linearen Matrixbüscheln $\left[\begin{smallmatrix} 0 & u \\ v & A \end{smallmatrix}\right]$ mit $d+1$ quadratischen Koeffizientenmatrizen der Größe $n+1$ (Transformationen, etc.) hat Komplexität $\mathcal{O}(dn^2)$. Nachdem maximal $2(n-1)$ Schritte anfallen, erhält man Gesamtkomplexität $\mathcal{O}(dn^4)$. Der Algorithmus von Cardon und Crochemore [CC80] hat Komplexität $\mathcal{O}(dn^3)$. Mit einem direkten Vergleich muss man vorsichtig sein. Letzterer funktioniert allgemeiner auch für reguläre Elemente, das

heißt, rationale formale Potenzreihen. Jedoch lässt sich die Idee hier direkt für größere Blöcke verallgemeinern, zum Beispiel für eine Block-Zerlegung $\mathcal{A}^{[k,k+1,\ldots,k+l]}$ für $k, l < n$ und sie kann teilweise für *nicht-reguläre* Elemente im freien Schiefkörper verwendet werden, zum Beispiel um das *Wortproblem* zu lösen, das Komplexität $\mathcal{O}(dn^6)$ hat. Abschnitt 4.2 ist dem Wortproblem gewidmet.

Bemerkung. Dieser Abschnitt ist zeitlich vor [Sch18a] entstanden und wird in einer erweiterten Version von [Sch17c] im „Journal of Symbolic Computation“ erscheinen.

Bemerkung 4.3.10. Tatsächlich kann man den Teil mit den Zeilen 12–15 im vorherigen Algorithmus auch weglassen, weil bei einem (polynomiellen) zulässigen linearen System der Dimension 2 $\mathbb{K}$-lineare Abhängigkeit der linken Familie äquivalent zur $\mathbb{K}$-linearen Abhängigkeit der rechten Familie ist, das heißt, der Fall in Zeile 12 würde (indirekt) in Zeile 16 erkannt.

4.4 Verfeinern von Pivotblöcken

Bereits in Beispiel 2.6.1 (Huas Identität) standen wir beim ZLS (2.6.3) vor der Aufgabe, einen Pivotblock zu verfeinern. Obwohl hier offensichtlich ist, welche Transformation notwendig ist, soll die Vorgehensweise in einer systematischen Art gezeigt werden. Davor aber werfen wir noch einen Blick darauf, *wie* dieser 2×2 Block entstanden ist, nämlich durch die Invertierung des durch das System

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & 0 & 1 \\ . & 1 & y \\ . & x & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ . \\ -1 \end{bmatrix} \right)$$

gegebene Element. Vertauscht man die Spalten 2 und 3 sieht man sofort, dass dieses Element das Produkt der beiden *atomaren* zulässigen linearen Systeme

$$\mathcal{A}_1 = \left(\begin{bmatrix} 1 \end{bmatrix}, \begin{bmatrix} x \end{bmatrix}, \begin{bmatrix} -1 \end{bmatrix} \right) \quad \text{und} \quad \mathcal{A}_2 = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} y & 1 \\ 1 & x \end{bmatrix}, \begin{bmatrix} . \\ -1 \end{bmatrix} \right)$$

ist. Wenden wir die minimale Inverse auf

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & 1 & 0 \\ . & y & 1 \\ . & 1 & x \end{bmatrix}, \begin{bmatrix} 0 \\ . \\ -1 \end{bmatrix} \right)$$

an, erhalten wir sofort ein *verfeinertes* (und minimales) ZLS —das sehr einfach in ein polynomielles umgeformt werden kann—, nämlich

$$\mathcal{A}'' = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} -1 & -x & -1 & . \\ . & -1 & -y & . \\ . & . & -1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Die Faktorisierung ist hier also sehr einfach. Nun zurück zur Verfeinerung. Um den zweiten Pivotblock im ZLS (2.6.3) —mit -1 skalierten Zeilen 1, 2 und 3—

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & 1 & x & . \\ . & y & 1 & . \\ . & 1 & 0 & x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right)$$

(gegebenenfalls) verfeinern zu können, suchen wir nach einer (zulässigen Transformation) (P, Q) der Form

$$(P, Q) = \left(\begin{bmatrix} 1 & . & . & . \\ . & \alpha_{2,2} & \alpha_{2,3} & . \\ . & \alpha_{3,2} & \alpha_{3,3} & . \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . & . \\ . & \beta_{2,2} & \beta_{2,3} & . \\ . & \beta_{3,2} & \beta_{3,3} & . \\ . & . & . & 1 \end{bmatrix} \right).$$

Insbesondere müssen die beiden Matrizen P und Q *invertierbar* sein, das heißt, wir benötigen die Bedingungen $\det(P) \neq 0$ und $\det(Q) \neq 0$, wie für die Faktorisierungsmatrizen (3.5.1). Um einen 1×1 Nullblock links unten in $(PAQ)_{2,2}$ erzeugen zu können, müssen wir das folgende *nicht-lineare* Gleichungssystem lösen:

$$\begin{aligned} \alpha_{2,2}\alpha_{3,3} - \alpha_{2,3}\alpha_{3,2} &= 1, \\ \beta_{2,2}\beta_{3,3} - \beta_{2,3}\beta_{3,2} &= 1, \\ \alpha_{3,2}\beta_{3,2} + \alpha_{3,3}\beta_{2,2} &= 0 \quad \text{für 1, und} \\ \alpha_{3,2}\beta_{2,2} &= 0 \quad \text{für } y. \end{aligned}$$

Die letzten beiden Gleichungen erhält man aus der Multiplikation der Transformationsblöcke mit den Koeffizientenmatrizen der Pivotblöcke (irrelevante Gleichungen sind auf der rechten Seite mit „$*$“ gekennzeichnet):

$$\begin{aligned} \begin{bmatrix} \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,2} & \alpha_{3,3} \end{bmatrix} \begin{bmatrix} . & 1 \\ 1 & . \end{bmatrix} \begin{bmatrix} \beta_{2,2} & \beta_{2,3} \\ \beta_{3,2} & \beta_{3,3} \end{bmatrix} &= \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \quad \text{für 1, und} \\ \begin{bmatrix} \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,2} & \alpha_{3,3} \end{bmatrix} \begin{bmatrix} 1 & . \\ . & . \end{bmatrix} \begin{bmatrix} \beta_{2,2} & \beta_{2,3} \\ \beta_{3,2} & \beta_{3,3} \end{bmatrix} &= \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \quad \text{für } y. \end{aligned}$$

Zur Lösung kann man wiederum Gröbner–Shirshov-Basen verwenden, siehe insbesondere Abschnitt 3.7 (Beispiele Faktorisierung). Im allgemeinen ist das bereits für Pivotblöcke der Größe 5 schwierig. Dazu kommt —im Vergleich zum Test auf Vollheit einer Matrix [CR99, Abschnitt 4]—, dass eine Lösung über dem algebraischen Abschluss $\overline{\mathbb{K}}$ des Grundkörpers nicht notwendigerweise eine über $\mathbb{K}$ ist.

Daher ist eine *der* zentralen Anforderungen an einen Minimierungsalgorithmus (wie er im nächsten Abschnitt entwickelt wird) das Respektieren der Verfeinerung eines zulässigen linearen Systems. Und *bevor* man invertiert, sollte man faktorisieren, weil man dann (gegebenenfalls) kleinere Pivotblöcke bekommt. Das heißt natürlich

nicht, dass das immer einfacher ist. Jedenfalls sollte man es mit *linearen* Techniken versuchen. Der Kreativität sind dabei keine Grenzen gesetzt. Selbst Varianten der in Bemerkung 2.3.12 beschriebenen Vorgehensweise könnte man sich überlegen ...

Bemerkung. Eine umfassende(re) Diskussion inklusive algorithmischer Aspekte wird in Kürze folgen [Sch18b].

4.5 Minimieren eines verfeinerten ZLS

Zunächst leiten wir —analog zu Definition 4.3.6— die *linken* beziehungsweise *rechten* Blockminimierungsgleichungen her. Dazu betrachten wir ein zulässiges lineares System $\mathcal{A} = (u, A, v)$ der Dimension n mit $m = \#_{\mathrm{pb}}(\mathcal{A}) \geq 2$ Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Für $1 \leq k < m$ transformieren wir dieses System mit der Blockzeilentransformation $(P, Q) = \big(P(\bar{T}, T), Q(\bar{U}, U)\big)_{\underline{k}}$, nämlich

$$
\begin{aligned}
PAQ &= \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{T} & T \\ . & . & I_{k+1:m} \end{bmatrix} \begin{bmatrix} A_{1:,1:} & A_{1:,k} & A_{1:,:m} \\ . & A_{k,k} & A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{U} & U \\ . & . & I_{k+1:m} \end{bmatrix} \\
&= \begin{bmatrix} A_{1:,1:} & A_{1:,k} & A_{1:,:m} \\ . & \bar{T}A_{k,k} & \bar{T}A_{k,:m} + TA_{:m,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{U} & U \\ . & . & I_{k+1:m} \end{bmatrix} \\
&= \begin{bmatrix} A_{1:,1:} & A_{1:,k}\bar{U} & A_{1:,k}U + A_{1:,:m} \\ . & \bar{T}A_{k,k}\bar{U} & \bar{T}A_{k,k}U + \bar{T}A_{k,:m} + TA_{:m,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \quad \text{und} \\
Pv &= \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{T} & T \\ . & . & I_{k+1:m} \end{bmatrix} \begin{bmatrix} v_{\underline{1:}} \\ v_{\underline{k}} \\ v_{\underline{:m}} \end{bmatrix} = \begin{bmatrix} v_{\underline{1:}} \\ \bar{T}v_{\underline{k}} + Tv_{\underline{:m}} \\ v_{\underline{:m}} \end{bmatrix}.
\end{aligned}
$$

Analog zu den linken Minimierungsgleichungen (4.3.3) erhalten wir eine *hinreichende* Bedingung für $(Q^{-1}s)_{\underline{k}} = 0^{n_k \times 1}$, nämlich die Existenz von Matrizen $T, U \in \mathbb{K}^{n_k \times n_{k+1:m}}$ und invertierbaren Matrizen $\bar{T}, \bar{U} \in \mathbb{K}^{n_k \times n_k}$, sodass

$$
\bar{T}A_{k,k}U + \bar{T}A_{k,:m} + TA_{:m,:m} = 0^{n_k \times n_{k+1:m}} \quad \text{und} \quad \bar{T}v_{\underline{k}} + Tv_{\underline{:m}} = 0^{n_k \times 1}
$$

gilt. Nachdem $\bar{T}$ invertierbar ist (als Diagonalblock einer invertierbaren Matrix P), ist diese Bedingung äquivalent zur Existenz von Matrizen $T', U \in \mathbb{K}^{n_k \times n_{k+1:m}}$ sodass

$$
A_{k,k}U + A_{k,:m} + \underbrace{\bar{T}^{-1}T}_{=:T'} A_{:m,:m} = 0^{n_k \times n_{k+1:m}} \quad \text{und} \quad v_{\underline{k}} + \underbrace{\bar{T}^{-1}T}_{=:T'} v_{\underline{:m}} = 0^{n_k \times 1} \tag{4.5.1}
$$

gilt. Mit der Blockspaltentransformation $(P,Q) = \big(P(\bar{T},T), Q(\bar{U},U)\big)^{\underline{k}}$ erhalten wir

$$\begin{aligned} PAQ &= \begin{bmatrix} I_{1:k-1} & T & . \\ . & \bar{T} & . \\ . & . & I_{k+1:m} \end{bmatrix} \begin{bmatrix} A_{1:,1:} & A_{1:,k} & A_{1:,:m} \\ . & A_{k,k} & A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} I_{1:k-1} & U & . \\ . & \bar{U} & . \\ . & . & I_{k+1:m} \end{bmatrix} \\ &= \begin{bmatrix} A_{1:,1:} & A_{1:,k} + TA_{k,k} & A_{1:,:m} + TA_{k,:m} \\ . & \bar{T}A_{k,k} & \bar{T}A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} I_{1:k-1} & U & . \\ . & \bar{U} & . \\ . & . & I_{k+1:m} \end{bmatrix} \\ &= \begin{bmatrix} A_{1:,1:} & A_{1:,1:}U + A_{1:,k}\bar{U} + TA_{k,k}\bar{U} & A_{1:,:m} + TA_{k,:m} \\ . & \bar{T}A_{k,k}\bar{U} & \bar{T}A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix} \end{aligned}$$

und damit —analog zu den rechten Minimierungsgleichungen (4.3.5)— eine *hinreichende* Bedingung für $(tP^{-1})_{\underline{k}} = 0^{1\times n_k}$, nämlich die Existenz von Matrizen $T, U' \in \mathbb{K}^{n_{1:k-1}\times n_k}$ sodass folgendes gilt:

$$A_{1:,1:}\underbrace{U\bar{U}^{-1}}_{=:U'} + A_{1:,k} + TA_{k,k} = 0^{n_{1:k-1}\times n_k}. \tag{4.5.2}$$

Bemerkung. Eine Variante des *linearen* Gleichungssystems (4.5.1) kommt auch in Lemma 4.2.2 beziehungsweise Satz 4.2.3 (Lineares Wortproblem) im vorherigen Abschnitt vor.

Bemerkung 4.5.3 (Erweitertes ZLS [Sch18a, Bemerkung 4.3])**.** In manchen Fällen ist es notwendig, ein *erweitertes* ZLS zu betrachten, um alle erforderlichen *linken* Minimierungsschritte ausführen zu können, zum Beispiel für $f^{-1}f$ wenn f vom Typ $(1,1)$ ist. Sei $\mathcal{A} = (u,A,v) = (1,A,\lambda)$ ein ZLS mit $m = \#_{\mathrm{pb}}(\mathcal{A}) \geq 2$ Pivotblöcken und $k = 1$. Die „erweiterte“ Blockzerlegung ist dann (die Blockzeile $A_{1:,1:}$ verschwindet)

$$\mathcal{A}^{[\underline{k}]} = \left(\left[\begin{array}{c|cc} 1 & A_{0,k} & . \\ \hline . & A_{k,k} & A_{k,:m} \\ . & . & A_{:m,:m} \end{array}\right], \left[\begin{array}{c} . \\ \hline . \\ v_{\underline{:m}} \end{array}\right] \right)$$

mit $A_{0,k} = [-1,0,\ldots,0]$. Die erste Zeile in $\mathcal{A}^{[\underline{1}]}$ wird nur indirekt (über zulässige Spaltenoperationen) verändert und bleibt daher skalar, man kann sie deswegen (gegebenenfalls) sehr einfach wieder eliminieren. Das ist in Beispiel 4.5.6 illustriert.

Notation. Gegeben ein zulässiges lineares System $\mathcal{A}$, bezeichnen wir mit $\tilde{\mathcal{A}} = \mathcal{A}^{[+0]}$ das erweiterte (zu $\mathcal{A}$ äquivalente) ZLS. Umgekehrt ist $\tilde{\mathcal{A}}^{[-0]} = \big(\mathcal{A}^{[+0]}\big)^{[-0]} = \mathcal{A}$. Die zusätzliche Zeile und Spalte wird mit 0 indiziert. Wurde $\tilde{\mathcal{A}}$ zulässig transformiert, bezeichnet $\tilde{\mathcal{A}}^{[-0]}$ ein ZLS.

Definition 4.5.4 (Minimierungsgleichungen und -transformationen [Sch18a, Definition 4.4])**.** Sei $\mathcal{A} = (u,A,v)$ ein ZLS der Dimension n mit $m = \#_{\mathrm{pb}}(\mathcal{A}) \geq 2$ Pivotblöcken der Größe $n_i = \dim_i(\mathcal{A})$. Für $k \in \{1,2,\ldots,m-1\}$ und die Blockzerlegung $\mathcal{A}^{[\underline{k}]}$ heißen die Gleichungen (4.5.1),

$$A_{k,k}U + A_{k,:m} + TA_{:m,:m} = 0^{n_k\times n_{k+1:m}} \quad \text{und} \quad v_{\underline{k}} + Tv_{\underline{:m}} = 0^{n_k\times 1}$$

Blocktransformation

$$\left(\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \alpha_{2,4} \\ 0 & 0 & \alpha_{3,3} & \alpha_{3,4} \\ 0 & 0 & \alpha_{4,3} & \alpha_{4,4} \end{bmatrix}, \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} & \beta_{1,4} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & \beta_{3,3} & \beta_{3,4} \\ 0 & 0 & \beta_{4,3} & \beta_{4,4} \end{bmatrix}\right)$$

Allgemeine Blockzeilentransformation, Definition 4.1.5

$$\left(\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \alpha_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}, \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} & \beta_{1,4} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}\right)$$

Spezielle Blockzeilentransformation

$$\left(\begin{bmatrix} 1 & . & \alpha_{1,3} & \alpha_{1,4} \\ . & 1 & \alpha_{2,3} & \alpha_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & \beta_{1,3} & \beta_{1,4} \\ . & 1 & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}\right)$$

Allgemeine Blockspaltentransformation, Definition 4.1.5

$$\left(\begin{bmatrix} 1 & . & \alpha_{1,3} & \alpha_{1,4} \\ . & 1 & \alpha_{2,3} & \alpha_{2,4} \\ 0 & 0 & \alpha_{3,3} & \alpha_{3,4} \\ 0 & 0 & \alpha_{4,3} & \alpha_{4,4} \end{bmatrix}, \begin{bmatrix} 1 & . & \beta_{1,3} & \beta_{1,4} \\ . & 1 & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & \beta_{3,3} & \beta_{3,4} \\ 0 & 0 & \beta_{4,3} & \beta_{4,4} \end{bmatrix}\right)$$

Spezielle Blockspaltentransformation

$$\left(\begin{bmatrix} 1 & . & \alpha_{1,3} & \alpha_{1,4} \\ . & 1 & \alpha_{2,3} & \alpha_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & \beta_{1,3} & \beta_{1,4} \\ . & 1 & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & 1 & . \\ 0 & 0 & . & 1 \end{bmatrix}\right)$$

Blockfaktorisierungstransformation, Bemerkung 3.7.8

$$\left(\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & . \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & . \\ 0 & 0 & \alpha_{3,3} & . \\ 0 & 0 & \alpha_{4,3} & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . & . \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \beta_{2,4} \\ 0 & 0 & \beta_{3,3} & \beta_{3,4} \\ 0 & 0 & \beta_{4,3} & \beta_{4,4} \end{bmatrix}\right)$$

Abbildung 4.2: Die *invertierbaren* (nicht notwendigerweise zulässigen) Transformationsmatrizen $P = (\alpha_{ij}), Q = (\beta_{ij}) \in \mathbb{K}^{(n+1)\times(n+1)}$ als Paar (P, Q) für ein zulässiges lineares System der Dimension 4 mit zwei Pivotblöcken der Größe 2. Die Invertierbarkeit muss gegebenenfalls mit $\det(P) \neq 0$ und $\det(Q) \neq 0$ sichergestellt werden.

bezüglich der speziellen Blockzeilentransformation $\big(P(T), Q(U)\big)_{\underline{k}}$ *linke Blockminimierungsgleichungen*. Sie werden mit $\mathcal{L}_{\underline{k}} = \mathcal{L}_{\underline{k}}(\mathcal{A})$ bezeichnet. Eine Lösung durch das Blockzeilenpaar (T, U) wird mit $\mathcal{L}_{\underline{k}}(T, U) = 0$ bezeichnet. Für $k \in \{2, 3, \ldots, m\}$ und die Blockzerlegung $\mathcal{A}^{[\underline{k}]}$ heißen die Gleichungen (4.5.2),

$$A_{1:,1:}U + A_{1:,k} + TA_{k,k} = 0^{n_{1:k-1} \times n_k}$$

bezüglich der speziellen Blockspaltentransformation $\big(P(T), Q(U)\big)^{\underline{k}}$ *rechte Blockminimierungsgleichungen*. Sie werden mit $\mathcal{R}_{\underline{k}} = \mathcal{R}_{\underline{k}}(\mathcal{A})$ bezeichnet. Eine Lösung durch das Blockspaltenpaar (T, U) wird mit $\mathcal{R}_{\underline{k}}(T, U) = 0$ bezeichnet. Siehe auch Definition 4.3.6.

Beispiel 4.5.5 (Nicht-kommutatives Kürzen, Verschmelzen von Atomen)**.** Wir greifen das Beispiel aus Abschnitt 3.4 wieder auf: Seien $f = 1 - xy$ und $g = (1 - zy)^{-1}$ gegeben durch die *minimalen* linearen zulässigen Systeme

$$\mathcal{A}_f = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -1 \\ . & 1 & y \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right) \quad \text{bzw.} \quad \mathcal{A}_g = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} y & 1 \\ 1 & z \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right).$$

A priori ist nicht klar, dass fg ein Atom ist. Laut Proposition 2.3.6 (Multiplikation Typ $(1, *)$) ist

$$\mathcal{A} = (u, A, v) = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & y & . \\ . & . & y & 1 \\ . & . & 1 & z \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right)$$

ein ZLS für fg. Laut Konstruktion ist die linke Teilfamilie $s_{\underline{3}} = (s_3, s_4)$ der linken Familie $s = A^{-1}v$ $\mathbb{K}$-linear unabhängig. Hier sieht man sofort, dass man Zeile 2 und Spalte 2 eliminieren kann, *nachdem* man Zeile 3 von Zeile 2 subtrahiert und Spalte 2 zu Spalte 4 addiert hat:

$$\mathcal{A}' = (u', A', v') = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -1 & -x \\ . & 1 & 0 & 0 \\ . & . & y & 1 \\ . & . & 1 & z \end{bmatrix}, \begin{bmatrix} . \\ 0 \\ . \\ 1 \end{bmatrix} \right).$$

Während es für die Bestimmung des linken ggT zweier Polynome genügt, Atome zu „kürzen“ (siehe das Beispiel in Abschnitt A.1), kann es im allgemeinen Fall passieren, dass (gewöhnungsbedürftigerweise) zwei Atome zu *einem* (neuen) Atom $h := fg$ „verschmelzen“. Daher gilt hier insbesondere weder $f \,|_{\mathrm{l}}\, fg$ noch $g \,|_{\mathrm{r}}\, fg$.

Im folgenden Beispiel schauen wir uns genauer an, welche Rolle die Faktorisierung spielt und wie wir die Anwendung möglicherweise nicht-linearer Techniken vermeiden können. Die einzelnen Schritte sind sehr detailliert erklärt und entsprechen (mit Ausnahme der Lösung von linearen Gleichungssystemen) denen des Algorithmus.

Beispiel 4.5.6 ([Sch18a, Beispiel 4.5])**.** Für $f = x^{-1}(1-xy)^{-1}$ und $g = x$ betrachten wir $h = fg$ gegeben durch das (nicht-minimale) ZLS (konstruiert laut Proposition 2.3.9)

$$\mathcal{A} = (u, A, v) = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} x & 1 & . & . \\ . & y & -1 & . \\ . & -1 & x & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right),$$

dessen Pivotblöcke *verfeinert* sind. Tatsächlich gibt es hier die (zulässige) Transformation (mit $T = 0$, $U = 1$ und *invertierbaren* Blöcken $\bar{T}, \bar{U} \in \mathbb{K}^{3\times 3}$)

$$(P, Q) = \left(\begin{bmatrix} 1 & 0 & 0 & . \\ 0 & 1 & 0 & . \\ 1 & 0 & 1 & T \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & . \\ 0 & 1 & 0 & . \\ -1 & 0 & 1 & U \\ . & . & . & 1 \end{bmatrix} \right),$$

deren Anwendung auf $\mathcal{A}$ zum ZLS

$$P\mathcal{A}Q = \mathcal{A}' = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} x & 1 & . & . \\ 1 & y & -1 & -1 \\ 0 & 0 & x & 0 \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right)$$

führt, in dem man Zeile 3 und Spalte 3 (und in weiterer Folge —nach einer entsprechenden Zeilenoperation— auch die letzte Zeile und Spalte) eliminieren kann.

Tatsächlich ist das aber viel einfacher möglich. Zunächst stellt man fest, dass die linke Teilfamilie $s_{\underline{2:3}}$ $\mathbb{K}$-linear unabhängig ist. Auch die rechte Teilfamilie $t_{\underline{1:2}}$ ist $\mathbb{K}$-linear unabhängig. Für die linke Familie bezüglich des ersten Pivotblocks betrachten wir das erweiterte ZLS (siehe auch Bemerkung 4.5.3)

$$\left[\begin{array}{c|cccc} 1 & -1 & . & . & . \\ \hline . & x & 1 & . & . \\ . & . & y & -1 & . \\ . & . & -1 & x & -x \\ . & . & . & . & 1 \end{array}\right] s = \left[\begin{array}{c} . \\ \hline . \\ . \\ . \\ 1 \end{array}\right]$$

von $\mathcal{A}$, die obere Zeile und die linke Spalte indizieren wir mit Null. Nun addieren wir Zeile 3 zu Zeile 1, subtrahieren Spalte 1 von Spalte 3 und addieren Spalte 1 zu Spalte 4:

$$\left[\begin{array}{c|cccc} 1 & -1 & . & 1 & -1 \\ \hline . & x & 0 & 0 & 0 \\ . & . & y & -1 & . \\ . & . & -1 & x & -x \\ . & . & . & . & 1 \end{array}\right] s = \left[\begin{array}{c} . \\ \hline 0 \\ . \\ . \\ 1 \end{array}\right].$$

Nun können wir Zeile 1 und Spalte 1 entfernen:

$$\left[\begin{array}{c|ccc} 1 & . & 1 & -1 \\ \hline . & y & -1 & . \\ . & -1 & x & -x \\ . & . & . & 1 \end{array}\right] s = \left[\begin{array}{c} . \\ \hline . \\ . \\ 1 \end{array}\right]. \tag{4.5.7}$$

Bevor wir den letzten (rechten) Minimierungsschritt durchführen, wandeln wir das erweiterte ZLS zurück in ein „normales“, indem wir die Spalten 1 und 2 vertauschen, die (neue) Spalte 1 mit -1 skalieren und von Spalte 3 subtrahieren:

$$\left[\begin{array}{c|ccc} 1 & -1 & . & 0 \\ \hline . & 1 & y & -1 \\ . & -x & -1 & 0 \\ . & . & . & 1 \end{array}\right] s = \left[\begin{array}{c} . \\ \hline . \\ . \\ 1 \end{array}\right].$$

(Gegebenenfalls können solange rechte Minimierungsschritte ausgeführt werden, bis man bei einem Pivotblock angelangt ist, der einen korrespondierenden Nicht-Null-Eintrag in Zeile 0 hat.) Nun kann man Zeile 0 und Spalte 0 wieder entfernen. Der letzte Schritt zu einem minimalen ZLS für $fg = (1 - yx)^{-1}$ ist trivial:

$$\begin{bmatrix} 1 & y & 0 \\ -x & -1 & 0 \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} 1 \\ . \\ 1 \end{bmatrix}.$$

Nach dem Entfernen der letzten Zeile und Spalte vertauscht man die verbleibenden beiden Zeilen um ein *standardisiertes* ZLS zu erhalten.

Zumindest eine Frage sollte sich jetzt ergeben haben: *Wie* prüft man —für einen gegebenen Blockindex k— die $\mathbb{K}$-lineare Unabhängigkeit der linken (bzw. rechten) Teilfamilie $s_{\underline{k:m}}$ (bzw. $t_{\underline{1:k}}$) im allgemeinen, vorausgesetzt $s_{\underline{k+1:m}}$ (bzw. $t_{\underline{1:k-1}}$) ist $\mathbb{K}$-linear unabhängig? Um diese zu beantworten ist ein wenig Vorbereitung notwendig.

Lemma 4.5.8 ([CR99, Lemma 1.2])**.** *Sei $f \in \mathbb{F}$ gegeben durch die lineare Darstellung $\pi_f = (u, A, v)$ der Dimension n. Dann ist $f = 0$ genau dann, wenn es invertierbare Matrizen $P, Q \in \mathbb{K}^{n\times n}$ gibt, sodass*

$$P\pi_f Q = \left(\begin{bmatrix} \tilde{u}_{\underline{1}} & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}_{1,1} & 0 \\ \tilde{A}_{2,1} & \tilde{A}_{2,2} \end{bmatrix}, \begin{bmatrix} 0 \\ \tilde{v}_{\underline{2}} \end{bmatrix} \right)$$

für quadratische Matrizen $\tilde{A}_{1,1}$ und $\tilde{A}_{2,2}$ ist.

Satz 4.5.9 (Linke Blockminimierung [Sch18a, Satz 4.8])**.** *Sei $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ ein ZLS der Dimension n mit $m = \#_{pb}(\mathcal{A}) \geq 2$ Pivotblöcken der Größen $n_i = \dim_i(\mathcal{A})$. Sei $k \in \{1, 2, \ldots, m-1\}$ so, dass die linke Teilfamilie $s_{\underline{k+1:m}}$ bezüglich der Blockzerlegung $\mathcal{A}^{[\underline{k}]}$ $\mathbb{K}$-linear unabhängig ist, während $s_{\underline{k:m}}$ $\mathbb{K}$-linear abhängig ist.*

Dann gibt es eine Blockzeilentransformation $(P,Q) = \big(P(\bar{T},T), Q(\bar{U},U)\big)_{\underline{k}}$, *sodass* $\tilde{\mathcal{A}} = P\mathcal{A}Q$ *die Form*

$$\begin{bmatrix} A_{1:,1:} & \tilde{A}_{1:,k'} & \tilde{A}_{1:,k''} & \tilde{A}_{1:,:m} \\ . & \tilde{A}_{k',k'} & 0 & 0 \\ . & \tilde{A}_{k'',k'} & \tilde{A}_{k'',k''} & \tilde{A}_{k'',:m} \\ . & . & . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} \tilde{s}_{\underline{1:}} \\ 0 \\ \tilde{s}_{\underline{k}''} \\ s_{\underline{:m}} \end{bmatrix} = \begin{bmatrix} . \\ . \\ . \\ v_{\underline{:m}} \end{bmatrix} \tag{4.5.10}$$

hat. Ist der Pivotblock $A_{k,k}$ verfeinert, *gibt es eine spezielle Blockzeilentransformation* $(P,Q) = \big(P(T), Q(U)\big)_{\underline{k}}$, *sodass die linken Blockminimierungsgleichungen*

$$A_{k,k}U + A_{k,:m} + TA_{:m,:m} = 0^{n_k \times n_{k+1:m}} \quad \textit{und} \quad v_{\underline{k}} + Tv_{\underline{:m}} = 0^{n_k \times 1}$$

erfüllt werden.

Beweis. Wir beziehen uns auf die Blockzerlegung

$$\mathcal{A}^{[\underline{k}]} = \left(\begin{bmatrix} u_{\underline{1:}} & . & . \end{bmatrix}, \begin{bmatrix} A_{1:,1:} & A_{1:,k} & A_{1:,:m} \\ . & A_{k,k} & A_{k,:m} \\ . & . & A_{:m,:m} \end{bmatrix}, \begin{bmatrix} . \\ . \\ v_{\underline{:m}} \end{bmatrix} \right).$$

Wegen der $\mathbb{K}$-linearen Abhängigkeit der linken Teilfamilie $s_{\underline{k:m}}$ gibt es eine invertierbare Matrix $\tilde{Q}$ mit den Blöcken $\bar{U}^\circ \in \mathbb{K}^{n_k \times n_k}$ und $U^\circ \in \mathbb{K}^{n_k \times n_{k+1:m}}$, sodass $(\tilde{Q}^{-1}s)_{n_{1:k-1}+1} = 0$ ist, das heißt, die erste Komponente in $s_{\underline{k}}$ eliminiert werden kann. Seien

$$A' = \begin{bmatrix} A_{k,k} & A_{k,:m} \\ . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} \bar{U}^\circ & U^\circ \\ . & I_{k+1:m} \end{bmatrix} \quad \text{und} \quad v' = \begin{bmatrix} . \\ v_{\underline{:m}} \end{bmatrix}.$$

Dann ist $\mathcal{A}' = (u', A', v')$ ein ZLS für $0 \in \mathbb{F}$ und wir können Lemma 4.5.8 anwenden, um eine Transformation

$$(P',Q') = \left(\begin{bmatrix} \bar{T} & T \\ T_{:m,k} & T_{:m,:m} \end{bmatrix}, \begin{bmatrix} \bar{U}' & U' \\ U_{:m,k} & U_{:m,:m} \end{bmatrix} \right)$$

zu erhalten, sodass $P'A'Q'$ rechts oben einen entsprechenden Nullblock hat und —ohne Beschränkung der Allgemeinheit— $P'v' = v'$ ist. Klarerweise können wir $U_{:m,:m} = I_{k+1:m}$ wählen. Und da $s_{\underline{:m}}$ $\mathbb{K}$-linear unabhängig ist, ist der Nullblock in

$$\begin{aligned} P'A' &= \begin{bmatrix} \bar{T} & T \\ T_{:m,k} & T_{:m,:m} \end{bmatrix} \begin{bmatrix} A'_{k,k} & A'_{k,:m} \\ . & A_{:m,:m} \end{bmatrix} \\ &= \begin{bmatrix} \bar{T}A'_{k,k} & \bar{T}A'_{k,:m} + TA_{:m,:m} \\ T_{:m,k}A'_{k,k} & T_{:m,k}A'_{k,:m} + T_{:m,:m}A_{:m,:m} \end{bmatrix} \end{aligned}$$

unabhängig von $T_{:m,k}$ und $T_{:m,:m}$, also können wir $T_{:m,k} = 0$ und $T_{:m,:m} = I_{k+1:m}$ wählen. Nun ist klar, dass die Spalten im linken unteren Block von

$$\begin{aligned} P'A'Q' &= \begin{bmatrix} \bar{T}A'_{k,k} & \bar{T}A'_{k,:m} + TA_{:m,:m} \\ . & A_{:m,:m} \end{bmatrix} \begin{bmatrix} \bar{U}' & U' \\ U_{:m,k} & I \end{bmatrix} \\ &= \begin{bmatrix} \bar{T}A'_{k,k}\bar{U}' + (\bar{T}A'_{k,:m} + TA_{:m,:m})U_{:m,k} & \bar{T}A'_{k,k}U' + \bar{T}A'_{k,:m} + TA_{:m,:m} \\ A_{:m,:m}U_{:m,k} & A_{:m,:m} \end{bmatrix} \end{aligned}$$

Linearkombinationen der Spalten von $A_{:m,:m}$ sind und wir deshalb auch $U_{:m,k} = 0$ annehmen können. Sei nun $\bar{U} = \bar{U}^\circ \bar{U}'$ und $U = \bar{U}^\circ U' + U^\circ$. Dann hat PAQ für die (für $k > 1$ zulässige) Blockzeilentransformation

$$(P, Q) = \left(\begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{T} & T \\ . & . & I_{k+1:m} \end{bmatrix}, \begin{bmatrix} I_{1:k-1} & . & . \\ . & \bar{U} & U \\ . & . & I_{k+1:m} \end{bmatrix} \right)$$

die gewünschte Form (4.5.10). Für den zweiten Teil müssen wir zuerst zeigen, dass *jeder* Eintrag in $s_{\underline{k}}$ durch eine Linearkombination von Komponenten in $s_{\underline{k+1:m}}$ eliminiert werden kann, das heißt, $n_{k''} = 0$ ist. Wir nehmen gegenteilig $n_{k''} > 0$ an. Aber dann hätte $\bar{T} A_{k,k} \bar{U}$ laut (4.5.10) einen rechten oberen Nullblock der Größe $(n_k - n_{k''}) \times n_{k''}$ und damit (nach entsprechender Permutation) einen linken unteren dieser Größe, was ein Widerspruch zur Annahme eines verfeinerten Pivotblocks $A_{k,k}$ wäre. Daher gibt es eine Matrix $U \in \mathbb{K}^{n_k \times n_{k+1:m}}$ sodass $s_{\underline{k}} - U[s_{\underline{k+1}}, \ldots, s_{\underline{m}}] = 0$ ist. Laut Annahme ist $v_{\underline{k}} = 0$. Nun können wir —so, wie in Lemma 3.2.3— Lemma 2.3.2 mit dem ZLS $(1, A_{:m,:m}, \lambda)$ und $B = -A_{k,k}U - A_{k,:m}$ (und $s_{\underline{:m}}$) anwenden. Deshalb gibt es eine Matrix $T \in \mathbb{K}^{n_k \times n_{k+1:m}}$, die $A_{k,k}U + A_{k,:m} + TA_{:m,:m} = 0$ erfüllt. Da die letzte Spalte von T Null ist, gilt auch $Tv_{\underline{:m}} = 0$. Mit $\bar{T} = \bar{U} = I_{n_k}$ ist (P, Q) die gesuchte spezielle Blockzeilentransformation. □

Bemerkung. Für den Beweis des zweiten Teils des Satzes kann man alternativ Lemma 4.5.8 verwenden. Zwar ist es viel allgemeiner, aber die Möglichkeit der Verwendung linearer Techniken ist nicht so klar ersichtlich.

Bemerkung. Man beachte, dass die linke Teilfamilie $(\tilde{s}_{\underline{k}''}, s_{\underline{:m}})$ nicht notwendigerweise $\mathbb{K}$-linear unabhängig ist. Gegebenenfalls kann man —nachdem man die Blockzeile und -spalte k' entfernt hat— den Satz erneut anwenden.

Bemerkung. Für $k = 1$ muss man gegebenenfalls ein erweitertes ZLS betrachten, siehe Bemerkung 4.5.3.

Bemerkung 4.5.11. Unter Voraussetzung der $\mathbb{K}$-linearen Unabhängigkeit der linken Teilfamilie $s_{\underline{k+1:m}}$ und eines verfeinerten Pivotblocks $A_{k,k}$ bedeutet der zweite Teil des vorherigen Satzes nichts geringeres als die Möglichkeit der Prüfung der $\mathbb{K}$-linearen (Un-)Abhängigkeit der linken Teilfamilie $s_{\underline{k:m}}$ mit *linearen* Techniken! Dass ein polynomielles ZLS trivialerweise verfeinert ist, haben wir uns also bereits implizit im Abschnitt 4.3 zunutze gemacht.

Satz 4.5.12 (Rechte Blockminimierung [Sch18a, Satz 4.10])**.** *Sei* $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ *ein ZLS der Dimension* n *mit* $m = \#_{pb}(\mathcal{A}) \geq 2$ *Pivotblöcken der Größen* $n_i = \dim_i(\mathcal{A})$. *Sei* $k \in \{2, 3, \ldots, m\}$ *so, dass die rechte Teilfamilie* $t_{\underline{1:k-1}}$ *bezüglich der Blockzerlegung* $\mathcal{A}^{[\underline{k}]}$ $\mathbb{K}$*-linear unabhängig ist, während* $t_{\underline{1:k}}$ $\mathbb{K}$*-linear abhängig ist. Dann gibt es eine Blockspaltentransformation* $(P, Q) = \big(P(\bar{T}, T), Q(\bar{U}, U)\big)^{\underline{k}}$, *sodass*

$\tilde{\mathcal{A}} = P\mathcal{A}Q$ die Form

$$\begin{bmatrix} u_{\underline{1}} & . & . & . \end{bmatrix} = \begin{bmatrix} t_{\underline{1}} & \tilde{t}_{\underline{2'}} & 0 & \tilde{t}_{\underline{3}} \end{bmatrix} \begin{bmatrix} A_{1,1} & \tilde{A}_{1,2'} & 0 & \tilde{A}_{1,3} \\ . & \tilde{A}_{2',2'} & 0 & \tilde{A}_{2',3} \\ . & \tilde{A}_{2'',2'} & \tilde{A}_{2'',2''} & \tilde{A}_{2'',3} \\ . & . & . & A_{3,3} \end{bmatrix} \tag{4.5.13}$$

hat. Ist der Pivotblock $A_{k,k}$ verfeinert, gibt es eine spezielle Blockspaltentransformation $(P, Q) = \big(P(T), Q(U)\big)^{\underline{k}}$, sodass die rechten Blockminimierungsgleichungen

$$A_{1:,1:}U + A_{1:,k} + TA_{k,k} = 0^{n_{1:k-1} \times n_k}$$

erfüllt werden.

Wenn man dann für die Minimierung abwechselnd linke und rechte Blockminimierungsschritte ausführt, das heißt, die Sätze 4.5.9 und 4.5.12 anwendet, muss man darauf achten, dass die jeweilige $\mathbb{K}$-lineare Unabhängigkeit der entsprechenden Teilfamilie gewährleistet ist. Das wird im folgenden Beispiel illustriert.

Beispiel 4.5.14. Sei $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ ein ZLS mit $m = 5$ Pivotblöcken. Für $k' = 2$ sei die linke Teilfamilie $s_{\underline{k+1:m}}$ $\mathbb{K}$-linear unabhängig und wir nehmen an, dass es eine spezielle Blockzeilentransformation $(P, Q) = \big(P(T), Q(U)\big)$ gibt, sodass die linken Blockminimierungsgleichungen erfüllt werden, das heißt $P\mathcal{A}Q$ die Form

$$\begin{bmatrix} A_{1,1} & A_{1,2} & \tilde{A}_{1,3} & \tilde{A}_{1,4} & \tilde{A}_{1,5} \\ . & A_{2,2} & 0 & 0 & 0 \\ . & . & A_{3,3} & A_{3,4} & A_{3,5} \\ . & . & . & A_{4,4} & A_{4,5} \\ . & . & . & . & A_{5,5} \end{bmatrix} \begin{bmatrix} \tilde{s}_{\underline{1}} \\ 0 \\ s_{\underline{3}} \\ s_{\underline{4}} \\ s_{\underline{5}} \end{bmatrix} = \begin{bmatrix} . \\ 0 \\ . \\ . \\ v_{\underline{5}} \end{bmatrix}$$

hat. Ist die rechte Teilfamilie $t_{\underline{1:3}}$ $\mathbb{K}$-linear unabhängig, so gilt das *nicht* notwendigerweise für die rechte Teilfamilie $t'_{\underline{1:3}}$ des verkleinerten ZLS $\mathcal{A}' = \big(P\mathcal{A}Q\big)^{[-\underline{k'}]}$. Das heißt, man muss Satz 4.5.12 auf $\mathcal{A}'$ mit $k = 3$ anwenden um das zu „erneut“ prüfen.

Bemerkung. Der Beweis des folgenden Algorithmus unterscheidet sich —abgesehen von der Verwendung eines erweiterten ZLS— nicht wesentlich von dem des Algorithmus 4.3.8 für polynomielle zulässige lineare Systeme, wo die Anzahl der Pivotblöcke mit der Dimension übereinstimmt. Die Rolle von Lemma 3.2.3 (und seiner Variante für die rechte Familie) übernimmt hier Satz 4.5.9 (beziehungsweise Satz 4.5.12).

Algorithmus 4.5.15 (Minimieren eines verfeinerten ZLS [Sch18a, Algorithmus 4.13])**.**

Eingabe: $\mathcal{A} = (u, A, v) = (1, A, \lambda)$ *verfeinertes ZLS (für ein Element* f*)*
mit $m = \#_{pb}(\mathcal{A}) \geq 2$ *Pivotblöcken der Größen* $n_i = \dim_i(\mathcal{A})$ *und*
$\mathbb{K}$*-linear unabhängigen Teilfamilien* $s_{\underline{m}}$ *und* $t_{\underline{1}}$*.*
Ausgabe: $\mathcal{A}' = (,,)$*, wenn* $f = 0$ *ist, oder*
ein minimales *verfeinertes ZLS* $\mathcal{A}' = (u', A', v') = (1, A', \lambda')$*, wenn* $f \neq 0$ *ist.*

1: $k := 2$
2: **while** $k \leq \#_{\mathbf{pb}}(\mathcal{A})$ **do**
3: $m := \#_{\mathbf{pb}}(\mathcal{A})$
4: $k' := m + 1 - k$
Ist die linke Teilfamilie $(s_{\underline{k'}}, \overbrace{s_{\underline{k'+1}}, \ldots, s_{\underline{m}}}^{\text{lin. unabhängig}})$ $\mathbb{K}$-linear abhängig?
5: **if** $\exists T, U \in \mathbb{K}^{n_k \times n_{k+1:m}}$ zulässig : $\mathcal{L}_{\underline{k'}}(\mathcal{A}) = \mathcal{L}_{\underline{k'}}(T, U) = 0$ **then**
6: **if** $k' = 1$ **then**
7: **return** $(,,)$
endif
8: $\mathcal{A} := \big(P(T)\mathcal{A}Q(U)\big)^{[-\underline{k'}]}$
9: **if** $k > \max\{2, \frac{m+1}{2}\}$ **then**
10: $k := k - 1$
endif
11: **continue**
endif
12: **if** $k' = 1$ **and** $\exists T, U \in \mathbb{K}^{n_k \times n_{k+1:m}} : \mathcal{L}_{\underline{k'}}(\mathcal{A}^{[+0]}) = \mathcal{L}_{\underline{k'}}(T, U) = 0$ **then**
13: $\tilde{\mathcal{A}} := \big(P(T)\mathcal{A}^{[+0]}Q(U)\big)^{[-\underline{k'}]}$
14: $\mathcal{A} := \tilde{\mathcal{A}}^{[-0]}$
15: **continue**
endif
Ist die rechte Teilfamilie $(\overbrace{t_{\underline{1}}, \ldots, t_{\underline{k-1}}}^{\text{lin. unabhängig}}, t_{\underline{k}})$ $\mathbb{K}$-linear abhängig?
16: **if** $\exists T, U \in \mathbb{K}^{n_{k-1:m} \times n_k}$ zulässig : $\mathcal{R}_{\underline{k}}(\mathcal{A}) = \mathcal{R}_{\underline{k}}(T, U) = 0$ **then**
17: $\mathcal{A} := \big(P(T)\mathcal{A}Q(U)\big)^{[-\underline{k}]}$
18: **if** $k > \max\{2, \frac{m+1}{2}\}$ **then**
19: $k := k - 1$
endif
20: **continue**
endif
21: $k := k + 1$
done
22: **return** $P\mathcal{A}$, mit P, sodass $Pv = [0, \ldots, 0, \lambda']^\top$

Beweis. Das zulässige lineare System $\mathcal{A}$ repräsentiert $f = 0$ genau dann, wenn $s_1 = (A^{-1}v)_1 = 0$ ist. Da alle Systeme äquivalent zu $\mathcal{A}$ sind, wird dieser Fall auch für $k' = 1$ erkannt, weil es laut Satz 4.5.9 eine *zulässige* Transformation gibt, sodass die ersten linken Blockminimierungsgleichungen erfüllt werden. Nun sei $f \neq 0$. Wir müssen zeigen, dass sowohl die linke Familie s' als auch die rechte Familie t' von

$\mathcal{A}' = (u', A', v')$ $\mathbb{K}$-linear unabhängig ist. Sei $m' = \#_{\mathrm{pb}}(\mathcal{A}')$ und für $k \in \{1, 2, \ldots, m'\}$ bezeichne

$$s'_{(k)} = (s'_{\underline{m'+1-k}}, s'_{\underline{m'+2-k}}, \ldots, s'_{\underline{m'}}) \quad \text{und} \quad t'_{(k)} = (t'_{\underline{1}}, t'_{\underline{2}}, \ldots, t'_{\underline{k}})$$

die linke beziehungsweise die rechte Teilfamilie. Laut Annahme sind $s'_{(1)}$ beziehungsweise $t'_{(1)}$ $\mathbb{K}$-linear unabhängig. Die Schleife beginnt mit $k = 2$. Nur wenn beide $s'_{(k)}$ und $t'_{(k)}$ jeweils $\mathbb{K}$-linear unabhängig sind, wird k inkrementiert. Im anderen Fall war ein linker (Satz 4.5.9) oder ein rechter (Satz 4.5.12) Minimierungsschritt erfolgreich und die Anzahl der Pivotblöcke des aktuellen ZLS ist echt kleiner als die des vorherigen. Nachdem k von unten beschränkt ist, bricht der Algorithmus in endlich vielen Schritten ab. Nun schauen wir uns die Zeilen 12–15 genauer an. Wenn es eine spezielle zulässige Blockzeilentransformation für das erweiterte ZLS für $k' = 1$ gibt, dann ist die Anzahl der Pivotblöcke vom ZLS gleich 2, weil sonst ein Widerspruch zur $\mathbb{K}$-linearen Unabhängigkeit der entsprechenden rechten Teilfamilie entstünde. Daher ist nach Zeile 14 nur mehr *ein* Pivotblock übrig und der Algorithmus stoppt. (Wie Zeile 0 und Spalte 0 von einem erweiterten ZLS entfernt werden, wird in Beispiel 4.5.6 illustriert.) Alle Transformationen sind derart, dass $\mathcal{A}'$ ein verfeinertes ZLS (und daher in *Standardform*) ist. Für $\#_{\mathrm{pb}}(\mathcal{A}') = 1$ ist laut Annahme a priori nur die linke (oder die rechte) Familie $\mathbb{K}$-linear unabhängig. Aber wäre das nicht auch der Fall für die jeweils andere Familie, ergäbe das —laut Satz 4.5.12 bzw. 4.5.9— einen Widerspruch zur Annahme eines verfeinerten Pivotblocks. □

Bemerkung. Auch hier gibt es —in Bezug auf die Vermeidung der mehrfachen Lösung von linearen Gleichungssystemen— noch Verbesserungspotential, was freilich nichts an der Laufzeitkomplexität ändern wird. Für eine Größe der Pivotblöcke von $\dim_i(\mathcal{A}) \approx \sqrt{n}$ und d Buchstaben bekäme man die Gesamtkomplexität $\mathcal{O}(dn^5)$, was zwischen dem Spezialfall (für polynomielle Systeme, Bemerkung 4.3.9) und dem Wortproblem (Abschnitt 4.2) liegt. Im Normalfall wird man ein aus zwei *minimalen* zulässigen linearen Systemen konstruiertes ZLS (zum Beispiel Addition und Multiplikation laut Proposition 2.3.1) minimieren. Dafür könnte man dann den hier vorgestellten Algorithmus anpassen.

Bemerkung. Die Lösung des Wortproblems für zwei durch *minimale* zulässige lineare Systeme gegebene Elemente ist *unabhängig* von der Verfeinerung. Wird der Algorithmus 4.5.15 auf ein ZLS angewendet, von dem man nicht weiß, ob es verfeinert ist, kann man in manchen Fällen trotzdem einfach feststellen, ob das ZLS $\mathcal{A}'$ minimal ist, zum Beispiel wenn $\dim(\mathcal{A}') = \#_{\mathrm{pb}}(\mathcal{A}')$ ist. Sind die Pivotblöcke größer, aber die rechte obere Struktur ist „feiner“, kann man stattdessen —für $f \neq 0$— versuchen, die Standardinverse $(\mathcal{A}')^{-1}$ zu minimieren. In konkreten Situationen gibt es sicher noch weitere Möglichkeiten, Minimalität zu erreichen beziehungsweise sie festzustellen. Hat man das Kapitel 3 (Faktorisieren) übersprungen, kann man das nun nachholen.

Bemerkung 4.5.16. Abgesehen von der eigentlichen Minimierung kann dieser Algorithmus verwendet werden, um festzustellen, ob ein Element f ein linker Faktor

eines Elementes fg ist. Das ist für die minimale Faktormultiplikation (Satz 3.5.2) relevant. Und klarerweise kann man mit diesem Algorithmus zwei Elemente auf Disjunktheit (Definition 2.4.1) testen. Das spielt dann besonders bei der Primärzerlegung (Satz 2.4.2) eine Rolle.

Zusammenfassend soll nun noch ein anderer Aspekt des Algorithmus 4.5.15 (beziehungsweise der Sätze 4.5.9 und 4.5.12) beleuchtet werden, der sich unmittelbar aus Proposition 2.1.8 und Bemerkung 4.5.11 erschließt. Die Bedeutung des folgenden Satzes wird klar, wenn man vor der Aufgabe steht, die $\mathbb{K}$-lineare (Un-)Abhängigkeit einer beliebigen Familie $(f_1, f_2, \ldots, f_n)$ über dem freien Schiefkörper zu prüfen und nicht (mehr) auf eine Darstellung als formale Potenzreihe zurückgreifen kann.

Satz 4.5.17 („Lineare“ Minimalitätscharakterisierung)**.** *Ein* verfeinertes *zulässiges lineares System $\mathcal{A} = (1, A, \lambda)$ für ein Element im freien Schiefkörper $\mathbb{F}$ mit $m = \#_{pb}(\mathcal{A}) \geq 2$ Pivotblöcken und $\lambda \neq 0$ ist genau dann minimal, wenn es weder für die linken Blockminimierungsgleichungen $\mathcal{L}_{\underline{k}}(\mathcal{A})$ für $k \in \{1, 2, \ldots, m-1\}$ noch für die rechten Blockminimierungsgleichungen $\mathcal{R}^{\underline{k}}(\mathcal{A})$ für $k \in \{2, 3, \ldots, m\}$ eine Lösung gibt.*

Beweis. Aus der Existenz einer Lösung folgt sofort Nicht-Minimalität, weil in diesem Fall —nach der entsprechenden Umformung— Zeilen und Spalten entfernt werden können. Und aus der Nicht-Minimalität folgt aus Proposition 2.1.8, dass entweder die linke oder die rechte Familie $\mathbb{K}$-linear abhängig ist. Ohne Beschränkung der Allgemeinheit sei es die linke $s = (s_{\underline{1}}, s_{\underline{2}}, \ldots, s_{\underline{m}})$ mit einem minimalen $k \in \{1, 2, \ldots, m-1\}$, sodass die linke Teilfamilie $(s_{\underline{k+1}}, \ldots, s_{\underline{m}})$ $\mathbb{K}$-linear *un*abhängig ist. Aufgrund der verfeinerten Pivotblöcke erfolgt nun aus Satz 4.5.9 die Existenz einer speziellen Blockzeilentransformation $(P, Q)_{\underline{k}}$ und damit die Lösung der k-ten linken Blockminimierungsgleichungen. □

Nachwort

Für ein Grundverständnis vom freien Schiefkörper sollte man sich zumindest zwei Fragen stellen: Gibt es für jedes Element eine (reine) lineare Darstellung? Und, kann man diese Darstellung über rationale Operationen „erreichen“? Antworten darauf findet man sehr kompakt in der Arbeit von Cohn und Reutenauer [CR99]. Die Frage nach der Minimalität taucht früher oder später ganz natürlich auf. Und in einer gewissen Weise sieht man in dieser Arbeit nicht mehr als die Spitze eines Eisberges.

Was sich jedenfalls verändert haben sollte, ist der Blick auf das Rechnen mit „normalen“ Brüchen und den Konzepten, die das erst ermöglichen. Und bei allen Unterschieden im Detail ist es erstaunlich, wieviele Parallelen es zwischen den ganzen Zahlen und den nicht-kommutativen Polynomen gibt.

Der Einstieg in die nicht-kommutative Welt ist sicherlich nicht einfach. Dabei kann der Kopf schon einmal schwirren. Aber die reale (nicht zu verwechseln mit der *reellen*) Welt ist durch und durch nicht-kommutativ: Von einfachen Bewegungen im Raum[1] bis hin zur Quantenphysik. Und in der mathematischen kennt man jedenfalls die (nicht-kommutative) Multiplikation zweier Matrizen. Setzt man in einer großen Matrix unabhängige Gaußsche Zufallszahlen ein, kann man die Catalan-Zahlen entdecken. (Wie, ist im Anhang A.4 skizziert.)

Aber er lohnt sich schon alleine deshalb, weil er einem die kommutative mit anderen Augen sehen lässt. Und zu entdecken gibt es ohnehin überall genug. Wusste man vorher nicht, wo man anfangen sollte, ist nun bereits der erste Schritt getan ...

Bevor es an einen kleinen Ausblick geht, sollte es noch einen Blick zurück geben. Aber nicht in Form einer langen Zusammenfassung, sondern *auf* etwas, das wahrscheinlich —ob all der Theoreme— untergegangen ist. Und zwar Lemma 2.3.2. Es gibt andere mit längeren und schwierigeren Beweisen. Es ist nicht so allgemein wie Lemma 4.5.8 [CR99, Lemma 1.2]. Aber *ohne* blieben von dieser Arbeit möglicherweise nicht viel mehr als einzelne Fragmente übrig, es ist eines der Herzstücke in der Anwendung *linearer* Techniken (für das Rechnen mit „nicht-kommutativen“ Brüchen). Dabei verknüpft es „nur“ zwei bereits bekannte Konzepte, nämlich das *linearer (nicht-voller) Matrizen* und das *minimaler zulässiger linearer Systeme*. So praktisch volle Matrizen für das Rechnen mit Elementen im freien Schiefkörper auch sind, sie entziehen sich

[1] Man denke an den unterschiedlichen Effekt, ein volles Glas Wasser auszuschütten und es dann jemanden zu reichen oder es zuerst jemanden zu reichen und dann auszuschütten.

jedem Versuch, *direkt* etwas (allgemein) zu zeigen.[2] Sollte also jemand an etwas rund um minimale lineare Darstellungen im freien Schiefkörper interessiert sein, das *nicht* hier steht, könnte dieses Lemma (oder die Idee des Beweises) vielleicht weiterhelfen.

Der „Hamiltonsche“ freie Schiefkörper?

Sei $\mathbb{H}$ der Schiefkörper der *Hamiltonschen Quaternionen*[3]. Kann man auch $\mathbb{H}\langle X\rangle$ in einen freien Schiefkörper einbetten? Ist man daran interessiert, sollte man unbedingt [CR94] und [Coh95, Abschnitt 6.3] im Detail lesen. Tatsächlich ist die Theorie von Cohn und Reutenauer noch viel allgemeiner, eher stellt sich die Frage, in wie weit man damit noch „praktisch“ rechnen kann. Statt $\mathbb{K}$-linearer Unabhängigkeit muss man dann zwischen *links-* und *rechts*-$\mathbb{H}$-linearer Unabhängigkeit unterscheiden. Die ersten konkreten Aufgabenstellungen könnten die algorithmische Lösung des Wortproblems und der Versuch einer „Linearisierung“ sein. Was umgesetzt werden kann und wie schwierig das gegebenenfalls ist, lässt sich derzeit noch nicht beurteilen. Vielleicht muss man zuerst einen Schritt „zurück“ gehen und mit der „Hamiltonschen“ freien assoziativen Algebra $\mathbb{H}\langle X\rangle$ beginnen um festzustellen, ob man auch dort direkt in den linearen Darstellungen arbeiten kann und ob sich die —in Abschnitt 3.3 präsentierte— Faktorisierung von Polynomen verallgemeinern lässt.

Das Rechnen mit nicht-kommutativen Wurzeln?

Wollte man dem —über „reguläre“ algebraische Elemente, das heißt, Elemente die eine Darstellung als formale Potenzreihe haben, hinaus— zumindest teilweise eine Bedeutung geben, müsste man sich jedenfalls mit den drei Definitionen algebraischer Abschlüsse von Schiefkörpern auseinandersetzten [Coh95, Abschnitt 8.1]: *charakteristisch algebraisch abgeschlossen* (CAC)[4], *polynomiell algebraisch abgeschlossen* (PAC) und *vollständig algebraisch abgeschlossen* (FAC).

Nachdem es *reguläre algebraische Systeme* (PAS)[5] [SS78, Abschnitt IV.1] oder [Coh85, Abschnitt 2.9] bereits gibt, stellt sich natürlich die Frage, *ob* beziehungsweise *wie* man diese so mit zulässigen linearen Systemen (ALS) verknüpfen kann, dass man „zulässige algebraische Systeme“ (AAS) beschreiben kann, mit denen sich tatsächlich „rechnen“ lässt. Aus meiner Sicht wäre das (im allgemeinen) nicht allzu aussichtsreich, man denke nur an das Wortproblem für allgemeine algebraische Elemente im Körper der komplexen Zahlen $\mathbb{C}$. Darüber nachzudenken scheint mir dagegen eine

[2] Wer schon einmal versucht hat, etwas für *lineare* volle Matrizen der Größe 3×3 zu zeigen, weiß, wovon ich spreche.

[3] Nicht zu verwechseln mit dem kürzbaren Monoid $\mathbb{H}$ in Abschnitt 3.4.

[4] Die Akronyme hier sind alle von der englischen Bezeichnung abgeleitet: characteristically/polynomially/fully algebraically closed.

[5] Englisch für „proper algebraic system“. Die Abweichung von der wörtlichen Übersetzung hängt mit der (Wahl der) Bezeichnung „reguläres lineares System“ (engl. PLS) für *reguläre* Elemente zusammen.

außerst inspirierende Quelle für Fragen zu sein, sowohl rein algebraisch als auch rein algorithmisch.

In diesem Sinne ist eine meiner naiven Fragestellungen, ob man (wenn man mit der Implementierung des freien Schiefkörpers am Computer halbwegs fertig ist) zumindest mit „periodischen“ zulässigen linearen Systemen (als Verallgemeinerung von „periodischen“ Kettenbrüchen) rechnen beziehungsweise dem einen Sinn geben kann. Voraussetzung dafür ist natürlich, dass man sich in der (vergleichsweise überschaubaren) Literatur nicht verliert. Als Ausgangspunkt(e) könnten unter anderem diese Arbeiten dienen: [Wed14], [Niv41], [EN44], [ML85], [Woo85], [Kol00] und/oder [Kol01]. Jedenfalls kommen darin die Hamiltonschen Quaternionen nicht zu kurz ...

Anhang A

Anwendungen

Die Möglichkeiten der Anwendung des Rechnens mit Elementen des freien Schiefkörpers (alias „freien Brüchen“) sollten so zahlreich sein, dass sie kaum vollständig aufgelistet werden können. Da ist es natürlich einfacher, wenn man überhaupt nur wenige kennt ...

Im folgenden Anhang B (Bemerkungen) findet man dann noch weitere Anknüpfungspunkte. Hier sollen nur zwei (nicht-kommutative algebraische) „Grundtechniken“ erwähnt werden: die Bestimmung des *linken ggT* zweier nicht-kommutativer Polynome in Abschnitt A.1 und die des *Ranges* einer Matrix über die „freie“ Gaußsche Elimination in Abschnitt A.2. Erstere sollte dabei helfen, *eigene* Techniken zu entwickeln (und gegebenenfalls zu implementieren). Letztere sollte einfach an die *lineare Algebra* mit all ihren Einsatzgebieten erinnern.

Bevor wir zu zwei konkreteren Anwendungen kommen, sollte noch erwähnt werden, dass es auch kontextspezifische Einschränkungen geben kann. So ist die (hier präsentierte) Faktorisierung von rationalen formalen Potenzreihen eben *nicht notwendigerweise* eine Faktorisierung von regulären Sprachen. Gemeinsamkeiten und Unterschiede (insbesondere in den Begriffsdefinitionen) herauszuarbeiten könnte durchaus für beide Seiten lohnenswert sein. Als konkreten Einstieg bieten sich [MSY02] und [BRS15], aber auch die Bemerkungen im Abschnitt B.3, an. Ist man mit der Theorie formaler Sprachen gut vertraut, ist [Coh75a] als Brücke zur (nicht-kommutativen) Algebra sehr empfehlenswert.

Konkret ist das Rechnen mit rationalen Funktionen (hier meist im Sinne von regulären Elementen) in der *Steuerungstheorie* relevant. Hier einen Bogen zwischen der „kommutativen“ Anwendung und der umfangreichen mathematischen „nicht-kommutativen“ Theorie zu spannen, ist schwierig. Der Abschnitt A.3 ist daher nur eine Art Minimalkompromiss, um die wichtigsten Begriffe einzuführen und ein paar Bemerkungen zum Arbeiten mit *matrixwertigen* Elementen loszuwerden.

Etwas, das vielleicht allgemein weniger bekannt ist, ist die *freie Wahrscheinlichkeitstheorie*. Das soll sich nach dem ersten Teil von Abschnitt A.4 ändern, in dem illustriert wird, *wie* man die Catalan-Zahlen in Zufallsmatrizen „findet“ und welch

verblüffenden Zusammenhang es zwischen der *kommutativen* und der *freien* Wahrscheinlichkeitstheorie (als zwei Vertreter von mehreren *nicht-kommutativen* Theorien) gibt. Dafür, dass einem das in ein paar Jahren helfen wird, wenn es die sechste Mobilfunkgeneration gibt, kann ich allerdings keine Garantie geben. Als Einstieg in dieses spannende Thema wäre [MAZd13] viel besser geeignet.

A.1 Linker größter gemeinsamer Teiler

Sind zwei Polynome $p, q \in \mathbb{K}\langle X\rangle \setminus \mathbb{K}$ gegeben, kann man den *linken* (bzw. *rechten*) *größten gemeinsamen Teiler* von p und q bestimmen, indem man ein zulässiges lineares System für $p^{-1}q$ (bzw. pq^{-1}) minimiert. Die Vorgehensweise soll nun an einem Beispiel illustriert werden. Sei $p = yx(1-yx)z = yxz - yxyxz$ und $q = y(1-xy)y = y^2 - yxy^2$, gesucht ist $h = \operatorname{lgcd}(p,q)$. Ein ZLS für $p^{-1}q$ (konstruiert laut Proposition 2.3.6 aus je einem minimalen für p^{-1} und q) ist

$$\begin{bmatrix}
z & 1 & . & . & . & . & . & . & . \\
. & x & -1 & . & . & . & . & . & . \\
. & -1 & y & -1 & . & . & . & . & . \\
. & . & . & x & -1 & . & . & . & . \\
. & . & . & . & y & -y & . & . & . \\
. & . & . & . & . & 1 & -x & 1 & . \\
. & . & . & . & . & . & 1 & -y & . \\
. & . & . & . & . & . & . & 1 & y \\
. & . & . & . & . & . & . & . & 1
\end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix}.$$

Ein *verfeinertes* zulässiges lineares System für p^{-1} erhält man durch die Polynomfaktorisierung (Abschnitt 3.3). Als ersten Schritt addieren wir Spalte 5 zu Spalte 6 und Zeile 6 zu Zeile 4,

$$\begin{bmatrix}
z & 1 & . & . & . & 0 & . & . & . \\
. & x & -1 & . & . & 0 & . & . & . \\
. & -1 & y & -1 & . & 0 & . & . & . \\
. & . & . & x & -1 & 0 & -x & 1 & . \\
. & . & . & . & y & 0 & 0 & 0 & 0 \\
. & . & . & . & . & 1 & -x & 1 & . \\
. & . & . & . & . & . & 1 & -y & . \\
. & . & . & . & . & . & . & 1 & y \\
. & . & . & . & . & . & . & . & 1
\end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix},$$

entfernen die Zeilen und Spalten 5 und 6,

$$\begin{bmatrix} z & 1 & . & . & . & . & . \\ . & x & -1 & . & . & . & . \\ . & -1 & y & -1 & . & . & . \\ . & . & . & x & -x & 1 & . \\ . & . & . & . & 1 & -y & . \\ . & . & . & . & . & 1 & y \\ . & . & . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix} \tag{A.1.1}$$

und merken uns den ersten (linken) Teiler $h_1 = y$, den wir eliminiert haben. (Beim nächsten Schritt mit einem größeren Block sieht man dann sofort, wie man diesen Teiler direkt aus dem ZLS „ablesen“ kann.) Nun gibt es zwei verschiedene Vorgehensweisen: Ist es nicht möglich einen „L“-Nullblock (wie zuvor) zu erzeugen, kann man versuchen, die obere Pivotblock-Struktur zu verändern um einen „Doppel-L“-Nullblock zu erzeugen. Hier ist das möglich, indem man Spalte 4 von Spalte 2 subtrahiert und Zeile 2 zu Zeile 4 addiert. Danach wendet man folgende (zulässige) Transformation

$$(P,Q) = \left(\begin{bmatrix} 1 & . & . & . & . & . & . \\ . & 1 & . & . & . & 1 & . \\ . & . & 1 & . & 1 & . & . \\ . & . & . & 1 & . & . & . \\ . & . & . & . & 1 & . & . \\ . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . \\ . & . & 1 & . & . & 1 & . \\ . & . & . & 1 & 1 & . & . \\ . & . & . & . & 1 & . & . \\ . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & 1 \end{bmatrix} \right)$$

an und erhält das ZLS

$$\begin{bmatrix} z & 1 & . & . & 0 & 0 & . \\ . & x & -1 & . & 0 & 0 & y \\ . & . & y & -1 & 0 & 0 & 0 \\ . & . & -1 & x & 0 & 0 & 0 \\ . & . & . & . & 1 & -y & . \\ . & . & . & . & . & 1 & y \\ . & . & . & . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix}. \tag{A.1.2}$$

Wie man diese Transformation (P,Q) erhält, ist im Prinzip in Abschnitt 4.5 beschrieben. Man kann direkt nach einer „Doppel-L“ Blocktransformation suchen. Im dritten Pivotblock von (A.1.2) sieht man sofort, dass ein weiterer (gemeinsamer) Faktor $h_2 = 1 - xy$ ist, weil in der zweiten Gleichung $xs_2 - h_2^{-1} = 0$ steht. Zur Erinnerung: Ein minimales ZLS für h_2 wäre zum Beispiel

$$\begin{bmatrix} 1 & -x & 1 \\ . & 1 & -y \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ -1 \end{bmatrix}.$$

Nachdem wir die Zeilen und Spalten $\{3,4,5,6\}$ im ZLS (A.1.2) entfernen, erhalten wir für $p^{-1}q$ das *minimale* ZLS

$$\begin{bmatrix} z & 1 & . \\ . & x & y \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}. \tag{A.1.3}$$

Also ist $h = h_1 h_2 = y(1-xy) = \mathrm{lgcd}(p,q)$. Die zweite Vorgehensweise ist, dass man, ausgehend von ZLS (A.1.1), einen rechten Minimierungsschritt bezogen auf Spalte 5 durchführt, danach einen linken bezogen auf die Zeilen 2 und 3 und zum Schluss noch einen rechten. Wieder erhält man das ZLS (A.1.3) (modulo zulässige Skalierung von Zeilen und Spalten), in dem der rechte Faktor y von q übrig bleibt. Damit ist $\mathrm{lgcd}(p,q) = y - yxy$. Details zur minimalen Polynommultiplikation finden sich in Abschnitt 3.2.

Bemerkung. Es kann vorkommen, dass —nachdem kein „L“-Minimierungsschritt mehr möglich ist— das ZLS noch *nicht* minimal ist, das heißt, ein zusätzlicher linker oder rechter Minimierungsschritt möglich ist. Die Details dazu finden sich im Abschnitt 4.5 (Minimieren eines allgemeinen ZLS). Um dieses subtile Phänomen richtig zu verstehen, muss man tiefer in die allgemeine Faktorisierungstheorie in Abschnitt 3.4 eintauchen. Hier genügt allerdings ein nochmaliger Blick auf das ZLS (A.1.3): *Beide* rechte Faktoren von p (hier xz) beziehungsweise q (hier y) können noch „direkt“ abgelesen werden. Mit anderen Worten: Der „Bruchstrich“ ist noch erkennbar. (Dieser würde gegebenenfalls bei einem weiteren Minimierungsschritt „verschwinden“).

A.2 Freie lineare Algebra

Oder: Gaußsche Elimination über dem freien Schiefkörper. Klassische Techniken können verwendet werden, um den (nicht-kommutativen) *Rang* [FR04] von Matrizen über dem freien Schiefkörper zu bestimmen, oder die Inverse (wenn möglich) oder Quasideterminanten [GGRW05] zu berechnen. In Bezug auf die Lösung von linearen Gleichungssystemen sei die *Bruhat Normalform* [Coh03b, Satz 9.2.2] als eine verallgemeinerte LU-Zerlegung erwähnt.

Das Wortproblem für Elemente, die nicht durch minimale zulässige lineare Systeme gegeben sind, kann gelöst werden, indem man den Rang der Linearisierung (der Differenz) durch Gaußsche Elimination ermittelt. Jeder Eintrag in dieser Matrix kann durch ein *minimales* zulässiges lineares System dargestellt werden. Für das „pathologische“ Beispiel 4.2.4 ist die Linearisierung

$$L = \begin{bmatrix} . & 1 & . & . \\ -1 & x & -x & . \\ 1 & . & x & -z \\ . & . & . & 1 \end{bmatrix}$$

nicht-voll, daher repräsentiert das entsprechende ZLS das Nullelement: Wendet man die (Zeilen-)Transformation

$$T = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ . & . & -z^{-1}x & 1 \end{bmatrix} \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & -x^{-1} & 1 & . \\ . & . & . & 1 \end{bmatrix} \begin{bmatrix} . & 1 & . & . \\ . & 1 & 1 & . \\ 1 & . & . & . \\ . & . & . & 1 \end{bmatrix}$$

von links an, erhält man

$$TL = \begin{bmatrix} -1 & x & -x & . \\ . & x & . & -z \\ . & . & . & x^{-1}z \\ . & . & . & 0 \end{bmatrix}.$$

Der Preis, den man zahlen muss, ist, dass *jede* Operation mit einem einzelnen Matrixeintrag nun eine Operation mit zulässigen linearen Systemen ist. Vielleicht ist es in manchen Fällen möglich, eine „bessere“ Eliminationstechnik zu verwenden, die die Struktur der Matrix berücksichtigt (eventuell kombiniert mit Spaltenoperationen) um das Rechnen mit den Darstellungen zu vereinfachen.

Bemerkung. Obwohl Lemma 2.1.3 die Existenz von invertierbaren Matrizen $P, Q \in \mathbb{K}^{n\times n}$ für eine nicht-volle *lineare* Matrix $L \in \mathbb{K}\langle X\rangle$ garantiert (siehe Satz 4.2.1, [CR99, Abschnitt 4]), sodass PLQ hohl ist, ist das Finden solcher Matrizen eine schwierige Aufgabe, weil alle möglichen (linken unteren) Nullblöcke der Größe $(n+1-i)\times i$ für $i = \{1, 2, \ldots, n\}$ getestet werden müssen. Während die Gaußsche Elimination immer zumindest eine Nullzeile (oder Nullspalte, wenn sie von rechts auf den Spalten durchgeführt wird) liefert (allerdings möglicherweise ebenfalls mit nicht-linearen Techniken). Interessant wäre natürlich, ob man eine „Klasse“ solcher Probleme beschreiben kann, in der sich die Größe der Pivotblöcke so beschränken lässt, dass man immer einfach minimieren kann.

A.3 Steuerungstheorie

Oder: Das Rechnen mit rationalen Funktionen. In diesem Abschnitt hat der Begriff „System“ *nichts* mit einem ZLS zu tun. Auch bei der anderen Notation (Systemmatrix, Buchstaben im Alphabet, etc.) gibt es Abweichungen.

In der *Steuerungstheorie* macht man es sich zunutze, dass man mit *Übertragungsfunktionen* von „Systemen“ (unter bestimmten Voraussetzungen) ganz einfach rechnen kann: Die Multiplikation (von Übertragungsfunktionen) entspricht dabei der *Serienschaltung*, die Addition der *Parallelschaltung* und eine bestimmte Kombination von rationalen Operationen (inklusive der Inversen) der *Rückkoppelung* [HD04, Kapitel 3].

Dabei spielt die *Minimalität* einer *Realisierung* beziehungsweise einer *linearen Darstellung* (siehe auch Abschnitt B.2) eine entscheidende Rolle in Bezug auf *Beobachtbarkeit* und *Steuerbarkeit* (siehe auch Abschnitt B.3) eines Systems. Solange man

nur mit *einem* Buchstaben im Alphabet $X = \{\xi\}$ arbeitet, ist der freie Schiefkörper $\mathbb{K}(\langle X\rangle)$ *kommutativ*. Für mehrere *kommutierende* Variablen $\xi_1, \xi_2, \ldots, \xi_d$ ist es im allgemeinen schwierig, *minimale* lineare Darstellungen zu konstruieren. Für bestimmte *kommutative* Polynome ist es möglich, siehe Proposition 2.2.6 und die ihr folgende Bemerkung.

Im folgenden betrachten wir (für $n, p, q \in \mathbb{N}$) das kontinuierliche und zeitinvariante System (für eine angewandte Sicht siehe zum Beispiel [HD04], für eine operatortheoretische [BGKR08])

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0, \\ x(0) &= 0 \end{aligned}$$

mit dem (zeitabhängigen) *Lösungs-* oder *Zustandsvektor* $x = x(t) = [x_1(t), \ldots, x_n(t)]$, dem *Ausgangsvektor* $y = y(t) = [y_1(t), \ldots, y_p(t)]$, dem *Steuervektor* $u = u(t) = [u_1(t), \ldots, u_q(t)]$, der *Systemmatrix* $A \in \mathbb{K}^{n\times n}$, der *Steuermatrix* $B \in \mathbb{K}^{n\times q}$, der *Beobachtungsmatrix* $C \in \mathbb{K}^{p\times n}$ und $D \in \mathbb{K}^{p\times q}$. Die (für $p, q > 1$ matrixwertige) *Übertragungsfunktion* ist dann

$$G(\xi) = \big(g_{ij}(\xi)\big) = D + C(\xi I - A)^{-1}B, \quad \xi \in \rho(A),$$

wobei $\rho(A)$ die Resolventenmenge von A bezeichnet.

Das Viertupel $\mathcal{R} = (\xi I - A, B, C, D)$ heißt *Realisierung*, siehe auch Definition B.2.1. Der einfachste Fall ist $p = q = 1$. Rein algebraisch ist aber auch das Rechnen mit matrixwertigen Übertragungsfunktionen (mit kommutierenden oder nichtkommutierenden Komponenten) nichts Besonderes. Die Frage nach der Minimalität der Systemmatrix (hier $\xi I - A$) dagegen schon. (Bis jetzt sind mir dazu in der Literatur weder entsprechende Fragen noch Diskussionen aufgefallen.)

Seien $F \in \mathbb{F}^{m\times p}$ und $G \in \mathbb{F}^{p\times q}$ für $\mathbb{F} = \mathbb{K}(\langle X\rangle)$ mit einem allgemeinen (nichtkommutativen) Alphabet X. Dann ist $H = FG \in \mathbb{F}^{m\times q}$ mit den Komponenten $h_{ij} \in \mathbb{F}$. Nun kann man laut Satz 2.4.2 (Primärzerlegung) jedes h_{ij} in seine *Primärkomponenten* zerlegen. Eine Systemmatrix der Dimension

$$n = \sum_{i=1}^{m}\sum_{j=1}^{q} \operatorname{rang}(h_{ij})$$

erhält man sehr einfach über die diagonale Summe der einzelnen Systemmatrizen der *minimalen* (reinen) linearen Darstellungen der jeweiligen Elemente h_{ij}. Für nicht-disjunkte Primärkomponenten (verschiedener Elemente) ließe sich eine *minimale* „Vereinigung“ mit einer linearen „matrixwertigen“ Darstellung, zum Beispiel

$$\left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -z \\ . & 1 & . \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . & . & . \\ 1 & . & -1 \\ . & 1 & . \end{bmatrix} \right)$$

für $H = [x, z, -x]$, konstruieren. Führt diese Vorgehensweise schon zu einer *minimalen* Realisierung (zum Beispiel für $D = 0$ und skalaren Matrizen B, C) für H? Wie müsste eine (neue) Darstellungsform beschaffen sein, sodass man damit praktikabel rechnen kann? Diese Fragestellungen könnte auch in Zusammenhang mit Abschnitt A.2 interessant sein, zum Beispiel wenn man die Inverse (einer invertierbaren Matrix $F \in \mathbb{F}^{m \times m}$) berechnen möchte.

A.4 Freie Wahrscheinlichkeitstheorie

Als Schnelleinstieg für jene, die mit der „klassischen“ Wahrscheinlichkeitstheorie vertraut sind, bietet sich die Einleitung von Biane [Bia98] an. Für jene, die mehr mit der Operatortheorie vertraut sind, auch die von Shlyakhtenko [Shl05]. Über beide Literaturverzeichnisse bekommt man dann gleich einen Eindruck über die Vielfalt rund um „nicht-kommutative“ Wahrscheinlichkeit. Detaillierte Einführungen bieten die Bücher von Nica und Speicher [NS06] beziehungsweise Hiai und Petz [HP00], ersteres insbesondere aus kombinatorischer Sicht.

Eine andere Art Schnelleinstieg wäre das „zwölfte Problem“ von Rota [Rot01]. Nachdem man tiefer in die nicht-kommutative Welt eingetaucht ist, kann man sich —nach dem Auftauchen— nämlich selbst fragen, ob man mit den gefundenen Erklärungen zufrieden ist.

Hier soll ein eher spielerischer Zugang —im Sinne eines Paddelns an der Oberfläche— über Zufallsmatrizen skizziert werden, der mit wenigen Programmzeilen in Octave [Oct18] nachvollzogen werden kann. Damit kann man sich sofort die Verteilung des eigenen nicht-kommutativen Lieblingspolynoms[1] „anschauen“, wenn man zum Beispiel „Halbkreis-Variablen“ (Gaussian Unitary Ensemble, GUE) einsetzt.

Der zweite Teil betrifft dann eine kleine konkrete Anwendung einer *minimalen* linearen Darstellung und ist *keinesfalls* als Einstieg geeignet, obwohl nur ein sehr spezieller (ja fast trivialer) Fall behandelt wird. Um dieses Beispiel —das man in gewisser Weise als Tor zur aktuellen Forschung sehen kann— nachvollziehen zu können, muss man sich jedenfalls in [BMS17] oder [HMS15] einlesen und das Buch [NS06] zum Nachschlagen bei der Hand haben. Tatsächlich ist der (algebraische) Weg, diese Technik für „einfache“ polynomielle Operatoren wie dem *Kommutator* oder dem *Antikommutator* anzuwenden noch weit. Das Ziel ist die Herleitung eines (nicht-linearen) Gleichungssystems, dessen „eindeutige“ Lösung die Verteilung „beschreibt“. Andere Wege, wie zum Beispiel den (analytischen) von Vasilchuk [Vas03] oder den (kombinatorischen) von Nica und Speicher [NS98], gibt es bereits.

Zunächst definieren wir ein paar Funktionen: `make_sa` um eine Matrix selbstadjungiert zu machen und „richtig“ zu skalieren, `norm_tr` um die normalisierte Spur einer Matrix zu berechnen und `catalan` um die Catalan-Zahlen [Slo18a] zu berechnen.

[1] Die Fragen, die auftauchen, wollte man das auch für reguläre nicht-kommutative Funktionen versuchen, sind zahlreich. Sehr schnell kommt man damit tief, sehr tief in die Operatortheorie ...

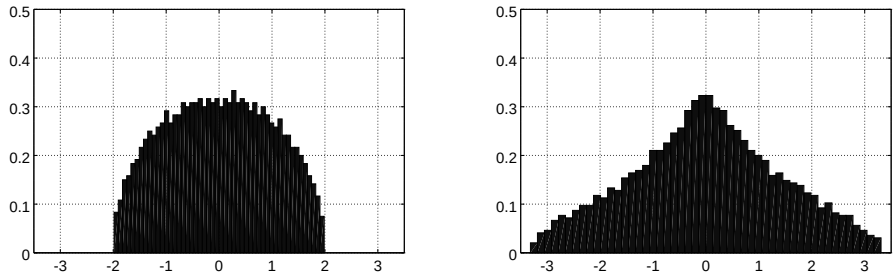


Abbildung A.1: Empirische Eigenwertverteilung einer GUE-Matrix (selbstadjungierte Matrix mit „richtig“ skalierten normal-verteilten *unabhängigen* Einträgen) der Größe $n = 1500$ (links) und des Antikommutators (rechts).

```
make_sa = @(A) (A + A') / sqrt(2*rows(A));
norm_tr = @(A) trace(A)/rows(A);
catalan = @(n) factorial(2*n) / (factorial(n+1) * factorial(n));
```

Danach erzeugen wir jeweils eine Zufallsmatrix (GUE) für die nicht-kommutativen Variablen unseres Alphabets:

```
n = 1500;
X = make_sa(randn(n));
Y = make_sa(randn(n));
```

Die Eigenwertverteilung einer Variable lässt sich folgendermaßen plotten:

```
figure(101);
hist(eig(X), 50, 50/4.0);
axis([-2.5 2.5 0 0.5]);
grid;
%set(gca(), 'fontsize', 20);
%print('eig_x.eps')
```

Damit kann man schließlich auch die Eigenwertverteilung eines nicht-kommutativen Polynoms, zum Beispiel des Antikommutators $p = p(x, y) = xy + yx$ inspizieren (die „Norm“ für das Histogramm muss entsprechend angepasst werden):

```
figure(102);
P = X*Y + Y*X;
hist(eig(P), 50, 50/6.5);
axis([-3.5 3.5 0 0.5]);
grid;
```

Die Ergebnisse sind in Abbildung A.1 zusammengefasst. Vergleiche auch mit [BMS17, Abbildung 1].

Bemerkung. Hier fehlen natürlich viele Details, die man aber in der bereits erwähnten Literatur findet. Um ein Gefühl dafür zu bekommen, was für „großes“ n passiert, sollte man zumindest Zufallsmatrizen der Größen $n \in \{1000, 2000, 4000\}$ erzeugen und einsetzen. Die Matrizen `X` und `Y` werden *frei*[2] für $n \to \infty$. Das heißt, dass alle *gemischten* Momente für *zentrierte* Zufallsvariablen verschwinden [NS06].

Bemerkung. Die *normalisierte Spur* $\varphi(x)$=`norm_tr(X)` kann man sich als einen *Erwartungswert* „$\mathbb{E}(x)$“ denken. Tatsächlich gibt es viele Parallelen zwischen der klassischen und der freien Wahrscheinlichkeitstheorie. Der Grund aber, dass ich hier vorsichtiger formuliere als es notwendig wäre, liegt darin, dass es doch einige Stolpersteine gibt, weil die zugrundeliegenden Konzepte unterschiedlich sind.

Schaut man sich die GUE-Zufallsmatrizen (alias „approximierte Halbkreis-Variablen“) genauer an, entdeckt man die Catalan-Zahlen

$$C_k = \frac{1}{k+1}\binom{2k}{k}.$$

Tabelle A.2 zeigt die ersten 6 Catalan-Zahlen und die ersten 12 Momente für GUE-Matrizen der Größen 750, 1500 und 3000 gegenübergestellt. In dieser Tabelle findet man auch die Momente der Normalverteilung, nämlich die Zahlenfolge 1, $3 = 1 \cdot 3$, $15 = 1 \cdot 3 \cdot 5$, $105 = 1 \cdot 3 \cdot 5 \cdot 7$, 945, 10395, ... [Slo18b]. Und diese Folge ist etwas versteckt und erschließt sich erst aus dem Zusammenhang zwischen der Anzahl der *Paarpartitionen* und der der *nicht-kreuzenden Paarpartitionen*. Mehr zu den Catalan-Zahlen findet man unter anderem in [Aig01].

```
Y_k = Y;
momente_Y = zeros(8,1);
for k=1:8
momente_Y(k) = norm_tr(Y_k);
Y_k = Y_k * Y;
endfor
gem_moment_XY = norm_tr(X*Y);
gem_moment_XYX = norm_tr(X*Y*X);
gem_moment_XYXY = norm_tr(X*Y*X*Y);
```

Was passiert mit den „gemischten“ Momenten $\varphi(xy)$, $\varphi(yx)$, $\varphi(xyx)$, $\varphi(xyxy)$, etc. für *freie* Variablen x und y (hier)? Und was muss man tun, um die „Freiheit“ zu sehen, wenn die Variablen x und y *nicht* zentriert sind? (Hier empfiehlt sich ein Blick ins Buch [NS06].)

Bemerkung. Will man in sein Polynom andere Zufallsvariablen (z.B. „Antikommutator“-Variablen) einsetzen, kann man sich diese entsprechend „zusammenbauen“. Allerdings muss man dabei beachten, dass man jede der „Zufallsmatrizen“ nur *einmal* verwendet, weil `X*Y+Y*X` und `X*Z+Z*X` *nicht frei* sind.

[2] „Frei“ ist eine Kurzform für „frei unabhängig“ als Pendant zur klassischen Unabhängigkeit von (kommutativen) Zufallsvariablen.

k	$(k-1)!!$	$C_{k/2}$	$n = 750$	$n = 1500$	$n = 3000$
1			0.001	0.000	0.000
2	1	1	0.998	1.001	1.000
3			-0.000	-0.002	0.001
4	3	2	1.993	2.006	1.999
5			-0.006	-0.010	0.004
6	15	5	4.974	5.024	4.993
7			-0.026	-0.042	0.0140
8	105	14	13.895	14.089	13.969
9			-0.088	-0.179	0.049
10	945	42	41.556	42.339	41.871
11			-0.240	-0.732	0.170
12	10395	132	130.150	133.290	131.470

Tabelle A.2: Die Momente der Normalverteilung, die Catalan Zahlen und die Momente $m_k = \varphi(Y^k)$ für GUE-Zufallsmatrizen.

Marchenko-Pastur via Linearisierung

Sei x eine selbstadjungierte nicht-kommutative Zufallsvariable mit Verteilung $\mu = \mu_x$. Die *Linearisierung* L_f einer nicht-kommutativen rationalen Funktion $f = f(x)$ kann verwendet werden, um die Verteilung von f zu berechnen [BMS17]. Für $f(x) = x^2 = -x(-1)^{-1}x$ ist eine (minimale) Linearisierung gegeben durch

$$L_f = \begin{bmatrix} \cdot & x \\ x & -1 \end{bmatrix}.$$

Für Definition und Konstruktion siehe Abschnitt B.1. Die operatorwertige Cauchy-Transformierte ist dann

$$\begin{aligned}
\mathcal{G}_{L_f}(B) &= \int_{\mathbb{R}} \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} - \begin{bmatrix} \cdot & t \\ t & -1 \end{bmatrix} \right)^{-1} \mathrm{d}\mu(t) \\
&= \int_{\mathbb{R}} \begin{bmatrix} b_{11} & b_{12} - t \\ b_{21} - t & b_{22} + 1 \end{bmatrix}^{-1} \mathrm{d}\mu(t) \\
&= \int_{\mathbb{R}} \frac{1}{b_{11}(b_{22}+1) - (b_{12}-t)(b_{21}-t)} \begin{bmatrix} b_{22}+1 & -b_{12}+t \\ -b_{21}+t & b_{11} \end{bmatrix} \mathrm{d}\mu(t) \\
&= -\int_{\mathbb{R}} \frac{1}{(\lambda_1 - t)(\lambda_2 - t)} \begin{bmatrix} b_{22}+1 & -b_{12}+t \\ -b_{21}+t & b_{11} \end{bmatrix} \mathrm{d}\mu(t).
\end{aligned}$$

Die Anwendung der Partialbruchzerlegung ergibt

$$\begin{aligned}\int_{\mathbb{R}} \frac{\alpha+\beta t}{(\lambda_1-t)(\lambda_2-t)}\,\mathrm{d}\mu(t) &= \int_{\mathbb{R}} \frac{a_1}{\lambda_1-t}\,\mathrm{d}\mu(t) + \int_{\mathbb{R}} \frac{a_2}{\lambda_2-t}\,\mathrm{d}\mu(t)\\ &= \frac{\alpha+\beta\lambda_1}{\lambda_2-\lambda_1}\,\mathcal{G}(\lambda_1) + \frac{\alpha+\beta\lambda_2}{\lambda_1-\lambda_2}\,\mathcal{G}(\lambda_2).\end{aligned}$$

Das heißt, dass sich unter Berücksichtigung der Bedingungen

$$\begin{aligned}\lambda_1+\lambda_2 &= b_{12}+b_{21} \quad \text{und}\\ \lambda_1\lambda_2 &= b_{12}b_{21} - b_{11}(b_{22}+1)\end{aligned}$$

die operatorwertige Cauchy-Transformierte durch die skalare Cauchy-Transformierte darstellen lässt. Für

$$B = \Lambda(z) = \begin{bmatrix} z & \cdot \\ \cdot & \cdot \end{bmatrix}$$

liefert der Eintrag $(1,1)$ der operatorwertigen Cauchy-Transformierten die Cauchy-Transformierte für $f(x)$. Die Bedingungen $\lambda_1+\lambda_2 = 0$ und $\lambda_1\lambda_2 = -z$ ergeben $\lambda_1 = \sqrt{z}$ und $\lambda_2 = -\sqrt{z}$.

$$\begin{aligned}\mathcal{G}_{x^2}(z) &= \Big(\mathcal{G}_{L_f}\big(\Lambda(z)\big)\Big)_{1,1}\\ &= \frac{-1}{\lambda_2-\lambda_1}\mathcal{G}(\lambda_1) + \frac{-1}{\lambda_1-\lambda_2}\mathcal{G}(\lambda_2)\\ &= \frac{1}{2\sqrt{z}}\mathcal{G}(\lambda_1) - \frac{1}{2\sqrt{z}}\mathcal{G}(\lambda_2).\end{aligned}$$

Für x mit einer Halbkreis-Verteilung heißt das:

$$\begin{aligned}\mathcal{G}_{x^2}(z) &= \frac{1}{2\sqrt{z}}\mathcal{G}(\sqrt{z}) - \frac{1}{2\sqrt{z}}\mathcal{G}(-\sqrt{z})\\ &= \frac{\sqrt{z}-\sqrt{z-4}}{4\sqrt{z}} - \frac{-\sqrt{z}+\sqrt{z-4}}{4\sqrt{z}}\\ &= \frac{z-\sqrt{z(z-4)}}{2z}.\end{aligned}$$

Vergleiche [NS06, Kapitel 12]. In der Formel (12.14) dort ist allerdings ein Druckfehler. Es sollte $\nu = (1-\lambda)\delta_0 + \tilde{\nu}$ für $0 \le \lambda \le 1$ heißen. Siehe [BLS96] oder [MP67].

Bemerkung. Auf diese Art die Cauchy-Transformierte eines rationalen Operators (in *einer* Variablen) zu berechnen ist zugegebenermaßen etwas umständlich. Allerdings tauchen selbst in diesem sehr einfachen Fall viele Fragen auf, die gerade beim Einarbeiten helfen können. Unter anderem muss man besonders auf die Wahl des Zweiges bei der Wurzel aufpassen, weil man hier zwei Punkte in *verschiedenen* Halbebenen der komplexen Zahlen hat. Und anhand dieses Beispiels kann man auch einfach zur R-Transformierten wechseln und sieht dann spätestens beim Versuch, die *operatorwertige* R-Transformierte zu berechnen, welche Hürden es noch gibt ...

Anhang B

Bemerkungen

Alternativ zu (den hier verwendeten) linearen Darstellungen beziehungsweise den zulässigen linearen Systemen gibt es noch andere Darstellungsmöglichkeiten (insbesondere für nicht-kommutative rationale Potenzreihen). Dazu gibt es in den Abschnitten B.1 (Linearisierung), B.2 (Realisierung) und B.3 (Reguläre lineare Systeme) weitere Information. Ein paar Anmerkungen zur Faktorisierung (und zur Abgrenzung von der Faktorisierung in Schiefpolynomringen) finden sich in Abschnitt B.4. Die Implementierung des „Rechnens mit freien Brüchen“ in Computer-Algebra-Systemen ist relativ aufwändig. Die mathematischen und algorithmischen Details findet man in den vorherigen Kapiteln. Ein paar Tipps in Abschnitt B.5 sollen bei der Strukturierung helfen.

B.1 Linearisierung

Sowohl im allgemeinen als auch im speziellen für reguläre Elemente kann man auch mit Linearisierungen (siehe folgende Definition) arbeiten. In diesem Zusammenhang sei Andersons „selbstadjungierter Linearisierungstrick“ [And13] erwähnt, ohne auf das Vorhandensein einer Involution näher einzugehen. Mit Linearisierungen kann man auch „rechnen“.

Definition B.1.1 (Linearisierung [BMS17], [CR99])**.** Eine *Linearisierung* von f ist eine Matrix $L = L_0 \otimes 1 + L_1 \otimes x_1 + \ldots + L_d \otimes x_d$, mit $L_\ell \in \mathbb{K}^{m\times m}$, der Form

$$L = \begin{bmatrix} c & u \\ v & A \end{bmatrix} \quad \in \mathbb{K}\langle X\rangle^{m\times m},$$

sodass A *voll* ist, das heißt *invertierbar* über $\mathbb{F}$, und $f = c - uA^{-1}v$ das Schur-Komplement ist. Ist $c = 0$, heißt L *reine* Linearisierung. Die *Größe* der Linearisierung ist $\operatorname{size}(L) = m$, die *Dimension* ist $\dim(L) = m - 1$.

Proposition B.1.2 ([BMS17, Proposition 3.2])**.** *Seien $A \in \mathbb{F}^{k\times k}$, $B \in \mathbb{F}^{k\times l}$, $C \in \mathbb{F}^{l\times k}$ und $D \in \mathbb{F}^{l\times l}$ gegeben und D invertierbar in $\mathbb{F}^{l\times l}$. Dann ist die Matrix $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ invertierbar in $\mathbb{F}^{(k+l)\times(k+l)}$ genau dann, wenn das* Schur-Komplement $A - BD^{-1}C$ *in $\mathbb{F}^{k\times k}$ invertierbar ist. In diesem Fall gilt:*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} . & . \\ . & D^{-1} \end{bmatrix} + \begin{bmatrix} I_k \\ -D^{-1}C \end{bmatrix} \left(A - BD^{-1}C\right)^{-1} \begin{bmatrix} I_k & -BD^{-1} \end{bmatrix}.$$

Bemerkungen. Sei $f \in \mathbb{F}$ gegeben durch die Linearisierung L. Dann ist $f = |L|_{1,1}$ die $(1,1)$-Quasideterminante von L [GGRW05]. Wenn man über *minimale* Linearisierungen spricht, muss man spezifizieren, welche Klassen von Matrizen man meint: Skalare Einträge in der ersten Zeile oder Spalte? Rein? Und, wenn zutreffend, selbstadjungiert?

Proposition B.1.3 ([Sch17b, Proposition 1.18])**.** *Sei*

$$L = \begin{bmatrix} c & u \\ v & A \end{bmatrix}$$

eine Linearisierung der Größe n für ein Element $f \in \mathbb{F}$ und ein weiteres Element $g \in \mathbb{F}$ über die reine Linearisierung

$$\tilde{L} = \begin{bmatrix} . & \tilde{u} \\ \tilde{v} & \tilde{A} \end{bmatrix} \quad \textit{mit} \quad \tilde{A} = \begin{bmatrix} c & u & -1 \\ v & A & . \\ -1 & . & . \end{bmatrix}, \quad \tilde{u} = [0, \ldots, 0, 1], \quad \tilde{v} = \tilde{u}^\top$$

der Größe $n+2$ definiert. Dann ist $g = f$.

Beweis. Unter Anwendung der Proposition B.1.2 —Schur-Komplement bezüglich des Block-Eintrages $(2,2)$— und $b = [-1, 0, \ldots, 0]$, kann die Inverse von $\tilde{A} = \left[\begin{smallmatrix} L & b^\top \\ b & . \end{smallmatrix}\right]$ geschrieben werden als

$$\tilde{A}^{-1} = \begin{bmatrix} L^{-1} & . \\ . & . \end{bmatrix} - \begin{bmatrix} -L^{-1}b^\top \\ 1 \end{bmatrix} \left(bL^{-1}b^\top\right)^{-1} \begin{bmatrix} -bL^{-1} & 1 \end{bmatrix}.$$

Daher folgt

$$\begin{aligned}
-\tilde{u}\tilde{A}^{-1}\tilde{v} &= -\left(\begin{bmatrix} . & . \end{bmatrix} - \left(bL^{-1}b^\top\right)^{-1} \begin{bmatrix} -bL^{-1} & 1 \end{bmatrix}\right) \begin{bmatrix} . \\ 1 \end{bmatrix} \\
&= \left(bL^{-1}b^\top\right)^{-1} \\
&= \left(b\left(\begin{bmatrix} . & . \\ . & A^{-1} \end{bmatrix} + \begin{bmatrix} 1 \\ -A^{-1}v \end{bmatrix} \left(c - uA^{-1}v\right)^{-1} \begin{bmatrix} 1 & -uA^{-1} \end{bmatrix}\right) b^\top\right)^{-1} \\
&= \left(\left(\begin{bmatrix} . & . \end{bmatrix} - \left(c - uA^{-1}v\right)^{-1} \begin{bmatrix} 1 & -uA^{-1} \end{bmatrix}\right) \begin{bmatrix} -1 \\ . \end{bmatrix}\right)^{-1} \\
&= c - uA^{-1}v.
\end{aligned}$$

□

Wenn die erste Zeile oder die erste Spalte einer Linearisierung für ein $f \in \mathbb{F}$ nicht-skalare Einträge hat, kann Proposition B.1.3 verwendet werden, um eine lineare Darstellung von f zu konstruieren. Umgekehrt, gegeben eine lineare Darstellung der Dimension n (von f), die in eine solche Form gebracht werden kann, lässt sich eine Linearisierung der Größe $n-1$ konstruieren.

Beispiel B.1.4. Für den Antikommutator $xy + yx$ ist ein minimales ZLS durch

$$\left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -y & . \\ . & 1 & . & -y \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

gegeben. Dreht man die Reihenfolge der Spalten um und multipliziert man die Systemmatrix mit -1, erhält man die Linearisierung

$$L'_{xy+yx} = \begin{bmatrix} . & . & . & . & 1 \\ . & . & y & x & -1 \\ . & y & . & -1 & . \\ . & x & -1 & . & . \\ 1 & -1 & . & . & . \end{bmatrix}$$

mit der Form wie in Proposition B.1.3 und erhält eine *minimale* (reine) Linearisierung des Antikommutators

$$L_{xy+yx} = \begin{bmatrix} . & y & x \\ y & . & -1 \\ x & -1 & . \end{bmatrix}.$$

B.2 Realisierung

Um hier die Gefahr des Sich-Verlierens zu vermeiden, gibt es nur eine Definition, die aber für $p = q = 1$ viele Ähnlichkeiten mit (nicht notwendigerweise reinen) linearen Darstellungen aufweist. Weitere Literatur findet man (aus Sicht der *Steuerungstheorie*) in [BGM05] und (aus Sicht der „nicht-kommutativen“ *Funktionentheorie*) in [KVV14].

Definition B.2.1 (Realisierung [HMV06])**.** Eine *Realisierung* einer Matrix $F \in \mathbb{F}^{p\times q}$ ist ein Viertupel (A, B, C, D) mit $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$, $A_\ell \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times q}$, $C \in \mathbb{K}^{p\times n}$ und $D \in \mathbb{K}^{p\times q}$ sodass A über dem freien Schiefkörper invertierbar ist und $F = D - CA^{-1}B$ gilt. Die *Dimension* der Realisierung ist $\dim(A, B, C, D) = n$.

Bemerkung. Eine Realisierung $\mathcal{R} = (A, B, C, D)$ kann auch in Blockform geschrieben werden:

$$L_{\mathcal{R}} = \begin{bmatrix} D & C \\ B & A \end{bmatrix} \quad \in \mathbb{K}\langle X\rangle^{(p+n)\times(q+n)}.$$

Hier ist die Definition so, dass $F = |L_{\mathcal{R}}|_{1',1'}$ die $(1,1)$-Block-Quasideterminante (in Bezug auf Block D) ist [GGRW05]. Für $A = -J + L_A(X)$ erhält man die *Deskriptor-Realisierung* in [HMV06]. Realisierungen in denen B und/oder C auch nicht-skalare Einträge enthalten, werden manchmal als „Schmetterlings-Realisierungen“ bezeichnet [HMV06]. Ein anderer Zugang zur Minimierung (über Modultheorie) von Realisierungen wird in [Vol18] beschrieben. In diesem Zusammenhang könnte auch [KL98] interessant sein.

B.3 Reguläres System

Für *reguläre Elemente* (*rationale formale Potenzreihen*) können minimale lineare Darstellungen alternativ über den Erweiterten Ho-Algorithmus [FM80] aus der Hankel-Matrix oder über die Minimierung einer gegebenen linearen Darstellung über den Algorithmus von Cardon und Crochemore [CC80] konstruiert werden. Letzterer ermittelt linear abhängige Zeilen in der *Steuerbarkeitsmatrix* und linear abhängige Spalten in der *Beobachtbarkeitsmatrix*. Die grundlegende Idee geht auf Schützenberger [Sch61] zurück. Steuerbarkeit und Beobachtbarkeit wird in [KFA69, Kapitel 10] diskutiert. Berstel und Reutenauer beschreiben eine Minimierung unter Verwendung *präfix-abgeschlossener* Mengen (des freien Monoids X^*) [BR11, Kapitel 2].

Erwähnenswert ist hier vielleicht noch, dass Schützenberger *Erkennbarkeit* (engl. *recognizable*) über *endliche deterministische Automaten* (engl. *DFA*) definiert [Sch61], während das nun über *lineare Darstellungen* üblich ist [BR11]. Jedenfalls kann man dann auch über Algorithmen zur Minimierung von Automaten nachdenken [BBCF10], vorausgesetzt man achtet auf den Koeffizientenring (-körper). Vorsicht: In diesem Kontext sind lineare Darstellungen der Form $\pi = (u, M, v)$ üblich, die hier als $\pi' = (u, I - M, v)$ gedacht werden müssen.

Das Pendant zu einem ZLS für diese Art von Darstellung ist dann ein *reguläres lineares System* (PLS für engl. *proper linear system*) [SS78, Abschnitt II.1]. Ein PLS $s = v + Ms$ ist also eine spezielle Form eines zulässigen linearen Systems $As = v$ mit $A = I - M$. Mit *Konjugationen* (der Systemmatrix) bleibt diese Form erhalten. Den (eindeutigen) Lösungsvektor s erhält man einfach über die *Quasiinverse* von M:

$$s = (I - M)^{-1}v = (I + M^+)v = (I + M + M^2 + M^3 + \ldots)v.$$

Um für ein reguläres lineares System der Dimension n die $\mathbb{K}$-lineare (Un-)Abhängigkeit der linken Familie („Steuerbarkeitsmatrix“) beziehungsweise rechten Familie („Beobachtbarkeitsmatrix“) zu prüfen, genügen die ersten $n - 1$ Matrixpotenzen von M [Coh95, Lemma 6.6.3]. Ruft man sich in Erinnerung, dass man eine Vektorraumstruktur (mit dem freien Monoid X^* als Basis) zur Verfügung hat, kann man sehr einfach minimieren. Das ist in [Sch17c, Abschnitt 4] illustriert. Mit dieser naiven Vorgehensweise kann man aber für $|X| \geq 2$ schnell an (praktische) Grenzen stoßen. Aber sie kann helfen, die Details des (optimalen) Algorithmus [CC80] besser zu verstehen.

So man eine Wahl hat, wird man von Fall zu Fall unterscheiden müssen, welcher Algorithmus der praktikabelste oder beste ist. Ein paar Aspekte werden in der

Bemerkung 4.3.9 angesprochen. An weiterer Literatur sollte noch die kleine Zusammenfassung über (nicht-kommutative) rationale formale Potenzreihen [Reu96b], die algebraische Sicht auf formale Sprachen [Coh75a] und das Buch [KS86] erwähnt werden.

B.4 Faktorisierung

Zunächst scheint eine Abgrenzung der hier präsentierten Faktorisierung (in freien assoziativen Algebren) zu der in Schiefpolynomringen (bzw. -bereichen), oder —allgemeiner— Ringen, die die Ore-Bedingung [Coh03b, Abschnitt 7.1], [Coh85, Abschnitt 0.8] oder [Lam99, Abschnitt 10.B] erfüllen, angebracht. In diesem Zusammenhang wären [HL13] und [LH18] als mögliche Ausgangspunkte zu nennen. Die Faktorisierung von Schiefpolynomen hat viele Verbindungen zu anderen Bereichen. Hier seien nur zwei Arbeiten beispielhaft genannt: [Ret10] und [GRW01]. In letzterer kommen auch Quasideterminanten [GGRW05] und der freie Schiefkörper vor. Für die Faktorisierung aus dem Blickwinkel von „polynomiellen Operatorbüscheln“ sollte noch [Isa73] erwähnt werden.

Polynomfaktorisierung —und damit sind wir wieder zurück in der freien assoziativen Algebra— aus der Sicht von Realisierungen (Abschnitt B.2) wird in [HKV17] diskutiert. Dort sollte man auch starten, wenn man an (matrixwertigen) „Nullstellen“ von Polynomen und deren Interpretation aus geometrischer Sicht interessiert ist. Ansonsten ist die Literatur diesbezüglich relativ überschaubar: Caruso [Car10] beschreibt Ideen von J. Davenport, die auf Homogenisierung beruhen. Ein Vergleich, wie sich diese verschiedenen Vorgehensweisen verhalten, steht noch aus. Ein paar Spezialfälle (z.B. variablen-disjunkte Faktorisierung) werden in [ARJ15] behandelt.

Für den hier präsentierten Weg gibt es schon ein paar praktische Anhaltspunkte in der Diplomarbeit von Birgit Janko [Jan18]: Irreduzibilität (über dem algebraischen Abschluss des Grundkörpers) lässt sich bis Rang 12 testen. Bei reduziblen Polynomen ist es bis Rang 17 möglich, die Ränge der Faktoren (einer Faktorisierung) zu bestimmen. Die Bestimmung konkreter Transformationsmatrizen freilich ist damit noch nicht sichergestellt. (In konkreten Fällen sollte man daher jedenfalls lineare Techniken probieren, siehe Bemerkung 3.7.5.)

Eine Aussage darüber, in wie weit spezielle Verfahren —wie zum Beispiel signaturbasierte Algorithmen [EF17]— bei der Berechnung von Gröbner–Shirshov-Basen helfen können, ist noch nicht möglich.

Die Faktorisierung *regulärer* Elemente auf Basis von Realisierungen wird zum Beispiel in [KVV09] und [BGKR08] diskutiert. A priori ist nicht klar, wie weit das hier präsentierte Konzept für reguläre Elemente (als ein Spezialfall der allgemeinen Faktorisierung im freien Schiefkörper) mit dem der *minimalen Faktorisierung* übereinstimmt. Zunächst müsste man einmal prüfen, welche der hier beschriebenen Phänomene (Unterschied zwischen äußerem Faktor und Teiler, Verschmelzen von Atomen, etc.) sich *wie* (eingeschränkt auf rationale formale Potenzreihen) verhalten und was man über matrixwertige Nullstellen beziehungsweise Singularitäten sagen kann.

B.5 Implementierung

Hier *muss* ich mit einer Warnung beginnen: Der Aufwand für eine Implementierung des freien Schiefkörpers in Computer-Algebra-Systemen für die „praktische“ Nutzbarkeit sollte keinesfalls unterschätzt werden. Denn sobald man in die Nähe von nicht-linearen Gleichungssystemen kommt und versucht sie —außer für konkrete einzelne kleine Beispiele— zu lösen, muss man tief, sehr tief in die (hoffentlich) teilweise verfügbaren Algorithmen eintauchen.

Spätestens hier kommt das Zitat (in der Einleitung) von Bergman voll zur Geltung. Schnell kommen dutzende (an sich einfache) Programme und Algorithmen zusammen, die im Gesamten eine entsprechende Komplexität entwickeln. (Hier ist noch nicht von den eigentlichen mathematischen Konzepten die Rede). Die grundlegende Schwierigkeit liegt wohl darin, dass man kaum je eine „Black-Box“ wird implementieren können, die alles vollautomatisch erledigt. Und das impliziert, dass man die Möglichkeit haben muss, auf verschiedenen Ebenen *manuell* einzugreifen (was wiederum dazu führt, dass man viel überprüfen und viele zusätzliche Prozeduren programmieren muss).

Neben der reinen Programmierzeit sollte es genug Zeit für das Nachvollziehen von Konzepten geben. Die „selbstverständlichsten“ Funktionen von (mathematischen) Computerprogrammen *muss* man dahingehend hinterfragen, welche Auswirkungen sie auf den Algorithmus haben. Insbesondere muss man auch über solche „Trivialitäten“ wie das Lösen von linearen Gleichungssystemen (in Bezug auf die verwendete Datenstruktur) nachdenken.

Insbesondere hier gilt, dass alles auf einmal nicht geht. Deshalb ist es wichtig, sich ein paar *konkrete* (erreichbare) Ziele zu setzen, wie zum Beispiel das Rechnen mit Polynomen (und darauf aufbauend die Faktorisierung). Oder: Das Rechnen mit regulären Elementen, bei denen man die Minimalität noch mit klassischen Algorithmen überprüfen kann. Oder: Die Beschränkung der Inversionshöhe (für die Bestimmung des linken größten gemeinsamen Teilers zweier Polynome) beziehungsweise eine Pivotblockgröße, die man noch systematisch auf mögliche Verfeinerungen untersuchen kann.

Zusätzlich zu den vielen Fragen, die (hoffentlich) bis jetzt —beim Lesen dieser Arbeit— aufgetaucht sind, wird es viele weitere (in Bezug auf die Implementierung) geben. Dafür sollte man genug Zeit zur Verfügung haben und auch in der Lage sein, zu improvisieren. „Nicht-Kommutativität“ ist zwar de facto überall (in Form der Matrixmultiplikation) in Softwaresystemen vorhanden, (allgemeine) nicht-kommutative algebraische Funktionalität ist eher selten. Das ist etwas, das man sich *vorher* gut überlegen muss, nämlich welches System man wählt und wo die Vor- und Nachteile liegen. (So etwas sollte man auch nicht auf Student*Innen abwälzen, es sei denn sie sind versiert im Programmieren und es geht genau um diese Problemstellung.)

Also los! Es gibt einiges zu tun. In der gesamten Arbeit gibt es verstreut —oft auch versteckt in Beispielen— jede Menge an kleineren Algorithmen, wie zum Beispiel die Bemerkung 2.3.12. Sich eine Liste anzulegen und zu überlegen, ob und wann man sie braucht, wäre eine gute Vorbereitung. Um das zu erleichtern ist das

Stichwortverzeichnis etwas dichter als üblich, auch um nach Beispielen suchen zu können. Die Palette an Beispielen ist relativ breit, dienen sie doch auch als Test für die Implementierung (einzelner Teile).

Die folgende grobe (mögliche) Strukturierung kann nur eine Orientierungshilfe sein. Und die eher lose aneinandergereihten Anmerkungen sind keinesfalls vollständig. Bewährt hat sich eine klare Trennung zwischen Datenstruktur (im Hintergrund) und „kompakter“ Darstellung (für das Arbeiten). So praktisch es ist, am Computer mit einer Liste an Matrizen zu programmieren, so ungeeignet ist es für einen Menschen, damit zu arbeiten (zumal in konkreten Beispielen die Nicht-Null-Einträge oft mit einer Hand abzählbar sind).

Lineare Matrixbüschel

Ein zulässiges lineares System der Dimension $n = 3$ für $f = (x - xyx)^{-1}$ als lineares (multivariantes) Matrixbüschel der Dimension $n{+}1$ (bezogen auf die Monome $(1, x, y)$) wäre

$$\mathcal{A} = (u, A, v) = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & 1 & . \\ . & y & -1 \\ . & -1 & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

$$\text{„=“} \begin{bmatrix} 0 & u \\ v & A \end{bmatrix} = \left(\left[\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{array}\right], \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \right).$$

Ohne viel zu beschreiben sollte klar sein, was alles an elementaren Funktionen notwendig ist. Klarerweise braucht man nur Koeffizientenmatrizen für Buchstaben, die auch verwendet werden. Das lässt sich alles programmieren, kann aber schnell dazu führen, dass man sich verzettelt. Schließlich muss man jede Menge an „Kleinigkeiten“ implementieren, wie zum Beispiel die rationalen Operationen (Proposition 2.3.1) —wobei hier *nicht* die Arithmetik gemeint ist, sondern die grundlegende Möglichkeit des Arbeitens mit Blöcken— oder auch das Extrahieren von Teilsystemen. Darüber, wieviele „Zwischenebenen“ praktikabel sind, muss man sich —abhängig von der Anwendung— Gedanken machen. Die Grundüberlegung ist, dass man —aufbauend auf eine Datenstruktur `LMMP` (engl. für „linear multivariate matrix pencil“)— verschiedene Pakete wie zum Beispiel `NCPOLY`, `NCRATS` oder `FSF` (engl. für „free skew field“) implementieren kann.

Freie assoziative Algebra

Mit den (nicht-kommutativen) Polynomen zu beginnen hat den Vorteil, dass man relativ schnell Ergebnisse sieht, das heißt, damit arbeiten und alle Zwischenschritte einfach überprüfen kann. Selbst ohne nicht-kommutativer (algebraischer) Funktionalität lassen sich die Monome und deren Koeffizienten extrahieren und man kann eine Ausgabeform programmieren, in der die Monome dann tatsächlich Wörter (bzw. „Strings“)

sind. Da verschiedene Minimierungsalgorithmen zur Verfügung stehen, die verhältnismäßig einfach (aber etwas technisch) implementiert werden können, lassen sie sich auch vergleichen. Mit ein paar weiteren Funktionen, wie dem Erzeugen von „generischen“ Transformationsmatrizen, um die Gleichungen für die Faktorisierung zu erstellen, kann man dann bereits viel machen ...

Rationale Potenzreihen

Wenn ein ZLS nicht in der Form $(u, I - M, v)$ ist, muss man es entsprechend transformieren, kann dann aber gut mit den Matrixpotenzen M^k arbeiten (natürlich vorausgesetzt, dass man deren Einträge —Polynome mit Monomen der Länge k— auch darstellen kann. Im Prinzip könnte man gleich die (zuvor) programmierte Version von $\mathbb{K}\langle X\rangle$ verwenden und hat dann eine Matrix voller „Listen an Matrizen“. Welche Laufzeitkomplexität würde sich für $M^k \in \mathbb{K}\langle X\rangle^{n\times n}$ ergeben? Natürlich kann man alternativ auch mit Listen von Tupeln (Koeffizient und Wort) arbeiten. Für ein Alphabet mit mehr als 2 Buchstaben stößt man so oder so schnell an die Grenzen ...

Freier Schiefkörper

Hat man sich schon im Detail mit den Spezialfällen (Polynome, reguläre Elemente) auseinandergesetzt und sie implementiert, kann man ein ZLS gegebenenfalls entsprechend umformen um so mehr Funktionalität (Anzeige, Berechnung bzw. „Approximation“ der linken oder rechten Familie, etc.) verwenden zu können. Im wesentlichen stehen drei Aufgaben(pakete) an:

Rationale Operationen: Schwierig ist das nicht, nur man muss alle möglichen Spezialfälle unterscheiden, damit man wirklich nur dann minimieren muss, wenn es unbedingt notwendig ist. Dabei ist es natürlich hilfreich, wenn man weiß, ob ein ZLS tatsächlich minimal ist (Flag) und welche Form es hat. Weiß man nicht, ob ein Pivotblock verfeinert ist, sollte man auch das kennzeichnen.

Minimierungsschritte: Dazu gehört die Analyse der Blockstruktur und die Anwendung eines „Blockgleichungslösers“ (für die *linearen* Minimierungsschritte). Die Minimierung selbst ist dann verhältnismäßig einfach. Wie man damit umgeht, wenn man die Minimalität *nicht* feststellen kann (bei regulären Elementen kann man sich helfen), muss man sich gut überlegen. Ein Spezialfall ist die Prüfung, ob zwei Elemente *gleich* sind (Wortproblem). Dabei spielt die Verfeinerung der Pivotblöcke keine Rolle.

Verfeinerung von Pivotblöcken und Faktorisierung: Im allgemeinen sind nichtlineare Techniken (Berechnung von Gröbner-Basen) notwendig. Je nachdem, welche Fälle (bei konkreten Berechnungen) auftreten, kann man Alternativen versuchen. Spätestens hier sollte man sich daran erinnern, dass die rechte obere Blockstruktur bei der Invertierung die neue Pivotblockstruktur ergibt. Hat man also rechts oben eine feine Struktur (so wie zum Beispiel bei Inversen von Polynomen), führen vielleicht *lineare* Techniken (keine „Überlappung“ von Zeilen- und Spaltentransformationen, Bemerkung 3.7.5) zum Erfolg.

Numerische Stolpersteine

Verwendet man Fließkomma-Zahlen für die Darstellung des Skalarenkörpers $\mathbb{K}$, muss man besonders vorsichtig in Bezug auf Rundungsfehler sein. Das gilt insbesondere bei der Lösung von linearen Gleichungssystemen. Ad hoc lässt sich nicht sagen, ob man etwas über deren *Konditionierung* aussagen kann. Gegebenenfalls muss man etwas tiefer in die numerische lineare Algebra eintauchen [Dem97]. Über die Verwendung von 96-bit Zahlen (long double) sollte man nachdenken. Eine vollständige Liste mit den Dingen, die zu beachten sind, kann es kaum geben. Stellvertretend sei ein „kleines“ Detail erwähnt: Wir betrachten das Monom xy. Wie kann man verhindern, dass — nach vielen einzelnen Berechnungsschritten— nicht folgendes ZLS

$$\begin{bmatrix} 1 & -\frac{1}{\varepsilon}x & . \\ . & 1 & -\varepsilon y \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}$$

mit $0 < \varepsilon \ll 1$ entsteht? Für das Rechnen mit einzelnen Beispielen (wie den hier vorgestellten) wird das keine große Rolle spielen. Solange man die linearen Darstellungen noch „sieht“, wird einem das auffallen. Gegebenenfalls muss man über eine Art „Normalisierung“ der Einträge nachdenken.

Rationale Ausdrücke

Das Grundprinzip ist sehr einfach: Die Struktur der Systemmatrix eines ZLS zu untersuchen um *rekursiv* Summanden, Faktoren und Inverse von Elementen zu detektieren. So lange, bis man bei Polynomen angekommen ist und diese konkret hinschreiben kann. Was man macht, wenn das nicht vollständig möglich ist (weil man nicht-lineare Gleichungssysteme lösen müsste, um Faktoren oder Summanden zu bestimmen) ist allerdings nicht so klar. Das Beispiel von der Einleitung von Kapitel 3 (Faktorisieren) sollte folgendermaßen angezeigt werden:

$$\begin{bmatrix} x & 1 & . \\ . & y & -1 \\ . & -1 & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \quad f = x^{-1}(1 - xy)^{-1}.$$

Und formt man es um, indem man Spalte 3 von Spalte 1 subtrahiert und Zeile 1 zu Zeile 3 addiert, sollte sich die Anzeige folgendermaßen verändern:

$$\begin{bmatrix} x & 1 & . \\ 1 & y & -1 \\ . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \quad f = (1 - yx)^{-1}x^{-1}.$$

Diese „Kleinigkeit“ hat es in sich. Die Motivation dazu ist dem Wunsch entsprungen, ganz „normale“ lineare Algebra (mit Matrizen über dem freien Schiefkörper) zu betreiben, siehe zum Beispiel Abschnitt A.2. Die lineare Darstellung wird dann bei Elementen in einer Matrix „ausgeblendet“.

Epilog

Um aus dem *theoretischen* freien Schiefkörper einen *angewandten* zu machen, muss man damit arbeiten können. Dabei ist es bereits eine *enorme* Vereinfachung, wenn man *manuell* einzelne Blockminimierungsschritte (am Computer) „kontrolliert“ ausführen kann. Im Zweifelsfall sollte man mit der Automatisierung nicht zu weit gehen und stattdessen zur Intervention (des Benutzers) auffordern.

Die Entwicklung der gesamten Theorie hier wäre *ohne* (experimenteller) Implementierung (in FriCAS [Fri18]) nicht möglich gewesen. Nicht, weil es einem das „händische“ Rechnen —oder das Vertiefen in die Theorie— ganz erspart, sondern weil es einen viel schneller an die Grenzen dessen bringt, was man (gerade noch) versteht. Wieviele tausende Zeilen ich programmiert und wieder verworfen habe, lässt sich nicht mehr sagen. Aber sollte ich die Zeit finden, werde ich noch einmal von vorne (mit der Implementierung) anfangen. Das mag —außer als Antwort auf die Frage, *warum* es den freien Schiefkörper (noch) nicht im Standard (als Black-Box) gibt— irrelevant erscheinen.

Doch es ist eher als Aufforderung gedacht, sich an eine (teilweise) Implementierung heranzuwagen (jedenfalls dann, wenn man gerne programmiert) um *selbst* ein tieferes Verständnis vom freien Schiefkörper zu bekommen. Denn es werden viele Fragen auftauchen, denen man sich dann *konkret* widmen kann. So kann man Schritt für Schritt in die —zum Teil sehr abstrakte— Theorie eintauchen. Bis man (hoffentlich) zu Fragen kommt, auf die man hier (in dieser Arbeit) keine Antworten findet ...

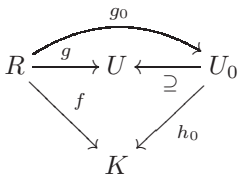
Anhang C

Cohns Konstruktion des freien Schiefkörpers

Der *freie Schiefkörper* $\mathbb{K}\langle\langle X\rangle\rangle$ ist das nicht-kommutative Pendant zum Körper der rationalen Funktionen $\mathbb{K}([X]) := \mathbb{K}(X)$, wenn man den *kommutativen* Polynomring $\mathbb{K}[X]$ durch den *nicht-kommutativen*, das heißt, die *freie assoziative Algebra* $\mathbb{K}\langle X\rangle$, (in den *nicht-kommutierenden* Variablen $X = \{x_1, x_2, \ldots, x_d\}$) ersetzt. Da es im nicht-kommutativen Fall verschiedene Quotienten(schief)körper geben kann [Lam99, Kapitel 4], muss zunächst einmal geklärt werden, was unter einem „universellen“ Quotientenkörper zu verstehen ist. Die folgende kurze Zusammenfassung ist aus [Coh06a, Kapitel 7] extrahiert. Eine etwas detailliertere findet sich in [Coh06b, Abschnitt 6].

C.1 Der universelle Quotientenkörper

Sei R ein (nicht notwendigerweise kommutativer) Ring (mit Eins) und $f : R \to K$ ein Homomorphismus in einen Körper K, sodass K durch das Bild von f erzeugt wird. Dann heißt das Paar (K, f) *epischer R-Körper*. (f ist ein *Epimorphismus* in der Kategorie der Ringe.) Ist f injektiv, so heißt (K, f) *Quotientenkörper* (von R). Ein epischer R-Körper (U, g) heißt *universell*, wenn es für jeden epischen R-Körper (K, f) einen lokalen Unterring[1] $U_0 \subseteq U$ mit $\mathrm{im}(g) \subseteq U_0$ und einen Homomorphismus $h_0 : U_0 \to K$ gibt, sodass $f = h_0 \circ g_0$ ist, wobei $g_0 = g$, aufgefasst als Abbildung $R \to U_0$. Als Diagramm:



[1] Ein Ring S heißt *lokal*, wenn das Komplement der Einheitengruppe $S \setminus S^\times$ ein Ideal in S bildet.

Der *freie Schiefkörper* $\mathbb{F} = \mathbb{K}(\langle X \rangle)$ ist der universelle Quotientenkörper der freien assoziativen Algebra $\mathbb{K}\langle X \rangle$. Der Begriff „free field“ geht auf Amitsur zurück [Ami66]. Eine quadratische Matrix A der Dimension n über R heißt *voll*, wenn für jede Faktorisierung $A = PQ$ mit $P \in R^{n \times m}$ und $Q \in R^{m \times n}$ gilt, dass $m \geq n$ ist (Definition 2.1.1, [CR99]). Die Konstruktion des freien Schiefkörpers erfolgt über die Invertierung *voller* Matrizen [Coh06a, Abschnitt 7.4]. *Nicht-volle* Matrizen werden singulär über einem R-Körper und bilden ein *Primmatrixideal* [Coh06a, Abschnitt 7.3].

Satz C.1.1 ([Coh06a, Spezialfall von Korollar 7.5.14])**.** *Sei X ein Alphabet und $\mathbb{K}$ ein kommutativer Körper. Die freie assoziative Algebra $\mathbb{K}\langle X \rangle$ hat einen universellen Quotientenkörper $\mathbb{F}$ über dem jede* volle *Matrix invertiert werden kann.*

C.2 Von der Allgemein- über die Normal- zur Standardform

Wie bereits in Bemerkung 2.1.11 erwähnt, definiert Cohn zulässige Systeme viel allgemeiner. Für jedes Element f im freien Schiefkörper $\mathbb{F} = \mathbb{K}(\langle X \rangle)$ gibt es ein „allgemeines“ zulässiges System, sodass f die *erste* Komponente dessen (eindeutig bestimmten) Lösungsvektors ist [Coh85, Abschnitt 7.1].

Die hier verwendete „lineare“ Variante in Definition 2.1.10 baut auf die (reinen) *linearen Darstellungen* von Cohn und Reutenauer [CR99] auf und kann sich somit die Vorteile beider Sichtweisen zunutze machen.

Für jedes Element f im freien Schiefkörper gibt es eine (reine) *lineare Darstellung* $\pi_f = (u, A, v)$ für ein $n \in \mathbb{N}$ mit $u^\top, v \in \mathbb{K}^{n \times 1}$ und einer *vollen* Matrix $A \in \mathbb{K}\langle X \rangle^{n \times n}$ mit (maximal) linearen Einträgen (in den Buchstaben des Alphabets X), sodass $f = uA^{-1}v$ ist. Zwei lineare Darstellungen heißen *äquivalent*, wenn sie das selbe Element repräsentieren [CR99].

Mit beliebigen Vertretern der sich aus dieser *Äquivalenzrelation* ergebenden Äquivalenzklassen zu arbeiten ist im allgemeinen sehr schwierig. Zwei *minimale* lineare Darstellungen $\pi_1 = (u_1, A_1, v_1)$ und $\pi_2 = (u_2, A_2, v_2)$ repräsentieren das selbe Element genau dann, wenn es *invertierbare* Matrizen P und Q über dem Grundkörper $\mathbb{K}$ gibt, sodass $(u_1, A_1, v_1) = \pi_1 = P\pi_2 Q = (u_2 Q, PA_2 Q, Pv_2)$ gilt [CR99]. Alle „minimalen“ Vertreter einer Äquivalenzklasse bilden die *Normalform* eines Elementes [CR94]. In diesem Fall lässt sich das Wortproblem sehr einfach lösen (Satz 4.2.3).

Die „verfeinerten“ Vertreter einer Äquivalenzklasse lassen sich —formuliert als zulässige lineare Systeme— sehr einfach minimieren (Algorithmus 4.5.15) beziehungsweise lässt sich deren Minimalität mit *linearen* Techniken feststellen (Satz 4.5.17). Und die „minimierten“ darunter bilden schließlich die *Standardform* (Definition 4.1.8). In wie weit eine weitere Spezialisierung (z.B. bei Kenntnis einer oder mehrerer Faktorisierungen oder für Elemente vom Typ $(1, *)$ oder $(*, 1)$) sinnvoll ist, muss sich erst noch zeigen ...

Appendix E

English Summary

Since all the main results are available in English already, we just provide an introduction here, mainly following [Sch18b]. The sections here correspond to the main chapters (in German) and serve as an extended table of contents to the main publications [Sch17b] (word problem, minimal inverse), [Sch17c] (polynomial factorization), [Sch17a] (general factorization theory) and [Sch18a] (constructing minimal linear representations).

For convenience we refer —additional to the internal reference— to the latest work whenever this is possible, for example "Theorem 2.5.13/[2.18]" means "[Sch18a, Theorem 2.18]".

Remark. Additional to the publications cited in the main papers, there is a perfect theoretical introduction to *fractions* [Coh84] and a very rich theoretical resource with focus on free associative algebras [Coh74].

E.1 Introduction

First of all, we need a suitable representation for the elements in the *free field* $\mathbb{F} = \mathbb{K}(\langle X \rangle)$ of the *free associative algebra* $\mathbb{K}\langle X \rangle$ over the *commutative* field $\mathbb{K}$ (for example the rational numbers $\mathbb{Q}$ or the real numbers $\mathbb{R}$) and the (finite) alphabet $X = \{x_1, x_2, \ldots, x_d\}$ (usually $X = \{x, y, z\}$). Here we use a special form of a *linear representation* of Cohn and Reutenauer [CR94], namely *admissible linear systems*.

To illustrate such a system, we consider a *linear* system of equations $As = v$ of dimension $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$, that is, we have n unknown components $s_1, s_2, \ldots, s_n$ in the solution vector s (and also v is a column vector with n rows). If A is invertible, we can write $s = A^{-1}v$. Now let $n = 1$ with $A = a \in \mathbb{Z}$ and $v \in \mathbb{Z}$ (integer entries). Then $s = a^{-1}v = \frac{v}{a}$ is a representation for a rational number $s \in \mathbb{Q}$ for $a \neq 0$. Now, given $s_1 = \frac{v_1}{a_1}$ and $s_2 = \frac{v_2}{a_2}$, we can compute the sum $s_1 + s_2 \in \mathbb{Q}$ by solving the linear

system $A's' = v'$,

$$\begin{bmatrix} a_1 & -a_1 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} s_1 + s_2 \\ s_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

(Notice the upper triangular form of the *system matrix* A', the "blocks" in the diagonal —here they have size 1×1— are called *pivot blocks*.) Usually we are interested in the *first* component of the solution vector s. If a_1 and a_2 are invertible, then A is invertible. In that case we call $As = v$ an *admissible linear system* (ALS for short). In other words: An ALS can represent a rational number. More general, one can view an ALS as a "generalized" fraction.

Important: A has to be "invertible" (for $n = 1$ we need $A \neq 0$, for $n > 1$ we need to clarify the meaning). One can extract the first component using the first identity (row) vector $u = e_1^\top = [1, 0, \ldots, 0]$, that is, the desired element $f = s_1 = uA^{-1}v$. The triple $\pi_f = (u, A, v)$ is called a *linear representation* of $f \in \mathbb{F}$. (Recall that usually we represent a rational number r by a tuple of integers (v, a), that is, $r = \frac{v}{a} = va^{-1} = a^{-1}v = 1 \cdot a^{-1}v$. So here we could write $\pi_r = (1, a, v)$.)

For the polynomial $f = xy + yx - yz \in \mathbb{K}\langle X\rangle$ an ALS $\mathcal{A}_f = (u, A, v)$ of dimension $n = 4$ is (the zeros are replaced by lower dots to emphasize the structure)

$$\begin{bmatrix} 1 & -x & -y & . \\ . & 1 & . & -y \\ . & . & 1 & z - x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xy + y(x - z) \\ y \\ x - z \\ 1 \end{bmatrix}.$$

Let $A = (a_{ij})$. The solution can be easily computed (starting from the bottom): $s_4 = 1$ and $s_i + a_{i,i+1}s_{i+1} + \ldots + a_{i,n}s_n = 0$ for $i = 3, 2, 1$. For this special form we have invertibility of A (already over $\mathbb{K}\langle X\rangle$). Is it possible to represent $f = xy+yx-yz$ by a smaller system? And, if necessary, *how* could one construct a *minimal* ALS? These are fundamental questions here, their (general) answering needs some patience.

Later we will define the *rank* of an element $f \in \mathbb{F}$ by the dimension of a *minimal* admissible linear system (for f). For a word/monomial, for example $g = xyz$, an ALS can easily be stated:

$$\begin{bmatrix} 1 & -x & . & . \\ . & 1 & -y & . \\ . & . & 1 & -z \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xyz \\ yz \\ z \\ 1 \end{bmatrix}.$$

Intuitively here it is somehow clearer (compared to the system for f before) that this ALS is minimal, but we have to make that more precise. The first goal will be to define "simple" *rational operations* on the level of these representations (systems), for example to scale, to add or to multiply elements (Proposition 2.3.1/[2.2]). That is not difficult but soon ponderous since the systems become bigger and bigger. And before we *invert* (take the reciprocal value of) an element, we have to ensure that this

is allowed. If a system is *minimal*, also that is easy. An ALS for the sum of $f_1 = 2x$ and $f_2 = 3y$ is

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & . & . \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 2 \\ . \\ 3 \end{bmatrix}.$$

What is the solution vector s? Is that system minimal for $f = f_1 + f_2 = 2x + 3y$?

Let $R = \mathbb{K}\langle X\rangle$. A (square) matrix $A \in R^{n\times n}$ is called *full*, if $A = PQ$ with $P \in R^{n\times m}$ and $Q \in R^{m\times n}$ implies $m \geq n$ [CR99]. To show that the full matrices over the *free associative algebra* are those which are invertible over the free field (and vice versa) is very difficult. For details we refer to [Coh85]. Important for us is that we can "address" *each* element f in the free field via a *linear representation* [CR99], that is, $\pi_f = (u, A, v)$ with (for some $n \in \mathbb{N}$) $u \in \mathbb{K}^{1\times n}$, full $A \in R^{n\times n}$ with entries of the form $\lambda_0 + \lambda_1 x_1 + \ldots + \lambda_d x_d$ with $\lambda_i \in \mathbb{K}$ and $x_i \in X$, $v \in \mathbb{K}^{n\times 1}$ and $f = uA^{-1}v$. If $u = [1, 0, \ldots, 0]$ we call π_f an *admissible linear system* and write $\mathcal{A}_f = \pi_f$.

Remark. The only non-invertible element in the rational numbers is zero. In our case, the non-invertible (square) matrices are the non-full matrices. Although the definition (of full matrices) is simple, testing *fullness* is very hard even for a linear matrix. An example for a non-full matrix is

$$A = \begin{bmatrix} z & . & . \\ x & . & . \\ y & -x & 1 \end{bmatrix} = \begin{bmatrix} z & 0 \\ x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ y & -x & 1 \end{bmatrix}.$$

Free Fractions

The main idea (of free fractions) is as simple as in the usage of "classical" fractions (for elements in $\mathbb{Q}$): *calculating*, *factorizing* and *minimizing* (or *cancelling*), for example

$$\frac{2}{3}\cdot\frac{3}{4} = \frac{6}{12} = \frac{2\cdot 3}{2\cdot 2\cdot 3} = \frac{1}{2} \quad \text{or}$$
$$\frac{1}{2} + \frac{3}{2} = \frac{4}{2} = \frac{2\cdot 2}{2} = 2.$$

At some point one stops this loop and uses the fraction (with *coprime* numerator and denominator, that is, their greatest common divisor is 1 or -1). However, the application (in our context) is not that easy. For a concrete expression like $f = x^{-1}zz^{-1}yz^{-1} = x^{-1}yz^{-1}$ one can find a simpler (and therefore a smaller ALS), for example

$$\begin{bmatrix} x & y \\ . & z \end{bmatrix} s = \begin{bmatrix} . \\ 1 \end{bmatrix}.$$

But what should one do with $g = x - \left(x^{-1} + (y^{-1} - x)^{-1}\right)^{-1}$ from Example 2.6.1/[5.1]?

Additionally, we need *minimal* admissible linear systems for the factorization, therefore we would run into troubles if we need the factorization for the minimization. The key idea to resolve this "dependencies" can be guessed already in the classical setting: One can remember the factorization of the numerator (for the product) and the denominator (for the sum and the product). The latter corresponds to the *standard form* (Definition 4.1.8/[3.8]).

There are a lot of definitions in Section 2.1/[1]. For an overview the mostly used will be introduced by examples. We take an element f in the free field $\mathbb{F}$ given by the admissible linear systems $\mathcal{A} = (u, A, v)$ of dimension $n = 4$. (For a rational number $r \in \mathbb{Q}$ we can write $\mathcal{A}_r = (1, a, v)$, that is, $r = 1 \cdot a^{-1} v = \frac{v}{a}$.) Recall that $f = uA^{-1}v$. If we write $s = A^{-1}v$, then f is the first component of the *solution vector* s in the system of "row" equations $As = v$. The n-tuple $(s_1, s_2, \ldots, s_n)$ of entries in s is called *left family*. (The column solution vector s and the left family are used synonymously.)

But before we take a closer look on these system of equations, we examine the "column" equations $u = tA$, in which f can be expressed as a linear combination of the components of the row solution vector $t = [t_1, t_2, \ldots, t_n]$. Here, underlined entries mean that they are *static*, that is, they must not be changed. If we describe (elementary) transformations in the following, they *always* refer to the system matrix A.

$$\underbrace{\begin{bmatrix}\underline{1} & \underline{0} & \underline{0} & \underline{0}\end{bmatrix}}_{u,\text{ left hand side}} = \underbrace{\begin{bmatrix}t_1 & t_2 & t_3 & t_4\end{bmatrix}}_{t = uA^{-1},\text{ right family}} \left.\begin{bmatrix}1-x & . & -x & -x \\ 1 & y & 1 & -2 \\ . & 1 & . & -x \\ . & . & . & 1\end{bmatrix}\right\} \begin{matrix}\text{dimension} \\ \dim(\mathcal{A}) = n\end{matrix}$$

The equations (starting from the left) are

$$\begin{aligned} 1 &= t_1(1-x) + t_2, \\ 0 &= t_2 y + t_3, \\ 0 &= -t_1 x + t_2 \quad \text{and} \\ 0 &= -t_1 x - 2t_2 - t_3 x + t_4. \end{aligned}$$

Instead of computing the solution t immediately, we will transform the system in such a way that this will be easier. Now we take a look on the system $As = v$:

$$\underbrace{\begin{bmatrix}1-x & . & -x & -x \\ 1 & y & 1 & -2 \\ . & 1 & . & -x \\ . & . & . & 1\end{bmatrix}}_{A,\text{ system matrix}} \underbrace{\begin{bmatrix}\underline{s_1} \\ s_2 \\ s_3 \\ s_4\end{bmatrix}}_{s = A^{-1}v,\text{ left family}} = \underbrace{\begin{bmatrix}. \\ -4 \\ . \\ 2\end{bmatrix}}_{v,\text{ right hand side}}$$

One equation, namely $s_4 = 2$, is especially easy to solve. Here we have $\kappa_1 s_1 + \kappa_2 s_2 + \kappa_3 s_3 + \kappa_4 s_4 = 1$ for $\kappa_1 = \kappa_2 = \kappa_3 = 0$ and $\kappa_4 = \frac{1}{2}$, therefore we write $1 \in L(\mathcal{A})$, the linear span (over $\mathbb{K}$) of the *left family*. (If there were not such a linear combination, we would write $1 \notin L(\mathcal{A})$.) We use an analogous notation for the linear span of the *right family* $R(\mathcal{A})$. Normally, we must distinguish between the element f and the representation $\mathcal{A}$. If $\mathcal{A}$ is *minimal* (which is the case here), we can define the *rank* of f as the *dimension* of $\mathcal{A}$, $\operatorname{rank}(f) := \dim(\mathcal{A})$. In this case we say "$f$ is of type $(*, 1)$" or $1 \in L(f)$ if $1 \in L(\mathcal{A})$ respectively "f is of type $(1, *)$" or $1 \in R(f)$ if $1 \in R(\mathcal{A})$.

Now we will transform this representation step by step such that the solution of both systems of equations, that is, the computation of s and t, becomes easier. Those families play a crucial role in characterizing minimality of a linear representation. However, the goal in fact will be, that we do not have to compute these solutions at all because, in general, this would not help us. Usually we write s and t (without its components) in "generic" form. The look "inside" (into the representation) is only for explanation. After the following transformation one should not forget this "inspection" and the computation of the "new" solutions s and t because this helps to understand the naming in *left* respectively *right* family.

Firstly we add 2-times row 4 to row 2 (for the solution vector t this means that we subtract 2-times t_2 from t_4). Then we exchange columns 2 and 3 (for s this means to exchange s_2 and s_3) and subtract (the new) column 2 from column 1. We collect these elementary transformations in the *admissible* transformation (P, Q), that is, the first component in the solution vector s does not change, with

$$P = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & 2 \\ . & . & 1 & . \\ . & . & . & 1 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} \underline{1} & \underline{0} & \underline{0} & \underline{0} \\ . & . & 1 & . \\ -1 & 1 & . & . \\ . & . & . & 1 \end{bmatrix}.$$

(Figure 4.1 on page 76 gives an overview of different transformation matrices.) Applying this transformation we obtain a new representation $\mathcal{A}' = (u', A', v') = P\mathcal{A}Q$,

$$\mathcal{A}' = (uQ, PAQ, Pv) = \left(\begin{bmatrix} \underline{1} & \underline{0} & \underline{0} & \underline{0} \end{bmatrix}, \begin{bmatrix} 1 & -x & . & -x \\ . & 1 & y & . \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 2 \end{bmatrix} \right).$$

The first component of the solution vector s is (still) $f = 2x - 2xyx$. Those who are not yet satisfied, can either subtract row 3 from row 1 or column 2 from column 4 and imagine our element alternatively as $x(2-2yx)$ or $(1-xy)2x$. This will be closer investigated in Section E.3 (factorizing). For polynomials we always find such a form with n (scalar) "pivot blocks" of size 1×1. This is not possible in general, but we will try to obtain small pivot blocks. Either by factorization or by "abstract" refinement (Section E.4). But we should not worry here. The examples in the beginning are such that we can easily minimize them by "hand" respectively check their minimality.

A last note concerning the system matrix A. We always write it in the compact form with (at most) *linear* entries (of non-commutative polynomials). In fact, A can also be interpreted as *linear matrix pencil* $A = (A_0, A_1, \ldots, A_d)$ with coefficient matrices $A_i \in \mathbb{K}^{n\times n}$ for an alphabet $X = \{x_1, \ldots, x_d\}$, also written as $A = A_0 + A_1x_1 + \ldots + A_dx_d$. For an implementation one can use a list of (square) matrices of size $n+1$. For the example $(x - xyx)^{-1}$ from the beginning of Section E.3 with respect to the monomials $(1, x, y)$ we have

$$\mathcal{A} = (u, A, v) = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & 1 & . \\ . & y & -1 \\ . & -1 & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

$$\text{“=”} \begin{bmatrix} 0 & u \\ v & A \end{bmatrix} = \left(\left[\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{array}\right], \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \right).$$

Left and Right Minimization Steps

For practical computations we repeatedly have to make admissible linear systems smaller. In concrete situations it is possible to minimize them. Later, in Section E.4 we will see that there are some subtle details behind the rather simple looking (*left* and *right*) “minimization steps”. Let us take a closer look on the example $2x + 3y$ from before:

$$\begin{bmatrix} 1 & -x & -1 & . \\ . & 1 & . & . \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 2 \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x + 3y \\ 2 \\ 3y \\ 3 \end{bmatrix}.$$

First we try a “left” minimization step, that is, eliminate a component of the left family. For that we subtract $\frac{2}{3}$-times row 4 from row 2 and add $\frac{2}{3}$-times column 2 to column 4:

$$\begin{bmatrix} 1 & -x & -1 & -\frac{2}{3}x \\ . & 1 & 0 & 0 \\ . & . & 1 & -y \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ 0 \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x + 3y \\ 0 \\ 3y \\ 3 \end{bmatrix}.$$

The second row reads $s_2 = 0$. That is, for the solution s_1 there is *no contribution* from (the new) s_2. Therefore we can remove the equation $s_2 = 0$ and the variable s_2 from our system of equations. Hence we get the following (not yet minimal) ALS for $2x + 3y$:

$$\begin{bmatrix} 1 & -1 & -\frac{2}{3}x \\ . & 1 & -y \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 3 \end{bmatrix}, \quad s = \begin{bmatrix} 2x + 3y \\ 3y \\ 3 \end{bmatrix}.$$

It is obvious that now it is possible to apply a "right" minimization step to eliminate t_2 (in the right family). In fact it is not necessary to compute the left or the right family at all to "minimize" (without checking minimality).

Minimality of a linear representation can be characterized by $\mathbb{K}$-linear independence of the entries of the column solution vector $s = A^{-1}v$ (the *left family*) and $\mathbb{K}$-linear independence of the entries of the row solution vector $t = uA^{-1}$ (the *right family*) [CR94, Proposition 4.7].

Since in general this is not easy to check we will investigate conditions (on the structure of the system matrix) in Section E.4 such that we can guarantee minimality if no more (block) row and column minimization steps are possible.

Sometimes a minimization is only possible in "blocks". Now we consider the ALS $\mathcal{A} = (u, A, v)$[1] for $ff^{-1} = 1$ with $f = xy - z$,

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & . & . \\ . & 1 & -y & . & . \\ . & . & 1 & -1 & . \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Here we can create an upper right block of zeros of size 3×2 in A by (as a first step) adding column 3 to column 4 and row 4 to row 2:

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & z & . \\ . & 1 & -y & . & -1 \\ . & . & 1 & 0 & 0 \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

And (as a second step) adding column 2 to column 5 and row 5 to row 1:

$$\mathcal{A}'' = \left(\begin{bmatrix} 1 & . & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z & 0 & 0 \\ . & 1 & -y & 0 & 0 \\ . & . & 1 & 0 & 0 \\ . & . & . & y & -1 \\ . & . & . & -z & x \end{bmatrix}, \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Now we can invert the lower 2×2 diagonal block (over the free field $\mathbb{F}$) and obtain $t''_4 = t''_5 = 0$ (due to the zeros in the corresponding entries in u). Hence we get the (non-minimal) ALS of dimension 3,

$$\mathcal{A}''' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & z \\ . & 1 & -x \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right).$$

[1] This ALS can be constructed in the following way: One starts with a *minimal* ALS of dimension 3 for the monomial xy (Proposition 2.2.1/[2.1]). Since xy is of type $(1, 1)$, one can immediately "add" z in the upper right entry of the system matrix to get a *minimal* ALS for $xy - z$. For the inverse we use the *minimal inverse* (Theorem 2.5.13/[2.18]). And finally, using the multiplication (Proposition 2.3.1/[2.2]) we obtain an ALS of dimension 5 for ff^{-1}.

Notice, that the lower entries in the right hand side are zero. Therefore a left block minimization step yields immediately the minimal system $\mathcal{A}''' = (1, [1], 1)$ for $1 \in \mathbb{F}$.

Notation. Given an ALS $\mathcal{A} = (u, A, v)$ with $v = [0, \ldots, 0, \lambda]^\top$ we write also write $\mathcal{A} = (1, A, \lambda)$.

Remark. The other case $f^{-1}f = 1$ is somewhat more difficult because we must not change the first component, which is „tied together“ with others, in the left family. The trick here is, to work with an "extended" ALS for $1 \cdot f^{-1}f$, using Proposition 2.3.1/[2.2] to multiply "1 from the left". For details and an illustration see Remark 4.5.3/[4.3] respectively Example 4.5.6/[4.5].

Remark. In some cases it is possible to do a left and a right minimization step *simultaneously*. This is used in Section A.1/[5] to compute the left greatest common divisor of two polynomials p and q by minimizing an ALS for $p^{-1}q$.

E.2 Calculating

One of the main parts of this section is the construction of a *minimal* admissible linear system for the inverse (of an element in the free field). The following (simple) construction (of an ALS for the inverse) is from Proposition 2.3.1/[2.2]. We assume that we have given the inverse of a monomial $f = xyz$ by the ALS $\mathcal{A}' = (u', A', v')$,

$$\begin{bmatrix} z & -1 & . \\ . & y & -1 \\ . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} z^{-1}y^{-1}x^{-1} \\ y^{-1}x^{-1} \\ x^{-1} \end{bmatrix}.$$

Checking also the $\mathbb{K}$-linear independence of the right family, the minimality is clear immediately. A (minimal) ALS for f is given by

$$\begin{bmatrix} . & z & -1 & . \\ . & . & y & -1 \\ -1 & . & . & x \\ . & 1 & . & . \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} xyz \\ 1 \\ z \\ yz \end{bmatrix},$$

with $-v$ in the upper *left* and u in the *lower* right part of the (new) system matrix. To get the form from Proposition 2.2.1/[2.1] we are already used to, we have to reverse the rows $1, 2, 3$ and columns $2, 3, 4$ and multiply the rows $1, 2, 3$ by -1. As a new system $\mathcal{A} = (u, A, v)$ for f we obtain

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & . & . \\ . & 1 & -y & . \\ . & . & 1 & -z \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right).$$

Here it is immediate, that $1 \in L(f)$ and $1 \in R(f)$, that is, f is *of type* $(1, 1)$. The application of the inverse from Proposition 2.3.1/[2.2] again yields an ALS for f^{-1},

however with dimension 5 already. Therefore an important (technical) task will be to check, *how* one can detect "special" forms (of the system matrices).

To be able to minimize, we would like to have a very "simple" structure, that is, the (diagonal) pivot blocks should be as small as possible. For the example here, an ALS for a monomial, this is respected by the *minimal inverse* (Theorem 2.5.13/[2.18]).

Preliminaries

For details see Section 2.1/[1] (or [Sch17b, Section 1]). The main definitions are Definition 2.1.4/[1.4] (linear representation), Definition 2.1.10/[1.9] (admissible linear system) and Definition 2.1.13/[1.11] (polynomial ALS).

Minimal Systems

The main idea is to start with minimal admissible linear systems and construct minimal ones for the *rational operations* (scalar multiplication, sum, product, inverse). We already have seen the *minimal monomial* (Proposition 2.2.1/[2.1]). More general it is possible to state *minimal* systems for a class of polynomials by a (generalized) "companion" system (Definition 2.2.3, [Sch17c, Section 3]).

Rational Operations

"Basic" rational operations (on the level of admissible linear systems) are easy to formulate (Proposition 2.3.1/[2.2]). For the multiplication we can provide alternative constructions yielding minimal admissible linear systems immediately in special cases, for example the *minimal polynomial multiplication* (Proposition 3.2.7/[2.16] or [Sch17c, Proposition 2.6]).

Disjoint Addition

For *disjoint* elements $f, g \in \mathbb{F}$ [CR99], that is, $\mathrm{rank}(f) + \mathrm{rank}(g) = \mathrm{rank}(f + g)$, the addition from Proposition 2.3.1/[2.2] is minimal. For further details we refer to the remarks after [Sch17a, Definition 2.2]. An important result of Cohn and Reutenauer is the *primary decomposition* (of elements in the free field) [CR99, Theorem 2.3], Theorem 2.4.2. .

Minimal Inverse

The derivation of the *minimal inverse* in Section 2.5/[Sch17b, Section 4] consists of two major steps (motivated in the beginning of this section): keeping the form for $f = (f^{-1})^{-1}$ and distinguishing different cases to ensure minimality. Notice especially the remark before [Sch17b, Theorem 4.20] how to transfer admissible linear systems into the appropriate form. The main result is Theorem 2.5.13/[2.18] respectively [Sch17b, Theorem 4.20].

Rational Identities

Using the minimal inverse and the rational operations (Proposition 2.3.1/[2.2]) one can already show non-trivial rational identities very systematically by "hand", illustrated in Example 2.6.1/[5.1].

E.3 Factorizing

Since the whole factorization theory originated from a "small" problem of the minimization of linear representations, it should lead as a thread through this section. Somehow this theory has become independent and is interesting now from a purely algebraic point of view since it enables to view the free field as a "ring". Not in the trivial sense, where each field is a ring, but using the richer "structure" by combining the non-commutative factorization theory and the embedding of non-commutative rings (to be more precise: *free ideal rings*, FIRs [Coh06a]) into their respective universal field of fraction. There are a lot of open questions, for example, is the free field a "similarity unique factorization domain"? Or, is the extension of the "classical" factorization theory (in free associative algebras) to the free field —assuming that polynomial atoms (and their inverse) remain irreducible— unique?

To not loose the thread, we come back to a simple example: Assume that we have given an element f by the admissible linear system $\mathcal{A}_f$,

$$\begin{bmatrix} x & 1 & . \\ . & y & -1 \\ . & -1 & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}.$$

By Proposition 2.3.9/[2.10] we construct an ALS $\mathcal{A}$ for fx, namely

$$\begin{bmatrix} x & 1 & . & . \\ . & y & -1 & . \\ . & -1 & x & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

Is $\mathcal{A}$ minimal? Now we repeat this step for f given by a different system $\mathcal{A}'_f$ and construct again a system $\mathcal{A}'$ for fx, namely

$$\begin{bmatrix} x & 1 & . & . \\ 1 & y & -1 & . \\ . & . & x & -x \\ . & . & . & 1 \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix},$$

in which one can read $\mathcal{A}'_f$ directly in the upper left 3×3 block of the system matrix. Here it is immediate that row/column 3 can be eliminated after adding column 3 to column 4. Therefore $\mathcal{A}'$ and hence $\mathcal{A}$ cannot be minimal. The connection to

factorization will become much clearer in Example 3.2.1/[Sch17c, Example 2.7], as soon as one verifies by the minimal inverse that $f = (pq)^{-1}$ for $p = x$ and $q = 1 - yx$.

The (lower left) 2×1 block of zeros in the system matrix of $\mathcal{A}_f$ becomes an upper right block of zeros in the system matrix of $\mathcal{A}_f^{-1}$, the standard inverse of $\mathcal{A}_f$,

$$\begin{bmatrix} 1 & -x & 1 & 0 \\ . & 1 & -y & 0 \\ . & . & -1 & -x \\ . & . & . & -1 \end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ . \\ 1 \end{bmatrix},$$

which is minimal here because f is of type $(0, 0)$ and $\mathcal{A}_f$ is minimal. And this upper right block of zeros is that one coming from multiplication $(1, *)$ or $(*, 1)$, see Proposition 2.3.6/[2.7] respectively Proposition 2.3.9/[2.10]. This yields a "natural" correspondence between factorizations and upper right zero block structure in the system matrix (assuming zero entries in the corresponding components of the right hand side).

In other words: One can find (non-trivial) factors of a polynomial by looking for "appropriate" transformations (of a *minimal* ALS). This is the main topic in Section 3.3/[Sch17c, Section 2]. If one factorizes a polynomial in two (not necessarily irreducible) factors, "their" admissible linear systems are *minimal*. The converse — and that is the core of Section 3.2— is also true. Example 3.3.1/[Sch17c, Example 3.7] could serve as an appetizer. There the polynomial factorization is used to compute the eigenvalues of a matrix via the factorization of its characteristic polynomial.

Preliminaries

The main definitions are in Section 3.1/[Sch17c, Section 1, Page 5].

Minimal Polynomial Multiplication

As an introduction one could take the multiplication of x and $1 - yx$ from Example 3.2.1/[Sch17c, Example 2.7]. To show minimality of the polynomial multiplication (Proposition 3.2.7/[2.16] or [Sch17c, Proposition 2.6]) some preparation is necessary. One of the key tools is (again) Cohn's [Coh95, Corollary 6.3.6] (Lemma 2.1.3/[1.3]). The proof for the minimal polynomial multiplication developed in Section 3.2/[Sch17a, Section 2] is slightly more general than that of [Sch17c, Proposition 2.6]. The main result is Proposition 3.2.7/[2.16].

Polynomial Factorization

The polynomial factorization theory depends on *minimal* (polynomial) admissible linear systems. How to obtain such systems directly is discussed in Section 2.2. How to construct them in general is discussed in Section 4.3. The main result is Theorem 3.3.6/[Sch17c, Theorem 2.14]. The practical application is by Proposition 3.3.7/[Sch17c, Proposition 2.15], a simple variant of [CR99, Theorem 4.1].

Factorization Theory

Once one can factorize polynomials (in the free associative algebra) on the level of *minimal* admissible linear systems, one can ask if it is possible to "factorize" general elements (in the free field) in a similar way in "generalized" atoms such that the "classical" factorization (of polynomials) does not change. A detailed discussion can be found in Section 3.4/[Sch17a, Section 3].

Minimal Factor Multiplication

Given two minimal admissible systems, under which conditions are the multiplications from Proposition 2.3.1/[2.2], Proposition 2.3.6/[2.7] and Proposition 2.3.9/[2.10] *minimal*? The answer is given in Theorem 3.5.2/[Sch17a, Theorem 4.2] within the (framework of the) general factorization theory.

General Factorization

Like in the general (minimal) multiplication (in Section 3.5) we have to distinguish several cases for the factorization in Theorem 3.6.6/[Sch17a, Theorem 4.8]. Looking for zero (lower left and upper right) blocks (of appropriate size) in the system matrix of a *minimal* ALS (similar to the polynomial factorization) is rather natural when we want to "reverse" the multiplication. The main difficulties however are far away from obvious and therefore one of the first steps in the general factorization theory in Section 3.4/[Sch17a, Section 3] is to define, *what* we mean by a "factor" (since in a field there are no non-zero non-units, that is, *each* non-zero element is invertible).

Examples Factorization

Polynomial factorization is illustrated in detail in Example 3.7.1/[Sch17c, Section 4]. The general factorization (of a regular element) is discussed briefly in Example 3.7.6/[Sch17a, Example 4.9].

E.4 Minimizing

The basic idea of the minimization (of a linear representation) with left and right minimization steps is surprisingly simple. If the block structure becomes coarser and a "look" is not sufficient any more, row and column transformations can be found by solving a *linear* system of equations. That is the essential content of Section 4.2/[Sch17b, Section 2] (word problem), the foundation stone of the whole theory. The naive idea was to solve "local" word problems, producing plenty of questions which —among other things— led to the factorization theory ...

But *when*, that is, under which conditions, is an admissible linear system (constructed out of two *minimal* ones by Proposition 2.3.1/[2.2]) minimal? If there are no more left or right "linear" minimization steps possible? Is it sufficient to find *one*

"finest" structure such that the system matrix is an upper block triangular matrix with a maximal number of (quadratic) diagonal blocks?

For polynomials (given by polynomial admissible linear systems) this can be done by a relatively simple algorithm which is formulated in Section 4.3. If one knows "all" factorizations of a polynomial, one also knows all "finest" pivot block structures of the *minimal* admissible linear systems of its inverse and one can continue to calculate "easily" because it is still rather simple to minimize.

Already in the beginning of Section E.2 (calculating) we have discussed assumptions on the construction of an ALS for the inverse of an element. In Section 4.4/[3] we investigate the connection between a factorization and the refinement of pivot blocks in the system of the inverse a little more thoroughly and describe the approach of the latter. One of the central question in Section 4.5/[4] is that of a *sufficient* condition for the minimization with *linear* techniques.

In fact one could develop a general minimization algorithm using non-linear systems of equations. However, these are usually difficult to solve. And if we do not know anything about the existence of a solution, we do not know anything about minimality. Therefore non-linear techniques should be avoided whenever this is possible by "keeping" a fine block structure.

Since the main goal of this section is to "minimize" addition and multiplication, some thoughts from this point of view should be summarized. That the factorization of an element does make sense for the multiplication is immediately clear: In case one can cancel factors. This is used for example to find the left greatest common divisor of two polynomials Section A.1/[Sch18a, Example 5.4]. But it is not that trivial since an atom might not necessarily lie "beside" its inverse. Additionally it can happen that two irreducible elements "fusion" to one (Section 3.4/[Sch17a, Section 3]) and therefore we need a refinement of pivot blocks "inside" an atom (irreducible element). But also from an additive point of view the factorization plays a crucial role because one needs "common" left and right factors of two summands only "once". Notice that there are also linear techniques for refinement, for example to bring an ALS to a suitable form for the minimal inverse (Theorem 2.5.13/[2.18] or [Sch17b, Theorem 4.20]).

Recall that here we operate *directly* in the (system matrix of the) linear representation and therefore we are *independent* of its regularity (that is, invertibility over the formal power series). And that has its price. The "classical" methods for the minimization of linear representations for *regular* elements work mainly *indirectly* by computing the left and right families, see for example Section B.3 respectively [Sch17b, Section 3].

Preliminaries and a Standard Form

To be able to formulate statements —in particular for the minimization— in a convenient way, we need some notation which formalizes what we have already used, namely to describe an ALS (and admissible transformations) in terms of *block* rows

and columns instead of (single) rows and columns. Then it is possible to define a *standard form* which plays an important role when we want to minimize admissible linear systems coming from addition or multiplication (later in Section 4.5/[4]). This is the first part in [Sch18a, Section 3]. To construct a *standard admissible linear system* out of a *minimal* ALS we need to "refine" it. This is the goal of Section 4.4, the second part in [Sch18a, Section 3].

Remark. For a polynomial p given by a *standard* ALS $\mathcal{A}$ (of dimension $n \geq 2$) the minimal inverse of $\mathcal{A}$ (of dimension $n-1$) is refined if and only if $\mathcal{A}$ is obtained by the minimal polynomial multiplication of its irreducible factors q_i in $p = q_1 q_2 \cdots q_m$. For a detailed discussion of polynomial factorization (in free associative algebras) we refer to Section 3.3/[Sch17c, Section 2].

The Word Problem

One of the difficulties in free fields is (that of) the *word problem*, that is, to check whether two admissible linear systems represent the *same* element. A solution to the word problem is [CR99, Theorem 4.1]. Unfortunately it is hard to apply practically already for two systems of dimension 3. If those systems are given by *minimal* admissible linear systems however, the word problem can be "linearized", that is, it is equivalent to the solution of a *linear* system of equations. For a detailed discussion we refer to Section 4.2/[Sch17b, Section 2]. The main result is Theorem 4.2.3/[Sch17b, Theorem 2.4]. The techniques used for the minimization in Section 4.3 and 4.5/[4] can be interpreted as solving "local" word problems. The other way around one can view the word problem as one "big" minimization step.

Minimizing a Polynomial ALS

For a *polynomial* ALS (which is refined by definition) the minimization can be done easily row by row respectively column by column. The details are in Section 4.3. The main idea is derived from the proof of the minimal polynomial multiplication (Proposition 3.2.7/[Sch17c, Proposition 2.6]).

Pivot Block Refinement

To be able to minimize an ALS using *linear* techniques only the pivot blocks have to be *refined*, that is, none can be (admissibly) transformed such that it splits in two (smaller) pivot blocks. For an illustration we refer to Section 4.4/[3]. A comprehensive discussion including algorithmic aspects will follow in [Sch18b].

Minimizing a Refined ALS

The core of the minimization is to establish the equivalence of minimality and the non-existence of solutions of certain *linear* systems of equations. Firstly we need to formalize what we have already done, namely to apply (left and right) minimization

steps (as "solutions" to *linear* systems of equations). This is somewhat technical (to implement) but rather simple. The other direction is difficult, namely to show that there is always a "linear" minimization step as long as the *refined* admissible linear system is not minimal. For the theoretical details we refer to Section 4.5/[4]. The main results are Algorithm 4.5.15 and Theorem 4.5.17 ("linear" characterization of minimality).

Literaturverzeichnis

[Aig01] M. Aigner. Catalan and other numbers: a recurrent theme. In *Algebraic combinatorics and computer science*, pages 347–390. Springer Italia, Milan, 2001.

[AL50] A. S. Amitsur and J. Levitzki. Minimal identities for algebras. *Proc. Amer. Math. Soc.*, 1:449–463, 1950.

[Ami66] S. A. Amitsur. Rational identities and applications to algebra and geometry. *J. Algebra*, 3:304–359, 1966.

[And13] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B):2103–2181, 2013.

[ARJ15] V. Arvind, G. Rattan, and P. Joglekar. On the complexity of noncommutative polynomial factorization. In *Mathematical foundations of computer science 2015. Part II*, volume 9235 of *Lecture Notes in Comput. Sci.*, pages 38–49. Springer, Heidelberg, 2015.

[BBCF10] J. Berstel, L. Boasson, O. Carton, and I. Fagnot. Minimization of automata. *ArXiv e-prints*, October 2010.

[Ber78] G. M. Bergman. The diamond lemma for ring theory. *Adv. in Math.*, 29(2):178–218, 1978.

[BGKR08] H. Bart, I. Gohberg, M. A. Kaashoek, and A. C. M. Ran. *Factorization of matrix and operator functions: the state space method*, volume 178 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2008. Linear Operators and Linear Systems.

[BGM05] J. A. Ball, G. Groenewald, and T. Malakorn. Structured noncommutative multidimensional linear systems. *SIAM J. Control Optim.*, 44(4):1474–1528, 2005.

[BHL17] J. P. Bell, A. Heinle, and V. Levandovskyy. On noncommutative finite factorization domains. *Trans. Amer. Math. Soc.*, 369(4):2675–2695, 2017.

[Bia98] P. Biane. Free probability for probabilists. *ArXiv Mathematics e-prints*, September 1998.

[BK00] L. A. Bokut′ and P. S. Kolesnikov. Gröbner-Shirshov bases: from inception to the present time. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 272(Vopr. Teor. Predst. Algebr i Grupp. 7):26–67, 345, 2000.

[BLS96] M. Bożejko, M. Leinert, and R. Speicher. Convolution and limit theorems for conditionally free random variables. *Pacific J. Math.*, 175(2):357–388, 1996.

[BMS17] S. T. Belinschi, T. Mai, and R. Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *J. Reine Angew. Math.*, 732:21–53, 2017.

[BR11] J. Berstel and C. Reutenauer. *Noncommutative rational series with applications*, volume 137 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2011.

[BRS15] P. C. Bell, D. Reidenbach, and J. Shallit. Factorization in formal languages. In *Developments in language theory*, volume 9168 of *Lecture Notes in Comput. Sci.*, pages 97–107. Springer, Cham, 2015.

[BS15] N. R. Baeth and D. Smertnig. Factorization theory: from commutative to noncommutative settings. *J. Algebra*, 441:475–551, 2015.

[Buc70] B. Buchberger. Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aequationes Math.*, 4:374–383, 1970.

[Car10] F. Caruso. Factorization of non-commutative polynomials. *ArXiv e-prints*, February 2010.

[CC80] A. Cardon and M. Crochemore. Détermination de la représentation standard d'une série reconnaissable. *RAIRO Inform. Théor.*, 14(4):371–379, 1980.

[CLO15] D. A. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015. An introduction to computational algebraic geometry and commutative algebra.

[Coh63] P. M. Cohn. Noncommutative unique factorization domains. *Trans. Amer. Math. Soc.*, 109:313–331, 1963.

[Coh72] P. M. Cohn. Generalized rational identities. In *Ring theory (Proc. Conf., Park City, Utah, 1971)*, pages 107–115. Academic Press, New York, 1972.

[Coh73] P. M. Cohn. The word problem for free fields. *J. Symbolic Logic*, 38:309–314, 1973.

[Coh74] P. M. Cohn. Progress in free associative algebras. *Israel J. Math.*, 19:109–151, 1974.

[Coh75a] P. M. Cohn. Algebra and language theory. *Bull. London Math. Soc.*, 7:1–29, 1975.

[Coh75b] P. M. Cohn. A correction and an addendum: "The word problem for free fields" (J. Symbolic Logic **38** (1973), 309–314). *J. Symbolic Logic*, 40(1):69–74, 1975.

[Coh77] P. M. Cohn. *Skew field constructions*. Cambridge University Press, Cambridge-New York-Melbourne, 1977. London Mathematical Society Lecture Note Series, No. 27.

[Coh82a] P. M. Cohn. Determinants on free fields. In *Algebraists' homage: papers in ring theory and related topics (New Haven, Conn., 1981)*, volume 13 of *Contemp. Math.*, pages 99–108. Amer. Math. Soc., Providence, R.I., 1982.

[Coh82b] P. M. Cohn. Ringe mit distributivem Faktorverband. *Abh. Braunschweig. Wiss. Ges.*, 33:35–40, 1982.

[Coh84] P. M. Cohn. Fractions. *Bull. London Math. Soc.*, 16(6):561–574, 1984.

[Coh85] P. M. Cohn. *Free rings and their relations*, volume 19 of *London Mathematical Society Monographs*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, second edition, 1985.

[Coh95] P. M. Cohn. *Skew fields*, volume 57 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. Theory of general division rings.

[Coh03a] P. M. Cohn. *Basic algebra*. Springer-Verlag London, Ltd., London, 2003. Groups, rings and fields.

[Coh03b] P. M. Cohn. *Further algebra and applications*. Springer-Verlag London, Ltd., London, 2003.

[Coh06a] P. M. Cohn. *Free ideal rings and localization in general rings*, volume 3 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.

[Coh06b] P. M. Cohn. Localization in general rings, a historical survey. In *Noncommutative localization in algebra and topology*, volume 330 of *London Math. Soc. Lecture Note Ser.*, pages 5–23. Cambridge Univ. Press, Cambridge, 2006.

[CR94] P. M. Cohn and C. Reutenauer. A normal form in free fields. *Canad. J. Math.*, 46(3):517–531, 1994.

[CR99] P. M. Cohn and C. Reutenauer. On the construction of the free field. *Internat. J. Algebra Comput.*, 9(3-4):307–323, 1999. Dedicated to the memory of Marcel-Paul Schützenberger.

[Dem97] J. W. Demmel. *Applied numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[EF17] C. Eder and J.-C. Faugère. A survey on signature-based algorithms for computing Gröbner bases. *J. Symbolic Comput.*, 80(part 3):719–784, 2017.

[EN44] S. Eilenberg and I. Niven. The "fundamental theorem of algebra" for quaternions. *Bull. Amer. Math. Soc.*, 50:246–248, 1944.

[Fli70] M. Fliess. Sur le plongement de l'algèbre des séries rationelles non commutatives dans un corps gauche. *C. R. Acad. Sci. Paris Sér. A-B*, 271:A926–A927, 1970.

[Fli74] M. Fliess. Matrices de Hankel. *J. Math. Pures Appl. (9)*, 53:197–222, 1974.

[FM80] E. Fornasini and G. Marchesini. On the problems of constructing minimal realizations for two-dimensional filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2):172–176, 1980.

[FR04] M. Fortin and C. Reutenauer. Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank. *Sém. Lothar. Combin.*, 52:Art. B52f, 12 pp. (electronic), 2004.

[Fri18] FriCAS *Computer Algebra System*, 2018. W. Hebisch, http://axiom-wiki.newsynthesis.org/FrontPage, svn co svn://svn.code.sf.net/p/fricas/code/trunk fricas.

[Gan65] F. R. Gantmacher. *Matrizenrechnung. Teil I. Allgemeine Theorie*. Zweite, Berichtigte Auflage. Hochschulbücher für Mathematik, Band 36. VEB Deutscher Verlag der Wissenschaften, Berlin, 1965.

[Gan66] F. R. Gantmacher. *Matrizenrechnung. Teil II. Spezielle Fragen und Anwendungen*. Zweite, Berichtigte Auflage. VEB Deutscher Verlag der Wissenschaften, Berlin, 1966.

[GGRW05] I. Gelfand, S. Gelfand, V. Retakh, and R. L. Wilson. Quasideterminants. *Adv. Math.*, 193(1):56–141, 2005.

[GHK06] A. Geroldinger and F. Halter-Koch. *Non-unique factorizations*, volume 278 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2006. Algebraic, combinatorial and analytic theory.

[GRW01] I. Gelfand, V. Retakh, and R. L. Wilson. Quadratic linear algebras associated with factorizations of noncommutative polynomials and noncommutative differential polynomials. *Selecta Math. (N.S.)*, 7(4):493–523, 2001.

[GW89] K. R. Goodearl and R. B. Warfield, Jr. *An introduction to noncommutative Noetherian rings*, volume 16 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1989.

[Hal80] P. R. Halmos. The heart of mathematics. *Amer. Math. Monthly*, 87(7):519–524, 1980.

[HD04] M. Horn and N. Dourdoumas. *Regelungstechnik*. Pearson Studium, 2004. Rechnerunterstützter Entwurf zeitkontinuierlicher und zeitdiskreter Regelkreise.

[Hig40] G. Higman. The units of group-rings. *Proc. London Math. Soc. (2)*, 46:231–248, 1940.

[HKV17] J. W. Helton, I. Klep, and J. Volčič. Geometry of free loci and factorization of noncommutative polynomials. *ArXiv e-prints*, August 2017.

[HL13] A. Heinle and V. Levandovskyy. Factorization of z-homogeneous polynomials in the first (q)-weyl algebra. *ArXiv e-prints*, January 2013.

[HMS15] J. W. Helton, T. Mai, and R. Speicher. Applications of realizations (aka linearizations) to free probability. *ArXiv e-prints*, November 2015.

[HMV06] J. W. Helton, S. A. McCullough, and V. Vinnikov. Noncommutative convexity arises from linear matrix inequalities. *J. Funct. Anal.*, 240(1):105–191, 2006.

[HP00] F. Hiai and D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[HS07] D. Herbera and J. Sánchez. Computing the inversion height of some embeddings of the free algebra and the free group algebra. *J. Algebra*, 310(1):108–131, 2007.

[HS15] D. Herbera and J. Sánchez. The inversion height of the free field is infinite. *Selecta Math. (N.S.)*, 21(3):883–929, 2015.

[Ikr91] Kh. D. Ikramov. Matrix pencils—theory, applications, numerical methods. In *Mathematical analysis, Vol. 29 (Russian)*, Itogi Nauki i Tekhniki, pages 3–106. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1991. Translated in J. Soviet Math. **6**4 (1993), no. 2, 783–853.

[Isa73] G. A. Isaev. Linear factorization of polynomial operator pencils. *Mat. Zametki*, 13:551–559, 1973.

[Jan18] B. Janko. Factorization of non-commutative Polynomials and Testing Fullness of Matrices. Diplomarbeit, TU Graz, February 2018.

[Jor89] D. A. Jordan. Unique factorisation of normal elements in noncommutative rings. *Glasgow Math. J.*, 31(1):103–113, 1989.

[KFA69] R. E. Kalman, P. L. Falb, and M. A. Arbib. *Topics in mathematical system theory*. McGraw-Hill Book Co., New York-Toronto, Ont.-London, 1969.

[KL98] I. Krupnik and P. Lancaster. Minimal pencil realizations of rational matrix functions with symmetries. *Canad. Math. Bull.*, 41(2):178–186, 1998.

[Kol00] P. S. Kolesnikov. The Makar-Limanov algebraically closed skew field. *Algebra Log.*, 39(6):662–692, 754–755, 2000.

[Kol01] P. S. Kolesnikov. On various definitions of algebraically closed skew fields. *Algebra Logika*, 40(4):396–414, 502, 2001.

[KS86] W. Kuich and A. Salomaa. *Semirings, automata, languages*, volume 5 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin, 1986.

[KVV09] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov. Singularities of rational functions and minimal factorizations: the noncommutative and the commutative setting. *Linear Algebra Appl.*, 430(4):869–889, 2009.

[KVV14] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov. *Foundations of free noncommutative function theory*, volume 199 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2014.

[Lam99] T. Y. Lam. *Lectures on modules and rings*, volume 189 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999.

[LH18] V. Levandovskyy and A. Heinle. A factorization algorithm for G-algebras and its applications. *J. Symbolic Comput.*, 85:188–205, 2018.

[MAZd13] R. R. Müller, G. Alfano, B. M. Zaidel, and R. de Miguel. Applications of large random matrices in communications engineering. *ArXiv e-prints*, October 2013.

[ML85] L. Makar-Limanov. Algebraically closed skew fields. *J. Algebra*, 93(1):117–135, 1985.

[Mor05] T. Mora. *Solving polynomial equation systems. II*, volume 99 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005. Macaulay's paradigm and Gröbner technology.

[MP67] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

[MSY02] A. Mateescu, A. Salomaa, and S. Yu. Factorizations of languages and commutativity conditions. *Acta Cybernet.*, 15(3):339–351, 2002.

[MV04] V. Mehrmann and H. Voss. Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. *GAMM Mitt. Ges. Angew. Math. Mech.*, 27(2):121–152 (2005), 2004.

[Niv41] I. Niven. Equations in quaternions. *Amer. Math. Monthly*, 48:654–661, 1941.

[NS98] A. Nica and R. Speicher. Commutators of free random variables. *Duke Math. J.*, 92(3):553–592, 1998.

[NS06] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[Oct18] *GNU Octave Scientific Programming Language*, 2018. https://www.gnu.org/software/octave/.

[Ore31] O. Ore. Linear equations in non-commutative fields. *Ann. of Math. (2)*, 32(3):463–477, 1931.

[Ret10] V. Retakh. From factorizations of noncommutative polynomials to combinatorial topology. *Cent. Eur. J. Math.*, 8(2):235–243, 2010.

[Reu96a] C. Reutenauer. Inversion height in free fields. *Selecta Math. (N.S.)*, 2(1):93–109, 1996.

[Reu96b] C. Reutenauer. A survey of noncommutative rational series. In *Formal power series and algebraic combinatorics (New Brunswick, NJ, 1994)*, volume 24 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 159–169. Amer. Math. Soc., Providence, RI, 1996.

[Reu99] C. Reutenauer. Malcev-Neumann series and the free field. *Exposition. Math.*, 17(5):469–478, 1999.

[Reu08] C. Reutenauer. Michel Fliess and non-commutative formal power series. *Internat. J. Control*, 81(3):336–341, 2008.

[Rot01] G.-C. Rota. Twelve problems in probability no one likes to bring up. In *Algebraic combinatorics and computer science*, pages 57–93. Springer Italia, Milan, 2001.

[Sch61] M. P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

[Sch17a] K. Schrempf. A Factorization Theory for some Free Fields. *ArXiv e-prints*, December 2017.

[Sch17b] K. Schrempf. Linearizing the Word Problem in (some) Free Fields. *ArXiv e-prints*, January 2017.

[Sch17c] K. Schrempf. On the Factorization of Non-Commutative Polynomials (in Free Associative Algebras). *ArXiv e-prints*, June 2017.

[Sch18a] K. Schrempf. A Standard Form in (some) Free Fields: How to construct Minimal Linear Representations. *ArXiv e-prints*, March 2018.

[Sch18b] K. Schrempf. Free Fractions: An Invitation to (applied) Free Fields. *In preparation*, 1:1–30, September 2018.

[Shl05] D. Shlyakhtenko. Notes on free probability theory. *ArXiv Mathematics e-prints*, April 2005.

[Slo18a] N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences*, 2018. http://oeis.org/A000108.

[Slo18b] N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences*, 2018. http://oeis.org/A001147.

[Sme15] D. Smertnig. Factorizations of elements in noncommutative rings: A survey. *ArXiv e-prints*, July 2015.

[SS78] A. Salomaa and M. Soittola. *Automata-theoretic aspects of formal power series*. Springer-Verlag, New York-Heidelberg, 1978. Texts and Monographs in Computer Science.

[Stu02] B. Sturmfels. *Solving systems of polynomial equations*, volume 97 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2002.

[Tay73] J. L. Taylor. Functions of several noncommuting variables. *Bull. Amer. Math. Soc.*, 79:1–34, 1973.

[Vas03] V. Vasilchuk. On the asymptotic distribution of the commutator and anticommutator of random matrices. *J. Math. Phys.*, 44(4):1882–1908, 2003.

[vdW73] B. L. van der Waerden. *Hamiltons Entdeckung der Quaternionen*. Vandenhoeck & Ruprecht, Göttingen, 1973. Erweiterte Fassung eines Vortrags, gehalten in Hamburg vor der Joachim Jungius-Gesellschaft der Wissenschaften am 26. Juni 1973, Veröffentlichung der Joachim Jungius-Gesellschaft der Wissenschaften.

[Vol18] J. Volčič. Matrix coefficient realization theory of noncommutative rational functions. *J. Algebra*, 499:397–437, 2018.

[Wed14] J. H. M. Wedderburn. On continued fractions in non-commutative quantities. *Ann. of Math. (2)*, 15(1-4):101–105, 1913/14.

[Woo85] R. M. W. Wood. Quaternionic eigenvalues. *Bull. London Math. Soc.*, 17(2):137–138, 1985.

Index

Q

R

S

T

U

V

W

Z