Das Konstruktionsproblem

Timo von Oertzen

Dissertation

zur Erlangung des Grades des
Doktors der Ingenieurwissenschaften
der Naturwissenschaftlich-Technischen Fakultät I
der Universität des Saarlandes.

Tage des Kolloquiums:	2. September 2003
Dekan:	Professor Dr. Philip Slusallek
1. Gutachter:	Professor Dr. Günter Hotz
2. Gutachter:	Professor Dr. Frank-Olaf Schreyer

Inhaltsverzeichnis

Abbildungsverzeichnis

Kapitel 1

Einleitung

1.1 Motivation

Die Beschäftigung mit der Mathematik ist schon in ihren Ursprüngen stark durch die Geometrie motiviert. Die direkten praktischen Umsetzungsmöglichkeiten in der Architektur, in der Konstruktion von Werkzeugen und natürlich auch in der Waffenherstellung boten einen solch unmittelbaren Anreiz, dass die Mathematik zunächst von der Geometrie weitgehend dominiert wurde. Die Loslösung der Mathematik von dieser direkten Motivation ergab sich erst mit der Zeit, und heute noch ist es in der Didaktik der Mathematik oft am einfachsten, Begeisterung für geometrische Probleme zu wecken.

Das in [18] vorgestellte Programm *Cedric* versucht, dieses Interesse zu nutzen und darauf aufzubauen. Mit dem Programm ist es möglich, geometrische Konstruktionen am Bildschirm zu erstellen und dort einfache Sätze unmittelbar ersichtlich zu machen. Dies wird besonders unterstützt durch die Dynamik, dass heißt die Möglichkeit, freie oder semifreie Punkte direkt mit der Maus zu greifen und auf dem Bildschirm zu verschieben. Der Rest der Konstruktion folgt dann der jeweiligen Bewegung. Darüber hinaus besteht in Cedric die Möglichkeit, eine numerische Verifikation oder sogar einen exakten, symbolischen Beweis einer Vermutung durchzuführen.

1.1.1 Klassische Dynamik

In den meisten dynamischen Geometriesystemen wie z.B. Cinderella (siehe [20]) oder Geolog (siehe [13]), die eine zeichnerische Eingabeoberfläche für Konstruktionen zur Verfügung stellen, gibt es die Möglichkeit, die freien oder semifreien Objekte auf dem Bildschirm zu verschieben. Bewegt man freie Objekte, so werden Zusammenhänge zwischen den Objekten oft schnell deutlich: drei Geraden, die sich immer in einem Punkt schneiden, fallen genauso schnell ins Auge wie drei Kreise, die durch einen Punkt gehen, oder im günstigen Fall auch drei Punkte, die auf einer Geraden liegen.

Für den didaktischen Einsatz ist diese relativ simple Eigenschaft geometrischer Systeme kaum zu überschätzen. Schüler, die einen geometrischen Satz auf diese Weise anschaulich betrachten können, gewinnen so einen ganz neuen, spielerischen Zugang zu der Materie. Der Spaß an der Sache weckt dabei natürlich

auch den Wunsch an weiteren Erfahrungen, so dass der Geometrieunterricht zum 'Selbstläufer' werden kann.

Auch in der Wissenschaft treten gelegentlich geometrische Konstruktionen auf, bei denen die direkte Vorstellungsgabe durch dynamische Geometriesysteme unterstützt werden kann. Gibt man eine komplizierte Konstruktion im Hinblick auf einen bestimmten Satz ein, so fallen einem möglicherweise schon bei den Hilfsobjekten für diese Konstruktion einfache Zusammenhänge auf, die sich als Lemmata für einen Gesamtbeweis als hilfreich erweisen können.

Betrachtet man den Begriff des Beweises genauer, so findet man zwei Haupteigenschaften von Beweisen: Den Nachweis der Korrektheit eines Satzes und eine Einsicht in dessen Bedeutung.

In erster Linie soll der Beweis sicherstellen, dass ein Satz korrekt ist, d.h. man erwartet eine Aussage, dass unter bestimmten Voraussetzungen ein bestimmter Zusammenhang immer gültig sein wird, unabhängig von nicht durch die Voraussetzungen berührten Parametern.

Bei geometrischen Sätzen lässt sich ein Satz in der Regel auf die Frage zurückführen, ob ein aus der Konstruktion gegebenes multivariates Polynom identisch null ist. Hat man eine obere Schranke für den Totalgrad des Polynoms und eine Routine zu dessen Auswertung, so kann man überprüfen, ob es sich bei dem Polynom um das Nullpolynom handelt, indem man hinreichend viele verschiedene Wertepaare in das Polynom einsetzt und überprüft, ob das Polynom an diesen Stellen null ist. Diese Art der Auswertung erfüllt ebenfalls die erste der oben genannten Forderungen an einen Beweis; es handelt sich aber nicht unbedingt um eine zufriedenstellende Lösung.

Unter Umständen ist man auch bereit, auf einen vollständigen Beweis zu verzichten und die erste Forderung an einen Beweis zu einer Wahrscheinlichkeitsaussage zu relaxieren. Mit der Methode der numerischen Einsetzung vieler Wertepaare kann auf diese Weise schnell mit einer sehr hohen Wahrscheinlichkeit für die Korrektheit eines Satzes garantiert werden. Obwohl dann eine formale Korrektheit nicht mehr nachgewiesen ist, besteht in der Praxis an solchen Verifikationen auch kein reeller Zweifel mehr.

Bei einer dynamischen Bewegung werden ebenfalls eine gewisse Anzahl von Parametern - durch die Bewegung der freien Punkte - eingegeben, und die Aussage jeweils überprüft, so dass man im Allgemeinen bei einer Erhaltung eines Zusammenhangs unter dynamischer Bewegung bereits von der Korrektheit des Satzes ausgeht, obwohl natürlich eine gewisse Fehlerwahrscheinlichkeit verbleibt. Dem numerischen Beweisen hat die Dynamik allerdings voraus, dass sie im Allgemeinen eine Einsicht in den Satz ermöglicht, die der reinen Auswertung von Datenpunkten natürlich fehlt.

	vollständige Korrektheit	nicht vollständige Korrektheit
lesbar	Mathematischer Beweis	Dynamik
schwer lesbar	Automatischer Beweis	Numerische Verifikation

Die Tabelle gibt einen schematische Überblick über diese vier verschiedenen Formen, eine Verifikation durchzuführen. Sie unterscheidet zwischen lesbaren und nicht lesbaren Verifikationen sowie zwischen vollständig korrekten und mit

hoher Wahrscheinlichkeit korrekten Verifikationen. Im Sinne dieser Einteilung ist die Dynamik für sich genommen bereits eine Form des Beweises.

Zwei einfache Forderungen erscheinen unmittelbar notwendig für eine dynamische Bewegung von geometrischen Objekten auf dem Bildschirm. Die erste ist eine feste Signatur, d.h. es sollte ohne Kenntnis der Geschichte der Bewegung möglich sein, anhand der freien Parameter das momentane Aussehen der Konstruktion zu ermitteln. Das zweite ist eine Stetigkeit der Bewegung, d.h. bei Bewegung eines freien Punktes um eine sehr kleine Strecke sollten sich auch die abhängigen Objekte nur ein kleines Stück bewegen.

Leider sind diese beiden Forderungen bereits unvereinbar (ein Beweis hierfür findet sich unter anderem in [20]). Verschiedene geometrische Systeme setzen hier unterschiedliche Prioritäten; z.B. ist die Dynamik von Cinderella stetig, während die von Geolog oder auch Cedric einer festen Signatur folgt.

1.1.2 Rückwirkende Dynamik

Bewegt man in einer Konstruktion freie oder semifreie Objekte, so bezeichnen wir die resultierende Bewegung der abhängigen Punkte als vorwärtsgerichtete Dynamik. Die oben angesprochenen Vorteile dieser Dynamik sind aber insoweit eingeschränkt, als es uns nicht möglich ist, die als abhängig konstruierten Objekte zu bewegen.

Diese Arbeit beschreibt eine Erweiterung des in [18] beschriebenen Programmes, mit der es möglich ist, auch ursprünglich abhängig konstruierte Punkte zu bewegen. Dafür nennt man dem System andere, ursprünglich freie Parameter, die dann automatisch so gewählt werden, dass der bewegte Punkt an der gewünschten Position zu liegen kommt. Unseres Wissens ist Cedric das einzige dynamische Geometriesystem, in dem diese rückwirkende Dynamik möglich ist.

Das einfachste Beispiel ist ein Kreuzungspunkt zweier Geraden, die durch jeweils zwei freie Punkte definiert sind. Bei der klassischen Dynamik ist der Kreuzungspunkt natürlich unbeweglich. Bei der rückwirkenden Dynamik ist es dagegen möglich, diesen Punkt zu bewegen; zwei der zuvor freien Punkte werden dann vom System entsprechend umgesetzt.

Diese Erweiterung ermöglicht im Allgemeinen eine noch tiefere Einsicht in Zusammenhänge in geometrische Konstruktionen, baut also die Vorteile der Dynamik im Allgemeinen noch weiter aus. Um sich des Vorteils bewusst zu werden, betrachte man folgendes Beispiel.

Gegeben sei ein Dreieck ABC und die beiden Höhen durch A und durch B. Den Höhenschnittpunkt wollen wir mit D bezeichnen. Angenommen wir lassen die beiden Punkte A und B fixiert und wollen den Punkt D frei bewegen. Entsprechend muss der Punkt C so bewegt werden, dass D immer der Höhenschnittpunkt des Dreiecks ABC bleibt; wie ist das zu erreichen? Im ersten Anlauf erscheint diese Frage nicht einfach. Tatsächlich aber verhält es sich hier wie bei vielen Problemen: Ist die Lösung erst einmal bekannt, so erscheint sie offensichtlich. Bewegt man den Punkt D mit Hilfe der rückwirkenden Dynamik und beobachtet die Bewegung des Punktes C, so fällt schnell auf, dass er sich ebenso verhält, wie sich D bei vorwärtsgerichteter Dynamik unter Bewegung von C verhalten würde. Tatsächlich wird im Nachhinein auch sofort klar,

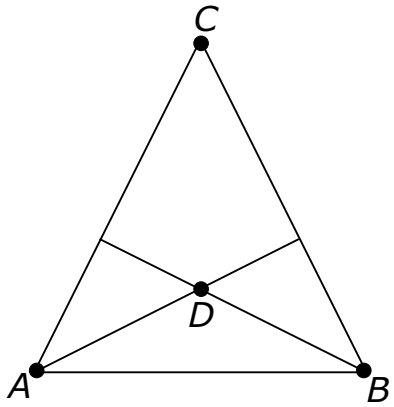

Abbildung 1.1: Der Höhenschnittpunkt im Dreieck
In den meisten dynamischen Geometriesystemen ist es möglich, den Punkt C zu bewegen, während D dieser Bewegung folgt. In Cedric kann man darüber hinaus auch den ursprünglich abhängigen Punkt D bewegen, wobei das System den Punkt C entsprechend versetzt.

dass wenn D der Höhenschnittpunkt von ABC ist, dass dann umgekehrt C der Höhenschnittpunkt von ABD ist.

Für den Geometrieunterricht ist die rückwirkende Dynamik ebenfalls von Vorteil, besonders bei der Beobachtung von Ortskurven. Die meisten Dreieckskonstruktionen, die in der Schule durchgenommen werden, beruhen auf dem Schnitt von Ortskurven, die sich durch Fixierung bestimmter Parameter im Dreieck ergeben. Nun ist es kein Problem, ein beliebiges Dreieck zu zeichnen, eine Größe (etwa eine Seitenhalbierende) einzuzeichnen, abzumessen und dann zu fixieren, wobei man einen der Dreieckspunkte auf die zugehörige eindimensionale Kurve einschränkt. Diese Kurve kann man dann durch Bewegung des Punktes sichtbar machen und erhält so unmittelbar eine Einsicht, auf was für einer Art von Ortskurve sich der dritte Dreieckspunkt bei fixierter Seitenhalbierenden bewegt. Führt man die gleichen Schritte für eine fixierte Höhe durch, so erhält man eine noch einfachere Ortskurve und auf diese Weise recht einfach eine Konstruktionsvorschrift für die Konstruktion eines Dreiecks aus gegebener Grundseite, Seitenhalbierenden und Höhe.

1.1.3 Konstruktionen

Die eben dargestellten Beispiele führen direkt zu einer weiteren Frage: Ist es möglich, die beschriebene Umkehrung der Konstruktion nicht nur numerisch approximativ, sondern auch symbolisch exakt durchzuführen? Wir stellen uns also die Frage, wie wir - gegeben die Punkte A, B und D - den Punkt C in unserem Beispiel konstruieren müssen, um die durch die Konstruktion gegebenen Nebenbedingungen aufrecht zu erhalten, also dass D der Höhenschnittpunkt des Dreiecks ABC ist. Dabei wollen wir uns wie im klassischen Konstruktionsproblem bei der Rekonstruktion von C auch nur auf Konstruktionen mit Zirkel

und Lineal beschränken. Die hier vorgestellte Erweiterung von Cedric ist in der Lage, solche Probleme zu lösen. Die theoretische Laufzeit ist wie bei vielen verwandten Problemen exponentiell, aber für viele interessante Instanzen ist der vorgestellte Algorithmus sehr schnell.

Auf der algebraischen Seite wissen wir, dass alle Koordinaten von mit Zirkel und Lineal konstruierten Punkten durch rationale Zahlen und Quadratwurzeln dargestellt werden können. Umgekehrt können wir auch Zahlen dieser Art wieder mit Zirkel und Lineal konstruieren, d.h. wir können zu einem gegebenen Ausdruck aus rationalen Zahlen und Wurzeln zwei Punkte konstruieren, die den Wert dieses Ausdrucks als Abstand haben. Wir benötigen dafür nur noch zwei weitere Punkte, deren Abstand die Länge Eins repräsentiert. Haben wir eine Konstruktionsaufgabe, d.h. eine vorgegebene Konstruktion, die nun aus einigen abhängigen Größen rekonstruiert werden soll, so besteht diese algebraisch aus einem System von Gleichungen, das identisch erfüllt werden muss.

Das Konstruktionsproblem lässt sich also bei gegebenem Polynom auf das Finden von Nullstellen dieses Polynoms zurückführen, die durch iterierte Anwendung von Quadratwurzeln über einem Grundkörper gebildet sind. Der Grundkörper ist hier eine transzendente Erweiterung der rationalen Zahlen, d.h. $\mathbb{Q}(X_1, \ldots, X_n)$. Für diese Aufgabe stellen wir in dieser Arbeit einen neuen Algorithmus vor, der an den Grundkörper nur die Voraussetzung stellt, dass Polynome über diesem Körper faktorisiert werden können. Unseres Wissens gibt es keinen anderen solchen Algorithmus. Die Laufzeit im schlechtesten Fall ist $O(d^{\log(d)} + T_f(d^2))$, wenn $T_f(d)$ die Kosten einer Faktorisierung eines Polynoms vom Grad d über dem Grundkörper sind. Dieser Fall tritt aber nur ein, wenn bestimmte zusätzliche Zusammenhänge zwischen den Koeffizienten des Eingabepolynoms bestehen; wir werden zeigen, dass dies in der Regel nicht der Fall ist, und die Laufzeit des Algorithmus im Normalfall in $O(\log(d) + T_f(d^2))$ liegt. Diese Laufzeit ist in der Praxis im Wesentlichen dominiert durch die Kosten, die die Faktorisierung eines quadratisch größeren Polynoms als das Eingabepolynom benötigt.

Über den rationalen Zahlen als Grundkörper stellten Susan Landau und Gary Lee Miller 1985 in [15] einen Algorithmus vor, der in polynomieller Zeit allgemein durch Wurzeln ausdrückbare Nullstellen finden konnte. Dieser Algorithmus hat auch theoretisch polynomielle Laufzeit, ist aber auf rationale Polynome beschränkt. Auch dieser Algorithmus basiert auf der Faktorisierung von Polynomen, die zu dieser Zeit noch nicht so weit fortgeschritten war. Es ist uns nicht bekannt, ob der Algorithmus jemals in der Praxis implementiert wurde.

Der in dieser Arbeit vorgestellte Algorithmus ist prinzipiell geeignet, sich in ähnlicher Form auch für die Suche nach anderen radikalischen oder irgendwie anders spezifizierten algebraischen Lösungen von Nullstellen anwenden zu lassen. In diesem Sinne versteht sich diese Arbeit also auch als Grundlagenforschung für kompakte Repräsentation besonderer algebraischer Elemente.

Wir definieren eine Darstellung für Ausdrücke mit Quadratwurzeln, die der Datenstruktur für diesen Algorithmus zu Grunde liegt. Wir werden feststellen, dass diese Darstellung für die meisten dieser Ausdrücke eindeutig ist; dieses Eindeutigkeitsresultat ist unseres Wissens ein neuer Beitrag zu diesem Feld.

Diese Eindeutigkeit erleichtert natürlich die Arbeit unseres Algorithmus; zudem könnte sie aber auch Auswirkung auf andere Bereiche haben, in denen es notwendig ist, algebraische Zahlen symbolisch vollständig zu repräsentieren. In vielen solchen Anwendungen wird das jeweilige Minimalpolynom der Zahl als Darstellung genutzt; in den Fällen, in denen die Zahl aber radikalisch darstellbar ist, ist eine solche Repräsentation natürlich deutlich kompakter und für viele Rechenanwendungen einfacher.

Weiterhin beweisen wir, dass die Galoisgruppen der zugehörigen Erweiterungen in der Regel eine eindeutige Gruppenstruktur besitzen, die wir explizit angeben und von der wir einige Eigenschaften zeigen. Während die Galoisgruppe für weniger komplexe Elemente meist in der Praxis recht gut bestimmt werden kann, ist dies für Elemente höheren Grades (d.h. solche, deren Minimalpolynom einen höheren Grad hat) oft schwierig; hier kann es von Vorteil sein, für beliebig komplexe Elemente Beispiele zur Verfügung zu haben, die nicht nur gut darstellbar sind (nämlich durch einen Ausdruck mit Quadratwurzeln), sondern deren Galoisgruppe auch eine bekannte Gruppenstruktur hat.

Der oben genannte Algorithmus dient zur Suche nach Nullstellen aus Quadratwurzeln univariater Polynome. Wir müssen dafür zunächst unser Gleichungssystem in ein polynomielles Gleichungssystem umwandeln und dann die Suche nach Lösungen durch Elimination auf den univariaten Fall zurückführen. Wir stellen hier neben klassischen, generellen Methoden zur Elimination auch ein Verfahren vor, das den geometrischen Ursprung des Gleichungssystem explizit ausnutzt und über Ortskurven vor Beginn der Rechnung numerisch Annahmen generiert, die das Gleichungssystem nicht unerheblich vereinfachen können.

1.2 Gliederung

Die vorliegende Arbeit beschreibt theoretische Resultate über durch Wurzeln ausdrückbare Elemente, den Algorithmus für das Finden solcher Lösungen von Gleichungssystemen, und der Umsetzung der rückwirkenden Dynamik.

Der *Euklidische Körper*, der alle durch Quadratwurzeln ausdrückbaren Erweiterungen eines Grundkörpers enthält, wird zunächst in Kapitel 2 definiert. Wir zeigen hier den Zusammenhang zwischen diesem Körper und dem Konstruktionsproblem und geben einen grundsätzlichen Überblick über den Lösungsweg für das Konstruktionsproblem.

In Kapitel 3 wird der Euklidische Körper näher betrachtet. Wir zeigen eine kompakte Repräsentation, in der alle Elemente dieses Körpers dargestellt werden können. Wir definieren die Eigenschaft der *Vollständigkeit* solcher Elemente (vollständig, da die zugehörige Galoisgruppe die 2-Sylow-Gruppe ist, also alle möglichen Elemente einer 2-Gruppe enthält), die von den meisten Elementen erfüllt wird und uns einige zusätzliche Aussagen über die Elemente erlaubt.

In Kapitel 4 gehen wir auf die Galoisgruppe von Körpererweiterungen mit Quadratwurzeln ein. Wir zeigen, dass die Galoisgruppe des Zerfällungskörpers des Minimalpolynoms eines vollständigen Elementes eine eindeutige Gruppenstruktur hat, und beweisen mit Hilfe dieser Struktur die Eindeutigkeit der kanonischen Darstellung für vollständige Elemente.

In Kapitel 5 stellen wir zunächst einen schnellen Testalgorithmus vor. Mit diesem Algorithmus ist es möglich, Polynomen schnell anzusehen, ob ihre Nullstellen durch Quadratwurzeln darstellbar sind. Für die konstruktive Berechnung dieser Nullstellen zeigen wir in diesem Kapitel, dass man die Suche nach Nullstellen auf Polynome kleineren Grades reduzieren kann, indem man spezielle Summen- und Differenzenpolynome berechnet. Für diese Berechnung führen wir einen effizienten Algorithmus an.

In Kapitel 6 gehen wir von einem Gleichungssystem in mehreren Unbekannten aus und führen dies auf das im vorigen Kapitel gelöste univariate Problem zurück. Dafür wandeln wir das ursprünglich auch Quadratwurzeln enthaltende System in ein polynomielles Gleichungssystem um und stellen dann klassische Verfahren vor, um die Anzahl der Gleichungen und die Anzahl der vorkommenden Variablen zu verringern. Neben diesen Verfahren stellen wir auch eine Vorgehensweise vor, mit der man unter konstruktionsspezifischen Annahmen das System ebenfalls vereinfachen kann; diese Annahmen können wir vorher über numerische Näherungen erzeugen.

In Kapitel 7 wenden wir die Numerik an, um die rückwirkende Dynamik, d.h. die Bewegung abhängiger Punkte, zu ermöglichen. Dabei passen wir spezielle Methoden an, damit die Umkehrung möglichst in Echtzeit ausgeführt werden kann.

Kapitel 8 fasst die Ergebnisse noch einmal kurz zusammen und nennt einige interessante, offene Probleme zu dem Themenbereich dieser Arbeit.

1.3 Notationen

Die folgende Tabelle gibt eine Übersicht über die in dieser Arbeit verwendeten Symbole. Die Mehrheit der Symbole ist kanonisch und nur der Vollständigkeit halber aufgelistet, während einzelne auch für diese Arbeit speziell definiert sind.

Zeichen	Erklärung
$\bigcup_{i\in I} M_i$	abzählbare Vereinigung von Mengen, d.h. $M_{i_1} \cup M_{i_2} \cup \dots$.
$\bigcap_{i\in I} M_i$	abzählbarer Schnitt von Mengen, d.h. $M_{i_1} \cap M_{i_2} \cap \dots$.
$\forall a \in M$	Für alle Elemente a aus M.
$\exists a \in M$	Es existiert ein Element a aus M.
$R[a]$	Ringadjunktion des Elements a an den Ring R, d.h. der kleinste Ring, der R und a umfasst.
$\mathbb{K}(a)$	Körperadjunktion des Elements a an den Körper $\mathbb{K}$, d.h. der kleinste Körper, der $\mathbb{K}$ und a umfasst.
$\{a \in M \mid Bed(a)\}$	Die Menge aller Elemente a aus M, die die Bedingung Bed erfüllen.
$\lvert p_1, p_2\rvert$	Der euklidische Abstand der Punkte p_1 und p_2.
$\overline{p_1 p_2}$	Die Gerade durch p_1 und p_2.
$a^\perp$	Für $a = (x, y)$ der Vektor $a^\perp = (y, -x)$, d.h. ein auf a senkrecht stehender Vektor.
$M_1 \times M_2$	Das kartesische Produkt von M_1 und M_2, d.h. die Menge aller Tupel (a, b) mit $a \in M_1$ und $b \in M_2$.
M^n	n mal das kartesische Produkt der Menge M mit sich selbst.
$M_1 \to M_2$	Die Menge der Funktionen von M_1 nach M_2.
$\xi^{(I)}$	Die durch I beschriebene Konjugierte von ξ nach Definition 3.2.1.
$\xi \sim \eta$	ξ ist zu η konjugiert nach Definition 3.2.1.
$[K_1 : K_2]$	Grad der Körpererweiterung von K_1 über K_2.
S_d	Die Permutationsgruppe von d Elementen.
Q_d	Die Quadratvertauschungsgruppe von 2^d Elementen nach Definition 4.2.1.
$G_1 \bullet G_2$	Das innere Produkt der Gruppen G_1 und G_2, d.h. die von $G_1 \bigcup G_2$ erzeugte Gruppe.
$\left(\begin{smallmatrix} 1 & \dots & n \\ s_1 & \dots & s_n \end{smallmatrix}\right)$	Das Element σ aus S_n mit $\sigma(i) = s_i$.
$G_1 \odot G_2$	Für $G_1 \subseteq S_{n_1}$ und $G_2 \subseteq S_{n_2}$ das Bild des natürlichen Homomorphismus von $G_1 \times G_2$ in die Gruppe $S_{n_1+n_2}$ nach Definition 4.3.1.
$\mathbb{F}_q$	Der endliche Körper mit q Elementen.
$\mathbb{Z}_n$	Die Menge $\{0, \dots, n-1\}$.
$\lvert M\rvert$	Die Mächtigkeit der Menge M.
$\det(A)$	Die Determinante der Matrix A.
I_n	Die Einheitsmatrix der Größe $n \times n$.

Kapitel 2

Aufgabenstellung

In diesem Kapitel beschäftigen wir uns zunächst mit der grundsätzlichen Problemstellung. Wir werden zeigen, auf welche Weise eine mit einer dynamischen Zeichenoberfläche formulierte Konstruktionsaufgabe sich in ein Gleichungssystem übersetzt und wie wir durch Quadratwurzeln ausdrückbare Lösungen solcher Systeme repräsentieren können. Die in diesem Kapitel besprochenen Themen sind Zusammenfassungen klassischer Geometrie.

Dieses aus der geometrischen Aufgabe gewonnene Gleichungssystem muss symbolisch und numerisch gelöst werden; auf die numerische Lösung werden wir in Kapitel 7 eingehen, während die zwischenliegenden Kapitel sich mit der exakten, symbolischen Lösung beschäftigen.

2.1 Der Euklidische Körper

Wir benötigen zunächst eine Repräsentation von Elementen, die aus endlichen Anwendungen der Körperoperationen und der Wurzeloperation aus $\mathbb{Q}$ hervorgehen. Diese Elemente bilden wieder einen Körper, und zwar einen Unterkörper der komplexen Zahlen; wir bezeichnen diesen Unterkörper als den *Euklidischen Körper*.

Für viele spätere Anwendungen benötigen wir nicht nur Wurzeloperationen über $\mathbb{Q}$, sondern auch über endlichen Körpern oder über Erweiterungen von $\mathbb{Q}$; wir definieren daher den Euklidischen Körper zunächst allgemein über einem vorgegebenen Grundkörper $\mathbb{K}$:

Definition 2.1.1. Sei $\mathbb{K}$ ein Körper. Es seien $\mathbf{k_i}$ definiert als

$$\begin{aligned} \mathbf{k_0} &= \mathbb{K} \\ \mathbf{k_{i+1}} &= \mathbf{k_i}[a \mid \exists b \in \mathbf{k_i} : \ b = a^2] \end{aligned}$$

Der *Euklidische Körper über* $\mathbb{K}$ sei dann

$$\mathbb{E}_{\mathbb{K}} := \bigcup_{i \in \mathbb{N}} \mathbf{k_i}$$

Es bleibt zu zeigen, dass $\mathbb{E}_\mathbb{K}$ und die $\mathbf{k_i}$ tatsächlich Körper sind. Folgendes Lemma zeigt dies und darüber hinaus auch, dass $\mathbb{E}_\mathbb{K}$ der kleinste Oberkörper von $\mathbb{K}$ ist, der unter Ziehen von Quadratwurzeln abgeschlossen ist.

Lemma 2.1.2. *Ist $\mathbb{K}$ ein Körper, so ist auch $\mathbb{E}_\mathbb{K}$ ein Körper, und zwar*

$$\mathbb{E}_\mathbb{K} = \bigcap \{\mathbf{K} \mid \mathbf{K} \supseteq \mathbb{K}, \forall a \in \mathbf{K}\ \exists b \in \mathbf{K} :\ b^2 = a\}$$

Beweis. Wir zeigen zunächst, dass die $\mathbf{k_i}$ aus Definition 2.1.1 Körper sind. Die unendliche Adjunktion von $\mathbf{k_i}$ zu $\mathbf{k_{i+1}}$ lässt sich als unendliche Vereinigung endlicher Ringerweiterungen vom Grad zwei schreiben, nämlich als

$$\mathbf{k_{i+1}} = \bigcup_{n\in\mathbb{N}} \bigcup_{a_1,\ldots,a_n\in\mathbf{k_i}} \mathbf{k_i}[\pm\sqrt{a_1}]\ldots[\pm\sqrt{a_n}]$$

Jede Ringerweiterung $\mathbf{k_i}[\sqrt{a}]$ ist identisch zur Körpererweiterung $\mathbf{k_i}(\sqrt{a})$, da $\frac{1}{\sqrt{a}} = \frac{\sqrt{a}}{a}$ ist, d.h. jedes Element der Vereinigung ist ein Körper. Die Vereinigung selbst ist dann ebenfalls abgeschlossen unter Körperoperationen, denn liegen zwei Elemente u und v in $\mathbf{k_i}$, so liegen u und v jeweils in zwei der vereinigten Körper, zu denen jeweils ein $n = n_1$ bzw. $n = n_2$ des ersten Vereinigungszeichens gehört. Spätestens bei $n = n_1 + n_2$ im ersten Vereinigungszeichen gibt es einen Körper, in dem sowohl u als auch v enthalten sind, und damit auch jede Körperverknüpfung der beiden Elemente. Also ist auch $\mathbb{E}_\mathbb{K}$ ein Körper.

Wir wollen jetzt noch zeigen, dass $\mathbb{E}_\mathbb{K}$ der kleinste Körper mit dieser Eigenschaft ist. Bezeichnen wir zunächst $\widetilde{\mathbb{E}_\mathbb{K}} = \bigcap \{\mathbf{K} \mid \mathbf{K} \supseteq \mathbb{K}, \forall a \in \mathbf{K}\ \exists b \in \mathbf{K}\ b^2 = a\}$. Wir haben gezeigt, dass $\mathbb{E}_\mathbb{K}$ ein Körper ist; nach Definition ist er abgeschlossen unter Wurzelziehen, ist also am Schnitt beteiligt. Daher gilt $\widetilde{\mathbb{E}_\mathbb{K}} \subseteq \mathbb{E}_\mathbb{K}$. Die Elemente von $\mathbb{E}_\mathbb{K}$ lassen sich aber aus einer Kette von Ring- und Wurzeloperationen darstellen, die nach den Forderungen auch in jedem der am Schnitt beteiligten $\mathbf{K}$ durchgeführt werden können; daraus folgt $\widetilde{\mathbb{E}_\mathbb{K}} \supseteq \mathbb{E}_\mathbb{K}$ und damit die Gleichheit. □

Über das Minimalpolynom eines Elementes definieren wir den *Grad* eines Elementes aus $\mathbb{E}_\mathbb{K}$:

Definition 2.1.3. Sei $\xi \in \mathbb{E}_\mathbb{K}$ und p das Minimalpolynom von ξ über $\mathbb{K}$. Dann sei der *Grad von ξ*:

$$\deg(\xi) = \deg(p)$$

Da es sich bei den Elementen des Euklidischen Körpers um sukzessive Anwendungen von Wurzeln über dem Grundkörper handelt, können wir folgende einfache Aussage über den Grad der Elemente machen:

Korollar 2.1.4. $\deg(\xi)$ *ist eine Zweierpotenz.*

Beweis. Nach dem Beweis zu Lemma 2.1.2 liegt jedes spezielle ξ in einer endlichen Kette von algebraischen Erweiterungen vom Grad zwei. Also ist der Grad dieser Körpererweiterung eine Zweierpotenz, und daher auch der Grad des Minimalpolynoms von ξ. □

Es stellt sich die Frage, inwieweit dieser Körper tatsächlich für geometrische Probleme interessant ist. Wir werden im folgenden Abschnitt kurz zeigen, dass der Euklidische Körper tatsächlich genau die durch Zirkel und Lineal konstruierbaren Objekte der Ebene ausmacht.

2.2 Konstruktionen mit Zirkel und Lineal und $\mathbb{E}$

In diesem Abschnitt wollen wir uns mit dem Zusammenhang zwischen Konstruktionen mit Zirkel und Lineal und dem Euklidischen Körper beschäftigen. Der Einfachheit halber beschränken wir uns bei den Konstruktionen auf die Repräsentation der Punkte; Geraden und Kreise kommen jeweils nur als Mittler zwischen solchen Punkten vor. Wir definieren jeden Punkt als Funktion anderer Punkte, und eine Konstruktion ist dann eine Menge von solchen Funktionen. Die Funktionen, die durch Konstruktion mit Zirkel und Lineal durchgeführt werden können, sind schnell aufgezählt:

Definition 2.2.1. Die Menge $\mathbf{K}$ der *konstruierbaren Punkte* wird rekursiv wie folgt definiert:

$$\begin{aligned}
\mathbf{K} &= \{null(), eins(), frei_{i,j}()\} \\
&\cup \{aufGerade_i(p_1, p_2) | p_1, p_2 \in \mathbf{K}\} \\
&\cup \{aufKreis_i(p_1, p_2) | p_1, p_2 \in \mathbf{K}\} \\
&\cup \{schnittGeradeGerade(p_1, p_2, p_3, p_4) | p_1, p_2, p_3, p_4 \in \mathbf{K}\} \\
&\cup \{schnittGeradeKreis(p_1, p_2, p_3, p_4) | p_1, p_2, p_3, p_4 \in \mathbf{K}\} \\
&\cup \{schnittKreisKreis(p_1, p_2, p_3, p_4) | p_1, p_2, p_3, p_4 \in \mathbf{K}\}
\end{aligned}$$

Die Indizes i und j an $frei$, $aufGeraden$ und $aufKreis$ sind natürliche Zahlen. Eine endliche Teilmenge $\{p_1, \ldots, p_n\} \subset \mathbf{K}$ von konstruierbaren Punkten, in denen die Argumente von p_i jeweils nur p_j mit $j < i$ sind, heißt *Konstruktion.*

Um eine sinnvolle Einbettung der konstruierbaren Punkte in ein Koordinatensystem zu ermöglichen, sind die beiden freien Punkte *null* und *eins* ausgezeichnet; der erste Punkt definiert den Ursprung des Koordinatensystems, der zweite Punkt liegt auf den Koordinaten $(1, 0)$ und definiert damit die Drehrichtung und die zugrunde liegende Streckeneinheit. Für die mit $frei$ bezeichneten freien Punkte und die semifreien Punkte $aufGerade$ und $aufKreis$ benötigen wir jeweils freie Parameter, die wir mit $X_1, X_2, \ldots$ bezeichnen. Die Indizes an $frei$ bzw. an $aufGerade$ und $aufKreis$ geben jeweils die Nummer der zugehörigen freien Parameter an. Wir suchen also eine Einbettungsfunktion der Menge der konstruierbaren Punkte in den Euklidischen Körper über $\mathbb{Q}(X_i | i \in \mathbb{N})$.

Definition 2.2.2. Die Einbettungsfunktion $\psi : \mathbf{K} \to \mathbb{E}^2_{\mathbb{Q}(X_i|i\in\mathbb{N})}$ sei definiert

als

$$\psi(f) = \begin{cases} (0,0) & f = null \\ (1,0) & f = eins \\ (X_i, X_j) & f = frei_{i,j} \\ \psi(p_1) + (1 - X_i)\psi(p_2) & f = aufGerade_i(p_1, p_2) \\ \psi(p_1) + \frac{1-X_i^2}{1+X_i^2}(\psi(p2) - \psi(p1)) & \\ \quad + \frac{2X_i}{1+X_i^2}(\psi(p2) - \psi(p1))^\perp & f = aufKreis_i(p_1, p_2) \\ gg(\psi(p_1), \psi(p_2), \psi(p_3), \psi(p_4)) & f = schnittGeradeGerade(p_1, ..., p_4) \\ gk(\psi(p_1), \psi(p_2), \psi(p_3), \psi(p_4)) & f = schnittGeradeKreis(p_1, ..., p_4) \\ kk(\psi(p_1), \psi(p_2), \psi(p_3), \psi(p_4)) & f = schnittKreisKreis(p_1, ..., p_4) \end{cases}$$

Die drei Funktionen gg, gk und kk beschreiben jeweils die Berechnung der Schnittpunkte in den Koordinaten.

Natürlich ist für jede konkrete Konstruktion jeweils nur eine endliche Anzahl n von Parametern nötig, also entsprechend auch nur eine endliche Erweiterung $\mathbb{Q}(X_1, \ldots, X_n)$ als Grundkörper.

Auf diese Weise können wir die Koordinaten eines konstruierbaren Punktes als Tupel von Elementen des Euklidischen Körpers schreiben. Tatsächlich zeigt folgendes Lemma, dass es auch umgekehrt zu jedem Tupel von Elementen aus $\mathbb{E}_{\mathbb{Q}(X_i|i\in\mathbb{N})}$ ein Element aus $\mathbf{K}$ gibt, so dass die Koordinaten dieses Punktes genau dem Tupel entsprechen. Es gilt also

Lemma 2.2.3. *ψ ist surjektiv.*

Beweis. Sei $(\xi_1, \xi_2) \in \mathbb{E}^2_{\mathbb{Q}(X_i|i\in\mathbb{N})}$. Es genügt, wenn wir zwei Punkte im Abstand ξ_1 und zwei Punkte im Abstand ξ_2 konstruieren können; wenn uns dies gelungen ist, zeichnen wir einen Kreis um *null* mit Radius ξ_1, einen Kreis durch *null* mit Radius ξ_2, eine Gerade durch die Punkte *null* und *eins* und eine Senkrechte zu dieser Geraden durch *null*. Den Kreis mit Radius ξ_1 schneiden wir mit der Geraden entlang der $X - Achse$, den anderen Kreis mit der Geraden entlang der $Y - Achse$ und erhalten so Punkte an $(\xi_1, 0)$ und $(0, \xi_2)$. Jetzt ergänzen wir nur noch den fehlenden Punkt des Rechtecks $null, (\xi_1, 0), (0, \xi_2)$ und erhalten so den erwünschten Punkt (ξ_1, ξ_2).

Als Element von $\mathbb{E}_{\mathbb{Q}(X_i|i\in\mathbb{N})}$ lässt sich ξ_i als Folge von Körperoperationen von $\mathbb{Q}(X_i|i \in \mathbb{N})$ und dem Ziehen von Wurzeln auffassen; es genügt also zu zeigen, dass wir alle X_i, alle Elemente von $\mathbb{Q}$ und die Operationen Addition, Subtraktion, Multiplikation, Division und Ziehen von Quadratwurzeln durchführen können. Da jedes Element $\frac{z}{n}$ von $\mathbb{Q}$ auch als z Additionen der Eins geteilt durch n Additionen der Eins aufgefasst werden kann, genügt es sogar, die Elemente der Grundmenge $\{0, 1, X_i\}$ für $i \in \mathbb{N}$ darstellen zu können. Dies gelingt mit *null*, *eins* und den freien Punkten.

Die Addition und die Subtraktion zweier Längen a und b sind kein Problem; haben wir a durch den Abstand zweier Punkte gegeben, so verlängern wir die Gerade und schlagen einen Kreis um einen der beiden Punkte mit Radius b. Je

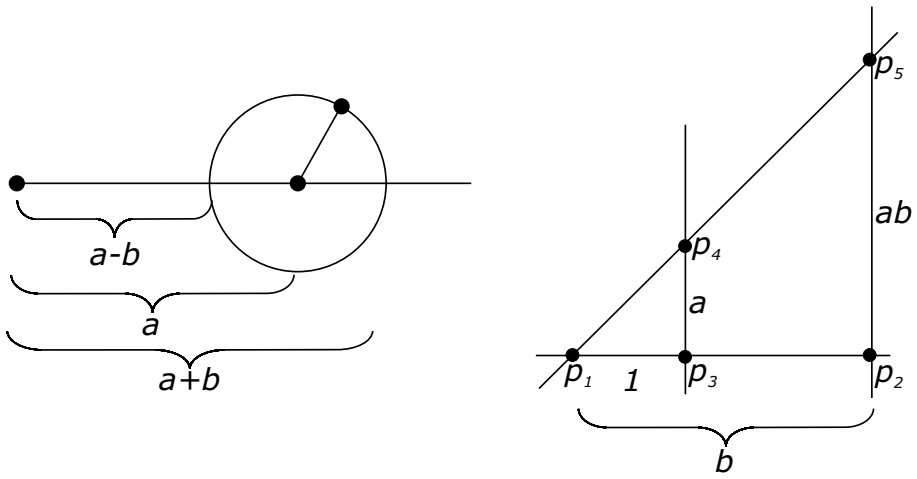

Abbildung 2.1: Konstruktion von Addition, Subtraktion und Multiplikation Im linken Bild werden die Strecken a und b subtrahiert bzw. addiert, im rechten Bild multipliziert.

nachdem, ob wir Addition oder Subtraktion erreichen wollen, wählen wir jetzt einen der beiden Schnittpunkte zwischen der Gerade und dem Kreis.

Multiplikation und Division simulieren wir durch den Strahlensatz: Wollen wir $a \cdot b$ konstruieren und haben b durch den Abstand zweier Punkte p_1 und p_2 gegeben, so schlagen wir einen Kreis um p_1 mit Radius 1 und markieren den Schnittpunkt p_3 mit der Geraden durch p_1 und p_2. Dann errichten wir auf dieser eine Senkrechte durch p_3 und tragen darauf den Abstand a ab, d.h. wir schlagen eine Kreis um p_3 mit Radius a und markieren den Schnittpunkt p_4 dieses Kreises mit der Senkrechten. Jetzt ziehen wir die Verbindungsgerade durch p_1 und p_4 und schneiden diese mit einer Senkrechten zu $\overline{p_1 p_2}$ durch p_2 in dem Punkt p_5. Nach dem Strahlensatz verhält sich $\frac{|p_2,p_5|}{b}$ wie $\frac{a}{1}$, d.h. $|p_2, p_5| = ab$.

Wollen wir umgekehrt $\frac{a}{b}$ berechnen, so zeichnen wir p_1, p_2 und p_3 genau wie eben, errichten jetzt aber zuerst einen Senkrechte zu $\overline{p1p2}$ und tragen darauf den Abstand a durch einen Punkt p_4 ab. Diesen Punkt p_4 und verbinden wir mit p_1. Die entstehende Gerade schneiden wir mit einer Senkrechten zu $\overline{p_1 p_2}$ durch den Punkt p_3 im Punkt p_5; wieder nach Strahlensatz verhält sich $\frac{a}{b} = \frac{|p5,p3|}{1}$, d.h. der Abstand von p_5 zu p_3 hat die gesuchte Länge.

Zuletzt wollen wir eine Wurzel aus a konstruieren, d.h. zu einem gegebenen Abstand $|p_1, p_2|$ eine Strecke der Länge b, so dass $b^2 = a$ ist. Dafür ziehen wir eine Strecke der Länge 1 von $|p_1, p_2|$ ab und nehmen den Mittelpunkt p_3 der verbleibenden Strecke. Bezeichne c den so gewonnenen Abstand von p_1 und p_3. Wir ziehen eine Parallele zu $\overline{p_1, p_2}$ im Abstand c und schneiden diese mit einem Kreis durch p_3 mit Radius $c + 1$ (d.h. dem Abstand $|p_3, p_2|$) im Punkt p_4. Weiterhin schneiden wir die Parallele mit einer Senkrechten zu $\overline{p1p2}$ durch p_3 im Punkt p_5. Wir erhalten damit für den Abstand $|p_4, p_5|$:

$$|p_4, p_5| = \sqrt{|p_3, p_4|^2 - |p_2, p_5|^2} = \sqrt{(c+1)^2 - c^2} = \sqrt{2c+1} = \sqrt{a}$$

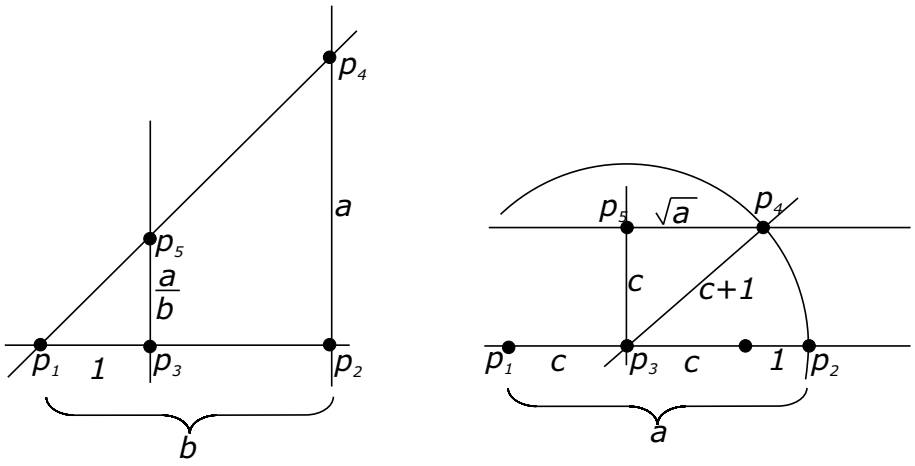

Abbildung 2.2: Konstruktion von Division und Wurzeloperator
Im linken Bild werden die Strecken a und b dividiert, während im rechten die Wurzel aus der Strecke a gezogen wird.

Damit haben wir alle Operationen simuliert, und die Aussage ist bewiesen. □

Wir haben also gezeigt, dass wir einerseits zu jedem durch Wurzeln ausdrückbaren Element auch eine Strecke entsprechender Länge konstruieren können, und dass wir andererseits jede durch Zirkel und Lineal konstruierbare Strecke durch Quadratwurzeln ausdrücken können.

Im Allgemeinen besteht eine Konstruktionsaufgabe aus mehreren Parametern, die konstruiert werden sollen. Im folgenden Abschnitt beschäftigen wir uns damit, wie wir von einer konkret gegebenen Konstruktionsaufgabe in Cedric zu einem Gleichungssystem mit Quadratwurzeln kommen.

2.3 Strategie zur Lösung des Konstruktionsproblems

Bei einer Konstruktion unterscheiden wir zwischen freien und abhängigen Parametern. Erstere werden durch den Benutzer festgelegt, indem er die entsprechenden Punkte mit der Maus an die gewünschten Stellen bewegt, während letztere durch das System auf Grundlage der Konstruktion berechnet werden. Bei der Vorbereitung einer Konstruktionsaufgabe tauscht der Benutzer den Status zweier Parameter aus, d.h. er bestimmt einen ursprünglich freien Parameter, der jetzt gebunden sein soll (d.h. vom System gesetzt wird), und dafür einen eigentlich abhängigen Parameter, der jetzt befreit sein soll. Bei der rückwirkenden Dynamik haben wir also 4 Gruppen von Parametern: unverändert freie oder abhängige Parameter und dann jeweils gleich viele gebundene und befreite Parameter.

	ursprünglich frei	ursprünglich abhängig
jetzt frei	freie Parameter	befreite Parameter
jetzt abhängig	gebundene Parameter	abhängige Parameter

So wird auf natürliche Weise ein Constraintsystem formuliert, das dann in der Bewegung erhalten bleiben soll.

Ist eine Größe durch eine Konstruktion mit Zirkel und Lineal aus den Variablen $X_1, \ldots, X_\nu$ hervorgegangen, so lässt sie sich durch einen Ausdruck aus $\mathbb{K} := \mathbb{Q}(X_1, \ldots, X_\nu)$ unter Abschluss von Quadratwurzeln darstellen, also aus $\mathbb{E}_\mathbb{K}$. Besteht die ursprüngliche Konstruktion aus n freien und m abhängigen Parametern, so können wir diese Konstruktion durch eine Funktion

$$f \in \mathbb{E}_\mathbb{K}^n \to \mathbb{E}_\mathbb{K}^m$$

beschreiben. Sei k mit $k < n$ und $k < m$ die Anzahl an befreiten bzw. gebundenen Parametern. In der Konstruktionsaufgabe sollen jetzt durch den Benutzer $n-k$ freie und k befreite Parameter bestimmt werden; nennen wir diese Werte $\xi_{k+1}, \ldots, \xi_n$ und $\eta_1, \ldots, \eta_k$. Dann besteht die Aufgabe darin, $\xi_1, \ldots, \xi_k$ und $\eta_{k+1}, \ldots, \eta_m$ zu finden, so dass

$$f(\xi_1, \ldots, \xi_n) = (\eta_1, \ldots, \eta_m)$$

gilt. Die vorgegebenen Werte $\xi_{k+1}, \ldots, \xi_n$ geben uns keine großen Probleme auf; wir setzen sie einfach in die Funktion f ein und erhalten für die restlichen k Eingabewerte eine neue Funktion mit dem Definitionsbereich $\mathbb{E}_\mathbb{K}^k$. Den Wertebereich dieser Funktion können wir auf die ersten k Einträge einschränken, denn die restlichen Werte können am Ende einfach durch Auswertung der Funktion berechnet werden. Diese Einschränkung bezeichnen wir mit F. Was bleibt, ist also bei vorgegebenen $\eta_1, \ldots, \eta_k$ die Werte $\xi_1, \ldots, \xi_k$ so zu bestimmen, dass

$$\begin{aligned} & F(\xi_1, \ldots, \xi_k) && = (\eta_1, \ldots, \eta_k) \\ \Leftrightarrow \quad & F(\xi_1, \ldots, \xi_k) - (\eta_1, \ldots, \eta_k) && = 0 \end{aligned}$$

In Kapitel 7 werden wir sehen, wie wir diese Funktion günstig numerisch approximieren können. Zunächst aber interessieren uns besonders die Lösungen, die sich symbolisch in $\xi_{k+1}, \ldots, \xi_n$ und $\eta_1, \ldots, \eta_k$ ausdrücken lassen, wobei in den Ausdrücken wieder Quadratwurzeln auftauchen dürfen. Durch Lemma 2.2.3 wissen wir, dass wir aus diesen Ausdrücken wieder eine Konstruktion mit Zirkel und Lineal erstellen können. Wir können dann also das Konstruktionsproblem lösen.

Folgende 3 Schritte führen zu diesem Ziel: Zunächst eliminieren wir die Quadratwurzeln aus F und formen das System in ein polynomielles Gleichungssystem um. Die Lösung dieses Systems führen wir dann auf die Lösung univariater Polynome zurück. Es bleibt dann die Frage, wie wir Lösungen eines univariaten Polynoms finden, die durch beliebige Anwendungen von Quadratwurzeln dargestellt werden können. Zur Lösung dieser Aufgabe müssen wir zunächst weiter ausholen und beschäftigen uns theoretisch näher mit dem Euklidischen Körper und den endlichen Erweiterungen, in denen einzelne Elemente des Euklidischen Körpers liegen.

Kapitel 3

Eigenschaften des Euklidischen Körpers

In diesem Kapitel beschäftigen wir uns mit der konkreten Repräsentation von Elementen des Euklidischen Körpers. Wir werden feststellen, dass wir jedes Element durch eine bestimmte Darstellung repräsentieren können, und dass diese Darstellung uns bei der Suche nach derartigen Nullstellen sehr hilfreich sein kann. Wir werden hier ein neues Resultat vorstellen, nämlich dass die Darstellung bei vielen Elementen, den so genannten *vollständigen* Elementen, eindeutig ist; die Eigenschaft der Vollständigkeit werden wir definieren und zeigen, dass die Elemente des Euklidischen Körpers generisch vollständig sind.

3.1 Der Darstellungssatz

In diesem Abschnitt wollen wir uns mit der Darstellung von Elementen des Euklidischen Körpers beschäftigen. Wir interessieren uns insbesondere für den Euklidischen Körper über $\mathbb{Q}$ und $\mathbb{Q}(X_1, \ldots, X_n)$, die jeweils Quotientenkörper von einfacheren Ringen sind. Betrachten wir also unseren Grundkörper $\mathbb{K}$ als Quotientenkörper eines Ringes $\mathbf{R}$.

Für die Elemente des Euklidischen Körpers $\mathbb{E}_\mathbb{K}$ gilt dann folgender Darstellungssatz:

Satz 3.1.1. *Sei $\mathbf{R}$ ein Ring und $\mathbb{K}$ der Quotientenkörper von $\mathbf{R}$, und sei $\xi \in \mathbb{E}_\mathbb{K}$. Sei weiterhin $2^d = \deg \xi$. Dann lässt sich ξ darstellen als*

$$\xi = \frac{1}{n}\left(k_0 + \sum_{i=1}^{d} k_i \sqrt{w_i}\right)$$

mit $n, k_0 \in \mathbf{R}$ und $w_i, k_i \in \mathbb{E}_\mathbb{K}$ mit $\deg(w_i) \leq 2^{i-1}$ und $\deg(k_i) \leq 2^{i-1}$ für $i = 1, \ldots, d$.

Die w_i und k_i sind Elemente der Ringerweiterung $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}]$. Alle Elemente dieser Ringerweiterung lassen sich wie ξ darstellen, und die darin vorkommenden Wurzeln sind identisch mit den w_j für $j < i$.

Beweis. Nach Definition entsteht ξ durch eine endlichen Folge von Ringoperationen und Wurzeloperationen mit Elementen aus $\mathbb{K}$. Wir betrachten die Folge der Anwendungen dieser Operationen; die Elemente direkt vor der k-ten Anwendung einer Wurzel bezeichnen wir mit w_k. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass $\sqrt{w_k} \notin \mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{k-1}}]$ ist; andernfalls lassen wir den Wurzeloperator weg und ersetzen ihn durch die Darstellung von $\sqrt{w_k}$ in $\mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{k-1}}]$.

Wir beweisen die Darstellung durch Induktion über d. Für $d = 0$ ist die Aussage klar. Sei die Aussage schon bewiesen für $d' < d$, also insbesondere für Elemente aus $\mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}]$. Der Körper $\mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_d}]$ ist jetzt eine Erweiterung vom Grad zwei über dem Ring $\mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}]$, d.h. jedes Element η vom Oberkörper und insbesondere ξ lässt sich darstellen als $\eta' + \eta''\sqrt{w_d}$ mit $\eta', \eta'' \in \mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}]$. Sowohl w_d als auch η' und η'' sind Elemente aus $\mathbb{K}[\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}]$, also nach Induktionsvoraussetzung darstellbar wie in der Behauptung mit $d - 1$ Wurzeln.

Den Nenner n von w_d können wir aus der Wurzel herausziehen und in das η'' mit eingehen lassen. Wir multiplizieren dann η' und η'' mit ihrem Hauptnenner und multiplizieren diesen zum äußeren n hinzu. □

Das beim ersten Hinschauen Erstaunliche dieser Darstellung ist die Gleichheit der Wurzeln $\sqrt{w_k}$ in der äußeren Summe und in der rekursiven Darstellung der k_i und w_i. Der Beweis macht klar, warum dies funktioniert; betrachtet man die Folge der Operationen zur Erstellung des Elementes (wenn man jeweils nur 'echte' Wurzeln, d.h. Wurzeln von Elementen, die keine Quadrate sind, zulässt), so erhält man auf diese Weise eine Folge von Erweiterungen, in denen die jeweiligen Elemente liegen.

Leider ist die durch diesen Satz gegebene Darstellung im Allgemeinen nicht eindeutig. Man betrachte dazu folgendes Beispiel:

Beispiel 3.1.2. Für $a, b, c \in \mathbf{R}$ ist

$$\begin{aligned} \sqrt{ab} + \sqrt{c(a+b) + 2c\sqrt{ab}} &= \sqrt{ac} + \sqrt{b(a+c) + 2b\sqrt{ac}} \\ &= \sqrt{bc} + \sqrt{a(b+c) + 2a\sqrt{bc}} \end{aligned}$$

Dies wird ersichtlich, wenn man die geschachtelte Wurzel $\sqrt{c(a+b) + 2c\sqrt{ab}}$ äquivalent umformt zu $\sqrt{c}\sqrt{a + 2\sqrt{ab} + b} = \sqrt{c}\left(\sqrt{a} + \sqrt{b}\right)$.

Die Eindeutigkeit ist schon deshalb nicht gegeben, weil wir einen Teil der jeweiligen k_i mit in die Wurzel $\sqrt{w_i}$ hineinziehen können. Da w_i in den Vorfaktoren k_j und Wurzeln w_j rekursiv wieder erscheint, ist das Verschieben von Faktoren in die Wurzeln nur insoweit möglich, wie diese Faktoren in allen Vorfaktoren auftauchen. Dies gilt allerdings nicht für die letzte Wurzel $\sqrt{w_d}$, die nur einmal ganz außen in der Darstellung auftaucht. Hier können wir den Faktor k_d auf jeden Fall bis auf Vorzeichen mit in die Wurzel ziehen:

Korollar 3.1.3. *Sei $\xi \in \mathbb{E}_{\mathbb{K}}$. Dann gibt es eine Darstellung von ξ wie im Darstellungssatz mit $k_d \in \{1, -1\}$. Verzichten wir darüber hinaus darauf, dass die rekursiv auftauchenden Wurzeln identisch sind und begnügen uns mit Gleichheit bis auf quadratische Faktoren in den entsprechenden Zwischenringen, so gibt es eine Darstellung mit $k_i \in \{0, 1, -1\}$ für $i < d$ und $k_d \in \{1, -1\}$.*

Beweis. $k_d \neq 0$ ist klar, da sonst der Grad des Elementes kleiner als 2^d wäre, da wir die zugehörige Körpererweiterung schon mit $d-1$ Erweiterungen vom Grad 2 erreichen könnten. Den Betrag von k_d und bei Verzicht einer unmittelbaren Identität der Wurzeln auch den Betrag der anderen k_i können wir in die jeweiligen Wurzeln mit hineinziehen. □

3.2 Das Minimalpolynom von ξ

In diesem Abschnitt werden wir darauf eingehen, in welchem Zusammenhang die Darstellung eines Elements des Euklidischen Körpers zu dessen Minimalpolynom steht. Wir werden feststellen, dass man alle Nullstellen des Minimalpolynoms durch Vertauschung der Vorzeichen in der durch den Darstellungssatz gegebenen Darstellung erhält.

Sei also $\xi \in \mathbb{E}_{\mathbb{K}}$. Wir interessieren uns für das Minimalpolynom p von ξ. Da $\deg(\xi) = \deg(p)$ ist, kennen wir schon den Grad dieses Polynoms; wie aber sehen seine anderen Nullstellen aus? Dafür definieren wir:

Definition 3.2.1. Sei $\xi \in \mathbb{E}_{\mathbb{K}}$ in einer Form wie im Darstellungssatz, d.h. $\xi = \frac{1}{n}\left(k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}\right)$. Sei weiterhin $(v_1, \ldots, v_d) \in \{0,1\}^d$. Dann heiße

$$\xi^{(v_1,\ldots,v_d)} = k_0 + \sum_{i=1}^{d}(-1)^{v_i} k_i^{(v_1,\ldots,v_{i-1})}\sqrt{w_i^{(v_1,\ldots,v_{i-1})}}$$

ein zu ξ *konjugiertes* Element. Die $k_i^{(v_1,\ldots,v_{i-1})}$ und die $w_i^{(v_1,\ldots,v_{i-1})}$ sind rekursiv ebenso definiert, wobei natürlich die v_i ohne Bedeutung sind, wenn ein Vorfaktor null ist. Wir notieren $\xi \sim \xi^{(v_1,\ldots,v_d)}$ für die Relation, dass ξ und $\xi^{(v_1,\ldots,v_d)}$ zueinander konjugiert sind.

Zwei Elemente heißen also konjugiert, wenn sie durch Permutation der Vorzeichen vor allen Auftreten der Wurzeln auseinander hervorgehen. Im folgenden Lemma stellen wir fest, dass die Konjugierten ihren Namen tatsächlich zu Recht tragen: Es handelt sich bei ihnen um die anderen Nullstellen des Minimalpolynoms von ξ.

Lemma 3.2.2. *Sei $\xi \in \mathbb{E}_{\mathbb{K}}$. Dann ist*

$$p = \prod_{\eta \sim \xi}(X - \eta)$$

das Minimalpolynom von ξ über $\mathbb{K}$.

Beweis. Das Polynom p ist normiert und hat Grad $2^d = \deg(\xi)$, d.h. es genügt zu zeigen, dass alle Koeffizienten von p in $\mathbb{K}$ liegen, dass also keine der Wurzeln mehr in den Koeffizienten vorkommen.

Durch den Darstellungssatz 3.1.1 wissen wir, dass sich ξ als $k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}$ darstellen lässt. Der Linearfaktor $(X - \xi)$ lässt sich also auffassen als ein Polynom in X, dessen Koeffizienten Ausdrücke in den Wurzeln $\sqrt{w_1}, \ldots, \sqrt{w_d}$ sind. Wir zeigen eine Verallgemeinerung der Behauptung: Sei $q(w_1, \ldots, w_d, X) := \eta_n(w_1, \ldots, w_d)X^n + \cdots + \eta_0(w_1, \ldots, w_d)$, wobei alle η_i Ausdrücke über den Wurzeln $w_1, \ldots, w_d$ sind. Die w_i kommen dabei natürlich auch wieder rekursiv in den w_j für $j > i$ vor. Dann gilt allgemein:

$$\prod_{(v_1,\ldots,v_d)\in\{0,1\}^d} q\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_d}\sqrt{w_d^{(v_1,\ldots,v_{d-1})}}, X\right) \in \mathbb{K}[X]$$

Wir zeigen diese Aussage per Induktion über die Anzahl der Wurzeln d. Ist $d = 0$, so ist jedes η_i ein Element aus $\mathbb{K}$, und die Aussage gilt nach Definition. Gelte die Aussage also schon für $d - 1$.

Die Wurzel $\sqrt{w_d}$ kommt in jedem η_i nur einmal vor. Wir zerlegen q in einen Anteil mit und eine Anteil ohne $\sqrt{w_d}$ als Faktor, d.h.

$$q(\sqrt{w_1}, \ldots, \sqrt{w_d}, X) = q_1(\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}, X) + q_2(\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}, X)\sqrt{w_d}$$

Für jedes Tupel $(v_1, ..., v_{d-1}) \in \mathbb{Z}_2^{d-1}$ gilt jetzt:

$$\begin{aligned}
& q\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_{d-1}}\sqrt{w_{d-1}^{(v_1,\ldots,v_{d-2})}}, \sqrt{w_d^{(v_1,\ldots,v_{d-1})}}, X\right) \\
& \cdot q\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_{d-1}}\sqrt{w_{d-1}^{(v_1,\ldots,v_{d-2})}}, -\sqrt{w_d^{(v_1,\ldots,v_{d-1})}}, X\right) \\
= \; & q_1\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_{d-1}}\sqrt{w_{d-1}^{(v_1,\ldots,v_{d-2})}}, X)\right)^2 \\
& - q_2\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_{d-1}}\sqrt{w_{d-1}^{(v_1,\ldots,v_{d-2})}}, X)\right)^2 w_d^{(v_1,\ldots,v_{d-1})}
\end{aligned}$$

Das w_d selbst ist ein Ausdruck in den Wurzeln mit kleineren Index. Insbesondere ist das Produkt dieser beiden Einsetzungen von q also nur noch ein Ausdruck in X und den Wurzeln $w_1, ..., w_{d-1}$. Wir bezeichnen dies Produkt als $q_3(w_1, \ldots, w_{d-1}, X)$. Betrachten wir das Produkt über alle Tupel $(v_1, ..., v_d)$, so bekommen wir 2^{d-1} solcher Paarungen, d.h. wir haben

$$\begin{aligned}
& \prod_{(v_1,\ldots,v_d)\in\{0,1\}^d} q\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_d}\sqrt{w_d^{(v_1,\ldots,v_{d-1})}}, X\right) \\
= \; & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} q_3\left((-1)^{v_1}\sqrt{w_1}, \ldots, (-1)^{v_{d-1}}\sqrt{w_{d-1}^{(v_1,\ldots,v_{d-2})}}, X\right)
\end{aligned}$$

Für diese Produkt gilt aber nach Induktionsvoraussetzung, dass es ein Element von $\mathbb{K}[X]$ ist. Damit ist unsere stärkere Behauptung gezeigt und somit insbesondere auch die Behauptung aus dem Lemma.

□

In diesem Lemma verwenden wir, dass wir den Grad von ξ kennen, um zu zeigen, dass das resultierende Polynome tatsächlich das Minimalpolynom von ξ ist. Die Aussage, dass das oben beschriebene Polynom Koeffizienten aus $\mathbb{K}$ hat, würde aber auch ohne die Gradaussage gelten. Diese Beobachtung kann in einigen Fällen wichtig sein, wenn ein Element mit Quadratwurzeln wie im Darstellungssatz gegeben ist, von dem aber unklar ist, ob der Grad des Elementes 2^d ist oder darunter liegt. Lassen wir die Gradaussage aus dem letzten Lemma einfach weg, dann erhalten wir folgendes Korollar:

Korollar 3.2.3. *Sei ξ wie im Darstellungssatz vom Grad 2^d, aber mit $d' \geq d$ verschiedenen Wurzeln. Seien $\eta_1, \ldots, \eta_{2^{d'}}$ die durch alle Vertauschungen der Vorzeichen vor den Wurzeln aus ξ hervorgehenden Elemente. Dann ist*

$$p = \prod_{i=1}^{2^{d'}} (X - \eta_i) \in \mathbb{K}[X]$$

und somit ein Vielfaches des Minimalpolynoms von ξ über $\mathbb{K}$.

Beweis. Der Beweis, dass $p \in \mathbb{K}[X]$ ist, ist im Beweis des obigen Lemmas enthalten. Damit muss es aber auch ein Vielfaches des Minimalpolynoms sein. □

Wir wollen ein weiteres Korollar zu Lemma 3.2.2 formulieren. Da die Konjugierten von ξ gerade die Nullstellen des Minimalpolynoms sind, ergeben offenbar alle elementarsymmetrischen Polynome, ausgewertet bei den Konjugierten, wieder ein Element des Grundkörpers, und daher natürlich überhaupt alle symmetrischen Polynome:

Korollar 3.2.4. *Sei ξ vom Grad 2^d mit den Konjugierten $\eta_1, \ldots, \eta_{2^d}$. Sei $p \in \mathbb{K}[X_1, \ldots, X_{2^d}]$ symmetrisch. Dann ist $p(\eta_1, \ldots, \eta_{2^d}) \in \mathbb{K}$.*

Beweis. Da das Polynom mit den Nullstellen $\eta_1, \ldots, \eta_{2^d}$ nur Koeffizienten aus $\mathbb{K}$ hat, sind also alle elementarsymmetrischen Polynome über $\eta_1, \ldots, \eta_{2^d}$ und damit alle symmetrischen Polynome in $\mathbb{K}$. □

Dieses Korollar sagt insbesondere, dass das Produkt aller Konjugierten einer Zahl ein Element des Grundkörpers ist. Dies gibt eine weitere Einsicht, warum der Nenner im Darstellungssatz nur eine Zahl aus $\mathbb{K}$ ist und keine Wurzeln enthält, obwohl wir im Euklidischen Körper auch durch einen Wurzelausdruck dividieren können. Denn erhalten wir dadurch einen Bruch mit einem Wurzelausdruck im Nenner, so erweitern wir diesen Bruch einfach um alle Konjugierten des Nenners; dadurch erhalten wir eine Darstellung des Bruchs mit einem Nenner aus $\mathbb{K}$.

3.3 Vollständige Elemente

Wie wir in Beispiel 3.1.2 gesehen haben, liefert der Darstellungssatz im Allgemeinen keine eindeutige Darstellung der Elemente von $\mathbb{E}_{\mathbb{K}}$. Selbst wenn wir alle Vorfaktoren k_i für $i > 0$ wie in Korollar 3.1.3 in die entsprechenden Wurzeln mit hineinziehen, haben wir immer noch Mehrdeutigkeiten. Es gibt aber eine Gruppe von Elementen des Euklidischen Körpers, deren Darstellung in der Tat eindeutig ist und die auch sonst interessante Eigenschaften besitzen, die so genannten *vollständigen* Elemente.

Die Eigenschaft der Vollständigkeit gilt für alle Elemente des Euklidischen Körpers, wenn nicht besondere Zusammenhänge zwischen den Faktoren k_i auftreten. Diese Eigenschaft ist wie folgt definiert:

Definition 3.3.1. Sei $\xi = k_0 + \sum_{i=1}^{d} \sqrt{w_i}$. Es sei

$$\mathbb{K}_i = \mathbb{K}(\sqrt{w'_j} \mid j \leq i, w'_j \sim w_j)$$

der Körper, in dem die Wurzeln aller Konjugierten von w_1 bis w_i adjungiert sind. ξ heißt *vollständig*, wenn $\forall i\ [\mathbb{K}_{i+1} : \mathbb{K}_i] = 2^{2^i}$.

Zu jedem $\sqrt{w_i}$ gibt es 2^i Konjugierte, d.h. $\mathbb{K}_{i+1}$ entsteht aus $\mathbb{K}_i$ durch 2^i Erweiterungen, die alle echte Erweiterungen sein müssen, um insgesamt auf eine Erweiterung vom Grad 2^{2^i} zu kommen. Das Element ist also vollständig in dem Sinn, dass keine Konjugierte keiner der in der Summe vorkommenden Wurzeln sich durch andere Konjugierte derselben oder kleinerer Wurzeln darstellen lässt.

Betrachten wir die Folge von Operationen, aus denen ein Element ξ entsteht. Dann besagt die Vollständigkeit von ξ, dass bei jedem Ziehen einer Quadratwurzel aus w_i die Wurzeln der Konjugierten von w_i voneinander unabhängig sind, dass also jede Wurzel aus einer Konjugierten von w_i eine Körpererweiterung vom Grad zwei ist.

Eine erste Folgerung aus der Vollständigkeit eines Elementes ist, dass die kleineren Wurzeln $\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}$ in w_i wieder vorkommen müssen; denn würde eine Wurzel $\sqrt{w_j}$ fehlen, so wären die beiden Konjugierten von w_i, die sich nur im Vorzeichen vor dieser Wurzel unterscheiden, identisch und könnten daher nicht mehr zwei unabhängige Erweiterungen vom Grad zwei darstellen. Um diese Eigenschaft formal zu fassen, definieren wir erstmal den Begriff *Vorkommen* einer Wurzel präzise:

Definition 3.3.2. Sei $\xi \in \mathbb{E}_{\mathbb{K}}$ in der Darstellung wie im Darstellungssatz und $\sqrt{w_i}$ eine Wurzel dieser Darstellung. Dann *kommt* $\sqrt{w_i}$ *in* ξ *vor*, wenn entweder $k_i \neq 0$ oder wenn $\sqrt{w_i}$ in einer anderen Wurzel $\sqrt{w_j}$ für $j < k$ im Sinne dieser Definition vorkommt und $k_j \neq 0$ ist.

Diese Definition bezieht sich auf eine bestimmte Darstellung eines Elementes. Dass eine Wurzel in einem Ausdruck vorkommt, bedeutet, dass der Vorfaktor vor dieser Wurzel in dem Ausdruck nicht überall 0 ist. Folgendes Lemma beweist, dass in vollständigen Elementen in allen w_j die Wurzeln w_i mit $i < j$ vorkommen:

Lemma 3.3.3. *Sei $\xi = k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}$ wie im Darstellungssatz. Ist ξ vollständig, dann kommt $\sqrt{w_i}$ in jeder Wurzel w_j für $j > i$ vor.*

Beweis. Wir beweisen die Behauptung per Induktion über j. Für $j = 1$ gibt es keine Wurzeln mit kleinerem Index, also gilt die Aussage unmittelbar. Für $j > 1$ zeigen wir, dass w_{j-1} in w_j vorkommt; nach Induktionsvoraussetzung kommen alle Wurzeln mit Indizes kleiner $j - 1$ in w_{j-1} vor, und dann auch in w_j.

Nehmen wir an, dass w_{j-1} nicht in w_j vorkommt. Dann wären die beiden Konjugierten von w_j, die gleich sind bis auf das Vorzeichen vor w_{j-1}, völlig identisch. Aus der Vollständigkeit folgt aber, dass alle Konjugierten aller Wurzeln nicht abhängig sind (da jede Konjugierte jeder Wurzel wieder eine echte Körpererweiterung ist). Dies ist aber ein Widerspruch dazu, dass zwei Konjugierte identisch sind. Also muss w_{j-1} in w_j vorkommen, und der Beweis ist abgeschlossen. □

Vollständigkeit vererbt sich auf die einzelnen w_i, wie wir mit folgendem Lemma feststellen:

Lemma 3.3.4. *Ist $\xi = k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}$ vollständig, dann ist für $1 \leq i \leq d$ auch w_i vollständig und vom Grad 2^{i-1}.*

Beweis. Da nach Lemma 3.3.3 alle $\sqrt{w_j}$ für $j < i$ in w_i vorkommen, sind die zu ξ entsprechend Definition 3.3.1 definierten Zwischenkörper $\mathbb{K}_j$ von ξ und die Zwischenkörper nach dieser Definition zu w_i bis auf den letzten Körper identisch, und für diese Zwischenkörper gilt die Unabhängigkeit nach Voraussetzung, und folglich nach Definition die Vollständigkeit von w_i.

Nach dem Darstellungssatz ist 2^{i-1} eine obere Schranke für den Grad von w_i; diese muss hier scharf sein, sonst könnten nicht alle 2^{i-1} Konjugierte von w_i verschieden sein. □

In [5] stellen Cox, Little und O'Shea ein Konzept von *generischen Eigenschaften* eines Polynoms dar. Dabei nennen sie eine Eigenschaft eines Polynoms generisch, wenn sie immer dann gilt, wenn nicht ein polynomieller Zusammenhang zwischen den Koeffizienten besteht. So hat z.B. ein Polynom vom Grad zwei generisch zwei unterschiedliche Nullstellen (nämlich immer dann, wenn der Ausdruck in der Wurzel der pq-Formel nicht null ergibt). Dass eine Eigenschaft generisch ist, ist ein formaler Ausdruck dafür, dass die Eigenschaft in 'fast allen Fällen' gilt.

Wir wollen parallel zu dieser Definition zeigen, dass ein Element aus $\mathbb{E}_{\mathbb{Q}}$ generisch vollständig ist, dass also 'fast alle' Elemente dieses Körpers vollständig sind. In unserem Fall sind nicht polynomielle Zusammenhänge zwischen den Elementen ausschlaggebend, sondern ob bestimmte Ausdrücke Quadrate anderer Ausdrücke sind. Wir wollen generisch hier also so verstehen, dass ein Element vollständig ist, wenn kein weiterer Zusammenhang zwischen den auftretenden k_i explizit gegeben ist. Wir betrachten also einen Quadratwurzelausdruck, in dem alle in der Darstellung des Darstellungssatzes vorkommenden Elemente des Grundkörpers symbolische Variablen sind. Tatsächlich sind solche generischen Elemente vollständig, wie folgendes Lemma zeigt:

Lemma 3.3.5. *Sei $\xi = k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}$ wie im Darstellungssatz, wobei alle in der Darstellung vorkommenden Elemente des Grundkörpers symbolische Variablen X_i sind. ξ ist dann also ein Element aus $\mathbb{E}_{\mathbb{K}}$ mit $\mathbb{K} = \mathbb{Q}(X_1, \ldots, X_N)$ für ein passendes N. Dann ist ξ vollständig.*

Beweis. Nehmen wir an, es gibt eine Stelle i in der Körperkette der $\mathbb{K}_i$ aus Definition 3.3.1, so dass $[\mathbb{K}_{i+1} : \mathbb{K}_i] < 2^{2^i}$, was die Definition der Vollständigkeit verletzen würde. Dann zerlegen wir diese Körpererweiterungen durch schrittweises Hinzufügen der Konjugierten von $\sqrt{w_{i+1}}$; eine dieser Erweiterungen muss vom Grad 1 sein (da die gesamte Kette der Erweiterungen sonst 2^i lang wäre). Hier ist also die nächste zu adjungierende Wurzel bereits im Körper enthalten, d.h. durch Körperoperationen über die anderen Wurzeln darstellbar. Wir betrachten die Kette dieser Operationen als einen Bruch, erweitern mit den anderen Konjugierten im Nenner und ziehen den so entstehenden wurzelfreien Nenner in die einzelnen Wurzeln, so dass wir einen Wurzelausdruck in Brüchen über $\mathbb{Q}[X_1, \ldots, X_N]$ erhalten. Jeder dieser Brüche sei in diesem Polynomring weitestmöglich gekürzt.

Es genügt zu zeigen, dass es überhaupt ein vollständiges Element aus $\mathbb{E}_{\mathbb{Q}}$ gibt. Denn dann muss die Einsetzung der zu diesem Element gehörenden Elemente des Grundkörpers in den eben beschriebenen symbolischen Ausdruck ebenfalls eine weitere Konjugierte erzeugen, was ein Widerspruch zur Vollständigkeit dieses Elementes darstellen würde.

Es genügt also zu zeigen, dass es über $\mathbb{E}_{\mathbb{Q}}$ überhaupt ein vollständiges Element gibt. Wir werden in Kapitel 4 sehen, dass es eine endliche und durch Normalteiler auflösbare Gruppe Q_d gibt, so dass jedes Element, dessen Zerfällungskörper Q_d als Galoisgruppe hat, vollständig ist. In [23] gibt Shafarevitsch einen Satz an, nach dem es zu jeder durch Normalteiler auflösbaren endlichen Gruppe eine Körpererweiterung über $\mathbb{Q}$ gibt, die diese Gruppe zur Galoisgruppe hat. Der in [23] geführte Beweis des Satzes war noch fehlerhaft, Shafarevitsch lieferte aber wenige Jahre später einen korrigierten Beweis nach. Für einen modernen Beweis sei der Leser auf [17] verwiesen.

Es gibt demnach vollständige Elemente vom Grad 2^d, womit das Lemma bewiesen ist. □

Durch dieses Lemma wird klar, dass Vollständigkeit der 'Normalfall' ist, während nicht vollständige Elemente nur bei Erfüllung bestimmter Zusatzbedingungen auftreten können. Viele der in dieser Arbeit gezeigten Eigenschaften von Elementen beruhen auf der Vollständigkeit.

Wir wollen aber auch für nicht vollständige Elemente möglichst viele Aussagen über die Darstellung im Darstellungssatz machen. Insbesondere kann man zeigen, dass es zu jedem Element mindestens eine Darstellung gibt, für die die Aussagen von Lemma 3.3.3 gelten, und insbesondere auch die Aussage, dass $\deg(w_i) = 2^{i-1}$ ist.

Lemma 3.3.6. *Sei $\mathbb{K}$ Quotientenkörper über einem Ring $\mathbf{R}$ und $\xi \in \mathbb{E}_{\mathbb{K}}$. Dann gibt es eine Darstellung der Form $\xi = \frac{1}{n}\left(k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}\right)$ wie im Darstellungssatz, in der für jedes w_i und für $j < i$ w_j in w_i vorkommt und $\deg(w_i) = 2^{i-1}$ ist.*

Beweis. Wir beweisen die Aussage durch Induktion über d. Für $d = 0$ ist die Aussage klar. Gelte sie also schon für alle $d' < d$.

Nehmen wir zunächst irgendeine Darstellung von ξ. Wir wollen diese Darstellung zunächst so umformen, dass $\deg(w_i) \leq \deg(w_{i+1})$ gilt.

Falls für irgendein i in dieser Darstellung $\deg(w_i) > \deg(w_{i+1})$ sein sollte, so kommt w_i in w_{i+1} nicht vor. Also können wir w_i und w_{i+1} vertauschen und umbenennen, ohne die vom Darstellungssatz geforderte Eigenschaft zu verletzen, dass jedes $w_i \in \mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}]$ ist.

Gehen wir also im Folgenden ohne Beschränkung der Allgemeinheit davon aus, dass $\deg(w_i) \leq \deg(w_{i+1})$ ist. Im Falle der Gleichheit gilt dieselbe Begründung wie eben, und wir wissen, dass w_i und w_{i+1} beide Elemente des Ringes $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}]$ sind. Wir betrachten die Summe der beiden, also $\sqrt{w_i} + \sqrt{w_{i+1}}$. Quadrieren wir diesen Term und ziehen aus dem Quadrat wieder die Wurzel, so erhalten wir $\sqrt{w_i} + \sqrt{w_{i+1}} = \sqrt{w_i + w_{i+1} + 2\sqrt{w_i w_{i+1}}}$.

Da $\sqrt{w_{i+1}}$ nach Forderung des Darstellungssatzes nicht Element des Ringes $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}, \sqrt{w_i}]$ ist, liegt auch das Produkt $\sqrt{w_i w_{i+1}}$ nicht in dem Ring $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}]$. Also ist $\sqrt{w_i w_{i+1}}$ eine echte Erweiterung dieses Ringes; wir definieren $\widetilde{w_i} = w_i w_{i+1}$ und $\widetilde{w_{i+1}} = w_i + w_{i+1} + 2\sqrt{\widetilde{w_i}}$. Da $\sqrt{\widetilde{w_i}}$ in $\widetilde{w_{i+1}}$ vorkommt, ist $\sqrt{\widetilde{w_{i+1}}}$ wiederum eine echte Erweiterung des Ringes $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}, \sqrt{\widetilde{w_i}}]$.

Wir haben also aus zwei Elementen vom Grad $\deg(w_i)$ ein Element vom Grad $\deg(w_i)$ gemacht und ein weiteres vom Grad $2\deg(w_i)$. Es bleibt zu zeigen, dass wir jedes spätere Auftreten von w_i und w_{i+1} auch durch die neu definierten Wurzeln ausdrücken können. Wir stellen fest, dass die beiden Ringe $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}, \sqrt{w_i}, \sqrt{w_{i+1}}]$ und $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}, \sqrt{\widetilde{w_i}}, \sqrt{\widetilde{w_{i+1}}}]$ nicht identisch sind; wir können aber ein Vielfaches von $\sqrt{w_i}$ mit den Elementen des neuen Ringes darstellen, und zwar wie folgt:

$$\begin{aligned}
\sqrt{\widetilde{w_{i+1}}}\left(\sqrt{\widetilde{w_i}} - w_i\right) &= \left(\sqrt{w_i} + \sqrt{w_{i+1}}\right)\left(\sqrt{\widetilde{w_i}} - w_i\right) \\
&= \left(\sqrt{w_i} + \sqrt{w_{i+1}}\right)\left(\sqrt{w_i w_{i+1}} - w_i\right) \\
&= \left(w_{i+1} - w_i\right)\sqrt{w_i}
\end{aligned}$$

Wir können also $\sqrt{w_i}$ durch den Quotienten $\frac{\sqrt{\widetilde{w_{i+1}}}\left(\sqrt{\widetilde{w_i}} - w_i\right)}{(w_{i+1} - w_i)}$ darstellen, in dem nur noch Elemente aus dem Ring $\mathbf{R}[\sqrt{w_1}, \ldots, \sqrt{w_{i-1}}, \sqrt{\widetilde{w_i}}, \sqrt{\widetilde{w_{i+1}}}]$ auftauchen. Erweitern wir mit allen Konjugierten des Nenners, so bleibt im Nenner nur eine rationale Zahl zurück, die wir aus den jeweiligen Wurzeln rausziehen und in den Nenner des Elementes mit eingliedern können. Die Wurzel $\sqrt{w_{i+1}}$ können wir dann einfach als $\sqrt{w_{i+1}} = \sqrt{\widetilde{w_{i+1}}} - \sqrt{w_i}$ darstellen.

Wir haben damit die Aussage über den Grad der Elemente gezeigt; die Aussage über das Vorkommen der kleineren Wurzeln in den größeren folgt dann unmittelbar, da sonst der jeweilige Grad nicht erreicht werden könnte. □

Kapitel 4

Galoistheorie für Quadratischen Erweiterungen

Im letzten Kapitel sind wir auf den Darstellungssatz und die Vollständigkeit von Elementen eingegangen; dabei haben wir in Lemma 3.3.5 bereits verwendet, dass es eine durch Normalteiler der Größe zwei auflösbare Galoisgruppe der Größe 2^{2^d-1} gibt. In diesem Kapitel werden wir diese Gruppe, die so genannte *Quadratvertauschungsgruppe*, einführen und einige ihrer Eigenschaften zeigen. Dadurch wird es uns insbesondere gelingen, die Eindeutigkeit der Darstellung für vollständige Elemente zu zeigen.

4.1 Quadratfixpolynome

Wir haben schon festgestellt, dass jedes symmetrische Polynom über den Konjugierten eines Elementes aus $\mathbb{E}_\mathbb{K}$ nach $\mathbb{K}$ zurückfällt. Im Allgemeinen ist die Galoisgruppe des Zerfällungskörpers des Minimalpolynoms aber kleiner als die gesamte Permutationsgruppe. Es gibt also noch andere, nicht symmetrische Polynome über den Konjugierten, die ebenfalls nach $\mathbb{K}$ zurückfallen.

Um diese zu finden, stellen wir zunächst folgendes einfache Lemma fest:

Lemma 4.1.1. *Sei $\xi \in \mathbb{E}_\mathbb{K}$. Sei $(v_1, \ldots, v_{d-1}) \in \{0,1\}^{d-1}$. Dann gilt:*

$$\xi^{(v_1,..,v_{d-1},0)} + \xi^{(v_1,...,v_{d-1},1)} \in \mathbb{K}[w_1, \ldots, w_{d-1}]$$

und

$$\xi^{(v_1,..,v_{d-1},0)} \cdot \xi^{(v_1,...,v_{d-1},1)} \in \mathbb{K}[w_1, \ldots, w_{d-1}]$$

Beweis. Sei $\xi = \xi_1 + \sqrt{w_d}$ mit $\xi_1, w_d \in \mathbb{K}[w_1, \ldots, w_{d-1}]$. Dann ist

$$\begin{aligned} \xi^{(v_1,..,v_{d-1},0)} + \xi^{(v_1,...,v_{d-1},1)} &= \xi_1^{(v_1,...,v_{d-1})} + \sqrt{w_d^{(v_1,...,v_{d-1})}} \\ &\quad + \xi_1^{(v_1,...,v_{d-1})} - \sqrt{w_d^{(v_1,...,v_{d-1})}} \\ &= 2\xi_1^{(v_1,...,v_{d-1})} \end{aligned}$$

und

$$\begin{aligned}\xi^{(v_1,..,v_{d-1},0)} \cdot \xi^{(v_1,...,v_{d-1},1)} &= \left(\xi_1^{(v_1,...,v_{d-1})} + \sqrt{w_d^{(v_1,...,v_{d-1})}}\right) \\ &\quad \cdot \left(\xi_1^{(v_1,...,v_{d-1})} - \sqrt{w_d^{(v_1,...,v_{d-1})}}\right) \\ &= \left(\xi_1^{(v_1,...,v_{d-1})}\right)^2 - w_d^{(v_1,...,v_{d-1})}\end{aligned}$$

□

Mit diesem Lemma können wir einige Polynome ermitteln, die, ausgewertet an den Konjugierten eines Elementes ξ, wieder in den Grundkörper zurückfallen. Wir nennen diese Polynome die *Quadratfixpolynome* eines Elementes ξ. Wir definieren die Quadratfixpolynome rekursiv:

Definition 4.1.2. $p(X) = X$ ist ein Quadratfixpolynom, und ist schon das Polynom $p \in \mathbb{K}[X_1, \ldots, X_n]$ ein Quadratfixpolynom, dann ist für $\circ \in \{\cdot, +\}$ auch $q(X_1, \ldots, X_{2n}) = p(X_1, \ldots, X_n) \circ p(X_{n+1}, \ldots, X_{2n})$ ein Quadratfixpolynom.

Beispiel 4.1.3. Die Quadratfixpolynome sind also X_1, $X_1 + X_2$, X_1X_2, $X_1 + X_2 + X_3 + X_4$, $(X_1 + X_2)(X_3 + X_4)$, $X_1X_2 + X_3X_4$, $X_1X_2X_3X_4$, usw.

Wie erwartet ergeben die Quadratfixpolynome ausgewertet bei den Konjugierten eines Elements von $\mathbb{E}_\mathbb{K}$ wieder ein Körperelement:

Lemma 4.1.4. *Sei $p \in \mathbb{K}[X_1, \ldots, X_{2^d}]$ ein Quadratfixpolynom und $\xi \in \mathbb{E}_\mathbb{K}$ vom Grad 2^d. Dann ist $p\left(\xi^{(0,...,0,0,0)}, \xi^{(0,...,0,0,1)}, \xi^{(0,...,0,1,0)}, \xi^{(0,...,0,1,1)}, \ldots, \xi^{(1,...,1)}\right) \in \mathbb{K}$. Die Konjugierten werden dabei in lexikographischer Ordnung aufgezählt.*

Beweis. Sei $p_d, p_{d-1}, \ldots, p_1,\ p_0 = p$ die Folge von Polynomen, aus denen p nach der Definition der Quadratfixpolynomen hervorgegangen ist, also jeweils $p_{i-1}(X_1, \ldots, X_{2^{d-i+1}}) = p_i(X_1, \ldots, X_{2^{d-i}}) \circ p_i(X_{2^{d-i}+1}, \ldots, X_{2^{d-i+1}})$ mit einer Operation $\circ \in \{\cdot, +\}$. Wir beweisen induktiv, dass alle in der Konstruktion auftretenden p_i genau die Konjugierten eines Elementes aus $\mathbb{K}[w_1, \ldots, w_i]$ in lexikographischer Ordnung berechnen. p_d ist die Identität und taucht für jede Konjugierte von ξ in der richtigen Reihenfolge genau einmal auf, d.h. der Induktionsanfang ist erfüllt. Gelte die Aussage für p_i. Dann ist $p_{i-1} = \eta_1 \circ \eta_2$, worin η_1 und η_2 zwei aufeinanderfolgende Konjugierte (beginnend mit der ersten und zweiten, dritten und vierten und so weiter) von einem $\eta \in \mathbb{K}[w_1, \ldots, w_i]$ sind. η_1 und η_2 sind also $\eta^{(v_1,...,v_{i-1},0)}$ und $\eta^{(v_1,...,v_{i-1},1)}$ für ein festes Tupel $(v_1, \ldots, v_{i-1})$. Nach Lemma 4.1.1 ist p_{i-1} ein Element aus $\mathbb{K}[w_1, \ldots, w_{i-1}]$, und wenn wir bei jedem Auftreten von p_{i-1} dieselbe Anordnung der Wurzeln zugrunde legen, ist p_{i-1} jeweils die Konjugierte mit der Zählung $(v_1, \ldots, v_{i-1})$, d.h. die Ordnung bleibt erhalten.

Für p_0 gilt die Aussage schließlich auch, und also ist $p_0 \in \mathbb{K}$, womit die Aussage bewiesen ist.

□

Betrachten wir die Galoisgruppe des Zerfällungskörpers von ξ, ergibt sich als Korollar

Korollar 4.1.5. *Die Galoisgruppe von* $\mathbb{K}[\xi^{(0,\ldots,0)},\ldots,\xi^{(1,\ldots,1)}]$ *lässt alle Quadratfixpolynome fest.*

Wie dieser Abschnitt zeigt, sind die Quadratfixpolynome Fixpolynome von jeder Körpererweiterung, die durch eine Folge von Quadratwurzelerweiterungen entsteht. Sie nehmen also für Erweiterungen vom Grad zwei eine ähnliche Rolle ein wie die symmetrischen Polynome bei allgemeinen galoischen Erweiterungen. Im folgenden Abschnitt nutzen wir die Quadratfixpolynome als ersten Zugang zur Galoisgruppe des Zerfällungskörpers eines ξ.

4.2 Die Quadratvertauschungsgruppe

Die Quadratfixpolynome sind nicht notwendigerweise alle Fixpolynome der Galoisgruppe des Zerfällungskörpers von ξ. In den Fällen aber, in denen dies der Fall ist, können wir die Galoisgruppe exakt feststellen. Wie wir zeigen werden, handelt es sich um folgende Untergruppe der Permutationsgruppe S_{2^d}:

Definition 4.2.1. Die Quadratvertauschungsgruppe Q_d sei eine Untergruppe der Permutationsgruppe S_{2^d}, die rekursiv wie folgt definiert wird:

$$
\begin{aligned}
Q_0 &= \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \\
Q_1 &= \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} \\
Q_{d+1} &= \left\{ \begin{pmatrix} 1 & \ldots & 2^{d+1} \\ 1 & \ldots & 2^{d+1} \end{pmatrix}, \begin{pmatrix} 1 & \ldots & 2^d & 2^d+1 & \ldots & 2^{d+1} \\ 2^d+1 & \ldots & 2^{d+1} & 1 & \ldots & 2^d \end{pmatrix} \right\} \bullet \\
&\quad \left\{ \begin{pmatrix} 1 & \ldots & 2^d & 2^d+1 & \ldots & 2^{d+1} \\ s_1 & \ldots & s_{2^d} & 2^d+1 & \ldots & 2^{d+1} \end{pmatrix} \mid \begin{pmatrix} 1 & \ldots & 2^d \\ s_1 & \ldots & s_{2^d} \end{pmatrix} \in Q_d \right\} \bullet \\
&\quad \left\{ \begin{pmatrix} 1 & \ldots & 2^d & 2^d+1 & \ldots & 2^{d+1} \\ 1 & \ldots & 2^d & s_1+2^d & \ldots & s_{2^d}+2^d \end{pmatrix} \mid \begin{pmatrix} 1 & \ldots & 2^d \\ s_1 & \ldots & s_{2^d} \end{pmatrix} \in Q_d \right\}
\end{aligned}
$$

Die Quadratvertauschungsgruppe Q_{d+1} verhält sich also jeweils auf der ersten und zweiten Hälfte der Indizes wie Q_d, und sie vertauscht die beiden Seiten miteinander. Die Quadratvertauschungsgruppe lässt die Quadratfixpolynome unverändert:

Lemma 4.2.2. *Sei p ein Quadratfixpolynom in 2^d Unbekannten und $\sigma \in Q_d$. Dann ist $p(X_1,\ldots,X_{2^d}) = p(X_{\sigma(1)},\ldots,X_{\sigma(2^d)})$.*

Beweis. Die Aussage gilt offenbar für $d = 0$. Gelte die Aussage schon für $d-1$. Wir schreiben das Quadratfixpolynom in 2^d Unbekannten als $p_d(X_1,..,X_{2^d}) = p_{d-1}(X_1,\ldots,X_{2^{d-1}}) \circ p_{d-1}(X_{2^{d-1}+1},\ldots,X_{2^d})$ mit $\circ \in \{+,\cdot\}$. Wir zeigen die Aussage für die Elemente jeder der drei erzeugenden Mengen der Rekursion aus Definition 4.2.1. Das Vertauschungselement der ersten Menge vertauscht die beiden p_{d-1} gegeneinander, was wegen der Kommutativität von $\circ$ den Wert des Polynoms nicht ändert. Elemente der zweiten und dritten erzeugenden Menge

führen innerhalb der jeweiligen p_{d-1} Vertauschungen durch, die nach Induktionsvoraussetzung den Wert unverändert lassen. Also bleibt p_d auch unter allen Kombinationen dieser Vertauschungen invariant, und die Behauptung ist gezeigt. □

Die Anzahl der Element von Q_d ist eine maximale Zweierpotenz in der Anzahl der Elemente von S_{2^d}, d.h. Q_d ist eine 2-Sylow-Untergruppe von S_{2^d}:

Lemma 4.2.3. *Q_d ist eine 2-Sylow-Untergruppe von S_{2^d}.*

Beweis. S_{2^d} hat $2^d!$ viele Elemente. In der Primfaktorzerlegung von $2^d!$ kommt der Faktor 2 genau $2^d - 1$ mal vor, d.h. wir müssen zeigen, dass die Ordnung von Q_d genau 2^{2^d-1} ist.

Dies gilt offenbar für Q_1; gelte die Aussage also für $d-1$. Die Größe von $|Q_d|$ ergibt sich als Produkt der Größen der drei erzeugenden Mengen aus Definition 4.2.1, da die die drei Gruppen bis auf das neutrale Element disjunkt sind und ihre Elemente gegeneinander vertauscht werden können. Also ist $|Q_d| = 2|Q_{d-1}|^2 = 2 \cdot \left(2^{2^{d-1}-1}\right)^2 = 2 \cdot 2^{2^d-2} = 2^{2^d-1}$. □

Sind die Quadratfixpolynome die einzigen Fixpolynome des Zerfällungskörpers von ξ, dann ist die Quadratvertauschungsgruppe die Galoisgruppe des Zerfällungskörpers von ξ. Dies gilt, da die Galoisgruppe eine Zweierpotenz sein muss, andererseits aber alle Elemente von Q_d enthalten muss (da ja alle Fixpolynome davon fest gelassen werden) und es keine größere Gruppe der Ordnung einer Zweierpotenz in S_{2^d} gibt.

Da alle 2-Sylow-Untergruppen von S_{2^d} durch innere Automorphismen auseinander hervorgehen, wir aber nach der Definition der Vollständigkeit wissen, dass vollständige Elemente einen Grad von 2^{2^d-1} haben, wissen wir bereits, dass die Galoisgruppe vollständiger Elemente die Struktur von Q_d hat. Ebenso wissen wir, dass Q_d als 2-Gruppe durch Normalteiler auflösbar ist (siehe z.B. in [19]), was wir für den Beweis von Lemma 3.3.5 vorausgesetzt haben. Im folgenden Abschnitt werden wir beweisen, dass bei lexikographischer Aufzählung der Konjugierten die Galoisgruppe vollständiger Elemente sogar genau Q_d ist.

Zunächst aber zeigen wir einige weitere allgemeine Eigenschaften von Q_d. Als Galoisgruppe eines Elementes aus $\mathbb{E}_\mathbb{K}$ muss Q_d einen Normalteiler der Ordnung 2 haben. Wir werden feststellen, dass dieser Normalteiler eindeutig bestimmt ist und aus dem neutralen Element und dem *normalen Element* besteht, welches wir wie folgt definieren:

Definition 4.2.4. Das normale Element σ_1 von Q_1 ist $\sigma_1 = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$, und ist $\sigma_{d-1} = \left(\begin{smallmatrix} 1 & \dots & 2^{d-1} \\ s_1 & \dots & s_{2^{d-1}} \end{smallmatrix}\right)$, dann ist

$$\sigma_d = \left(\begin{smallmatrix} 1 & \dots & 2^{d-1} & 2^{d-1}+1 & \dots & 2^d \\ s_1 & \dots & s_{2^{d-1}} & s_1 + 2^{d-1} & \dots & s_{2^{d-1}} + 2^{d-1} \end{smallmatrix}\right)$$

Das normale Element operiert also auf der ersten Hälfte der Indizes wie das normale Element der darunterliegenden Quadratvertauschungsgruppe, und verhält sich auf der zweiten Hälfte parallel wie auf der ersten. Es ist also

$$\begin{aligned}\sigma_2 &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 1 & 4 & 3 & 5 & 4 & 8 & 7 \end{pmatrix} \\ &\vdots\end{aligned}$$

Wir stellen fest, dass dieses Element tatsächlich invariant ist unter allen inneren Automorphismen von Q_d, und dass das Element darüber hinaus das einzige nicht triviale Element mit dieser Eigenschaft ist (d.h. das einzige außer dem neutralen Element):

Lemma 4.2.5. *σ_d ist das einzige nicht triviale Element von Q_d, das invariant unter allen inneren Automorphismen von Q_d ist.*

Beweis. Wir müssen drei Dinge zeigen; zunächst, dass σ_d in Q_d liegt, zweitens, dass σ_d invariant ist unter allen inneren Automorphismen von Q_d und drittens, dass es das einzige Element mit dieser Eigenschaft ist. Wir zeigen alle drei Behauptungen per Induktion über d. Das Element σ_1 liegt in Q_1, es ist invariant unter den beiden existierenden inneren Automorphismen und ist das einzige nicht triviale Element von Q_1.

Gelten die Aussagen schon für Q_{d-1}. Wir bezeichnen für jedes Element $\upsilon \in Q_{d-1}$ mit $\upsilon^{(1)} \in Q_d$ das Element, das wie υ auf der ersten Hälfte der Elemente operiert und die zweite Hälfte der Elemente fest lässt, und $\upsilon^{(2)} \in Q_d$ das entsprechende für die zweite Hälfte. Mit $\tau \in Q_d$ bezeichnen wir das Element, dass die rechte Hälfte der Elemente mit der linken vertauscht.

Nach Induktionsvoraussetzung ist $\sigma_{d-1} \in Q_{d-1}$; dann ist aber $\sigma_d \in Q_d$, denn $\sigma_d = \sigma_{d-1}^{(1)}\sigma_{d-1}^{(2)}$.

Um zu zeigen, dass σ_d invariant ist unter allen inneren Automorphismen, genügt es, dies für τ und für die Elemente $\upsilon^{(1)}$ bzw. $\upsilon^{(2)}$ für jedes $\upsilon \in Q_{d-1}$ zu zeigen. Dass dies ausreicht, sieht man wie folgt: Q_d wird von diesen Elementen erzeugt, und gilt für zwei Elemente a und b $a\sigma_d a^{-1} = b\sigma_d b^{-1} = \sigma_d$, so gilt es offenbar auch für ab, da $ab\sigma_d(ab)^{-1} = a\sigma_d a^{-1} = \sigma_d$.

Für $\upsilon \in Q_{d-1}$ und $i \in \{1, 2\}$ ist tatsächlich $\upsilon^{(i)}\sigma_d \left(\upsilon^{(i)}\right)^{-1} = \sigma_d$, denn nach Induktionsvoraussetzung gilt dies für die durch das i festgelegte Seite, und die andere Seite wird durch das $\upsilon^{(i)}$ bzw. dessen Inverse nicht beeinflusst. Außerdem gilt offenbar $\tau\sigma_d\tau^{-1} = \sigma_d$, da σ_d auf beiden Hälften exakt die gleiche Operation durchführt.

Es bleibt zu zeigen, dass σ_d das einzige Element mit der beschriebenen Eigenschaft ist. Angenommen, es gäbe ein weiteres Element ϕ, das invariant unter allen inneren Automorphismen ist. Nehmen wir o.B.d.A. an, ϕ sei auf der linken Hälfte der Elemente nicht wie σ_d, d.h. es existiert ein Index $i < 2^{d-1}$ mit $\phi(i) \neq \sigma_d(i)$ (Sind ϕ und σ_d auf der linken Seite gleich, dann müssen sie auf der rechten Seite verschieden sein, und die folgenden Begründungen gelten unter Vertauschung der Begriffe 'links' und 'rechts'). Da ϕ invariant ist unter allen inneren Automorphismen, ist es insbesondere invariant unter allen $\{\upsilon^{(1)} | \upsilon \in Q_{d-1}\}$. Da alle $\upsilon^{(1)}$ die rechte Seite konstant lassen, bedeutet dies, dass die linke Seite von ϕ invariant ist unter allen inneren Automorphismen von Q_{d-1}. Da aber

die linke Seite von ϕ nach Konstruktion von Q_d ein Element von Q_{d-1} ist, ist dies ein Widerspruch zur Induktionsvoraussetzung, nach der σ_{d-1} das einzige Element von Q_{d-1} ist, das invariant unter allen inneren Automorphismen ist. □

4.3 Der Eindeutigkeitssatz

In diesem Abschnitt werden wir eine weitere und bedeutende Eigenschaft von vollständigen Elementen zeigen, nämlich dass der Darstellungssatz für vollständige Elemente eine eindeutige Darstellung liefert, wenn wir die Vorfaktoren k_i jeweils mit in die Wurzeln w_i hineinziehen. Dafür wollen wir zunächst zeigen, dass Q_d bei lexikographischer Ordnung der Konjugierten die Galoisgruppe eines vollständigen Elementes ξ ist.

Dafür interessieren wir uns für die Elemente der Galoisgruppe, die ξ selbst fix lassen, d.h. für die Galoisuntergruppe, die zum kleinsten Zwischenkörper zwischen $\mathbb{E}_\mathbb{K}$ und dem Zerfällungskörper des Minimalpolynoms von ξ gehört. Zur Definition dieser Untergruppe führen wir zunächst noch eine Notation ein.

Definition 4.3.1. Seien die Permutationen $\sigma_1 = \left(\begin{smallmatrix} 1 & \cdots & n_1 \\ s_{1,1} & \cdots & s_{1,n_1} \end{smallmatrix} \right) \in S_{n_1}$ und $\sigma_2 = \left(\begin{smallmatrix} 1 & \cdots & n_2 \\ s_{2,1} & \cdots & s_{2,n_2} \end{smallmatrix} \right) \in S_{n_2}$, dann sei $\sigma_1 \odot \sigma_2 \in S_{n_1+n_2}$ mit

$$\sigma_1 \odot \sigma_2 = \left(\begin{smallmatrix} 1 & \cdots & n_1 & n_1+1 & \cdots & n_1+n_2 \\ s_{1,1} & \cdots & s_{1,n_1} & s_{2,1}+n_1 & \cdots & s_{2,n_2}+n_1 \end{smallmatrix} \right)$$

Entsprechend notieren wir $G \odot H = \{g \odot h | g \in G, h \in H\}$. $\odot$ liefert also das Kreuzprodukt der beiden Gruppen und den kanonischen Isomorphismus in die Permutationsgruppe mit passend vielen Elementen.

Mit dieser Notation können wir jetzt die Gruppe G_d definieren, von der wir später zeigen werden, dass es sich um die mit $\mathbb{K}(\xi)$ korrespondierende Galoisuntergruppe handelt:

Definition 4.3.2. Die ξ-Untergruppe G_d von Q_d sei

$$G_d = Q_0 \odot Q_0 \odot Q_1 \odot Q_2 \odot \cdots \odot Q_{d-1}$$

G_d lässt also die ersten beiden Elemente fest, operiert auf den nächsten beiden wie Q_1, auf den folgenden vier wie Q_2 und so weiter. Wir stellen einige Eigenschaften von G_d fest:

Lemma 4.3.3. *G_d ist Untergruppe von Q_d und enthält genau die Elemente σ, für die $\sigma(1) = 1$ gilt. $|G_d| = 2^{2^d-d-1}$, also $\frac{|Q_d|}{|G_d|} = 2^d$. Ist Q_d die Galoisgruppe des Zerfällungskörpers von ξ, dann ist der zu G_d gehörende Zwischenkörper genau $\mathbb{K}(\xi)$, und er ist der einzige Zwischenkörper vom Grad 2^d über dem Grundkörper, der ξ enthält.*

Beweis. Der erste Teil der Aussage ist trivial. Die Mächtigkeit von G_d ergibt sich aus dem Produkt der Mächtigkeiten der Q_i, also $|Q_d| = |Q_0| \cdot |Q_0| \cdot |Q_1| \cdot \cdots \cdot |Q_{d-1}|$. Da $|Q_i| = 2^{2^i-1}$ ist, gilt

$$|G_d| = 2^{2^1-1} \cdot \cdots \cdot 2^{2^{d-1}-1} = 2^{2^{d-1}+\cdots+2^1-(d-1)} = 2^{2^d-d-1}$$

Entsprechend ist dann $\frac{|Q_d|}{|G_d|} = 2^{2^d-1-(2^d-d-1)} = 2^d$.

Die Untergruppe jedes Zwischenkörpers, der ξ enthält, muss ξ fest lassen, d.h. für alle Element σ dieser Gruppe muss $\sigma(1) = 1$ gelten. Also muss diese Untergruppe ihrerseits Untergruppe von G_d sein; außer G_d selbst haben aber alle Untergruppen von G_d kleinere Ordnung, womit die Eindeutigkeitsaussage gezeigt wäre. Da ξ nach Voraussetzung selbst schon ein Element vom Grad 2^d ist, muss dieser Zwischenkörper $\mathbb{K}(\xi)$ sein.

□

Der zu G_d gehörende Zwischenkörper $\mathbb{K}(\xi)$ enthält außer ξ auch alle $\sqrt{w_i}$ der Darstellung von ξ. Unmittelbar ist dies nicht klar, da ξ ja nur eine Summe von gewichteten $\sqrt{w_i}$ ist. Folgendes Lemma beweist diesen Umstand:

Lemma 4.3.4. *Sei $\xi \in \mathbb{E}_\mathbb{K}$ in der durch den Darstellungssatz beschriebenen Form. Dann gilt für alle w_i: $\sqrt{w_i} \in \mathbb{K}(\xi)$.*

Beweis. Wir nehmen wegen Lemma 3.3.6 o.B.d.A. an, dass $\deg(w_d) = 2^{d-1}$. Wir beweisen durch Induktion über d. Offenbar gilt die Aussage für $d = 1$.

Gelte die Aussage schon für $d-1$. Schreiben wir jetzt für ein ξ vom Grad d $\xi = \xi_1 + \sqrt{w_d}$, so sind darin ξ_1 und w_d jeweils Elemente vom Grad höchstens $d-1$, genau genommen $\deg(\xi_1) \leq d-1$ und $\deg(w_d) = d-1$ mit denselben Wurzeln $\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}$. Wir können also die Induktionsvoraussetzung anwenden und ξ_1 und w_d durch die enthaltenen Wurzeln darstellen, wodurch wir $\mathbb{K}(\xi_1) \subseteq \mathbb{K}(\sqrt{w_1}, \ldots, \sqrt{w_{d-1}}) = \mathbb{K}(w_d)$ erhalten. Es ist also $\mathbb{K}(\xi) \subseteq \mathbb{K}(\sqrt{w_d})$, da offenbar $\mathbb{K}(w_d) \subseteq \mathbb{K}(\sqrt{w_d})$ ist, also auch $\mathbb{K}(\xi_1) \subseteq \mathbb{K}(w_d) \subseteq \mathbb{K}(\sqrt{w_d})$, und ξ die Summe von ξ_1 und $\sqrt{w_d}$ ist.

Der Körper $\mathbb{K}(\xi)$ lässt sich als Vektorraum der Dimension 2^d über $\mathbb{K}$ auffassen. $\sqrt{w_d}$ ist aber ebenfalls ein Element vom Grad 2^d, d.h. $\mathbb{K}(\sqrt{w_d})$ ist ebenfalls ein Vektorraum der Dimension 2^d. Zwei Vektorräume, die die gleiche Dimension haben, und von denen der eine den anderen umschließt, müssen aber bereits identisch sein; es gilt $\mathbb{K}(\xi) = \mathbb{K}(\sqrt{w_d})$. Wir können also $\sqrt{w_d}$ durch ξ darstellen; subtrahieren wir $\sqrt{w_d}$ danach von ξ, bleibt nur ξ_1 über, dessen Wurzeln nach Induktionsvoraussetzung durch ξ_1 und damit dann auch durch ξ darstellbar sein müssen.

□

Wir können also aus ξ alle einzelnen Wurzeln $\sqrt{w_i}$ isolieren. Die Aussage gilt natürlich auch dann noch, wenn wir statt $\xi = \frac{1}{n}\left(k_0 + \sum_{i=1}^{d} k_d\sqrt{w_d}\right)$ nur einen Teil der Summe haben, wenn die Summe also schon vor dem Grad d abbricht. In Definition 3.3.1 der Vollständigkeit haben wir auch die Zwischenkörper $\mathbb{K}_i$ definiert, die alle Konjugierten aller Wurzeln bis zur i-ten Wurzel enthielten; mit folgendem Korollar können wir zeigen, dass diese $\mathbb{K}_i$ wirklich Zwischenkörper zwischen $\mathbb{K}$ und dem Zerfällungskörper sind:

Korollar 4.3.5. *Sei $\xi = k_0 + \sum_{i=1}^{d} \sqrt{w_i}$ und sei $\eta_j = k_0 + \sum_{i=1}^{j} \sqrt{w_i}$ für $j \leq d$. Dann ist $\mathbb{K}_j = \mathbb{K}(\eta'_j \mid \eta'_j \sim \eta_j)$, und es gilt $\mathbb{K} = \mathbb{K}_0 \subseteq \mathbb{K}_1 \subseteq \cdots \subseteq \mathbb{K}_d = \mathbb{K}(\xi'|\xi' \sim \xi)$.*

Beweis. Es ist nach Definition $\mathbb{K}_j = \mathbb{K}(\sqrt{w_i'} \mid i < j, w_i' \sim w_i)$, und nach Lemma 4.3.4 ist es egal, ob wir die einzelnen Wurzeln oder die Summe adjungieren. Da dies jeweils auch für $\mathbb{K}_{i+1}$ gilt, ist jeweils $\mathbb{K}_i \subseteq \mathbb{K}_{i+1}$. □

Mit Hilfe dieses Korollars können wir jetzt zeigen, dass die Vollständigkeit eines Elementes bereits eine hinreichende Bedingung dafür ist, dass Q_d die Galoisgruppe des Zerfällungskörpers ist:

Satz 4.3.6. *ξ ist vollständig genau dann, wenn Q_d die Galoisgruppe des Zerfällungskörpers des Minimalpolynoms von ξ ist.*

Beweis. Ist Q_d die Galoisgruppe des Zerfällungskörpers, so ist der Zerfällungskörper eine Erweiterung vom Grad 2^{2^d-1}, und da alle $\mathbb{K}_i$ Zwischenkörper sind, muss der Grad von $\mathbb{K}_{i+1}$ über $\mathbb{K}_i$ genau 2^i sein, und folglich ist ξ vollständig.

Für die andere Richtung betrachten wir eine Folge $\{e\} = U_d \subset U_{d-1} \subset \cdots \subset U_1$ von Untergruppen von Q_d. Die Gruppe U_d enthält nur das neutrale Element von Q_d. Die Gruppe U_{d-1} enthält außerdem noch alle Vertauschungen benachbarter Elemente aus Q_d, also die Vertauschung 1 gegen 2, 3 gegen 4 und so weiter. U_{d-2} wiederum enthält U_{d-1} und darüber hinaus alle Vertauschungen von benachbarten Paaren aus Q_d, also 1 und 2 gegen 3 und 4 und so weiter. Definieren wir τ_i als die Vertauschung $\begin{pmatrix} 1 & \dots & 2^{i-1} & 2^{i-1}+1 & \dots & 2^i \\ 2^{i-1}+1 & \dots & 2^i & 1 & \dots & 2^{i-1} \end{pmatrix}$ und verwenden wieder $\odot$ für die Einbettung des äußeren Produkts in die passende Permutationsgruppe, so können wir U_i formal schreiben als:

$$U_i = U_{i+1} \bullet \left(\underbrace{\tau_{d-i} \odot \cdots \odot \tau_{d-i}}_{2^i \text{ mal}} \right)$$

Wir zeigen jetzt durch vollständige Induktion, dass U_i die jeweils mit dem Zwischenkörper $\mathbb{K}_i$ korrespondierende Untergruppe der Galoisgruppe des Zerfällungskörpers des Minimalpolynoms von ξ ist. Für $i = d$ ist dies offensichtlich, denn offenbar gehört nur das neutrale Element zur Galoisgruppe, da hier der ganze Körper fest gelassen wird. Sei $d > i \geq 1$ und gelte die Aussage schon für $i+1$. Da ξ vollständig ist, ist $[\mathbb{K}_{i+1} : \mathbb{K}_i] = 2^{2^i}$, d.h. für keine Konjugierte von w_i lässt sich $\sqrt{w_i}$ durch die Wurzeln der anderen Konjugierten darstellen. Also ist die Erweiterung eine Folge von 2^i galoischen Erweiterungen vom Grad 2, nämlich jeweils der Adjunktion von $\sqrt{w_i'}$ für eine Konjugierte von w_i. Die zugehörigen Galoisgruppen dieser Erweiterungen sind natürlich zweielementige Gruppen. Bei solchen Erweiterungen ist immer das nicht triviale Element der Galoisgruppe derjenige Körperautomorphismus, der $\sqrt{w_i'}$ durch $-\sqrt{w_i'}$ ersetzt. Bei den Konjugierten von ξ drückt sich dies durch eine Permutation der Nullstellen aus, in der genau die zu der entsprechenden Konjugierten gehörenden Nullstellen als Block gegen den folgenden Block ausgetauscht werden, was in der Konstruktion von U_i einem der τ_{d-i} entspricht. Die Galoisgruppe von $\mathbb{K}_i$ ist also die Galoisgruppe von $\mathbb{K}_{i+1}$, erweitert um die 2^i vielen Nachbarblock-Vertauschungen.

Letztlich ist $U_1 = Q_d$ die zu $\mathbb{K}$ gehörende Galoisgruppe.

□

Mit Hilfe von G_d zeigen wir jetzt die Eindeutigkeit der Darstellung von ξ für den Fall, dass ξ vollständig ist. Dazu benötigen wir noch folgendes Lemma:

Lemma 4.3.7. *Es gibt genau eine Folge von Untergruppen $G_d = U_0 \subset \cdots \subset U_d = Q_d$ mit $\frac{|U_{i+1}|}{|U_i|} = 2$.*

Beweis. Wir zeigen, dass die U_i von der Form $U_i = Q_i \odot Q_i \odot Q_{i+1} \odot \cdots \odot Q_{d-1}$ sind, d.h. wir zeigen, dass die so definierten U_i Gruppen zwischen G_d und Q_d sind und jeweils U_{i+1} doppelt so groß ist wie U_i. Definitionsgemäß gilt dies für U_0.

Gelte die Aussage also schon für U_i. Sei $\sigma \in U_{i+1}$, aber $\sigma \notin U_i$. Wir können o.B.d.A. annehmen, dass σ sich auf der rechten Hälfte der Elemente wie die Einheit verhält; insoweit dies nicht der Fall ist, können wir σ mit einem Element aus U_i multiplizieren, das sich links wie die Einheit verhält und rechts das Inverse der rechten Seite von σ ist. Beschränken wir uns auf die linke Seite, so können wir wieder o.B.d.A. annehmen, dass sich σ auch auf dem zweiten Viertel (und damit auf den gesamten drei hinteren Vierteln) wie die Einheit verhält. Dies setzen wir fort, bis wir die beiden Q_{i+1} erreichen. Es gibt nur noch ein Element, das Q_{i+2} von $Q_{i+1} \odot Q_{i+1}$ unterscheidet, nämlich die Vertauschung der rechten und der linken Seite. Damit σ also nicht in U_i liegt, muss σ sich auf den ersten Elementen wie diese Vertauschung verhalten. Damit ist $U_i \bullet \{\sigma, e\}$ aber genau das oben definierte U_{i+1}. □

Durch Zusammensetzen dieser Lemmata erhalten wir das gewünschte Eindeutigkeitsresultat:

Satz 4.3.8. *Ist $\xi \in \mathbb{E}_{\mathbb{K}}$ vollständig, dann ist die im Darstellungssatz beschriebene Darstellung von ξ eindeutig.*

Beweis. Ist ξ vollständig, so ist Q_d die Galoisgruppe von ξ und damit G_d die Galoisgruppe des Zerfällungskörpers über $\mathbb{K}(\xi)$. Durch die eindeutige Folge von Untergruppen U_i sind auch die Zwischenkörper $K_d = \mathbb{K}(\xi) \supset K_{d-1} \supset \cdots \supset K_0 = \mathbb{K}$ eindeutig, und da die $\mathbb{K}\left(\sqrt{w_1}, \ldots, \sqrt{w_i}\right)$ solche verschiedenen Zwischenkörper ausmachen, ist $K_i = \mathbb{K}\left(\sqrt{w_1}, \ldots, \sqrt{w_i}\right)$. Für jedes w_i ist dabei die Darstellung als $\xi_1 + \xi_2\sqrt{w_{i-1}}$ mit ξ_1 und $\xi_2 \in K_{i-1}$ eindeutig, und damit insbesondere auch die Darstellung von ξ selbst. □

Bemerkung 4.3.9. Die in Satz 4.3.8 gezeigte Eindeutigkeit der Darstellung gilt natürlich nur, solange wir keine Verschiebung der Vorfaktoren k_i in die Wurzeln betrachten. Streng genommen besagt der Satz, dass die in Lemma 3.1.3 beschriebene Darstellung eindeutig ist, d.h. die Darstellung, die die k_i immer bis auf Vorzeichen mit in die Wurzel zieht.

4.4 Beispiel einer Galoisgruppe

Wir wollen zum Abschluss dieses Kapitels ein Beispiel betrachten, an dem man sich Q_d, ihre Untergruppen und die zugehörigen Zwischenkörper verdeutlichen kann. Wir betrachten dafür ein weitgehend generisches Element vom Grad 4,

nämlich das Element $\xi = \sqrt{a} + \sqrt{b + \sqrt{a}}$. Auf den konstanten Summanden und auf die Vorfaktoren vor $\sqrt{a}$ verzichten wir der Einfachheit halber.

Die anderen Konjugierten von $\xi = \xi_1$ sind $\xi_2 = \sqrt{a} - \sqrt{b + \sqrt{a}}$, weiterhin $\xi_3 = -\sqrt{a} + \sqrt{b - \sqrt{a}}$ und $\xi_4 = -\sqrt{a} - \sqrt{b - \sqrt{a}}$. Das Minimalpolynom p von ξ ist das Polynom mit diesen vier Konjugierten als Nullstellen, also

$$p = X^4 - (2a + 2b)X^2 + 4aX - 2ab + b^2 + a^2 - a$$

Wir haben durch den Verzicht auf ein paar der Faktoren Einschränkungen in der Generalität von ξ gemacht, aber trotz dieser Einschränkungen ist ξ noch vollständig. Die Definition nachzuprüfen ist etwas aufwändig; wir können aber wesentlich einfacher feststellen, dass der Zerfällungskörper von p eine Körpererweiterung vom Grad acht ist, und da Q_2 vom Grad acht und eine 2-Sylowgruppe von S_8 ist, muss ξ vollständig sein. Dass der Zerfällungskörper von p Grad acht über dem Grundkörper $\mathbb{K} = \mathbb{Q}(a, b)$ hat, stellen wir wie folgt fest:

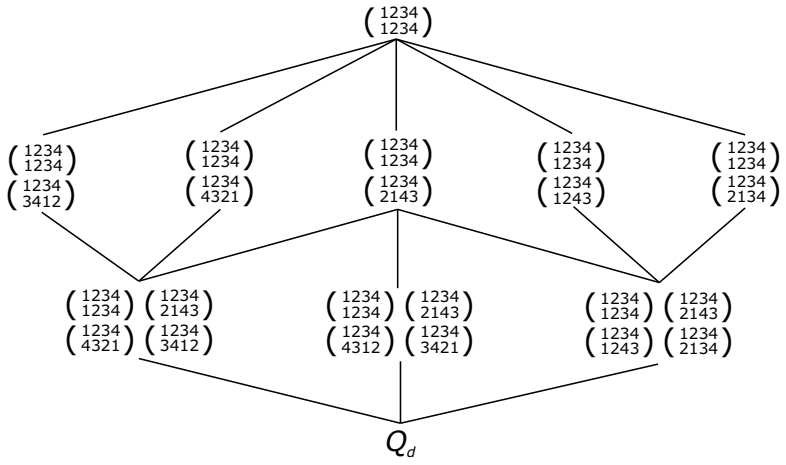

Abbildung 4.1: Die Untergruppen von Q_2 als Baum
Die Untergruppe der Ordnung 2 in der Mitte und die Untergruppen der Ordnung vier sind nicht triviale Normalteiler.

Lemma 4.4.1. *Der Zerfällungskörper $\mathbb{K}(\xi_1, \xi_2, \xi_3, \xi_4)$ von p hat Grad acht über $\mathbb{Q}(a, b)$.*

Beweis. $\sqrt{b^2 - a} \in \mathbb{K}(\xi_1, \xi_2, \xi_3, \xi_4)$, denn $(\xi_1 - \xi_2)(\xi_3 - \xi_4) = 4\sqrt{b^2 - a}$. Wie man nachrechnet, ist p aber über $\mathbb{K}(\sqrt{b^2 - a})$ immer noch irreduzibel, d.h. der Zerfällungskörper muss eine Erweiterung vom Grad mindestens vier über diesem Zwischenkörper sein, und damit mindestens eine Erweiterung vom Grad acht über $\mathbb{K}$. Mehr als Grad acht geht aber nicht, da acht die größte Zweierpotenz ist, die noch Teiler von 4! ist. □

ξ ist also vollständig und hat infolgedessen Q_2 als Galoisgruppe. Wir betrachten alle acht Elemente von Q_2:

$$Q_2 = \{ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} \}$$

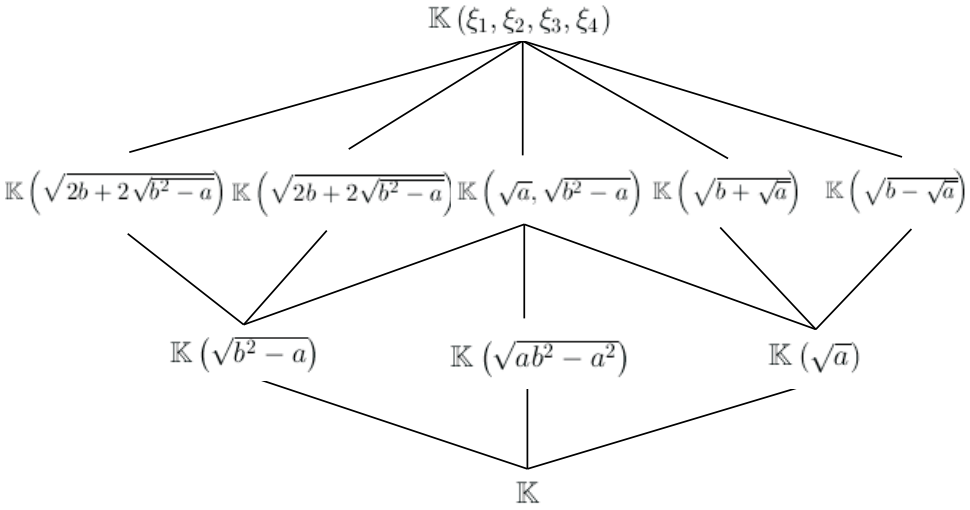

Abbildung 4.2: Die Zwischenkörper zwischen $\mathbb{Q}$ und dem Zerfällungskörper des Minimalpolynoms von $\sqrt{a} + \sqrt{b + \sqrt{a}}$
Die Zwischenkörper sind angeordnet wie die korrespondierenden Untergruppen von Q_2 in Abbildung 4.1.

Die ersten sechs Elemente sind selbstinvers und bilden daher jeweils mit dem neutralen Element eine Untergruppe der Ordnung zwei, außer dem neutralen Element selbst natürlich, das die triviale Untergruppe bildet. Das vierte Element ist das normale Element, und entsprechend ist die von diesem Element gebildete Untergruppe als einzige Untergruppe der Ordnung zwei ein Normalteiler von Q_2. Die ersten vier Elemente, die Elemente 1,4,7,8 und 1,4,5,6 bilden jeweils eine Untergruppe der Ordnung vier, und alle diese drei Untergruppen sind Normalteiler. Abbildung 4.1 zeigt die Anordnung der Gruppen.

Mit jeder Untergruppe korrespondiert ein Zwischenkörper zwischen $\mathbb{K}$ und dem Zerfällungskörper. Um eine bessere Einsicht in die Struktur der Zwischenkörper zu ermöglichen, wollen wir die Zuordnung an unserem Beispiel vollständig durchführen.

Die Zuordnung der trivialen Gruppen ist klar. Die rechte der drei Untergruppen der Ordnung vier bildet jeweils die Elemente im rechten und im linken Teil auf den gleichen Teil ab, so dass $\xi_1 + \xi_2$ und $\xi_3 + \xi_4$ konstant bleiben; diese beiden Summen entsprechen der Wurzel $\sqrt{a}$, d.h. dies ist der Zwischenkörper $\mathbb{K}_1$ aus Definition 3.3.1. Die beiden Körper darüber sind jeweils die Erweiterungen um die zweite Wurzel mit beiden Konjugierten, d.h. $\sqrt{b + \sqrt{a}}$ und $\sqrt{b - \sqrt{a}}$.

Die linke der drei Untergruppen der Ordnung vier vertauscht die Elemente eins und zwei gegen drei und vier oder beide synchron gegeneinander, d.h. die

Zahl $(\xi_1 - \xi_2)(\xi_3 - \xi_4) = \sqrt{b^2 - a}$ wird fest gelassen. Das Produkt dieser beiden Wurzeln, $\sqrt{ab^2 - a^2}$, wird dann von der mittleren Untergruppe fest gelassen.

Der Normalteiler mit zwei Elementen in der Mitte von Abbildung 4.1 lässt diese beiden Wurzeln separat fest. Die Untergruppe der Ordnung zwei ganz links lässt $\xi_1 + \xi_3 = -(\xi_2 + \xi_4) = \sqrt{2b + 2\sqrt{b^2 - a}}$ fest, die zweite Untergruppe von links lässt $\xi_2 + \xi_3 = -(\xi_1 + \xi_4) = \sqrt{2b - 2\sqrt{b^2 - a}}$ fest. Abbildung 4.2 gibt eine Übersicht über die Zwischenkörper in gleicher Anordnung wie die zugehörigen Untergruppen der Galoisgruppe in Abbildung 4.1.

Kapitel 5

Auflösen univariater Polynome durch Wurzelausdrücke

In diesem Kapitel wollen wir jetzt einen neuen Algorithmus einführen, um zu einem gegebenen Polynom $p \in \mathbb{K}[X]$ eventuell existierende Nullstellen $\xi \in \mathbb{E}_\mathbb{K}$ finden zu können. Wenn p eine solche Nullstelle hat, so ist das Minimalpolynom von ξ ein Faktor von p, und nach Korollar 2.1.4 wissen wir, dass der Grad des Minimalpolynom eine Zweierpotenz ist. Wir faktorisieren p zunächst über $\mathbb{K}$ in seine irreduziblen Anteile und ignorieren alle Faktoren, deren Grad keine Zweierpotenz ist. Bei den restlichen Faktoren suchen wir dann nach Nullstellen aus $\mathbb{E}_\mathbb{K}$.

Zunächst werden wir einen Test einführen, mit dem wir für ein gegebenes Polynom mit Grad einer Zweierpotenz schnell feststellen können, ob es Minimalpolynom eines Elementes $\xi \in \mathbb{E}_\mathbb{K}$ ist. Im zweiten Abschnitt beschreiben wir eine konstruktive Zerlegung des Elementes, die uns ermöglicht, die Darstellung von ξ rekursiv zu berechnen. Die dafür benötigten Algorithmen werden in den dann folgenden Abschnitten näher eingeführt.

5.1 Schnelles Testen über einem Erweiterungskörper von $\mathbb{Q}$

Wie schon erwähnt, können wir ein gegebenes Polynom $p \in \mathbb{K}[X]$ zunächst faktorisieren und alle Faktoren ignorieren, deren Grad nicht eine Zweierpotenz ist. Ein verbleibendes irreduzibles Polynom p mit einer Zweierpotenz muss aber nicht notwendig Nullstellen aus $\mathbb{E}_\mathbb{K}$ haben. Man betrachte als Beispiel das Polynom $p = X^4 + 3X^3 + 1$. p ist irreduzibel über $\mathbb{Q}$; faktorisiert man p aber über $\mathbb{F}_5$, so zerfällt p in die Faktoren $p = (X + 4)(X^3 - X^2 - X - 1)$. Für jedes Minimalpolynom eines $\xi \in \mathbb{E}_\mathbb{Q}$ ist es aber eine notwendige Bedingung, dass der Grad aller Faktoren über einem Primkörper ebenfalls eine Zweierpotenz ist, wie folgendes Lemma zeigt:

Lemma 5.1.1. *Sei $\mathbb{K} = \mathbb{Q}(Y_1, \ldots, Y_n)$ eine transzendente Erweiterung von*

$\mathbb{Q}$, weiterhin q eine Primzahl und $\mathbb{K}_q = \mathbb{F}_q(Y_1, \ldots, Y_n)$. Sei φ der natürliche Homomorphismus von $\mathbb{Z}(Y_1, \ldots, Y_n) \to \mathbb{F}_q(Y_1, \ldots, Y_n)$ und $p \in \mathbb{K}[X]$ das Minimalpolynom eines $\xi \in \mathbb{E}_\mathbb{K}$. Sei p' ein Vielfaches von p vom selben Grad aus $\mathbb{Z}[Y_1, \ldots, Y_n][X]$. Dann ist der Grad aller Faktoren von $\varphi(p')$ über $\mathbb{K}_q$ eine Zweierpotenz.

Beweis. Es genügt zu zeigen, dass alle Nullstellen von $\varphi(p') \in \mathbb{E}_{\mathbb{K}_q}$ sind, denn nach Korollar 2.1.4 müssen alle irreduziblen Polynome mit Nullstellen aus $\mathbb{E}_{\mathbb{K}_q}$ Zweierpotenzen als Grad haben. Wir betrachten das Element $\xi' \in \mathbb{E}_{\mathbb{K}_q}$, das wir aus einer beliebigen Darstellung von ξ nach dem Darstellungssatz erhalten, wenn wir den Nenner n von ξ auf 1 setzen und φ auf jedes Element von $\mathbb{Z}$ in der Darstellung anwenden. Jeder Ausdruck in den Nullstellen, der ein Element aus $\mathbb{K}$ darstellt, ist ein Ausdruck in diesen Elementen, und die Homomorphieeigenschaft von φ bleibt erhalten; also ist insbesondere die Auswertung des Polynoms p bzw. p' von ξ, die ja null ergibt, ein Ausdruck in diesen Elementen aus $\mathbb{Z}$; wenden wir φ auf diese Elemente an (d.h. wir werten $\varphi(p')$ an der Stelle ξ' aus), erhalten wir wegen der Homomorphieeigenschaft wieder null. Dies können wir mit jeder Konjugierten von ξ durchführen und erhalten so mit Vielfachheit alle Nullstellen von $\varphi(p')$, und die Aussage ist bewiesen. □

Wir können auf diese Weise schnell testen, ob ein irreduzibles Polynom über einer algebraischen Erweiterung von $\mathbb{Q}$ durch Quadratwurzeln auflösbar ist, indem wir das Polynom modulo einiger Primzahlen faktorisieren. Entsteht dabei ein Faktor, dessen Grad keine Zweierpotenz ist, hat das Polynom keine durch Quadratwurzeln darstellbaren Nullstellen.

Dies ist aber offenbar nur eine notwendige Bedingung. Es stellt sich nun die umgekehrte Frage, wie viele positive Tests benötigt werden, um mit einer hohen Wahrscheinlichkeit davon ausgehen zu können, dass p tatsächlich eine Nullstelle in $\mathbb{E}_\mathbb{K}$ hat.

Nehmen wir also an, wir haben ein Polynom p, dessen Nullstellen nicht durch Quadratwurzeln darstellbar sind. Sei G die Galoisgruppe des Zerfällungskörpers von p. Faktorisieren wir p modulo einer Primzahl q, so haben die jeweiligen Zerfällungskörper von $p \mod q$ eine Untergruppe von G als Galoisgruppe. Wenn G eine Untergruppe U hat, deren Größe keine Zweierpotenz ist, so ist nach dem Satz von Chebotarev (siehe in [25]) die Dichte der Primzahlen, die einen zu U gehörenden Faktor erzeugen, gleich $\frac{|U|}{|G|}$.

Die Grenzen für die umgekehrte Wahrscheinlichkeit, für Primzahlen bis zu einer bestimmten Größe eine solche Untergruppe U zu finden, sind theoretisch nicht sehr günstig und bleiben weit hinter der wirklichen Effektivität des Testes zurück. Solche Grenzen wurden von Lagarias und Odlyzko in [14] gegeben. In der Praxis funktioniert der Test aber auch bei wenigen kleinen Primzahlen schon sehr effektiv; faktorisiert man ein gegebenes, über $\mathbb{Q}$ irreduzibles Polynom nur über den ersten 25 Primzahlen (d.h. den Primzahlen mit einem Betrag kleiner 100), so wird der Test in der Praxis bereits eine nahezu fehlerfreie Aussage ermöglichen. Wir haben zu diesem Zweck 2174 zufällige, irreduzible Polynome getestet, deren Grad eine Zweierpotenz zwischen 4 und 512 war (knapp 2000 Polynome vom Grad 4, jeweils knapp 50 vom Grad 8 bis 64, und jeweils 3

vom Grad 128, 256 und 512). Die Polynome wurden modulo aller ersten 25 Primzahlen faktorisiert, und es wurde jeweils gezählt, wie viele der Primzahlen eine Faktorisierung ergab, in der mindestens bei einem Faktor der Grad keine Zweierpotenz war.

Die Ergebnisse dieses Experimentes finden sich in folgender Tabelle. In den Spalten stehen die verschiedenen Polynomgrade, während die Zeilen die Anzahl an Primzahlen angibt, für die der schnelle Test ein negatives Ergebnis zurückgab. In den einzelnen Zellen steht dann die Anzahl der Polynome, auf die dies zutraf.

	4	8	16	32	64	128	256	512
0	1	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0
2	10	0	0	0	0	0	0	0
3	25	0	0	0	0	0	0	0
4	50	0	0	0	0	0	0	0
5	142	0	0	0	0	0	0	0
6	245	0	0	0	0	0	0	0
7	307	0	0	0	0	0	0	0
8	350	0	0	0	0	0	0	0
9	320	0	0	0	0	0	0	0
10	211	0	0	0	0	0	0	0
11	136	0	0	0	0	0	0	0
12	109	0	0	0	0	0	0	0
13	41	0	0	0	0	0	0	0
14	19	2	0	0	0	0	0	0
15	6	7	0	0	0	0	0	0
16	2	5	0	0	0	0	0	0
17	0	9	0	0	0	0	0	0
18	0	9	0	0	0	0	0	0
19	0	7	1	0	0	0	0	0
20	0	7	0	0	0	0	0	0
21	0	2	6	0	0	0	0	0
22	0	0	10	2	0	0	0	0
23	0	0	10	8	3	0	0	0
24	0	0	16	13	13	0	0	1
25	0	0	4	24	32	3	3	2
Σ	1975	48	47	47	48	3	3	3
μ	8,2	17,6	23,0	24,3	24,6	25,0	25,0	24,7

In der vorletzten Zeile stehen jeweils die Summen der Polynome mit dem zur Spalte gehörenden Grad und in der letzten Zeile die mittlere Anzahl an Primzahlen, bei denen der Test ein negatives Resultat zurückgab.

Man erkennt, dass selbst bei Polynomen vom Grad vier nahezu jedes Polynom mindestens bei einer der ersten 25 Primzahlen abgelehnt wurde. Schaut

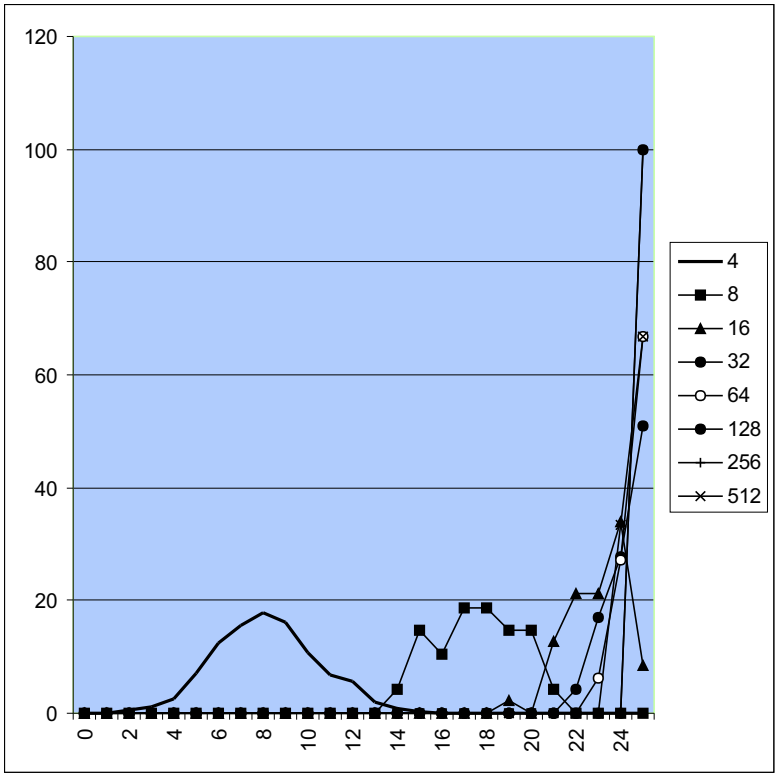

Abbildung 5.1: Diagramm zu den Ergebnissen der Versuche mit dem schnellen Testalgorithmus

man sich das zugehörige Diagramm 5.1 (und darin insbesondere die Kurve am weitesten links, die zu Polynomen vom Grad 4 gehört und auf den meisten Daten beruht) an, so lassen die Daten ungefähr eine Binomialverteilung vermuten. Um ein Gefühl für die Wahrscheinlichkeit zu bekommen, dass der Test fehlschlägt, wollen wir einmal anhand der Daten annehmen, es handele sich tatsächlich um eine Binomialverteilung mit dem Mittelwert von 8,17, den man bei den 1975 Polynomen vom Grad vier erhält. Unter diesen Annahmen wäre die Wahrscheinlichkeit, innerhalb einer solchen Verteilung zufällig 25 mal innerhalb der ersten Primzahlen nur Faktoren mit Zweierpotenz als Grad zu bekommen, 0,0051 %. Bei den Polynomen höheren Grades liegt die Wahrscheinlichkeit noch um Größenordnungen darunter. Wir können also wohl davon ausgehen, dass unser Test mit den ersten 25 Primzahlen ausgeführt hinreichend gut ist.

Wie die Tabelle zeigt, ergab eines der Polynome bei der Faktorisierung bei allen 25 Primzahlen nur Faktoren mit Zweierpotenz. Dieses zufällig erzeugte Polynom war $X^4 - 87X^3 + 5X^2 - 2X + 92$. Mit Hilfe der in den folgenden Abschnitten beschriebenen Mitteln berechnet man, dass tatsächlich $\frac{1}{4}\left(87 + \sqrt{7913} + \sqrt{14754 + 166 \cdot \sqrt{7913}}\right)$ eine Nullstelle dieses Polynoms ist; also hat der Test bei allen 2174 zufällig erzeugten Polynomen ausnahmslos das richtige Ergebnis zurückgeliefert.

Require: p irreduzibel und $\deg(p)$ Zweierpotenz
  for $i := 1$ bis 25 **do**
    $\bar{p} := p \mod prim[i]$
    for all $\widetilde{p}$ Faktor von $\bar{p}$ **do**
      if $\deg(\widetilde{p})$ keine Zweierpotenz **then**
        return $false$
      end if
    end for
  end for
  return $true$

Abbildung 5.2: Schnelles Testen auf durch Wurzeln ausdrückbare Nullstellen.

5.2 Vorbereitung zur konstruktiven Berechnung

Wir können jetzt also schnell feststellen, ob ein gegebenes Polynom Minimalpolynom eines $\xi \in \mathbb{E}_{\mathbb{K}}$ ist. Der folgende Abschnitt beschäftigt sich jetzt damit, diese Nullstelle - d.h. ihre Darstellung nach dem Darstellungssatz - konstruktiv und deterministisch zu finden.

Sei also p Minimalpolynom eines $\xi \in \mathbb{E}_{\mathbb{K}}$. Wir zerlegen ξ in den Teil vor der obersten Wurzel und in die oberste Wurzel, also $\xi = \xi_1 + \sqrt{\xi_2}$. ξ_1 und ξ_2 sind ebenfalls Elemente von $\mathbb{E}_{\mathbb{K}}$. Wählen wir die Darstellung aus Lemma 3.3.6, so ist $\deg(\xi_1) \leq \frac{1}{2}\deg(\xi)$ und $\deg(\xi_2) = \frac{1}{2}\deg(\xi)$. Können wir also die Minimalpolynome für ξ_1 und ξ_2 bestimmen, so haben wir unser Problem auf zwei einfachere Probleme zurückgeführt und können diesen Prozess wiederholen, bis $\deg(\xi) = 1$ ist.

ξ_1 und ξ_2 lassen sich als Summe geeigneter Konjugierten von ξ darstellen. Der folgende Satz nutzt diesen Umstand aus und liefert uns eine Aussage über die Minimalpolynome von ξ_1 und ξ_2.

Sei $\deg(\xi) = 2^d$, $(v_1, \ldots, v_d) \in \{0,1\}^d$ und $\xi^{(v_1,\ldots,v_d)}$ eine Konjugierte von ξ, die wie in Definition 3.2.1 definiert dann ein anderes Vorzeichen als ξ vor der Wurzel w_i hat, wenn $v_i = 1$ ist. Wir stellen für die Minimalpolynome von ξ_1 und ξ_2 fest:

Satz 5.2.1. *Sei $\xi \in \mathbb{E}_{\mathbb{K}}$ und $\xi = \xi_1 + \sqrt{\xi_2}$. Dann ist*

$$p_1 = \prod_{(v_1,\ldots,v_{d-1}) \in \{0,1\}^{d-1}} X - \left(\xi^{(v_1,\ldots,v_{d-1},0)} + \xi^{(v_1,\ldots,v_{d-1},1)}\right)$$

eine Potenz des Minimalpolynoms von $2\xi_1$ und

$$p_2 = \prod_{(v_1,\ldots,v_{d-1}) \in \{0,1\}^{d-1}} X - \left(\xi^{(v_1,\ldots,v_{d-1},0)} - \xi^{(v_1,\ldots,v_{d-1},1)}\right)^2$$

das Minimalpolynom von $4\xi_2$.

Beweis. Es ist

$$\begin{aligned} p_1 = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - \left(\xi^{(v_1,\ldots,v_{d-1},0)} + \xi^{(v_1,\ldots,v_{d-1},1)}\right) \\ = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - \xi_1^{(v_1,\ldots,v_{d-1})} - \sqrt{\xi_2^{(v_1,\ldots,v_{d-1})}} \\ & \qquad - \xi_1^{(v_1,\ldots,v_{d-1})} + \sqrt{\xi_2^{(v_1,\ldots,v_{d-1})}} \\ = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - 2\,\xi_1^{(v_1,\ldots,v_{d-1})} \end{aligned}$$

Sei $2^{d'} = \deg(\xi_1)$. Für $d' < d$ sind einige der Terme $\xi_1^{(v_1,\ldots,v_{d-1})}$ identisch, und wir können umformen:

$$\begin{aligned} p_1 = & \prod_{(v_1,\ldots,v_{d'-1})\in\{0,1\}^{d'-1}} \left(X - 2\,\xi_1^{(v_1,\ldots,v_{d'-1})}\right)^{2^{d-d'}} \\ = & \left(\prod_{(v_1,\ldots,v_{d'-1})\in\{0,1\}^{d'-1}} X - 2\,\xi_1^{(v_1,\ldots,v_{d'-1})}\right)^{2^{d-d'}} \end{aligned}$$

Der Term in den Klammern ist das Produkt aller Konjugierten von $2\xi_1$, also das Minimalpolynom. Entsprechend gilt für p_2:

$$\begin{aligned} p_2 = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - \left(\xi^{(v_1,\ldots,v_{d-1},0)} - \xi^{(v_1,\ldots,v_{d-1},1)}\right)^2 \\ = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - \left(\xi_1^{(v_1,\ldots,v_{d-1})} + \sqrt{\xi_2^{(v_1,\ldots,v_{d-1})}}\right. \\ & \qquad \left. - \xi_1^{(v_1,\ldots,v_{d-1})} + \sqrt{\xi_2^{(v_1,\ldots,v_{d-1})}}\right)^2 \\ = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - \left(2\,\sqrt{\xi_2^{(v_1,\ldots,v_{d-1})}}\right)^2 \\ = & \prod_{(v_1,\ldots,v_{d-1})\in\{0,1\}^{d-1}} X - 4\xi_2^{(v_1,\ldots,v_{d-1})} \end{aligned}$$

Der letzte Term ist das Produkt alle Konjugierten von $4\xi_2$, also das Minimalpolynom.

□

5.3 Algorithmus für das Kombinationspolynom

Wir können jetzt also unser Problem der Suche nach Nullstellen auf ein einfacheres Problem zurückführen, wenn wir ein Polynom mit den Summen bzw. den Differenzen der jeweils passenden beiden Konjugierten finden. Leider kennen wir die Nullstellen nicht explizit, sondern nur das Minimalpolynom p für unsere Nullstellen; das ist jedoch hinreichend, wenn es uns gelingt, ein Polynom

p_+ bzw. p_- zu finden, dessen Nullstellen gerade die Summen bzw. die Differenzen zweier Nullstellen von p sind. In diesem Fall wird p_+ das Polynom p_1 und p_- das Polynom p_2 als Faktor haben; wenn wir also das Summen- bzw. Differenzenpolynom faktorisieren, erhalten wir dadurch die gesuchten Polynome.

5.3.1 Allgemeines Kombinationspolynom

Der nachfolgende Algorithmus ist etwas mächtiger, als das für unser spezielles Problem notwendig wäre. Wir werden ihn später auf unser spezielles Problem einschränken, zunächst aber lösen wir das folgende Problem in voller Allgemeinheit:

Definition 5.3.1. Seien $p_1, \ldots, p_m \in R[X]$ Polynome vom Grad $n_1, \ldots, n_m$ über einem Ring und $f \in R[X_1, \ldots, X_m]$ eine über ein Polynom beschriebene Funktion. Seien weiterhin $\xi_0^{(j)}, \ldots, \xi_{n_j-1}^{(j)}$ die Nullstellen von p_j. Dann heiße

$$K_f(p_1, \ldots, p_m) := \prod_{\substack{(i_1,\ldots,i_m) \in \\ \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m}}} X - f(\xi_{i_1}^{(1)}, \ldots, \xi_{i_m}^{(m)})$$

das f-Kombinationspolynom von $p_1, \ldots, p_m$.

Für unsere spezielle Anwendung werden wir uns später auf den Sonderfall $m = 2$ und $f(\xi, \eta) = \xi + \eta$ beschränken und dann $K_f(p, p)$ berechnen. Das allgemeine Problem lässt sich aber mit dem gleichen Arbeitsaufwand mit erledigen und ist allein für sich betrachtet interessant.

Wir suchen also eine Möglichkeit, dieses Polynom K_f zu bestimmen. Zunächst stellen wir über den Grad von K_f fest:

Lemma 5.3.2. *Es gilt*

$$\deg\left(K_f(p_1, \ldots, p_m)\right) = \prod_{i=1}^{m} \deg\left(p_i\right)$$

Beweis. $\deg(p_i)$ ist gleich der Anzahl der Nullstellen von p_i, also n_i; das Produkt in der Definition von K_f durchläuft alle Tupel aus $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m}$, also das Produkt über alle n_i.

□

Wenn wir mit $\eta_0, \ldots, \eta_{l-1}$ die Nullstellen von K_f bezeichnen (wobei $l = \deg(K_f)$ ist), dann können wir die Koeffizienten von K_f über die elementarsymmetrischen Polynome ausdrücken, also

$$K_f = X^l + \sigma_1 X^{l-1} + \cdots + \sigma_l$$

mit

$$\begin{aligned}
\sigma_1(\eta_0,\ldots,\eta_{l-1}) &:= \sum_{i=0}^{l-1}\eta_i \\
\sigma_2(\eta_0,\ldots,\eta_{l-1}) &:= \sum_{(i_1<i_2)\in\mathbb{Z}_l^2}\eta_{i_1}\eta_{i_2} \\
\sigma_3(\eta_0,\ldots,\eta_{l-1}) &:= \sum_{(i_1<i_2<i_3)\in\mathbb{Z}_l^3}\eta_{i_1}\eta_{i_2}\eta_{i_3} \\
&\vdots \\
\sigma_l(\eta_0,\ldots,\eta_{l-1}) &:= \sum_{(i_1<\ldots<i_l)\in\mathbb{Z}_l^l}\eta_{i_1}\ldots\eta_{i_l} = \prod_{i=0}^{l-1}\eta_i
\end{aligned}$$

Um die Ordnung der Indizes zu eliminieren, nehmen wir einige Terme mehrfach in die Summe auf und teilen dann entsprechend durch die Häufigkeit, mit der jeder Term auftritt. Damit können wir das k-te elementarsymmetrische Polynom wie folgt aufschreiben:

$$\sigma_k(\eta_0,\ldots,\eta_{l-1}) := \frac{1}{k!}\sum_{(i_1\neq\cdots\neq i_k)\in\mathbb{Z}_l^k}\eta_{i_1}\ldots\eta_{i_k}$$

Dabei notieren wir im Folgenden der Einfachheit halber $(i_1 \neq \cdots \neq i_k)$ für ein k-Tupel aus paarweise verschiedenen Elementen. Wir definieren jetzt eine Folge von Funktion c_k, die im Wesentlichen die Summe aus der Definition von σ_k enthält, allerdings mit variablen Exponenten an jedem η_i:

Definition 5.3.3. Es sei

$$c_k : \mathbb{N}^k \to \mathbb{Z}[\eta_0,\ldots,\eta_{l-1}]$$

mit

$$c_k(n_1,\ldots,n_k) = \sum_{(i_1\neq\cdots\neq i_k)\in\mathbb{Z}_l^k}\eta_{i_1}^{n_1}\ldots\eta_{i_k}^{n_k}$$

Korollar 5.3.4. *c_k ist symmetrisch, d.h.*

$$\forall\tau\in S_k :\ c_k(n_1,\ldots,n_k) = c_k(n_{\tau(1)},\ldots,n_{\tau(k)})$$

Beweis. Folgt unmittelbar aus der Kommutativität der Multiplikation in der Definition von c_k. □

Die Definition der c_k wirkt im ersten Moment äußerst willkürlich. Ursprünglich waren wir an den c_k mit nur Einsen für die k Argumente interessiert, um daraus die elementarsymmetrischen Polynome und somit die Koeffizienten von K_f berechnen zu können. Wir werden aber feststellen, dass wir die $c_k(1,\ldots,1)$ sehr effizient durch die $c_1(n)$ ausdrücken können; für diese wiederum finden sich aber - je nach der Funktion f - oft einfache Darstellungen über die Koeffizienten der Polynome $p_1,\ldots,p_n$. Besonders für die Spezialfälle des Summen- und Differenzenpolynoms finden wir hier schön einfache Darstellungen.

Wir zeigen jetzt also zunächst, wie man die c_k rekursiv darstellen kann.

Satz 5.3.5. *Es gilt:*

$$\begin{aligned} c_k(n_1,\ldots,n_k) &= c_1(n_k)\cdot c_{k-1}(n_1,\ldots,n_{k-1}) \\ &\quad -\sum_{j=1}^{k-1} c_{k-1}(n_1,\ldots,n_j+n_k,\ldots,n_{k-1}) \end{aligned}$$

Beweis. Es ist

$$\begin{aligned} c_k(n_1,\ldots,n_k) &= \sum_{\substack{(i_1\neq\cdots\neq i_k)\\ \in\mathbb{Z}_l^k}} \eta_{i_1}^{n_1}\ldots\eta_{i_k}^{n_k} \\ &= \sum_{\substack{(i_1\neq\cdots\neq i_{k-1})\\ \in\mathbb{Z}_l^{k-1}}} \eta_{i_1}^{n_1}\ldots\eta_{i_{k-1}}^{n_{k-1}}\cdot \sum_{\substack{i_k\in\mathbb{Z}_l\\ i_k\notin\{i_1,\ldots,i_{k-1}\}}} \eta_{i_k}^{n_k} \\ &= \sum_{\substack{(i_1\neq\cdots\neq i_{k-1})\\ \in\mathbb{Z}_l^{k-1}}} \eta_{i_1}^{n_1}\ldots\eta_{i_{k-1}}^{n_{k-1}}\cdot \left(\sum_{i_k\in\mathbb{Z}_l}\eta_{i_k}^{n_k} - \sum_{j=1}^{k-1}\eta_{i_j}^{n_k}\right) \\ &= \left(\sum_{i_k\in\mathbb{Z}_l}\eta_{i_k}^{n_k}\right)\left(\sum_{\substack{(i_1\neq\cdots\neq i_{k-1})\\ \in\mathbb{Z}_l^{k-1}}} \eta_{i_1}^{n_1}\ldots\eta_{i_{k-1}}^{n_{k-1}}\right) \\ &\quad -\sum_{j=1}^{k-1}\sum_{\substack{(i_1\neq\cdots\neq i_{k-1})\\ \in\mathbb{Z}_l^{k-1}}} \eta_{i_1}^{n_1}\ldots\eta_{i_j}^{n_j+n_k}\ldots\eta_{i_{k-1}}^{n_{k-1}} \\ &= c_1(n_k)\cdot c_{k-1}(n_1,\ldots,n_{k-1}) \\ &\quad -\sum_{j=1}^{k-1} c_{k-1}(n_1,\ldots,n_j+n_k,\ldots,n_{k-1}) \end{aligned}$$

□

Die elementarsymmetrischen Polynome und damit die Koeffizienten von K_f sind durch die $c_k(1,\ldots,1)$ leicht zu berechnen; wir interessieren uns also hauptsächlich für diese Argumente von c_k. Für diese Argumente lässt sich die Rekursionsformel noch weiter vereinfachen, wobei wir nur Rekursionen im Index von c und im letzten der Argumente benötigen. Für diese gilt folgendes Korollar:

Korollar 5.3.6. *Es gilt:*

$$c_k(1,\ldots,1,n) = c_1(n)\cdot c_{k-1}(1,\ldots,1) - (k-1)\cdot c_{k-1}(1,\ldots,1,n+1)$$

Beweis.

$$\begin{aligned} c_k(1,\ldots,1,n) &= c_1(n)\cdot c_{k-1}(1,\ldots,1) - \sum_{j=1}^{k-1} c_{k-1}(1,\ldots,1+n,\ldots,1) \\ &= c_1(n)\cdot c_{k-1}(1,\ldots,1) - \sum_{j=1}^{k-1} c_{k-1}(1,\ldots,1,1+n) \\ &= c_1(n)\cdot c_{k-1}(1,\ldots,1) - (k-1)\cdot c_{k-1}(1,\ldots,1,n+1) \end{aligned}$$

□

Die $c_k(1,\ldots,1,n)$ bezeichnen wir als *Gaertner-Polynome* und schreiben sie als $g(k,n) := c_k(1,\ldots,1,n)$. Wir rollen die Rekursionsgleichung im zweiten Argument auf und erhalten folgende kompaktere Darstellung:

Lemma 5.3.7. *Es gilt:*

$$\begin{aligned} g(k,n) &= (-1)^{k-1}(k-1)!\cdot g(1,k+n-1) \\ &\quad + \sum_{i=0}^{k-2}(-1)^i \frac{(k-1)!}{(k-1-i)!} g(k-1-i,1)\cdot g(1,n+i) \end{aligned}$$

Beweis. Wir machen eine Induktion über das k im Argument. Offenbar gilt die Aussage für $k=1$, denn die Summe ergibt dann null, und der Vorterm ist genau $g(k,n)$. Gelte die Aussage für alle $k' < k$, dann gilt:

$$\begin{aligned} g(k,n) &= g(k-1,1)\cdot g(1,n) - (k-1)\cdot g(k-1,n+1) \\ &= g(k-1,1)\cdot g(1,n) - (k-1)(-1)^{k-2}(k-2)!\cdot g(1,k+n-1) - \\ &\quad (k-1)\sum_{i=0}^{k-3}(-1)^i \frac{(k-2)!}{(k-2-i)!} g(k-2-i,1)\cdot g(1,n+1+i) \\ &= (-1)^{k-1}(k-1)!\cdot g(1,k+n-1) + \\ &\quad \frac{(k-1)!}{(k-1-0)!} g(k-1-0,1)\cdot g(1,n+0) + \\ &\quad \sum_{i=1}^{k-2}(-1)^i \frac{(k-1)!}{(k-1-i)!} g(k-1-i,1)\cdot g(1,n+i) \\ &= (-1)^{k-1}(k-1)!\cdot g(1,k+n-1) + \\ &\quad \sum_{i=0}^{k-2}(-1)^i \frac{(k-1)!}{(k-1-i)!} g(k-1-i,1)\cdot g(1,n+i) \end{aligned}$$

□

Es gelingt uns also, die $g(k,n)$ durch eine relativ einfache Rekursion zu berechnen, wenn wir die $g(1,n)$ voraussetzen. Die $g(1,n)$ $(= c_1(n))$ müssen jetzt noch jeweils abhängig von der Funktion f berechnet werden. Wir trennen uns hier von dem allgemeinen Fall und betrachten die beiden für uns interessanten Sonderfälle p_+ und p_-.

5.3.2 Das Differenzenpolynom

Um die Nullstellen eines Polynoms p im Quadratwurzelkörper berechnen zu können, interessieren wir uns für die Differenzen zweier Nullstellen von p, d.h. für $K_f(p,p)$ mit $f(\xi_1,\xi_2)=\xi_1-\xi_2$. Wir notieren dieses Kombinationspolynom mit p_-, also $p_- := K_f(p,p)$ mit der Subtraktionsfunktion für f. Wir stellen zuerst einige Eigenschaften dieses Polynoms fest.

Lemma 5.3.8. *p_- hat einen Faktor X^d, und in $\frac{p_-}{X^d}$ sind alle Koeffizienten mit ungeraden Exponenten in X null.*

Beweis.

$$\begin{aligned}
p_- &= \prod_{(i,j)\in\mathbb{Z}_d^2} X-(\xi_i-\xi_j)\\
&= \prod_{(i<j)\in\mathbb{Z}_d^2} X-(\xi_i-\xi_j)\cdot\prod_{(i>j)\in\mathbb{Z}_d^2} X-(\xi_i-\xi_j)\cdot\prod_{i\in\mathbb{Z}_d} X-(\xi_i-\xi_i)\\
&= \left(\prod_{(i<j)\in\mathbb{Z}_d^2} X-(\xi_i-\xi_j)\cdot\prod_{(i<j)\in\mathbb{Z}_d^2} X+(\xi_i-\xi_j)\right)\cdot X^d\\
&= \left(\prod_{(i<j)\in\mathbb{Z}_d^2} X^2-(\xi_i-\xi_j)^2\right)\cdot X^d
\end{aligned}$$

□

Die $g(k,1)$ des letzten Abschnittes 5.3.1 entsprechen den Koeffizienten von p_- bis auf einen multiplikativen Faktor. Daher sind auch alle $g(k,1)$ für ungerade k null.

Wenden wir jetzt die Rekursionsformel für $g(k,n)$ an, so erhalten wir eine einfache Darstellung für die $g(k,1)$:

Lemma 5.3.9. *Für ungerade k ist $g(k,1)=0$, und für gerade k gilt:*

$$g(k,1) = -(k-1)!\left(g(1,k)+\sum_{\substack{i=0\\ i \text{ ungerade}}}^{k-2}\frac{1}{(k-1-i)!}g(k-1-i,1)g(1,1+i)\right)$$

Beweis. Der 1. Fall ist bereits in Lemma 5.3.8 bewiesen. Nehmen wir an, k ist gerade. Dann gilt nach Lemma 5.3.7

$$g(k,1) = (-1)^{k-1}(k-1)!\cdot g(1,k)+\sum_{i=0}^{k-2}(-1)^i\frac{(k-1)!}{(k-1-i)!}g(k-1-i,1)\cdot g(1,1+i)$$

Ist in der Summe i gerade, so wird $k-1-i$ ungerade, und damit $g(k-1-i,1)=0$. Ist i ungerade, so ist $(-1)^i=-1$. Da k gerade ist, ist $(-1)^{k-1}$ ebenfalls -1. Klammert man das $(k-1)!$ noch aus, erhält man die Behauptung. □

Um die Koeffizienten von p_- zu berechnen, müssen wir die $g(k,1)$ noch durch $k!$ teilen; der einfacheren Berechnung wegen nehmen wir diese Division in die Rekursion mit auf und definieren $\gamma_1(n) := \frac{1}{n!}g(n,1)$, wodurch wir für γ_1 folgende Rekursion erhalten:

Lemma 5.3.10. *Für ungerade n ist $\gamma_1(n)$ null, und für gerade n gilt:*

$$\gamma_1(n) = -\frac{1}{n}\left(g(1,n) + \sum_{\substack{i=0 \\ i \text{ ungerade}}}^{n-2} \gamma_1(n-1-i)g(1,1+i)\right)$$

Beweis. Der Fall für n ungerade ist trivial. Sei also n gerade, dann gilt

$$\begin{aligned}
\gamma_1(n) &= \frac{1}{n!}g(n,1) \\
&= -\frac{1}{n!}(n-1)! \\
&\quad \cdot \left(g(1,n) + \sum_{\substack{i=0 \\ i \text{ ungerade}}}^{n-2} \frac{1}{(n-1-i)!}g(n-1-i,1)g(1,1+i)\right) \\
&= -\frac{1}{n}\left(g(1,n) + \sum_{\substack{i=0 \\ i \text{ ungerade}}}^{n-2} \gamma_1(n-1-i)g(1,1+i)\right)
\end{aligned}$$

□

Mit diesem Lemma können wir γ_1 durch die $g(1,n)$ berechnen; es bleibt die Frage, wie wir diese durch die Koeffizienten des ursprünglichen Polynoms ausdrücken.

Die Koeffizienten von p sind ihrerseits die elementarsymmetrischen Polynome der ξ_i. Wir bezeichnen diese ebenfalls mit σ_k. Zusätzlich definieren wir die Potenzsummen über die Nullstellen ξ_i:

Definition 5.3.11. Die Potenzsumme s_k über den Nullstellen ξ_i von p mit $d = \deg(p)$ sei:

$$s_k := \sum_{i=1}^{d} \xi_i^k$$

Die s_k sind symmetrische Polynome und lassen sich daher durch die elementarsymmetrischen Polynome darstellen, und zwar durch die so genannten *Newton-Gleichungen*:

$$\begin{aligned}
s_k = &\quad \sigma_1 s_{k-1} - \cdots - (-1)^{k-1}\sigma_{k-1}s_1 - (-1)^k \cdot k \cdot \sigma_k && \text{für } 1 \le k \le d \\
s_k = &\quad \sigma_1 s_{k-1} - \cdots - (-1)^{d-1}\sigma_{d-1}s_{k-d+1} - (-1)^d \sigma_d s_{k-d} && \text{für } k > d
\end{aligned}$$

Die Newton-Gleichungen finden sich z.B. in [5]. Mit den s_k können wir die $g(1,n)$ schreiben als:

Lemma 5.3.12. *Sei p wie oben und $d = \deg(p)$. Dann ist*

$$g(1,n) = ds_n\left(1 + (-1)^n\right) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} s_{n-k}s_k$$

Insbesondere ist dieser Term für ungerade n gleich null.

Beweis.

$$\begin{aligned}
g(1,n) &= \sum_{(i,j)\in\mathbb{Z}_d^2} (\xi_i - \xi_j)^n \\
&= \sum_{(i,j)\in\mathbb{Z}_d^2} \left(\left(\xi_i^n + (-1)^n \xi_j^n\right) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} \xi_i^{n-k}\xi_j^k \right) \\
&= ds_n(1 + (-1)^n) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} \sum_{(i,j)\in\mathbb{Z}_d^2} \xi_i^{n-k}\xi_j^k \\
&= ds_n(1 + (-1)^n) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} \sum_{i=0}^{d-1}\xi_i^{n-k} \sum_{j=0}^{d-1}\xi_j^k \\
&= ds_n(1 + (-1)^n) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} s_{n-k}s_k
\end{aligned}$$

Für ungerade n gilt darüber hinaus:

$$\begin{aligned}
g(1,n) &= ds_n(1 + (-1)^n) + \sum_{k=1}^{n-1}(-1)^k \binom{n}{k} s_{n-k}s_k \\
&= ds_n(1-1) + \sum_{k=1}^{\frac{n-1}{2}}(-1)^k \binom{n}{k} s_{n-k}s_k + \sum_{k=\frac{n-1}{2}+1}^{n-1}(-1)^k \binom{n}{k} s_{n-k}s_k \\
&= \sum_{k=1}^{\frac{n-1}{2}}(-1)^k \binom{n}{k} s_{n-k}s_k + \sum_{k=1}^{\frac{n-1}{2}}(-1)^{n-k} \binom{n}{n-k} s_k s_{n-k} \\
&= \sum_{k=1}^{\frac{n-1}{2}}(-1)^k \binom{n}{k} s_{n-k}s_k - \sum_{k=1}^{\frac{n-1}{2}}(-1)^k \binom{n}{k} s_{n-k}s_k \\
&= 0
\end{aligned}$$

□

Fassen wir also kurz zusammen: Wir suchen ein Polynom p_-, das als Nullstellen die Differenzen der Nullstellen des ursprünglichen Polynoms p hat. Mit

Hilfe der Newton-Gleichungen können wir die Potenzsummen dieser Nullstellen durch die Koeffizienten von p beschreiben. Damit berechnen wir die $g(1,n)$ und daraus über die Rekursionsgleichungen in Lemma 5.3.12 die γ_1, die ihrerseits die Koeffizienten des gesuchten Polynoms p_- sind. Da p_- einen Faktor X^d hat, brauchen wir die letzten d Werte für γ_1 nicht zu berechnen; davor reichen uns die geraden Koeffizienten, d.h. insgesamt nur $\frac{n(n-1)}{2}$ Koeffizienten.

5.3.3 Das Summenpolynom

Für die Nullstellen im Quadratwurzelkörper wollen wir zusätzlich zu den Differenzen auch ein Polynom bestimmen, das die Summen der Nullstellen des Ursprungspolynoms beschreibt, um die letzte der Wurzeln zu eliminieren. Konkret wollen wir also für $f(\xi_1,\xi_2) = \xi_1 + \xi_2$ das Polynom $K_f(p,p)$ berechnen. Bei diesem Polynom werden nicht wieder wie beim Differenzenpolynom mehr als die Hälfte der Koeffizienten null, so dass es im ersten Moment so scheint, als müssten wir n^2 viele Koeffizienten berechnen. Tatsächlich zerfällt $K_f(p,p)$ allerdings in drei Faktoren; einer für die Nullstellen $2\xi_i$ für $i = j$ und zwei für die Nullstellen $\xi_i + \xi_j$ für $i < j$ und $i > j$. Die letzten beiden Faktoren sind identisch, und tatsächlich genügt es uns, einen von diesen zu berechnen. Wir wollen hier diesen entscheidenden Faktor mit p_+ bezeichnen. Formal ist also:

Definition 5.3.13. Sei R ein beliebiger Ring, $p \in R[X]$ und $\xi_0, \ldots, \xi_{n-1}$ die Nullstellen von p_j. Dann sei

$$p_+ := \prod_{(i<j)\in\mathbb{Z}_n^2} X - (\xi_i + \xi_j)$$

Der Grad von p_+ ist die Summe aller Zahlen zwischen 1 und $n-1$, d.h. $\frac{n(n-1)}{2}$. Wenn wir mit η_i wieder die Nullstellen von p_+ bezeichnen, bleiben die restlichen Definitionen genau wie im Abschnitt 5.3.1 über das allgemeine Kombinationspolynom. Mit Hilfe der rekursiven Beschreibung von g können wir also die Koeffizienten von p_+ durch die $g(1,n)$ beschreiben. Wie zuvor beim Differenzenpolynom stellt sich die Frage, wie wir diese wiederum durch die Koeffizienten von p beschreiben können. Setzen wir die Potenzsummen s_k wieder wie im Abschnitt 5.3.2 voraus, so wird dies Problem durch folgendes Lemma gelöst:

Lemma 5.3.14. *Sei p wie oben und $d = deg(p)$. Dann ist*

$$g(1,n) = \frac{2ds_n - 2^n s_n + \sum_{k=1}^{n-1} \binom{n}{k} s_k s_{n-k}}{2}$$

Beweis. Wir lösen zunächst die Summe $\sum_{(i,j)\in\mathbb{Z}_d^2}(\xi_i+\xi_j)^n$ auf. Es gilt

$$\begin{aligned}
\sum_{(i,j)\in\mathbb{Z}_d^2}(\xi_i+\xi_j)^n &= \sum_{(i,j)\in\mathbb{Z}_d^2}\left(\xi_i^n+\xi_j^n+\sum_{k=1}^{n-1}\binom{n}{k}\xi_i^k\xi_j^{n-k}\right)\\
&= \sum_{i=0}^{d-1}\sum_{j=0}^{d-1}\xi_i^n+\sum_{i=0}^{d-1}\sum_{j=0}^{d-1}\xi_j^n+\sum_{k=1}^{n-1}\binom{n}{k}\sum_{i=0}^{d-1}\xi_i^k\sum_{j=0}^{d-1}\xi_j^{n-k}\\
&= 2d\sum_{i=0}^{d-1}\xi_i^n+\sum_{k=1}^{n-1}\binom{n}{k}\sum_{i=0}^{d-1}\xi_i^k\sum_{j=0}^{d-1}\xi_j^{n-k}\\
&= 2ds_n+\sum_{k=1}^{n-1}\binom{n}{k}s_ks_{n-k}
\end{aligned}$$

Die Summe $\sum_{(i,j)\in\mathbb{Z}_d^2}(\xi_i+\xi_j)^n$ durchläuft alle i und j und lässt sich in drei Summen zerlegen, nämlich in zwei gleiche Summen, die die gemischten Terme beschreiben, und die Terme, in denen $i=j$ ist. Seien η_i die Nullstellen von p_+, dann können wir die $g(1,n)$ umgekehrt ausdrücken als:

$$\begin{aligned}
g(1,n) &= \sum_{i=1}^{deg(p_+)}\eta_i^n\\
&= \sum_{(i<j)\in\mathbb{Z}_d^2}(\xi_i+\xi_j)^n\\
&= \frac{\sum_{(i,j)\in\mathbb{Z}_d^2}(\xi_i+\xi_j)^n-\sum_{i=0}^{d-1}(\xi_i+\xi_i)^n}{2}\\
&= \frac{2ds_n-2^ns_n+\sum_{k=1}^{n-1}\binom{n}{k}s_ks_{n-k}}{2}
\end{aligned}$$

□

Parallel zum Differenzenpolynom nehmen wir auch hier die Division der Fakultät noch in die Rekursion mit hinein und definieren $\gamma_2=\frac{1}{n!}g(1,n)$. Auch hier gewinnen wir eine einfache Rekursion für γ_2:

Lemma 5.3.15. *Es gilt*

$$\gamma_2(n)=\frac{1}{n}(-1)^{n-1}g(1,n)+\frac{1}{n}\sum_{i=0}^{n-2}(-1)^i\gamma_2(n-1-i)g(1,1+i)$$

Beweis.

$$\begin{aligned}
\gamma_2(n) &= \frac{1}{n!}g(n,1) \\
&= \frac{1}{n!}\Bigg((n-1)!(-1)^{n-1}g(1,n) \\
&\qquad + \sum_{i=0}^{n-2}(-1)^i\frac{(n-1)!}{(n-1-i)!}g(n-1-i,1)g(1,1+i)\Bigg) \\
&= \frac{1}{n}(-1)^{n-1}g(1,n) + \frac{1}{n}\sum_{i=0}^{n-2}(-1)^i\frac{(n-1-i)!\gamma_2(n-1-i)}{(n-1-i)!}g(1,1+i) \\
&= \frac{1}{n}(-1)^{n-1}g(1,n) + \frac{1}{n}\sum_{i=0}^{n-2}(-1)^i\gamma_2(n-1-i)g(1,1+i)
\end{aligned}$$

Die oberste Einsetzung ist nach Lemma 5.3.7. □

Wir können also bei der Addition ebenfalls mit Hilfe der s_k die Koeffizienten γ_2 von p_+ berechnen. Für die praktische Implementierung benötigen wir beide Polynome, weshalb wir hier die beiden Ergebnisses dieses und des vorherigen Abschnittes im nun folgenden Abschnitt zu einer Methode zusammenführen.

5.3.4 Implementierung

Fassen wir die gesamte Prozedur noch einmal zusammen: Mit Hilfe der Koeffizienten von p berechnen wir zunächst mit den Newton-Gleichungen die Potenzsummen s_k. Damit drücken wir für die Addition und die Subtraktion die jeweiligen $g(1,n)$ aus; zur besseren Unterscheidung nennen wir diese beiden hier β_1 (für die Subtraktion) und β_2 (für die Addition). Mit diesen Werten können wir dann γ_1 und γ_2 berechnen und erhalten damit die Koeffizienten für die beiden gesuchten Polynome.

Der Pseudocode für den Algorithmus zur Berechnung der γ findet sich in Abbildung 5.3.

Lemma 5.3.16. *Der Algorithmus berechnet γ_1 und γ_2 korrekt und benötigt $O(d_1^4)$ arithmetische Operationen im Grundkörper, wenn d_1 der Grad des Eingabepolynoms ist.*

Beweis. Die Korrektheit folgt aus den Sätzen dieses Kapitels.

Die erste FOR-Schleife durchläuft d_1 mal einen Berechnungsschritt mit einer Summe mit d_1 Summanden, benötigt also $O(d_1^2)$ viele Schritte. Die zweite FOR-Schleife durchläuft weniger als $2d_2$ viele Schritte mit einer Summe von ebenfalls d_1 Summanden; da $d_2 = \frac{d_1(d_1-1)}{2} \in O(d_1^2)$ ist, braucht diese Schleife also $O(d_1^3)$ viele Schritte. Die dritte und vierte Schleife durchlaufen jeweils d_2 viele Schritte, wobei die Summe im Inneren maximal $2d_2$ viele Summanden hat, d.h. diese beiden Schleifen brauchen beide $O(d_1^4)$ viele Schritte.

□

Require: Polynom $p = X^{d_1} + \sum_{i=1}^{d_1} \sigma_i X^{d_1-i}$
  $d_1 := \deg(p)$
  $d_2 := \frac{d_1(d_1-1)}{2}$
  for $k = 1$ **to** d_1 **do**
    $s_k := \sigma_1 s_{k-1} - \cdots - (-1)^{k-1}\sigma_{k-1}s_1 - (-1)^k \cdot k \cdot \sigma_k$
  end for
  for $k = d_1 + 1$ **to** $2d_2$ **do**
    $s_k := \sigma_1 s_{k-1} - \cdots - (-1)^{d_1-1}\sigma_{d_1-1}s_{k-d_1+1} - (-1)^{d_1}\sigma_{d_1}s_{k-d_1}$
  end for
  for $n = 1$ **to** d_2 **do**
    $\beta_1(2n-1) := 0$
    $\beta_1(2n) := 2d_1 s_{2n} + \sum_{k=1}^{2n-1}(-1)^k \binom{2n}{k} s_{2n-k}s_k$
    $\beta_2(n) := \frac{2d_1 s_n - 2^n s_n + \sum_{k=1}^{n-1}\binom{n}{k}s_k s_{n-k}}{2}$
  end for
  for $k = 1$ **to** d_2 **do**
    $\gamma_1(2k-1) := 0$
    $\gamma_1(2k) := -\frac{1}{2k}\left(\beta_1(2k) + \sum_{i=0}^{2k-2}\gamma_1(2k-1-i)\beta_1(1+i)\right)$
    $\gamma_2(k) := \frac{1}{k}\left((-1)^{k-1}\beta_2(k) + \sum_{i=0}^{k-2}(-1)^i\gamma_2(k-1-i)\beta_2(1+i)\right)$
  end for
  return (γ_1, γ_2)

Abbildung 5.3: Algorithmus für die Berechnung der Kombinationspolynome Die Listen γ_1 und γ_2 stellen jeweils die Koeffizienten des Summen- und des Differenzenpolynoms dar.

Die $g(k,1)$ als Polynome über $g(1,n)$ und im Spezialfall auch über σ_i sind konstant und können auch im Voraus berechnet werden.

Mit diesem Algorithmus und der Faktorisierung gelingt es uns jetzt, das Finden von Nullstellen im Euklidischen Körper auf immer einfachere Probleme zurückzuführen, bis es im linearen Fall schließlich trivial ist. In den letzten beiden Abschnitten dieses Kapitels werden wir jetzt noch einige lokale Vereinfachungen machen.

5.4 Faktorisierung des Summen- und Differenzenpolynoms

Wir verfügen jetzt also über einen Algorithmus, der effizient das Summen- bzw. Differenzenpolynom aus dem Minimalpolynom p von ξ errechnet. Wir wissen, dass für $\xi = \xi_1 + \sqrt{\xi_2}$ mit $\deg(\xi_1) \leq \frac{1}{2}\deg(\xi)$ und $\deg(\xi_2) = \frac{1}{2}\deg(\xi)$ das Minimalpolynom von ξ_1 als Faktor des Summenpolynoms und das Minimalpolynom von ξ_2 als Faktor des Differenzenpolynoms nach Substitution von X^2 durch X auftritt. Es genügt also, die beiden Polynome zu faktorisieren, mit allen Faktoren von höchstens halbem Grad rekursiv fortzufahren, und am Ende alle möglichen Ergebnissen für ξ_1 und ξ_2 zu kombinieren und zu testen. Tatsächlich ist dies für manche Beispiele nötig; ist ξ aber vollständig, so können wir den richtigen Faktor in beiden Polynomen eindeutig identifizieren und können dabei sogar die Faktorisierung selbst vereinfachen. Auch für nicht vollständige Elemente lassen sich in günstigen Fällen einige Schritte sparen.

Um dies zu zeigen, betrachten wir zunächst das Summenpolynom näher.

5.4.1 Eigenschaften des Summenpolynoms

Sei $\xi \in \mathbb{E}_\mathbb{K}$ mit Grad 2^d. Das Summenpolynom des Minimalpolynoms dieses Elementes wollen wir wieder mit p_+ bezeichnen, d.h. $p_+ := \prod_{(I<J)\in\left(\mathbb{Z}_2^d\right)^2} X - \left(\xi^{(I)} + \xi^{(J)}\right)$, worin $<$ die lexikographische Ordnung von Tupeln aus $\mathbb{Z}^d$ beschreibt und $\xi^{(0,\ldots,0)}, \ldots, \xi^{(1,\ldots,1)}$ wie zuvor die Konjugierten von ξ darstellt. Es gilt folgendes Lemma:

Lemma 5.4.1. *p_+ zerfällt in Faktoren vom Grad $2^{d-1}, 2^d, \ldots, 2^{2d-2}$.*

Beweis. Der Einfachheit halber betrachten wir das Summenpolynom ohne Ordnung der Summanden, also $p_+^2 := \prod_{(I\neq J)\in\left(\mathbb{Z}_2^d\right)^2} X - \left(\xi^{(I)} + \xi^{(J)}\right)$. Hier müssen wir uns keine Sorge um die Reihenfolge der Summanden machen. Da dieses Polynom das Quadrat von p_+ ist, kommt jeder in p vorkommende Faktor in gerader Potenz vor und ist ebenfalls ein Faktor von p_+.

Wir definieren jetzt Polynome p_k über deren Nullstellen, die jeweils eine Teilmenge der Nullstellen von p sind. Wir zeigen dann im Anschluss, dass die Grade dieser Teiler p_k wie in der Behauptung sind und die Polynome Koeffizienten aus $\mathbb{K}$ haben. Diese Polynome p_k für $1 \leq k \leq d$ sind:

$$p_k := \prod_{I\in\mathbb{Z}_2^d, J\in\mathbb{Z}_2^{d-k}} \left(X - \left(\xi^{(I)} + \xi^{(i_1,\ldots,i_{k-1},1-i_k,j_1,\ldots,j_{d-k})}\right)\right)$$

p_k durchläuft also in dem ersten der beiden Summanden von der Summe $\xi^{(I)} + \xi^{(i_1,\ldots,i_{k-1},1-i_k,j_1,\ldots,j_{d-k})}$ alle Konjugierten. Im zweiten Summanden sind alle Vorzeichen bis zur k-ten Wurzel wie im ersten Summanden, das Vorzeichen vor der k-ten Wurzel ist vertauscht, und die restlichen Vorzeichen durchlaufen alle Paarungen. Der Faktor p_d ist das Polynom, das wir schon vom Satz 5.2.1 her kennen.

Der Grad von p_k ist 2^{2d-k}. Alle Nullstellen werden doppelt durchlaufen, d.h. p_k ist ein Quadrat eines Faktors von p_+ mit Grad 2^{2d-k-1}. Der Index k läuft von 1 bis d, d.h. die Faktoren haben die geforderten Grade; es bleibt nur zu zeigen, dass $p_k \in \mathbb{K}[X]$ gilt.

In der Summe der beiden Konjugierten können wir die jeweils ersten $k-1$ Wurzeln zusammenfassen, denn die w_j für $j \leq k$ sind in beiden Summanden gleich. w_k selbst ist auch in beiden Summanden gleich, und da vor beiden k-ten Wurzeln verschiedene Vorzeichen stehen, fällt $\sqrt{w_k}$ weg. Bei den restlichen w_j für $j > k$ liegen verschiedene Konjugierte vor, d.h. im Allgemeinen bleiben diese Wurzeln stehen. Die Summe hat also folgende Form:

$$\xi^{(I)} + \xi^{(i_1,\ldots,i_{k-1},1-i_k,j_1,\ldots,j_{k-1})} = \sum_{m=0}^{k-1} (-1)^{i_m} 2\sqrt{w_m^{(i_1,\ldots,i_{m-1})}} +$$

$$\sum_{m=k+1}^{d} (-1)^{i_m} \sqrt{w_m^{(i_1,\ldots,i_{m-1})}} + (-1)^{j_{m-k}} \sqrt{w_m^{(i_1,\ldots,i_{k-1},1-i_k,j_1,\ldots,j_{m-k})}}$$

Man überzeugt sich leicht, dass alle Wurzeln mit allen Vorzeichen durchlaufen werden. Nach Korollar 3.2.3 ist $p_k \in \mathbb{K}[X]$, womit die Aussage gezeigt ist. □

Bemerkung 5.4.2. Es gibt einen deutlich einfacheren Beweis für dieses Lemma; nach Konstruktion sind alle Nullstellen von p_+ durch Wurzeln auflösbar. Also muss p_+ in Faktoren mit Zweierpotenzen als Grad zerfallen. Der Grad von p_+ ist $2^{d-1} \cdot (2^d - 1) = 2^{d-1} + \cdots + 2^{2d-2}$. Bedenkt man, dass zwei gleiche Zweierpotenzen zur nächstgrößeren Zweierpotenz zusammengefasst werden können, ist diese Zerlegung eindeutig, und der Satz ist ebenfalls bewiesen.

Leider sind die Faktoren im Allgemeinen nicht unbedingt irreduzibel, wie folgendes Beispiel zeigt:

Beispiel 5.4.3. Sei $p = X^4 - 4X^3 - 4X^2 + 4X + 1$. Man überzeugt sich, dass p Minimalpolynom von $\xi = 1 + \sqrt{3} + \sqrt{2 + \sqrt{3}}$ ist. Das Summenpolynom von p ist $p_+ = X^6 - 12X^5 + 40X^4 - 132X^2 + 16X + 32$. Dieses Polynom zerfällt in drei Faktoren vom Grad zwei, nämlich $p_+ = (X^2 - 4X - 2)(X^2 - 4X - 8)(X^2 - 4X + 2)$, was bedeutet, dass der nach Lemma 5.4.1 existierende Faktor vom Grad 4 hier in zwei Faktoren zerfallen ist.

Wir stellen jedoch fest: Das Produkt der Terme $\sqrt{2 + \sqrt{3}}$ und $\sqrt{2 - \sqrt{3}}$ ist $\sqrt{4 - 3} = 1$. Das heißt, dass $\mathbb{Q}(\sqrt{3}, \sqrt{2 + \sqrt{3}}) = \mathbb{Q}(\sqrt{3}, \sqrt{2 - \sqrt{3}})$ ist, und also ξ nicht vollständig ist.

Tatsächlich wird der folgende Satz zeigen, dass bei vollständigen Elementen alle in Lemma 5.4.1 genannten Faktoren bis auf den kleinsten Faktor irreduzibel sind. Das hilft uns auf zweierlei Weisen: Erstens können wir uns bei der Faktorisierung von p_+ auf Faktoren vom Grad $\frac{1}{2}\deg(\xi)$ beschränken, was also bei irreduziblen größeren Faktoren nur ein einziger Faktor ist, und zweitens müssen wir den Algorithmus nicht rekursiv für viele kleinere Faktoren aufrufen, sondern können uns auf den einen existierenden von hinreichend kleinem Grad beschränken.

Dafür zeigen wir zunächst folgendes Lemma:

Satz 5.4.4. *Sei ξ vollständig und $\deg(\xi) = 2^d$. Seien $\xi^{(I)}$ und $\xi^{(J)}$ zwei Konjugierte von ξ, wobei I und J entweder ganz identisch sind oder sich in einem Index kleiner d (also vor dem letzten Index) unterscheiden. Dann sind I und J bis auf Vertauschung der beiden durch die Summe der Konjugierten $\xi^{(I)} + \xi^{(J)}$ bereits eindeutig festgelegt.*

Beweis. Sind I und J identisch, ist die Summe der beiden Konjugierten genau $2\xi^{(I)}$ und damit ebenfalls vollständig; eine Darstellung über die Summe zweier anderer Konjugierter würde sich aber spätestens in der letzten Wurzel unterscheiden, d.h. es gäbe zwei Darstellungen von $2\xi^{(I)}$, was ein Widerspruch zur Vollständigkeit ist.

Seien I und J also verschieden und $k < d$ der kleinste Index mit $i_k \neq j_k$. Die Darstellung der Summe muss nicht eindeutig sein, aber da die Konjugierten aller Wurzeln unabhängig sind (wegen der Vollständigkeit von ξ), wissen wir, dass es eine Darstellung der Summe gibt, in der dieselben Wurzeln wie in $\xi^{(I)}$ und $\xi^{(J)}$ vorkommen. Wir betrachten diese Darstellung.

Bis zu dieser Stelle k haben alle Wurzeln der Summe dieselben Vorzeichen wie $\xi^{(I)}$, die k-te Wurzel selbst hebt sich weg. Nehmen wir o.B.d.A. an, $i_k = 0$ und $j_k = 1$. Nach Voraussetzung ist $k < d$; da in der obersten Wurzel w_d nach Lemma 3.3.3 alle Wurzeln vorkommen, kommen in der Summe noch zwei verschiedene Konjugierte von $\sqrt{w_d}$ vor, jeweils von einem der Summanden eine. Welche Konjugierte zu welchem Summanden gehört, können wir am Vorzeichen vor $\sqrt{w_k}$ erkennen. Die Vorzeichen aller Wurzeln zwischen k und d können wir dann an den Vorzeichen in der entsprechenden Konjugierten in $\sqrt{w_k}$ ablesen.

Also sind durch die Summe bereits die Vorzeichen vor den Wurzeln der beiden Summanden bis auf Vertauschung eindeutig bestimmt. □

Wir haben diesen Satz bewiesen, um eine Beziehung zwischen den Zwischenkörpern, die eine Summe zweier Konjugierter enthalten, und ihrer Galoisgruppen herzustellen. Denn wegen Satz 5.4.4 muss jede Galoisgruppe, die mit einem Zwischenkörper mit $\xi_i + \xi_j$ korrespondiert, i und j als Paar fest lassen. Denn würde ein Element der Galoisgruppe i und j nicht als Paar fest lassen, so würde dieses Element aufgefasst als Körperautomorphismus wegen des letzten Satzes die Summe nicht konstant lassen, und die Galoisgruppe würde nicht mit einem Zwischenkörper korrespondieren, der diese Summe enthält. Dies ermöglicht uns den folgenden Satz:

Satz 5.4.5. *Sei ξ vollständig und p_+ das Summenpolynom des Minimalpolynoms von ξ. Dann zerfällt p_+ in einen möglicherweise reduziblen Faktor vom Grad 2^{d-1} und weitere irreduzible Faktoren vom Grad $2^d, \ldots, 2^{2d-2}$.*

Beweis. Dass p_+ in besagte Faktoren zerfällt, ist in Lemma 5.4.1 gezeigt. Es muss jetzt nur noch gezeigt werden, dass alle Faktoren bis auf den kleinsten irreduzibel sind, wenn ξ vollständig ist. Dafür nutzen wir den Beweis von Lemma 5.4.1, der uns die Nullstellen von p_k gibt. Im Beweis von Lemma 5.4.1 war p_k ein quadratischer Faktor von p_+^2; der Einfachheit halber bezeichnen wir hier mit p_k den einfachen Faktor von p_+.

Wir beweisen die Irreduzibilität von p_k für $k < d$, indem wir zeigen, dass p_k Minimalpolynom einer seiner Nullstellen ist.

Jedes p_k hat die Nullstelle

$$\left(\xi_{\underbrace{(0,\ldots,0)}_{d}} + \xi_{\underbrace{(0,\ldots,0,1}_{k},\underbrace{0,\ldots,0)}_{d-k}} \right)$$

Zählen wir die Konjugierten lexikographisch auf, wie wir das für die Notation der Galoisgruppe getan haben, so entspricht dieses der Summe $\xi_1 + \xi_{2^{d-k}+1}$. Nach Satz 5.4.4 ist die Summe nur über diese beiden Summanden darstellbar, d.h. die Galoisgruppe desjenigen Zwischenkörpers, der $\xi_1+\xi_{2^{d-k}+1}$ enthält, muss die Elemente 1 und $2^{d-k}+1$ fest lassen oder gegeneinander vertauschen. Wir suchen diejenige Untergruppe von Q_d, die dies erfüllt.

Wir suchen dafür zunächst alle Elemente σ von Q_d, für die $\sigma(1) = 1$ und $\sigma(2^{d-k}+1) = 2^{d-k}+1$ gilt. Diese sind gegeben durch folgende Gruppe:

$$\begin{aligned} G &= (Q_0 \odot Q_0 \odot Q_1 \odot \cdots \odot Q_{d-k-1}) \\ &\odot (Q_0 \odot Q_0 \odot Q_1 \odot \cdots \odot Q_{d-k-1}) \\ &\odot (Q_{d-k+1} \odot \cdots \odot Q_{d-1}) \end{aligned}$$

Der obere Block reicht von der Stelle 1 bis zur Stelle 2^{d-k} und stellt dort dieselbe Gruppe dar, die wir schon in der Definition 4.3.2 gesehen haben und die hier das erste Element fest lässt. Ein genauso großer Block folgt, der das Element $2^{d-k}+1$ fest lässt. Die restlichen Gruppen erlauben alle Vertauschungen von Q_d, die nicht mit den ersten beiden Blöcken interagieren.

Die Größe von G ist das Produkt der Größen aller durch $\odot$ verknüpften Gruppen, also

$$\begin{aligned}
|G| &= \left(\prod_{i=1}^{d-k-1} 2^{2^i-1}\right)^2 \cdot \prod_{i=d-k+1}^{d-1} 2^{2^i-1} \\
&= 2^{2\left(-(d-k-1)+\sum_{i=1}^{d-k-1} 2^i\right)} \cdot 2^{(d-k+1-d-1)+\sum_{i=d-k+1}^{d-1} 2^i} \\
&= 2^{(k-2d+1)+\left(\sum_{i=1}^{d-k-1} 2^i\right)+\left(\sum_{i=1}^{d-1} 2^i\right)-2^{d-k}} \\
&= 2^{(k-2d+1)+2^{d-k}-1+2^d-1-2^{d-k}} \\
&= 2^{2^d+k-2d-1}
\end{aligned}$$

Die Gruppenelemente, die eins und $2^{d-k}+1$ gegeneinander vertauschen, erzeugen wir einfach aus derselben Gruppe und der Vertauschung der ersten beiden Blöcke, d.h. die Größe der minimalen Galoisuntergruppe der Summe $\xi_1 + \xi_{2^{d-k}+1}$ ist das Doppelte der Größe von $|G|$, also 2^{2^d+k-2d}. Dies sind genau die Elemente, die unsere Anforderung erfüllen, die Summe $\xi_1 + \xi_{2^{d-k}+1}$ unverändert zu lassen, also die zu $\mathbb{K}(\xi_1+\xi_{2^{d-k}+1})$ korrespondierende Galoisgruppe. Den Grad dieser Körpererweiterung und damit den Grad des Minimalpolynoms von $\xi_1 + \xi_{2^{d-k}+1}$ erhalten wir, wenn wir die Gesamtgröße der Galoisgruppe Q_d durch diese Größe teilen. Dies ergibt dann 2^{2d-k-1}. Wie im Beweis zu 5.4.1 festgestellt, ist dies genau der Grad von p_k; also hat p_k den selben Grad wie das Minimalpolynom einer seiner Nullstellen und ist damit irreduzibel. □

Nach Satz 5.2.1 wissen wir, dass der kleinste Faktor zwar reduzibel sein kann, dann aber eine Potenz des eigentlich gesuchten Minimalpolynoms des Ausdrucks vor der letzten Wurzel in ξ ist. Ist also ξ vollständig, so finden wir den für uns interessanten Faktor von p_+ schon beim Suchen von quadratischen Faktoren (da alle anderen Faktoren irreduzibel und von verschiedenen Graden sind, ist der Faktor der einzige mehrfach auftretende), oder aber wir müssen nur nach dem einzigen existierenden Faktor vom Grad 2^{d-1} suchen. Faktorisieren wir mit Hilfe einer Abbildung in einen endlichen Körper, so können wir uns dort durch die Beschränkung auf Faktoren vom Grad $\leq 2^{d-1}$ viel Arbeit erleichtern. Trotz der Ergebnisse dieses Abschnittes müssen wir auch nach eventuellen kleineren Faktoren suchen, da wir nicht im Voraus wissen, ob ξ vollständig ist; da aber die meisten Elemente vollständig sind, werden wir im Allgemeinen keine kleineren Faktoren finden.

Ähnliches (und sogar noch mehr) gilt auch für das Differenzenpolynom, mit dem wir uns im nächsten Abschnitt beschäftigen werden.

5.4.2 Eigenschaften des Differenzenpolynoms

Sei $\xi \in \mathbb{E}_\mathbb{K}$ mit Grad 2^d. Das Differenzenpolynom des Minimalpolynoms dieses Elementes wollen wir wieder parallel zum Summenpolynom mit p_- bezeichnen, d.h. $p_- := \prod_{(I\neq J)\in\left(\mathbb{Z}_2^d\right)^2} X-(\xi^{(I)}-\xi^{(J)})$, worin $\xi^{(0,\ldots,0)},\ldots,\xi^{(1,\ldots,1)}$ wie zuvor die Konjugierten von ξ darstellen. Wir erinnern uns, dass das Differenzenpolynom den Grad $2^d(2^d-1)$ hat (also doppelt so groß ist wie das Summenpolynom), aber alle Koeffizienten bei ungeraden Exponenten null sind.

Es gilt auch hier ein allgemein gültiges Zerfällungslemma, nämlich:

Lemma 5.4.6. *p_- zerfällt in Faktoren vom Grad $2^d, 2^{d+1}, \ldots, 2^{2d-1}$.*

Beweis. Wie beschränken uns hier auf den einfachen Beweis ähnlich wie in der Bemerkung 5.4.2. Nach Konstruktion sind alle Nullstellen von p_- durch Wurzeln auflösbar. Also muss p_- in Faktoren mit Zweierpotenzen als Grad zerfallen. Der Grad von p_- ist $2^d \cdot (2^d - 1) = 2^d + \cdots + 2^{2d-1}$. Bedenkt man, dass zwei gleiche Zweierpotenzen zur Nächsthöheren zusammengefasst werden können, ist diese Zerlegung eindeutig, und der Satz ist bewiesen. □

Auch hier können wir durch Beschränkung auf vollständige Elemente feststellen, dass die oben genannten Faktoren irreduzibel sind; und in diesem Fall sogar alle, auch der Faktor mit dem kleinsten Grad. Dies liegt besonders daran, dass wir die Eindeutigkeit der Differenz zweier Konjugierter etwas stärker fassen können als die Eindeutigkeit der Summe von Lemma 5.4.4:

Satz 5.4.7. *Sei ξ vollständig und $2^d = \deg(\xi)$. Seien $\xi^{(I)}$ und $\xi^{(J)}$zwei Konjugierte von ξ, wobei I und J nicht identisch sind. Dann sind I und J durch die Differenz der Konjugierten $\xi^{(I)} - \xi^{(J)}$ bereits eindeutig festgelegt.*

Beweis. Sei k der kleinste Index mit $i_k \neq j_k$. Die Darstellung der Differenz muss nicht eindeutig sein, aber da die Konjugierten aller Wurzeln wegen der Vollständigkeit von ξ unabhängig sind, wissen wir, dass es eine Darstellung der Summe gibt, in der dieselben Wurzeln wie in $\xi^{(I)}$ und $\xi^{(J)}$ vorkommen. Wir betrachten diese Darstellung.

Bis zur Stelle k heben sich alle Wurzeln weg. An der Stelle k trennen wir die beiden Wurzeln wieder voneinander. Wir können feststellen, ob i_k oder j_k eins ist; im ersteren Fall hat das Vorzeichen der Differenz gegenüber ξ gewechselt, im letzteren nicht. Da alle Wurzeln und insbesondere $\sqrt{w_k}$ in der obersten Wurzel $\sqrt{w_d}$ nach Lemma 3.3.3 wieder vorkommt, kommen zwei Konjugierte von $\sqrt{w_d}$ in der Differenz vor, und wir können die Vorzeichen der restlichen Wurzeln wieder zuordnen. Die Vorzeichen vor den Wurzeln von $\xi^{(I)}$ und $\xi^{(J)}$ sind dadurch eindeutig bestimmt, und damit auch die Konjugierten selbst. □

Wie beim Summenpolynom können wir mit diesem Lemma die Differenzen zu den Galoisgruppen des ersten Zwischenkörpers, der die Differenz enthält, zuordnen. Auf diese Weise gelingt wieder eine Aussage über den Grad des Elements und damit auch des Faktors von p_-, in dem die Differenz vorkommt:

Satz 5.4.8. *Sei ξ vollständig und p_- das Differenzenpolynom des Minimalpolynoms von ξ. Dann zerfällt p_- in irreduzible Faktoren vom Grad $2^d, \ldots, 2^{2d-1}$.*

Beweis. Die allgemeine Zerfällung ist bereits in Lemma 5.4.6 gezeigt. Es muss jetzt nur noch gezeigt werden, dass alle Faktoren irreduzibel sind, wenn ξ vollständig ist. Parallel zum Beweis von Lemma 5.4.1 stellen wir fest, dass

$$p_k := \prod_{I \in \mathbb{Z}_2^d, J \in \mathbb{Z}_2^{d-k}} \left(X - \left(\xi^{(I)} - \xi^{(i_1, \ldots, i_{k-1}, 1-i_k, j_1, \ldots, j_{d-k})} \right) \right)$$

ein Faktor von p_- vom Grad 2^{2d-k} ist und eine Nullstelle

$$\left(\xi^{\underbrace{(0,\ldots,0)}_{d}} - \xi^{\underbrace{(0,\ldots,0,1}_{k},\underbrace{0,\ldots,0)}_{d-k}}\right)$$

besitzt. In lexikographischer Aufzählung der Konjugierten entspricht dies $\xi_1 - \xi_{2^{d-k}+1}$. Wir zeigen, dass p_k Minimalpolynom dieser Nullstelle ist.

Nach Satz 5.4.7 ist keine andere Differenz von Nullstellen dieser Differenz gleich, also muss die Galoisuntergruppe jedes Zwischenkörpers, der $\xi_1 - \xi_{2^{d-d}+1}$ enthält, diese Elemente einzeln fest lassen (einzeln im Gegensatz zum Summenpolynom, wo die beiden Elemente auch gegeneinander vertauscht werden durften; dies würde hier zu einer Änderung des Vorzeichens führen). Die Galoisgruppe, die diese Elemente fest lässt, ist das G aus dem Beweis zu Satz 5.4.5. Den Grad des korrespondierenden kleinsten Zwischenkörpers erhalten wir durch Division der Ordnung von Q_d durch die Ordnung von G, also $\frac{2^{2^d-1}}{2^{2^d+k-2d-1}} = 2^{2d-k}$. Dies entspricht aber dem Grad von p_k, also ist p_k Minimalpolynom der Differenz und damit irreduzibel. □

Also zerfällt auch das Differenzenpolynom bei vollständigen Elementen in Faktoren aufsteigender Zweierpotenzen, wobei wir diesmal sogar feststellen können, dass auch der kleinste Faktor irreduzibel ist. Dies ist auch einleuchtend, bedenkt man, dass es nach Lemma 3.3.6 eine Darstellung geben muss, in der die letzte Wurzel vollen Grad hat und vollständige Elemente eine eindeutige Darstellung haben (und mithin in dieser Darstellung die letzte Wurzel vollen Grad hat).

In der Praxis müssen wir, ähnlich wie beim Summenpolynom, trotzdem auch nach anderen Faktoren vom Grad 2^d des Differenzenpolynoms suchen (bzw. vom Grad 2^{d-1}, wenn wir X^2 durch X substituiert haben). Im Allgemeinen (nämlich bei vollständigen Elementen) werden wir aber keine solchen Faktoren antreffen. Die Suche nach Faktoren vom Grad kleiner als 2^{d-1} ist nicht nötig, da ξ nach Lemma 3.3.6 in jedem Fall - vollständig oder nicht - auch eine Darstellung besitzt, in der ξ_2 vollen Grad hat.

Dieser Abschnitt hat uns gezeigt, dass wir im Allgemeine nach der Faktorisierung des Summen- und Differenzenpolynoms bei beiden nur jeweils einen Kandidaten haben, an dem wir fortsetzen müssen, und dieser eine Kandidat dann auch relativ leicht zu finden ist.

5.5 Einfache Sonderfälle

Bei einigen speziellen Problemen ist das Finden von Nullstellen in $\mathbb{E}_\mathbb{K}$ deutlich einfacher. Da dies naturgemäß eher die Probleme kleineren Grades betrifft und wir in der Rekursion in jedem Fall früher oder später auch solche Polynome antreffen, ist es von besonderer Bedeutung, Sonderfälle effizient lösen zu können. Daher stellen wir in diesem Abschnitt ein paar Sonderfälle speziell heraus und betrachten Lösungswege für diese Fälle.

Require: $p = aX^2 + bX + c$
  if $a == 0$ **then**
    return $\{\frac{-c}{b}\}$
  else
    return $\{\frac{-b \pm \sqrt{b^2-4ac}}{2a}\}$
  end if

Abbildung 5.4: Algorithmus für Polynome vom Grad ≤ 2

Require: p mit nur geraden Exponenten
  $\widetilde{p} := p\left(\sqrt{X}\right)$
  $\xi :=$ **löse**$(\widetilde{p})$
  return $\{\sqrt{\xi}\}$

Abbildung 5.5: Algorithmus für Polynome mit nur geraden Exponenten
löse *stellt die allgemeine Lösungsroutine dar, die weiter unten vorgestellt wird.*

5.5.1 Lineare und quadratische Polynome

Ein ganz einfacher Fall ist natürlich, wenn p von vornherein nur Grad eins hat. In diesem Fall hat $p = X - a$ natürlich eine Nullstelle in $\mathbb{E}_\mathbb{K}$, nämlich a. Dieser Fall ist sozusagen der Rekursionsanfang des allgemeinen Problems.

Auch bei quadratischen Polynomen haben wir wegen der klassischen Formel für quadratische Polynome kein Problem, die stets vorhandenen Lösungen in $\mathbb{E}_\mathbb{K}$ exakt anzugeben; Für $p = X^2 + aX + b$ sind die beiden Lösungen für $p(X) = 0$ durch $X = -\frac{a}{2} \pm \sqrt{\frac{a^2-4b}{4}}$ gegeben. Sollte p also vom Grad kleiner gleich zwei sein, können wir diese Lösungen direkt zurückgeben.

5.5.2 Polynome mit nur geraden Exponenten

Falls in der durch den Darstellungssatz gegebenen Darstellung von ξ die Faktoren vor allen Wurzeln außer der letzten null sind, d.h. ξ die Form $\sqrt{\xi_2}$ hat, dann ist das Minimalpolynom von ξ das Minimalpolynom von ξ_2, wenn man X^2 durch X substituiert. In diesem Fall wäre das Summenpolynom konstant null, während der interessante Faktor des Differenzenpolynoms wieder genau p ergeben würde. Wir können also von vornherein X^2 durch X substituieren und mit dem einfacheren Polynom fortfahren, und die Quadratwurzel außen um ξ_2 später wieder hinzufügen.

Dieser Fall ist natürlich nicht allzu häufig; andererseits ist der Test auf diesen Fall extrem simpel und spart gegebenenfalls mit dem ersten Rekursionsschritt den größten Teil der Arbeit ein.

Require: $p = X^{2^d} + a_1 X^{2^d-1} + \cdots + a_{2^d}$
$\widetilde{p} := p\left(X - \frac{a_1}{2^d}\right)$
$\xi :=$ **löse**$(\widetilde{p})$
return $\{\xi - \frac{a_1}{2^d}\}$

Abbildung 5.6: Algorithmus zur Elimination des zweithöchsten Koeffizienten

5.5.3 Elimination des konstanten Anteils

In der durch den Darstellungssatz gegebenen Darstellung von ξ existiert auch ein Summand k_0, der vor allen Wurzeln steht. Wir stellen fest:

Lemma 5.5.1. *Sei $\xi = \frac{1}{n}\left(k_0 + \sum_{i=1}^{d} k_i\sqrt{w_i}\right)$ und p das Minimalpolynom von ξ. Dann ist der zweithöchste Koeffizient von p genau $-\frac{2^d k_0}{n}$.*

Beweis. Nach Lemma 3.2.2 sind die Nullstellen des Minimalpolynoms genau alle Konjugierten von ξ. Also ist der zweithöchste Koeffizient die negative Summe aller Konjugierten. Dabei heben sich die Wurzeln gegenseitig weg, d.h. wir haben nur $\frac{k_0}{n}$ als Beitrag jeder Konjugierten zur Summe. Da es genau 2^d Konjugierte gibt, ist der zweithöchste Koeffizient also $-\frac{2^d k_0}{n}$. □

Auf diese Weise können wir den konstanten Summanden aus dem Minimalpolynom ablesen. Substituieren wir X in p durch $\left(X + \frac{k_0}{n}\right)$, so verschieben wir die Nullstellen um diesen Term; wir erhalten ein neues Polynom, das $\frac{1}{n}\left(\sum_{i=1}^{d} k_i\sqrt{w_i}\right)$ als Nullstellen hat, also immer noch ein Element vom selben Grad, aber ohne konstanten Summanden.

Diese Operation bringt uns zwei Vorteile: Erstens ist der zweithöchste Koeffizient des Polynoms nach der Verschiebung null, was die Berechnungen in den anderen Algorithmen vereinfacht. Zum Zweiten ist es natürlich möglich, dass alle anderen Vorfaktoren vor den Wurzeln bis auf die höchste Wurzel null sind, und wir daher jetzt die in Unterabschnitt 5.5.2 beschriebene Vereinfachung anwenden können.

Allein die Iteration dieser ersten drei Sonderfälle kann einige Probleminstanzen von vornherein ohne viel Mühe lösen. Betrachten wir dazu folgendes Beispiel.

Beispiel 5.5.2. Sei

$$p = X^8 + 16X^7 + 100X^6 + 304X^5 + 450X^4 + 272X^3 + 28X^2 - 16X - 4$$

Der zweithöchste Koeffizient ist nicht null. Wir teilen diesen Koeffizienten durch den Grad des Polynoms und bringen auf diese Weise in Erfahrung, dass $\frac{k_0}{n} = -2$ ist. Substituieren wir X durch $X - 2$, so erhalten wir

$$X^8 - 12X^6 + 50X^4 - 84X^2 + 44$$

In diesem Polynom sind alle Exponenten gerade, d.h. die Nullstellen sind die Wurzeln der Nullstellen von $X^4 - 12X^3 + 50X^2 - 84X + 44$. Hier haben wir wieder einen zweithöchsten Koeffizienten ungleich null; teilen wir diesen durch vier, so stellen wir fest, dass der konstante Anteil der Nullstellen dieses Polynoms drei ist. Wir substituieren wieder X durch $X + 3$ und erhalten dann

$$X^4 - 4X^2 - 1$$

Hier können wir wieder X^2 durch X ersetzen und erhalten schließlich die quadratische Gleichung $X^2 - 4X - 1 = 0$, die wir mit Hilfe der Formel für Grad zwei zu $2 \pm \sqrt{5}$ lösen. Wieder die Rekursion rückwärts durchlaufend fügen wir zunächst ein Wurzelzeichen um diese Lösung hinzu, addieren dann die Drei aus dem zweiten Schritt, fügen wieder eine Wurzel hinzu und subtrahieren zuletzt die Zwei aus dem ersten Reduktionsschritt. Insgesamt bekommen wir dann die Lösungen

$$-2 \pm \sqrt{3 \pm \sqrt{2 \pm \sqrt{5}}}$$

Wir überprüfen, dass dieser Ausdruck und damit auch all seine Konjugierten tatsächlich Nullstellen von p sind. Angesichts des relativ komplex erscheinenden Anfangspolynoms p ergibt hier die Anwendung der Sonderfälle eine effiziente Lösung.

5.5.4 Polynome vom Grad vier

Bekanntermaßen sind auch die Nullstellen eines Polynoms vom Grad vier immer durch Radikale ausdrückbar, allerdings nicht notwendigerweise durch Quadratwurzeln. In diesem Fall bietet sich eine etwas andere Vorgehensweise an als für die Fälle höheren Grades. Dabei expandieren wir ein Element vom Grad vier symbolisch und vergleichen das Ergebnis mit dem vorgegebenen Polynom.

Wir betrachten ein Element aus $\mathbb{E}_{\mathbb{K}}$ vom Grad vier. Wir wollen o.B.d.A. annehmen, dass der konstante Summand dieses Elements null ist; andernfalls können wir den konstanten Faktor wie in Unterabschnitt 5.5.3 beschrieben eliminieren. Das Element soll echt vom Grad vier sein, d.h. der Faktor vor der zweiten Wurzel ist nicht null; ob der Faktor positiv oder negativ ist, ist nicht interessant, da zu beiden Fällen eine Konjugierte vorliegen wird. Wir können also ein Element vom Grad vier durch vier Elemente a, b, c und d aus $\mathbb{K}$ wie folgt ausdrücken:

$$\xi = a\sqrt{b} + \sqrt{c + d\sqrt{b}}$$

Hierin genügt es noch, wenn $a \in \{0, -1, 1\}$; jeden anderen Betrag können wir in die Wurzel mit hineinziehen und entsprechend aus d herausdividieren. Weiter können wir $a = 0$ ebenfalls ausschließen, wenn wir voraussetzen, dass wir den in Unterabschnitt 5.5.2 beschriebenen Sonderfall schon behandelt haben; denn dieser wäre eingetreten, wenn $a = 0$ ist. Also ist $a \in \{-1, 1\}$. Offenbar gibt es dann aber unabhängig von dem Wert von ξ eine Konjugierte von ξ, in

Require: $p = X^4 + k_2X^2 + k_1X + k_0$
  $q := -k_1X^3 + (k_2^2 - 4k_0)X^2 + 2k_1k_2X + k_1^2$
  faktorisiere q
  if q **hat einen Linearfaktor** $(X - x)$ **then**
    $d := x$
    $b := -\frac{k_1}{4x}$
    $c := -\frac{k_2}{2} - b$
    return $\{\sqrt{b} + \sqrt{c + d\sqrt{b}}\}$
  else
    return $\{\}$
  end if

Abbildung 5.7: Algorithmus zum Finden von Nullstellen mit Quadratwurzeln bei Polynomen vom Grad vier

der $a = 1$ ist, und da es uns genügt, irgendeine Konjugierte von ξ zu finden, nehmen wir o.B.d.A. $a = 1$ an.

Nehmen wir das Produkt aus den vier Linearfaktoren $X - \eta$ für die Konjugierten η von ξ, so erhalten wir symbolisch das Minimalpolynom:

$$P = X^4 + (-2c - 2b)\,X^2 + (4bd)\,X + \left(c^2 + b^2 - 2bc - d^2b\right)$$

Nehmen wir an, wir haben ein Eingabepolynom p, das ebenfalls vom Grad vier ist. Unsere Aufgabe ist es, nach Nullstellen von p aus dem Euklidischen Körper zu suchen. Wir drücken auch p symbolisch aus und setzen p und P koeffizientenweise gleich. Ist $p = X^4 + k_2X^2 + k_1X + k_0$, so erhalten wir damit das Gleichungssystem

$$\begin{aligned} k_2 &= -2c - 2b \\ k_1 &= 4bd \\ k_0 &= c^2 + b^2 - 2bc - d^2b \end{aligned}$$

Wir lösen die oberste Gleichung nach c auf und erhalten $c = -\frac{k_2}{2} - b$. Setzen wir dies in die beiden anderen Gleichungen ein, so haben wir noch

$$\begin{aligned} k_1 &= 4bd \\ k_0 &= 2bk_2 + 4b^2 - d^2b + \frac{1}{4}k_2^2 \end{aligned}$$

Lösen wir jetzt die obere Gleichung nach b auf und setzen sie in die untere ein, so erhalten wir die Gleichung

$$k_0 = \frac{k_1k_2}{2d} + \frac{k_1^2}{4d^2} - \frac{dk_1}{4} + \frac{k_2^2}{4}$$

Wir können $d = 0$ ausschließen, denn ansonsten wäre auch k_1 null, und wir hätten ebenfalls den in 5.5.2 beschriebenen Sonderfall[1]. Also können wir die Gleichung mit $4d^2$ multiplizieren und erhalten die Form

$$0 = -4k_0d^2 + 2dk_1k_2 + k_1^2 - d^3k_1 + d^2k_2^2$$

Wir wollen d berechnen; daher substituieren wir $X = d$, und erhalten:

$$0 = -k_1X^3 + (k_2^2 - 4k_0^2)X^2 + 2k_1k_2X + k_1^2$$

Wir bezeichnen das Polynom auf der rechten Seite als q_p. Es gilt

Lemma 5.5.3. *Ist $p = X^4 + k_2X^2 + k_1X + k_0 \in \mathbb{K}[X]$ mit $k_1 \neq 0$ Minimalpolynom eines Elements aus $\mathbb{E}_\mathbb{K}$, dann hat $q_p \in \mathbb{K}[X]$ einen linearen Faktor $(X - x)$, und mit $d = x$, $b = \frac{k_1}{4x}$ und $c = -\frac{k_2}{2} - b$ ist die Nullstelle von p bis auf Konjugation $\sqrt{b} + \sqrt{c + d\sqrt{b}}$.*

Beweis. Wenn p Minimalpolynom eines Elements aus $\mathbb{E}_\mathbb{K}$ ist, dann ist dieses Element vom Grad vier und erfüllt also alle oben beschriebenen Gleichungen. q_p muss dann mit d eine Lösung aus $\mathbb{K}$ haben. Die Größen b und c ergeben sich dann durch Einsetzen in die jeweiligen vorigen Gleichungen. □

Wir können also die Nullstellen von Polynomen vom Grad vier, die durch Quadratwurzeln darstellbar sind, über die Faktorisierung von q_p finden. Falls q_p keinen linearen Faktor aus $\mathbb{K}$ hat oder eine der anderen Bedingungen nicht erfüllt sind, dann hat p keine Nullstelle aus $\mathbb{E}_\mathbb{K}$. Sind die Bedingungen erfüllt, dann können wir die Nullstelle aus b, c und d berechnen und testen, ob dies tatsächlich eine Nullstelle von q ist.

Insgesamt ist dies deutlich einfacher als die Berechnung des Summen- und Differenzenpolynoms, die im Wesentlichen vom Grad sechs wären und daher auch schwieriger zu faktorisieren wären.

5.5.5 Polynome vom Grad acht

Ab einem Grad von acht ist die eben bei Grad vier verwendete Methode deutlich schwieriger. Die Berechnung des Summen- und des Differenzenpolynoms ist aber auch symbolisch noch möglich. Auf diese Weise kann also dieser Schritt des Algorithmus im Wesentlichen in einer festen Vorverarbeitung geschehen[2]. Das Vorgehen in diesem Sonderfall ist direkt mit dem im Hauptalgorithmus identisch, daher verzichten wir hier auf nähere Erläuterungen.

[1] Dies erscheint im ersten Moment erstaunlich, da ξ dann ja $\sqrt{b} + \sqrt{c}$ wäre und somit nicht ein Ausdruck unter einer Wurzel. Dies ist aber eine Voraussetzung dafür, dass der Sonderfall aus Abschnitt 5.5.2 zutrifft. Tatsächlich ist ξ in diesem Fall aber nicht mehr vollständig und hat eine zweite Darstellung, nämlich $\sqrt{b + c + 2\sqrt{bc}}$.

[2] Die vollständigen symbolischen Summen- und Differenzenpolynome für Polynome vom Grad acht füllen jeweils etwa 13 Seiten. Wir verzichten daher hier darauf, das Polynom vollständig anzugeben.

5.6 Zusammengesetzter Algorithmus

Wir wollen in diesem Abschnitt die Ergebnisse des Kapitels zusammentragen und einen vollständigen Algorithmus zum Finden durch Quadratwurzeln ausdrückbarer Nullstellen angeben. Die Faktorisierung über dem Grundkörper $\mathbb{K}$ setzen wir dabei voraus. An einigen Stellen benötigen wir keine volle Faktorisierung, sondern nur die irreduziblen Faktoren bis zu einem bestimmten Grad d; wir notieren den Faktorisierungsaufruf dann einfach mit dieser Obergrenze als zweitem Argument.

Zuerst einmal die Schritte im Überblick: Wir erhalten ein univariates Polynom $p \in \mathbb{K}[X]$. Wir wollen alle Nullstellen dieses Polynoms angeben, die in $\mathbb{E}_\mathbb{K}$ liegen.

Dazu faktorisieren wir p zunächst mit Hilfe des Faktorisierungsalgorithmus. Alle Faktoren, deren Grad keine Zweierpotenz ist, verwerfen wir. Bleiben ansonsten keine Faktoren über, so hat nach Korollar 2.1.4 das Polynom p keine Nullstellen in $\mathbb{E}_\mathbb{K}$, und wir geben die leere Menge zurück. Für alle Linearfaktoren $X - a$ fügen wir a der Lösungsmenge zu. Alle quadratischen Faktoren lösen wir mit Hilfe der pq-Formel.

Für Faktoren vom Grad vier oder höher wissen wir nicht im Voraus, ob der Faktor durch Quadratwurzeln darstellbare Nullstellen hat. Ist $\mathbb{K}$ eine Erweiterung von $\mathbb{Q}$, so nutzen wir jetzt zunächst den schnellen Test aus Abschnitt 5.1, indem wir den Faktor modulo der ersten 25 Primzahlen weiter faktorisieren. Erhalten wir dabei einen Faktor, dessen Grad nicht eine Zweierpotenz ist, so sind wir sicher, dass wir keine Nullstellen aus $\mathbb{E}_\mathbb{K}$ finden werden, und können abbrechen. Andernfalls besteht natürlich immer noch eine Restmöglichkeit, dass der Faktor trotzdem keine durch Quadratwurzeln darstellbare Nullstellen hat; in diesem Fall würden wir aber im weiteren Verlauf des Algorithmus irgendwann auf diese Tatsache stoßen.

Auf die Faktoren, die diese Vortests erfolgreich durchlaufen haben, wenden wir jetzt, soweit möglich, die Sonderfallbetrachtung aus Unterabschnitt 5.5.2 und 5.5.3 an. Können wir damit den Faktor auf ein Polynom vom Grad vier reduzieren, wenden wir den Sonderfall aus Unterabschnitt 5.5.4 an.

Falls wir auf einen Faktor keinen weiteren der Sonderfälle anwenden können, starten wir mit diesem Faktor den allgemeinen Algorithmus; wir wollen diesen Faktor mit $\widetilde{p}$ bezeichnen. Mit dem in Abschnitt 5.3.4 beschriebenen Algorithmus berechnen wir das Summenpolynom p_+ und das Differenzenpolynom p_- von $\widetilde{p}$. Die Polynome p_+ und p_- faktorisieren wir jetzt, wobei wir uns auf Faktoren vom Grad $\leq 2^{d-1}$ bzw. beim Differenzenpolynom auf Faktoren exakt vom Grad 2^d beschränken. Im Allgemeinen rechnen wir damit, nur jeweils einen Faktor mit passendem Grad zu erhalten. Nur für den Fall, dass die Nullstellen von $\widetilde{p}$ trotz des vorher absolvierten Tests doch nicht durch Quadratwurzeln darstellbar sind oder wenn das zwar der Fall ist, die Nullstelle aber nicht vollständig ist, können wir verschiedene Faktoren erhalten, die unseren Gradbedingungen genügen. In diesem Fall halten wir einen Faktor von p_+ fest, wissen jetzt aber nicht, welcher Faktor von p_- zu dieser Darstellung gehört; wir müssen also alle Faktoren von p_- ausprobieren.

Ist $\widetilde{p}$ Minimalpolynom eines Elementes ξ aus $\mathbb{E}_\mathbb{K}$ und ist $\xi = \xi_1 + \sqrt{\xi_2}$, so sind

Require: $p \in \mathbb{K}[X]$
  faktorisiere p
  $lsg = \{\}$
  for all $\widetilde{p}$ **Faktor von** p**,** $\deg(\widetilde{p}) = 2^d$ **do**
    if $\deg(\widetilde{p}) \leq 2$ **then**
      $lsg := lsg \cup$ **löseGrad2**$(\widetilde{p})$
      nächster Faktor
    end if
    $\widetilde{p}$:=**eliminiereKonstante**$(\widetilde{p})$
    if $\widetilde{p}$ **hat nur gerade Exponenten then**
      $lsg :=\ lsg \cup$**reduziereWennAlleExponentenGrade**$(\widetilde{p})$
      nächster Faktor
    end if
    if schnellerTest$(\widetilde{p}) == false$ **then**
      nächster Faktor
    end if
    if $\deg(\widetilde{p}) == 4$ **then**
      $lsg := lsg \cup$ **löseGrad4**$(\widetilde{p})$
      nächster Faktor
    end if
    (p_+, p_-) :=**berechneSummenUndDifferenzenPolynom**$(\widetilde{p})$
    faktorisiere p_+**,** 2^{d-1}
    faktorisiere p_-**,** 2^d
    $lsg_1 :=$ **löse**$(\widetilde{\mathbf{p_+}})$
    for all $\widetilde{p_+}$ **Faktor von** p_+**,** $\deg(\widetilde{p_+}) \leq 2^{d-1}$ **do**
      $lsg_1 := lsg_1 \cup \{\frac{\xi}{2} | \xi \in$ **löse**$(\widetilde{p_+})\}$
    end for
    if $p_-(\sqrt{X})$ **hat einen Faktor** $\widetilde{p_-}$ **mit** $\deg(\widetilde{p_-}) = 2^{d-1}$ **then**
      $lsg_2 := \{\sqrt{\frac{\xi}{4}} | \xi \in$ **löse**$(\widetilde{p_-})\}$
      if $lsg_2 == \emptyset$ **then**
        return $\{\}$
      else
        $\xi_2 := \sqrt{\frac{\xi}{4}}$ **für ein** $\xi \in lsg_2$
      end if
    else
      return $\{\}$
    end if
    for all $\xi_1 : \xi_1 \sim \xi_1' \in lsg_1$ **do**
      if $\widetilde{p}(\xi_1 + \xi_2) == 0$ **then**
        $lsg := lsg \cup \{\xi_1 + \xi_2\}$
      end if
    end for
  end for
  return lsg

Abbildung 5.8: Algorithmus zum Finden von Nullstellen mit Quadratwurzeln

die jeweiligen Faktoren von p_+ und p_- nach Satz 5.2.1 die Minimalpolynome von $2\xi_1$ und $4\xi_2$, wenn wir X^2 im Differenzenpolynom durch X ersetzen. Wir rufen den Algorithmus mit diesen beiden Polynomen rekursiv auf, wobei wir uns soweit wie möglich wieder der vereinfachenden Sonderfälle bedienen.

Beim Zusammensetzen der beiden Teile ξ_1 und ξ_2 müssen wir bedenken, dass wir nicht wissen, welche jeweilige Konjugierte von ξ_1 mit welcher Konjugierten von ξ_2 kombiniert werden muss. Wir müssen also für eine Darstellung von ξ_2 alle möglichen Darstellungen von ξ_1 ausprobieren und den Kandidaten jeweils durch Einsetzen in $\widetilde{p}$ testen. Sollte sich dabei keine der Kombinationen als Nullstelle von $\widetilde{p}$ erweisen, so haben wir gezeigt, dass $\widetilde{p}$ keine Nullstelle in $\mathbb{E}_\mathbb{K}$ hat; denn ansonsten müssten nach dem oben Gesagten eine Kombination erfolgreich sein. Natürlich ist es möglich, dass bereits einer der beiden rekursiven Aufrufe keine Lösung zurückgegeben hat; in diesem Fall können wir ebenfalls mit einer leeren Menge als Rückgabe abbrechen (es sei denn, wir hatten mehrere Faktoren von p_+ oder p_-, die unsere Gradanforderungen erfüllten).

Der vollständige Algorithmus findet sich in Abbildung 5.8. Dabei bezeichnet der Aufruf **nächster Faktor** einen Sprung ans Ende der äußeren **FOR**-Schleife, in der die Faktoren von p durchlaufen werden. Außerdem werden die vorher angegebenen Algorithmen aufgerufen, und zwar mit den Bezeichnungen **löseGrad2** (Abbildung 5.4), **eliminiereKonstante** (Abbildung 5.6), **reduziereWennAlleExponentenGrade** (Abbildung 5.5), **löseGrad4** (Abbildung 5.7) und **berechneSummenUndDifferenzenPolynom** (Abbildung 5.3).

Betrachten wir die Laufzeit des Algorithmus. Zunächst benötigen wir eine Faktorisierung des Eingabepolynoms p. Sei d der Grad eines der Faktoren, dann schließt sich (wenn nicht einer unserer Sonderfälle zuschlägt) ein Aufruf des Berechnungsalgorithmus für das Summen- und Differenzenpolynom an. Die beiden resultierenden Polynome sind vom Grad $\frac{d(d-1)}{2}$ und müssen jetzt wieder faktorisiert werden; diese Faktorisierungen machen gewöhnlich den Hauptteil der Laufzeit aus. Mit den Ergebnissen wird der Algorithmus rekursiv aufgerufen, wobei die Polynome jeweils nur noch den Grad $\frac{d}{2}$ haben. Im schlimmsten Fall (wenn die Nullstellen nicht vollständig sind) müssen wir dabei $\frac{d(d-1)2}{2d} = d-1$ rekursive Aufrufe machen. Hat p aber eine vollständige Nullstelle aus $\mathbb{E}_\mathbb{K}$, so benötigen wir nach den Sätzen 5.4.8 und 5.4.5 nur zwei rekursive Aufrufe.

Die Laufzeit für die Berechnung der γ wollen wir mit $T_\gamma(d)$ bezeichnen, und die Laufzeit einer Faktorisierung eines Polynoms vom Grad d mit $T_f(d)$. Nach Lemma 5.3.16 benötigt die Berechnung der γ asymptotisch $O(d^4)$ Operationen in $\mathbb{K}$. Die Laufzeit der Faktorisierung ist natürlich abhängig von $\mathbb{K}$. Wir wollen aber annehmen, dass die Faktorisierung wenigstens linear ansteigende Laufzeit hat, so dass wir für $a \geq 1$ die Abschätzung $a\,T_f(b) \leq T_f(ab)$ machen können. Die gesamte Laufzeit für einen Schritt unseres Algorithmus fassen wir in $t(d)$ zusammen, d.h. wir definieren

$$t(d) = T_f(d) + T_\gamma(d) + 2T_f\left(\frac{d(d-1)}{2}\right)$$

Auch für $t(d)$ können wir jetzt $a\,t(b) \leq t(ab)$ abschätzen.

Wir wollen die Laufzeitabschätzung zunächst für den ungünstigsten Fall

durchführen, in dem wir tatsächlich in jedem Schritt $d-1$ rekursive Aufrufe machen müssen. In diesem Fall erhalten wir folgendes Lemma:

Lemma 5.6.1. *Sei $p \in \mathbb{K}[X]$ ein Polynom vom Grad d. Bezeichne $T(d)$ die Laufzeit für den Algorithmus zur Suche nach Nullstellen von p aus $\mathbb{E}_{\mathbb{K}}$, dann gilt*

$$T(d) \in O\left(d^{\log(d)} t(d)\right)$$

Beweis. Es ist

$$\begin{aligned}
T(d) &\leq t(d) + dT\left(\frac{d}{2}\right) \\
&\leq t(d) + dt\left(\frac{d}{2}\right) + d^2 T\left(\frac{d}{4}\right) \\
&\vdots \\
&\leq \sum_{i=0}^{\log(d)-1} d^i t\left(\frac{d}{2^i}\right) \\
&\leq \sum_{i=0}^{\log(d)-1} \left(\frac{d}{2}\right)^i t(d) \\
&\in O\left(d^{\log(d)} t(d)\right)
\end{aligned}$$

□

Diese Abschätzung ist aber nur in dem unwahrscheinlichen Fall relevant, wenn das Summen- und Differenzenpolynom in jedem Schritt maximal ungünstig zerfällt. Aus den Abschnitten 5.4.1 und 5.4.2 wissen wir, dass dies in der Regel nicht der Fall ist. Wir wollen also eine weitere Abschätzung machen, in der wir davon ausgehen, dass wir den Algorithmus mit einem Polynom aufrufen, welches tatsächlich durch Quadratwurzeln ausdrückbare Nullstellen hat, und dass diese Nullstellen vollständig sind.

Lemma 5.6.2. *Sei $p \in \mathbb{K}[X]$ ein Polynom vom Grad d mit vollständigen Nullstellen aus $\mathbb{E}_{\mathbb{K}}$. Bezeichne $T(d)$ die Laufzeit für den Algorithmus zur Suche nach Nullstellen von p aus $\mathbb{E}_{\mathbb{K}}$, dann gilt*

$$T(d) \in O\left(\log(d)\, t(d)\right)$$

Beweis. Außer $t(d)$ werden wir genau zwei Aufrufe des Algorithmus mit halbem Grad haben (jetzt nur zwei, da wir annehmen, dass die Nullstellen existieren und vollständig sind). Schreiben wir diese Rekursion auf, so erhalten wir:

$$\begin{aligned}
T(d) &\leq t(d) + 2T\left(\frac{d}{2}\right) \\
&\leq t(d) + 2t\left(\frac{d}{2}\right) + 4T\left(\frac{d}{4}\right) \\
&\vdots \\
&\leq \sum_{i=0}^{\log(d)-1} 2^i t\left(\frac{d}{2^i}\right) \\
&\leq \sum_{i=0}^{\log(d)-1} t(d) \\
&\in O\left(\log(d) t(d)\right)
\end{aligned}$$

□

Zusammengefasst ist die Laufzeit des Algorithmus zur Berechnung von mit Quadratwurzeln ausdrückbaren vollständigen Nullstellen eines univariaten Polynoms vom Grad d also in $O\left(\log(d)\left(T_f(d) + T_\gamma(d) + 2T_f\left(\frac{d(d-1)}{2}\right)\right)\right)$. Existieren keine solche Nullstellen oder sind diese nicht vollständig, kann der Algorithmus maximal eine Laufzeit von $O\left(d^{\log(d)}\left(T_f(d) + T_\gamma(d) + 2T_f\left(\frac{d(d-1)}{2}\right)\right)\right)$ haben.

Kapitel 6

Quadratwurzellösungen von Systemen von Gleichungen

Bisher sind wir davon ausgegangen, die durch Quadratwurzeln erzeugten Nullstellen eines univariaten Polynoms zu finden, was eine Konstruktionsaufgabe löst, wenn genau ein Parameter aus genau einem anderen konstruiert werden soll und sich der Constraint zwischen diesen beiden Größen polynomiell aufschreiben lässt. Im Allgemeinen haben wir aber mehrere Variablen und auch mehrere Gleichungen. Außerdem sind alle diese Gleichungen ihrerseits aus Konstruktionen hervorgegangen, d.h. wir haben auch in den Gleichungen Quadratwurzeln stehen. Wir wollen uns jetzt dem Problem zuwenden, ein solches Gleichungssystem auf die oben besprochene Auflösung eines univariaten Polynoms zurückzuführen. Dazu zeigen wir im ersten Abschnitt dieses Kapitels zunächst, wie wir aus dem allgemeinen Gleichungssystem mit Quadratwurzeln ein polynomielles Gleichungssystem erstellen können. Die folgenden drei Abschnitte beschäftigen sich mit der Anpassung klassischer Methoden zur Reduzierung eines multivariaten Gleichungssystems auf eine univariate polynomielle Gleichung; allgemeine Formen dieser Methoden finden sich z.B. in den Büchern von Cox, Little und O'Shea [5, 6]. Im letzten Abschnitt gehen wir auf eine neue Reduktionsmöglichkeit ein, die besonderen Nutzen aus den geometrischen Ursprüngen des Gleichungssystems zieht.

6.1 Reduktion auf polynomielle Gleichungssysteme

Bedingt durch die Konstruktion sind alle Gleichungen durch Anwendung der Körperoperationen und des Quadratwurzeloperators aus Elementen von $\mathbb{Q}$ und Variablen für die freien Parameter hervorgegangen. Alle freien Parameter, die in der Konstruktionsaufgabe keinem Constraint unterliegen, bezeichnen wir als $Y_1, \ldots, Y_r$ und betrachten sie als dem Grundkörper zugehörig, d.h. $\mathbb{K} = \mathbb{Q}(Y_1, \ldots, Y_r)$. Bezeichnen wir die anderen Variablen mit $X_1, \ldots, X_m$, so besteht also jede Gleichung unseres Constraintsystems aus Elementen aus $\mathbb{E}_{\mathbb{K}(X_1,\ldots,X_m)}$. Wir beschäftigen uns zunächst damit, diese Gleichungen auf polynomielle Gleichungen zurückzuführen.

Dies gelingt uns mit Hilfe des Korollars 3.2.3. Nach diesem Korollar müssen

wir die Gleichungen nur mit allen Konjugierten multiplizieren, um jede Gleichung in einen rationalen Ausdruck aus $\mathbb{K}(X_1, \ldots, X_m)$ umzuwandeln. Dabei fügen wir jeder Gleichung noch Lösungen hinzu und damit möglicherweise auch dem gesamten System, d.h. wir müssen nach Berechnen der Lösungen für das polynomielle System alle erhaltenen Lösungen noch kontrollieren.

Jetzt müssen wir uns noch um die Nenner kümmern; dafür kürzen wir die Brüche zunächst so weit wie möglich und lösen dann das Gleichungssystem aus den Zählern. Die Lösungen, die wir dann erhalten, setzen wir in die Nenner ein und verwerfen alle Lösungen, bei denen ein Nenner null wird.

Fassen wir die Schritte noch einmal kurz zusammen. Wir haben zu Beginn ein System von n Gleichungen $q_1, \ldots, q_n \in \mathbb{E}_{\mathbb{K}(X_1,\ldots,X_m)}$. Wir multiplizieren jedes q_i mit all seinen Konjugierten. Die erhaltenen rationalen Ausdrücke kürzen wir und betrachten dann nur noch die Zähler $p_1, \ldots, p_n \in \mathbb{K}[X_1, \ldots, X_m]$. Dieses polynomielle Gleichungssystem lösen wir dann. Alle durch Quadratwurzeln ausdrückbaren Lösungen dieses Systems setzen wir in das ursprüngliche Gleichungssystem ein und überprüfen, ob jeder Ausdruck identisch null ist.

Es bleibt der Zwischenschritt, das polynomielle Gleichungssystem auf die Lösung univariater Polynome zurückzuführen. Verschiedene Möglichkeiten, dieses Problem anzugehen, betrachten wir in den folgenden Abschnitten.

6.2 Schrittweise Reduktion

Wir suchen jetzt für Polynome $p_1, \ldots, p_n \in \mathbb{K}[X_1, \ldots, X_m]$ eine gemeinsame Nullstelle aus $\mathbb{E}_{\mathbb{K}}^m$. Betrachten wir folgendes simple Beispiel.

Beispiel 6.2.1. Sei

$$\begin{aligned} p_1 &= X^2 - Y^2 - 1 \\ p_2 &= X^2 - Y^2 + X - 2 \end{aligned}$$

Wir suchen Werte für X und Y so, dass $p_1(X, Y) = p_2(X, Y) = 0$ ist. Wir können die obere Gleichung univariat auflösen und erhalten $X = \sqrt{Y^2 + 1}$. Setzen wir dies in die untere Gleichung ein, so erhalten wir $1 = \sqrt{Y^2 + 1}$ und stellen daher fest, dass $Y = 0$ und $X = 1$ eine Lösung ist.

Was dieses Gleichungssystem für uns so einfach gemacht hat, war der Umstand, dass wir die obere Gleichung bezüglich X univariat auflösen konnten und hier schon einen Ausdruck mit Quadratwurzeln erhielten, in dem Y noch symbolisch vorkam. Dies konnten wir einfach in die zweite Gleichung einsetzen und hatte somit das Gleichungssystem um eine Variable und eine Gleichung reduziert.

Für den ersten Auflösungsschritt haben wir p_1 also als Polynom in $\mathbb{K}(Y)[X]$ aufgefasst und dort univariat gelöst. Verallgemeinernd kann man alle Gleichungen durchgehen und überprüfen, ob man diese Gleichung nach einer ihrer Variablen univariat auflösen kann. Den Test können wir mit dem Algorithmus aus Abschnitt 5.1 durchführen. Gelingt dies, so setzen wir die Lösung in alle

anderen Gleichungen ein und erhalten dadurch wieder ein System mit Quadratwurzeln, in dem aber eine Variable und eine Gleichung eliminiert sind. Diese Systeme formen wir wieder wie im Abschnitt zuvor in ein polynomielles Gleichungssystem um und suchen nach einer weiteren Gleichung, die wir in einer ihrer Variablen univariat auflösen können.

Es stellt sich die Frage, ob bei einer gemeinsamen Nullstelle aus $\mathbb{E}_{\mathbb{K}}^n$ des Gleichungssystems notwendig immer mindestens eine der Gleichung nach einer der Variablen univariat durch Quadratwurzeln auflösbar ist. Wir stellen an folgendem Beispiel fest, dass dies leider nicht der Fall ist:

Beispiel 6.2.2. Sei

$$\begin{aligned} p_1 &= X^3 - Y^3 + Y^2 \\ p_2 &= X^3 + Y^3 - 1 \end{aligned}$$

Wir stellen fest, dass sowohl p_1 als auch p_2 irreduzibel sowohl über $\mathbb{Q}(X)$ als auch über $\mathbb{Q}(Y)$ sind. Da beide Polynome in beiden Variablen den Grad drei haben, haben sie also in $\mathbb{Q}(X)$ bzw. $\mathbb{Q}(Y)$ jeweils keine Nullstelle, die durch Quadratwurzeln ausdrückbar ist. Tatsächlich überzeugt man sich aber, dass $X = 0$ und $Y = 1$ eine Lösung des Systems ist, also das System mit $(0, 1)$ eine Lösung in $\mathbb{E}_{\mathbb{Q}}$ hat.

Es ist also möglich, dass wir keine einzelne Gleichung univariat durch Quadratwurzeln auflösen können, trotzdem das Gesamtsystem aber durch Quadratwurzeln auflösbar ist. Die Herangehensweise über das einzelne Auflösen nach Variablen führt demnach nicht in allen Fällen zum Ziel. Finden wir also keine einzeln auflösbare Gleichung, wenden wir eine allgemeinere, klassische Methode an, die aufwändiger ist, im Falle einer existierenden Lösung aber garantiert zum Ziel führt. Diese Methode arbeitet über Resultanten und wird im nächsten Abschnitt vorgestellt.

6.3 Reduktion über verallgemeinerte Resultanten

Der in diesem Abschnitt beschriebene Weg zur Reduktion arbeitet über die Methode der *verallgemeinerten Resultanten*, wie sie z.B. in [5] beschrieben wird.

Wir bezeichnen die klassische Resultante zweier Polynome p_1 und p_2 bezüglich einer Unbekannten X mit $Res(p_1, p_2, X)$. Die verallgemeinerte Resultante von n Polynomen definieren wie folgt:

Definition 6.3.1. Seien $p_1, \ldots, p_n \in \mathbb{Q}[X_1, \ldots, X_m]$. Seien $u_2, \ldots, u_n$ neue Variablen. Wir betrachten die Resultante $q := Res(p_1, u_2 \cdot p_2 + \cdots + u_n \cdot p_n, X_1)$. Die verallgemeinerten Resultanten $R(p_1, \ldots, p_n, X_1)$ von $p_1, \ldots, p_n$ sind alle Koeffizienten von q aufgefasst als Polynom in $u_2, \ldots, u_n$ über $\mathbb{Q}[X_1, \ldots, X_m]$.

Mit diesen verallgemeinerten Resultanten beweisen wir jetzt das Eliminations- und Expansionstheorem für Nullstellen aus $\mathbb{E}_{\mathbb{K}}$.

Satz 6.3.2. *Sei $\xi_1, \ldots, \xi_m \in \mathbb{E}_\mathbb{K}$ so, dass $p_i(\xi_1, \ldots, \xi_m) = 0\ \forall i$. Dann gilt für alle verallgemeinerten Resultanten $r \in R(p_1, \ldots, p_n, X_1)$: $r(\xi_2, \ldots, \xi_m) = 0$.*

Beweis. Es genügt zu zeigen, dass $Res(p_1, u_2 \cdot p_2 + \cdots + u_n \cdot p_n, X_1) = 0$ gilt. Denn ist diese Resultante als Polynom in $u_2, \ldots, u_n$ gleich null, so muss jeder Koeffizient null sein, und die Koeffizienten sind nach Definition die verallgemeinerten Resultanten.

Betrachten wir die beiden Polynome p_1 und $u_2 \cdot p_2 + \cdots + u_n \cdot p_n$ unter Einsetzung der $\xi_2, \ldots, \xi_n$, so wird jedes p_i zu einem univariaten Polynom in X_1. Wir wissen, dass alle diese univariaten Polynome eine gemeinsame Nullstelle haben, nämlich nach Voraussetzung ξ_1. Dass heißt aber schon, dass alle p_i einen gemeinsamen Faktor haben; diesen Faktor können wir aus dem Polynom $u_2 \cdot p_2 + \cdots + u_n \cdot p_n$ ausklammern und erhalten damit auch einen gemeinsamen Faktor der beiden Polynome p_1 und $u_2 \cdot p_2 + \cdots + u_n \cdot p_n$, deren Resultante wir bilden. Haben aber die beiden Polynome einer Resultanten einen gemeinsamen Faktor, dann ist die Resultante ebenfalls null. □

Für das Expansionstheorem benötigen wir als Voraussetzung, dass die Leitkoeffizienten der Polynome bezüglich X_1 keine gemeinsamen Nullstellen haben; ein Bruch dieser Voraussetzung führt allerdings nur zu einigen unnötigen Nullstellen, die wir im Anschluss zu prüfen haben. Mit diesen Voraussetzungen beweisen wir das Expansionstheorem für $\mathbb{E}_\mathbb{K}$:

Satz 6.3.3. *Sei $(\xi_2, \ldots, \xi_m) \in \mathbb{E}_\mathbb{K}^{m-1}$ eine Lösung für alle verallgemeinerten Resultanten $r \in R(p_1, \ldots, p_n, X_1)$ und sei $(\xi_2, \ldots, \xi_m)$ nicht gemeinsame Nullstelle der Leitkoeffizienten der Polynome p_1 und $u_2 \cdot p_2 + \cdots + u_n \cdot p_n$ bezüglich X_1. Dann ist $GCD\,(p_1(X_1, \xi_2, \ldots, \xi_m), \ldots, p_n(X_1, \xi_2, \ldots, \xi_m)) \neq 1$.*

Beweis. Die Koeffizienten von $Res(p_1, u_2 \cdot p_2 + \cdots + u_n \cdot p_n)$ sind die verallgemeinerten Resultanten und nach Voraussetzung alle null. Die Resultante kann aber nur null sein, wenn entweder die Leitkoeffizienten der beiden Polynome beide null sind, oder die beiden Polynome einen gemeinsamen, nicht konstanten Faktor haben. Ersteres ist nach Voraussetzung ausgeschlossen; also haben die beiden Polynome einen gemeinsamen Faktor g. Da die u_i in p_1 nicht vorkommen, muss g ein Faktor des Inhaltes von $u_2 \cdot p_2 + \cdots + u_n \cdot p_n$ bezüglich der u_i sein und damit ein Faktor von jedem p_i für $2 \leq i \leq n$. Von p_1 ist g ebenfalls Faktor, d.h. g ist Faktor aller p_i, und die Aussage ist bewiesen. □

Für einige Fälle kann es hilfreich sein, statt der verallgemeinerten Resultanten besser paarweise von alle Polynomen die klassischen Resultanten zu berechnen. Tatsächlich gilt dann auch für diese Menge das Expansionstheorem, nicht aber das eben für die verallgemeinerte Resultante gezeigte Eliminationstheorem. Um sich davon zu überzeugen, betrachte man folgendes Beispiel:

Beispiel 6.3.4. Sei

$$\begin{aligned} p_1 &= (y - z + 1)x^2 + (y - z + 3)x + (y - z + 2) \\ p_2 &= \left((y - z)^2 + 1\right) x^2 + (y - z + 4)x + (y - z + 3) \\ p_3 &= x^2 + (y - z + 5)x + (y - z + 6) \end{aligned}$$

Zunächst sieht man schnell, dass alle Leitkoeffizienten bezüglich x paarweise teilerfremd sind. Man stellt fest, dass für $y = 1$ und $z = 1$ die drei Polynome zu $p_1 = (x + 1)(x + 2)$, $p_2 = (x + 1)(x + 3)$ und $p_3 = (x + 2)(x + 3)$ werden, also alle paarweise gemeinsame Faktoren haben, so dass $y = 1$ und $z = 1$ eine Partiallösung der paarweisen klassischen Resultanten $Res(p_1, p_2)$, $Res(p_1, p_3)$ und $Res(p_2, p_3)$ darstellt. Tatsächlich aber haben offenbar nicht alle drei Gleichungen einen gemeinsamen Faktor für diese Einsetzung.

Noch weiter ist es sogar so, dass sich für jede Partiallösung auf $y = z$ die oben genannten Polynome ergeben, so dass sogar an unendlich vielen Stellen die Aussage des Expansionstheorems nicht gilt.

Das Expansionstheorem besagt nur, dass es einen gemeinsamen Faktor der $p_1, \ldots, p_n$ gibt, wenn wir die Partiallösung $\xi_2, \ldots, \xi_m$ einsetzen. Ob dieser Faktor eine Nullstelle aus $\mathbb{E}_{\mathbb{K}}$ hat, muss dann noch mit dem univariaten Algorithmus überprüft werden.

Wir erhalten also durch die generalisierten Resultanten eine Lösungsweg für die Aufgabe, ein gegebenes polynomielles Gleichungssystem zu lösen. Wir berechnen dafür die generalisierten Resultanten und eliminieren eine der Variablen, berechnen für das resultierende System wieder die generalisierten Resultanten mit der nächsten Variable, bis wir schließlich nur noch ein System mit univariaten Gleichungen erhalten. Wenn dieses System eine gemeinsame Lösung hat, dann ist dies auch eine Lösung des kleinsten gemeinsamen Teilers des Systems; wir berechnen diesen also und lösen dieses univariate Polynom mit dem Algorithmus aus dem letzten Kapitel. Gelingt uns dies, so setzen wir die Lösung in das System aus dem Schritt zuvor ein und erhalten jetzt auch dort univariate Polynome, die wir auf dieselbe Weise lösen.

6.4 Reduktion über Multiresultanten

Die Methode der Reduktion über verallgemeinerte Resultanten hat den Vorteil, dass man in jedem Schritt wieder schnellere, nicht allgemeine Reduktionsmethoden probieren kann. Außerdem kann man nach der Reduktion auf ein univariates Problem die univariaten Lösungen in das bivariate System des Schrittes zuvor einsetzen (da wir ja symbolisch exakte Lösungen zur Verfügung haben) und erhält auf diese Weise wieder ein univariates System. Das heißt, wir müssen nur eine Variable nach der anderen reduzieren.

Trotzdem sind generalisierte Resultanten im Allgemeinen recht umständlich, da man durch die u_i viele unnötige Variablen hinzubekommt. Mit Multiresultanten (siehe z.B. [6]) ist es möglich, alle Unbekannten bis auf eine in einem Schritt zu eliminieren, ohne zusätzliche Variablen einführen zu müssen.

Wir definieren zunächst den Begriff der Multiresultanten und zeigen dann einige ihrer Eigenschaften.

Seien $f_1, \ldots, f_n$ Polynome in n Unbekannten, die wir mit $X_0, \ldots, X_{n-1}$ bezeichnen wollen, über einem Erweiterungskörper $\mathbb{K}$ von $\mathbb{Q}$. Die Multiresultante dieser Polynome bezüglich X_0 soll ein univariates Polynom in X_0 sein, so dass zu jeder gemeinsamen Nullstelle $(\xi_0, \ldots, \xi_{n-1})$ von $f_1, \ldots, f_n$ die erste Komponente ξ_0 Nullstelle der Multiresultante ist. Umgekehrt soll die Multiresultante

möglichst wenig weitere Nullstellen haben.

In [6] wird beschrieben, wie man zu n Polynomen in $n-1$ Unbekannten eine Matrix findet, deren Determinante 0 ist, wenn die Polynome eine gemeinsame Nullstelle haben. Wir betrachten nun eine unserer Unbekannten X_0 als einen Parameter, d.h. wir betrachten die f_i als Polynome in $n-1$ Unbekannten über dem Körper $\mathbb{K}(X_0)$. Man sagt, wir *verstecken* die Variable X_0. Auf diese Weise wird die Determinante der genannten Matrix ein Element aus $\mathbb{K}(X_0)$. Die Nullstellen dieser Determinante in X_0 umfassen dann auch die ersten Komponenten aller gemeinsamen Nullstellen der f_i.

Um diese Matrix zu definieren, betrachten wir die f_i jetzt als Elemente des Polynomrings $\mathbb{K}(X_0)[X_1, \ldots, X_{n-1}]$. Um die Monome leichter notieren zu können, schreiben wir in Zukunft einfach X^α für $X_1^{\alpha_1} \cdot \cdots \cdot X_{n-1}^{\alpha_{n-1}}$. Dabei soll $|\alpha| = \alpha_1 + \cdots + \alpha_{n-1}$ der Grad des Monoms sein.

Zunächst führen wir eine neue Variable X_n ein, um die f_i zu homogenisieren, d.h. dafür zu sorgen, dass alle Monome in f_i denselben Totalgrad $\deg(f_i)$ haben. Sind also die $f_i = \sum_{j=1}^{k_i} a_j X^{\alpha_{i,j}}$, so definieren wir

$$F_i = \sum_{j=1}^{k_i} a_j X^{\alpha_{i,j}} X_n^{\deg(f_i)-|\alpha|}$$

Die Koeffizienten a_j sind aus dem Grundkörper $\mathbb{K}(X_0)$. Man erkennt, dass für die Einsetzung $X_n = 1$ die F_i wieder zu den f_i evaluieren; wir brauchen aber diese homogene Form, um die oben beschriebene Matrix zu finden.

Sei $d_i = \deg(f_i)$ und $d = \sum_{i=1}^{n}(d_i - 1) + 1$. Dann stellen wir fest:

Lemma 6.4.1. *Jedes Monom X^α mit $|\alpha| \geq d$ wird für mindestens ein i durch $X_i^{d_i}$ geteilt.*

Beweis. Wäre dies nicht so, hätten alle X_i höchstens Grad $d_i - 1$, und das Monom hätte damit höchstens Grad $d - 1$. □

Durch diese Eigenschaft können wir die Menge alle Monome vom Grad d folgendermaßen aufteilen:

Definition 6.4.2. Sei S die Menge aller Monome X^α vom Grad $|\alpha| = d$. Wir definieren eine disjunkte Aufteilung dieser Menge in Mengen $S_1, \ldots, S_n$ wie folgt:

$$\begin{aligned}
S_1 &= \{X^\alpha : |\alpha| = d, X_1^{d_1} \text{ teilt } X^\alpha\} \\
S_2 &= \{X^\alpha : |\alpha| = d, X_2^{d_2} \text{ teilt } X^\alpha\} - S_1 \\
&\vdots \\
S_n &= \{X^\alpha : |\alpha| = d, X_n^{d_n} \text{ teilt } X^\alpha\} - S_1 - \cdots - S_{n-1}
\end{aligned}$$

Nach 6.4.1 liegt jedes Monom vom Grad d in einem der S_i. Wir betrachten jetzt folgendes Gleichungssystem:

$$\begin{aligned} \frac{X^\alpha}{X_1^{d_1}} F_1 &= 0 \ \forall X^\alpha \in S_1 \\ &\vdots \\ \frac{X^\alpha}{X_n^{d_n}} F_n &= 0 \ \forall X^\alpha \in S_n \end{aligned} \tag{6.4.1}$$

Jede dieser Gleichungen ist offenbar homogen vom Grad d.

Die Definition der S_n und dieses Gleichungssystems erscheint im ersten Moment etwas willkürlich. Tatsächlich aber erhalten wir auf diese Weise Zugriff auf die gesuchte Multiresultante. Denn notieren wir die Matrix A aus den Koeffizienten des Gleichungssystems 6.4.1 so, dass jede Gleichung zu einer Zeile gehört und jede Spalte einem Monom vom Grad d der Gleichungen zugeordnet wird, so stellen wir für diese Matrix A Folgendes fest:

Lemma 6.4.3. *A ist quadratisch mit der Größe $\binom{d+n-1}{n-1}$, und $\det(A)$ ist ein Vielfaches der gesuchten Multiresultante, d.h. $\det(A)$ ist ein Polynom in X_0, unter dessen Nullstellen alle ersten Komponenten aller Nullstellen des Gleichungssystems $f_1 = \cdots = f_n = 0$ sind.*

Beweis. Jede der obigen Gleichungen ist homogen vom Grad d, d.h. die Anzahl der Spalten entspricht genau der Anzahl N der Monome vom Grad d. Nach Konstruktion ist aber jedes dieser Monome in genau einem der S_i, d.h. die Anzahl der Zeilen ist ebenfalls N. Dass N der behauptete Binomialkoeffizient $\binom{d+n-1}{n-1}$ ist, zeigen wir per Induktion über n. Offenbar gilt dies für jedes d und für $n = 1$, da es genau ein homogenes Monom in einer Variable vom Grad d gibt und $\binom{d}{0}$ definitionsgemäß 1 ist. Gelte die Aussage schon für $n-1$ Variablen. Die erste Variable unseres Monoms kann irgendeinen Grad i zwischen 0 und d haben; die anderen $n-1$ Variablen haben dann zusammen noch den Grad $d-i$. Nach Induktionsvoraussetzung ist also $N = \sum_{i=0}^{d} \binom{i+n-2}{n-2}$. Diese Summe ist gleich $\binom{d+n-1}{n-1}$, und die Aussage über die Größe von A ist bewiesen.

Dass $\det(A)$ ein Polynom ist, sieht man leicht, da alle Einträge von A Polynome sind und die Determinante von A wiederum ein Polynom in ihren Einträgen ist.

Es bleibt zu zeigen, dass wenn $(\xi_0, \ldots, \xi_{n-1})$ eine gemeinsame Nullstelle des ursprünglichen Gleichungssystems $f_1 = \cdots = f_n = 0$ ist, dass dann ξ_0 eine Nullstelle der Determinante ist. Setzen wir ξ_0 in die f_i ein, so ist $(\xi_1, \ldots, \xi_{n-1})$ eine gemeinsame Nullstelle dieser Polynome. Setzen wir noch $\xi_n = 1$, so ist dann auch $F_i(\xi_0, \ldots, \xi_n) = 0$ für alle i. Setzt man das ξ_0 für alle X_0 in der Matrix ein, so erhält man eine Matrix $\widetilde{A}$ mit Einträgen aus $\mathbb{K}$. Es ist zu zeigen, dass die Determinante von $\widetilde{A}$ gleich null ist. Wir tun dies, indem wir zeigen, dass das Gleichungssystem $\widetilde{A}b = 0$ eine nichttriviale Lösung hat.

Wir betrachten den Vektor $b_1, \ldots, b_N$, indem jedes b_i das zur i-ten Spalte korrespondierende Monom unter der Einsetzung $(\xi_1, \ldots, \xi_n)$ ist. Offenbar ist dieser Vektor eine Lösung des Gleichungssystems $\widetilde{A}b = 0$, da die jeweiligen Ergebnisse ein Vielfaches der F_i ausgewertet an $(\xi_0, \ldots, \xi_n)$ darstellen. Außerdem

ist X_n^d eines der Monome, und da $\xi_n = 1$ ist, ist mindestens ein b_j ist ungleich null, und der Vektor daher keine triviale Lösung. Also ist die Determinante von $\widetilde{A}$ gleich null, d.h. $\det(A)$ hat bei $X_0 = \xi_0$ eine Nullstelle, und die Aussage ist bewiesen. □

Die Determinante von A hat zusätzlich zu den gesuchten Nullstellen ξ_0 auch noch weitere Nullstellen, die nicht zu Lösungen des Gleichungssystems korrespondieren. In [16] benennt Macaulay jedoch einen Minor A' von A, dessen Determinante (wenn sie nicht konstant null ist) ein Teiler von $\det(A)$ darstellt. Zu jeder Nullstelle von $\frac{\det(A)}{\det(A')}$ gehört dann auch eine gemeinsame, nicht triviale Nullstelle des homogenen Gleichungssystems $F_1 = \cdots = F_n = 0$. Falls ξ_n in dieser Lösung nicht null ist, so lässt sich die Lösung jeweils zu einer gewünschten Lösung des Systems $f_1 = \cdots = f_n = 0$ umformen, indem wir die ξ_i jeweils durch ξ_n dividieren.

Die Matrix A' findet man durch Streichung aller Spalten und Zeilen von A, die mit Monomen korrespondieren, die nur genau von einem $X_i^{d_i}$ geteilt werden. Wir bezeichnen diese Monome als *reduziert*:

Definition 6.4.4. Ein Monom X^α heißt reduziert, wenn es nur durch genau ein $X_i^{d_i}$ geteilt wird. Der Minor von A, der durch Streichung aller mit reduzierten Monomen korrespondierenden Spalten und Zeilen entsteht, bezeichnen wir als A'.

Wir sind also bis auf zwei Probleme am Ziel: Erstens müssen wir uns um den Fall kümmern, wenn $\det(A')$ für alle X_n gleich null ist. Zweitens müssen wir noch mit Lösungen des homogenen Gleichungssystems umgehen, deren Homogenisierungsvariable null ist (man bezeichnet diese Lösungen auch als *unendlich ferne* Lösungen). Letzteres ist in unserem Sonderfall nicht weiter problematisch, da wir die zu unendlich fernen Lösungen gehörenden ξ_0 einfach austesten können. Laufzeittechnisch macht das in der Praxis nicht viel Unterschied, da die sowieso seltenen unendlich fernen Lösungen i.A. wohl noch seltener in $\mathbb{E}_\mathbb{K}$ liegen, und nur für solche Lösungen interessieren wir uns[1].

Das erste Problem ist kritischer, da es für $\det(A') = 0$ auch passieren kann, dass $\det(A)$ selbst null ist, obwohl die Multiresultante im Allgemeinen nicht konstant null ist. Wir lösen dieses Problem mit einem Trick von Canny (siehe [2]). Ist $\det(A') = 0$, so führen wir eine neue Variable u ein und berechnen statt der Resultante von $F_1, \ldots, F_n$ die Resultante von $F_1 - uX_1^{d_1}, \ldots, F_n - uX_n^{d_n}$ wie zuvor über die Determinante von A_u, deren Einträge jetzt auch das u enthalten. Wir stellen fest:

Lemma 6.4.5. *$\det(A'_u)$ mit den u als Einträgen ist nicht konstant null und teilt $\det(A_u)$. Sei r der kleinste Exponent von u in $\det(A_u)$. Dann ist r auch der kleinste Exponent von $\det(A'_u)$, und der Quotient der beiden zugehörigen Koeffizienten ist die Multiresultante.*

[1] Genau genommen sind Lösungen aus $\mathbb{E}_\mathbb{K}$ insgesamt eine Seltenheit. Wenn eine solche Lösung auftritt, so ist dies i.A. eine Folge der speziellen, zu Grunde liegenden geometrischen Konstruktion. Hat also unser System eine durch Quadratwurzeln darstellbare Lösung, so können wir in der Praxis davon ausgehen, dass nicht zusätzlich auch noch die Koeffizienten höchsten Grades eine solche Lösung haben.

Beweis. Man überzeugt sich schnell, dass man eine Ordnung der Monome so festlegen kann, dass $A_u = A - uI_N$ und $A'_u = A' - uI_{N'}$ ist, wobei die I jeweils die Einheitsmatrizen entsprechender Größe bezeichnen. Letzterer Term ist das charakteristische Polynom von A' und somit nicht konstant null. Damit gilt für diese Resultante aber der Satz von Macaulay, und entsprechend teilt die Determinante $\det(A'_u)$ die Determinante $\det(A_u)$, und der Quotient ergibt die Multiresultante. Für die Monome dieses Quotienten mit Exponenten ≥ 1 in u interessieren wir uns nicht. Der konstante Koeffizient des Quotienten dagegen (der existiert, da die Multiresultante der F_i nicht konstant null ist) ist aber offensichtlich exakt der Quotient aus den Koeffizienten mit geringster Ordnung in $\det(A - uI_N)$ und $\det(A' - uI_{N'})$. □

6.5 Reduktion über Ortskurven

In den vorangegangenen Abschnitten haben wir uns mit klassischen Eliminationsmöglichkeiten für polynomielle Gleichungssysteme beschäftigt, ohne viel Bezug auf den geometrischen Hintergrund zu nehmen. In diesem Abschnitt führen wir eine andere Methode ein, um das Gleichungssystem zu vereinfachen. Die Grundidee ist die folgende: Unser Gleichungssystem rührt von einer geometrischen Konstruktion her. Die Variablen der Polynome sind vorgegebene Größen, Parameter semifreier Punkte oder - was uns hier vornehmlich interessiert - die beiden Koordinaten ursprünglich freier Punkte, die durch die Konstruktionsaufgabe gebunden worden sind. Sind die zu konstruierenden Größen vorgegeben, so liegen die gebundenen Punkte i.A. fest; lassen wir jedoch eine spezielle konstruierte Größe parametrisch (d.h. auf der algebraischen Seite, dass wir eine der Bestimmungsgleichungen ignorieren), so erhalten wir im Allgemeinen eine Bewegung der gebundenen Punkte auf einem eindimensionalen Unterraum. Angenommen ein solcher Punkt bewegt sich entlang einer konstruierbaren Kurve, also einem Kreis oder einer Geraden. Dann können wir dies feststellen, indem wir den freien Parameter einige Werte durchlaufen lassen und das System numerisch berechnen (die Methoden hierfür werden in Kapitel 7 beschrieben). Natürlich erhalten wir auf diese Weise nur einen Verdacht, wie die Kurve aussehen könnte; aber mit folgender Technik können wir beweisen, dass es sich um eine solche Kurve handelt, und unter Umständen Nutzen daraus ziehen.

Seien $Y_1, \ldots, Y_m$ die Parameter der Konstruktion (d.h. entweder von vornherein freie Parameter oder aber durch die Konstruktionsaufgabe frei gemachte Parameter). Seien $X_1, \ldots, X_n$ die gebundenen Parameter und

$$\begin{aligned} Y_1 &= f_1(X_1, \ldots, X_n) \\ &\vdots \\ Y_n &= f_n(X_1, \ldots, X_n) \end{aligned}$$

das Gleichungssystem zur Bestimmung der X_i. Jedes f_i ist also Element von $\mathbb{E}_{\mathbb{Q}(Y_{n+1},\ldots,Y_m,X_1,\ldots,X_n)}$.

Wir wählen jetzt zufällige Werte für die Y_i und lassen dabei numerisch eines der Y_j für $1 \leq j \leq n$ einen gewissen Wertebereich durchlaufen. Dabei testen wir, ob einer der ursprünglichen Punkte sich auf einer konstruierbaren Kurve bewegt.

Nehmen wir an, beim Durchlaufen verschiedener Werte für Y_1 gewinnen wir die Vermutung, dass (X_1, X_2) sich auf einer konstruierbaren Kurve, z.B. einem Kreis, bewegt. Wir bezeichnen mit (m_1, m_2) den Mittelpunkt dieses Kreises und mit r den Radius. Weiterhin sei λ ein Parameter, der bestimmt, wo auf dem Kreis sich ein Punkt befindet. Wenn wir annehmen, dass bei parametrischer Bewegung von Y_1 die Punkte X_1, X_2 auf einem Kreis liegen, so leiten wir folgende Vermutungen ab:

- m_1, m_2 und r hängen nicht von Y_1 ab.
- Beschreiben wir (X_1, X_2) durch die Parameter des Kreises und λ, so hängen $Y_2, \ldots, Y_n$ nicht von λ ab.

Um diese Vermutungen zu nutzen, drücken wir (X_1, X_2) durch m_1, m_2, r und λ aus. Die konkrete Parametrisierung hat folgende Form:

- Für Kreise: $(X_1, X_2) = g(m_1, m_2, r, \lambda) = \left(m_1 + r \cdot \frac{1-\lambda^2}{1+\lambda^2}, m_2 + r \cdot \frac{2\lambda}{1+\lambda^2}\right)$
- Für Geraden: $(X_1, X_2) = g(m, r, \lambda) = (\lambda, m + \lambda r)$

Im Falle der Geraden kann es passieren, dass sich (X_1, X_2) entlang einer Geraden parallel zur 2. Achse des Koordinatensystems bewegt, was die obige Darstellung unmöglich macht. Wir können aber der numerischen Näherung der Geraden leicht ansehen, ob die Steigung einen Betrag größer als 1 hat; in diesem Fall vertauschen wir einfach die Achsen und wählen $g(m, r, \lambda) = (m + \lambda r, \lambda)$, was das Problem löst.

Im Folgenden werden wir der einfacheren Darstellung wegen nicht beide Fälle durchgehen, sondern den Fall voraussetzen, dass sich (X_1, X_2) auf einem Kreis bewegt. Wir setzen $g(m_1, m_2, r, \lambda)$ in die Gleichungen $f_1, \ldots, f_n$ ein und betrachten das Gleichungssystem

$$\begin{aligned} Y_1 &= f_1(g(m_1, m_2, r, \lambda), X_3, \ldots, X_n) \\ &\vdots \\ Y_n &= f_n(g(m_1, m_2, r, \lambda), X_3, \ldots, X_n) \end{aligned}$$

Wir formen das Gleichungssystem (in dem bisher noch Wurzeln und Nenner vorkommen) in ein polynomielles Gleichungssystem um und erhalten polynomielle Gleichungen $F_1, \ldots, F_n \in \mathbb{Q}(Y_1, \ldots, Y_m)[m_1, m_2, r, \lambda, X_3, \ldots, X_n]$ mit $F_1 = \cdots = F_n = 0$.

Der folgende Satz hilft uns, unsere durch die Numerik gewonnenen Annahmen nutzbar zu machen.

Satz 6.5.1. *Seien $F_1, \ldots, F_n$ wie oben. Gebe es konstruierbare Parameter m_1, m_2, r und λ bzw. m, r und λ so, dass der Punkt $(X_1, X_2) = g(m_1, m_2, r, \lambda)$ bzw. $(X_1, X_2) = g(m, r, \lambda)$ mit einem der oben genannten g ist. Seien weiterhin die Parameter bis auf λ unabhängig von Y_1 und $Y_2, \ldots, Y_n$ unabhängig von λ. Dann gilt:*

1. *Beim Einsetzen jedes Elementes aus $\mathbb{Q}$ in Y_1 sind die Lösungen für die Parameter der Gleichung identisch.*

2. *Betrachten wir die Polynome $F_2, \ldots, F_n$ als Polynome über λ und existiert eine Lösung, so sind die Koeffizienten aller Gleichungen $F_2, \ldots, F_n$ aufgefasst als Polynome in λ für diese Lösung null.*

Beweis. Im Wesentlichen ist der Satz nur eine Formalisierung der vorher schon genannten Tatsachen. Da Y_1 unabhängig von den Parametern der Lösung ist, müssen die gleichen Lösungen herauskommen, egal welches Y_1 wir vorgeben. Das Y_1 bestimmt nur die Lage des Punktes auf der Ortskurve, die Parameter der Ortskurve selbst sind unabhängig davon.

Da die anderen Gleichungen unabhängig von λ sind, müssen alle Polynome $F_2, \ldots, F_n$ für jede Einsetzung von λ null werden, d.h. jeder Koeffizient vor λ muss null sein. □

Die Gleichungen $F_2, \ldots, F_n$ splitten sich durch diesen Satz in eine Vielzahl von Gleichungen auf, nämlich in die Koeffizienten der Gleichungen in λ.

Gleichzeitig erlaubt uns der Satz, zur Berechnung von m_1, m_2, r und λ in der obersten Gleichung den Parameter Y_1 durch ein beliebiges Element von $\mathbb{Q}$ zu ersetzen. Wir können durch verschiedene Ersetzungen beliebig viele verschiedene Gleichungen erzeugen. Abgesehen von Faktoren von F_1, in denen Y_1 nicht vorkommt, können wir die erste Gleichung völlig außer Acht lassen.

Beim Kreis als Ortskurve werden wir als eine Lösung für r immer null erhalten (m_1 und m_2 liegen dann so, dass genau die alte Lage von X_1 und X_2 herauskommt, während λ bedeutungslos wird). Durch die Numerik haben wir die Annahme gewonnen, dass sich (X_1, X_2) auf einem Kreis von positivem Radius bewegt, d.h. die Lösung $r = 0$ können wir ignorieren.

Wir haben auf diese Weise keine Variable des Gleichungssystems eliminiert, sondern einen Parameter des Grundkörpers, und i.A. die Anzahl der Gleichungen erhöht. Es ist nicht unmittelbar klar, warum diese Umformungen überhaupt einen Gewinn bringen sollen. Tatsächlich beruht dies auf der Annahme, dass der eliminierte Parameter das Problem grundlegend vereinfacht, und die höhere Anzahl an Gleichungen eine 'Auswahl' von simpleren Gleichungen ermöglicht.

Desweiteren erhalten wir ein interessantes Seitenprodukt: Finden wir tatsächlich eine Lösung für das modifizierte Gleichungssystem, so beweist dies unsere ursprüngliche Annahme, dass ein Punkt sich unter Freilassung eines bestimmten Parameters auf einer von diesem Parameter unabhängigen Ortskurve bewegt. Wir lassen den Computer an dieser Stelle also einen nicht vorher fest verankerten Satz selbstständig finden, durch einen dem menschlichen Ausprobieren ähnlichen Prozess (der numerischen Überprüfung) verifizieren und schließlich symbolisch beweisen. Auf diese Weise gehen wir einen Schritt weiter

in den mathematischen Fähigkeiten eines Programmes: Wir haben ein Programm, das nicht nur ausprobiert oder automatisch beweist, sondern selbstständig neue und für eine größere Aufgabe relevante Sätze findet. Natürlich findet dies nur in einem kleinen, abgegrenzten Bereich statt und muss auch nicht notwendig zu einer Vereinfachung führen, es bleibt aber nichtsdestoweniger ein interessanter Umstand. Darüber hinaus verbessern wir in gewisser Weise die Qualität unserer Ausgabe; finden wir eine zweite Ortskurve, auf der (X_1, X_2) liegt, wenn man einen anderen Parameter frei lässt, so lässt sich die Position von (X_1, X_2) als Schnitt zweier konstruierbarer Ortskurven angeben, was im Allgemeinen konstruktionstechnisch leichter durchführbar sein sollte als die Umsetzung eines Ausdrucks mit Körperoperationen und Quadratwurzeln (obwohl es natürlich sein kann, dass der Mittelpunkt des Kreises bzw. die Punkte zur Fixierung einer Gerade ihrerseits kompliziert zu konstruieren sind).

Wir wollen das ganze Verfahren an einem Beispiel verdeutlichen.

Beispiel 6.5.2. Wir betrachten ein Dreieck aus den Punkten $A = (0, 0)$, $B = (1, 0)$ und $C = (a, b)$. Wir konstruieren die Seitenhalbierende über $\overline{BC}$, indem wir den Abstand des Mittelpunktes $M = \left(\frac{a+1}{2}, \frac{b}{2}\right)$ zu A berechnen. Weiterhin konstruieren wir die Höhe h des Dreiecks über $\overline{AB}$. Die Konstruktionsaufgabe soll jetzt daraus bestehen, den Punkt C aus der Seitenhalbierenden s und der Höhe h zu bestimmen.

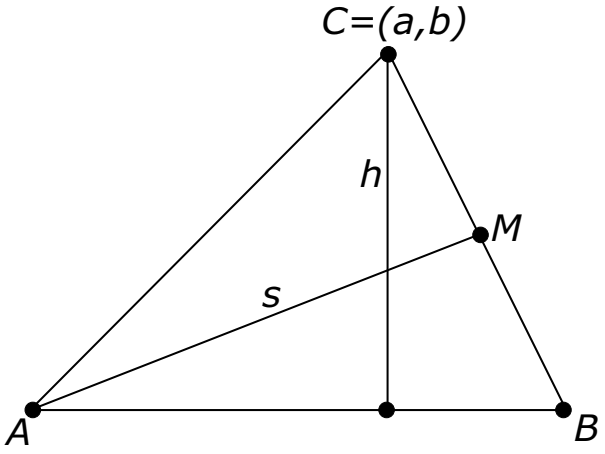

Abbildung 6.1: Darstellung der Beispielaufgabe
Es soll C aus s und h konstruiert werden.

Schreiben wir die Konstruktion als Gleichungssystem in den freien Parametern a und b, so erhalten wir

$$\begin{aligned} s &= \sqrt{\left(\frac{a+1}{2}\right)^2 + \left(\frac{b}{2}\right)^2} \\ h &= b \end{aligned}$$

Natürlich ist dies ein einfaches Beispiel, in dem es kein Problem wäre, die untere Gleichung in die obere einzusetzen, die Wurzel aufzulösen und die verbleibende quadratische Gleichung zu lösen. Wir erhielten dann $C = (\sqrt{4s^2 - h^2} - 1, h)$ und hätten damit die Konstruktionsaufgabe gelöst. Wir wollen aber zur Verdeutlichung die Methode der Ortskurve anwenden.

Halten wir den Parameter h fest und durchlaufen einige Werte für s, so bewegt sich C auf einer Geraden (nämlich der Parallelen zur X-Achse in einem Abstand von h). Wir setzen $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda \\ m + \lambda r \end{pmatrix}$ und erhalten das Gleichungssystem

$$\begin{aligned} s &= \sqrt{\left(\frac{\lambda + 1}{2}\right)^2 + \left(\frac{m + \lambda r}{2}\right)^2} \\ h &= m + \lambda r \end{aligned}$$

Die obere Gleichung lassen wir außer Acht, während wir die andere in die polynomielle Gleichung $(m - h) + \lambda r = 0$ umformen. Gemäß Satz 6.5.1 ist jeder Koeffizient bezüglich λ eine eigene Gleichung, die null werden muss, d.h. wir haben $m - h = r = 0$ und haben damit die Gerade bereits vollständig bestimmt.

Wenden wir uns der interessanteren Ortskurve zu, nämlich der Ortskurve, die wir erhalten, wenn wir verschiedene Werte für h wählen und s konstant lassen. Wir stellen fest, dass sich C dann auf einem Kreis bewegt. Während die zuvor berechnete Ortskurve keine Überraschung war, ist diese Beobachtung nicht ganz trivial; mit bloßem Auge beobachtet man, dass der Mittelpunkt des Kreises weder auf A noch auf B liegt, und dass der Radius keine der bisher konstruierten Strecken darstellt.

Wir setzen jetzt die Ortskurve als Darstellung der Koordinaten von C ein und erhalten $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} m_1 + r \cdot \frac{1-\lambda^2}{1+\lambda^2} \\ m_2 + r \cdot \frac{2\lambda}{1+\lambda^2} \end{pmatrix}$. Dies ergibt dann das Gleichungssystem

$$\begin{aligned} s &= \sqrt{\left(\frac{m_1 + r \cdot \frac{1-\lambda^2}{1+\lambda^2} + 1}{2}\right)^2 + \left(\frac{m_2 + r \cdot \frac{2\lambda}{1+\lambda^2}}{2}\right)^2} \\ h &= m_2 + r \cdot \frac{2\lambda}{1 + \lambda^2} \end{aligned}$$

Eliminieren wir in der oberen Gleichung die Wurzel, so erhalten wir als polynomielle Gleichung:

$$\begin{aligned} 0 = \big(r^2 - 2r - 2m_1 r + m_1^2 + 2m_1 + m_2^2 - 4s^2 + 1\big) \quad & \lambda^4 \quad + \\ (4m_2 r) \quad & \lambda^3 \quad + \\ \big(2r^2 + 2m_1^2 + 4m_1 + 2m_2^2 - 8s^2 + 2\big) \quad & \lambda^2 \quad + \\ (4m_2 r) \quad & \lambda \quad + \\ \big(r^2 + 2r + 2m_1 r + m_1^2 + 2m_1 + m_2^2 - 4s^2 + 1\big) & \end{aligned}$$

Jeder einzelne Koeffizient ist eine eigene Gleichung zur Bestimmung der Parameter des Kreises. Durch den Parameter bei λ und λ^3 bietet sich wie erwartet $r = 0$ als Lösung an, die wir aber ausschließen können; also wissen wir, dass $m_2 = 0$ ist. Setzen wir dies ein, so erhalten wir für die Koeffizienten bei λ^4, λ^2 und λ^0:

$$\begin{aligned} 0 &= \big(r^2 - 2r - 2m_1 r + m_1^2 + 2m_1 - 4s^2 + 1\big) \\ 0 &= \big(2r^2 + 2m_1^2 + 4m_1 - 8s^2 + 2\big) \\ 0 &= \big(r^2 + 2r + 2m_1 r + m_1^2 + 2m_1 - 4s^2 + 1\big) \end{aligned}$$

Subtrahieren wir die Hälfte der mittleren Gleichung von der untersten, so erhalten wir $0 = 2r(1 + m_1)$. Da wir $r = 0$ ausschließen können, erhalten wir $m_1 = -1$ und bei Einsetzung z.B. in die mittlere Gleichung $r = 2s$. Die Berechnung von r und m_1 kann natürlich genauso einfach über Resultanten erfolgen, oder auch durch Auflösen einer Gleichung nach einer der beiden Größen.

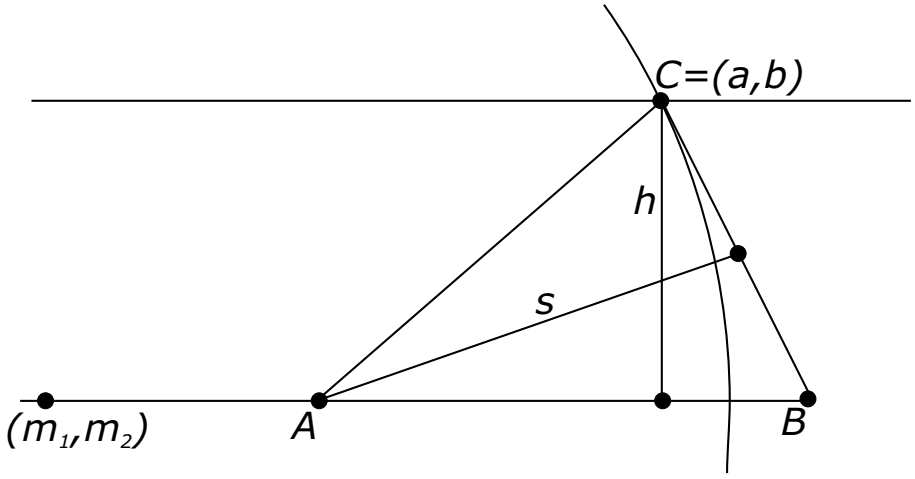

Abbildung 6.2: Lösung der Beispielaufgabe
C liegt auf dem Schnitt einer Parallelen zu $\overline{AB}$ und einem Kreis um (m_1, m_2) mit Radius $2s$

Wir haben also festgestellt, dass der Punkt C sich bei fixierter Höhe entlang einer Parallelen zur X-Achse mit Abstand h und bei fixierter Seitenhalbierenden

entlang eines Kreises um einen Punkt auf der Verlängerung von A und B im selben Abstand von A wie B bewegt, dessen Radius das Doppelte der Länge der Seitenhalbierende ist. Sind beide Größen gegeben, befindet sich der Punkt auf dem Schnittpunkt dieser beiden Kurven, d.h. es ist gar nicht mehr nötig, das jeweilige λ noch zu berechnen.

Der Rechenaufwand für diese Methode ist etwas größer als er für ein direktes Einsetzen wäre, dafür erhalten wir aber eine deutlich schönere Lösung als die schlichte Angabe $C = (\sqrt{4s^2 - h^2} - 1, h)$, die wir bei einer direkten Berechnung erhalten hätten.

Kapitel 7

Numerische Umkehrung

Bisher haben wir uns mit symbolischen Lösungen von Konstruktionsaufgaben beschäftigt. Diese exakte Rechnung liefert eine Konstruktionsvorschrift für die Umkehrung, was natürlich aus mathematischer Sicht ein sehr interessanter Aspekt ist. Darüber hinaus hat es aber auch einen großen Reiz, eine solche Umkehrung direkt am Bildschirm verfolgen zu können.

Dementsprechend beschäftigt sich dieses Kapitel mit der rückwirkenden Dynamik, mit deren Hilfe ermöglicht werden soll, auch ursprünglich abhängig konstruierte Punkte auf der Zeichenoberfläche zu bewegen. Die Konstruktion soll dabei natürlich erhalten bleiben, so dass einige ursprünglich freie Parameter vom System so angepasst werden müssen, dass die erwünschte Position von den abhängigen Punkte eingenommen wird. Diese Eigenschaft der rückwirkenden Dynamik ist auf dem Gebiet der dynamischen Geometriesysteme völlig neu.

Es ist kaum genug zu betonen, welch hohe Bedeutung die rückwirkende Dynamik für die Schaffung von Einsicht in Konstruktionen hat. Sie macht oft Lösungen offensichtlich, die man sonst erst mühsam erarbeiten müssten. Dies gilt natürlich einmal bei der Suche nach Konstruktionsvorschriften, wie wir es zum Beispiel in Abschnitt 6.5 gesehen haben; konstruiert man ein Dreieck und in diesem Dreieck eine Seitenhalbierende und eine Höhe, fixiert danach die Höhe und bewegt dafür den zugehörigen Dreieckspunkt semifrei, so fällt unmittelbar ins Auge, dass sich der Punkt entlang einer Geraden bewegt. Genauso verhält es sich, wenn wir den selben Punkt unter Fixierung der Seitenhalbierende bewegen: Jetzt bewegt der Punkt sich leicht erkennbar auf einem Kreis. Es ist dann natürlich deutlich einfacher, zu erraten, durch welche Parameter diese Kurven wieder bestimmt sein könnte, als ohne diese Hilfsmittel zu begreifen, auf welche Weise ein Dreieck aus einer Grundseite, einer Höhe und einer Seitenhalbierenden zu konstruieren ist.

Man erinnere sich aber auch an das Beispiel aus der Einleitung, in der die rückwirkende Bewegung des Höhenschnittpunktes eines Dreiecks beschrieben wurde (siehe Abbildung 7.1). Hier hatten wir ein Dreieck und den Schnittpunkt der Höhen konstruiert. Bewegte man jetzt diesen Schnittpunkt und lässt dafür die Bewegung eines der Eckpunkte durch das System bestimmen, so fällt schnell ins Auge, dass sich dieser Eckpunkt jetzt seinerseits wie der Höhenschnittpunkt

verhält - was im Nachhinein dann auch rechnerisch sofort einsichtig wird. Wir erkennen auf diese Weise also mit Hilfe der rückwirkenden Dynamik geometrische Sätze, die uns ohne eine dynamische Betrachtung wahrscheinlich nicht aufgefallen wären.

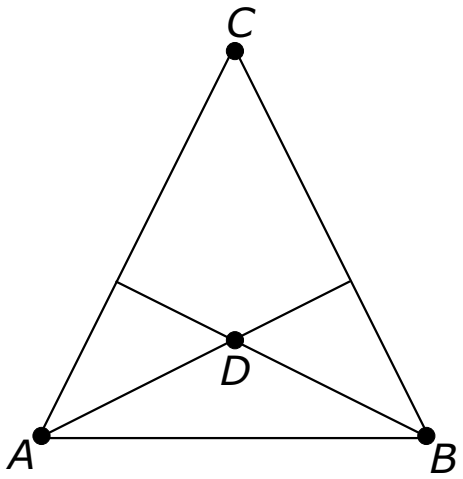

Abbildung 7.1: Bewegung des Höhenschnittpunktes

Darüber hinaus ist natürlich oft dann noch eine lokale Umkehrung der Konstruktion möglich, wenn eine Lösung des symbolischen Konstruktionsproblems nicht besteht; dann nämlich, wenn das korrespondierende Gleichungssystem zwar eine Lösung hat, diese sich aber nicht durch Quadratwurzeln ausdrücken lässt. In diesem Teilaspekt ist die rückwirkende Dynamik also eine Erweiterung der symbolischen Umkehrung.

Diese Kapitel beschäftigt sich mit der für die rückwirkende Dynamik angewandten Methodik. Wir führen aus, wie wir hierfür das Newtonverfahren anpassen, und motivieren durch die Dynamik eine bekannte Erweiterung des Newtonverfahrens neu, mit der das Verfahren unabhängig von der Wahl eines günstigen Startwertes wird.

7.1 Methodik der rückwirkenden Dynamik

Erinnern wir uns an die grundsätzliche Aufgabenstellung aus Kapitel 2. Wir hatten in einer Konstruktion zwischen freien Parametern und abhängigen Parameter unterschieden; zur Vorbereitung der rückwirkenden Dynamik tauscht dann der Benutzer den Status zweier Parameter aus, so dass ein zuvor freier Parameter jetzt vom System kontrolliert wird und andererseits ein zuvor abhängiger Parameter vom Benutzer festgelegt werden darf. Die ursprüngliche Konstruktion haben wir mit einer Funktion $f \in \mathbb{E}_{\mathbb{Q}}^n \to \mathbb{E}_{\mathbb{Q}}^m$ bezeichnet, die wir dann auf die Eingabewerte bzw. die Zielwerte einschränkten, die für die Umkehrung von Bedeutung waren. Wir erhielten dadurch schließlich folgendes Gleichungssystem:

$$F(\xi_1, \ldots, \xi_k) - (\eta_1, \ldots, \eta_k) = 0$$

Wir wollen die Funktion auf der linken Seite mit $g := F(X_i, \ldots, X_k) - (\eta_1, \ldots, \eta_k)$ bezeichnen. Im Folgenden schreiben wir der Einfachheit halber nur noch X für den Vektor $(X_1, \ldots, X_k)$.

Für die symbolische Umkehrung mussten wir diese Funktion explizit bestimmen, d.h. wir mussten F als ein System von rationalen Funktionen darstellen, in denen eventuell auch Wurzelzeichen vorkamen. Natürlich wird diese Repräsentation schnell groß und entsprechend aufwändig zu berechnen. Die numerische Umkehrung sollte dagegen so weit wie möglich in Echtzeit ablaufen, damit die Bewegung als solche beobachtbar bleibt. Diesem Ziel ist es sehr abträglich, wenn wir für die Dynamik das symbolische Gleichungssystem jedesmal aufs Neue erstellen müssen. Natürlich könnte diese Berechnung einmal vor Beginn der Bewegung (also bei der Neuverteilung der Freiheitsgrade durch den Benutzer) erfolgen, aber selbst dann noch ist die Auswertung dieser Funktionen langwieriger, als wenn wir eine direkte Berechnung der Koordinaten durch eine Hintereinanderausführung der einzelnen geometrischen Funktionen durchführen.

Diese direkte Auswertung der Funktion $F(X) = (f_1(X), \ldots, f_k(X))$ steht uns zudem natürlich sowieso schon zur Verfügung, denn auch bei der normalen Dynamik werden zur Berechnung der abhängigen Größen die Koordinaten aller Punkte sukzessive (in der Reihenfolge ihrer Eingabe) ausgerechnet. Wir benötigen allerdings nicht nur die Funktion g: Denn zur Lösung des Nullstellenproblems $g(X) = F(X) - (\eta_1, \ldots, \eta_k) = 0$ bietet es sich an, das Newtonverfahren zu verwenden, und dafür benötigen wir nicht nur die Funktion g selbst, sondern auch noch die Ableitungen von g nach den jeweiligen Parametern. In den einzelnen Komponenten unterscheidet sich g dabei nur durch den konstanten Summanden von F; entsprechend sind die partiellen Ableitungen von g identisch mit den partiellen Ableitungen von F. Aus diesen Ableitungen bilden wir dann die *Jacobi-Matrix*, d.h. die Matrix

$$J := \begin{pmatrix} \frac{\partial f_1}{\partial X_1} & \cdots & \frac{\partial f_k}{\partial X_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial X_k} & \cdots & \frac{\partial f_k}{\partial X_k} \end{pmatrix}$$

Um die partiellen Ableitungen zu erhalten, erweitern wir die Berechnung der Koordinaten aller Punkte zusätzlich noch um die Berechnung der Ableitung dieser Koordinaten nach jedem der gebundenen Parameter. Hierfür benötigen wir jeweils die Koordinaten und die gewünschten Ableitungen für alle vorangegangenen Punkte, um die Ableitung der neuen Größe nach Anwendung der Kettenregel für die Ableitung von Hintereinanderausführungen von Funktionen anwenden zu können.

Beispiel 7.1.1. Um die Berechnung klar zu machen, wollen wir noch einmal das Beispiel aus Abbildung 7.1 bemühen. Aus den sechs ursprünglich freien Parametern a_1, a_2, b_1, b_2, c_1 und c_2 (den Koordinaten der drei Eckpunkte A, B und C) berechnen wir die Koordinaten $d_1(a_1, \ldots, c_2)$ und $d_2(a_1, \ldots, c_2)$ des Höhenschnittpunktes D. Für die Umkehrung haben wir die Koordinaten c_1 und

c_2 gebunden und dafür d_1 und d_2 befreit. Die für das Newtonverfahren benötigte Jacobi-Matrix hat also die Form

$$J = \begin{pmatrix} \frac{\partial d_1(a_1,\ldots,c_2)}{\partial c_1} & \frac{\partial d_2(a_1,\ldots,c_2)}{\partial c_1} \\ \frac{\partial d_1(a_1,\ldots,c_2)}{\partial c_2} & \frac{\partial d_2(a_1,\ldots,c_2)}{\partial c_2} \end{pmatrix}$$

Wir berechnen die Koordinaten aller Punkte in der Reihenfolge der Punkte. Für die Koordinaten des Punktes A sind die Koordinaten direkt als Argumente gegeben. Zusätzlich bestimmen wir die partiellen Ableitungen der Koordinaten von A nach c_1 und c_2, wobei diese hier aber natürlich null sind. Für B erhalten wir die Koordinaten genauso, und auch hier sind alle partiellen Ableitungen null. Für C erhalten wir die Koordinaten ebenfalls direkt, während die vier Ableitungen entweder null oder eins sind: $\frac{\partial c_1}{\partial c_1}$ und $\frac{\partial c_2}{\partial c_2}$ sind jeweils eins, $\frac{\partial c_1}{\partial c_2}$ und $\frac{\partial c_2}{\partial c_1}$ dagegen null.

Die Koordinaten des Punktes D sind $D = (d_1(a_1,\ldots,c_2), d_2(a_1,\ldots,c_2))$. Für die Ableitungen nutzen wir die Ableitungsregel für Ableitungen von Hintereinanderausführung, also z.B. für d_1 nach c_1

$$\frac{\partial d_1(a_1,\ldots,c_2)}{\partial c_1} = \frac{\partial c_1}{\partial c_1} \cdot \frac{\partial d_1}{\partial c_1}(a_1,\ldots,c_2)$$

Die Ableitung $\frac{\partial c_1}{\partial c_1}$ haben wir vorher schon berechnet (in unserem Beispiel ist das natürlich trivial, aber in einer allgemeinen Situation kann hier auch die Ableitung einer komplexeren Funktion stehen). Genauso berechnen wir auch die anderen drei partiellen Ableitungen für die Jacobi-Matrix.

Auf diese Weise haben wir also eine Möglichkeit, die Funktion g und die Jacobi-Matrix effizient zu berechnen. Findet jetzt eine Bewegung auf dem Bildschirm statt, so approximieren wir mit Hilfe des Newtonverfahrens die gebundenen Parameter so, dass alle befreiten Parameter auf den gewünschten Positionen zu liegen kommen.

Beim Newtonverfahren beginnen wir mit einem Startvektor $X^{(0)}$ und definieren eine Folge von Vektoren $X^{(i)}$ rekursiv über den sogenannten *Newton-Schritt*:

$$X^{(i+1)} = X^{(i)} - J(X^{(i)})^{-1} g(X^{(i)})$$

Die Matrix $J(X^{(i)})$ notiert dabei die Jacobi-Matrix an der Stelle $X^{(i)}$. Wir berechnen also für jedes $X^{(i)}$ den Wert $g(X^{(i)})$ und die Jacobi-Matrix, invertieren diese und ziehen das Produkt von der letzten Position $X^{(i)}$ ab. Ist der Startwert $X^{(0)}$ günstig gewählt, d.h. befindet er sich in einem bestimmten Gebiet um eine Nullstelle der Funktion g, so konvergiert die Folge $X^{(i)}$ gegen diese Nullstelle. Man bezeichnet diesen Bereich um die Nullstelle herum als *Konvergenzbereich* der Nullstelle. Nähere Information über das Newtonverfahren finden sich z.B. in [27].

Das Newtonverfahren ist für unsere Anwendung auch deshalb so günstig, da wir in den meisten Fällen einen sehr guten Startwert $X^{(0)}$ zur Verfügung haben. Denn die Bewegungen der Punkte auf dem Bildschirm gehen meist nur über relativ kleine Entfernungen, weshalb die jeweils vorangegangene Situation

in der Regel ein solcher guter Startwert ist. Gehen wir also davon aus, dass der Benutzer nur kleine Bewegungen ausführt, wird das Newtonverfahren mit der alten Situation als Startwert immer sehr schnell hinreichend gut konvergieren (in der Praxis meist in zwei bis drei Schritten).

Falls der Benutzer allerdings eine sehr ruckartige Bewegung vollführt, kann es passieren, dass die vorangegangene Stellung nicht mehr im Konvergenzbereich des Newtonverfahrens liegt. Hier bietet sich auf elementare Weise eine Methode an, die nicht auf gute Startwerte für das Newtonverfahren angewiesen ist: Anstatt die Änderung der Parameter von der bekannten Situation zur Zielsituation in einem Schritt durchzuführen, interpolieren wir die Punkte auf der Bewegungslinie und approximieren die Lösungen dazwischen jeweils mit dem Newtonverfahren (ohne direkten Bezug zur Geometrie ist diese Methode zur Verstärkung des Netwonverfahrens z.B. in [1] zu finden).

Bezeichnen wir mit $g_0 := F(X) - (\eta'_1, \ldots, \eta'_k)$ die Funktion bezüglich der letzten bekannten Situation und mit $g_1 := F(X_i, \ldots, X_k) - (\eta_1, \ldots, \eta_k) = g$ die Funktion bezüglich der neuen Situation, von der wir eine Nullstelle finden wollen. Jetzt definieren wir für $0 \leq \lambda \leq 1$:

$$g_\lambda(X) := (1 - \lambda)g_0 + \lambda g_1$$

offenbar ist die Funktion für $\lambda = 0$ und $\lambda = 1$ wohldefiniert. Für λ dicht bei null haben wir einen guten Startwert für das Newtonverfahren, denn eine Nullstelle von g_0 ist uns durch die letzte bekannte Position gegeben. Lassen wir also λ in hinreichend kleinen Schritten anwachsen, so können wir jeweils eine Nullstelle von g_λ mit dem letzten Schritt als Startwert finden, bis wir schließlich $\lambda = 1$ erreichen und unsere gesuchte Funktion $g_1 = g$ approximieren.

In der Bewegung auf dem Bildschirm entspricht dieser Vorgang einem Abschreiten der Verbindungsstrecke von der alten Lage eines Punktes zur neuen Lage in kleinen Schritten. Natürlich müssen wir nicht zwingend den kürzesten Weg zwischen den beiden Punkten nehmen, sondern können stattdessen λ auch entlang jedes anderen Pfades von null nach eins bewegen. Zu jedem solchen Pfad gehört dann auch eine Familie von Funktionen g_λ und ein stetiger Pfad von Nullstellen dieser Funktionen. Hat unsere die Konstruktion beschreibenden Funktion F auf diesem Pfad keine Singularitäten, so zeigt das folgende Lemma, dass es eine hinreichend kleine Schrittweite größer null gibt, mit der die Zwischenlösungen in jedem Schritt im Konvergenzbereich der jeweils nachfolgenden Funktion sind:

Lemma 7.1.2. *Sei g_λ wie oben und X_λ jeweils eine Nullstelle von g_λ.*

Sei $p : [0, 1] \to \mathbb{C}$ ein Pfad von 0 nach 1 und $p'(t) := X_{p(t)}$ der zugehörige stetige Pfad von Nullstellen von $g_{p(t)}$.

Falls F auf diesem Pfad p' keine Singularität hat, dann existiert eine Folge von Werten $t_1, \ldots, t_l \in [0, 1]$ mit $t_1 = 0$ und $t_l = 1$ so, dass $X_{p(t_i)}$ jeweils im Konvergenzbereich des Newtonverfahrens für $g_{p(t_{i+1})}$ liegt.

Beweis. Da F auf p' keine Singularität hat, gilt das auch für g_λ, insbesondere an der Stelle X_λ. Also hat g_λ um die Nullstelle X_λ einen Konvergenzradius echt größer null, den wir mit μ_λ bezeichnen. Sei $\mu := \inf_{0 \leq \lambda \leq 1}(\mu_\lambda)$. Da jedes Element

der Menge größer null und die Menge kompakt ist, so ist folglich auch $\mu > 0$. Zerteilt man den Pfad p' in μ große Stücke und nimmt von jedem Schritt den zugehörigen Ursprungswert in p, so erhält man die geforderte Folge. □

Praktisch kennen wir diese Schrittweite μ nicht. Außerdem wäre es auch nicht nötig, den ganzen Pfad entlang in so kleinen Schritten zurückzulegen (bedingt durch die kleinen Bewegungsweiten in der Bewegung auf dem Bildschirm reicht, wie gesagt, meistens sogar ein einziger Schritt). Wir suchen den richtigen Wert für das μ mit einer logarithmischen Suche; zunächst versuchen wir direkt, von der bekannten Nullstelle von g_0 aus direkt g_1 zu approximieren. Gelingt uns das nicht in einer festgelegten Anzahl von Schritten, so suchen wir zunächst eine Nullstelle von $g_{\frac{1}{2}}$, dann von $g_{\frac{1}{4}}$ und so weiter, bis wir schließlich nach n Halbierungen die Nullstelle von $g_{2^{-n}}$ finden. Von dieser Stelle aus gehen wir jetzt wieder genauso vor, wobei wir hier bereits mit einem n starten, das um eins kleiner ist als das erfolgreiche n des letzten Schrittes.

Beschränkt man sich auf reellwertige Pfade, so kann es natürlich vorkommen, dass wir auf dem Pfad der Nullstellen auf eine Singularität treffen. Lassen wir aber auch Pfade von null nach eins durch die komplexen Zahlen zu, so finden wir immer Pfade, auf denen keine Singularitäten vorkommen. Bei jedem speziellen Pfad könnte man natürlich unglücklich auf eine solche Singularität treffen; in unserem Algorithmus für die rückwirkende Dynamik probieren wir daher eine Familie von Pfaden aus, falls der erste Pfad nicht zum Erfolg führt.

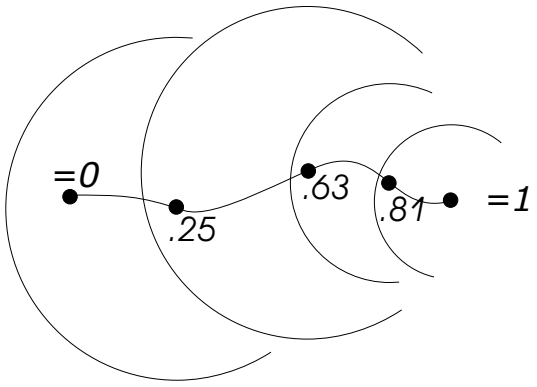

Abbildung 7.2: Schrittweise Annährung an die Nullstelle von g_1
Alle eingezeichneten Punkte sind jeweils Nullstellen von g_λ für das jeweils am Punkt notierte λ. Der Punkt für $\lambda = 0$ ist zu Beginn bekannt. Die Kreise um die Punkte deuten die jeweiligen zugehörigen Konvergenzradien der Nullstellen von g_λ an.

Abbildung 7.2 stellt den Prozess an einem Beispiel schematisch dar. Ganz links liegt die Nullstelle von $g_0 = F(X) - (\eta'_1, \ldots, \eta'_k)$, die durch die vorangegangene Situation bekannt ist. Der Punkt ganz rechts ist die gesuchte Nullstelle von $g_1 = F(X) - (\eta'_1, \ldots, \eta'_k)$. Um jeden eingezeichneten Punkt herum ist der Konvergenzradius der eingezeichneten Nullstelle von g_λ angedeutet. In dem dargestellten Fall wurde zunächst mit dem Punkt $\lambda = 0$ als Startwert eine

Nullstelle von g_1 gesucht; da der Konvergenzbereich den Punkt $\lambda = 0$ aber nicht einschließt, schlug dieser Versuch fehl, ebenso wie die Suche nach einer Nullstelle von $g_{0.5}$. Von $g_{0.25}$ schließlich wurde eine Nullstelle gefunden. Von dort aus wurde eine Nullstelle von $g_{0.625}$ (der Hälfte der verbleibenden Strecke zwischen 0.25 und 1.0) gesucht und gefunden. Mit dieser Stelle als Startwert wurde zunächst wieder $g_1 = 0$ zu lösen versucht, der Konvergenzradius des Punktes bei $\lambda = 1$ reichte aber nach wie vor noch nicht aus. Allerdings lag der letzte Punkt im Konvergenzradius von $g_{0.8125}$ (der Mitte zwischen 0.625 und 1.0), und von hier aus schließlich wurde auch die Nullstelle von g_1 gefunden.

Kapitel 8

Zusammenfassung und Ausblick

8.1 Zusammenfassung

Wir haben in dieser Arbeit das Konstruktionsproblem formuliert und eine numerische und eine symbolische Lösung dieses Problems vorgestellt. Die numerische Lösung, die so genannte rückwirkende Dynamik, ermöglicht es, auf einer interaktiven Zeichenoberfläche alle Punkte - sowohl Ursprungspunkte der Konstruktion als auch abhängige Punkte - frei zu bewegen und die resultierende Bewegung der übrigen Punkte zu beobachten. Cedric ist unseres Wissens das erste dynamische Geometriesystem, das eine solche rückwirkende Dynamik anbietet.

Für die exakte Lösung haben wir einen Algorithmus vorgestellt, der isolierte Nullstellen eines Gleichungssystems findet, die durch Quadratwurzeln ausdrückbar sind. Das Gleichungssystem kann dabei selbst Quadratwurzeln enthalten.

Wir haben den Euklidischen Körper vorgestellt und einen Darstellungssatz formuliert, der uns eine minimale Repräsentation in der Anzahl verschiedener Wurzeln von Elementen des Euklidischen Körpers erlaubt. Wir haben für diese Darstellung ein Eindeutigkeitsresultat für die meisten der Elemente dieses Körpers gezeigt.

Zur Lösung des Gleichungssystems haben wir drei klassische Methoden angepasst und eine weitere, neue Methode eingeführt, um das Gleichungssystem auf die Lösung eines univariaten Polynoms zurückzuführen.

Für die Lösung des univariaten Problems haben wir einen Algorithmus vorgestellt, der über dem Grundkörper nur die Faktorisierung benötigt und dann alle durch Quadratwurzeln ausdrückbaren Nullstellen eines vorgegebenen Polynoms finden kann.

Wir haben gezeigt, dass dieser Algorithmus bei polynomieller Faktorisierung für Polynome vom Grad d maximal $O(d^{\log(d)})$ Schritte benötigt und im Regelfall - bei vollständigen Elementen - sogar nur d Faktorisierungen eines Polynoms vom Grad kleiner oder gleich d^2 benötigt, also in polynomieller Zeit läuft. Bei exponentieller Faktorisierung ist die Laufzeit in jedem Fall durch diesen Anteil dominiert.

8.2 Ausblick

Es ist eine interessante Frage, inwieweit sich der Algorithmus effizient auch auf die Suche nach durch andere Radikale ausdrückbaren Nullstellen erweitern lässt. Dafür bedarf es eines komplizierteren Kombinationspolynoms, dessen Berechnung aber grundsätzlich mit den in Abschnitt 5.3.1 eingeführten Methoden möglich ist. Es wäre interessant, ob man dem Polynom schnell ansehen kann, welches Kombinationspolynom man zuerst anwenden sollte.

Unter Umständen lässt sich dieselbe Methode auch auf nicht einfache Radikale erweitern. Im Prinzip ist das Quadratwurzelzeichen $\sqrt{a}$ eine abkürzende Schreibweise für alle Nullstellen von $X^2 - a$, und ebenso jedes n-te Radikal $\sqrt[n]{a}$ für die Nullstellen von $X^n - a$. Betrachtet man z.B. Polynome vom Grad fünf, so haben diese im Allgemeinen keine durch Radikale ausdrückbare Nullstellen mehr. Man kann aber durch Elimination des konstanten Summanden alle Polynome vom Grad fünf in eine Normalform des Aufbaus $X^5 + A \cdot X^3 + B \cdot X^2 + C \cdot X + D$ bringen; führt man für die Nullstellen solcher Polynome ein eigenes Zeichen ein, so lassen sich auch die Nullstellen von Polynomen vom Grad fünf unter Umständen ebenfalls leichter repräsentieren.

Es wäre hier auch interessant zu untersuchen, ob wirklich die vier Parameter A, B, C und D benötigt werden, oder ob sich Nullstellen von Polynomen vom Grad fünf etwa schon durch Nullstellen einfacherer Polynome, z.B. $X^5 + aX + b$ repräsentieren lassen (wie sich ja Polynome vom Grad vier auch durch Nullstellen der Polynome $X^3 - a$ und $X^2 - a$ schon vollständig repräsentieren lassen). Es schließt sich die Frage an, ob es vielleicht eine für die Praxis relevante Untermenge der Polynome vom Grad fünf gibt, für die noch einfachere Darstellungen möglich sind, oder wie diese Vereinfachungen bei Polynomen höherer Grade in Frage kommt.

Der in Kapitel 7 beschriebene Ansatz der rückwirkenden Dynamik hat einen unschönen Aspekt bei der Bewegung semifreier, d.h. durch eine Nebenbedingung auf einen eindimensionalen Unterraum beschränkter Punkte. Der freie Parameter muss im Moment entlang einer künstlichen Achse (etwa der X-Achse entlang der Horizontalen) festgehalten werden, was besonders dann unelegant ist, wenn der eindimensionale Unterraum senkrecht zu dieser Achse verläuft. Zur Zeit arbeiten wir daran, diese künstlichen Achsen interaktiv der gerade durchgeführten Bewegung anzupassen. Dies ist nicht ganz einfach, da für jeden semifreien Punkt eine solche Achse festgelegt werden muss, die aber andererseits wieder in sinnvoller Beziehung zu den anderen Achsen sein sollte.

Der in 6.5 vorgestellte Ansatz für die Reduktion von Gleichungssystemen über Ortskurven ist nur ein erster Schritt in die Richtung, nicht nur die Geometrie durch Algebra zu erfassen, sondern andersherum algebraische Problem mit Hilfe der Geometrie besser lösen zu können (wie hier die Vereinfachung von Gleichungssystemen durch numerisches Vorerfassen von Ortskurven). Es ist eine interessante Überlegung, wie weit dieser Ansatz beim Lösen des Konstruktionsproblems helfen kann. Besonders bemerkenswert ist hier, dass das System über die Numerik zu Annahmen kommt, wie das Problem anzugehen sein könnte, also sozusagen eigene, nicht vorgegebene Lösungsideen entwickelt. Es wäre interessant, diesen Ansatz auszudehnen, um allgemein symbolische Berechnun-

gen und insbesondere automatisches Beweisen durch numerische Schaffung von Annahmen zu beschleunigen. Im Prinzip ermöglichen die Näherungslösungen das, was dem scharfen Betrachten der Aufgabenstellung durch den Menschen entspricht.

8.3 Danksagung

Ich danke besonders Professor Hotz, einmal für die Betreuung dieser Arbeit und seine Mühen bei der Korrektur, zum Zweiten und insbesondere aber für seine Zeit in unzähligen Gesprächen, seine Geduld und sein Vertrauen in mich.

Des Weiteren danke ich Professor Schreyer für die Unmengen an Zeit, die er in Gesprächen mit mir über diese Arbeit geopfert hat, und für die Durchsicht und Zweitkorrektur dieser Arbeit.

Professor Mehlhorn danke ich für seine Bereitschaft, meine Arbeit als Zweitbetreuer mit unterstützt zu haben, und insbesondere für seine Zeit in Gesprächen und Vorträgen.

Besonderer Dank gilt Tobias Gärtner, der nicht nur bei der Entstehung vieler Teile dieser Arbeit inhaltlich und emotional beteiligt war, sondern der mich auch durch seine gründliche Durchsicht und seine hilfreiche Kritik an dieser Arbeit sehr unterstützt hat.

Für weitere Durchsicht der Arbeit und eine Vielzahl von Motivationen danke ich Manuel Bodirsky, Jan Schwinghammer, Dr. Julia von Oertzen, Martin Struwe und Markus Zacharski.

Meinen Freunden und meiner Familie gilt mein Dank sowieso. Ohne Euch wäre nichts von dieser Arbeit auch nur im Entferntesten möglich gewesen.

Kapitel A

Anhang

Im Anhang an diese Arbeit wollen wir einige Beispiele für die Anwendung des Konstruktionsalgorithmus betrachten. Für die Berechnung wurde der Algorithmus einmal vollständig in Java (1.3.1) implementiert. Um einige verwendete Algorithmen (z.B. die Faktorisierung) in modernen Versionen nutzen zu können, wurde der algebraische Teil, d.h. der Algorithmen zum Finden radikalischer Nullstellen, noch einmal in Maple (V. 7.0) implementiert.

In den hier angegebenen Beispielen wurde die geometrische Konstruktion in Java in ein algebraisches Problem umgewandelt, dieses nach Maple portiert und dort mit der Maple-Implementierung gelöst. Die Rechenzeiten für beide Teile sind jeweils getrennt angegeben, gemessen jeweils auf einem Pentium mit 2.8 GHz.

Die Lösung aller Konstruktionsaufgaben läuft im Prinzip nach dem selben Muster ab: In der Aufgabe selber ist eine gewünschte Konstruktion angegeben, die aus gegebenen Größen (nennen wir diese *Zielgrößen*) des Objektes erstellt werden soll; ein einfaches Beispiel wäre ein Dreieck, dass aus vorgegebener Seite und zwei Winkeln konstruiert werden soll. Es wird nun zunächst die Konstruktion in allgemeiner Form erstellt (d.h. wir konstruieren irgendein Dreieck) und fügen darin die in der Aufgabe spezifizierten Zielgrößen ein (d.h. wir benennen die beiden Winkel und die Seite des Dreiecks). Das System stellt jetzt diese Zielgrößen symbolisch als Ausdrücke mit Quadratwurzeln über den Grundelementen dar (in unserem Beispiel erhalten wir also einen Ausdruck für die Dreiecksseite und die beiden Winkel in Abhängigkeit von den Koordinaten der drei Dreieckspunkte).

Diese Darstellungen der Zielgrößen ist jetzt ein Gleichungssystem; die Konstruktionsaufgabe ist gelöst, wenn dieses Gleichungssytem durch Ausdrücke mit Quadratwurzeln nach den Grundgrößen (also z.B. den Punkten des Dreiecks) auflösen können. Diese Ausdrücke sind dann eine Repräsentation der auszuführenden Konstruktion, die eine kompakte Notation relativ komplexer Handlunsanweisungen erlaubt. Lemma 2.2.3 beschreibt, wie diese Notation in eine Folge von Operationen mit Zirkel und Lineal umgewandelt werden kann.

A.1 Konstruktion eines Dreiecks aus zwei Winkeln und einer Seite

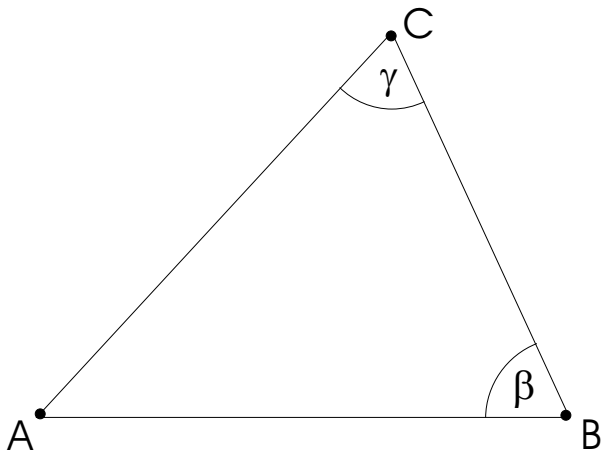

Abbildung A.1: Konstruktion eines Dreiecks aus zwei Winkeln und einer Seite Der Punkt C wird aus A, B und den beiden Winkeln β und γ rekonstruiert.

Wir betrachten zunächst das eben eingeführte simple Beispiel, um die Art der Lösungen eines Konstruktionsproblems zu erläutern. Sei ABC ein Dreieck. Von dem Dreieck ist die Seite AB, der Tangens des Winkels β bei B und der Tangens des Winkels γ bei C gegeben; aus diesen Größen soll das Dreieck konstruiert werden.

Wir zeichnen also zunächst das Dreieck; die Lage der gegebenen Seite ist willkürlich, wir können also zwei Punkte des Dreiecks auf feste Koordinaten setzen. Entsprechend legen wir A auf den Koordinatenursprung $(0, 0)$ und B auf $(1, 0)$. Die Lage des Punktes C lassen wir offen und setzen seine Koordinaten symbolisch mit zwei Variablen $C = (x, y)$ an.

Jetzt spezifizieren wir den Winkel β und γ; das System repräsentiert diese Winkel über ihren Tangenswert und berechnet die Werte $\tan(\beta)$ und $\tan(\gamma)$ als symbolische Ausdrücke über den Koordinaten x und y (Laufzeit 120 ms):

$$\begin{aligned} \tan(\beta) &= -\frac{y}{1-x} \\ \tan(\gamma) &= -\frac{y}{y^2+x^2-y} \end{aligned}$$

Wir schreiben im Folgenden kurz b für $\tan(\beta)$ und g für $\tan(\gamma)$.

Die verbleibende Aufgabe ist, dieses Gleichungssystem nach x und y unter Verwendung von Quadratwurzeln aufzulösen. Die Aufgabe ist in diesem Fall trivial (der Vollständigkeit halber sind alle Konstruktion, auch die trivial lösbaren, nach Maple portiert, dort auf ein univariates Problem zurückgeführt und mit

dem allgemeinen Algorithmus zum Finden von Quadratwurzeln aus Kapitel 5 gelöst worden). Die Lösungen dieser Gleichungen ist (Laufzeit 40 ms):

$$\begin{aligned} x &= -\frac{-\tan(\gamma) - 2\tan(\gamma)\tan(\beta)^2 + \tan(\beta) \pm (\tan(\beta) + \tan(\gamma))}{2\tan(\gamma)(\tan(\beta)^2 + 1)} \\ y &= -\frac{\tan(\beta)^2 + \tan(\gamma)\tan(\beta) \pm \tan(\beta)(\tan(\beta) + \tan(\gamma))}{2\tan(\gamma)(\tan(\beta)^2 + 1)} \end{aligned}$$

Das $\pm$ entsteht durch eine Wurzel aus eins (in beiden Ausdrücken die selbe Wurzel). Nimmt man hier ein negatives Vorzeichen, erhält man $x = 1$ und $y = 0$, d.h. mit $B = C$ ein degeneriertes Dreieck. Der andere Fall vereinfacht sich zu

$$\begin{aligned} x &= -\frac{-\tan(\gamma)\tan(\beta)^2 + \tan(\beta)}{\tan(\gamma)(\tan(\beta)^2 + 1)} \\ y &= -\frac{\tan(\beta)^2 + \tan(\gamma)\tan(\beta)}{\tan(\gamma)(\tan(\beta)^2 + 1)} \end{aligned}$$

Die Ausdrücke beschreiben nun vermöge der in Lemma 2.2.3 beschriebenen Isomorphie eine Konstruktion von x und y aus den Tangenswerten von β und γ; wir stellen fest, dass diese Konstruktion (von dem degenerierten Fall abgesehen) eindeutig bestimmt ist.

A.2 Konstruktion eines Dreiecks aus einer Seite, einem Winkel und einer Seitenhalbierenden

Wir wollen jetzt auf die selbe Art und Weise ein Dreieck aus einer Seite, der Länge der Winkelhalbierenden und dem der Seite gegenüberliegendem Winkel konstruieren. Diese Konstruktion birgt ohne den Einsatz von Computeralgebrasystemen schon ein wenig mehr Probleme.

Wir zeichnen zunächst wieder ein Dreieck ABC, diesmal mit der Seitenhalbierenden von A zur Mitte von BC und dem Winkel am Punkt C. Wir bezeichnen mit s die Länge der Seitenhalbierenden und mit γ den Winkel.

Sei $A = (0, 0)$, $B = (1, 0)$ und $C = (x, y)$. Die Länge der Seitenhalbierenden und der Tangens von γ berechnen sich dann in einer Laufzeit von 230 ms zu:

$$\begin{aligned} s &= \frac{\sqrt{x^2 + y^2 + 1 + 2x}}{2} \\ \tan(\gamma) &= \frac{y}{x^2 + y^2 - x} \end{aligned}$$

Für x und y gibt es dann folgende erfüllende Lösung mit Quadratwurzeln (Laufzeit 50 ms):

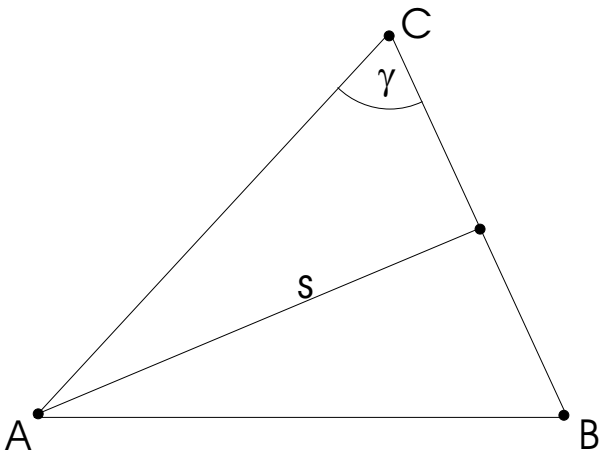

Abbildung A.2: Konstruktion eines Dreiecks aus einer Seite, einem Winkel und einer Seitenhalbierenden

Der Punkt C wird aus A, B, γ und der Länge der Seitenhalbierenden s rekonstruiert.

$$
\begin{aligned}
x &= \frac{24s^2\tan(\gamma)^2 - 6\tan(\gamma)^2 - 2 + 4\sqrt{5s^2\tan(\gamma)^2 - \tan(\gamma)^2 - 4s^4\tan(\gamma)^2 + s^2}}{2(9\tan(\gamma)^2 + 1)} \\
y &= \sqrt{4s^2 - x^2 - 1 - 2x}
\end{aligned}
$$

Dies ergibt zwei Lösungen für x und für beide jeweils zwei Lösungen für y, einmal ober- und einmal unterhalb der Seite AB. Bei der zweiten Lösung nimmt der Winkel γ den Komplementärwert an, so dass außerhalb des Dreiecks die beiden Seiten AC und BC den gleichen Winkel einschließen wie bei der anderen Lösung innerhalb des Dreiecks.

A.3 Konstruktion eines Dreiecks aus drei Höhen

Bei der nächsten Konstruktion wollen wir ein Dreieck aus den drei Höhen rekonstruieren; da wir diesmal keine Seite vorgegeben haben, fixieren wir einen Eckpunkt als Nullpunkt des Koordinatensystems und setzen die Richtung einer Dreiecksseite als x-Achse fest.

Wir konstruieren also zunächst unser Dreieck ABC mit $A = (0, 0)$ und $B = (z, 0)$. Den Punkt C setzen wir wieder völlig frei auf $C = (x, y)$. Die drei Höhen h_a, h_b und h_c berechnen sich dann zu (Lauzeit 211 ms):

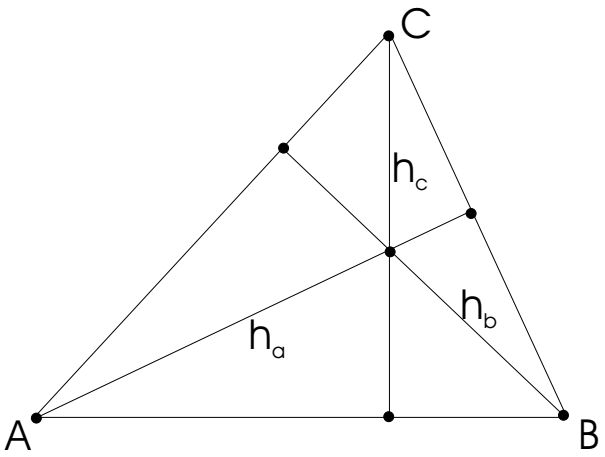

Abbildung A.3: Konstruktion eines Dreiecks aus drei Höhen
Das Dreieck A, B, C wird aus den Längen der drei Höhen konstruiert.

$$\begin{aligned} h_a &= \frac{zy\sqrt{x^2+y^2-2zx+z^2}}{x^2+y^2-2zx+z^2} \\ h_b &= \frac{zy\sqrt{x^2+y^2}}{x^2+y^2} \\ h_c &= \pm y \end{aligned}$$

Haben wir jetzt andersherum h_a, h_b und h_c gegeben, so ist offenbar $y = \pm h_c$. Die anderen beiden Größen haben beide den gleichen Nenner

$$\begin{aligned} n = \; (\; & (h_b h_c - h_a h_b - h_a h_c)(h_b h_c + h_a h_b - h_a h_c) \\ & (h_b h_c + h_a h_b + h_a h_c)(-h_b h_c + h_a h_b - h_a h_c) \quad)^{\frac{1}{2}} \end{aligned}$$

insgesamt ist ihr Wert (Laufzeit 78 ms):

$$\begin{aligned} z &= \frac{2h_a^2 h_b^2 h_c}{n} \\ x &= \frac{h_c\left(-h_c^2 h_b^2 + h_b^2 h_a^2 + h_c^2 h_a^2\right)}{n} \end{aligned}$$

Es ergeben sich insgesamt vier Lösungen: 2 Durch die beiden Zweige der Quadratwurzel, die eine Spiegelung um die Y-Achse (also die Gerade parallel zu AB durch A) ausmachen, und die Spiegelung des dritten Punktes um die X-Achse; insgesamt erhalten wir also ein eindeutiges Dreieck in vier verschiedenen Spiegelungen.

A.4 Konstruktion eines Dreiecks aus einer Seite und der Eulergraden

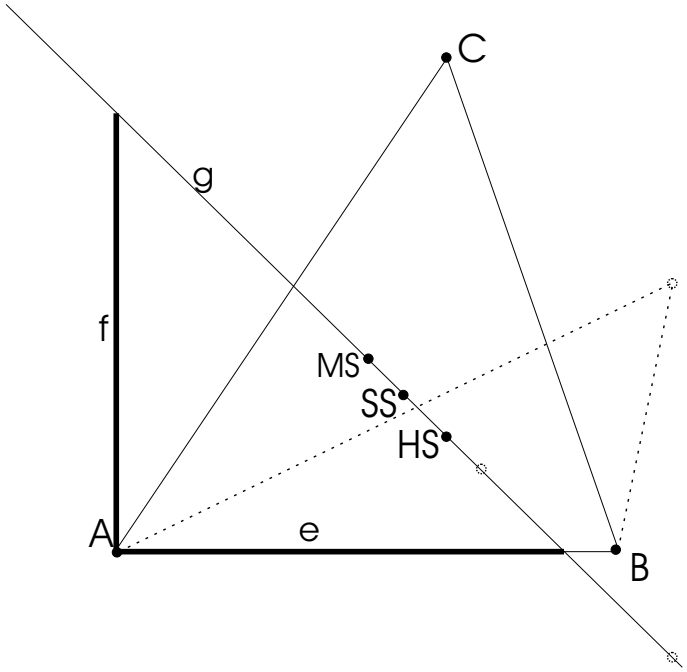

Abbildung A.4: Konstruktion eines Dreiecks aus einer Seite und der Eulergeraden

Ist eine Seite gegeben, so genügt die Eulergerade für die Konstruktion des Dreiecks mit Zirkel und Lineal.

Wir wollen ein Dreieck konstruieren, von dem wir eine Seite und die Eulergeraden gegeben haben, wobei die Eulergerade definiert ist als die Gerade durch den Schnittpunkt der Höhen (HS), der Seitenhalbierenden (SS) und der Mittelsenkrechten (MS) (für einen Beweis, dass diese drei Punkte auf einer Geraden liegen, siehe z.B. [18]). Wir zeichnen für diese Konstruktion wieder ein allgemeines Dreieck ABC und zeichnen die Eulergerade ein; Ihre Lage wollen wir durch den Schnitt mit der X-Achse und der Y-Achse festlegen, d.h. in unserem Fall durch den Schnitt mit einer Geraden durch AB und einer Senkrechten dazu durch A.

Wir legen für A wieder die Koordinaten $(0, 0)$ und für $B = (1, 0)$ fest; C legen wir wie zuvor mit Parametern auf den Punkt (x, y). Sei e die Entfernung des Schnittes von g mit der X-Achse zum Ursprung und f die Entfernung des Schnittes von g mit der Y-Achse. In Abhängigkeit von x und y ist dies (Rechenzeit 250 ms):

$$
\begin{aligned}
e &= \frac{-xy^2 - x^3 + x}{-y^2 - 3x^2 + 3x} \\
f &= \frac{-xy^2 - x^3 + x}{-2xy + y}
\end{aligned}
$$

Wir lösen jetzt nach x und y:

$$
\begin{aligned}
x &= \frac{e^2 - f^2 + 4f^2e + \sqrt{e^4 - 2f^2e^2 + 12f^2e^3 + f^4 - 4f^4e + 4f^4e^2 - 12f^2e^4}}{2(f^2 + e^2)} \\
y &= (\quad f \cdot \left(-f^2 - 3e^2 + 2f^2e + 6e^3\right) \\
&\quad -f\sqrt{e^4 - 2f^2e^2 + 12f^2e^3 + f^4 - 4f^4e + 4f^4e^2 - 12f^2e^4} \quad) \\
&\quad \cdot \frac{1}{2e(f^2 + e^2)}
\end{aligned}
$$

Die zweite mögliche Lösung ist in der Zeichnung mit den gestrichelten Linien und Punkt angedeutet (der Schnitt der Mittelsenkrechten liegt am selben Punkt wie zuvor).

A.5 Ein Drei-Hebel-System

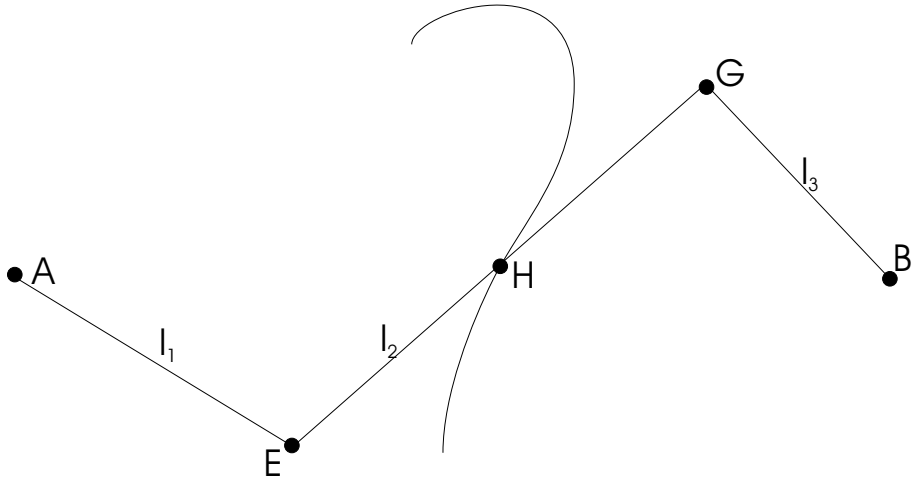

Abbildung A.5: Robotik-Konstruktion aus drei Hebelarmen
Die Kurve in der Mitte ist die Ortskurve der von H erreichbaren Punkte. Aus der Position von H lässt sich die Einstellungen der anderen Hebel nicht mit Zirkel und Lineal konstruieren.

Als nächstes wollen wir ein Beispiel aus der Robotik betrachten: Gegeben seien drei mit Gelenken aneinander befestigten Hebel mit festen Längen l_1, l_2 und l_3. Die Endpunkte der Kette seien jeweils fest eingespannt. Drehen wir

den ersten Hebel um seinen festen Punkt, so bewegen sie die anderen beiden entsprechend mit; wir interessieren uns für die Kurve des Mittelpunktes des mittleren Hebels. Offenbar können wir das Hebelsystem mit Zirkel und Lineal konstruieren; geht es aber auch, aus einer Koordinate der Lage des Mittelpunktes durch eine Konstruktion mit Zirkel und Lineal die Stelle zu bestimmen, an der der erste Hebel sein muss?

Setzen wir einen Fixpunkt des Hebelsystems bei $A = (0, 0)$ und den zweiten bei $B = (1, 0)$. Die Längen der ersten Hebel belegen wir mit den Variablen l_1, l_2 und l_3. Der Punkt E soll der Endpunkt des ersten Hebels sein, d.h. er muss semifrei mit Parameter λ auf einem Kreis um A mit Radius l_1 liegen; wir verwenden zur Darstellung des Punktes die Kreisparametrisierung $E = (l_1 \frac{-2\lambda}{1+\lambda^2}, l_1 \frac{1-\lambda^2}{1+\lambda^2}$. Um E schlagen wir einen Kreis mit Radius l_2 und schneiden diesen mit einem Kreis um B mit Radius l_3; den Schnittpunkt wollen wir G nennen. Der Mittelpunkt H von E und G ist dann der von uns gesuchte Punkt. Wir betrachten zunächst die x-Koordinate dieses Punktes (die Berechnung dauerte 1331 ms); wir haben hier einen Anteil innerhalb der ersten Wurzel:

$$\begin{aligned} x_1 &= -\left(l_3^2 + l_3^2\lambda^2 - 4l_3\lambda^2 + 4\lambda^2 - l_1^2\lambda^2 - l_1^2 + 2l_1l_2 + 2l_1l_2\lambda^2 - l_2^2\lambda^2 - l_2^2\right) \\ &\quad \cdot \left(l_3^2 + l_3^2\lambda^2 + 4\lambda^2 - 4l_3\lambda^2 - l_1^2\lambda^2 - l_1^2 - 2l_1l_2 - 2l_1l_2\lambda^2 - l_2^2\lambda^2 - l_2^2\right) \end{aligned}$$

und einen Anteil vor der ersten Wurzel:

$$x_2 = (-2\lambda^2 - l_3 + l_3\lambda^2)(3l_3^2\lambda^2 - 12l_3\lambda^2 - l_2^2\lambda^2 + l_1^2\lambda^2 + 12\lambda^2 + 3l_3^2 + l_1^2 - l_2^2)$$

Mit einem multiplikativen Faktor vor der Wurzel und dem Nenner ergeben diese beiden den folgenden Ausdruck für die x-Koordinate von H:

$$x = \frac{(-2\lambda + 2l_3\lambda)\sqrt{x_1} + x_2}{4(4\lambda^2 + 4\lambda^4 - 4l_3\lambda^2 - 4l_3\lambda^4 + l_3^2 + 2l_3^2\lambda^2 + l_3^2\lambda^4)}$$

Löst man in dieser Gleichung die Quadratwurzel auf und faktorisiert, so erhält man folgendes Lösungspolynom für λ:

$$
\begin{aligned}
0 \;=\;& (4\lambda^2 - 4l_3\lambda^2 + l_3^2\lambda^2 + l_3^2)\cdot \\
& (96xl_3\lambda^2 + 192x\lambda^4 + 128x^2\lambda^4 + 8x\lambda^2 l_1^2 l_3 + 8x\lambda^4 l_3 l_2^2 + 8x\lambda^4 l_3 l_2^2 - 8x\lambda^4 l_1^2 l_3 \\
& -8x\lambda^6 l_1^2 l_3 - 192xl_3\lambda^4 + 64x^2\lambda^6 - 24x\lambda^4 l_3^3 - 24x\lambda^6 l_3^3 + 32x\lambda^4 l_1^2 + 16x\lambda^6 l_1^2 \\
& +24x\lambda^2 l_3^3 - 8x\lambda^2 l_3 l_2^2 - 128x^2\lambda^4 l_3 + 48x^2\lambda^2 l_3^2 - 64x^2\lambda^6 l_3 + 48x^2\lambda^4 l_3^2 \\
& +16x^2\lambda^6 l_3^2 - 32x\lambda^4 l_2^2 - 16x\lambda^6 l_2^2 - 48xl_3^2\lambda^2 + 64\lambda^2 x^2 + 96xl_3^2\lambda^4 + 24l_3^3 x \\
& -8l_2^2 l_3 x + 16l_3^2 x^2 + 8l_3 l_1^2 x - 16x\lambda^2 l_2^2 - 64x^2\lambda^2 l_3 + 16x\lambda^2 l_1^2 + 192x\lambda^6 \\
& -288xl_3\lambda^6 + 144xl_3^2\lambda^6 + 3l_1^4\lambda^2 + l_1^4 + 9l_3^4 + l_2^4 + 16\lambda^4 + 144\lambda^6 - 8l_3 l_2^2\lambda^2 \\
& +24l_3 l_2^2\lambda^6 - 2l_3^2 l_2^2\lambda^2 - 2l_3^2 l_2^2\lambda^4 - 6l_3^2 l_2^2\lambda^6 + 16l_3 l_2^2\lambda^4 - 8l_2^2\lambda^2 - 32l_2^2\lambda^4 \\
& -6l_1^2 l_2^2\lambda^4 - 2l_1^2 l_2^2\lambda^6 + 40l_1^2 l_3\lambda^2 + 16l_1^2 l_3\lambda^4 - 24l_1^2 l_3\lambda^6 - 14l_1^2 l_3^2\lambda^2 - 14l_1^2 l_3^2\lambda^4 \\
& +6l_1^2 l_3^2\lambda^6 - 6l_1^2 l_2^2\lambda^2 - 24l_2^2\lambda^6 + 3l_2^4\lambda^2 + 3l_2^4\lambda^4 + l_2^4\lambda^6 + 96l_3\lambda^4 - 288l_3\lambda^6 \\
& +40l_3^2\lambda^2 - 128l_3^2\lambda^4 + 216l_3^2\lambda^6 - 6l_3^2 l_2^2 - 8l_3^3\lambda^2 + 48l_3^3\lambda^4 - 72l_3^3\lambda^6 - 5l_3^4\lambda^2 \\
& -5l_3^4\lambda^4 + 9l_3^4\lambda^6 - 8l_1^2\lambda^2 + 16l_1^2\lambda^4 + 24l_1^2\lambda^6 - 2l_1^2 l_2^2 + 6l_1^2 l_3^2 + 3l_1^4\lambda^4 + l_1^4\lambda^6)
\end{aligned}
$$

Der erste Faktor liefert für reellwertige l_3 nur komplexwertige λ und enthält also keine reelle Lösung; der zweite Faktor ist irreduzibel und vom Grad 6 bezüglich λ, diese Polynom kann also nicht durch Quadratwurzeln gelöst werden; wir haben damit also bewiesen, dass es zu der gegebenen Konstruktion keine Umkehrung mit Zirkel und Lineal gibt.

A.6 Konstruktion des regelmäßigen Fünfecks

Wir wollen jetzt ein regelmäßiges Fünfeck konstruieren. Wir nutzen dafür einige triviale Symmetrien: Zunächst betrachten wir eine Seite $\overline{AB}$ des Fünfecks, wobei $A = (0, 0)$ und $B = (1, 2)$ sein soll. Sei C ein semifreier Punkt mit Parameter λ auf der Mittelsenkrechten von A und B. Wir nehmen eine Parallele von AB durch C und schneiden diese mit einem Kreis um B mit Radius $\overline{AB}$ im Punkt D; um D wiederum schlagen wir einen Kreis, ebenfalls mit Radius AB, und schneiden diesen mit der Mittelsenkrechten von AB im Punkt E. Offenbar sind A, B, D und E dann vier Punkte eines gleichseitigen Fünfecks (den fünften Punkt erhält man z.B. durch Schnitt der Parallelen durch C mit einem Kreis um A mit Radius AB). Das Fünfeck wird regelmäßig, wenn jetzt noch zwei Diagonalen gleich lang sind; wir nehmen also die Differenz der Diagonalen $\overline{AD}$ und $\overline{AE}$. Die Berechnung dieser Größen nimmt 246 ms in Anspruch und ergibt folgenden Ausdruck:

$$\overline{AG} - \overline{AI} = \sqrt{5\lambda^2 - 5\sqrt{4-\lambda^2} + 5\lambda\sqrt{-\lambda^2 - 2\sqrt{4-\lambda^2}}\sqrt{2} - 2\sqrt{10 + 5\sqrt{4-\lambda^2}}}$$

Lösen wir die Quadratwurzeln auf, so erhalten wir

$$0 = 16\lambda^4 - 79\lambda + 69$$

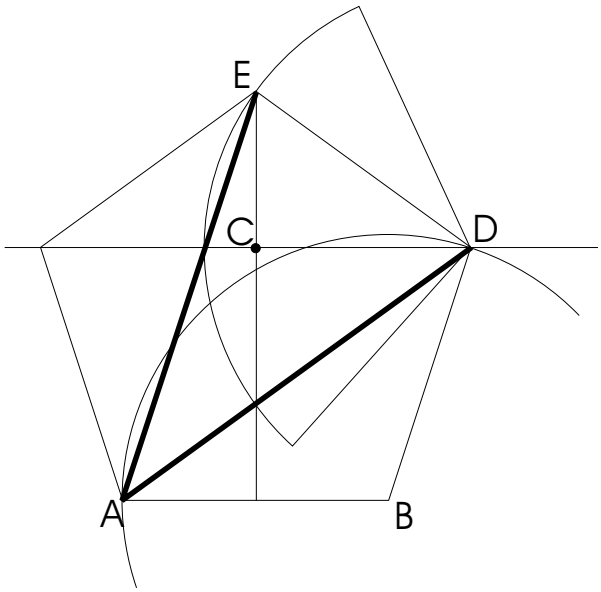

Abbildung A.6: Konstruktion eines regelmäßigen Fünfecks
Das Fünfeck wird über die Gleicheit der Diagonalen in einem gleichseitigen Fünfeck konstruiert; alternative Lösungen sind die Spiegelung an AB und der Drudenfuß.

Mit den Lösungen

$$\lambda = \pm\frac{\sqrt{158 \pm 10\sqrt{73}}}{8}$$

Die Werte mit den + in der Wurzel ergeben das übliche regelmäßige Fünfeck (ober- oder unterhalb von AB). Die anderen beiden Werte ergeben den 'Drudenfuß', d.h. ebenfalls ein regelmäßiges Fünfeck, aber mit sich paarweise schneidenden Seiten.

A.7 Kombinierte Konstruktion

Wir wollen ein weiteres Beispiel mit einer Folge von Konstruktionen betrachten. Sei $ABCD$ ein Quadrat und $E = (x, y)$ ein weiterer Punkt. Sei F der Höhenschnittpunkt in ABE und G und H die Umkreismittelpunkte von DAE und BCE. Offenbar liegen diese beiden Punkte auf der Halbierenden des Quadrates (parallel zu AB), da diese Gerade den Mittelsenkrechten der Dreiecke DAE (auf DA) und BCE (auf BC) entsprechen. Seien a und b die beiden Abstände von G zu DA und von H zu BC.

Wir wollen den Punkt E aus dem Quadrat und den beiden Abständen a und b berechnen. Für a und b bekommen wir folgende Ausdrücke (715 ms):

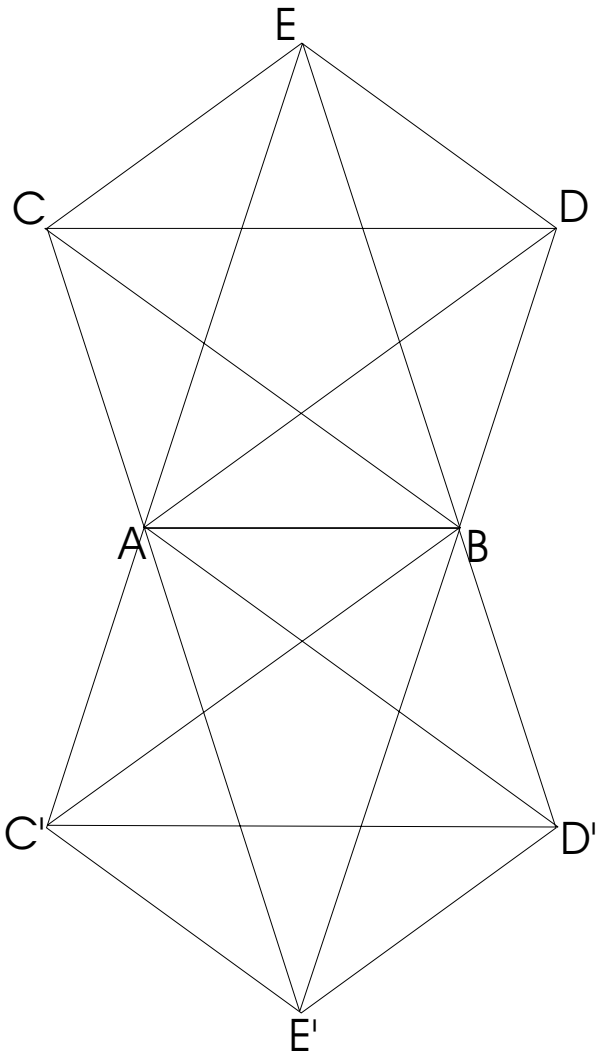

Abbildung A.7: Mögliche Lösungen für das regelmäßige Fünfeck

$$
\begin{aligned}
a &= \frac{\sqrt{5}(2y^3+xy^2-7y^2+3xy+2x^2y-3x^2+x^3)}{2(y^2-4xy+4x^2)} \\
b &= \frac{\sqrt{5}(2y^3+xy^2-10y^2+15y-5xy+2x^2y-5x+x^3)}{2(y^2-4xy+4x^2)}
\end{aligned}
$$

Berechnen wir die Resultante dieser beiden Gleichungen bezüglich x und faktorisieren, so erhalten wir neben zwei Faktoren für $x = 0$ und $x = 1$:

$$
\begin{aligned}
0 \;=\; & 130a + 80ax^2 - 70b + 24b^2x^2a - 8b^3x^2 - 8a^2b - 16a^3x - 80bx^2 \\
& +8a^3 + 140bx - 180ax - 16b^2xa + 8a^3x^2 - 24bx^2a^2 + 32a^2xb
\end{aligned}
$$

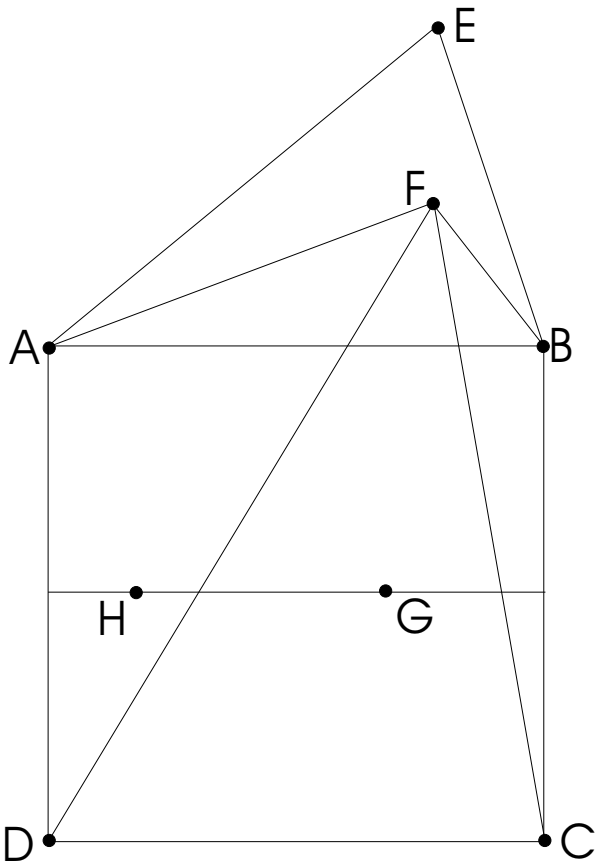

Abbildung A.8: Kombinierte Konstruktion
F ist der Höhenschnitt von ABE, G und H jeweils die Schwerpunkte von ADF und BCF. Aus den Abständen von G und H zu AD und BC lässt sich der Punkt E mit Zirkel und Lineal konstruieren.

Diese lösen sich (ohne Verwendung von Vereinfachungen innerhalb von 345s) zu:

$$
\begin{aligned}
x \quad = (\quad & 4\sqrt{5}b^2 + 15\sqrt{5} + 10\sqrt{5}a^2 - 35b + 45a + 4b^2a - 14\sqrt{5}ba + 4a^3 - 8a^2b \\
& +(\ -200\sqrt{5}a - 32\sqrt{5}b^4a + 32\sqrt{5}a^4b + 96\sqrt{5}b^3a^2 + 200\sqrt{5}b \\
& \quad -800\sqrt{5}b^2a - 160\sqrt{5}a^3 + 1800ba - 700b^2 + 640a^3b - 96\sqrt{5}a^3b^2 \\
& \quad +800\sqrt{5}a^2b - 700a^2 - 1160b^2a^2 - 60b^4 + 160\sqrt{5}b^3 + 640b^3a - 60a^4)^{\frac{1}{2}} \quad) \\
& \cdot \frac{1}{40a + 5\sqrt{5} - 4b^3 + 5\sqrt{5}b^2 + 12b^2a - 20\sqrt{5}ba + 4a^3 - 12a^2b - 40b + 10\sqrt{5}a^2}
\end{aligned}
$$

Die Konstruktion von y läuft parallel; auf diese Weise haben wir eine Konstruktion von E aus den Größen a und b erstellt.

A.8 Konstruktion des Icosaeders

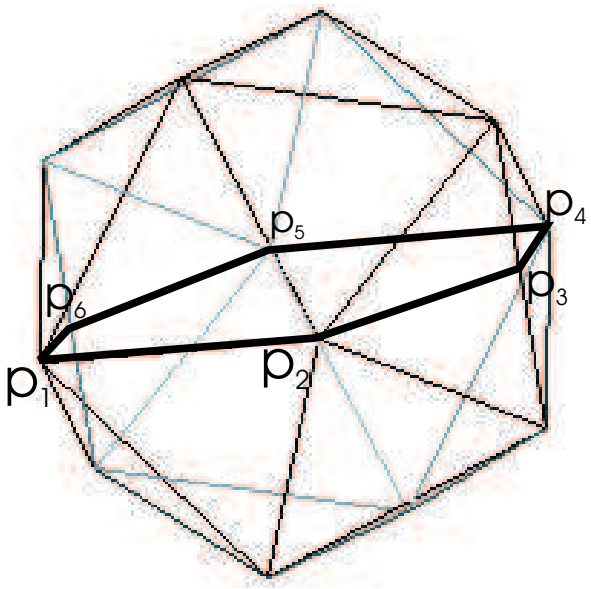

Abbildung A.9: Icosaeder mit Schnittfläche
Gelingt es, das fett gezeichnete (unregelmäßige) Sechseck zu konstruieren, so ist die Konstruktion des gesamten Icosaeders einfach, indem man ein identisches Sechseck senkrecht zu dem ersten aufrichtet.

Ein Icosaeder ist ein 20-Seitiger regelmäßiger Körper, bestehend aus 20 Dreiecken. Der Icosaeder findet sich in Euklids Orginalwerkt 'Die Elemente, Band X' [7]; wir wollen hier eine mit Hilfe des Nullstellen-Algorithmus erstellte Konstruktion vorstellen. Um den Icosaeder (seiner Natur nach dreidimensional) zu konstruieren, betrachten wir die Äquatorebene der dem Icosaeder umbeschriebenen Kugel durch zwei gegenüberliegende Kanten. Sei a die Kantenlänge des Icosaeders, dann besteht der Schnitt aus einem Sechseck mit zwei gegenüberliegenden Kanten der Länge a, die jeweils an beiden Enden den Äquatorkreis berühern (wir nennen diese Punkte p_1, p_2, p_4 und p_5, und vier Kanten der Länge h (wobei h die Höhe eines der Icosaeder-Dreiecke ist), deren Treffpunkte p_3 und p_6 innerhalb des Kreises liegen. Betrachten wir senkrecht zu dieser Ebene durch p_3 und p_6 eine weitere Schnittebene des Icosaeders, so enthält diese das selbe Sechseck, wobei p_3 und p_6 die Mittelpunkte der langen Kanten (die senkrecht auf der ursprünglichen Ebene stehen) darstellen. Wir stellen also fest, dass der Abstand $\overline{p_1 p_5}$ gleich dem Abstand $\overline{p_3 p_6}$ seien muss. Geben wir den Kugelradius vor, bestimmt das aber das Sechseck bereits; wir haben also die 3-dimensionale Konstruktionsaufgabe in eine zweidimensionale umgewandelt.

Setzen wir $A = (0, 0)$ als den Umkreismittelpunkt und $p_1 = (1, 2)$. p_2 setzen wir semifrei (mit Parameter λ) auf den Umkreis. Dadurch haben wir auch p_4 (p_1 gegenüberliegend) und p_5 (p_2 gegenüberliegend) konstruiert. Wir konstruieren jetzt die Höhe des gleichseitigen Dreiecks mit Seitenlänge $\overline{p_1 p_2}$ und schlagen mit

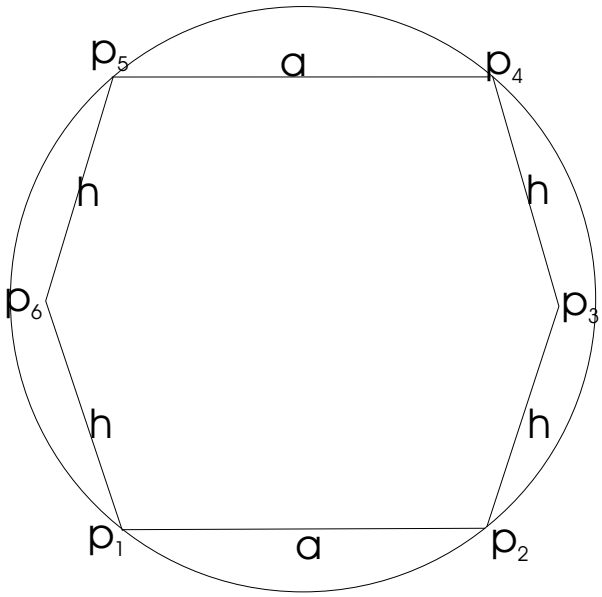

Abbildung A.10: Die Schnittfläche aus dem Icosaeder in der Ebene

diesem Radius vier Kreise um p_2, p_4, p_1 und p_5; ein Schnittpunkt der ersten beiden und ein Schnittpunkt der letzteren beiden liefern uns jeweils p_3 und p_6. Wir messen die Distanzen $\overline{p_1p_5}$ und $\overline{p_3p_6}$ und erhalten in 15930 ms:

$$\overline{p_1p_5} - \overline{p_3p_6} = \frac{\sqrt{-2\sqrt{5} + 2\sqrt{4\lambda^2 - 1 - 2\sqrt{3\lambda^2 - 1}}}}{\sqrt{1+\lambda^2}}$$

setzen wir diese Differenz auf 0 und eliminieren die Wurzeln, so erhalten wir einen Faktor $\lambda^4 - 7\lambda^2 - 9$, den wir lösen müssen; die Lösung ist dann

$$\lambda = \frac{\sqrt{14 - 2\sqrt{85}}}{2}$$

A.9 Konstruktion des regelmäßigen Siebzehnecks

Die Konstruktion des regelmäßigen Siebzehnecks verwenden wir die Tatsache, dass die Lage der siebzehnten Einheitswurzeln in $\mathbb{C}$ ein Siebzehneck bilden. Wir brauchen also nur den Realteil einer Nullstellen des Polynoms $X^{17}-1$ (zusätzlich zur 1) zu erzeugen; diese Polynom zerfällt in $(X-1)(X^{16}+\cdots+1)$, wodurch wir den Punkt 1 als Startpunkt für das Siebzehneck schonmal zur Verfügung haben. Ist x eine Lösung dieses Polynoms, so ist $z = \frac{x+\frac{1}{x}}{2}$ der Realteil dieser Einheitswurzel (denn da $|z| = 1$, ist $\overline{z} = \frac{1}{z}$). Wir dividieren also den Restfaktor durch X^8 und substituieren $2Z = X + \frac{1}{X}$. Wir erhalten dann das Polynom

$$Z^8 + Z^7 - 7Z^6 - 6Z^5 + 15Z^4 + 10Z^3 - 10Z^2 - 4Z + 1$$

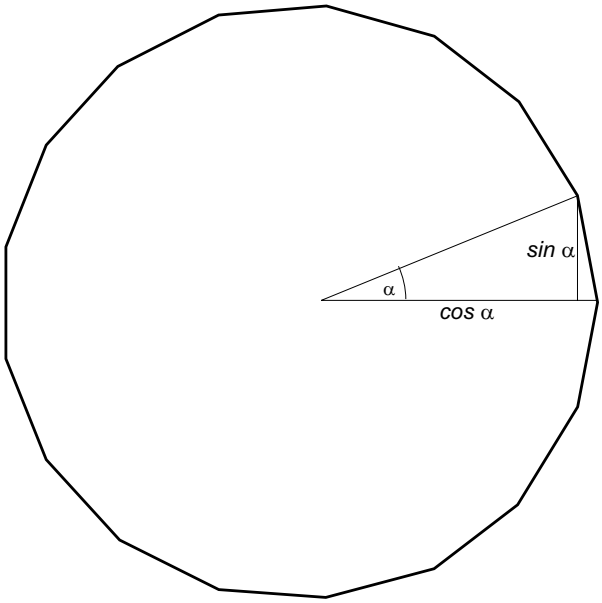

Abbildung A.11: Das regelmäßige Siebzehneck
Die Konstruktion des regelmäßigen Siebzehnecks gelingt über die Konstruktion der 17ten Einheitswurzel.

Zu diesem Polynom finden wir innerhalb von 185s die durch Quadratwurzeln ausdrückbaren Nullstellen:

$$\frac{1}{8}\left(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}+2\sqrt{17+3\sqrt{17}-\sqrt{170+38\sqrt{17}}}\right)$$

Dieser Ausdruck beschreibt jetzt eine Konstruktion für das regelmäßige Siebzehneck: Beginnend mit dem Einheitskreis konstruiert man die Hälfte des obigen Ausdrucks und trägt durch die X-Achse mit diesem Abstand zur Null eine Senkrechte durch die X-Achse; Die Schnittpunkte dieser Senkrechten mit dem Einheitskreis gibt die Position des zweiten (und 17.) Punktes des Siebzehnecks an, wodurch die restlichen Punkte leicht zu konstruieren sind.

Literaturverzeichnis

[1] E.L. Allgower, K. Georg: Numerical Path Following. In P.G. Ciarlet, J.L. Lions (Edt): Techniques of Scientific Computing (part 2), s. 3-203. North-Holland, 1997.

[2] J. Canny: Generalised characteristic polynomials, Journal of Symbolic Computation 9, s. 241-250, 1990.

[3] S. C. Chou: GEO-Prover - A Geometry Theorem Prover Developed at UT.

[4] S. C. Chou, X.S. Gao, J.Z. Zhang: Machine Proofs in Geometry. World Scientific, Singapore, 1994.

[5] D. Cox, J. Little, D. O'Shea: Ideals, Varieties, and Algorithms. Springer-Verlag, 1997.

[6] D. Cox, J. Little, D. O'Shea: Using Algebraic Geometry. Springer-Verlag, 1998.

[7] Euklid: Die Elemente Band XIII, §16. Herausgegeben von C. Thaler, Akademische Verlagsgesellschaft m.b.H. Leipzig, 1933

[8] E. W. Elcock: Representation of Knowledge in a Geometry Machine. Machine Intelligence 8,1977.

[9] G. Fischer: Analytische Geometrie. Vieweg-Verlag, 1979.

[10] J. v. z. Gathen, J. Gerhard: Modern Computer Algebra. Cambridge University Press, 1999.

[11] K.O. Geddes, S.R. Czapor, G. Labahn: Algorithms for Computer Algebra. Kluwer, 1992.

[12] D. Hilbert: Grundlagen der Geometrie. Teubner-Verlag.

[13] G. Hollander: Geolog. Dümmler Verlag, 1993.

[14] J. C. Lagarias, A.M. Odlyzko: Effective version of the Chebotarev density theorem. In A. Fröhlich (Edt): Algebraic Number Fields, s. 409 - 464. Academic Press, 1977.

[15] S. Landau, G. Miller: Solvability by Radicals is in Polynomial Time. Journal of Computer and Systems Sciences, vol. 30, No. 2 (1985), s. 179-208.

[16] F. Macaulay: On some formulas in elimination, Proceedings of London Mathematic Society 3, s. 3-27, 1902.

[17] J. Neukirch, K. Schmidt, A. Wingberg: Cohomology of Number Field, s. 476-507. Springer-Verlag, 2000.

[18] T. v. Oertzen: Cedric - ein automatisches geometrisches Beweissystem. Diplomarbeit, 1999.

[19] H.J. Reifen, G. Scheja, U. Vetter: Algebra. Verlag Bibliographisches Institut AG, 1984.

[20] J. Richter-Gebert, U. H. Kortenkamp: The Interactive Geometry Software Cinderella. Springer-Verlag, 1999.

[21] E. Rohnert: Eine Benutzerschnittstelle zum automatischen Beweisen von Sätzen aus der Geometrie, Universität des Saarlandes, 1997.

[22] H. R. Schwarz: Numerische Mathematik. Teubner-Verlag, 1997.

[23] I.R. Shafarevitsch: Construction of fields of algebraic numbers with given solvable Galois group. Izv. Akad. Nauk. SSSR. Ser. Mat. 18, 1954, s. 525-578 (in Russisch), oder Amerk. Math. Soc. Transl. 4, 1956, s. 185-237 (in Englisch)

[24] L. Smith: Linear Algebra 3rd Edition. Springer-Verlag, 1998.

[25] N. Tschebotareff: Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören. Math. Annalen 95, s. 191-228, 1925.

[26] W. T. Wu: Geometric Theorems Proving in Geometry.

[27] E. Zeidler, W. Hackbusch, R. Schwarz: Teubner-Taschenbuch der Mathematik. Teubner-Verlag, 1996.

Anmerkung zur Literatur

Der in [23] beschriebene Beweis enthielt einen Fehler bezüglich der Primzahl zwei; Shafarevitsch schloss diese Lücke wenige Jahre später. Für einen vollständigen und modernen Beweis sei der Leser auf [17] verwiesen.