

# Skript zum Kurs Geometrie mit dem Computer Sommersemester 2005

Hans-Gert Gräbe, Univ. Leipzig

21. Juli 2005

## 0 Einführung

Die (synthetische) Geometrie ist eine sehr alte mathematische Disziplin und stand lange Zeit wohl für Mathematik schlechthin, ehe ihr dieser Platz durch eine stürmische Entwicklung der Mathematik in den letzten 200 Jahren von anderen Disziplinen streitig gemacht wurde. Natürlich hat sich in dieser Zeit auch die Geometrie weiter entwickelt. Teildisziplinen wie Differentialgeometrie oder Algebraische Geometrie untersuchen komplizierte geometrische Gebilde und haben zu wichtigen Einsichten über die Struktur von Raum (und Zeit) geführt. Die elementare Geometrie ist darüber, vollkommen zu unrecht, in die zweite Reihe gerückt. Das findet insbesondere seinen Ausdruck im Curriculum der Schule, in dem (elementar)geometrische Fragestellungen nur noch in geringem Umfang auftauchen.

Andererseits faszinieren solche Aufgaben immer wieder durch die Einfachheit, mit der relevante Probleme formuliert werden können, sowie den Scharfsinn und die Tiefgründigkeit der Argumentation, die zu deren Beantwortung herangezogen werden müssen. Sie sind damit für Hobbymathematiker, interessierte Schüler eingeschlossen, immer wieder eine Fundgrube von Problemen und Ideen, an denen man die eigene Argumentationskraft trainieren und verbessern kann. Die Vielfalt der Argumentationsmuster, die dabei zum Einsatz kommen, lassen eine Mechanisierung derartiger Beweisansätze als schier unmöglich erscheinen.

Besonders **Fragen der Konstruierbarkeit mit Zirkel und Lineal** haben Mathematiker verschiedener Epochen immer wieder fasziniert. So gehören die beiden großen Fragestellungen aus der antiken Mathematik nach der Verdopplung eines Würfels und der Dreiteilung eines beliebigen Winkels mit diesen Instrumenten zu den wohl auch außerhalb der Mathematik bekanntesten geometrischen Problemen. Trotz der Einfachheit der Fragestellung ließ sich deren Unlösbarkeit erst exakt nachweisen, als ein entsprechender **algebraischer Apparat**, in diesem Fall die Körpertheorie, entwickelt wurde. Eine solche *Methode der „Symbolisierung“ geometrischer Sachverhalte* in der Sprache der Algebra erlaubte es C.F. Gauß im Jahre 1796, die Konstruierbarkeit eines 17-Ecks nachzuweisen. Die entsprechenden Argumente sind heute in den meisten Standardwerken zur (höheren) Algebra als Anwendungsbeispiele dieser Theorie enthalten.

Ein exaktes Studium der mit der Konstruierbarkeit verbundenen Fragestellungen kommt um **eine ordentliche Fundierung, eine Axiomatisierung der Geometrie** nicht herum. Auch hier lassen sich die entsprechenden Ansätze bis in die Antike hinein, etwa zu den Büchern des Euklid, verfolgen. Mathematiker hat dabei immer interessiert, geometrische Aussagen und Konstruktionen mit möglichst geringen Voraussetzungen herzuleiten bzw. auszuführen. Die aus der Schule bekannte Geometrie setzt dabei das umfangreichste Instrumentarium voraus. Neben Punkten, Geraden, Parallelen, Längen und Winkelgrößen gibt es auch noch Strecken, Strahlen und Halbebenen, wozu auf jeder Geraden  $g$  (auf konsistente Weise) eine Ordnungsrelation zur Verfügung stehen muss, die

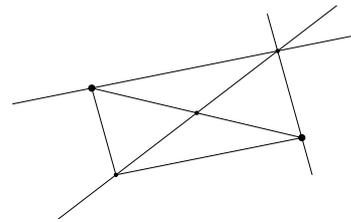
es erlaubt, für drei Punkte  $A, B, C \in g$  zu entscheiden, ob  $C$  zwischen  $A$  und  $B$  liegt. Geometrische Aussagen, die von Strahlen, Halbebenen und dieser *Zwischenrelation* Gebrauch machen, werden der **Ordnungsgeometrie** zugeordnet.

Da algebraische Verfahren, die wir zum Mechanisieren ausschließlich heranziehen werden, mit solchen Ordnungsrelationen nicht gut zusammenspielen, wollen wir derartige geometrische Aussagen im Weiteren ausklammern.

Damit wird der Kreis der betrachteten geometrischen Aufgaben aber nur etwas eingeschränkt, da viele Konfigurationen, in denen Strecken vorkommen, diese Ordnungsrelation in Wirklichkeit nicht ausnutzen. So kann man etwa den Mittelpunkt einer Strecke  $AB$  bestimmen, ohne zu wissen, wo auf der Geraden  $g = g(AB)$  links oder rechts ist, indem nach dem aus der Schule bekannten Verfahren die Kreise  $c(A, B)$  (mit Mittelpunkt  $A$  und Peripheriepunkt  $B$ ) und  $c(B, A)$  zum Schnitt gebracht und deren zwei Schnittpunkte miteinander verbunden werden. Der Schnittpunkt dieser Verbindungsgeraden mit  $g$  ist der zu konstruierende Mittelpunkt.

Eine solche Geometrie, welche nur von Punkten, Geraden, Parallelen, Längen und Winkelgrößen (und damit auch Senkrechten und Kreisen) gebrauch macht, wird als **Euklidische, Bewegungs- oder Kongruenzgeometrie** bezeichnet.

Allerdings benötigt man ein so umfangreiches Arsenal von Hilfsmitteln zur Konstruktion des Mittelpunkts einer Strecke nicht wirklich. Man kann den Mittelpunkt einer Strecke  $AB$  auch bestimmen, indem man einen dritten Punkt  $C$  beliebig wählt und das Parallelogramm  $ACBD$  konstruiert. Die Mitte der Strecke  $AB$  ist dann genau der Diagonalschnittpunkt in diesem Parallelogramm.



Affine Geometrie: Konstruktion des Mittelpunkts einer Strecke

Wir haben dafür die Möglichkeit der Euklidischen Geometrie, Längen (und Winkel) vorgegebener Größe in einem vorgegebenen Punkt anzutragen, nicht benötigt, sondern einzig die Möglichkeit, zu vorgegebenen Geraden Parallelen konstruieren zu können. Eine Geometrie, die nur mit Punkten, Geraden und Parallelen auskommt, bezeichnet man als **affine Geometrie**. Im Mittelpunkt dieser Geometrie stehen der Strahlensatz, Teilverhältnisse und Eigenschaften des Parallelogramms.

Noch allgemeinere Sätze der **projektiven Geometrie** erhält man, wenn man auch auf die Verwendung von Parallelen verzichtet. Derartige Sätze sind invariant unter projektiven Transformationen, d.h. solchen, die man in der Malerei bei der Übertragung einer weiten Landschaft auf die Staffelei antrifft, wenn sich die ehemals parallelen Geraden im Bild auf der *Horizontlinie* schneiden. Eine solche projektive Transformation  $\pi$  überträgt eine geometrische Konfiguration von einer Ebene  $\varepsilon$  (im Raum) auf eine andere Ebene  $\varepsilon'$  nach folgendem Verfahren:

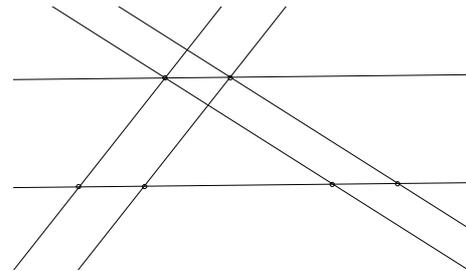
Wähle ein Projektionszentrum  $Z$  außerhalb der beiden Ebenen aus. Den Bildpunkt  $A' = \pi(A) \in \varepsilon'$  zu einem Original  $A \in \varepsilon$  findet man als den Schnittpunkt von  $g(AZ)$  mit  $\varepsilon'$ .

Offensichtlich gehen bei dieser Konstruktion Geraden in Geraden über. In der Tat, die Geraden  $g(AZ)$  für  $A \in g$  spannen eine Ebene auf, so dass die Bildpunkte auf der Schnittgeraden dieser Ebene mit  $\varepsilon'$  liegen. Allerdings besitzt nicht jeder Punkt  $A$  der Urbildebene  $\varepsilon$  einen Bildpunkt, denn die Gerade  $g(AZ)$  kann ja parallel zu  $\varepsilon'$  verlaufen. Die entsprechenden Punkte  $A$  mit dieser Eigenschaft liegen genau auf der Schnittgeraden von  $\varepsilon$  mit der Parallelen zu  $\varepsilon'$  durch  $Z$ . Diese Gerade bezeichnet man als die *Ausnahme Gerade* auf  $\varepsilon$ . Ihre Punkte werden in die „unendlich ferne“ Gerade von  $\varepsilon'$  abgebildet. Insbesondere sind die Bilder zweier Geraden, die sich in  $\varepsilon$  auf dieser Ausnahme Geraden schneiden, parallel zueinander. Genauso gibt es auf  $\varepsilon'$  eine Ausnahme Gerade.

Die Abbildung  $\pi$  ist jenseits der beiden Ausnahmegeraden eineindeutig. Erweitert man  $\varepsilon$  bzw.  $\varepsilon'$  jeweils durch Hinzunahme einer *Ferngeraden* zur projektiven Ebene  $\bar{\varepsilon}$  bzw.  $\bar{\varepsilon}'$ , wobei die jeweilige Ferngerade Bild bzw. Urbild der Ausnahmegeraden der anderen Ebene ist, so wird die Abbildung  $\pi$  sogar im Ganzen eineindeutig. Aussagen der projektiven Geometrie enthalten also typischerweise Formulierungen der Art „... die Geraden schneiden sich oder sind parallel zueinander ...“.

Wie kann man nun eine solche Vielfalt von Ansätzen unter einen Hut bringen? Zunächst waren es Mathematiker am Ende des 19. Jahrhunderts, vor allem F. Klein und D. Hilbert, die einen Zusammenhang zwischen dem Umfang der eingesetzten Konzepte und Transformationsgruppen herausfanden. Die aus der Schule bekannte Phrase „eindeutig bis auf Kongruenz“ besagt genau dies. Aussagen der Bewegungsgeometrie, etwa die Konstruktion eines Dreiecks aus vorgegebenen drei Streckenlängen, sind immer nur eindeutig bis auf Kongruenztransformationen möglich. Form und Größe des Dreiecks sind eindeutig bestimmt, seine Lage in der Zeichenebene kann durch Drehung, Verschiebung und Spiegelung weitgehend frei gewählt werden. Die zugehörige Bewegungsgruppe ist die  $O(2)$ , die Gruppe der orthogonalen Transformationen der Ebene. Streckenlängen und Winkelgrößen bleiben dabei erhalten, so dass orthonormale Koordinatensysteme bei solchen Transformationen in orthonormale Koordinatensysteme überführt werden. Solche Koordinatensysteme bezeichnen wir auch als *karthesische Koordinaten*.

Aussagen der affinen Geometrie bleiben unter weitergehenden Transformationen erhalten. Die zugehörige Gruppe ist die  $Gl(2)$ , die orthonormale Koordinatensysteme in schiefwinklige überführt und auch die „Längen“ der Einheitsstrecken nicht erhält (aber Längen gibt es ja nicht). Allerdings kann man durch Parallelogramme wenigstens Strecken vorgegebener Länge auf parallelen (und mit einem transitiven Ansatz damit auch auf derselben) Geraden abtragen, was Grundlage für (unabhängige) Koordinaten auf wenigstens jeder der beiden Achsen ist. Natürlich muss ein exakt arbeitender Mathematiker hier auch einen Eindeutigkeitsatz beweisen. Wie lautet der Satz und wie geht der Beweis?



Affine Geometrie: Abtragen einer Streckenlänge auf derselben Geraden

Schließlich gibt es noch weitergehende Transformationen, unter denen Aussagen der projektiven Geometrie erhalten bleiben. Die zugehörige Gruppe  $PSl(2)$  ist mit projektiven oder homogenen Koordinaten verbunden, auf die hier zunächst nicht eingegangen werden soll.

Unsere hauptsächlichen Arbeitsmittel werden die Einführung von Koordinaten und Methoden der analytischen Geometrie sein. Es stellt sich dabei heraus, dass es ein solcher Ansatz gestattet, konstruktive, also algorithmische Ansätze auf der Seite der Geometrie mit Hilfe eines informatiktheoretischen Hilfsmittels, des Unterprogramms, in einem symbolischen Kontext in vielen Fällen so auszuwerten, dass sich daraus ein im mathematischen Sinne exakter Beweis ergibt.

Das Vorhandensein eines Koordinatensystems werden wir dabei als gegeben voraussetzen. Da hierfür allein die Festlegung einer Einheitsstrecke und deren Übertragbarkeit an alle Orte und Richtungen der Ebene gewährleistet sein muss, stellt das wenigstens für Problemstellungen innerhalb der Euklidischen Geometrie keine Einschränkung dar. Das Vorhandensein eines Koordinatensystems kann allerdings aus noch viel allgemeineren Annahmen heraus abgeleitet werden. Diese Frage steht im Zentrum der axiomatischen Einführung der Geometrie und wird deshalb in den entsprechenden Lehrbüchern umfassend abgehandelt. Insbesondere in der Monographie [5] von W.-T. Wu sind dazu interessante Ausführungen enthalten, in denen auch Koordinatensysteme über nichtkommutativen Zahlbereichen eine Rolle spielen. Wir werden darauf nicht näher eingehen.

Eine weitere praktische Anwendung der Koordinatenmethode wird uns allerdings interessieren, denn sie ist auch die Basis für die bildliche Darstellung geometrischer Konfigurationen in Grafik-Software, so dass dieser Kurs gegenüber früheren einen stärkeren informatik-praktischen Bezug

haben wird. Wir werden parallel zu den mathematischen Fragen auch die Modellierung in entsprechender Dynamischer Geometrie-Software (DGS) studieren, wozu wir den im Hause entwickelten Java-basierten „GeoFuchs“ als Grundlage heranziehen werden. Er ist noch nicht sehr vollkommen, deshalb wird auch andere DGS, insbesondere das Programm „Zirkel und Lineal“ von René Grothmann aus Eichstätt, Verwendung finden. Die grundlegende Struktur all dieser Programme ist aber ähnlich.

## 1 Einige Sätze aus der ebenen Geometrie

In diesem Kapitel wollen wir zunächst einige einfache und weniger einfache Sätze aus der ebenen Geometrie kennenlernen bzw. uns wieder ins Gedächtnis zurückrufen.

Dies soll zum einen ausreichendes Material für die weiteren Betrachtungen zur Verfügung stellen, an dem sich zu entwickelnde algorithmische Ansätze werden demonstrieren lassen, und zum anderen die Vielfalt geometrischer Argumente noch einmal demonstrieren, die im Rahmen einer Mechanisierung unter einen gemeinsamen Hut zu bringen sind.

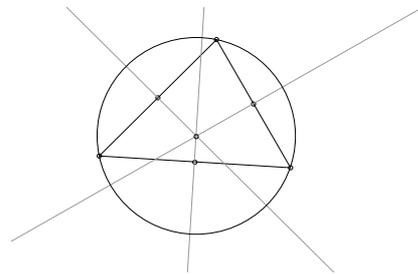
Außerdem sollen wichtige Begriffe, die beim Beweisen geometrischer Sachverhalte eine Rolle spielen, beispielhaft demonstriert werden.

### 1.1 Sätze über die Ecktransversalen im Dreieck

#### Satz vom Schnittpunkt der Mittelsenkrechten

Verwendet den Begriff der *Ortslinie*: Eine Mittelsenkrechte besteht aus genau den Punkten, die von zwei gegebenen Punkten  $A, B$  den gleichen Abstand haben. Aus dem Beweis ergibt sich dann außerdem, dass der Schnittpunkt  $O$  von allen drei Eckpunkten gleichweit entfernt ist, also der Umkreismittelpunkt sein muss.

Eine Ortslinie verbindet eine geometrische (Mittelsenkrechte als Gerade) mit einer logischen ( $P$  mit  $|AP| = |BP|$ ) Eigenschaft. Ihre Beweiskraft entwickeln Ortslinien aus dem Zusammenspiel beider Seiten.



Die Mittelsenkrechten eines Dreiecks schneiden sich in einem Punkt, dem Umkreismittelpunkt

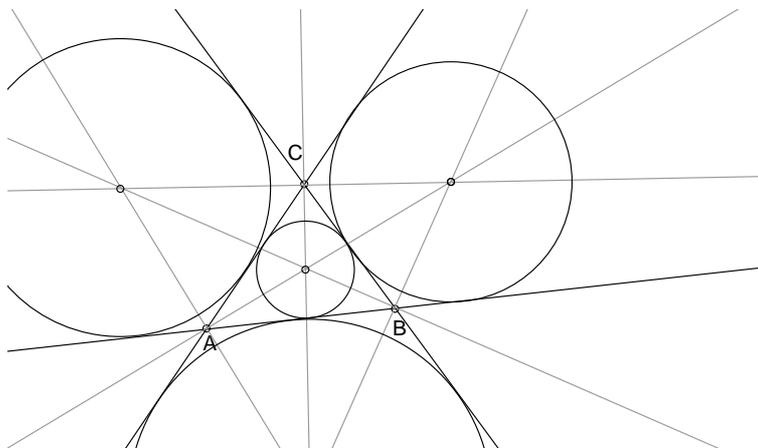
**Beweis:** Sei  $ABC$  das gegebene Dreieck,  $D, E, F$  die Mittelpunkte der Seiten  $\overline{BC}, \overline{AC}, \overline{AB}$  und  $M$  der Schnittpunkt der beiden Mittelsenkrechten  $m_{AB}$  und  $m_{BC}$ . Wir zeigen, dass  $M$  dann auch auf der dritten Mittelsenkrechten liegt.

$$\begin{aligned} M \in m_{AB} &\Rightarrow |MA| = |MB| \\ M \in m_{BC} &\Rightarrow |MB| = |MC| \end{aligned}$$

Daraus folgt  $|MA| = |MC|$ , also  $M \in m_{AC}$ .  $\square$

#### Satz vom Schnittpunkt der Winkelhalbierenden

Auch hier spielen Ortslinien eine Rolle: Eine Winkelhalbierende besteht aus genau den Punkten, die von zwei gegebenen Geraden den gleichen Abstand haben. Aus dem Beweis ergibt sich dann außerdem, dass der Schnittpunkt  $W$  von allen drei Dreiecksseiten gleichweit entfernt ist, also der Inkreismittelpunkt sein muss.



Inkreis und Ankreise eines Dreiecks

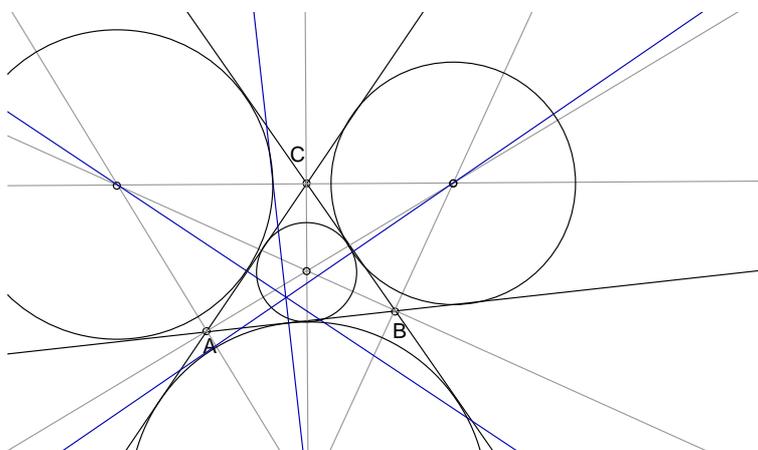
Allerdings ist das ungenau, denn der genannte geometrische Ort besteht aus den Punkten auf dem Geraden-Paar, das die Halbierenden von Innen- und Außenwinkel bilden. Innenwinkel und Außenwinkel sind allerdings Begriffe der Ordnungsgeometrie und deshalb von unserem Standpunkt aus nicht zu unterscheiden. Die drei Winkelhalbierendenpaare haben neben dem Inkreismittelpunkt noch die drei Mittelpunkte der Ankreise gemeinsam.

Man beachte die Ähnlichkeit zum Begriff des *Parallelenpaars* als dem geometrischen Ort aller Punkte, die von einer gegebenen Gerade einen vorgegebenen Abstand haben.

Zwei Geraden schneiden sich (normalerweise) immer in genau einem Punkt. Wenn drei Geraden durch einen gemeinsamen Punkt gehen oder parallel sind, so liegt schon eine besondere Situation vor. Solche Geraden nennt man *konkurrent*.

Umgekehrt geht durch zwei Punkte (normalerweise) immer genau eine Gerade. Wenn drei Punkte auf einer gemeinsamen Geraden liegen, so liegt ebenfalls eine besondere Situation vor. Solche Punkte nennt man *kollinear*.

Ein nicht so einfaches Beispiel für Geraden am Dreieck, die durch einen gemeinsamen Punkt gehen, kann man als Nebeneffekt der Konstruktion der Ankreise beobachten: Um einen Kreis zu zeichnen brauchen wir neben dem Mittelpunkt einen Punkt auf der Peripherie, im Fall des Ankreises also den Lotfußpunkt aus dem Ankreismittelpunkt auf die zugehörige Dreiecksseite. Diese drei Lote gehen durch einen gemeinsamen Punkt!

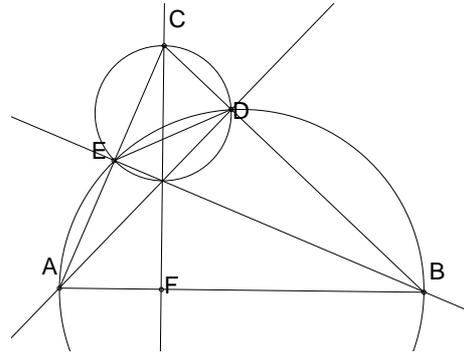


Die Lote aus den Ankreiszentren gehen durch einen gemeinsamen Punkt

**Satz vom Höhenschnittpunkt**

Beweis über Mehrfachanwendung von Thaleskreis und Peripheriewinkelsatz.

Eine weitere Anwendung des Sehnensatzes liefert die zusätzliche Eigenschaft, dass der Höhenschnittpunkt die Höhen so teilt, dass die Produkte aus den beiden Höhenabschnitten jeweils gleichgroß sind.



**1.2 Der Satz von Ceva**

All diese Sätze kann man aus einem allgemeinen Prinzip über Teilverhältnisse von Transversalen am Dreieck herleiten.

**Satz 1 (Satz des Ceva)** *Drei Ecktransversalen des Dreiecks  $\triangle ABC$  mögen die gegenüberliegenden Seiten in den Punkten  $D, E, F$  schneiden. Diese drei Ecktransversalen gehen genau dann durch einen Punkt, wenn*

$$\frac{|BD|}{|DC|} \cdot \frac{|CE|}{|EA|} \cdot \frac{|AF|}{|FB|} = 1$$

*gilt.*

Beweis durch Flächenzerlegung.

In diesem Beweis spielen Verhältnisse der Art  $t = TV(A, B; F) = |AF| / |FB|$  für kollineare Punkte  $A, B, F$  eine wichtige Rolle, die auch als *Teilverhältnis* bezeichnet werden. Für Punkte innerhalb  $\overline{AB}$  bestimmt es die Lage von  $F$  eindeutig. Verwendet man gerichtete Streckenlängen, so kann man dieses Teilverhältnis auch für Punkte  $F$  auf  $AB$  definieren, die außerhalb der Strecke liegen. Für Punkte jenseits von  $A$  ergeben sich Werte  $-1 < t < 0$ , für Punkte jenseits von  $B$  ergibt sich  $t < -1$ . Stets bestimmt der Wert von  $t$  die Lage von  $F$  eindeutig. Die Ausnahmen  $F = B$  sowie  $t = 1$  lassen sich durch Hinzunahme eines Werts  $t = \infty$  sowie eines Fernpunkts auf der Geraden einordnen. Die Punkte der (projektiven) Geraden  $AB$  werden so durch  $t \in \mathbb{P}^1$  parametrisiert, wobei als Bezugsgrößen die Punkte  $A(t = 0), B(t = \infty)$  und der Fernpunkt ( $t = -1$ ) dienen.

Als Folgerung erhält man neue Beweise der Sätze vom Schnittpunkt der Seitenhalbierenden und vom Höhenschnittpunkt.

**Aufgabe 1**

- a) *Zeigen Sie, dass man aus dem Satz des Ceva auch den Satz vom Höhenschnittpunkt herleiten kann, indem Sie die Längen der Seitenabschnitte durch geeignete trigonometrische Formeln ausdrücken.*
- b) *Zeigen Sie, dass die Transversalen zu den Berührungspunkten des Inkreises sich in einem Punkt schneiden. (Hinweis: Berechnen Sie die Länge der einzelnen Tangentenabschnitte.)*

Weitere Folgerung:

*Die Transversalen zu den drei Berührungspunkten der Ankreise schneiden sich in einem Punkt. (Dieser Punkt heißt auch Nagelscher Punkt)*

Wir betrachten dazu verschiedene Beziehungen zwischen den Tangentenabschnitten zu den Berührungspunkten der drei Ankreise:

$$\begin{array}{rcl} x + y + z & = & r + s + t \\ x + y + t & = & z + r + s \\ \hline \Rightarrow z & = & t \end{array}$$

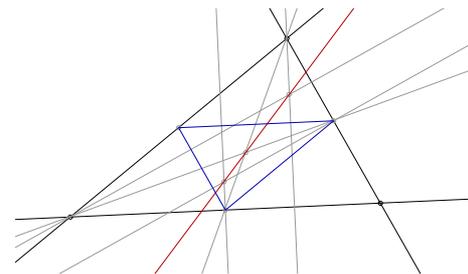
Analog gilt  $x = r$ ,  $y = s$ .

### 1.3 Weitere Sätze am Dreieck

**Satz 2 (Eulersche Gerade)** *In einem Dreieck liegen der Höhenschnittpunkt  $H$ , der Schwerpunkt  $S$  und der Umkreismittelpunkt  $M$  auf einer Geraden.  $S$  teilt  $\overline{HM}$  im Verhältnis 2:1.*

Im Beweis werden Dreiecke in Ähnlichkeitslage verwendet.

Zwei Dreiecke  $ABC$  und  $A'B'C'$  heißen in Ähnlichkeitslage, wenn einander entsprechende Seiten jeweils zueinander parallel sind.

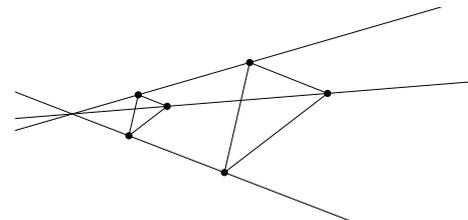


Die Eulersche Gerade

Dann kann man schlussfolgern, dass das eine Dreieck aus dem anderen durch eine Zentralstreckung hervorgeht. Diese Aussage ist Gegenstand des folgenden Satzes.

**Satz 3 (Der affine Satz von Desargue)**

1. Sind  $\triangle ABC$  und  $\triangle A'B'C'$  in Ähnlichkeitslage, d.h.  $AB \parallel A'B'$ ,  $AC \parallel A'C'$  und  $BC \parallel B'C'$ , so sind  $AA'$ ,  $BB'$  und  $CC'$  konkurrent oder parallel.
2. Sind umgekehrt  $AA'$ ,  $BB'$  und  $CC'$  konkurrent oder parallel und  $AB \parallel A'B'$ ,  $AC \parallel A'C'$ , so gilt auch  $BC \parallel B'C'$ .



Der (affine) Satz von Desargue

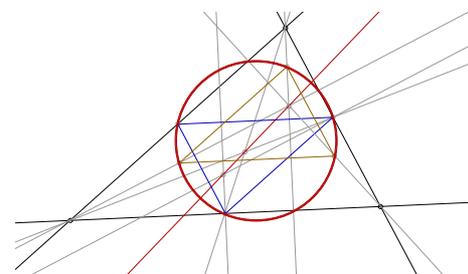
**Aufgabe 2** *Leiten Sie diesen Satz aus dem Strahlensatz her.*

Der Satz von Desargue spielt in der Fundierung der Geometrie als „schwache Version des Strahlensatzes“ eine wichtige Rolle. Er ist schwächer als der Strahlensatz und führt damit zu einer umfassenderen als der affinen Geometrie. Details finden sich im Buch [5].

**Satz 4 (Der Satz vom Feuerbachschen Kreis)**

*Der Mittelpunkt  $N$  von  $HM$  ist der Mittelpunkt eines Kreises, auf dem neun ausgezeichnete Punkte des Dreiecks  $\triangle ABC$  liegen, und zwar*

- die drei Seitenmitten,
- die drei Höhenfußpunkte und
- die drei Mitten der oberen Höhenabschnitte



Der Feuerbachsche Kreis

**Beweis:** Die beiden Dreiecke, die durch die Seitenmitten bzw. die Mitten der oberen Höhenabschnitte aufgespannt werden, sind in Ähnlichkeitslage mit dem Faktor  $(-1)$ , also zueinander

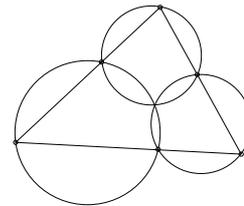
kongruent. Da dabei  $H$  als Höhenschnittpunkt des Höhendreiecks in  $M$  als Höhenschnittpunkt des Mittendreiecks übergeht, ist die Mitte  $N$  der Strecke  $MH$  gerade das Zentrum dieser Punktspiegelung, d.h. Drehung um  $180^\circ$ . Das Dreieck  $ABC$  geht bei Streckung um den Faktor  $-\frac{1}{2}$  mit Zentrum  $S$  in das Mittendreieck über, dessen Umkreismittelpunkt  $M$  also in den Umkreismittelpunkt des Mittendreiecks. Bild von  $M$  bei dieser Streckung ist aber gerade  $N$ . Also geht ein Kreis mit Zentrum in  $N$  die genannten 6 Punkte. Weiter entsprechen sich bei der Punktspiegelung mit Zentrum in  $N$  Seitenmitte und gegenüberliegende Mitte des oberen Höhenabschnitts. Die Verbindungsgerade geht also durch das Streckungszentrum  $N$  und ist ein Durchmesser des Feuerbachkreises. Aus dem Satz des Thales folgt schließlich, dass auch der Höhenfußpunkt auf dem Feuerbachkreis liegt.  $\square$

Der Feuerbachkreis hat eine weitere, mit elementargeometrischen Mitteln nur schwer zu beweisende Eigenschaft: Er berührt den Inkreis des Dreiecks  $\triangle ABC$  und dessen drei Ankreise.

### 1.4 Sätze am Dreieck aus der neueren Forschung

Wir erwähnen hier nur zwei Sätze, die im Zusammenhang mit dem mechanisierten Theorembeweisen in der Geometrie immer wieder als Beispiele herangezogen werden.

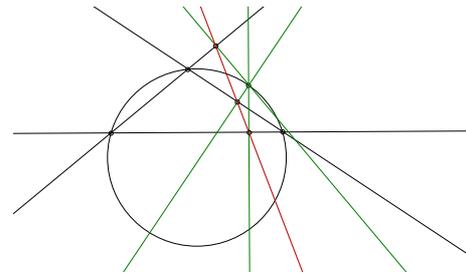
**Satz 5 (Der Miquelsche Punkt)**  $P, Q, R$  seien Punkte auf den Seiten des Dreiecks  $\triangle ABC$ . Zeichnet man durch jede Ecke und die beiden Punkte, welche auf den zu dieser Ecke inzidenten Seiten liegen, Kreise, so gehen diese durch einen gemeinsamen Punkt.



Der Miquelsche Punkt

Der Beweis folgt fast unmittelbar aus dem Satz über das Sehnenviereck und dessen Umkehrung.

**Satz 6 (Die Simsonsche Gerade)** Fällt man von einem Punkt  $P$  außerhalb eines Dreiecks  $\triangle ABC$  die Lote auf die Dreiecksseiten oder deren Verlängerungen, so liegen die drei Fußpunkte der Lote genau dann auf einer Geraden, wenn  $P$  auf dem Umkreis des Dreiecks  $\triangle ABC$  liegt.



Die Simsonsche Gerade

Beweis: Seien die drei Fußpunkte mit  $A_1, B_1, C_1$  bezeichnet (d.h.  $A_1$  ist der Fußpunkt des Lots von  $P$  auf  $BC$  usw.). Der Thaleskreis über  $\overline{PC}$  liefert  $|\angle B_1A_1P| = |\angle B_1CP|$ . Der Thaleskreis über  $\overline{PA}$  liefert  $|\angle B_1C_1P| = |\angle B_1AP|$ . Im Sehnenviereck  $ABCP$  ist  $|\angle APC| = 180^\circ - \beta$ . Wegen zweier rechter Winkel ist aber auch  $|\angle A_1PC_1| = 180^\circ - \beta$ . Folglich ist  $|\angle A_1B_1C_1| = |\angle AB_1C| = 180^\circ$ ,  $A_1, B_1, C_1$  also kollinear.  $\square$

Zu beiden Sätzen gäbe es noch eine Menge zu sagen. So kann man etwa zu jedem Punkt  $P$  im Inneren des Dreiecks  $\triangle ABC$  Punkte  $A_1, B_1, C_1$  so auf den Dreiecksseiten finden, dass  $P$  der zugehörige Miquelsche Punkt ist. Ein solches Dreieck bekommt man (Thalesatz!) insbesondere dann, wenn man von  $P$  aus die Lote auf die drei Dreiecksseiten fällt. Dieses Dreieck wird auch als das zum Punkt  $P$  gehörende *Fußpunktdreieck* bezeichnet.

**Aufgabe 3**

a) [2, 1.91] Zeigen Sie, dass die Seiten des Fußpunktdreiecks von  $P$  die Längen

$$\frac{ax}{2r}, \frac{by}{2r}, \frac{cz}{2r}$$

haben, wobei  $a = |BC|$ ,  $b = |AC|$ ,  $c = |AB|$  die Längen der Seiten des Dreiecks  $\triangle ABC$ ,  $r$  dessen Umkreisradius und  $x = |AP|$ ,  $y = |BP|$ ,  $z = |CP|$  die Abstände von den Eckpunkten zu  $P$  sind.

b) Zeigen Sie, dass für den Flächeninhalt des Fußpunktdreiecks  $\triangle A_1B_1C_1$  bzgl.  $P$

$$F(A_1B_1C_1) = \frac{r^2 - |MP|^2}{4r^2} \cdot F(ABC)$$

gilt, wobei  $M$  der Umkreismittelpunkt ist.

Das eben betrachtete Fußpunktdreieck entartet zu einer Geraden, wenn  $P$  auf dem Umkreis des Dreiecks  $\triangle ABC$  liegt und ergibt dann genau die Simsonsche Gerade. Von Satz über die Simsonsche Gerade gilt übrigens auch die Umkehrung:

**Satz 7 (Simsonsche Gerade II)** Sind die Lotfußpunkte  $A_1, B_1, C_1$  von einem Punkt  $P$  auf die Seiten des Dreiecks  $\triangle ABC$  oder deren Verlängerungen kollinear, so liegt  $P$  auf dem Umkreis des Dreiecks.

**Aufgabe 4**

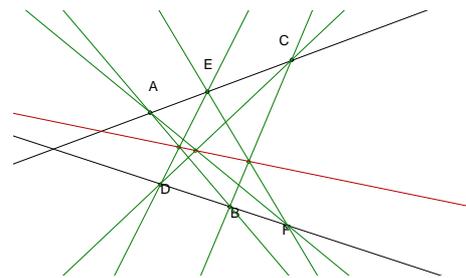
- a) Leiten Sie aus der Flächeninhaltsformel für das Fußpunktdreieck (vorige Aufgabe) einen zweiten Beweis für den Satz über die Simsonsche Gerade her.
- b) Beweisen Sie die Umkehrung des Satzes von der Simsonschen Geraden.

Weitere interessante Sätze, die an dieser Stelle vielleicht noch zu nennen wären (alle aus [2]): das Schmetterlings-Theorem oder der Satz von Morley.

**1.5 Sätze der projektiven Geometrie**

Wir wollen dieses Kapitel mit einigen Sätzen aus der projektiven Geometrie beschließen, die ob der verwendeten Mittel (meist nur genügend verzwickte Geradenkonfigurationen) einen ganz speziellen Reiz besitzen.

**Satz 8 (Theorem von Pappus)** Sind  $A, C, E$  und  $B, D, F$  jeweils kollineare Punkte, so sind auch die Schnittpunkte  $AB \cap DE, BC \cap EF$  und  $CD \cap AF$  kollinear. Die entstehende Gerade wird als Pappus-Gerade bezeichnet.



Pappus-Gerade

Beweis: Wir beweisen diesen Satz zuerst in einer speziellen Situation, in der zwei der drei Verbindungegeraden zueinander parallel sind:

$$AB \parallel DE, BC \parallel EF \Rightarrow CD \parallel AF$$

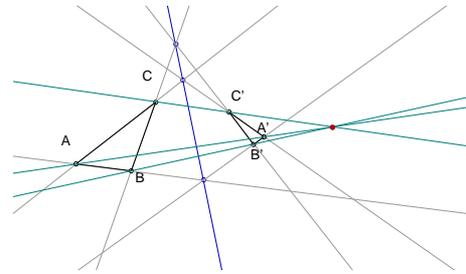
Der Beweis ergibt sich hier unmittelbar aus dem Strahlensatz und seinen Umkehrungen. Für den allgemeinen Beweis verwenden wir eine projektive Transformation, bei der die Gerade  $g(AB \cap DE, BC \cap EF)$  die Ausnahme Gerade ist, d.h. in der Bildebene in die Ferngerade übergeht.

Damit sind für das Bild unserer Geradenkonfiguration die Voraussetzungen des bewiesenen Spezialfalls erfüllt und die Bilder von  $g(CD)$  und  $g(AF)$  sind zueinander parallel. Die Urbilder schneiden sich damit auf der Ausnahmegeraden, so dass die drei Schnittpunkte kollinear sind.  $\square$

Übrigens ist auch der Satz von Desargue eigentlich ein projektiver Satz.

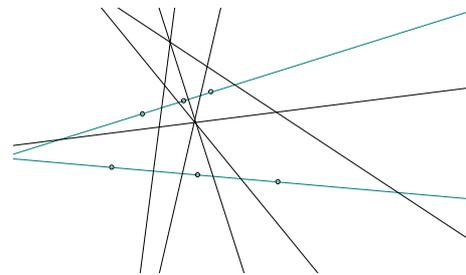
**Satz 9 (Allgemeiner Satz von Desargue)** Sind die Schnittpunkte  $AB \cap A'B'$ ,  $AC \cap A'C'$  und  $BC \cap B'C'$  kollinear, so gehen die drei Geraden  $AA'$ ,  $BB'$  und  $CC'$  durch einen gemeinsamen Punkt.

Beweis: Wir betrachten eine projektive Transformation, die die Gerade durch die drei Schnittpunkte auf die Ferngerade abbildet. Dann haben wir gerade die Situation des affinen Satzes von Desargue vorliegen.  $\square$



Allgemeiner Satz von Desargue

Eine interessante Fragestellung, die wir zum selben Thema hier nur aufwerfen wollen, entsteht aus dem Vergleich verschiedener Pappus-Geraden. Sind  $A_1, A_2, A_3$  und  $B_1, B_2, B_3$  jeweils kollinear, so führen die verschiedenen Permutationen der Punkte  $B_1, B_2, B_3$  zu insgesamt 6 solchen Geraden (Permutationen der Punkte auf der anderen Geraden haben keinen Einfluss: Ist  $(\sigma, \tau)$  ein Paar von Permutationen der Punkte  $(A)$  und  $(B)$ , so liefert die Permutation  $(1, \sigma^{-1}\tau)$  dieselbe Pappus-Gerade). Es stellt sich heraus, dass drei dieser Geraden durch einen gemeinsamen Punkt gehen.



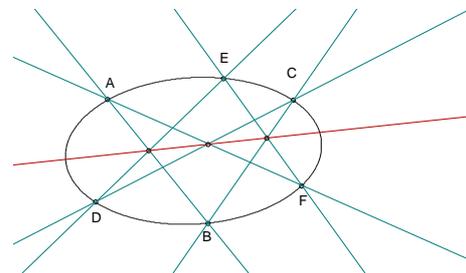
Die 6 Pappusgeraden zu einer Punktconfiguration

Der Satz von Pappus ist ein Spezialfall eines noch allgemeineren Satzes der projektiven Geometrie. Wir betrachten dazu eine Konfiguration aus sechs Punkten der Ebene  $A, B, C, D, E, F$ , für die  $X = AB \cap DE, Y = BC \cap EF$  und  $Z = CD \cap AF$  kollinear sind. In einer solchen Konfiguration sind die ersten fünf Punkte frei wählbar, wobei auch die Lage von  $X$  bestimmt wird. Die Richtung der Geraden  $g = g(EF)$  bestimmt  $Y$  eindeutig und somit auch  $Z = CD \cap XY$ .  $F$  ergibt sich dann als  $F = AZ \cap g$ . Eine solche Punktekonfiguration bezeichnet man als ein *Pascalsches Sechseck*.

Führt man Koordinaten ein, so stellt sich heraus, dass sich solche Punktekonfigurationen analytisch recht einfach charakterisieren lassen. Es gilt der folgende

**Satz 10 (Satz von Pascal)** Sechs Punkte spannen genau dann ein Pascalsches Sechseck auf, wenn sie auf einer Kurve zweiten Grades liegen.

Auf einen vollständigen Beweis dieses Satzes muss hier verzichtet werden, da hierfür noch Kurven zweiten Grades einzuführen wären. Man kann aber wie im Beweis des Satzes von Pappus einen wesentlichen Teil durch eine projektive Transformation aus einer spezielleren Situation herleiten, die Gegenstand der folgenden Aufgabe ist.



Pascalsches Sechseck

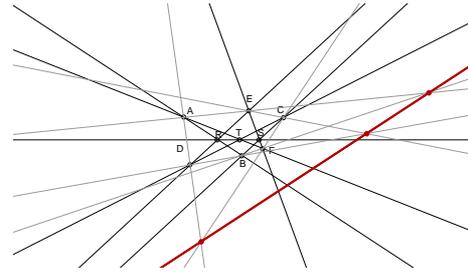
**Aufgabe 5** Beweisen Sie den folgenden Spezialfall des Satzes von Pascal:

Liegen die Punkte  $A, B, C, D, E, F$  auf einem Kreis und gilt  $AB \parallel CD$  und  $BC \parallel EF$ , so ist auch  $CD \parallel EF$ .

Eine wichtige Folgerung aus dem Satz von Pascal ist die aus der ursprünglichen Definition nicht ersichtliche Tatsache, dass jede Permutation von Punkten, die ein Pascalsches Sechseck aufspannen, wieder ein solches bilden. Das liefert Aussagen der ebenen Geometrie, zu deren Formulierung keine Kurven zweiter Ordnung benötigt werden. Solche Sätze heißen in [5] *Sätze vom Pascal-Typ*. Ein solcher wäre z.B. die folgende Aussage:

Sind  $R = AB \cap DE, S = BC \cap EF, T = CD \cap AF$  kollinear, so sind auch  $X = AD \cap CF, Y = BD \cap CE, Z = BF \cap AE$  kollinear.

Um dieses Bild zu zeichnen müssen die Punkte in geeigneter Reihenfolge eingeführt werden:  $R, T$  und  $A$  können frei gewählt werden. Die Punkte  $S \in RT, B \in AR, C \in BS, D \in CT$  und  $E \in DR$  ergeben sich als (semifreie) Gleiter auf Geraden, der letzte Punkt  $F = ES \cap AT$  als Schnittpunkt zweier Geraden.



Satz vom Pascal-Typ

## 1.6 Zur Dualität von Punkten und Geraden in der projektiven Geometrie

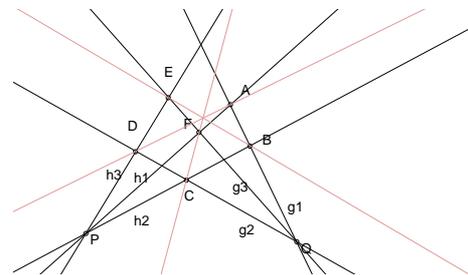
In vielen geometrischen Aussagen über Punkte und Geraden kann man die Worte „Punkt“ und „Gerade“ vertauschen und bekommt einen ebenfalls gültigen geometrischen Satz. Die einfachsten Aussagen dieser Art sind

- Es gibt genau eine Gerade durch zwei (voneinander verschiedene) Punkte.
- Zwei voneinander verschiedene Geraden haben genau einen Schnittpunkt (oder sind parallel).

Die Sonderrolle zueinander paralleler Geraden kann man aufheben, wenn man von der affinen zur projektiven Ebene übergeht, die man durch Hinzufügen der Punkte auf einer Ausnahmegereaden erhält, die „unendlich weit“ entfernt liegen, so dass zwei parallele Geraden genau einen gemeinsamen Punkt auf dieser Ferngeraden haben.

Als Beispiele für solche „dualen“ Sätze betrachten wir zunächst den folgenden Satz:

**Satz 11 (Dualer Satz von Pappus)** Seien  $g_1, g_2, g_3$  und  $h_1, h_2, h_3$  jeweils konkurrente Geraden, die sich in  $P$  bzw.  $Q$  schneiden und  $A, \dots, F$  die Schnittpunkte  $A = h_1 \cap g_1, B = g_1 \cap h_2, C = h_2 \cap g_2, D = g_2 \cap h_3, E = h_3 \cap g_3$  und  $F = g_3 \cap h_1$ . Die drei Verbindungsgeraden  $AD, BE$  und  $CF$  gehen durch einen gemeinsamen Punkt.

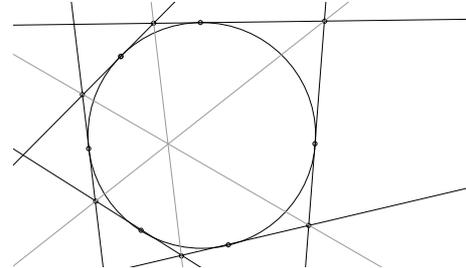


Dualer Satz von Pappus

**Aufgabe 6** Durch eine projektive Transformation kann man die Punkte  $P$  und  $Q$  auf die Ferngerade legen und bekommt so einen (gleichwertigen) Satz über zwei Tripel paralleler Geraden. Formulieren und beweisen Sie diese Aussage.

Da bei projektiven Transformationen zwar keine Kreise erhalten bleiben, aber Quadriken in Quadriken übergehen, ist auch der Satz von Pascal eine Aussage der projektiven Geometrie. Vertauscht man hier die Rolle von „Punkt“ und „Gerade“, so erhält man den Satz von Brianchon.

**Satz 12 (Satz von Brianchon)** Die Geraden  $a, b, c, d, e, f$  mögen eine Quadrik berühren, so dass sich „benachbarte“ Geraden in den Punkten  $A, B, C, D, E, F$  schneiden (d.h.  $ABCDEF$  ist ein Tangentensechseck). In jedem solchen Tangentensechseck gehen die Diagonalen  $AD, BE$  und  $CF$  zwischen gegenüber liegenden Eckpunkten durch einen gemeinsamen Punkt.



Der Satz von Brianchon

## 2 Die Koordinatenmethode

### 2.1 Grundlegende geometrische Zusammenhänge in koordinatengeometrischer Interpretation

Für die Visualisierung geometrischer Konfigurationen spielt die Darstellung durch Koordinaten eine zentrale Rolle. Im klassischen Zugang der ebenen Geometrie werden dazu Punkte  $P$  durch Koordinaten  $(p_1, p_2)$  im Punktraum  $\mathbb{A}_2$  dargestellt und Darstellungen anderer geometrischer Objekte daraus abgeleitet. Geraden können etwa durch zwei Punkte, ein Kreis durch Zentrum und Peripheriepunkt gegeben werden.

Eine kompakte Geradendarstellung ergibt sich durch Tripel  $g = (g_1, g_2, g_3)$ , das für die Gerade aus den Punkten  $\{ (x_1, x_2) : g_1x_1 + g_2x_2 + g_3 = 0 \}$  steht. Ein solches Tripel bezeichnet man als *homogene Koordinaten* der Geraden  $g$ . Zueinander proportionale Tripel beschreiben dieselbe Gerade  $g$  und für (echte) Geraden dürfen  $g_1$  und  $g_2$  nicht gleichzeitig verschwinden.

Die wichtigsten geometrischen Eigenschaften von Punkten und Geraden spiegeln sich dann in den folgenden Formeln wider:

- $A, B, C$  sind *kollinear*, d.h. liegen auf einer gemeinsamen Geraden  $g$  genau dann, wenn das homogene lineare Gleichungssystem

$$\begin{aligned} g_1a_1 + g_2a_2 + g_3 &= 0 \\ g_1b_1 + g_2b_2 + g_3 &= 0 \\ g_1c_1 + g_2c_2 + g_3 &= 0 \end{aligned}$$

eine (nichttriviale) Lösung in  $(g_1, g_2, g_3)$  besitzt, d.h. wenn

$$\begin{pmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{pmatrix} = 0$$

gilt.

- Analog gehen drei Geraden  $g, h, k$  durch einen gemeinsamen Punkt genau dann, wenn das lineare Gleichungssystem

$$\begin{aligned} g_1x_1 + g_2x_2 + g_3 &= 0 \\ h_1x_1 + h_2x_2 + h_3 &= 0 \\ k_1x_1 + k_2x_2 + k_3 &= 0 \end{aligned}$$

eine Lösung in  $(x_1, x_2)$  besitzt. Das ist genau dann der Fall, wenn die zugehörige Koeffizientenmatrix denselben Rang wie die erweiterte Koeffizientenmatrix hat. Da dieser Rang höchstens 2 sein kann, muss also

$$\det \begin{pmatrix} g_1 & g_2 & g_3 \\ h_1 & h_2 & h_3 \\ k_1 & k_2 & k_3 \end{pmatrix} = 0$$

gelten. Ist der Rang der Koeffizientenmatrix gleich 2, so hat das System dann eine (eindeutig bestimmte) Lösung. Ist ihr Rang dagegen gleich 1, d.h. sind ihre drei Zeilen  $(g_1, g_2)$ ,  $(h_1, h_2)$  und  $(k_1, k_2)$  zueinander proportional, so sind die drei Geraden  $g, h, k$  zueinander parallel, schneiden sich also “im Unendlichen” oder fallen zusammen.

Geraden, die durch einen gemeinsamen Punkt gehen oder zueinander parallel sind, bezeichnet man als *konkurrente Geraden*.

- Für die Parameter einer Geraden durch zwei Punkte  $A, B$  erhalten wir aus der Zwei-Punkte-Gleichung

$$(g_1, g_2, g_3) = (b_2 - a_2, a_1 - b_1, a_2 b_1 - a_1 b_2)$$

- Zwei Geraden  $g, h$  sind parallel genau dann, wenn  $g_1 h_2 - h_1 g_2 = 0$  gilt, d.h. ihre Normalenvektoren  $(g_1, g_2)$  und  $(h_1, h_2)$  zueinander parallel sind.
- Die Parameter der Parallelen  $h$  zu  $g$  durch einen Punkt  $P$  ergeben sich durch Adjustieren des Absolutglieds von  $g$  als

$$(h_1, h_2, h_3) = (g_1, g_2, -(g_1 p_1 + g_2 p_2)).$$

- Die Koordinaten des Schnittpunkts  $P$  zweier Geraden  $g, h$  berechnet sich als Lösung des entsprechenden Gleichungssystems nach der Cramerschen Regel zu

$$(p_1, p_2) = \left( \frac{g_2 h_3 - g_3 h_2}{d}, \frac{g_3 h_1 - g_1 h_3}{d} \right) \quad \text{mit} \quad d = g_1 h_2 - h_1 g_2$$

- Ein Punkt  $X$  auf der Geraden  $g = AB$  hat bei bekanntem Verhältnis  $u = \frac{|AX|}{|AB|}$  die Koordinaten

$$(x_1, x_2) = ((1 - u) a_1 + u b_1, (1 - u) a_2 + u b_2)$$

Dieses Verhältnis  $u$  bestimmt die Lage von  $X$  auf  $g$  relativ zu  $A, B$  eindeutig. Wir bezeichnen  $u = GP(X; A, B)$  als *Gleiterparameter*. Liegt  $X$  im Inneren der Strecke  $\overline{AB}$ , so gilt  $0 < u < 1$ , für Punkte  $X$  jenseits von  $B$  gilt  $u > 1$  und für Punkte jenseits von  $A$  schließlich  $u < 0$ . Zum früher eingeführten Teilverhältnis  $TV(A, B; X) = \frac{|AX|}{|XB|}$  besteht der Zusammenhang  $TV(A, B; X) = \frac{u}{1-u}$ .

Auch Begriffe aus der Euklidischen Geometrie lassen sich symbolisch durch entsprechende Koordinaten ausdrücken:

- So ergibt sich der Abstand zwischen den Punkten  $A, B$  aus der Formel

$$d(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

Da es sich dabei nicht um einen arithmetischen Ausdruck handelt, wollen wir statt dessen mit dem *Abstandsquadrat*  $sqrdist(A, B) = d(A, B)^2$  arbeiten.

- Zwei Geraden  $g, h$  sind orthogonal genau dann, wenn ihre Normalenvektoren  $(g_1, g_2)$  und  $(h_1, h_2)$  senkrecht aufeinander stehen, d.h. für das entsprechende Skalarprodukt

$$g_1 h_1 + g_2 h_2 = 0$$

gilt.

- Schließlich lässt sich das Lot  $h$  von  $P$  auf die Gerade  $g$  ausdrücken als

$$(h_1, h_2, h_3) = (g_2, -g_1, g_1 p_2 - g_2 p_1).$$

## 2.2 Homogene Punktkoordinaten

Bei der Betrachtung der Konkurrenz dreier Geraden können wir statt nach Lösungen  $(x_1, x_2)$  des inhomogenen Gleichungssystems

$$g_1 x_1 + g_2 x_2 + g_3 = 0$$

$$h_1 x_1 + h_2 x_2 + h_3 = 0$$

$$k_1 x_1 + k_2 x_2 + k_3 = 0$$

auch nach Lösungen  $(x_1, x_2, x_3)$  des homogenen Gleichungssystems

$$\begin{aligned} g_1x_1 + g_2x_2 + g_3x_3 &= 0 \\ h_1x_1 + h_2x_2 + h_3x_3 &= 0 \\ k_1x_1 + k_2x_2 + k_3x_3 &= 0 \end{aligned} \quad (*)$$

mit  $x_3 = 1$  fragen. Da Lösungen homogener Gleichungssysteme durch einen skalaren Faktor variiert werden können, reicht die Existenz von Lösungen mit  $x_3 \neq 0$  aus.

Solche Koordinaten  $X = (x_1, x_2, x_3)$  bezeichnet man als *homogene* oder *projektive Punktkoordinaten*. Sie sind — wie die homogenen Geradenkoordinaten — nur bis auf einen skalaren Faktor verschieden Null eindeutig bestimmt, wobei den affinen Koordinaten  $(x_1, x_2)$  die projektiven Koordinaten  $(x_1, x_2, 1)$  entsprechen. Letztere bezeichnen wir auch als *normierte Koordinaten* und schreiben  $\bar{X}$ .  $X$  liegt auf der Geraden  $g$  genau dann, wenn

$$g_1x_1 + g_2x_2 + g_3x_3 = 0$$

gilt. An dieser Formel sieht man schon, dass Punkt- und Geradenkoordinaten in zueinander dualer Weise eingehen, was die früher beschriebene Dualität von Punkten und Geraden in Sätzen der projektiven Geometrie plausibel macht. Die Punkte, für deren homogene Koordinate  $x_3 = 0$  gilt, liegen auf der Ferngeraden, denn deren homogene Koordinaten lauteten ja gerade  $(0 : 0 : 1)$ .

Wir bezeichnen diese Erweiterung der affinen Ebene  $\mathbb{A}_2$  um die Punkte der Ferngeraden als projektive Ebene  $\mathbb{P}_2$ . Bezeichnen wir dual dazu mit  $\mathbb{P}_2^*$  den Raum der projektiven Geraden, so lassen sich die weiter oben untersuchten geometrischen Beziehungen nennerfrei durch Skalar-, Vektor- und Spatproduktoperationen im  $\mathbb{R}^3$  beschreiben.

- $A, B, C$  in homogenen Punktkoordinaten sind kollinear  $\Leftrightarrow \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = 0$

Notation:  $|A \ B \ C| = 0$  (Spatprodukt)

- Analog sind drei Geraden  $g, h, k$  konkurrent  $\Leftrightarrow |g \ h \ k| = 0$
- Punkt  $P$  und Gerade  $g$  sind inzident  $\Leftrightarrow p_1g_1 + p_2g_2 + p_3g_3 = 0$

Notation:  $P * g = 0$  (Skalarprodukt)

- Für den Schnittpunkt  $P$  zweier Geraden  $g, h$  können wir die frühere Formel nennerfrei interpretieren:

$$\begin{aligned} P &= (g_2h_3 - g_3h_2, g_3h_1 - g_1h_3, g_1h_2 - g_2h_1) = \left( \begin{vmatrix} g_2 & g_3 \\ h_2 & h_3 \end{vmatrix}, \begin{vmatrix} g_3 & g_1 \\ h_3 & h_1 \end{vmatrix}, \begin{vmatrix} g_1 & g_2 \\ h_1 & h_2 \end{vmatrix} \right) \\ &= (g_1, g_2, g_3) \times (h_1, h_2, h_3) = g \times h \end{aligned}$$

Das sind genau die Koordinaten des Vektorprodukts zweier Vektoren im  $\mathbb{R}^3$ .

- Die Gleichung einer Geraden durch zwei in homogenen Koordinaten gegebene (verschiedene) Punkte  $A, B$  lautet analog

$$g = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) = A \times B$$

- $A, B, C$  sind kollinear  $\Leftrightarrow A$  ist inzident zur Geraden durch  $B$  und  $C \Leftrightarrow A * (B \times C) = 0$ . dies stimmt wegen des bekannten Zusammenhangs

$$|A \ B \ C| = A * (B \times C)$$

zwischen Spat-, Vektor- und Skalarprodukt im  $\mathbb{R}^3$  mit obiger Determinantenformel überein.

Homogene Punkt- bzw. Geradenkoordinaten sind genau dann *nicht zulässig*, wenn sich  $(0 : 0 : 0)$  ergibt. Aus der Formel für die Koordinaten des Schnittpunkts zweier Geraden  $g, h$  ist ersichtlich, dass sich nicht zulässige Koordinaten genau dann ergeben, wenn die Koordinaten von  $g$  und  $h$  proportional sind, d.h. wenn  $g$  und  $h$  identisch sind.

Analog ergeben sich nicht zulässige Geradenkoordinaten für die Verbindungsgerade zweier Punkte  $A$  und  $B$  genau dann, wenn  $A = B$  gilt.

Auch Parallelität und Teilverhältnisse kann man ausdrücken, wenn berücksichtigt wird, dass diese Größen nicht projektiv invariant sind, d.h. bei ihrer Definition die Ferngerade  $l_0 = (0 : 0 : 1)$  eine Rolle spielen muss:

- Zwei Geraden  $g, h$  sind parallel  $\Leftrightarrow$  sie schneiden sich auf der Ferngeraden, d.h.  $|g \ h \ l_0| = 0$ . Das stimmt mit unserer weiter oben hergeleiteten Formel überein.
- Die Gerade  $h$  durch  $P$ , die parallel zu  $g$  verläuft, ergibt sich als Verbindung von  $g \times l_0 = (-g_2 : g_1 : 0)$  und  $P: h = P \times (g \times l_0)$ .
- Alle Senkrechten zur Geraden  $g$  gehen durch den gemeinsamen Fernpunkt  $O_g = (g_1 : g_2 : 0)$ , so dass sich die Senkrechte  $h$  zu  $g$  durch  $P$  als  $h = P \times O_g = (-p_3 g_2 : p_3 g_1 : p_1 g_2 - p_2 g_1)$  in Übereinstimmung mit der früher gefundenen Darstellung ergibt.

Mit Parallelen kann man aus einem Standardframe ein ganzes affines Koordinatensystem gewinnen. Als *Standardframe* bezeichnet man die Punkte  $e_1 = (1 : 0 : 0), e_2 = (0 : 1 : 0), e_3 = (0 : 0 : 1), u = (1 : 1 : 1)$ . Geometrisch sind  $e_3$  der Ursprung des (affinen) Koordinatensystems,  $e_1, e_2$  die Fernpunkte in  $x$ - bzw.  $y$ -Richtung (diese bestimmen zusammen mit  $e_3$  die  $x$ - und  $y$ -Richtung) und  $u$  der Einheitspunkt mit den (affinen) Koordinaten  $(1, 1)$ . Parallelen zur  $x$ - bzw.  $y$ -Achse durch  $u$  schneiden die  $y$ - bzw.  $x$ -Achse in den Koordinaten-Einheiten und auch andere Teilpunkte auf den Achsen mit ganzzahligen Koordinaten lassen sich leicht konstruieren.

**Teilverhältnisse:** Die Mittelpunktsformel

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

gilt für homogene Koordinaten nicht mehr, da sie selbst nicht homogen ist. Die richtige Modifikation lautet

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \frac{b_3}{2} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \frac{a_3}{2} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

und allgemeiner für einen Gleiter  $M$  auf  $AB$  mit Parameter  $u$

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = u b_3 \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + (1 - u) a_3 \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Die Ferngerade bildet insgesamt eine Art „schwarzes Loch“, denn liegt einer der beiden Randpunkte dort, so auch  $M$  (es gilt stets  $m_3 = a_3 b_3$ ). Wenn beide Randpunkte  $A$  und  $B$  auf der Ferngeraden liegen, so ergeben sich für  $M$  unzulässige Koordinaten.

Die beiden Koeffizienten in der Darstellung  $M = u b_3 A + (1 - u) a_3 B$  ergeben in der Summe nicht mehr unbedingt 1, so dass sich ein Gleiter  $M$  auf  $AB$  auch als

$$M = \mu_A A + \mu_B B$$

darstellen lässt. In der Tat, ist  $g = (g_1, g_2, g_3)$  die Gerade durch  $A$  und  $B$ , also  $g * A = g * B = 0$ , so gilt  $g * M = 0$  auch für jede solche Linearkombination  $M$ . Andererseits ergeben Paare  $(\mu_A, \mu_B)$ , welche sich nur durch einen skalaren Faktor unterscheiden, verschiedene homogene Koordinaten

für denselben Punkt  $M$ , so dass die Punkte auf  $AB$  durch das Verhältnis  $\mu_A : \mu_B \in \mathbb{P}^1$  eindeutig charakterisiert werden (das Verhältnis  $\infty$  ist möglich und entspricht dem Punkt  $A$ ). Dieses Verhältnis ist für  $a_3 = b_3 = 1$  gerade das früher eingeführte Teilverhältnis  $TV(A, B; M)$ , wie sich durch Übergang zu normierten Koordinaten auch für  $M$  leicht nachrechnen lässt.

Allgemein gilt

$$TV(A, B; M) = \frac{\mu_A/\mu_B}{a_3/b_3},$$

denn aus  $M = \mu_A A + \mu_B B$  folgt für normierte Koordinaten  $m_3 \overline{M} = \mu_A a_3 \overline{A} + \mu_B b_3 \overline{B}$  mit  $m_3 = \mu_A a_3 + \mu_B b_3$ . Das Teilverhältnis ist also keine projektive Invariante, sondern hängt von der speziellen Wahl der homogenen Koordinaten von  $A$  und  $B$  ab. Invariant wird erst das *Doppelverhältnis*

$$DV(A, B; M, N) = TV(A, B; M) : TV(A, B; N),$$

wobei  $N = \nu_A A + \nu_B B$  ein weiterer Punkt auf der Geraden  $AB$  ist:

$$DV(A, B; M, N) = \frac{\mu_A/\mu_B}{\nu_A/\nu_B}$$

Durch das Verhältnis  $\mu_A/\mu_B$  werden also die Punkte  $M \in AB$  eindeutig parametrisiert, aber ihre genaue Lage nicht absolut beschrieben, sondern nur relativ zueinander.

Hält man allerdings einen weiteren Punkt auf der Geraden fest, so beschreibt das Teilverhältnis die Lage auch absolut. Das Teilverhältnis kann etwa als Doppelverhältnis mit  $N = F$ , dem Fernpunkt auf der Geraden  $AB$ , geschrieben werden: Wegen  $F = -b_3 A + a_3 B$  ergibt sich

$$TV(A, B; M) = -DV(A, B; M, F).$$

Hieraus folgt, dass projektive Transformationen, welche die Ferngerade fest lassen (also genau die affinen Transformationen), auch teilverhältnisinvariant sind.

### 2.3 Zur Algorithmisierung geometrischer Konstruktionen. Analytische Geometrie mit dem Computer

Wir können auf der Basis der im Abschnitt 2.1 hergeleiteten Beziehungen nun in einer klassischen imperativen Programmiersprache (die an dieser Stelle noch nicht über die Fähigkeit zur Symbolverarbeitung verfügen muss) Funktionen schreiben, die in der Lage sind, Beziehungen in durch konkrete Koordinatenwerte vorgegebenen geometrischen Konfigurationen zu überprüfen oder gesuchte Größen auszurechnen. Entsprechende Funktionen sind auch die Grundlage für dynamische Geometriesysteme, mit denen entsprechende Konfigurationen grafisch dargestellt werden können.

Die elementaren geometrischen Objekte Punkt und Gerade setzen wir dazu als Klassen `Point` und `Line` um, die in Java etwa als

```
public class Point {
    public Scalar x,y;

    public Point () { }
    public Point(Scalar x, Scalar y) { this.x=x; this.y=y; }
    ...
};

public class Line {
    public Scalar a,b,c;

    public Line () {}
```

```

    public Line(Scalar a, Scalar b, Scalar c) {
        if (iszero(a) & iszero(b) & iszero(c))
            throw new GeoException("Gerade mit Nullkoordinaten");
        this.a=a; this.b=b; this.c=c;
    }
    ...
};

```

definiert werden können und Punkte  $P(x_1, x_2)$  bzw. Geraden  $g = \{(x_1, x_2) : g_1x_1 + g_2x_2 + g_3 = 0\}$  darstellen. `Scalar` ist dabei der Grundbereich, aus dem die Werte der Koordinaten kommen, also in Java etwa der interne Typ `double` oder etwas Selbstgestricktes wie `Rational` oder `Complex`. Wir wollen annehmen, dass es sich um einen Körper handelt, also für diesen Datentyp die arithmetischen Operationen  $+$   $-$   $*$   $/$  sowie ein boolesches Prädikat `boolean iszero()` definiert sind.

Geometrische Grundkonstruktionen können wir in diesem Kontext als Funktionen auffassen, die aus gegebenen Objekten neue konstruieren, und in einer weiteren Klasse `GeoFunctions` als Klassenfunktionen bündeln.

1) Die Gerade durch zwei Punkte  $P$  und  $Q$

```

public static Line pp_line(Point p, Point q) {
    return new Line(q.y-p.y, p.x-q.x, p.y*q.x-p.x*q.y);
}

```

$P$  und  $Q$  sind dabei als formale Parameter vom Typ `Point` Container für die aktuellen Koordinaten, der Rückgabewert der Funktion vom Typ `Line` der Container für die berechneten Koordinaten des davon abhängenden Objekts.

2) Analog können wir den Schnittpunkt zweier Geraden berechnen, wobei die zu definierende Funktion mit einer Ausnahme abbricht, wenn kein bzw. kein eindeutig bestimmter Schnittpunkt existiert.

```

public static Point intersection_point(Line g, Line h) {
    Scalar d = g.a*h.b-g.b*h.a;
    if (iszero(d)) throw new GeoException("Geraden sind parallel");
    return new Point((g.b*h.c - g.c*h.b)/d, (g.c*h.a - g.a*h.c)/d);
}

```

Auch hier sind  $g$  und  $h$  formale Parameter, diesmal vom Typ `Line`.

3) Für das Lot  $l$  von einem Punkt  $P$  auf eine Gerade  $g$  erhalten wir analog

```

public static Line ortho_line(Point p, Line g) {
    return new Line(g.b, -g.a, g.a*p.y - g.b*p.x);
}

```

und für die Parallele zu einer Geraden  $g$  durch einen Punkt  $P$

```

public static Line par_line(Point p, Line g) {
    return new Line(g.a, g.b, -(g.a*p.x + g.b*p.y));
}

```

Das *Abstandsquadrat* ergibt sich schließlich als

```

public static Scalar sqrdist(Point p, Point q) {
    return new Scalar((p.x-q.x)*(p.x-q.x) + (p.y-q.y)*(p.y-q.y));
}

```

4) Neben freien Punkten, die mit dem Punktkonstruktor erzeugt werden können, sind auch Punkte auf vorgegebenen Geraden (*Geradengleiter*) oder Kreisen (*Kreisgleiter*) interessant. Einen Geradengleiter auf einer durch zwei Punkte gegebenen Geraden kann man etwa durch ein variables Teilverhältnis festlegen:

```
public static Point varpoint(Point P, Point Q, Scalar u) {
    return new Point((1-u)*p.x+u*q.x,(1-u)*p.y+u*q.y);
}
```

Insbesondere liefert

```
Point midpoint(Point P,Point Q) { return varpoint(P,Q,new Scalar(1/2)); }
```

den Mittelpunkt der Strecke  $PQ$ .

5) Komplexere geometrische Konstruktionen (Makros) können aus nacheinander ausgeführten Grundkonstruktionen zusammengesetzt werden. Dem entsprechen auf der Seite der Programmiersprachen zusammengesetzte Funktionen. So findet man etwa den Fußpunkt des Lots vom Punkt  $P$  auf die Gerade  $a$  als

```
public static Point pedalpoint(Point P, Line a) {
    return intersection_point(ortho_line(P,a),a);
}
```

**Aufgabe 7** *Geben Sie entsprechende Funktionen an*

- `Line p_bisector(Point A, Point B)`  
*für die Mittelsenkrechte (perpendicular bisector) auf der Seite AB,*
- `Line altitude(Point A, Point B, Point C)`  
*für die Höhe durch A im Dreieck ABC,*
- `Line median(Point A, Point B, Point C)`  
*für die Seitenhalbierende, die durch A im Dreieck ABC verläuft.*

6) Schließlich kann man testen, ob für gewisse Konfigurationen geometrische Bedingungen erfüllt sind. So kann man etwa testen, ob zwei gegebene Geraden  $g$  und  $h$  parallel oder orthogonal sind, indem man prüft, ob  $g_1h_2 - g_2h_1$  bzw.  $g_1h_1 + g_2h_2$  verschwindet, oder ob ein Punkt  $P$  auf einer Geraden  $g$  liegt. Entsprechende Funktionen haben folgende Spezifikation:

```
public static boolean is_parallel(Line g, Line h) {
    return iszero(g.a*h.b-h.a*g.b);
}
```

bzw.

```
public static boolean is_orthogonal(Line g, Line h) {
    return iszero(g.a*h.a+g.b*h.b);
}
```

```
public static boolean is_point_on_line(Point P, Line g) {
    return iszero(g.a*P.x+g.b*P.y+g.c);
}
```

Auch kompliziertere Bedingungen, die wir im letzten Paragraph hergeleitet hatten, kann man auf diese Weise prüfen, so z.B., ob drei gegebene Punkte  $P, Q, R$  *kollinear* oder drei gegebene Geraden  $a, b, c$  *konkurrent* sind. Die entsprechenden Funktionen `is_collinear` und `is_concurrent` lassen sich leicht angeben.

**Aufgabe 8** Führen Sie die Implementierung entsprechender Funktionen wie oben begonnen in Java zu Ende.

Mit diesem Arsenal kann man die Koordinaten auch komplizierterer geometrischer Objekte bestimmen und entsprechende geometrische Sätze in konkreten Konfigurationen überprüfen.

Beispiel 1: Der Satz vom Schnittpunkt der Mittelsenkrechten. Die Funktion

```
static boolean CircumCenter_Test1(Point A, Point B, Point C) {
    return is_concurrent(p_bisector(A,B), p_bisector(B,C), p_bisector(C,A));
}
```

prüft, ob für ein Dreieck, das durch seine Eckpunkt(koordinaten)  $A, B, C$  gegeben ist, die drei Mittelsenkrechten durch einen gemeinsamen Punkt gehen. Alternativ kann man wie im elementargeometrischen Beweis dieses Satzes auch zuerst die Koordinaten des Schnittpunkt zweier der Mittelsenkrechten bestimmen und dann zeigen, dass dieser Punkt auf der dritten Mittelsenkrechten liegt:

```
static boolean CircumCenter_Test2(Point A, Point B, Point C) {
    Point M = intersection_point(p_bisector(A,B), p_bisector(B,C));
    return on_line(M, p_bisector(C,A));
}
```

Letztere Funktion lässt sich zu eine Test erweitern, der zeigt, dass dieser Schnittpunkt der Mittelpunkt des Umkreises des Dreiecks  $\triangle ABC$  ist. Dazu ist

$$sqrdist(M, A) = sqrdist(M, B) = sqrdist(M, C)$$

zu testen, d.h. die letzte Zeile durch

```
return iszero(sqrdist(M,A) - sqrdist(M,B)) &
        iszero(sqrdist(M,B) - sqrdist(M,C));
```

zu ersetzen.

**Aufgabe 9** Geben Sie analoge Testfunktionen für die Sätze vom Schnittpunkt der Höhen bzw. der Seitenhalbierenden. Lassen Sie auch testen, dass die Produkte aus den Höhenabschnitten im Dreieck gleich sind bzw. der Schnittpunkt der Seitenhalbierenden diese im Verhältnis 2:1 teilt.

Beispiel 2: Der Satz von der Eulerschen Geraden:

```
static boolean EulerLine_Test(Point A, Point B, Point C) {
    Point M = intersection_point(p_bisector(A,B), p_bisector(B,C));
    Point H = intersection_point(altitude(A,B,C), altitude(B,C,A));
    Point S = intersection_point(median(A,B,C), median(B,C,A));
    return is_collinear(M,H,S);
}
```

**Aufgabe 10** Modifizieren Sie diese Testfunktion so, dass sie den Satz vom Feuerbachschen Kreis testet. (Feuerbach.Test(Point A, Point B, Point C))

Testen Sie den (affinen) Satz von Desargue. (Desargue\_Test(Point A, Point B, Point C))

## 2.4 Zum grundsätzlichen Aufbau einer dynamischen Geometrie-Software (DGS)

Wir wollen uns an dieser Stelle auf die Betrachtung von Punkten und Geraden in der Ebene  $\mathcal{E}$  beschränken und die grundlegenden informatischen Begriffe entwickeln, die für das Verständnis einer DGS erforderlich sind, sowie deren Verhältnis zu anderen Begriffen und Konzepten der Informatik insgesamt herausarbeiten.

Im letzten Abschnitt hatten wir bereits gesehen, dass sich das Attribute und Methoden bündelnde Klassen- und Instanzenkonzept des objektorientierten Programmierens gut für DGS eignet. Es erlaubt die Kapselung von durch Koordinaten gegebener geometrischer Gebilde in neuen Sinneinheiten.

Die in der Informatik übliche Unterscheidung von abstrakter Identität eines Objekts und dessen sich im Laufe der Zeit ändernden Objektzustands spielt für DGS eine wichtige Rolle. Wie bei Variablen haben wir dabei zu unterscheiden zwischen dem Objekt als Container des Zustands (dieser wird in den **Attributen** des Objekt dargestellt) und dem sich über die Zeit ändernden Zustand selbst (also den **Attributwerten**). Eine Zustandsänderung als Änderung der Attributwerte des Objekt wird in einer DGS praktisch etwa durch die Bewegung von freien Punkten oder Gleitern bewirkt. Ist  $g.c$  also ein Attribut eines Objekts  $g$  mit Werten in einem Bereich  $C$ , so wird in der Attributdeklaration

```
public CType c;
```

von  $g$  Speicherplatz für das Attribut  $g.c$  reserviert, der dann konkrete Werte  $g.c \in C$  aufnehmen kann. Dieser Wert kann sich über die Zeit ändern, was als  $g.c(t) \in C$  für einen Zeitparameter  $t \in T$  oder gleich als **Attributwertfunktion**  $g.c : T \rightarrow C$  dargestellt werden kann. Diese Mehrdeutigkeit der Notation ist aus dem jeweiligen Kontext her aufzulösen.

Bewegen geometrischer Objekte kann die Bewegung anderer geometrischer Gebilde zur Folge haben kann; Zustandsänderungen propagieren also durch ein Netz von Abhängigkeiten. Diese Abhängigkeiten werden aber durch dieselben Formeln für zu verschiedenen Zeitpunkten vorliegende Zustände derselben Objekte beschrieben, so dass es sich um Abhängigkeiten zwischen den Objekten selbst handelt. Wir führen deshalb die folgende begriffliche Unterscheidung ein:

**Definition 1** Die Klassen **Punkt** und **Gerade** bezeichnen wir als **geometrische Typen**, Instanzen einer solchen Klassen als **geometrische Objekte**. Jeden Objektzustand eines solchen geometrischen Objekts bezeichnen wir als **spezielles geometrisches Objekt**.

Der Punkt  $P$  ist als geometrisches Objekt also eine Instanz der Klasse **Punkt**, der im Laufe seines Lebenszyklus für verschiedene spezielle Punkte  $P(t)$  steht, je nachdem, wo er sich gerade „befindet“. Er wird durch Attribute charakterisiert, deren Werte den jeweiligen Objektzustand beschreiben, der sich im Laufe der Zeit ändern kann. Diese Attribute sind einerseits Bewegungsattribute wie die Koordinaten von  $P$ , also  $(x_P, y_P)$  oder homogen  $(x_P : y_P : z_P)$ , andererseits Zustandsattribute wie die Farbe, Bezeichnung etc. Wir wollen uns im Weiteren nur für die Bewegungsattribute interessieren, da allein diese das geometrische Verhalten von  $P$  oder einer Geraden  $g$  beschreiben. Wie oben haben wir zu unterscheiden zwischen Attribut  $g.c$ , konkreten Attributwerten  $g.c \in C_L$  oder  $g.c(t) \in C_L$  für  $t \in T$  und der Attributwertfunktion  $g.c : T \rightarrow C_L$ , welche die gesamte zeitlichen Existenz von  $g$  beschreibt.

$C_L$  ist dabei der Bereich der (für Geraden zulässigen) Koordinatenwerte. Er hängt nicht von  $g$  selbst, sondern nur vom geometrischen Typ  $g \in \mathbf{Gerade}$  und natürlich vom gewählten **Koordinatenmodell** ab. Das Modell der homogenen Geradenkoordinaten lässt sich als Menge von Äquivalenzklassen

$$C_L = (K^3 \setminus \{(0, 0, 0)\}) / \sim$$

bzgl. der Relation

$$(x, y, z) \sim (x', y', z') \Leftrightarrow \exists c \in K^* : x' = cx, y' = cy, z' = cz$$

oder kurz als Menge der nichttrivialen Orbits der Aktion der multiplikativen Gruppe  $K^*$  auf  $K^3$  beschreiben. Ein solches Koordinatenmodell basiert stets auf einem **Grundbereich**  $K$ , aus welchem die Werte der Koordinaten kommen. Wir wollen stets voraussetzen, dass  $K$  ein algebraisch abgeschlossener Körper mit  $\text{char}(K) = 0$  ist.

Neben dem Bereich  $C_L$  der Geradenkoordinaten haben wir noch den Bereich  $C_P$  der Punktkoordinaten. Ist allgemein  $X$  ein Typ, so bezeichnen wir mit  $C(X)$  diesen *Wertebereich* des Typs  $X$ .

**Definition 2** Wir bezeichnen das (informatische) Attribut  $g.c$  vom Typ  $C_L$  als Koordinatenattribut und jeden Wert  $g.c(t) \in C_L$  als spezielle Koordinaten des speziellen geometrischen Objekts  $g(t)$ .

DGS sind als grafische Anwendungen sinnvoller Weise nach dem Model-View-Controller-Konzept aufgebaut, wobei der Controller mit dem View vereinigt sein kann. Mausaktionen werden vom View über den Controller an das Model weitergegeben, dort die entsprechenden Berechnungen aktualisiert, und schließlich über den Controller (oder auch direkt über Event-Steuerung) an den View der Befehl zu einem (re)paint gegeben. In diesem Kontext ist  $g$  als geometrisches Objekt auf der Model-Seite, die speziellen geometrischen Objekte  $g(t)$  auf der View-Seite zu finden.

Geometrische Objekte werden durch entsprechende Konstruktionswerkzeuge Schritt für Schritt erzeugt, welche auf der Seite der Informatik als Funktionen daherkommen. Für Funktionen haben wir zwischen Funktionsdefinition und Funktionsaufruf zu unterscheiden, wobei in einer Funktionsdefinition formale Parameter als Platzhalter für geometrische Objekte als Aufrufargumente auftreten. Die Dynamik eines so neu konstruierten geometrischen Objekts ergibt sich aus der Dynamik der geometrischen Objekte, welche an der Konstituierung beteiligt waren, und der Spezifik des Konstruktionswerkzeugs selbst.

Diese Werkzeuge sind prototypisch von zwei verschiedenen Arten.

**Prototyp 1:** `Point F = pedalpoint(Point P, Line a)`

Mit diesem Aufruf des Werkzeugs `pedalpoint` wird aus zwei vorhandenen geometrischen Objekten  $P \in \text{Point}$  und  $a \in \text{Line}$  ein neues geometrisches Objekt  $F \in \text{Point}$  erzeugt, dessen Dynamik  $F.c : T \rightarrow C_P$  sich aus den Dynamiken  $P.c : T \rightarrow C_P$  und  $a.c : T \rightarrow C_L$  als Zusammensetzung

$$F.c = \phi \circ (P.c \times a.c) : T \xrightarrow{P.c \times a.c} C_P \times C_L \xrightarrow{\phi} C_P$$

ergibt, wobei  $\phi : C_P \times C_L \rightarrow C_P$  beschreibt, wie aus den Koordinaten von  $P$  und  $a$  diejenigen von  $F$  zu berechnen sind.

**Prototyp 2:** `Point M = varpoint(Point A, Point B, X u)`

Hier hängt  $M$  noch zusätzlich von einem Parameter  $u$  ab, dessen Natur  $X$  wir nun näher spezifizieren wollen. Ein Vergleich mit Prototyp 1 zeigt, dass  $u$  hier in derselben Doppelbedeutung wie geometrische Objekte auch auftritt: Einerseits als abstrakte Identität eines Stellparameters, welcher selbst eine zeitliche Dynamik hat, und andererseits als Eingangsparameter, welcher die Dynamik von  $M$  beeinflusst.  $u$  bestimmt diese Dynamik aber nur relativ zu  $A$  und  $B$ , denn ein Bewegen eines dieser Punkte ohne Änderung des Stellparameters  $u$  ändert die Lage des speziellen geometrischen Objekts  $M$  ebenfalls.

$X$  steht also für einen weiteren Datentyp eines Stellparameters **SP**, dessen Wertebereich im Koordinatenmodell der homogenen Punktkoordinaten  $\mathbb{P}^1$  liegt. In Analogie zu den geometrischen Objekten bezeichnen wir einen solchen Typ als **Parametertyp** und Instanzen  $u$  dieses Typs als **Parameterobjekte**. Ein solches Parameterobjekt hat einerseits eine abstrakte Identität und andererseits einen zeitlichen Verlauf  $u.c : T \rightarrow C_S$  mit Werten aus dem (skalaren) Parameterbereich

$C_S$  ( $S$  wie Stellparameter). Diese Funktion bezeichnen wieder als *Parameterfunktion*, einen konkreten Wert  $u.c(t) \in C_S$  als *speziellen Parameter*. Wir können uns die Entkopplung von der visuellen Darstellung als „Bewegen mit der Maus“ etwa als Schieberegler vorstellen, dessen Schieben den Punkt  $M$  auf der Geraden  $AB$  bewegt.

Eine solche Trennung ist auch bei freien Punkten sinnvoll, da für diese ebenfalls eine mittelbare Steuerung der Bewegung denkbar wäre<sup>1</sup>. Allerdings haben freie Punkte zwei Freiheitsgrade, so dass hierfür noch ein zweiter Parametertyp  $MP$  mit einem Parameterbereich  $C_M$  ( $M$  wie Mausparameter) benötigt wird.

Mit diesen zusätzlichen Definitionen können wir nun eine einheitliche Definition eines Konstruktionswerkzeugs geben, welche davon ausgeht, dass genügend Parameterobjekte zur Verfügung stehen, aber selbst das Anlegen eines einzigen freien Punktes durch ein Konstruktionswerkzeug geschieht (und so ist es ja praktisch auch). Die gesamte Dynamik der Konstruktion ist in den Parameterobjekten gekapselt. Allerdings werden auch „konstante“ Parameterobjekte benötigt, wenn wir z. B. den Mittelpunkt einer Strecke als `varpoint(A,B,1/2)` konstruieren wollen. Diese Implementierungseinheiten von Parameterobjekten sollen hier aber nicht diskutiert werden.

**Definition 3** Seien  $T_1, \dots, T_n$  geometrische oder Parametertypen,  $T_a$  ein geometrischer Typ und  $C_1, \dots, C_n, C_a$  die zugehörigen Wertebereiche. Als **Konstruktionswerkzeug**  $w$  bezeichnen wir eine (informatische) Funktion der Signatur

$$w : (T_1 \times \dots \times T_n) \rightarrow T_a$$

zusammen mit einer mathematischen Funktion

$$w.c : (C_1 \times \dots \times C_n) \rightarrow C_a,$$

so dass für spezielle Objekte  $o_i \in T_i$  der richtigen Typen und  $o_a = w(o_1, \dots, o_n) \in T_a$

$$o_a.c = w.c(o_1.c, \dots, o_n.c)$$

gilt.

Eine weitere Feinheit haben wir dabei noch nicht berücksichtigt: Ist z.B.

`Line g = pp_line(Point A, Point B)`

das Konstruktionswerkzeug, welches zu zwei gegebenen Punkten die Gerade durch diese Punkte konstruiert, so ist für spezielle Punkte  $A$  und  $B$  diese Gerade nur definiert, wenn  $A.c \neq B.c$  gilt.

$$\text{pp\_line.c} : C_P \times C_P \longrightarrow C_L$$

ist also nur eine partiell definierte Funktion. Die Ausnahmemenge wird durch ein boolesches Prädikat

$$\text{pp\_line.NDG} : C_P \times C_P \longrightarrow \text{Boolean}$$

bestimmt, welches in unserem Fall die einfache Form  $\text{pp\_line.NDG}(x, y) = (x \neq y)$  hat. Hierbei ist `Boolean` der Wertebereich des Datentyps `boolean`.

**Definition 4** Zu jedem Konstruktionswerkzeug  $w$  gibt es weiter eine **Nichtdegenerationsbedingung**

$$w.NDG : (C_1 \times \dots \times C_n) \rightarrow \text{Boolean},$$

so dass  $w$  genau auf denjenigen speziellen Tupeln  $(o_1, \dots, o_n)$  definiert ist, für welche

$$w.NDG(o_1.c, \dots, o_n.c) = \text{true}$$

gilt.

---

<sup>1</sup>Das ist auch praktisch so: Die Koordinaten des Mauszeigers werden auf der View-Seite in Fensterkoordinaten abgegriffen und auf der Modell-Seite in Weltkoordinaten umgerechnet.

**Beispiele:**

`freePoint: Param2 → Point`

legt einen freien Punkt an.

`pp_line: Point×Point → Line`

erzeugt die Gerade durch zwei gegebene Punkte. Die zugehörige NDG ist `pp_line.NDG(A, B) = (A ≠ B)` für  $A, B \in C_P$ .

`intersection_point: Line×Line → Point`

erzeugt den Schnittpunkt zweier Geraden.

Die zugehörige NDG ist `intersection_point.NDG(a, b) = (a ≠ b)` für  $a, b \in C_L$ .

Konstruktionswerkzeuge werden eingesetzt, um mit ihnen eine geometrische Konfiguration schrittweise aufzubauen, wobei neu erzeugte geometrische Objekte von bereits vorhandenen sowie von Parameterobjekten abhängen. Diese Abhängigkeitsverhältnisse lassen sich durch einen (endlichen) gerichteten azyklischen Graphen (DAG) darstellen und intern durch Ereignispropagation modellieren. Auch auf dieser Ebene wollen wir zwischen Definition und Aufruf unterscheiden, da die Abhängigkeitsverhältnisse zwischen Ein- und Ausgabegrößen der Konstruktionswerkzeuge auf der Ebene der geometrischen Typen und nicht der geometrischen Objekte liegen.

**Definition 5** Als **Konfiguration** bezeichnen wir deshalb einen Abhängigkeitsgraphen  $\Gamma = (T, E)$  zwischen Typen. Ein Tupel  $O = (o_\tau \in \tau \mid \tau \in T)$  dazu passender spezieller Objekte der korrekten Typen bezeichnen wir als **Realisierung der Konfiguration**  $\Gamma$ .

Wir bezeichnen weiter  $\Gamma = (T, E)$  als **Parameterkonfiguration**, wenn  $T$  ausschließlich Parametertypen enthält.

Damit können wir nun den Begriff der Konstruktion schrittweise herleiten.

**Definition 6** Sei  $\Gamma = (T, E)$  ein Konfiguration und  $w : (T_1 \times \dots \times T_n) \rightarrow T_a$  ein Konstruktionswerkzeug.

$w$  ist auf  $\Gamma$  **anwendbar**, wenn es eine Belegung  $f : [1 \dots n] \rightarrow T$  mit  $f(i) = T_i$  gibt.

Als **Konstruktionsschritt** bezeichnen wir in diesem Fall die Sequenz folgender Aktionen:

1. Auswahl einer solchen Belegung  $f$ ,
2. Ergänzung von  $\Gamma$  zu  $\Gamma' = (T \cup \{T_a\}, E \cup \{(T_i, T_a), i = 1, \dots, n\})$ .

Ist  $O = (o_\tau \in \tau \mid \tau \in T)$  eine Realisierung von  $\Gamma$  und  $o_i = o_{f(i)}$ , so sagen wir, dass  $w$  auf  $(o_1, \dots, o_n)$  **ausführbar** ist, wenn  $w.NDG(o_1.c, \dots, o_n.c) = \text{true}$  gilt, und bezeichnen in diesem Fall die Sequenz folgender Aktionen als **Ausführung des Konstruktionsschritts**:

1. Anwendung des Konstruktionswerkzeugs zur Erzeugung von  $o = w(o_1, \dots, o_n) \in T_a$ ,
2. Ergänzung von  $O$  zu  $O' = O \cup \{o\}$ .

In all diesen Definitionen ist es möglich, den Begriff des Konstruktionswerkzeugs so zu fassen, dass statt eines einzelnen Objekts  $o_a$  gleich ein ganzes Tupel  $(o_{a_1}, \dots, o_{a_m})$  von Objekten konstruiert wird, deren Typen  $(T_{a_1}, \dots, T_{a_m})$  vorgegeben sind. Ein solches Konstruktionswerkzeug bezeichnen wir als **verallgemeinertes Konstruktionswerkzeug**.

**Definition 7** Als **Konstruktion**  $\mathcal{K}$  bezeichnen wir eine Folge von Konstruktionsschritten, welche die **Startkonfiguration**  $\Gamma_S = (T_S, E_S)$  durch endlich viele konsekutive Konstruktionsschritte

$$\Gamma_S = \Gamma_0 \rightarrow \Gamma_1 \rightarrow \dots \rightarrow \Gamma_N = \Gamma_E \quad (*)$$

mit der **Endkonfiguration**  $\Gamma_E = (T_E, E_E)$  verbindet.

Eine Realisierung  $O_S = (o_\tau \in \tau \mid \tau \in T_S)$  von  $\Gamma_S$  nennen wir **zulässig** für  $\mathcal{K}$ , wenn die konsekutiven Konstruktionsschritte sämtlich ausführbar sind, und bezeichnen die Herstellung der Realisierung  $O_E = (o_\tau \in \tau \mid \tau \in T_E)$  von  $\Gamma_E$  durch sukzessives Anwenden der Konstruktionsschritte als **Anwenden der Konstruktion** auf  $O_S$ . Wir schreiben auch kurz  $O_E = \mathcal{K}(O_S)$ .

Die Folge der Konstruktionsschritte, in welchen die Konfiguration  $\Gamma$  entsteht, bezeichnen wir als **Konstruktionsbeschreibung**.

Als **Zeichnung** bezeichnen wir schließlich die Anwendung einer Konstruktion auf eine Realisierung einer Parameterkonfiguration (also ein „weißes Blatt“).

Die Definition zeigt, dass zusammen mit (\*) auch jede Teilfolge

$$\Gamma_S = \Gamma_0 \rightarrow \Gamma_1 \rightarrow \dots \rightarrow \Gamma_i$$

eine Konstruktion ist. Wir bezeichnen sie als *Teil- oder Zwischenkonstruktion*  $\mathcal{K}^{(i)}$ .

Die dynamischen Freiheitsgrade einer Konstruktion werden allein durch deren Parameterobjekte bestimmt, sind also bereits in der Startkonfiguration festgelegt.

Es ist sinnvoll, als Belegung  $f$  nicht nur injektive Funktionen zuzulassen. So kann etwa das Konstruktionswerkzeug `altitude(A, B, C)`, welches die Höhe durch  $A$  im Dreieck  $\triangle ABC$  konstruiert, auch für  $A = B$  sinnvoll angewendet werden, um eine Senkrechte in  $B$  zu errichten. Die Nichtdegenerationsbedingung dieses Werkzeugs ist `altitude.NDG(A.c, B.c, C.c) = (B.c ≠ C.c)`.

Jede Konstruktionsbeschreibung ist zugleich ein verallgemeinertes Konstruktionswerkzeug, was die Einführung des Konzepts von **Makros** ermöglicht: Für ein verallgemeinertes Konstruktionswerkzeug  $w : T_1 \times \dots \times T_n \rightarrow U_1 \times \dots \times U_m$  kann man  $w$  aus dem Abbildungsgraphen

$$\tilde{w} = id \times w : T_1 \times \dots \times T_n \rightarrow T_1 \times \dots \times T_n \times U_1 \times \dots \times U_m$$

von  $w$  durch Projektion auf die  $U$ -Komponenten gewinnen. Dasselbe gilt für die Koordinatenabbildungen  $\tilde{w}.c$ . Ist nun

$$\Gamma_S = \Gamma_0 \xrightarrow{\tilde{w}_1} \Gamma_1 \xrightarrow{\tilde{w}_2} \dots \xrightarrow{\tilde{w}_N} \Gamma_N = \Gamma_E$$

die Folge von Konstruktionsschritten, welche eine Konstruktion  $\mathcal{K}$  beschreibt, so beschreibt

$$\tilde{w}_K = \tilde{w}_N \circ \dots \circ \tilde{w}_1 : \prod (\tau \in T_S) \longrightarrow \prod (\tau \in T_E)$$

das zugehörige verallgemeinerte Konstruktionswerkzeug. Die zugehörige Nichtdegenerationsbedingung ergibt sich analog als **und**-Verknüpfung aus den Abbildungen

$$\tilde{w}_K^{(i)}.NDG : \prod (C(\tau) : \tau \in T_S) \xrightarrow{\tilde{w}_{i-1}.c \circ \dots \circ \tilde{w}_1.c} \prod (C(\tau) : \tau \in T(\Gamma_{i-1})) \xrightarrow{\tilde{w}_i.NDG} \text{Boolean}$$

für  $i = 1, \dots, N$ , welche die Ausführbarkeit der  $i$ -ten Teilkonstruktion bestimmt.

Konstruktionsbeschreibungen werden wir in der formalen Notation eines **geometrischen Linearprogramms** (GLP) angeben. Ein solches Programm enthält in den ersten Zeilen die Startkonfiguration. Danach folgt die Angabe der einzelnen Konstruktionsschritte, wobei die Angabe der Belegung des entsprechenden Konstruktionswerkzeugs über Bezeichner für die konstruierten geometrischen Objekte ohne Vorwärtsreferenzen erfolgt.

Die Konstruktion eines durch drei freie Punkte aufgespannten Dreiecks lässt sich damit wie folgt beschreiben:

```

MP  $u_1, u_2, u_3$ ;
Point A = freePoint( $u_1$ );
Point B = freePoint( $u_2$ );
Point C = freePoint( $u_3$ );
Line a = pp_line(B,C);
Line b = pp_line(A,C);
Line c = pp_line(A,B);

```

Sind in solchen GLP Abkürzungen wie folgt erlaubt

```

Point F = intersection_point(pp_line(B,C),ortho_line(A,pp_line(B,C)));

```

so sprechen wir von der *schwachen GLP-Notation*. Sind als Parameter in einem Konstruktionswerkzeug nur Bezeichner zugelassen, so sprechen wir von der *strengen GLP-Notation*.

In obigem Beispiel ist  $F$  der Lotfußpunkt der Höhe durch  $A$  im Dreieck  $\triangle ABC$ . Jedes GLP in schwacher Notation kann ohne Schwierigkeiten durch Einführung *anonymer geometrischer Objekte* in die strenge Notation überführt werden (und dies wird intern bei der geometrischen Darstellung einer solchen Konfiguration auch ausgeführt). Obiger Schritt kann wie folgt in eine Sequenz von Konstruktionsschritten *entschachtelt* werden:

```

Line _l1 = pp_line(B,C);
Line _l2 = ortho_line(A,_l1);
Point F = intersection_point(_l1,_l2);

```

Dies kann sogar automatisch geschehen, wenn man zulässt, dass „dasselbe geometrische Objekt mehrfach konstruiert wird“.

**Aufgabe 11** *Erläutern Sie, wie die Sache mit dem „mehrfach konstruierten geometrischen Objekt“ wirklich ist.*

### 3 Geometrie und Formeln

#### 3.1 Symbolische analytische Geometrie

Ist das geometrische Objekt  $o = w(o_1, \dots, o_n)$  mit dem Konstruktionswerkzeug  $w$  konstruiert worden, so ergibt sich der Objektzustand  $o(t)$  von  $o$  zum Zeitpunkt  $t$  als  $o(t) = w(o_1(t), \dots, o_n(t))$  aus den Objektzuständen  $o_1(t), \dots, o_n(t)$  der Ausgangsobjekte zum Zeitpunkt  $t$ . Koordinatenmäßig sind sie über die mathematische Funktion  $w.c$  verbunden, welche die aktuellen Koordinaten von  $o$  vermöge  $o.c(t) = w.c(o_1.c(t), \dots, o_n.c(t))$  aus den aktuellen Korodinen von  $o_1, \dots, o_n$  berechnet. Dabei sind jedesmal dieselben Rechnungen mit je anderen Werten auszuführen.

Wir können die *wiederholte* Berechnung von  $o.c(t)$  für verschiedene  $t$  längs des zugehörigen GLP vermeiden, wenn wir  $w.c$  in einem System implementieren, das symbolische Formelmanipulationen erlaubt, dann  $o.c$  für „allgemeine“  $o_1, \dots, o_n$  berechnen und  $o.c(t)$  durch Variablenspezifikation aus dieser Formel gewinnen.

Entsprechende Implementierungen stehen für Maple, MuPAD, Reduce und Mathematica mit den GEOPROVER-Paketen [3] zur Verfügung. Die Pakete implementieren einen gemeinsamen Sprachstandard für Geometriebeweise, dessen Syntax in der GeoCode-Tabelle des SymbolicData-Projekts [4] fixiert ist und unterscheiden in der Version 1.3 (noch) nicht zwischen  $w$  und  $w.c$ . Kleinere Syntaxunterschiede zwischen den Sprachen der Target-CAS machen außerdem ein gemeinsames Oberformat und entsprechende Übersetzungswerkzeuge erforderlich. Hier ist eine Migration nach XML in Arbeit.

Aber weiter in unserem Text: Wir führen zunächst das Konzept des allgemeinen geometrischen Objekts ein. Dies ist ein „typisches“ Objekt eines geometrischen Typs, dessen Koordinaten nicht aus dem Grundbereich  $K$  gewählt sind, sondern aus einer genügend allgemeinen transzendenten Erweiterung von  $K$ , so dass jedes spezielle geometrische Objekt des vorgegebenen Typs daraus durch Variablenspezifikation hergestellt werden kann.

Sei also  $X = [x_\alpha : \alpha \in I]$  ein Satz von abzählbar vielen Variablen und  $R = K[x_\alpha : \alpha \in I]$  der Polynomring über  $K$  in diesen Variablen. Ist  $C = C(T)$  der Wertebereich eines geometrischen Typs  $T$  über dem Grundbereich  $K$ , so bezeichnen wir mit  $\overline{C} = C \otimes_K R$  den Wertebereich, den man durch Erweiterung  $K \rightarrow R$  des Grundbereichs erhält.

Betrachten wir als Beispiel eine Gerade  $g$  in ihren verschiedenen Bedeutungen. Als Folge spezieller geometrischer Objekte  $g(t)$  wird deren Lage durch  $g.c(t) \in C_G$  beschrieben. Als symbolisches Objekt  $\overline{g}$  ist die Gerade ein „spezielles geometrisches Objekt mit allgemeinen Koordinaten“, also  $\overline{g}.c \in \overline{C}_G$ , wobei die Koordinaten zusätzlich von den Variablen aus  $X$  abhängen. Die Koordinaten sollen so allgemein gewählt werden, dass sich jedes spezielle geometrische Objekt durch eine geeignete Variablenspezifikation herstellen lässt und es damit auch eine Folge spezieller Belegungen der Variablen  $X \rightarrow X(t)$  gibt, so dass  $g.c(t) = \text{subs}(X = X(t), \overline{g}.c)$  gilt.

**Definition 8** Ein spezielles geometrisches Objekt  $\overline{o}$  des Typs  $T$  mit  $o.c \in \overline{C}$  bezeichnen wir als universelles geometrisches Objekt vom Typ  $T$ , wenn sich aus  $\overline{o}$  alle speziellen geometrischen Objekte des Typs  $T$  durch Variablenspezifikation<sup>2</sup> herstellen lassen.

Ist  $T$  eine endliche Menge geometrischer Typen, so bezeichnen wir ein Tupel  $\overline{O} = (\overline{o}_\tau : \tau \in T)$  als Tupel universeller Objekte, wenn es aus universellen geometrischen Objekten  $\overline{o}_\tau$  der richtigen Typen besteht, die unabhängig voneinander spezifiziert werden können. Genauer gesagt verlangen wir, dass sich aus dem Tupel  $\overline{O}$  jedes Tupel

$$\prod (o_\tau : \tau \in T) \in \prod (C(\tau) : \tau \in T)$$

<sup>2</sup>Genau gesagt muss gefordert werden, dass dies nicht nur für spezielle Objekte mit Koordinaten aus  $K$  gilt, sondern auch für Objekte, deren Koordinaten in Erweiterungen von  $K$  liegen.  $R$  hat die universelle Eigenschaft, dass jede (endliche) Erweiterung von  $K$  als Spezialisierung der Erweiterung  $K \rightarrow R$  hergestellt werden kann. Auf diese Feinheiten soll hier nicht eingegangen werden. Die Idee eines solchen *universellen geometrischen Objekts* folgt dem Konzept des allgemeinen Punkts einer Varietät in der algebraischen Geometrie.

spezieller geometrischer Objekte durch Variablenspezifikation gewinnen lässt.

Insbesondere nennen wir eine Realisierung  $\overline{O}$  einer Konfiguration  $\Gamma = (T, E)$  eine universelle Realisierung, wenn  $\overline{O}$  zugleich ein Tupel universeller Objekte ist<sup>3</sup>.

Beschreibt  $w : T_1 \times \dots \times T_n \rightarrow T_a$  einen Konstruktionsschritt, so können wir diesen mit einem Tupel  $\overline{O} = (\overline{o}_1, \dots, \overline{o}_n)$  universeller Objekte ausführen. Da alle von uns bisher implementierten Funktionen Punkte oder Geraden zurückliefern, deren Koordinaten sich rational in den Koordinaten der als Aufrufparameter verwendeten Objekte ausdrücken lassen, erhalten wir für die Koordinaten des Ergebnisobjekts dieses Konstruktionsschritts rationale Ausdrücke in den Koordinaten von  $\overline{O}$ . Diese Ausdrücke stellen *universelle Formeln* in dem Sinne dar, dass jede Ausführung des Konstruktionsschritts mit einer zulässigen speziellen Realisierung  $O = (o_1, \dots, o_n)$  der Startkonfiguration, also einer solchen mit  $w.NDG(o_1.c, \dots, o_n.c) = \text{true}$ , zu einem speziellen Ergebnisobjekt führt, dessen Koordinaten sich auch durch jede Variablenspezifikation aus den universellen Formeln bestimmen lassen, durch welche sich  $O$  aus  $\overline{O}$  ergibt.

Ist

$$\Gamma_S = \Gamma_0 \xrightarrow{w_1} \Gamma_1 \xrightarrow{w_2} \dots \xrightarrow{w_N} \Gamma_N = \Gamma_E$$

eine Konstruktion, so können wir diese Überlegungen auf jeden einzelnen Konstruktionsschritt anwenden. Starten wir mit einer universellen Realisierung  $\overline{O}_S$  der Startkonfiguration  $\Gamma_S$ , so liefert  $w_1.c$  universelle Formeln für die Koordinaten des in  $\Gamma_1$  neu konstruierten Objekts  $o$ .

Die universelle Formel von  $w_2$  ergibt sich, wenn statt  $o$  auch ein universelles Objekt  $\overline{o}$  verwendet wird. Die Koordinaten von  $w_2 \circ w_1(\overline{O})$  ergeben sich, wenn in dieser universellen Formel die Spezifikation  $\overline{o} \rightarrow o$  ausgeführt wird. Dies bedeutet, in der universellen Formel von  $w_2$  die Variablen aus  $\overline{o}$  durch die universellen Formeln zu ersetzen, welche die Koordinaten von  $o$  beschreiben. Kurz, in der universellen Formel von  $w_2$  haben wir einzelne Variablen durch rationale Ausdrücke zu ersetzen.

Dies setzt sich über die anderen Konstruktionsschritte fort, so dass sich die Koordinaten der Realisierung  $\overline{O}_E$  der Endkonfiguration aus den universellen Formeln des Schritts  $w_N$  ergeben, indem dort für die Variablen rationale Ausdrücke in den Variablen von  $\overline{O}_S$  eingesetzt werden, die sich Schritt für Schritt nach demselben Prinzip in den vorangegangenen Konstruktionsschritten ergeben haben. Ersetzt man aber in einem rationalen Ausdruck einzelne Variablen durch andere rationale Ausdrücke, so entsteht als Ergebnis wieder ein rationaler Ausdruck (oder, wenn sich der Nullausdruck als Hauptnenner ergibt, ein Fehler). Durch eine solche Konstruktion wird der Bereich der rationalen Ausdrücke also nicht verlassen.

Bestimmen wir etwa die Koordinaten des Umkreismittelpunkts des Dreiecks  $\triangle ABC$  als Schnittpunkt der entsprechenden Mittelsenkrechten  $m, n$ , so bedeutet dies auf dem betrachteten symbolischen Niveau, dass beim Aufruf von `intersection_point(p_bisector(B,C), p_bisector(C,A))` in der universellen Formel

$$\left( \frac{m_2 n_3 - m_3 n_2}{m_1 n_2 - m_2 n_1}, \frac{m_3 n_1 - m_1 n_3}{m_1 n_2 - m_2 n_1} \right),$$

nach der sich die Koordinaten des Schnittpunkts der Geraden  $m, n$  berechnen, die Komponenten von  $m, n$  durch die entsprechenden rationalen Ausdrücke

$$\text{p\_bisector.c}(B.c, C.c) = \left( b_1 - c_1, b_2 - c_2, \frac{-b_1^2 - b_2^2 + c_1^2 + c_2^2}{2} \right)$$

bzw.

$$\text{p\_bisector.c}(C.c, A.c) = \left( -a_1 + c_1, -a_2 + c_2, \frac{a_1^2 + a_2^2 - c_1^2 - c_2^2}{2} \right)$$

zu ersetzen sind. Diese Ausdrücke sind ihrerseits während des Aufrufs

<sup>3</sup>Was nichts anderes bedeutet als dass wir die Vorgeschichte  $E$  der Entstehung von  $\Gamma$  vergessen.

```
p_bisector(B,C) == ortho_line(midPoint(B,C),pp_line(B,C))
```

in einem Substitutions- und Simplifikationsprozess (der Tiefe 2) nach demselben Prinzip entstanden. Wir erhalten in diesem Fall als universelle Formel für die Koordinaten des Umkreismittelpunkts  $M$

$$M.c := \left( \begin{array}{l} \frac{a_1^2 b_2 - a_1^2 c_2 + a_2^2 b_2 - a_2^2 c_2 -}{b_1^2 c_2 + b_2^2 c_2 - b_2 c_1^2 - b_2 c_2^2} \quad , \quad \frac{-a_1^2 b_1 + a_1^2 c_1 + a_1 b_1^2 + a_1 b_2^2 -}{a_1 c_1^2 - a_1 c_2^2 - a_2^2 b_1 + a_2^2 c_1 -} \\ \frac{2 a_1 b_2 - 2 a_1 c_2 - 2 a_2 b_1 + 2 a_2 c_1 +}{2 b_1 c_2 - 2 b_2 c_1} \quad , \quad \frac{b_1^2 c_1 + b_1 c_1^2 + b_1 c_2^2 - b_2^2 c_1}{2 a_1 b_2 - 2 a_1 c_2 - 2 a_2 b_1 + 2 a_2 c_1 +} \\ \frac{}{2 b_1 c_2 - 2 b_2 c_1} \end{array} \right)$$

wenn wir vom universellen Dreieck  $\triangle ABC$  mit  $A.c = (a_1, a_2)$ ,  $B.c = (b_1, b_2)$ ,  $C.c = (c_1, c_2)$  als Startkonfiguration ausgehen.

Der Simplifikationsprozess rationaler Funktionen liefert eine einfachere rationale Funktion, die für alle Variablenspezifikationen, welche für die unsimplifizierte Funktion nicht zu einem Fehlerabbruch führen, denselben Wert hat wie diese. Der Definitionsbereich der simplifizierten rationalen Funktion kann allerdings größer sein als derjenige der Ausgangsfunktion.

Für diesen Schnittpunkt  $M$  können wir nun untersuchen, ob er tatsächlich der Umkreismittelpunkt ist, also eine gewisse geometrische Eigenschaft hat. Dazu wäre etwa zu untersuchen, ob die universellen Formeln

```
sqrdist(M,A)-sqrdist(M,B);
sqrdist(M,A)-sqrdist(M,C);
```

welche diese Eigenschaft kodieren, jeweils zu Null simplifizieren.

**Definition 9** Seien  $T_1, \dots, T_n$  geometrische Typen.

Eine (informatische) Funktion

$$\phi : T_1 \times \dots \times T_n \rightarrow \text{boolean}$$

zusammen mit einer (mathematischen) Funktionen

$$\phi.c : C(T_1) \times \dots \times C(T_n) \rightarrow \text{Boolean}$$

bezeichnen wir als geometrische Eigenschaft.

Das eben beschriebene Beispiel ist die typische Konstellation für eine ganze Klasse geometrischer Sätze: Wenn sich eine gewisse Menge geometrischer Objekte in einer durch eine Konstruktion beschriebenen Abhängigkeitsrelation zueinander befindet, dann gilt für diese Objekte eine zusätzliche geometrische Eigenschaft. Sätze dieser Struktur bezeichnen wir als *geometrische Sätze vom konstruktiven Typ*.

Weitere Beispiele siehe `geo-1.txt` oder das Maple-Arbeitsblatt `geo-1.mws`.

Abschließend sei angemerkt, dass die bisherigen Betrachtungen immer vom Modell der affinen Punktkoordinaten ausgingen. Für das Modell mit homogenen Punktkoordinaten lässt sich sogar erreichen, dass alle universellen Formeln polynomial sind, da durch Skalieren mit einem entsprechenden Faktor in den universellen Formeln immer Nennerfreiheit erreicht werden kann. Da die Klasse der polynomialen Ausdrücke nicht verlassen wird, wenn man in polynomialen Ausdrücken Variablen durch Polynome ersetzt, gelten dieselben Ausführungen wie oben, wenn man „rationale Ausdrücke“ durch „polynomiale Ausdrücke“ ersetzt. Dabei können überhaupt keine rechnerischen Ausnahmen mehr auftreten, sondern nur die geometrische Ausnahme, dass sich als universelle Koordinaten eines Objekts  $(0 : 0 : 0)$  ergibt.

### 3.2 Der Mechanisierungssatz für geometrische Sätze vom konstruktiven Typ

Untersuchen wir nun die Beweiskraft der ausgeführten symbolischen Berechnungen für Sätze vom konstruktiven Typ näher. Sei  $\mathcal{K}$  die Konstruktion, welche die Endkonfiguration beschreibt, in welcher eine gewisse geometrische Eigenschaft gelten soll,  $\overline{O}$  eine universelle Realisierung der Startkonfiguration von  $\mathcal{K}$ ,  $O$  eine zulässige spezielle Realisierung dieser Startkonfiguration, welche durch die Variablenspezifikation  $X \rightarrow X_0$  aus  $\overline{O}$  entsteht, für die also  $O.c = \text{subs}(X = X_0, \overline{O}.c)$  gilt. Wir wollen zunächst annehmen, dass die Konstruktion auf  $\overline{O}$  ausführbar ist.

$\mathcal{K}$  ist aus einzelnen Konstruktionsschritten aufgebaut, so dass sich ein induktiver Zugang aufdrängt. Seien deshalb  $\mathcal{K}^{(i-1)}$  und  $\mathcal{K}^{(i)}$  Teilkonstruktionen sowie  $w_i$  der sie verbindende Konstruktionsschritt. Wenden wir  $\mathcal{K}^{(i-1)}$  auf  $\overline{O}$  an, so entsteht eine Realisierung  $\overline{O}_{i-1}$  von  $\Gamma_{i-1}$ . Ähnlich sind  $O$  und  $O_{i-1}$  verbunden. Wir nehmen als Induktionsvoraussetzung an, dass die Koordinaten von  $\overline{O}_{i-1}$  durch universelle Formeln beschrieben werden, also  $O_{i-1}.c = \text{subs}(X = X_0, \overline{O}_{i-1}.c)$  gilt. Anwenden von  $w_i$  auf  $\overline{O}_{i-1}$  liefert eine Realisierung  $\overline{O}_i = \tilde{w}_i(\overline{O}_{i-1})$  von  $\Gamma_i$ , wobei sich die Koordinaten des neu konstruierten Objekts aus der universellen Formel von  $w_i$  ergeben, indem wie oben eine universelle Realisierung von  $\Gamma_{i-1}$  zu  $\overline{O}_{i-1}$  spezifiziert wird. Die Wirkung ist dieselbe wie beim Aufruf der (informatischen) Funktion  $w_i$ , wo die in der Definition verwendeten Aufrufparameter durch Werte, in diesem Fall die berechneten universellen Formeln, zu ersetzen sind. Konstruiert man von der *speziellen* Realisierung  $O_{i-1}$  ausgehend  $w_i(O_{i-1})$ , so wird genau derselbe Aufrufmechanismus für die Berechnung der Koordinaten des neuen *speziellen* Objekts aus den *speziellen* Koordinaten von  $O_{i-1}$  aufgerufen. Die entscheidende Eigenschaft ist also

$$w_i.c(O_{i-1}.c) = w_i.c(\text{subs}(X = X_0, \overline{O}_{i-1}.c)) = \text{subs}(X = X_0, w_i.c(\overline{O}_{i-1}.c)),$$

oder in Worten, dass es egal ist, erst die universelle Formel für  $w_i(\overline{O}_{i-1})$  zu berechnen und dann spezielle Koordinaten einzusetzen oder spezielle Koordinaten bereits in  $\overline{O}_{i-1}$  einzusetzen (was nach Induktionsvoraussetzung  $O_{i-1}$  ergibt) und daraus  $w_i(O_{i-1})$  zu berechnen. Impliziter Kern der Argumentation ist also die Kommutativität des folgenden Diagramms

$$\begin{array}{ccccc} \overline{O}.c & \xrightarrow{\tilde{w}^{(i-1)}.c} & \overline{O}_{i-1}.c & \xrightarrow{\tilde{w}_i.c} & \overline{O}_i.c \\ \downarrow s & & \downarrow s & & \downarrow s \\ O.c & \xrightarrow{\tilde{w}^{(i-1)}.c} & O_{i-1}.c & \xrightarrow{\tilde{w}_i.c} & O_i.c \end{array}$$

wobei  $s$  für die Variablensubstitution  $X \rightarrow X_0$  steht.

Die Zulässigkeit von  $O$  für  $\mathcal{K}^{(i)}$  lässt sich durch die Bedingung

$$\bigwedge_{j \leq i} \tilde{w}_j.\text{NDG}(O_{j-1}.c) = \text{true}$$

beschreiben, die eine Konjunktion geometrischer Eigenschaften ist.

Nachdem wir nun verstehen, was genau das symbolische Ausführen einer Konstruktion bedeutet und welche Rolle die dabei produzierten universellen Formeln spielen, gilt es den Charakter des symbolischen Auswertens einer geometrischen Eigenschaft zu verstehen. Sei also

$$\phi : T_1 \times \dots \times T_n \rightarrow \text{boolean}$$

eine solche geometrische Eigenschaft. Rufen wir  $\phi$  mit speziellen geometrischen Objekten über dem Grundbereich  $K$  auf, so erhalten wir einen der Werte **true** oder **false** zurück. Das gilt nicht mehr, wenn wir denselben Aufruf in einer symbolischen Umgebung absetzen, denn der entstehende *boolesche Ausdruck* wird normalerweise noch Variablen enthalten und kann erst nach Belegung der Variablen mit konkreten Werten vollständig ausgewertet werden. Rufen wir also  $\phi$  mit einem Tupel

symbolischer Objekte auf, so ist das Ergebnis nicht aus dem Wertebereich `Boolean`, sondern aus dem Bereich `BooleanExpression`:

$$\phi.c : \overline{C}(T_1) \times \dots \times \overline{C}(T_n) \rightarrow \text{BooleanExpression}$$

Rufen wir  $\phi$  mit einem Tupel *universeller* Objekte auf, so ist das Ergebnis der zu  $\phi$  assoziierte *universelle boolesche Ausdruck*.

Analog zu obigem kommutativem Diagramm erhalten wir das folgende Diagramm

$$\begin{array}{ccccc} \overline{O}.c & \xrightarrow{\tilde{w}^{(N)}.c} & \overline{O}_N.c & \xrightarrow{\phi.c} & \phi.c(\overline{O}_N.c) \\ \downarrow s & & \downarrow s & & \downarrow s \\ O.c & \xrightarrow{\tilde{w}^{(N)}.c} & O_N.c & \xrightarrow{\phi.c} & \phi.c(O_N.c) \end{array}$$

wobei  $\phi.c(\overline{O}_N.c) \in \text{BooleanExpression}$  der boolesche Ausdruck ist, welcher sich aus einem universellen booleschen Ausdruck von  $\phi$  durch Spezifikation des universellen Tupels zu  $\overline{O}_N$  ergibt.

Die oben geführten Beweise verwenden wiederum implizit die Kommutativität dieses Diagramms, d. h. dass der konkrete boolesche Wert  $\phi.c(O_N.c) \in \text{Boolean}$  gleichberechtigt als Ergebnis der Anwendung von  $\phi$  auf  $O_N$  oder aus  $\phi.c(\overline{O}_N.c)$  durch die Variablenspezifikation  $s$  und nachfolgende Simplifikation des nun variablenfreien booleschen Ausdrucks zu einem booleschen Wert gewonnen werden kann.

In allen bisherigen Beispielen lässt sich die Behauptung  $\phi$  als Verschwinden des Werts einer gewissen Berechnung formulieren, d. h. in der Form

$$\phi.c : C(T_1) \times \dots \times C(T_n) \xrightarrow{\overline{\phi}} K \xrightarrow{\text{iszero}} \text{Boolean}$$

faktorisieren, was sich nach Grundbereichserweiterung  $K \rightarrow R$  zu

$$\phi.c : \overline{C}(T_1) \times \dots \times \overline{C}(T_n) \xrightarrow{\overline{\phi}} R \xrightarrow{\text{iszero}} \text{BooleanExpression}$$

verändert. Zu  $\overline{\phi}$  gehörte stets eine polynomiale universelle Formel, aus welcher der in allen Fällen wenigstens rationale Ausdruck  $\overline{\phi}(\overline{O}_N.c)$  durch Spezifikation des universellen Tupels zu  $\overline{O}_N$  gewonnen wurde. Eine weitere Simplifikation vereinfachte dann bereits diesen rationalen *Ausdruck* zu Null, woraus folgt, dass  $\text{iszero} \circ \overline{\phi}(\overline{O}_N.c)$  bereits als boolescher *Ausdruck* unabhängig von Variablenbelegungen zu `true` simplifiziert.

Von ähnlicher Struktur sind die Zulässigkeitsbedingungen  $\tilde{w}_j.\text{NDG}(O_{j-1}.c) = \text{true}$  für  $O$ , die sich bisher stets als  $(\text{not iszero}) \circ \tilde{w}_j.\text{DG}(O_{j-1}.c)$  darstellen ließen, wobei  $\tilde{w}_j.\text{DG}$  dieselbe Struktur hat wie  $\overline{\phi}$  in obigen Betrachtungen.

Kern der gesamten Argumentation ist also die Kommutativität der oben angeführten Diagramme, was äquivalent zur Vertauschbarkeit von Variablensubstitution und Berechnung der entsprechenden universellen Formeln ist. In allen bisher betrachteten Beispielen ist dies durch den polynomialen oder rationalen Charakter der universellen Formeln gewährleistet, da Variablensubstitutionen operationstreu sind, d. h. mit den arithmetischen Operationen in  $R$  bzw. im Quotientenkörper  $Q(R)$  kommutieren. Wir führen deshalb die folgenden Begriffe ein.

**Definition 10** *Ein Konstruktionswerkzeug  $w$ , dessen universelle Formel aus rationalen oder sogar polynomialen Ausdrücken in den Variablen eines universellen Tupels besteht, bezeichnen wir als rational bzw. polynomial.*

*Geometrische Bedingungen  $\phi$ , die sich wie oben in der Form  $\phi.c = \text{iszero} \circ \overline{\phi}$  darstellen lassen, so dass die universelle Formel von  $\overline{\phi}$  polynomial ist, bezeichnen wir als polynomiale Bedingung.*

Lässt sich  $w.\text{NDG}$  als  $(\text{not iszero}) \circ w.\text{DG}$  wie oben darstellen<sup>4</sup>, wobei die universelle Formel von  $w.\text{DG}$  polynomial ist, bezeichnen wir  $w$  als Konstruktionswerkzeug mit polynomialen Degenerationsbedingungen.

Eine Konstruktion  $\mathcal{K}$  bezeichnen wir als polynomial bzw. rational, wenn sie aus polynomialen bzw. rationalen Konstruktionswerkzeugen mit polynomialen Degenerationsbedingungen aufgebaut ist.

Als geometrischen Satz vom konstruktiven Typ bezeichnen wir eine rationale Konstruktion  $\mathcal{K}$  zusammen mit einer auf der Endkonfiguration von  $\mathcal{K}$  gegebenen polynomialen Bedingung  $\phi$ .

Wir sagen, dass der Satz gilt, wenn für jede zulässige spezielle Realisierung der Startkonfiguration von  $\mathcal{K}$  die Bedingung  $\phi$  auf der zugehörigen Realisierung der Endkonfiguration zu `true` auswertet.

Wie ausgeführt ist zu beachten, dass diese Definitionen nicht nur vom Charakter der Konstruktionswerkzeuge abhängen, sondern auch vom verwendeten Koordinatenmodell. Im Modell der affinen Punktkoordinaten etwa sind einige der bisher betrachteten Konstruktionswerkzeuge nur rationale Werkzeuge, im Modell der homogenen Punktkoordinaten dagegen alle Konstruktionswerkzeuge polynomial.

Alle bisher betrachteten geometrischen Sätze waren Sätze vom konstruktiven Typ. Ein solcher Satz  $(\mathcal{K}, \phi)$  wird durch die folgenden Formeln begleitet: Universelle Realisierungen  $\overline{O} \subset \overline{O}'$  von Start- und Endkonfiguration von  $\mathcal{K}$ , universelle Formeln für die Ausführung  $\overline{O}_E = \mathcal{K}(\overline{O})$  von  $\mathcal{K}$  auf  $\overline{O}$ , die universelle Formel  $\Phi = \overline{\phi}(\overline{O}'.c)$  der Behauptung  $\phi$  (enthält Variablen aus  $\overline{O}'$ ) sowie das Resultat der Substitution  $\Phi' = \overline{\phi}(\overline{O}_E.c) = \Phi|_{\overline{O}' \rightarrow \overline{O}_E}$  (enthält nur noch Variablen aus  $\overline{O}$ ).

Für die ausgeführten symbolischen Rechnungen gibt es folgende Alternativen:

1.  $\overline{O}$  selbst ist für  $\mathcal{K}$  nicht zulässig.

Sei  $i$  minimal mit der Eigenschaft, dass  $\overline{O}$  für alle  $\mathcal{K}^{(j)}$  mit  $j \geq i$  nicht zulässig ist. Die zugehörige universelle Formel der Degenerationsbedingung von  $\mathcal{K}^{(i)}$  ist  $\psi_i = w_i.\text{DG}(\overline{O}_{i-1}.c)$ , wobei sich wegen der Minimalität von  $i$  die Realisierung  $\overline{O}_{i-1}$  berechnen lässt. Degeneriertheit von  $\overline{O}$  bedeutet, dass  $\psi_i$  als rationale Funktion in den Variablen von  $\overline{O}$  zu Null simplifiziert<sup>5</sup>.

Es gibt also auch keine zulässigen speziellen Realisierungen — die Voraussetzungen des Satzes sind widersprüchlich.

Dieser Sachverhalt lässt sich seinerseits als geometrischer Satz vom konstruktiven Typ mit der Behauptung  $w_i.\text{DG}$  formulieren.

2. Die zurückgegebene rationale Funktion  $\Phi'$  simplifiziert zu Null.

Dann gilt die geometrische Aussage für alle zulässigen speziellen Realisierungen  $O$  der Startkonfiguration, da sich der Wert von  $\overline{\phi}$  auf  $O$  durch Variablenspezifikation  $\overline{O} \rightarrow O$  aus  $\Phi'$  ergibt. Der Satz ist gültig.

3. Die zurückgegebene rationale Funktion  $\Phi'$  simplifiziert nicht zu Null.

Dann gilt die Aussage für fast alle zulässigen speziellen Realisierungen der Startkonfiguration nicht. Der Satz ist in der formulierten Form nicht allgemeingültig.

Die Aussage gilt nur unter Zusatzbedingungen, die analytisch das Verschwinden des Zählerpolynoms des simplifizierten Ausdrucks nach sich ziehen müssen.

Gelingt es, diese Bedingung als geometrische Eigenschaft zu identifizieren, dann lässt sich ein entsprechender geometrischer Satz formulieren. Er ist aber nicht mehr vom konstruktiven Typ.

<sup>4</sup>Mehrere und-verknüpfte solche Bedingungen lassen sich durch Aufmultiplizieren zu einer zusammenfassen, da ein Produkt nur dann verschieden Null ist, wenn alle Faktoren verschieden Null sind. Das ist allerdings eine theoretische Aussage. In praktischen Anwendungen arbeitet man aus Performancegründen mit mehreren Polynomen.

<sup>5</sup>Da die universelle Formel von  $w_i.\text{DG}$  ein polynomialer Ausdruck ist, ist  $\psi_i$  immer definiert, wenn es die universellen Formeln von  $\overline{O}_{i-1}$  sind.

Eine spezielle Realisierung  $O$  der Startkonfiguration ist dabei zulässig, wenn auf ihr alle Konstruktionsschritte ausgeführt werden können, d. h. wenn  $\tilde{w}_K^{(i)}.NDG(O.c)$  für alle  $i = 1, \dots, N$  den Wert **true** liefert. Auch das können wir an den universellen Formeln von  $w_i.DG$  ablesen, indem wir das Tupel universeller Objekte jeweils durch die universellen Formeln von  $\overline{O}_{i-1}$  ersetzen. *Universelle Degenerationsbedingungen* von  $\mathcal{K}$  sind also genau die Zähler  $\psi_i$  der rationalen Funktionen  $w_i.DG(\overline{O}_{i-1}.c)$ . Als Formeln mit Variablen aus  $\overline{O}$  können sie während des *Aufrufs* des jeweiligen Konstruktionswerkzeugs zur Berechnung der universellen Formeln von  $\overline{O}_i$  aus denen von  $\overline{O}_{i-1}$  als Seiteneffekt mit berechnet und in einer für  $\mathcal{K}$  global definierten Liste aufgesammelt werden.

Wir haben damit den folgenden Mechanisierungssatz bewiesen:

**Satz 13** (*Über das mechanisierte Beweisen geometrischer Sätze vom konstruktiven Typ*)

Sei  $(\mathcal{K}, \phi)$  ein geometrischer Satz vom konstruktiven Typ,

- $\overline{O}$  eine universelle (und zulässige – sonst wäre  $\mathcal{K}$  widersprüchlich) Realisierung der Startkonfiguration von  $\mathcal{K}$ ,
- $\overline{O}_E$  das Ergebnis der Anwendung von  $\mathcal{K}$  auf  $\overline{O}$ ,
- $\Phi' = \overline{\phi}(\overline{O}_E.c) \in Q(R)$  und
- $\psi_i \in R$  die universellen Degenerationsbedingungen der einzelnen Teilkonstruktionen und  $\Psi = \psi_1 \cdot \dots \cdot \psi_N \in R$ .

Der Satz ist genau dann gültig, d. h. gilt für alle zulässigen speziellen Realisierungen  $O$  der Startkonfiguration, wenn  $\Phi'$  als rationale Funktion zu Null vereinfacht werden kann.

Zulässige spezielle Realisierungen  $O$  sind genau diejenigen, welche aus  $\overline{O}$  durch eine Variablenspezifikation  $X \rightarrow X_0$  gewonnen werden können, für die  $\text{subs}(X = X_0, \psi_i) \neq 0$  für  $i = 1, \dots, N$  oder kurz  $\text{subs}(X = X_0, \Psi) \neq 0$  gilt.

## 4 Geometrische Sätze vom Gleichungstyp

Der bisher betrachtete Mechanisierungsansatz für Geometrietheoreme ging davon aus, dass man einen gegebenen Satz der Geometrie in eine konstruktive Form überführen kann. Das ist aber eine für geometrische Aussagen eher untypische Situation. Die meisten geometrischen Sätze gehen von einer bestimmten (konstruktiv gegebenen) geometrischen Konfiguration  $\mathcal{K}$  aus und behaupten dann:

*Wenn in  $\mathcal{K}$  zusätzlich gewisse geometrische Eigenschaften erfüllt sind, dann ergeben sich daraus gewisse andere geometrische Eigenschaften als Konsequenzen.*

So behauptet etwa der **Satz des Ceva**:

Gegeben sei ein Dreieck  $\triangle ABC$  und Punkte  $P, Q, R$  auf den Dreiecksseiten mit Teilverhältnissen  $u, v, w$ .

Die Ecktransversalen  $AP, BQ$  und  $CR$  gehen genau dann durch einen gemeinsamen Punkt, wenn

$$u v w = (1 - u)(1 - v)(1 - w)$$

gilt.

Eine wortwörtliche Übertragung der Formulierung in die im vorigen Abschnitt entwickelte Sprache lautet:

```
Point A = freePoint(MP(a1,a2));
Point B = freePoint(MP(b1,b2));
Point C = freePoint(MP(c1,c2));
Point P = varpoint(B,C,SP(u));
Point Q = varpoint(C,A,SP(v));
Point R = varpoint(A,B,SP(w));
```

wobei  $MP(a, b)$  bzw.  $SP(u)$  universelle Parameterobjekte vom Typ MP bzw. SP bezeichnen.

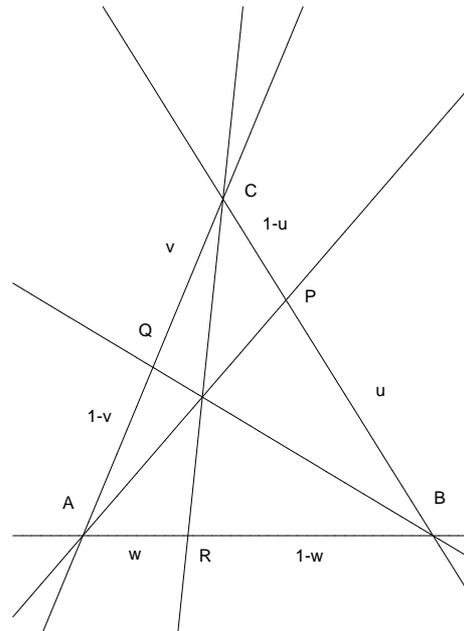
Wir werden im Weiteren die folgende kürzere Notation der GEOPROVER-Version 1.3 verwenden, die sich ausschließlich auf die Koordinatenattribute bezieht.

```
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2);
P:=varpoint(B,C,u); Q:=varpoint(C,A,v); R:=varpoint(A,B,w);

pre:=is_concurrent(pp_line(A,P),pp_line(B,Q),pp_line(C,R));
con:=u*v*w-(1-u)*(1-v)*(1-w);
```

In der gegebenen Konfiguration, die generisch durch die 6 Koordinaten von  $A, B, C$  und die Teilverhältnisse  $u, v, w$  beschrieben wird, ist also zu zeigen, dass jede Nullstelle des Polynoms  $pre$  (der Voraussetzung = *presumptio*) auch eine Nullstelle des Polynoms  $con$  (der Schlussfolgerung = *conclusio*) ist. In diesem Fall lässt sich  $pre$  in die drei Faktoren

$$\begin{aligned} f_1 &= -a_1 b_2 + a_1 c_2 + a_2 b_1 - a_2 c_1 - b_1 c_2 + b_2 c_1, \\ f_2 &= a_1 b_2 - a_1 c_2 - a_2 b_1 + a_2 c_1 + b_1 c_2 - b_2 c_1, \\ f_3 &= 2 u v w - u v - u w + u - v w + v + w - 1 \end{aligned}$$

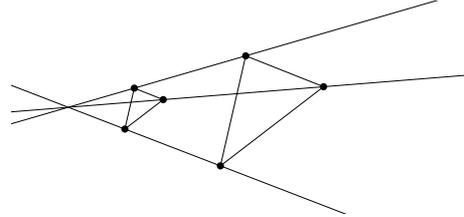


Der Satz von Ceva

zerlegen, von denen nur der dritte mit *con* übereinstimmt, so dass es auch Spezifikationen gibt, in denen *pre* verschwindet, *con* jedoch nicht.

Analysiert man die beiden anderen Faktoren genauer, so erkennt man, dass sie genau der Bedingung `is_collinear(A,B,C)` entsprechen. Eine Spezifikation, in der dieses Polynom verschwindet, wäre also eine *degenerierte Lage*, so dass wir sehen, dass die beiden Polynome *pre* und *con* “im Wesentlichen” übereinstimmen. Der Satz des Ceva gilt also für alle (nicht degenerierten) Dreiecke.

Ähnlich können wir auch die Aussage des **Satzes von Desargue** formulieren. Hier sind 6 Punkte  $A, B, C, A', B', C'$  gegeben. Die Voraussetzung besteht aus drei geometrischen Bedingungen, dass in den Dreiecken  $ABC$  und  $A'B'C'$  einander entsprechende Geraden zueinander parallel sind. Die Behauptung lautet, dass dann die Geraden  $AA', BB'$  und  $CC'$  konkurrent sind:



Der Satz von Desargue

```
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2);
A1:=Point(d1,d2); B1:=Point(e1,e2); C1:=Point(f1,f2);

pre:={ is_parallel(pp_line(A,B),pp_line(A1,B1)),
        is_parallel(pp_line(B,C),pp_line(B1,C1)),
        is_parallel(pp_line(A,C),pp_line(A1,C1))};

con:=is_concurrent(pp_line(A,A1),pp_line(B,B1),pp_line(C,C1));
```

Die Voraussetzung entspricht hier also einer Menge von drei Polynomen

$$\{ a_1 d_2 - a_1 e_2 - a_2 d_1 + a_2 e_1 - b_1 d_2 + b_1 e_2 + b_2 d_1 - b_2 e_1, \\ b_1 e_2 - b_1 f_2 - b_2 e_1 + b_2 f_1 - c_1 e_2 + c_1 f_2 + c_2 e_1 - c_2 f_1, \\ a_1 d_2 - a_1 f_2 - a_2 d_1 + a_2 f_1 - c_1 d_2 + c_1 f_2 + c_2 d_1 - c_2 f_1 \},$$

die Behauptung dem deutlich komplizierteren Polynom

$$\begin{aligned} & -a_1 b_1 c_2 d_2 + a_1 b_1 c_2 e_2 + a_1 b_1 d_2 f_2 - a_1 b_1 e_2 f_2 + a_1 b_2 c_1 d_2 - a_1 b_2 c_1 f_2 - a_1 b_2 c_2 e_1 + \\ & a_1 b_2 c_2 f_1 - a_1 b_2 d_2 f_1 + a_1 b_2 e_1 f_2 - a_1 c_1 d_2 e_2 + a_1 c_1 e_2 f_2 + a_1 c_2 d_2 e_1 - a_1 c_2 e_2 f_1 - \\ & a_1 d_2 e_1 f_2 + a_1 d_2 e_2 f_1 - a_2 b_1 c_1 e_2 + a_2 b_1 c_1 f_2 + a_2 b_1 c_2 d_1 - a_2 b_1 c_2 f_1 - a_2 b_1 d_1 f_2 + \\ & a_2 b_1 e_2 f_1 - a_2 b_2 c_1 d_1 + a_2 b_2 c_1 e_1 + a_2 b_2 d_1 f_1 - a_2 b_2 e_1 f_1 + a_2 c_1 d_1 e_2 - a_2 c_1 e_1 f_2 - \\ & a_2 c_2 d_1 e_1 + a_2 c_2 e_1 f_1 + a_2 d_1 e_1 f_2 - a_2 d_1 e_2 f_1 + b_1 c_1 d_2 e_2 - b_1 c_1 d_2 f_2 - b_1 c_2 d_1 e_2 + \\ & b_1 c_2 d_2 f_1 + b_1 d_1 e_2 f_2 - b_1 d_2 e_2 f_1 + b_2 c_1 d_1 f_2 - b_2 c_1 d_2 e_1 + b_2 c_2 d_1 e_1 - b_2 c_2 d_1 f_1 - \\ & b_2 d_1 e_1 f_2 + b_2 d_2 e_1 f_1 - c_1 d_1 e_2 f_2 + c_1 d_2 e_1 f_2 + c_2 d_1 e_2 f_1 - c_2 d_2 e_1 f_1 \end{aligned}$$

Dass jede gemeinsame Nullstelle der Polynome *pre* eine Nullstelle von *con* ist, ist hier nicht mehr so einfach zu erkennen. Der Satz ist jedoch offensichtlich ein Satz der affinen Geometrie, so dass wir zur Koordinatisierung ein beliebiges schiefwinkliges Koordinatensystem verwenden können, bzw. alternativ bei vorgegebenem Koordinatensystem durch eine geeignete affine Transformation die Konfiguration Eigenschaften erhaltend in eine solche transformieren, für die

```
A:=Point(0,0); B:=Point(0,1); C:=Point(1,0);
```

gilt. Mit diesen speziellen Koordinaten erhalten wir

$$\text{pre} = \{d_1 - e_1, -e_1 - e_2 + f_1 + f_2, -d_2 + f_2\}$$

und als Behauptung

$$\text{con} = -d_1 e_1 f_2 - d_1 e_2 f_2 + d_1 f_2 + d_2 e_1 f_1 + d_2 e_1 f_2 - d_2 e_1.$$

*pre* erlaubt es, die Variablen in abhängige und unabhängige zu unterteilen und die abhängigen durch die unabhängigen auszudrücken. Das entsprechende lineare Gleichungssystem lässt sich einfach lösen (MuPAD)

```
sol:=solve(pre);
```

$$\{[d_1 = f_1 - e_2 + f_2, d_2 = f_2, e_1 = f_1 - e_2 + f_2]\}$$

In die Behauptung eingesetzt erhalten wir

```
normal(subs(con,sol[1]));
```

0

woran wir erkennen, dass jede gemeinsame Nullstelle von *pre* auch eine Nullstelle von *con* ist.

**Aufgabe 12** Formulieren und untersuchen Sie auf dieselbe Weise den Satz des Pappus.

Wir sehen also, dass die algebraische Formulierung geometrischer Sätze normalerweise mit der universellen Realisierung einer konstruktiv erzeugten Konfiguration  $\mathcal{K}$  startet, zusätzliche geometrische Voraussetzungen in einer Menge  $pre = \{f_1, \dots, f_m\} \subset R$  von Polynomen kodiert, und die Behauptung ebenfalls in einem Polynom  $con = g \in R$  ausgedrückt werden kann, so dass für jede Variablenspezifikation  $X \rightarrow X_0$ , für die  $pre|_{X \rightarrow X_0} = 0$  gilt (und die evtl. noch einer Nichtdegenerationsbedingung  $\Psi|_{X \rightarrow X_0} \neq 0$  genügt), auch  $con|_{X \rightarrow X_0} = 0$  folgt.

## 4.1 Nichtlineare geometrische Bedingungen

Mit diesem Ansatz können wir auch nichtlineare geometrische Bedingungen erfassen. Diese Bedingungen treten auf, wenn sich schneidende geometrische Linien mehrere Schnittpunkte haben (Schnitt zweier Kreise, Schnitt von Kreis und Gerade) oder wenn es mehrere Geraden mit einer gegebenen Eigenschaft gibt (Parallelenpaar zu einer gegebenen Geraden in gegebenem Abstand, Winkelhalbierendenpaar eines gegebenen Winkels). Diese können ihrem Wesen nach nicht konstruktiver Natur sein, da nicht in jedem Schritt ein *eindeutig bestimmtes* geometrisches Objekt konstruiert wird.

Betrachten wir zunächst, wie sich die eben aufgelisteten geometrischen Bedingungen durch die Koordinaten der an ihnen beteiligten geometrischen Objekte beschreiben lassen.

### Punkte auf einem Kreis

Ähnlich wie eine Gerade kann man einen Kreis durch die Koordinaten seines Mittelpunkts  $M = (x_M, y_M)$  und eines Punkts  $P = (x_P, y_P)$  auf der Peripherie beschreiben. Statt des Radius  $r$  verwenden wir dessen Quadrat  $r^2 = (x_P - x_M)^2 + (y_P - y_M)^2$ , den *Sqradius*, um nicht bereits an dieser Stelle Wurzelausdrücke zu erzeugen.

Als Objekt der analytischen Geometrie wird ein solcher Kreis durch seine Gleichung

$$(x - x_M)^2 + (y - y_M)^2 - r^2 = (x^2 + y^2) + c_2 x + c_3 y + c_4$$

mit  $c_2 = -2x_0$ ,  $c_3 = -2y_0$ ,  $c_4 = x_0^2 + y_0^2 - r^2$  gegeben, die durch die drei Parameter  $c = (c_2, c_3, c_4)$  eindeutig charakterisiert werden kann. Umgekehrt lassen sich aus den Kreisparametern  $(c_2, c_3, c_4)$  auch Zentrum  $(x_M, y_M)$  und Sqradius  $r^2$  unmittelbar polynomial zurückgewinnen.

Im GEOPROVER-Paket stehen dafür die Funktionen

```
circle_center(c:Circle):Point und
circle_sqradius(c:Circle):Scalar
```

sowie der Konstruktor

```
pc_circle(M:Point, P:Point):Circle
```

zur Verfügung.

Bestimmen wir nun den Kreis, der durch drei vorgegebene Punkte  $A, B, C$  verläuft. Formeln für die Koordinaten des Mittelpunkts  $M = (m_1, m_2)$  dieses Kreises können wir aus den Gleichungen

```
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2); M:=Point(m1,m2);
polys:={sqrdist(M,A)-sqrdist(M,B), sqrdist(M,A)-sqrdist(M,C)};
```

bestimmen, wenn wir für  $A, B, C$  universelle Punkte einsetzen. Das entsprechende Gleichungssystem

$$\begin{aligned} 2 \cdot m_1 \cdot (-a_1 + b_1) + 2 \cdot m_2 \cdot (-a_2 + b_2) + a_1^2 + a_2^2 - b_1^2 - b_2^2 &= 0, \\ 2 \cdot m_1 \cdot (-a_1 + c_1) + 2 \cdot m_2 \cdot (-a_2 + c_2) + a_1^2 + a_2^2 - c_1^2 - c_2^2 &= 0 \end{aligned}$$

ist linear in  $(m_1, m_2)$  und besitzt eine eindeutige Lösung genau dann, wenn die zugehörige Determinante, die wieder einmal genau mit der Kollinearitätsbedingung  $\text{is\_collinear}(A, B, C)$  übereinstimmt, verschieden von 0 ist.

```
solve(polys, {m1, m2});
```

ergibt dann die gesuchten Mittelpunktskoordinaten, die wir nun fest als Prozedur zur weiteren Verwendung einbrennen können. Alternativ hätten wir den Mittelpunkt natürlich auch wie bisher berechnen können:

```
M:=intersection_point(p_bisector(B,C), p_bisector(C,A));
```

Für den Kreis durch drei vorgegebene Punkte  $A, B, C$  erhalten wir

$$\left( \begin{array}{c} \frac{-a_1^2 b_2 + a_1^2 c_2 - a_2^2 b_2 + a_2^2 c_2 + a_2 b_1^2 + a_2 b_2^2 - a_2 c_1^2 - a_2 c_2^2 - b_1^2 c_2 - b_2^2 c_2 + b_2 c_1^2 + b_2 c_2^2}{D}, \\ \\ \frac{a_1^2 b_1 - a_1^2 c_1 - a_1 b_1^2 - a_1 b_2^2 + a_1 c_1^2 + a_1 c_2^2 + a_2^2 b_1 - a_2^2 c_1 + b_1^2 c_1 - b_1 c_1^2 - b_1 c_2^2 + b_2^2 c_1}{D}, \\ \\ \frac{-a_1^2 b_1 c_2 + a_1^2 b_2 c_1 + a_1 b_1^2 c_2 + a_1 b_2^2 c_2 - a_1 b_2 c_1^2 - a_1 b_2 c_2^2 - a_2^2 b_1 c_2 + a_2^2 b_2 c_1 - a_2 b_1^2 c_1 + a_2 b_1 c_1^2 + a_2 b_1 c_2^2 - a_2 b_2^2 c_1}{D} \end{array} \right)$$

mit  $D = a_1 b_2 - a_1 c_2 - a_2 b_1 + a_2 c_1 + b_1 c_2 - b_2 c_1 = \text{is\_collinear}(A, B, C)$ . Die Parameter dieses Kreises sind rationale Funktionen der Parameter der Punkte  $A, B, C$ , wobei im Nenner wiederum ein Polynom  $D$  auftritt. Um solche Nenner zu vermeiden, wollen wir ähnlich wie für

Geraden auch für Kreise homogene Koordinaten verwenden, d.h. einen Kreis durch ein Tupel  $c = (c_1 : c_2 : c_3 : c_4)$  charakterisieren, das für die Punktmenge

$$\{(x, y) : c_1(x^2 + y^2) + c_2 x + c_3 y + c_4 = 0\}$$

steht.

Wir wollen als nächstes charakterisieren, wann 4 Punkte  $A, B, C, D$  auf einer gemeinsamen Kreislinie liegen. Mit den oben berechneten Koordinaten des Umkreismittelpunkts  $M$  von  $ABC$  ergibt sich die gesuchte Bedingung aus

```
D_:=Point(d1,d2);
con:=sqrdist(M,A)-sqrdist(M,D_);
```

als Zähler  $p_1 = \text{numer}(\text{con})$  der (gekürzten) rationalen Funktion  $\text{con}$ . Dieses Zählerpolynom  $p$  ist ein Polynom vierten Grades, das die entsprechende Bedingung beschreibt. Wir können auch dieses Polynom extrahieren und als Prozedur `is_concyclic` zur weiteren Verwendung in unsere Programmsammlung aufnehmen.

**Satz 14** Vier Punkte  $A = (a_1, a_2), B = (b_1, b_2), C = (c_1, c_2), D = (d_1, d_2)$  liegen genau dann auf einem gemeinsamen Kreis, wenn

$$\begin{aligned} & a_1^2 b_1 c_2 - a_1^2 b_1 d_2 - a_1^2 b_2 c_1 + a_1^2 b_2 d_1 + a_1^2 c_1 d_2 - a_1^2 c_2 d_1 - a_1 b_1^2 c_2 + a_1 b_1^2 d_2 - \\ & a_1 b_2^2 c_2 + a_1 b_2^2 d_2 + a_1 b_2 c_1^2 + a_1 b_2 c_2^2 - a_1 b_2 d_1^2 - a_1 b_2 d_2^2 - a_1 c_1^2 d_2 - a_1 c_2^2 d_2 + \\ & a_1 c_2 d_1^2 + a_1 c_2 d_2^2 + a_2^2 b_1 c_2 - a_2^2 b_1 d_2 - a_2^2 b_2 c_1 + a_2^2 b_2 d_1 + a_2^2 c_1 d_2 - a_2^2 c_2 d_1 + \\ & a_2 b_1^2 c_1 - a_2 b_1^2 d_1 - a_2 b_1 c_1^2 - a_2 b_1 c_2^2 + a_2 b_1 d_1^2 + a_2 b_1 d_2^2 + a_2 b_2^2 c_1 - a_2 b_2^2 d_1 + \\ & a_2 c_1^2 d_1 - a_2 c_1 d_1^2 - a_2 c_1 d_2^2 + a_2 c_2^2 d_1 - b_1^2 c_1 d_2 + b_1^2 c_2 d_1 + b_1 c_1^2 d_2 + b_1 c_2^2 d_2 - \\ & b_1 c_2 d_1^2 - b_1 c_2 d_2^2 - b_2^2 c_1 d_2 + b_2^2 c_2 d_1 - b_2 c_1^2 d_1 + b_2 c_1 d_1^2 + b_2 c_1 d_2^2 - b_2 c_2^2 d_1 \end{aligned}$$

gilt.

Diese Bedingung kann man auch als Determinante schreiben: Die vier Punkte  $A, B, C, D$  liegen genau dann auf einem gemeinsamen Kreis, wenn es eine nichttriviale Lösung  $x = (x_1, x_2, x_3, x_4)$  des Gleichungssystems

$$\begin{aligned} x_1(a_1^2 + a_2^2) + x_2 a_1 + x_3 a_2 + x_4 &= 0 \\ x_1(b_1^2 + b_2^2) + x_2 b_1 + x_3 b_2 + x_4 &= 0 \\ x_1(c_1^2 + c_2^2) + x_2 c_1 + x_3 c_2 + x_4 &= 0 \\ x_1(d_1^2 + d_2^2) + x_2 d_1 + x_3 d_2 + x_4 &= 0 \end{aligned}$$

gibt, d.h. wenn

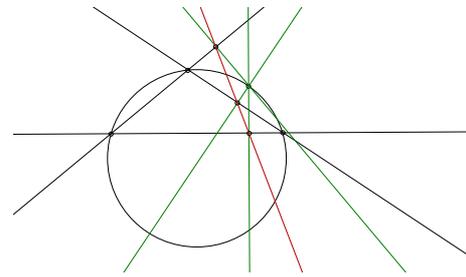
$$\begin{vmatrix} (a_1^2 + a_2^2) & a_1 & a_2 & 1 \\ (b_1^2 + b_2^2) & b_1 & b_2 & 1 \\ (c_1^2 + c_2^2) & c_1 & c_2 & 1 \\ (d_1^2 + d_2^2) & d_1 & d_2 & 1 \end{vmatrix} = 0$$

gilt.

Damit können wir bereits den **Satz von der Simsonschen Geraden** beweisen. Seien  $A, B, C, D$  vier Punkte wie eben betrachtet und

```
R:=pedalpoint(D,pp_line(B,C));
S:=pedalpoint(D,pp_line(A,C));
T:=pedalpoint(D,pp_line(A,B));
```

die Fußpunkte der Lote aus  $D$  auf die drei Seiten des Dreiecks  $ABC$  oder deren Verlängerungen. Der Satz von der Simsonschen Geraden behauptet, dass für einen Punkt  $D$  auf dem Umkreis von  $ABC$  die Lotfußpunkte  $R, S, T$  kollinear sind.



Simsonsche Gerade

Die Voraussetzung wird genau durch das oben hergeleitete Polynom  $p_1$  vierten Grades charakterisiert. Die Behauptung

```
con:=is_collinear(R,S,T);
```

ist eine rationale Funktion, deren Zähler ein Polynom 8. Grades mit 576 Summanden enthält und dessen Nenner aus dem Produkt

$$\text{sqrdist}(A, B) \cdot \text{sqrdist}(B, C) \cdot \text{sqrdist}(A, C)$$

besteht. Das Zählerpolynom lässt sich seinerseits in das Produkt

$$\text{collinear}(A, B, C)^2 \cdot p_1$$

zerlegen, woraus der geforderte Beweis nunmehr leicht abzuleiten ist. Aus dem Beweis folgt, dass nicht nur der Satz selbst richtig ist, sondern auch dessen Umkehrung gilt: Sind die Lotfußpunkte  $R, S, T$  kollinear, so liegt  $D$  auf dem Umkreis des Dreiecks  $ABC$ .

Auf ähnliche Weise kann man auch den folgenden Satz herleiten.

**Satz 15 (Satz des Ptolemäus)** *Das Viereck  $ABCD$  ist genau dann ein Sehnenviereck, wenn für die Seitenlängen  $p, q, r, s$  und die Diagonallängen  $t, u$  die Beziehung*

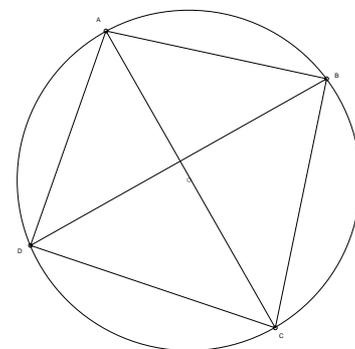
$$pr + qs = tu$$

*gilt.*

Dieser Satz macht in seiner Formulierung zum einen von der speziellen Lage der Punkte auf der Kreislinie Gebrauch (ansonsten kann man die Vierecksseiten nicht von den Diagonalen unterscheiden) und verwendet zum anderen mit den Seitenlängen Größen, die sich nicht polynomial durch die Koordinaten der Eckpunkte ausdrücken lassen. Eine zur Ausgangsbedingung äquivalente Aussage, die besser in unser Konzept passt, erhält man, wenn man die Behauptung wie folgt umformt:

$$(pr)^2 + (qs)^2 - (tu)^2 = 2pqrs$$

$$((pr)^2 + (qs)^2 - (tu)^2)^2 - (2pqrs)^2 = 0$$



Satz des Ptolemäus

Ausmultiplizieren liefert die Gleichung

$$p^4 r^4 - 2 p^2 q^2 r^2 s^2 - 2 p^2 r^2 t^2 u^2 + q^4 s^4 - 2 q^2 s^2 t^2 u^2 + t^4 u^4 = 0,$$

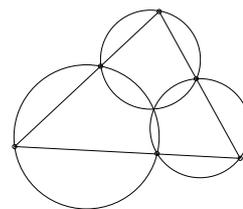
in der nur noch Quadrate von Seitenlängen vorkommen und die zudem nunmehr symmetrisch in den drei Paaren gegenüberliegender Strecken ist.

Ersetzen wir diese Quadrate durch die entsprechenden Abstandsquadrate (Rechnungen mit MuPAD)

```
pp:=(p*r)^2+(q*s)^2-(t*u)^2 - (2*p*q*r*s)^2;
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2); D_:=Point(d1,d2);
rules:={
p^2 = sqrdist(A,B), q^2 = sqrdist(B,C), r^2 = sqrdist(C,D_),
s^2 = sqrdist(D_,A), t^2 = sqrdist(A,C), u^2 = sqrdist(B,D_)};
p2:=normal(subs(pp,rules));
```

so erhalten wir ein Polynom  $p_2$  8. Grades mit 962 Summanden, das im Wesentlichen das Quadrat des Polynoms  $p_1$  ist ( $p_2 = -4 p_1^2$ ). MuPAD 3.1 braucht für die Faktorisierung von  $p_2$  mehr als 80 s.

Betrachten wir schließlich noch den **Satz vom Miquelschen Punkt**. Die zugehörige allgemeine geometrische Konfiguration (Dreieck  $ABC$  mit Punkten  $R, S, T$  auf den Dreiecksseiten) lässt sich wie folgt angeben:



Miquelscher Punkt

```
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2);
R:=varpoint(B,C,u1); S:=varpoint(A,C,u2); T:=varpoint(A,B,u3);
```

Zum Beweis des Satzes bestimmen wir (zur Vereinfachung wieder mit einem speziell positionierten Koordinatensystem) die Koordinaten des Schnittpunkts  $P = (x_1, x_2)$  zweier der Kreise und zeigen, dass dieser Punkt auch auf dem dritten Kreis liegt:

```
A:=Point(0,0); B:=Point(0,b2); C:=Point(c1,c2);
R:=varpoint(B,C,u1); S:=varpoint(A,C,u2); T:=varpoint(A,B,u3);
P:=Point(x1,x2);
polys:={is_concyclic(A,S,T,P), is_concyclic(B,R,T,P)};
sol:=solve(polys,{x1,x2},IgnoreSpecialCases);
```

Wir erhalten zwei Lösungen, deren Koordinaten jeweils rationale Funktionen in den Parametern sind. Eine der beiden Lösungen ist der Punkt  $T$ , von dem wir bereits wissen, dass er zu beiden Kreisen gehört. Der zweite Punkt ist der gesuchte Schnittpunkt  $P$  (seine rationalen Koordinaten haben 38 bzw. 17 Zählerterme und 42 Nennerterme). Er liegt auf dem dritten Kreis, wie die folgende Rechnung zeigt.

```
P0:=normal(subs(P,sol[1]));
is_concyclic(C,R,S,P0);
```

Es ist nicht verwunderlich, dass die Schnittpunktkoordinaten von  $P$  rationale Funktionen in den Parametern sind, wenn einer der beiden Schnittpunkte bekannt ist, da die Schnittpunktbedingung auf eine Gleichung zweiten Grades hinausläuft.

**Aufgabe 13** Finden Sie allgemein Formeln für folgende Funktionen:

- `other_cl_point(c:Circle,l:Line,P:Point):Point`, die zu gegebenem Kreis  $c$  und gegebener Gerade  $l$  den zweiten Schnittpunkt von  $c$  und  $l$  findet, wenn  $P$  der andere Schnittpunkt ist.
- `other_cc_point(c1:Circle,c2:Circle,P:Point):Point`, die zu zwei gegebenen Kreisen  $c_1$  und  $c_2$  den zweiten Schnittpunkt von  $c_1$  und  $c_2$  findet, wenn  $P$  der andere Schnittpunkt ist.

Im Gegensatz zu obigem Beispiel lassen sich im Allgemeinen die Schnittpunktkoordinaten zweier Kreise oder von Kreis und Gerade nicht rational durch die Ausgangskordinaten ausdrücken. So erhalten wir etwa für die Schnittpunkte zweier Kreise mit den Radien  $r_1$  und  $r_2$  und dem Mittelpunktsabstand  $2u$  als Bedingung an die Koordinaten des Schnittpunkts  $P = (x_1, x_2)$  das Gleichungssystem

```
P:=Point(x1,x2);
polys:={
  on_circle(P,pc_circle(Point(-u,0),Point(-u-r1,0))),
  on_circle(P,pc_circle(Point(u,0),Point(u+r2,0)))};
{-r1^2 + u^2 + 2u x1 + x1^2 + x2^2, -r2^2 + u^2 - 2u x1 + x1^2 + x2^2},
```

zu dessen Lösung

```
sol:=solve(polys,{x1,x2},IgnoreSpecialCases);
```

$$x_1 = \frac{r_1^2 - r_2^2}{4u}, \quad x_2 = \pm \frac{\sqrt{-r_1^4 + 2r_1^2 r_2^2 + 8r_1^2 u^2 - r_2^4 + 8r_2^2 u^2 - 16u^4}}{4u}$$

bereits Wurzelausdrücke erforderlich sind.

## Winkel und Winkelhalbierende

(Orientierte) Winkel  $\alpha$  sind als eigenständige Objekte im GEOPROVER-Paket 1.3 noch nicht enthalten und werden sinnvoll als Tripel von Punkten `Angle(A,B,C)` aufzufassen sein, wobei  $B$  der Scheitelpunkt ist und die Orientierung durch  $\angle(\overrightarrow{BA}, \overrightarrow{BC})$  gegeben wird. Davon zu unterscheiden ist das **Winkelmaß**  $m(\alpha) = \text{p3\_angle}(A,B,C)$  als Maß des orientierten Winkels zwischen den Geraden  $g = BA$  und  $h = BC$ .

Als Maß werden wir den Tangens der Winkelgröße verwenden. Zwischen diesem und den Anstiegen  $\tan(\alpha_g) = -g_1/g_2$  bzw.  $\tan(\alpha_h) = -h_1/h_2$  der Geraden  $g = (g_1, g_2, g_3)$  und  $h = (h_1, h_2, h_3)$  besteht folgender Zusammenhang:

$$\tan(\angle(g, h)) = \tan(\alpha_h - \alpha_g) = \frac{\tan(\alpha_h) - \tan(\alpha_g)}{1 + \tan(\alpha_h) \tan(\alpha_g)} = \frac{g_1 h_2 - g_2 h_1}{g_1 h_1 + g_2 h_2}$$

Dieser Ausdruck ist nicht definiert, wenn der Nenner verschwindet, also die beiden Geraden  $g$  und  $h$  senkrecht aufeinander stehen. Wir können den Wert ähnlich den Teilverhältnissen auch als Element  $(g_1 h_2 - g_2 h_1) : (g_1 h_1 + g_2 h_2) \in \mathbb{P}^1$  interpretieren.

Die Bedingung, dass ein Punkt  $P(x_1, x_2)$  auf einer der beiden Winkelhalbierenden liegt, ergibt sich damit aus folgender Rechnung

```
A:=Point(a1,a2); B:=Point(b1,b2); C:=Point(c1,c2);
normal(p3_angle(A,B,P) - p3_angle(P,B,C));
```

als universelle rationale Formel, von der wiederum nur das Verschwinden des Zählers von Interesse ist. Die entsprechende Bedingung ist ein Polynom vom Grad 2 in  $x_1$  und  $x_2$ , das für allgemeine Punkte  $A, B, C$  56 Terme enthält.

Dieser Ausdruck beschreibt zwei zueinander senkrechte Geraden  $w_1$  und  $w_2$ , deren Parameter sich allerdings, wie die Koordinaten der Schnittpunkte zweier Kreise, im Allgemeinen nicht rational durch die Koordinaten der Punkte  $A, B, C$  ausdrücken lassen. Legen wir mit  $B:=\text{Point}(0,0)$  den Scheitel des Winkels in den Ursprung, so vereinfacht sich das Polynom  $z$  zu

$$(a_2 c_1 + c_2 a_1) x_1^2 + (2 a_2 c_2 - 2 a_1 c_1) x_2 x_1 + (-c_2 a_1 - a_2 c_1) x_2^2,$$

und für die Koordinaten eines Punktes  $P$  auf der Winkelhalbierenden ergibt sich

```
solve(z,x1,IgnoreSpecialCases);
```

$$x_1 = \frac{a_1 c_1 - a_2 c_2 \pm \sqrt{a_2^2 c_2^2 + a_1^2 c_1^2 + a_2^2 c_1^2 + c_2^2 a_1^2}}{a_2 c_1 + c_2 a_1} x_2.$$

Das GEOPROVER-Paket enthält deshalb nur eine Prozedur `on_bisector(P,A,B,C)`, mit der geprüft werden kann, ob ein gegebener Punkt  $P$  auf einer der beiden Winkelhalbierenden liegt.

**Aufgabe 14** Vier Punkte  $A, B, C, D$  liegen nach der Umkehrung des Peripheriewinkelsatzes auf einem Kreis, wenn  $B$  und  $D$  auf demselben Bogen  $\widehat{AC}$  liegen und  $|\angle ABC| = |\angle ADC|$  gilt.

Zeigen Sie, dass diese Bedingung genau auf das oben hergeleitete Polynom vierten Grades führt.

Diskutieren Sie den Fall, dass  $B$  und  $D$  auf unterschiedlichen Bögen  $\widehat{AC}$  liegen. (Hinweis: Beachten Sie, dass  $\tan(-\alpha) = -\tan(\alpha)$  gilt, da mit orientierten Winkeln gerechnet wird.)

Um trotzdem den **Satz vom Schnittpunkt der Winkelhalbierenden** zu beweisen, können wir wieder zunächst die Koordinaten des Schnittpunkts  $P = (x_1, x_2)$  von zwei der drei Winkelhalbierenden bestimmen und dann prüfen, ob  $P$  auch auf der dritten Winkelhalbierenden liegt. Zur Vereinfachung der Rechnung fixieren wir ein spezielles Koordinatensystem, in dem die Eckpunkte des Dreiecks die folgenden Koordinaten haben:

```
A:=Point(1,0); B:=Point(0,0); C:=Point(c1,c2); P:=Point(x1,x2);
```

```
polys:={on_bisector(P,B,A,C), on_bisector(P,C,B,A)};
```

Zur Bestimmung der Koordinaten von  $P$  erhalten wir das Gleichungssystem

$$\begin{aligned} -x_1^2 c_2 + 2 x_2 c_1 x_1 + c_2 x_2^2 &= 0 \\ -2 x_2 + 2 x_2 c_1 + 2 x_2 x_1 - 2 x_2 c_1 x_1 - c_2 x_2^2 + c_2 - 2 x_1 c_2 + x_1^2 c_2 &= 0 \end{aligned}$$

aus zwei Gleichungen zweiten Grades, für welche sowohl Maple (9.5) als auch MuPAD in der Standardeinstellung nunmehr die Lösung in RootOf-Notation zurückgeben. Hier das Ergebnis von Maple:

```
sol:=solve(polys,{x1,x2});
P1:=subs(sol,P);
```

$$P_1 = \left[ \frac{c_2 - 2 U_1 + 2 c_1 U_1}{2(c_2 - U_1)}, U_1 \right]$$

mit

$$U_1 = \text{RootOf} \left( \left( \begin{array}{l} 4c_2 Z^4 + (-8c_1^2 - 8c_2^2 + 8c_1) Z^3 + \\ (-4c_1 c_2 + 4c_1^2 c_2 - 4c_2 + 4c_2^3) Z^2 + 4c_2^2 Z - c_2^3 \end{array} \right), Z \right)$$

Hinter  $U_1$  verbergen sich vier Lösungen, deren einzelne Untersuchung in dieser Form schwierig ist. Expandiert man den ROOTOF-Ausdruck, so liefert Maple eine (noch) übersichtliche Form mit vier doppelt geschachtelten Wurzelausdrücken

```
sol1:=allvalues(sol);
PValues:=map(u->simplify(subs(u,P)),sol1);
```

während MuPAD (2.5) diese spezielle Form nicht erkennt und vier Lösungen mit je etwa 70 000 Termen produziert. Allerdings kann man in der gegenwärtigen solve-Ausgabestruktur von MuPAD daraus nicht mehr ohne Weiteres die Koordinatenwerte für  $P$  extrahieren.

```
sol:=solve(polys,{x1,x2},IgnoreSpecialCases,MaxDegree=4);
```

Aus der Maple-Rechnung können wir zum Beweis des Satzes über die Schnittpunkte der Winkelhalbierenden nur die vier für  $P$  berechneten Werte in die Bedingung `on_bisector(P,A,C,B)` einsetzen und überprüfen, ob der so entstehende (nicht mehr nur rationale) Ausdruck zu Null vereinfacht:

```
map(u->on_bisector(u,A,C,B),PValues);
```

[0, 0, 0, 0]

Allerdings lassen sich geschachtelte Wurzeln nicht immer so problemlos vereinfachen. Außerdem steht eine Interpretation des Ergebnisses ähnlich dem Mechanisierungssatz für Geometriesätze vom konstruktiven Typ noch aus. Die Rechnungen verlassen den Polynomring  $R$  in Richtung algebraischer Erweiterungen.

Ein ähnliches Ergebnis erhält man mit Mathematica:

```
C//Unprotect;
A=point[1,0]; B:=point[0,0]; C:=point[c1,c2]; P:=point[x1,x2];
polys={onbisector[P,B,A,C], onbisector[P,C,B,A]}
sol=Solve[#==0 & /@ polys,{x1,x2}]
onbisector[P/. sol[[#]],A,C,B]& /@ {1,2,3,4}
```

Bei der Formulierung sind zwei Besonderheiten zu beachten: (1) Mathematica verlangt in der Solve-Funktion eine Liste von Gleichungen als erstes Argument, so dass in der Liste `polys` jedes Polynom  $f$  durch  $f = 0$  zu ersetzen ist, und (2) wurde die Infix-Notation `/@` von `Map` sowie `/.` von `Subs` an mehreren Stellen verwendet.

Es ergibt sich nun die Frage, ob auch ohne explizite Auflösung der Gleichung für  $x_2$  in Wurzelausdrücke nachgewiesen werden kann, dass jede der vier Lösungen auch Nullstelle der Gleichung

```
on_bisector(P,A,C,B);
```

$$2c_1^3 x_2 - c_1^3 - 2c_1^2 c_2 x_1 - 2c_1^2 x_1 x_2 + 2c_1^2 x_1 + 2c_1 c_2^2 x_2 - c_1 c_2^2 + 2c_1 c_2 x_1^2 - 2c_1 c_2 x_2^2 - c_1 x_1^2 + c_1 x_2^2 - 2c_2^3 x_1 + 2c_2^2 x_1 x_2 + 2c_2^2 x_1 - 2c_2 x_1 x_2$$

ist. Derartige Fragen wollen wir im nächsten Abschnitt genauer formulieren und studieren.

Hier sei noch angemerkt, dass man Winkelhalbierende mit einem anderen Konzept auch als geometrische Objekte vom konstruktiven Typ einführen kann.

**Aufgabe 15** Zeigen Sie auf folgende Weise, dass sich die Winkelhalbierenden in einem Punkt schneiden:

$A$  und  $B$  seien gegeben und  $D$  sei der Schnittpunkt der Winkelhalbierenden aus  $A$  und  $B$ . Dann ist  $C$  der Schnittpunkt der Geraden, die aus  $AB$  durch Spiegelung an  $AD$  bzw.  $BD$  hervorgehen.  $CD$  ist genau dann Winkelhalbierende, wenn  $AC$  bei Spiegelung an  $CD$  in  $BC$  übergeht.

Sie können dazu die GEOPROVER-Funktion `sym_line` verwenden.

## 4.2 Geometrische Sätze vom Gleichungstyp

Alle unsere Geometrietheoreme hatten die folgende Gestalt: Gegeben ist

- eine konstruktiv gegebene (rationale oder polynomiale) geometrische Konfiguration  $\mathcal{K}$  mit  $\overline{O}$  als universeller Realisierung der Startkonfiguration und daraus abgeleitet universelle Formeln für die Endkonfiguration  $\overline{O}_E = \mathcal{K}(\overline{O})$ , was wir im Weiteren als *allgemeine geometrisch Konfiguration* (AGK) bezeichnen;
- eine Menge von polynomialen geometrischen Bedingungen mit Polynomen

$$F = \{f_1, \dots, f_m\} \subset R$$

in den Variablen von  $\overline{O}$  als universellen Formeln, die weitere implizite Abhängigkeiten zwischen den Elementen der Konfiguration  $\Gamma_E$  beschreiben und die wir als *allgemeine geometrische Voraussetzungen* (AGV) bezeichnen;

- eine (oft nicht explizit bekannte) polynomiale Nichtdegenerationsbedingung mit universeller Formel  $h \in R$  für die Gültigkeit der Schlussfolgerung, die wenigstens die Ausführbarkeit von  $\mathcal{K}$  garantiert, sowie
- eine polynomiale geometrische Bedingung mit universeller Formel  $g \in R$ , welche die Folgerung des Satzes algebraisch kodiert.

**Definition 11** Als Geometriesatz vom Gleichungstyp bezeichnen wir dann eine Aussage der folgenden Gestalt:

Jede durch eine Substitution  $X = \mathbf{a}$  erzeugte spezielle Realisierung  $O_E = \overline{O}_E(X = \mathbf{a})$ , die nicht degeneriert ist, für welche also  $h(\mathbf{a}) \neq 0$  gilt, und welche die AGV erfüllt, also  $\mathbf{a}$  eine gemeinsame Nullstelle der Polynome  $f \in F$  ist, beschreibt auch eine Nullstelle von  $g$ , d. h. erfüllt die geometrische Folgerung.

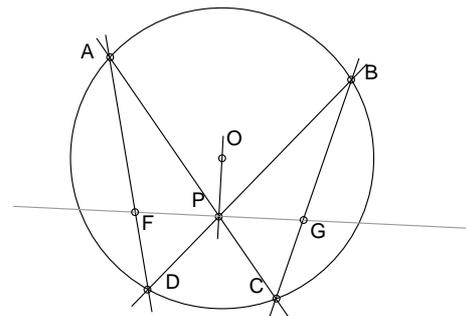
$$\forall \mathbf{a} = (a_1, \dots, a_n) \in K^n \left( \bigwedge_{f \in F} f(\mathbf{a}) = 0 \right) \wedge (h(\mathbf{a}) \neq 0) \Rightarrow (g(\mathbf{a}) = 0)$$

Wir schreiben für einen solchen Satz auch kurz  $(F/h \Rightarrow g)$ .

Betrachten wir zum Abschluss dieses Punktes einen weiteren geometrischen Satz, den wir als Satz vom Gleichungstyp formulieren können. Es handelt sich um das folgende

**Schmetterlingstheorem** [2, S. 31]:

Auf einer Kreislinie mit dem Mittelpunkt  $O$  seien vier Punkte  $A, B, C, D$  gegeben.  $P$  sei der Schnittpunkt von  $AC$  und  $BD$ . Die Senkrechte auf der Verbindungslinie  $OP$  schneide die Seiten  $AD$  und  $BC$  (die ‘‘Schmetterlingsflügel’’) in den Punkten  $F$  und  $G$ . Dann gilt  $|PF| = |PG|$ .



Wir wollen bei der Herleitung einer algebraischen Form dieses Satzes zugleich zwischen (in einem später zu präzisierenden Sinne) “unabhängigen” Variablen  $u_i$ , deren Wert frei wählbar ist, und “abhängigen” Variablen  $x_i$ , deren Wert sich eindeutig oder endlich vieldeutig aus den algebraischen Bedingungen und bereits fixierten Werten ergibt, unterscheiden.

Sei dazu  $O$  der Koordinatenursprung und  $A = (u_4, 0)$  auf der  $x$ -Achse gelegen.  $B, C$  und  $D$  seien drei weitere Punkte auf dem Kreis. Je eine Koordinate kann frei gewählt werden, die andere ergibt sich aus der Kreisgleichung. Die Punkte seien also  $B = (u_1, x_1)$ ,  $C = (u_2, x_2)$  und  $D = (u_3, x_3)$ .  $P = (x_4, x_5)$  ergibt sich dann eindeutig als Schnittpunkt, ebenso wie die Senkrechte  $h$  und die Punkte  $F = (x_6, x_7)$  und  $G = (x_8, x_9)$ . Insgesamt erhalten wir folgende algebraische Formulierung

```
unprotect(O): unprotect(D):
O:=Point(0,0); A:=Point(u4,0);
B:=Point(u1,x1); C:=Point(u2,x2); D:=Point(u3,x3);
P:=Point(x4,x5);
F:=Point(x6,x7); G:=Point(x8,x9);
c0:=pc_circle(O,A);
h:=ortho_line(P,pp_line(O,P));

polys:= { on_circle(B,c0), on_circle(C,c0), on_circle(D,c0),
          is_collinear(B,D,P), is_collinear(A,C,P),
          on_line(F,h), is_collinear(A,D,F),
          on_line(G,h), is_collinear(B,C,G)};

con:=sqrdist(P,F)-sqrdist(P,G);
```

$$\begin{aligned}
polys = \{ & u_1^2 + x_1^2 - u_4^2, u_2^2 + x_2^2 - u_4^2, u_3^2 + x_3^2 - u_4^2, \\
& u_1 x_3 - u_1 x_5 - u_3 x_1 + u_3 x_5 + x_4 x_1 - x_4 x_3, u_4 x_2 - u_4 x_5 + u_2 x_5 - x_4 x_2, \\
& - x_4 x_6 - x_5 x_7 + x_5^2 + x_4^2, u_4 x_3 - u_4 x_7 + u_3 x_7 - x_6 x_3, \\
& - x_4 x_8 - x_5 x_9 + x_5^2 + x_4^2, u_1 x_2 - u_1 x_9 - u_2 x_1 + u_2 x_9 + x_8 x_1 - x_8 x_2 \} \\
con = & x_6^2 - 2x_4 x_6 + x_7^2 - 2x_5 x_7 - x_8^2 + 2x_4 x_8 - x_9^2 + 2x_5 x_9
\end{aligned}$$

Das Gleichungssystem hat eine recht überschaubare Struktur. Die Anzahl der abhängigen Variablen  $x_i$  entspricht genau der Anzahl der Gleichungen in der Menge  $polys$ , die linear in  $x_4, \dots, x_9$  und quadratisch in  $x_1, x_2, x_3$  sind. Maple 9 liefert entsprechend eine „Lösung“ mit drei ROOTOF-Symbolen

```
vars:={seq(cat(x,i),i=1..9)};
sol:=[solve(polys,vars)];
nops(sol);
```

mit der es aber diesmal nichts mehr anfangen kann:

```
map(u->normal(subs(u,con)),sol);
```

Fordert man explizite Lösungen für die  $x_i, i = 1, \dots, 9$  als Ausdrücke in den  $u_i, i = 1, 2$ , die sich in diesem Fall als einfache Quadratwurzeln berechnen lassen, so liefert Maple zwar 8 Lösungen zurück, kann aber auch mit ihnen nichts mehr anfangen, selbst wenn der mächtige Simplifikationsalgorithmus `evala` für algebraische Ausdrücke angeworfen wird.

```
sol1:=map(allvalues,sol);
```

```
nops(sol1);
map(u->normal(subs(u,con)),sol1);
map(u->evala(subs(u,con)),sol1);
```

Der lineare Teil des Gleichungssystems kann „per Hand“ gelöst, aber auch durch eine stärker konstruktiv formulierte AGK eliminiert werden. Wählen wir etwa  $F$  und  $G$  als Geradengleiter und  $P$  als Schnittpunkt der Geraden  $AC$  und  $BD$  sowie  $F$  und  $G$  als Schnittpunkte mit  $h$ , so benötigen wir nur 3 Gleichungen, um das Problem zu formulieren

```
# ==> Butterfly_b
O:=Point(0,0); A:=Point(u4,0);
B:=Point(u1,x1); C:=Point(u2,x2); D:=Point(u3,x3);

P:=intersection_point(pp_line(A,C), pp_line(B,D));
c0:=pc_circle(O,A);
h:=ortho_line(P,pp_line(O,P));
F:=intersection_point(pp_line(A,D),h);
G:=intersection_point(pp_line(B,C),h);

polys:= { on_circle(B,c0), on_circle(C,c0), on_circle(D,c0)};

con:=numer(sqrdist(P,F)-sqrdist(P,G));
```

Allerdings verlagert sich damit das Problem nur, denn der Ausdruck für  $con$  ist nun deutlich komplizierter geworden und enthält (nach Maple-Zählung)  $\text{length}(con) = 548\,465$  Terme.

```
vars:={x1,x2,x3};
sol:=[solve(polys,vars)];
nops(sol);
map(u->normal(subs(u,con)),sol);
sol1:=map(allvalues,sol);
nops(sol1);
map(u->normal(subs(u,con)),sol1);
```

Maple kann auch in dieser einfachen Formulierung die Behauptung nicht verifizieren, obwohl die Lösungsmenge  $sol_1$  nur einfache Quadratwurzeln enthält. Erst der stärkere Simplifikationsoperator `evala` vereinfacht das Polynom, welches die geometrische Behauptung kodiert, zu Null, wobei die Variante `sol` deutlich schneller abgearbeitet wird als die Variante `sol1`.

Dasselbe gilt für Mathematica und erst recht MuPAD. Einzig Reduce kann eine leicht modifizierte Algebraisierung mit 5 Gleichungen ( $F$  und  $G$  werden als Geradengleiter auf  $AD$  bzw.  $BC$  angesetzt, siehe `Butterfly_c`) in vernünftiger Zeit zu einem Ende bringen.

Das Schmetterlingstheorem lässt allerdings auch eine konstruktive Formulierung zu, wenn wir zur Definition der Punkte  $B, C, D$  eine (rationale) Parametrisierung des Kreises verwenden. Diese ergibt sich, wenn wir durch einen bekannten Punkt auf der Kreislinie Sekanten zeichnen und die Koordinaten des jeweils zweiten Schnittpunkts in Abhängigkeit vom Anstieg dieser Sekante ausdrücken. Wir hatten das weiter oben bereits am Beispiel der Funktion `other_cl_point` besprochen. Für den Einheitskreis  $x^2 + y^2 = 1$  und den Referenzpunkt  $(-1, 0)$  ergibt sich für die Koordinaten eines Kreisgleiters folgende Formel

```
O:=Point(0,0);
A:=Point(-1,0);
kreis:=pc_circle(O,A);
line:=pp_line(A,Point(0,r));
other_cl_point(A,kreis,line);
```

$$\left[ \frac{r^2 - 1}{r^2 + 1}, \frac{2r}{r^2 + 1} \right]$$

Sind  $O$  und  $A$  Punkte in allgemeiner Lage, so ergeben sich nur leicht kompliziertere Formeln, die in der GEOPROVER-Funktion `circle_slider(O,A,u)` implementiert sind. Das entsprechende algebraische Problem kann mit dem GEOPROVER-Paket und jedem der Target-CAS gelöst werden.

```
O:=Point(0,0);
A:=Point(u4,0);

B:=circle_slider(O,A,u1);
C:=circle_slider(O,A,u2);
D:=circle_slider(O,A,u3);
P:=intersection_point(pp_line(A,C), pp_line(B,D));
c0:=pc_circle(O,A);
h:=ortho_line(P,pp_line(O,P));
F:=intersection_point(pp_line(A,D),h);
G:=intersection_point(pp_line(B,C),h);

con:=normal(sqrdist(P,F)-sqrdist(P,G));
```

### 4.3 Geometrietheoreme vom linearen Typ

Unsere generelle Vorgehensweise zum Beweis geometrischer Sätze vom Gleichungstyp wird sein, mögliche Belegungen der Variablen der universellen Realisierung der Startkonfiguration „schrittweise“ zu denken. Wir werden zunächst einen Teil von ihnen, die „unabhängigen Variablen“, mit beliebigen (zulässigen) Werten belegen, und daraus die (endlich vielen) Belegungen für die verbleibenden „abhängigen“ Variablen, welche die zusätzlichen Gleichungen  $f \in F$  respektieren, bestimmen.

Untersuchen wir zunächst den einfachsten Fall von Geometrietheoremen vom Gleichungstyp – solche, die auf lineare Gleichungssysteme in den abhängigen Variablen führen. Sätze vom konstruktiven Typ lassen sich als solche Geometrietheoreme formulieren und meist gilt auch die umgekehrte Aussage.

Betrachten wir als Beispiel den **Satz von Desargue**:

```
A:=Point(u1,u2); B:=Point(u3,u4); C:=Point(u5,u6);
P:=Point(u7,u8); Q:=Point(u9,x1); R:=Point(x2,x3);

polys:={is_parallel(pp_line(A,C),pp_line(P,R)),
        is_parallel(pp_line(B,C),pp_line(Q,R)),
        is_parallel(pp_line(A,B),pp_line(P,Q))};

con:= is_concurrent(pp_line(A,P),pp_line(B,Q),pp_line(C,R));
```

Die drei Gleichungen zur Bestimmung der abhängigen Parameter sind linear in  $x_1, x_2, x_3$ . Als Lösung ergeben sich rationale Funktionen in  $\mathbf{u}$ , die, in die Behauptung *con* eingesetzt, diese zu Null vereinfachen (MuPAD).

```
sol:=solve(polys,{x1,x2,x3});
map(sol,u->normal(subs(con,u)));
```

$$\begin{aligned}
 x_1 &= \frac{u_2 u_7 - u_1 u_8 - u_2 u_9 + u_3 u_8 - u_4 u_7 + u_4 u_9}{u_3 - u_1} \\
 x_2 &= \frac{u_1 u_9 - u_3 u_7 + u_5 u_7 - u_5 u_9}{u_1 - u_3} \\
 x_3 &= \frac{u_2 u_7 - u_1 u_8 - u_2 u_9 + u_3 u_8 - u_6 u_7 + u_6 u_9}{u_3 - u_1}
 \end{aligned}$$

Bei der Auswahl der unabhängigen und abhängigen Parameter sind wir dem konstruktiven Ansatz gefolgt. Statt dessen könnte man auch versuchen, drei andere der Parameter als abhängig zu deklarieren, etwa

```
P:=Point(u7,x1); Q:=Point(u8,x2); R:=Point(u9,x3);
polys1:={ is_parallel(pp_line(A,C),pp_line(P,R)),
           is_parallel(pp_line(B,C),pp_line(Q,R)),
           is_parallel(pp_line(A,B),pp_line(P,Q))};
```

Auch hier erhalten wir drei Gleichungen, die linear in den als abhängig deklarierten Variablen  $x_1, x_2, x_3$  sind. Allerdings ist die entsprechende Lösungsmenge leer:

```
sol:=solve(polys1,{x1,x2,x3});
```

Geht man etwas vorsichtiger zu Werke, löst zunächst die beiden ersten Gleichungen nach  $x_1$  und  $x_2$  auf und setzt diese Werte dann in die dritte Gleichung ein, um diese schließlich nach  $x_3$  aufzulösen, so erkennt man, dass dort  $x_3$  bereits nicht mehr enthalten ist, sondern eine Gleichung vierten Grades in  $\mathbf{u}$  entstanden ist, die Parameter also nicht unabhängig sind wie angenommen.

```
sol:=solve({op(polys1,1..2)},{x1,x2});
normal(subs(polys[3],sol[1]));
```

$$\frac{\begin{pmatrix} u_2 u_7 u_3^2 + u_4 u_5^2 u_8 + u_1^2 u_6 u_9 - u_6 u_7 u_3^2 + u_3^2 u_6 u_9 - u_1^2 u_4 u_9 - u_2 u_9 u_3^2 - u_4 u_7 u_5^2 + \\ u_4 u_1^2 u_8 - u_6 u_1^2 u_8 + u_2 u_7 u_5^2 - 2 u_4 u_1 u_8 u_5 + u_6 u_1 u_8 u_3 - u_2 u_5^2 u_8 + u_1 u_6 u_5 u_8 - \\ u_3 u_6 u_5 u_8 - u_2 u_3 u_1 u_8 + u_2 u_5 u_1 u_8 + u_2 u_5 u_3 u_8 - u_3 u_1 u_4 u_7 + u_3 u_4 u_7 u_5 - \\ 2 u_3 u_2 u_7 u_5 + u_3 u_1 u_6 u_7 - 2 u_3 u_1 u_6 u_9 + u_3 u_1 u_2 u_9 + u_5 u_1 u_4 u_7 - u_5 u_1 u_6 u_7 - \\ u_5 u_1 u_2 u_9 + u_5 u_6 u_7 u_3 + u_5 u_2 u_9 u_3 + u_1 u_3 u_4 u_9 + u_1 u_5 u_4 u_9 - u_3 u_5 u_4 u_9 \end{pmatrix}}{(u_3 - u_5)(u_1 - u_5)}$$

Untersuchen wir nun allgemein, was man über Geometrietheoreme vom Gleichungstyp aussagen kann, in denen die abhängigen Parameter  $\mathbf{x}$  linear von den unabhängigen Parametern  $\mathbf{u}$  abhängen. Solche Geometrietheoreme wollen wir als *Sätze vom linearen Typ* bezeichnen.

Das entsprechende Gleichungssystem hat die Gestalt

$$\sum_{j=1}^m c_{ij}(\mathbf{u})x_j + c_{i0}(\mathbf{u}) = 0, \quad i = 1, \dots, m$$

Seine Lösbarkeit hängt ganz wesentlich von der Determinante

$$D(\mathbf{u}) := \det|c_{ij}(\mathbf{u})|_{i,j=1,\dots,m}$$

ab. Genauer hat das Gleichungssystem für jede Belegung der Parameter  $\mathbf{u}$ , für die  $D(\mathbf{u}) \neq 0$  gilt, eine eindeutige Lösung, die man nach der Cramerschen Regel in der Form  $x_i = x_i(\mathbf{u}) = D_i(\mathbf{u})/D(\mathbf{u}) \in k(\mathbf{u})$  ausdrücken kann.

Für obiges Beispiel lässt sich die Koeffizientenmatrix wie folgt extrahieren (MuPAD):

```
vars:=[x1,x2,x3];
l:=map([op(polys)],p->map(vars,u->coeff(p,u,1)));
KoeffMat:=Dom::Matrix()(l);
```

$$\begin{pmatrix} u_3 - u_1 & 0 & 0 \\ u_3 - u_5 & u_4 - u_6 & -u_3 + u_5 \\ 0 & u_2 - u_6 & -u_1 + u_5 \end{pmatrix}$$

Deren Determinante ist ein Polynom mit der Faktorzerlegung

`factor(linalg::det(KoeffMat));`

$$(u_1 - u_3)(u_1 u_4 - u_2 u_3 - u_1 u_6 + u_2 u_5 + u_3 u_6 - u_4 u_5)$$

Der zweite Faktor kodiert die geometrische Nichtdegenerationsbedingung `is_collinear(A,B,C)` und ließ sich in obiger Formel `sol` offensichtlich herauskürzen. Der erste Faktor kodiert die Nichtdegenerationsbedingung  $u_1 \neq u_3$ , denn aus  $u_1 = u_3$  folgt zwingend  $u_7 = u_9$ , so dass in diesem Fall die „unabhängigen“ Variablen nicht mehr unabhängig belegt werden können.

Für  $D(\mathbf{u}) \neq 0$  erhalten wir universelle Formeln  $x_i(\mathbf{u})$  für die abhängigen Variablen, die wir wie im Fall der Sätze vom konstruktiven Typ in die Behauptung  $g(\mathbf{x}, \mathbf{u})$  einsetzen können. Die so entstehende rationale Funktion  $g(\mathbf{x}(\mathbf{u}), \mathbf{u})$  in den Variablen  $\mathbf{u}$  können wir vereinfachen und erhalten ähnliche Alternativen:

1) Führt die Simplifikation auf das Nullpolynom, so ist für alle zulässigen Spezifikationen der unabhängigen Variablen (d.h. solche, für die  $D(\mathbf{u}) \neq 0$  gilt) der Satz richtig.

2) Ergibt die Simplifikation eine nicht verschwindende rationale Funktion  $\frac{P(\mathbf{u})}{Q(\mathbf{u})} \neq 0$ , so ist der Satz für fast alle Spezifikationen falsch und höchstens unter der zusätzlichen Voraussetzung  $P(\mathbf{u}) = 0$  richtig. Er gilt also dann nur, wenn zwischen den als unabhängig angenommenen Variablen eine Abhängigkeit besteht.

3) Ist schließlich bereits  $D(\mathbf{u})$  identisch Null, so gibt es eine  $k(\mathbf{u})$ -lineare Kombination der Zeilen der Koeffizientenmatrix zum Nullvektor

$$\sum_i \alpha_i(\mathbf{u}) c_{ij}(\mathbf{u}) = 0, \quad j = 1, \dots, m$$

Dann ist entweder  $\sum_i \alpha_i(\mathbf{u}) c_{i0}(\mathbf{u}) \neq 0$  eine nichttriviale algebraische Relation zwischen den unabhängigen Variablen oder aber auch die erweiterte Koeffizientenmatrix hat über  $k(\mathbf{u})$  einen Rang kleiner als  $m$ , d.h. eine der abhängigen Variablen  $x_i$  kann frei gewählt werden. In beiden Fällen sind die Voraussetzungen nicht eingehalten.

**Satz 16** (Über das mechanisierte Beweisen geometrischer Sätze vom linearen Typ):

Sei  $(F \Rightarrow g)$  ein Satz vom linearen Typ und  $(\mathbf{u}, \mathbf{x})$  eine Aufteilung der Variablen der universellen Realisierung  $\overline{O}$  in bzgl.  $F$  unabhängige und abhängige.

Ist  $D(\mathbf{u})$  nicht identisch Null, so ist  $\mathbf{u}$  eine maximale Menge unabhängiger Variablen und der Satz gilt unter der Nichtdegenerationsbedingung  $D(\mathbf{u}) \neq 0$  genau dann, wenn  $g(x(\mathbf{u}), \mathbf{u})$  als rationale Funktion in  $k(\mathbf{u})$ , die durch Substitution der eindeutig bestimmten Lösung  $x = x(\mathbf{u}) \in k(\mathbf{u})^m$  von  $F$  in  $g(\mathbf{x}, \mathbf{u})$  entsteht, identisch Null ist.

Für Sätze vom linearen Typ ist die Nichtdegenerationsbedingung nicht Teil der Formulierung des Satzes und damit der Gültigkeitsbereich des Satzes nicht vorgegeben. Dies ist typisch für geometrische Sätze vom Gleichungstyp:

**Definition 12** Ein geometrischer Satz vom Gleichungstyp  $(F \Rightarrow g)$  heißt im Allgemeinen gültig, wenn es eine polynomiale Nichtdegenerationsbedingung  $h \in R$  gibt, so dass  $(F/h \Rightarrow g)$  gültig ist.

Wie im Beispiel des Satzes von Desargue kann es sein, dass die Bedingung  $D(\mathbf{u}) \neq 0$  eine Reihe geometrisch relevanter Fälle ausschließt. Die Ausnahmemenge wird jedoch durch eine algebraische Bedingung auf die unabhängigen Variablen beschrieben, ist also „klein“. Oft kann durch ein weiteres geometrisches Argument der Satz für einen relevanten Teil der Ausnahmemenge auf einen

nicht degenerierten Fall zurückgeführt werden. Sind etwa in der obigen Formulierung des Satzes von Desargues die Geraden  $AB$  und  $PQ$  parallel zur  $y$ -Achse, aber  $A, B, C$  nicht kollinear, so kann durch eine Drehung die Realisierung der Konfiguration in eine andere Realisierung überführt werden, die sich nicht in einer Ausnahmelage befindet.

**Aufgabe 16** *Beweisen Sie auf diesem Weg den Satz des Apollonius: In einem rechtwinkligen Dreieck liegen die drei Seitenmitten und der Höhenfußpunkt auf der Hypotenuse auf einem gemeinsamen Kreis.*

**Aufgabe 17** *Beweisen Sie auf diese Weise den Satz von Pappus.*

## 5 Gröbnerbasen und generisch gültige Geometrietheoreme

### 5.1 Ideale und Nullstellenmengen.

#### Korrespondenzsatz und Zerlegungssatz

Zur Untersuchung der Nullstellenmengen (nicht linearer) polynomialer Gleichungssysteme  $F = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})\}$  geht man – ähnlich zum Untervektorraum, der von den Zeilenvektoren der Koeffizientenmatrix eines linearen Gleichungssystems erzeugt wird – zu invarianten Objekten, den Idealen im Ring  $R = k[\mathbf{x}]$ , über.

**Definition 13** Als von den Polynomen  $f \in F$  erzeugtes Ideal  $Id(F)$  bezeichnet man die Menge aller polynomialen Linearkombinationen

$$Id(F) = \left\{ \sum_{f \in F} r_f f : r_f \in R \right\}$$

von Elementen aus  $F$ .

Da  $R$  ein Ring ist, kann man alle angegebenen Operationen ausführen. Es entsteht eine Menge, die abgeschlossen ist bzgl. Addition, Subtraktion und Vervielfachung mit Elementen  $r \in R$ . Diese Eigenschaft charakterisiert Ideale zugleich. Zur Bestimmung der Nullstellenmenge ist es unerheblich, ob man von der Menge  $F$  oder aber vom Ideal  $Id(F)$  ausgeht.

Auf den Idealen eines Ringes  $R$  sind folgende Operationen möglich:

Summe zweier Ideale  $I_1, I_2 \in \mathcal{I}(R)$ :

$$I_1 + I_2 = \{a + b \in R : a \in I_1, b \in I_2\}$$

Produkt zweier Ideale  $I_1, I_2 \in \mathcal{I}(R)$ :

$$I_1 \cdot I_2 = \left\{ \sum_k a_k \cdot b_k \in R : a_k \in I_1, b_k \in I_2 \right\}$$

Durchschnitt  $I_1 \cap I_2$  zweier Ideale  $I_1, I_2 \in \mathcal{I}(R)$ .

Idealquotient eines Ideals  $I \in \mathcal{I}(R)$  bzgl. eines Polynoms  $f \in R$ :

$$I : f = \{g \in R : f \cdot g \in I\}$$

Stabiler Idealquotient eines Ideals  $I \in \mathcal{I}(R)$  bzgl. eines Polynoms  $f \in R$ :

$$I : f^\infty = \{g \in R : \exists n f^n \cdot g \in I\}$$

Eine Menge, die sich als Nullstellengebilde

$$V(F) = V(Id(F)) = \{ \mathbf{a} = (a_1, \dots, a_n) \in \mathbf{C}^n : \forall f \in F f(\mathbf{a}) = 0 \}$$

eines Systems von Polynomen  $F$  darstellen lässt, bezeichnen wir als *affine Varietät*.

Im Gegensatz zur linearen Algebra können *verschiedene* Ideale dieselbe affine Varietät definieren. Das hängt damit zusammen, dass Ideale eigentlich Nullstellenmengen mit deren Vielfachheiten verkörpern. So besteht etwa aus Sicht der Nullstellen kein Unterschied zwischen den Mengen  $F_1 = \{x + 1, y - 1\}$  und  $F_2 = \{(x + 1)^2, (y - 1)^3\}$ , aus Sicht der Ideale schon. Sei deshalb  $rad(I)$  das größte Ideal, welches dieselbe Nullstellenmenge  $V(I)$  wie  $I$  (über einem algebraisch abgeschlossenem Körper) hat. Dieses Ideal bezeichnet man auch als das *Radikal* von  $I$ . Ideale, die

mit ihrem Radikal übereinstimmen, bezeichnen wir auch als *Radikalideale*. Nach dem Hilbertschen Nullstellensatz gilt

$$\text{rad}(I) = \{f \in R : \exists k > 0 f^k \in I\}.$$

Der Hilbertsche Nullstellensatz behandelt den in gewissem Sinne prototypischen Spezialfall, dass  $V(I) = \emptyset$  gilt:

**Satz 17 (Hilberts Nullstellensatz)** *Ist  $K$  ein algebraisch abgeschlossener Körper und  $V(I) = \emptyset$ , so gilt  $1 \in I$ , d.h.,  $I$  ist das triviale Eins-Ideal.*

Dieser Satz verallgemeinert den Fundamentalsatz der Algebra, dass jedes nicht triviale Polynom in  $\mathbb{C}$  eine Nullstelle besitzt.

Ist umgekehrt  $Z \subset \mathbb{A}^n$  eine Teilmenge des affinen Raums, so bezeichnen wir mit

$$I(Z) = \{f \in R : \forall z \in Z f(z) = 0\}$$

das Radikalideal der Polynome, die auf  $Z$  verschwinden.

Zwischen affinen Varietäten und Radikalidealen besteht eine eindeutige inklusionsumkehrende Korrespondenz, die in der Vorlesung “Gröbnerbasen und deren Anwendungen” als Korrespondenzsatz bewiesen wird:

**Satz 18 (Korrespondenzsatz)**  *$V$  und  $I$  sind zueinander inverse, inklusionsumkehrende Korrespondenzen zwischen den affinen Varietäten im  $\mathbb{A}^n$  und den Radikalidealen in  $R$ , die Durchschnitte der Nullstellenmengen in Idealsummen und Vereinigungen der Nullstellenmengen in Idealdurchschnitte überführen, d.h.*

1. die Bilder unter  $V$  sind genau die affinen Varietäten im  $\mathbb{A}^n$ ,
2. die Bilder unter  $I$  sind genau die Radikalideale in  $R$ ,
3. für jede Teilmenge  $W \subset \mathbb{A}^n$  ist  $V(I(W))$  die kleinste affine Varietät, die  $W$  umfasst (deren affiner Abschluss),
4. für jedes Ideal  $J \subset R$  gilt  $I(V(J)) = \text{rad}(J)$ ,
5.  $I_1 \subseteq I_2 \Rightarrow V(I_1) \supseteq V(I_2)$ ,
6.  $V_1 \subseteq V_2 \Rightarrow I(V_1) \supseteq I(V_2)$ .
7.  $V(I_1 + I_2) = V(I_1) \cap V(I_2)$  und
8.  $V(I_1 \cdot I_2) = V(I_1 \cap I_2) = V(I_1) \cup V(I_2)$ .

Ohne Beweis. Siehe Vorlesung “Gröbnerbasen und deren Anwendungen”.

**Definition 14** *Eine affine Varietät  $V$  heißt irreduzibel, wenn sie sich nicht als Vereinigung zweier echt kleinerer Varietäten darstellen lässt, d.h. wenn*

$$V = V_1 \cup V_2 \Rightarrow V = V_1 \text{ oder } V = V_2$$

*gilt.*

Ist  $V$  eine solche irreduzible Varietät und verschwindet das Produkt  $fg$  auf ganz  $V$ , so muss bereits einer der Faktoren auf  $V$  verschwinden. In der Tat, wegen

$$(V \cap V(f)) \cup (V \cap V(g)) = V \cap (V(f) \cup V(g)) = V \cap V(fg) = V$$

wäre das eine Zerlegung in kleinere Varietäten. Das Verschwindungsideal  $I = I(V)$  ist also ein Primideal.

**Satz 19** In einem Polynomring lässt sich jedes Radikalideal als Durchschnitt endlich vieler Primideale darstellen.

Jede affine Varietät  $V$  ist die Vereinigung von endlich vielen irreduziblen Varietäten  $V = \cup V_i$ .

Ohne Beweis. Siehe Vorlesung “Gröbnerbasen und deren Anwendungen”.

Die Teilvarietäten  $V_i$  bezeichnet man auch als die *irreduziblen Komponenten* von  $V$ . Wird etwa die Varietät  $V = V(f)$  durch eine Gleichung  $f$  gegeben und ist  $f = p_1^{a_1} \cdot \dots \cdot p_k^{a_k}$  die Zerlegung dieses Polynoms in Primfaktoren, so lässt sich  $V$  zerlegen als

$$V(f) = \bigcup_{i=1}^k V(p_i).$$

Es kann sein, dass geometrische Sätze nur auf einigen Komponenten eines Nullstellengebildes gültig sind, während andere Komponenten degenerierten Lagen entsprechen.

Eine vollständige Zerlegung in irreduzible Komponenten ist meist schwierig. Oft genügt eine teilweise Zerlegung in degenerierte und nicht degenerierte Teile.

## 5.2 Idealtest, Normalformen und Gröbnerbasen

Die Frage, ob ein Polynom  $g$  auf den gemeinsamen Nullstellen der Polynome  $f \in F$  verschwindet, lässt sich also als Idealthaltenseinsproblem „ $g \in Id(F)$ ?“ umformulieren, da Ideale  $Id(F)$ , die aus geometrischen Fragestellungen entstehen, meist bereits Radikalideale sind.

Bei der algorithmischen Umsetzung solcher Idealthaltenseinstests spielen Normalformen und Gröbnerbasen eine zentrale Rolle. Fixieren wir dazu zunächst wieder die erforderliche Begrifflichkeit:

Sei  $R = k[x_1, \dots, x_n] = k[\mathbf{x}]$  ein Polynomring und  $T := T(\mathbf{x}) = T(x_1, \dots, x_n)$  das zugehörige Termmonoid, welches mit einer Termordnung, also einer linearen, monotonen, noetherschen Ordnungsrelation versehen ist. Die Terme eines Polynoms  $0 \neq f(\mathbf{x}) = \sum_{i=0}^N c_i \mathbf{x}^{\alpha_i} \in R$  seien so angeordnet, dass in der fixierten Termordnung  $\mathbf{x}^{\alpha_i} > \mathbf{x}^{\alpha_j}$  für  $i < j$  gilt. Bezeichne weiter  $T(f) := \{\mathbf{x}^{\alpha_i}, i = 0, \dots, N\}$  die Menge der in der Darstellung von  $f$  auftretenden Terme. Dann können wir die folgenden Begriffe definieren:

- den Leitterm  $lt(f) := \mathbf{x}^{\alpha_0}$ ,
- den Leitkoeffizienten  $lc(f) := c_0$ ,
- das Leitmonom  $lm(f) := lc(f) \cdot lt(f)$ ,
- das Reduktum  $red(f) := f - lm(f)$ .

Für eine Menge  $B \subset R$  von Polynomen bilden die Vielfachen der Leiterte  $lt(b), b \in B$ , das Monoid-Ideal  $\Sigma(B) := (lt(b) : b \in B) \subset T$ . Mit diesen Begriffen kann man eine Erweiterung der Division mit Rest von Polynomen einführen, indem ein Polynom  $f \in R$  so lange reduziert wird, so lange einer der Terme durch den Leitterm eines der Polynome  $b \in B$  teilbar ist. Das Ergebnis dieser Reduktion bezeichnet man als (vollständige) *Normalform*  $r = NF(f, B)$  des Polynoms  $f$  bzgl.  $B$ . Die Terme  $t \in \Sigma(B)$  bezeichnet man dabei auch als *Nichtstandardterme*, da sie weiter reduziert werden können. Die Terme  $t \in N(B) = T \setminus \Sigma(B)$  heißen *Standardterme*.

**Satz 20** Sei  $R = k[x_1, \dots, x_n]$  ein Polynomring über einem Körper  $k$ ,  $<$  eine Termordnung auf  $T(x_1, \dots, x_n)$  und  $B = \{b_1, \dots, b_m\} \subset R$  eine endliche Menge von Polynomen. Dann gibt es einen Algorithmus, der für jedes Polynom  $f \in R$  nach endlich vielen Schritten eine Darstellung

$$f = v_1 b_1 + \dots + v_m b_m + r$$

mit  $v_1, \dots, v_m, r \in R$  produziert, in der  $r = 0$  gilt oder  $r$  eine Linearkombination von Standardtermen bzgl.  $B$  ist und  $lt(f) \geq lt(v_i b_i)$  für alle  $i$  gilt.

Dieser Normalform-Algorithmus ist in den verschiedenen CAS (oft nur für eine eingeschränkte Auswahl von Termordnungen) implementiert. In MuPAD muss dazu in den Datentyp `poly` umgewandelt werden. Für weitere Anwendungen setzen wir dabei gleich den Koeffizientenbereich auf einen sehr allgemeinen Wert und die Termordnung auf `LexOrder`.

```
DEF:=Dom::ExpressionField(normal):
NF:=proc(f,polys,vars)
begin
  expr(groebner::normalf(poly(f,vars,DEF),
    map(polys,u->poly(u,vars,DEF)), LexOrder));
end_proc:
```

Bei der Berechnung einer Normalform wird  $f$  nur um ein Element aus dem Ideal  $I = Id(B)$  abgeändert. Dafür können wir die aus der Zahlentheorie bekannte Notation der Restklassen verallgemeinern:

**Definition 15** *Zwei Polynome  $f, g \in R$  heißen kongruent modulo dem Ideal  $I$ , wenn  $f - g \in I$  gilt. Wir schreiben in diesem Fall  $f \equiv g \pmod{I}$ .*

Mit Kongruenzen kann man genauso rechnen, wie wir das von Restklassen ganzer Zahlen gewohnt sind. Insbesondere gilt  $f \in Id(B)$ , wenn wir  $NF(f, B) = 0$  nachweisen können. Die Umkehrung dieser Aussage gilt nicht allgemein, da für beliebige  $B$  eine solche Normalform nicht eindeutig bestimmt ist, sondern vom Reduktionsweg abhängen kann.

Diese Mehrdeutigkeit wird beseitigt, wenn man von  $B$  zu einer Gröbnerbasis  $G = GBasis(B)$  übergeht. Gröbnerbasen zeichnen sich durch eine Reihe zueinander äquivalenter Eigenschaften aus:

**Satz 21 (Charakterisierungssatz für Gröbnerbasen)** *Folgende Bedingungen an eine Basis  $G$  eines Ideals  $I \subset R$  sind äquivalent:*

1.  $G$  ist eine Gröbnerbasis von  $I$ , d.h. es gilt  $\Sigma(I) = \Sigma(G)$ .
2. Für jedes Element  $f \in I$  und jede Reduktionsstrategie gilt  $NF(f, G) = 0$ .
3. Jedes Element  $f \in I$  hat eine Darstellung

$$f = \sum_{g_i \in G} h_i g_i \quad \text{mit} \quad \forall i \ (lt(f) \geq lt(h_i g_i)).$$

4. Die Standardterme  $N(G)$  bilden eine Vektorraumbasis des Faktorringes  $R/I$ , d.h. jedes Element  $f \in R$  besitzt eine eindeutige Darstellung

$$f \equiv \sum_{m \in N(G)} c_m m \pmod{I}$$

mit  $c_m \in k$ .

Ohne Beweis. Siehe Vorlesung “Gröbnerbasen und deren Anwendungen”.

Für derartige Basen gilt also  $NF(f, G) = 0 \Leftrightarrow f \in I$  und das Ergebnis der Normalformberechnung ist unabhängig vom gewählten Reduktionsweg.

Die Berechnung von Gröbnerbasen ist ebenfalls in den verschiedenen CAS (für eine eingeschränkte Auswahl von Termordnungen) implementiert. Wir wollen den in MuPAD verfügbaren allgemeinen Algorithmus `groebner` für unsere Zwecke konfigurieren (Grundbereich `ExpressionField` und Termordnung `LexOrder` – beides aus Effizienzgründen nicht die Defaulteinstellung).

```

DEF:=Dom::ExpressionField(normal):
GBasis:=proc(polys,vars)
begin
  expr(groebner::gbasis(map(polys,u->poly(u,vars,DEF)),LexOrder));
end_proc:

```

Für geometrische Sätze vom Gleichungstyp können wir damit das folgende hinreichende Beweiskriterium formulieren:

**Satz 22** *Gilt  $NF(g \cdot h, GBasis(F)) = 0$ , so ist der Satz  $[F/h \Rightarrow g]$  gültig.*

Wenden wir dieses hinreichende Kriterium zum Beweis eines geometrischen Satzes vom Gleichungstyp an.

**Beispiel 1 Arnon-Beispiel** ([1, S. 267]): *Gegeben sei ein Quadrat  $ABCD$  und der Punkt  $P$  auf der Parallelen zu  $BD$  durch  $C$  mit  $|BP| = |BD|$ . Sei ferner  $Q$  der Schnittpunkt von  $CD$  und  $BP$ . Man zeige, dass dann  $|DP| = |DQ|$  gilt.*

Eine Lösung in einem speziellen Koordinatensystem geht von drei Variablen aus:

```

unprotect(D):
A:=Point(0,0); B:=Point(1,0); C:=Point(1,1); D:=Point(0,1);
P:=Point(x1,x2); Q:=varpoint(C,D,x3);

polys:={ on_line(P,par_line(C,pp_line(B,D))),
         sqrdist(B,D)-sqrdist(B,P), on_line(Q,pp_line(B,P))};
con:=sqrdist(D,Q)-sqrdist(D,P);

```

Wir erhalten als Idealbasis

$$\text{polys} = \{x_1 + x_2 - 2, -x_1^2 + 2x_1 - x_2^2 + 1, 1 - x_2x_3 - x_1\}$$

und als Behauptung

$$\text{con} = -x_1^2 - x_2^2 + 2x_2 + x_3^2 - 2x_3$$

Die folgende Rechnung zeigt den Unterschied zwischen der Verwendung der Ausgangspolynome und einer Gröbnerbasis:

```

vars:=[x1,x2,x3];
NF(con,polys,vars);

```

$$x_3^2 - 2x_3 + 4x_2 - 5$$

da bzgl. der lexikografischen Termordnung im Ausgangssystem  $polys$  nur Vielfache von  $x_1$  als Leitterme auftreten. Mit der zugehörigen Gröbnerbasis

```

gb:=GBasis(polys,vars);

```

$$\{x_3^2 - 4x_3 + 1, 2x_2 + x_3 - 3, 2x_1 - x_3 - 1\}$$

liefert die entsprechende Normalform

```

NF(con,gb,vars);

```

Null und damit einen Beweis der Behauptung.

Allerdings ist es schwierig, die genauen Koordinaten der Punkte  $P$  und  $Q$  zu bestimmen, denn das Polynom  $x_3^2 + 2x_3 - 2$  aus der Gröbnerbasis hat *zwei* Nullstellen für  $x_3$  (aus denen sich dann jeweils  $x_1$  und  $x_2$  aus den beiden folgenden Gleichungen eindeutig ergibt). Hat diese zweite Lösung eine geometrische Bedeutung oder ist sie degeneriert? Ein genauerer Blick auf die geometrische Konfiguration zeigt, dass wir die zweite Lösung schlichtweg übersehen haben.

Allerdings war dieses Beispiel besonders einfach, weil in ihm alle Variablen durch Gleichungen gebunden waren. Für allgemeinere Beweisschemata läuft der vorgestellte Ansatz meist nicht so glatt durch.

**Beispiel:** Betrachten wir noch einmal den *Satz vom Schnittpunkt der Winkelhalbierenden*. In dessen Beweisschema treten vier Variable auf.

```
A:=Point(0,0); B:=Point(1,0); C:=Point(c1,c2); P:=Point(x1,x2);
polys:={on_bisector(P,A,B,C), on_bisector(P,B,C,A)};
con:=on_bisector(P,C,A,B);
```

Eine Gröbnerbasis (bzgl. der Ordnung DegRevLex)

```
vars:=[x1,x2,c1,c2];
gb:=GBasis(polys,vars);
```

enthält 5 Polynome. Leider kann damit nicht bewiesen werden, dass der Satz gilt, denn die Normalform von  $con$  reduziert nicht zu Null.

```
NF(con,gb,vars);
```

$$2x_2 - c_2 - 2c_1x_2 - 2x_1x_2 + 2c_2x_1$$

Allerdings reduziert das folgende Produkt zu Null:

```
NF(((c1-1)^2+c2^2)*con,gb,vars);
```

Damit gilt der Satz  $[F/h \Rightarrow g]$  unter der Nichtdegenerationsbedingung  $h = (c_1 - 1)^2 + c_2^2$ . Deren Herkunft bleibt an dieser Stelle natürlich im Dunkeln. Über den reellen Zahlen entspricht  $h = 0$  dem Fall  $c_1 = 1, c_2 = 0$ , also der degenerierten Situation  $B = C$ , in welcher die Gerade  $BC$  als Schenkel der beiden Winkel in  $polys$  entartet.

Eine spezielle Anwendung von Gröbnerbasen ergibt sich aus folgender Überlegung: Ist  $g(\mathbf{x})$  ein Polynom, das auf allen gemeinsamen Nullstellen von  $F(\mathbf{x})$  verschwindet und  $t$  eine neue Variable, so hat offensichtlich das System  $F = 1 - t \cdot g = 0$  keine gemeinsamen Lösungen. Das ist aber nach dem Hilbertschen Nullstellensatz genau dann der Fall, wenn das von diesen Polynomen erzeugte Ideal das Einsideal ist, was wiederum durch eine Gröbnerbasisberechnung geprüft werden kann. Dieser **indirekte Beweisansatz** ergibt für obiges Beispiel mit

```
vars:=[t,x1,x2,c1,c2];
gb:=GBasis([op(polys),1-t*con],vars);
```

jedoch nicht das Einsideal, sondern eine Gröbnerbasis mit sechs Elementen. Diese enthält aber wenigstens das Polynom  $h = (c_1 - 1)^2 + c_2^2$ , so dass für gemeinsame Lösungen von  $F = 1 - t \cdot g = 0$  stets  $h = 0$  gilt. Das kann durch  $h$  als Nichtdegenerationsbedingung ausgeschlossen werden, so dass wieder  $(F/h \Rightarrow g)$  bewiesen ist. Diesen Zugang können wir zu folgendem hinreichenden Kriterium verallgemeinern:

**Satz 23** Ist  $t$  eine neue Variable und  $h(x) \in J = I(F, 1 - t \cdot g)$ , so ist  $[F/h \Rightarrow g]$  gültig.

**Aufgabe 18** Diskutieren Sie dieselben Fragen bzgl. des folgenden Beweisansatzes:

```
A:=Point(0,0); B:=Point(1,0); C:=Point(c1,c2); P:=Point(x1,x2);
polys:={on_bisector(P,A,B,C), on_bisector(P,C,A,B)};
con:=on_bisector(P,B,C,A);
```

Hinweis: Durch die spezielle Wahl der Koordinaten von  $A$  und  $B$  kann der degenerierte Fall  $A = B$  hier nicht auftreten und  $[polys \Rightarrow con]$  ist ohne weitere Voraussetzungen gültig.

### 5.3 Eliminationsideale und Gröbnerbasen

Zur Berechnung eines solchen Polynoms  $h(x)$  muss ein  $t$ -freies Polynom in  $J$  gefunden werden. Meist fordert man, dass  $h$  gänzlich nur „unabhängige“ Variablen enthält. Beide Fragen führen auf ein Eliminationsproblem:

Gegeben ist ein Polynomring  $R = k[\mathbf{u}, \mathbf{x}]$  und ein Ideal  $J \subset R$ . Gesucht sind alle Polynome in  $J$ , die keine der Variablen  $x \in \mathbf{x}$  enthalten, also die Menge

$$J' = J \cap k[\mathbf{u}].$$

Es stellt sich heraus, dass diese Menge ein Ideal im Ring  $k[\mathbf{u}]$  ist, welches als *Eliminationsideal* bezeichnet wird.

Basen für Eliminationsideale können aus Gröbnerbasen für spezielle Eliminationstermordnungen abgelesen werden. Als *Eliminationsordnung* für  $\mathbf{u}$  bezeichnet man jede Termordnung, für welche

$$t_1 \in T(\mathbf{u}), t_2 \in T(\mathbf{x} \cup \mathbf{u}) \setminus T(\mathbf{u}) \Rightarrow t_1 < t_2$$

gilt. Insbesondere ist die lexikographische Termordnung eine Eliminationsordnung für jedes Anfangssegment von Variablen.

#### Satz 24 (Eliminationssatz für Gröbnerbasen)

Ist  $G = GBasis(F)$  eine (min. reduzierte) Gröbnerbasis des Polynomsystems  $F \subset R = k[\mathbf{u}, \mathbf{x}]$  bzgl. einer Eliminationsordnung für  $\mathbf{u}$ , so ist

$$G' = \{g \in G : lt(g) \in T(\mathbf{u})\}$$

eine (min. reduzierte) Gröbnerbasis des Eliminationsideals  $J' = Id(F) \cap k[\mathbf{u}]$ .

Ohne Beweis. Siehe Vorlesung „Gröbnerbasen und deren Anwendungen“.

### 5.4 \* Der Gröbnerfaktorisierer

Enthält ein Ideal  $I$  ein Polynom  $f = f_1 \cdot f_2$ , das sich in zwei Faktoren zerlegen lässt, die beide nicht im Ideal  $I$  enthalten sind, so entspricht dem die Aufspaltung der Nullstellenmenge  $Z = V(I)$  in die beiden Teile  $V(I + Id(f_1)) \cup V(I + Id(f_2))$ . Auch wenn jeder der beiden Teile noch aus mehreren Komponenten von  $Z$  bestehen kann, so kann eine solche partielle Zerlegung hilfreich sein, da wir das Geometrietheorem auf jeder einzelnen Komponente verifizieren können bzw. vielleicht eine Komponente (und den Grund dafür) entdecken, auf der das Theorem nicht gültig ist. Führt man solche Probefaktorisierungen auf den Basiselementen und auch während der Rechnungen des Gröbneralgorithmus aus, so erhält man im Ergebnis aus dem Ausgangssystem  $F$  eine Liste von

Gröbnerbasen  $G_i$ , so dass  $V(F) = \cup_i V(G_i)$  gilt. Diese Variante des Gröbneralgorithmus bezeichnet man als den *Gröbnerfaktorisierer*. Er ist in einigen der CAS ebenfalls implementiert. Da in einer optimierten Variante auf dem zweiten Zweig  $f_1 \neq 0$  berücksichtigt wird, kann man in einigen Implementierungen dem Gröbnerfaktorisierer bereits beim Start (optional) Nichtdegenerationsbedingungen übergeben.

Maple	<code>gsolve(polys, vars)</code>
MuPAD	<code>groebner::factor(gbasis(polys, T0), T0)</code> (eingeschränkt)
Reduce	<code>groebnerf(B[, nondegs])</code> (bzgl. TORDER)
Reduce (CALI)	<code>groebfactor(polys, nondegs)</code> (bzgl. Currentring)

Obwohl dies nicht garantiert werden kann, ist die dabei produzierte Zerlegung für eine lexikografische Termordnung in praktischen Anwendungen oft nahe an der Zerlegung in irreduzible Komponenten.

Betrachten wir noch einmal das Beispiel des Schnittpunkts der Winkelhalbierenden (Rechnungen mit Reduce):

```
A:=Point(0,0); B:=Point(1,0); C:=Point(c1,c2); P:=Point(x1,x2);
polys:={on_bisector(P,A,B,C), on_bisector(P,B,C,A)};
con:=on_bisector(P,C,A,B);

vars:={x1,x2,c1,c2};
res:=geo_groebfactor(polys,vars,{});
for each x in res collect con mod x;
```

$$\left\{ \begin{aligned} &2x_1x_2 - 2x_1c_2 + 2x_2c_1 - 2x_2 + c_2, \\ &x_1^2c_2 - 2x_1c_1c_2 - x_2^2c_2 + 2x_2c_1^2 - 2x_2c_1 + c_1c_2, \\ &4x_2^4c_2 - 8x_2^3c_1^2 + 8x_2^3c_1 - 8x_2^3c_2^2 + 4x_2^2c_1^2c_2 - 4x_2^2c_1c_2 + 4x_2^2c_2^3 \\ &\quad - 4x_2^2c_2 + 4x_2c_2^2 - c_2^3, \\ &2x_1c_1c_2^2 - x_1c_2^2 + 2x_2^3c_2 - 4x_2^2c_1^2 + 4x_2^2c_1 - 2x_2^2c_2^2 - 2x_2c_1^2c_2 \\ &\quad + 2x_2c_1c_2 - 2x_2c_2 - c_1c_2^2 + c_2^2 \end{aligned} \right\},$$

$$\left\{ \begin{aligned} &c_1^2 - 2c_1 + c_2^2 + 1, \\ &x_1^2c_2 - 2x_1x_2c_1 + 2x_1x_2 - 2x_1c_2 - x_2^2c_2 + 2x_2c_1 - 2x_2 + c_2 \end{aligned} \right\}$$

Die faktorisierte Gröbnerbasis besteht aus zwei Komponenten, wovon auf der ersten  $con \equiv 0$  gilt. Die andere enthält in der Gröbnerbasis u.a. wieder das Polynom  $h = c_1^2 - 2c_1 + c_2^2 + 1$ , entspricht also dem degenerierten Fall, in dem  $B = C$  gilt.

## 5.5 Unabhängige und abhängige Variablen

In obigen Beispielen haben wir bereits begonnen, auf einer heuristischen Basis zwischen unabhängigen und abhängigen Variablen zu unterscheiden. Dies wollen wir nun präzisieren. Im konstruktiven Fall war eine wesentliche Voraussetzung für die Gültigkeit des entsprechenden Mechanisierungssatzes die algebraische Unabhängigkeit der Variablen gewesen. Ist  $J \subset R = k[x_1, \dots, x_n] = R[\mathbf{x}]$  ein Relationenideal, so wollen wir eine Teilmenge  $\mathbf{u} \subset \mathbf{x}$  als *unabhängig* modulo  $J$  bezeichnen, wenn

$$J \cap k[\mathbf{u}] = (0)$$

gilt, d.h. wenn es keine polynomiale Relation in  $J$  zwischen den ausgewählten Variablen gibt. Der Nachweis, dass eine vorgegebene Variablenmenge unabhängig ist, wird somit auf ein Eliminationsproblem reduziert, das mit der Berechnung einer Gröbnerbasis beantwortet werden kann: Ist  $G$  eine Gröbnerbasis von  $J$  bzgl. einer Eliminationsordnung für  $\mathbf{u}$  und  $\{lt(g), g \in G\} \cap T(\mathbf{u}) = \emptyset$ , so ist  $\mathbf{u}$  eine modulo  $J$  unabhängige Variablenmenge.

Ist  $J$  ein Primideal, so haben alle maximalen unabhängigen Variablenmengen die gleiche Kardinalität. Diese Anzahl (der unabhängigen Parameter) bezeichnet man auch als die Dimension des Ideals  $J$  bzw. der Nullstellenmenge  $V(J)$ . Ist  $\mathbf{u}$  eine solche maximale unabhängige Variablenmenge und  $x_i$  eine weitere Variable, so hängt  $x_i$  über  $J$  algebraisch von den Variablen  $\mathbf{u}$  ab, d.h. für alle  $x_i$  existieren Polynome

$$f_i(x_i, \mathbf{u}) = \sum_j c_{ij}(\mathbf{u})x_i^j \in J$$

$x_i$  kann man dann zwar im Allgemeinen nicht rational durch die unabhängigen Parameter ausdrücken, jedoch ergibt sich für fast alle Parameterbelegungen (genauer: solche, für die es ein  $j$  mit  $c_{ij}(\mathbf{u}) \neq 0$  gibt) eine Bestimmungsgleichung für  $x_i$  und damit nur endlich viele zulässige Parameterspezifikationen  $(\mathbf{x}_0, \mathbf{u}_0)$  für eine vorgegebene Spezifikation  $\mathbf{u} \mapsto \mathbf{u}_0$ . Wenn wir  $\mathbf{u}_0$  so wählen, dass sogar  $lc(f_i)(\mathbf{u}_0) \neq 0$ , also  $\deg(f_i) = \deg(f_i|_{u=\mathbf{u}_0})$  für alle  $i$  gilt, lässt sich auch die Anzahl dieser Lösungen voraussagen. Dazu unten mehr.

## 5.6 Generisch gültige Geometrietheoreme

Im Fall der Geometrietheoreme vom konstruktiven und vom linearen Typ haben wir lineare Algebra über dem Körper  $k(\mathbf{u})$  der rationalen Funktionen in den unabhängigen Parametern getrieben. Wir wollen deshalb nun auch für den allgemeinen Fall im Ring  $S = k(\mathbf{u})[\mathbf{x}]$  der Polynome in  $\mathbf{x}$  mit rationalen Funktionen in  $\mathbf{u}$  als Koeffizienten rechnen statt wie bisher im Ring  $R = k[\mathbf{u}, \mathbf{x}]$ . Wir wollen dabei voraussetzen, dass  $\mathbf{u}$  eine maximale unabhängige Variablenmenge modulo  $J = Id(F)$  ist.

Betrachten wir zunächst einige Beispiele.

**Beispiel 1:** Satz vom Miquelschen Punkt

```
A:=Point(0,0); B:=Point(1,0); C:=Point(u1,u2);
R:=varpoint(B,C,u3); S:=varpoint(A,C,u4); T:=varpoint(A,B,u5);
P:=Point(x1,x2);
```

```
polys:={ is_concyclic(A,S,T,P), is_concyclic(B,R,T,P)};
con:= is_concyclic(C,R,S,P);
```

Wir wollen die Gröbnerbasis von  $J = Id(polys)$  im Ring  $S = k(\mathbf{u})[x_1, x_2]$  berechnen:

```
vars:=[x1,x2];
gb:=GBasis(polys,vars);
```

Diese Gröbnerbasis hat eine recht einfache Struktur

$$\{x_2^2 - p_1(\mathbf{u})x_2, x_1 - p_2(\mathbf{u})x_2 - u_5\},$$

so dass das zugehörige Gleichungssystem zwei Lösungen  $(x_1, x_2) \in \mathbb{A}_{k(\mathbf{u})}^2$  mit  $x_{2,1} = 0$  und  $x_{2,2} = p_1(\mathbf{u})$  besitzt. Eine Lösung entspricht dem Schnittpunkt  $F$  der beiden Kreise, die andere dem Punkt  $P$  und erfüllt die Gleichung  $con$ :

```
sol:=solve(gb,vars,IgnoreSpecialCases);
map(sol,u->normal(subs(con,u)));
```

Die Lösung besteht in diesem Fall also aus zwei Komponenten, wobei der Satz auf einer Komponente gilt, auf der zweiten (aus nahe liegenden Gründen) dagegen nicht. Es handelt sich dabei jedoch nicht um eine degenerierte Lage wie in früheren Beispielen, sondern um einen essentiell auszuschließenden Fall.

**Beispiel 2:** Satz vom Schnittpunkt der Winkelhalbierenden

```
A:=Point(0,0); B:=Point(1,0); C:=Point(u1,u2); P:=Point(x1,x2);
polys:={on_bisector(P,B,A,C), on_bisector(P,C,B,A)};
con:= on_bisector(P,A,C,B);
```

```
gb:=GBasis(polys,vars);
NF(con,gb,vars);
```

Als Gröbnerbasis erhalten wir hier

$$\left[ x_1 - \frac{2}{u_2 - 2u_1u_2}x_2^3 + \frac{4u_1^2 - 4u_1 + 2u_2^2}{u_2^2 - 2u_1u_2^2}x_2^2 + \frac{2u_1^2 - 2u_1 + 2}{u_2 - 2u_1u_2}x_2 - \frac{u_1 - 1}{2u_1 - 1}, \right. \\ \left. x_2^4 - \frac{2u_1^2 - 2u_1 + 2u_2^2}{u_2}x_2^3 - \frac{u_2^2}{4}x_2^2 + u_2x_2 - (-u_1^2 + u_1 - u_2^2 + 1) \right]$$

In  $S = k(\mathbf{u})[x_1, x_2]$  ist das davon erzeugte Ideal nulldimensional und hat genau vier Nullstellen in  $\mathbb{A}_K^2$ , dem zweidimensionalen affinen Raum über  $K = \overline{k(\mathbf{u})}$ , dem algebraischen Abschluss von  $k(\mathbf{u})$ . Diese entsprechen den vier („generischen“) Schnittpunkten der möglichen Auswahlen der Halbierenden von Innen- und Außenwinkel von  $\angle ABC$  und  $\angle BCA$ . Jeder von ihnen liegt auf der („generischen“) Halbierenden entweder des Innen- oder des Außenwinkels von  $\angle CAB$ , da  $\text{NF}(\text{con}, \text{gb}) = 0$  und folglich  $\text{con} \in \text{Id}(\text{polys}) \cdot S$  gilt. Die Probleme mit der degenerierten Situation  $B = C$  treten nicht auf.

Es bleibt zu untersuchen, was das mit der ursprünglichen geometrischen Fragestellung zu tun hat.

**Beispiel 3:** Die Simonsche Gerade

```
M:=Point(0,0);
A:=Point(0,u1); B:=Point(u2,x2); C:=Point(u3,x3); P:=Point(u4,x4);
R:=varpoint(A,B,x5); S:=varpoint(B,C,x6); T:=varpoint(A,C,x7);
```

```
polys:={ is_orthogonal(pp_line(A,B),pp_line(P,R)),
         is_orthogonal(pp_line(A,C),pp_line(P,T)),
         is_orthogonal(pp_line(B,C),pp_line(P,S)),
         sqrdist(M,A)-sqrdist(M,P),
         sqrdist(M,B)-sqrdist(M,P),
         sqrdist(M,C)-sqrdist(M,P)};
```

```
con:= is_collinear(R,S,T);
vars:=[x5,x6,x7,x2,x3,x4];
gb:=GBasis(polys,vars);
NF(con,gb,vars);
```

Auch in diesem Fall ist das Ideal  $\text{Id}(\text{polys}) \subset k(\mathbf{u})[\mathbf{x}]$  nulldimensional. Aus der Gröbnerbasis

$$\{ x_2^2 - (u_1^2 - u_2^2), \\ x_3^2 - (u_1^2 - u_3^2), \\ x_4^2 - (u_1^2 - u_4^2), \\ (2u_1u_3)x_7 - u_3x_4 + u_4x_3 - (u_1u_3 - u_1u_4), \\ (2u_1u_2)x_5 - u_2x_4 + u_4x_2 - (u_1u_2 - u_1u_4), \\ (2u_1^2u_2 - 2u_1^2u_3)x_6 + u_2x_4x_3 + u_3x_4x_2 - u_4x_3x_2 - (u_1^2u_2 - u_1^2u_3 + u_1^2u_4 - u_2u_3u_4) \}$$

lesen wir ab, dass dieses System 8 „generische“ Lösungen besitzt und  $con$  auf allen diesen Punkten verschwindet, d.h. der Satz von der Simonschen Geraden „generisch“ (d.h. in  $\mathbb{A}_K^2$ ) gilt. Gröbnerbasisberechnungen bzgl. des kompletten Satzes von Variablen dauern deutlich länger und liefern überdies keinen Beweis des Satzes, weil degenerierte Komponenten mitgeschleppt worden sind. Der Gröbnerfaktorisierer zerlegt das Beispiel in 9 Komponenten, wovon 8 degenerierten Situationen entsprechen und auf zwei dieser Komponenten  $con$  nicht verschwindet (Rechnungen mit Reduce).

In allen bisher betrachteten Beispielen besteht die Gröbnerbasis aus Gleichungen vom Grad 1 und 2 in den abhängigen Variablen und erlaubt es, die abhängigen Variablen durch die unabhängigen auszudrücken, auch wenn es im Gegensatz zum linearen Fall mehrere, aber stets endlich viele Lösungen für  $\mathbf{x}$  gibt, die sich durch komplizierte *universelle Formeln* in den Parametern  $\mathbf{u}$  ausdrücken lassen. Diese Formeln „leben“ nicht mehr in  $k(\mathbf{u})$ , sondern in einer algebraischen Erweiterung dieses Funktionenkörpers. In manchen Fällen, wie etwa beim Miquelschen Punkt, ist darüber hinaus der Satz nicht in allen diesen „generischen Lösungen“ richtig, sondern einige müssen ausgeschlossen werden.

Wenn  $\mathbf{u}$  eine maximale unabhängige modulo  $J$  Variablenmenge ist, so ist in jedem Fall das Erweiterungsideal  $J' = J \cdot S$  ein nulldimensionales Ideal und  $S/J'$  ein endlichdimensionaler  $k(\mathbf{u})$ -Vektorraum. Dessen Dimension gibt an, wie viele Lösungen  $\mathbf{x}$  es über dem „generischen“ Tupel  $\mathbf{u}$  gibt. Beides kann man aus einer Gröbnerbasis  $G' = \mathbf{GBasis}_S(F)$  ablesen: Dimension Null liegt vor, wenn die Anzahl der Standardterme  $\Omega = N_S(G')$  endlich ist, die vektorraumdimension stimmt dann mit dieser Anzahl überein, da  $\Omega$  eine  $k(\mathbf{u})$ -Vektorraumbasis von  $S/J'$  ist.

Untersuchen wir zunächst den Fall, dass der geometrische Satz auf allen diesen „generischen Lösungen“ gilt. Um dies zu testen, haben wir  $NF_S(con, G') = 0$  für die über  $S$  berechnete Gröbnerbasis  $G'$  geprüft. Der folgende Satz gibt eine erste Auskunft, was dieses Ergebnis mit unserer ursprünglichen geometrischen Fragestellung zu tun hat.

**Satz 25 (Spezialisierungssatz)** Sei  $F \subset S = k[\mathbf{u}, \mathbf{x}]$  eine Menge von Polynomen,  $J = \text{Id}(F)$  das von ihnen erzeugte Ideal,  $\mathbf{u}$  eine maximal unabhängige modulo  $J$  Variablenmenge,  $S = k(\mathbf{u})[\mathbf{x}]$  und  $G' = \mathbf{GBasis}_S(F)$  die Gröbnerbasis des nulldimensionalen Ideals  $J' = J \cdot S$ .

Dann gilt für fast alle Spezifikationen  $\mathbf{u} \mapsto \mathbf{u}_0$  der Parameter mit speziellen Werten:

- $F_0 = F|_{\mathbf{u}=\mathbf{u}_0}$  ist ein nulldimensionales algebraisches Gleichungssystem,
- $G_0 = G'|_{\mathbf{u}=\mathbf{u}_0}$  ist dessen Gröbnerbasis über  $S_0 = S|_{\mathbf{u}=\mathbf{u}_0} = k[\mathbf{x}]$  und
- $NF_S(g(\mathbf{x}, \mathbf{u}), G') = 0$  impliziert  $NF_{S_0}(g(\mathbf{x}, \mathbf{u}_0), G_0) = 0$ .

„Fast alle“ bedeutet in diesem Zusammenhang die Existenz eines Polynoms  $h(\mathbf{u})$ , so dass die Aussage höchstens für  $\mathbf{u}_0$  mit  $h(\mathbf{u}_0) = 0$  nicht gilt.

Dieser Satz folgt sofort aus der Existenz des Buchbergeralgorithmus zur Berechnung der Gröbnerbasis  $G'$ . Wählt man die Spezifikation  $\mathbf{u} \mapsto \mathbf{u}_0$  so, dass keiner der Faktoren  $t(\mathbf{u}) \in k(\mathbf{u})$ , durch die Koeffizienten während der Rechnungen geteilt wurden, und keiner der Leitkoeffizienten der  $g \in G'$  verschwindet, so ergibt der Gröbnertrace der Berechnung von  $G'$  aus  $F$  bei Spezialisierung einen Gröbnertrace für die Berechnung von  $G_0$  aus  $F_0$ . Als  $h(\mathbf{u})$  kann man nun einfach das Produkt all dieser polynomialen Bedingungen nehmen.

Angewendet auf Beweisschemata vom Gleichungstyp ergibt sich folgender Satz:

**Satz 26 (Gültigkeit geometrischer Sätze vom Gleichungstyp “im Allgemeinen”)**

Sei  $[F(\mathbf{x}, \mathbf{u}) \Rightarrow g(\mathbf{x}, \mathbf{u})]$  ein Satz vom Gleichungstyp,  $\mathbf{u}$  eine für diesen Satz maximale unabhängige Teilmenge der Variablen und  $S = k(\mathbf{u})[\mathbf{x}]$ .

Gilt  $NF_S(g, \mathbf{GBasis}_S(F)) = 0$ , so gibt es eine (effektiv konstruierbare) Nichtdegenerationsbedingung  $h(\mathbf{u})$ , so dass  $g$  auf allen gemeinsamen Nullstellen  $(\mathbf{x}_0, \mathbf{u}_0)$  von  $F$  mit  $h(\mathbf{u}_0) \neq 0$  verschwindet, also der Satz  $[F/h \Rightarrow g]$  gilt.

Auch die Frage, ob  $\mathbf{u}$  wirklich eine maximale unabhängige Teilmenge der Variablen ist, kann an Hand der Gröbnerbasis  $G'$  entschieden werden. Ist  $G' = \{1\}$  (bzw. enthält ein Polynom aus  $k[\mathbf{u}]$ ), so ist  $J' = Id_S(F)$  das triviale Ideal und  $\mathbf{u}$  war in Wirklichkeit algebraisch abhängig modulo  $J$ . Ansonsten ist  $\mathbf{u}$  algebraisch unabhängig modulo  $J$  und maximal genau dann, wenn  $S/J'$  nulldimensional ist, d.h.  $G'$  zu jeder  $\mathbf{x}$ -Variablen ein Polynom enthält, dessen Leitern eine reine Potenz in dieser Variablen ist.

$h(\mathbf{u})$  im Rahmen fertiger Gröbnerpakete aufzusammeln ist schwierig. Allerdings folgt allein aus der Existenz von  $h \in k[\mathbf{u}]$ , dass eine „zufällige“ Wahl von  $\mathbf{u}_0$  „normalerweise“ zulässig ist. Deshalb auch „Gültigkeit im Allgemeinen“.

Es ist auch plausibel, dass wenigstens für ein Radikalideal  $J'$  aus dem Nichtverschwinden von  $NF_S(g, \mathbf{GBasis}_S(F))$  folgt, dass  $g$  auf wenigstens einer der „allgemeinen“ Nullstellen von  $F$  nicht verschwindet und folglich nicht allgemeingültig ist.

## 5.7 Eine geometrische Interpretation des Spezialisierungssatzes

Um den Zusammenhang zwischen diesen „allgemeinen“ Nullstellen und dem (uns eigentlich interessierenden) Nullstellengebilde  $V_R(F)$  besser zu verstehen, müssen wir den Zusammenhang zwischen Nullstellen  $V_R(F)$  von  $F$  über  $R = k[\mathbf{x}, \mathbf{u}]$  und  $V_S(F)$  über  $S = k(\mathbf{u})[\mathbf{x}]$  genauer studieren.

Der Polynomring  $S$  enthält  $R$  als Unterring. Das von  $F$  in  $S$  erzeugte Ideal  $J' = Id_S(F) = J \cdot S$  bezeichnet man als das *Erweiterungsideal* von  $J$ . Umgekehrt können wir mit einem Ideal  $J' \subset S$  das *Kontraktionsideal*  $J' \cap R$  in  $R$  verbinden, das aus allen rationalen Kombinationen der Erzeugenden besteht, „in denen sich die Nenner wegekürzen“. Sicherlich ist  $J \subset J' \cap R$ .

Die Elemente des Rings  $S$  sind Polynome in den Variablen  $\mathbf{x}$  mit rationalen Funktionen in  $\mathbf{u}$  als Koeffizienten. Durch Hauptnennerbildung überzeugt man sich leicht, dass man jedes Element aus  $S$  in der Form  $\frac{z(\mathbf{x}, \mathbf{u})}{n(\mathbf{u})}$  mit  $z \in R$  und  $n \in U = k[\mathbf{u}] \setminus \{0\}$  darstellen kann. Die Menge  $U$  ist *multiplikativ*, d.h.  $n_1, n_2 \in U$  impliziert  $n_1 n_2 \in U$ , womit die üblichen Rechenregeln für Brüche Anwendung finden können. Wir schreiben deshalb auch

$$S = U^{-1}R := \left\{ \frac{z}{n} : z \in R, n \in U \right\}$$

und nennen den Ring  $S$  die *Lokalisierung* von  $R$  nach der multiplikativen Menge  $U$ .

Zunächst wollen wir das Erweiterungsideal  $J' = J \cdot S$  genauer beschreiben. Dessen Elemente lassen sich in der Form

$$u = \sum_{f \in F} \frac{z_f}{n_f} f = \frac{\sum h_f f}{n}$$

mit  $z_f, h_f \in R$ ,  $n_f, n \in U$  darstellen. Da  $\sum h_f f$  genau die Elemente aus  $J$  sind, erkennen wir, dass  $J'$  aus genau den Elementen von  $S$  besteht, die sich in der Form  $\frac{z}{n}$  mit  $z \in J, n \in U$  darstellen lassen. Entsprechend besteht das Kontraktionsideal  $J' \cap R$  aus allen solchen Elementen, die außerdem noch polynomial auch in  $\mathbf{u}$  sind, d.h. für die zusätzlich  $z$  vollständig durch  $n$  teilbar ist:

$$J' \cap R = \{r \in R : \exists n \in U (n \cdot r \in J)\}.$$

Das erinnert an die Definition des Idealquotienten. Allerdings kommen die Kofaktoren hier nicht aus einem Ideal.

Wir hatten gesehen, dass jedes Radikalideal  $J$  als Durchschnitt endlich vieler Primideale  $J = \bigcap P_\alpha$  darstellbar ist. Eine bzgl.  $J$  unabhängige Variablenmenge  $\mathbf{u}$  muss allerdings bzgl. einer dieser

Primkomponenten, die ja größer sind als  $J$ , nicht mehr unabhängig sein. Wir erkennen das daran, ob  $P_\alpha \cap U = \emptyset$  gilt (in diesem Fall bleibt die Variablenmenge unabhängig) oder nicht (in diesem Fall gibt es eine algebraische Beziehung  $h_\alpha(\mathbf{u}) \in P_\alpha$  zwischen den eigentlich unabhängigen Variablen  $\mathbf{u}$ ). Komponenten der ersten Art nennen wir *generisch*, Komponenten der zweiten Art *speziell*. Zerlegen wir das Nullstellengebilde der Voraussetzungen eines Geometrietheorems in seine Komponenten, so entsprechen letztere gewissen, durch  $h_\alpha(\mathbf{u}) = 0$  beschriebenen, degenerierten Situationen.

**Satz 27**

(1) Sei  $J$  ein Radikalideal und  $J = \cap P_\alpha$  dessen Zerlegung in Primkomponenten. Dann gilt

$$J' = J \cdot S = \bigcap (P_\alpha \cdot S).$$

(2)

$$P_\alpha \cdot S = \begin{cases} \text{ein Primideal } Q_\alpha \subset S \text{ mit } Q_\alpha \cap R = P_\alpha & \text{wenn } P_\alpha \cap U = \emptyset \\ (1) & \text{wenn } P_\alpha \cap U \neq \emptyset \end{cases}$$

Beweis: (1) Wegen  $J \subset \cap P_\alpha$  folgt  $J \cdot S \subset \cap (P_\alpha \cdot S)$ . Sei umgekehrt  $u \in \cap (P_\alpha \cdot S)$ . Nach geeigneter Hauptnennerbildung hat es eine Darstellung der Form  $u = \frac{z}{n}$  mit  $z \in \cap P_\alpha = J$ ,  $n \in U$ , also  $u \in J \cdot S$ .

(2) Sei  $0 \neq s \in P_\alpha \cap U$ . Dann ist  $1 = \frac{s}{s} \in P_\alpha \cdot S$ .

Sei  $P_\alpha \cap U = \emptyset$  und  $Q_\alpha := P_\alpha \cdot S$ .

$Q_\alpha$  ist ein Primideal: Aus

$$u_1 u_2 = \frac{p_1}{n_1} \frac{p_2}{n_2} = \frac{p}{n} \in Q_\alpha$$

mit  $p_1, p_2 \in R$ ,  $p \in P_\alpha$ ,  $n, n_1, n_2 \in U$  folgt  $n \cdot p_1 p_2 \in P_\alpha$  und wegen  $n \notin P_\alpha$  und dessen Primidealeigenschaft schließlich  $p_1 p_2 \in P_\alpha$ .

$Q_\alpha \cap R = \{u \in R : \exists n \in U (n \cdot u \in P_\alpha)\}$ . Wie eben folgt dann bereits  $u \in P_\alpha$ .  $\square$

**Folgerung 1**

$$J' \cap R = \cap \{P_\alpha : P_\alpha \cap U = \emptyset\}$$

Ist insbesondere  $g$  ein Polynom in  $R$ ,  $F \subset R$  eine Menge von Polynomen und  $G' = \text{GBasis}_S(F)$ , so gilt  $NF_S(g, G') = 0$  genau dann, wenn  $g$  auf allen generischen Komponenten von  $V_R(F)$  verschwindet.

Wir sagen in diesem Fall, dass der geometrische Satz bzgl. der unabhängigen Variablenmenge  $\mathbf{u}$  *generisch richtig* ist, was folgendes heißt:

**Satz 28 (Generische Gültigkeit geometrischer Sätze vom Gleichungstyp)**

Sei  $[F(\mathbf{x}, \mathbf{u}) \Rightarrow g(\mathbf{x}, \mathbf{u})]$  ein Satz vom Gleichungstyp,  $\mathbf{u}$  eine für diesen Satz maximale unabhängige Teilmenge der Variablen und  $S = k(\mathbf{u})[\mathbf{x}]$ .

Gilt  $NF_S(g, \text{GBasis}_S(F)) = 0$ , so ist  $[F \Rightarrow g]$  auf allen generischen Komponenten von  $V_R(F)$  richtig.

Die Aussage  $g = 0$  ist höchstens auf speziellen Komponenten von  $V(F)$  falsch, auf denen aber die Variablen  $\mathbf{u}$  nicht mehr unabhängig sind. Jede solche Komponente enthält im definierenden Ideal ein Polynom  $h_\alpha(\mathbf{u})$ . Das Produkt dieser Polynome können wir als Nichtdegenerationsbedingung nehmen.

Ist umgekehrt  $NF_S(g, G') \neq 0$ , so verschwindet  $g$  auf einer der generischen Komponenten nicht, d.h. kann durch keine Nichtdegenerationsbedingung, die nur die unabhängigen Parameter enthält, gerettet werden.

Generell, wenn  $NF_S(g, G')$  nicht verschwindet, wie oben im Beispiel *Miquelscher Punkt*, liefern uns Untersuchungen in  $S$  Aussagen, die das Verhalten auf speziellen Komponenten ausblenden. In diesem Beispiel hat  $V(J)$  zwei generische Komponenten. Eine entspricht dem Schnittpunkt  $F$  der Kreise durch  $A, F, E$  und  $B, F, D$ , die andere dem gesuchten Schnittpunkt  $P$ . Auf der ersten Komponente gilt der Satz nicht, auf der zweiten sehr wohl. Der Gröbnerfaktorisierer zeigt, dass  $V(J)$  daneben noch eine weitere spezielle Komponenten hat, welche die Bedingung  $c_1 = 0$  enthält und einer degenerierten Lagen entspricht. In einem solchen Fall ist aber der Kreis und damit auch der Schnittpunkt  $P$  nicht mehr eindeutig bestimmt, so dass wir auch nicht prüfen können, ob er auf dem dritten Kreis liegt.

## Literatur

- [1] S.-C. Chou. Proving elementary geometry theorems using Wu's algorithm. In *Contemp. Math.*, volume 19, pages 243 – 286. AMS, Providence, Rhode Island, 1984.
- [2] H.S.M. Coxeter and S.L. Greitzer. *Geometry revisited*. Toronto – New York, 1967.
- [3] The GeoProver package for mechanized (plane) geometry theorem proving, version 1.3a, 2003. see <http://www.informatik.uni-leipzig.de/~graebe>.
- [4] The SymbolicData project, 2000–2002. see <http://www.symbolicdata.org>.
- [5] W.-T. Wu. *Mechanical Theorem Proving in Geometries*. Number 1 in Texts and Monographs in Symbolic Computation. Springer, Wien, 1994.

## 6 Aufgaben

Beweisen Sie die folgenden Geometrietheoreme durch Zurückführung auf Sätze vom konstruktiven Typ oder vom Gleichungstyp. Versuchen Sie, ob sie auch elementargeometrische Beweise finden können.

1. Beweisen Sie den Satz vom Höhenfußpunktdreieck: Ist  $\Delta ABC$  ein Dreieck mit den Höhenfußpunkten  $D, E, F$ , so wird der Winkel  $\angle DEF$  von der Höhe durch  $E$  halbiert.  
Finden Sie auch die entsprechenden Nicht-Degenerations-Bedingungen.
2. Zeigen Sie, dass die 6 Fußpunkte der Lote von den Höhenfußpunkten  $D, E, F$  auf die gegenüberliegenden Dreiecksseiten (oder deren Verlängerungen) auf einem gemeinsamen Kreis, dem *Taylorkreis*, liegen.
3. Beweisen Sie die Umkehrung des Satzes von der Simonschen Geraden: Sind die Lotfußpunkte  $R, S, T$  von einem Punkt  $P$  auf die Seiten des Dreiecks  $\Delta ABC$  oder deren Verlängerungen kollinear, so liegt  $P$  auf dem Umkreis des Dreiecks.
4. Beweisen Sie auch folgende Verallgemeinerung:  
Ist  $M$  der Umkreismittelpunkt des Dreiecks  $\Delta ABC$  und  $P, R, S, T$  wie oben, so hängt der Flächeninhalt  $F(\Delta RST)$  nur von  $|MP|$  ab. Finden Sie eine genaue Formel für diesen Flächeninhalt.
5. Beweisen Sie die Flächeninhaltsformel

$$F(\Delta ABC) = \frac{abc}{4R},$$

wobei  $a, b, c$  die Seitenlängen und  $R$  der Umkreisradius ist.

6. Beweisen Sie die Flächeninhaltsformel

$$F(\Delta ABC) = \rho \cdot \frac{a+b+c}{2},$$

wobei  $a, b, c$  die Seitenlängen und  $\rho$  der Inkreisradius ist.

Wie ist die Formel für die Ankreisradien  $\rho_A, \rho_B$  und  $\rho_C$  zu modifizieren?

7. Gegeben sei ein Kreis  $k$ , die Tangente von  $B$  an diesen Kreis,  $A$  der Berührungspunkt,  $M$  der Mittelpunkt von  $AB$  und  $D$  ein Punkt auf dem Kreis  $k$ .  $C$  sei der Schnittpunkt von  $DM$  mit  $k$ ,  $E$  der Schnittpunkt von  $BD$  mit  $k$  und  $F$  der Schnittpunkt von  $BC$  mit  $k$ . Zeigen Sie, dass  $EF$  parallel zu  $AB$  ist.
8. Beweisen Sie, dass die Lote durch die drei Ankreiszentren auf die jeweilige Dreiecksseite durch einen gemeinsamen Punkt gehen.