

Skript zum Kurs  
Geometrie mit dem Computer  
Sommersemester 2018

Hans-Gert Gräbe, Univ. Leipzig  
<http://bis.informatik.uni-leipzig.de/HansGertGraebe>

11. Juli 2018

## 0 Einführung

Die (synthetische) Geometrie ist eine sehr alte mathematische Disziplin und stand – neben grundlegenden Fragen der Zahlentheorie – lange Zeit wohl für Mathematik schlechthin, ehe ihr dieser Platz durch eine stürmische Entwicklung und Ausdifferenzierung der Mathematik in den letzten 200 Jahren von anderen Disziplinen streitig gemacht wurde. Natürlich hat sich in dieser Zeit auch die Geometrie weiterentwickelt. Teildisziplinen wie Differentialgeometrie oder Algebraische Geometrie untersuchen komplizierte, stark nichtlineare 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 welchem (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 erforderlich sind. Sie bieten damit für Hobbymathematiker, interessierte Schülerinnen und Schüler eingeschlossen, immer wieder eine Fundgrube von Problemen und Ideen, an denen die eigene Argumentationskraft trainiert und verbessert werden 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 Carl Friedrich 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 genauer ausgeführt.

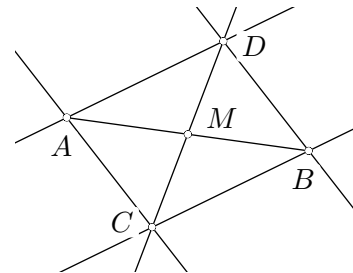
Ein exaktes Studium der mit der Konstruierbarkeit verbundenen Fragestellungen kommt um **eine ordentliche Fundierung, eine Axiomatisierung der Geometrie** nicht herum. Auch hier lassen sich die entsprechenden Ansätze bis in die Antike hinein, etwa zu den Büchern des Euklid, verfolgen. Mathematiker hat 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, werden wir derartige geometrische Aussagen im Weiteren aus unseren Betrachtungen ausklammern.

Damit wird der Kreis der zu untersuchenden geometrischen Problemstellungen 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 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.



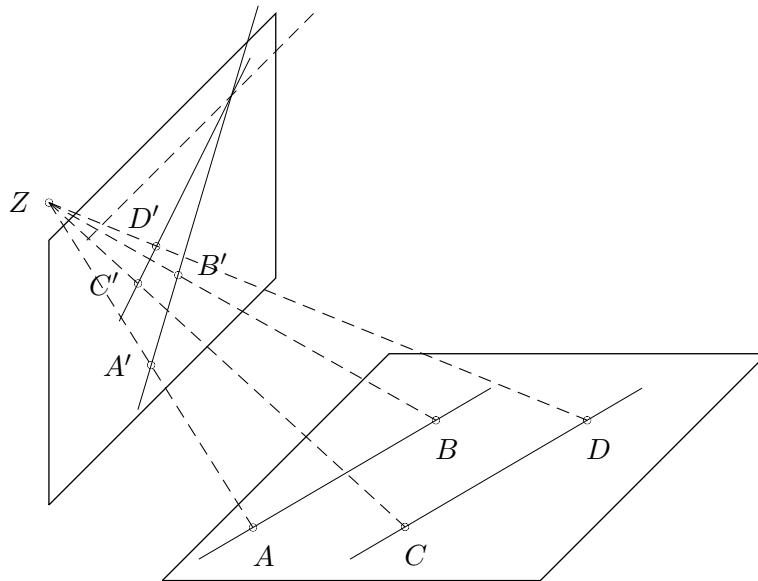
Affine Geometrie: Konstruktion des Mittelpunkts einer Strecke

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. Eine etwas eingeschränktere Geometrie erhalten wir, wenn wir Winkel-treue, also Winkelgrößen und deren Erhaltung, nicht aber die Erhaltung von Streckenlängen fordern. Diese Geometrie bezeichnet man als **Ähnlichkeitsgeometrie**.

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

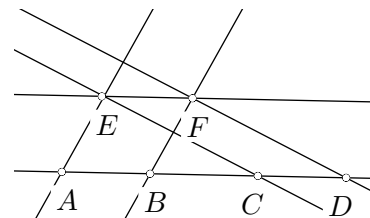


Projektive Abbildung – Die Bildgeraden paralleler Geraden schneiden sich auf der Ausnahmegeraden der Bildebene

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 Ausnahme Geraden 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 Ausnahme Geraden 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 Felix Klein und David 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 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 Gruppe der affinen Transformationen, die orthonormale Koordinatensysteme in schiefwinklige überführt und auch die „Längen“ der Einheitsstrecken nicht erhält (aber Längen gibt es in der affinen Geometrie 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 Eindeutigkeitssatz 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 der projektiven Transformationen 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 informatik-theoretischen 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 auch einen starken informatik-praktischen Bezug haben wird. Wir werden parallel zu den mathematischen Fragen auch die Modellierung in entsprechender Dynamischer Geometrie-Software (DGS) studieren, wozu wir das Programm GEOGEBRA <http://www.geogebra.org> einsetzen werden. Dies ist keine Einschränkung, da die grundlegende Struktur all dieser Programme ähnlich ist.

## 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

**Satz 1** Die Mittelsenkrechten  $m_{AB}$ ,  $m_{AC}$  und  $m_{BC}$  eines Dreiecks  $ABC$  schneiden sich in einem Punkt.

Der Beweis verwendet den Begriff der *Ortslinie*:

Die Mittelsenkrechte  $m_{AB}$  besteht aus genau den Punkten  $P$  der Ebene, die von  $A$  und  $B$  den gleichen Abstand haben.

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.

*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. Im Folgenden bezeichnet  $d(M, A)$  den Abstand der beiden Punkte.

$$M \in m_{AB} \Rightarrow d(M, A) = d(M, B)$$

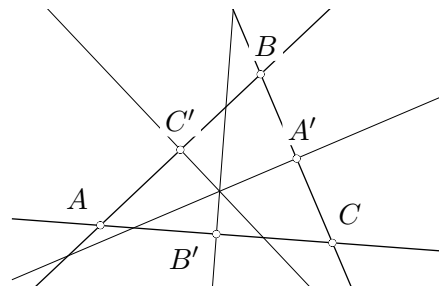
$$M \in m_{BC} \Rightarrow d(M, B) = d(M, C)$$

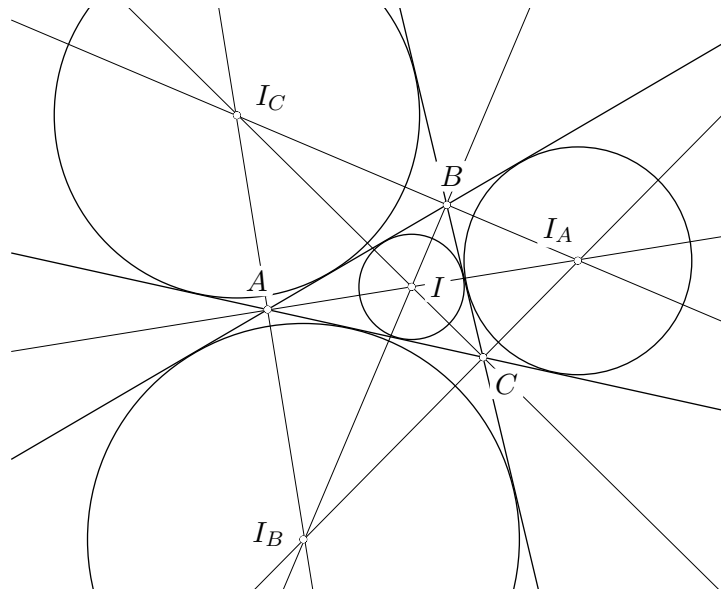
Daraus folgt  $d(M, A) = d(M, C)$ , also  $M \in m_{AC}$ .  $\square$

Aus dem Beweis ergibt sich außerdem, dass der Schnittpunkt  $M$  von allen drei Eckpunkten gleichweit entfernt ist, also der Umkreismittelpunkt sein muss.

#### Satz vom Schnittpunkt der Winkelhalbierenden

**Satz 2** Die Winkelhalbierenden  $w_A, w_B$  und  $w_C$  eines Dreiecks  $ABC$  schneiden sich in einem Punkt.





Inkreis und Ankreise eines Dreiecks

Auch hier spielen Ortslinien eine Rolle:

Die Winkelhalbierende  $w_A$  besteht aus genau den Punkten  $P$ , die von den beiden Schenkeln des Winkels mit Scheitel  $A$  gleichen Abstand haben.

Gehen wir wieder von einem Dreieck  $ABC$  aus, so kann hierzu zunächst das GEOGEBRA-Werkzeug *Winkelhalbierende des durch drei Punkte gegebenen Winkels*  $\angle ABC$  mit Scheitel  $B$  verwendet werden und wir argumentieren für den Schnittpunkt  $I$  der Winkelhalbierenden durch  $A$  und  $B$  wie eben, wobei  $d(I, AB)$  den Abstand von  $I$  zum Schenkel  $AB$  des Winkels bezeichnet.

$$\begin{aligned} I \in w_A &\Rightarrow d(I, AB) = d(I, AC) \\ I \in w_B &\Rightarrow d(I, BA) = d(I, BC) \end{aligned}$$

Aus dem Beweis ergibt sich dann außerdem, dass der Schnittpunkt  $I$  von allen drei Dreiecksseiten gleich weit entfernt, also der Inkreismittelpunkt ist.

Hier sind allerdings subtile Fragen zu berücksichtigen, denn eigentlich geht dieser Winkelbegriff von Schenkeln und damit Strahlen aus und ist damit ein Begriff der Ordnungsgeometrie. Zur Konstruktion der Winkelhalbierenden in GEOGEBRA können drei Punkte *oder* zwei Geraden angegeben werden. Im zweiten Fall werden *zwei* Winkelhalbierende konstruiert, da sich für zwei Geraden Winkel und Nebenwinkel nicht unterscheiden lassen. Auch die Beschreibung des geometrischen Orts ist zu präzisieren:

Der geometrische Ort der Punkte, die von zwei sich schneidenden Geraden gleichen Abstand haben, ist das *Winkelhalbierendenpaar*.

Die beiden Geraden dieses Paares kann man erst in einer Ordnungsgeometrie voneinander unterscheiden – die Innenwinkelhalbierende im Dreieck schneidet die Gegenseite in einem inneren Punkt, die Außenwinkelhalbierende ist parallel (im gleichschenkligen Dreieck) oder schneidet die Verlängerung der Gegenseite in einem äußeren Punkt.

Verwenden wir Winkelhalbierendenpaare und  $d(I, AB)$  als Abstand von  $I$  zur Geraden  $AB$ , so haben  $w_A$  und  $w_B$  insgesamt vier Schnittpunkte. Ist  $I$  einer von ihnen, so geht die bisherige Argumentation durch und es kann gezeigt werden, dass  $I$  dann auch auf dem Winkelhalbierendenpaar  $w_C$  liegt. Wir haben neben dem Inkreismittelpunkt  $I$  so auch die Mittelpunkte der drei Ankreise  $I_A, I_B$  und  $I_C$  gefunden.

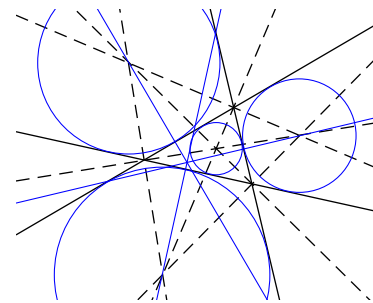
**Satz 3** Die Winkelhalbierendenpaare  $w_A, w_B$  und  $w_C$  eines Dreiecks  $ABC$  haben vier Punkte gemeinsam, den Inkreismittelpunkt  $I$  sowie die Mittelpunkte  $I_A, I_B$  und  $I_C$  der drei Ankreise.

Man beachte die Ähnlichkeit zum Begriff des *Parallelenpaares* 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!

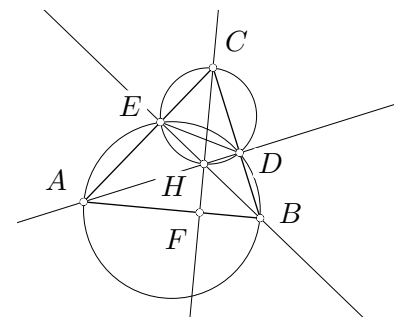


**Satz vom Höhenschnittpunkt**

**Satz 4** Die Höhen eines Dreiecks  $ABC$  gehen durch einen gemeinsamen Punkt, das Orthozentrum.

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.



Ein anderer Beweis des Satzes: Die Parallelen zu den Dreiecksseiten durch die gegenüberliegenden Eckpunkte spannen ein Dreieck  $A'B'C'$  auf. Zeigen Sie, dass die Höhen im Dreieck



$ABC$  gerade die Mittelsenkrechten im Dreieck  $A'B'C'$  sind.

Noch ein Beweis des Satzes: Die Fußpunkte der Höhen im Dreieck  $ABC$  spannen ein Dreieck  $DEF$  auf, das *Höhenfußpunktdreieck*. Zeigen Sie, dass die Höhen im Dreieck  $ABC$  gerade die (Innen)-Winkelhalbierenden im Dreieck  $DEF$  sind.

## 1.2 Der Satz von Ceva

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

**Satz 5 (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:* Beweis durch Flächenzerlegung. Wir bezeichnen dazu die Streckenlängen wie im folgenden Bild mit  $a_1, \dots, c_2$ . Gehen die drei Transversalen durch einen gemeinsamen Punkt  $S$ , so gilt

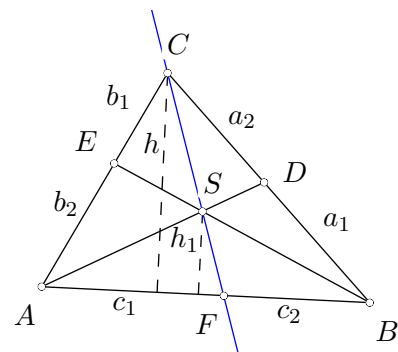
$$\begin{aligned} \frac{c_1}{c_2} &= \frac{F(ASF)}{F(BSF)} = \frac{F(ACF)}{F(BCF)} \\ &= \frac{F(ACF) - F(ASF)}{F(BCF) - F(BSF)} = \frac{F(ASC)}{F(BSC)} \end{aligned}$$

und analog

$$\frac{b_1}{b_2} = \frac{F(BSC)}{F(BSA)}, \quad \frac{a_1}{a_2} = \frac{F(ASB)}{F(ASC)},$$

also insgesamt

$$\frac{a_1}{a_2} \cdot \frac{b_1}{b_2} \cdot \frac{c_1}{c_2} = 1.$$



Ist umgekehrt

$$\frac{a_1}{a_2} \cdot \frac{b_1}{b_2} \cdot \frac{c_1}{c_2} = 1$$

und  $S$  der Schnittpunkt der Transversalen durch  $A$  und  $B$ , so schneidet die Gerade  $CS$  die Dreiecksseite  $AB$  in einem Punkt  $F'$  und dieser teilt die Strecke  $\overline{AB}$  in Abschnitte der Längen  $c'_1$  und  $c'_2$ . Nach der bereits bewiesenen Richtung gilt für die drei Transversalen  $AD, BE, CF'$  durch  $S$

$$\frac{a_1}{a_2} \cdot \frac{b_1}{b_2} \cdot \frac{c'_1}{c'_2} = 1$$

und damit

$$TV(A, B; F) = \frac{c_1}{c_2} = \frac{c'_1}{c'_2} = TV(A, B; F').$$

Hier bezeichnet  $TV(A, B; F)$  das *Teilverhältnis*, in welchem  $F$  die Strecke  $\overline{AB}$  teilt. Die Größe des Teilverhältnisses bestimmt die Lage des Punkts auf  $\overline{AB}$  eindeutig, so dass  $F = F'$  folgt.  $\square$

Der hier geführte Beweis ist erst einmal nur für Transversalen schlüssig, die sich im Inneren des Dreiecks  $ABC$  schneiden. Wir hatten aber schon gesehen, dass  $t = TV(A, B; F) = \frac{c_1}{c_2}$  in Wirklichkeit eine affine Größe ist, deren Bestimmung ohne Längenmessung auskommt. Durch Abtragung rationaler (und reeller) Vielfacher einer Basislänge auf der Geraden  $AB$  kann  $t$  aus der Beziehung  $\overrightarrow{AF} = t \cdot \overrightarrow{FB}$  der gleich- ( $t > 0$ ) oder entgegengerichteten ( $t < 0$ ) kollinearen Vektoren bestimmt werden.

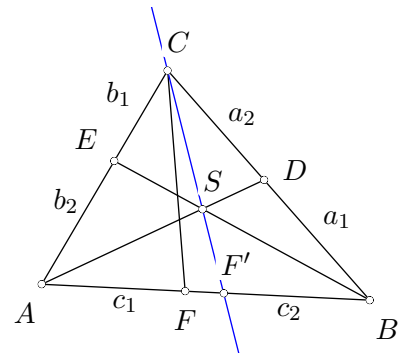
Für Punkte innerhalb der Strecke  $\overline{AB}$  gilt  $t > 0$  und  $t$  wächst monoton im Bereich  $0 < t < +\infty$ , wenn  $F$  von  $A$  in Richtung  $B$  wandert. Liegt  $F$  auf  $AB$ , aber außerhalb der Strecke  $\overline{AB}$ , so ergeben sich jenseits von  $A$  (monoton wachsend) Werte  $-1 < t < 0$ , für Punkte jenseits von  $B$  (ebenfalls monoton wachsend) ergibt sich  $-\infty < t < -1$ . 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  $F_{AB}(t = -1)$  dienen.

Ersetzt man in der Formulierung des Satzes von Ceva die Quotienten der Streckenlängen durch die Teilverhältnisse

$$\frac{|BD|}{|DC|} \cdot \frac{|CE|}{|EA|} \cdot \frac{|AF|}{|FB|} \Rightarrow TV(B, C; D) \cdot TV(C, A; E) \cdot TV(A, B; F),$$

so erhält man eine Formulierung des Satzes von Ceva, die für beliebige Punkte  $D \in BC, E \in AC, F \in AB$  gilt.

Als Folgerung aus dem Satz von Ceva ergeben sich neue Beweise der Sätze vom Schnittpunkt der Seitenhalbierenden und vom Höhenschnittpunkt.



*Aufgabe:*

- 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.
- Zeigen Sie, dass sich die Transversalen zu den Berührungspunkten des Inkreises in einem Punkt schneiden.
- Zeigen Sie, dass sich die Transversalen zu den Berührungspunkten der Ankreise in einem Punkt schneiden. Dieser Punkt heißt auch *Nagelscher Punkt*.

### 1.3 Weitere Sätze am Dreieck

Zwischen affiner und Euklidischer Geometrie lässt sich eine weitere Klasse von geometrischen Sätzen abgrenzen, in denen nur die Winkeltreue der Abbildungen gefordert wird. Neben den Abbildungen der Euklidischen Geometrie der Ebene (Verschiebungen, Drehungen Spiegelungen) sind dabei zusätzlich noch Streckungen zu berücksichtigen. Die entsprechende Klasse geometrischer Sätze wird auch als *Ähnlichkeitsgeometrie* bezeichnet.

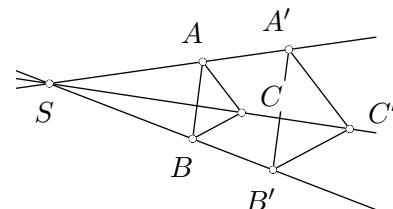
Bei einer zentrischen Streckung mit Zentrum  $S$  und Streckungsfaktor  $t$  wird ein Punkt  $A$  auf einen Punkt  $A'$  auf der Geraden  $SA$  mit  $\overrightarrow{SA'} = t \cdot \overrightarrow{SA}$  abgebildet. Ist  $B$  ein weiterer Punkt, so gilt  $AB \parallel A'B'$  nach Strahlensatz. Geraden werden bei Streckungen also auf parallele Bildgeraden abgebildet.

Man kann fragen, ob es umgekehrt zu zwei Dreiecken  $ABC$  und  $A'B'C'$ , in denen zugehörige Seiten zueinander parallel sind, stets eine Streckung gibt, mit welcher das eine Dreieck in das andere überführt werden kann. Dreiecke (und allgemeiner geradlinig begrenzte Figuren) mit der Eigenschaft, dass entsprechende (man sagt auch: homologe) geradlinige Stücke zueinander parallel sind, bezeichnet man als zueinander *in Ähnlichkeitslage* liegend.

Wenn es eine solche Streckung gibt, so ist deren Zentrum  $S$  ein gemeinsamer Punkt der Geraden  $AA'$ ,  $BB'$  und  $CC'$ . Dazu muss natürlich gezeigt werden, dass für Dreiecke in Ähnlichkeitslage diese drei Verbindungsgeraden durch einen gemeinsamen Punkt gehen. Diese Aussage ist Gegenstand von Teil 1 des folgenden Satzes:

#### Satz 6 (Affiner Satz von Desargue)

- 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 die Geraden  $AA'$ ,  $BB'$  und  $CC'$  konkurrent.
- Sind umgekehrt  $AA'$ ,  $BB'$  und  $CC'$  konkurrent und  $AB \parallel A'B'$ ,  $AC \parallel A'C'$ , so gilt auch  $BC \parallel B'C'$ .



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

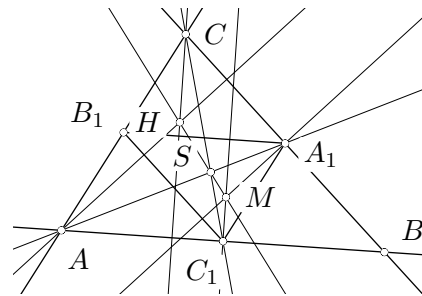
Der Satz von Desargue spielt in der Fundierung der Geometrie als „schwache Version des Strahlensatzes“ eine wichtige Rolle. Er ist schwächer als der Strahlensatz und führt damit zu einer umfassenderen als der affinen Geometrie. Details finden sich im Buch [5].

Mit diesem Konzept der Ähnlichkeitsabbildungen lassen sich weitere geometrische Sätze elegant beweisen.

**Satz 7 (Eulersche Gerade)** *In einem Dreieck liegen Höhenschnittpunkt  $H$ , Schwerpunkt  $S$  und Umkreismittelpunkt  $M$  auf einer Geraden.  $S$  teilt  $\overline{HM}$  im Verhältnis 2:1.*

*Beweis:*  $A_1, B_1, C_1$  seien die Mitten der Seiten  $\overline{BC}$ ,  $\overline{AC}$  und  $\overline{AB}$ ,  $\triangle A_1B_1C_1$  also das *Mittendreieck* des Dreiecks  $ABC$ .

Dieses Mittendreieck liegt in Ähnlichkeitslage zum Ausgangsdreieck, geht also durch eine Streckung um einen Punkt  $S$  aus dem Ausgangsdreieck hervor. Durch dieses Streckungszentrum verlaufen alle Verbindungsgeraden zwischen Urbild und Bild, also insbesondere  $AA_1, BB_1$  und  $CC_1$ , so dass  $S$  der Schnittpunkt der Seitenhalbierenden sein muss (und wir zugleich ein weiteres Mal gezeigt haben, dass die Seitenhalbierenden eines Dreiecks durch einen gemeinsamen Punkt gehen).



Da die Seiten des Mittendreiecks gerade halb so lang sind wie die Seiten des Ausgangsdreiecks, kann auch der Streckungsfaktor zu  $t = -\frac{1}{2}$  bestimmt werden. Daraus ergibt sich unmittelbar, dass  $S$  die Seitenhalbierenden im Verhältnis 2:1 teilt.

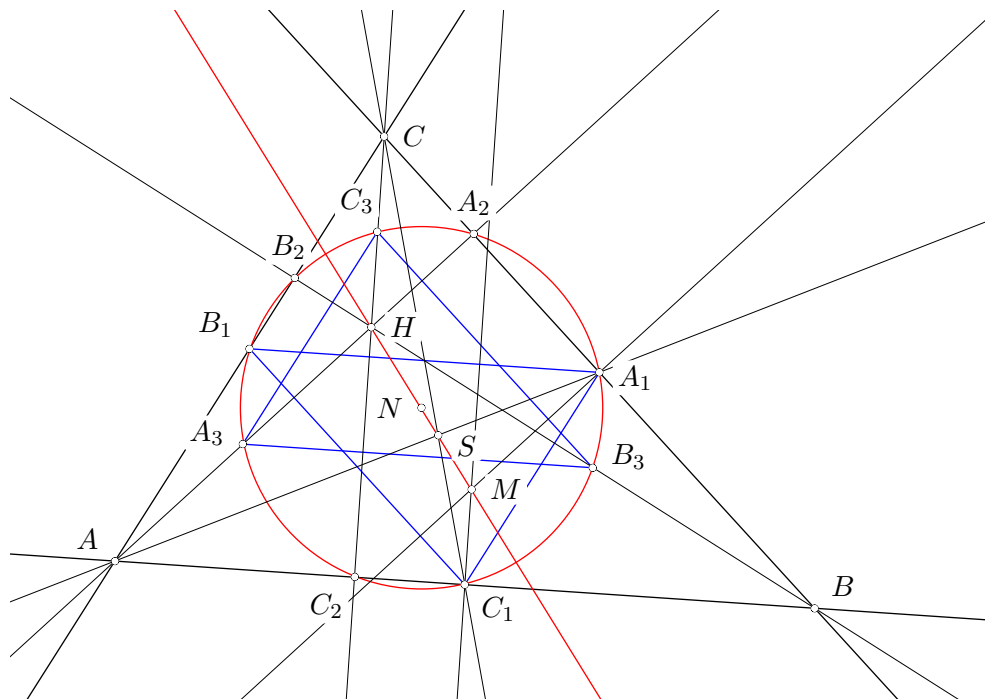
Bei der Streckung um  $S$  geht weiterhin der Höhenschnittpunkt  $H$  des Ausgangsdreiecks in den Höhenschnittpunkt  $H' = M$  des Mittendreiecks über. Die Strecke  $\overline{HH'} = \overline{HM}$  geht also ebenfalls durch das Streckungszentrum  $S$  und wird von  $S$  im Verhältnis 2:1 geteilt.  $\square$

Die im Beweis konstruierte Figur mit zwei Dreiecken in Ähnlichkeitslage kann noch um ein weiteres Dreieck in Ähnlichkeitslage ergänzt werden – das Dreieck, welches von den Mitten der oberen Höhenabschnitte aufgespannt wird. Die genauere Analyse dieser Dreiecke führt auf einen Kreis, auf dem neun markante Punkte des Dreiecks  $ABC$  liegen. In der deutschsprachigen Literatur wird dieser Kreis als Feuerbachkreis, in der englischsprachigen Literatur als 9-Punkte-Kreis (nine point circle) bezeichnet.

**Satz 8 (Der Satz vom Feuerbachschen Kreis)** *Der Mittelpunkt  $N$  von  $\overline{HM}$  ist der Mittelpunkt eines Kreises, auf dem neun ausgezeichnete Punkte des Dreiecks  $ABC$  liegen, und zwar*

- die drei Seitenmitten  $A_1, B_1$  und  $C_1$ ,
- die drei Höhenfußpunkte  $A_2, B_2$  und  $C_2$  und
- die drei Mitten der oberen Höhenabschnitte  $A_3, B_3$  und  $C_3$ .

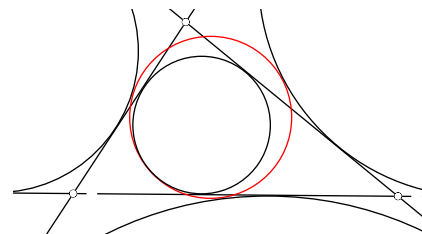
*Beweis:* Die beiden Dreiecke, die durch die Seitenmitten bzw. die Mitten der oberen Höhenabschnitte aufgespannt werden, sind in Ähnlichkeitslage mit dem Faktor  $(-1)$ , also zueinander kongruent. Da dabei  $H$  als Höhenschnittpunkt des Höhendreiecks in  $M$  als Höhenschnittpunkt des Mittendreiecks übergeht, ist die Mitte  $N$  der Strecke  $\overline{MH}$  gerade das Zentrum



Eulersche Gerade und Feuerbachscher Kreis

der zugehörigen Streckung, die in diesem Fall eine Drehung um  $N$  um  $180^\circ$  ist. Das Dreieck  $ABC$  geht bei der 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  $N$  durch die genannten sechs Punkte. Weiter entsprechen sich bei der Punktspiegelung mit Zentrum in  $N$  Seitenmitte  $C_1$  und gegenüberliegende Mitte  $C_3$  des oberen Höhenabschnitts. Die Verbindungsgerade  $C_1C_3$  geht also durch das Streckungszentrum  $N$  und ist ein Durchmesser des Feuerbachkreises. Aus dem Satz des Thales folgt schließlich, dass auch der Höhenfußpunkt  $C_2$  auf dem Feuerbachkreis liegt. Dasselbe gilt für  $A_2$  und  $B_2$ .  $\square$

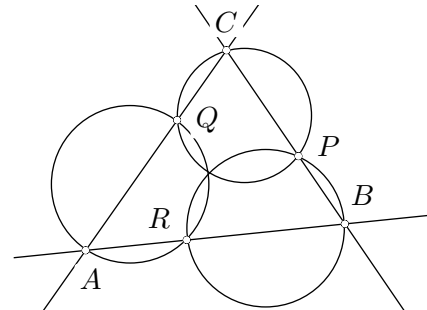
Der Feuerbachkreis hat eine weitere, mit elementargeometrischen Mitteln nur schwer zu beweisende Eigenschaft: Er berührt den Inkreis und die drei Ankreise des Ausgangsdreiecks.



### 1.4 Miquelscher Punkt und Simonsche Gerade

#### Satz 9 (Miquelscher Punkt)

$P, Q, R$  seien Punkte auf den Seiten des Dreiecks  $ABC$ . 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 Beweis ergibt sich unmittelbar aus dem Satz über gegenüberliegende Winkel im Sehnenviereck und dessen Umkehrung.

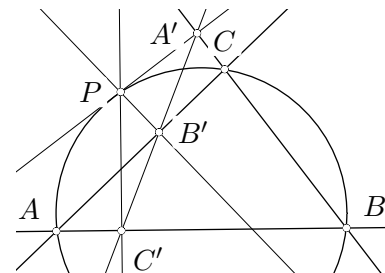
#### Satz 10 (Simonsche Gerade)

Fällt man von einem Punkt  $P$  außerhalb eines Dreiecks  $ABC$  die Lote auf die Dreiecksseiten oder deren Verlängerungen, so liegen die drei Fußpunkte  $A', B'$  und  $C'$  der Lote genau dann auf einer Geraden, wenn  $P$  auf dem Umkreis des Dreiecks  $ABC$  liegt.

*Beweis:* Für beliebige  $P$  liefert der Thaleskreis über  $\overline{PC}$  die Winkelgleichheit  $|\angle B'A'P| = |\angle B'CP|$  und der Thaleskreis über  $\overline{PA}$  analog  $|\angle B'C'P| = |\angle B'AP|$ .

Liegt  $P$  auf dem Umkreis, so ist  $ABCP$  ein Sehnenviereck und es gilt  $|\angle APC| = 180^\circ - \beta$ . Wegen zweier rechter Winkel gilt aber auch  $|\angle A'PC'| = 180^\circ - \beta$ . Folglich ist  $|\angle A'B'C'| = |\angle AB'C| = 180^\circ$ , die Punkte  $A', B', C'$  sind also kollinear.

Sind  $A', B', C'$  kollinear, so schließt man rückwärts längs derselben Schlusskette, dass  $ABCP$  ein Sehnenviereck ist.  $\square$



Zu beiden Sätzen gäbe es noch eine Menge zu sagen. So kann man etwa zu jedem Punkt  $P$  im Inneren des Dreiecks  $ABC$  Punkte  $A', B', C'$  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:*

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  $ABC$ ,  $r$  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  $A'B'C'$  bzgl.  $P$

$$F(A'B'C') = \frac{r^2 - |MP|^2}{4r^2} \cdot F(ABC)$$

gilt, wobei  $M$  der Umkreismittelpunkt ist.

Das eben betrachtete Fußpunktdreieck entartet zu einer Geraden, wenn  $P$  auf dem Umkreis des Dreiecks  $ABC$  liegt und ergibt dann genau die Simsonsche Gerade.

*Aufgabe:* Leiten Sie aus der Flächeninhaltsformel für das Fußpunktdreieck (vorige Aufgabe) einen zweiten Beweis für den Satz über die Simsonsche Gerade her.

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 ausüben.

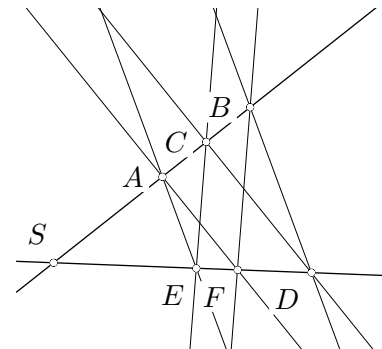
**Satz 11 (Theorem von Pappus)** Sind  $A, B, C$  und  $D, E, F$  jeweils kollineare Punkte, so sind auch die Schnittpunkte  $G = AE \cap BD$ ,  $H = BF \cap CE$  und  $K = CD \cap AF$  kollinear. Diese Gerade wird als Pappus-Gerade bezeichnet.

*Beweis:*

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

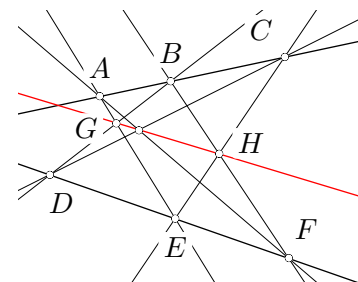
Ist in obiger Situation  $AE \parallel BD$  und  $BF \parallel CE$ , so gilt auch  $CD \parallel AF$ .

Der Beweis ergibt sich unmittelbar aus dem Strahlensatz: Wegen  $AE \parallel BD$  ist  $\frac{|SA|}{|SB|} = \frac{|SE|}{|SD|}$  und wegen  $BF \parallel CE$  ist  $\frac{|SB|}{|SF|} = \frac{|SC|}{|SE|}$ . Daraus ergibt sich  $\frac{|SA|}{|SF|} = \frac{|SC|}{|SD|}$  und weiter  $CD \parallel AF$ .



Für den allgemeinen Beweis führen wir eine projektive Transformation  $\pi$  unserer Ebene  $\varepsilon$  in eine Bildebene  $\varepsilon'$  aus, bei der die Gerade  $GH$  die Ausnahme Gerade in  $\varepsilon$  ist, also in die Ferngerade von  $\varepsilon'$  abgebildet wird. Dazu muss das Projektionszentrum  $Z$  nur so gewählt werden, dass  $\varepsilon$  parallel zur Ebene durch  $Z, G, H$  verläuft. Das ist immer möglich.

Dann gilt für die Bildpunkte aber  $A'E' \parallel B'D'$  und  $B'F' \parallel C'E'$ , womit die Voraussetzungen des bewiesenen Spezialfalls für die Bildkonfiguration in  $\varepsilon'$  erfüllt sind.



Also gilt auch  $C'D' \parallel A'F'$ , d.h.  $K$  wird unter  $\pi$  ebenfalls auf die Ferngerade abgebildet. Damit muss aber auch  $K$  auf der Ausnahmegeraden liegen. Die drei Schnittpunkte  $G$ ,  $H$  und  $K$  sind kollinear.  $\square$

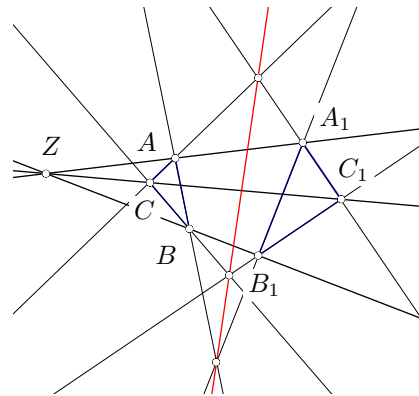
*Aufgabe:* Führen Sie den Beweis des Spezialfalls aus, wenn der Schnittpunkt  $S$  nicht existiert, die beiden Ausgangsgeraden also parallel sind.

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

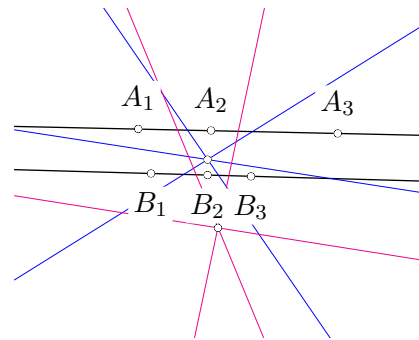
**Satz 12 (Allgemeiner Satz von Desargue)**

Für sechs Punkte  $A, B, C, A_1, B_1, C_1$  gilt:  
 Die Schnittpunkte  $AB \cap A_1B_1$ ,  $AC \cap A_1C_1$  und  $BC \cap B_1C_1$  sind genau dann kollinear, wenn die drei Geraden  $AA_1$ ,  $BB_1$  und  $CC_1$  konkurrent sind, also durch einen gemeinsamen Punkt  $Z$  gehen oder parallel sind.

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

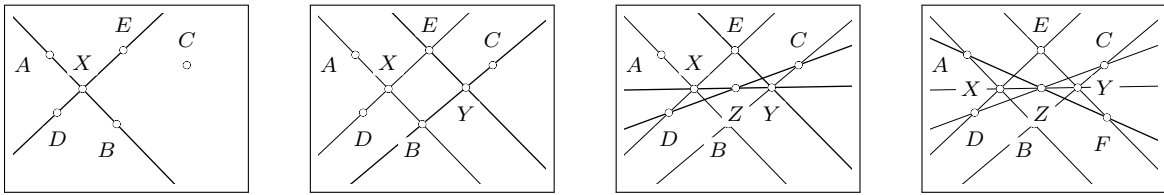


Eine interessante Fragestellung, die wir zum selben Thema hier nur aufwerfen wollen, entsteht aus dem Vergleich verschiedener Pappus-Geraden. Sind  $A_1, A_2, A_3$  und  $B_1, B_2, B_3$  jeweils kollinear, so führen die verschiedenen Permutationen der Punkte  $B_1, B_2, B_3$  zu insgesamt sechs solchen Geraden. Die 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 und die anderen drei Geraden durch einen anderen gemeinsamen Punkt gehen.



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 können die ersten fünf Punkte frei gewählt werden, wobei auch die Lage von  $X$  bestimmt wird (Bild 1).  $Y$  als Gleiter auf  $BC$  bestimmt die Richtung der Geraden  $g = EF$  (Bild 2) und die Lage von  $Z$  als Schnittpunkt  $CD \cap XY$  (Bild 3) eindeutig.  $F$  ergibt sich dann als Schnittpunkt  $AZ \cap g$  (Bild 4). Eine solche Punktekonfiguration bezeichnet man als *Pascalsches Sechseck*, die Gerade durch  $X, Y, Z$  als die zugehörige *Pascalgerade*.



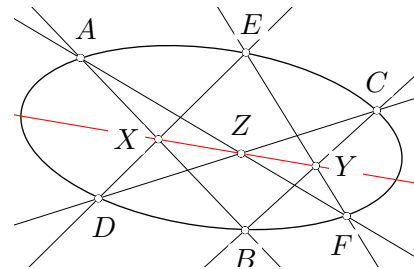


Konstruktion eines Pascalschen Sechsecks  $ABCDEF$

Mit dem weiteren Konzept von Kurven zweiten Grades (Quadriken) stellt sich heraus, dass sich solche Punktekonfigurationen analytisch recht einfach charakterisieren lassen. Es gilt der folgende

**Satz 13 (Satz von Pascal)** *Sechs Punkte bilden genau dann ein Pascalsches Sechseck, wenn sie auf einer Kurve zweiten Grades liegen.*

So wie durch drei gegebene, voneinander verschiedene (nicht kollineare) Punkte stets eindeutig ein Kreis konstruiert werden kann, existiert durch fünf (allgemein gelegene) Punkte stets genau eine Quadrik. Pascalsche Sechsecke verallgemeinern damit die Notation des Sehnenvierecks.



$q$  sei die Quadrik durch  $A, B, C, D, E$ .  $q$  lässt sich als ebene Schnittfigur eines dreidimensionalen Doppelkegels mit Spitze  $Z$  darstellen. Durch eine projektive Transformation kann  $q$  in einen Kreis überführt werden (man nehme  $Z$  als Projektionszentrum und eine Bildebene senkrecht zur Kegelachse), so dass sich der Satz von Pascal auf den folgenden Spezialfall reduzieren lässt:

**Satz 14** *Sind  $A, B, C, D, E$  fünf Punkte auf einem Kreis  $k$ ,  $F$  ein weiterer Punkt und  $X, Y, Z$  wie oben die Schnittpunkte der entsprechenden Verbindungsgeraden, so gilt:  $X, Y, Z$  sind genau dann kollinear, wenn  $F$  auf  $k$  liegt.*

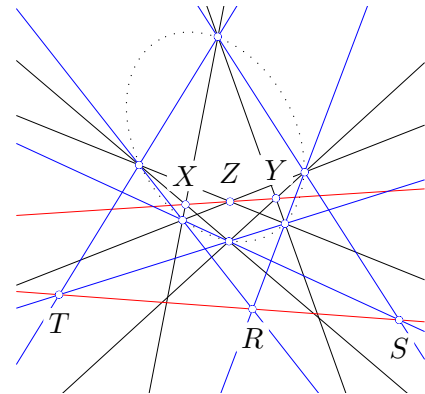
Auf einen vollständigen Beweis dieses Satzes muss hier verzichtet werden.

*Aufgabe:* Beweisen Sie den folgenden Spezialfall des Spezialfalls: Liegen die Punkte  $A, B, C, D$  und  $E$  auf einem Kreis  $k$  und gilt  $AB \parallel CD$  sowie  $BC \parallel EF$ , so ist  $CD \parallel EF$  genau dann, wenn  $F$  auf  $k$  liegt.

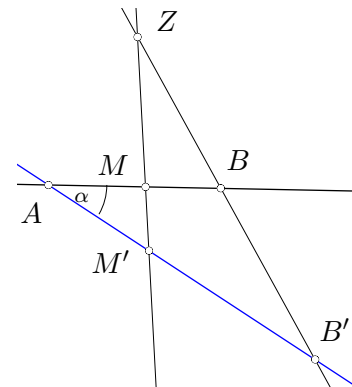
Für beliebige sechs Punkte  $A, B, C, D, E, F$  bezeichnen wir die Gerade durch  $X = AB \wedge DE$  und  $Y = BC \wedge EF$  als deren *Pascalgerade*. Die Punkte bilden genau dann ein Pascalsches Sechseck, wenn auch  $Z = CD \wedge AF$  auf dieser Pascalgeraden liegt.

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 weitere Aussagen der ebenen (projektiven) Geometrie, zu deren Formulierung keine Kurven zweiter Ordnung benötigt werden. Solche Sätze heißen in [5] *Sätze vom Pascal-Typ*. Ein solcher Satz wäre z. B. die folgende Aussage:

**Satz 15** Sind  $X = AB \cap DE$ ,  $Y = BC \cap EF$ ,  $Z = CD \cap AF$  kollinear (also  $ABCDEF$  ein Pascalsches Sechseck), so sind auch  $R = AD \cap CF$ ,  $S = BD \cap CE$ ,  $T = BF \cap AE$  kollinear.



Wir hatten gesehen, dass unter affinen Transformationen das Teilverhältnis erhalten bleibt. Dies gilt für projektive Transformationen nicht mehr. Betrachten wir etwa ein gleichseitiges Dreieck  $ZAB$ , eine Gerade  $g$  durch  $A$ , die mit  $g = AB$  einen Winkel der Größe  $\alpha$  bildet, und einen Punkt  $M$  auf  $g$  mit dem Teilverhältnis  $TV(A, B; M) = t$ . Sei weiter  $\pi$  die Projektion mit Zentrum  $Z$ , die  $g$  auf  $g'$  und dabei  $M$  auf  $M'$  und  $B$  auf  $B'$  abbildet. Ist etwa  $M$  der Mittelpunkt der Strecke  $\overline{AB}$ , so gilt  $TV(A, B; M) = 1$ , aber in diesem Bild offensichtlich  $TV(A', B'; M') < 1$ .



*Aufgabe:*

- Finden Sie eine Formel für  $t' = TV(A', B'; M')$  in Abhängigkeit von  $t$  und  $\alpha$ .
- Zeigen Sie, dass  $t'$  für verschiedene Lagen von  $g'$  bei gegebenem  $t$  jeden Wert annehmen kann. Wann ist insbesondere  $t' < 0$ , wenn  $t > 0$  gilt?
- Sei  $N$  ein weiterer Punkt auf  $g$  und  $N'$  dessen Bildpunkt auf  $g'$ . Zeigen Sie, dass für das *Doppelverhältnis*

$$DV(A, B; M, N) = \frac{TV(A, B; M)}{TV(A, B; N)}$$

stets

$$DV(A, B; M, N) = DV(A', B'; M', N')$$

unabhängig von der Lage von  $g'$  gilt.

Das Doppelverhältnis ist auch im allgemeinen Fall eine projektive Invariante. Ist  $g = AB$ ,  $M$  ein Punkt auf  $g$  und  $F_g$  der Fernpunkt von  $g$ , so gilt

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

da der Fernpunkt der Geraden gerade dem Teilverhältnis  $TV(A, B; F_g) = -1$  entspricht. In diesem Sinne ergibt sich die affine Geometrie als „projektive Geometrie mit fixierter Ferngeraden“.

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

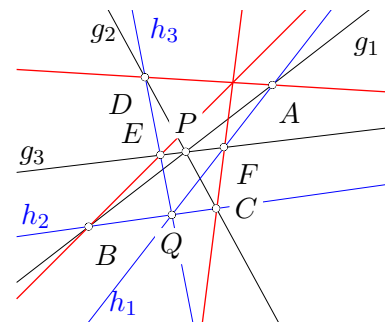
In vielen geometrischen Aussagen über Punkte und Geraden kann man die Worte „Punkt“ und „Gerade“ vertauschen und bekommt einen ebenfalls gültigen geometrischen Satz. Die einfachsten Aussagen dieser Art sind

- 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 16 (Dualer Satz von Pappus)** *Seien die Geraden  $g_1, g_2, g_3$  und  $h_1, h_2, h_3$  jeweils konkurrent und  $A, \dots, F$  die Schnittpunkte  $A = h_1 \wedge g_1, B = g_1 \wedge h_2, C = h_2 \wedge g_2, D = g_2 \wedge h_3, E = h_3 \wedge g_3$  und  $F = g_3 \wedge h_1$ . Die drei Verbindungsgeraden  $AD, BE$  und  $CF$  sind konkurrent.*



*Aufgabe:* Sei  $P$  der gemeinsame Schnittpunkt von  $g_1, g_2, g_3$  und  $Q$  der gemeinsame Schnittpunkt von  $h_1, h_2, h_3$ . 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.

*Aufgabe:* Formulieren Sie den dualen Satz zum Satz von Desargue? Warum bezeichnet man den Satz von Desargue als *selbstdual*?

Einen praktischen Ansatz für den Beweis dualer Sätze bietet die *Polargerade* eines Punktes  $A$  bzgl. eines gegebenen Kreises  $k$  (bzw. allgemeiner einer gegebenen Quadrik). Sei dazu  $A$  zunächst außerhalb von  $k$  gelegen, so dass wir die Tangenten aus  $A$  an  $k$  bestimmen können und deren Berührungspunkte  $A_1$  und  $A_2$ . Als Polargerade  $p_A$  bezeichnet man die Verbindungsgerade dieser Berührungspunkte. Liegt  $A$  auf der Kreislinie, so fallen die Punkte  $A_1$  und  $A_2$  zusammen und wir definieren  $p_A$  als die Tangente in  $A$  an den Kreis.

Umgekehrt kann zu einer Geraden  $g$ , die  $k$  in  $X$  und  $Y$  schneidet, der *Polarpunkt*  $P_g$  als der Schnittpunkt der Tangenten an  $k$  in  $X$  und  $Y$  bestimmt werden. Offensichtlich ist die Polargerade von  $P_g$  gerade wieder  $g$  und umgekehrt der Polarpunkt der Polargeraden  $p_A$  gerade wieder der Punkt  $A$ . Man bezeichnet deshalb diese eineindeutige Beziehung zwischen (zunächst Teilmengen von) Punkten und Geraden der Ebene auch kurz als *Polare*. GEOGEBRA stellt ein entsprechendes Konstruktionswerkzeug zur Verfügung.

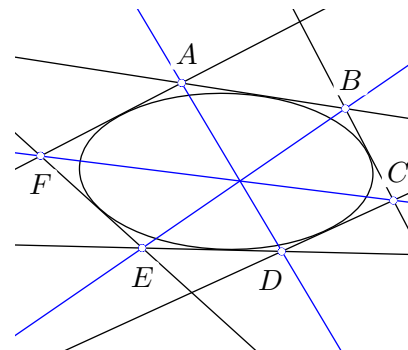
*Aufgabe:* Zeigen Sie, dass die Polargeraden kollinearere Punkte konkurrent und die Polarpunkte konkurrender Geraden kollinear sind.

Mit dieser Information können wir nun auch die Polargerade für einen Punkt  $A$  innerhalb des Kreises  $k$  bestimmen. Wir nehmen dazu zwei Geraden  $g_1$  und  $g_2$  durch  $P$  – die damit  $k$  garantiert schneiden – und bestimmen deren Polarpunkte  $G_1$  und  $G_2$ . Deren Verbindungsgerade ist die gesuchte Polargerade  $p_A$ . Umgekehrt können wir mit einer Geraden  $g$  verfahren, die  $k$  nicht schneidet: Wir nehmen zwei Punkte  $A_1$  und  $A_2$  auf dieser Geraden – die damit garantiert außerhalb von  $k$  liegen – und bestimmen deren Polargeraden  $a_1$  und  $a_2$ . Der Schnittpunkt dieser Polargeraden ist gerade der Polarpunkt  $P_g$  der Geraden  $g$ .

Auch der Satz von Pascal lässt eine duale Version zu, welche eine Aussage über das Schnittverhalten der Polaren der sechs Punkte eines Pascalschen Sechsecks liefert:

**Satz 17 (Satz von Briancon)** *Die Geraden  $a, b, c, d, e$  und  $f$  mögen eine Quadrik berühren, so dass sich „benachbarte“ Tangenten in den Punkten  $A, B, C, D, E$  und  $F$  schneiden (d. h.  $ABCDEF$  ist ein Tangentensechseck). In jedem solchen Tangentensechseck gehen die Diagonalen  $AD, BE$  und  $CF$  zwischen gegenüberliegenden Eckpunkten durch einen gemeinsamen Punkt.*

Ein elementargeometrischer Beweis dieses Satzes für den Fall eines Kreises (auf den der allgemeine Satz durch eine projektive Transformation zurückgeführt werden kann) ist in [2, Abschnitt 4.9] enthalten.



## 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  $(p_x, p_y)$  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  $g = (g_1, g_2, g_3)$ , welches für die Gerade  $\{(p_x, p_y) : g_1 p_x + g_2 p_y + g_3 = 0\}$  steht. Ein solches Tripel bezeichnet man als *homogene Koordinaten* der Geraden  $g$ . Zueinander proportionale Tripel beschreiben dieselbe Gerade  $g$  – wir schreiben deshalb auch  $g = (g_1 : g_2 : g_3)$  – und für (echte) Geraden dürfen  $g_1$  und  $g_2$  nicht gleichzeitig verschwinden. Es gibt genau eine „unechte“ Gerade, diese hat die homogenen Koordinaten  $l_0 = (0 : 0 : 1)$ . Wir sehen später, dass dies genau die *Ferngerade* der affinen Ebene ist.

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

$$\begin{aligned} g_1 a_x + g_2 a_y + g_3 &= 0 \\ g_1 b_x + g_2 b_y + g_3 &= 0 \\ g_1 c_x + g_2 c_y + g_3 &= 0 \end{aligned}$$

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

$$\det \begin{pmatrix} a_x & a_y & 1 \\ b_x & b_y & 1 \\ c_x & c_y & 1 \end{pmatrix} = 0$$

gilt.

- Analog sind drei Geraden  $g, h, k$  *konkurrent*, d. h. gehen durch einen gemeinsamen Punkt  $P = (p_x, p_y)$  genau dann, wenn das lineare Gleichungssystem

$$\begin{aligned} g_1 p_x + g_2 p_y + g_3 &= 0 \\ h_1 p_x + h_2 p_y + h_3 &= 0 \\ k_1 p_x + k_2 p_y + k_3 &= 0 \end{aligned}$$

eine Lösung in  $(p_x, p_y)$  besitzt. Das ist genau dann der Fall, wenn die zugehörige Koeffizientenmatrix denselben Rang hat wie die erweiterte Koeffizientenmatrix. 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.

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

$$(g_1, g_2, g_3) = (b_y - a_y, a_x - b_x, a_y b_x - a_x b_y)$$

- 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

$$(h_1, h_2, h_3) = (g_1, g_2, -(g_1 p_x + g_2 p_y)) .$$

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

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

- Ein Punkt  $P$  auf der Geraden  $g = AB$  hat die Koordinaten

$$(p_x, p_y) = ((1 - u) a_x + u b_x, (1 - u) a_y + u b_y)$$

für ein geeignetes  $u \in \mathbb{R}$ . Diese Beziehung ergibt sich aus der Vektorgleichung von Ortsvektoren

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AP} = \overrightarrow{OA} + u \overrightarrow{AB} = \overrightarrow{OA} + u (\overrightarrow{OB} - \overrightarrow{OA}) = (1 - u) \overrightarrow{OA} + u \overrightarrow{OB}$$

und gilt für alle Punkte  $P \in g(AB)$ , wobei  $u$  aus der Beziehung  $\overrightarrow{AP} = u \overrightarrow{AB}$  eindeutig bestimmt ist. Wir bezeichnen  $u = GP(A, B; P)$  als *Gleiterparameter*. Liegt  $P$  im Inneren der Strecke  $\overline{AB}$ , so gilt  $0 < u < 1$ , für Punkte  $P$  jenseits von  $B$  gilt  $u > 1$  und für Punkte jenseits von  $A$  schließlich  $u < 0$ . Zum früher eingeführten Teilverhältnis  $TV(A, B; P)$  besteht der Zusammenhang

$$TV(A, B; P) = \frac{u}{1 - u} .$$

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

$$d(A, B) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2} .$$

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  $g$  als

$$(h_1, h_2, h_3) = (g_2, -g_1, g_1 p_y - g_2 p_x)$$

ausdrücken.

## 2.2 Homogene Punktkoordinaten

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

$$\begin{aligned} g_1 p_x + g_2 p_y + g_3 &= 0 \\ h_1 p_x + h_2 p_y + h_3 &= 0 \\ k_1 p_x + k_2 p_y + k_3 &= 0 \end{aligned}$$

auch nach Lösungen  $(p_x, p_y, p_z)$  des homogenen Gleichungssystems

$$\begin{aligned} g_1 p_x + g_2 p_y + g_3 p_z &= 0 \\ h_1 p_x + h_2 p_y + h_3 p_z &= 0 \\ k_1 p_x + k_2 p_y + k_3 p_z &= 0 \end{aligned}$$

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

Solche Koordinaten  $P = (p_x, p_y, p_z)$  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  $(p_x, p_y)$  die projektiven Koordinaten  $(p_x, p_y, 1)$  entsprechen. Letztere bezeichnen wir auch als *normierte Koordinaten* und schreiben  $\bar{P}$ .  $P$  liegt auf der Geraden  $g$  genau dann, wenn

$$g_1 p_x + g_2 p_y + g_3 p_z = 0$$

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  $p_z = 0$  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  $\mathbb{P}^2$ . Die weiter oben untersuchten geometrischen Beziehungen lassen sich dann nennerfrei durch Skalar-, Vektor- und Spatproduktoperationen im  $\mathbb{R}^3$  beschreiben.

- $A, B, C$  in homogenen Punktkoordinaten sind kollinear genau dann, wenn

$$\det \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} = 0$$

erfüllt ist

Notation:  $\text{sp}(A, B, C) = 0$  (Spatprodukt)

- Analog sind drei Geraden  $g, h, k$  konkurrent genau dann, wenn  $\text{sp}(g, h, k) = 0$  gilt.
- Punkt  $P$  und Gerade  $g$  sind inzident genau dann, wenn  $p_x g_1 + p_y g_2 + p_z g_3 = 0$  gilt.

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

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

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

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

$$g = (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x) = A \times B$$

- $A, B, C$  sind kollinear genau dann, wenn  $A$  inzident zur Geraden durch  $B$  und  $C$  ist. Dies gilt genau dann, wenn  $A * (B \times C) = 0$  erfüllt ist.

Dies stimmt wegen des bekannten Zusammenhangs  $\text{sp}(A, B, C) = A * (B \times C)$  zwischen Spat-, Vektor- und Skalarprodukt im  $\mathbb{R}^3$  mit obiger Determinantenformel überein.

Homogene Punkt- bzw. Geradenkoordinaten sind genau dann *nicht zulässig*, wenn sich die Koordinaten  $(0 : 0 : 0)$  ergeben. 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 genau dann, wenn sie sich auf der Ferngeraden schneiden, d. h. wenn  $\text{sp}(g, h, l_0) = 0$  gilt. Das stimmt mit unserer weiter oben hergeleiteten Formel überein.

Die Koordinaten des Fernpunkts  $F_g$  der Geraden  $g$  ergeben sich aus der Formel

$$F_g = g \times l_0 = (-g_2 : g_1 : 0).$$



- Die Gerade  $h$  durch  $P$ , die parallel zu  $g$  verläuft, ergibt sich als Verbindung des Fernpunkts  $F_g = (-g_2 : g_1 : 0)$  der Geraden  $g$  und  $P$  zu  $h = P \times F_g$ .
- 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

$$h = P \times O_g = (-p_z g_2 : p_z g_1 : p_x g_2 - p_y g_1)$$

in Übereinstimmung mit der früher gefundenen Darstellung ergibt.  $O_g$  wird auch als *Orthogonalpunkt* von  $g$  bezeichnet.

Mit Parallelen kann man aus einem Standardframe ein ganzes affines Koordinatensystem gewinnen. Als *Standardframe* bezeichnet man ein Punkte-Quadrupel  $(E_0, E_1, E_2, E_3)$  der projektiven Ebene, von denen keine drei auf einer Geraden liegen. Als Ursprung  $E_0$ , Fernpunkt  $E_1$  der  $x$ -Achse, Fernpunkt  $E_2$  der  $y$ -Achse und Einheitspunkt  $E_3$  bestimmen diese vier Punkte ein Koordinatensystem so dass

$$E_0 = (0 : 0 : 1), E_1 = (1 : 0 : 0), E_2 = (0 : 1 : 0), E_3 = (1 : 1 : 1)$$

gilt. Der Einheitspunkt  $E_3$  mit den (affinen) Koordinaten  $(1, 1)$  bestimmt die beiden Koordinateneinheiten, da die Parallelen zur  $x$ - bzw.  $y$ -Achse durch  $E_3$  die  $y$ - bzw.  $x$ -Achse in den Koordinaten-Einheiten schneiden.

### 2.3 Teilverhältnis und Doppelverhältnis in homogenen Koordinaten

**Teilverhältnisse:** Sei  $M$  ein Punkt auf  $AB$  mit dem Gleiterparameter  $GP(A, B; M) = u$ , also

$$\begin{pmatrix} m_x \\ m_y \end{pmatrix} = (1 - u) \begin{pmatrix} a_x \\ a_y \end{pmatrix} + u \begin{pmatrix} b_x \\ b_y \end{pmatrix}.$$

Diese Gleichung wird in homogenen Koordinaten zu

$$\begin{pmatrix} m_x \\ m_y \\ 1 \end{pmatrix} = (1 - u) \begin{pmatrix} a_x \\ a_y \\ 1 \end{pmatrix} + u \begin{pmatrix} b_x \\ b_y \\ 1 \end{pmatrix}$$

und müsste für andere Koordinaten derselben Punkte zu

$$\begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} = (1 - u) b_z \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + u a_z \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix}$$

modifiziert werden. Wir erhalten also eine Darstellung  $M = \mu_A A + \mu_B B$ , in der die skalaren Faktoren  $\mu_A = (1 - u) b_z$  und  $\mu_B = u a_z$  für denselben Punkt  $M$  von der Wahl der homogenen Koordinaten der Punkte  $A$  und  $B$  abhängen. Die Darstellung ist nicht eindeutig. Für die Darstellung von Teilverhältnissen muss in homogenen Koordinaten also ein größerer Aufwand

getrieben werden, was aber nicht weiter verwundern kann – wir hatten bereits gesehen, dass Teilverhältnisse unter projektiven Transformationen nicht invariant sind.

Betrachten wir dazu Darstellungen  $M = \mu_A A + \mu_B B$  mit  $\mu_A, \mu_B \in K$  genauer. Jeder solche Punkt  $M$  liegt auf der Geraden  $g = AB$ , denn es gilt

$$\text{sp}(M, A, B) = \mu_A \text{sp}(A, A, B) + \mu_B \text{sp}(B, A, B) = 0$$

wegen der Linearität des Spatprodukts.

Bei *fixierten* homogenen Koordinaten von  $A$  und  $B$  ergeben Paare  $(\mu_A, \mu_B)$ , welche sich nur durch einen skalaren Faktor unterscheiden, (verschiedene homogene Koordinaten für) denselben Punkt  $M$ . Die Punkte auf  $AB$  werden also durch das Verhältnis  $\mu = (\mu_B : \mu_A) \in \mathbb{P}^1$  eindeutig charakterisiert (das Verhältnis  $\infty$  ist möglich und entspricht dem Punkt  $B$ ). Man sagt auch, dass dies eine *Parametrisierung* der Punkte der Geraden  $AB$  durch  $\mu \in \mathbb{P}^1$  ist. Dabei entspricht  $A$  immer dem Parameter  $(0 : 1)$  und  $B$  dem Parameter  $(1 : 0)$ .

Für *andere* homogene Koordinaten derselben Punkte  $A$  und  $B$  führt dasselbe Verhältnis  $\mu = (\mu_B : \mu_A) \in \mathbb{P}^1$  allerdings zu anderen Punkten  $M$  auf  $g$ . Andere homogene Koordinaten von  $A$  und  $B$  ergeben eine andere Parametrisierung der Geraden  $AB$ .

Sind  $A$  und  $B$  in *normierten* Koordinaten gegeben und weiter  $(\mu_B : \mu_A)$  so normiert, dass  $\mu_A + \mu_B = 1$  gilt (und damit auch  $M$  in normierten Koordinaten angegeben ist), so ergibt sich mit  $\mu_A = 1 - u$ ,  $\mu_B = u$  gerade die affine Formel  $M = (1 - u)A + uB$  und das Teilverhältnis  $TV(A, B; M) = \frac{u}{1-u} = \mu$ . Die Normierung  $\mu_A + \mu_B = 1$  ist genau dann möglich, wenn  $\mu_A + \mu_B \neq 0$  gilt. Das einzige Verhältnis, welches dieser Bedingung nicht genügt, ist  $\mu_F = (\mu_B : \mu_A) = (-1 : 1)$ . Für den zugehörigen Punkt  $F$  ergibt sich  $TV(A, B; F) = -1$ . Diesen Punkt hatten wir bereits früher als den Fernpunkt  $F_g$  der Geraden  $g$  identifiziert. Wir schließen, dass  $F_g = B - A$  gilt, wenn  $A$  und  $B$  in normierten Koordinaten gegeben sind. Allgemein gilt  $F_g = -b_z A + a_z B$ .

Sind  $A$  und  $B$  in allgemeinen homogenen Koordinaten gegeben, so gilt

$$TV(A, B; M) = \frac{\mu_B b_z}{\mu_A a_z},$$

denn aus  $M = \mu_A A + \mu_B B$  folgt für normierte Koordinaten

$$m_z \overline{M} = \mu_A a_z \overline{A} + \mu_B b_z \overline{B} \quad \text{mit} \quad m_z = \mu_A a_z + \mu_B b_z.$$

Erst das *Doppelverhältnis*  $DV(A, B; M, N)$  für einen weiteren Punkt  $N = \nu_A A + \nu_B B$  auf der Geraden  $AB$  ist unabhängig von der Parametrisierung. Es gilt

$$DV(A, B; M, N) = \frac{TV(A, B; M)}{TV(A, B; N)} = \frac{(\mu_B b_z) : (\mu_A a_z)}{(\nu_B b_z) : (\nu_A a_z)} = \frac{\mu_B : \mu_A}{\nu_B : \nu_A}.$$

Die Lage eines Punkts  $M$  auf  $AB$  kann aus  $\mu_M$  ohne Bezug auf die Parametrisierung also nur relativ zu einem dritten zu fixierenden Punkt  $E \in AB$  durch  $DV(A, B; M, E)$  eindeutig bestimmt werden.

Wie im Fall eines Standardsframes im  $\mathbb{P}^2$ , welches erst durch vier Punkte eindeutig bestimmt ist, ist ein Standardframe im  $\mathbb{P}^1$  – also eine Parametrisierung – erst durch die Angabe der Koordinaten von drei Punkten  $A = (0 : 1)$ ,  $B = (1 : 0)$  und  $E = (1 : 1)$  eindeutig bestimmt.

Fordern wir etwa, dass  $E$  der Mittelpunkt der Strecke  $\overline{AB}$  ist, so ergibt sich  $TV(A, B; E) = 1 = \mu_E$  und damit auch für alle anderen Punkte

$$\mu_M = \frac{\mu_M}{\mu_E} = DV(A, B; M, E) = \frac{TV(A, B; M)}{TV(A, B; E)} = TV(A, B; M).$$

Zur Einbettung der affinen in die projektive Geometrie wird allerdings statt  $E$  gewöhnlich der Fernpunkt  $F = F_g$  der Geraden  $g = AB$  als weiterer Punkt fixiert. In normierten Koordinaten für  $A$  und  $B$  ergibt sich  $F = (-1 : 1)$  und damit  $TV(A, B; M) = -DV(A, B; M, F_g)$ .

Die hier beschriebenen Vorteile homogener Koordinaten veranlassen einige Designer von DGS, diese zur Darstellung von Punkten zu verwenden. Mit Blick auf die Abweichungen von den aus der Schule bekannten Notationen der analytischen Geometrie sowie der Probleme vor allem der Darstellung von Teilverhältnissen werden wir im Weiteren zunächst mit affinen Punktkoordinaten arbeiten.

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

Wir können auf der Basis der im Abschnitt 2.1 hergeleiteten Beziehungen 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 double x,y;

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

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

    public Line () {}
    public Line(double a, double b, double 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  $P(p_x, p_y)$  bzw. Geraden

$$g = \{(p_x, p_y) : g_1 p_x + g_2 p_y + g_3 = 0\}$$

darstellen. `double` kann dabei durch einen beliebigen anderen Grundbereich für die Werte der Koordinaten ersetzt werden. Wir wollen im Weiteren nur 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) {
    double 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 double sqrdist(Point p, Point q) {
    return (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, double 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, 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:* 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:* 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  $ABC$  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:* 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:* Modifizieren Sie diese Testfunktion so, dass sie den Satz vom Feuerbachschen Kreis testet.

*Aufgabe:* Testen Sie den Satz von Desargue.

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

Wir wollen uns an dieser Stelle zunächst 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 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**). 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.

Das Bewegen geometrischer Objekte kann die Bewegung anderer geometrischer Gebilde zur Folge haben – Zustandsänderungen propagieren also durch einen gerichteten azyklischen Graphen von Abhängigkeiten. Diese Abhängigkeiten werden durch Berechnungsvorschriften beschrieben, nach denen die Koordinaten des abhängigen Objekts aus denen der Vorgängerobjekte bestimmt werden, so dass es sich um Abhängigkeiten zwischen den Objekten selbst handelt. Wir führen deshalb die folgende begriffliche Unterscheidung ein:

**Definition 1** *Die Klassen **Point** und **Line** bezeichnen wir als **geometrische Typen**, Instanzen einer solchen Klassen als **geometrische Objekte**. Jedes solche geometrische Objekt hat eine abstrakte Identität, auf die wir über den Namen des Objekts zugreifen, so dass wir voraussetzen, dass der Name eines geometrischen Objekts nicht veränderbar ist. Jeden Zustand eines solchen geometrischen Objekts im Laufe seines Lebenszyklus bezeichnen wir als **spezielle Realisierung des geometrischen Objekts**. Eine solche spezielle Realisierung ist also eine konkrete Ausprägung eines geometrischen Objekts in Raum und Zeit.*

Für jeden geometrischen Typ  $T$  spezifizieren wir ein spezielles Attribut  $c$ , welches die Koordi-

naten des jeweiligen Objekts enthält, und bezeichnen dieses als *Koordinatenattribut*. Ist etwa  $g \in \text{Line}$  ein Objekt vom Typ Gerade, so steht  $g.c$  für das Koordinatenattribut von  $g$ . Wie oben haben wir zu unterscheiden zwischen

- dem Attribut  $g.c$  selbst als informatischem Begriff,
- der Berechnungsvorschrift, nach welcher sich der Attributwert aus anderen Attributwerten berechnet und
- konkreten Attributwerten  $g.c \in C(L)$ .

Aus dem Kontext der weiteren Ausführungen wird deutlich, in welchem Sinne  $g.c$  jeweils aufzufassen ist.  $C(L)$  ist dabei der Bereich der (für Geraden zulässigen) Koordinatenwerte. Dieser hängt nicht von  $g$  selbst ab, sondern nur vom geometrischen Typ  $g \in \text{Line}$  und natürlich vom gewählten **Koordinatenmodell**. 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.

In praktischen Implementierungen wird  $C(L) = K^3$  gesetzt, also mit Tripeln gearbeitet, und die weitere Struktur nur semantisch berücksichtigt.

DGS sind als grafische Anwendungen sinnvollerweise nach dem MVC-Konzept aufgebaut. Mausaktionen werden vom Controller an das Modell 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 Modell-Seite, die spezielle Realisierung auf der View-Seite zu finden.

Geometrische Objekte werden durch entsprechende Konstruktionswerkzeuge schrittweise erzeugt, welche auf der Seite der Informatik als Funktionen daherkommen. Für jede solche Konstruktion haben wir zwischen der Beschreibung dieser Konstruktion und deren Ausführung (und dynamischen Veränderung) zu unterscheiden. Eine Konstruktion muss erst vollständig beschrieben sein, ehe sie (erfolgreich oder auch erfolglos) ausgeführt werden kann. Wir müssen also – wie in anderen Bereichen der Informatik auch – zwischen *Designzeit* und *Laufzeit* unterscheiden.

Zur Designzeit wird eine Konstruktionsbeschreibung mit Hilfe vorhandener Konstruktionswerkzeuge nach entsprechenden Konstruktionsregeln erstellt. Die (zur Laufzeit wesentliche) Dynamik neu konstruierter geometrischer Objekte ergibt sich aus der Dynamik der geometrischen Objekte, welche an der Konstituierung beteiligt waren, und der Spezifik des Konstruktionswerkzeugs selbst. Diese ist in der Berechnungsvorschrift kodiert, nach welcher sich



der Attributwert des Koordinatenattributs des neuen Objekts aus den Attributwerten der Koordinatenattribute der Vorgängerobjekte berechnet. Diese Berechnungsvorschrift wird zur Designzeit mit dem Koordinatenattribut des neuen Objekts verbunden und zur Laufzeit ausgeführt. Wir finden hier also die klassische informatische Unterscheidung zwischen Definition und Ausführung einer Berechnungsvorschrift wieder.

Konstruktionswerkzeuge sind prototypisch von zwei verschiedenen Arten.

**Prototyp 1:** `Point H = 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  $H \in \text{Point}$  erzeugt. Der Attributwert des Koordinatenattributs  $H.c$  von  $H$  berechnet sich dabei aus den Attributwerten der Koordinatenattribute  $P.c$  und  $a.c$  nach einer Berechnungsvorschrift  $\phi : C(P) \times C(L) \rightarrow C(P)$  als  $H.c = \phi(P.c, a.c)$ . Die Dynamik von  $H$  ist also durch die Dynamiken von  $P$  und  $a$  *vollständig* determiniert.

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

Hier hängt  $M$  noch zusätzlich von einem Stellparameter  $u$  ab, dessen Natur  $X$  wir nun näher spezifizieren wollen. Ein Vergleich mit Prototyp 1 zeigt, dass  $u$  in derselben Doppelbedeutung wie ein geometrisches Objekt auch auftritt: Einerseits als abstrakte Identität eines Stellparameters, welcher selbst eine zeitliche Dynamik hat, und andererseits als Eingangsparemeter, welcher die Dynamik von  $M$  beeinflusst. Die im Vergleich zu Prototyp 1 scheinbar eigenständige Dynamik von  $M$  kann also im Stellparameter gekapselt werden, so dass auch hier die Dynamik von  $M$  wie im Prototyp 1 *vollständig* durch die Dynamiken von  $A$ ,  $B$  und  $u$  determiniert ist.

$X$  steht also für einen weiteren Datentyp eines Stellparameters **SP** mit eigenem (eindimensionalem) Wertebereich  $C(S)$ . 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 mit Werten aus dem Parameterbereich  $C(S)$ .  $u.c \in C(S)$  bezeichnen wir in Analogie zur Notation für geometrische Objekte als *spezielle Realisierung des Parameterobjekts*.

Wir können uns die Entkopplung von der visuellen Darstellung als „Bewegen mit der Maus“ etwa als Schieberegler vorstellen, dessen Schieben eine Bewegung von  $M$  auf der Geraden  $AB$  nach sich zieht. Eine solche Trennung ist auch bei freien Punkten sinnvoll, da für diese ebenfalls eine mittelbare Steuerung der Bewegung denkbar wäre. 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. 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.

**Definition 2** 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**  $w$  bezeich-

nen wir eine (informatische) Funktion der Signatur

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

zusammen mit einer Berechnungsvorschrift

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

so dass sich für Objekte  $o_i$  vom Typ  $T_i$  und  $o_a = w(o_1, \dots, o_n)$  vom Typ  $T_a$  der Attributwert des Koordinatenattributs  $o_a.c$  einer speziellen Realisierung von  $o_a$  als

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

aus den Attributwerten des Koordinatenattributs einer speziellen Realisierung der vorgegebenen Objekte  $o_1, \dots, o_n$  berechnet.

Weiter sei

$$\tilde{w} = \text{id} \times w : (T_1 \times \dots \times T_n) \rightarrow (T_1 \times \dots \times T_n \times T_a)$$

der Graph von  $w$ , also die Abbildung, welche die Aufrufargumente mit in das Rückgabepaket aufnimmt, und  $\tilde{w}.c$  entsprechend der Graph von  $w.c$ .

$w.c$  beschreibt die Dynamik von  $o_a$  in Abhängigkeit der Dynamiken von  $o_1, \dots, o_n$ .

Eine weitere Feinheit haben wir 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 Realisierungen der Punkte  $A$  und  $B$  diese Gerade nur definiert, wenn diese nicht zusammenfallen. Die Berechnungsvorschrift

$$\text{pp\_line.c} : C(P) \times C(P) \rightarrow C(L)$$

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

$$\text{pp\_line.DG} : C(P) \times C(P) \rightarrow \text{Boolean}$$

bestimmt, welches in unserem Fall die einfache Form  $\text{pp\_line.DG}(c_1, c_2) = \text{isequal}(c_1, c_2)$  hat. Hierbei ist `Boolean` der Wertebereich des Datentyps `boolean`. In diesem (und so auch in den meisten Fällen) ist `isequal` sogar eine *geometrische Eigenschaft*, denn sie kann als (mathematische) Berechnungsvorschrift der informatischen Funktion

$$\text{isequal} : \text{Point} \times \text{Point} \rightarrow \text{boolean}$$

interpretiert werden, nach der berechnet wird, ob die Attributwerte der Koordinatenattribute spezieller Realisierungen zweier Punkte zusammenfallen.

**Definition 3** Zu jedem Konstruktionswerkzeug  $w$  gibt es also weiter eine **Degenerationsbedingung**

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

so dass die Berechnungsvorschrift  $w.c$  genau dann auf einer Realisierung von  $(o_1, \dots, o_n)$  ausgeführt werden kann, wenn

$$w.DG(o_1.c, \dots, o_n.c) = \text{false}$$

gilt.

**Beispiele:**

`freePoint: MP → Point`

legt einen freien Punkt an. Die zugehörige Degenerationsbedingung für  $m \in C(M)$  ist leer: `freePoint.DG(m) = false`. (Präziser: Die Degenerationsbedingung hängt vom gewählten Modell für Mausparameter ab)

`pp_line: Point×Point → Line`

erzeugt die Gerade durch zwei gegebene Punkte. Die zugehörige Degenerationsbedingung für  $c_1, c_2 \in C(P)$  lautet `pp_line.DG(c1, c2) = isequal(c1, c2)`.

`intersection_point: Line×Line → Point`

erzeugt den Schnittpunkt zweier Geraden. Die zugehörige Degenerationsbedingung für  $a, b \in C(L)$  ist `intersection_point.DG(a, b) = iszero(a1 b2 - b1 a2)`. Diese Bedingung lässt sich ebenfalls als geometrische Bedingung `is_parallel(oa, ob)` der zugehörigen geometrischen Objekte anschreiben.

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 sind bei der (Neu-)Berechnung der Attributwerte der Koordinatenattribute zu berücksichtigen, was sich intern durch Ereignispropagation modellieren lässt.

Diese Abhängigkeiten lassen sich durch einen (endlichen) gerichteten azyklischen Graphen (DAG) darstellen, der aus der Konstruktionsbeschreibung extrahiert werden kann, also bereits zur Designzeit bekannt ist. Als *Konstruktionsbeschreibung* wird deshalb der prinzipielle Vorgang des Erstellens einer geometrischen Konfiguration zur Definitionszeit bezeichnet, als *Realisierung der Konfiguration* eine Folge spezieller Realisierungen geometrischer Objekte, welche nach dieser Beschreibung zur Konstruktionszeit erzeugt werden, sofern dies überhaupt möglich ist. Es kann Konstruktionsbeschreibungen geben, die sich nicht realisieren lassen, etwa weil sie in einem Zwischenschritt stets zu degenerierten Lagen führen, in denen der nächste Konstruktionsschritt nicht mehr ausführbar ist. Die Aussage über die prinzipielle Nichtrealisierbarkeit einer geometrischen Konfiguration ist ihrerseits wieder ein geometrischer Satz.

**Definition 4** Als *Konfiguration* bezeichnen wir einen azyklischen gerichteten Abhängigkeitsgraphen  $\Gamma = (O, E)$  mit Knotenmenge  $O = (o_1, \dots, o_m)$  und Kantenmenge  $E$ , wobei die Knotenmenge eine Menge von geometrischen und Parameterobjekten ist. Wir setzen voraus, dass  $(o_i, o_j) \in E \Rightarrow i < j$  gilt, also nur „spätere“ von „früheren“ Objekten abhängen.  $T(o)$  bezeichnet den Typ des Objekts  $o \in O$ . Weiter setzen wir voraus, dass Parameterobjekte keine eingehenden Kanten haben, also zwischen Parameterobjekten keine Abhängigkeiten bestehen.

Wir bezeichnen  $\Gamma = (A, \emptyset)$  als *Parameterkonfiguration*, wenn  $A$  ausschließlich Parameterobjekte enthält.

Damit können wir nun den Begriff der Konstruktionsbeschreibung schrittweise herleiten.

**Definition 5** Sei

- $O = (o_1, \dots, o_m)$  eine Sequenz geometrischer und Parameterobjekte und  $\Gamma = (O, E)$  ein Konfiguration,
- $w : (T_1 \times \dots \times T_n) \rightarrow T_{m+1}$  ein Konstruktionswerkzeug und
- $u \in O^n$  mit  $T(u_i) = T_i$  eine Auswahlfunktion auf  $O$  für eine typgerechte Belegung der Aufrufparameter. ( $w$  ist **auf**  $\Gamma$  **anwendbar**, wenn es eine solche Belegung gibt.)

Die Ergänzung von  $O$  um das geometrische Objekt  $o_{m+1} = w(u_1, u_2, \dots, u_n)$  vom Typ  $T_a$  und die Ergänzung von  $\Gamma$  zu  $\Gamma' = (O \cup \{o_{m+1}\}, E \cup \{(u_i, o_{m+1}), i = 1, \dots, n\})$  bezeichnen wir als **Konstruktionsschritt**.

Es ist sinnvoll, als Belegung  $u$  auch Belegungen mit Dopplungen zuzulassen. So kann etwa das Konstruktionswerkzeug `altitude(A, B, C)`, welches die Höhe durch  $A$  im Dreieck  $ABC$  konstruiert, auch für  $A = B$  sinnvoll angewendet werden, um eine Senkrechte in  $B$  zu errichten.

Für die Laufzeiteigenschaft eines Konstruktionsschritts halten wir fest: Ein Konstruktionsschritt ist auf einer speziellen Realisierung von  $\Gamma$  **ausführbar**, wenn  $w.DG(u_{1.c}, \dots, u_{n.c}) = \text{false}$  gilt. Wir wollen annehmen, dass die Ausführung der Konstruktion mit einer Ausnahme abbricht, wenn der Schritt nicht ausführbar ist.

In all diesen Definitionen ist es möglich, den Begriff des Konstruktionsschritts und des Konstruktionswerkzeugs so zu fassen, dass statt eines einzelnen geometrischen Objekts  $o_{m+1}$  gleich ein ganzes Tupel  $(o_{m+1}, \dots, o_{m+k})$  neuer Objekte der Typen  $(T_{m+1}, \dots, T_{m+k})$  konstruiert wird. Ein solches Konstruktionswerkzeug bezeichnen wir als **verallgemeinertes Konstruktionswerkzeug**.

**Definition 6** Als **Konstruktionsbeschreibung**  $\mathcal{K}$  bezeichnen wir eine Folge von Konstruktionsschritten, welche die **Startkonfiguration**  $\Gamma_S = (O_S, E_S)$  mit  $O_S = (o_1, \dots, o_m)$  durch endlich viele konsekutive Konstruktionsschritte

$$\tilde{w}^{(N)} : \Gamma_S = \Gamma_0 \xrightarrow{\tilde{w}_1} \Gamma_1 \xrightarrow{\tilde{w}_2} \dots \xrightarrow{\tilde{w}_N} \Gamma_N = \Gamma_E \quad (\mathcal{K})$$

in die **Endkonfiguration**  $\Gamma_E = (O_E, E_E)$  mit  $O_E = (o_1, \dots, o_{m+N})$  überführt. Wir schreiben auch kurz  $O_E = \mathcal{K}(O_S)$ .

Die dynamischen Freiheitsgrade einer Konstruktionsbeschreibung werden allein durch deren Parameterobjekte bestimmt, sind also bereits in der Startkonfiguration festgelegt, da nach Definition in einem Konstruktionsschritt ausschließlich geometrische Objekte erzeugt werden.

Die Definition zeigt, dass zusammen mit  $(\mathcal{K})$  auch jede Teilfolge

$$\tilde{w}^{(i)} : \Gamma_S = \Gamma_0 \xrightarrow{\tilde{w}_1} \Gamma_1 \xrightarrow{\tilde{w}_2} \dots \xrightarrow{\tilde{w}_i} \Gamma_i$$

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

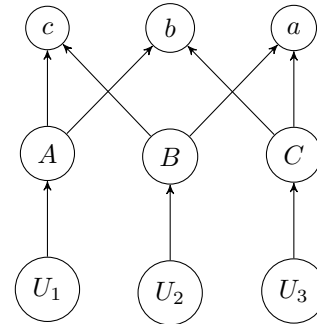
Konstruktionsbeschreibungen werden wir in der formalen Notation eines **geometrischen Linearprogramms** (GLP) angeben. Ein solches Programm enthält in den ersten Zeilen die

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

```

MP U1(ax, ay);
MP U2(bx, by);
MP U3(cx, cy);
Point A = freePoint(U1);
Point B = freePoint(U2);
Point C = freePoint(U3);
Line a = pp_line(B,C);
Line b = pp_line(A,C);
Line c = pp_line(A,B);
    
```



Die ersten drei Zeilen enthalten die dynamischen Parameter, die restlichen sechs Zeilen beschreiben die Abhängigkeit der geometrischen Objekte von den zuvor konstruierten. Auf der rechten Seite ist der zugehörige Abhängigkeitsgraph dargestellt.

Sind in solchen GLP als Abkürzungen geschachtelte Funktionsaufrufe 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  $ABC$ . Jedes GLP in schwacher Notation kann durch Einführung *anonymer geometrischer Objekte* in die strenge Notation überführt werden. 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);
    
```

Untersuchen wir nun genauer, wann eine Konstruktionsbeschreibung  $\mathcal{K}$  auf einer (zulässigen) speziellen Realisierung der Startkonfiguration  $O_S$  ausführbar ist. Die Konstruktionsbeschreibung ist genau dann *nicht* ausführbar, wenn für ein  $i > 0$  gilt, dass die Konstruktionsbeschreibung  $\mathcal{K}^{(i-1)}$  ausführbar ist,  $\mathcal{K}^{(i)}$  aber nicht mehr. Sei  $u^{(i)}$  die Parameterauswahl der Anwendung des Werkzeugs  $w_i$  im Schritt  $i$ . Wir können (und wollen)  $u^{(i)} \in O_E^n$  annehmen, da alle vor Schritt  $i$  konstruierten Objekte auch in der Endkonfiguration enthalten sind. Als Degenerationsbedingung im Schritt  $i$  ergibt sich  $w_i.DG(u^{(i)}.c)$ , als vollständige Degenerationsbedingung die Disjunktion

$$\bigvee_{i>0} w_i.DG(u^{(i)}.c)$$

dieser Bedingungen, wobei davon ausgegangen wird, dass dieser Ausdruck *kurz* (short) nach wachsendem  $i$  ausgewertet wird, da die in  $w_i.DG(u^{(i)}.c)$  eingehenden Argumente zur Auswertung dieser Formel vorher berechnet werden, die entsprechenden Konstruktionsschritte auf dieser speziellen Realisierung der Konfiguration dazu ausführbar sein müssen.

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

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

von  $w$  durch Projektion auf die  $U$ -Komponenten gewinnen. Dasselbe gilt für die Berechnungsvorschrift  $\tilde{w}.c$ . 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 der Konstruktionsbeschreibung  $\mathcal{K}$ , so ist

$$\tilde{w}_K = \tilde{w}_N \circ \cdots \circ \tilde{w}_1 : \prod (T(o) \mid o \in O_S) \rightarrow \prod (T(o) \mid o \in O_E)$$

das zugehörige verallgemeinerte Konstruktionswerkzeug. Die Degenerationsbedingung  $\tilde{w}_K.DG$  ergibt sich wie oben als oder-Verknüpfung der einzelnen  $w_i.DG(u^{(i)}.c)$

**Die Simonsche Gerade.** Sehen wir uns als komplexeres Beispiel die Konstruktionsbeschreibung zum Satz über die Simonsche Gerade an.

```

MP U1(ax, ay);
MP U2(bx, by);
MP U3(cx, cy);
SP S(u);
Point A = freePoint(U1);
Point B = freePoint(U2);
Point C = freePoint(U3);
Line a = pp_line(B,C);
Line b = pp_line(A,C);
Line c = pp_line(A,B);
Circle k = P3.circle(A,B,C);
Point P = circle_slider(k,S);
Point A1 = pedalpoint(P,a);
Point B1 = pedalpoint(P,b);
Point C1 = pedalpoint(P,c);

```

Damit ist die Konstruktionsbeschreibung  $K$  gegeben, mit welcher die *Endkonfiguration*  $\Gamma_E = (O_E, E_E)$  erzeugt wird, in der der Satz über die Simonsche Gerade  $is\_collinear(A1, B1, C1)$  gelten soll. Genauer besagt der entsprechende Satz, dass *in jeder nicht degenerierten speziellen Realisierung dieser geometrischen Konstruktion* (so dass also die drei Punkte  $A_1$ ,  $B_1$  und  $C_1$  konstruiert werden können) die drei Punkte  $A_1$ ,  $B_1$  und  $C_1$  auf einer gemeinsamen Geraden liegen.

In einer DGS müssen zur bildlichen dynamischen Darstellung die Koordinaten der speziellen Realisierungen geometrischer Objekte nach dieser Konstruktionsbeschreibung nach jedem Mausevent neu berechnet werden. Da mit den eingeführten geometrischen Objekten weitere Attribute wie Farbe, Linienstärke, Bezeichnung usw. verbunden sind, müssen die entsprechenden Objekte in einer zur Laufzeit zu verwaltenden *Symboltabelle* aufgesammelt werden.

Interpretierte Sprachen sind hierfür besser geeignet als compilierte, da dort bereits eine solche Symboltabelle angelegt ist. In compilierenden Programmiersystemen wie Java wird die Symboltabelle im Prinzip nur zur Übersetzungszeit benötigt. Im Design eines DGS auf dieser Basis muss also eine solche Symboltabelle (mit ihrer zentralen Eigenschaft der eindeutigen Repräsentation) softwareseitig angelegt und verwaltet werden.

Andererseits legt das vielfache Ausführen der mit einer Konstruktion verbundenen stets gleichen Berechnungsvorschriften auf verschiedenen (Maus)-Eingabewerten nahe, den entsprechenden Code zur Laufzeit „on the fly“ zu compilieren, wenn ein solches Feature vom verwendeten Programmiersystem angeboten wird. Weiter ist zu berücksichtigen, dass spezielle Parameterwerte zu degenerierten Situationen führen können, in denen einzelne spezielle Realisierungen von Objekten der jeweiligen Konstruktion nicht existieren und damit davon abhängende spezielle Realisierungen weiterer geometrischer Objekte auf einem ganzen Zweig des Abhängigkeitsgraphen nicht konstruiert werden können. Als Vorstufe davon können einzelne spezielle Realisierungen von Objekte zwar konstruierbar sein, aber außerhalb der Zeichenfläche liegen. Dies muss innerhalb der DGS abgefangen werden.

### 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 berechnet sich der Wert des Koordinatenattributs einer speziellen Realisierung von  $o$  aus den aktuellen Werten der Koordinatenattribute einer speziellen Realisierung von  $o_1, \dots, o_n$  als  $o.c = w.c(o_1.c, \dots, o_n.c)$ . Dabei sind jedesmal dieselben Rechnungen  $w.c$  mit je anderen Werten, also eine allgemeine Berechnungsvorschrift auszuführen.

Ist allgemeiner

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

eine Konstruktionsbeschreibung, so berechnen sich die Koordinatenwerte einer speziellen Realisierung der Endkonfiguration  $O_E = \mathcal{K}(O_S)$  – Ausführbarkeit vorausgesetzt – nach einer Berechnungsvorschrift, die sich als Zusammensetzung  $\tilde{w}_N.c \circ \dots \circ \tilde{w}_1.c$  der Berechnungsvorschriften der einzelnen Konstruktionsschritte ergibt.

Damit entsteht die Frage, ob sich diese häufig auszuführenden Berechnungsvorschriften durch gleichwertige, aber effizienter auszuführende ersetzen lassen. Aus Sicht der Informatik wäre insbesondere über eine compilierte Form nachzudenken, was eine DGS mit entsprechenden Fähigkeiten zur inkrementellen Übersetzung erfordert. Dieser Ansatz soll uns hier nicht weiter interessieren.

Wir wollen stattdessen eingeschränkte Klassen von Berechnungsvorschriften betrachten, auf die sich unsere Anwendungen bisher immer reduzieren ließen: Die Koordinaten von  $o.c$  sollen sich durch arithmetische Operationen aus den Koordinaten von  $o_1.c, \dots, o_n.c$  berechnen lassen.

*Beispiel:* Schnittpunkt der Mittelsenkrechten. Siehe `geo-0.txt`.

Die entsprechenden Berechnungsvorschriften für die Koordinatenattributwerte können in Maxima wie folgt angeschrieben werden:

```
pp_line(A,B):=
  if is(A=B) then error("Punkte fallen zusammen")
  else Line(part(b,2)-part(a,2),part(a,1)-part(b,1),
    part(a,2)*part(b,1)-part(a,1)*part(b,2));

midpoint(A,B):=
  Point(1/2*(part(a,1)+part(b,1)),1/2*(part(a,2)+part(b,2)));

ortho_line(P,a):=
  block([u:first(a), v:second(a)], Line(v,-u,u*second(p)-v*first(p)));

p_bisector(B,C):=
  ortho_line(midpoint(B,C),pp_line(B,C));

intersection_point(a,b):=
  block([d,d1,d2],
  d:expand(part(a,1)*part(b,2)-part(b,1)*part(a,2)),
```



```

d1:=expand(part(a,3)*part(b,2)-part(b,3)*part(a,2)),
d2:=expand(part(a,1)*part(b,3)-part(b,1)*part(a,3)),
if d=0 then error("Lines are parallel"),
Point(-d1/d,-d2/d));

```

```
on_line(P,a):=ratsimp(part(p,1)*part(a,1)+part(p,2)*part(a,2)+part(a,3));
```

Wir setzen nun die Koordinaten der einzelnen Punkte mit Unbestimmten an und erhalten daraus die Berechnungsvorschrift für die Mittelsenkrechte von  $\overline{AB}$ .

```

/* start configuration by Point constructor */
A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy);
/* construction step 1, variables as input */
ma:p_bisector(B,C);

```

$$[2b_x - 2c_x, 2b_y - 2c_y, c_y^2 + c_x^2 - b_y^2 - b_x^2]$$

Ähnliche Vereinfachungen ergeben sich für die Berechnungsvorschrift des Schnittpunkts  $M$  der Mittelsenkrechten  $m_a$  und  $m_b$ :

```

mb:p_bisector(A,C);
/* construction step 2, formulas as input */
M:intersection_point(ma,mb);

```

$$\left[ \begin{array}{l} \frac{(b_y - c_y) \left( (a_x - c_x) \left( \frac{a_x}{2} + \frac{c_x}{2} \right) + (a_y - c_y) \left( \frac{a_y}{2} + \frac{c_y}{2} \right) \right) - (a_y - c_y) \left( (b_x - c_x) \left( \frac{b_x}{2} + \frac{c_x}{2} \right) + (b_y - c_y) \left( \frac{b_y}{2} + \frac{c_y}{2} \right) \right)}{a_x b_y - a_y b_x - a_x c_y + a_y c_x + b_x c_y - b_y c_x}, \\ \frac{(b_x - c_x) \left( (a_x - c_x) \left( \frac{a_x}{2} + \frac{c_x}{2} \right) + (a_y - c_y) \left( \frac{a_y}{2} + \frac{c_y}{2} \right) \right) - (a_x - c_x) \left( (b_x - c_x) \left( \frac{b_x}{2} + \frac{c_x}{2} \right) + (b_y - c_y) \left( \frac{b_y}{2} + \frac{c_y}{2} \right) \right)}{a_x b_y - a_y b_x - a_x c_y + a_y c_x + b_x c_y - b_y c_x} \end{array} \right]$$

```
ratsimp(M);
```

$$\left[ \begin{array}{l} \frac{-a_x^2 b_y + a_x^2 c_y - a_y^2 b_y + a_y^2 c_y + a_y b_x^2 + a_y b_y^2 - a_y c_x^2 - a_y c_y^2 - b_x^2 c_y - b_y^2 c_y + b_y c_x^2 + b_y c_y^2}{2a_x b_y - 2a_y b_x - 2a_x c_y + 2a_y c_x + 2b_x c_y - 2b_y c_x}, \\ \frac{-a_x^2 b_x + a_x^2 c_x + a_x b_y^2 + a_x b_y^2 - a_x c_x^2 - a_x c_y^2 - a_y^2 b_x + a_y^2 c_x - b_x^2 c_x + b_x c_x^2 + b_x c_y^2 - b_y^2 c_x}{2a_x b_y - 2a_y b_x - 2a_x c_y + 2a_y c_x + 2b_x c_y - 2b_y c_x} \end{array} \right]$$

Nun bestimmen wir noch die Berechnungsvorschrift für  $m_c$  und den Test der Bedingung, dass  $M$  auf  $m_c$  liegt:

```
mc:=p_bisector(A,B);
/* check the conjecture as step 3, formulas as input */
result:=on_line(M,mc);
```

0

Alle Berechnungsvorschriften präsentieren sich als komplizierte rationale Ausdrücke in den Eingangsvariablen. Der letzte rationale Ausdruck vereinfacht zu null.

Da in eine Konstruktion nur endlich viele Parameterobjekte eingehen und diese die einzige Quelle der Dynamik der Konstruktion sind, lässt sich auch im allgemeinen Fall jede Berechnungsvorschrift durch arithmetische Operationen aus Kernen aufbauen, die nur Koeffizienten aus einem effektiv berechenbaren Grundkörper  $k$  und Koordinatenwerte  $\gamma_a$  der Parameterobjekte enthalten.

Ersetzen wir diese Kerne durch Variablen  $x_a$ , so wird die Berechnungsvorschrift jeder einzelnen Koordinate durch einen rationalen Ausdruck  $A(x_a) \in k(x_a)$  gegeben und wir können nach Vereinfachungsmöglichkeiten dieses Ausdrucks zu einem Ausdruck  $A'(x_a)$  in  $k(x_a)$  fragen. Die Berechnungsvorschrift lässt sich aus  $A$  durch die Substitution  $s = \{x_a \rightarrow \gamma_a\}$  zurückgewinnen. Nach Definition von  $k(x_a)$  sind  $A$  und  $A'$  (bis auf mögliche Unterschiede im Definitionsbereich, wenn die Ausdrücke Nenner enthalten) berechnungsäquivalent. Im Rahmen eines CAS kann die Umwandlung eines rationalen Ausdrucks in eine Berechnungsvorschrift durch Ersetzen der Variablen durch Zahlenwerte mit dem Substitutionsoperator `ev` erreicht werden, da bei der Auswertung des Ausdrucks mit Zahlenwerten die entsprechenden Operationen als Funktionsaufrufe interpretiert werden. Dabei kann es allerdings geschehen, dass die Berechnung mit einem Fehler (Ausnahme) abbricht.

Wir werden die Möglichkeiten dieses Ansatzes zunächst an einigen Beispielen betrachten und uns später für die Beweiskraft der entsprechenden Rechnungen interessieren. Dazu benötigen wir zunächst Implementierungen der Berechnungsvorschriften der verschiedenen Konstruktionswerkzeuge in einem CAS, die in der Lage sind, diese Berechnungen auch auf symbolischen Eingabewerten auszuführen. Wir hatten gesehen, dass sich bei diesen Berechnungen rationale Ausdrücke in  $k(x_a)$  als Koordinatenwerte ergeben, die sich durch Termumformungen weiter vereinfachen lassen. Da für rationale Ausdrücke mit der *rationalen Normalform* stets eine „einfachste“ Darstellung existiert, die sich in allen CAS durch entsprechende Transformationsfunktionen (in Maxima durch die Funktion `ratsimp`) effizient berechnen lassen, wollen wir weiter voraussetzen, dass in den Berechnungsfunktionen bei symbolischen Eingabewerten stets diese rationale Normalform berechnet wird.

Entsprechende Implementierungen stehen für Maple, Maxima, 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.

**Beispiele:** Siehe `geo-1.txt`

Die Ausdrücke  $A$  bzw.  $A'$  sind *universelle Formeln* der zum Werkzeug  $w$  gehörenden Berechnungsvorschrift  $w.c$ . Wir wollen diesen Begriff nun präziser einführen.

Betrachten wir dazu den Konstruktionsschritt  $o = w(o_1, \dots, o_n)$ , der aus gegebenen Objekten  $o_1, \dots, o_n$  das neue Objekt  $o$  erzeugt. Da in die Berechnung  $w.c$  die Koordinaten der aktuellen Realisierungen der Aufrufobjekte eingehen, muss zunächst ein Mechanismus eingeführt werden, diese Parameter durch Variable zu ersetzen. Mit diesem Ersetzungsmechanismus erzeugen wir *universelle Realisierungen* der gegebenen Objekte  $o_1, \dots, o_n$ , aus denen sich alle anderen Realisierungen ableiten lassen.

Die Wertebereiche  $C(T)$  geometrischer Typen lassen sich in allen bisher betrachteten Beispielen auf Tupel von Zahlen aus einem Grundbereich  $K$  zurückführen, wenn wir

1. gewisse Tupel als *degeneriert* ausschließen, etwa das Tupel  $\text{Line}(0, 0, 0)$  für Geraden, und
2. berücksichtigen, dass verschiedene Tupel die gleiche spezielle geometrische Konfiguration beschreiben.

Letzteres ist allerdings für die Untersuchung geometrischer Zusammenhänge auf Allgemeingültigkeit nicht relevant, wenn die Aussage als

$$\text{(die geometrische Aussage) gilt für alle nicht degenerierten Tupel } (\gamma_a) \in K^N \quad (\text{A})$$

angeschrieben werden kann. Für die konkreten Anwendungen muss  $k \subset K$  gelten, die Operationen in  $K$  aber nicht unbedingt effektiv sein, da (A) mathematisch deduktiv zu beweisen ist.

Um die Zuordnung zwischen Formeln und speziellen geometrischen Realisierungen mit Koordinaten aus  $K$  genauer zu beschreiben, fixieren wir einen Satz  $X = (x_a : a \in I)$  von abzählbar vielen Variablen, den Polynomring  $R = k[x_a : a \in I]$  über  $k$  und  $Q(R)$  als dessen Quotientenkörper.

Wenn wir Realisierungen geometrischer Objekte mit Variablen statt Zahlen zulassen, so bedeutet dies praktisch, dass nicht nur Elemente aus  $K$ , sondern aus  $R$  an allen Stellen als Koordinaten erlaubt werden. Da die Berechnungsvorschriften nur arithmetische Operationen enthalten, sind sie auch auf solchen Eingaben – sofern es sich nicht um über  $w.DG$  beschriebene degenerierte Lagen handelt – ausführbar und die Ergebnisse haben Koordinaten aus  $Q(R)$ .

**Definition 7** Als universelle Realisierung eines Objekts  $o$  vom Typ  $T$  bezeichnen wir jede Realisierung  $o.u$  über  $R$  mit einem Satz „frischer“ Variablen, so dass sich für jede spezielle Realisierung von  $o$  die Koordinatendarstellung durch geeignete Variablenspezifikation<sup>1</sup>  $s = \{x_a \rightarrow \gamma_a\}$  mit  $\gamma_a \in K$  ergibt:  $o.c = \text{subs}(o.u, s)$ .

„Frische“ Variablen heißt dabei: Ist

$$\text{var}(o.u) = \{a \in I \mid x_a \text{ kommt in } o.u \text{ wirklich vor}\}$$

<sup>1</sup>Genauer muss gefordert werden, dass sich nicht nur spezielle Realisierungen von  $o$  mit Koordinaten aus  $K$  ableiten lassen, sondern auch Realisierungen, deren Koordinaten in Erweiterungen von  $K$  liegen. Auf diese Feinheiten soll hier nicht eingegangen werden. Die Idee einer *universellen Realisierung eines geometrischen Objekts* folgt dem Konzept des allgemeinen Punkts einer Varietät in der algebraischen Geometrie.

die Menge der Indizes der in  $o.u$  wirklich vorkommenden Variablen, so sind diese Mengen für alle universellen Realisierungen aller Objekte paarweise disjunkt. Da in einer endlichen Beschreibung nur eine endliche Anzahl universeller Realisierungen endlich vieler Objekte verwendet werden kann,  $I$  aber als abzählbar unendlich vorausgesetzt wurde, ist eine solche Variablenauswahl immer möglich.

Eine universelle Realisierung eines geometrischen Objekts  $o$  ist selbst wieder eine spezielle Realisierung von  $o$  über dem erweiterten Koeffizientenbereich  $R$ .

Wir können nun den Konstruktionsschritt  $o = w(o_1, \dots, o_n)$  mit universellen Realisierungen  $o_1.u, \dots, o_n.u$  von  $O = (o_1, \dots, o_n)$  ausführen. Zur Ausführbarkeit des Konstruktionsschritts muss die Bedingung  $w.DGF = w.DG(o_1.u, \dots, o_n.u)$  untersucht werden. Für spezielle Realisierungen mit Koordinaten aus  $K$  wertet diese Bedingung stets zu einem der booleschen Werte **true** oder **false** aus. Das ist für universelle Realisierungen nicht mehr der Fall, da der entstehende Ausdruck  $w.DGF$  Variablen enthält, also eine boolesche Formel und kein boolescher Ausdruck mehr ist. Der Wertebereich dieser Formeln ist also **BooleanExpression** und nicht mehr **Boolean**. Als *Tautologie* bezeichnen wir jede boolesche Formel, die für alle Variablenbelegungen zu **true** auswertet. Wir wollen im Weiteren voraussetzen, dass  $w.DGF$  keine Tautologie ist.

Wenn in der Berechnungsvorschrift für  $w$  nur arithmetische Operationen verwendet werden, erhalten wir für die Koordinaten  $w.uc = o.c$  des Ergebnisses dieses Konstruktionsschritts rationale Ausdrücke in  $\{x_a \mid a \in \cup_i \text{var}(o_i.u)\}$ . Diese Ausdrücke stellen *universelle Formeln* in dem Sinne dar, dass jede Ausführung des Konstruktionsschritts mit einer *zulässigen* speziellen Realisierung von  $O$  zu einer speziellen Realisierung von  $o$  führt, deren Koordinaten sich durch die Variablenspezifikation  $s = \{x_a \rightarrow \gamma_a\}$  aus  $w.uc$  berechnen lassen, für die  $o_i.c = \text{subs}(o_i.u, s)$ ,  $i = 1, \dots, n$ , gilt. Ausführbarkeit bedeutet dabei, dass  $w.DG(o_1.c, \dots, o_n.c) = \text{subs}(w.DGF, s) = \text{false}$  gilt.

**Definition 8**  $w.uc$  bezeichnen wir als die universelle Formel des Konstruktionswerkzeugs  $w$ ,  $w.DGF$  als dessen universelle Degenerationsbedingung.

Ist

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

die Konstruktionsbeschreibung einer Konstruktion  $\mathcal{K}$ , so können wir diese Überlegungen auf jeden einzelnen Konstruktionsschritt anwenden.

Sei dazu  $\Gamma_S = (O_S, E_S)$  mit  $O_S = (o_1, \dots, o_m)$  die Startkonfiguration,  $\Gamma_E = (O_E, E_E)$  mit

$$O_E = (o_1, \dots, o_{m+N})$$

die Endkonfiguration und universelle Realisierungen

$$A_E = (o_1.u, \dots, o_{m+N}.u)$$

für die Objekte aus  $O_E$  fixiert (mit „frischen Variablen“ für die gegebenen geometrischen Objekte  $o_1, \dots, o_m$  und die neu zu konstruierenden Objekte  $o_{m+i}$ ,  $i = 1, \dots, N$ ).

Wir beginnen mit der universellen Realisierung  $o_1.u, \dots, o_m.u$  der Startkonfiguration und analysieren die Berechnungen der daraus im Zuge der Konstruktion  $K$  Schritt für Schritt zu konstruierenden Realisierungen  $o_{m+i}.uc$  der Objekte  $o_{m+i}$ ,  $i = 1, \dots, N$ .

$o_{m+1}.\mathbf{uc}$  berechnet sich aus der universellen Formel  $w_1.\mathbf{uc}$ , ist aber selbst keine universelle Realisierung mehr, sondern ergibt sich aus der universellen Realisierung  $o_{m+1}.\mathbf{u}$  durch die Substitution

$$s_1 = \{x_a \rightarrow w_1.\mathbf{uc}_a\} \text{ für } x_a \in \text{var}(o_{m+1}.\mathbf{u}),$$

wobei  $w_1.\mathbf{uc}_a$  den zum Index  $a$  gehörenden Teil von  $w_1.\mathbf{uc}$  bezeichnet. Es werden also die Variablen in  $o_{m+1}.\mathbf{u}$  durch die Formeln  $w_1.\mathbf{uc}$  ersetzt:

$$o_{m+1}.\mathbf{uc} = w_1.\mathbf{uc} = \mathbf{subs}(o_{m+1}.\mathbf{u}, s_1).$$

Die Berechnung ist ausführbar, wenn  $w_1.\mathbf{DGF}$  keine Tautologie ist.

In der universellen Formel  $w_2.\mathbf{uc}$  wird statt der konstruierten Realisierung  $o_{m+1}.\mathbf{uc}$  die universelle Realisierung  $o_{m+1}.\mathbf{u}$  verwendet.  $o_{m+2}.\mathbf{uc}$  ergibt sich also, wenn in der universellen Formel  $w_2.\mathbf{uc}$  die Substitution  $s_1$  ausgeführt wird:

$$o_{m+2}.\mathbf{uc} = \mathbf{subs}(w_2.\mathbf{uc}, s_1)$$

Kurz, in der universellen Formel  $w_2.\mathbf{uc}$  haben wir die Variablen  $x_a \in \text{var}(o_{m+1}.\mathbf{u})$  durch rationale Ausdrücke zu ersetzen. Die Berechnung ist ausführbar, wenn  $\mathbf{subs}(w_2.\mathbf{DGF}, s_1)$  keine Tautologie ist, was wir voraussetzen wollen.

$o_{m+2}.\mathbf{uc}$  ergibt sich auch als Folge der Substitutionen

$$s_2 = \{x_b \rightarrow w_2.\mathbf{uc}_b\} \text{ für } x_b \in \text{var}(o_{m+2}.\mathbf{u}),$$

mit der die Variablen in der universellen Realisierung  $o_{m+2}.\mathbf{u}$  durch die universellen Formeln  $w_2.\mathbf{uc}$  ersetzt werden, und  $s_1$ , mit der die Variablen  $x_a \in \text{var}(o_{m+1}.\mathbf{u})$  ersetzt werden.

$$o_{m+2}.\mathbf{uc} = \mathbf{subs}(o_{m+2}.\mathbf{u}, s_2, s_1)$$

Dies setzt sich über die anderen Konstruktionsschritte fort, so dass sich die Koordinaten der Realisierung der Endkonfiguration  $O_E$  aus der universellen Realisierung von  $O_E$  durch die Substitutionen  $s_N, s_{N-1}, \dots, s_1$  ergeben, mit denen rückwärts Schritt für Schritt für die freien Variablen der universellen Realisierungen der abhängigen Objekte die rationalen Ausdrücke eingesetzt werden, welche sich im jeweils vorherigen Konstruktionsschritt 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.

Als universelle Degenerationsbedingung der gesamten Konstruktion  $\mathcal{K}$  ergibt sich

$$\mathcal{K}.\mathbf{DGF} = \bigvee_{i=1}^N \mathbf{subs}(w_i.\mathbf{DGF}, s_{i-1}, \dots, s_1).$$

Wir setzen wieder voraus, dass dies keine Tautologie ist, da sonst  $\mathcal{K}$  auf keiner Startkonfiguration ausführbar wäre.

**Definition 9** *Eine Realisierung  $B_E = (o_1.\mathbf{u}, \dots, o_m.\mathbf{u}, o_{m+1}.\mathbf{uc}, \dots, o_{m+N}.\mathbf{uc})$  der Endkonfiguration  $O_E = \mathcal{K}(O_S)$ , die mit einer universellen Realisierung von  $O_S$  startet, bezeichnen wir als universelle Realisierung der Konstruktion  $\mathcal{K}$ , die dabei berechneten Formeln  $o_i.\mathbf{uc}$  der Realisierungen der abgeleiteten Objekte als deren universelle Formeln und  $\mathcal{K}.\mathbf{DGF}$  als universelle Degenerationsbedingung der Konstruktion.*

Sehen wir uns als Beispiel noch einmal die Konstruktion  $\mathcal{K}$  des Umkreismittelpunkts  $M$  des Dreiecks  $ABC$  als Schnittpunkt der Mittelsenkrechten  $m = m_a$ ,  $n = m_b$  an. Wir überführen dazu zunächst den Aufruf

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

in ein GLP:

```
Start(Point A,B,C);
Line g1:=p_bisector(B,C);
Line g2:=p_bisector(A,C);
Point M:=intersection_point(g1,g2);
```

Als universelle Realisierung der Startkonfiguration setzen wir  $A.u = (a_x, a_y)$ ,  $B.u = (b_x, b_y)$ ,  $C.u = (c_x, c_y)$  und ergänzen dies mit  $g_1.u = (m_1, m_2, m_3)$ ,  $g_2.u = (n_1, n_2, n_3)$ ,  $M.u = (m_x, m_y)$  zu einer universellen Realisierung der Endkonfiguration.

Auf dem betrachteten symbolischen Niveau bedeutet die Ausführung der Konstruktion auf der universellen Startkonfiguration, dass in der universellen Formel

$$w.\text{uc} = \left( \frac{m_2 n_3 - m_3 n_2}{m_1 n_2 - m_2 n_1}, \frac{m_3 n_1 - m_1 n_3}{m_1 n_2 - m_2 n_1} \right)$$

des Konstruktionswerkzeugs `intersection_point`, nach der sich die Koordinaten des Schnittpunkts der universellen Realisierungen  $g_1.u$  und  $g_2.u$  der beiden Geraden  $g_1, g_2$  berechnen, die Variablen von  $m_i, n_i$  durch die entsprechenden rationalen Ausdrücke

$$s_1 = \left( m_1 = b_x - c_x, m_2 = b_y - c_y, m_3 = \frac{1}{2} (-b_x^2 - b_y^2 + c_x^2 + c_y^2) \right)$$

bzw.

$$s_2 = \left( n_1 = -a_x + c_x, n_2 = -a_y + c_y, n_3 = \frac{1}{2} (a_x^2 + a_y^2 - c_x^2 - c_y^2) \right)$$

zu ersetzen sind. Diese Ausdrücke sind ihrerseits während des Aufrufs

```
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  $M.\text{uc}$  für den Umkreismittelpunkt  $M$  die früher berechnete Formel.

Die entsprechenden universellen Degenerationsbedingungen lauten

$$\text{intersection\_point.DGF} = \text{iszero}(m_1 n_2 - m_2 n_1)$$

sowie

$$\text{p\_bisector.DGF} = (u_x = v_x) \wedge (u_y = v_y).$$

Daraus ergibt sich

$$\begin{aligned} \mathcal{K}.DGF = & \text{iszero}((b_x - c_x)(a_y - c_y) - (a_x - c_x)(b_y - c_y)) \\ & \vee ((b_x = c_x) \wedge (b_y = c_y)) \vee ((a_x = c_x) \wedge (a_y = c_y)), \end{aligned}$$

was sich zu

$$\text{iszero}((b_x - c_x)(a_y - c_y) - (a_x - c_x)(b_y - c_y)) \quad (\text{B})$$

vereinfachen lässt, da für  $(b_x = c_x) \wedge (b_y = c_y) = \text{true}$  die erste Klammer zu **true** auswertet. Dasselbe gilt für den ersten und dritten Ausdruck der oder-Verknüpfung.

(B) ist die universelle Formel der geometrischen Bedingung `is_collinear(A,B,C)` und als Verschwinden eines polynomialen Ausdrucks in den Koordinaten der universellen Realisierungen von  $A, B, C$  angeschrieben. Formeln ähnlicher Gestalt ergeben sich aus anderen geometrischen Bedingungen. Untersuchen wir etwa, ob  $M$  tatsächlich der Umkreismittelpunkt ist, so sind die booleschen Formeln

$$(\text{sqrdist}(M,A)=\text{sqrdist}(M,B)) \text{ and } (\text{sqrdist}(M,A)=\text{sqrdist}(M,C));$$

auf obiger Konfiguration  $\mathcal{K}$  auszuwerten. Auch diese können als Verschwinden polynomialer Ausdrücke

$$\text{iszero}(\text{sqrdist}(M,A)-\text{sqrdist}(M,B)) \text{ and } \text{iszero}(\text{sqrdist}(M,A)-\text{sqrdist}(M,C));$$

angeschrieben werden.

Für obige universelle Realisierung der Endkonfiguration  $(A, B, C, g_1, g_2, M)$  ergibt sich

$$a_x^2 - 2m_x a_x + a_y^2 - 2m_y a_y - b_x^2 + 2m_x b_x - b_y^2 + 2m_y b_y$$

als universelle Formel für `sqrdist(M,A)-sqrdist(M,B)`, die zu null vereinfacht, wenn die oben berechneten Ausdrücke für  $M$  eingesetzt werden.

**Definition 10** 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 Berechnungsvorschrift

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

und einer Degenerationsbedingung

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

so dass die Berechnungsvorschrift auf allen Tupeln  $(c_1, \dots, c_n)$  ausführbar ist, für die

$$\phi.DG(c_1, \dots, c_n) = \text{false}$$

gilt, bezeichnen wir als geometrische Eigenschaft.

Sind  $o_1 \in T_1, \dots, o_n \in T_n$  Objekte der richtigen Typen und  $o_1.\mathbf{u}, \dots, o_n.\mathbf{u}$  deren universelle Realisierungen, so bezeichnen wir

$$\phi.\mathbf{uc} = \phi.c(o_1.\mathbf{u}, \dots, o_n.\mathbf{u}) \in \text{BooleanExpression}$$

als die zugehörige universelle Formel dieser geometrischen Eigenschaft und

$$\phi.DGF = \phi.DG(o_1.\mathbf{u}, \dots, o_n.\mathbf{u}) \in \text{BooleanExpression}$$

als die zugehörige universelle Degenerationsbedingung dieser Eigenschaft.

Sowohl die zu beweisenden Aussagen der bisher betrachteten geometrischen Sätze als auch die Degenerationsbedingungen sind solche geometrischen Eigenschaften.

Wir können damit die typische Konstellation für eine ganze Klasse geometrischer Sätze genauer beschreiben:

*Wenn sich eine gewisse Menge geometrischer Objekte in einer durch eine Konstruktionsbeschreibung  $\mathcal{K}$  definierten Abhängigkeitsrelation befindet, dann gilt für diese Objekte eine zusätzliche geometrische Eigenschaft  $\phi$ .*

Genauer: Für jede zulässige spezielle Realisierung  $C_E$  der Endkonfiguration von  $\mathcal{K}$ , also eine solche mit  $\mathcal{K}.DG(C_E) = \text{false}$  (Ausführbarkeit der Konstruktionsvorschrift), ist  $\phi.DG(C_E) = \text{false}$  (Bestimmtheit der Schlussfolgerung) und  $\phi.c(C_E) = \text{true}$ :

$$\forall C_E \left( \text{not } \mathcal{K}.DG(C_E) \Rightarrow (\text{not } \phi.DG(C_E) \wedge \phi.c(C_E)) \right) \quad (\text{C})$$

(C) besteht aus zwei Teilen

$$\forall C_E \left( \text{not } \mathcal{K}.DG(C_E) \Rightarrow \text{not } \phi.DG(C_E) \right) \quad (\text{C.1})$$

und

$$\forall C_E \left( \text{not } \mathcal{K}.DG(C_E) \Rightarrow \phi.c(C_E) \right), \quad (\text{C.2})$$

wobei (C.1) in der Kontraposition auch als

$$\forall C_E \left( \phi.DG(C_E) \Rightarrow \mathcal{K}.DG(C_E) \right) \quad (\text{C.3})$$

angeschrieben werden kann. Dies ist selbst wieder ein – gewöhnlich wesentlich einfacher zu beweisender – geometrischer Satz; meist ist die Degenerationsbedingung  $\phi.DG$  sowieso leer. Wir wollen voraussetzen, dass dieser Satz gilt, so dass sich unsere Behauptung auf (C.2) reduziert, was äquivalent auch als

$$\forall C_E \left( \mathcal{K}.DG(C_E) \vee \phi.c(C_E) \right) \quad (\text{C.4})$$



angeschrieben werden kann – eine spezielle Realisierung der gegebenen Konfiguration ist entweder degeneriert für die Konstruktionsvorschrift *oder* die Schlussfolgerung gilt. Sätze dieser Struktur bezeichnen wir als *geometrische Sätze vom konstruktiven Typ*.

Abschließend sei angemerkt, dass die bisher betrachteten Beispiele 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  $(0 : 0 : 0)$  für die Koordinaten einer speziellen Realisierung ergibt.

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

Untersuchen wir nun die Beweiskraft der ausgeführten symbolischen Berechnungen für Sätze vom konstruktiven Typ näher. Sei dazu  $\mathcal{K}$  eine Konfiguration, die durch eine Konstruktionsbeschreibung

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

erzeugt wird, wie bisher  $O_S = (o_1, \dots, o_m)$  die Start- und  $O_E = (o_1, \dots, o_{m+N})$  die Endkonfiguration.

Wir fixieren wieder universelle Realisierungen  $A_E = (o_1.\mathbf{u}, \dots, o_{m+N}.\mathbf{u})$  der Objekte aus  $O_E$  und unterteilen die Menge der dabei eingeführten Variablen in zwei disjunkte Teilmengen  $X = (X_\alpha)$  und  $Y = (Y_\alpha)$ , wobei die Variablen aus  $X$  in den universellen Realisierungen  $A_S = (o_1.\mathbf{u}, \dots, o_m.\mathbf{u})$  der unabhängigen Objekte vorkommen und die Variablen aus  $Y$  in den universellen Realisierungen  $o_{m+1}.\mathbf{u}, \dots, o_{m+N}.\mathbf{u}$  der abgeleiteten Objekte. Wir setzen weiter  $S = k[X, Y]$  und  $R = k[X]$  für die Polynomringe in den entsprechenden Variablenmengen.

Sei weiter  $B_E = (o_1.\mathbf{u}, \dots, o_m.\mathbf{u}, o_{m+1}.\mathbf{uc}, \dots, o_{m+N}.\mathbf{uc})$  die Realisierung (universelle Formeln) der Endkonfiguration  $O_E$ , die aus der universellen Startkonfiguration  $A_S$  mit der Konstruktion  $\mathcal{K}$  erzeugt wurde, und  $\phi$  eine geometrische Eigenschaft, die auf dieser Konfiguration „allgemein“ gelten soll.

Dazu ist zu zeigen, dass für jede zulässige spezielle Realisierung  $C_E$  der Endkonfiguration  $\phi(C_E) = \mathbf{true}$  gilt:

$$\forall C_E \left( \mathcal{K}.\mathbf{DG}(C_E) \vee \phi.\mathbf{c}(C_E) \right) \quad (\text{C.4})$$

Sei dazu  $C_S = (o_1.\mathbf{c}, \dots, o_m.\mathbf{c})$  eine zulässige spezielle Realisierung der Startkonfiguration und  $C_E = (o_1.\mathbf{c}, \dots, o_m.\mathbf{c}, o_{m+1}.\mathbf{c}, \dots, o_{m+N}.\mathbf{c})$  die aus  $C_S$  durch  $\mathcal{K}$  erzeugte spezielle Realisierung der Endkonfiguration  $O_E$ .

$B_E$  entsteht aus  $A_E$  durch die Substitutionen  $s_N, \dots, s_1$  wie im letzten Abschnitt beschrieben, die zu einer Variablensubstitution  $\{y_\alpha \rightarrow y_\alpha(X)\}$  der abhängigen Variablen  $y_\alpha \in Y$

durch rationale Ausdrücke  $y_\alpha(X) \in Q(R)$  zusammengefasst werden können, die wir kurz als  $Y \rightarrow Y(X)$  anschreiben. Diese Variablensubstitution kann eindeutig zu einem Ringhomomorphismus  $\Phi : S \rightarrow Q(R)$  fortgesetzt werden. Das Bild  $\Phi(S)$  kann genauer beschrieben werden: Ist  $U \subset R$  die Menge der Nenner, die in den Formeln  $y_\alpha(X)$  vorkommen, so kommen in  $r \in \Phi(S)$  nur Nenner vor, die sich als Produkte von Elementen aus  $U$  darstellen lassen.

Sei weiter  $X \rightarrow X_0$  die Variablensubstitution, die  $A_S$  in  $C_S$  überführt, so dass also  $o_i.c = \text{subs}(o_i.u, X \rightarrow X_0)$  für  $i = 1, \dots, m$  gilt. Auch diese Substitution kann eindeutig zu einem Homomorphismus  $\Phi_0 : R \rightarrow K$  fortgesetzt werden, allerdings nicht auf ganz  $Q(R)$ .  $X \rightarrow X_0$  kann nicht auf solche Elemente  $\frac{f(X)}{g(X)} \in Q(R)$  fortgesetzt werden, für die  $g(X_0) = 0$  in  $K$  gilt. Wir bezeichnen die Substitution  $X \rightarrow X_0$  als *nicht zulässig*, wenn sie sich nicht auf ganz  $\Phi(S)$  fortsetzen lässt. Das ist genau dann der Fall, wenn es ein  $u(X) \in U$  mit  $u(X_0) = 0$  gibt, denn als Nenner treten in  $\Phi(S)$  nur Produkte von Polynomen aus  $U$  auf. Nach Konstruktion ist dann aber  $X_0$  Nullstelle des Nenners einer der Formeln  $y_\alpha(X) \in Q(R)$ . Wird die Koordinate  $y_\alpha$  in der Konstruktion von  $o_{m+i}.c$  verwendet, so ist  $o_{m+i}.c$  nicht definiert,  $\text{subs}(w_i.\text{DGF}, s_{i-1}, \dots, s_1) = \text{true}$  und damit auch

$$\mathcal{K}.\text{DG}(C_E) = \text{subs}(\mathcal{K}.\text{DGF}, X \rightarrow X_0) = \text{true}.$$

$C_S$  ist also auch eine für  $\mathcal{K}$  degenerierte Lage im früher definierten Sinne.

Zum Beweis der geometrischen Aussage, dass für jede zulässige spezielle Realisierung  $C_E$  der Endkonfiguration  $\phi(C_E) = \text{true}$  gilt, betrachten wir das folgende kommutative Diagramm:

$$\begin{array}{ccccc} A_E & \xrightarrow{Y \rightarrow Y(X)} & B_E & \xrightarrow{X \rightarrow X_0} & C_E \\ \downarrow \phi.c & & \downarrow \phi.c & & \downarrow \phi.c \\ \phi.c(A_E)(X, Y) & \xrightarrow{Y \rightarrow Y(X)} & \phi.c(B_E)(X) & \xrightarrow{X \rightarrow X_0} & \phi.c(C_E) \end{array} \quad (\text{BS.0})$$

In allen bisher betrachteten Beispielen gehörten die booleschen Formeln  $\phi.c(A_E)(X, Y)$  zur Prädikatenlogik erster Stufe mit Atomen der Bauart  $\text{iszero}(p)$  mit  $p \in S$ , und Vereinfachungen im Bereich `BooleanExpression` wurden stets in diesem Sinne interpretiert. Nach unseren bisherigen Überlegungen müssten wir eigentlich  $p \in Q(S)$  voraussetzen, aber da ein rationaler Ausdruck genau dann verschwindet, wenn dessen Zähler verschwindet, können (und werden) wir uns auf polynomiale Ausdrücke  $p \in S$  beschränken. Wir setzen im Weiteren voraus, dass booleschen Formeln für geometrische Eigenschaften und Degenerationsbedingungen stets zu dieser Klasse  $\mathcal{B}$  von Ausdrücken gehören.

Im Diagramm (BS.0) wird die Berechnungsvorschrift  $\phi.c$  der geometrischen Eigenschaft auf drei verschiedene geometrische Konfigurationen angewandt:

- Auf die universelle Realisierung  $A_E(X, Y)$  der Endkonfiguration, welche die Abhängigkeiten der Konstruktion nicht berücksichtigt, – es ergibt sich die universelle Formel

$$\phi.\text{uc} = \phi.c(A_E)(X, Y)$$

der geometrischen Eigenschaft.

- Auf die universellen Formeln  $B_E(X)$  der Endkonfiguration von  $\mathcal{K}$ , die sich aus  $A_E$  durch die Substitution  $Y \rightarrow Y(X)$  ergeben. Dabei gilt

$$\phi.c(B_E)(X) = \Phi(\phi.c(A_E)(X, Y)),$$

wenn  $\Phi$  auf  $\mathcal{B}$  operationstreu fortgesetzt wird. Das linke Diagramm kommutiert.

- Auf die spezielle Realisierung  $C_E$  der Endkonfiguration von  $\mathcal{K}$ , die sich aus  $B_E$  durch die Substitution  $X \rightarrow X_0$  ergibt. Dabei gilt

$$\phi.c(C_E) = \Phi_0(\phi.c(B_E)(X)),$$

wenn  $\Phi_0$  auf  $\mathcal{B}$  operationstreu fortgesetzt wird. Das rechte Diagramm kommutiert.

Ist die geometrische Eigenschaft  $\phi.c(A_E)$  selbst bereits atomar, hat also die universelle Formel  $\phi.uc = \phi.c(A_E)$  die Gestalt  $\phi.uc = \mathbf{iszero}(\psi)$  mit  $\psi \in S$ , so bezeichnen wir  $\psi.uc = \psi$  als *polynomiale universelle Formel* der geometrischen Eigenschaft  $\phi$ .

Wir erhalten damit aus (BS.0) das folgende Beweisschema-Diagramm

$$\begin{array}{ccccc} A_E & \xrightarrow{Y \rightarrow Y(X)} & B_E & \xrightarrow{X \rightarrow X_0} & C_E \\ \downarrow \psi & & \downarrow \psi & & \downarrow \psi \\ \psi(A_E) & \xrightarrow{Y \rightarrow Y(X)} & \psi(B_E) & \xrightarrow{X \rightarrow X_0} & \psi(C_E) \end{array} \quad (\text{BS.1})$$

in welchem nur noch rationale Ausdrücke vorkommen und das aus denselben Gründen wie oben kommutativ ist. Insbesondere ergibt sich  $\psi(B_E)$  aus der polynomialen universellen Formel  $\psi.uc(X, Y) = \psi(A_E)$  der geometrischen Eigenschaft  $\phi$  durch die Substitution  $Y \rightarrow Y(X)$ . Eine weitere Simplifikation vereinfachte dann in den bisher betrachteten Beispielen bereits diesen rationalen *Ausdruck* zu null, woraus folgt, dass  $\mathbf{iszero}(\psi(B_E))$  bereits als boolesche *Formel* unabhängig von Variablenbelegungen (und damit für *alle* Variablenbelegungen) zu **true** simplifiziert.

Auch die Degenerationsbedingung  $\mathcal{K}.DGF$  gehört zur Klasse  $\mathcal{B}$ , wenn dies für die Degenerationsbedingungen der einzelnen Konstruktionswerkzeuge gilt – sie ergibt sich aus den Degenerationsbedingungen der einzelnen Konstruktionswerkzeuge, wenn dort  $Y \rightarrow Y(X)$  ersetzt wird:

$$\begin{aligned} \mathcal{K}.DG(B_E) &= \mathcal{K}.DGF(X) = \mathbf{subs} \left( \bigvee_{i=1}^N w_i.DGF(X, Y), Y \rightarrow Y(X) \right) \\ \mathcal{K}.DG(C_E) &= \mathbf{subs}(\mathcal{K}.DGF(X), X \rightarrow X_0) \end{aligned}$$

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 11** *Einen Konstruktionswerkzeug  $w$  bezeichnen wir als rational bzw. polynomial, wenn dessen universelle Formel  $w.uc$  aus rationalen oder sogar polynomialen Ausdrücken besteht und für die universelle Degenerationsbedingung  $w.DGF \in \mathcal{B}$  gilt.*

*Eine Konstruktionsbeschreibung  $\mathcal{K}$  bezeichnen wir als polynomial bzw. rational, wenn sie aus Konstruktionsschritten aufgebaut ist, die mit polynomialen bzw. rationalen Konstruktionswerkzeugen angeschrieben sind.*

Eine geometrische Bedingung  $\phi \in \mathcal{B}$  bezeichnen wir als polynomiale Bedingung.

Als geometrischen Satz vom rationalen konstruktiven Typ bezeichnen wir eine rationale Konstruktionsbeschreibung  $\mathcal{K}$  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  $C_S$  der Startkonfiguration von  $\mathcal{K}$  die Bedingung  $\phi$  auf der zugehörigen Realisierung  $C_E$  der Endkonfiguration zu **true** ausgewertet.

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.

Für einen Satz  $(\mathcal{K}, \phi)$  vom rationalen konstruktiven Typ, dessen geometrische Behauptung durch eine atomare Formel  $\phi.\text{uc} = \text{iszero}(\psi)$  angeschrieben ist, haben wir die folgenden Formeln:

- Eine universelle Realisierung  $A_S(X) = (o_1.\mathbf{u}, \dots, o_m.\mathbf{u})$  der Startkonfiguration,
- eine universelle Realisierung  $A_E(X, Y) = (o_1.\mathbf{u}, \dots, o_{m+N}.\mathbf{u})$  der Endkonfiguration,
- universelle Formeln  $B_E(X) = \text{subs}(A_E, Y \rightarrow Y(X))$  für die Realisierung von  $\mathcal{K}$  auf der universellen Realisierung  $A_S$  der Startkonfiguration,
- die universelle Formel  $\Psi(X, Y) = \psi(A_E)$  der Behauptung  $\phi$  sowie
- das Resultat der Substitution  $\Psi'(X) = \psi(B_E) = \text{subs}(\Psi(X, Y), Y \rightarrow Y(X))$ .

Für die ausgeführten symbolischen Rechnungen gibt es folgende Alternativen:

1.  $A_S$  selbst ist für  $\mathcal{K}$  nicht zulässig, d. h.  $\mathcal{K}.\text{DGF}(X)$  vereinfacht als **BooleanExpression** zu **true**. Es gibt also auch keine zulässigen speziellen Realisierungen – die Voraussetzungen des Satzes sind widersprüchlich.

Sei in diesem Fall  $i$  minimal mit der Eigenschaft, dass  $\text{subs}(w_i.\text{DGF}, Y \rightarrow Y(X))$  zu **true** vereinfacht. Damit kann die Teilkonstruktion  $\mathcal{K}^{(i-1)}$  ausgeführt werden und  $A_S$  ist für  $\mathcal{K}^{(i-1)}$  zulässig.

Die Nichtausführbarkeit der Konstruktion lässt sich damit als geometrischer Satz vom konstruktiven Typ mit der Konstruktionsbeschreibung  $\mathcal{K}^{(i-1)}$  und der Behauptung  $w_i.\text{DG}$  formulieren: *Wenn man wie angegeben bis zum Schritt  $i - 1$  konstruiert, dann landet man immer in einer degenerierten Lage des Konstruktionswerkzeugs  $w_i$ .*

2. Die zurückgegebene rationale Funktion  $\Psi'(X)$  simplifiziert zu null.

Dann gilt die geometrische Aussage für alle zulässigen speziellen Realisierungen  $C_E$  der Endkonfiguration, da sich der Wert von  $\psi$  auf  $C_E$  durch Variablenspezifikation  $X \rightarrow X_0$  aus  $\Psi'$  ergibt. Der Satz ist allgemeingültig.

3. Die zurückgegebene rationale Funktion  $\Psi'(X)$  simplifiziert *nicht* zu null.

Dann gilt die Aussage für fast alle zulässigen speziellen Realisierungen der Endkonfiguration *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 vom konstruktiven Typ.

In vielen Fällen ist unklar, was „zulässig“ bedeutet, da die genauen Degenerationsbedingungen nicht bekannt sind. Sehen wir uns deshalb an, was wir über die Degenerationsbedingungen „im Allgemeinen“ wissen. Die Disjunktion

$$\bigvee_{i=1}^N w_i.\text{DGF}(X, Y) \in \mathcal{B}$$

der universellen Degenerationsbedingungen  $w_i.\text{DGF}(X, Y)$  der in der Konstruktion  $\mathcal{K}$  verwendeten Werkzeuge  $w_i$  können wir in die konjunktive Normalform

$$\bigvee_{i=1}^N w_i.\text{DGF}(X, Y) = b_1 \wedge b_2 \wedge \cdots \wedge b_k$$

überführen, wobei  $b_i$  Disjunktionen atomarer Ausdrücke in  $\mathcal{B}$  sind. Dann gilt für die universelle Degenerationsbedingung  $\mathcal{K}.\text{DGF}(X)$  der Konstruktion  $\mathcal{K}$

$$\mathcal{K}.\text{DGF}(X) = \Phi \left( \bigvee_{i=1}^N w_i.\text{DGF}(X, Y) \right) = \Phi(b_1) \wedge \Phi(b_2) \wedge \cdots \wedge \Phi(b_k),$$

wobei jedes der  $b_i$  eine Disjunktion von Atomen der Gestalt  $\text{iszero}(f(X, Y))$  für ein  $f \in S$  ist und  $\Phi(b(X, Y))$  für  $b(X, Y \rightarrow Y(X))$  steht. Ist  $\mathcal{K}.\text{DGF}(X)$  selbst nicht trivial (gilt also die obige Alternative 1 nicht), so muss einer der Ausdrücke  $\Phi(b_i)$  nicht trivial sein. Nehmen wir an, dies gilt für  $i = 1$  und es ist

$$b_1(X, Y) = \text{iszero}(f_1) \vee \text{iszero}(f_2) \vee \cdots \vee \text{iszero}(f_s) = \text{iszero}(f)$$

mit  $f_1, \dots, f_s \in S$  und  $f = f_1 \cdot f_2 \cdot \dots \cdot f_s$ . Dann gilt aber auch  $\Phi(b_1) = \text{iszero}(\Phi(f))$ , wobei  $\Phi(f) = f(X, Y \rightarrow Y(X)) \in Q(R)$  eine rationale Funktion ist, deren Zähler  $g(X) \in R$  nicht verschwindet.

Simplifiziert die rationale Funktion  $\Psi'(X)$  zu null, so gilt

$$\forall C_E \left( \mathcal{K}.\text{DG}(C_E) \vee \phi.c(C_E) \right) \tag{C.4}$$

und damit auch die schwächere Aussage

$$\forall C_E \left( b_1(C_E) \vee \phi.c(C_E) \right).$$

Es gilt also  $\phi.c(C_E) = \text{true}$  für alle Belegungen  $X \rightarrow X_0$  mit  $g(X_0) \neq 0$ . Wenn wir  $g \in R$  nicht explizit kennen, so können wir doch die folgende Aussage treffen:

**Satz 18 (Satz über die „allgemeine“ Gültigkeit eines Satzes vom rationalen konstruktiven Typ)**

Vereinfacht unter den obigen Voraussetzungen die rationale Funktion  $\Psi'(X) \in Q(R)$  zu null, so existiert ein Polynom  $g \in R$ , so dass für jede Substitution  $X \rightarrow X_0$  mit  $g(X_0) \neq 0$

1.  $C_E = \text{subs}(B_E, X \rightarrow X_0)$  zulässig ist und
2.  $\phi.c(C_E)$  gilt, der geometrische Satz für die spezielle Realisierung  $C_E$  also gültig ist.

Da  $g(X) = 0$  eine Hyperfläche im  $\mathbb{A}^N$  definiert, also ein Objekt der Dimension  $N - 1$ , gilt der Satz „fast überall“.

Genauer haben wir den folgenden Mechanisierungssatz bewiesen:

**Satz 19 (Satz über das mechanisierte Beweisen geometrischer Sätze vom rationalen konstruktiven Typ)**

Sei  $(\mathcal{K}, \phi)$  ein geometrischer Satz vom rationalen konstruktiven Typ mit einer atomaren Formel  $\phi.uc = \text{iszero}(\psi)$  als Aussage,

- $A_E(X, Y)$  eine universelle Realisierung der Endkonfiguration,
- $B_E(X) = \text{subs}(A_E, Y \rightarrow Y(X))$  das Ergebnis der Anwendung von  $\mathcal{K}$  auf eine universelle Realisierung  $A_S$  der Startkonfiguration,
- $\Psi' = \text{subs}(\psi(A_E), Y \rightarrow Y(X)) \in Q(R)$  der rationale Ausdruck, der durch Substitution der universellen Formeln der Konstruktion  $\mathcal{K}$  in die universelle Realisierung  $A_E$  entsteht, und

$$\mathcal{K}.DGF(X) = \Phi \left( \bigvee_{i=1}^N w_i.DGF(X, Y) \right)$$

die universelle Degenerationsbedingung der Konstruktion  $\mathcal{K}$ .

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

Zulässige spezielle Realisierungen  $C_S$  sind genau diejenigen, welche aus  $A_S = B_S$  durch eine Variablenspezifikation  $X \rightarrow X_0$  gewonnen werden können, für die  $\text{subs}(\mathcal{K}.DGF(X), X \rightarrow X_0) = \text{false}$  gilt.

Ist  $\mathcal{K}.DGF(X)$  als Boolesche Formel trivial, so gibt es keine zulässigen speziellen Realisierungen. Andernfalls existiert ein nicht triviales Polynom  $g \in R$ , so dass jede Variablenspezifikation  $X \rightarrow X_0$  mit  $g(X_0) \neq 0$  zu einer zulässigen speziellen Realisierung führt.

## 4 Weitere geometrische Sätze und Werkzeuge

### 4.1 Vor- und Nachbedingungen

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.

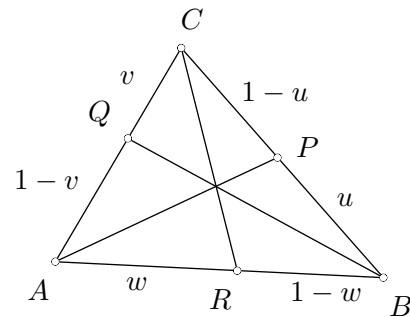
So behauptet etwa der **Satz des Ceva**:

Gegeben sei ein Dreieck  $ABC$  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.



In der im vorigen Abschnitt entwickelten Terminologie lässt sich  $\mathcal{K}$  wie folgt beschreiben:

```
Start(Point A,B,C; SP U,V,W);
Point P = varpoint(B,C,U);
Point Q = varpoint(C,A,V);
Point R = varpoint(A,B,W);
```

Da wir uns nur für die universellen Formeln der Konfiguration  $\mathcal{K}$  interessieren, werden wir im Weiteren die folgende kürzere Notation der GEOPROVER-Version 1.3 verwenden, die sich ausschließlich auf die Koordinatenattribute bezieht und Parameter direkt für freie Punkte (zwei Variablen) und Gleiter (eine Variable) einführt. Die universellen Formeln der Konfiguration  $\mathcal{K}$  ergeben sich dann unmittelbar als

```
A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy);
P:varpoint(B,C,u); Q:varpoint(C,A,v); R:varpoint(A,B,w);
```

Hierbei ist `varpoint` so skaliert, dass `subs(P, u = 0) = B` und `subs(P, u = 1) = C` gilt. Für die universelle Degenerationsbedingung ergibt sich  $\mathcal{K}.DGF = \text{is\_collinear}(A.u, B.u, C.u)$ .

Die weitere geometrische Voraussetzung sowie die Behauptung des Satzes von Ceva sind dann in den Formeln

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

kodiert.

In der gegebenen Konfiguration  $\mathcal{K}$ , in deren universellen Formeln die sechs Koordinaten von  $A, B, C$  und die Teilverhältnisse  $u, v, w$  als Variablen vorkommen, ist also zu zeigen, dass jede Nullstelle des Polynoms  $poly$  (der Voraussetzung) auch eine Nullstelle des Polynoms  $con$  (der Schlussfolgerung = conclusio) ist.

In diesem Fall lässt sich  $poly$  als  $-f_1^2 \cdot f_2$  mit

$$\begin{aligned} f_1 &= a_x b_y - a_x c_y - a_y b_x + a_y c_x + b_x c_y - b_y c_x \\ f_2 &= 2uvw - uv - uw + u - vw + v + w - 1 \end{aligned}$$

darstellen. Also gilt

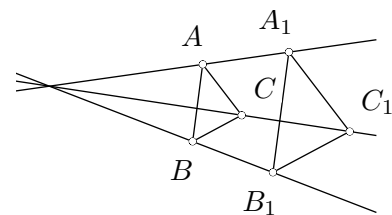
$$\text{iszero}(poly) \Leftrightarrow \text{iszero}(f_1) \vee \text{iszero}(f_2).$$

Analysiert man die beiden anderen Faktoren genauer, so erkennt man, dass  $f_1$  die universelle Formel der Bedingung  $\text{is\_collinear}(A, B, C)$  ist und  $f_2 = con$  gilt, so dass

$$\text{iszero}(poly) \Leftrightarrow \mathcal{K}.DGF \vee \text{iszero}(con)$$

gilt. Damit ist der Satz von Ceva entsprechend der im letzten Abschnitt entwickelten Methodik bewiesen.

Ähnlich können wir auch die Aussage des **affinen Satzes von Desargue** formulieren. Hier sind 6 Punkte  $A, B, C, A_1, B_1, C_1$  gegeben. Die Voraussetzung besteht aus drei geometrischen Bedingungen, dass in den Dreiecken  $ABC$  und  $A_1B_1C_1$  einander entsprechende Geraden zueinander parallel sind. Die Behauptung lautet, dass dann die Geraden  $AA_1, BB_1$  und  $CC_1$  konkurrent sind.



Die universellen Formeln der Beschreibung der zu Grunde liegenden Konfiguration

$$\mathcal{K} = \text{Start}(\text{Point}A, B, C, A_1, B_1, C_1)$$

bestehen aus genau den universellen Realisierungen der 6 Punkte  $A, B, C, A_1, B_1, C_1$

```
A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy);
A1:Point(dx,dy); B1:Point(ex,ey); C1:Point(fx,fy);
```

Die Voraussetzung lässt sich in den drei universellen Formeln

```
polys: [
  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))
];
```

kodieren, die Behauptung in der universellen Formel



```
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

$$\begin{aligned} & a_x d_y - a_x e_y - a_y d_x + a_y e_x - b_x d_y + b_x e_y + b_y d_x - b_y e_x \\ & b_x e_y - b_x f_y - b_y e_x + b_y f_x - c_x e_y + c_x f_y + c_y e_x - c_y f_x \\ & a_x d_y - a_x f_y - a_y d_x + a_y f_x - c_x d_y + c_x f_y + c_y d_x - c_y f_x, \end{aligned}$$

die Behauptung dem deutlich komplizierteren Polynom

$$\begin{aligned} & -a_x b_x c_y d_y + a_x b_x c_y e_y + a_x b_x d_y f_y - a_x b_x e_y f_y + a_x b_y c_x d_y - a_x b_y c_x f_y \\ & - a_x b_y c_y e_x + a_x b_y c_y f_x - a_x b_y d_y f_x + a_x b_y e_x f_y - a_x c_x d_y e_y + a_x c_x e_y f_y \\ & + a_x c_y d_y e_x - a_x c_y e_y f_x - a_x d_y e_x f_y + a_x d_y e_y f_x - a_y b_x c_x e_y + a_y b_x c_x f_y \\ & + a_y b_x c_y d_x - a_y b_x c_y f_x - a_y b_x d_x f_y + a_y b_x e_y f_x - a_y b_y c_x d_x + a_y b_y c_x e_x \\ & + a_y b_y d_x f_x - a_y b_y e_x f_x + a_y c_x d_x e_y - a_y c_x e_x f_y - a_y c_y d_x e_x + a_y c_y e_x f_x \\ & + a_y d_x e_x f_y - a_y d_x e_y f_x + b_x c_x d_y e_y - b_x c_x d_y f_y - b_x c_y d_x e_y + b_x c_y d_y f_x \\ & + b_x d_x e_y f_y - b_x d_y e_y f_x + b_y c_x d_x f_y - b_y c_x d_y e_x + b_y c_y d_x e_x - b_y c_y d_x f_x \\ & - b_y d_x e_x f_y + b_y d_y e_x f_x - c_x d_x e_y f_y + c_x d_y e_x f_y + c_y d_x e_y f_x - c_y d_y e_x f_x \end{aligned}$$

mit 48 Summanden vom Grad 4. Dass jede gemeinsame Nullstelle der Polynome *polys* 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 Algebraisierung ein beliebiges schiefwinkliges Koordinatensystem verwenden können, bzw. alternativ bei vorgegebenem Koordinatensystem durch eine geeignete affine Transformation die Konfiguration unter Erhaltung aller geometrischen Eigenschaften 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

$$\text{polys} = [d_x - e_x, -e_x - e_y + f_x + f_y, -d_y + f_y]$$

und als Behauptung

$$\text{con} = -d_x e_x f_y - d_x e_y f_y + d_x f_y + d_y e_x f_x + d_y e_x f_y - d_y e_x.$$

*polys* ist ein homogenes lineares Gleichungssystem mit Elementen aus dem Polynomring  $S = k[d_x, d_y, e_x, e_y, f_x, f_y]$ , welches sich einfach lösen lässt (Maxima)

```
sol:solve(polys,[dx,dy,ey]);
```

$$[[d_x = f_x, d_y = f_y + f_x - e_x, e_y = f_y + f_x - e_x]]$$

Wir haben drei der Parameter  $Y = (d_x, d_y, e_y)$  durch die anderen Parameter  $X = (f_x, f_y, e_x)$  ausgedrückt. In die Behauptung eingesetzt erhalten wir

```
expand(subst(sol[1],con));
```

0

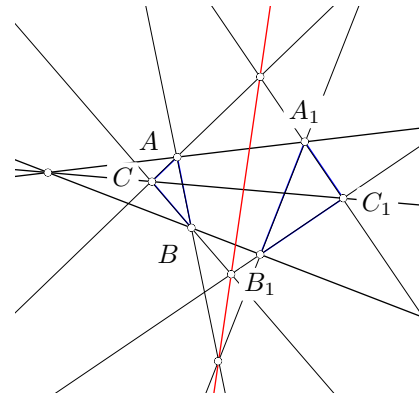
Der entstehende Ausdruck vereinfacht bereits als Polynom in  $R = k[f_x, f_y, e_x]$  zu null, woraus die Behauptung des Satzes von Desargue für alle zulässigen speziellen Realisierungen folgt.

Für den **allgemeinen Satz von Desargue** können wir ähnlich argumentieren. Wir beginnen wieder mit der Konfiguration

```
A:Point(ax,ay);
B:Point(bx,by);
C:Point(cx,cy);
A1:Point(dx,dy);
B1:Point(ex,ey);
C1:Point(fx,fy);
```

und erweitern diese um die universellen Formeln der Schnittpunktkoordinaten

```
X:intersection_point(pp_line(A,B),pp_line(A1,B1));
Y:intersection_point(pp_line(A,C),pp_line(A1,C1));
Z:intersection_point(pp_line(B,C),pp_line(B1,C1));
```



Voraussetzung und Behauptung des Satzes lassen sich dann in den rationalen Ausdrücken

```
poly: is_concurrent(pp_line(A,A1),pp_line(B,B1),pp_line(C,C1));
con: is_collinear(X,Y,Z);
```

kodieren. Während *poly* ein irreduzibles Polynom vierten Grades mit 48 Termen ist, hat *con* eine deutlich komplexere Struktur. Maxima stellt Zähler und Nenner von *con* als Polynome in rekursiver Normalform mit jeweils 8 Termen dar. Expandiert man die Ausdrücke mit `expand` in die distributive Normalform, so ergeben sich 1290 Zähler- und 378 Nennerterme. Der Nenner ist das Produkt der Degenerationsbedingungen

$$\text{is\_parallel}(\text{pp\_line}(A, B), \text{pp\_line}(A_1, B_1)) \text{ usw.}$$

des für  $X, Y, Z$  angewendeten Konstruktionswerkzeugs.

Da der Satz ein Satz (mindestens) der affinen Geometrie ist, können wir durch Auswahl des Koordinatensystems wieder die Rechnung vereinfachen. Wir setzen

```
A1:Point(0,0); B1:Point(0,1); C1:Point(1,0);
```

und erhalten nun

$$\text{con} = \frac{(a_y b_x - a_x c_y + a_x b_x c_y - a_y b_x c_x + a_x b_y c_y - a_y b_x c_y) (a_x b_y - a_y b_x - a_x c_y + a_y c_x + b_x c_y - b_y c_x)}{(b_x + b_y - c_x - c_y) (a_x - b_x) (a_y - c_y)}$$

Der Nenner ist wieder gerade die Degenerationsbedingung des Werkzeugs in den reduzierten Koordinaten, der Zähler lässt sich als  $\text{poly} \cdot \text{is\_collinear}(A, B, C)$  darstellen. Damit gilt also

auch hier  $poly(X_0) = 0 \Leftrightarrow con(X_0) = 0$  für jede nicht degenerierte spezielle Realisierung  $X \rightarrow X_0$  der Grundkonfiguration.

*Aufgabe:* Formulieren und untersuchen Sie auf dieselbe Weise den Satz von Pappus.

Wir sehen also, dass die algebraische Formulierung allgemeinerer geometrischer Sätze mit den Formeln der universellen Realisierung  $B_E$  einer konstruktiv erzeugten Basiskonfiguration  $\mathcal{K}$  startet, zusätzliche geometrische Voraussetzungen in einer Menge  $F = \{f_1, \dots, f_m\} \subset R$  von Polynomen kodiert, und die Behauptung ebenfalls in einem Polynom  $g \in R$  ausgedrückt werden kann.

Den geometrischen Satz zu beweisen bedeutet zu zeigen, dass für jede zulässige Variablenspezifikation  $X \rightarrow X_0$ , für die  $F(X_0) = 0$  gilt, auch  $g(X_0) = 0$  folgt:

$$\forall X_0 \in K^n \text{ (not } \mathcal{K}.\text{DGF}(X_0) \Rightarrow (F(X_0) = 0 \Rightarrow g(X_0) = 0)) \quad (\text{G.0})$$

Wir wollen wie oben  $\mathcal{K}.\text{DGF}$  auf eine Formel  $\text{iszero}(h)$  mit  $h \in R$  reduzieren, womit sich (G.0) in folgende zueinander äquivalente Aussagen umformen lässt:

$$\begin{aligned} & \forall X_0 \in K^n (h(X_0) \neq 0 \Rightarrow (F(X_0) = 0 \Rightarrow g(X_0) = 0)) \\ \Leftrightarrow & \forall X_0 \in K^n (h(X_0) = 0 \vee F(X_0) \neq 0 \vee g(X_0) = 0) \\ \Leftrightarrow & \forall X_0 \in K^n (F(X_0) = 0 \Rightarrow (h(X_0) = 0 \vee g(X_0) = 0)) \\ \Leftrightarrow & \forall X_0 \in K^n (F(X_0) = 0 \Rightarrow h(X_0) \cdot g(X_0) = 0) \end{aligned}$$

Zum Beweis des Satzes muss also untersucht werden, was man über das Verschwinden von Polynomen  $g, h \in R$  auf Nullstellenmengen

$$V(F) = \{X_0 \in K^n \mid f_1(X_0) = \dots = f_m(X_0) = 0\}$$

aussagen kann. Die letzte Bedingung ist dann äquivalent zu  $V(F) \subset V(g \cdot h)$ . Für derartige Sätze schreiben wir auch kurz  $(F \Rightarrow g)_h$ . Wenn die Degenerationsbedingung  $h$  ohne Bedeutung ist, dann schreiben wir auch kurz  $(F \Rightarrow g)$ .

In den oben betrachteten Fällen war  $m = 1$ . Derartige Sätze wollen wir als *Sätze vom Äquivalenztyp* bezeichnen.

**Definition 12** Als Satz vom Äquivalenztyp bezeichnen wir eine rationale Konstruktionsbeschreibung  $\mathcal{K}$  zusammen mit einer auf der Endkonfiguration von  $\mathcal{K}$  gegebenen Aussage  $(F \Rightarrow g)_h$ , in der  $F$  aus nur einem Element  $f \in R$  besteht.

Der Beweis eines solchen Satzes lässt sich auf die Aussage  $V(f) \subset V(g \cdot h)$  reduzieren, wobei  $h \in R$  die Zulässigkeit von speziellen Konfigurationen kodiert. Diese Bedingung lässt sich in Anwendungen oft im Zuge des Beweisverfahrens finden, das auf die Faktorisierung der Polynome  $f$  und  $g$  hinausläuft.

## 4.2 Geometrische Sätze vom linearen Typ

Betrachten wir noch einmal den **affinen Satz von Desargue** in seiner ursprünglichen Koordinateninterpretation

```

A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy);
A1:Point(dx,dy); B1:Point(ex,ey); C1:Point(fx,fy);

polys: [
  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));

```

Auch in diesem Fall lassen sich die drei Gleichungen *polys* verwenden, um die Variablen  $Y = (e_y, f_x, f_y)$  durch die restlichen Variablen  $U = (a_x, a_y, b_x, b_y, c_x, c_y, d_x, d_y, e_x)$  auszudrücken. Wir können die Variablen  $X$  der universellen Realisierung der Basiskonfiguration bzgl. der Gleichungen *polys* also aufteilen in  $X = (Y, U)$ , wobei  $U$  unabhängige Variablen sind und  $Y$  vermöge der Beziehungen *polys* durch  $U$  ausgedrückt werden können. Eine genauere Analyse zeigt, dass *polys* sogar linear in  $Y$  ist, so dass wir das entsprechende Gleichungssystem mit linearer Algebra über  $k(U)$  lösen können, wenn wir die Polynome in  $S = k(U)[Y]$  statt in  $R = K[U, Y]$  betrachten.

```

uvars: [ax, ay, bx, by, cx, cy, dx, dy, ex];
yvars: [ey, fx, fy];
sol:solve(pre, yvars);

```

$$e_y = \frac{a_x d_y - a_y d_x + a_y e_x - b_x d_y + b_y d_x - b_y e_x}{a_x - b_x}$$

$$f_x = \frac{a_x e_x - b_x d_x + c_x d_x - c_x e_x}{a_x - b_x}$$

$$f_y = \frac{a_x d_y - a_y d_x + a_y e_x - b_x d_y + c_y d_x - c_y e_x}{a_x - b_x}$$

Wir erhalten wie im letzten Abschnitt rationale Ausdrücke als universelle Formeln  $Y = Y(U)$  für die abhängigen Variablen. Setzen wir diese universellen Formeln für die abhängigen Variablen in die universelle Formel der Behauptung ein, so erhalten wir für `subs(con, Y → Y(U))` das Nullpolynom als Ausdruck in  $k(U)$

```
ratsimp(subst(sol, con));
```

0

Allerdings hatten wir dabei Glück mit der Auswahl der Parameter. Hätten wir

```

uvars: [ax, ay, bx, by, cx, cy, dx, ex, fx];
yvars: [dy, ey, fy];

```

gewählt, so wäre *polys* ebenfalls linear über  $k(U)$  gewesen, aber

```
sol:solve(pre, yvars);
```

liefert in diesem Fall [] – das lineare Gleichungssystem besitzt keine Lösungen. Um die Gründe genauer zu verstehen, wollen wir die entsprechende Koeffizientenmatrix extrahieren und analysieren.

```
KoeffMat(polys, vars) :=
  apply(matrix, map(lambda([p], map(lambda([u], coeff(p, u, 1))), vars)), polys));
```

In diesem Fall ergibt sich

```
KM:KoeffMat(pre, yvars);
```

$$KM = \begin{pmatrix} a_x - b_x & b_x - a_x & 0 \\ a_x - c_x & 0 & c_x - a_x \\ 0 & b_x - c_x & c_x - b_x \end{pmatrix}$$

und  $\det(KM) = 0$ . Im ersten Fall dagegen ergibt sich

$$KM = \begin{pmatrix} b_x - a_x & 0 & 0 \\ 0 & a_y - c_y & c_x - a_x \\ b_x - c_x & b_y - c_y & c_x - b_x \end{pmatrix}$$

Für die Determinante dieser Koeffizientenmatrix gilt

$$\begin{aligned} \det(KM) &= a_x^2 c_y - a_x^2 b_y - a_y b_x^2 + b_x^2 c_y + a_x a_y b_x - a_x a_y c_x + a_x b_x b_y \\ &\quad - 2 a_x b_x c_y + a_x b_y c_x + a_y b_x c_x - b_x b_y c_x \\ &= -(a_x - b_x) (a_x b_y - a_y b_x - a_x c_y + a_y c_x + b_x c_y - b_y c_x) \end{aligned}$$

Für jede Belegung  $U \rightarrow U_0$  mit  $\det(KM(U_0)) \neq 0$  hat  $\mathbf{subs}(polys, U \rightarrow U_0)$  als Gleichungssystem in  $Y$  eine eindeutig bestimmte Lösung  $Y_0$ , die wir mit der Cramerschen Regel bestimmen können, und jede Belegung  $X \rightarrow X_0$  mit  $polys(X_0) = 0$  lässt sich auf diese Weise als  $X_0 = (Y_0, U_0)$  darstellen. Wegen  $con(Y_0, U_0) = \mathbf{subs}(con(Y(U), U), U \rightarrow U_0)$  kann die geometrische Behauptung wieder auf das Verschwinden von  $con(Y(U), U) \in k(U)$  als rationalem Ausdruck zurückgeführt werden.

Wir haben den Satz also unter der Nichtdegenerationsbedingung  $d(U) = \det(KM) \in k[U]$  bewiesen. Diese lässt sich in unserem Beispiel geometrisch als

$$d(U) = (a_x - b_x) \cdot \mathbf{is\_collinear}(A, B, C)$$

darstellen. Der zweite Faktor ist eine Degenerationsbedingung für die Konfiguration, der erste Faktor enthält eine zusätzliche Restriktion. Wir haben den geometrischen Satz also nur für solche speziellen Realisierungen bewiesen, für die zusätzlich  $a_x \neq b_x$  gilt. Für spezielle Realisierungen mit  $a_x = b_x$  ist der geometrische Satz durch unsere Rechnung noch nicht bewiesen und zusätzliche Argumente erforderlich.

Untersuchen wir nun allgemein, was man über geometrische Sätze aussagen kann, in denen sich die Variablen in den universellen Formeln der Konfiguration  $\mathcal{K}$  so als  $X = (Y, U)$  aufteilen lassen, dass das System der Voraussetzungen  $F = \{f_1, \dots, f_m\}$  gerade ein quadratisches

lineares Gleichungssystem in  $Y$  über  $k(U)$  bildet. Solche geometrischen Sätze bezeichnen wir als *Sätze vom linearen Typ*.

Das entsprechende Gleichungssystem hat die Gestalt

$$f_i = \sum_{j=1}^m c_{ij}(U) y_j + c_{i0}(U) = 0, \quad i = 1, \dots, m$$

Seine Lösbarkeit hängt ganz wesentlich von der Determinante

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

ab. Genauer, das Gleichungssystem hat auch im allgemeinen Fall für jede Belegung  $U \rightarrow U_0$ , für die  $D(U_0) \neq 0$  gilt, eine eindeutige Lösung, die man nach der Cramerschen Regel aus den universellen Formeln  $y_i = y_i(U) = \frac{D_i(U)}{D(U)} \in k(U)$  durch die Substitution  $U \rightarrow U_0$  gewinnen kann, und jede Belegung  $X \rightarrow X_0$  mit  $F(X_0) = 0$  lässt sich auf diese Weise als  $X_0 = (Y_0, U_0)$  darstellen.

Diese universellen Formeln können wir wie für Sätze vom rationalen konstruktiven Typ in die Behauptung  $g(Y, U)$  einsetzen und die so entstehende rationale Funktion  $g'(U) = g(Y(U), U)$  vereinfachen. Wir erhalten ähnliche Alternativen:

1) Führt die Simplifikation von  $g'(U)$  auf das Nullpolynom, so gilt für alle Spezifikationen  $U \rightarrow U_0$  mit  $D(U_0) \neq 0$  auch  $g(Y_0, U_0) = 0$  und der Satz ist richtig.

2) Ergibt die Simplifikation eine nicht verschwindende rationale Funktion  $g'(U) = \frac{P(U)}{Q(U)}$ , so ist der Satz für fast alle Spezifikationen  $U \rightarrow U_0$  der unabhängigen Variablen falsch und höchstens unter der zusätzlichen Voraussetzung  $P(U) = 0$  richtig. Er gilt also dann nur, wenn zwischen den (als unabhängig angenommenen) Variablen  $U$  eine (weitere) Abhängigkeit besteht.

3) Ist schließlich bereits  $D(U)$  das Nullpolynom, so gibt es eine  $k(U)$ -lineare Kombination der Zeilen der Koeffizientenmatrix zum Nullvektor

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

Dan gilt aber

$$\sum_i \alpha_i(U) f_i = \sum_i \alpha_i(U) c_{i0}(U) = g(U) \in k[U].$$

Ist auch  $g(U)$  das Nullpolynom, so besteht eine lineare Abhängigkeit zwischen den Polynomen  $f_1, \dots, f_m$  als Elemente aus  $S$ . Ist  $g(U) \neq 0$ , so folgt aus den Polynomen  $f_1, \dots, f_m$  eine algebraische Abhängigkeitsrelation zwischen den als „unabhängig“ deklarierten Variablen  $U$ . In beiden Fällen sind die Voraussetzungen nicht eingehalten.

Wir haben damit den folgenden Satz bewiesen:

**Satz 20 (Über das mechanisierte Beweisen geometrischer Sätze vom linearen Typ)**

Sei  $\mathcal{K}$  die polynomiale Konfiguration eines geometrischen Satzes mit universellen Formeln in den Variablen  $X = (Y, U)$ ,  $F \subset k[Y, U]$  eine Menge von  $m$  Polynomen, welche die geometrischen Voraussetzungen des Satzes kodieren und  $g(Y, U)$  die algebraische Kodierung der geometrischen Behauptung.

Ist  $(F \Rightarrow g)$  ein Satz vom linearen Typ bzgl. der Variablenaufteilung  $X = (Y, U)$  und  $D(U)$  nicht das Nullpolynom, so gilt der Satz unter der Nichtdegenerationsbedingung  $D(U)$  genau dann, wenn  $g'(U)$  als rationale Funktion in  $k(U)$ , die durch Substitution der eindeutig bestimmten Lösung  $Y = Y(U) \in k(U)^m$  von  $F$  in  $g(Y, U)$  entsteht, identisch null ist.

Wie im Beispiel des Satzes von Desargue kann es sein, dass die Bedingung  $D(U) \neq 0$  eine Reihe geometrisch relevanter Fälle ausschließt. Wird die Ausnahmemenge durch eine algebraische Bedingung beschrieben, die nicht geometrisch invariant ist, so hilft oft die Wahl eines speziellen Koordinatensystems weiter. So kann in unserem Fall  $a_x - b_x = 0$  vermieden werden, indem das (bisher nicht weiter spezifizierte) Koordinatensystem so gewählt wird, dass  $A(0, 0)$ ,  $B(0, 1)$  gilt. Eine geometrisch invariante Bedingung wie `is_collinear(A, B, C)` ist dagegen in jedem Koordinatensystem auf gleiche Weise gültig oder nicht.

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

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

### 4.3 Sätze am Kreis

Mit dem in diesem Kapitel entwickelten Ansatz können wir auch nichtlineare geometrische Bedingungen erfassen. Diese Bedingungen treten auf, wenn 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.

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

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

$$(x - m_x)^2 + (y - m_y)^2 - r^2 = (x^2 + y^2) + c_2 x + c_3 y + c_4$$

mit  $c_2 = -2m_x$ ,  $c_3 = -2m_y$ ,  $c_4 = m_x^2 + m_y^2 - r^2$  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  $(m_x, m_y)$  und der Sqradius  $r^2$  unmittelbar polynomial zurückgewinnen.

Im GEOPROVER-Paket stehen dafür die Funktionen

```
die Funktionen circle_center(c:Circle):Point
und circle_sqradius(c:Circle):Scalar,
der Konstruktor pc_circle(M:Point, P:Point):Circle
und die boolesche Funktion on_circle(P:Point, c:Circle)
```

zur Verfügung. `pc_circle` bestimmt die Koordinaten eines Kreises mit Zentrum  $M$  und einem gegebenen Punkt  $A$  auf der Peripherie und `on_circle` gibt für einen Punkt  $P(p_x, p_y)$  den Wert des Ausdrucks  $(p_x^2 + p_y^2) + c_2 p_x + c_3 p_y + c_4$  zurück.

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

```
A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy); M:Point(mx,my);
polys:[ sqrdist(M,A)-sqrdist(M,B), sqrdist(M,A)-sqrdist(M,C) ];
```

bestimmen, wobei für  $A, B, C$  universelle Realisierungen eingesetzt sind. Das entsprechende Gleichungssystem

$$\begin{aligned} 2m_x(-a_x + b_x) + 2m_y(-a_y + b_y) + a_x^2 + a_y^2 - b_x^2 - b_y^2 &= 0, \\ 2m_x(-a_x + c_x) + 2m_y(-a_y + c_y) + a_x^2 + a_y^2 - c_x^2 - c_y^2 &= 0 \end{aligned}$$

ist linear in  $(m_x, m_y)$  und besitzt eine eindeutige Lösung genau dann, wenn die zugehörige Determinante, die wieder einmal genau mit der Kollinearitätsbedingung `is_collinear(A,B,C)` übereinstimmt, verschieden von null ist.

```
solve(polys, [mx,my]);
```

ergibt dann die universellen Formeln für die Mittelpunktskoordinaten, die wir 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

$$\begin{aligned} M &= \left( \frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \frac{\Delta_3}{\Delta} \right) \text{ mit} \\ \Delta_1 &= -a_x^2 b_y + a_x^2 c_y - a_y^2 b_x + a_y^2 c_x + a_y b_x^2 + a_y b_y^2 - a_y c_x^2 - a_y c_y^2 - b_x^2 c_y \\ &\quad - b_y^2 c_x + b_y c_x^2 + b_y c_y^2 \\ \Delta_2 &= a_x^2 b_x - a_x^2 c_x - a_x b_x^2 - a_x b_y^2 + a_x c_x^2 + a_x c_y^2 + a_y^2 b_x - a_y^2 c_x + b_x^2 c_x \\ &\quad - b_x c_x^2 - b_x c_y^2 + b_y^2 c_x \\ \Delta_3 &= -a_x^2 b_x c_y + a_x^2 b_y c_x + a_x b_x^2 c_y + a_x b_y^2 c_x - a_x b_y c_x^2 - a_x b_y c_y^2 - a_y^2 b_x c_y \\ &\quad + a_y^2 b_y c_x - a_y b_x^2 c_x + a_y b_x c_x^2 + a_y b_x c_y^2 - a_y b_y^2 c_x \\ \Delta &= a_x b_y - a_x c_y - a_y b_x + a_y c_x + b_x c_y - b_y c_x \end{aligned}$$

Die Parameter dieses Kreises sind rationale Funktionen der Parameter der Punkte  $A, B, C$ , wobei im Nenner das Polynom  $\Delta = \text{is\_collinear}(A, B, C)$  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. Als homogene Koordinaten für einen Kreis  $c = \text{p3\_circle}(A, B, C)$  ergibt sich damit  $c = (\Delta : \Delta_1 : \Delta_2 : \Delta_3)$  mit den oben berechneten Polynomen.

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

```
D:Point(dx,dy);
on_circle(P,p3_circle(A,B,C));
```

als Polynom vierten Grades, das die entsprechende Bedingung beschreibt.

**Satz 21** Vier Punkte  $A = (a_x, a_y), B = (b_x, b_y), C = (c_x, c_y), D = (d_x, d_y)$  liegen genau dann auf einem gemeinsamen Kreis, wenn das Polynom

$$\begin{aligned}
 p_1 = & a_x^2 b_x c_y - a_x^2 b_x d_y - a_x^2 b_y c_x + a_x^2 b_y d_x + a_x^2 c_x d_y - a_x^2 c_y d_x - a_x b_x^2 c_y + \\
 & a_x b_x^2 d_y - a_x b_y^2 c_y + a_x b_y^2 d_y + a_x b_y c_x^2 + a_x b_y c_y^2 - a_x b_y d_x^2 - a_x b_y d_y^2 - a_x c_x^2 d_y - \\
 & a_x c_y^2 d_y + a_x c_y d_x^2 + a_x c_y d_y^2 + a_y^2 b_x c_y - a_y^2 b_x d_y - a_y^2 b_y c_x + a_y^2 b_y d_x + a_y^2 c_x d_y - \\
 & a_y^2 c_y d_x + a_y b_x^2 c_x - a_y b_x^2 d_x - a_y b_x c_x^2 - a_y b_x c_y^2 + a_y b_x d_x^2 + a_y b_x d_y^2 + a_y b_y^2 c_x - \\
 & a_y b_y^2 d_x + a_y c_x^2 d_x - a_y c_x d_x^2 - a_y c_x d_y^2 + a_y c_y^2 d_x - b_x^2 c_x d_y + b_x^2 c_y d_x + b_x c_x^2 d_y + \\
 & b_x c_y^2 d_y - b_x c_y d_x^2 - b_x c_y d_y^2 - b_y^2 c_x d_y + b_y^2 c_y d_x - b_y c_x^2 d_x + b_y c_x d_x^2 + b_y c_x d_y^2 - \\
 & b_y c_y^2 d_x
 \end{aligned}$$

verschwindet.

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

$$\begin{aligned}
 x_1 (a_x^2 + a_y^2) + x_2 a_x + x_3 a_y + x_4 &= 0 \\
 x_1 (b_x^2 + b_y^2) + x_2 b_x + x_3 b_y + x_4 &= 0 \\
 x_1 (c_x^2 + c_y^2) + x_2 c_x + x_3 c_y + x_4 &= 0 \\
 x_1 (d_x^2 + d_y^2) + x_2 d_x + x_3 d_y + x_4 &= 0
 \end{aligned}$$

gibt, d. h. wenn

$$\begin{vmatrix}
 (a_x^2 + a_y^2) & a_x & a_y & 1 \\
 (b_x^2 + b_y^2) & b_x & b_y & 1 \\
 (c_x^2 + c_y^2) & c_x & c_y & 1 \\
 (d_x^2 + d_y^2) & d_x & d_y & 1
 \end{vmatrix} = 0$$

gilt. Wir können diese universelle Formel als Prozedur

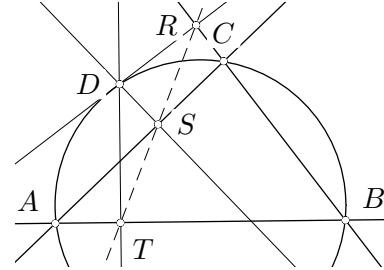
$$\text{is\_concylic}(A, B, C, D) = \text{on\_circle}(P, \text{p3\_circle}(A, B, C))$$

zur weiteren Verwendung in unsere Programmsammlung aufnehmen.

Damit können wir bereits den **Satz von der Simonschen Geraden** als Satz vom Äquivalenztyp beweisen. Seien dazu  $A, B, C, D$  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 Simonschen Geraden behauptet, dass für einen Punkt  $D$  auf dem Umkreis von  $ABC$  die Lotfußpunkte  $R, S, T$  kollinear sind und umgekehrt.



Wir gehen wieder von den universellen Formeln einer Konfiguration  $\mathcal{K}$  mit freien Punkten  $A, B, C, D$  und daraus konstruierten Lotfußpunkten  $R, S, T$  aus.

Die universelle Formel der geometrischen Eigenschaft, dass  $D$  auf dem Umkreis von  $\triangle ABC$  liegt, ist genau das oben hergeleitete Polynom  $p_1$  vierten Grades. Die universelle Formel der Behauptung

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

ist eine rationale Funktion, deren Zähler ein Polynom  $p_2$  8. Grades mit 576 Summanden ist und dessen Nenner  $p_3$  aus dem Produkt

$$p_3 = \text{sqrdist}(A, B) \cdot \text{sqrdist}(B, C) \cdot \text{sqrdist}(A, C)$$

besteht.

Über dem Grundbereich  $k = \mathbb{R}$  gilt  $\text{sqrdist}(B, C) = 0 \Leftrightarrow B = C$ , so dass  $\text{sqrdist}(B, C) = 0$  eine geometrisch degenerierte Lage kodiert. Das ist über dem Grundbereich  $k = \mathbb{C}$  nicht mehr so, so dass eine Geometrie über diesem Grundbereich andere Eigenschaften hat als die uns vertrauten. Suchen wir nach der Quelle der Nenner  $\text{sqrdist}(B, C) = b_x^2 - 2b_x c_x + b_y^2 - 2b_y c_y + c_x^2 + c_y^2$ , so werden wir schnell im Konstruktionswerkzeug fündig, mit dem  $R$  konstruiert wurde. In der Tat lässt sich ein Lot nur dann fällen, wenn  $B \neq C$  ist, da sonst die Richtung der Basisgeraden unbestimmt ist. Zur Bestimmung des Lotfußpunkts muss aber weiterhin das Lot die Basisgerade schneiden – ein in der reellen Geometrie offensichtlicher Tatbestand, der in der komplexen Geometrie nicht mehr erfüllt ist: Die Berechnungen zeigen, dass das Lot auf  $BC$  genau dann parallel zu dieser Basisgeraden ist, wenn  $\text{sqrdist}(B, C) = 0$  gilt. Das ergibt sich auch aus unseren früheren Betrachtungen zu homogenen Koordinaten und der Einbettung der affinen in die projektive Ebene – die Lotgerade  $h$  durch einen Punkt  $P$  zu einer gegebenen Geraden  $g = (g_1 : g_2 : g_3)$  kann als Gerade durch  $P$  und den Orthogonalpunkt  $O_g = (g_1 : g_2 : 0)$  der Geraden  $g$  bestimmt werden. Diese ist genau dann parallel zu  $g$ , wenn  $h$  durch den Fernpunkt  $F_g = (-g_2 : g_1 : 0)$  der Geraden  $g$  geht, wenn also  $O_g = F_g$  gilt. Das ist aber genau dann der Fall, wenn die entsprechenden homogenen Koordinaten proportional sind, wenn also  $g_1^2 + g_2^2 = 0$  gilt.  $g$  ist also eine Gerade, die auf sich selbst senkrecht steht. Solche (notwendig komplexen) Geraden bezeichnet man auch als *isotrope* Geraden. Im Affinen gibt es

keine Lotfußpunkte auf solche Geraden. Die korrekte Degenerationsbedingung des Werkzeugs `pedalpoint(D,g)` ist also in der Tat  $g_1^2 + g_2^2 = 0$ .

$p_2$  lässt sich seinerseits in das Produkt

$$p_2 = \text{is\_collinear}(A, B, C)^2 \cdot p_1$$

zerlegen, woraus der geforderte Beweis nunmehr leicht abzuleiten ist, dass  $D$  genau dann auf dem Umkreis des Dreiecks  $ABC$  liegt, wenn die drei Fußpunkte  $R, S, T$  auf einer Geraden liegen.

#### 4.4 Schnittpunkte von zwei Kreisen. Die Potenzgerade

Im Allgemeinen lassen sich 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  $P = (p_x, p_y)$  das Gleichungssystem

```
P:Point(px,py);
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)))
];
```

$$[-r_1^2 + u^2 + 2u p_x + p_x^2 + p_y^2, -r_2^2 + u^2 - 2u p_x + p_x^2 + p_y^2],$$

zu dessen Lösung

```
sol:solve(polys,[px,py]);
```

$$p_x = \frac{r_1^2 - r_2^2}{4u}, \quad p_y = \pm \frac{\sqrt{(r_1 + r_2 - 2u)(r_1 + r_2 + 2u)(r_1 - r_2 + 2u)(r_2 - r_1 + 2u)}}{4u}$$

bereits Wurzelausdrücke erforderlich sind.

Das ist auch verständlich, denn es kann vorkommen, dass zwei Kreise in der reellen Geometrie keine gemeinsamen Punkte haben – dies ist genau dann der Fall, wenn der Radikand des Wurzelausdrucks negativ wird. Wir betrachten dazu in GEOGEBRA zwei Kreise, bestimmen deren Schnittpunkte und konstruieren deren Verbindungsgerade. Verändern wir die Größe eines Kreises, so dass keine (reellen) Schnittpunkte mehr existieren, verschwindet auch die Gerade. Verändern wir die Größe weiter, so dass wieder (reelle) Schnittpunkte existieren, so werden diese und alle davon abhängenden Objekte auch wieder angezeigt. Im Algebrafenster sehen wir, dass die entsprechenden geometrischen Stücke in der Zwischenzeit als *undefiniert* behandelt werden.

Interessanterweise ist allerdings  $p_x$  ein *rationaler* Ausdruck, so dass auch im Fall eines negativen Radikanden die imaginären Schnittpunkte  $P_1$  und  $P_2$  dieselbe *reelle*  $x$ -Koordinate haben. Mehr noch, verbindet man diese beiden imaginären Punkte, so ergibt sich (in den hier gewählten Koordinaten) eine *reelle* Gerade senkrecht zur  $x$ -Achse. Diese Gerade wird auch als Potenzgerade der beiden Kreise bezeichnet.

**Definition 13** Sind  $k_1$  und  $k_2$  zwei Kreise mit den Mittelpunkten  $M_1$  und  $M_2$  und den Radien  $r_1$  und  $r_2$ , so bezeichnet man die Menge derjenigen Punkte  $P$ , für die

$$\text{sqrdist}(M_1, P) - \text{sqrdist}(M_2, P) = r_1^2 - r_2^2$$

oder gleichbedeutend

$$\text{sqrdist}(M_1, P) - r_1^2 = \text{sqrdist}(M_2, P) - r_2^2$$

gilt, als die Potenzgerade  $\text{radical\_axis}(k_1, k_2)$  dieser beiden Kreise.

Die Größe  $p(P, k_1) = \text{sqrdist}(M_1, P) - r_1^2$  bezeichnet man auch als die Potenz des Punkts  $P$  bzgl. des Kreises  $k_1$ .

Die Potenzgerade zweier Kreise ist also der geometrische Ort aller Punkte, die bzgl. der beiden Kreise dieselbe Potenz haben. Nach Satz des Pythagoras ist für einen Punkt  $P$  außerhalb des Kreises  $k$  die Potenz  $p(P, k)$  gerade gleich dem Quadrat der Länge eines der Tangentenabschnitte von  $P$  an den Kreis  $k$ .

Ist  $P$  ein solcher Punkt und  $X$  der Lotfußpunkt von  $P$  auf die Zentrale  $M_1M_2$ , so liegt nach dem Satz des Pythagoras auch  $X$  auf der Potenzgeraden. Die Potenzgerade ist also genau die Senkrechte auf der Zentralen im Punkt  $X$ . Schneiden sich die beiden Kreise, so liegen die Schnittpunkte ebenfalls auf der Potenzgeraden. Da sich die Koordinaten von  $X$  rational durch die Kreisparameter ausdrücken lassen, ist auch  $\text{radical\_axis}(k_1, k_2)$  ein rationales Konstruktionswerkzeug, obwohl sich die Koordinaten der Schnittpunkte der beiden Kreise selbst nicht rational durch die Kreisparameter ausdrücken lassen. Die Geradengleichung der Potenzgeraden der beiden Kreise ergibt sich sofort aus den Kreisgleichungen

$$\begin{aligned} k_1 : (x^2 + y^2) + a_1 x + a_2 y + a_3 &= 0 \\ k_2 : (x^2 + y^2) + b_1 x + b_2 y + b_3 &= 0 \end{aligned}$$

als deren Differenz  $(a_1 - b_1)x + (a_2 - b_2)y + a_3 - b_3 = 0$ .

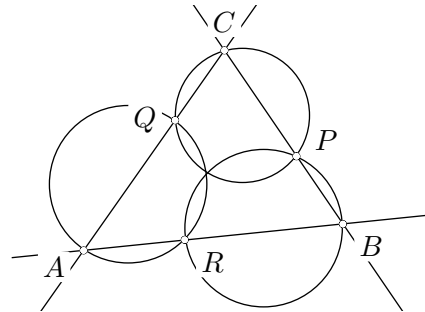
*Aufgabe:*

- a) Geben Sie eine Konstruktionsbeschreibung für die Potenzgerade, wenn sich die beiden Kreise nicht schneiden.
- b) Zeigen Sie, dass die drei Potenzgeraden, die sich zu drei gegebenen Kreisen paarweise konstruieren lassen, konkurrent sind.

#### 4.5 Kreise mit vorgegebenem Schnittpunkt

Betrachten wir dazu den **Satz vom Miquel-schen Punkt**. Die polynomiale geometrische Grundkonfiguration (Dreieck  $ABC$  mit Punkten  $R, S, T$  auf den Dreiecksseiten) lässt sich wie folgt beschreiben:

```
A:Point(ax,ay);
B:Point(bx,by);
C:Point(cx,cy);
P:varpoint(B,C,u1);
Q:varpoint(A,C,u2);
R:varpoint(A,B,u3);
```



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

```
A:Point(0,0); B:Point(0,by); C:Point(cx,cy);
P:varpoint(B,C,u1); Q:varpoint(A,C,u2); R:varpoint(A,B,u3);
c1:p3_circle(R,Q,A); c2:p3_circle(R,P,B); c3:p3_circle(P,Q,C);
```

```
S:Point(sx,sy);
polys:[ on_circle(S,c1), on_circle(S,c2) ];
sol:solve(polys,[sx,sy]);
```

Wir erhalten (nach längerer Rechnung mit Maxima) zwei Lösungen, deren Koordinaten jeweils rationale Funktionen in den Parametern sind. Eine der beiden Lösungen ist der Punkt  $R$ , von dem wir bereits wissen, dass er zu beiden Kreisen gehört. Der zweite Punkt ist der gesuchte Schnittpunkt  $S$  (seine rationalen Koordinaten haben 10 bzw. 26 Zählerterme und 42 Nennerterme). Er liegt auf dem dritten Kreis, wie die folgende Rechnung zeigt.

```
S0:ev(S,sol[2]);
on_circle(S0,c3);
```

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

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

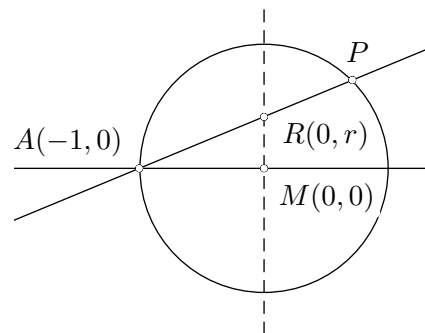
- `other_cl_point(P:Point,c:Circle,l:Line):Point`, die zu gegebenem Kreis  $c$  und gegebener Gerade  $l$  den zweiten Schnittpunkt von  $c$  und  $l$  findet, wenn  $P$  der andere Schnittpunkt ist.
- `other_cc_point(P:Point,c1:Circle,c2:Circle):Point`, die zu zwei gegebenen Kreisen  $c_1$  und  $c_2$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

## 4.6 Kreisgleiter

Sätze über Punkte auf einer Kreislinie  $k(M, r)$  lassen sich oft auch als Sätze vom konstruktiven Typ anschreiben, wenn eine rationale Parametrisierung der Punkte auf dem Kreis als Kreisgleiter verwendet wird. 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 wie im Beispiel der Funktion `other_cl_point` ausdrücken.

Für den Einheitskreis  $x^2 + y^2 = 1$  und den Referenzpunkt  $(-1, 0)$  ergibt sich für die Koordinaten eines Kreisgleiters folgende Formel

```
M:Point(0,0); A:Point(-1,0);
kreis:pc_circle(M,A);
R:Point(0,r);
line:pp_line(A,R);
P:other_cl_point(A,kreis,line);
```



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

Jeder Punkt  $P \neq A$  auf dem Kreis kann auf diese Weise eindeutig dargestellt werden. Dass eine *vollständige* Parametrisierung der Kreislinie durch Werte  $r \in \mathbb{R}$  nicht möglich ist, hat topologische Gründe – die Kreislinie ist kompakt, die Gerade  $\mathbb{R}$  dagegen nicht. Auch in diesem Fall kann der fehlende Punkt  $A$  mit  $r = \infty$  in Verbindung gebracht werden, so dass eine vollständige Parametrisierung der Kreislinie durch Punkte  $(r : s) \in \mathbb{P}^1$  auf der (kompakten) projektiven Geraden möglich ist – in homogenen Koordinaten  $P = (r^2 - s^2 : 2rs : r^2 + s^2)$ .

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

Damit können wir einen konstruktiven Beweis des **Satzes von der Simonschen Geraden** geben:

```
M:Point(0,0);
A:Point(0,1);
B:circle_slider(M,A,u1);
C:circle_slider(M,A,u2);
D:circle_slider(M,A,u3);
R:pedalpoint(D,pp_line(B,C));
S:pedalpoint(D,pp_line(A,C));
T:pedalpoint(D,pp_line(A,B));
is_collinear(R,S,T);
```

#### 4.7 Winkelhalbierende

Die Größe des Winkels  $\alpha$  zwischen zwei Geraden  $g = BA$  und  $h = BC$  kann über die Anstiege  $\tan(\alpha_g) = -g_1/g_2$  bzw.  $\tan(\alpha_h) = -h_1/h_2$  der Geraden  $g = (g_1, g_2, g_3)$  und  $h = (h_1, h_2, h_3)$  bestimmt werden. Es 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. Dies ist – ähnlich der Darstellung  $y = mx + n$  einer Geraden, die Parallelen zur  $y$ -Achse ausschließt – keine geometrische, sondern eine rein algebraische Degenerationsbedingung, die aus dem Defizit des Modells herrührt. 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. Im GEOPROVER-Paket kann eine solche Winkelgröße über das Konstruktionswerkzeug `l2_angle(g, h)` bestimmt werden.

Für die Geraden  $g = (g_1, g_2, g_3)$  und  $h = (h_1, h_2, h_3)$  ergeben sich die Geradenkoordinaten einer Winkelhalbierenden  $w = (w_1, w_2, w_3)$  als Nullstellen von

```
sys:l2_angle(g,w)-l2_angle(w,h);
```

$$\frac{g_1 w_2 - g_2 w_1}{g_1 w_1 + g_2 w_2} + \frac{h_1 w_2 - h_2 w_1}{h_1 w_1 + h_2 w_2}$$

In den Formeln spielt die dritte Komponente der Geradenkoordinaten keine Rolle, da es beim Winkelvergleich nur um Geradenanstiege geht.  $w_3$  kann aus der Bedingung `is_concurrent(g, h, w)` als

$$w_3 = \frac{g_1 h_3 w_2 - g_2 h_3 w_1 - g_3 h_1 w_2 + g_3 h_2 w_1}{g_1 h_2 - g_2 h_1}$$

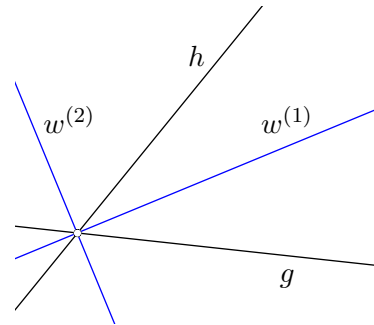
bestimmt werden. Im Weiteren wollen wir annehmen, dass die Geraden durch den Ursprung gehen, also  $g_3 = h_3 = w_3 = 0$  gilt. Für  $(w_1, w_2)$  ergibt sich

```
eqn:expand(num(ratsimp(sys)));
solve(eqn,w2);
```

$$w_2 = w_1 \frac{(g_1 h_2 + g_2 h_1) (w_2^2 - w_1^2) + 2 (g_1 h_1 - g_2 h_2) w_1 w_2 - (g_1 h_1 - g_2 h_2) \pm \sqrt{(g_1^2 + g_2^2) (h_1^2 + h_2^2)}}{g_1 h_2 + g_2 h_1}.$$

was auf die homogenen Geradenkoordinaten

$$\left( (g_1 h_2 + g_2 h_1), -(g_1 h_1 - g_2 h_2) \pm \sqrt{(g_1^2 + g_2^2) (h_1^2 + h_2^2)}, 0 \right)$$



führt. 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 Ausgangsgeraden ausdrücken lassen. Auch lassen sich die beiden Winkelhalbierenden algebraisch nicht voneinander unterscheiden.

*Aufgabe:* Weisen Sie nach, dass die beiden durch die Formeln gegebenen Geraden  $w^{(1)}$  und  $w^{(2)}$  senkrecht aufeinander stehen.

Als polynomiale geometrische Eigenschaft lässt sich jedoch die Bedingung ausdrücken, dass ein Punkt  $P(p_x, p_y)$  auf einer der beiden Winkelhalbierenden liegt, wie sich aus folgender Rechnung ergibt:

```
A:Point(ax,ay); B:Point(bx,by); C:Point(cx,cy); P:Point(px,py);
sys:ratsimp(p3_angle(A,B,P) - p3_angle(P,B,C));
pol2:expand(num(sys));
```

Hierbei steht `p3_angle(A,B,C)` für den Tangens des Winkels  $\angle(BA, BC)$ , der mit dem Konstruktionswerkzeug `l1_angle(pp_line(B,A), pp_line(B,C))` bestimmt werden kann.

Von dieser Formel ist wiederum nur das Verschwinden des Zählers von Interesse; dieses Polynom vom Grad 2 in  $p_x, p_y$ , das für allgemeine Punkte  $A, B, C$  56 Terme enthält, ist die universelle Formel der polynomialen geometrischen Eigenschaft, dass  $P$  auf einer der Winkelhalbierenden von  $\angle ABC$  liegt. Im GEOPROVER-Paket ist diese universelle Formel als `on_bisector(P, A, B, C)` implementiert.

*Aufgabe:* 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.)

Zum Beweis des **Satzes vom Schnittpunkt der Winkelhalbierenden** bestimmen wir Bedingungen für die Koordinaten des Schnittpunkts  $P = (p_x, p_y)$  von zwei der drei Winkelhalbierenden und prüfen, ob  $P$  auch auf der dritten Winkelhalbierenden liegt.

Zur Vereinfachung der Rechnung gehen wir von der Konfiguration

```
Start(Point A(0,0), B(1,0), C, P)
```

aus, in der wir das Koordinatensystem so gelegt haben, dass  $B$  im Ursprung liegt und  $A$  im Einheitspunkt der  $x$ -Achse. Wir starten wieder mit einer universellen Realisierung dieser Konfiguration

```
A:Point(0,0); B:Point(1,0); C:Point(cx,cy); P:Point(px,py);
```

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

```
polys:[ on_bisector(P,B,A,C), on_bisector(P,C,B,A) ];
```

$$\left[ c_y p_x^2 - 2 c_x p_x p_y - c_y p_y^2, c_y - 2 p_y + c_y p_x^2 - c_y p_y^2 + 2 c_x p_y - 2 c_y p_x + 2 p_x p_y - 2 c_x p_x p_y \right]$$



aus zwei Polynomen zweiten Grades in  $(p_x, p_y)$ , dessen vier zu erwartende gemeinsame Nullstellen  $(p_x^{(i)}, p_y^{(i)})$ , jeweils in die universelle Formel

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

$$2c_x^3 p_y - 2c_x^2 c_y p_x + c_x^2 c_y - 2c_x^2 p_x p_y - 2c_x^2 p_y + 2c_x c_y^2 p_y + 2c_x c_y p_x^2 - 2c_x c_y p_y^2 + 2c_x p_x p_y - 2c_y^3 p_x + c_y^3 + 2c_y^2 p_x p_y - 2c_y^2 p_y - c_y p_x^2 + c_y p_y^2$$

eingesetzt, diese zu null vereinfachen sollten. Wir können die Variablen der universellen Realisierung also wieder in unabhängige  $U = (c_x, c_y)$  und abhängige  $Y = (p_x, p_y)$  unterteilen, Formeln  $Y = Y(U)$  bestimmen und versuchen, den – diesmal allerdings nicht rationalen – Ausdruck  $\text{subs}(con, Y = Y(U))$  weiter zu vereinfachen. Die Beweiskraft dieses Herangehens ist im nächsten Kapitel gesondert zu diskutieren. Im Folgenden wollen wir zunächst deren prinzipielle Zweckmäßigkeit untersuchen.

```
yvars:[px,py];
sol:solve(polys,yvars);
```

Das kann Maxima von sich aus nicht. Eine genauere Inspektion der Gleichungen *polys* zeigt, dass beide den quadratischen Term  $p_x^2 - p_y^2$  enthalten, die Differenz der beiden Gleichungen als zu einer linearen Gleichung in  $p_x, p_y$  führt. Diese kann nach  $p_y$  aufgelöst werden:

```
step1:solve(polys[1]-polys[2],[py]);
```

$$\left\{ \left[ p_y = -\frac{c_y - 2c_y p_x}{2c_x + 2p_x - 2} \right] \right\}$$

Setzen wir dies in das erste Polynom ein, so erhalten wir einen rationalen Ausdruck und mit dessen Zähler eine polynomiale Bestimmungsgleichung 4. Grades für  $p_x$ :

```
step2:ratsimp(ev(polys[1],step1[1]));
step3:factor(num(step2));
step4:solve(step3,[px]);
```

$$c_y (4p_x^4 - 8p_x^3 + 4(c_x - c_x^2 - c_y^2 + 1)p_x^2 + 4(c_x^2 - c_x + c_y^2)p_x - c_y^2)$$

Zu dieser Gleichung findet Maxima Lösungen in geschachtelten Wurzelausdrücken

```
step4:solve(step3,[px]);
```

$$\left\{ \left[ p_x = \frac{1}{2} - \frac{1}{2} \sqrt{2c_x^2 - 2c_x + 2c_y^2 - 2\sqrt{c_x^4 - 2c_x^3 + 2c_x^2 c_y^2 + c_x^2 - 2c_x c_y^2 + c_y^4 + c_y^2 + 1}} \right], \dots \right\},$$

die nun in die Behauptung *con* eingesetzt werden können:

```
step5:ev(ev(con,step1[1]),step4[1]);
ratsimp(step5);
```

0

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

Es ergibt sich die Frage, ob auch ohne explizite Bestimmung der Nullstellen des Polynomsystems  $polys$  als Wurzelausdrücke nachgewiesen werden kann, dass jede der vier Lösungen auch Nullstelle der Gleichung  $con$  ist. Derartige Fragen wollen wir im nächsten Kapitel genauer formulieren und studieren.

*Anmerkung:* Winkelhalbierende kann man auch mit einem anderen Konzept als geometrische Objekte vom konstruktiven Typ einführen kann. Wir verwenden dazu die Eigenschaft, dass  $P$  genau dann auf der Winkelhalbierenden des Winkels  $\angle ABC$  liegt, wenn die Gerade  $BC$  mit der Geraden  $\text{sym\_line}(c, l)$  zusammenfällt, die sich bei der Spiegelung von  $c = \text{pp\_line}(B, A)$  an  $l = \text{pp\_line}(B, P)$  ergibt.  $\text{sym\_line}(c, l)$  ist aber ein polynomiales Konstruktionswerkzeug. Wir gehen dazu von der Konfiguration

```
Start(Point A(-1,0), B(1,0), P)
```

aus. Ist  $P$  der Schnittpunkt der Winkelhalbierenden durch  $A$  und  $B$ , so ist  $C$  der Schnittpunkt zweier gespiegelter Geraden.

```
a:sym_line(pp_line(A,B),pp_line(A,P));
b:sym_line(pp_line(A,B),pp_line(B,P));
C:intersection_point(a,b);
```

Nun können wir prüfen, ob  $P$  auf einer der Winkelhalbierenden durch  $C$  liegt:

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

0

*Aufgabe:* Leiten Sie die universellen Formeln des Konstruktionswerkzeugs `sym_line` her.

#### 4.8 Tangenten an Kreise (Optional)

Die Tangente an einen Kreis  $c$  in einem Punkt  $P \in c$  ergibt sich als Senkrechte im Punkt  $P$  auf dem Berührungsradius. Die homogenen Koordinaten dieser Geraden können also wie folgt berechnet werden:

```
P:Point(px,py);
c:[c1,c2,c3,c4];
l:ortho_line(P,pp_line(circle_center(c),P));
```

$$\left[ c_2 + 2p_x c_1, 2p_y c_1 + c_3, -2p_y^2 c_1 - p_y c_3 - p_x c_2 - 2p_x^2 c_1 \right]$$

Wegen

$$\text{on\_circle}(P, c) = c_1 (p_x^2 + p_y^2) + c_2 p_x + c_3 p_y + c_4 = 0$$

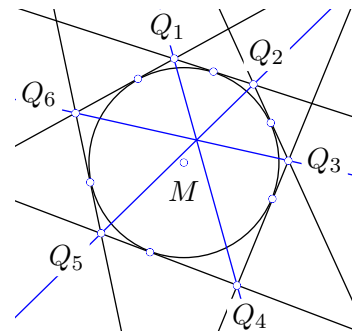
kann die dritte Komponente weiter vereinfacht werden zu  $p_y c_3 + p_x c_2 + 2c_4$ . Diese Formel können wir in eine neue Prozedur gießen.

```
tangent_line:=
  if on_circle(P,c)!=0 then error("Point not on the circle")
  else Line(2*part(c,1)*part(P,1)+part(c,2),
    2*part(c,1)*part(P,2)+part(c,3),
    part(c,2)*part(P,1)+part(c,3)*part(P,2)+2*part(c,4)
  );
```

Eine ähnliche Formel gilt auch für Tangenten an allgemeine Quadriken.

Mit dieser Funktion lässt sich ein konstruktives Beweisschema für den *Satz von Brianchon* für einen Kreis erstellen:

*Sind  $t_1, \dots, t_6$  sechs Tangenten an einen Kreis  $c$  und  $Q_i = t_i \cap t_{i+1}$  jeweils die Schnittpunkte „benachbarter“ Tangenten (mit  $t_7 = t_1$ ), so gehen die drei Geraden  $Q_1Q_4$ ,  $Q_2Q_5$  und  $Q_3Q_6$  durch einen gemeinsamen Punkt.*



```
M:Point(0,0); P[1]:Point(1,0); c:pc_circle(M,P[1]);
makelist(P[i]:circle_slider(M,P[1],u[i]),i,2,6);
makelist(t[i]:tangent_line(P[i],c),i,1,6); t[7]:t[1];
makelist(Q[i]:intersection_point(t[i],t[i+1]),i,1,6);
makelist(l[i]:pp_line(Q[i],Q[i+3]),i,1,3);
is_concurrent(l[1],l[2],l[3]);
```

Untersuchen wir nun, wann eine Gerade  $l$  einen Kreis  $c$  berührt. Wir stellen zunächst die Gleichungen auf, aus denen sich die Koordinaten der Schnittpunkte  $P = (p_x, p_y) \in l \cap c$  bestimmen lassen

```
P:Point(px,py); c:Circle(c1,c2,c3,c4); l:Line(l1,l2,l3);
polys:[ on_circle(P,c), on_line(P,l) ];
```

$$\text{polys} = \left[ c_1 p_x^2 + c_2 p_x + c_1 p_y^2 + c_3 p_y + c_4, l_3 + l_1 p_x + l_2 p_y \right]$$

Die zweite Gleichung ist linear in  $p_y$ . Lösen wir diese Gleichung auf, setzen das Ergebnis in die erste Gleichung ein und beseitigen Nenner

```
s:solve(polys[2],[py]);
poly2:l2^2*ratsimp(ev(polys[1],s));
```

$$p_x^2 c_1 (l_1^2 + l_2^2) + p_x (c_2 l_2^2 - c_3 l_1 l_2 + 2 c_1 l_1 l_3) + (c_4 l_2^2 - c_3 l_2 l_3 + c_1 l_3^2),$$

so erhalten wir ein Polynom zweiten Grades für die  $x$ -Koordinaten der beiden Schnittpunkte von  $c$  und  $l$ .  $l$  ist genau dann Tangente, wenn diese beiden Schnittpunkte zusammenfallen, wenn also die Diskriminante

```
Discr:coeff(poly2,[px],1)^2-4*coeff(poly2,[px],2)*coeff(poly2,[px],0);
```

dieses quadratischen Polynoms verschwindet.  $Discr$  enthält noch einen geometrisch irrelevanten Faktor  $l_2^2$ , so dass wir als Berührbedingung (nach einiger Termumformung) gerade

$$4 c_1 ((-c_1 l_3 + c_2 l_1 + c_3 l_2) l_3 - c_4 (l_1^2 + l_2^2)) + (c_2 l_2 - c_3 l_1)^2$$

erhalten. Dieses Polynom stimmt mit der Bedingung überein, die sich aus einer mehr geometrischen Argumentation ergibt: Eine Berührsituation liegt genau dann vor, wenn der Fußpunkt des Lots aus dem Kreismittelpunkt von  $c$  auf die Gerade  $l$  (im Allgemeinen ist das die Mitte der Sehne, die  $c$  aus  $l$  ausschneidet) selbst auf dem Kreis liegt:

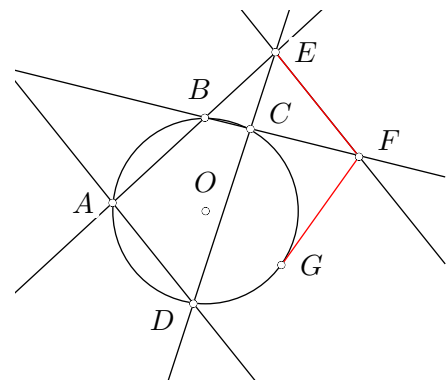
```
on_circle(pedalpoint(circle_center(c),l),c);
```

Diese Bedingung können wir ebenfalls in eine neue Prozedur gießen.

```
is_cl_tangent:=block([c1,c2,c3,c4,l1,l2,l3],
  c1:part(c,1), c2:part(c,2), c3:part(c,3), c4:part(c,4),
  l1:part(l,1), l2:part(l,2), l3:part(l,3),
  ratsimp(4*c1*((-c1*l3+c2*l1+c3*l2)*l3-c4*(l1^2+l2^2))+(c2*l2-c3*l1)^2
);
```

Damit können wir folgenden Satz beweisen:

**Satz 22 (Beispiel Chou 106)** Seien  $A, B, C, D$  vier Punkte auf einem Kreis  $c$ ,  $E$  der Schnittpunkt von  $AB$  und  $CD$ ,  $F$  der Schnittpunkt von  $BC$  und der Parallelen zu  $AD$  durch  $E$  und  $G$  der Berührungspunkt der Tangente aus  $F$  an den Kreis  $c$ . Dann ist  $|EF| = |FG|$ .



Hier ist die Grundkonstruktion für ein entsprechendes Beweisschema. Der Kreis  $c$  ist eindeutig bestimmt, wenn wir den Kreismittelpunkt  $O$  und  $A$  auf der Peripherie fixieren.

Zur Reduktion der Zahl der Variablen wählen wir ein spezielles Koordinatensystem.  $B, C, D$  und  $G$  können nun als Kreisgleiter gewählt werden. Für deren universelle Formeln führen wir die Variablen  $u_2, u_3, u_4$  und  $x_1$  ein.

```

O:Point(0,0); A:Point(1,0); B:circle_slider(0,A,u2);
C:circle_slider(0,A,u3); D:circle_slider(0,A,u4);
E:intersection_point(pp_line(B,A),pp_line(C,D));
F:intersection_point(par_line(E,pp_line(A,D)), pp_line(B,C));
G:circle_slider(0,A,y1);

```

Die weitere Voraussetzung, dass  $GE$  Tangente an den Kreis  $c$  ist, kann im Polynom

```
p1:is_cl_tangent(pc_circle(0,A),pp_line(F,G));
```

die Behauptung des Satzes als

```
p2:ratsimp(sqrdist(E,F)-sqrdist(F,G));
```

kodiert werden. Wir haben einen Satz vom Äquivalenztyp, den wir wieder mit Faktorisierung beweisen können:

```

p2a:num(p2);
expand((p2a)^2/4-p1);

```

0

Auch für die Bedingung, dass sich zwei Kreise berühren, kann eine einfache Bedingung hergeleitet werden. In diesem Fall ist die Potenzgerade genau die gemeinsame Tangente. Also können wir die Berührbedingung zweier Kreise auf die von Kreis und Gerade zurückführen:

```
is_cc_tangent(c1,c2)==is_cl_tangent(c1,radical_axis(c1,c2))
```

Damit können wir ein Beweisschema für eine interessante Eigenschaft des Feuerbachschen Kreises formulieren:

**Satz 23** *Der Feuerbachkreis berührt den Inkreis und die drei Ankreise des Dreiecks.*

Wir gehen dazu von der Konfiguration aus, die wir zum Beweis des Satzes vom Schnitt der Winkelhalbierenden verwendet hatten.

```

A:Point(1,0); B:Point(0,0);
C:Point(cx,cy); P:Point(px,py);
Q:pedalpoint(P,pp_line(A,B));

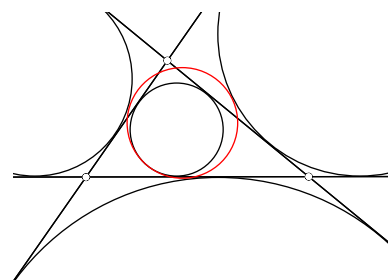
```

$Q$  als Lotfußpunkt von  $P$  auf  $AB$  wird später zur Konstruktion des jeweiligen Berührkreises benötigt. Wir ergänzen diese Konfiguration um die weiteren für die Konstruktion des Feuerbachkreises erforderlichen geometrischen Objekte

```

M:intersection_point(p_bisector(A,B), p_bisector(B,C));
H:intersection_point(altitude(A,B,C),altitude(B,C,A));
N:midpoint(M,H);
c1:pc_circle(N,midpoint(A,B));

```



$P$  als Zentrum des In- oder eines der Ankreise ist einer der vier Schnittpunkte der Winkelhalbierenden, was wir als zusätzliche Bedingung

```
polys:[ on_bisector(P,B,A,C), on_bisector(P,C,B,A) ];
```

anschreiben können, die Behauptung lautet, dass der zugehörige Kreis den Feuerbachkreis  $c_1$  berührt:

```
con:is_cc_tangent(pc_circle(P,Q),c1);
```

Das entsprechende Gleichungssystem kann wie weiter oben gelöst werden. Auch hier ergibt sich, dass Maxima den resultierenden Ausdruck zu null vereinfachen kann, wenn die komplizierten Wurzelausdrücke für  $(p_x, p_y)$  in  $con$  eingesetzt werden.

```
step1:solve(polys[1]-polys[2],[py]);
step2:ratsimp(ev(polys[1],step1[1]));
step3:num(step2);
step4:solve(step3,[px]);
step5:ev(ev(con,step1[1]),step4[1]);
ratsimp(step5);
```

0

Eine genauere Untersuchung der Beweiskraft dieses Ansatzes steht allerdings noch aus und soll im nächsten Kapitel durchgeführt werden.

*Aufgabe:* (1) Erstellen Sie ein Werkzeug `pascal_line(A, B, C, D, E, F)`, mit dem zu sechs beliebigen Punkten die Pascalgerade, also die Gerade durch die beiden Punkte  $AE \cap BD$  und  $BF \cap CE$ , konstruiert werden kann.

(2) Zeigen Sie, dass im Kontext des Satzes von Pappus ( $A, B, C$  und  $D, E, F$  jeweils kollinear) auch der Punkt  $AF \cap CD$  auf der Pascalgeraden (die in diesem Fall mit der Pappusgeraden zusammenfällt) liegt.

(3) Zeigen Sie, dass im Kontext des Satzes von Pappus je drei der sechs Pappusgeraden der Konfiguration durch einen gemeinsamen Punkt gehen.

(4) Beweisen Sie folgenden Satz:

**Satz 24** Sind  $X = AB \cap DE$ ,  $Y = BC \cap EF$ ,  $Z = CD \cap AF$  kollinear (also  $ABCDEF$  ein Pascalsches Sechseck), so sind auch  $R = AD \cap CF$ ,  $S = BD \cap CE$ ,  $T = BF \cap AE$  kollinear.

## 5 Geometrische Sätze vom Gleichungstyp

Im letzten Kapitel hatten wir geometrische Sätze studiert, deren algebraische Formulierung mit den universellen Formeln  $B_E$  einer konstruktiv erzeugten Basiskonfiguration  $\mathcal{K}$  startet, zusätzliche geometrische Voraussetzungen, eine Nichtdegenerationsbedingung sowie die Behauptung in einer Menge  $F = \{f_1, \dots, f_m\} \subset R$  von Polynomen, einem Polynom  $h \in R$  sowie einem Polynom  $g \in R$  kodiert, so dass der Beweis des geometrischen Satzes auf eine Aussage  $V(F) \subset V(h \cdot g)$  über Nullstellen von Polynomen bzw. rationalen Funktionen zurückgeführt werden kann.

### 5.1 Die Problemstellung

Wir wollen dies wie folgt präzisieren und voraussetzen, dass folgende Größen gegeben sind:

- eine konstruktive *rationale* geometrische Konfiguration  $\mathcal{K}$  mit universellen Formeln  $B_E(X)$  der Endkonfiguration, die aus einer universellen Startkonfiguration  $A_S(X)$  gewonnen wurde, was wir im Weiteren als *allgemeine geometrische Basiskonfiguration* (AGK) bezeichnen;
- eine Menge polynomialer<sup>2</sup> geometrischer Bedingungen mit Polynomen

$$F = \{f_1, \dots, f_m\} \subset R = k[X]$$

in den Variablen  $X = (x_1, \dots, x_n)$  als universelle Formeln, die weitere implizite Abhängigkeiten zwischen den Elementen von  $B_E$  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.

Der Beweis des geometrischen Satzes kann dann auf den Beweis von  $V(F) \subset V(h \cdot g)$  über dem Koordinatengrundbereich  $K/k$  zurückgeführt werden, den wir als algebraisch abgeschlossen vorausgesetzt hatten.

### 5.2 Geometriesätze vom Gleichungstyp und Ideale

Obwohl die explizite Bestimmung von  $V(F)$  schwierig ist und den Übergang zu algebraischen Erweiterungen  $K/k$  erfordert, lassen sich eine Reihe solcher Sätze bereits durch algebraische Überlegungen beweisen, die nur Umformungen von Polynomen über dem Grundkörper  $k$ , also nur rationale Operationen erfordern.

---

<sup>2</sup>Enthält  $B_E(X)$  rationale Formeln, so sind die universellen Formeln darauf aufbauender geometrischer Bedingungen auch rationale Ausdrücke. Rationale Ausdrücke verschwinden aber genau dann, wenn deren Zähler in einer rationalen Normalform verschwindet. Eine solche Reduktion der Problemstellung wollen wir hier stets voraussetzen.

In der linearen Algebra wird bei der Betrachtung linearer Gleichungssysteme von den Zeilenvektoren der zugehörigen Matrix zum Unterraum übergegangen, der von diesen Zeilenvektoren erzeugt ist. Dieses mathematische Objekt ist invariant bzgl. Zeilenumformungen, die etwa im Gauß-Algorithmus vorgenommen werden, womit die Invarianz der Lösungsmenge unter derartigen Transformationen gesichert wird.

Ähnlich gehen wir in der algebraischen Geometrie vom Polynomsystem  $F$  zum Ideal

$$I(F) = \{p_1f_1 + p_2f_2 + \dots + p_mf_m \mid p_1, \dots, p_m \in R\}$$

über, das von diesen Polynomen erzeugt wird. Das Ideal  $I(F)$  ist die Menge aller polynomialen Kombinationen der Elemente  $f \in F$ . Damit gilt  $V(F) = V(I)$  und jedes Polynom  $g \in I$  verschwindet auf allen Nullstellen  $X_0 \in V(F)$ .

**Satz 25** Ist  $I = I(F) \subset R$  das von  $F$  erzeugte Ideal und  $g \in I$ , so gilt der Satz ( $F \Rightarrow g$ ).

Ideale  $I \subset R$  haben die Eigenschaften

$$\begin{aligned} a, b \in I &\Rightarrow a + b \in I \\ a \in I, r \in R &\Rightarrow ra \in I. \end{aligned}$$

Die Modulo-Relation für Polynome ist definiert durch

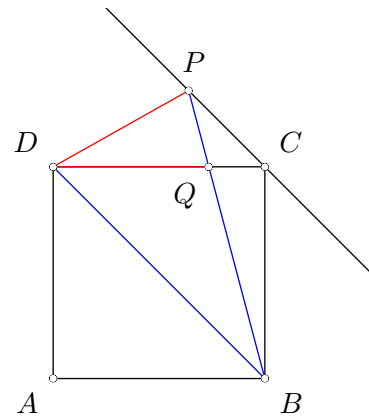
$$g_1 \equiv g_2 \pmod{I} \Leftrightarrow g_1 - g_2 \in I$$

und ist eine Kongruenzrelation auf  $R$ , welche die bekannte Modulo-Relation auf den ganzen Zahlen verallgemeinert.  $g \in I$  und  $g \equiv 0 \pmod{I}$  sind äquivalente Aussagen.

Betrachten wir dazu zunächst das folgende Beispiel von **Arnon** ([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.



Für die Basiskonfiguration wählen wir ein Koordinatensystem, so dass die Eckpunkte des Quadrats die Koordinaten

A:Point(0,0); B:Point(1,0);  
C:Point(1,1); D:Point(0,1);

haben.

Diese Basiskonfiguration erweitern wir um einen freien Punkt  $P$  und den Geradengleiter  $Q \in CD$ .



```
P:Point(px,py); Q:varpoint(C,D,qs);
```

Die universellen Formeln der Basiskonfiguration enthalten die drei Variablen  $X = (p_x, p_y, q_s)$ . Die AGV lassen sich durch ein System  $F = \{f_1, f_2, f_3\}$  von drei Polynomen anschreiben

```
polys: [
  on_line(P,par_line(C,pp_line(B,D))),
  sqrdist(B,D)-sqrdist(B,P),
  on_line(Q,pp_line(B,P))
];
```

$$[p_x + p_y - 2, -p_x^2 + 2p_x - p_y^2 + 1, 1 - p_y q_s - p_x],$$

aus denen sich alle zulässigen Werte  $X_0$  mit  $F(X_0) = 0$  bestimmen lassen.

```
xvars: [px,py,qs];
sol:solve(polys,xvars);
```

$$\left[ \begin{array}{l} \left[ p_x = \frac{3 + 2\sqrt{3}}{\sqrt{3} + 1}, p_y = -\frac{2 + \sqrt{3}}{3\sqrt{3} + 5}, q_s = \sqrt{3} + 2 \right], \\ \left[ p_x = \frac{2\sqrt{3} - 3}{\sqrt{3} - 1}, p_y = -\frac{\sqrt{3} - 2}{3\sqrt{3} - 5}, q_s = 2 - \sqrt{3} \right] \end{array} \right]$$

Sehen wir uns den Lösungsweg genauer an: Lösen wir die erste Gleichung nach  $p_y = 2 - p_x$  auf und setzen dies in die zweite und dritte Gleichung ein, so ergibt sich

$$[-2p_x^2 + 6p_x - 3, p_x q_s - 2q_s - p_x + 1]$$

Das erste Polynom ergibt nach Reskalierung  $f_4 = p_x^2 - 3p_x - \frac{3}{2} \in I = I(F)$ , das zweite eine rationale Formel  $q_s = \frac{p_x - 1}{p_x - 2}$ . Aus der ersten Gleichung lassen sich die zwei (nicht mehr rationalen) Werte für  $p_x$  bestimmen. Setzt man jeden dieser Werte in die zweite Gleichung ein, so kann dazu stets ein eindeutig bestimmter Wert  $q_s$  ermittelt werden.

Die Behauptung

```
g:sqrdist(D,Q) - sqrdist(D,P);
```

$$g = -p_x^2 - p_y^2 + 2p_y + q_s^2 - 2q_s$$

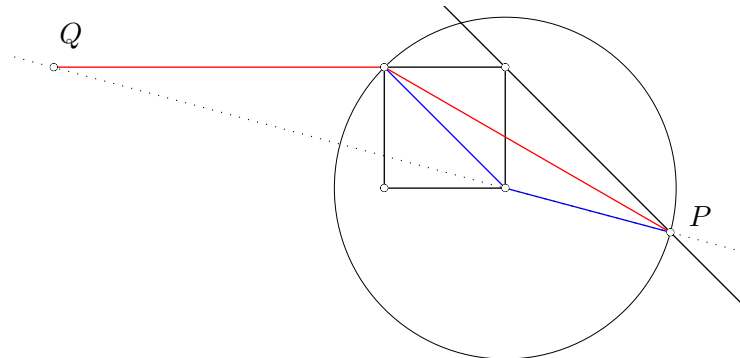
gilt für all diese Werte, wie die Rechnung

```
map(lambda([u],ratsimp(ev(g,u))),sol);
```

$$[0,0]$$

zeigt. Jedoch wurde dabei wiederum direkt mit algebraischen Zahlen gerechnet. Es sei noch angemerkt, dass die Rechnung *zwei* Lösungen für  $P$  liefert. Das ist auch plausibel, weil  $|BP| = |BD|$  irgend etwas mit dem Kreis um  $B$  durch  $D$  zu tun hat und dieser Kreis natürlich *zwei* Schnittpunkte mit der Parallelen durch  $C$  hat.

In der Abbildung ist die zweite Lösung zu sehen. Wären Sie ohne unsere Rechnungen auf diesen zweiten Fall gekommen?



*Aufgabe:* Formulieren und beweisen Sie die Aussage des Arnon-Beispiels als Satz vom Äquivalenztyp, der mit einer Variablen und einer Gleichung als AGV auskommt.

### 5.3 Ideale und Gröbnerbasen

Gehen wir vom Ideal

$$I = I(F) = (f_1 = p_x + p_y - 2, f_2 = -p_x^2 + 2p_x - p_y^2 + 1, f_3 = 1 - p_y q_s - p_x)$$

aus, das vom Polynomsystem  $F$  erzeugt wird, so können wir aus dem Polynom  $f_1$  die algebraische Ersetzungsregel

$$p_y \equiv 2 - p_x \pmod{I} \quad (\text{R.1})$$

herleiten. Damit können wir  $p_y$  in  $f_2$  und  $f_3$  eliminieren.

$$\begin{aligned} f_2 &\equiv f_4 = -2p_x^2 + 6p_x - 3 \pmod{I} \\ f_3 &\equiv f_5 = p_x q_s - 2q_s - p_x + 1 \pmod{I}. \end{aligned}$$

Mit  $f_2, f_3 \in I$  folgt auch  $f_4, f_5 \in I$ . Aus  $f_4$  können wir analog die algebraische Ersetzungsregel

$$p_x^2 \equiv 3p_x - \frac{3}{2} \pmod{I} \quad (\text{R.2})$$

herleiten. Wir können diese Ersetzungsregeln verwenden, um in  $g(p_x, p_y, q_s)$  alle monomialen Vielfachen von  $p_y$  und  $p_x^2$  zu eliminieren.

```
g1: expand(ev(g, py = 2 - px));
g2: expand(ev(g1, px^2=3*px-3/2));
```

$$\begin{aligned} g_1 &= -2p_x^2 + 2p_x + q_s^2 - 2q_s \\ g_2 &= q_s^2 - 2q_s - 4p_x + 3 \end{aligned}$$

Es gilt  $g \equiv g_1 \equiv g_2 \pmod{I}$  und  $g_2$  hat die Gestalt  $g_2 = -4p_x + (q_s^2 - 2q_s + 3)$ . Aus der dritten Gleichung

$$f_5 = p_x q_s - 2q_s - p_x + 1 = q_s(p_x - 2) - p_x + 1$$

ergibt sich  $q_s = \frac{p_x - 1}{p_x - 2}$ , also keine polynomiale Abhängigkeit. Einsetzen in  $g_2$  liefert

```
g3:ratsimp(ev(g2,qs=(px-1)/(px-2)));
```

$$g3 = -\frac{-9 + 24p_x - 18p_x^2 + 4p_x^3}{p_x^2 - 4p_x + 4}$$

Hier können wir zunächst formal  $p_x^3$  und dann noch  $p_x^2$  mit der Regel (R.2) ersetzen:

```
g3a:ratsimp(ev(g3,px^3=3*px^2-3/2*px));
g3b:ratsimp(ev(g3a,px^2=3*px-3/2));
```

und erhalten im letzten Schritt null. Allerdings haben wir diese Rechnungen nicht im Ring  $R$  ausgeführt und können uns deshalb auch nicht auf das Rechnen modulo des Ideals  $I$  berufen.

Nun gilt aber

$$(p_x - 2)(p_x - 1) = p_x^2 - 3p_x + 2 = f_4 + \frac{1}{2} \equiv \frac{1}{2} \pmod{I}.$$

und damit für  $f_5 = p_x q_s - 2q_s - p_x + 1 = q_s(p_x - 2) + (1 - p_x) \in I$  auch

$$(p_x - 1)f_5 \equiv \frac{1}{2}q_s - (p_x - 1)^2 = \frac{1}{2}q_s - p_x^2 + 2p_x - 1 \equiv \frac{1}{2}q_s - p_x + \frac{1}{2} \pmod{I},$$

damit  $f_6 = q_s - 2p_x + 1 \in I$  und schließlich  $q_s \equiv 2p_x - 1 \pmod{I}$ . Wir können also auch die folgenden Rechnungen ausführen:

```
g1:expand(ev(g,py = 2 - px));
g2:expand(ev(g1,qs=2*px - 1));
g3:expand(ev(g2,px^2=3*px-3/2));
```

Es ergibt sich  $g \equiv g_1 \equiv g_2 \equiv g_3 = 0 \pmod{I}$ , also  $g \in I$ . Damit ist der geometrische Satz bewiesen.

Im Zuge der Untersuchungen haben wir mit

$$G = \left\{ p_y + p_x - 2, q_s - 2p_x + 1, p_x^2 - 3p_x + \frac{3}{2} \right\}$$

eine andere Basis des Ideals  $I = I(F)$  konstruiert, die unmittelbar erlaubt, alle monomialen Vielfachen von  $p_y$ ,  $q_s$  und  $p_x^2$  zu ersetzen und so zu jedem Polynom  $g(p_x, p_y, q_s)$  ein Polynom  $g' = ap_x + b$  mit  $g \equiv g' \pmod{I}$  zu konstruieren. Eine solche Basis  $G$  heißt *Gröbnerbasis* (bzgl. der lexikografischen Termordnung), das Polynom  $g'$  *Normalform* von  $g$  bzgl. der Basis  $G$ .

Gröberbasen und Normalformen existieren für jedes Ideal und sind in Maxima im **grobner**-Paket implementiert. Es ist zu beachten, dass bei der Berechnung von Gröbnerbasen die distributive Darstellung von Polynomen (in der Standardeinstellung bzgl. der lexikographischen Termordnung) verwendet wird. Das Gröbnerpaket implementiert deshalb eine eigene Polynomarithmetik, die von der Reihenfolge der Variablen abhängt. Deshalb muss bei jedem Kommando die Liste der Variablen angegeben werden. Das jeweilige Ergebnis und auch die Rechenzeiten können mit der Wahl dieser Variablenreihenfolge stark variieren.

Wir verwenden die folgenden Funktionsdefinitionen, um Gröbnerbasen effektiv zu berechnen, ohne sich mit der speziellen Maxima-Datenstruktur für Polynome beschäftigen zu müssen.

```
NF(f, polys, vars) := poly_normal_form(f, polys, vars);
GBasis(polys, vars) := poly_reduced_grobner(polys, vars);
```

Zum besseren Verständnis des Geschehens ist es sinnvoll, an einigen Stellen die Anordnung der Terme eines Polynoms  $f \in R$  entsprechend der Termordnung zu studieren und den größten Term des Polynoms  $f$  bzgl. dieser Termordnung, das *Leitmonom*  $lm(f)$  zu extrahieren. Dies ist mit den folgenden beiden Funktionen möglich:

```
grob_order(p,v) := block([poly_return_term_list: true], poly_add(p,0,v));
LM(p,v) := first(grob_order(p,v));
```

Damit können wir den Beweis für das Arnon-Beispiel nun kurz wie folgt führen:

```
xvars:=[qs,py,px];
G:GBasis(polys,xvars);
g1:NF(g,G,xvars);
```

$$G = [p_y + p_x - 2, 2p_x^2 - 6p_x + 3, -q_s + 2p_x - 1]$$

$$g_1 = 0$$

Allgemein gilt

**Satz 26** *Ist  $X$  eine geordnete Liste der Variablen, die in den universellen Formeln  $B_E$  vorkommen,  $G = \text{GBasis}(F, X)$  eine Gröbnerbasis von  $F$  und  $\text{NF}(g, G, X) = 0$ , so gilt der Satz ( $F \Rightarrow g$ ).*

Das Arnon-Beispiel war besonders einfach, weil in ihm alle Variablen durch Gleichungen gebunden sind. Für allgemeinere Beweisschemata läuft der vorgestellte Ansatz meist nicht so glatt durch.

Sehen wir uns als Nächstes noch einmal die Argumentation zum Satz vom Schnittpunkt der Winkelhalbierenden an. Wir hatten ausgehend von der polynomialen Grundkonfiguration

```
A:Point(0,0); B:Point(1,0); C:Point(cx,cy); P:Point(px,py);
```

die AGV als Polynomsystem  $F = \{f_1, f_2\}$

```
polys:[ on_bisector(P,B,A,C), on_bisector(P,C,B,A) ]
```

und die Behauptung als

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

angeschrieben. Berechnen wir hier die Gröbnerbasis

```
vars:[px,py,cx,cy];
G:GBasis(polys,vars);
```

$$\begin{aligned}
& 2c_x^2 p_y - 2c_y c_x p_x - 2c_x p_y + c_y c_x + \underline{c_y p_x^2} - c_y p_y^2 \\
& \frac{1}{2} c_y - p_y + c_x p_y - c_y p_x + \underline{p_x p_y} \\
& 2c_x p_y^2 - \frac{1}{2} c_x c_y^2 - \frac{1}{2} c_y^2 p_x + c_y p_y^3 + \frac{1}{2} c_y^2 - 2c_x^2 p_y^2 - c_y^2 p_y^2 - c_y p_y + c_x c_y p_y \\
& \quad + \underline{c_x c_y^2 p_x} - c_x^2 c_y p_y \\
& c_x^2 c_y p_y^2 - 2c_x^2 p_y^3 - c_x c_y p_y^2 + 2c_x p_y^3 + c_y^3 p_y^2 - \frac{1}{4} c_y^3 - 2c_y^2 p_y^3 + c_y^2 p_y + \underline{c_y p_y^4} - c_y p_y^2
\end{aligned}$$

Aus zwei Polynomen sind hier vier geworden. Eine Gröbnerbasis kann deutlich mehr Elemente enthalten als die Ausgangsbasis. In der Ausgabe werden die Terme nicht lexikographisch angeordnet, so dass die Leitertme nicht sofort zu erkennen sind, die in den Ersetzungsregeln verwendet werden. `map(lambda([u],LM(u,vars)),G)` zeigt die Leitertme an; sie sind in obigen Formeln unterstrichen.

`NF(g,G,vars);`

0

Diese Rechnung beweist den Satz vom Schnittpunkt der Winkelhalbierenden, denn auch hier gilt  $g \in I(F)$ .

*Weiteres Beispiel:* Auf diese Weise können wir auch beweisen, dass der Feuerbachkreis den Inkreis und die drei Ankreise des Dreiecks berührt. Die AGV und die Gröbnerbasis sind dieselben wie oben, die Basiskonfiguration und die geometrische Behauptung  $g$ , die wir aus dem letzten Kapitel übernehmen, etwas komplexer:  $M$  und  $H$  sind Umkreismittelpunkt und Höhenschnittpunkt im Dreieck  $ABC$ ,  $N$  als Mittelpunkt von  $\overline{MH}$  der Mittelpunkt des Feuerbachkreises, der u.a. durch die Seitenmitte von  $\overline{AB}$  geht. Damit kann der Feuerbachkreis als  $c_1$  angeschrieben werden. In- und Ankreise haben als Mittelpunkt  $P$ , der Lotfußpunkt aus  $P$  auf die Gerade  $AB$  ist ein Punkt auf der Peripherie des entsprechenden Berührkreises.

```

M:intersection_point(p_bisector(A,B), p_bisector(B,C));
H:intersection_point(altitude(A,B,C),altitude(B,C,A));
N:midpoint(M,H);
c1:pc_circle(N,midpoint(A,B));
Q:pedalpoint(P,pp_line(A,B));

g:is_cc_tangent(pc_circle(P,Q),c1);

```

`NF(g,G,vars);`

0

**Satz 27 (Dualer Satz von Pappus)** Seien  $g_1, g_2, g_3$  und  $h_1, h_2, h_3$  jeweils konkurrente Geraden und  $A, \dots, F$  die Schnittpunkte  $A = h_1 \wedge g_1, B = g_1 \wedge h_2, C = h_2 \wedge g_2, D = g_2 \wedge h_3, E = h_3 \wedge g_3$  und  $F = g_3 \wedge h_1$ . Die drei Verbindungsgeraden  $AD, BE$  und  $CF$  sind konkurrent.

*Aufgabe:* Beweisen Sie diesen Satz.

#### 5.4 Weitere Beweisansätze

Wir haben allerdings eine für die Rechnung günstige Auswahl getroffen. Wären wir von einem Schnittpunkt der Winkelhalbierendenpaare durch  $B$  und  $C$  gestartet und hätten zeigen wollen, dass dieser auch auf dem Winkelhalbierendenpaar durch  $A$  liegt, so wären wir zunächst gescheitert. Auch diesmal enthält die Gröbnerbasis vier Elemente, aber  $g$  kann damit nicht zu null reduziert werden.

```
polys:[ on_bisector(P,C,B,A), on_bisector(P,A,C,B) ];
g:on_bisector(P,B,A,C);
G:GBasis(polys,vars);
g1:NF(g,G,vars);
```

$$g_1 = 2p_y - c_y - 2c_x p_y + 2c_y p_x - 2p_x p_y$$

Wegen  $g \equiv g_1 \not\equiv 0 \pmod{I}$  gilt sogar  $g \notin I$ . Dies ergibt sich aus der Gröbnerbasen-Eigenschaft von  $G$ . Allerdings reduziert

```
NF(((cx-1)^2+cy^2)*g,G,vars);
```

zu null. Für das Polynom  $h = (c_x - 1)^2 + c_y^2$  gilt also  $h \cdot g \in I$ . Der Satz  $(F \Rightarrow g)_h$  ist zum Satz  $(F \Rightarrow g \cdot h)$  äquivalent und damit unser Satz unter der Nichtdegenerationsbedingung  $h$  bewiesen. Dies gilt auch allgemein.

**Satz 28** *Ist  $G = \text{GBasis}(F, X)$  eine Gröbnerbasis von  $F$  und  $\text{NF}(h^k g, G, X) = 0$  für ein  $k \geq 0$ , so gilt der Satz  $(F \Rightarrow g)_h$ .*

Die Frage nach der geometrischen Relevanz dieser Nichtdegenerationsbedingung bleibt noch zu diskutieren.  $h$  ist genau die geometrische Bedingung  $\text{sqrdist}(B, C) = 0$ . Über den reellen Zahlen entspricht dies  $B = C$ , so dass die Gerade  $BC$  als ein Schenkel entartet. Im Allgemeinen bedeutet  $\text{sqrdist}(B, C) = 0$  aber nur, dass die Gerade  $l = BC$  isotrop ist, also senkrecht auf sich selbst steht. In unserer ersten Rechnung waren wir von den Winkelhalbierenden durch  $A$  und  $B$  ausgegangen und die spezielle Wahl des Koordinatensystems hatte dafür gesorgt, dass  $AB$  nicht isotrop war. Das „o. B. d. A.“ ist also möglicherweise doch nicht ganz unbedenklich, denn die Eigenschaft einer Geraden, isotrop zu sein, ist natürlich vom Koordinatensystem unabhängig. Wenn eine der drei Seiten des Dreiecks nicht isotrop ist, dann können wir unseren Beweis stets mit dieser Seite starten. Es bleibt also offen, wie das mit den Winkelhalbierenden für ein Dreieck ist, dessen Seiten sämtlich isotrop sind. Diesen Fall haben wir auch mit unserer ersten Berechnung (durch die spezielle Wahl der Koordinaten für  $A$  und  $B$ ) nicht untersucht.

Wir wollen diese komplizierten Fragen der geometrischen Interpretation von auf algebraischem Weg gefundenen Nichtdegenerationsbedingungen hier nicht weiter verfolgen und uns der zweiten Frage zuwenden, wie derartige Nichtdegenerationsbedingungen algebraisch gefunden werden können. Eine systematische Suche könnte zum Beispiel nach der Menge

$$\{h \in R : h \cdot g \in I\}$$

oder allgemeiner

$$\left\{ h \in R : \exists k \in \mathbb{N} \ h^k \cdot g \in I \right\}$$

aller Polynome fragen, die den „Beweis retten“. Es stellt sich heraus, dass diese Mengen auch Ideale sind und als Idealquotient bzw. stabiler Idealquotient berechnet werden können.

Ein weiterer **indirekter Beweisansatz** ergibt sich aus folgender Überlegung: Ist  $g$  ein Polynom, das auf allen gemeinsamen Nullstellen von  $F$  verschwindet, und  $t$  eine neue Variable, so hat offensichtlich das System  $F = 1 - t \cdot g = 0$  keine gemeinsamen Lösungen. Ist umgekehrt  $X_0$  eine Lösung mit  $F(X_0) = 0$ , aber  $g(X_0) \neq 0$ , so hat  $F = 1 - t \cdot g = 0$  die Lösung  $(t_0, X_0)$  mit  $t_0 = \frac{1}{g(X_0)}$ .

$V(F, 1-t \cdot g) = \emptyset$  ist aber nach dem Hilbertschen Nullstellensatz äquivalent dazu, dass das Ideal  $I = I(F, 1-t \cdot g) \subset k[t, X]$  das Einsideal ist, was wiederum durch eine Gröbnerbasisberechnung geprüft werden kann. Für obiges Beispiel ergibt sich in der ersten Variante

```
polys:[ on_bisector(P,B,A,C), on_bisector(P,C,B,A) ];
g:on_bisector(P,A,C,B);
GBasis(append(polys,[1-t*g]), [t,px,py,cx,cy]);
```

unmittelbar das Einsideal, also  $V(F) \subset V(g)$ .

**Satz 29** *Ist  $G = GBasis([F, 1 - t \cdot g], [t, X]) = (1)$ , so gilt der Satz  $(F \Rightarrow g)$ .*

In der zweiten Variante

```
polys:[ on_bisector(P,C,B,A), on_bisector(P,A,C,B) ];
g:on_bisector(P,B,A,C);
gb:GBasis(append(polys,[1-t*g]), [t,px,py,cx,cy]);
```

ergibt sich leider nicht das Einsideal, sondern eine komplizierte Gröbnerbasis mit folgendem Polynom allein in den Variablen  $c_x, c_y$ , von deren Unabhängigkeit wir ausgegangen waren:

```
poly_elimination_ideal (gb, 3, vars);
```

$$[c_x^2 - 2c_x + c_y^2 + 1]$$

Dies ist genau das oben verwendete Polynom  $h \in R$ , für das wir dann  $V(F) \subset V(h \cdot g)$  zeigen konnten. Hier wurde dieses Polynom allerdings systematisch konstruiert: Es ergab sich  $h \in I(F, 1 - t \cdot g)$  und damit

$$V(I(F, 1 - t \cdot g)) = V(F) \cap V(1 - t \cdot g) \subset V(h)$$

– jede gemeinsame Nullstelle  $X_0$  von  $F$ , die keine Nullstelle von  $g$  ist, liegt im Degenerationslokalus  $V(h)$ . Dieses Ergebnis gilt auch allgemein

**Satz 30** *Ist  $G = GBasis([F, 1 - t \cdot g], [t, X])$  und  $NF(h, G, [t, X]) = 0$  für ein  $h \in R$ , so gilt der Satz  $(F \Rightarrow g)_h$ .*

*Beweis:* Aus  $\text{NF}(h, G, [t, X]) = 0$  und damit  $h \in I(F, 1 - t \cdot g)$  folgt die Existenz einer polynomialen Darstellung

$$h(X) = \sum_i p_i(X, t) \cdot f_i(X) + p(X, t) \cdot (1 - t \cdot g(X)). \quad (*)$$

in  $k[t, X]$ . Gäbe es nun ein  $X_0 \in \mathbb{A}^n$  mit  $F(X_0) = 0$ , aber  $g(X_0) \neq 0$ , so können wir diesen Wert in die Darstellung (\*) einsetzen:

$$h(X_0) = \sum_i p_i(X_0, t) \cdot f_i(X_0) + p(X_0, t) \cdot (1 - t \cdot g(X_0)) = p(X_0, t) \cdot (1 - t \cdot g(X_0))$$

wegen  $f(X_0) = 0$  für alle  $f \in F$ . Auf der linken Seite steht ein Polynom vom Grad 0 in  $t$ , auf der rechten Seite ein Polynom von positivem Grad in  $t$ . Die Annahme der Existenz eines solchen  $X_0$  führt also auf einen Widerspruch.  $\square$

Allgemein ist das Ideal der  $t$ -freien Polynome  $I(F, 1 - t \cdot g) \cap k[X]$  gerade der *stabile Idealquotient*

$$I : g^\infty = \{h \in R : \exists n > 0 (g^n \cdot h \in I)\}$$

und kann als Eliminationsideal berechnet werden. Im obigen Beispiel ergibt sich

```
poly_ideal_saturation1 (polys, g, vars);
```

$$[(-c_y p_y^2) - 2 c_x p_x p_y + 2 p_x p_y + 2 c_x p_y - 2 p_y + c_y p_x^2 - 2 c_y p_x + c_y, c_y^2 + c_x^2 - 2 c_x + 1]$$

was mit

```
poly_elimination_ideal (gb, 1, vars);
```

übereinstimmt. Dies ist im Wesentlichen die Menge  $\{h \in R : V(g \cdot h) \supset V(F)\}$  derjenigen Polynome, für welche der Satz  $(F \Rightarrow g)_h$  gilt.

## 5.5 Unabhängige und abhängige Variablen

Wir haben 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  $I \subset R = k[X]$  ein Ideal, so heißt eine Teilmenge  $U \subset X$  der Variablen *unabhängig* bzgl.  $I$ , wenn  $I \cap k[U] = (0)$  gilt, d. h.  $I$  kein nicht triviales Polynom allein in den  $U$ -Variablen enthält.

Der Nachweis, dass eine vorgegebene Variablenmenge unabhängig ist, kann nach Definition auf ein Eliminationsproblem reduziert werden, das mit der Berechnung einer Gröbnerbasis beantwortet werden kann: Ist  $G$  eine Gröbnerbasis von  $I$  bzgl. einer Eliminationsordnung für  $U$ , so ist

$$G' = \{g \in G : \text{lt}(g) \in k[U]\}$$

eine Gröbnerbasis von  $I \cap k[U]$ .  $U$  ist also genau dann unabhängig bzgl.  $I$ , wenn  $G' = \emptyset$  gilt.



Da degenerierte Lagen in diesem Sinne stets durch Bedingungen auf die ansonsten frei wählbaren Parameter charakterisiert sind, sind vor allem Polynome  $h \in k[U]$  als Nichtdegenerationsbedingungen von geometrischer Relevanz. Genau ein solches Polynom haben wir für unser Beispiel oben (zweite Variante des Schnitts der Winkelhalbierenden) bereits in  $G$  selbst gefunden.

## 5.6 Generisch gültige Geometrietheoreme

Im Fall der Geometrietheoreme vom konstruktiven und vom linearen Typ haben wir die Menge der Variablen  $X = (Y, U)$  in unabhängige und abhängige unterteilt und lineare Algebra über dem Körper  $k(U)$  der rationalen Funktionen in  $U$  getrieben. Wir wollen deshalb nun auch für den allgemeinen Fall im Ring  $S = k(U)[Y]$  der Polynome in  $Y$  mit rationalen Funktionen in  $U$  als Koeffizienten rechnen statt wie bisher im Ring  $R = k[X] = k[Y, U]$ . Wir wollen dabei voraussetzen, dass  $U$  eine maximale bzgl.  $I = I(F)$  unabhängige Variablenmenge ist.

Betrachten wir zunächst einige Beispiele.

**Beispiel 1: Satz vom Miquelschen Punkt.** Wir wählen als AGK die folgende Basis-konfiguration, wobei wir im Weiteren einheitlich Koordinaten freier Punkte wie  $C(c_x, c_y)$  und Gleiterparameter wie  $p_s$  für den Geradengleiter  $P$  bezeichnen wollen. Gegeben ist  $\triangle ABC$ ,  $P, Q, R$  als Geradengleiter auf den jeweiligen Trägergeraden der Dreiecksseiten und  $S$  als weiterer freier Punkt, der durch die AGV als Schnittpunkt der beiden Kreise durch  $A$  und  $B$  ausgezeichnet wird.

```
A:Point(0,0); B:Point(1,0); C:Point(cx,cy);
P:varpoint(B,C,ps); Q:varpoint(A,C,qs); R:varpoint(A,B,rs);
S:Point(sx,sy);
```

```
polys:[ is_concyclic(A,Q,R,S), is_concyclic(B,P,R,S) ];
g: is_concyclic(C,P,Q,S);
```

Die beiden AGV-Polynome legen nahe, die Variablen  $Y = (s_x, s_y)$  als abhängig zu betrachten, die restlichen als unabhängig. Dies kann an dieser Stelle zwar nicht begründet werden, doch wir rechnen erst einmal los – vielleicht ergibt sich wie im Fall der Sätze vom linearen Typ eine Begründung im Nachhinein.

```
yvars:[sx,sy]; uvars:[cx,cy,ps,qs,rs];
G:GBasis(polys,yvars);
NF(g,G,yvars);
```

Leider haben wir keinen Erfolg, die Normalform reduziert nicht zu null. Der Grund ist allerdings klar, denn von den *zwei* Möglichkeiten für  $S$  ist eine der bereits bekannte Schnittpunkt  $R$  der beiden Kreise, der natürlich nicht auf dem dritten Kreis liegt.  $V(F) = V_1 \cup V_2$  besteht also aus zwei Komponenten und der Satz gilt nur auf einer von ihnen.

Schauen wir uns die Struktur der Polynome dieser Gröbnerbasis genauer an:

```
map(lambda([u],grob_order(u,yvars)),G);
map(lambda([u],LM(u,yvars)),G);
```

Da die Gröbnerbasis (nach Division durch den jeweiligen Leitkoeffizienten aus  $k(U)$ ) die einfache Struktur

$$[s_x - p_1(U) s_y - r_s, s_y^2 - p_2(U) s_y]$$

hat, lassen sich die beiden Lösungen  $(s_x, s_y) \in \mathbb{A}_{k(U)}^2$  des zugehörigen Gleichungssystems mit  $s_{y,1} = 0$  und  $s_{y,2} = p_2(U)$  leicht bestimmen. Eine Lösung entspricht dem Schnittpunkt  $R$  der beiden Kreise, die andere dem Punkt  $S$  und erfüllt die Gleichung  $g$ :

```
sol:solve(G,yvars);
ratsimp(ev(g,sol[1]));
ratsimp(ev(g,sol[2]));
```

Die Nullstellenmenge  $V(F) \subset \mathbb{A}_{k(U)}^2$  besteht in diesem Fall also aus zwei (nulldimensionalen über  $k(U)$ ) Komponenten, wobei der Satz auf einer Komponente gilt, auf der zweiten (aus naheliegenden 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** In der zweiten Variante unserer bisherigen Rechnungen ergibt sich mit der natürlichen Einteilung in abhängige  $Y = (p_x, p_y)$  und unabhängige  $U = (c_x, c_y)$  Variablen

```
A:Point(0,0); B:Point(1,0); C:Point(cx,cy); P:Point(px,py);
polys:[ on_bisector(P,C,B,A), on_bisector(P,A,C,B) ];
g:on_bisector(P,B,A,C);

yvars:[px,py]; uvars:[cx,cy];
G:GBasis(polys,yvars);
```

$$p_x + p_1(U) p_y^3 + p_2(U) p_y^2 + p_3(U) p_y + p_4(U) \\ p_y^4 + q_1(U) p_y^3 + q_2(U) p_y^2 + q_3(U) p_y + q_4(U)$$

Gegenüber der Maxima-Ausgabe wurde wieder durch die Leitkoeffizienten der Polynome dividiert, um die Essenz der Struktur zu verdeutlichen.

Auch in diesem Fall ist das von  $F$  in  $S = k(U)[p_x, p_y]$  erzeugte Ideal  $I' = I(F) \cdot S = I(G)$  null-dimensional und hat genau vier Nullstellen in  $\mathbb{A}_L^2$ , dem zweidimensionalen affinen Raum über  $L = \overline{k(U)}$ , dem algebraischen Abschluss von  $k(U)$ . 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

```
NF(g,G,yvars);
```

zu null reduziert und in  $S$  folglich  $g \in I'$  gilt. Die Probleme mit der degenerierten Situation  $\text{sqrdist}(B, C) = 0$  treten nicht auf.

**Beispiel 3: Die Simsonsche Gerade** Als AGK wählen wir einen durch das Zentrum  $M$  und einen Punkt  $A$  auf der Peripherie definierten Kreis  $c$ , weitere freie Punkte  $B, C, D$  sowie Geradengleiter  $R, S, T$ .

```
M:Point(0,0); A:Point(0,1); c:pc_circle(M,A);
B:Point(bx,by); C:Point(cx,cy); D:Point(dx,dy);
R:varpoint(B,C,rs); S:varpoint(A,C,ss); T:varpoint(A,B,ts);
```

Als AGV ergeben sich sechs Bestimmungsgleichungen

```
polys:[
  on_circle(B,c), on_circle(C,c),on_circle(D,c),
  is_orthogonal(pp_line(A,B),pp_line(D,T)),
  is_orthogonal(pp_line(A,C),pp_line(D,S)),
  is_orthogonal(pp_line(B,C),pp_line(D,R))
];
```

und als Behauptung das Polynom

```
g: is_collinear(R,S,T);
```

Die entsprechende Setzung der Variablen und Gröbnerbasisberechnung

```
yvars:[rs,ss,ts,by,cy,dy]; uvars:[bx,cx,dx];
G:GBasis(polys,yvars);
NF(g,G,yvars);
```

reduziert  $g$  wieder zu null. Mit

```
map(lambda([u],grob_order(u,yvars)),G);
```

kann die Gestalt der Gröbnerbasis studiert werden. Auch in diesem Fall ist das Ideal  $I' = I(F) \cdot k(U)[Y]$  nulldimensional. Aus der Gröbnerbasis lesen wir ab, dass dieses System 8 „generische“ Lösungen besitzt<sup>3</sup> und  $g$  auf allen diesen Punkten verschwindet, d. h. der Satz von der Simsonschen Geraden „generisch“ (d. h. in  $\mathbb{A}_L^2$ ) gilt.

*Aufgabe:* Beweisen Sie auf dieselbe Weise die Umkehrung des Satzes von der Simsonschen Geraden.

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  $Y$  gibt, die sich durch komplizierte *universelle Formeln* in den Parametern  $U$  ausdrücken lassen. Diese Formeln „leben“ nicht mehr in  $k(U)$ , 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.

<sup>3</sup>Dies ergibt sich aus dem spezifischen proof scheme – zu jeder der Wahlen  $(b_x, c_x, d_x)$  gibt es je *zwei* Werte für  $(b_y, c_y, d_y)$ , zusammen also 8 Komponenten.

Wenn  $U$  eine bzgl.  $I = I(F)$  maximale unabhängige Variablenmenge ist, so ist in jedem Fall das Erweiterungsideal  $I' = I \cdot S$  ein nulldimensionales Ideal und  $S/I'$  ein endlichdimensionaler  $k(U)$ -Vektorraum. Dessen Dimension gibt an, wie viele Lösungen  $Y$  es über dem „generischen“ Tupel  $U$  gibt. Beides kann man aus einer Gröbnerbasis  $G' = \mathbf{GBasis}(F, Y)$  ablesen: Dimension null liegt vor, wenn die Anzahl der Standardterme  $N(G', Y) \subset T(Y)$  endlich ist, die Vektorraumdimension stimmt dann mit dieser Anzahl überein, da  $N(G', Y)$  eine  $k(U)$ -Vektorraumbasis von  $S/I'$  ist. Diese „Punkte“ im affinen Raum über  $k(U)$  werden in der algebraischen Geometrie auch als *allgemeine Punkte* der zugehörigen Varietät bezeichnet.

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

**Satz 31 (Gültigkeit geometrischer Sätze vom Gleichungstyp „im Allgemeinen“)**

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

Sei weiter  $G' = \mathbf{GBasis}(F, Y)$  eine über  $S$  berechnete Gröbnerbasis.

Gilt  $\mathbf{NF}(g, G', Y) = 0$ , so gibt es eine (effektiv konstruierbare) Nichtdegenerationsbedingung  $h(U)$ , so dass  $g$  auf allen gemeinsamen Nullstellen  $(Y_0, U_0)$  von  $F$  mit  $h(U_0) \neq 0$  verschwindet, also der Satz  $(F \Rightarrow g)_h$  gilt.

*Beweis:*  $\mathbf{NF}(g, G', Y) = 0$  bedeutet  $g \in I' = I(F) \cdot S$  als Ideal in  $S$ . Also gibt es eine Darstellung

$$g = \sum_i p_i(Y, U) f_i(Y, U) \quad (\text{G})$$

in  $S$  mit  $p_i(Y, U) \in k(U)[Y]$ . Bilden wir den Hauptnenner  $h(U)$  zu allen Nennern in allen Koeffizienten dieser Polynome, so gilt  $p_i(Y, U) = \frac{q_i(Y, U)}{h(U)}$  mit Polynomen  $q_i(Y, U) \in k[U][Y] = R$  und folglich

$$h(U) \cdot g = \sum_i q_i(Y, U) f_i(Y, U) \text{ in } R$$

und damit  $h(U) \cdot g \in I$ .  $\square$

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

Eine Darstellung (G) kann in manchen Gröbnerpaketen explizit berechnet werden, woraus sich ein Kandidat für  $h(U)$  durch Hauptnennerbildung extrahieren lässt. Allerdings folgt allein aus der Existenz von  $h \in k[U]$ , dass eine „zufällige“ Wahl von  $U_0$  „normalerweise“ zulässig ist. Genauer: Der Satz gilt auf einer nichttrivialen Zariski-offenen (und damit dichten) Teilmenge der Parameter  $U$ . Deshalb auch „Gültigkeit im Allgemeinen“.

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

## 5.7 Eine geometrische Interpretation des Satzes über die generische Gültigkeit von Sätzen von Gleichungstyp

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  $R = k[Y, U]$  und  $V = V_S(F)$  über  $S = k(U)[Y]$  genauer studieren.

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

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

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

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

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

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

Verschiedene Ideale können dasselbe Nullstellengebilde haben, etwa gilt  $V(f) = V(f^2)$  für ein  $f \in R^4$ . Es existiert jedoch eine eindeutige Korrespondenz zwischen Nullstellenmengen (über einem algebraisch abgeschlossenen Körper) und *Radikalidealen*, siehe dazu den Korrespondenzsatz im Anhang oder in der Vorlesung *Gröbnerbasen und Anwendungen*.

Jedes solche Radikalideal  $I$  kann als Durchschnitt endlich vieler Primideale  $I = \cap P_\alpha$  dargestellt werden, was dual zur Zerlegung der Nullstellengebilde  $V = V(I)$  in irreduzible Komponenten  $V = \cup_\alpha V_\alpha$  mit  $V_\alpha = V(P_\alpha)$  ist. Eine bzgl.  $I$  unabhängige Variablenmenge  $U$  muss allerdings bzgl. einer dieser Primkomponenten, die ja größer sind als  $I$ , nicht mehr unabhängig sein. Wir erkennen das daran, ob  $P_\alpha \cap N = \emptyset$  gilt (in diesem Fall bleibt die Variablenmenge unabhängig) oder nicht (in diesem Fall gibt es eine algebraische Beziehung  $h_\alpha(U) \in P_\alpha$  zwischen den eigentlich unabhängigen Variablen  $U$ ). Komponenten der ersten Art nennen wir *generisch*, Komponenten der zweiten Art *speziell*. Können wir das Nullstellengebilde der Voraussetzungen eines Geometrietheorems in seine Komponenten zerlegen, so entsprechen letztere gewissen, durch  $h_\alpha(U) = 0$  beschriebenen, degenerierten Situationen.

<sup>4</sup>Ideale entsprechen in diesem Sinne Nullstellen „mit Vielfachheiten“.

**Satz 32**

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

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

(2)

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

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

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

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

$Q_\alpha$  ist ein Primideal: Aus

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

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

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

**Folgerung 1**

$$I' \cap R = \cap \{P_\alpha : P_\alpha \cap N = \emptyset\}$$

Ist insbesondere  $F \subset R$  eine Menge von Polynomen,  $G' = \mathbf{GBasis}(F, Y)$  und  $g \in R$  ein Polynom, so gilt  $\mathbf{NF}(g, G', Y) = 0$  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  $U$  generisch richtig ist, was folgendes heißt:

**Satz 33 (Generische Gültigkeit geometrischer Sätze vom Gleichungstyp)**

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

Sei weiter  $G' = \mathbf{GBasis}(F, Y)$  eine über  $S$  berechnete Gröbnerbasis.

Gilt  $\mathbf{NF}(g, G', Y) = 0$ , 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  $U$  nicht mehr unabhängig sind. Jede solche Komponente enthält im definierenden Ideal ein Polynom  $h_\alpha(U)$ . Das Produkt dieser Polynome können wir als Nichtdegenerationsbedingung nehmen.

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

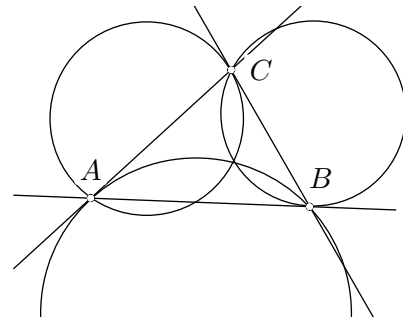
Generell, wenn  $\text{NF}(g, G', Y)$  nicht verschwindet, wie oben im Beispiel *Miquelscher Punkt*, liefern uns Untersuchungen in  $S$  Aussagen, die das Verhalten auf speziellen Komponenten ausblenden. In diesem Beispiel hat  $V(I)$  zwei generische Komponenten. Eine entspricht dem Schnittpunkt  $R$  der Kreise durch  $A, Q, R$  und  $B, P, R$ , die andere dem gesuchten Schnittpunkt  $S$ . Auf der ersteren Komponente gilt der Satz nicht, auf der zweiten sehr wohl. Der Gröbnerfaktorisierer zeigt, dass  $V(I)$  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  $S$  nicht mehr eindeutig bestimmt, so dass wir auch nicht prüfen können, ob er auf dem dritten Kreis liegt.

## 5.8 Weitere Beispiele

Betrachten wir als weiteres Beispiel den **Satz vom Brocardschen Punkt**.

**Satz 34** *Zum Dreieck  $ABC$  betrachten wir den Kreis durch  $A$ , der  $BC$  in  $C$  berührt, den Kreis durch  $B$ , der  $AC$  in  $A$  berührt und den Kreis durch  $C$ , der  $AB$  in  $B$  berührt. Diese drei Kreise gehen durch einen gemeinsamen Punkt, den Brocardschen Punkt.*

Wir beginnen mit einem speziellen Koordinatensystem, in dem die Eckpunkte  $A, B, C$  des Dreiecks sowie weitere freie Punkte  $M_1, M_2, M_3$  und  $P$  gegeben sind, die durch die AGV als Mittelpunkte der drei Berührkreise  $c_1, c_2, c_3$  sowie als Schnittpunkt von  $c_1$  und  $c_2$  spezifiziert werden.



```
A:Point(0,0); B:Point(1,0); C:Point(cx,cy); P:Point(px,py);
M1:Point(m1x,m1y); M2:Point(m2x,m2y); M3:Point(m3x,m3y);
c1:pc_circle(M1,A); c2:pc_circle(M2,B); c3:pc_circle(M3,C);
```

Die AGV  $F = \{f_1, \dots, f_8\}$  bestehen aus 8 Polynomen in 10 Variablen, wovon drei die Berührbedingungen und 5 die Lage der Punkte  $A, B, C, P$  auf den jeweiligen Kreisen kodieren.

```
polys1:[
  is_c1_tangent(c1,pp_line(A,C)),
  is_c1_tangent(c2,pp_line(A,B)),
  is_c1_tangent(c3,pp_line(B,C))
];

polys2:[
  on_circle(B,c1), on_circle(C,c2), on_circle(A,c3),
  on_circle(P,c1), on_circle(P,c2)
];
```

Als Behauptung  $g$  ergibt sich

```
g:= on_circle(P,c3);
```

$$-c_x^2 + 2 m_{3x} c_x - c_y^2 + 2 m_{3y} c_y + p_x^2 - 2 m_{3x} p_x + -p_y^2 - 2 m_{3y} p_y$$

Mit Blick auf die 8 Gleichungen setzen wir  $U = (c_x, c_y)$  und die restlichen acht Variablen als abhängige Variablenmenge  $Y$ .

```
yvars:[px,py,m1y,m2y,m3y,m1x,m2x,m3x]; uvars:[cx,cy];
G:GBasis(append(polys1,polys2),yvars);
NF(g,G,yvars);
```

Die Rechnung crasht mit Maxima. Leider reduzierte  $g$  auch im Erfolgsfall nicht zu null, was aber auch nicht überraschend ist, da – wie im Satz vom Miquelschen Punkt – neben der intendierten Lösung für  $P$  auch noch  $P = B$  Nullstelle des Polynomsystems  $F$  ist.

Eine genauere Analyse zeigt außerdem, dass die drei Berührbedingungen  $polys_1$  vollständige Quadrate sind, also unnötig komplizierte Polynome ergeben. Ist allgemein die Gerade  $l = RS$  die Tangente an einen Kreis  $c$  mit dem Berührungspunkt  $R$ , so hat auch der Berührungspunkt  $R'$  der zweiten Tangente von  $S$  an den Kreis  $c$  rationale Koordinaten, so dass die quadratische Gleichung, aus welcher wir die Berührbedingung als deren Diskriminante gewonnen hatten, zwei *rationale* Lösungen hat. Damit muss die Diskriminante aber ein vollständiges Quadrat sein.

```
M:Point(0,0); R:Point(rx,ry); S:Point(sx,sy);
h:is_cl_tangent(pc_circle(M,R),pp_line(R,S));
factor(h);
```

$$4 (r_x^2 - s_x r_x + r_y^2 - s_y r_y)^2$$

Wir können die drei Berührbedingungen durch die folgenden einfacheren Bedingungen ersetzen, dass der Berühr-Radius senkrecht auf der Tangenten steht.

```
polys1b:[
  is_orthogonal(pp_line(M1,A),pp_line(A,C)),
  is_orthogonal(pp_line(M2,B),pp_line(B,A)),
  is_orthogonal(pp_line(M3,C),pp_line(C,B))
];
```

Die Gröbnerbasis dieses Systems

```
G:GBasis(append(polys1b,polys2),yvars);
```

enthält ein Polynom  $p_y^2 - q(U) p_y$ , das sich wieder in zwei Linearfaktoren zerlegen lässt. Die folgende Rechnung zeigt, dass  $V_S(G)$  damit wieder in zwei Komponenten zerfällt. Auf der einen, im Fall  $P = B$ , gilt der Satz nicht, auf der anderen ist er gültig.



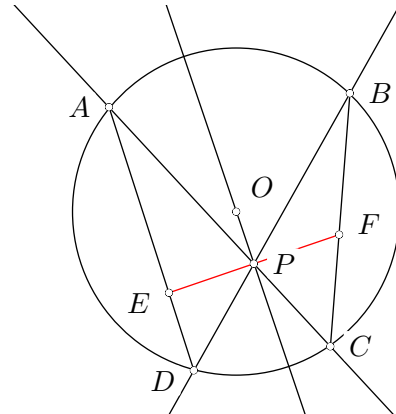
```
sol:solve(G,yvars);
map(lambda([u],ratsimp(ev(g,u))),sol);
```

$$[c_x^2 - 2c_x + c_y^2 + 1, 0]$$

*Aufgabe:* Geben Sie einen konstruktiven Beweis dieses Satzes.

Betrachten wir zum Abschluss dieses Punktes einen weiteren geometrischen Satz, den wir als Satz vom Gleichungstyp formulieren können. Es handelt sich um das **Schmetterlingstheorem**, vgl. [2, S. 31]:

**Satz 35** *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  $E$  und  $F$ . Dann gilt  $|PE| = |PF|$ .*



Zur Formulierung als geometrischer Satz vom Gleichungstyp sei  $O$  der Koordinatenursprung und  $A = (1, 0)$  auf der  $x$ -Achse gelegen.  $B, C$  und  $D$  seien drei weitere Punkte mit den universellen Realisierungen  $B = (b_x, b_y)$ ,  $C = (c_x, c_y)$  und  $D = (d_x, d_y)$ , die zusätzlich auf dem Kreis  $c$  um  $O$  durch  $A$  liegen sollen. Dies ergibt drei Bestimmungsgleichungen, so dass wir  $b_y, c_y, d_y \in Y$  und  $b_x, c_x, d_x \in U$  wählen.  $P = (p_x, p_y)$  ist der Schnittpunkt von  $AC$  und  $BD$ , was zwei Bestimmungsgleichungen für  $p_x, p_y$  ergibt, so dass wir  $p_x, p_y \in Y$  wählen. Für die Punkte  $F = (f_x, f_y)$  und  $G = (g_x, g_y)$  ergeben sich ebenfalls je zwei Schnittbedingungen, so dass wir ebenfalls  $f_x, f_y, g_x, g_y \in Y$  setzen.

Insgesamt erhalten wir folgende algebraische Formulierung

```
O:Point(0,0); A:Point(1,0); c:pc_circle(O,A);
B:Point(bx,by); C:Point(cx,cy); D:Point(dx,dy);
P:Point(px,py); E:Point(ex,ey); F:Point(fx,fy);
l:ortho_line(P,pp_line(O,P));
```

```
polys: [
  on_circle(B,c), on_circle(C,c), on_circle(D,c),
  is_collinear(B,D,P), is_collinear(A,C,P),
  on_line(E,l), is_collinear(A,D,E),
  on_line(F,l), is_collinear(B,C,F)
];
```

```
g:sqrdist(P,E)-sqrdist(P,F);
```

Das Gleichungssystem hat eine recht überschaubare Struktur

$$\begin{aligned} polys = [ & b_x^2 + b_y^2 - 1, c_x^2 + c_y^2 - 1, d_x^2 + d_y^2 - 1, \\ & b_x d_y - b_y d_x - b_x p_y + b_y p_x + d_x p_y - d_y p_x, \\ & c_y - p_y + c_x p_y - c_y p_x, p_x^2 - e_x p_x + p_y^2 - e_y p_y, \\ & d_y - e_y + d_x e_y - d_y e_x, p_x^2 - f_x p_x + p_y^2 - f_y p_y, \\ & b_x c_y - b_y c_x - b_x f_y + b_y f_x + c_x f_y - c_y f_x ] \\ \\ g = & e_x^2 - 2 p_x e_x + e_y^2 - 2 p_y e_y - f_x^2 + 2 p_x f_x - f_y^2 + 2 p_y f_y \end{aligned}$$

Maximas GBasis-Algorithmus kommt damit aber nicht mehr zurecht.

```
yvars:=[ex,ey,fx,fy,px,py,by,cy,dy]; uvars:=[bx,cx,dx];
G:=GBasis(polys,yvars);
```

Mit einer stärker konstruktiv formulierten AGK können einige der Polynome aus der AGV eliminiert werden. Wählen wir etwa  $E$  und  $F$  als Geradengleiter und  $P$  als Schnittpunkt der Geraden  $AC$  und  $BD$ , so benötigen wir nur 5 Gleichungen, um das Problem zu formulieren

```
O:Point(0,0); A:Point(1,0); c:pc_circle(O,A);
B:Point(bx,by); C:Point(cx,cy); D:Point(dx,dy);
E:varpoint(A,D,es); F:varpoint(B,C,fs);
P:intersection_point(pp_line(A,C), pp_line(B,D));
h:ortho_line(P,pp_line(O,P));

polys:[
  on_line(E,h), on_line(F,h),
  on_circle(B,c), on_circle(C,c), on_circle(D,c)
];

g:num(ratsimp(sqrdist(P,E)-sqrdist(P,F)));

yvars:[es,fs,by,cy,dy]; uvars:[bx,cx,dx];
G:GBasis(polys,yvars);
```

Auch diese Rechnung geht mit Maxima leider nicht durch.

*Aufgabe:* Geben Sie ein konstruktives Beweisschema für diesen Satz an, indem Sie  $B, C, D$  als Kreisgleiter ansetzen, und untersuchen Sie, ob dieser Beweis durchgeht.

## 6 Anhang: Gröbnerbasen und deren Eigenschaften

### 6.1 Ideale und Nullstellenmengen.

#### Korrespondenzsatz und Zerlegungssatz

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

**Definition 14** Als von den Polynomen  $f \in F$  erzeugtes Ideal  $I(F)$  bezeichnet man die Menge aller polynomialen Linearkombinationen

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

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  $I(F)$  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

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

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  $\text{rad}(I)$  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 36 (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 37 (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 Idealdurchschnitten 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 15** 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  $f g$  auf ganz  $V$ , so muss bereits einer der Faktoren auf  $V$  verschwinden. In der Tat, wegen

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

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

**Satz 38** 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

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

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

## 6.2 Idealtest, Normalformen und Gröbnerbasen

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

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

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

- den Leitterm  $lt(f) := X^{\alpha_0}$ ,
- 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 Leiterterme  $lt(b), b \in B$ , 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 39** 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.

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

**Definition 16** 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  $f \in I(B)$ , 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  $G = \mathbf{GBasis}(B)$  übergeht. Gröbnerbasen zeichnen sich durch eine Reihe zueinander äquivalenter Eigenschaften aus:

**Satz 40 (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 Faktorrings  $R/I$ , 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.

### 6.3 Eliminationsideale und Gröbnerbasen

An verschiedenen Stellen sind Polynome in einem Ideal  $I$  zu finden, die gewissen Variablen nicht enthalten oder zu zeigen, dass es keine derartigen nichttrivialen Polynome gibt. Derartige Fragen führen auf ein Eliminationsproblem:

Gegeben ist ein Polynomring  $R = k[Y, U]$  und ein Ideal  $I \subset R$ . Gesucht sind alle Polynome in  $I$ , die keine der Variablen  $x \in Y$  enthalten, also die Menge

$$I' = I \cap k[U].$$

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

Basen für Eliminationsideale können aus Gröbnerbasen für spezielle Eliminationstermordnungen abgelesen werden. Als *Eliminationsordnung* für  $U$  bezeichnet man jede Termordnung, für welche

$$t_1 \in T(U), t_2 \in T(Y \cup U) \setminus T(U) \Rightarrow t_1 < t_2$$

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

#### Satz 41 (Eliminationssatz für Gröbnerbasen)

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

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

eine (min. reduzierte) Gröbnerbasis des Eliminationsideals  $I' = I(F) \cap k[U]$ .

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

Ist  $I$  ein Primideal, so haben alle maximalen unabhängigen Variablenmengen die gleiche Kardinalität. Diese Anzahl (der unabhängigen Parameter) bezeichnet man auch als die *Dimension*

des Ideals  $I$  bzw. der Nullstellenmenge  $V(I)$ . Ist  $U$  eine solche maximale unabhängige Variablenmenge und  $x_i$  eine weitere Variable, so hängt  $x_i$  über  $I$  algebraisch von den Variablen  $U$  ab, d. h. für alle  $x_i$  existieren Polynome

$$f_i(x_i, U) = \sum_j c_{ij}(U)x_i^j \in I$$

$x_i$  kann man dann zwar im Allgemeinen nicht rational durch die unabhängigen Parameter ausdrücken, jedoch ergibt sich für fast alle Parameterbelegungen  $U_0$  (genauer: solche, für die es ein  $j$  mit  $c_{ij}(U_0) \neq 0$  gibt) eine Bestimmungsgleichung für  $x_i$  und damit nur endlich viele zulässige Parameterspezifikationen  $X_0 = (Y_0, U_0)$  für eine vorgegebene Spezifikation  $U_0$ . Wenn wir  $U_0$  so wählen, dass sogar  $lc(f_i)(U_0) \neq 0$ , also  $\deg(f_i(Y, U)) = \deg(f_i(Y, U_0))$  für alle  $i$  gilt, lässt sich auch die Anzahl dieser Lösungen voraussagen.

### 6.4 Nichtdegenerationsbedingungen und Idealquotienten

Sei  $X = (Y, U)$  eine Aufspaltung der Variablen und  $U$  eine maximale unabhängige bzgl.  $I = I(F)$  Teilmenge von  $X$ . Im Allgemeinen ist es schwierig, eine Zerlegung  $I = \cap P_\alpha$  zu bestimmen und die speziellen Komponenten, also solche mit  $P_\alpha \cap k[u] \neq (0)$ , zu erkennen. Außerdem kann eine solche Komponente verschiedene  $h_\alpha(U)$  enthalten, so dass die Frage steht, ob es so etwas wie eine „einfachste“ Nichtdegenerationsbedingung gibt und wie sich diese ggf. bestimmen lässt. Wir wollen dazu die Menge aller Degenerationsbedingungen  $h \in S$ , den *Degenerationslocus*  $DL(F/g)$  des Geometrietheorems ( $F \Rightarrow g$ ), näher beschreiben.

Aus der Äquivalenz

$$(F = 0, h \neq 0 \Rightarrow g = 0) \Leftrightarrow (F = 0 \Rightarrow g \cdot h = 0)$$

und dem Hilbertschen Nullstellensatz folgt zunächst für  $I = I(F)$ , dass

$$DL(F/g) = \{h \in S : \exists n (gh)^n \in I\}$$

selbst wieder ein Ideal ist. Dieses Ideal ist eng mit dem stabilen Idealquotienten bzgl.  $g$  verbunden:

$$DL(F/g) = \{h \in S : \exists n h^n \in I : g^\infty\} = \text{rad}(I : g^\infty).$$

Ist  $I$  selbst ein Radikalideal, so gilt sogar

$$DL(F/g) = I : g^\infty.$$

**Satz 42** *Der Degenerationslocus  $DL(F/g)$  eines geometrischen Satzes ( $F \Rightarrow g$ ) stimmt mit dem Radikal des stabilen Idealquotienten  $J = I(F) : g^\infty$  überein.*

Der Degenerationslocus muss dabei nicht unbedingt eine Hyperfläche ( $h = 0$ ) sein, sondern kann kleinere Dimension haben. Wegen  $V(h) \supset V(J)$  für  $h \in J$  kann jedes Element aus  $J$  als Nichtdegenerationsbedingung verwendet werden. Insbesondere gibt es keine „minimale“ Nichtdegenerationsbedingung, wenn das Ideal  $J$  kein Hauptideal ist.

Ist  $I = \cap P_\alpha$  die Zerlegung des Radikalideals  $I$  in Primkomponenten, so gilt  $I : g^\infty = \cap (P_\alpha : g^\infty)$  und

$$P_\alpha : g^\infty = \begin{cases} (1) & \text{für } g \in P_\alpha \\ P_\alpha & \text{sonst} \end{cases}$$



Folglich besteht der Degenerationslokus aus genau den Komponenten von  $I$ , auf denen  $g$  nicht verschwindet.

Als *Degenerationslocus im engeren Sinne* bzgl. einer maximalen Teilmenge unabhängiger Variablen  $U$  bezeichnet man das Eliminationsideal

$$DL_U(F/g) = DL(F/g) \cap k[U].$$

Alle anderen Polynome in  $DL(F/g)$  beschreiben, welche Bedingungen an die abhängigen Variablen in einem solchen degenerierten Fall noch bestehen bleiben. Jedes Polynom  $h(U) \in DL_U(F/g)$  kann als Degenerationsbedingung für den (dann richtigen) Satz  $(F/h \Rightarrow g)$  verwendet werden.

Der stabile Idealquotient  $I(F/g)$  lässt sich als Eliminationsideal aus  $I(F(X)) + Id(1 - t \cdot g(X)) \cap k[X]$  berechnen, wobei  $t$  eine neue Variable ist.

Allerdings werden mit diesem Zugang nicht alle *geometrischen* Nichtdegenerationsbedingungen gefunden, sondern nur solche, unter denen die *algebraische* Aussage nicht mehr gilt.

## 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.

## Anhang: Aufgaben

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

1. Beweisen Sie den Satz vom Höhenfußpunktdreieck: Ist  $\triangle ABC$  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. (Chou, 133) Zeigen Sie allgemeiner: Sind  $M, N$  Punkte auf den Seiten  $AC, AB$  eines Dreiecks  $ABC$ , so dass sich die Geraden  $BM, CN$  auf der Höhe  $AD$  ( $D$  ist der Höhenfußpunkt) dieses Dreiecks schneiden, so ist  $AD$  die Winkelhalbierende des Winkels  $\angle MDN$ .
3. 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.
4. Beweisen Sie die Umkehrung des Satzes von der Simsonschen Geraden: Sind die Lotfußpunkte  $P, Q, R$  von einem Punkt  $D$  auf die Seiten des Dreiecks  $\triangle ABC$  oder deren Verlängerungen kollinear, so liegt  $D$  auf dem Umkreis des Dreiecks.
5. Beweisen Sie auch folgende Verallgemeinerung:  
Ist  $M$  der Umkreismittelpunkt des Dreiecks  $\triangle ABC$  und  $D, P, Q, R$  wie oben, so hängt der Flächeninhalt  $F(\triangle PQR)$  nur von der Länge der Strecke  $MP$  ab.  
Finden Sie eine genaue Formel für diesen Flächeninhalt.
6. Beweisen Sie die Flächeninhaltsformel

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

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

7. Beweisen Sie die Flächeninhaltsformel

$$F(\triangle ABC) = \rho \cdot \frac{a+b+c}{2},$$

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?

8. (Chou, 115) 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.
9. Beweisen Sie, dass die Lote durch die drei Ankreiszentren auf die jeweilige Dreiecksseite durch einen gemeinsamen Punkt gehen.

10. (Chou, 315) Die Sehnen  $PA$ ,  $PB$ ,  $PC$  eines gegebenen Kreises sind die Durchmesser der drei Kreise  $c_{PA}$ ,  $c_{PB}$ ,  $c_{PC}$ . Der Kreis  $c_{PA}$  schneidet die Geraden  $PB$ ,  $PC$  in den Punkten  $D$ ,  $G$ , der Kreis  $c_{PB}$  schneidet  $PA$ ,  $PC$  in  $E$ ,  $H$ , der Kreis  $c_{PC}$  schneidet  $PA$ ,  $PB$  in  $F$ ,  $I$ .

Zeigen Sie, dass die drei Geraden  $DG$ ,  $EH$ ,  $FI$  konkurrent sind.

11. (Chou, 99)  $A$  and  $B$  seien die Schnittpunkte zweier Kreise  $c_O$  und  $c_P$ . Durch  $A$  verläuft eine Gerade, welche die Kreise in  $C$  and  $D$  schneidet.  $G$  ist der Mittelpunkt der Strecke  $CD$ .  $BG$  schneidet die Kreise  $c_O$  und  $c_P$  in  $E$  and  $F$ .

Zeigen Sie, dass  $G$  der Mittelpunkt von  $F$  und  $E$  ist.

12. (Chou, 37)  $PT$  and  $PB$  seien die beiden Tangenten von  $P$  an einen Kreis,  $T$  und  $B$  resp. die Berührungspunkte.  $AB$  sei der Durchmesser dieses Kreises durch  $B$ .

Zeigen Sie, dass  $AP$  das Lot aus  $T$  auf  $AB$  halbiert.

13. (Chou 28) In einem Sechseck  $ABCDEF$  sind die Geraden  $CF$ ,  $AB$ ,  $DE$  sowie die Geraden  $BE$ ,  $CD$ ,  $AF$  jeweils konkurrent.

Zeigen Sie, dass dann die Geraden  $AD$ ,  $EF$ ,  $BC$  ebenfalls konkurrent sind.

14. (BWM 2017) In einem konvexen Tangentenviereck  $A_1A_2A_3A_4$  sei  $M$  der Mittelpunkt des Inkreises, der die Seiten des Vierecks berührt. Weiter sei  $g_1$  die Gerade durch  $A$ , die senkrecht auf  $A_1M$  steht und entsprechend  $g_2, g_3, g_4$ . Die Geraden  $g_1, \dots, g_4$  bestimmen ein weiteres Viereck  $B_1B_2B_3B_4$ , wobei  $B_1$  der Schnittpunkt von  $g_1$  und  $g_2$  ist; entsprechend bezeichnet  $B_2, B_3$  und  $B_4$  den Schnittpunkt von  $g_2$  und  $g_3$ ,  $g_3$  und  $g_4$  sowie  $g_4$  und  $g_1$ . Beweisen Sie, dass sich die Diagonalen des Vierecks  $B_1B_2B_3B_4$  im Punkt  $M$  schneiden.