Friedrich-Alexander-Universität
Erlangen-Nürnberg

DISSERTATION

2011 **Thomas Rainer Werner**

Rational families of circles and bicircular quartics

Rationale Kreisscharen und bizirkulare Quartiken

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur
Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von
Thomas Rainer Werner
aus Lichtenfels

Als Dissertation genehmigt von
der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 18. Juli 2012

Vorsitzender der Prüfungskommission: Prof. Dr. Rainer Fink

Erstberichterstatter: Prof. Dr. Wolf P. Barth

Zweitberichterstatter: Prof. Dr. Wulf-Dieter Geyer

Abstract

This dissertation deals with special plane algebraic curves, with so called bicircular quartics. These are curves of degree four that have singularities in the circular points at infinity of the complex projective plane $\mathbb{P}_2(\mathbb{C})$. The main focus lies on real curves, i.e. such curves that are invariant under complex conjugation. Many of the statements on bicircular quartics presented in this work are well known since the end of the 19$^{\text{th}}$ century, but the way of proving at that time did not fully employ the language of the then well developed projective geometry.

The primary goal of this text is the formulation of the classical statements on bicircular quartics in modern language. In order to achieve this the theoretical framework is built beginning with the space of circles in the language of projective geometry. Within the space of circles are discussed at first linear and then quadratic families of circles. In the following the theorems on bicircular quartics and their degenerate form, the circular cubics, are proved by means of geometrical statements on the mentioned families of circles in the projective space of circles. An ancillary goal of this work is the provision of tools that facilitate an easy depiction of bicircular quartics and rational families of circles with the help of a computer.

Zusammenfassung

Diese Dissertation beschäftigt sich mit speziellen ebenen algebraischen Kurven, mit sogenannten bizirkularen Quartiken. Das sind Kurven vierten Grades, die Singularitäten in den unendlich fernen Kreispunkten der komplex-projektiven Ebene $\mathbb{P}_2(\mathbb{C})$ besitzen. Das Hauptaugenmerk liegt dabei auf reellen Kurven, d.h. solchen Kurven, die unter der komplexen Konjugation invariant sind. Viele der in dieser Arbeit vorgestellten Aussagen über bizirkulare Quartiken sind bereits seit Ende des 19. Jahrhunderts wohl bekannt, die Beweisführung von damals nutzte jedoch nicht konsequent die Sprache der seinerzeit bereits entwickelten projektiven Geometrie aus.

Das primäre Ziel dieser Arbeit ist die Formulierung der klassischen Aussagen über bizirkulare Kurven in moderner Sprache. Zu diesem Zwecke wird das theoretische Gerüst beginnend beim Raum der Kreise in der Sprache der projektiven Geometrie aufgebaut. Im Raum der Kreise werden zunächst lineare, dann quadratische Kreisscharen diskutiert. Im Folgenden werden die Theoreme über bizirkulare Quartiken und über ihre Entartungsform, die zirkularen Kubiken, mit der Hilfe von geometrischen Aussagen über die oben genannten Kreisscharen im projektiven Raum der Kreise bewiesen. Ein untergeordnetes Ziel dieser Arbeit ist die Bereitstellung von Werkzeugen, die eine einfache Darstellung von bizirkularen Kurven und von rationalen Kreisscharen am Rechner ermöglichen.

Acknowledgement

I thank my advisor Mr. Prof. Dr. Wolf Barth for his numerous precious suggestions and hints before and during the development of this text. Moreover I thank him for his patience and for the opportunity to work under his leadership as a scientific assistant.

Above all I thank my wife Šárka for her love, her understanding and her support.

Danksagung

Ich bedanke mich bei meinem Betreuer Herrn Prof. Dr. Wolf Barth für seine zahlreichen wertvollen Ratschläge und Hinweise vor und während der Entstehung dieses Textes. Weiterhin danke ich ihm für seine Geduld und dafür, dass ich unter seiner Leitung als wissenschaftliche Hilfskraft tätig sein konnte.

Vor allem bedanke ich mich bei meiner Ehefrau Šárka für ihre Liebe, ihr Verständnis und ihre Unterstützung.

Contents

1 Introduction

1.1 Main results

Projective space of circles

Circles and lines are in general treated as different geometric objects. The distiction between the cases $\nu = 1$ (circle) and $\nu = 0$ (line) in equations of the form

$$\nu(x^2 + y^2) - 2\xi x - 2\eta y + \zeta = 0$$

can be found in many sources dealing with circles and with the inversion map, because circles can be mapped onto lines and vice versa.

The unification of circles and straight lines has been adumbrated, for example, in Pedoe's book [44, p. 29] on circles by the introduction of a parameter λ such that

$$(x^2 + y^2) - 2\left(\frac{\xi + \lambda\xi'}{1+\lambda}\right) x - 2\left(\frac{\xi + \eta\xi'}{1+\lambda}\right) y + \left(\frac{\zeta + \lambda\zeta'}{1+\lambda}\right) = 0,$$

where the forbidden value $\lambda = -1$ stands for the straight line hidden in this family of circles. But he does not write this equation explicitly as the linear combination of two circles

$$(1+\lambda)\,(x^2 + y^2) - 2\,(\xi + \lambda\xi')\,x - 2\,(\xi + \eta\xi')\,y + (\zeta + \lambda\zeta') = 0,$$

where the exceptional case is possible.

The first result of this paper is the description of the space of circles with the means of projective geometry, i.e. the extension from the affine space of circles $\mathfrak{Circ}$ to the projective space of circles $\mathbb{P}(\mathfrak{Circ})$. Straight lines belong to the projective space of circles in a natural way. One of the benefits of this unification is the easy description of inversions as linear maps on $\mathbb{P}(\mathfrak{Circ})$ and in particular the description of the inversion about the unit circle as a transposition of two coordinate axes.

Solution of the general equation of degree four

It is well known that it is possible to solve the general quartic equation in one variable by radicals. This can be done by solving appropriate equations of lesser degree in order to construct field extensions corresponding to the subnormal series

$$S_4 \triangleright V_4 \triangleright S_2 \triangleright 1$$

of the Galois group S_4 of the polynomial of fourth degree that is to be solved.

The second result is the solution of the general equation of degree four by applying the theory of bicircular quartics. It is shown that the results developed for bicircular quartics in dimension two can be generalized into other dimensions including dimension one. It is possible to obtain a straightforward procedure for solving this equation. An important observation is that the resultant of degree three, which is needed for the solution of the equation, comes directly from the equation determining the (one-dimensional) inversions that leave the given quartic polynomial invariant.

Geometric and computational results

A result of this work concerning the geometry of the space of bicircular quartics is the description of the action of inversions on bicircular quartics as an induced action of the inversion on the space of circles. This induced action can be easily explained with the three circles form of bicircular quartics.

Another result is the reformulation of the framework describing circles, the space of circles and families of circles from an entirely projective point of view. Many parts of the description can be used directly for computation, for example for the determination of the two inversions that map two given circles onto each other. The main advantage of this description is that one does no longer have to take care of exceptional cases, where circles are mixed up with straight lines.

Finally a complete list of possible configurations of singularities of irreducible and reducible bicircular quartic curves is given. It is shown how singular points can be identified in the C-Q-form and in the three circles form of a bicircular quartic. Some well known curves of this type are presented and discussed.

1.2 Motivation

Quadratic families of circles

From Pappus of Alexandria we know that the Greek mathematician Apollonius of Perga gave a general solution for the task that is nowadays known as the

Definition 1.1 (Problem of Apollonius)
Let $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ be three given circles. Find all circles that are touching these given circles.

The original problem can be split up into two steps. The first step consists in finding all circles that are tangent to two given circles. In the second step only those circles that are touching the third given circle are picked out from the intermediate solution.

The first step is worth to be treated as a problem itself. It turns out that in general the circles tangent to two given circles $\mathcal{C}_1$ and $\mathcal{C}_2$ are divided into two families. Each of these families corresponds to one of the two circles of inversion that map $\mathcal{C}_1$ and $\mathcal{C}_2$ onto each other. This correspondence is established by the fact that all circles of one family are orthogonal to the corresponding circle of inversion. The locus of their centers is a conic section. This property has been used by Adriaan van Roomen (Adrianus Romanus) to find a solution to the Problem of Apollonius by intersecting two hyperbolas.

The roles in the last problem can be reversed. The circles $\mathcal{C}_1$ and $\mathcal{C}_2$ then become the envelope of a family of cicles whose centers lie on a given conic section and are orthogonal to a given circle. One might want to examine how variation of the orthogonal circle changes the envelope. The pair of circles, i.e. the envelope of the family of circles, then generally deforms into a bicircular quartic. In this more general case we do not have two but four families of circles having the same envelope. These families can be represented by a conic section in the space of circles and are therefore called *rational families of circles.*

The doubling of the cube and the angle trisection are classical problems. They can not be solved by straightedge and compass alone, but a solution is possible by using special curves. The first one, the so called Delian problem, can be successfully attacked with the utilization of the Cissoid of Diocles, a certain circular cubic. The second problem can be solved by using a so called trisectrix, like the Trisectrix of Maclaurin or the Limaçon trisectrix. The first of these curves is a circular cubic, while the second one is a bicircular quartic.

Inversive geometry

The concept of inversion about a circle was introduced in the first half of the 19^{th} century. It is usually defined as a map on $\mathbb{R}^2$. This leads to several problems, for example that the center of the circle of inversion does not have an image point in $\mathbb{R}^2$. One can overcome this limitation by extending the real plane by an infinitely distant point.

It is possible to describe inversions of the real plane by maps on $\mathbb{C}$. They are of the form

$$z \mapsto \frac{a\overline{z} + b}{c\overline{z} + d}$$

with restrictions on the parameters a, b, c, d. We will see that a general inversion can be reduced to the inversion $z \mapsto \frac{1}{\overline{z}}$ about the unit circle. Therefore an inversion is of the form

$$z \mapsto \frac{\beta\overline{z} + \alpha\overline{\alpha} - \beta\overline{\beta}}{\overline{z} - \overline{\beta}}.$$

α stands for the radius of the circle of inversion and β for the location of its center.

Inversions on $\mathbb{R}^2$ are somehow dissatisfying. Firstly the way how the image points are constructed is not a truly real construction. A distinction has to be made, whether the original point lies inside or outside the circle of inversion. Secondly the identification $\mathbb{R}^2 \cong \mathbb{C}$ introduces complex conjugation in order describe the way how inversions act on the real plane. Hence the describing maps on $\mathbb{C}$ are anticonformal and supersede the natural inversions on $\mathbb{C}$, where the inversion about the unit circle is $z \mapsto \frac{1}{z}$. Thirdly it simply makes no sense to regard inversions as maps on $\mathbb{R}^2$. They have to be treated projectively and, above all, they do not map points on points, but circles on circles. The last argument becomes obvious when one recalls that Möbius transformations have exactly this property of mapping lines and circles on lines and circles. Within this context inversions in $\mathbb{P}(\mathbb{C})$ are only special Möbius transformations.

Inversions can be defined in any dimension. They are usually introduced as maps in two dimensions. The key to understanding inversions is the fact that the are not mapping points onto points, but circles onto circles. Inversions act on something that is called the *space of circles*. Points are treated as circles of radius zero and lines are contained in the projective closure of this space. The inversion about a given circle maps its center onto the line at infinity and vice versa. In the same way it is possible to define a space of hyperspheres in any dimension. The corresponding inversions map hyperplanes and hyperspheres onto hyperplanes and hyperspheres, i.e. generalized hyperspheres onto generalized hyperspheres.

Systematic aspects

The image under inversion of a generalized circle is again a generalized circle. This allows us to say that the set of generalized circles is closed under inversion. The simplest curves not contained in this set are the conic sections. So which set of curves is closed under inversion and contains all conic sections? We will see that this is the set of rational bicircular quartics. It also includes all rational circular cubics and all curves of first and second order. Rational bicircular quartics form an important class of curves. Members of it have already been studied by ancient mathematicians, for example the hippopedes, a special case of spiric sections. All these rational curves are contained in the set of bicircular quartics that is also closed under inversion. A general spiric section is a special case of this kind of curve.

We can sum up the above discussion with the statement that the simplest sets of curves that are closed under inversion are:

- the set of nullcircles (circles with radius zero),
- the set of generalized circles,
- the set of rational bicircular quartics,
- the set of bicircular quartics.

We will see that nullcircles are circles with a finite singularity and rational bicircular quartics are bicircular quartics with a finite singularity.

The concept of bicircular quartics can be defined in arbitrary dimension, because the algebraic equation of a bicircular quartic is built from the equations of a circle, a line and a conic section. These are elementary objects that exist in every dimension $d \in \mathbb{N}$. Thus we can build up a similar equation from hyperspheres, hyperplanes and quartics. The resulting hypersurface may be called a bispherical (d-dimensional) quartic. It is possible to define the action of an inversion about a hypersphere on these objects. Many properties transpose easily between dimensions. The concept makes sense even in dimension $d = 1$.

On the other hand bicircular quartics can be seen as something very similar to the union of two circles. Their definition can be generalized to that of n-circular $2n$-tics in a straightforward way. Circles fit into this concept as circular quadric curves. The two aspects d-dimensionality and n-circularity can be combined with ease. Multicircular hyperspheres can be seen as universal geometric objects, just like hyperplanes, hyperspheres and quadric hypersurfaces.

1.3 History and sources

Ancient mathematics

The earliest known appearance of a bicircular quartic is the *hippopede* in a work of Eudoxus of Cnidus (408-355 BC) about spherical astronomy. This curve is not the plane curve with the same name but a spherical curve that has the same figure-eight shape as a plane hippopede. More detailed information about his construction are found in [54, p. 295f].

Proclus has written that a mathematician known as Perseus considered the intersection of a torus and a plane which is parallel to the equatorial plane of the torus. These intersections are called spiric sections. They are a special case of toric sections, where the intersecting plane is no longer required to be parallel to the axis of the torus. Toric sections are bicircular quartics, hence spiric sections were the first curves of this type which were investigated. The best known spiric sections are the *Cassinian curves*, another special class are the so called *hippopedes*. The most prominent bicircular quartic is the *lemniscate of Bernoulli* which is a hippopede and a Cassinian curve at the same time. Defining curves as the intersection of a torus and a plane resembles to the fact discovered by Apollonius of Perga (≈262-190 BC), that ellipse, parabola and hyperbola arise as intersections of a cone with a plane.

Other bicircular curves were known to the ancient Greek mathematicians even before the Spiric sections that are dated around 150BC. One of them is the Conchoid of Nicomedes which was invented before 200BC. It can be used to solve the Delian problem, as is shown in [54, p. 390ff]. This curve is a circular cubic. Another such curve that is also solving the problem of doubling the cube was constructed around 200BC by Diocles. The construction of the Cissoid of Diocles together with the solution of the problem are found in [54, p. 442ff].

From the Renaissance to the Industrial Revolution

Albrecht Dürer (1471-1528) proposed in his work *Underweysung der Messung mit dem Zirkel und Richtscheyt* a construction of the *cardioid*, which he called *Spinnenkurve (spider-line)*. The name comes from the drawing where he constructs the curve which resembles the legs of a spider. The construction makes use of the fact that the cardioid is an epicycloid where a circle is rolling without slip on the outside of a circle of equal size. The cardioid is a rational bicircular quartic.

When we look at the corresponding epitrochoids, i.e. where the fixed and the rolling circle have the same radius, we obtain the *limaçons (of Pascal)*. This class of curves is named after Étienne Pascal (1588-1651), even though it seems unlikely that he was the first person ever to have studied them, since it is assumed that Eudoxus of Cnidus and Apollonius of Perga already worked on astronomical models which should describe the apparent motion of the celestial bodies. Their models were much more complicated than a simple system with just one deferent and epicycle, so at least the cardioid should be known since ancient times.

In their most general form bicircular quartics may arise as intersections of a cyclide with a plane. A torus is just a special type of cyclide, hence toric sections are curves of this type. The name *cyclide* was coined by Charles Dupin (1784-1873) who discovered them in 1803. Further investigations on cyclides were carried out by Jean Gaston Darboux (1842-1917) in 1873. However, we should be aware that defining bicircular quartics as cyclidic sections is not appropriate. Cyclides are better understood as the analogon to bicircular quartic curves in three dimensions and not vice versa. Many properties of cyclides can be found in [18]. The text also covers the figures arising from the intersection of a cyclide and a sphere. John Casey called these spherical curves *sphero-quartics*. With the radius of the intersecting sphere going to infinity these sphero-quartics become bicircular quartics.

Irish mathematics in the 19th century

George Salmon (1819-1904) features cyclides in his book *A treatise on the analytic geometry of three dimensions (1862)*. Ten years earlier, the first edition [49] of his famous *Treatise on the higher plane curves* contains only a short chapter about quartics. Bicircular quartic curves do not appear in it at all. The second edition [50] which was published in 1873 already features 13 pages dealing exclusively with this topic.

In the period between these two editions bicircular quartic curves were studied at the Trinity College in Dublin. The most elaborate text about this subject is [17]. This text *On Bicicircular Quartics* was written by John Casey (1820-1891) and appeared in the Transactions of the Royal Irish Academy in 1871. Many parallels can be found between this work and [18] which was published in the same year. The text about bicircular quartics however seems to be written before the text about cyclides, because even though they were published in the same year the publishing journals date the first one to 1867 and the second one to 1871.

Important sources

Since the inclusion of bicircular quartics in the later editions of Salmon's treatise [50] bicircular quartic curves and their degenerate form, circular cubic curves, had become a standard topic covered by textbooks about plane curves. Fields of research on this kind of curves ranged from the investigation of their mechanical generation to the identification of (geo-)metrical invariants. At the beginning of the 20$^{\text{th}}$ century they were also discussed in textbooks for electrical engineers on circuits powered by an alternating current.

The book [15] can be used as a starting point. It is a textbook on plane algebraic curves and gives an introduction into the algebraic geometry of plane curves. It does not treat bicircular quartics at all, but in some sense fills the gap between older works on plane curves and contemporary books on algebraic geometry.

The books [7] and [34] can be used as a stepping stone into the past. The notation used in these two works is more similar to the way like things were written down in the time of Salmon and Casey. Both of them feature chapters dedicated to circular cubics and bicircular quartics. The techniques mentioned in [58] may be useful. It is quite probable that the author of a 19$^{\text{th}}$ century mathematical text took this knowledge as granted.

A reader interested in the mathematical-historical aspect of these curves should refer to [41] written by Gino Loria. It features an extensive catalog of curves that have been named. But this work is more that a mere list of known curves. It contains a lot of references to even older works that are not listed in the Zentralblatt due to their age.

Basic information about circular cubics and bicircular quartics and also special cases of them be found in the sources listed below. These items are listed in chronological order and the headlines should just give a convenient short hand notation of the book titles. For complete information please refer to the corresponding entries in the bibliography.

Basset (1901): Cubic and quartic curves

[7] is a general textbook about plane curves and contains in particular

- p. 74-96: circular cubics in general, Trisectrix of Maclaurin, Logocyclic Curve[1], Cissoid of Diocles
- p. 133-161: bicircular quartics in general
- p. 162-203: special quartic curves

Loria (1902): Ebene Kurven

[41] is a book focused on curves with a name. It contains a vast number of references and covers

- p. 31-49: circular cubics, Cissoid of Diocles, generalizations of the cissoid, ophiurides
- p. 58-74: right strophoid, oblate strophoids, generalizations of the strophoid[2], conchoids of de Sluze,
- p. 81-93: trisectrix curves (t. of Maclaurin, t. of Catalan, t. of de Longchamps), duplicatrix curves
- p. 109-132: bicircular quartics in general, spiric and toric sections, Conchoid of Nicomedes
- p. 161-170: Cartesian ovals
- p. 193-199: Cassinian curves
- p. 316-343: multiplicatrix curves, mediatrix curves, sectrix curves
- p. 357-367: anallagmatic curves

[1]Logocyclic curve is only another name for the right strophoid.

[2]Loria names these curves *panstrophoids*.

Ganguli (1919): Plane curves

[34] is a textbook for post-gradual students on the general theory of plane curves and consists of two parts. The pages listed below are all from the second part which covers cubic and quartic curves.

- p. 216-228: circular cubics in general, right strophoid, oblate strophoids, trisectrix of Maclaurin, cissoid of Diocles
- p. 276-305: bicircular quartics in general
- p. 306-315: circular cubics as degenerate bicircular quartics
- p. 316-328: Cassinian curves, Cartesian ovals, lemniscate of Bernoulli, limaçons, cardioid, conchoid of Nicomedes
- p. 329-334: additional comments on circular cubics and bicircular quartics

Wieleitner (1943): Algebraische Kurven

[58] consists of two parts and shortly covers several general topics about plane curves. It explains the notation that was used at that time and already appeared in earlier works. The curves of our interest are treated in the first part of this book. The page numbers refer to the reprint from the year 1943.

- p. 57-60: circular cubics
- p. 60-64: bicurcular quartics

Fladt (1962): Analytische Geometrie spezieller ebener Kurven

[33] provides a convenient overview over the geometrical generation of bicircular quartics, their geometrical properties and cases of their degeneration. It also lists classes of these curves that have a special name.

- p. 253-262: bicircular quartics in general
- p. 262-265: rational bicircular quartics
- p. 265-270: spiric and toric sections, lemniscates of Booth[3]
- p. 270-273: Cassinian curves
- p. 273-282: Cartesian ovals
- p. 282-289: circular conchoids of Pascal

[3]These curves are also known under the name *hippopedes*.

Bartl (1979): Analytische Geometrie

[5] consists of four volumes. It was never published as a book and only a few copies of the original manuscript exist. It contains general information about algebraic curves and also an overview about curves with names. In many cases it also features geometrical and mechanical constructions of these curves.

- p. 20-27: trisection of the angle, conchoid of Nikomedes, trisectrix of Maclaurin, limaçon of Pascal
- p. 28-31: the Delian problem, cissoid of Diocles
- p. 552-559: list of special cubic curves
- p. 560-570: cissoid of Diocles, cissoid of Peano, ophiurides
- p. 573-581: right strophoid, oblique stophoids, focal curve of Quetelet, strophoid of Moivre, nephroid of Freet[4], panstrophoids
- p. 587-592: conchoids of de Sluze
- p. 605-615: trisectrix curves, trisectrix of Cramer, cubic of Sperber, trisectrix of Maclaurin,
- p. 623: circular folium
- p. 646-648: circular cubics in general[5]
- p. 687-694: bicircular quartics in general
- p. 707-711: rational and rational bicircular quartics
- p. 737-749: list of special quartic curves
- p. 800: bicircular quartic curves
- p. 824-832: inverse curves of conic sections
- p. 859-860: (bicircular) quartic of Teixeira
- p. 870: sesqui-sectrix (of van der Schouten)
- p. 891-937: circular conchoids, limaçon of Pascal, limaçons, cardioid, Cartesian ovals, toric sections, spiric sections (of Perseus), lemiscates of Booth, Cassinian curves, lemiscate of Bernoulli, cyclides of Dupin

[4]Bartl mentions two curves with this name. One of them is a circular cubic, the other a tricircular sextic.

[5]Bartl calls these curves *cataspirica*.

2 Algebraic geometry

2.1 Basic terms

Affine and projective space

Definition 2.1 (Affine space $\mathbb{K}^n$)
Let $\mathbb{K}$ be a field and $n \in \mathbb{N}$. Then the set of n-tuples $(x_1, x_2, \ldots, x_n)$ with $x_i \in \mathbb{K}$ is called the n-dimensional affine space $\mathbb{K}^n$ over $\mathbb{K}$.

$\mathbb{K}^n$ is an n-dimensional vector space over $\mathbb{K}$. Some sources write this as $\mathbb{A}_n(\mathbb{K}) \cong \mathbb{K}^n$.

Most people that are learning something about geometry are taught the Euclidean geometry of the real plane $\mathbb{R}^2$. This is quite sufficient for everyday tasks like for example deciding whether given angle is a right angle.[1] But Euclidean geometry has two major drawbacks with respect to theoretic considerations:

- the intersection of two straight lines usually consists of exactly one point, but sometimes (when the lines are parallel) they do not intersect at all
- the number of common points of a given circle and a variable line can be zero, one or two depending on the distance d from the center of the circle to the line and the radius r of the circle (with the three cases corresponding to $r < d$, $r = d$ and $r > d$ respectively)

The first issue was adressed by Renaissance thinkers when they introduced a concept called perspective. Parallel lines are perceived as intersecting in a point "very, very far away from the observer". Every direction corresponds to one of these fictitious points. This concept has become precise through projective geometry. The second issue is an analogy to the problem that the number of solutions of a quadratic equation depends on the signum of the discriminant. This problem can be easily solved by considering complex numbers instead of real numbers.

> **Remark**
> It is not possible to rely solely on intuition, because there is no intuitive way to discuss questions of higher order properly. Two given conics seem to have at most four real points of intersection. Variation of the conics can make two or more of these points to join each other into some higher order intersection.
>
> Circles are a special kind of conics. They always have at most two real points in common (not four). This indicates that the other two (invisible) points of intersection are complex in nature (if they exist at all). Are these two complex points amalgated into a double

[1] [30, book 1, definition 10]

point of intersection or are they two different points? This can be answered of course with an argument utilizing complex conjugation. But what happens when the centers of the intersecting circles coincide, i.e. where and how do concentric circles intersect?

Definition 2.2 (Projective space $\mathbb{P}_n(\mathbb{K})$)
Let $\mathbb{K}$ be a field and $n \in \mathbb{N}$. Then the n-dimensional projective space over $\mathbb{K}$ is the set of equivalence classes

$$\mathbb{P}_n(\mathbb{C}) := (\mathbb{K}^{n+1} \setminus \{0, \ldots, 0\}) \;/ \sim$$

with respect to the equivalence relation

$$\mathbf{x} \sim \mathbf{y} \Longleftrightarrow \mathbb{K}\mathbf{x} = \mathbb{K}\mathbf{y}.$$

The equivalence class $\mathbb{K}\mathbf{x}$ of a point $\mathbf{x} = (x_0, x_1, \ldots, x_n) \neq (0, \ldots, 0)$ is denoted by

$$(x_0 : x_1 : \cdots : x_n).$$

We have already indicated that the projective space is an extension of the affine space.

Lemma 2.3 (Embedding of affine space)
Let $\mathbb{K}^n$ and $\mathbb{P}_n(\mathbb{K})$ be the n-dimensional affine and projective space. Then the map

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{P}_n(\mathbb{K}), \\ (x_1, \ldots, x_n) & \mapsto & (1 : x_1 : \cdots : x_n) \end{array} \right.$$

with inverse map

$$\varphi^{-1} : \left\{ \begin{array}{ccc} \mathbb{P}_n(\mathbb{K}) & \to & \mathbb{K}^n, \\ (x_0 : x_1 : \cdots : x_2) & \mapsto & (\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}). \end{array} \right.$$

is a bijection between $\mathbb{K}^n$ and $\{(x_0 : x_1 : \cdots : x_n) \in \mathbb{P}_n(\mathbb{K}) \mid x_0 \neq 0\}$.

Definition 2.4 (The hyperplane/line at infinity)
The set

$$\mathcal{H}_0 = \{(x_0 : x_1 : \cdots : x_n) \in \mathbb{P}_n(\mathbb{K}) \mid x_0 = 0\}$$

is called hyperplane at infinity*.*

In the complex projective plane $\mathbb{P}_2(\mathbb{C})$ the set

$$\mathcal{L}_0 = \{(x_0 : x_1 : x_2) \in \mathbb{P}_2(\mathbb{C}) \mid x_0 = 0\}$$

is called line at infinity*.*

Most passages of this text are written in the language of projective geometry. Some parts use the notation of affine geometry, when it seems to be appropriate. As a general distinction between these two notations we choose the letters x, y, z for affine and the indexed letters $x_0, x_1, \ldots$ for projective coordinates.

Algebraic curves

Definition 2.5 (Monomial)
Let $X_1, \dots, X_n$ be indeterminates and $a_1, \dots, a_n$ non-negative integers. Then the product

$$\prod_{i=1}^{n} X_i{}^{a_i} = X_1{}^{a_1} \dots X_n{}^{a_n}$$

is called a monomial. *Writing $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{a} = (a_1, \dots, a_n)$ we can write this monomial in short as $\mathbf{X}^{\mathbf{a}}$. It has the* degree

$$\deg(\mathbf{X}^{\mathbf{a}}) = |\,\mathbf{a}\,| = \sum_{i=1}^{n} a_i.$$

Definition 2.6 (Polynomial)
Let R be a ring, $r_1, \dots, r_k \in R$ elements of the ring and $X_1, \dots, X_n$ indeterminates. Then a polynomial *in n indeterminates over R is a linear combination*

$$r_1\mathbf{X}^{\mathbf{a_1}} + \dots + r_k\mathbf{X}^{\mathbf{a_k}}$$

of monomials. The degree of a polynomial is the maximum over the degrees of its monomials.

The set of all such polynomials is itself a ring, the ring $R[X_1, \dots, X_n]$ of polynomials in n variables over R.

Definition 2.7 (Homogeneous polynomial)
The polynomial

$$r_1\mathbf{X}^{\mathbf{a_1}} + \dots + r_k\mathbf{X}^{\mathbf{a_k}}$$

is called homogeneous *of degree d, if exists a natural number d such that for all $j \in \{1, \dots, k\}$ holds $|\,\mathbf{a_j}\,| = d$.*

The set $R_d[X_1, \dots, X_n]$ of all homogeneous polynomials of degree d over R is also a ring.

For our purposes the ring R will always be a field, because we want to evaluate polynomials at certain points of the vector space $\mathbb{K}^n$.

Definition 2.8 (Evaluation of a polynomial)
Let $p \in \mathbb{K}[X_1, \dots, X_n]$ be a polynomial. Then the evaluation *of*

$$p = c_1\mathbf{X}^{\mathbf{a_1}} + \dots + c_k\mathbf{X}^{\mathbf{a_k}}$$

at the point $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{K}^n$ is the sum

$$p(\mathbf{x}) := p(\mathbf{X})|_{\mathbf{X}=\mathbf{x}} = c_1\mathbf{x}^{\mathbf{a_1}} + \dots + c_k\mathbf{x}^{\mathbf{a_k}}.$$

A point $\mathbf{x} \in \mathbb{K}^n$ is called a zero of the polynomial $p \in \mathbb{K}[X_1, \dots, X_n]$ if $p(\mathbf{x}) = 0$. Of course, $\mathbf{x}$ is then also a zero of any polynomial in $\mathbb{K}p$.

Definition 2.9 (Affine hypersurface)
The ideal $\mathbb{K}p \subset \mathbb{K}[X_1, \ldots, X_n]$ *is called the* affine hypersurface *defined by the polynomial* $p \in \mathbb{K}[X_1, \ldots, X_n]$.

The zero set of p (which is the same for all non-zero polynomials in $\mathbb{K}p$) is also called the hypersurface defined by p. However, we have to be careful, because the polynomials $p = X_1$ and $q = {X_1}^2$ have the same zero set, but the ideals $\mathbb{K}p$ and $\mathbb{K}q$ are different. We always want to treat the hypersurfaces defined by p and by q as different.

Definition 2.10 (Projective hypersurface)
Let $p \in \mathbb{K}_d[X_0, \ldots, X_n]$ *be a non-constant homogeneous polynomial of degree* d*. Then the ideal* $\mathbb{K}p$ *is called the* projective hypersurface *defined by* p.

Definition 2.11 (Algebraic curve)
An affine (or projective) algebraic curve is an affine (or projective) 2-dimensional hypersurface.

When the degree of the defining polynomial is d, then the curve is called of degree or of order d and has the generic name d-tic. Curves of degree $1, 2, 3, 4, \ldots$ are known as lines, quadrics, cubics, quartics and so on.

Lemma 2.12 (Irreducible hypersurfaces)
If the polynomial p *factors into irreducible factors* $p = p_1 \cdot \cdots \cdot p_n$ *then the hypersurface defined by* p *is the union of the hypersurfaces defined by* $p_1, \ldots, p_n$*. A hypersurface defined by an irreducible polynomial is said to be irreducible.*

We have already stated in lemma (2.3) that the affine space of dimension n is embedded in the n-dimensional projective space. This technique of homogenization and dehomogenization can be used for polynomials, too. We can transform every polynomial in n variables into a polynomial in $(n + 1)$ variables in the following way.

Definition 2.13 (Homogenization of a polynomial)
Let $p(X_1, \ldots, X_n) \in \mathbb{K}[X_1, \ldots, X_n]$ *be a polynomial of degree* d*. Then the homogeneous polynomial* $p^* = {X_0}^d \cdot p(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0})$ *is called the* homogenization *of* p.

The result p^* is a homogeneous polynomial of the same degree as the original polynomial p. The homogenization of a polynomial can be reversed.

Definition 2.14 (Dehomogenization of a polynomial)
Let $p(X_0, X_1, \ldots, X_n) \in \mathbb{K}_d[X_0, \ldots, X_n]$ *be a homogeneous polynomial of degree* d*. Then the polynomial* $p_* = p(1, X_1, \ldots, X_n)$ *is called the* dehomogenization*of* p.

We should remark here that $p = (p^*)_*$ always holds, but in general for a given homogeneous polynomial $p \in \mathbb{K}[X_0, \ldots, X_n]$ the situation $p \neq (p_*)^*$ can occur. For instance, with $p = X_0$ we obtain $(p_*)^* = 1^* = 1 \neq X_0$. Hence we sometimes have to apply an appropriate coordinate transformation in order to ensure that X_0 does not divide p.

2.2 Geometric tools

Points on an algebraic curve

Usually the points on the hypersurface defined by the polynomial p are characterized by their property, that they are zeros of all polynomials in $\mathbb{K}p$. This is a special case of the more general

Definition 2.15 (Multiplicity of a point on a curve)
Let $\mathbf{x} \in \mathbb{P}_2(\mathbb{K})$ be a point and C an algebraic curve defined by the polynomial p. Then the multiplicity $\mathrm{mult}_{\mathbf{x}}(C)$ *of $\mathbf{x}$ on C is defined as the smallest natural number m such that the m-th Taylor polynomial of p in $\mathbf{x}$ is not the zero polynomial.*

In this context points on the curve C are exactly the points $\mathbf{x}$ with $\mathrm{mult}_{\mathbf{x}}(C) > 0$. The definition is equivalent to defining a m-fold point of the curve through the property, that all partial derivations of p up to the order $(m-1)$ are vanishing at $\mathbf{x}$ and that there is at least one partial derivation of m-th order not vanishing in $\mathbf{x}$. p is counted as the 0-th derivation of itself.

Definition 2.16 (Smooth and singular points)
A point $\mathbf{x}$ on a curve C is called regular *or* smooth*, if* $\mathrm{mult}_{\mathbf{x}}(C) = 1$*. If* $\mathrm{mult}_{\mathbf{x}}(C) > 1$*, the point is called* singular*.*

A curve without singular points is called smooth. If one of the points on the curve is a singularity, then the curve is called singular.

In the following we will always take the field $\mathbb{K}$ as the field $\mathbb{C}$ of complex numbers.

Definition 2.17 (Local ring)
For every point $\mathbf{x} \in \mathbb{P}_2(\mathbb{C})$ the local ring *at $\mathbf{x}$ is the ring of convergent power series centered at $\mathbf{x}$. We denote it by $\mathcal{O}_{\mathbf{x}}(\mathbb{C}^2)$.*

When a point $\mathbf{x}$ lies on the curves C and D we also say that the curves intersect in $\mathbf{x}$ or that they have $\mathbf{x}$ in common.

Definition 2.18 (Intersection multiplicity)
Let C and D be projective curves defined by $p, q \in \mathbb{C}[X_0, \ldots, X_n]$. After an appropriate coordinate transformation X_0 divides neither p nor q and $\mathbf{x}$ does not lie on the line at infinity. Then the intersection multiplicity *of C and D at the point $\mathbf{x}$ is defined as*

$$\mathrm{mult}_{\mathbf{x}}(C, D) = \dim_{\mathbb{C}} \left(\mathcal{O}_{\mathbf{x}}(\mathbb{C}^2) / \langle p_*, q_* \rangle \right),$$

where $\dim_{\mathbb{C}}$ *denotes the dimension of the residue class ring as a complex vector space and $\langle p_*, q_* \rangle$ is the ideal generated by p_* and q_*.*

The order of a curve, the multiplicity of a point on a curve and the intersection multiplicity of two curves are projective invariants.

Theorem 2.19 (Properties of the intersection multiplicity)
For the intersection multiplicity of curves C, C' and D holds

$$\begin{aligned} (i) \quad & \mathrm{mult}_{\mathbf{x}}(C, -D) = \mathrm{mult}_{\mathbf{x}}(C, D), \\ (ii) \quad & \mathrm{mult}_{\mathbf{x}}(C, C + D) = \mathrm{mult}_{\mathbf{x}}(C, D), \\ (iii) \quad & \mathrm{mult}_{\mathbf{x}}(C \cdot C', D) = \mathrm{mult}_{\mathbf{x}}(C, D) + \mathrm{mult}_{\mathbf{x}}(C', D). \end{aligned}$$

Definition 2.20 (Tangent)
Let $\mathbf{x}$ be a point on the curve C with multiplicity $\mathrm{mult}_{\mathbf{x}}(C) = m$. Then a straight line L running through $\mathbf{x}$ with intersection multiplicity $\mathrm{mult}_{\mathbf{x}}(C, L) > m$ is called tangent *to C in $\mathbf{x}$.*

Tangents can touch the curve in more than one point. In such a case the line is a called a double tangent, triple tangent and so on. Because $\mathbb{C}$ is algebraically closed we can make use of

Theorem 2.21 (Bezout's theoreom)
Two curves C_1 and C_2 in $\mathbb{P}_2(\mathbb{C})$ of degree d_1 and d_2 without a common component intersect in exactly $(d_1 \cdot d_2)$ points when counted with multiplicity.

Definition 2.22 (Image of an algebraic curve)
Let $p \in \mathbb{C}_d[X_0, X_1, X_2]$ be a homogeneous polynomial of degree d, $\mathcal{C} = \langle p \rangle$ be a projective algebraic curve and $\phi : \mathbb{P}_2(\mathbb{C}) \to \mathbb{P}_2(\mathbb{C})$ be a birational map on the projective plane. Then the image of the curve $\mathcal{C}$ is the projective curve

$$\phi(\mathcal{C}) = \left\langle p \circ \phi^{-1} \right\rangle .$$

Remark
The composition $p \circ \phi^{-1}$ is again a homogeneous polynomial. For a point $\mathbf{x}$ on $\mathcal{C}$, we have

$$p \circ \phi^{-1}(\phi(\mathbf{x})) = p(\phi^{-1} \circ \phi(\mathbf{x})) = p(x) = 0,$$

thus $\phi(\mathbf{x})$ lies on $\phi(\mathcal{C})$.

Dual curve and Plücker equations

A line in $\mathbb{P}_2(\mathbb{C})$ is defined by the linear equation

$$a_0 x_0 + a_1 x_1 + a_2 x_2 = 0. \tag{2.1}$$

We may also take equation (2.1) as a condition on $(a_0 : a_1 : a_2)$. For a given $(x_0 : x_1 : x_2)$ it describes the family of lines going through $(x_0 : x_1 : x_2)$. Each tuple $(a_0 : a_1 : a_2)$ uniquely defines a straight line in projective space.

Definition 2.23 (Dual curve)
The set of tangents to an algebraic curve C forms and algebraic curve, the so called dual curve *C^* of C.*

The dual curve of the dual curve is the original curve. The degree of C^* is called the class of C. This is the number of tangents one can in general draw to C through an arbitrary point $\mathbf{x} \in \mathbb{P}_2(\mathbb{C})$.

Each double tangent of C corresponds to a double point of C^* and vice versa. Each point of inflection C corresponds to a cusp of C^* and vice versa.

For a smooth curve C of degree c the degree of C^* is $d(d-1)$. For singular curves we can use the Plücker formulas.

Theorem 2.24 (Plücker formulas)
Let C be an algebraic curve and C^* its dual curve. Denote by d and d^* the degree of C resp. C^*, by δ and δ^* the number of double points of C resp. C^* with distinct tangents and by κ and κ^* the number of cusps of C resp. C^*. Then

$$\begin{aligned} d^* &= d(d-1) - 2\delta - 3\kappa, \\ \kappa^* &= 3d(d-2) - 6\delta - 8\kappa \end{aligned}$$

and the dual equations

$$\begin{aligned} d &= d^*(d^*-1) - 2\delta^* - 3\kappa^*, \\ \kappa &= 3d^*(d^*-2) - 6\delta^* - 8\kappa^*. \end{aligned}$$

C and C^* have the same genus g which can be computed as

$$g = \frac{1}{2}(d-1)(d-2) - \delta - \kappa$$

or equivalently as

$$g = \frac{1}{2}(d^*-1)(d^*-2) - \delta^* - \kappa^*.$$

Definition 2.25 (Pedal curve)
Let C be an algebraic curve and $P \in \mathbb{P}_2(\mathbb{C})$ be a point. For every tangent $T \in C^*$ to C we write $\pi_T(P)$ for the intersection point of T and the line through P and perpendicular to T. Then the set of these points $\pi_T(P)$ is the pedal curve *of C with respect to the pedal point P.*

The dual curve C^* of a given curve C is the pedal curve of the inverse of C. This is shown in [15].

3 Circles

3.1 Real and complex circles

The term *circle* is used in different contexts, with different meanings. In the geometry of the real plane it is a set of points fulfilling a certain condition, in the theory of algebraic curves in projective space it is an equivalence class of polynomials. Even though it may seem superfluent to discuss these aspects in detail, since many people think of circles as a subject that is not hard to understand, we take a discussion of these aspects as a starting point.

Classical definition

The definition, that can be considered the standard one for the vast majority of people is

Definition 3.1 (Circle in $\mathbb{R}^2$)
Let $C \in \mathbb{R}^2$ be a point and $r \in \mathbb{R}_+$ a positive real number. The locus

$$\mathcal{C} = \left\{ X \in \mathbb{R}^2 \mid d(C, X) = r \right\}$$

of points X, which lie at distance r from C, is called a circle. *r is then called the* radius *and C the* center *of this circle.*

This definition corresponds to definition 15 in the first book of Euclid's Elements ([30]) and suffices for virtually all practical purposes, when circles in $\mathbb{R}^2$ are considered. In Euclid's terminology the object defined by us is the circumference of a circle, where circle stands for the area enclosed by its circumference. In [20] this distinction is found in the definitions XXXII to XXXIV of the first book.

Let $\mathfrak{Circ}(\mathbb{R}^2)$ be the set of circles in the real plane. We know that the map

$$\begin{array}{rcl} \mathbb{R}^2 \times \mathbb{R}_+ & \to & \mathfrak{Circ}(\mathbb{R}^2) \\ (C, r) & \mapsto & \mathcal{C} = \{X \in \mathbb{R}^2 \mid d(C, X) = r\} \end{array}$$

is bijective. Each circle in $\mathbb{R}^2$ is uniquely determined by the coordinates of its center C and its radius r. This means that the set $\mathfrak{Circ}(\mathbb{R}^2)$ of circles in $\mathbb{R}^2$ is a three-parameter family over $\mathbb{R}$. The main drawback of this description of a circle is that the definition of the euclidean distance

$$d(X, X') = \sqrt{(x - x')^2 + (y - y')^2}$$

contains a square root. Because we always have $r > 0$, we do not have to worry about ambiguities of the sign of r and may use the condition

$$d(C, X)^2 = r^2$$

in definition (3.1) instead.

Definition 3.2 (Equation of a circle)
The circle with center $C = (c_1, c_2)$ and radius $r > 0$ is the set of solutions $(x, y) \in \mathbb{R}^2$ of the equation

$$(x - c_1)^2 + (y - c_2)^2 = r^2. \tag{3.1}$$

Circles in $\mathbb{C} \cong \mathbb{R}^2$

The classical definition (3.1) remains unproblematic, when we use it to define circles in $\mathbb{C}$. We identify the complex numbers with the real plane by taking their real and imaginary part as real coordinates, i.e.

$$x = \Re(z) \text{ and } y = \Im(z).$$

This induces the distance function

$$d(z_1, z_2) = |z_1 - z_2|$$

on $\mathbb{C}$ from the distance function on $\mathbb{R}$. The circles in $\mathbb{C}$ are thus identified with the circles in $\mathbb{R}^2$ as follows.

Definition 3.3 (Circle in $\mathbb{C}$)
Let $c \in \mathbb{C}$ be a complex number and $r \in \mathbb{R}_+$ a positive real number. The locus

$$\mathcal{C} = \left\{ z \in \mathbb{C} \mid d(z, c)^2 = r^2 \right\} \in \mathfrak{Circ}(\mathbb{C})$$

of complex numbers z lying at the distance r from c is called a circle *in $\mathbb{C}$. r is called the* radius *and c the* center *of this circle.*

This definition contains a subtle difficulty, because a circle is actually determined by an equation between five real numbers. Basically the equation (3.1) of a circle from definition (3.2) is of the form

$$(x - x_C)^2 + (y - y_C)^2 = d^2. \tag{3.2}$$

It contains the two coordinates x_C and y_C of the center, the two coordinates x and y of the variable point on the circle and the radius d. For an arbitrary circle these numbers are at first undetermined, and we need three more equations

$$y_C = c_1, \; y_C = c_2, \; d = r$$

to fix the center as C and the radius as r in order to pick out a certain circle $\mathcal{C}$.

Equation (3.2) can be expressed as

$$|z - c|^2 = r^2. \tag{3.3}$$

No problem arises with equation (3.2) in $\mathbb{R}^2$, because altogether we have four equations between five real numbers to determine a circle. This system of nonlinear equations has a (over $\mathbb{R}$) one-dimensional set of solutions in $\mathbb{R}^2$, if the equations are sensible. But for circles in $\mathbb{C}$, equation (3.3) is an equation involving only three complex numbers. Accordingly, after having chosen the center c and the radius r, we have only one complex variable z left and we should expect, that equation (3.3) has a set of solutions of dimension 0 over $\mathbb{C}$.

Over $\mathbb{R}$ a circle in $\mathbb{C}$ is one-dimensional. Circles in $\mathbb{C}$ therefore violate the naive rule

$$\dim_{\mathbb{R}} X = 2 \cdot \dim_{\mathbb{C}} X.$$

In our case, the reason for the observed discrepance is the fact that

$$|z|^2 = z\overline{z}$$

turns equation (3.3) into a equation of real numbers. It is not a truly complex equation.

Description using a Möbius transformation

In the last paragraph we have seen that the definition

$$\mathcal{E} : |z| = 1$$

of the unit circle may be confusing. With $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ another description of the unit circle is

$$\mathcal{E} = \left\{ z \in \mathbb{C} \mid \frac{z-i}{z+i} \in i\overline{\mathbb{R}} \right\}. \tag{3.4}$$

This relation expresses the fact that the unit circle has the segment between $-i$ and i as a diameter. Due to the Theorem of Thales the points on the unit circle are exactly[1] the points z, where the triangle formed by z, i and $-i$ has a right angle at z. Hence, the quotient of the chords $(z-i)$ and $(z+i)$ must be purely imaginary.

We want to give a different explanation and therefore define the map

$$f(z) := \frac{z-i}{z+i},$$

which is a Möbius transformation. Möbius transformations acting on $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ map circles and lines onto circles and lines. Moreover, such a map is defined by the image of a given triple of points. We can take for example the triple -1, i, 1 of points on the unit circle. Their images are $f(-1) = i$, $f(i) = 0$, $f(1) = -i$, respectively. This implies that f maps the unit circle onto the imaginary line. As a Möbius transformation is an automorphism of $\overline{\mathbb{C}}$, we are free to define the unit circle as the preimage of the imaginary axis with respect to the map f.

Compliance with the classical definition

Equation (3.4) includes the classical definition of the unit circle $\mathcal{E}$ as the locus of points with distance 1 from the origin. We divide

$$\frac{z-i}{z+i} \in i\overline{\mathbb{R}}$$

[1]The critical point $-i$ of this definition corresponds to $\frac{-i-i}{-i+i} = \infty \in i\overline{\mathbb{R}}$.

by the imaginary unit

$$\frac{z-i}{iz-1} \in \overline{\mathbb{R}}$$

and enlarge the fraction with the complex conjugate of its denominator, which yields

$$\frac{(z-i)(\overline{iz}-1)}{(iz-1)(\overline{iz-1})} = \frac{(z-i)(-i\overline{z}-1)}{(iz-1)(-i\overline{z}-1)}.$$

After multiplication we obtain

$$\frac{-iz\overline{z}-z-\overline{z}+i}{-i^2z\overline{z}-iz+i\overline{z}+1}$$

and use the formulas

$$2\Re(z) = z+\overline{z}$$

and

$$2i\Im(z) = z-\overline{z}$$

to rewrite equation (3.4) as

$$\mathcal{E} = \left\{ z \in \mathbb{C} \mid \frac{-2\Re(z)+i(1-|z|^2)}{(1+|z|^2)+2\Im(z)} \in \overline{\mathbb{R}} \right\}.$$

Because the denominator of this fraction is a real number, this fraction is real, if and only if the nominator is also a real number. Hence equation (3.4) is equivalent to the equation

$$|z|^2 = 1.$$

Circles in $\mathbb{C}^2$ and $\mathbb{P}_2(\mathbb{C})$

Although we will quite frequently return to circles in $\mathbb{R}^2$, we are actually interested in circles in $\mathbb{C}^2$, where we can use powerful tools like Bezout's Theorem, for instance. Starting from equation (3.2),

$$(x-c_1)^2+(y-c_2)^2 = r^2,$$

we now allow the radius to be any complex number. The indeterminates x and y may also be complex. We indicate this by the replacements x_1 for x and x_2 for y. Subtracting r^2 and grouping monomials with respect to their degree turns the equation into

$$({x_1}^2+{x_2}^2)-2c_1x_1-2c_2x_2+({c_1}^2+{c_2}^2-r^2) = 0.$$

With

$$c_3 = {c_1}^2+{c_2}^2-r^2$$

we may rewrite this as

$$ {x_1}^2+{x_2}^2-2c_1x_1-2c_2x_2+c_3 = 0. \tag{3.5}$$

With this equation we can define any circle in $\mathbb{C}^2$ for $c_1, c_2, c_3 \in \mathbb{C}$, but it may also be used to define a circle in $\mathbb{R}^2$. In that case $c_1, c_2, c_3 \in \mathbb{R}$ and $c_3 - {c_1}^2 - {c_2}^2 > 0$. In the following the word *circle* always stands for *complex circle*. The coefficients c_1, c_2, c_3, however, are assumed to be real numbers throughout this text.

Definition 3.4 (Equation of a circle in $\mathbb{P}_2(\mathbb{C})$)
In the projective plane $\mathbb{P}_2(\mathbb{C})$ a circle has the equation

$$x_1{}^2 + x_2{}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 + c_3x_0{}^2 = 0. \tag{3.6}$$

This is just the homogenization of equation (3.5).

3.2 The space of circles

Real and imaginary circles, nullcircles

Equation (3.6) contains three independent parameters. Two of them, c_1 and c_2, are the coordinates of the center C. The third parameter c_3 has no immediate geometrical equivalent, but is connected to the radius r of the circle by

$$r^2 = c_3 - c_1{}^2 - c_2{}^2. \tag{3.7}$$

Maybe it is worth to be noticed here that the size of the radius r, which was a fundamental value in the classical definition (3.1), has been replaced by its square r^2. We will see that r^2 is the truly fundamental value and that the properties of a complex circle and its relations to other geometrical objects are determined by r^2 rather than r.

For the real part of a circle equation (3.7) leads to three different situations. In the case $r^2 > 0$, the intersection of a complex circle and the real plane is a circle in $\mathbb{R}^2$ and nothing unusual happens. We call this kind of circle a *real circle*. If $r^2 < 0$, the equation has no real solutions and we say the circle to be *imaginary*. The intermediate case $r^2 = 0$ describes a circle with vanishing radius. The real part of the complex circle only consists of the center C. This form of circle is called a *nullcircle*. While the polynomial

$$(x_1 - c_1x_0)^2 + (x_2 - c_2x_0)^2$$

of a nullcircle is irreducible in $\mathbb{R}[x_0, x_1, x_2]$, it is clearly reducible in $\mathbb{C}[x_0, x_1, x_2]$.

Theorem 3.5 (Factorization of nullcircles)
A nullcircle in $\mathbb{P}_2(\mathbb{C})$ decomposes into a pair of complex conjugated lines. Each of these lines runs through the center of the nullcircle and one of the circular points at infinity.

Proof
Let

$$(x_1 - c_1x_0)^2 + (x_2 - c_2x_0)^2 = \Big((x_1 - c_1x_0) + i(x_2 - c_2x_0)\Big) \cdot \Big((x_1 - c_1x_0) - i(x_2 - c_2x_0)\Big)$$

be the equation of the given nullcircle. The first linear factor is a line through $(1 : c_1 : c_2)$ and $(0 : 1 : i)$, the second linear factor is a line through $(1 : c_1 : c_2)$ and $(0 : 1 : -i)$. □

This decomposition allows us to call circles with $r^2 \neq 0$ *regular* circles and nullcircles *singular* circles, because the center of a nullcircle is a double point.

The space of circles

The parameters c_1, c_2, c_3 in equation (3.6) can be regarded as the coordinates of a point in a three-dimensional complex vector space. We want to call this space the space of complex circles $\mathfrak{Circ}(\mathbb{P}_2(\mathbb{C}))$ and abbreviate this to $\mathfrak{Circ}$. We distinguish this space from $\mathbb{C}^3$ by marking its elements with a small circle, thus the triple of the parameters c_1, c_2, c_3 will be written as $(c_1, c_2, c_3)^\circ$. Moreover we denote the coordinate axes of the space of circles by the letters ξ, η and ζ.

Definition 3.6 (Space of circles)
The space of complex circles $\mathfrak{Circ}$ is a three dimensional complex vector space together with the bijective map

$$\chi : \begin{cases} \mathfrak{Circ} & \rightarrow & \mathbb{C}[x_0, x_1, x_2], \\ (c_1, c_2, c_3)^\circ & \mapsto & {x_1}^2 + {x_2}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 + c_3{x_0}^2 = 0. \end{cases}$$

With $\mathcal{C}$ being the circle with the equation

$${x_1}^2 + {x_2}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 + c_3{x_0}^2 = 0.$$

we write $\mathcal{C}^\circ$ as short hand for $\chi^{-1}(\mathcal{C})$, i.e. $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$.

Definition 3.7 (Representant of a circle)
The point $\mathcal{C}^\circ$ from above is called the representant *of the circle $\mathcal{C}$.*

Points in $\mathbb{C}^2$ can be interpreted as nullcircles. The origin $(0, 0, 0)^\circ$ of $\mathfrak{Circ}$, for instance, corresponds to the circle with center $(0, 0)$ and radius 0.

Theorem 3.8 (Embedding of $\mathbb{C}^2$ into $\mathfrak{Circ}$)
The affine complex plane $\mathbb{C}^2$ is embedded into the space of circles $\mathfrak{Circ}$.

Proof
The circle $\mathcal{C}$ with representant $\mathcal{C}^\circ = (\xi, \eta, \zeta)^\circ$ on the paraboloid

$$\Pi : \xi^2 + \eta^2 - \zeta = 0 \tag{3.8}$$

is a nullcircle and corresponds to the point $C = (\xi, \eta) \in \mathbb{C}^2$. The point $P = (x_1, x_2) \in \mathbb{C}^2$ corresponds to the nullcircle $\mathcal{P}$ represented by $\mathcal{P}^\circ = (x_1, x_2, {x_1}^2 + {x_2}^2)^\circ \in \mathfrak{Circ}$. □

Remark
The points on the line at infinity of $\mathbb{P}_2(\mathbb{C})$ are not embedded as nullcircles in $\mathfrak{Circ}$.

Definition 3.9 (Paraboloid Π of nullcircles)
Representants of a nullcircle are of the form $(c_1, c_2, {c_1}^2 + {c_2}^2)^\circ$. All of them lie on the paraboloid $\Pi : \xi^2 + \eta^2 - \zeta = 0$ which is called the paraboloid of nullcircles.

A picture of the real part of the paraboloid Π is shown in 3.1. Points below Π correspond to real, points above Π to imaginary circles.

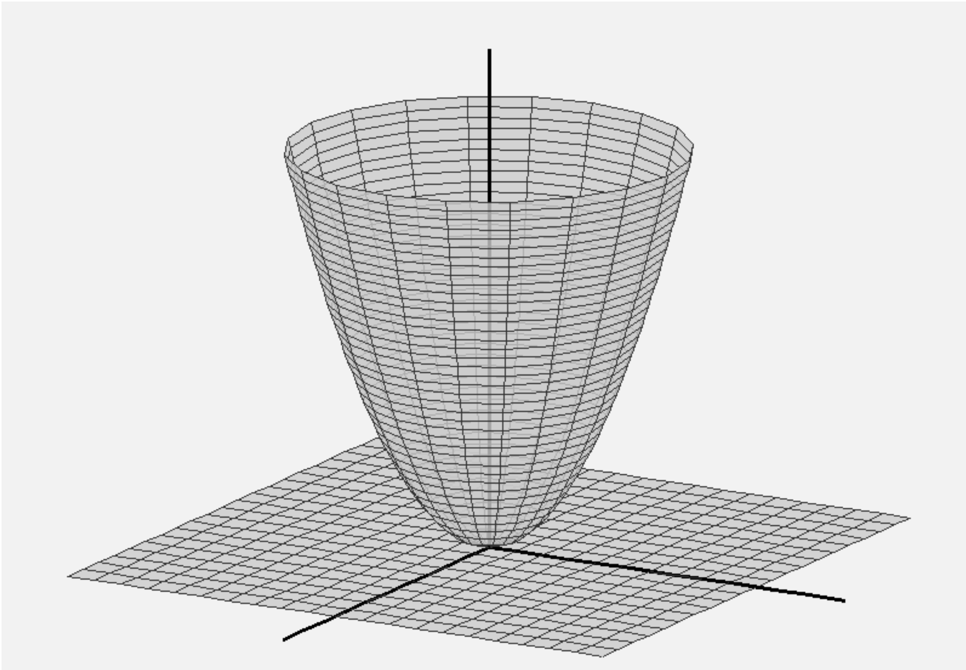


Figure 3.1: Paraboloid of nullcircles in $\mathfrak{Circ}$

Tangential planes of Π

The $\xi\eta$-plane $\zeta = 0$ consists of circles with an equation of the form

$$x_1{}^2 + x_2{}^2 - 2\xi x_0 x_1 - 2\eta x_0 x_2 = 0.$$

We can easily see that each of these circles runs through the origin $(1 : 0 : 0)$ of the complex projective plane. The $\xi\eta$-plane itself is tangent to Π in $(0, 0, 0)^\circ$. This is a special case of a first fundamental observation. We want to remark that the plane T touching Π in the point $(c_1, c_2, c_1{}^2 + c_2{}^2)^\circ$ has the equation

$$T : 2c_1\xi + 2c_2\eta - \zeta - c_1{}^2 - c_2{}^2 = 0. \tag{3.9}$$

Theorem 3.10 (Tangential planes of Π)
Let T the plane tangent to Π in the point $(c_1, c_2, c_1{}^2 + c_2{}^2)^\circ$. Then
(i) all circles in T contain the point (c_1, c_2),
(ii) all circles containing (c_1, c_2) lie on T.

Proof
Let $(\xi, \eta, \zeta)^\circ$ represent the circle

$$x_1{}^2 + x_2{}^2 - 2\xi x_0 x_1 - 2\eta x_0 x_2 + \zeta x_0{}^2 = 0.$$

We can check whether the point $(x_0 : x_1 : x_2) = (1 : c_1 : c_2)$ lies on this arbitrary circle. This holds exactly if

$$c_1{}^2 + c_2{}^2 - 2\xi c_1 - 2\eta c_2 + \zeta = 0,$$

which is equivalent to the condition (3.9) for $(\xi, \eta, \zeta)^\circ$ to be lying on T. □

The polar plane of a representant $\mathcal{C}^\circ$

Since all points of $\mathbb{C}^2$ are embedded into the space of circles, this in particular holds for the points on a given circle $\mathcal{C}$. Let $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ be the representant of this circle and $\mathcal{P} = (\xi, \eta, \xi^2 + \eta^2)^\circ \in \Pi$ a nullcircle whose center $P = (\xi, \eta) \in \mathbb{C}^2$ lies on $\mathcal{C}$. From theorem (3.10) follows that $\mathcal{C}^\circ$ lies on the plane T tangent to Π in $\mathcal{P}^\circ$. Hence, the line $\mathcal{C}^\circ\mathcal{P}^\circ$ is tangent to Π. Conversely, if $\mathcal{C}^\circ\mathcal{P}^\circ$ is tangent to Π, then $\mathcal{C}^\circ$ lies on T. From this observation follows immediately

Theorem 3.11 (Points/nullcircles on a circle)
The cone consisting of lines that run through the representant $\mathcal{C}^\circ$ of a circle $\mathcal{C}$ and are tangent to Π touches the paraboloid exactly in the representants $\mathcal{P}^\circ$ of nullcircles whose centers P lie on $\mathcal{C}$.

Definition 3.12 (Polar plane of a representant)
Let $\mathcal{C}$ be a circle with representant $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$. Then the plane

$$P_{\mathcal{C}^\circ} : 2c_1\xi + 2c_2\eta - \zeta - c_3 = 0 \tag{3.10}$$

is called the polar plane $P_{\mathcal{C}^\circ}$ *of $\mathcal{C}^\circ$ (with respect to Π).*

By comparing the equation of the polar plane with the equation of the tangential plane in a nullcircle we immediately obtain

Corollary 3.13 (Polar plane of a nullcircle)
In the case that $\mathcal{C}^\circ$ is a nullcircle the polar plane of $\mathcal{C}^\circ$ is the plane tangent to Π in $\mathcal{C}^\circ$.

Theorem 3.14 (Symmetry)
Let $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ be representants of circles. $\mathcal{D}^\circ$ lies on the polar plane $P_{\mathcal{C}^\circ}$ if and only if $\mathcal{C}^\circ$ lies on the polar plane $P_{\mathcal{D}^\circ}$.

Proof
For $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ we have

$$\mathcal{D}^\circ \in P_{\mathcal{C}^\circ} \iff 2c_1d_1 + 2c_2d_2 - c_3 - d_3 = 0$$

and also

$$2d_1c_1 + 2d_2c_2 - d_3 - c_3 = 0 \iff \mathcal{C}^\circ \in P_{\mathcal{D}^\circ}.$$

Because the equations are equivalent, the statement holds. □

Definition 3.15 (Conjugated representants)
Let $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ be representants of circles. $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ are said to be conjugated*, when*

$$2c_1d_1 + 2c_2d_2 - c_3 - d_3 = 0. \tag{3.11}$$

Theorem 3.16 (Intersection of the polar plane and Π)
Let $\mathcal{C}$ be a circle with representant $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$. Then the intersection of Π and the polar plane of $\mathcal{C}^\circ$ is the set of nullcircles that represent the points on $\mathcal{C}$.

Proof
For a point on Π holds $\zeta = \xi^2 + \eta^2$. We substitute the right hand side for ζ in the equation (3.10) of the polar plane, obtaining

$$2c_1\xi + 2c_2\eta - \xi^2 - \eta^2 - c_3 = 0.$$

This is equivalent to the equation of the circle $\mathcal{C}$, hence the center (ξ, η) of every nullcircle on the polar plane of $\mathcal{C}$ lies on this circle. $\square$

3.3 Elementary geometry

We shall see in this section that the chosen construction of the space of circles allows us to express geometric relations between circles in simple, linear formulas.

Polarity and orthogonality

Definition 3.17
Let $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ be representants of circles. Then

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 2c_1d_1 + 2c_2d_2 - c_3 - d_3 \tag{3.12}$$

is called the polarity *of $\mathcal{D}^\circ$ with respect to $\mathcal{C}^\circ$.*

Remark
The polarity $P(\mathcal{C}^\circ, \mathcal{D}^\circ)$ is symmetrical in $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$, i.e. $P(\mathcal{C}^\circ, \mathcal{D}^\circ) = P(\mathcal{D}^\circ, \mathcal{C}^\circ)$.

Definition 3.18 (Polarization function)
Let $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ be representants of circles. The function

$$P : \begin{cases} \mathfrak{Circ}^2 & \to \mathbb{C} \\ (\mathcal{C}^\circ, \mathcal{D}^\circ) & \mapsto 2c_1d_1 + 2c_2d_2 - c_3 - d_3 \end{cases}$$

is called polarization function.

Remembering equation (3.7) we obtain

Lemma 3.19
For the representant $\mathcal{C}^\circ$ of a circle $\mathcal{C}$ with radius $r_\mathcal{C}$

$$P(\mathcal{C}^\circ, \mathcal{C}^\circ) = -2{r_\mathcal{C}}^2.$$

Two real circles $\mathcal{C}, \mathcal{D}$ with centers C resp. D are orthogonal, if

$$d(C, D)^2 = {r_\mathcal{C}}^2 + {r_\mathcal{D}}^2,$$

that is, if the square of the distance of the centers and the squares of the radii fulfil the Theorem of Pythagoras.

Definition 3.20 (Orthogonality)
Two complex circles $\mathcal{C}$ and $\mathcal{D}$ with representants $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ are called orthogonal*, if*

$$(c_1 - d_1)^2 + (c_2 - d_2)^2 = ({c_1}^2 + {c_2}^2 - c_3) + ({d_1}^2 + {d_2}^2 - d_3),$$

that is, if

$$2c_1d_1 + 2c_2d_2 - c_3 - d_3 = 0.$$

We can summarize the connection between polarity and orthogonality like this.

Corollary 3.21 (Polarity of orthogonal circles)
The following statements about two circles $\mathcal{C}$ and $\mathcal{D}$ with representants $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ are equivalent:

(i) $\mathcal{C}^\circ$ *and* $\mathcal{D}^\circ$ *are conjugated,*
(ii) $P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 0$,
(iii) $\mathcal{C}$ *and* $\mathcal{D}$ *cut each other orthogonally.*

For a particular choice of $\mathcal{C}$ we can express this geometrically as

Corollary 3.22 (Orthogonal circles have conjugated representants)
The polar plane $P_{\mathcal{C}^\circ}$ of $\mathcal{C}^\circ$ consists exactly of the points representing circles that are orthogonal to $\mathcal{C}$.

An interesting special case of orthogonality is concerning nullcircles.

Theorem 3.23 (Self-orthogonal circles)
A circle is a nullcircle, if and only if it is orthogonal to itself.

Proof
$\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ represents a nullcircle if and only if it lies on Π, i.e. if $c_3 = {c_1}^2 + {c_2}^2$. This is exactly the same condition as

$$P(\mathcal{C}^\circ, \mathcal{C}^\circ) = 2{c_1}^2 + 2{c_2}^2 - 2c_3 = 0. \qquad \square$$

Again we can rewrite this in more geometrical terms.

Corollary 3.24
$\mathcal{C}^\circ$ represents a nullcircle exactly if it lies on its own polar plane $P_{\mathcal{C}^\circ}$.

Associated circles

In some situations, when we encounter an imaginary circle, we will make use of the *associated* real circle instead.

Definition 3.25 (Associated circles)
Two circles $\mathcal{C}$ and $\mathcal{D}$ are called associated*, when they have the same center*

$$(c_1, c_2) = (d_1, d_2)$$

and the squares of their radii add up to zero, i.e.

$$r_{\mathcal{C}}{}^2 = {c_1}^2 + {c_2}^2 - c_3 = -({d_1}^2 + {d_2}^2 - d_3) = -r_{\mathcal{D}}{}^2.$$

Remark
When a pair of regular circles is associated, then one of them is real and the other imaginary.

Corollary 3.26 (Self-associated)
A circle is a nullcircle, if and only if it is associated with itself.

Theorem 3.27 (Associated circles are orthogonal)
Let $\mathcal{C}$ and $\mathcal{D}$ be associated circles. Then they are orthogonal.

Proof
We take the representants $\mathcal{C}^\circ = (c_1, c_2, c_3)$ and $\mathcal{D}^\circ = (d_1, d_2, d_3)$. Then their polarity is

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 2c_1d_1 + 2c_2d_2 - c_3 - d_3.$$

$\mathcal{C}$ and $\mathcal{D}$ being associated is numerically expressed by

$${c_1}^2 + {c_2}^2 - c_3 = -({d_1}^2 + {d_2}^2 - d_3),$$

which is equivalent to

$$c_3 = {c_1}^2 + {c_2}^2 + {d_1}^2 + {d_2}^2 - d_3.$$

We substitute this expression in the formula of the polarity and obtain

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = -(c_1 - d_1)^2 - (c_2 - d_2)^2.$$

Because $\mathcal{C}$ and $\mathcal{D}$ are concentric, $P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 0$. $\square$

Corollary 3.28
Two circles are associated, if and only if they are concentric and orthogonal.

Angle of intersection

Two representants $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ are conjugated, when $P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 0$ holds. For real circles this property is equivalent to the geometrical condition that the distance $d(C, D)$ between the centers C and D of $\mathcal{C}$ and $\mathcal{D}$ and the radii $r_\mathcal{C}$ and $r_\mathcal{D}$ suffice

$$d(C, D)^2 = {r_\mathcal{C}}^2 + {r_\mathcal{D}}^2.$$

In this case the circles $\mathcal{C}$ and $\mathcal{D}$ intersect at right angle.

The equivalence

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 0 \Longleftrightarrow \mathcal{C} \perp \mathcal{D}$$

resembles to the relation in $\mathbb{R}^n$, where for two vectors v and w

$$\langle v, w \rangle = 0 \Longleftrightarrow v \perp w.$$

This similarity suggests the question, whether the number $P(\mathcal{C}^\circ, \mathcal{D}^\circ)$ retains a geometric meaning in the case, when it does not vanish. For real circles we have

$$2c_1d_1 + 2c_2d_2 - c_3 - d_3 = {r_\mathcal{C}}^2 + {r_\mathcal{D}}^2 - d^2.$$

When the circles are intersecting, we can apply the law of cosines and get

$$d^2 = {r_\mathcal{C}}^2 + {r_\mathcal{D}}^2 - 2r_\mathcal{C}r_\mathcal{D}\cos(\alpha),$$

where α is the angle of intersection between $\mathcal{C}$ and $\mathcal{D}$. Using equation (3.7) we arrive at

$$2c_1d_1 + 2c_2d_2 - c_3 - d_3 = 2r_\mathcal{C}r_\mathcal{D}\cos(\alpha)$$

or

$$\cos(\alpha) = \frac{2c_1d_1 + 2c_2d_2 - c_3 - d_3}{2r_\mathcal{C}r_\mathcal{D}}.$$

We square this equation

$$\cos^2(\alpha) = \frac{(2c_1d_1 + 2c_2d_2 - c_3 - d_3)^2}{4({c_1}^2 + {c_2}^2 - c_3)({d_1}^2 + {d_2}^2 - d_3)}$$

and rewrite the right hand side as

$$\cos^2(\alpha) = \frac{P(\mathcal{C}^\circ, \mathcal{D}^\circ)^2}{P(\mathcal{C}^\circ, \mathcal{C}^\circ)P(\mathcal{D}^\circ, \mathcal{D}^\circ)}. \tag{3.13}$$

Since we agreed to restrain the parameters of representants to real values, the right hand side of equation (3.13) will always be a real number.

Remark
If we write the scalar product of two vectors $v, w \in \mathbb{R}^n$ as

$$\langle v, w\rangle = \sum_{i=1}^{n} v_i w_i,$$

then for the angle α between v and w holds

$$\cos^2(\alpha) = \frac{\langle v, w\rangle^2}{\langle v, v\rangle\,\langle w, w\rangle}.$$

This is essentially the same as equation (3.13), only that the polarity function is replaced by the corresponding symmetric function for real vectors.

4 Transformations

4.1 Translations, rotations and dilations

We shall now determine how the representant of a circle is affected by the transformations of $\mathbb{C}^2$ mentioned in the title of this section. This will enable us to give simplified proofs, because we can deduce theorems for the general case from statements, that we have shown for some normalized situation only.

Translations

Let $\mathcal{C}$ be a circle with representant $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$. The polynomial p that defines this circle is

$$p(\mathbf{x}) = {x_1}^2 + {x_2}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 + c_3{x_0}^2.$$

The translation

$$\tau : \mathbf{x} \mapsto \mathbf{x} + \mathbf{t}$$

about the vector $\mathbf{t} = (t_1, t_2)$ in $\mathbb{C}^2$ maps the circle $\mathcal{C}$ onto the circle $\tau(\mathcal{C})$ with center $(c_1 + t_1, c_2 + t_2)$ and radius r. We verify this now with the tool from above.

The translation τ in $\mathbb{C}^2$ is represented in $\mathbb{P}_2(\mathbb{C})$ by the linear map

$$\tau : \begin{cases} \mathbb{P}_2(\mathbb{C}) & \rightarrow & \mathbb{P}_2(\mathbb{C}) \\ \mathbf{x} & \mapsto & \mathbf{y} \end{cases}$$

with

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ t_1 & 1 & 0 \\ t_2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}.$$

We compute the inverse map

$$\tau^{-1}(\mathbf{y}) = \begin{pmatrix} 1 & 0 & 0 \\ -t_1 & 1 & 0 \\ -t_2 & 0 & 1 \end{pmatrix} \mathbf{y},$$

which gives

$$\mathbf{x} = \begin{pmatrix} y_0 \\ y_1 - t_1y_0 \\ y_2 - t_2y_0 \end{pmatrix}.$$

Hence, we begin with

$$P(\tau^{-1}(\mathbf{y})) = 0.$$

The left hand side is

$$(y_1 - t_1y_0)^2 + (y_2 - t_2y_0)^2 - 2c_1y_0(y_1 - t_1y_0) - 2c_2y_0(y_2 - t_2y_0) + c_3{y_0}^2 = 0$$

and after expanding the monomials and sorting them in lexicographical order

$${y_1}^2 + {y_2}^2 - 2(t_1 + c_1)y_0y_1 - 2(t_2 + c_2)y_0y_2 + ({t_1}^2 + {t_2}^2 + 2c_1t_1 + 2c_2t_2 + c_3){y_0}^2 = 0.$$

This equation describes a circle with center

$$\tau(c_1, c_2) = (c_1 + t_1, c_2 + t_2).$$

Its radius is equal to the radius of $\mathcal{C}$.

Theorem 4.1 (Action of a translation on $\mathfrak{Circ}$)
The image $\tau(\mathcal{C})$ of a circle $\mathcal{C}$ with representant $\mathcal{C}^\circ$ has the representant

$$\tau(\mathcal{C}^\circ) := (\tau(\mathcal{C}))^\circ = (c_1 + t_1, c_2 + t_2, c_3 + {t_1}^2 + {t_2}^2 + 2c_1t_1 + 2c_2t_2).$$

Rotations

A rotation ϱ about the angle α with center $(0, 0)$ in $\mathbb{C}^2$ is described in $\mathbb{P}_2(\mathbb{C})$ by the linear map

$$\varrho : \begin{cases} \mathbb{P}_2(\mathbb{C}) & \to & \mathbb{P}_2(\mathbb{C}) \\ \mathbf{x} & \mapsto & \mathbf{y} \end{cases}$$

with

$$\mathbf{y} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C & -S \\ 0 & S & C \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix},$$

where

$$C = \cos(\alpha) \quad \text{and} \quad S = \sin(\alpha).$$

We compute the inverse map

$$\varrho^{-1}(\mathbf{y}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C & S \\ 0 & -S & C \end{pmatrix} \mathbf{y},$$

which gives

$$\mathbf{x} = \begin{pmatrix} y_0 \\ Cy_1 + Sy_2 \\ -Sy_1 + Cy_2 \end{pmatrix}.$$

Again, we begin with

$$P(\varrho^{-1}(\mathbf{y})) = 0.$$

The left hand side is

$$(Cy_1 + Sy_2)^2 + (-Sy_1 + Cy_2)^2 - 2c_1y_0(Cy_1 + Sy_2) - 2c_2y_0(-Sy_1 + Cy_2) + c_3{y_0}^2 = 0$$

and after expanding the monomials and sorting them in lexicographical order

$$y_1{}^2 + y_2{}^2 - 2(Cc_1 - Sc_2)y_0y_1 - 2(Sc_1 + Cc_2)y_0y_2 + c_3y_0{}^2 = 0.$$

This equation describes a circle with center

$$\varrho(c_1, c_2) = (\cos(\alpha)c_1 - \sin(\alpha)c_2, \sin(\alpha)c_1 + \cos(\alpha)c_2).$$

Its radius is equal to the radius of $\mathcal{C}$.

Theorem 4.2 (Action of a rotation on $\mathfrak{Circ}$)
The image $\varrho(\mathcal{C})$ of a circle $\mathcal{C}$ with representant $\mathcal{C}^\circ$ has the representant

$$\varrho(\mathcal{C}^\circ) := (\varrho(\mathcal{C}))^\circ = (\cos(\alpha)c_1 - \sin(\alpha)c_2, \sin(\alpha)c_1 + \cos(\alpha)c_2, c_3).$$

Dilations

A dilation δ about the factor D and center $(0,0)$ in $\mathbb{C}^2$ is described in $\mathbb{P}_2(\mathbb{C})$ by the linear map

$$\delta : \begin{cases} \mathbb{P}_2(\mathbb{C}) & \rightarrow & \mathbb{P}_2(\mathbb{C}) \\ \mathbf{x} & \mapsto & \mathbf{y} \end{cases}$$

with

$$\mathbf{y} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}.$$

We compute the inverse map

$$\delta^{-1}(\mathbf{y}) = \begin{pmatrix} D & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{y},$$

which gives

$$\mathbf{x} = \begin{pmatrix} Dy_0 \\ y_1 \\ y_2 \end{pmatrix}.$$

Like before, we begin with

$$P(\delta^{-1}(\mathbf{y})) = 0.$$

The left hand side is

$$y_1{}^2 + y_2{}^2 - 2c_1Dy_0y_1 - 2c_2Dy_0y_2 + c_3D^2y_0{}^2 = 0.$$

Here we do not even have to sort the monomials in order to see that this equation describes a circle with center

$$\delta(c_1, c_2) = (D \cdot c_1, D \cdot c_2)$$

and radius $D \cdot r$.

Theorem 4.3 (Action of a dilation on $\mathfrak{Circ}$)
The image $\delta(\mathcal{C})$ of a circle $\mathcal{C}$ with representant $\mathcal{C}^\circ$ has the representant

$$\delta(\mathcal{C}^\circ) := (\delta(\mathcal{C}))^\circ = (Dc_1, Dc_2, D^2c_3).$$

4.2 Inversions

The unit inversion ε

We have seen in the previous paragraphs that translations, rotations and dilations map circles on circles. Moreover they map lines on lines, because they are linear maps. In this section we shall discuss a class of maps that map lines and circles on lines and circles.

Definition 4.4 (Inversion about the unit circle)
Let $\mathcal{E}$ be a the unit circle in $\mathbb{C}^2$ with representant $\mathcal{E}^\circ = (0,0,-1)^\circ$. The inversion ε *about $\mathcal{E}$ in $\mathbb{C}^2$ is described by the quadratic map*

$$\varepsilon : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} {x_1}^2 + {x_2}^2 \\ x_0 x_1 \\ x_0 x_2 \end{pmatrix}$$

in $\mathbb{P}_2(\mathbb{C})$. The action of the inversion on $\mathbb{C}^2$ is induced via the embedding of $\mathbb{C}^2$ in $\mathbb{P}_2(\mathbb{C})$ given in lemma (2.3).

The map ε is not defined in $(1:0:0)$ and on $\mathcal{L}_0$. Every other inversion can be derived from ε. This is possible, because every circle in $\mathbb{C}^2$ is the image of the unit circle $\mathcal{E}$ under the composition of a dilation and a translation. We will write $\tau\delta$ for the composition $\tau \circ \delta$ throughout this text.

Theorem 4.5
Let $\mathcal{C}$ be a circle with representant $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$. Then

$$\mathcal{C} = \tau\delta(\mathcal{E})$$

with the dilation

$$\delta : \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{{c_1}^2 + {c_2}^2 - c_3}\,\xi \\ \sqrt{{c_1}^2 + {c_2}^2 - c_3}\,\eta \\ ({c_1}^2 + {c_2}^2 - c_3)\zeta \end{pmatrix}$$

and the translation

$$\tau : \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \mapsto \begin{pmatrix} \xi + c_1 \\ \eta + c_2 \\ \zeta + 2c_1\xi + 2c_2\eta + {c_1}^2 + {c_2}^2 \end{pmatrix}.$$

Proof
We start with $\mathcal{E}^\circ = (0,0,-1)^\circ$ and calculate $\delta(\mathcal{E}^\circ) = (0,0,-{c_1}^2 - {c_2}^2 + c_3)^\circ$. Then

$$\tau\delta(\mathcal{E}^\circ) = (0 + c_1, 0 + c_2, -{c_1}^2 - {c_2}^2 + c_3 + 0 + 0 + {c_1}^2 + {c_2}^2)^\circ = (c_1, c_2, c_3)^\circ. \qquad \square$$

General inversions

Definition 4.6 (Inversion about an arbitrary circle)
For $\mathbf{x} \in \mathbb{P}_2(\mathbb{C})$ the inversion about the circle $\mathcal{C} = \tau\delta(\mathcal{E})$ is

$$\iota_{\mathcal{C}} : \mathbf{x} \mapsto \tau\delta\varepsilon\delta^{-1}\tau^{-1}(\mathbf{x}).$$

In contrast to the maps studied in the previous section the inversion is not bijective on $\mathbb{C}^2$.

Theorem 4.7
The inversion about a circle $\mathcal{C}$ is not defined in its center C.

Proof
The unit inversion would map the origin $(1 : 0 : 0)$ onto $(0 : 0 : 0)$. This is not allowed, hence the inversion about $\mathcal{C}$ is undefined in the point $C = \tau\delta((1 : 0 : 0))$. □

Theorem 4.8 (Inversion about a nullcircle)
The inversion about a nullcircle is not defined.

Proof
The construction in definition (4.6) fails, when $\mathcal{C}$ is a nullcircle. The dilation δ would map the whole complex plane onto the center C. In consequence the inverse map δ^{-1} does not exist. □

Corollary 4.9
Only regular circles can be circles of inversion.

At the end of this section we will discuss another reason why no inversions about nullcircles exist. For the moment we take for granted that there is now way to refine our definition of inversion in order to make inversions about nullcircles possible.

The following two theorems are fundamental for inversions. They tell us that inversions and reflections behave in a very similar way.

Theorem 4.10 (Inversions are involutions I)
In $\mathbb{C}^2 \setminus \{C\}$ the inversion about the circle $\mathcal{C} = \tau\delta(\mathcal{E})$ is an involution.

Proof
Let $\mathbf{x} = (x_0 : x_1 : x_2)$ be a point in $\mathbb{P}_2(\mathbb{C}) \setminus \mathcal{L}_0$. Then

$$\varepsilon^2(\mathbf{x}) = \Big({x_0}^2({x_1}^2 + {x_2}^2) : ({x_1}^2 + {x_2}^2)x_0x_1 : ({x_1}^2 + {x_2}^2)x_0x_2\Big),$$

and we have

$$\varepsilon^2(\mathbf{x}) \sim \mathbf{x}.$$

From this follows that ε in an involution on $\mathbb{C}^2 \setminus \{(0, 0)\}$. For any other inversion $\iota_{\mathcal{C}}$ we have

$${\iota_{\mathcal{C}}}^2 = (\tau\delta\varepsilon\delta^{-1}\tau^{-1})^2 = \tau\delta\varepsilon\delta^{-1}\tau^{-1}\tau\delta\varepsilon\delta^{-1}\tau^{-1} = \tau\delta\varepsilon^2\delta^{-1}\tau^{-1} = id$$

and see immediately that $\iota_{\mathcal{C}}$ is an involution on $\mathbb{C}^2 \setminus \{\tau\delta(0, 0)\}$. $\tau\delta(0, 0)$ is the center C of the circle of inversion $\mathcal{C} = \tau\delta(\mathcal{E})$, $\iota_{\mathcal{C}}$ is not defined there. □

Theorem 4.11 (Fixed points of an inversion)
Every point on the circle $\mathcal{C}$ is a fixed point of the inversion $\iota_{\mathcal{C}}$. There are no other fixed points.

Proof
The fixed points of the unit inversion ε have to satisfy the equation

$$x_0{}^2 = x_1{}^2 + x_2{}^2.$$

This is exactly the equation of the unit circle $\mathcal{E}$. □

Modification of the construction

The given definition of the inversion about an arbitrary circle $\mathcal{C}$ contains an ambiguity. The dilation δ is not uniquely defined, because instead of the suitable dilation factor

$$D = \sqrt{c_1{}^2 + c_2{}^2 - c_3}$$

we can also choose $-D$. (Both dilations map the unit circle $\mathcal{E}$ onto the same circle $\tau^{-1}(\mathcal{C})$.) Although this is not a serious problem, it would be more pleasing, if the circle of inversion determined the construction in definition (4.6) uniquely. We achieve this by using the dilation δ^2 with dilation factor $D^2 = (-D)^2$, because no matter in which way we choose the dilation δ, its square δ^2 is always the same.

Definition 4.12 (General inversion)
The inversion about $\mathcal{C} = \tau\delta(\mathcal{E})$ is the map $\tau\delta^2\varepsilon\tau^{-1}$.

We only have to show

Theorem 4.13
Let δ be a dilation with center at the origin and dilation factor D. Then

$$\delta^2\varepsilon = \delta\varepsilon\delta^{-1}.$$

Proof

$$\delta\varepsilon\delta^{-1}((x_0 : x_1 : x_2)) = \delta\varepsilon((Dx_0 : x_1 : x_2)) = \delta((x_1{}^2 + x_2{}^2 : Dx_0x_1 : Dx_0x_2)) =$$
$$= (x_1{}^2 + x_2{}^2 : D^2x_0x_1 : D^2x_0x_2) = \delta^2((x_1{}^2 + x_2{}^2 : x_0x_1 : x_0x_2)) = \delta^2\varepsilon((x_0 : x_1 : x_2)).$$ □

The use of δ^2 will prove as an advantage in section (5.4).

Inversion of algebraic curves

Let C be an algebraic curve with defining polynomial $p \in \mathbb{C}[x_0, x_1, x_2]$. We shall examine how inversions act on algebraic curves. The most basic observation might be that an inversion in general doubles the order of an algebraic curve, since every monomial of degree d is substituted by a sum of monomials of degree $2d$. When we keep definition (4.6), this will always be the case.

Theorem 4.14
The image of an algebraic curve $\mathcal{C}$ of degree d under inversion is an algebraic curve of degree $2d$.

Proof
Let p be a homogeneous polynomial with degree $d = k + l + m$ and $\mathcal{C} = \langle p \rangle$. The inversion ε maps the monomial $({x_0}^k + {x_1}^l + {x_2}^m)$ of degree d to

$$\varepsilon({x_0}^k + {x_1}^l + {x_2}^m) = ({x_1}^2 + {x_2}^2)^k (x_0 x_1)^l (x_0 x_2)^k,$$

which is a monomial of degree $2d$. □

The inversion about a circle has the points on this circle as fixed points, thus the image of the circle as algebraic curve should have the circle as a component.

Example
The inversion ε maps the unit circle

$$\mathcal{E} : {x_1}^2 + {x_2}^2 - x_0^2 = 0$$

onto

$${x_0}^2 {x_1}^2 + {x_0}^2 {x_0}^2 - ({x_1}^2 + {x_2}^2)^2 = 0.$$

This can be factorized to

$$({x_0}^2 - {x_1}^2 - {x_2}^2)({x_1}^2 + {x_2}^2) = 0,$$

so the image is the union of the unit circle and the nullcircle with center at the origin. When we apply ε a second time, we obtain the identity map on $\mathbb{C}^2$, but the map ε^2 maps $\mathcal{E}$ onto

$$({x_1}^2 + {x_2}^2 - {x_0}^2)({x_1}^2 + {x_2}^2)^2 {x_0}^2 = 0.$$

and we see, that the component ${x_1}^2 + {x_2}^2 = 0$ does not vanish.

Because of definition (4.6) this problem is not caused by some special choice of our example; it arises with every inversion. We have to treat the appearing additional components as computational artefacts which do not belong to the actual image of the unit circle. In order to make things behave like we expect we have to modify definition (4.4) by punching out these exceptional components of a curve $\mathcal{C}$. The artefacts are the line at infinity

$$\mathcal{L}_0 : x_0 = 0$$

and the nullcircle with center at the origin

$$\mathcal{O} : {x_1}^2 + {x_2}^2 = 0.$$

The inversion is not defined there, because it maps points on $\mathcal{L}_0$ and $\mathcal{O}$.

Definition 4.15 (Unit inversion of a curve)
Let

$$\mathcal{C} = (\mathcal{O})^k \cup (\mathcal{L})^l \cup \tilde{\mathcal{C}}$$

be the decomposition of a curve $\mathcal{C}$ into components, where $\tilde{\mathcal{C}}$ does contain neither $\mathcal{O}$ nor $\mathcal{L}_0$. Let $\mathcal{D} = \varepsilon(\tilde{\mathcal{C}})$ be the image of $\tilde{\mathcal{C}}$ under the inversion about the unit circle and

$$\mathcal{D} = (\mathcal{O})^{k'} \cup (\mathcal{L})^{l'} \cup \tilde{\mathcal{D}}$$

be the analogous decomposition of $\mathcal{D}$. Then the curve $\tilde{\mathcal{D}}$ is the image of the curve $\tilde{\mathcal{C}}$ under the inversion about the unit circle.

Remark
$\tilde{\mathcal{C}}$ may contain the origin as a point, but it must not contain the origin as the component $\mathcal{O}$.

This way of handling inversions is very cumbersome. Later we will discuss a new approach that does not suffer from the appearance of exceptional components.

Theorem 4.16 (Inversions are involutions II)
Let $\mathcal{C}$ be a curve without exceptional components. Then $\varepsilon^2(\mathcal{C}) = \mathcal{C}$.

Proof
We have already seen that the inversion ε maps on the monomial ${x_0}^k {x_1}^l {x_2}^m$ to

$$\varepsilon({x_0}^k {x_1}^l {x_2}^m) = ({x_1}^2 + {x_1}^2)^k (x_0 x_1)^l (x_0 x_2)^m.$$

We keep the exceptional factors (we will prove below that this is allowed) and compute

$$\varepsilon^2({x_0}^k {x_1}^l {x_2}^m) = ({x_1}^2 + {x_1}^2)^{k+l+m} {x_0}^{2k+l+m} {x_1}^l {x_2}^m.$$

Now we see that

$$\varepsilon^2({x_0}^k {x_1}^l {x_2}^m) = {x_0}^{k+l+m} ({x_1}^2 + {x_1}^2)^{k+l+m} ({x_0}^k {x_1}^l {x_2}^m).$$

The exceptional factors are the same for every monomial of a homogeneous polynomial, thus a homogeneous polynomial P of degree d has the image

$$\varepsilon^2(P) = ({x_1}^2 + {x_1}^2)^d {x_0}^d \cdot P,$$

which is P after elimination of the exceptional factors.

We have to prove that

$$\varepsilon(x_0 P) = \varepsilon(({x_1}^2 + {x_2}^2)P) = \varepsilon(P).$$

But one can check by a simple calculation that

$$\varepsilon(x_0 P) = ({x_1}^2 + {x_2}^2)\varepsilon(P)$$

and

$$\varepsilon(({x_1}^2 + {x_2}^2)P) = {x_0}^2({x_1}^2 + {x_2}^2)\varepsilon(P). \square$$

4.3 Properties of the inversion

We have seen that inversions are involutions and that they in general double the order of an algebraic curve. This is not always the case. Circles, for example, and also bicircular quartics, as we will see later, do not show this behavoir.

Images of lines and circles under inversion

We start our examination with the action of inversions on straight lines and on circles. Our studies are limited to the action of the unit inversion ε, because the inversion about any given circle of inversion can be done by transforming this circle into the unit circle, performing ε and reverting the transformation.

Theorem 4.17 (Image of a line through the origin)
Let $\mathcal{L} : a_1x_1 + a_2x_2 = 0$ be a line through the origin. It has the image

$$\varepsilon(\mathcal{L}) : a_1x_1 + a_2x_2 = 0.$$

Hence a line through the origin is mapped onto itself.

Proof

$$\varepsilon(\mathcal{L}) : a_1x_0x_1 + a_2x_0x_2 = x_0(a_1x_1 + a_2x_2) = 0$$

is the union of $\mathcal{L}_0$ and $\mathcal{L}$. The exceptional component $\mathcal{L}_0$ is discarded. □

Theorem 4.18 (Image of a line not through the origin)
Let $\mathcal{L} : a_1x_1 + a_2x_2 + a_0x_0 = 0,\ a_0 \neq 0$ be a line not through the origin. It has the image

$$\varepsilon(\mathcal{L}) : a_0({x_1}^2 + {x_2}^2) + a_1x_0x_1 + a_2x_0x_2 = 0,$$

which is a circle through the origin with center $(-\frac{a_1}{2a_0}, -\frac{a_2}{2a_0})$.

Proof
The equation of the image is

$$\varepsilon(\mathcal{L}) : a_1x_0x_1 + a_2x_0x_2 + a_0({x_1}^2 + {x_2}^2) = 0$$

and we do not have to strip off exceptional components. Because $a_0 \neq 0$, we may divide

$$\varepsilon(\mathcal{L}) : {x_1}^2 + {x_2}^2 - 2\left(\frac{-a_1}{2a_0}\right)x_0x_1 - 2\left(\frac{-a_2}{2a_0}\right)x_0x_2 = 0.$$

This equation describes a circle in $\mathbb{C}^2$ with center $\left(-\frac{a_1}{2a_0}, -\frac{a_2}{2a_0}\right)$. The origin satisfies its equation. □

Theorem 4.19 (Image of a circle through the origin)
Let $\mathcal{C} : {x_1}^2 + {x_2}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 = 0$ be a circle. The image of $\mathcal{C}$ is

$$\varepsilon(\mathcal{C}) : -2c_1x_1 - 2c_2x_2 + x_0 = 0.$$

This is a line not running through the origin.

Proof
The equation we obtain after applying the inversion is

$$\varepsilon(\mathcal{C}) : (x_0x_1)^2 + (x_0x_2)^2 - 2c_1({x_1}^2 + {x_2}^2)x_0x_1 - 2c_2({x_1}^2 + {x_2}^2)x_0x_2 = 0.$$

This contains $\mathcal{L}_0$ and $\mathcal{O}$ as exceptional components. After removal of these remains

$$\varepsilon(\mathcal{C}) : x_0 - 2c_1x_1 - 2c_2x_2 = 0,$$

which describes a straight line that does not contain the origin. □

Theorem 4.20 (Image of a circle not through the origin)
Let $\mathcal{C} : {x_1}^2 + {x_2}^2 - 2c_1x_0x_1 - 2c_2x_0x_2 + c_3{x_0}^2 = 0$ be a circle not running through the origin, i.e. with $c_3 \neq 0$. Then $\mathcal{C}$ has the image

$$\varepsilon(\mathcal{C}) : c_3({x_1}^2 + {x_2}^2) - 2c_1x_0x_1 - 2c_2x_0x_2 + {x_0}^2 = 0,$$

which also is a circle not running through the origin.

Proof
Once again we apply the inversion. The raw result is

$$\varepsilon(\mathcal{C}) : (x_0x_1)^2 + (x_0x_2)^2 - 2c_1({x_1}^2 + {x_2}^2)x_0x_1 - 2c_2({x_1}^2 + {x_2}^2)x_0x_2 + c_3({x_1}^2 + {x_2}^2)^2 = 0.$$

Like before this contains $\mathcal{O}$, but no longer $\mathcal{L}_0$. After removal of the nullcircle $\mathcal{O}$ the final result is

$$\varepsilon(\mathcal{C}) : c_3({x_1}^2 + {x_2}^2) - 2c_1x_1 - 2c_2x_2 + x_0 = 0,$$

which is a circle, because $c_3 \neq 0$. Obviously it does not contain the origin. □

In the next section we shall develop a tool to derive the last four propositions from one single theorem. This requires a uniform treatment of lines and circles. Until then we are going to collect other important properties.

Invariant circles

A straight line through the center of inversion is mapped onto itself. We can also characterize it as a line that cuts the circle of inversion orthogonally. We will see that the second point of view provides more insight.

Theorem 4.21 (Invariant circles)
A circle $\mathcal{C}$ that cuts the circle $\mathcal{I}$ orthogonally is mapped onto itself by the inversion ι about $\mathcal{I}$. If a circle $\mathcal{C}$ is invariant under ι, either $\mathcal{C} \perp \mathcal{I}$ or $\mathcal{C} = \mathcal{I}$.

Proof
Let $\mathcal{C}$ be the circle with the representant $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ and be orthogonal to the unit circle $\mathcal{E}$. This means that their representants are conjugated. Because $\mathcal{E}$ is represented by $\mathcal{E}^\circ = (0, 0, -1)^\circ$, we have

$$P_{\mathcal{E}^\circ} : \zeta = 1.$$

Thus $c_3 = 1$ and $\mathcal{C}$ does not run through the origin. From theorem (4.20) we know that

$$\varepsilon(\mathcal{C}) : c_3({x_1}^2 + {x_2}^2) - 2c_1x_0x_1 - 2c_2x_0x_2 + {x_0}^2 = 0,$$

but as the coefficient $c_3 = 1$, this proves that $\varepsilon(\mathcal{C}) = \mathcal{C}$.

Let $\mathcal{C}$ be a circle that is invariant under ε. It must not run through the center of $\mathcal{E}$, because the inversion would map it onto a straight line then. From theorem (4.20) now follows that $c_3 = 1$ or $c_3 = -1$. In the first case $\mathcal{C}^\circ$ is conjugated to $\mathcal{E}^\circ$. According to theorem (3.21) this is equivalent to $\mathcal{C} \perp \mathcal{E}$. In the second case holds $(c_1, c_2) = (-c_1, -c_2)$ which then implies $\mathcal{C} = \mathcal{E}$. □

Commuting inversions

We can deduce an important theorem from the preceding statement. It is known that the reflections about two perpendicular lines commute. We will see now that this is also true for inversions about circles.

Corollary 4.22 (Commuting inversions)

Let $\mathcal{C}$ and $\mathcal{D}$ be orthogonal circles. Then the corresponding inversions $\iota_\mathcal{C}$ and $\iota_\mathcal{D}$ commute, i.e.

$$\iota_\mathcal{C}\iota_\mathcal{D} = \iota_\mathcal{D}\iota_\mathcal{C}.$$

Conversely, commuting inversions $\iota_\mathcal{C} \neq \iota_\mathcal{D}$ have orthogonal circles of inversion.

Proof
We can assume without loss of generality that $\mathcal{D} = \mathcal{E}$.

In theorem (4.21) we have shown that $\mathcal{C}$ is invariant under $\varepsilon = \iota_\mathcal{D}$. Thus the inversion $\iota_\mathcal{C}$ is invariant under ε, since it is completely defined by $\mathcal{C}$. This can be written as

$$\varepsilon\iota_\mathcal{C}\varepsilon^{-1} = \iota_\mathcal{C}$$

or equivalently as

$$\varepsilon\iota_\mathcal{C} = \iota_\mathcal{C}\varepsilon.$$

The converse part is proved by observing that commuting of ε and $\iota_\mathcal{C}$ implies invariance of $\iota_\mathcal{C}$ under ε. This further implies that $\mathcal{C}$ is invariant under ε. Hence $\mathcal{C} \perp \mathcal{E}$, because we assumed $\mathcal{C} \neq \mathcal{E}$. □

Angles

Theorem 4.23 (Preservation of angles)

Let ι be an inversion and $\mathcal{C}$ and $\mathcal{D}$ be two circles with representants $\mathcal{C}^\circ = (c_1, c_2, c_3)^\circ$ resp. $\mathcal{D}^\circ = (d_1, d_2, d_3)^\circ$ with $c_3 \neq 0$ and $d_3 \neq 0$ that cut each other with intersection angle α. Then their images $\iota(\mathcal{C})$ and $\iota(\mathcal{D})$ under inversion also intersect with angle α.

Proof
As the inversion ι we may assume ε. Neither $\mathcal{C}$ nor $\mathcal{D}$ contains the origin. Then equation (3.13) tells us that

$$\cos^2(\alpha) = \frac{P(\mathcal{C}^\circ, \mathcal{D}^\circ)^2}{P(\mathcal{C}^\circ, \mathcal{C}^\circ)P(\mathcal{D}^\circ, \mathcal{D}^\circ)}.$$

We start our calculation with the images of the representatives. According to theorem (4.20)

$$\iota(\mathcal{C})^\circ = \left(-\frac{c_1}{c_3}, -\frac{c_2}{c_3}, \frac{1}{c_3}\right)^\circ$$

and

$$\iota(\mathcal{D})^\circ = \left(-\frac{d_1}{d_3}, -\frac{d_2}{d_3}, \frac{1}{d_3}\right)^\circ.$$

After that we take equation (3.12) and compute

$$P\left(\iota(\mathcal{C})^\circ, \iota(\mathcal{C})^\circ\right) = 2\frac{{c_1}^2}{{c_3}^2} + 2\frac{{c_2}^2}{{c_3}^2} - 2\frac{1}{c_3} = \frac{1}{{c_3}^2}(2{c_1}^2 + 2{c_2}^2 - 2c_3),$$

$$P\left(\iota(\mathcal{D})^\circ, \iota(\mathcal{D})^\circ\right) = 2\frac{{d_1}^2}{{d_3}^2} + 2\frac{{d_2}^2}{{d_3}^2} - 2\frac{1}{d_3} = \frac{1}{{d_3}^2}(2{d_1}^2 + 2{d_2}^2 - 2d_3)$$

and

$$P\left(\iota(\mathcal{C})^\circ, \iota(\mathcal{D})^\circ\right) = 2\frac{c_1 d_1}{c_3 d_3} + 2\frac{c_2 d_2}{c_3 d_3} - \frac{1}{c_3} - \frac{1}{d_3} = \frac{1}{c_3 d_3}(2c_1 d_1 + 2c_2 d_2 - c_3 - d_3).$$

This can be summarized as

$$P\left(\iota(\mathcal{C})^\circ, \iota(\mathcal{C})^\circ\right) = \frac{1}{{c_3}^2} P\left(\mathcal{C}^\circ, \mathcal{C}^\circ\right),$$

$$P\left(\iota(\mathcal{D})^\circ, \iota(\mathcal{D})^\circ\right) = \frac{1}{{d_3}^2} P\left(\mathcal{D}^\circ, \mathcal{D}^\circ\right)$$

and

$$P\left(\iota(\mathcal{C})^\circ, \iota(\mathcal{D})^\circ\right) = \frac{1}{c_3 d_3} P\left(\mathcal{C}^\circ, \mathcal{D}^\circ\right).$$

When we now finally compute the intersection angle of $\iota(\mathcal{C})$ and $\iota(\mathcal{D})$ using equation (3.13), the fractions cancel out. This completes the proof. □

It is possible to include the special cases $c_3 = 0$ and $d_3 = 0$ in the previous theorem. Since we will see in the next section that there is a more uniform theory of the inversion of lines and circles, we want to accept this small gap at this point of the discussion. The angle of intersection is indeed preserved even in the special case that we have not treated here.

We stated earlier that only regular circles allow an inversion. We argumented that the inversion about a nullcircle could not be reduced to the inversion about the unit circle. This argument is weak, because it sounds like that we are just not competent enough to work out an alternative definition for this degenerate case. We will see now that there is in fact no possible way to do so.

Theorem 4.24 (Inversions about associated circles)
Let $\mathcal{C}$ and $\mathcal{D}$ be two associated circles with center C. Then the inversions $\iota_\mathcal{C}$ and $\iota_\mathcal{D}$ are connected by

$$\iota_\mathcal{C} = \delta_{-1}\iota_\mathcal{D},$$

where δ_{-1} is the inversion[1] *about the center C.*

Proof
We assume $\mathcal{C} = \mathcal{E}$ and $\mathcal{D} = \mathcal{E}'$ and compute

$$\varepsilon\varepsilon' : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \varepsilon \begin{pmatrix} {x_1}^2 + {x_2}^2 \\ -x_0 x_1 \\ -x_0 x_2 \end{pmatrix} = \begin{pmatrix} {x_0}^2({x_1}^2 + {x_2}^2) \\ -x_0 x_1({x_1}^2 + {x_2}^2) \\ -x_0 x_2({x_1}^2 + {x_2}^2) \end{pmatrix}.$$

We see that

$$\varepsilon\varepsilon'(\mathbf{x}) \sim \delta_{-1}(\mathbf{x}),$$

i.e. $\varepsilon\varepsilon'$ is the inversion δ_{-1} about the origin. ε and ε' are inversions and therefore involutions. This enables us to write

$$\varepsilon = \varepsilon\varepsilon'\varepsilon' = \delta_{-1}\varepsilon'.$$ □

[1]Here *inversion* about the point C stands for the dilation with center C and factor -1.

Corollary 4.25 (No inversion about a self-associated circle)
No inversion about a self-associated circle exists.

Proof
We assume that such an inversion ι would exist. We write ι' for the inversion about the associated circle, but we see that $\iota' = \iota$, because the circle is self-associated. Hence

$$\iota^2 = \iota\iota' = \delta_{-1} \neq \mathrm{id},$$

which is a contradiction to $\iota^2 = \mathrm{id}$. □

Self-associated circles are exactly the nullcircles, so we have finally shown that inversions about nullcircles are not defined.

5 Projective space of circles

5.1 The projective space of circles $\mathbb{P}(\mathfrak{Circ})$

Motivation

In geometrical investigations degenerate cases arise. Let us take for example

Theorem 5.1 (Circle through three points)
Let P_1, P_2, P_3 be three points in $\mathbb{C}^2$ that do not lie on a straight line. Then a uniquely defined circle running through them exists.

Apparently this is only possible if these points are not collinear, because in that exceptional case they do not define a circle but a straight line. From the viewpoint of an engineer it is clear that any small random displacement of three collinear points deforms their locus from a line into a circle. On the other hand it is easy to imagine that when the curvature of a circle decreases to zero, i.e. when its radius increases beyond all bounds, it just turns into a straight line.

We seek to treat lines as degenerate case of circles. Fortunately treating lines as infinitely big circles can be worked out precisely within the framework of the space of circles. Before that we should start with a special case in $\mathbb{R}^2$ in order to point out how problems and counterintuitive situations may arise.

Example
Let

$$P_1 = (-1, 0),\ P_2 = (0, -t) \text{ and } P_3 = (1, 0)$$

be three points and $t \in \mathbb{R}$ a real parameter. We denote the circle through these points by $\mathcal{C}_t$. For $t \neq 0$, we obtain the equation

$$\mathcal{C}_t : x^2 + y^2 + \left(t - \frac{1}{t}\right) y - 1 = 0.$$

The left hand side

$$p_t(x, y) = x^2 + y^2 + \left(t - \frac{1}{t}\right) y - 1$$

is the polynomial that defines the circle $\mathcal{C}_t$. According to equation (3.7) the radius of this circle satisfies

$${r_{\mathcal{C}_t}}^2 = t^2 + \frac{1}{t^2} - 1.$$

We see from this, that $\mathcal{C}_t$ is always real and that $r_{\mathcal{C}_t} \geq 1$. Moreover,

$$\lim_{t \to 0} r_{\mathcal{C}_t} = \lim_{t \to \infty} r_{\mathcal{C}_t} = \infty.$$

In the origin p_t always has the value (-1), regardless of t. Supposed that we can trust our intuition the limit of the circle $\mathcal{C}_t$ for $t \to 0$ is the x-axis. Unfortunately this implies a discontinuity in the values of the polynomial

$$f(t) = p_t(0,0),$$

because $f(t) = 1$ for all $t \neq 0$, but $f(0) = 0$.

For $t \neq 0$, we may modify p_t by multiplying with t.

$$\mathcal{C}_t : t(x^2 + y^2) + (t^2 - 1)y - t = 0$$

is an equivalent description of the circle through P_1, P_2 and P_3. Now the value of left hand side at the origin is $(-t)$, which strongly suggests that the origin will lie on the (perhaps critical) circle $\mathcal{C}_0$. For $t = 0$, the equation of the circle becomes

$$\mathcal{C}_0 : -y = 0,$$

which indeed describes the x-axis.

The other critical limit is $t \to \infty$. Here we also expect that the circle degenerates into the x-axis. This expectation is much harder to explain, because the point $P_2 = (0, -\infty)$ does not lie on the x-axis. In order to support our intuition we have to divide the original equation by t, obtaining

$$\mathcal{C}_t : \frac{1}{t}(x^2 + y^2) + \left(1 - \frac{1}{t^2}\right) y - \frac{1}{t} = 0.$$

In this way we arrive at the equation

$$\mathcal{C}_\infty : y = 0.$$

This result implies that $P_2 = (0, -\infty)$ lies on the x-axis, doesn't it?

For $t \to -\infty$ we also obtain $\mathcal{C}_{-\infty} : y = 0$, hence $P_2 = (0, +\infty)$ lies on the x-axis, too. This seems to be a little confusing, because $(0, \infty)$ and $(0, -\infty)$ should also lie on the y-axis, of course. But then the x-axis and the y-axis would have the three distinct points $(0, 0)$, $(0, \infty)$ and $(0, -\infty)$ in common. Even if we assume that $(0, \pm\infty)$ is in fact only one point, we are still in trouble. it seems that as an infinitely big circle the x-axis, which is a straight line, intersects with the y-axis, which is a different straight line, in two distict points. Under normal circumstances, i.e. in Euclidean geometry, we would have shown intuitively that the x-axis and the y-axis are the same.

Definition

It would be nice, if we could handle exceptional situations like that from the example without individual treatment. We are following a similar procedure like that which we have already successfully applied in order to extend affine space to infinity. We start with the triple $\mathcal{C}^\circ = (\xi, \eta, \zeta)^\circ$ which describes the circle

$$\mathcal{C} : {x_1}^2 + {x_2}^2 - 2\xi x_0 x_1 - 2\eta x_0 x_2 + \zeta {x_0}^2 = 0$$

in $\mathbb{P}_2(\mathbb{C})$. This equation is homogeneous in the variables x_0, x_1, x_2, but inhomogeneous in the parameters ξ, η, ζ.

In our example the circle was required to run through three given points. This turns ξ, η, ζ into variables that have to satisfy a given system of linear equations. We introduce an additional coordinate axis ν and define the equivalence class $(\nu : \xi : \eta : \zeta)$ in $\mathbb{P}_3(\mathbb{C})$ in the usual way. Now we only have to explain which kind of circle this object represents.

Definition 5.2 (Equivalence classes of circles)
The tuple $\mathcal{C}^\circ = (\nu : \xi : \eta : \zeta)^\circ$ represents the circle

$$\mathcal{C} : \nu({x_1}^2 + {x_2}^2) - 2\xi x_0 x_1 - 2\eta x_0 x_2 + \zeta {x_0}^2 = 0.$$

Obviously the equivalence relation of the representants in $\mathbb{P}_3(\mathbb{C})$ is the identity relation in the underlying space of circles.

Definition 5.3 (Projective space of circles)
The set of equivalence classes $(\xi_0 : \xi_1 : \xi_2 : \xi_3)^\circ$ is called projective space of circles $\mathbb{P}(\mathfrak{Circ})$.

The introduction of projective coordinates is an important step towards a more powerful theory. For a given equation of a circle all non-zero multiples of this equation describe the same circle. Projective coordinates reflect this equivalence in a natural way. It makes no sense to cripple the description of circles through equations of the form

$$\nu({x_1}^2 + {x_2}^2) - 2\xi x_0 x_1 - 2\eta x_0 x_2 + \zeta {x_0}^2 = 0 \tag{5.1}$$

by reducing it to the cases $\nu = 1$, where it describes a proper circle, and $\nu = 0$, where it describes a straight line. When treating linear families of circles the projective description will prove superior to the notation used in [44, p. 14]. Pedoe has no choice, because he has to begin the equation of a circle with the term (x^2+y^2). Later in the text ([44, p. 27]) when he introduces Joachimsthal's[1] ratio formula the chosen ratio of division $\lambda : 1$ makes sure that all equations will contain $1 \cdot (x^2 + y^2)$. This limitation is not necessary when representants are taken from the projective space of circles.

Center and radius, infinitely large circles

Theorem 5.4 (Embedding of the space of circles)
The space of circles $\mathfrak{Circ}$ is embedded in the projective space of circles $\mathbb{P}(\mathfrak{Circ})$ as the subspace $(1 : \xi : \eta : \zeta)$. The bijection is defined in analogy to theorem (2.3) as the canonical embedding of $\mathbb{C}^3$ in $\mathbb{P}_3(\mathbb{C})$.

For a representant $(\xi_0 : \xi_1 : \xi_2 : \xi_3)^\circ$ with $\xi_0 \neq 0$ we can easily see, that the corresponding circle has its center at $\left(\frac{\xi_1}{\xi_0}, \frac{\xi_2}{\xi_0}\right)$ and the radius $\frac{{\xi_1}^2+{\xi_2}^2-\xi_0\xi_3}{{\xi_0}^2}$. In $\mathbb{P}_2(\mathbb{C})$ holds

Theorem 5.5 (Center of a circle)
Let $\mathcal{C}$ be the circle with the representant $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$. Then its center is

$$C = (c_0 : c_1 : c_2).$$

With this new tool it becomes clear that the representant $\mathcal{L} = (0 : \xi_1 : \xi_2 : \xi_3)^\circ$ on the hyperplane at infinity $\mathcal{H}_0$ corresponds to the equation

$$\mathcal{L} : -2\xi_1 x_0 x_1 - 2\xi_2 x_0 x_2 + \xi_3 {x_0}^2 = 0.$$

[1] The surname of Ferdinand Joachimsthal (1818-1861) is misspelled as Joachimstahl in the cited book.

This describes the union of the straight line

$$-2\xi_1 x_1 - 2\xi_2 x_2 + \xi_3 x_0 = 0$$

and the line at infinity.

Now we have an explanation for the strange behaviour of the coordinate axes in our example. The x-axis as an infinitely large circle is in fact the union of the x-axis and the line at infinity. This causes the two distinct intersection points. One intersection is the origin, the other is the intersection of the y-axis and the line at infinity.

By the means of this construction we can treat circles and lines (more) uniformly. Geometrically it gives us control over how the radius and the center of a circle go to infinity simultaneously, when it is deformed into a straight line.

5.2 The action of transformations on $\mathbb{P}(\mathfrak{Circ})$

Translations, rotations and dilations

We already put down the laws that describe the induced action in $\mathfrak{Circ}$ of a translation, rotation or dilation in $\mathbb{C}^2$. We want to do the same for the induced action on $\mathbb{P}(\mathfrak{Circ})$.

Theorem 5.6 (Translation)
The translation τ that maps the origin onto the point $(t_0 : t_1 : t_2)$ acts on $\mathbb{P}(\mathfrak{Circ})$ as

$$\tau : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} {t_0}^2\xi_0 \\ {t_0}^2\xi_1 + t_0 t_1 \xi_0 \\ {t_0}^2\xi_2 + t_0 t_2 \xi_0 \\ {t_0}^2\xi_3 + 2t_0t_1\xi_1 + 2t_0t_2\xi_2 + ({t_1}^2 + {t_2}^2)\xi_0 \end{pmatrix}.$$

Proof
$(\xi_0 : \xi_1 : \xi_2 : \xi_3)^\circ$ represents the equation

$$\xi_0({x_1}^2 + {x_2}^2) - 2\xi_1 x_0 x_1 - 2\xi_2 x_0 x_2 + \xi_3 {x_0}^2 = 0.$$

We take

$$\tau^{-1} : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} t_0 x_0 \\ t_0 x_1 - t_1 x_0 \\ t_0 x_2 - t_2 x_0 \end{pmatrix}$$

and substitute this in the equation. We obtain

$$\begin{aligned} \xi_0\Big((t_0x_1 - t_1x_0)^2 + (t_0x_2 - t_2x_0)^2\Big) \quad &- \quad 2\xi_1(t_0x_0)(t_0x_1 - t_1x_0) \\ &- \quad 2\xi_2(t_0x_0)(t_0x_2 - t_2x_0) \quad + \quad \xi_3(t_0x_0)^2 = 0, \end{aligned}$$

which is the same as

$$\begin{aligned} {t_0}^2\xi_0({x_1}^2 + {x_2}^2) \quad &- \quad 2({t_0}^2\xi_1 + t_0t_1\xi_0)x_0x_1 \\ &- \quad 2({t_0}^2\xi_2 + t_0t_2\xi_0)x_0x_2 \\ &+ \quad \Big({t_0}^2\xi_3 + 2t_0t_1\xi_1 + 2t_0t_2\xi_2 + ({t_1}^2 + {t_2}^2)\xi_0\Big){x_0}^2 \quad = \quad 0. \qquad \square \end{aligned}$$

Theorem 5.7 (Dilation)
Let δ be a dilation with center at the origin and factor D. Then it acts on $\mathbb{P}(\mathfrak{Circ})$ as

$$\delta : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} \xi_0 \\ D\xi_1 \\ D\xi_2 \\ D^2\xi_3 \end{pmatrix}.$$

Proof
There is no significant difference to the action on $\mathfrak{Circ}$ here. $\square$

Much more important is the following transformation.

Theorem 5.8 (Square of a dilation)
Let δ be a dilation with center at the origin that maps the unit circle $\mathcal{E}$ with representant $\mathcal{E}^\circ = (1:0:0:-1)^\circ$ onto a circle whose radius r satisfies

$$r^2 = r_{\mathcal{C}}{}^2 = \frac{c_1{}^2 + c_2{}^2 - c_0c_3}{c_0{}^2}$$

for a given circle $\mathcal{C}$ with representant $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$. Then the square δ^2 acts on $\mathbb{P}(\mathfrak{Circ})$ as

$$\delta^2 : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} c_0{}^4\xi_0 \\ c_0{}^2(c_1{}^2 + c_2{}^2 - c_0c_3)\xi_1 \\ c_0{}^2(c_1{}^2 + c_2{}^2 - c_0c_3)\xi_2 \\ (c_1{}^2 + c_2{}^2 - c_0c_3)^2\xi_3 \end{pmatrix}.$$

Proof
The inverse of δ^2 is the map

$$(\delta^2)^{-1} : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} (c_1{}^2 + c_2{}^2 - c_0c_3)x_0 \\ c_0{}^2x_1 \\ c_0{}^2x_2 \end{pmatrix}.$$

We substitute this in the equation

$$\begin{aligned} \xi_0\Big((c_0{}^2x_1)^2 + (c_0{}^2x_2)^2\Big) \quad & - \quad 2\xi_1\Big((c_1{}^2 + c_2{}^2 - c_0c_3)x_0(c_0{}^2x_1)\Big) \\ & - \quad 2\xi_2\Big((c_1{}^2 + c_2{}^2 - c_0c_3)x_0(c_0{}^2x_2)\Big) \\ & + \quad \xi_3(c_1{}^2 + c_2{}^2 - c_0c_3)^2x_0{}^2 \qquad = \quad 0. \end{aligned}$$

This may be rearranged into

$$\begin{aligned} c_0{}^4\xi_0(x_1{}^2 + x_2{}^2) \quad & - \quad 2c_0{}^2(c_1{}^2 + c_2{}^2 - c_0c_3)\xi_1x_0x_1 \\ & - \quad 2c_0{}^2(c_1{}^2 + c_2{}^2 - c_0c_3)\xi_2x_0x_2 \\ & + \quad (c_1{}^2 + c_2{}^2 - c_0c_3)^2\xi_3x_0{}^2 \qquad = \quad 0. \end{aligned}$$

$\square$

The unit inversion

For describing the action of inversions we currently use four separate theorems in order to cover all cases. The possibility to map a circle onto a straight line prohibited us to give an explicit formula for the action of an inversion on $\mathfrak{Circ}$.

In $\mathbb{P}(\mathfrak{Circ})$ this separation between circles and lines is no longer needed. We will give a formula for the action of a general inversion on $\mathbb{P}(\mathfrak{Circ})$. The four theorems in section (4.3) will follow immediately from theorem (5.9).

Theorem 5.9 (Inversion about the unit circle)
The inversion in $\mathbb{P}(\mathfrak{Circ})$ about the unit circle $\mathcal{E}$ is the map

$$\varepsilon : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} \xi_3 \\ \xi_1 \\ \xi_2 \\ \xi_0 \end{pmatrix}.$$

Proof
Inversions are involutions, hence

$$\varepsilon^{-1} : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} {x_1}^2 + {x_2}^2 \\ x_0 x_1 \\ x_0 x_2 \end{pmatrix}.$$

We substitute this and obtain

$$\begin{array}{rlr} \xi_0(x_0{x_1}^2 + x_0{x_2}^2) & - \ 2\xi_1({x_1}^2+{x_2}^2)(x_0x_1) & \\ & - \ 2\xi_2({x_1}^2+{x_2}^2)(x_0x_2) & \\ & + \ \xi_3({x_1}^2+{x_2}^2)^2 & = \ 0. \end{array}$$

This is

$$({x_1}^2+{x_2}^2)\Big(\xi_3({x_1}^2+{x_2}^2) - 2\xi_1 x_0 x_1 - 2\xi_2 x_0 x_2 + \xi_0 {x_0}^2\Big) = 0. \qquad \square$$

General inversions

It seems to be futile to give an explicit formula for the action of a general inversion in $\mathbb{P}(\mathfrak{Circ})$, because we already showed that we can always reduce the general case to the inversion ε about the unit circle. But without an explicit formula we will not be able to completely unify lines and circles, since at this point the inversion is only defined about finite circles.

Let $\mathcal{C} = \tau\delta(\mathcal{E})$ be a circle with representant $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ and $\iota_\mathcal{C}$ be the inversion about $\mathcal{C}$. We know from theorem (4.13) that

$$\iota_\mathcal{C} = \tau\delta\varepsilon\delta^{-1}\tau^{-1} = \tau\delta^2\varepsilon\tau^{-1}.$$

The final result is

Theorem 5.10 (General inversion in $\mathbb{P}(\mathfrak{Circ})$)

$$\iota_\mathcal{C} : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} ({c_1}^2+{c_2}^2)\xi_0 & -2c_0c_1\xi_1 & -2c_0c_2\xi_2 & +{c_0}^2\xi_3 \\ c_1c_3\xi_0 & +(-{c_1}^2+{c_2}^2-c_0c_3)\xi_1 & -2c_1c_2\xi_2 & +c_0c_1\xi_3 \\ c_2c_3\xi_0 & -2c_1c_2\xi_1 & +({c_1}^2-{c_2}^2-c_0c_3)\xi_2 & +c_0c_2\xi_3 \\ {c_3}^2\xi_0 & -2c_1c_3\xi_1 & -2c_2c_3\xi_2 & +({c_1}^2+{c_2}^2)\xi_3 \end{pmatrix}.$$

Proof
We can show this by calculation of
$$\tau\delta^2\varepsilon\tau^{-1}(\xi),$$
but we omit the intermediate results. After applying the theorems (5.6) and (5.8) and cancelling out some terms we arrive at

$$\begin{pmatrix} {c_0}^6\Big(& ({c_1}^2+c_2)^2\xi_0 & -2c_0c_1\xi_1 & -2c_0c_2\xi_2 & +{c_0}^2\xi_3\Big) \\ {c_0}^6\Big(& c_1c_3\xi_0 & +(-{c_1}^2+{c_2}^2-c_0c_3)\xi_1 & -2c_1c_2\xi_2 & +c_0c_1\xi_3\Big) \\ {c_0}^6\Big(& c_2c_3\xi_0 & -2c_1c_2\xi_1 & +({c_1}^2-{c_2}^2-c_0c_3)\xi_2 & +c_0c_2\xi_3\Big) \\ {c_0}^6\Big(& {c_3}^2\xi_0 & -2c_1c_3\xi_1 & -2c_2c_3\xi_2 & +({c_1}^2+{c_2}^2)\xi_3\Big) \end{pmatrix}.$$

For $c_0 \neq 0$ we can reduce this by the common factor ${c_0}^6$. □

From theorem (5.10) we see that

Theorem 5.11 (Linearity of inversions in $\mathbb{P}(\mathfrak{Circ})$)
Inversions are linear maps on $\mathbb{P}(\mathfrak{Circ})$.

For $\iota_{\mathcal{C}}$ we can write
$$\iota_{\mathcal{C}} : \mathbf{x} \mapsto M\mathbf{x},$$
where the transformation matrix is

$$M = \begin{pmatrix} c_0c_3 & -2c_0c_1 & -2c_0c_2 & {c_0}^2 \\ c_1c_3 & -2{c_1}^2 & -2c_1c_2 & c_0c_1 \\ c_2c_3 & -2c_1c_2 & -2{c_2}^2 & c_0c_2 \\ {c_3}^2 & -2c_1c_3 & -2c_2c_3 & c_0c_3 \end{pmatrix} + ({c_1}^2+{c_2}^2-c_0c_3)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We check by an easy computation that

$$M^2 = ({c_1}^2+{c_2}^2-c_0c_3)^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This implies

Theorem 5.12 (Inversions are involutions)
An inversion in $\mathbb{P}(\mathfrak{Circ})$ is an involution.

Because for a nullcircle
$${c_1}^2+{c_2}^2-c_0c_3 = 0,$$
we can also immediately deduce a statement in $\mathbb{P}(\mathfrak{Circ})$ that we have already shown in $\mathfrak{Circ}$.

Corollary 5.13
No inversion about a nullcircle exists.

5.3 Elementary geometry in $\mathbb{P}(\mathfrak{Circ})$

We have to rediscuss topics like polarity and intersection angle, because the definitions in section (3.3) have been put down for finite circles only.

The polar plane in $\mathbb{P}(\mathfrak{Circ})$

Let

$$\mathcal{L} : a_1x_1 + a_2x_2 + a_0x_0 = 0$$

be a line in $\mathbb{P}_2(\mathbb{C})$. In $\mathbb{P}(\mathfrak{Circ})$ $\mathcal{L}$ is represented by

$$\mathcal{L}^\circ = (0 : a_1 : a_2 : -2a_0)^\circ.$$

The polar plane of $\mathcal{L}^\circ$ in $\mathfrak{Circ}$ should contain the representants of circles $\mathcal{C}$ that cut $\mathcal{L}$ orthogonally. These circles have their center on $\mathcal{L}$ and hence

$$P_{\mathcal{L}^\circ} : a_1\xi + a_2\eta + a_0 = 0.$$

This equation imposes the condition

$$C = (\xi, \eta) \in \mathcal{L}$$

on the center C, the radius of the circle is arbitrary. Now we only have to homogenize this equation.

Definition 5.14 (Polar plane of a line)
The line

$$\mathcal{L} : 2a_1x_1 + 2a_2x_2 - a_3x_0 = 0$$

with representant $\mathcal{L}^\circ = (0 : a_1 : a_2 : a_3)^\circ$ *has the polar plane*

$$P_{\mathcal{L}^\circ} : 2a_1\xi_1 + 2a_2\xi_2 - a_3\xi_0 = 0.$$

For a circle $\mathcal{C}$ with representant $\mathcal{C}^\circ = (1 : c_1 : c_2 : c_3)^\circ$ and center $C = (1 : c_1 : c_2)$ the condition $C \in \mathcal{L}$ is equivalent to $\mathcal{C}^\circ \in P_{\mathcal{L}^\circ}$. This means that the polar plane we have just defined contains all circles whose centers lie on $\mathcal{L}$. These are the circles that intersect with the line $\mathcal{L}$ at a right angle. On the polar plane $P_{\mathcal{C}^\circ}$ in the projective space of circles we also find straight lines with representants $(0 : \xi_1 : \xi_2 : \xi_3)^\circ$. They suffice

$$2a_1\xi_1 + 2a_2\xi_2 = 0.$$

Hence they are of the form $(0 : a_2 : -a_1 : \xi_3)^\circ$ and represent the lines

$$2a_2x_1 - 2a_1x_2 - \xi_3x_0 = 0$$

in $\mathbb{P}_2(\mathbb{C})$. These lines are orthogonal to $\mathcal{L}$.

We see from the above that the polar plane can be defined for lines in the projective space of circles. Since we have observed in section (3.3) that being conjugated is a symmetric relation and since the polar plane of a line contains circles, we should be able to give a useful definition of the polar plane for all elements of $\mathbb{P}(\mathfrak{Circ})$.

Definition 5.15 (Polar plane in $\mathbb{P}(\mathfrak{Circ})$)
Let $\mathcal{C}$ be a projective circle with representant $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$. Then the polar plane of $\mathcal{C}^\circ$ is

$$P_{\mathcal{C}^\circ} : 2c_1\xi_1 + 2c_2\xi_2 - c_3\xi_0 - c_0\xi_3 = 0.$$

Polarity and angles

Definition 5.16 (Polarity)
Let $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ and $\mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ$ be the representants of two projective circles. Then the symmetric function

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 2c_1d_1 + 2c_2d_2 - c_3d_0 - c_0d_3$$

is called polarity *of $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$.*

Theorem 5.17 (Polarity and orthogonality)
For two projective circles $\mathcal{C}$ and $\mathcal{D}$ is equivalent:
(i) $P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 0$,
(ii) *$\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ are conjugated,*
(iii) *$\mathcal{C}$ and $\mathcal{D}$ are orthogonal.*

Theorem 5.18 (Angle)
For the intersection angle α of the circles $\mathcal{C}$ and $\mathcal{D}$ represented by $\mathcal{C}^\circ, \mathcal{D}^\circ \in \mathbb{P}(\mathfrak{Circ})$ holds

$$\cos^2(\alpha) = \frac{P(\mathcal{C}^\circ, \mathcal{D}^\circ)^2}{P(\mathcal{C}^\circ, \mathcal{C}^\circ)P(\mathcal{D}^\circ, \mathcal{D}^\circ)}.$$

Proof
We have to check three cases. The intersection can be between two circles, between a circle and a line or between two lines.

For two regular circles $\mathcal{C}$ and $\mathcal{D}$ we can choose the representants $\mathcal{C}^\circ = (1 : c_1 : c_2 : c_3)^\circ$ and $\mathcal{D}^\circ = (1 : d_1 : d_2 : d_3)^\circ$. The polarity

$$P(\mathcal{C}^\circ, \mathcal{D}^\circ) = 2c_1d_1 + 2c_2d_2 - c_3 - d_3$$

in $\mathbb{P}(\mathfrak{Circ})$ reduces to the polarity in $\mathfrak{Circ}$.

In the second case a circle $\mathcal{C}$ and a straight line $\mathcal{L}$ intersect. The representants are $\mathcal{C}^\circ = (1 : c_1 : c_2 : c_3)^\circ$ and $\mathcal{L}^\circ = (0 : a_1 : a_2 : a_3)^\circ$. Let L be the point on $\mathcal{L}$, where the perpendicular from the center of $\mathcal{C}$ onto $\mathcal{L}$ intersects. In $\mathbb{C}^2$ the line $\mathcal{L}$ and the perpendicular $\mathcal{P}$ are described by the equations

$$\mathcal{L} : 2a_1x_1 + 2a_2x_2 = a_3 \quad \text{and} \quad \mathcal{P} : a_2x_1 - a_1x_2 = a_2c_1 - a_1c_2.$$

The solution $L = (x_1, x_2)$ of this system of linear equations is

$$L = \left(\frac{a_1a_3 + 2{a_2}^2c_1 - 2a_1a_2c_2}{2({a_1}^2 + {a_2}^2)}, \frac{a_2a_3 + 2{a_1}^2c_2 - 2a_1a_2a_1}{2({a_1}^2 + {a_2}^2)} \right).$$

For the cosine of the intersection angle α, the distance $d(C, L)$ from the point C to the line L and the radius $r_\mathcal{C}$ of the circles holds

$$\cos^2(\alpha) = \frac{d(C, L)^2}{{r_\mathcal{C}}^2}. \tag{5.2}$$

We compute the distance

$$d(C,L)^2 = \left(\frac{a_1a_3 - 2{a_1}^2c_1 - 2a_1a_2c_2}{2({a_1}^2 + {a_2}^2)}\right)^2 + \left(\frac{a_2a_3 - 2{a_2}^2c_2 - 2a_1a_2c_1}{2({a_1}^2 + {a_2}^2)}\right)^2$$

which after expanding and cancelling the common factor $({a_1}^2 + {a_2}^2)$ of nominator and denominator becomes

$$d(C,L)^2 = \frac{(2a_1c_1 + 2a_2c_2 - a_3)^2}{4({a_1}^2 + {a_2}^2)}.$$

We use this result in equation (5.2) and obtain

$$\cos^2(\alpha) = \frac{(2a_1c_1 + 2a_2c_2 - a_3)^2}{2{r_{\mathcal{C}}}^2 \cdot 2({a_1}^2 + {a_2}^2)}.$$

Finally let $\mathcal{K}$ and $\mathcal{L}$ be two lines with representants $\mathcal{K}^\circ = (0 : a_1 : a_2 : a_3)^\circ$ and $\mathcal{L}^\circ = (0 : b_1 : b_2 : b_3)^\circ$. We can calculate their intersection angle easily as the angle between their normal vectors

$$\vec{v}_{\mathcal{K}} = (a_1, a_2) \text{ and } \vec{v}_{\mathcal{L}} = (b_1, b_2).$$

Here we use the scalar product which gives us

$$\cos^2(\alpha) = \frac{(a_1b_1 + a_2b_2)^2}{({a_1}^2 + {a_2}^2)({b_1}^2 + {b_2}^2)}. \qquad \square$$

5.4 The inversive group

Reflections and inversions

We have integrated lines into the space of circles. The same unification is possible for the inversions about circles and the reflections about straight lines. Both sorts of maps are involutions and preserve angles. While inversions map lines and circles on lines and circles, reflections keep lines and circles seperate.

Unfortunately we can not extend our definition (4.6) of inversions about circles to reflections about straight lines, because we can not express a line $\mathcal{L}$ as an image $\mathcal{L} = \tau\delta(\mathcal{E})$ of the unit circle $\mathcal{E}$, where δ is a dilation with the origin as center and τ a translation. But we can treat a reflection as the limit of a series of inversions. In the proof of theorem (5.10) we calculated the action of an arbitrary inversion about the circle $\mathcal{C}$ with representant $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$.

We can write a reflection $\iota_{\mathcal{L}}$ about the line $\mathcal{L}$ with representant $\mathcal{L}^\circ = (0 : c_1 : c_2 : c_3)^\circ$ as the limit

$$\iota_{\mathcal{L}} = \lim_{c_0 \to 0} \iota_{\mathcal{C}}$$

of a series of inversions. These are defined by a series of circles of inversion that are given by a series of representants $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ with fixed c_1, c_2 and c_3 and variable c_0. It is possible to show with elementary methods that for all points in $\mathbb{P}(\mathbb{C})$ and for all series of c_0 converging to zero the series of image points under the series of inversions converge to the image point under the reflection about $\mathcal{L}$. We want to omit such a proof here and state

Theorem 5.19 (Action of a reflection on $\mathbb{P}(\mathfrak{Circ})$)
Let $\mathcal{L}$ be a line with representant $\mathcal{L}^\circ = (0 : c_1 : c_2 : c_3)^\circ$. Then the reflection $\iota_{\mathcal{L}}$ about $\mathcal{L}$ acts on $\mathbb{P}(\mathfrak{Circ})$ as

$$\iota_{\mathcal{L}} : \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} ({c_1}^2 + {c_2}^2)\xi_0 \\ (-{c_1}^2 + {c_2}^2)\xi_1 - 2c_1c_2\xi_2 \\ -2c_1c_2\xi_1 + ({c_1}^2 - {c_2}^2)\xi_2 \\ {c_3}^3\xi_0 - 2c_1c_3\xi_1 - 2c_2c_3\xi_2 + ({c_1}^2 + {c_2}^2)\xi_3 \end{pmatrix}.$$

We see immediately from this theorem that the image of a straight line under this kind of map is always a straight line and the image of a circle is always a circle. Moreover the reflection about $\mathcal{L}$ fixes the line $\mathcal{L}$, because

$$\iota_{\mathcal{L}} : \begin{pmatrix} 0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ (-{c_1}^2 + {c_2}^2)c_1 - 2c_1c_2 \cdot c_2 \\ -2c_1c_2 \cdot c_1 + ({c_1}^2 - {c_2}^2)c_2 \\ -2c_1c_3 \cdot c_1 - 2c_2c_3 \cdot c_2 + ({c_1}^2 + {c_2}^2)c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -({c_1}^2 + {c_2}^2)c_1 \\ -({c_1}^2 + {c_2}^2)c_2 \\ -({c_1}^2 + {c_2}^2)c_3 \end{pmatrix}.$$

Remark
It is known from elementary geometry that the reflections about two orthogonal lines commute. This has become a special case within our theory. Inversions commute, when the circles of inversions are orthogonal.

Connection to Möbius transformations

In theorem (5.10) we have shown, that inversions are linear maps on $\mathbb{P}(\mathfrak{Circ})$. Since we have also shown that the map

$$\begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mapsto \begin{pmatrix} {c_1}^2 + {c_2}^2 & -2c_0c_1 & -2c_0c_2 & +{c_0}^2 \\ c_1c_3 & -{c_1}^2 + {c_2}^2 - c_0c_3 & -2c_1c_2 & +c_0c_1 \\ c_2c_3 & -2c_1c_2 & +{c_1}^2 - {c_2}^2 - c_0c_3 & +c_0c_2 \\ {c_3}^2 & -2c_1c_3 & -2c_2c_3 & {c_1}^2 + {c_2}^2 \end{pmatrix} \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

describes inversions about regular circles as well as reflections about straight lines, we can state the following

Lemma 5.20
Every translation, rotation and dilation in $\mathbb{P}(\mathfrak{Circ})$ can be described as a composition of (generalized) inversions.

Proof
For translations and rotations this is obvious, because we may choose appropriate reflections about lines. For a dilation δ with center $C = (c_1, c_2)$ and factor D we may choose the inversions about the circles $\mathcal{C}_1$ with center C and ${r_{\mathcal{C}_1}}^2 = 1$ and the circle $\mathcal{C}_2$ with center C and ${r_{\mathcal{C}_2}}^2 = D$. Then $\delta = \iota(\mathcal{C}_2)\iota(\mathcal{C}_1)$. $\square$

Our results for generalized inversions are very similar to those for Möbius transformations. Möbius transformations are automorphisms of the Riemann sphere $\overline{\mathbb{C}}$. In this space they map generalized circles onto generalized circles, they are suitable to describe translations, rotations dilations and inversions. Moreover they form a group, the so called Möbius group $Möb(\mathbb{C})$, which is isomorphic to $PSL_2(\mathbb{C})$. In some sense we can assume $\overline{\mathbb{C}}$ as the real part of the space $\mathbb{P}_2(\mathbb{C})$ and the Möbius transformations as the induced action on the Riemann sphere by generalized inversions in $\mathbb{P}_2(\mathbb{C})$. More information about Möbius transformations is found in [4].

The transformation matrix M

The matrix

$$M = \begin{pmatrix} {c_1}^2 + {c_2}^2 & -2c_0c_1 & -2c_0c_2 & +{c_0}^2 \\ c_1c_3 & -{c_1}^2 + {c_2}^2 - c_0c_3 & -2c_1c_2 & +c_0c_1 \\ c_2c_3 & -2c_1c_2 & +{c_1}^2 - {c_2}^2 - c_0c_3 & +c_0c_2 \\ {c_3}^2 & -2c_1c_3 & -2c_2c_3 & {c_1}^2 + {c_2}^2 \end{pmatrix}$$

that represents inversions as linear maps has the eigenvalue

$$({c_1}^2 + {c_2}^2 - c_0c_3)$$

with algebraic and geometric multiplicity 3 and the eigenspace generated by

$$\mathcal{C}_1^\circ = \begin{pmatrix} 2c_1 \\ c_3 \\ 0 \\ 0 \end{pmatrix}^\circ, \ \mathcal{C}_2^\circ = \begin{pmatrix} 2c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix}^\circ \text{ and } \mathcal{C}_3^\circ = \begin{pmatrix} -c_0 \\ 0 \\ 0 \\ c_3 \end{pmatrix}^\circ$$

and the eigenvalue

$$-({c_1}^2 + {c_2}^2 - c_0c_3)$$

with algebraic and geometric multiplicity 1 and the corresponding eigenvector

$$\mathcal{C}^\circ = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix}^\circ .$$

The last eigenvector represents the circle of inversion. The polarity of these four circles is

$$P(\mathcal{C}^\circ, {\mathcal{C}_i}^\circ) = 0 \text{ for } i = 1, 2, 3$$

and also

$$P({\mathcal{C}_1}^\circ, {\mathcal{C}_2}^\circ) = 0.$$

We recall from theorem (5.17) that two circles are orthogonal, when their polarity vanishes. Because ${\mathcal{C}_1}^\circ$, ${\mathcal{C}_2}^\circ$ and ${\mathcal{C}_3}^\circ$ belong to the same eigenspace, we can modify ${\mathcal{C}_3}^\circ$ in a way, such that it remains in this eigenspace, but additionally is orthogonal[2] to the other two

[2] It is important to keep in mind that not the vectors of the representants, but the represented circles are to be orthogonal to each other.

eigencircles. The result is

$$\mathcal{C}_3{}^\circ = \begin{pmatrix} 2{c_1}^2 + 2{c_2}^2 - c_0c_3 \\ c_1c_3 \\ c_2c_3 \\ {c_3}^2 \end{pmatrix}^\circ .$$

From the above we deduce

Theorem 5.21 (Inversion in diagonal form)
Let M be the describing matrix of the inversion about $\mathcal{C}$ in $\mathbb{P}(\mathfrak{Circ})$. Then exists a regular matrix

$$T = (\mathcal{C}^\circ, \mathcal{C}_1{}^\circ, \mathcal{C}_2{}^\circ, \mathcal{C}_3{}^\circ)\,,$$

such that

$$T^t M T = ({c_1}^2 + {c_2}^2 - c_0c_3) \cdot \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

6 Linear families of circles

6.1 Lines in $\mathbb{P}(\mathfrak{Circ})$

In the preceding chapter we discussed the unification of circles and lines and called the conjunction *generalized circles*. We succeeded in developing an exact method for treating reflections about lines as inversions about infinitely large circles. In this chapter we shall finally examine the geometry of the space $\mathbb{P}(\mathfrak{Circ})$ itself.

The representants of circles are basically points in the space of circles $\mathfrak{Circ}$ or, if we speak of generalized circles, points in $\mathbb{P}(\mathfrak{Circ})$. This enables us to define a set of circles with certain properties through its set of representants. We already know such sets. The plane that touches the paraboloid Π of nullcircles in the point $\mathcal{C}^\circ$, for example, contains the representants of all those circles that are orthogonal to $\mathcal{C}$, i.e. run through the center of $\mathcal{C}$.[1] A more general example is the polar plane $P_{\mathcal{C}^\circ}$ of a given representant $\mathcal{C}^\circ$. It contains the representants $\mathcal{D}^\circ$ of circles with $\mathcal{C} \perp \mathcal{D}$. The tangential plane is, as we remarked in corollary (3.13), the special case of the polar plane $P_{\mathcal{C}^\circ}$, when the point $\mathcal{C}^\circ$ represents a nullcircle and thus lies on Π.

Linear families of circles

The simplest non-empty set of points in $\mathbb{P}(\mathfrak{Circ})$ that we can choose is the set $\{\mathcal{C}^\circ\}$. It represents the circle with given center and square of the radius. If we drop the second condition, we obtain a one parameter family of circles with a given center but variable radius.

Definition 6.1 (Concentric family of circles)
Let $C = (c_1, c_2) \in \mathbb{C}^2$ be a point and $\mathfrak{F}$ be the family of circles with center C. Then the set of representants

$$\mathfrak{F}^\circ = \{\mathcal{C}^\circ \in \mathbb{P}(\mathfrak{Circ}) \mid \mathcal{C}^\circ = (\lambda_0 : \lambda_0 c_1 : \lambda_0 c_2 : \lambda_1)^\circ,\ (\lambda_0 : \lambda_1) \in \mathbb{P}(\mathbb{C})\}$$

is a line in $\mathbb{P}(\mathfrak{Circ})$. The represented circles form the concentric family $\mathfrak{F}$.

Remark
Every concentric family contains the representant $(0 : 0 : 0 : 1)^\circ$. This is the representant of the line $\mathcal{L}_0$ at infinity.

Concentric families are only a special case of linear families of circles.

[1]These two properties are equivalent for nullcircles.

Definition 6.2 (Linear family of circles)
Let $\mathcal{C}$ and $\mathcal{D}$ be two different circles with representants $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$. The line

$$\mathfrak{L}^\circ := \{\lambda_0\mathcal{C}^\circ + \lambda_1\mathcal{D}^\circ, (\lambda_0 : \lambda_1) \in \mathbb{P}(\mathbb{C})\}$$

in $\mathbb{P}(\mathfrak{Circ})$ represents the linear family *$\mathfrak{L}$ of circles generated by $\mathcal{C}$ and $\mathcal{D}$.*

If at least one of the generalized circles $\mathcal{C}$ and $\mathcal{D}$ is not a straight line, we want to call the family *regular* and otherwise *singular*.

Corollary 6.3
It is always possible to write a concentric family of circles as the set of linear combinations

$$\mathfrak{L}^\circ := \{\lambda_0{\mathcal{C}_0}^\circ + \lambda_1{\mathcal{L}_0}^\circ, (\lambda_0 : \lambda_1) \in \mathbb{P}(\mathbb{C})\},$$

where $\mathcal{C}_0$ is the nullcircle of the family and $\mathcal{L}_0$ the line at infinity.

A linear family of circles is also called a *pencil* of circles. We are going to deal with linear as well as non-linear families of circles later, hence we will stick with the notation *linear family.*

Lemma 6.4
A regular linear family of circles contains exactly one line.

Proof
For $(c_0, d_0) \neq (0, 0)$ the equation

$$\lambda_0 c_0 + \lambda_1 d_0 = 0$$

has exactly one solution, namely

$$(\lambda_0 : \lambda_1) = (d_0 : -c_0) \in \mathbb{P}(\mathbb{C}).$$

Hence $d_0\mathcal{C}^\circ - c_0\mathcal{D}^\circ$ represents the only line in the linear family generated by $\mathcal{C}$ and $\mathcal{D}$. □

Definition 6.5 (Axis)
The uniquely defined straight line of a regular linear family of circles is called axis *of the family.*

Theorem 6.6 (Axis of a linear family)
Let $\mathfrak{L}$ be a regular linear family of circles generated by $\mathcal{C}$ and $\mathcal{D}$ with representants $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ and $\mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ$, where $(c_0 : d_0) \in \mathbb{P}(\mathbb{C})$. Then the axis $\mathcal{L}$ of $\mathfrak{L}$ has the representant

$$\mathcal{L}^\circ = d_0\mathcal{C}^\circ - c_0\mathcal{D}^\circ.$$

Proof
Be $\mathcal{L}^\circ = (l_0 : l_1 : l_2 : l_3)^\circ$. Then $l_0 = d_0c_0 - c_0d_0 = 0$. □

Corollary 6.7
Every linear family of circles contains at least one line.

When we take lines in $\mathfrak{Circ}$, then every linear family of circles is automatically regular, because we can not choose lines as generators. Nonetheless each of these families implicitly contains an exceptional circle with infinitely large radius. Its representant is the infinitely distant point on the line representing the family. This example underlines that it is useful to work with representants from $\mathbb{P}(\mathfrak{Circ})$. In [44] Pedoe uses representants from $\mathfrak{Circ}$. Hence he can write the equation of the (radical) axis of two circles $\mathcal{C}$ and $\mathcal{D}$ simply as the line represented by $(\mathcal{C}^\circ - \mathcal{D}^\circ)$. In his notation the term $(x^2 + y^2)$ always cancels out. This seems to be an advantage now, but soon we will see that it is rather a drawback.

A special case of linear family arises, when its generators are concentric. At a first glance a concentric family does not seem to contain a line. But if we broaden our view to generalized circles, then the axis of every concentric family is the line at infinity. It is even possible that a linear family of circles does not contain any circle at all, but only lines. This happens when both generating circles are lines, which means that the generated family is singular. Singular families can not be treated properly with the notation used in [44], of course.

Base points

Lemma 6.8
If $P = (p_0 : p_1 : p_2) \in \mathbb{P}_2(\mathbb{C})$ lies on two different circles of a linear family $\mathfrak{L}$, then P lies on all circles of $\mathfrak{L}$.

Proof
For $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ with defining polynomial p, $\mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ$ with defining polynomial q and arbitrary $(\lambda_0 : \lambda_1) \in \mathbb{P}(\mathbb{C})$ the representant

$$\lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ$$

belongs to the linear family defined by $\mathcal{C}$ and $\mathcal{D}$. Now we assume $P \in \mathcal{C} \cap \mathcal{D}$, which means

$$p(p_0, p_1, p_2) = q(p_0, p_1, p_2) = 0.$$

But then also

$$\lambda_0 p(p_0, p_1, p_2) + \lambda_1 q(p_0, p_1, p_2) = 0,$$

hence P lies on the circle represented by

$$\lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ. \qquad \square$$

Definition 6.9 (Base point)
A point $P \in \mathbb{P}_2(\mathbb{C})$ is called base point *of the linear family $\mathfrak{L}$ of circles, if P lies on every circle in $\mathfrak{L}$.*

From lemma (6.8) now follows that the base points of a linear family $\mathfrak{L}$ are determined by its pair of generators $\mathcal{C}$ and $\mathcal{D}$.

Theorem 6.10 (Base points of a linear family)
Let $\mathfrak{L}$ be a linear family of circles generated by $\mathcal{C}$ and $\mathcal{D}$. Then the base points of $\mathfrak{L}$ are the intersection points $\mathcal{C} \cap \mathcal{D}$.

Proof
The set of common points $\mathcal{C} \cap \mathcal{D}$ is, according to lemma (6.8), contained in the set of base points. If $\mathfrak{L}$ had another base point $P \notin \mathcal{C} \cap \mathcal{D}$, then P would lie on every circle in $\mathfrak{L}$, in particular on $\mathcal{C}$ and $\mathcal{D}$, which is a contradiction. □

A linear family of circles in $\mathbb{P}_2(\mathbb{C})$ has two base points in general, but these points may also coincide. This happens, when the family is generated by circles that touch each other. Without formal proof we shall be convinced that in such a family all circles touch each other in the single base point.

Even though we restrict ourselves to circles that have representants with real coordinates, the two generators of a linear family do not have to intersect in points with real coordinates. The base points of such a linear family can be a pair of real points, a pair of complex conjugated points or one single real point. A more thorough discussion of the real part follows in section (6.2).

Nullcircles

Another important property of a linear family is whether it contains nullcircles. A nullcircle of $\mathfrak{L}$ is represented by a point in the intersection of $\mathfrak{L}^\circ$ and Π. There are two different situations for this. Either the line intersects in two distinct points or the line touches Π. In the first case the family contains two, in the latter case only one nullcircle. When we are studying linear families in $\mathfrak{Circ}$, the first case further divides in two subcases, depending on whether $\mathfrak{L}^\circ$ intersects Π in two finite points or in one finite point and one point at infinity.

Again the above only holds for complex circles, because we are using Bezout's Theorem. When we impose our restriction that the representants have real coordinates, the intersection of $\mathfrak{L}^\circ$ and Π may be a pair of real points, a pair of complex conjugated points or one single real point in $\mathbb{P}(\mathfrak{Circ})$. It was said before that a more thorough discussion of the real part follows.

We can not understand the meaning of *intersecting with* Π *in a point at infinity* within the naive concept of nullcircles in $\mathbb{P}_2(\mathbb{C})$. In $\mathbb{P}(\mathfrak{Circ})$ a line can intersect with Π at infinity only in the point $(0:0:0:1)^\circ$, where Π touches the hyperplane at infinity. This point represents the line at infinity, thus the line at infinity is also a nullcircle.

6.2 The conjugated family

Definition

Every circle $\mathcal{C}$ of a linear family $\mathfrak{L}$ defines the polar plane $P_{\mathcal{C}^\circ}$ of its representant $\mathcal{C}^\circ$ in the projective space of circles $\mathbb{P}(\mathfrak{Circ})$. From corollary (3.22) we know that $P_{\mathcal{C}^\circ}$ contains exactly the points representing circles perpendicular to $\mathcal{C}$.

Theorem 6.11 (Circle perpendicular to a linear family)
If a circle is perpendicular to two different circles of a linear family $\mathfrak{L}$, then it is perpendicular to all circles of $\mathfrak{L}$.

Proof
Let $\mathcal{C}$ and $\mathcal{D}$ be two different circles in the linear family $\mathfrak{L}$ and $\mathcal{P}$ be a circle perpendicular to $\mathcal{C}$ and $\mathcal{D}$. Then the representant $\mathcal{P}^\circ$ lies on both polar planes $P_{\mathcal{C}^\circ}$ and $P_{\mathcal{D}^\circ}$ and we have

$$P(\mathcal{C}^\circ, \mathcal{P}^\circ) = P(\mathcal{D}^\circ, \mathcal{P}^\circ) = 0.$$

Every circle in $\mathfrak{L}$ has a representant of the form $\lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ$ with $(\lambda_0, \lambda_1) \in \mathbb{P}(\mathbb{C})$, thus

$$P(\lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ, \mathcal{P}^\circ) = \lambda_0 P(\mathcal{C}^\circ, \mathcal{P}^\circ) + \lambda_1 P(\mathcal{D}^\circ, \mathcal{P}^\circ) = 0.$$

Hence $\mathcal{P}^\circ$ lies on the polar plane of every point on $\mathfrak{L}^\circ$ and is therefore perpendicular to every circle in $\mathfrak{L}$. $\square$

This observation can also be expressed in geometrical terms.

Corollary 6.12
Let $\mathcal{C}$ and $\mathcal{D}$ be two different circles that generate the linear family $\mathfrak{L}$ of circles. Let $\mathcal{P}^\circ$ be the representant of an arbitrary circle $\mathcal{P} \in \mathfrak{L}$. Then the intersection $P_{\mathcal{C}^\circ} \cap P_{\mathcal{D}^\circ}$ is also contained in the polar plane $P_{\mathcal{P}^\circ}$.

Since $\mathbb{P}(\mathfrak{Circ})$ is a projective space, the intersection of arbitrary planes is never empty.

Definition 6.13 (Conjugated family of circles)
Let $\mathfrak{L}$ be a linear family of circles. The line

$$\bigcap_{\mathcal{C}^\circ \in \mathfrak{L}^\circ} P_{\mathcal{C}^\circ}$$

that lies on all polar planes of representants $\mathcal{C}^\circ$ in $\mathfrak{L}^\circ$ represents the conjugated family *of $\mathfrak{L}$.*

It is clear from the above that one does not have to take the intersection over all representants $\mathcal{C}^\circ \in \mathfrak{L}^\circ$, but that it suffices to take two different representants on $\mathfrak{L}^\circ$.

A simple, but important fact is

Lemma 6.14 (Existence of the conjugated family)
Let $\mathcal{C}$ and $\mathcal{D}$ be different circles. Then infinitely many circles exist that are perpendicular to both of them. In particular there are always at least two circles with this property.

This lemma implies that the conjugated family is indeed a linear family of circles.

Theorem 6.15 (Generation of the conjugated family)
Let $\mathcal{C}$ and $\mathcal{D}$ be two different circles that are both perpendicular to all circles of the linear family $\mathfrak{L}$. Then the linear family generated by $\mathcal{C}$ and $\mathcal{D}$ contains exactly the circles that are perpendicular to the circles in $\mathfrak{L}$.

Proof
Lemma (6.14) ensures the existence of such circles $\mathcal{C}$ and $\mathcal{D}$. For any circle $\mathcal{P} \in \mathfrak{L}$ we have

$$P(\mathcal{P}^\circ, \mathcal{C}^\circ) = P(\mathcal{P}^\circ, \mathcal{D}^\circ) = 0.$$

Every circle in the linear family of circles that is generated by $\mathcal{C}$ and $\mathcal{D}$ has a representant of the form $\lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ$ with $(\lambda_0, \lambda_1) \in \mathbb{P}(\mathbb{C})$. We observe, that

$$P(\mathcal{P}^\circ, \lambda_0 \mathcal{C}^\circ + \lambda_1 \mathcal{D}^\circ) = \lambda_0 P(\mathcal{P}^\circ, \mathcal{C}^\circ) + \lambda_1 P(\mathcal{P}^\circ, \mathcal{D}^\circ) = 0. \quad \square$$

The last two theorems hold, because the polarity function is linear in both arguments. In the proofs we can see that theorem (6.11) depends on the linearity of P in the first argument, theorem (6.15) depends on the linearity of P in the second argument.

Theorem 6.16 (Symmetry)
For two linear families $\mathfrak{K}$ and $\mathfrak{L}$ are equivalent:
(i) $\mathfrak{L}$ *is the conjugated family of* $\mathfrak{K}$,
(ii) $\mathfrak{K}$ *is the conjugated family of* $\mathfrak{L}$.

Proof
The equivalence is obvious. The conditions can be written as
(i) $\forall \mathcal{C} \in \mathfrak{K}\ \forall \mathcal{D} \in \mathfrak{L} : \mathcal{C} \perp \mathcal{D}$,
(ii) $\forall \mathcal{D} \in \mathfrak{L}\ \forall \mathcal{C} \in \mathfrak{K} : \mathcal{C} \perp \mathcal{D}$.
Both conditions are obviously equivalent to

$$\forall (\mathcal{C}, \mathcal{D}) \in \mathfrak{K} \times \mathfrak{L} : \mathcal{C} \perp \mathcal{D}. \qquad \square$$

Corollary 6.17
Let $\mathfrak{K}$ and $\mathfrak{L}$ be conjugated linear families of circles. Then any two of the following statements imply the third:
(i) $\mathcal{C} \in \mathfrak{K}$,
(ii) $\mathcal{D} \in \mathfrak{L}$,
(iii) $\mathcal{C} \perp \mathcal{D}$.

A linear family of circles and its conjugated family are closely connected. Corollary (6.17) immediately implies

Theorem 6.18 (Base points and nullcircles of conjugated families)
Let $\mathfrak{K}$ and $\mathfrak{L}$ be linear families of circles. Then every base point of $\mathfrak{K}$ is a nullcircle of $\mathfrak{L}$ and every nullcircle of $\mathfrak{K}$ is a base point of $\mathfrak{L}$.

Proof
A base point P of $\mathfrak{K}$ lies on every circle of $\mathfrak{K}$ and is, when regarded as nullcircle $\mathcal{P}$, perpendicular to every circle of $\mathfrak{K}$. Hence it is a circle of $\mathfrak{L}$. If, on the other hand, $\mathcal{P}$ is a nullcircle of $\mathfrak{K}$, then every circle perpendicular to $\mathcal{P}$ runs through its center P. Hence the point P lies on every circle of $\mathfrak{L}$. $\square$

Regular configurations

The real part of a linear family $\mathfrak{K}$ of circles and its conjugated family $\mathfrak{L}$ can appear in a number of different ways. We should recall that, when seen as generalized circles in $\mathbb{P}_2(\mathbb{C})$, there is no difference between regular and singular configurations. We can use two properties in order to distinguish different types of linear families of circles. Firstly, whether a family is regular or not, and secondly, how many real base points respectively nullcircles the family contains. We have shown that a regular family may contain none, one or two real nullcircles. These are at the same time the base points of the conjugated family.

When we assume $\mathfrak{K}$ to have two different nullcircles, then the conjugated family $\mathfrak{L}$ has two base points. This case is shown in figure (6.1). The situation remains essentially the

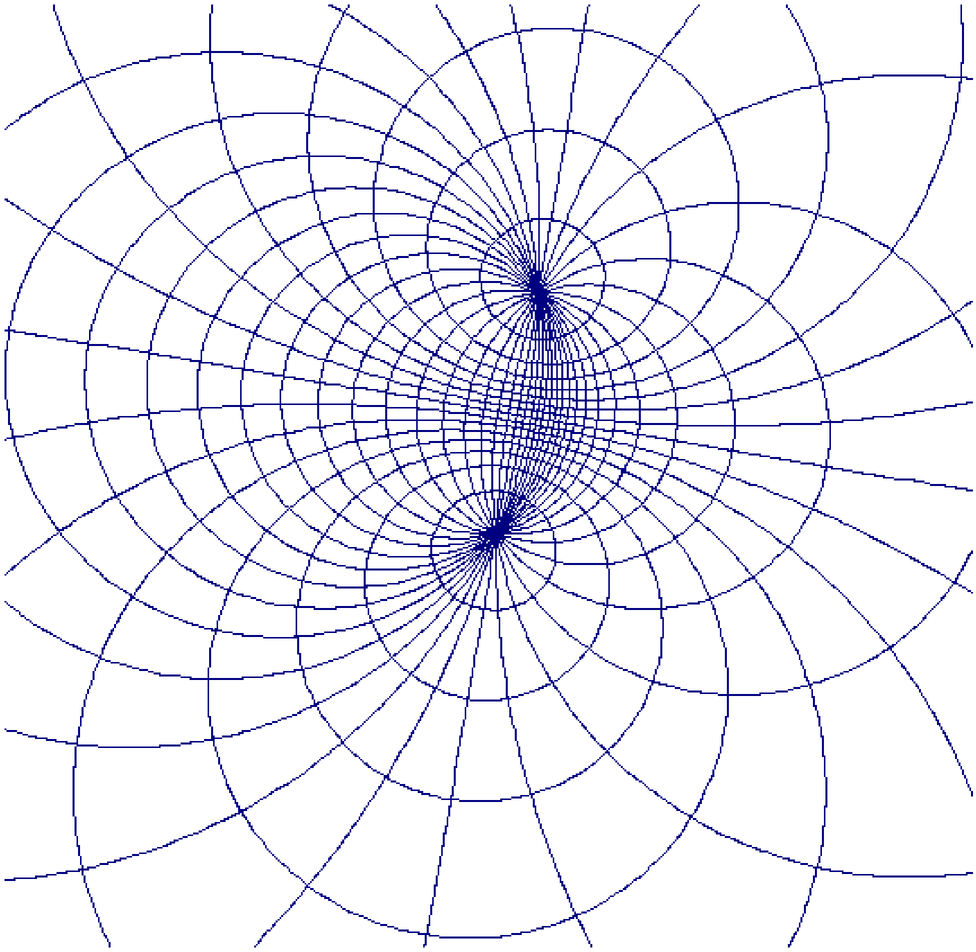


Figure 6.1: Two conjugated families of circles, one with two real base points and one without real base points

same after exchanging the roles of $\mathfrak{K}$ and $\mathfrak{L}$, i.e. by assuming $\mathfrak{L}$ as the original and $\mathfrak{K}$ as the conjugated family. When $\mathfrak{K}$ has only one base point, then this point is also the only base point of the conjugated family $\mathfrak{L}$. We know that this base point is the only nullcircle of both families. This configuration is depicted in figure (6.2).

Singular configurations

For singular linear families we have shown that they contain the line at infinity. When we generalize the idea of base points and nullcircles to singular families we arrive at two different sorts of singular linear families. The first type has two different base points, while the second type has only one base point.

Definition 6.19 (Radial family)
A singular linear family of circles with a finite base point is called a radial family *of circles.*

Theorem 6.20 (Conjugated family of a radial family)

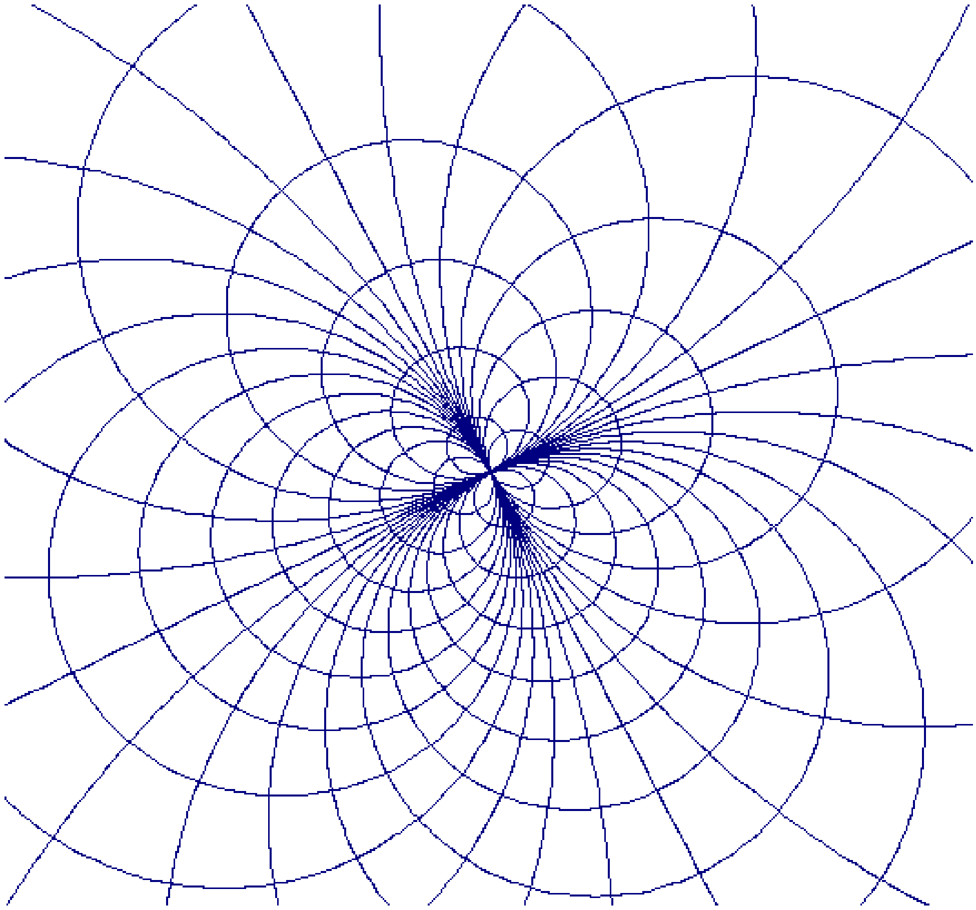


Figure 6.2: Two conjugated families of circles, both with one base point

The conjugated family of a radial family of circles with finite base point P is a concentric family whose members have their center at P.

A potential second base point of a radial family must also be a nullcircle of the concentric family. But there is only one finite nullcircle in this family of circles. Hence the other nullcircle has to be the line at infinity. We can immediately deduce

Corollary 6.21
The line at infinity belongs to every concentric family of circles and is a base point of every radial family of circles.

This important and counterintuitive observation also arises with the third class of singular families.

Definition 6.22 (Parallel family)
A linear family of circles without a finite base point is called a parallel family *of circles.*

We prefer consistence here, even though it is of course confusing that a parallel family of circles does not contain regular circles at all. A parallel family of circles consists only of straight lines.

Theorem 6.23 (Conjugated family of a parallel family)
The conjugated family of a parallel family of circles is again a parallel family of circles.

Since every linear family of circles contains at least one base point, we have

Theorem 6.24 (Base point of a parallel family)
The only base point of a parallel family, which is at the same time also its only nullcircle, is the line at infinity.

Corollary 6.25
Every singular family has the line at infinity as a base point.

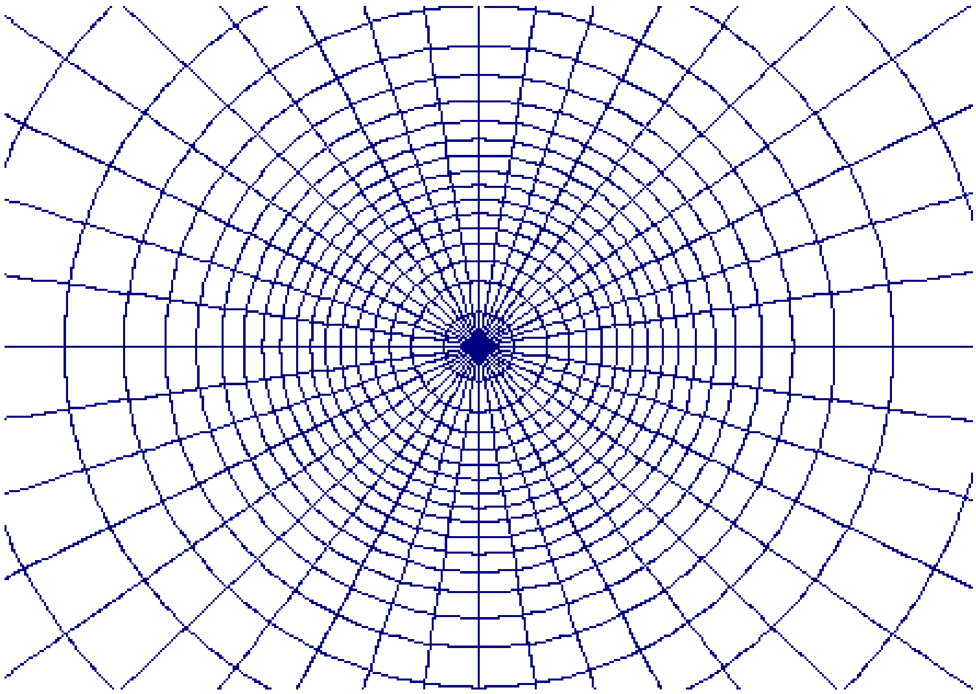


Figure 6.3: Two conjugated families of circles, radial and concentric

The real part of the configurations of singular families are shown for a radial family and the conjugated concentric family in figure (6.3) and for a pair of conjugated parallel families in (6.4). These singular cases are even more common than the regular cases, because the radial-concentric pair and the parallel-parallel pair is what we know as polar coordinates and cartesian coordinates respectively. A pair of conjugated regular families of circles in general position leads to so called bipolar coordinates.

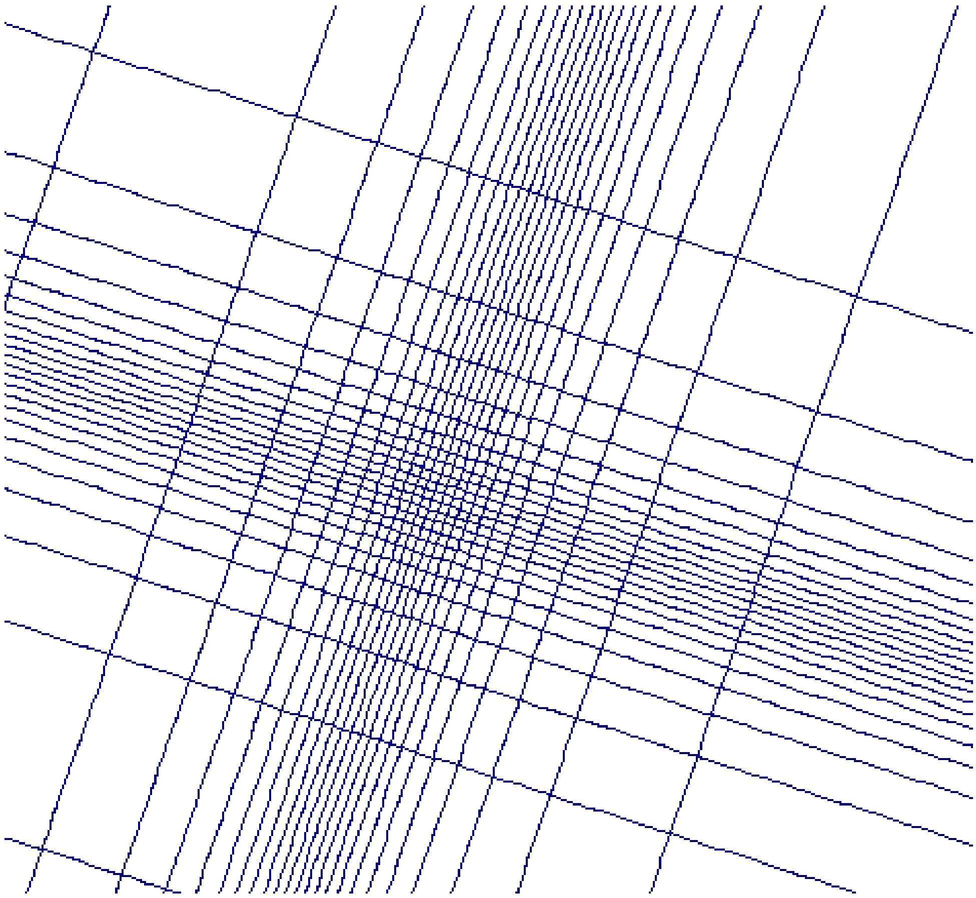


Figure 6.4: Two conjugated parallel families of circles

7 Inversion of a linear family

7.1 Inversion about a circle of the family

We shall examine in this section which effect an inversion has on linear families of circles. Because inversions are bijective linear maps on $\mathbb{P}(\mathfrak{Circ})$, they have the following basic properties:

- Inversions map linear families of circles on linear families of circles.
- Since inversions preserve angles, a pair of conjugated linear families is mapped on a pair of conjugated families.
- An inversion does not change the number of base points and nullcircles of a linear family.

Stabilizing inversions

At the beginning we want to identify the inversions that stabilize a linear family.

Definition 7.1 (Stabilizer)
The bijective map φ is said to stabilize *a set X, when $\varphi(X) = X$ holds. The set of maps that stabilize X is called* stabilizer *of X.*

From theorem (4.21) we already know that the stabilizer of a regular circle $\mathcal{C}$ decomposes into the circle itself and the set of circles that is represented by the polar plane $P_{\mathcal{C}^\circ}$ of its representant $\mathcal{C}^\circ$.

Lemma 7.2 (Trivial action on a linear family)
Let $\mathfrak{K}$ and $\mathfrak{L}$ be two conjugated families and ι be the inversion about the circle $\mathcal{I}$. Then holds

$$\mathcal{I} \in \mathfrak{L} \Longleftrightarrow \forall \mathcal{K} \in \mathfrak{K} : \iota(\mathcal{K}) = \mathcal{K},$$

i.e. the restriction of ι on $\mathfrak{K}$ is the identity map.

Proof
It is obvious that $\mathcal{I} \in \mathfrak{L}$ implies that $\mathcal{I}$ intersects all circles of $\mathfrak{K}$ at right angle. In consequence all circles of $\mathfrak{K}$ are invariant under ι.

If conversely for every circle $\mathcal{K} \in \mathfrak{K}$ holds $\iota(\mathcal{K}) = \mathcal{K}$, then according to theorem (4.21) the circle $\mathcal{I}$ is either identical or orthogonal to all circles of $\mathfrak{K}$. It can not be identical to all these circles and thus belongs to the conjugated family $\mathfrak{L}$. $\square$

Theorem 7.3 (Stabilizing inversions)
Let $\mathfrak{K}$, $\mathfrak{L}$, ι and $\mathcal{I}$ be as in lemma (7.2). Then $\iota(\mathfrak{K}) = \mathfrak{K}$ holds, if and only if $\mathcal{I} \in (\mathfrak{K} \cup \mathfrak{L})$.

Proof
The case that every circle of $\mathfrak{K}$ is individually invariant under ι is equivalent to $\mathcal{I} \in \mathfrak{L}$ according to theorem (7.2). Now let us assume that there exists a circle $\mathcal{K}$ with $\iota(\mathcal{K}) \neq \mathcal{K}$. Because $\iota(\mathcal{K}) \in \mathfrak{K}$, the circles $\mathcal{K}$ and $\iota(\mathcal{K})$ intersect in the base points of $\mathfrak{K}$. These intersection points also lie on $\mathcal{I}$, hence $\mathcal{I}$ runs through the base points of $\mathfrak{K}$. If $\mathfrak{K}$ has two base points, we have already shown that $\mathcal{I} \in \mathfrak{K}$. If $\mathfrak{K}$ has only one base point, then $\mathcal{I}$ has not only this point in common with $\mathcal{K}$ and $\iota(\mathcal{K})$, but also must have the same tangent as $\mathcal{K}$ and $\iota(\mathcal{K})$ in this point. Hence $\mathcal{I} \in \mathfrak{K}$ in this case, too.

If on the other hand $\mathcal{I} \in \mathfrak{K}$, then the intersection points of $\mathcal{I}$ and $\mathcal{K}$ are base points of $\mathfrak{K}$. At the same time these points must also lie on $\iota(\mathcal{K})$, the image circle $\iota(\mathcal{K})$ is therefore contained in $\mathfrak{K}$, if $\mathfrak{K}$ has two base points. If it has only one base point, we again take into account the common tangent of $\mathcal{I}$ and $\mathcal{K}$ which is also a tangent to $\iota(\mathcal{K})$. Hence $\iota(\mathcal{K}) \in \mathfrak{K}$. $\square$

Representation of the inversion

For a linear family $\mathfrak{K}$ of circles and a circle $\mathcal{I}$ of this family the inversion ι about $\mathcal{I}$ induces an action on $\mathfrak{K}$. In order to examine this action we choose generators $\mathcal{C}$ and $\mathcal{D}$, such that

$$\mathfrak{K}^\circ = \{\lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ, (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\}.$$

The inversion ι acts linearly on $\mathbb{P}(\mathfrak{Circ})$, in particular it acts linearly on $\mathfrak{K}^\circ$. Therefore the action of ι on $\mathfrak{K}^\circ$ is already determined by its action on the two generators $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ of the family.

Additionally we do not have to regard the action of ι on entire $\mathbb{P}(\mathfrak{Circ})$. There is a bijection between $\mathbb{P}(\mathbb{C})$ and the points on $\mathfrak{K}^\circ \subset \mathbb{P}(\mathfrak{Circ})$. It suffices to examine the action of ι on the projective coefficient $(\lambda : \mu)$.

Theorem 7.4 (Inversive action from within a linear family)
Let

$$\mathfrak{K}^\circ = \{\lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ, (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\}$$

be a line in $\mathbb{P}(\mathfrak{Circ})$, $\mathcal{I} \in \mathfrak{K}$ a circle and ι the inversion about $\mathcal{I}$. Then the action of ι on $\mathfrak{K}$ is described by the action of ι on $\mathfrak{K}^\circ$ via the map

$$\lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ \mapsto \lambda'\mathcal{C}^\circ + \mu'\mathcal{D}^\circ$$

with

$$\begin{pmatrix} \lambda' \\ \mu' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix},$$

where the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *contains the coefficients of the images*

$$\iota(\mathcal{C}^\circ) = a\mathcal{C}^\circ + c\mathcal{D}^\circ$$

and

$$\iota(\mathcal{D}^\circ) = b\mathcal{C}^\circ + d\mathcal{D}^\circ.$$

Proof
It has been discussed above that the action of ι is entirely determined by the coefficients $(a : c)$ and $(b : d)$ of the generators $\iota(\mathcal{C}^\circ)$ and $\iota(\mathcal{D}^\circ)$. It is clear then that the linear map ι may be described by the given matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. □

Corollary 7.5
The matrix that describes the action of ι on the set $\mathbb{P}(\mathbb{C})$ of coefficients is regular.

Proof
The matrix can not be singular, because if it was, its columns would be linear dependent. This would be the case, if either $a = c = 0$ or $b = d = 0$ or if the columns were multiples of each other. The first two cases are impossible because of $(a : c) \in \mathbb{P}(\mathbb{C})$ and $(b : d) \in \mathbb{P}(\mathbb{C})$. The third case would yield $\iota(\mathcal{C}^\circ) = \iota(\mathcal{D}^\circ)$. That implies $\mathcal{C}^\circ = \mathcal{D}^\circ$, which is also impossible, because $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ generate $\mathfrak{K}^\circ$ and are therefore different. □

The action of the inversion ι on the set of coefficents $\mathbb{P}(\mathbb{C})$ is better known by the term *Möbius transformation*. The identification $\mathbb{P}(\mathbb{C}) \cong \overline{\mathbb{C}}$ is done by the isomorphism

$$\begin{array}{ccc} \mathbb{P}(\mathbb{C}) & \to & \overline{\mathbb{C}} \\ (\lambda : \mu) & \mapsto & \frac{\lambda}{\mu}, \end{array}$$

which is the standard correspondence between projective and affine coordinates. We will refer to this possibility of description. We conclude this paragraph with

Definition 7.6 (Describing matrix)
Let $\mathfrak{K}$ and ι be as above. Then the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ that describes the action of ι on $\mathfrak{K}$ is called the describing matrix *of ι on $\mathfrak{K}$.*

Fixed points – eigencircles

In the last paragraph we have introduced the possibility to describe an inversion about a circle of a linear family of circles by the means of Möbius transformations. An elementary result about these maps (which is shown in [4] for example) is

Theorem 7.7 (Fixed points of Möbius transformations)
Let φ be a Möbius transformation. Then equivalent are:
(i) φ is the identity map,
(ii) φ has (at least) three fixed points.

We are interested in the number of eigenvectors of the describing matrix of ι. ι is an inversion about a circle of $\mathfrak{K}$ and hence not the identity map on $\mathfrak{K}^\circ$. Moreover we showed in (7.5) that the matrix is regular, hence all of its eigenvectors are different from 0. The number of eigenvectors is the same as the number of fixed points of ι in $\mathfrak{K}^\circ$ and the number of circles in $\mathfrak{K}$ that are invariant under ι.

Theorem 7.8 (Invariant circles of the family)
Let $\mathfrak{K}$ be a linear family of circles, $\mathcal{I}$ a circle of $\mathfrak{K}$ and ι the inversion about $\mathcal{I}$. Then there are exactly two circles in $\mathfrak{K}$, that are invariant under ι.

Proof
Obviously one circle in $\mathfrak{K}$ that is invariant under ι is $\mathcal{I}$ itself. Because $\mathcal{I}$ is not a nullcircle, the line $\mathfrak{K}^\circ$ does not lie on the polar plane $P_{\mathcal{I}^\circ}$ of $\mathcal{I}^\circ$. Hence there exists exactly one intersection point $\mathcal{J}^\circ$ of the line and the polar plane. $\mathcal{J}^\circ$ represents a circle in $\mathfrak{K}$ that is orthogonal to $\mathcal{I}$ and thus invariant under ι. □

Corollary 7.9 (Eigenvectors of the describing matrix)
Let ι and $\mathfrak{K}$ be like above. Then the describing matrix of the action of ι on $\mathfrak{K}$ has two different eigenvectors.

Corollary 7.10 (Fixed points of ι)
The representation of the inversion ι in $\mathbb{P}(\mathbb{C})$ has exactly two fixed points.

The above gives rise to the following

Definition 7.11 (Eigencircles)
The circles of a linear family $\mathfrak{K}$ that are invariant under the inversion ι are called eigencircles *of ι in $\mathfrak{K}$.*

It is obvious that the action of ι on $\mathfrak{K}$ can be described in its simplest form, when the generators of $\mathfrak{K}$ are eigencircles of ι.

Theorem 7.12 (Diagonalization)
Let $\mathfrak{K}$, ι and $\mathcal{I}$ be like above. Let further $\mathcal{J}$ be the second eigencircle of ι in $\mathfrak{K}$. When the linear family is expressed as

$$\mathfrak{K}^\circ = \{\lambda \mathcal{I}^\circ + \mu \mathcal{J}^\circ, (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\},$$

the describing matrix of ι is diagonal.

Lemma 7.13 (Orthogonal circle in the family)
Let $\mathfrak{K}$ be a linear family of circles and $\mathfrak{L}$ be its conjugated family. For every circle $\mathcal{C} \notin \mathfrak{L}$ exists a unique circle $\mathcal{P} \in \mathfrak{K}$ with $\mathcal{C} \perp \mathcal{P}$.

Proof
If the polar plane $P_{\mathcal{C}^\circ}$ of $\mathcal{C}^\circ$ contained the line $\mathfrak{K}^\circ$, $\mathcal{C}^\circ$ would lie on $\mathfrak{L}^\circ$. Hence the plane does not contain $\mathfrak{K}^\circ$, which means that $P_{\mathcal{C}^\circ}$ and $\mathfrak{K}^\circ$ have exactly one point in common. This point is the representant $\mathcal{P}^\circ$ of the circle in question. □

This proof has already been carried out in the less general situation of theorem (7.8), where the circle $\mathcal{C}$ belonged to the family $\mathfrak{K}$. In that situation the circle $\mathcal{C}$ was a circle of inversion and hence a regular circle. In the current situation $\mathcal{C}$ could be a nullcircle of $\mathfrak{K}$. If it was the only nullcircle and hence also a base point of the family, then every circle of $\mathfrak{K}$ would cut the singular circle $\mathcal{C}$ orthogonally and there would be no longer uniqueness. But this exceptional case is prevented by the condition $\mathcal{C} \notin \mathfrak{L}$, because if $\mathcal{C}$ was the only nullcircle of $\mathfrak{K}$, it would be the only nullcircle of $\mathfrak{L}$ at the same time.

Corollary 7.14
The inversion about a regular circle $\mathcal{C}$ of the linear family $\mathfrak{K}$ has the circle $\mathcal{C}$ and the uniquely defined circle of $\mathfrak{K}$ that cuts $\mathcal{C}$ orthogonally as eigencircles.

7.2 Inversion about a general circle

We know from theorem (7.3) that a general inversion does not map a given linear family of circles onto itself. We still can define an induced action of this inversion on the family.

Projection onto a linear family

In general an inversion maps a circle of a linear family of circles onto a circle that does not belong to this family. In order to define an induced action on a family of circles we have to project the action of the inversion onto the family in some reasonable way.

One could think of the circle $\mathcal{P}$ from lemma (7.13) as a potential projection of $\mathcal{C}$ onto $\mathfrak{K}$. Despite the uniqueness of the circle $\mathcal{P}$ we can not choose this map as the projection, because it is not the identity map on $\mathfrak{K}$. Hence we have to choose a different way of projecting.

At first we begin with the line $\mathfrak{K}^\circ$. The projection

$$\pi : \mathbb{P}(\mathfrak{Circ}) \to \mathfrak{K}^\circ$$

has to fulfil the identity relation

$$\forall \mathcal{K}^\circ \in \mathfrak{K}^\circ : \pi(\mathcal{K}^\circ) = \mathcal{K}^\circ.$$

For any representant $\mathcal{K}^\circ$ in $\mathfrak{K}^\circ$ there is a unique $(\lambda : \mu) \in \mathbb{P}(\mathbb{C})$ with

$$\mathcal{K}^\circ = \lambda \mathcal{C}^\circ + \mu \mathcal{D}^\circ.$$

We define the projection of $\mathcal{K}^\circ$ onto $\mathfrak{K}^\circ$ as the circle

$$\pi(\mathcal{K}^\circ) = \lambda' \mathcal{C}^\circ + \mu' \mathcal{D}^\circ, \tag{7.1}$$

where $(\lambda' : \mu') \in \mathbb{P}(\mathbb{C})$ solves the equation

$$\mathcal{K}^\circ = \lambda' \mathcal{C}^\circ + \mu' \mathcal{D}^\circ.$$

This definition obviously yields

$$(\lambda', \mu') = (\lambda, \mu)$$

and seems to be much ado about nothing. For a fixed pair of generators of $\mathfrak{K}$ the given projection onto $\mathfrak{K}$ translates into solving the linear equation (7.1) and in its current form it has a solution, if and only if $\mathcal{K}^\circ \in \mathfrak{K}^\circ$. The point here is that an inhomogeneous linear equation – like equation (7.1) – does not need to have a solution, but any linear equation $Mx = b$ with coefficient matrix M and the vector b as right hand side can be transformed into the corresponding normal equation $M^t M x = M^t b$ by multiplication with the transposed coefficient matrix M^t from the left. The normal equation is always solvable. When the original equation is solvable, then the solution of the original equation and the normal equation are identical.

For a circle $\mathcal{K} \notin \mathfrak{K}$ equation (7.1) is not solvable. Hence we define an appropriate normal equation and define the projection onto $\mathfrak{K}$ in terms of the solution of this normal equation. We have to take care that the solution of the normal equation also solves (7.1) in the special case, when the original equation was already solvable.

Definition 7.15 (Projection onto a linear family)
Let $\mathcal{K}$ be a circle and $\mathfrak{K}$ a linear family with generators $\mathcal{C}$ and $\mathcal{D}$. The coefficent matrix $M = (\mathcal{C}^\circ, \mathcal{D}^\circ)$ contains the representants of the generators as column vectors. Then the projection *of $\mathcal{K}$ onto $\mathfrak{K}$ is represented by*

$$\pi(\mathcal{K}^\circ) = \lambda \mathcal{C}^\circ + \mu \mathcal{D}^\circ,$$

where (λ,μ) is the solution of the normal equation

$$M^t M \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = M^t \mathcal{K}^\circ. \tag{7.2}$$

Induced action on a linear family

We already examined the action of an inversion on a linear family, when the circle of inversion belonged to the family. In order to extend our description to general inversions we have to compensate the fact that in general the image of a circle of the family does not belong to the family. We achieve this by projecting this image back onto the family.

Definition 7.16 (Induced action of an inversion)
Let ι be the inversion about a regular circle $\mathcal{I}$ and $\mathfrak{K}$ be a linear family of circles. The inversion ι induces an action on $\mathfrak{K}$ by the map

$$\begin{array}{ccc} \mathfrak{K} & \longrightarrow & \mathfrak{K} \\ \mathcal{K} & \longmapsto & (\pi \circ \iota)(\mathcal{K}), \end{array}$$

where π is the projection onto $\mathfrak{K}$.

Theorem 7.17 (Linearity of the induced action)
The induced action of an inversion on a linear family of circles is a linear map.

Proof
The induced action is the composition of the maps ι and π. Both are linear maps on $\mathbb{P}(\mathfrak{Circ})$, hence their composition is also linear on $\mathfrak{K}$. □

Corollary 7.18 (Induced action in matrix form)
The representant $\mathcal{K}^\circ = \lambda \mathcal{C}^\circ + \mu \mathcal{D}^\circ \in \mathfrak{K}$ has the image $\lambda' \mathcal{C}^\circ + \mu' \mathcal{D}^\circ$ with

$$\begin{pmatrix} \lambda' \\ \mu' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix}.$$

The coefficients of the matrix are defined by the images $\mathcal{C}'$ and $\mathcal{D}'$ of the generating circles $\mathcal{C}$ and $\mathcal{D}$ of the family $\mathfrak{K}$ via

$$(\mathcal{C}')^\circ = a\mathcal{C}^\circ + b\mathcal{D}^\circ \text{ and } (\mathcal{D}')^\circ = c\mathcal{C}^\circ + d\mathcal{D}^\circ.$$

Eigencircles of the induced action

We saw in the previous section that the action induced by the inversion about a circle of a linear family of circles onto this family has two eigencircles: the circle of inversion itself and its orthogonal circle in the family. This situation becomes only slightly more complicated in the general case.

The circle $\mathcal{O} \in \mathfrak{K}$ that is orthogonal to the circle $\mathcal{I}$ of inversion is obviously invariant under the induced action of the inversion. We will see that the induced action always has two different eigencircles.

Theorem 7.19 (Eigencircles of the induced action)
Let $\mathcal{I}$ be a regular circle and $\mathfrak{K}$ a linear family of circles. If $\mathcal{I}$ does not belong to the conjugated family of $\mathfrak{K}$ then the induced action of the inversion ι about $\mathcal{I}$ on the family $\mathfrak{K}$ leaves two different circles invariant. These circles are the orthogonal circle of $\mathcal{I}$ in $\mathfrak{K}$ and the projection $\pi(\mathcal{I})$ of $\mathcal{I}$ onto $\mathfrak{K}$.

Proof
It remains to show that $\pi(\mathcal{I})$ is invariant under the induced action and that $\pi(\mathcal{I})$ is not orthogonal to $\mathcal{I}$.

We can assume the circle of inversion to be the unit circle $\mathcal{E}$. In the basis $\mathcal{E}^\circ, {\mathcal{C}_1}^\circ, {\mathcal{C}_2}^\circ, {\mathcal{C}_3}^\circ$ of orthogonal circles with representants

$$\begin{aligned}
\mathcal{E}^\circ &= (1:0:0:-1)^\circ,\\
{\mathcal{C}_1}^\circ &= (1:0:0:1)^\circ,\\
{\mathcal{C}_2}^\circ &= (0:1:0:0)^\circ,\\
{\mathcal{C}_3}^\circ &= (0:0:1:0)^\circ
\end{aligned}$$

the inversion about $\mathcal{E}$ is represented by the linear map

$$\mathcal{X}^\circ \mapsto \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathcal{X}^\circ.$$

The polar plane $P_{\mathcal{E}^\circ}$ of the unit circle is generated by ${\mathcal{C}_1}^\circ, {\mathcal{C}_2}^\circ$ and ${\mathcal{C}_3}^\circ$. We can assume that (after applying an appropriate transformation) the line $\mathfrak{K}^\circ$ representing the linear family $\mathfrak{K}$ of circles intersects the plane $P_{\mathcal{E}^\circ}$ in ${\mathcal{C}_1}^\circ$.

As the generators of $\mathfrak{K}$ we can choose the circle $\mathcal{C} = \mathcal{C}_1$ with representant ${\mathcal{C}_1}^\circ$ and some other circle $\mathcal{D} \in \mathfrak{K}$ with representant $\mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ$.

We can write $\mathcal{D}^\circ$ in the chosen basis as

$$\mathcal{D}^\circ = \frac{d_0 - d_3}{2}\mathcal{I}^\circ + \frac{d_0 + d_3}{2}{\mathcal{C}_1}^\circ + d_1{\mathcal{C}_2}^\circ + d_2{\mathcal{C}_3}^\circ$$

and hence its image under the inversion ι is

$$\iota(\mathcal{D})^\circ = \frac{-d_0 + d_3}{2}\mathcal{I}^\circ + \frac{d_0 + d_3}{2}{\mathcal{C}_1}^\circ + d_1{\mathcal{C}_2}^\circ + d_2{\mathcal{C}_3}^\circ.$$

A short check shows that $\iota(\mathcal{D})^\circ = (d_3 : d_1 : d_2 : d_0)^\circ$ which is what we expected. This can also be written as

$$\iota(\mathcal{D})^\circ = \mathcal{D}^\circ + (d_3 - d_0)\mathcal{I}^\circ. \tag{7.3}$$

The projection back onto $\mathfrak{K}$ yields

$$(\pi \circ \iota)(\mathcal{D})^\circ = \pi(\mathcal{D}^\circ) + (d_3 - d_0)\pi(\mathcal{I}^\circ) = \mathcal{D}^\circ + (d_3 - d_0)\pi(\mathcal{I}^\circ).$$

With $\pi(\mathcal{I}^\circ) = \lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ$ we can calculate $(\pi \circ \iota)\pi(\mathcal{I}^\circ)$ as

$$(\pi \circ \iota)\pi(\mathcal{I}^\circ) = (\pi \circ \iota)(\lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ)$$

and use equation (7.3) and the linearity of ι and π in order obtain

$$(\pi \circ \iota)\pi(\mathcal{I}^\circ) = (1 + \mu(d_3 - d_0))\pi(\mathcal{I}^\circ).$$

Thus the circle represented by $\pi(\mathcal{I}^\circ)$ is invariant under the action induced by ι onto $\mathfrak{K}$.

In order to prove that $\pi(\mathcal{I})$ is not orthogonal to $\mathcal{I}$ it suffices to show that it is different from $\mathcal{C}$, since this is the only circle in $\mathfrak{K}$ with this property. This uniqueness also implies that $\mathcal{D} \not\perp \mathcal{I}$, a fact that can be equivalently expressed as $d_0 \neq d_3$.

We have to show that $\mu \neq 0$ in the linear combination

$$\pi(\mathcal{I}^\circ) = \lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ.$$

We obtain (λ, μ) by solving the normal equation (7.2), in this case

$$\begin{pmatrix} 2 & (d_0 + d_3) \\ (d_0 + d_3) & \sum_{i=0}^{3} {d_i}^2 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} O \\ d_0 - d_3 \end{pmatrix}.$$

The determinant Δ of the left hand side is

$$\Delta = 2{d_1}^2 + 2{d_2}^2 + (d_0 - d_3)^2.$$

Because of $d_0 \neq d_3$ we see immediately that $\Delta > 0$. Hence the normal equation has a unique solution and $\pi(\mathcal{I})$ is well defined.

Solving the equation for μ now yields

$$\mu = \frac{2(d_0 - d_3)}{\Delta}$$

and from the same argument as above follows $\mu \neq 0$. □

Open questions concerning the induced action

The induced action is ill defined for a couple of reasons.

- In general the circles $\mathcal{C}$ and its projection $\pi(\mathcal{C})$ onto $\mathfrak{K}$ are mapped onto different circles of $\mathfrak{K}$. This means that the composed maps $\pi \circ \iota \circ \pi$ and $\pi \circ \iota$ act identically on $\mathfrak{K}$, but differently on $\mathbb{P}(\mathfrak{Circ})$. In other words: the set of circles that are projected onto the same circle $\pi(\mathcal{C})$ as some given circle $\mathcal{C}$ is mapped by ι onto a set of circles with different projections onto $\mathfrak{K}$. It would be more pleasing, if the set of images under inversion would again share the same image under the projection $(\pi \circ \iota)(\pi(\mathcal{C}))$ onto the linear family.

- The induced action is not an involution. We would expect from an induced action that it shares this important property with the original action. For a circle $\mathcal{K}$ of the family $\mathfrak{K}$ the induced action $(\pi \circ \iota)(\mathcal{K})$ is identical to the map $(\pi \circ \iota)(\pi(\mathcal{K}))$, because $\pi(\mathcal{K}) = \mathcal{K}$. Unfortunately the circles $\iota(\mathcal{K})$ and $\mathcal{K}$ are mapped onto different circles of $\mathfrak{K}$ by $\pi \circ \iota$ and $\pi \circ \iota \circ \pi$.

- Some circles can not be projected onto the family $\mathfrak{K}$. The projection onto a linear family is not defined in terms of $\mathbb{P}(\mathfrak{Circ})$, but of the underlying $\mathbb{C}^4$. Thus the projection as we defined it maps some circles onto the "null-representant" $(0:0:0:0)^\circ$. We can construct examples of this. Let $\mathcal{C}$ be a circle not on $\mathfrak{K}$. Then $\mathcal{C}$ and $\pi(\mathcal{C})$ generate a linear family of circles. This family contains a circle with representant $\mathcal{C}^\circ - \pi(\mathcal{C})^\circ$ which is projected onto $\pi(\mathcal{C})^\circ - \pi^2(\mathcal{C})^\circ = (0:0:0:0)^\circ$.

- The induced action of a general inversion ι can not be represented by an inversion about a circle of the linear family. It would be very comfortable, if we could describe the induced action of an inversion by an inversion about a circle of the family. Moreover the circle of this inversion should be the projection of the original circle $\mathcal{I}$ of inversion onto the family.

It is still an open question how to overcome these difficulties. Most likely these problems can be solved by an appropriate choice of the projection onto a linear family. The way of projection that was chosen in this text does not respect the special geometry of the space $\mathbb{P}(\mathfrak{Circ})$. Instead it uses the common normal equation which is meaningful in Euclidean geometry of $\mathbb{R}^n$ (or $\mathbb{C}^n$). Recalling the discussion about the similarities between the polarity function and the scalar product at the end of section (3.3) the projection of circles onto a linear family should be an appropriate adaption of the orthogonal projection in Euclidean space.

7.3 Normalized inversion

Definition and basic properties

Because inversions are linear in $\mathbb{P}(\mathfrak{Circ})$, for any representant $\lambda_0\mathcal{C}_0{}^\circ + \lambda_1\mathcal{C}_1{}^\circ \in \mathfrak{K}$ its image under inversion is

$$\iota(\lambda_0\mathcal{C}_0{}^\circ + \lambda_1\mathcal{C}_1{}^\circ) = \lambda_0\iota(\mathcal{C}_0)^\circ + \lambda_1\iota(\mathcal{C}_1)^\circ.$$

We have already shown that an inversion given by the linear map of theorem (5.11) is an involution in $\mathbb{P}(\mathfrak{Circ})$. We will modify this formula slightly in order to make this map an involution on the representants in the underlying space $\mathbb{C}^4$.

Definition 7.20 (Normalized projective inversion)
Let $\mathcal{I}$ be a regular circle with representant $\mathcal{I}^\circ$ and M the matrix from theorem (5.11). Then the normalized inversion *about $\mathcal{I}$ is the map*

$$\xi \mapsto \frac{2M}{P(\mathcal{I}^\circ, \mathcal{I}^\circ)}\xi.$$

Lemma 7.21
Every normalized inversion is a involution not only on $\mathbb{P}(\mathfrak{Circ})$, but also on the underlying space $\mathbb{C}^4$.

Proof
We have seen before that for a general inversion with representing matrix M the square M^2 is $\frac{1}{4}P(\mathcal{I}^\circ,\mathcal{I}^\circ)^2$ times the identity matrix. Therefore it is obvious that $\left(\frac{2M}{P(\mathcal{I}^\circ,\mathcal{I}^\circ)}\right)^2$ is the identity matrix. □

The normalized inversion has some useful properties. Because we always have to do computations with some representants of circles, it is very comfortable to have a description that is an involution not only on the represented circles but also on the representants themselves.

Theorem 7.22 (The normalized inversion is well defined)
Let $\mathcal{C}_1{}^\circ$ and $\mathcal{C}_2{}^\circ$ be representants of the regular circle $\mathcal{C}$. Then the corresponding describing matrices of the normalized inversion about $\mathcal{C}$ are equal.

Proof
Because $\mathcal{C}_1{}^\circ, \mathcal{C}_2{}^\circ \in \mathbb{P}(\mathfrak{Circ})$ represent the same circle $\mathcal{C}$, there exists a projective number $(\lambda_1 : \lambda_2) \in \mathbb{P}(\mathbb{C})$ with

$$\lambda_1\mathcal{C}_1{}^\circ = \lambda_2\mathcal{C}_2{}^\circ.$$

By M_1 and M_2 we denote the describing matrices of the non-normalized inversions. From (5.11) we see immediately that these matrices suffice

$$\lambda_1{}^2 M_1 = \lambda_2{}^2 M_2.$$

The polarity function shows the same behavior:

$$\lambda_1{}^2 P(\mathcal{C}_1{}^\circ,\mathcal{C}_1{}^\circ) = \lambda_2{}^2 P(\mathcal{C}_2{}^\circ,\mathcal{C}_2{}^\circ).$$

Thus the factors λ_i cancel out in the quotients

$$\frac{2M_1}{P(\mathcal{C}_1{}^\circ,\mathcal{C}_1{}^\circ)} = \frac{2\lambda_1{}^2 M_1}{\lambda_1{}^2 P(\mathcal{C}_1{}^\circ,\mathcal{C}_1{}^\circ)} = \frac{2\lambda_2{}^2 M_2}{\lambda_2{}^2 P(\mathcal{C}_2{}^\circ,\mathcal{C}_2{}^\circ)} = \frac{2M_2}{P(\mathcal{C}_2{}^\circ,\mathcal{C}_2{}^\circ)}$$

and the representing matrix of the normalized inversion does not depend on the chosen representant of the circle $\mathcal{C}$ of inversion. □

The relation between inversion and normalized inversion is very much that same as between Möbius transformations with coefficient matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $GL(2,\mathbb{C})$ and those with coefficient matrix in $SL(2,\mathbb{C})$. The second kind may be in analogy called normalized Möbius transformations.

Inversions between two given circles

Theorem 7.23 (Polarity and normalized inversions)
Let $\mathcal{C}, \mathcal{D}$ be arbitrary circles and $\mathcal{I}$ be a regular circle with ι being the normalized inversion about $\mathcal{C}$. Then for their representants holds:
(i) $P(\iota(\mathcal{C}^\circ),\mathcal{I}^\circ) = -P(\mathcal{C}^\circ,\mathcal{I}^\circ)$,
(ii) $P(\iota(\mathcal{C}^\circ),\iota(\mathcal{D}^\circ)) = P(\mathcal{C}^\circ,\mathcal{D}^\circ)$.

Proof
Assume the circle of inversion to be the unit circle $\mathcal{E}$. The usual inversion ε the unit circle is already normalized. For $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ$ we have $\varepsilon(\mathcal{C}^\circ) = (c_3 : c_1 : c_2 : c_0)^\circ$ and therefore

$$P(\varepsilon(\mathcal{C}^\circ), \mathcal{E}^\circ) = (c_0 - c_3) = -(c_3 - c_0) = -P(\mathcal{C}^\circ, \mathcal{E}^\circ).$$

For $\mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ$ we have $\varepsilon(\mathcal{D}^\circ) = (d_3 : d_1 : d_2 : d_0)^\circ$ and the second equation holds due to

$$P(\varepsilon(\mathcal{C}^\circ), \varepsilon(\mathcal{D}^\circ)) = 2c_1d_1 + 2c_2d_2 - c_0d_3 - c_3d_0 = P(\mathcal{C}^\circ, \mathcal{D}^\circ). \qquad \square$$

The normalized inversion can be used for interesting geometric operations.

Theorem 7.24 (Inversions between two circles (special case))
Let $\mathcal{C}$ and $\mathcal{D}$ be two different circles with representants $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ fulfilling

$$P(\mathcal{C}^\circ, \mathcal{C}^\circ) = P(\mathcal{D}^\circ, \mathcal{D}^\circ) \neq 0.$$

Let $\mathcal{I}^\circ = \mathcal{C}^\circ + \mathcal{D}^\circ$ and $\mathcal{J}^\circ = \mathcal{C}^\circ - \mathcal{D}^\circ$ be the representants of the circles $\mathcal{I}$ and $\mathcal{J}$.

If $\mathcal{C}$ and $\mathcal{D}$ are not touching each other, then $\mathcal{I}$ and $\mathcal{J}$ are regular circles and for the normalized inversions $\iota_\mathcal{I}$ and $\iota_\mathcal{J}$ holds

$$\iota_\mathcal{I}(\mathcal{C}^\circ) = -\mathcal{D}^\circ,\ \iota_\mathcal{I}(\mathcal{D}^\circ) = -\mathcal{C}^\circ,\ \iota_\mathcal{J}(\mathcal{C}^\circ) = \mathcal{D}^\circ,\ \iota_\mathcal{J}(\mathcal{D}^\circ) = \mathcal{C}^\circ.$$

If $\mathcal{C}$ and $\mathcal{D}$ are touching each other then one of the circles $\mathcal{I}$ and $\mathcal{J}$ is a nullcircle. The inversion about the other circle then still has the above property.

Proof
The situation can be transformed to the case where $\mathcal{C}$ is the unit circle and $\mathcal{D}$ a straight line parallel to the x-axis. Then we have

$$\mathcal{C}^\circ = (1 : 0 : 0 : -1)^\circ \ \text{ and } \ \mathcal{D}^\circ = (0 : 0 : 1 : c_3)^\circ$$

and $P(\mathcal{C}^\circ, \mathcal{C}^\circ) = P(\mathcal{D}^\circ, \mathcal{D}^\circ) = 2$. We obtain $\mathcal{I}^\circ = (1 : 0 : 1 : -1 + c_3)^\circ$ which is a nullcircle, when $c_3 = 2$, and $\mathcal{J}^\circ = (1 : 0 : -1 : -1 - c_3)^\circ$ which is a nullcircle, when $c_3 = -2$. Both cases occur, if and only if $\mathcal{C}$ and $\mathcal{D}$ are touching each other.

From theorem (5.11) and definition (7.20) follows that $\iota_\mathcal{I}$ is the map

$$\mathcal{X}^\circ \mapsto \frac{1}{2 - c_3} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 - c_3 & 0 & 0 \\ -1 + c_3 & 0 & -c_3 & 1 \\ (-1 + c_3)^2 & 0 & -2(-1 + c_3) & 1 \end{pmatrix} \mathcal{X}^\circ$$

and a short calculation now yields $\iota_\mathcal{I}(\mathcal{C}^\circ) = -\mathcal{D}^\circ$ and $\iota_\mathcal{I}(\mathcal{D}^\circ) = -\mathcal{C}^\circ$.

In the same way we can show the statements for $\iota_\mathcal{J}$ which is described by

$$\mathcal{X}^\circ = \frac{1}{2 + c_3} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 + c_3 & 0 & 0 \\ 1 + c_3 & 0 & c_3 & 1 \\ (1 + c_3)^2 & 0 & -2(1 + c_3) & 1 \end{pmatrix} \mathcal{X}^\circ.$$

Obviously in each of the exceptional cases $c_3 = 2$ and $c_3 = -2$ only one of the inversions is possible. $\square$

Corollary 7.25
The inversions about $\mathcal{I}$ and $\mathcal{J}$ map $\mathcal{C}$ onto $\mathcal{D}$ and vice versa.

Even in the case

$$0 \neq P(\mathcal{C}^\circ, \mathcal{C}^\circ) \neq P(\mathcal{D}^\circ, \mathcal{D}^\circ) \neq 0$$

we can multiply each representant by an appropriate non-zero factor to make the values of the polarity functions equal. These factors λ and μ may be chosen in two different ways.

Theorem 7.26 (Inversions between two circles (general case))
Let $\mathcal{C}$ and $\mathcal{D}$ be two regular circles that are not touching each other. Then there exist exactly two inversions that map these circles onto each other.

Proof
For regular circles $\mathcal{C}, \mathcal{D}$ the values $P(\mathcal{C}^\circ, \mathcal{C}^\circ)$ and $P(\mathcal{D}^\circ, \mathcal{D}^\circ)$ of the polarity function are different from zero. Thus there exist exactly two solutions $(\lambda : \mu) \in \mathbb{P}(\mathbb{C})$ of the equation

$$P(\lambda\mathcal{C}^\circ, \lambda\mathcal{C}^\circ) = P(\mu\mathcal{D}^\circ, \mu\mathcal{D}^\circ),$$

which can be written as

$$\lambda^2 P(\mathcal{C}^\circ, \mathcal{C}^\circ) = \mu^2 P(\mathcal{D}^\circ, \mathcal{D}^\circ).$$

These solutions are connected by the relation that $(\lambda : \mu)$ solves the system, if and only if $(\lambda : -\mu)$ solves the system. They are always different from each other, because

$$\mathcal{C} \text{ regular} \Longrightarrow P(\mathcal{C}^\circ, \mathcal{C}^\circ) \neq 0 \Longrightarrow \mu \neq 0$$

and

$$\mathcal{D} \text{ regular} \Longrightarrow P(\mathcal{D}^\circ, \mathcal{D}^\circ) \neq 0 \Longrightarrow \lambda \neq 0.$$

From theorem (7.24) follows that there exist two inversions between the representants $\lambda\mathcal{C}^\circ$ and $\mu\mathcal{D}^\circ$ and therefore also between the circles $\mathcal{C}$ and $\mathcal{D}$. □

From the above discussion follows that for any two regular circles $\mathcal{C}$ and $\mathcal{D}$ that are not touching each other there is a pair of uniquely defined inversions that map $\mathcal{C}$ and $\mathcal{D}$ onto each other. We may calculate the circles of inversion by choosing an arbitrary pair $\mathcal{C}^\circ$ and $\mathcal{D}^\circ$ of representants and calculating the representants of the circles $\mathcal{I}$ and $\mathcal{J}$ of inversion as

$$\mathcal{I}^\circ = \lambda\mathcal{C}^\circ + \mu\mathcal{D}^\circ$$

and

$$\mathcal{J}^\circ = \lambda\mathcal{C}^\circ - \mu\mathcal{D}^\circ,$$

where $(\lambda : \mu) \in \mathbb{P}(\mathbb{C})$ solves

$$\lambda^2 P(\mathcal{C}^\circ, \mathcal{C}^\circ) = \mu^2 P(\mathcal{D}^\circ, \mathcal{D}^\circ).$$

Corollary 7.27
In the special case that both of the circles $\mathcal{C}$ and $\mathcal{D}$ are real, we may write the solution explicitly as square roots, i.e.

$$\mathcal{I}^\circ = \sqrt{P(\mathcal{D}^\circ, \mathcal{D}^\circ)}\mathcal{C}^\circ + \sqrt{P(\mathcal{C}^\circ, \mathcal{C}^\circ)}\mathcal{D}^\circ$$

and

$$\mathcal{J}^\circ = \sqrt{P(\mathcal{D}^\circ, \mathcal{D}^\circ)}\mathcal{C}^\circ - \sqrt{P(\mathcal{C}^\circ, \mathcal{C}^\circ)}\mathcal{D}^\circ.$$

Remark
The circles $\mathcal{I}$ and $\mathcal{J}$ obviously meet $\mathcal{C}$ and $\mathcal{D}$ in their points of intersection. It is clear that these circles of inversion must divide the intersection angles between $\mathcal{C}$ and $\mathcal{D}$ in half. In the special case that the given circles are in fact straight lines the circles of inversion are also lines and therefore the angle bisectors of the intersection angle. A disbelieving reader may carry out the necessary calculations in order to assure himself that in the special case where $\mathcal{C}$ and $\mathcal{D}$ are lines, the procedure mentioned in the last corollary indeed simplifies to the known method for calculating angle bisectors.

8 Bicircular Quartics

8.1 Inverse and pedal curve of a conic section

In the preceding chapters we have investigated, how inversions act on lines and circles. We did this for single curves first, then for entire families of circles. We have seen that the inversion about a circle maps straight lines and circles onto straight lines and circles. For a comprehensive description of inversions it was necessary to unify these two kinds of curves. We achieved this by the introduction of $\mathbb{P}(\mathfrak{Circ})$, where lines are nothing else than a special type of circles.

We used the calculational tools available in this projective space to prove some fundamental theorems about (generalized) circles. Some of these propositions will be useful for us in the following chapters where we want to examine the properties of bicircular quartics. These curves arise, for example, as the image of a conic section under inversion about a circle. The definition of the term *bicircular quartic* will be given in the next section of the text.

Inverse of a conic section

The image of a generalized circle under inversion is a generalized circle. It is obvious that a singular conic section, i.e. a union of two lines, will be mapped onto the union of two circles. This includes the special case, when the image is the union of a circle and a line or the union of two lines. This observation is only a special case of the following obvious

Lemma 8.1 (Image of a reducible curve)
The image of a reducible curve is reducible.

We want to know what happens to a general conic section under inversion. Earlier in the text we have shown theorem (4.14) which says that an inversion doubles the degree of an algebraic curve. This means that the image of a general conic under inversion is a curve of forth degree. The image of of a conic section under reflection about a straight line is a conic section. Interpreted as an inversion about a generalized circle the image is a reducible quartic curve and contains the line at infinity two times as an exceptional component.

Definition 8.2 (Equation of a conic section)
Let C be a general conic section. Its homogeneous equation

$$C : a_{00}{x_0}^2 + a_{11}{x_1}^2 + a_{22}{x_2}^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + 2a_{12}x_1x_2 = 0 \tag{8.1}$$

contains six parameters a_{ij}.

We can write any inversion about a proper circle as a concatenation of translations, dilations and the inversion ε about the unit circle. Since translations and dilations will change neither the order of a curve nor the fact that it eventually has the circular points at infinity as singularities, it suffices to look at the image of a conic section under ε. We obtain

$$\varepsilon(C) : a_{00}({x_1}^2 + {x_2}^2)^2 + 2(a_{01}x_0x_1 + a_{02}x_0x_2)({x_1}^2 + {x_2}^2) + {x_0}^2(a_{11}{x_1}^2 + 2a_{12}x_1x_2 + a_{22}{x_2}^2) = 0.$$

For $a_{00} \neq 0$ this is the equation of a bicircular quartic. For $a_{00} = 0$ we obtain a cubic running through the circular points at infinity, a so called circular cubic. Seen as a quartic curve the image $\varepsilon(C)$ in this case is the union of a circular cubic and the line at infinity. From this point of view it is also singular in the circular points at infinity. The figures (8.1) and (8.2) show these two cases.

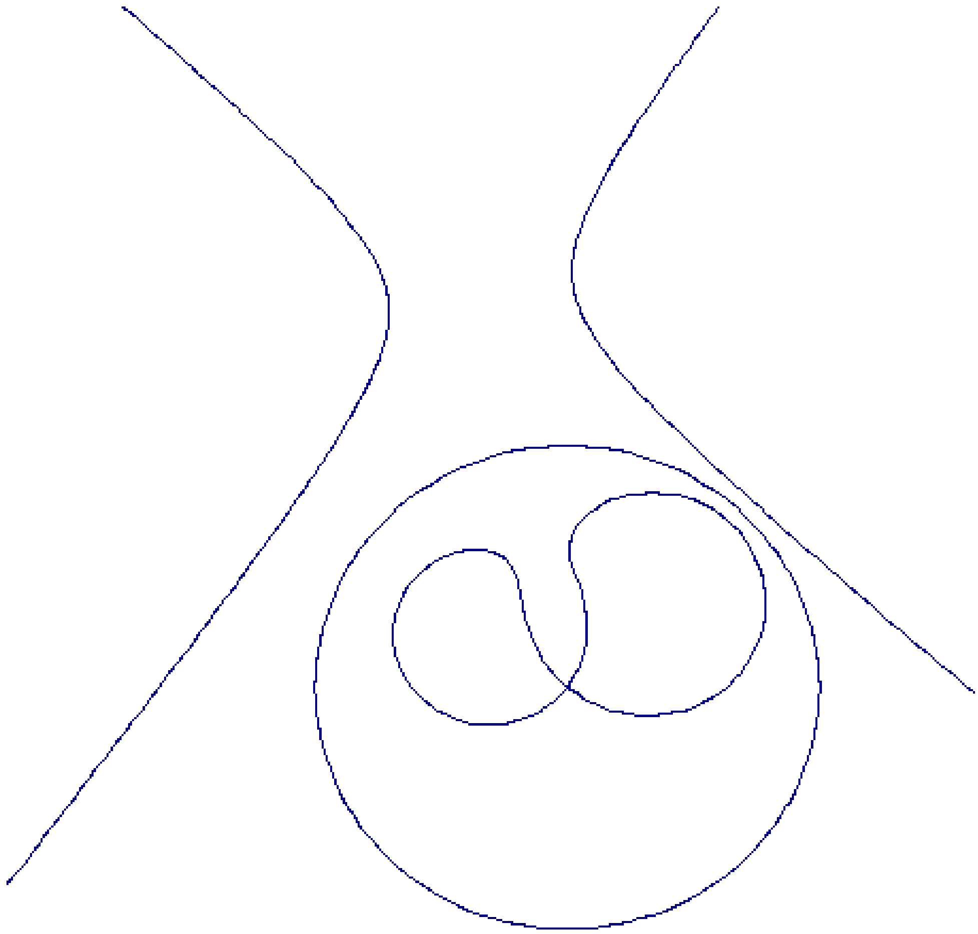


Figure 8.1: Inverse curve of a hyperbola – the center of the circle of inversion does not lie on the curve

In general the image of a conic section under inversion is a bicircular quartic. Inverting the image curve reproduces the original curve of degree two. One might ask: Which kind

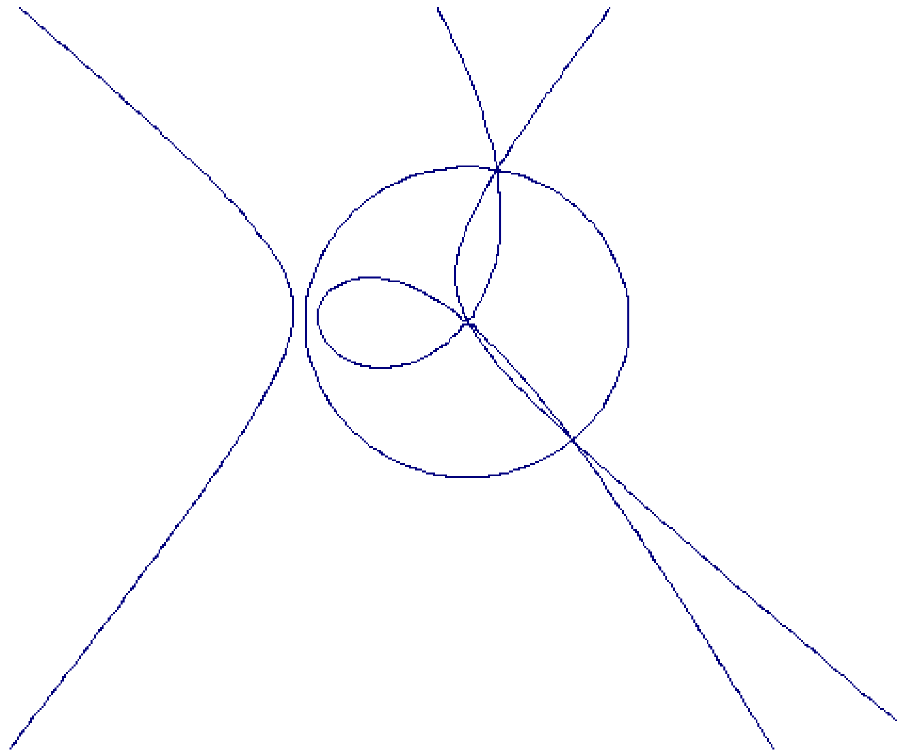


Figure 8.2: Inverse curve of a hyperbola – the center of the circle of inversion lies on the curve

of curve is the image of a bicircular quartic under inversion? We will discuss later that the image under an arbitrary inversion is again a bicircular quartic.

Pedal curve of a conic section

The pedal curve of an ellipse with respect to the origin of the coordinate system can be calculated in an elementary way. We assume its axes to be parallel to the coordinate axes of the and its center to be at (u, v). With the length of the semi-axes named a and b the ellipse is given by the equation

$$\frac{(x-u)^2}{a^2} + \frac{(y-v)^2}{b^2} = 1$$

and the tangent in the point (x_0, y_0) as

$$\frac{(x-u)(x_0-u)}{a^2} + \frac{(y-v)(y_0-v)}{b^2} = 1.$$

We are looking for the points where a line through the origin intersects with the tangent orthogonally. This leads to the system of equations

$$\begin{aligned}
&b^2(x_0-u)^2 + a^2(y_0-v)^2 - a^2b^2 = 0,\\
&b^2(x_0-u)(x-u) + a^2(y_0-v)(y-v) - a^2b^2 = 0,\\
&a^2(y_0-v)x - b^2(x_0-u)y = 0.
\end{aligned}$$

Elimination of x_0 and y_0 from these equations leads to the equation

$$(x^2+y^2)^2-(2ux+2vy)(x^2+y^2)+(u^2-a^2)x^2+2uvxy+(v^2-b^2)y^2=0$$

of the pedal curve. We see that the pedal curve of an ellipse is a bicircular quartic.

8.2 Bicircular Quartics

Definition

In the last section we already mentioned the term bicircular quartic. This section will catch up for clarification. The attribute *bicircular* describes that the curve in question is closely related to the union of two circles. Actually it is a generalization of this, but has many properties in common with such a union. One of this properties is the fact that a union of circles has singularities at the circular points $(0:1:\pm i)$ at infinity. We take this as

Definition 8.3 (Bicircular quartic)
A bicircular quartic *is a quartic curve with singularities in the circular points at infinity.*

This definition enables us to determine the possible shape of the equation of a bicircular quartic. We take the equation

$$\begin{aligned} q(x_0,x_1,x_2) \;=\; & a_{400}{x_0}^4 + \\ & + 4a_{310}{x_0}^3x_1 + 4a_{301}{x_0}^3x_2 + \\ & + 6a_{220}{x_0}^2{x_1}^2 + 12a_{211}{x_0}^2x_1x_2 + 6a_{202}{x_0}^2{x_2}^2 + \\ & + 4a_{130}x_0{x_1}^3 + 12a_{121}x_0{x_1}^2x_2 + 12a_{112}x_0x_1{x_2}^2 + 4a_{103}x_0{x_2}^3 + \\ & + a_{040}{x_1}^4 + 4a_{031}{x_1}^3x_2 + 6a_{022}{x_1}^2{x_2}^2 + 4a_{013}x_1{x_2}^3 + a_{004}{x_2}^4 = 0 \end{aligned} \tag{8.2}$$

of a general quartic curve and check the conditions

$$\begin{aligned} \left.\frac{\partial q}{\partial x_0}\right|_{(0,1,\pm i)} &= 0, \\ \left.\frac{\partial q}{\partial x_1}\right|_{(0,1,\pm i)} &= 0, \\ \left.\frac{\partial q}{\partial x_2}\right|_{(0,1,\pm i)} &= 0. \end{aligned}$$

The first condition immediately yields

$$\begin{aligned} a_{130} &= 3a_{112}, \\ a_{103} &= 3a_{121}, \end{aligned}$$

the second condition implies

$$\begin{aligned} a_{040} &= 3a_{022}, \\ a_{013} &= 3a_{031} \end{aligned}$$

and the third

$$\begin{aligned} a_{031} &= 3a_{013}, \\ a_{004} &= 3a_{022}. \end{aligned}$$

We obtain the following restrictions for the coefficients a_{ijk} of the terms of degree three and four in x_1 and x_2.

$$\begin{gathered} a_{130} = 3a_{112}, \\ a_{103} = 3a_{121} \\ \text{and} \\ a_{040} = 3a_{022} = a_{004}, \\ a_{031} = 0, \quad a_{013} = 0. \end{gathered} \tag{8.3}$$

After renaming the coefficients we can write the general equation of a bicircular quartic as

$$\begin{aligned} a({x_1}^2 + {x_2}^2)^2 + 2x_0(b_1x_1 + b_2x_2)({x_1}^2 + {x_2}^2) + & \\ {x_0}^2\Big(c_{20}{x_1}^2 + 2c_{11}x_1x_2 + c_{02}{x_2}^2 + 2c_{10}x_0x_1 + 2c_{01}x_0x_2 + c_{00}{x_0}^2\Big) &= 0. \end{aligned} \tag{8.4}$$

Remark
The reader might be surprised by the factor 2 instead of 4 in front of the second expression. We will see later that this greatly simplifies calculations.

The 9-tuple

$$q = (a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00})$$

of coefficients of this equation can be multiplied with any number $\lambda \in \mathbb{C} \setminus \{0\}$ without changing the curve.

Definition 8.4 (Space of bicircular curves)
The projective space $\mathbb{P}_8(\mathbb{C})$ together with the bijection

$$\begin{aligned} \mathbb{P}_8(\mathbb{C}) &\rightarrow \text{set of bicircular quartics} \\ (a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00}) &\mapsto \text{equation (8.4)} \end{aligned}$$

is called the space of bicircular quartics*.*

For the sake of simplicity we want to call the equation of a bicircular curve a *bicircular equation*. In analogy to the projective space of circles we can define the projective space of bicircular quartics as the nine-dimensional affine space of coefficients together with the usual equivalence relation

$$q \sim q' \Longleftrightarrow \mathbb{C}q = \mathbb{C}q'.$$

Hence the space of bicircular quartics forms an eight-dimensional projective space over $\mathbb{C}$ and any projective vector

$$(a : b_1 : b_2 : c_{20} : c_{11} : c_{02} : c_{10} : c_{01} : c_{00}) \in \mathbb{P}_8(\mathbb{C})$$

defines a bicircular quartic given by equation (8.4). In certain cases the curve may have finite singularities or be reducible. A circle, for instance, can be interpreted as a bicircular quartic with the parameters

$$a = b_1 = b_2 = c_{11} = 0 \text{ and } c_{20} = c_{02}.$$

The resulting equation

$${x_0}^2\Big(c_{20}({x_1}^2 + {x_2}^2) + 2c_{10}x_0x_1 + 2c_{01}x_0x_2 + c_{00}{x_0}^2\Big) = 0$$

is the union of the circle represented by $(c_{20} : -c_{10} : -c_{01} : c_{00})^\circ \in \mathbb{P}(\mathfrak{Circ})$ and two times the line at infinity.

Multicircular curves

Equation (8.4) reflects the property of the curve that it has the the circular points as double points. We can define more generally a n-circular curve of degree $2n$ through

$$\begin{aligned} & p_0({x_1}^2+{x_2}^2)^n + x_0 p_1({x_1}^2+{x_2}^2)^{n-1} + \dots \\ & \dots + {x_0}^{n-2} p_{n-2}({x_1}^2+{x_2}^2)^2 + {x_0}^{n-1} p_{n-1}({x_1}^2+{x_2}^2) + {x_0}^n p_n \quad = \quad 0, \end{aligned} \tag{8.5}$$

where p_i is a homogeneous polynomial of degree i in x_1 and x_2 for $i = 0, \dots, (n-1)$ and p_n is a homogeneous polynomial of degree n in x_0, x_1 and x_2. It has the circular points at infinity as n-fold points. Equation (8.5) contains $(n+1)^2$ coefficients, which are connected by the usual equivalence relation. Hence

Corollary 8.5
The space of n-circular curves of degree $2n$ is $n(n+2)$-dimensional over $\mathbb{C}$.

Within this more general framework the case $n = 1$ treats circles as unicircular curves of degree 2. The projective space of circles $\mathbb{P}(\mathfrak{Circ})$ is a three-dimensional projective complex vector space. For $n = 2$ we obtain the eight-dimensional space of bicircular quartics.

Theorem 8.6 (n-circular curve as an image under inversion)
The image of a general curve of degree n under inversion is a n-circular curve of degree $2n$.

Proof
Be

$$\sum_{i+j+k=n} c_{ijk} {x_0}^i {x_1}^j {x_2}^k = 0$$

the equation of a curve of degree n. Then the equation of the image curve under the inversion ε about the unit circle $\mathcal{E}$ is

$$\sum_{i+j+k=n} c_{ijk} ({x_1}^2+{x_2}^2)^i (x_0 x_1)^j (x_0 x_2)^k = 0.$$

This equation can be easily sorted with respect to the index i in order to express the coefficients of the polynomials p_m in equation (8.5) by the coefficients c_{ijk}. □

C-Q-form of a bicircular quartic

We may rewrite equation (8.4) in different ways. First of all we may use completion of the square in order to combine the terms containing the factor $({x_1}^2+{x_2}^2)$. This is achieved in two steps. We assume $a \neq 0$, such that we can multiply the whole equation by $\frac{1}{a}$. After rescaling we have

$$({x_1}^2+{x_2}^2)^2 + 2x_0(b_1x_1+b_2x_2)({x_1}^2+{x_2}^2) + \dots = 0.$$

We rearrange the terms as

$$\Big({x_1}^2+{x_2}^2+x_0(b_1x_1+b_2x_2)\Big)^2 - {x_0}^2(b_1x_1+b_2x_2)^2 + \dots = 0.$$

We rearrange the first square once more and obtain

$$\Big((x_1+\frac{b_1}{2}x_0)^2+(x_2+\frac{b_2}{2}x_0)^2-\frac{1}{4}({b_1}^2{x_0}^2+{b_2}^2{x_0}^2)\Big)^2.$$

After sorting all terms into their place the equation becomes

$$\begin{array}{l}\Big((x_1+\frac{b_1}{2}x_0)^2+(x_2+\frac{b_2}{2}x_0)^2\Big)^2+\\ +{x_0}^2\Big((c_{20}-\frac{3}{2}{b_1}^2-\frac{1}{2}{b_2}^2){x_1}^2+2(c_{11}-b_1b_2)x_1x_2+\\ (c_{02}-\frac{1}{2}{b_1}^2-\frac{3}{2}{b_2}^2){x_2}^2+2(c_{10}-\frac{1}{4}b_1({b_1}^2+{b_2}^2))x_0x_1+\\ 2(c_{01}-\frac{1}{4}b_2({b_1}^2+{b_2}^2))x_0x_2+(c_{00}-\frac{1}{8}({b_1}^2+{b_2}^2)^2){x_0}^2\Big) \quad = \quad 0.\end{array}$$

Hence we can write bicircular quartics with $a \neq 0$ as

$$\mathcal{C}^2+{x_0}^2Q=0$$

with $\mathcal{C}$ being the equation of the nullcircle with center at $(-\frac{b_1}{2a},-\frac{b_2}{2a})$ and Q being a quadric. This new description contains eight parameters, two from the nullcircle $\mathcal{C}$ and six from the quadric Q. By a translation we can move the center of $\mathcal{C}$ into the origin. By a rotation around the origin we can make the principal axes of Q parallel to the coordinate axes. Thus we can simplify the equation of any bicircular quartic by a simple transformation of coordinates to

$$({x_1}^2+{x_2}^2)^2+{x_0}^2\Big(\hat{c}_{20}{x_1}^2+\hat{c}_{02}{x_2}^2+2\hat{c}_{10}x_0x_1+2\hat{c}_{01}x_0x_2+\hat{c}_{00}{x_0}^2\Big)=0.$$

We want to abbreviate this equation by

$${\mathcal{C}_0}^2+{x_0}^2Q_0=0. \tag{8.6}$$

Obviously we may vary the radius of $\mathcal{C}$, because such a change can be absorbed by changing the quadric accordingly. Thus we may introduce a parameter t representing the radius of the circle, obtaining

$$\begin{array}{l}({x_1}^2+{x_2}^2+t{x_0}^2)^2+\\ +{x_0}^2\Big((\hat{c}_{20}-2t){x_1}^2+(\hat{c}_{02}-2t){x_2}^2+2\hat{c}_{10}x_0x_1+2\hat{c}_{01}x_0x_2+(\hat{c}_{00}-t^2){x_0}^2\Big) \quad = \quad 0.\end{array} \tag{8.7}$$

This is the general form of equation (8.6), where we denote

$$\mathcal{C}_t=({x_1}^2+{x_2}^2+t{x_0}^2)^2$$

and

$$Q_t=\Big((\hat{c}_{20}-2t){x_1}^2+(\hat{c}_{02}-2t){x_2}^2+2\hat{c}_{10}x_0x_1+2\hat{c}_{01}x_0x_2+(\hat{c}_{00}-t^2){x_0}^2\Big)$$

in order to write the equation of a bicircular quartic as

$${\mathcal{C}_t}^2+{x_0}^2Q_t=0. \tag{8.8}$$

Later we will use this equation for the determination of double tangents by variation of the parameter t.

Definition 8.7 (C-Q-form)
Equation (8.8) is called the C-Q-form *of a bicircular quartic.*

We have derived the C-Q-form under the assumption that $a \neq 0$. Circular cubics can not be written in this form.

8.3 Envelope of a rational family of circles

Parametrization of a rational family of circles

A *rational family* of circles is a family of circles, whose representants form a rational curve in $\mathbb{P}(\mathfrak{Circ})$. In this text we want to reserve this expression for a special case of rational families.

Definition 8.8 (Rational family of circles)
A rational family *of circles is a family of circles whose representants form a conic section in* $\mathbb{P}(\mathfrak{Circ})$.

Remark
In [44] Pedoe calls this a conic system of circles.

A conic section in $\mathbb{P}(\mathfrak{Circ})$ is defined by a total of eight parameters. Three parameters arise from the plane in which it lies, five points in this plane define the conic section as a plane quadric curve. For our studies we choose the following rational parametrization of the curve. We recall

Theorem 8.9 (Parametrization of a quadric in $\mathbb{P}_2(\mathbb{C})$)
Let P_1, P_2 and P_0 be points in $\mathbb{P}_2(\mathbb{C})$ not lying on a straight line and $k \neq 0$ a complex number. Then the set

$$Q = \left\{ {\lambda_0}^2 P_1 + 2k\lambda_0\lambda_1 P_0 + {\lambda_1}^2 P_2, (\lambda_0 : \lambda_1) \in \mathbb{P}(\mathbb{C}) \right\}$$

forms a non-singular quadric in $\mathbb{P}_2(\mathbb{C})$.

Proof
Because the points are not collinear, we may choose the coordinate system in such a way that the points P_1, P_2, P_0 have the coordinates

$$P_1 = (0:1:0),\ P_2 = (0:0:1),\ P_0 = (1:0:0).$$

Thus the variable point on Q lies on the conic section

$$k^2 x_1 x_2 - {x_0}^2 = 0.$$

The gradient $(-2x_0, k^2x_2, k^2x_1)$ does not vanish on $\mathbb{P}_2(\mathbb{C})$, hence the conic section is non-singular. □

The points P_1 and P_2 lie on Q and belong to the parameter values $(\lambda_0 : \lambda_1) = (1:0)$ and $(\lambda_0 : \lambda_1) = (0:1)$ respectively. The point P_0 determines the direction of the tangents to Q in P_1 and P_2.

Lemma 8.10
The lines P_0P_1 and P_0P_2 touch the conic section Q in P_1 and in P_2 respectively.

Proof
We carry out the proof for the point P_1. The line P_0P_1 is defined by the equation $x_2 = 0$. In P_1 the gradient of the left hand side has the value $(0 : 0 : 1)$. Evaluation of the gradient of the describing polynomial of Q in P_1 yields $(0 : 0 : k^2)$. Since $k \neq 0$ these values are equal, which shows that P_0P_1 is tangent to Q in P_1. $\square$

The parametrization of theorem (8.9) is nothing else than the weighted quadratic Bezier curve with control points P_1, P_2 and P_0.

Definition 8.11 (Parametrization of a conic section in $\mathbb{P}(\mathfrak{Circ})$)
Let $\mathcal{C}_1{}^\circ, \mathcal{C}_2{}^\circ, \mathcal{D}^\circ \in \mathbb{P}(\mathfrak{Circ})$ be representants of circles that do not all belong to the same linear family of circles. Then the set of representants

$$\left\{\mathcal{X}^\circ | \mathcal{X}^\circ = \lambda_1{}^2\mathcal{C}_1{}^\circ + 2k\lambda_1\lambda_2\mathcal{D}^\circ + \lambda_2{}^2\mathcal{C}_2{}^\circ, (\lambda_1 : \lambda_2) \in \mathbb{P}(\mathbb{C})\right\}$$

is a non-singular conic section in $\mathbb{P}(\mathfrak{Circ})$.

This conic section lies in the plane defined by $\mathcal{C}_1{}^\circ$, $\mathcal{C}_2{}^\circ$ and $\mathcal{D}^\circ$. Since the polarity function is linear we can easily describe this plane in a slightly different way.

Lemma 8.12
The conic section

$$\left\{\mathcal{X}^\circ | \mathcal{X}^\circ = \lambda_1{}^2\mathcal{C}_1{}^\circ + 2k\lambda_1\lambda_2\mathcal{D}^\circ + \lambda_2{}^2\mathcal{C}_2{}^\circ, (\lambda_1 : \lambda_2) \in \mathbb{P}(\mathbb{C})\right\}$$

lies in the polar plane of the unique circle that is orthogonal to $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{D}$.

Proof
Let $\mathcal{O}^\circ$ be the intersection of the polar planes of $\mathcal{C}_1{}^\circ$, $\mathcal{C}_2{}^\circ$ and $\mathcal{D}^\circ$. $\mathcal{O}^\circ$ is uniquely defined, because $\mathcal{C}_1{}^\circ$, $\mathcal{C}_2{}^\circ$ and $\mathcal{D}^\circ$ are not collinear. $\mathcal{O}$ is the circle that cuts the given circles $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{D}$ at right angle.

The proposition follows immediately from the linearity of the polarity function. For a circle $\mathcal{X}$ represented by the point $\mathcal{X}^\circ$ on the conic section holds

$$P(\mathcal{O}^\circ, \mathcal{X}^\circ) = \lambda_1{}^2 P(\mathcal{O}^\circ, \mathcal{C}_1{}^\circ) + 2k\lambda_1\lambda_2 P(\mathcal{O}^\circ, \mathcal{D}^\circ) + \lambda_2{}^2 P(\mathcal{O}^\circ, \mathcal{C}_2{}^\circ) = 0.$$

Thus $\mathcal{O}$ is orthogonal to $\mathcal{X}$. $\square$

The representants $\mathcal{C}_1{}^\circ$ and $\mathcal{C}_2{}^\circ$ lie on the conic section and the lines $\mathcal{C}_1{}^\circ\mathcal{D}^\circ$ and $\mathcal{C}_2{}^\circ\mathcal{D}^\circ$ are tangent to it in these points.

Envelope of a rational family of circles

A conic section in $\mathbb{P}(\mathfrak{Circ})$ is determined by three points through the rational parametrization

$$\lambda_1{}^2\mathcal{C}_1{}^\circ + 2k \cdot \lambda_1\lambda_2\mathcal{D}^\circ + \lambda_2{}^2\mathcal{C}_2{}^\circ$$

with $(\lambda_1 : \lambda_2) \in \mathbb{P}(\mathbb{C})$ and $k \in \mathbb{C}$. For each of the given circles $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{D}$ we can choose different representants. This affects the parametrization, of course.

The representants $\mathcal{C}_1{}^\circ$ and $\mathcal{C}_2{}^\circ$ belong to this conic section. These are the points, where the tangents from $\mathcal{D}^\circ$ touch the conic section. The coefficient k determines the shape of the conic section. For a more convenient way of writing we introduce

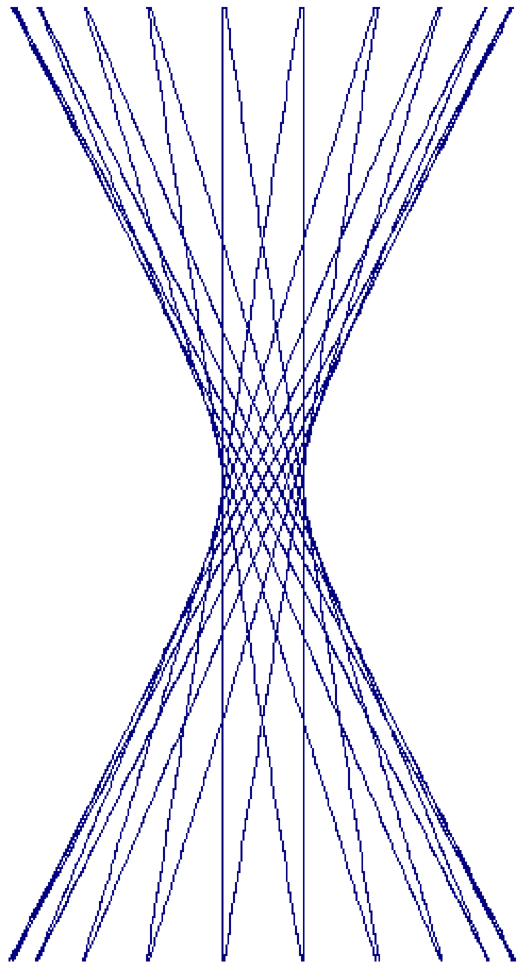


Figure 8.3: Hyperbola as the envelope of a rational family of lines

Definition 8.13 (Product of circles)
Let $\mathcal{C}$ and $\mathcal{D}$ be circles with representants

$$\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ \text{ and } \mathcal{D}^\circ = (d_0 : d_1 : d_2 : d_3)^\circ.$$

Then

$$\mathcal{C} \cdot \mathcal{D}$$

stands for the curve defined by the equation

$$\Big(c_0({x_1}^2+{x_2}^2)-2c_1x_0x_1-2c_2x_0x_2+c_3{x_0}^2\Big) \cdot \Big(d_0({x_1}^2+{x_2}^2)-2d_1x_0x_1-2d_2x_0x_2+d_3{x_0}^2\Big) = 0.$$

With this tools we can easily calculate the envelope of a rational family of circles.

Theorem 8.14 (Envelope of a rational family of circles)
The family of circles $\mathcal{X}$ represented by

$$\mathcal{X}^\circ = {\lambda_1}^2{\mathcal{C}_1}^\circ + 2k \cdot \lambda_1\lambda_2\mathcal{D}^\circ + {\lambda_2}^2{\mathcal{C}_2}^\circ$$

has the bicircular quartic B with the equation

$$B : \mathcal{C}_1 \cdot \mathcal{C}_2 - k^2 \cdot \mathcal{D}^2 = 0 \tag{8.9}$$

as envelope.

Proof
Let $p(\lambda, \mu)$ be the polynomial defining the circle $\mathcal{X}$ for the parameter $(\lambda : \mu)$. We write p_λ and p_μ for the partial derivatives

$$\frac{\partial p}{\partial \lambda} = 2\lambda \mathcal{C}_1 + 2\mu k \mathcal{D} \quad \text{and} \quad \frac{\partial p}{\partial \mu} = 2\lambda k \mathcal{D} + 2\mu \mathcal{C}_2.$$

Then the point P lies on the envelope if and only if p, p_λ and p_μ vanish in P for some $(\lambda : \mu) = (\lambda_P : \mu_P)$. The polynomial p is homogeneous of degree 2 in λ and μ. From Euler's homogeneous function theorem follows

$$2p = \lambda p_\lambda + \mu p_\mu.$$

We show first that if P lies on the quartic B then it is a zero of p_λ and p_μ for some parameter $(\lambda : \mu)$. If $\mathcal{D}$ vanishes in P then so does $\mathcal{C}_1\mathcal{C}_2$ because of equation (8.9). Thus we may assume without loss of generality that $\mathcal{C}_1$ also vanishes in P. In this case p_λ vanishes in P and p_μ vanishes in P for $(\lambda : \mu) = (1 : 0)$. If $\mathcal{D}$ does not vanish in P then p_λ vanishes there for $(\lambda : \mu) = (k\mathcal{D}(P) : -\mathcal{C}_1(P))$. The case $(0 : 0)$ can not occur, because $k\mathcal{D}(P)$ is a product of two nonzero factors. We substitute $(\lambda : \mu)$ into p_μ and obtain

$$p_\mu(P) = 2k^2\mathcal{D}(P)\mathcal{D}(P) - 2\mathcal{C}_1(P)\mathcal{C}_2(P).$$

From equation (8.9) follows that $p_\mu(P)$ is zero. This shows that every point of B belongs to the envelope.

Now we prove that all points on the envelope fulfil equation (8.9) and therefore lie on B. Again we start with the case $\mathcal{D}(P) = 0$. Because $(\lambda : \mu) \neq (0 : 0)$, we may assume $\lambda \neq 0$ without loss of generality. p_λ then vanishes in P exactly if $\mathcal{C}_1(P) = 0$. But then P lies on B, since it is also a zero of $(\mathcal{C}_1\mathcal{C}_2 - k^2\mathcal{D}^2)$. In the case $\mathcal{C}(P) \neq 0$ we use that p_λ vanishes in P for the parameter $(\lambda : \mu) = (k\mathcal{D}(P) : -\mathcal{C}_1(P))$, which is always different from $(0 : 0)$. We substitute this into p_μ and like before obtain

$$p_\mu(P) = 2k^2\mathcal{D}(P)\mathcal{D}(P) - 2\mathcal{C}_1(P)\mathcal{C}_2(P).$$

We assumed that $p_\mu(P) = 0$, thus P belongs to B. $\square$

The proof shows us a severe problem. Any different choice of one of the representants $\mathcal{C}_1{}^\circ$, $\mathcal{C}_2{}^\circ$ and $\mathcal{D}^\circ$ changes the shape of rational family of circles. In order to avoid this the equation has to be homogeneous with respect to the representants. This is ensured by choosing in some sense the unit representant from the set of all possible representants of the circle $\mathcal{C}$.

Definition 8.15 (Unit representant of a circle)
Let $\mathcal{C}^\circ = (c_0 : c_1 : c_2 : c_3)^\circ \in \mathbb{P}(\mathfrak{Circ})$ be a representant of a circle. Then we want to call

$$|\mathcal{C}^\circ| = |c_0| + |c_1| + |c_2| + |c_3|$$

the absolute value *of $\mathcal{C}^\circ$ and*

$$\hat{\mathcal{C}}^\circ = \left(\frac{c_0}{|\mathcal{C}^\circ|} : \frac{c_1}{|\mathcal{C}^\circ|} : \frac{c_2}{|\mathcal{C}^\circ|} : \frac{c_3}{|\mathcal{C}^\circ|} \right)^\circ$$

the unit representant *of $\mathcal{C}^\circ$.*

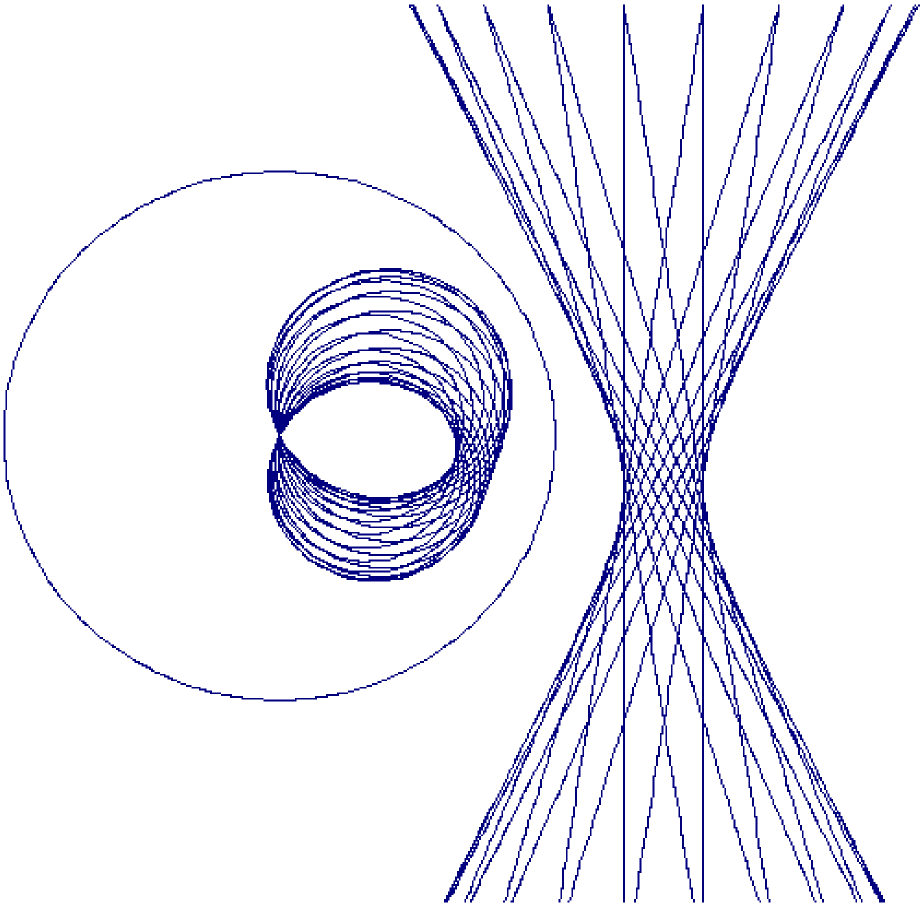


Figure 8.4: Hyperbola as the envelope of a rational family of lines and the inverse family of circles with its envelope - a singular bicircular quartic

Remark
This definition of the unit representant of a circle is rather arbitrary. There are many other ways to define the absolute value of a representant. Our choice is derived from the L^1-norm, but any other L^n-norm could be used instead.

We now have a rational parametrization that is homogeneous in the representants,

$$\lambda_1{}^2\hat{\mathcal{C}}_1^\circ + 2k \cdot \lambda_1\lambda_2\hat{\mathcal{D}}^\circ + \lambda_2{}^2\hat{\mathcal{C}}_2^\circ, \tag{8.10}$$

with $(\lambda_1 : \lambda_2) \in \mathbb{P}(\mathbb{C})$ and $k \in \mathbb{C}$.

The representants $\hat{\mathcal{C}}_1^\circ$ and $\hat{\mathcal{C}}_2^\circ$ belong to this conic section. These are the points, where the tangents from $\hat{\mathcal{D}}^\circ$ touch the conic section. The coefficient k determines the shape of the conic section in $\mathbb{P}(\mathfrak{Circ})$. The values $k = \pm 1$ lead to a parabola. The parametrization makes sense only if the circles $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{D}$ do not all belong to one single linear family of circles.

We have shown now that the envelope of a rational family of circles is a bicircular quartic in general. There are of course also degenerate cases, for instance, when the three representants $\mathcal{C}_1{}^\circ$, $\mathcal{C}_2{}^\circ$ and $\mathcal{D}^\circ$ are collinear in $\mathbb{P}(\mathfrak{Circ})$. It is also possible that the envelope consists of a pair of lines or forms a conic section in $\mathbb{P}(\mathbb{C})$. One important degenerate case is when the three representants are taken from $\mathfrak{Circ}$ and for the parameter k holds $k = \pm 1$. Then the terms of fourth order cancel out and the resulting envelope is a circular cubic. We will not investigate in this direction here.

An important observation is that we can write bicircular quartics in a new way.

Definition 8.16 (Three circles form)
The equation

$$B : \mathcal{C}_1\mathcal{C}_2 - k^2\mathcal{D}^2 = 0$$

is called the three circles form *of a bicircular quartic.*

9 Inversion of Bicircular Quartics

9.1 Bicircular quartics as envelopes

In this chapter we are going to examine the influence of inversions on bicircular quartics. Before we focus on this we want to answer a question that is to the closest connected with the end of the previous chapter. We showed that the envelope of a rational family of circles is a bicircular quartic. We will now show that the converse is also true. We want to prove that every bicircular quartic is also the envelope of a rational family of circles.

Geometric properties of the C-Q-form

We showed in section (8.2) that every bicircular quartic B can be rewritten in the form of equation (8.8). We called this the C-Q-form

$$B : {\mathcal{C}_t}^2 - {x_0}^2 Q_t = 0,$$

where t is the square of the radius of $\mathcal{C}_t$. We deduce some important geometric results from this equation and start from the quite obvious

Lemma 9.1 (Common points of B, $\mathcal{C}_t$ and Q_t)
For a point $P \in \mathbb{P}_2(\mathbb{C})$ any two of the following statements imply the third.
(i) P *lies on* B*,*
(ii) P *lies on* $\mathcal{C}_t$*,*
(iii) P *lies on* ${x_0}^2 Q_t$*.*

The particular implication $(i) \wedge (ii) \Rightarrow (iii)$ of this lemma can also be expressed in the following words.

Lemma 9.2
All common points of B *and* $\mathcal{C}_t$ *also lie on* ${x_0}^2 Q_t$*.*

Moreover we see from equation (8.8) that

Theorem 9.3 (Intersection multiplicities in the C-Q-form)
Let

$$B : {\mathcal{C}_t}^2 - {x_0}^2 Q_t = 0$$

be the C-Q-form of a bicircular quartic B*. Then for any point* $P \in \mathbb{P}_2(\mathbb{C})$

$$\operatorname{mult}_P(B, {x_0}^2 Q_t) = 2 \cdot \operatorname{mult}_P(B, \mathcal{C}_t).$$

Proof
We assume that no two curves of B, $\mathcal{C}_t$ and Q_t have a common component and that no line through the point $(1:0:0) \in \mathbb{P}_2(\mathbb{C})$ contains more than one intersection point of these curves. Then for the intersection multiplicity holds

$$\mathrm{mult}_P(B, {x_0}^2 Q_t) = \mathrm{mult}_P(B, B - {\mathcal{C}_t}^2) = \mathrm{mult}_P(B, {\mathcal{C}_t}^2 - B) = \mathrm{mult}_P(B, {\mathcal{C}_t}^2)$$

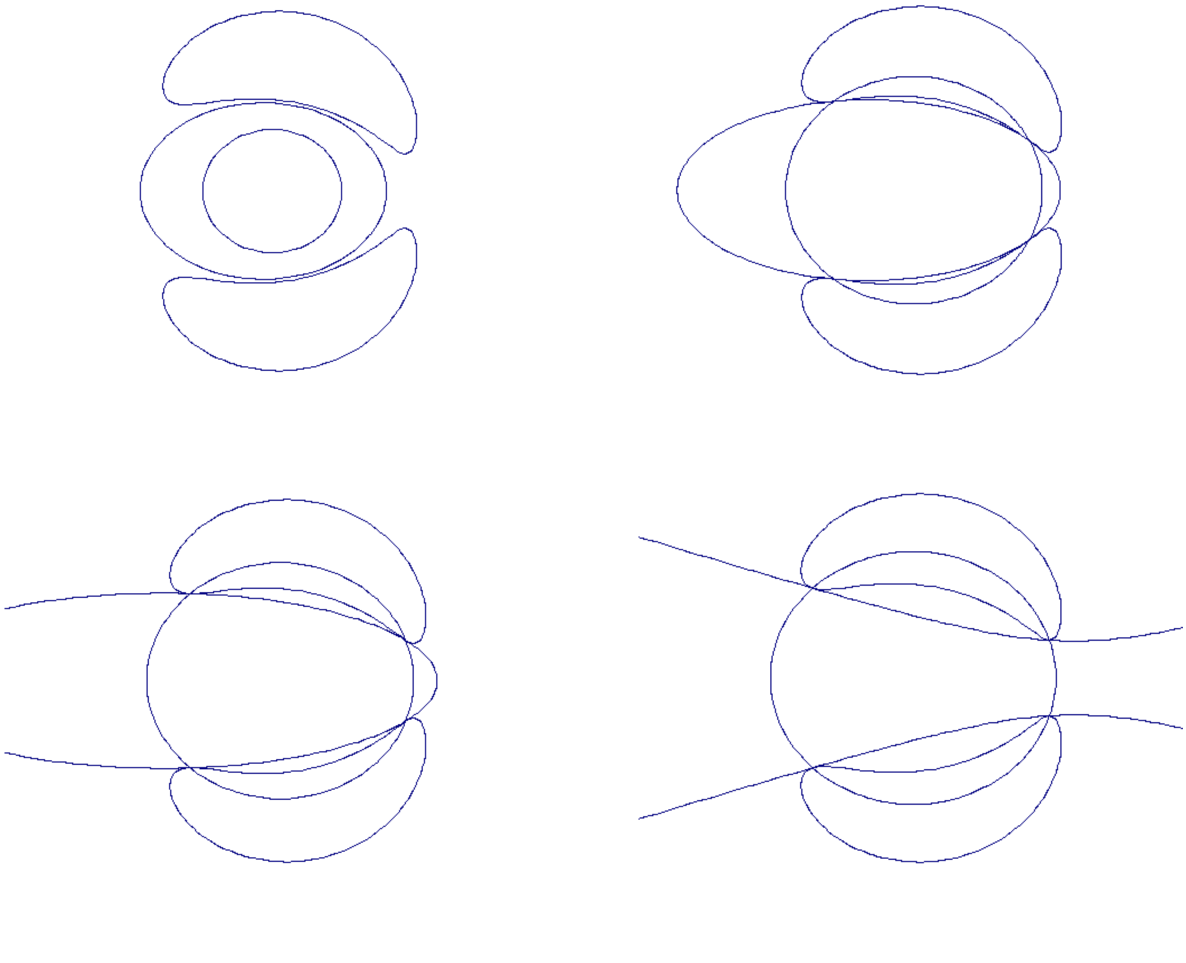
due to the definition of the intersection multipicity as the dimension of a ring. The equation

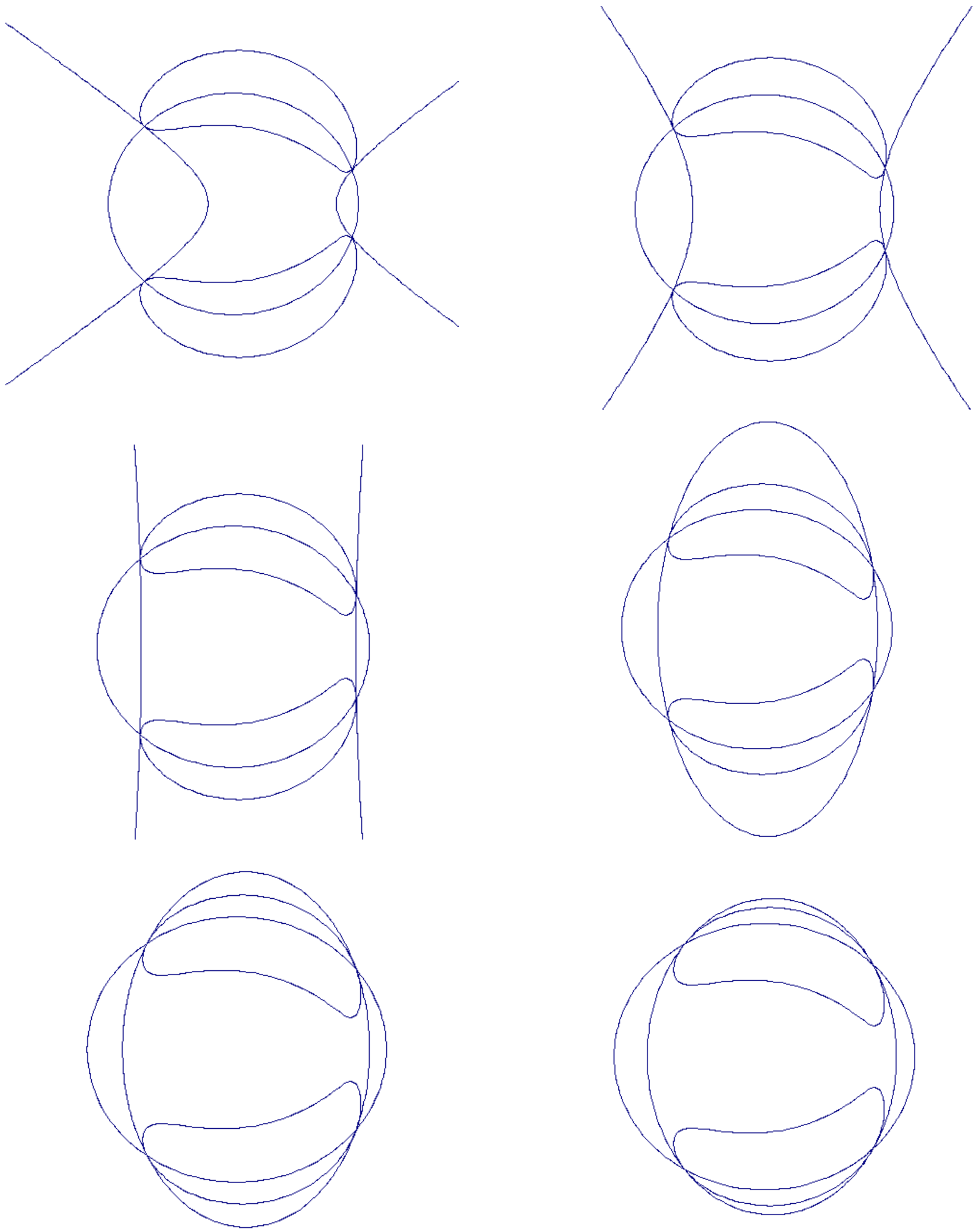
$$\mathrm{mult}_P(B, {\mathcal{C}_t}^2) = 2 \cdot \mathrm{mult}_P(B, \mathcal{C}_t)$$

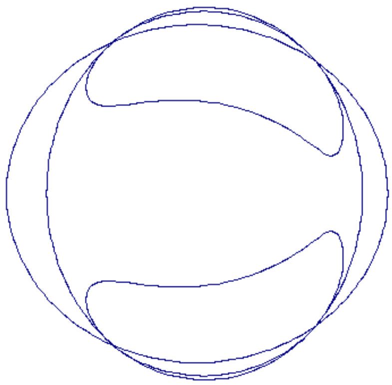
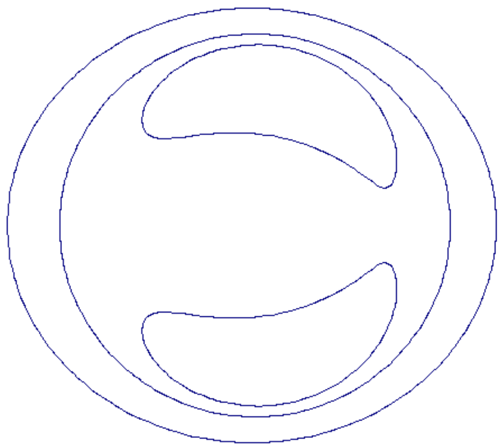
completes the proof. □

This theorem tells us that every intersection point of the bicircular quartic B and the variable circle $\mathcal{C}_t$ is at the same time a point where the quadric Q_t touches B.

The following figures show the variation of the quadric Q_t for different values of t. In the first and the last image the circle C_t does not intersect the bicircular curve in real, but in complex points.

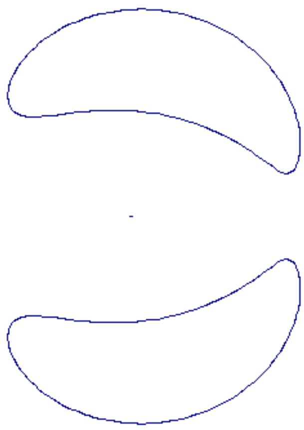
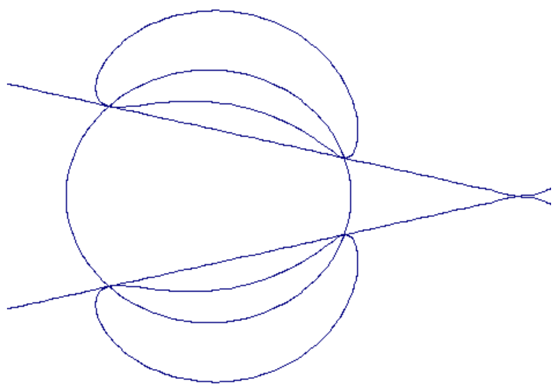


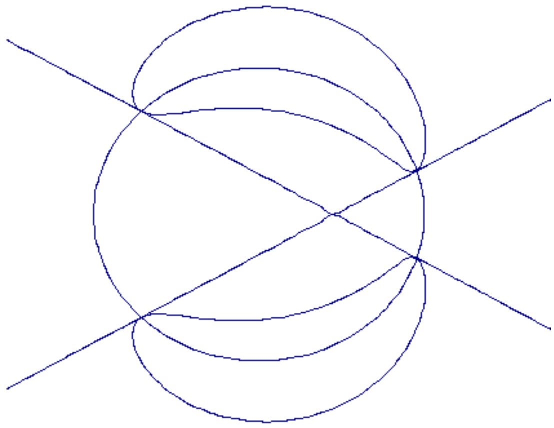
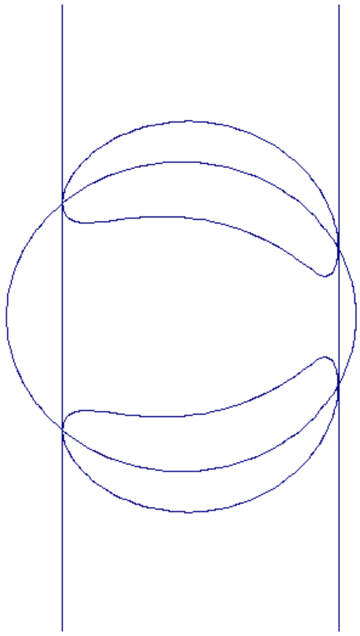


Degeneration of the quadric Q_t

We see in the preceding figures that there are four special values for t where the quadric Q_t degenerates. The first value corresponds to the transition, when the quadric becomes real and contains exactly one real point. The other values belong to the change of type from an ellipse to a hyperbola between the third and the fourth image, the change of the opening direction of the hyperbola between the fourth and the fifth image and the change of type between image number nine and ten. In all four cases the quadric splits into two lines that touch the bicircular quartic in two different points. In the first image there is nothing to see, because the lines are imaginary and only their real intersection point is visible.

We recall the equation (8.7) of a bicircular quartic in C-Q-form as

$$\begin{aligned}&({x_1}^2+{x_2}^2+t{x_0}^2)^2+\\&+{x_0}^2\Big((c_{20}-2t){x_1}^2+(c_{02}-2t){x_2}^2+2c_{10}x_0x_1+2c_{01}x_0x_2+(c_{00}-t^2){x_0}^2\Big) \quad = \quad 0.\end{aligned}$$

The quadric Q_t is determined by the coefficient matrix

$$\begin{pmatrix} c_{20}-2t & 0 & c_{10} \\ 0 & c_{02}-2t & c_{01} \\ c_{10} & c_{01} & c_{00}-t^2 \end{pmatrix}$$

which is depending on the parameter t. This matrix has the determinant

$$\begin{aligned}\Delta_t \quad = \quad & -4t^4 \\ & +(2c_{20}+2c_{02})\,t^3 \\ & +(4c_{00}-c_{20}c_{02})\,t^2 \\ & +({c_{10}}^2+{c_{01}}^2-2c_{20}c_{00}-2c_{02}c_{00})\,t \\ & +(c_{20}c_{02}c_{00}-c_{20}{c_{01}}^2-c_{02}{c_{10}}^2)\,.\end{aligned} \tag{9.1}$$

Δ_t is a polynomial of fourth degree in t, thus it has four complex zeros when counted with multiplicity. For each of these zeros the determinant Δ_t of the quadric vanishes and Q_t splits into the union of two straight lines for these values of t.

Definition 9.4 (Double tangents)
A line L is called a double tangent *of the curve C, if either there are at least two distinct points P, Q with* $\mathrm{mult}_P(L,C) \geq 2$ *and* $\mathrm{mult}_Q(L,C) \geq 2$ *or there is at least one point P with* $\mathrm{mult}_P(L,C) \geq 4$.

The second condition is apparently the degeneration of the first, where the two points of contact lie infinitesimally close to each other. The curve C is a bicircular quartic throughout this text and is cut by L in four points. These points group into two points of intersection multiplicity 2 in the first case or into a single point with intersection multiplicity 4. Because of this the we can read the signs $\geq$ of inequality as equalities $=$.

Theorem 9.5 (Double tangents of a bicircular quartic)
When the quadric Q_t in the C-Q-form is reducible, then the components of Q_t are double tangents to the bicircular quartic B.

Proof
We have shown in theorem (9.3) that the quadric Q_t touches B in the points where C_t intersects B. This remains true, when Q_t splits into a pair of lines. Each component of the quadric intersects B in four points when counted with multiplicity. But because the intersection multiplicity in these points is at least 2, there are either two points with multiplicity 2 or one point of multiplicity 4, making the component a double tangent to B.
□

Determination of the rational family of circles

We will see that there is a generic way to construct a rational family of circles which has a given bicircular quartic as envelope. The double tangents to a given bicircular quartic are straight lines, thus they are generalized circles. We can use this fact to construct a special three circles form from the C-Q-form of a given bicircular quartic. All we need is a slightly technical

Lemma 9.6
Let B be a bicircular quartic with

$$B : {\mathcal{C}_t}^2 - {x_0}^2 Q_t = 0,$$

let Δ_t be the coefficient matrix of the quadric Q_t and let τ be a zero of Δ_t. The components of Q_τ be $L_{\tau,1}$ and $L_{\tau,2}$. Then the circle represented by

$$\lambda^2 x_0 \hat{L}_1^\circ + 2k\lambda\mu\hat{\mathcal{C}}_\tau^\circ + \mu^2 x_0 \hat{L}_2^\circ \in \mathbb{P}(\mathfrak{Circ})$$

with $k = \frac{|\mathcal{C}_\tau{}^\circ|}{\sqrt{|x_0 L_1{}^\circ||x_0 L_2{}^\circ|}}$ touches B in two points for all $(\lambda : \mu) \in \mathbb{P}(\mathbb{C})$.

Proof
The rational family of circles

$$\lambda^2 x_0 \hat{L}_1^\circ + 2k\lambda\mu\hat{\mathcal{C}}_\tau^\circ + \mu^2 x_0 \hat{L}_2^\circ \in \mathbb{P}(\mathfrak{Circ})$$

has the bicircular quartic

$$x_0 \hat{L}_1 x_0 \hat{L}_2 - k^2 \cdot \hat{\mathcal{C}}_\tau^2 = 0$$

as envelope. Substitution of the value of k shows that this equation is equivalent to

$$x_0 L_1 x_0 L_2 - {\mathcal{C}_\tau}^2 = 0. \qquad \square$$

Corollary 9.7 (Rational family of circles with given envelope)
With the notion of the preceding lemma the rational family of circles represented by

$$\left\{\lambda^2 x_0 \hat{L}_1^\circ + 2k\lambda\mu\hat{\mathcal{C}}_\tau^\circ + \mu^2 x_0 \hat{L}_2^\circ \in \mathbb{P}(\mathfrak{Circ}) : (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\right\}$$

with $k = \frac{|\mathcal{C}_\tau{}^\circ|}{\sqrt{|x_0 L_1{}^\circ||x_0 L_2{}^\circ|}}$ *has the bicircular quartic* B *as envelope.*

From the above investigations we can deduce

Corollary 9.8 (Bicircular quartics are envelopes)
Every bicircular quartic is the envelope of some rational family of circles.

But moreover holds

Theorem 9.9 (Four rational families of circles)
There are in general four rational families of circles that have a given bicircular quartic as envelope.

Proof
The determinant Δ_t of the quadric Q_t is a polynomial of fourth degree in t. It has four zeros in general and we want to call these zeros α, β, γ and δ.

These four zeros give rise to four different C-Q-forms. The quadrics $Q_\alpha, Q_\beta, Q_\gamma$ and Q_δ are reducible by definition and split into the lines $L_{\alpha,1}, L_{\alpha,2}$ and so on.

Let us now without loss of generality assume that the rational families of circles represented by

$$\left\{\lambda^2 x_0 \hat{L}_{\alpha,1}^\circ + 2k\lambda\mu\hat{\mathcal{C}}_\alpha^\circ + \mu^2 x_0 \hat{L}_{\alpha,2}^\circ \in \mathbb{P}(\mathfrak{Circ}) : (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\right\}$$

and

$$\left\{\lambda^2 x_0 \hat{L}_{\beta,1}^\circ + 2k\lambda\mu\hat{\mathcal{C}}_\beta^\circ + \mu^2 x_0 \hat{L}_{\beta,2}^\circ \in \mathbb{P}(\mathfrak{Circ}) : (\lambda : \mu) \in \mathbb{P}(\mathbb{C})\right\}$$

are identical. This implies that the conic sections have the same intersection points with the plane at infinity of $\mathbb{P}(\mathfrak{Circ})$. These common points are $L_{\alpha,1}^\circ, L_{\alpha,2}^\circ, L_{\beta,1}^\circ$ and $L_{\beta,2}^\circ$. Comparing the coefficients of equation (8.8) shows that

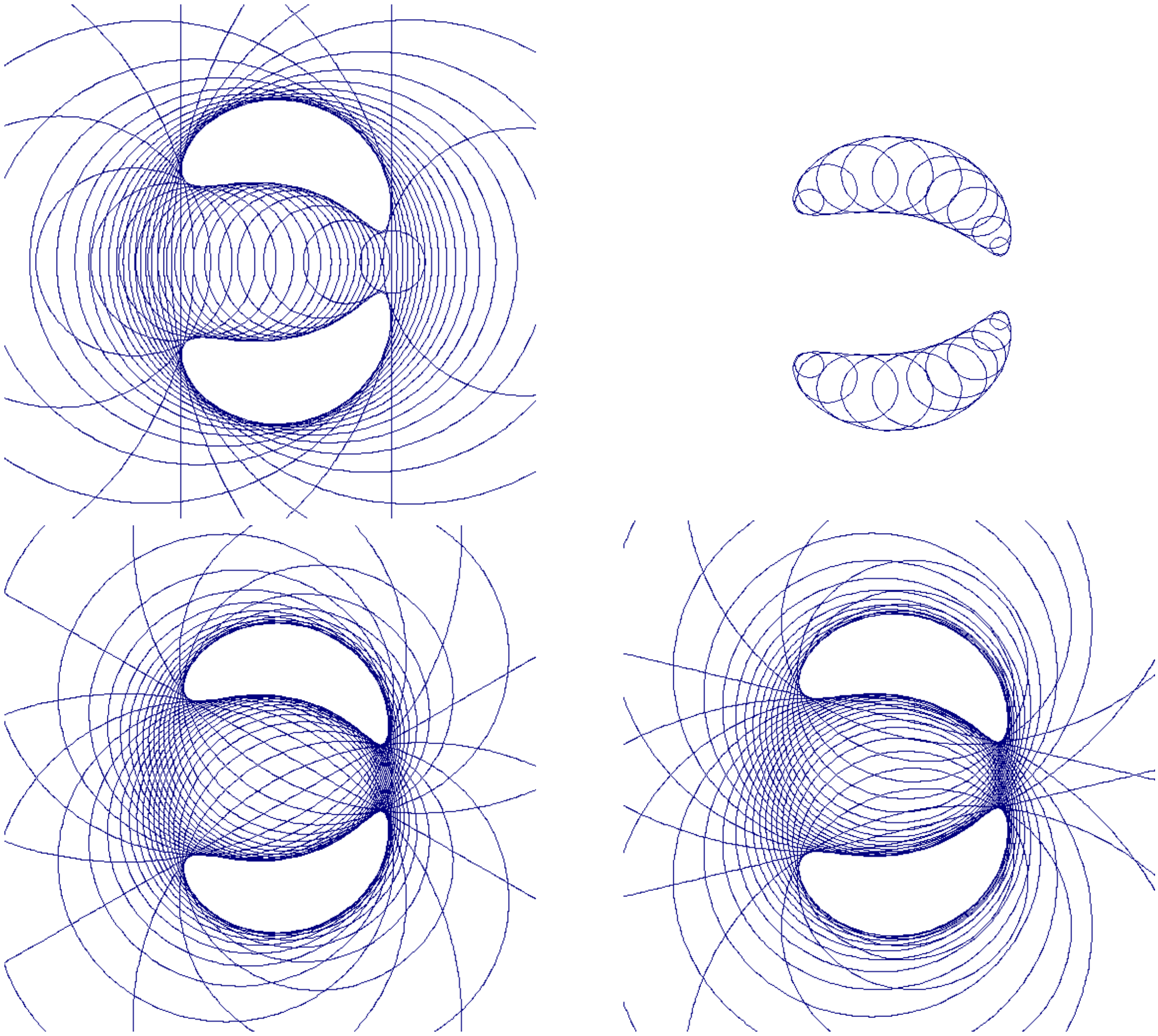
$$\left\{L_{\alpha,1}^\circ, L_{\alpha,2}^\circ\right\} \neq \left\{L_{\beta,1}^\circ, L_{\beta,2}^\circ\right\} \quad \text{for} \quad \alpha \neq \beta.$$

Therefore

$$\left\{L_{\alpha,1}^\circ, L_{\alpha,2}^\circ, L_{\beta,1}^\circ, L_{\beta,2}^\circ\right\}$$

forms a set of at least three points lying on a conic section in $\mathbb{P}(\mathfrak{Circ})$ and at the same time lying on the plane at infinity. This is possible only if the conic section lies entirely in the plane at infinity. This contradicts the fact that $\mathcal{C}_\alpha$ is a proper circle, hence the above assumption is false and the two conic sections in $\mathbb{P}(\mathfrak{Circ})$ are different. □

The following images show the four rational families of circles that have the bicircular quartic from our example as envelope.



9.2 Inversion of bicircular quartics

We have shown in section (8.1) that the inverse curve of a general quadric is a bicircular quartic. Now we want to examine how inversions act on bicircular quartics.

The image under the unit inversion

We want to investigate the behaviour of bicircular quartics under inversions. In order to do so we recall the equation (8.4) of a bicircular quartic as

$$\begin{aligned} & a({x_1}^2+{x_2}^2)^2 + 2x_0(b_1x_1+b_2x_2)({x_1}^2+{x_2}^2)+ \\ & {x_0}^2\Big(c_{20}{x_1}^2+2c_{11}x_1x_2+c_{02}{x_2}^2+2c_{10}x_0x_1+2c_{01}x_0x_2+c_{00}{x_0}^2\Big) \quad = \quad 0. \end{aligned}$$

We also recall the inversion ε about the unit circle $\mathcal{E}$ as

$$\varepsilon : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} {x_1}^2 + {x_2}^2 \\ x_0 x_1 \\ x_0 x_2 \end{pmatrix}.$$

Theorem 9.10 (Image of a bicircular quartic under the unit inversion)
Let B be a bicircular quartic. The the image $\varepsilon(B)$ of B under the inversion ε about the unit circle $\mathcal{E}$ is

$$\begin{aligned} &c_{00}({x_1}^2 + {x_2}^2)^2 + 2x_0(c_{10}x_1 + c_{01}x_2)({x_1}^2 + {x_2}^2) + \\ &{x_0}^2\Big(c_{20}{x_1}^2 + c_{11}x_1x_2 + c_{02}{x_2}^2 + 2b_1x_0x_1 + 2b_2x_0x_2 + a{x_0}^2\Big) = 0. \end{aligned}$$

Proof
The proof is carried out by calculation. Application of $\varepsilon(B)$ on equation (8.4) gives

$$\begin{aligned} a({x_0}^2{x_1}^2 + {x_0}^2{x_2}^2)^2 + 2({x_1}^2 + {x_2}^2)(b_1x_0x_1 + b_2x_0x_2)({x_0}^2{x_1}^2 + {x_0}^2{x_2}^2) + \\ ({x_1}^2 + {x_2}^2)^2\Big(c_{20}{x_0}^2{x_1}^2 + 2c_{11}x_0x_1x_0x_2 + \\ c_{02}{x_0}^2{x_2}^2 + 2c_{10}({x_1}^2 + {x_2}^2)x_0x_1 + \\ 2c_{01}({x_1}^2 + {x_2}^2)x_0x_2 + c_{00}({x_1}^2 + {x_2}^2)^2\Big) &= 0. \end{aligned}$$

The factor $\left({x_1}^2 + {x_2}^2\right)^2$ of the left side can be ignored, because this is just an arteficial component of the image. Thus we obtain

$$\begin{aligned} a{x_0}^4 + 2(b_1x_0x_1 + b_2x_0x_2){x_0}^2 + \\ (c_{20}{x_0}^2{x_1}^2 + 2c_{11}x_0x_1x_0x_2 + \\ c_{02}{x_0}^2{x_2}^2 + 2c_{10}({x_1}^2 + {x_2}^2)x_0x_1 + \\ 2c_{01}({x_1}^2 + {x_2}^2)x_0x_2 + c_{00}({x_1}^2 + {x_2}^2)^2\Big) &= 0. \end{aligned}$$

The terms of this equation can be sorted in the same way as the equation of B. The equation of $\varepsilon(B)$ becomes

$$\begin{aligned} &c_{00}({x_1}^2 + {x_2}^2)^2 + 2x_0(c_{10}x_1 + c_{01}x_2)({x_1}^2 + {x_2}^2) + \\ &{x_0}^2\Big(c_{20}{x_1}^2 + c_{11}x_1x_2 + c_{02}{x_2}^2 + 2b_1x_0x_1 + 2b_2x_0x_2 + a{x_0}^2\Big) = 0. \end{aligned}$$

We see immediately that the inversion about the unit circle maps the bicircular quartic with coefficient vector

$$(a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00})$$

onto the bicircular quartic with coefficient vector

$$(c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}, b_1, b_2, a). \qquad \square$$

The calculation of this proof also shows that the factor $({x_1}^2 + {x_2}^2)^2$ can always be factored out after the substitution step. We have already met this behaviour in section (4.2). On the other hand the nullcircle $({x_1}^2 + {x_2}^2)$ might be indeed a true component of the image curve, e.g. because the line at infinity is a component of the original curve.

For the inversion of circles we could avoid this problem completely by taking each circle as a whole. The inversion then acts on its representant in $\mathbb{P}(\mathfrak{Circ})$ instead of the coordinates of its equation. The same is possible for bicircular quartics, too. We can state

Corollary 9.11 (Inversion of bicicular curves)
The unit inversion ε acts on the space of bicircular curves as

$$(a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00}) \mapsto (c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}, b_1, b_2, a).$$

Circular cubics

The method of generalization from the inversion ε about the unit circle to an arbitrary inversion that we have developed in section (4.2) can also be applied in the case of bicircular quartics. We can express any inversion about a circle as the composition of a translation, a dilation and the special inversion ε. We will deduce the effect of a general inversion on a bicircular quartic from the properties of ε.

The points at infinity in the space of circles do not represent true circles, but straight lines. We unified these two kinds of curves as generalized circles. In the same way the space of bicircular quartics consists of points that represent generalized bicircular quartics.

Definition 9.12 (Circular cubic)
The point $(a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00})$ in the space of bicircular quartics represents a proper bicircular quartic exactly if $a \neq 0$. For $a = 0$ it represents a circular cubic.

Remark
For $a = 0$ the equation (8.4) contains the factor x_0. Thus circular cubics can be regarded as an important example of bicircular quartics having an exceptional component.

We can handle circular cubics in the same way as we interpreted straight lines as reducible circles that have the line at infinity as one component. For this unfication we do not even have to adapt the definition of a bicircular quartic as a quartic curve that has singularities in the two circular points. When interpreted as a curve of degree four a circular cubic is singular in these points, because they are the intersection points of the cubic component and the line at infinity. Hence they are singularities. From the action of the the inversion ε on the space of bicircular quartics we immediately deduce

Theorem 9.13 (Center of inversion on bicircular quartic)
The inverse curve of a bicircular quartic is a circular cubic if and only if the center of the circle of inversion lies on the original curve.

Proof
It is sufficient to prove this for the inversion about the unit circle. The original curve runs through the origin exactly if $c_{00} = 0$. But if this is true, then the inverse curve is a circular cubic.
□

Inversions and the three circles form

We have shown in theorem (8.14) that the envelope of the rational family of circles represented by

$$\mathcal{X}^\circ = {\lambda_1}^2 \hat{\mathcal{C}}_1^\circ + 2k \cdot \lambda_1 \lambda_2 \hat{\mathcal{D}}^\circ + {\lambda_2}^2 \hat{\mathcal{C}}_2^\circ$$

is the bicircular quartic B with the equation

$$B : \hat{\mathcal{C}}_1 \cdot \hat{\mathcal{C}}_2 - k^2 \cdot \hat{\mathcal{D}}^2 = 0.$$

In theorem (9.9) we have given a constructive proof that every bicircular quartic can be written in this form and that there are in general four different ways to do so. We therefore take the bicircular quartic B as given in the form

$$B : \mathcal{C}\mathcal{C}' - k^2 \mathcal{D}^2 = 0,$$

where $\mathcal{C}, \mathcal{C}'$ and $\mathcal{D}$ are the equations of three circles that do not belong to one linear family of circles. This means that the representants

$$\begin{aligned}\mathcal{C}^\circ &= (c_0 : c_1 : c_2 : c_3)^\circ,\\ \mathcal{C}'^\circ &= (c_0' : c_1' : c_2' : c_3')^\circ,\\ \mathcal{D}^\circ &= (d_0 : d_1 : d_2 : d_3)^\circ\end{aligned}$$

do not lie on a straight line in $\mathbb{P}(\mathfrak{Circ})$. According to theorem (5.9) the inversion ε about the unit circle maps these representants to

$$\begin{aligned}\varepsilon(\mathcal{C})^\circ &= (c_3 : c_1 : c_2 : c_0)^\circ,\\ \varepsilon(\mathcal{C}')^\circ &= (c_3' : c_1' : c_2' : c_0')^\circ,\\ \varepsilon(\mathcal{D})^\circ &= (d_3 : d_1 : d_2 : d_0)^\circ.\end{aligned}$$

The image points do not lie on a straight line in $\mathbb{P}(\mathfrak{Circ})$, thus the corresponding set

$$\left\{ {\lambda_1}^2 \varepsilon(\mathcal{C})^\circ + 2k \cdot \lambda_1 \lambda_2 \varepsilon(\mathcal{D})^\circ + {\lambda_2}^2 \varepsilon(\mathcal{C}')^\circ, (\lambda, \mu) \in \mathbb{P}(\mathbb{C}) \right\}$$

is also a rational family of circles. It has the bicircular quartic $\tilde{B}$ with the equation

$$\tilde{B} : \varepsilon(\mathcal{C})\varepsilon(\mathcal{C}') - k^2 \cdot \varepsilon(\mathcal{D})^2 = 0$$

as envelope.

Theorem 9.14 (Inversion of the three circles form)
Let the bicircular quartics B and $\tilde{B}$ be like in the preceding paragraph. Then

$$\varepsilon(B) = \tilde{B}.$$

Proof
The equality is obvious, because the formal inversion of circles by substitution of the variables leads to the cancellation of the factor $({x_1}^2 + {x_2}^2)$. The same procedure applied on bicircular quartics contains the cancellation of the factor $({x_1}^2 + {x_2}^2)^2$. Hence the statement is true. □

We can express this fact in words. The envelope $\tilde{B}$ of the rational family defined by the images of the circles in the three circles form is the image of the envelope B of the original rational family under inversion. This means that it does not matter in which order we compute the envelope or the image under inversion. Both possible ways will lead to the same result.

This seems quite obvious when one takes into account that the inversion is preserving intersection angles. Every circle that belongs to the original rational family in general touches the curve B in two different points. The image of this circle thus also touches the image curve $\varepsilon(B)$ in two points. There remains a small gap in this argumentation, because we have to show that all of these image circles belong to the same rational family of circles, i.e. that their representants are not scattered among the four possible rational families of circles sharing the same envelope $\varepsilon(B)$. We can close this gap by argumenting geometrically in $\mathbb{P}(\mathfrak{Circ})$ that the induced action of ε on the space of circles maps planes onto planes.

Corollary 9.15
The action

$$\varepsilon : (a, b_1, b_2, c_{20}, c_{11}, c_{02}, c_{10}, c_{01}, c_{00}) \mapsto (c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}, b_1, b_2, a)$$

of the inversion about the unit circle on the space of bicircular quartics is induced by its action on $\mathbb{P}(\mathfrak{Circ})$ via the three circles form.

The action of the inversion ε about the unit circle on the space of bicircular curves is determined completely by its action on the space of circles.

Inversion of multicircular curves

We recall equation (8.5)

$$\begin{aligned} &p_0({x_1}^2 + {x_2}^2)^n + x_0 p_1({x_1}^2 + {x_2}^2)^{n-1} + \dots \\ &\dots + {x_0}^{n-2} p_{n-2}({x_1}^2 + {x_2}^2)^2 + {x_0}^{n-1} p_{n-1}({x_1}^2 + {x_2}^2) + {x_0}^n p_n \quad = \quad 0 \end{aligned}$$

of a n-circular $2n$-tic. Circles are examples for $n = 1$, bicircular quartics for $n = 2$. One prominent example for a tricircular curve of sixth order is Watt's curve. Its equation is

$$(x^2 + y^2)^3 - 2d^2(x^2 + y^2)^2 + (4a^2y^2 + d^2)(x^2 + y^2) - 4a^2b^2y^2 = 0.$$

We have demonstrated in theorem (8.6) that the inverse curve of a general n-tic is a n-circular curve of order $2n$. For $n \geq 2$ not all such curves can be images under this special kind of inversion. Hence an inversion must map a general n-circular $2n$-tic on a curve with degree bigger than n. As before we may restrict our investigations to the effect of the inversion ε about the unit circle.

Theorem 9.16 (Inversion of a n-circular $2n$-tic)
Let C be a n-circular curve of order $2n$. Then the image $\varepsilon(C)$ of this curve under inversion about the unit circle is again a n-circular $2n$-tic.

Proof
We take the equation

$$C : \sum_{i=0}^{n} p_i {x_0}^i ({x_1}^2 + {x_2}^2)^{n-i} = 0$$

of C from above. Then

$$\varepsilon(C) : \sum_{i=0}^{n} \varepsilon(p_i) \varepsilon(x_0)^i \varepsilon({x_1}^2 + {x_2}^2)^{n-i} = 0,$$

which is

$$\varepsilon(C) : \sum_{i=0}^{n} \varepsilon(p_i)({x_1}^2 + {x_2}^2)^i {x_0}^{2(n-i)}({x_1}^2 + {x_2}^2)^{n-i} = 0.$$

The factor $({x_1}^2 + {x_2}^2)^n$ is an artifact, hence the equation of the inverse curve is

$$\varepsilon(C) : \sum_{i=0}^{n} {x_0}^{2(n-i)} \varepsilon(p_i) = 0. \tag{9.2}$$

If we write out every polynomial p_i as

$$p_i = \sum_{\alpha+\beta+\gamma=i} c^i_{\alpha,\beta,\gamma} {x_0}^\alpha {x_1}^\beta {x_2}^\gamma,$$

then its image under inversion is

$$\varepsilon(p_i) = \sum_{\alpha+\beta+\gamma=i} c^i_{\alpha,\beta,\gamma} ({x_1}^2 + {x_2}^2)^\alpha (x_0 x_1)^\beta (x_0 x_2)^\gamma.$$

We can rearrange the coefficients into

$$\varepsilon(p_i) = \sum_{j=0}^{i} q_j {x_0}^{i-j} ({x_1}^2 + {x_2}^2)^j$$

where each q_j is a homogeneous polynomial of degree j.

Now we substitute the last expression into equation (9.2), obtaining

$$\varepsilon(C) : \sum_{i=0}^{n} {x_0}^{2(n-i)} \sum_{j=0}^{i} q_j {x_0}^{i-j} ({x_1}^2 + {x_2}^2)^j = 0.$$

Rewriting the sum as

$$\sum_{i=0}^{n} \sum_{j=0}^{i} {x_0}^{2(n-i)} q_j {x_0}^{i-j} ({x_1}^2 + {x_2}^2)^j,$$

changing inner and outer sum

$$\sum_{j=0}^{n} \sum_{i=j}^{n} {x_0}^{2(n-i)} q_j {x_0}^{i-j} ({x_1}^2 + {x_2}^2)^j$$

and extracting some factors independent of i yields

$$\sum_{j=0}^{n} {x_0}^{n-j} ({x_1}^2 + {x_2}^2)^j \sum_{i=j}^{n} q_j {x_0}^{n-i-j}.$$

After shifting the inner summation index we arrive at

$$\sum_{j=0}^{n} {x_0}^{n-j} ({x_1}^2 + {x_2}^2)^j \sum_{i=0}^{n-j} q_{i+j} {x_0}^{n-i} = 0$$

and reversion of the outer sum shows us that the equation of the image curve is

$$\varepsilon(C) : \sum_{j=0}^{n} \left(\sum_{i=0}^{j} q_{i+j} {x_0}^{n-i} \right) {x_0}^j ({x_1}^2 + {x_2}^2)^{n-j} = 0.$$

This equation has the form of equation (8.5), hence the inverse curve is again a n-circular curve of order $2n$. □

Bicircular curves are anallagmatic curves

We have shown in the preceding chapter that we are always able to construct a rational family of circles which has a given bicircular quartic as envelope. We proved in theorem (9.9) that four such rational families exist for a general bicircular quartic.

Definition 9.17 (Orthogonal circle of a rational family)
Let

$$\left\{ {\lambda_1}^2 \hat{\mathcal{C}}_1^\circ + 2k \cdot \lambda_1 \lambda_2 \hat{\mathcal{D}}^\circ + {\lambda_2}^2 \hat{\mathcal{C}}_2^\circ, (\lambda : \mu) \in \mathbb{P}(\mathbb{C}) \right\}$$

be a rational family of circles. The circle $\mathcal{O}$ that is orthogonal to the circles $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{D}$ is called the orthogonal circle *of the rational family.*

The naming in this definition makes sense, because

Theorem 9.18 (Orthogonal circle of a rational family)
The orthogonal circle of a rational family of circles is orthogonal to all circles of the family.

Proof
Let

$$\mathcal{X}^\circ = {\lambda_1}^2 \hat{\mathcal{C}}_1^\circ + 2k \cdot \lambda_1 \lambda_2 \hat{\mathcal{D}}^\circ + {\lambda_2}^2 \hat{\mathcal{C}}_2^\circ$$

represent a circle of the rational family. We can use the same argument as in the proof of theorem (6.11), i.e. the linearity of the polarity function in the first argument, to show that

$$P(\mathcal{X}^\circ, \mathcal{O}^\circ) = 0. \qquad \square$$

Every circle of a rational family of circles is perpendicular to the orthogonal circle of the family. This implies

Corollary 9.19
A rational family of circles is invariant under inversion about its orthogonal circle.

Remark
An even stronger statement holds. Every circle of the family is invariant under this inversion.

Since an inversion preserves angles of intersection, we can deduce from the invariance of the rational family the invariance of its envelope. Hence

Corollary 9.20
Let B be a bicircular quartic and $\mathcal{O}$ the orthogonal circle of a rational family of circles that has B as envelope. Then B is invariant under inversion about $\mathcal{O}$.

Curves with this property have a special name.

Definition 9.21 (Anallagmatic curve)
An algebraic curve that is invariant under inversion about a circle is called anallagmatic.

Hence bicircular quartics (and also circular cubics) are anallagmatic curves.

Theorem 9.22 (Geometric relations of the stabilizing inversions)
Let B be a bicircular quartic. If B is invariant under the inversion about the two circles $\mathcal{O}_1$ and $\mathcal{O}_2$ then these circles intersect at right angle.

Proof
For both circles of inversions we can find the corresponding rational family of circles. Let $\mathfrak{K}_1$ be the family that has $\mathcal{O}_1$ as orthogonal circle and $\mathfrak{K}_2$ be the family that has $\mathcal{O}_2$ as orthogonal circle.

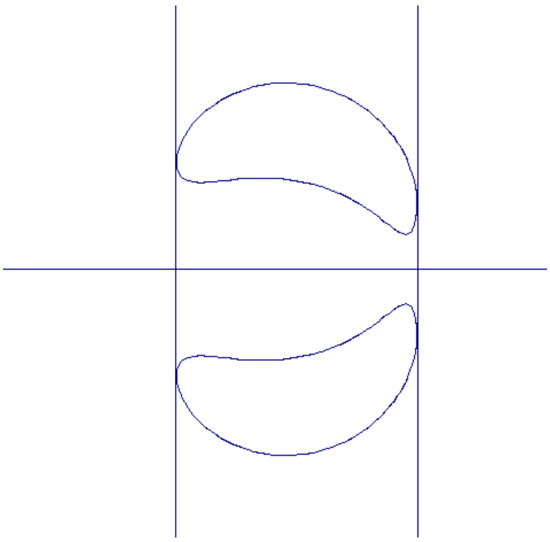
The inversion about $\mathcal{O}_1$ leaves B invariant. Thus it also leaves the rational family $\mathfrak{K}_2$ invariant. This invariance is not circle by circle, but only as a whole set. The invariance of the family now implies the invariance of its orthogonal circle $\mathcal{O}_2$, thus $\mathcal{O}_1 \perp \mathcal{O}_2$. □

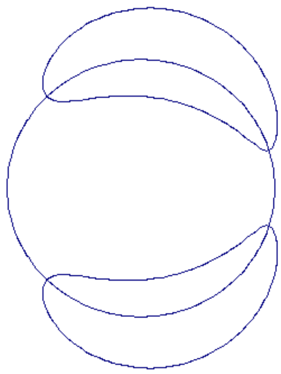
Theorem 9.23 (Invariance of a bicircular quartic under inversion)
A general bicircular quartic is invariant under four different inversions. There is no fifth such inversion.

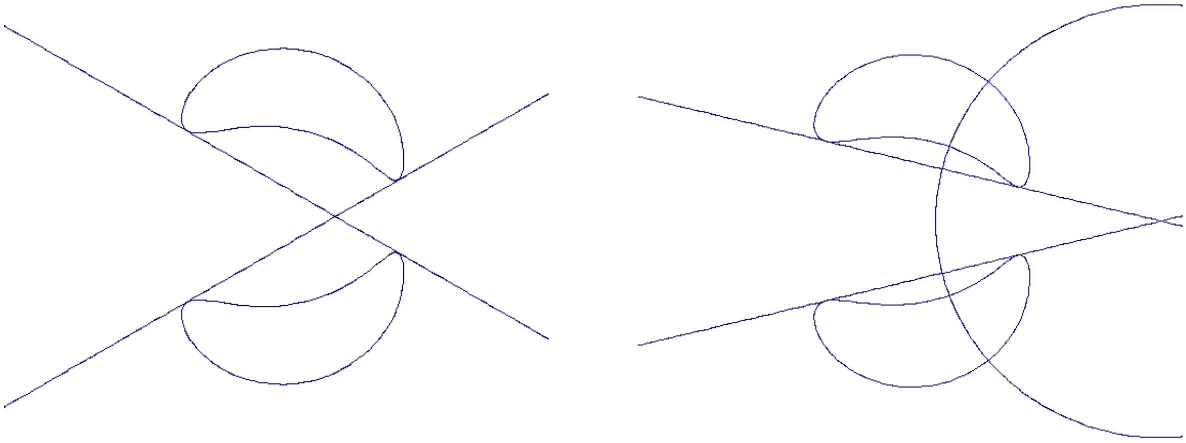
Proof
Let $\iota_1, \ldots, \iota_4$ be four inversions that leave B invariant and $\mathcal{O}_1, \ldots, \mathcal{O}_4$ the corresponding circles of inversion. Then a fifth circle of inversion would have to be perpendicular to all of them at the same time, which is impossible. □

A general bicircular curve is the envelope of four different rational families of circles and it is invariant under four different inversions. The four circles of inversion are mutually orthogonal, which implies that they can not be all real. From the orthogonality follow a lot of other properties. One of them concerns the four centers of the circles of inversion. The orthocenter of the triangle formed by three of them is the fourth center.

The following figures show the circles of inversions for our example curve together with the corresponding double tangents. The double tangents intersect in the center of the circle of inversion. In the second picture the double tangents are imaginary, in the third picture the circle of inversion is imaginary.







9.3 Singularities

Reducibility and number of singularities

It is easy to show that an irreducible quartic curve may have up to three isolated singularities. We want to restrict our analysis of reducible curves to those without multiple components. In this case the number of (isolated) singularities can not exceed six. This maximum can only be reached, if the quartic splits into four lines that are in general position.

We have defined bicircular quartics as curves with singularities in the circular points at infinity, so these curves have at least two and at most six singular points. We will show that examples of bicircular quartics exist for each one of these values.

We begin with the case, when the bicircular quartic is reducible.

Lemma 9.24
A reducible quartic curve decomposes into irreducible components in one of the following ways.

- *two components: a line and a cubic or two quadrics*
- *three components: two lines and a quadric*
- *four components: four lines*

We are going to discuss each of these cases in detail. Because we are studying curves defined by real polynomials, the complex conjugate of a singularity of a curve C is also a singularity of C and both points have the same multiplicity on the curve.

A line and a cubic

The circular points at infinity can not be both singularities of the irreducible cubic component, because an irreducible cubic has at most one singularity. Hence they are singular, because they are intersection points of the cubic and the line at infinity. This line can not touch the cubic there. If it touched the cubic in one of the points, it would touch it in the second point, too. This would make the line a double tangent of the cubic, but cubics do not have double tangents. Hence the line at infinity intersects the cubic in a third point which can not be a singular point of the cubic, because the line and the cubic can only intersect in three points when counted with multiplicity. It is however possible that the cubic has a finite singular point elsewhere.

This case allows three or four singularities. Two of them, namely the circular points at infinity, are always simple intersection points of two curves that are regular there. A possible third singularity comes from the cubic.

Two quadrics

Since the quadrics are irreducible, they do not have singular points themselves. All singularities have to be points of intersection and the circular points at infinity are such points. This means that both quadric components are circles. It is possible that the circular points are simple intersections, in that case there are two more points of intersection which may be different from each other or may coincide. On the other hand the circular points may be double points of intersection, i.e. the quadrics touch each other there and have no other common point.

This case allows two, three or four singularites. Two singularities means that the quadric touch each other in the circular points. Three singularities arise, when the circular points are true intersection points and the quadrics touch each other in a third singularity of the quartic. We count four singularities, when the quadrics intersect in the circular points at infinity and in two other points.

Two lines and a quadric

None of the components can have singularities, thus all singularities of the quartic are intersection points. The two lines do not intersect in one of the circular points, because they can not intersect in the other at the same time. From this immediately follows that the quadric component is a circle. There are two possible configurations of intersection. Either one line intersects the quadric in both circular points and the other does not run through these two points or each line meets the quadric in one of the circular points.

In the first case there are four singularities, if the other line touches the quadric, or five singularities, if the other line meets the quadric in two distinct points. In the second case there are three singular points, if each lines touches the quadric in one of the circular points at infinity. The third singularity is then the intersection of the lines. The quartic has three or five singularities, when each line meets the quadric in two distinct points. In general the number is five, but it drops to three in case that the common point of the two lines lies on the quadric.

Four lines

The four lines can not intersect in one single point, because the circular points must be both singularities. The lines can not intersect in exactly two or exactly five different points.

If they intersect in three different points, then one line must be a multiple component of the quartic. Thus there may be four or six points of intersection.

There are four singularities, if three lines meet in one point and the other line cuts the first three lines in three different points. Two of these other intersections have to be the circular points. Six singularities arise from the generic configuration and two of them are the circular points at infinity.

Irreducible bicircular quartics

In this case the curve has two singularities by definition. A third singularity may exist. It has to be finite, because the line at infinity can not intersect a bicircular quartic in three singular points. The type of the singularities is limited to A_1 or A_2 and the circular points of infinity are always of the same type.

Possible types of singularities

A very limiting aspect with regard to the type of these singularities is the following

Lemma 9.25
If the equation of an algebraic curve C is containing only real coefficients, then the complex conjugate of a singularity of C is also a singularity of C and both are of the same type.

A complete list which and how many singularities a quartic curve may have is given in [8]. Since a bicircular quartic has two singularities of the same type, there are the following possible configurations of singularities for a bicircular quartic.[1] An irreducible bicircular quartic can have this singular points:

- ${A_1}^2$
- ${A_2}^2$
- ${A_1}^3$
- ${A_1}^2A_2$
- $A_1{A_2}^2$
- ${A_2}^3$

For reducible bicircular quartics these configurations are possible:

- line and cubic: ${A_1}^3, {A_1}^4, {A_1}^3A_2$
- two quadrics: ${A_3}^2, {A_1}^2A_3, {A_1}^4$
- two lines and quadric: ${A_1}^2D_4, A_1{A_3}^2, {A_1}^3A_3, {A_1}^5$
- four lines: ${A_1}^3D_4, {A_1}^6$

Only the type ${A_1}^3$ can occur in both cases.

[1] We follow the notation of Bauer, where the expression ${A_1}^2A_2$ means that the curve has two singularities of type A_1 and one singularity of type A_2.

Singularities in the C-Q-form and the three circles form

The C-Q-form of a bicircular quartic contains important information about the type of singularities at the circular points at infinity.

Theorem 9.26 (Singularities in the C-Q-form)
Let B be a bicircular quartic with the C-Q-form

$$\mathcal{C}^2 - {x_0}^2 Q = 0.$$

In the case that B is irreducible the circular points at infinity are singularities of the type A_2, if and only if Q is a circle, else they are of the type A_1.

Proof
We can apply a coordinate transformation in order to center the circle $\mathcal{C}$ at the origin and make the axes of the conic section parallel to the axes of the coordinate system. Because we can vary the radius of $\mathcal{C}$, we may choose it to be a nullcircle. Hence

$$\mathcal{C} : {x_1}^2 + {x_2}^2 = 0$$

and

$$Q : a_{11}{x_1}^2 + a_{22}{x_2}^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + a_{00}{x_0}^2 = 0.$$

In our case the decision between the two possible types of the singularities in the circular points can be easily made by examining the Hessian H of the bicircular quartic, i.e. the matrix containing the second partial derivatives.

By a elementary calculation we obtain

$$H = \begin{pmatrix} B_{00} & -2x_0(3a_{01}x_0 + 2a_{11}x_1) & -2x_0(3a_{02}x_0 + 2a_{22}x_2) \\ -2x_0(3a_{01}x_0 + 2a_{11}x_1) & 12{x_1}^2 + 4{x_2}^2 - 2a_{11}{x_0}^2 & 8x_1x_2 \\ -2x_0(3a_{02}x_0 + 2a_{22}x_2) & 8x_1x_2 & 4{x_1}^2 + 12{x_2}^2 - 2a_{22}{x_0}^2 \end{pmatrix}$$

where $B_{00} = -2a_{11}{x_1}^2 - 2a_{22}{x_2}^2 - 12x_0(a_{00}x_0 + a_{01}x_1 + a_{02}x_2)$. In the circular points this matrix becomes

$$H_{(0:1:\pm i)} = \begin{pmatrix} -2a_{11} + 2a_{22} & 0 & 0 \\ 0 & 8 & \pm 8i \\ 0 & \pm 8i & -8 \end{pmatrix}$$

and we see immediately that the case $a_{11} = a_{22}$ leads to a matrix of rank 1 whereas $a_{11} \neq a_{22}$ leads to a matrix of rank 2. These cases indicate a singularity of type A_2 or A_1 respectively.
□

Theorem 9.27 (Singularities in the three circles form)
Let B be a bicircular quartic and

$$B : \mathcal{C}_1\mathcal{C}_2 - k^2\mathcal{D}^2 = 0$$

its three circles form. If a point P lies on all three circles $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{D}$ then it is a singularity of B.

Proof
The partial derivative of the polynomial defining B with respect to the variable x_i is

$$\frac{\partial B}{\partial x_i} = \frac{\partial \mathcal{C}_1}{\partial x_i} \cdot \mathcal{C}_2 + \mathcal{C}_1 \cdot \frac{\partial \mathcal{C}_2}{\partial x_i} + 2k^2 D \cdot \frac{\partial D}{\partial x_i}$$

for $i = 0, 1, 2$. Because $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{D} = 0$ in a point lying an all three circles, the partial derivatives are also vanishing there.
□

From the previous paragraphs follows immediately that we can use some properties of bicircular quartics for the description of irreducible singular quartics with two or three singularities. The curves are of rational or elliplic type, i.e. they are of genus 0 or 1. The singularities can not lie on a straight line and can be of type A_1 and A_2 only. In particular there are always two singularities of the same kind. The statement *Q_t is a circle* is equivalent to *Q_t is running through the circular points at infinity.* These points play a special role for the classical interpretation of points in the projective space, where $x_0 = 0$ means *infinitely far away* and circles are conic sections with axes of equal length. But this choice is arbitrary in the projective $\mathbb{P}_2(\mathbb{C})$ and there are many possible affine spaces. We can choose two arbitrary points as the circular points at infinity.

Let π be a projective map on $\mathbb{P}(\mathbb{C})$ that maps the singular points of C of the same type onto the circular points at infinity. Then the resulting curve must be a bicircular curve, because it is a quartic curve with the circular points as singularities. All theorems that are not relying on metric arguments thus remain true after the application of the inverse map π^{-1}. An example is the three circles form. Let Q_0, Q_1, Q_2 be quadrics with at least two common points. From the above we can deduce that the envelope of the rational family

$$\{\lambda^2 Q_1 + 2k\lambda\mu Q_2 + \mu^2 Q_0, (\lambda, \mu) \in \mathbb{P}(\mathbb{C})\}$$

is the quartic curve $Q_1 Q_2 - k^2 Q_0$ and that it has singularities at the common points of the quadrics. The inversion about a conic section and its basic properties are discussed in [39].

10 Classical bicircular curves

10.1 Means for classifying bicircular curves

The classification of bicircular quartic can be done in different ways.

Algebraic properties

The most basic division of bicircular quartics into classes can be made with respect to their reducibility. We have seen that bicircular quartics are either irreducible or reducible and that they may decompose into up to four irreducible components. We have already listed the reducible cases without multiple components in the previous section.

- the line at infinity and an irreducible circular cubic
- two irreducible circles
- an irreducible circle and two lines, either one of them being the line at infinity and the other not running through the circular points or the components of a reducible circle
- two reducible circles or one reducible circle, the line at infinity and a line not running through the circular points

These reducible type are exceptional cases.

Theorem 10.1 (General bicircular quartic)
The general bicircular quartic is irreducible.

Proof
The space of bicircular quartics is $\mathbb{P}_8(\mathbb{C})$. Any of the reducible cases has a lower number of parameters[1] so their union can not fill up the entire space of bicircular quartics. □

The irreducible bicircular quartics can be divided into elliptic and rational curves. This depends on the number of their singular points and will be treated later. The general C-Q-form

$$\mathcal{C}_t{}^2 - x_0{}^2 Q_t = 0$$

was used earlier, because the quadric Q_t is reducible for special values of the parameter t. The determinant Δ_t of the quadric is a polynomial of fourth degree in t. We can divide bicicular quartics into classes where Δ_t has

[1] These numbers of parameters are 7, 6, 5, 4, respectively.

- four simple zeros,
- two simple zeros and one double zero,
- two double zeros,
- one single zero and one triple zero,
- one fourfold zero.

We have shown at the end of the previous section that these zeros correspond one-to-one with the inversions that leave the bicircular quartic invariant. Since every zero of Δ_t corresponds one-to-one with a (potential) circle whose inversion leaves the curve invariant, we can state the following

Lemma 10.2
Multiple zeros of Δ_t do not correspond to circles of inversion, but to nullcircles.

Proof
Since two potential circles $\mathcal{O}_1$ and $\mathcal{O}_2$ of inversion are intersecting at right angle, they have to become nullcircles, when $\mathcal{O}_1 = \mathcal{O}_2$. □

Singularities

It is obvious that the singularities of a bicircular quartic can be used for classification. The configuration of singularities of a bicircular quartic to some extent depends also on its decomposition into components. A D_4 singularity arises only in two cases: either when the bicircular quartic splits into the line at infinity, a line not through any of the circular points and a circle which can be irreducible or reducible or when the quartic splits into an irreducible circle and a reducible circle with its center lying on the first one.

When the curve is irreducible we can use the singularities in the circular points for classification. They can only be of the type A_1 or A_2. The number of singularities can not exceed three, hence we can divide irreducible bicircular quartics into two categories. Those with three singularities are always rational curves. This follows immediately from the formula for the genus of the curve. This means that there exists a rational bijection between the curve and the projective line. We have seen that the image of a conic section under inversion is a bicircular quartic (or a circular conic) in almost all cases, but that not all bicircular quartics are such image curves.

Theorem 10.3 (Inverse curve of a conic section)
The inverse curve of a conic section is a bicircular quartic that has a singularity at the center of the circle of inversion.

Proof
It is sufficient to examine the action of the inversion about the unit circle. We have calculated that an image curve has the equation

$$\varepsilon(C) : a_{00}({x_1}^2 + {x_2}^2)^2 + 2(a_{01}x_0x_1 + a_{02}x_0x_2)({x_1}^2 + {x_2}^2) + {x_0}^2(a_{11}{x_1}^2 + 2a_{12}x_1x_2 + a_{22}{x_2}^2) = 0.$$

We see immediately that the partial equations with respect to x_1 and x_2 are both equal to 0 at the point $(1 : 0 : 0)$ which is the center of the circle of inversion. □

Theorem 10.4 (Inverse curve of an irreducible conic section)
The inverse curve of an irreducible quadric with the center of the circle of inversion not lying on the quadric is an irreducible bicircular quartic with three singular points and the center of inversion is one of these singularities.

If the center of the circle of inversion lies on the quadric, then the inverse curve is (the union of the line at infinity and) an irreducible circular cubic. and the center of inversion is the singularity of the cubic.

On the other hand it is quite obvious that

Theorem 10.5
When the center of inversion is located at the finite singularity of a rational bicircular quartic, then its inverse curve is a conic section.

Proof
We assume the finite singularity to be at the origin. Then the polynomial

$$a({x_1}^2+{x_2}^2)^2+x_0(b_1x_1+b_2x_2)({x_1}^2+{x_2}^2)+{x_0}^2(c_{11}{x_1}^2+2c_{12}x_1x_2+c_{22}{x_2}^2+2c_{01}x_0x_1+2c_{02}x_0x_2+c_{00}{x_0}^2)$$

defining the given bicircular curve has a zero at the origin, hence $a_{00} = 0$. Because the origin is a singular point of the curve, the first partial derivations of the defining polynomial also vanish there, yielding $a_{01} = a_{02} = 0$. Recalling theorem (9.10) we see immediately that the inverse curve is the union of a quadric and two times the line at infinity. □

We can distinguish between different types of bicircular quartics by the number of their singularities. This number can only be two or three. At the same time it determines, whether the quartic is an elliptic or a rational curve. Another criterion for classification is the type of the occurring singularities. We saw that instead of calculating the Hessian matrix in the circular points we can simply read off the type of the singularities at the circular points from the C-Q-form of the curve.

Geometry in $\mathfrak{Circ}$ and $\mathbb{P}(\mathfrak{Circ})$

Since bicircular quartics are envelopes, it is possible to classify them according to the corresponding rational family of circles. This rational family can be represented by

- an ellipse,
- a parabola or
- a hyperbola

in the space of circles.

The position of this conic section with respect to the paraboloid Π of nullcircles has an influence on the shape of the envelope. In particular one can examine the real part and divide the classes into subclasses depending on the number of real intersections of the conic section and Π. This has been done in [52] for rational families of circles in $\mathfrak{Circ}$ that are symmetrical to one of the coordinate axes.

10.2 List of classical bicircular curves

Examples of bicircular quartics were already known to ancient geometers. In this section we will describe in which way the special types of classical curves are derived from bicircular curves. As a starting point we want to recall once more the C-Q-form

$$B : ({x_1}^2 + {x_2}^2)^2 + {x_0}^2(c_{20}{x_1}^2 + c_{02}{x_2}^2 + c_{10}x_0x_1 + x_{01}x_0x_2 + c_{00}{x_0}^2) = 0.$$

Most sources related to these classical curves give equations in their affine form. Therefore we want to write the (normalized) equation of a general bicircular quartic as

$$B : (x^2 + y^2)^2 + ax^2 + by^2 + cx + dy + e = 0 \tag{10.1}$$

throughout this section.

Toric sections

A conic section is the figure that arises from the intersection of a cone and a plane. A toric section is essentially the same, but with a torus as the three-dimensional body to be intersected.

Definition 10.6 (Torus)
A torus *is an algebraic surface of degree four with an equation of the form*

$$T : (x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4R^2(x^2 + y^2) = 0.$$

The origin is the center *of T, the xy-plane is called its* equatorial plane *and the z-axis the axis of the torus.*

Definition 10.7 (Toric section)
Let T be a torus and P a plane in $\mathbb{P}_2(\mathbb{C})$ and C the intersection of T and P. The plane curve C is a called a toric section.

Special toric sections, the so called spiric sections, were investigated by Greek mathematicians in the second century BC. We will come back to them later. In this text the term toric section stands for the most general class of this curves.

A torus has a rotational symmetry with respect to the line running through its center and being perpendicular to its equatorial plane. It is symmetric with respect to its center, its equatorial plane and all planes that are perpendicular to the equatorial plane and are running through the center of the torus. This implies

Theorem 10.8 (Symmetry of toric sections)
Toric sections always have at least one axis of symmetry.

Proof
Let E be the equatorial plane of the torus T and P the plane intersecting T. When E and P are not parallel, then there is exactly one plane Σ perpendicular to these planes and running through the symmetry center of the torus. T and P are both symmetrical with respect to Σ, hence their intersection is symmetrical with respect to the line σ which is the intersection of P and Σ. □

Theorem 10.9 (Equation of a general toric section)
The equation of an arbitrary toric section T can be brought into the form

$$\begin{array}{lll} B: & (x^2+y^2)^2 & \\ & +2\left(\zeta^2-R^2-r^2\right)x^2+2\left(\zeta^2+(1-2\cos^2(\alpha))R^2-r^2\right)y^2 & \\ & -8R^2\cos(\alpha)\sin(\alpha)\zeta y+(\zeta^2+R^2-r^2)^2-4R^2\sin^2(\alpha)\zeta^2 & = \; 0 \end{array} \tag{10.2}$$

by an appropriate transformation of coordinates.

Proof
The shape of a toric section does not change if we rotate the plane P around the axis of the torus T. Thus we can assume from now on that P is perpendicular to the yz-plane.

We can also always perform a rotation about the x-axis, such that the plane P is mapped onto a plane parallel to the xy-plane. This means that we obtain an equation for the toric section by intersecting the image of T under this rotation with the image of P.

The rotation can be written as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where α is the angle between P and the equatorial plane of T.

Because of $(y')^2+(z')^2=y^2+z^2$ we have

$$T' : (x^2+y^2+z^2+R^2-r^2)^2-4R^2(x^2+(\cos(\alpha)y+\sin(\alpha)z)^2)=0.$$

With ζ being the distance between the plane P and the center of T we have to intersect T' with the plane

$$P' : z=\zeta.$$

This yields

$$\begin{array}{lll} B: & (x^2+y^2)^2+2(x^2+y^2)(\zeta^2+R^2-r^2)+(\zeta^2+R^2-r^2)^2 & \\ & -4R^2\left(x^2+\cos^2(\alpha)y^2+2\cos(\alpha)y\cdot\sin(\alpha)\zeta+\sin^2(\alpha)\zeta^2\right) & = \; 0 \end{array}$$

which is equivalent to

$$\begin{array}{lll} B: & (x^2+y^2)^2 & \\ & +2\left(\zeta^2-R^2-r^2\right)x^2+2\left(\zeta^2+(1-2\cos^2(\alpha))R^2-r^2\right)y^2 & \\ & -8R^2\cos(\alpha)\sin(\alpha)\zeta y+(\zeta^2+R^2-r^2)^2-4R^2\sin^2(\alpha)\zeta^2 & = \; 0. \end{array} \quad \square$$

We see immediately from equation (10.2) that it describes a bicircular quartic and that it is invariant under the substitution $x \mapsto -x$, hence

Corollary 10.10 (Toric sections are bicircular quartics)
Every toric section is a bicircular quartic and has an axis of symmetry.

The equation (10.1) of a given bicircular quartic has a special form, if this curve is a toric section. We can choose the coordinate system such that the line of symmetry is the x-axis. Hence the general equation of a toric section is

$$B : (x^2+y^2)^2+ax^2+by^2+dy+e=0.$$

Central toric sections and Villarceau circles

Toric sections can have a second axis of symmetry. From the last proof we see that this happens, when

$$\cos(\alpha)\sin(\alpha)\zeta = 0.$$

For now we want to discuss the case $\zeta = 0$. In this case the plane P runs through the center of the torus T. The other possibility where α is the multiple of a right angle will be discussed later.

Definition 10.11 (Central toric section)
A toric section where the center of the torus T lies on the cutting plane P is called a central toric section.

If the center of T lies on P, then T as well as P are symmetrical to this point. But then the toric section defined by them is also symmetrical with respect to the center of the torus as a point on P. Since the curve has always an axis of symmetry, this center of symmetry has to lie on the axis and thus exists also a second axis of symmetry perpendicular to the first.

Theorem 10.12 (Symmetry of central toric sections)
Central toric sections have a center of symmetry and two mutually perpendicular axes of symmetry.

From the above immediately follows that central toric sections can be written as

$$B : (x^2 + y^2)^2 + ax^2 + by^2 + e = 0.$$

We recall equation (10.2)

$$\begin{aligned} B:\ & (x^2+y^2)^2 \\ & 2\left(\zeta^2 - R^2 - r^2\right)x^2 + 2\left(\zeta^2 + (1-2\cos^2(\alpha))R^2 - r^2\right)y^2 \\ & -8R^2\cos(\alpha)\sin(\alpha)\zeta y + (\zeta^2 + R^2 - r^2)^2 - 4R^2\sin^2(\alpha)\zeta^2 \quad = \quad 0 \end{aligned}$$

of a general toric section. For $\zeta = 0$ we obtain

$$T : (x^2 + y^2)^2 - 2\left(R^2 + r^2\right)x^2 + 2\left((1 - 2\cos^2(\alpha))R^2 - r^2\right)y^2 + (R^2 - r^2)^2 = 0,$$

which is the equation of a central toric section.

Theorem 10.13 (Splitting central toric sections)
A central toric section splits into two circles in three cases. This happens, when the intersecting plane is

- *the equatorial plane of the torus,*
- *perpendicular to the equatorial plane of the torus,*
- *touching the torus in two isolated points.*

Proof
Because of the twofold axial symmetry of central toric sections we may assume that the centers of its components $\mathcal{C}_1$and $\mathcal{C}_2$ lie on the x-axis. Now we have

$$B = \mathcal{C}_1 \cdot \mathcal{C}_2$$

with

$$\mathcal{C}_i = (x^2 + y^2) - 2\eta_i y + \zeta_i.$$

The comparison of the coefficients yields five equations:

$$\begin{array}{lrrcl} (1) & x(x^2+y^2): & \eta_1 + \eta_2 & = & 0, \\ (2) & x^2: & \zeta_1 + \zeta_2 + 4\eta_1\eta_2 & = & -2R^2 - 2r^2, \\ (3) & y^2: & \zeta_1 + \zeta_2 & = & 2R^2 - 4\cos^2(\alpha)R^2 - 2r^2, \\ (4) & x: & \eta_1\zeta_2 + \eta_2\zeta_1 & = & 0, \\ (5) & 1: & \zeta_1\zeta_2 & = & (R^2 - r^2)^2. \end{array}$$

Equation (1) reflects the axial symmetry $\eta := \eta_1 = -\eta_2$. This turns equation (4) into

$$\eta(\zeta_1 - \zeta_2) = 0.$$

This is solved, if $\eta = 0$ or $\zeta := \zeta_1 = \zeta_2$.

For $\eta = 0$ we obtain $\cos(\alpha) = \pm 1$ from equation (2) and (3). This means that P is the equatorial plane and the toric section splits into two concentric circles. For $\eta \neq 0$ we deduce from equation (5) that

$$\zeta = \pm(R^2 - r^2).$$

In equation (2) the choice $\zeta = R^2 - r^2$ implies $\eta = \pm R$ where a change of sign only exchanges the circles $\mathcal{C}_1$ and $\mathcal{C}_2$. In this case we see from equation (3) that $\cos(\alpha) = 0$. The plane P is perpendicular to the equatorial plane of T. The choice $\zeta = r^2 - R^2$ leads to $\eta = \pm r$ in equation (2). This together with equation (3) gives

$$\cos^2(\alpha) = 1 - \frac{r^2}{R^2}.$$

This is equivalent to $\sin(\alpha) = \pm\frac{r}{R}$. In this case the plane P touches the torus T in two isolated points. □

Definition 10.14 (Villarceau circles)
Let T be a torus and P be a plane touching T in two points. Then the toric section generated by P intersecting T splits into two circles that are called Villarceau circles.

These circles were described by Yvon Villarceau in [55]. It might be surprising to find these skew circles on every torus at first sight. But their existence follows from

Theorem 10.15
Let S be an algebraic surface and P be a plane touching S in a regular point X. Then the curve generated by the intersection of P and S has a singularity in X.

So if the plane P touches the torus T in two distinct points, then the toric section has (at least) four singularities altogether. This implies that this bicircular quartic is reducible. It can not have a line as a component, because the torus does not contain lines, hence it must split into a pair of circles. From the above follows

Corollary 10.16
Through each (regular) point on a torus run four circles. One is parallel and one perpendicular to the equatorial plane of the torus, the other two are Villarceau circles.

A treatise with a more general setting is [40], it also gives some additional references on this topic.

Spiric sections, Cassinian curves, hippopedes

The earliest occurrence of toric sections as such is unknown. It is assumed that they were studied after the description of conic sections as the intersection of a plane and a cone. Secondary sources name Perseus as the discoverer of the so called spiric lines. Virtually nothing is known about him, he probably lived in the second century BC.

Definition 10.17 (Spiric section)
A spiric section *is a toric section where the intersecting plane is perpendicular to the equatorial plane of the torus.*

The so called spiric sections are therefore special toric sections. The earliest source about this type of curves is attributed to Eudoxus of Cnidus (≈ 410 BC – 350 BC). His work about planetary movement contains the description of a hippopede, a special type of a spiric section. Unfortunately all works of Eudoxus are lost. It is worth to mention that the hippopede he constructed is not a plane curve but a curve on a sphere. Corresponding plane curves can be derived from the spherical one by stereographic projection or by inversion about an appropriate sphere.

The equation of a general spiric section can be easily derived from equation (10.2). Because α is a right angle we obtain

$$B : (x^2+y^2)^2 + 2\left(\zeta^2 - R^2 - r^2\right)x^2 + 2\left(\zeta^2 + R^2 - r^2\right)y^2 + (\zeta^2 + R^2 - r^2)^2 - 4R^2\zeta^2 = 0.$$

This is an equation of the form

$$B : (x^2+y^2)^2 + ax^2 + by^2 + e = 0,$$

thus spiric sections have two axes of symmetry, just like central toric sections. Nevertheless these kinds of toric sections are different as real curves.

Theorem 10.18 (Central toric sections and spiric sections)
A given central toric section with $\alpha \neq \frac{\pi}{2}$ can not be constructed as a spiric section and a spiric section with $\zeta \neq 0$ can not be constructed as a central toric section.

Proof
It suffices to compare the coefficient of the monomial y^2. If there was a way to switch between the two kinds of toric sections, we would be able to find a pairs (α, ζ) such that

$$(1 - 2\cos^2(\alpha))R^2 - r^2 = \zeta^2 + R^2 - r^2.$$

This simplifies to

$$-2R^2\cos^2(\alpha) = \zeta^2$$

having a real solution only for the trivial case, where $\alpha = \frac{\pi}{2}$ and $\zeta = 0$. □

There a two classes of spiric sections that have especially attracted the attention of the mathematical community. The first one are the so called Cassini ovals.

Definition 10.19 (Cassinian curve)
A Cassinian curve *is a spiric section where $\zeta = r$.*

This definition immediately implies

Corollary 10.20 (Equation of a Cassinian curve)
The equation of a Cassinian curve can be brought into the form

$$B : (x^2 + y^2)^2 - 2R^2x^2 + 2R^2y^2 + (R^2 - r^2)^2 - 4R^2r^2 = 0$$

by an appropriate transformation of coordinates.

Cassinian curves have a well known property that follows from the form of their equation.

Theorem 10.21 (Focal property of Cassinian curves)
Two points F_1 and F_2 called foci *and a constant b exist for a given Cassinian curve B with*

$$d(P, F_1) \cdot d(P, F_2) = b^2$$

for all points P on the curve B.

Proof
We assume the foci to have the coordinates

$$F_1 = (-a, 0), \quad F_2 = (a, 0).$$

We square the equation describing the focal property and obtain that the point $P = (x, y)$ belongs to the geometrical locus defined by this property if and only if the pair (x, y) suffices

$$\left((x + a)^2 + y^2\right)\left((x - a)^2 + y^2\right) = b^4.$$

Multiplication on the left hand side and reordering leads to

$$(x^2 + y^2)^2 - 2a^2x^2 + 2a^2y^2 + a^4 = b^4$$

which is the equation of a Cassinian curve. □

The other important class of spiric sections are the so called hippopedes.

Definition 10.22 (Hippopede)
A hippopede *is a spiric section where $\zeta = R - r$.*

When a spiric section is a hippopede, the intersecting plane P touches the torus T. This implies that a hippopede has always three singularities, hence

Corollary 10.23
Hippopedes are rational curves.

The general equation of a hippopede can be brought into the form

$$(x^2 + y^2)^2 - 4Rrx^2 + 2R(R - r)y^2 = 0.$$

Curves described by this equation always have a singular point at the origin. For $r = R$ they split into the union of two circles touching each other at the origin and having centers at $(\pm R, 0)$.

Theorem 10.24
The inverse of an irreducible central conic section about a circle centered at the center of the conic is a hippopede.

Proof
We can apply a transformation of coordinates to map every irreducible central conic section onto a quadric with an equation of the form

$$ax^2 + by^2 + 1 = 0$$

with and the circle of inversion onto the unit circle. If $a = b$ the given conic is a circle with center at the origin, hence the its image under inversion is also circle with its center at the origin. If $a \neq 0$, the inverse curve has the equation

$$(x^2 + y^2)^2 + ax^2 + by^2 = 0.$$

This is the equation of a hippopede with

$$4R^2 = b - a \quad \text{and} \quad 4r^2 = \frac{a^2}{b - a}. \qquad \square$$

One prominent curve is connecting these two classes of spiric sections. The *Lemniscate of Bernoulli* is a Cassini oval as well as a hippopede. It is the inverse curve of the rectangular hyperbola.

Cartesian ovals – Limaçons

An important kind of bicircular curves are those with singularities of type A_2 at the circular points. We have shown that these curves have a C-Q-form where the quadric Q is also a circle. One well known example is the cardioid. Its equation is usually given as

$$(x^2 + y^2 + 2ax)^2 - 4a^2(x^2 + y^2) = 0.$$

A translation turns this into the C-Q-form

$$(x^2 + y^2)^2 - 6a^2x^2 - 6a^2y^2 + 8a^3x + 5a^4 = 0.$$

Definition 10.25 (Cartesian ovals)
A Cartesian oval *or* oval of Descartes *is a bicircular quartic where the circular points at infinity are singularities of type A_2.*

We have shown in theorem (9.26) that if the curve B is a Cartesian oval, then equation (10.1) turns into

$$B : (x^2 + y^2)^2 + b(x^2 + y^2) + cx + dy + e = 0.$$

The cardioid belongs to a subset of these curves called *Limaçons*. They are a two parameter family given by the equation

$$(x^2 + y^2 - ax)^2 = b^2(x^2 + y^2).$$

Two special cases are:

- $a = b$: cardioid,
- $a = 2b$: trisectrix limaçon.

Circular cubics – Conchoids of de Sluze

Bicircular quartics may degenerate into circular cubics. The most famous circular cubics are contained in the family of curves that is called the *conchoids of de Sluze*. These curves are given by the equation

$$(x - 1)(x^2 + y^2) = ax^2$$

with parameter a. They always have a singularity at the origin making them rational cubic curves. For $a = 0$ the cubic splits into a nullcircle and the line $x = 1$ which is the asymptote of the curves for other values of a. For special values of a we obtain, for example, these important curves:

- $a = -1$: cissoid of Diocles,
- $a = -2$: right strophoid (also called Newton's strophoid),
- $a = -4$: trisectrix of Maclaurin.

11 Applications of bicircular curves

11.1 Electrical networks

At the beginning of the 20^{th} century bicircular quartics appeared in text books on electrical engineering. The impedance of a circuit fed by a sinusoidal voltage generally depends on the frequency of this voltage. The impedance can be expressed as a complex number. In the simplest cases the possible values of the impedance lie on a circle. More complicated circuits lead to more complicated curves, e.g. to circular cubics and bicircular quartics. The reader may refer to [11], for instance.

Calculation of impedance

In electrical circuits with a constant current I the voltage U across a resistor of resistance R is given by the formula

$$U = R \cdot I.$$

With $G = \frac{1}{R}$ being the conductance of the resistor we can also write

$$I = G \cdot U.$$

This formula remains true for resistors in circuits with variable current, but voltage and current are no longer proportional to each other at capacitors and inductors. The usage of variable current makes it necessary to take into account effects that arise from the fluctuation of the current. The voltage U across a capacitor and the current I flowing through it are related by the differential equation

$$I = C \cdot \frac{dU}{dt}$$

with C being the capacitance of the capacitor. For

$$U(t) = U_0 \cdot e^{i\omega t}$$

we have

$$I(t) = i\omega C \cdot U(t).$$

The current I through a inductor and the voltage across it are related by the differential equation

$$U = L \cdot \frac{dI}{dt}$$

with L being the inductance of the inductor. For

$$I(t) = I_0 \cdot e^{i\omega t}$$

we have

$$U(t) = i\omega L \cdot I(t).$$

Therefore we can assign a physical quantity

$$Z = R + iX$$

called impedance to capacitors and resistors. For resistors the imaginary part X, called reactance, is zero and Z coincides with the resistance R. The reciprocal quantity of the impedance is called admittance Y. It can be written as

$$Y = G + iE$$

where G is the conductance from above and E the so called susceptance.

For sinusoidal voltage $U(t) = U_0 \cdot e^{i\omega t}$ and current $I(t) = I_0 \cdot e^{i\omega t}$ the admittance of a capacitor with capacitance C is

$$Y_C = i\omega C$$

and the impedance Z_L of an inductor with inductance L is

$$Z_L = i\omega L.$$

The behavior of impedance and admittance with respect to series and parallel circuits is the same as that of resistance and conductance. This means, for example, that a series of the impedances Z_1 and Z_2 has the total impedance $Z = Z_1 + Z_2$.

RC- and RL-circuit

One simple question regarding electric circuits containing resistors, capacitors and inductors is the following. Given that the voltage

$$U_0(t) = U_0 \cdot e^{i\omega t}$$

lies across the series of a resistor with resistance R and a capacitor with capacitance C, which voltage $U_C(t)$ lies across the capacitor?

The answer can be given by calculating the total impedance of the series

$$Z = R + \frac{1}{i\omega C} = \frac{i\omega RC + 1}{i\omega C},$$

then the current

$$I(t) = \frac{U(t)}{Z} = \frac{U(t) \cdot i\omega C}{i\omega RC + 1}$$

and finally the voltage

$$U_C(t) = Z_C I(t) = \frac{U(t)}{i\omega RC + 1}$$

across the capacitor from this current.

The total impedance Z plays a crucial role in this calculation. When we want to examine the behaviour of Z for varying ω, we may scale the physical quantities in such a way that
$$Z = 1 + \frac{c}{i\omega}$$
for some real constant c. We can interpet this as the image of the real line in the Riemann sphere under the Möbius transformation
$$z \mapsto w = 1 + \frac{c}{iz}.$$
This is a straight line through $w = 1$ parallel to the imaginary axis.

If we substitute the capacitor with an inductor of the inductance L and pose the same question as before, we obtain
$$Z = R + i\omega L,$$
then
$$I_C(t) = \frac{U(t)}{Z} = \frac{U(t)}{R + i\omega L}$$
and finally
$$U_L(t) = Z_L I(t) = \frac{i\omega L U(t)}{R + i\omega L}.$$
Again we can rescale our axes and obtain the Möbius transformation
$$z \mapsto w = \frac{iz}{c + iz}.$$
$z = 0$ has the image $w = 0$, $z = c$ the image $w = \frac{i}{1+i} = \frac{i+1}{2}$ and $z = \infty$ maps to $w = 1$. Thus the image is a circle with center $w = \frac{1}{2}$.

RC- and RL-cascade

Electrical circuits can have the form of a so called cascade. In such a circuit a certain arrangement of elements is repeated. One important example are voltage multipliers, which are built as a cascade of rectifiers (usually diodes) and capacitors. Our interest is focused on those circuits, where the RC- and RL-circuit are repeated. We want to call them the RC- and the RL-cascade respectively.

The total impedance of an RC-cascade is
$$Z = R_1 + \frac{iX_1(R_2 + iX_2)}{iX_1 + R_2 + iX_2}$$
where $X_1 = -\frac{1}{\omega C_1}$ and $X_2 = -\frac{1}{\omega C_2}$. This can be written as
$$Z = R_1 + \frac{i\omega R_2 C_2 + 1}{-\omega^2 R_2 C_1 C_2 + i\omega(C_1 + C_2)}.$$

We ignore for now the offset R_1 and concentrate on the remaining part of this expression. Like before we can rescale the physical quantities to bring it into the form

$$w = \frac{iaz + 1}{-bz^2 + icz}.$$

It is shown in [11] that $z \mapsto w$ maps the real axis onto a rational bicircular quartic.

For the RL-cascade the formula for the total impedance is obiously also

$$Z = R_1 + \frac{iX_1(R_2 + iX_2)}{iX_1 + R_2 + iX_2},$$

this time with $X_1 = \omega L_1$ and $X_2 = \omega L_2$. After rescaling the term can be written as

$$w = \frac{-az^2 + ibz}{icz + 1}$$

which is in some sense the reciprocal of the expression derived for the RC-cascade. Again, the image of the real line under the map $z \mapsto w$ is a rational bicircular quartic.

In [11] this topic is discussed more thoroughly. Cases of degeneration are covered, such as when the bicircular quartic turns into a circular cubic or a conic. Because the curves are always rational, they have a finite singularity. The dependency of the type of this singularity is also covered in the mentioned work.

11.2 Generalization in different dimensions

Bicircular quartics as a general concept

The equation of a general bicircular quartic can be easily adapted to higher dimensions. It suffices to replace the term

$$({x_1}^2 + {x_2}^2)$$

by

$$\sum_{i=1}^{d} {x_i}^2.$$

In three dimensions this leads to the equation

$$({x_1}^2 + {x_2}^2 + {x_3}^2)^2 + x_0(a_1x_1 + a_2x_2 + a_3x_3)({x_1}^2 + {x_2}^2 + {x_3}^2) + {x_0}^2Q = 0$$

where Q is a homogeneous polynomial of degree 2. With the abbreviation

$$\varrho^2 = \sum_{i=1}^{n} {x_i}^2$$

we can write the equation of a bicircular quartic as

$$p_0(\varrho^2)^2 + x_0p_1\varrho^2 + {x_0}^2p_2$$

in analogy to equation (8.5), where p_i is a homogeneous polynomial of degree i. This equation now makes sense in all positive dimensions and we may call the objects it describes a *bispherical quartic hypersurfaces*. We have already seen an equation of this form in three dimensions. The torus was defined by

$$(x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4R^2(x^2 + y^2) = 0.$$

The terms can be easily rearranged into

$$(\varrho^2)^2 + 2(R^2 - r^2)\varrho^2 - 4R^2(x^2 + y^2) + (R^2 - r^2)^2 = 0,$$

hence the torus is an example of a bispherical surface of degree 4. It is obvious that this extension works in arbitrary dimensions, because quadrics exist in arbitrary dimensions.

A possibility to extend the equation of a bispherical hypersurface is to vary its degree. Equation (8.5) then becomes

$$\begin{array}{l} p_0(\varrho^2)^n + x_0 p_1 (\varrho^2)^{n-1} + \dots \\ \dots + {x_0}^{n-2} p_{n-2} (\varrho^2)^2 + {x_0}^{n-1} p_{n-1} \varrho^2 + {x_0}^n p_n \quad = \quad 0, \end{array} .$$

It describes an n-spherical hypersurfaces of degree $2n$. The image of a n-spherical hypersurface under d-dimensional inversion is a surface of the same type. Each of these hypersurfaces is in general invariant under $d+2$ inversions. The inversion about a hypersphere can be defined in any dimension in analogy to definition (4.12). It is clear what the translation τ and the dilation δ^2 are. The inversion about the unit hypersphere is

$$(x_0 : x_1 : \dots : x_d) \mapsto (\sum_{i=1}^{d} {x_i}^2 : x_0 x_1 : \dots : x_0 x_d).$$

Cyclides – bispherical surfaces

The naming bispherical quartic surface is derived from the expression bicircular quartic curve. This kind of surfaces is better known as *cyclides*. The intersection of the plane at infinity and a cyclide is the absolute conic and all points of this intersection are double points. The union of two spheres is a special case of a cyclide, hence the name bispherical.

This kind of surface was first examined by the French mathematician Charles Dupin. The so called *Dupin cyclides* are envelopes of a one-parameter family of spheres whose centers lie on a given conic section in space and are orthogonal to a given fixed sphere. A general cyclide is defined by equation (8.5). Like Dupin cyclides it can be regarded as the envelope of a family of spheres that are orthogonal to a given sphere. The main difference is that these spheres have their centers on a quadric surface, whereas the quadric is degenerated to a conic section in the case of Dupin cyclides. We want to call the quadric on which the centers of the spheres lie the *deferent quadric* of the cyclide.

The prototype of a Dupin cyclide is the torus. It is the envelope of a family of spheres that have their centers on a circle and are orthogonal to the plane containing this circle. This plane is the equatorial plane of the torus. Other Dupin cyclides can be derived by inversion of a torus about a sphere with its center lying on the equatorial plane of the torus. In fact all Dupin cyclides are images of a torus under an inversion of this type. The *deferent quadric* of the original torus is a circle, but this is no longer true after the inversion. The same phenomenon is already observable in two dimensions.

From the construction follows that Dupin cyclides have a plane of symmetry. It is the equatorial plane of the original torus and it makes sense to call it the equatorial plane of the Dupin cyclide. This plane contains the deferent quadric of the cyclide. Planar sections of cyclides are bicircular quartics. It is also possible to intersect a sphere and a cyclide. The resulting curves are spherical bicircular quartics like the Hippopede of Eudoxus. These curves are trated in [18] and are named sphero-quartics.

The image of a cyclide under inversion is again a cyclide. If the center of inversion lies on the surface, then the image under inversion is interpretable as a cyclide that contains the plane at infinity as a component. The remaining part is a cubic surface that contains the absolute quadric. We could call this a spherical quadric. All proper cyclides can be expressed by a three dimensional analogon of the C-Q-form. This equation allows similar deductions like for bicircular quartics. Again we can introduce a parameter t for the variation of the radius of the sphere C. The quadric Q is varied simultaneously and its determinant is a polynomial of fifth degree in t. Hence this quadric degenerates to a cone for five values of t when counted with multiplicity. For bicircular quartic this led to four inversions about circles that left the curve invariant. For cyclides this means that they are in general invariant under five inversions about a sphere. The same argument leads to six inversions in four dimensions and so on, that is $(d+2)$ inversions of invariance for d-dimensional bispherical quartics.

Polynomials of degree four

The last statement of the previous paragraph holds also in $\mathbb{P}(\mathbb{C})$. This means that a one-dimensional bispherical quartic is invariant under three inversions about a one-dimensional sphere. The interpretation of these inversions as permutations of the zeros of a polynomial of degree four leads to a geometric solution of such polynomials, and in particular to a different one than given in [31].

We have to clarify what bispherical quartics look like in one-dimensional space. We write equation (8.5) explicitly for the dimension $d = 1$ as

$$c_4 {x_1}^4 + c_3 x_0 {x_1}^3 + c_2 {x_0}^2 {x_1}^2 + c_1 {x_0}^3 x_1 + c_0 {x_0}^4 = 0.$$

The left hand side of this equation is a homogeneous polynomial of degree four. A one-dimensional bispherical quartic is therefore nothing else than the set of zeros of this quartic polynomial when taken with multiplicity. The analogon of a circular cubic is the set of zeros of a cubic polynomial.

In the same way we want to give sense to the term one-dimensional sphere. The equation

$$\nu({x_1}^2 + {x_2}^2) - 2\xi x_0 x_1 - 2\eta x_0 x_2 + \zeta {x_0}^2 = 0$$

of a two-dimensional sphere, a circle, can be reduced by one dimension by setting $x_2 = 0$. We obtain

$$a_2 {x_1}^2 + a_1 x_0 x_1 + a_0 {x_0}^2 = 0$$

as the equation of a general one-dimensional sphere. This means that a sphere in $\mathbb{P}(\mathbb{C})$ is a pair of points or a double point. It is obvious that all quadrics in $\mathbb{P}(\mathbb{C})$ are spheres.

We can bring the equation of a bispherical quartic into C-Q-form. The result is

$$x_1{}^4 + x_0{}^2(b_2 x_1{}^2 + 2b_1 x_0 x_1 + b_0 x_0{}^2) = 0.$$

Again we can introduce a parameter t:

$$(x_1{}^2 + t)^2 + x_0{}^2((b_2 - 2t)x_1{}^2 + 2b_1 x_0 x_1 + (b_0 - t^2)x_0{}^2) = 0.$$

The quadric Q_t has the determinant

$$\Delta_t = \begin{vmatrix} (b_2 - 2t) & b_1 \\ b_1 & (b_0 - t^2) \end{vmatrix} = 2t^3 - b_2 t^2 - 2b_0 t + b_0 b_2 - b_1{}^2.$$

This polynomial of third degree has three zeros. For these values of t the quadric Q_t degenerates into a double point. Each of the degenerate quadrics gives rise to a hypersphere of inversion in $\mathbb{P}(\mathbb{C})$. This results in three inversions that map the given bispherical quartic onto itself. For later reference we explicitly state

Corollary 11.1 (Inversion in $\mathbb{P}(\mathbb{C})$)
An inversion in the projective line is the composition of a translation, a dilation and the unit inversion

$$(x_0 : x_1) \mapsto (x_1{}^2 : x_0 x_1) = (x_1 : x_0).$$

The geometry of bicircular quartics applied to $\mathbb{P}(\mathbb{C})$ now gives us a method how to solve the general quartic equation. The usual way to achieve this is to find appropriate field extensions for the subnormal series

$$S_4 \rhd V_4 \rhd S_2 \rhd 1$$

of the Galois group S_4, the permutation group of the four zeros. V_4 stands for the Klein four-group $S_2 \times S_2$ and S_2 for the permutation group of two elements. The corresponding quotients in the series are S_3, S_2 and S_2.

The first step in this series is the big one. We need to find the zeros of an appropriate polynomial of third degree. In our case this polynomial is the determinant Δ_t. Its zeros can be used to transform the quartic equation into a system of two quadratic equations. Since the roots of the quartic equation are given by a bispherical quartic, the three inversions leaving this hypersurface invariant are the equivalents of the Galois automorphisms mapping pairs of zeros onto each other. Their action on the zeros is the action of the Klein four-group V_4. The quotient S_3 can be understood as the six possibilities to exchange the three inversions in any way.

We have to determine the corresponding hypersphere of inversion of every root t_i of the determinant Δ_t. Its center $C_i = (1 : c_i)$ lies at the double zero of the quadric equation

$$Q_{t_i} : (b_2 - 2t_i)x_0{}^2 + 2b_1 x_0 x_1 + (b_0 - t_i{}^2)x_0{}^2 = 0.$$

This point is easily calculated as

$$c_i = \frac{-b_1}{b_2 - 2t_i}.$$

The hypersphere of inversion is orthogonal to the hypersphere $\mathcal{C}_t$. The radius of the latter squared is $-t_i$, hence the square of the radius r_i of the hypersphere of inversion is

$$r_i^{\,2} = t_i^{\,2} + c_i^{\,2} = t_i^{\,2} + \frac{b_1^{\,2}}{(b_2 - 2t_i)^2}.$$

After applying the translation

$$x \mapsto x - c_i$$

the quartic equation is no longer depressed, but it is turned into a quasi-symmetric function. This is a quartic equation of the form

$$x^4 + d_1x^3 + d_2x^2 + md_1x + m^2 = 0.$$

Equations of this type can be solved by solving two quadratic equations. One can show that

$$m = r_i^{\,2}.$$

A quasi-symmetric equation is solved by dividing the equation by x^2 and setting

$$z := x + \frac{m}{x}. \tag{11.1}$$

This results in the quadratic equation

$$z^2 + d_1z + (d_2 - 2m) = 0$$

and the resubstitution of its solutions z_j into

$$x^2 - xz_j + m = 0.$$

The pair of equations reflects the two steps in the subnormal series of the Galois group from the Klein four-group to the trivial group.

The cubic polynomial Δ_t that we derived from properties of bispherical hypersurfaces in $\mathbb{P}(\mathbb{C})$ can be called the *resolvent cubic* of our problem. We remember that the depressed quartic equation

$$x^4 + b_2x^2 + 2b_1x + b_0$$

has the resolvent cubic

$$\Delta_t = 2t^3 - b_2t^2 - 2b_0t + (b_0b_2 - b_1^{\,2}).$$

One remarkable aspect of this solution of the quartic equation is the appearance of the transformation given in equation (11.1). It is called the Zhukovsky transformation, named after the Russian mathematician Nikolai Yegorovich Zhukovsky. It played an important role in understanding the principle of lift in aerodynamics. The transformation depends on two parameters and maps a circle onto the cross-section of an airfoil. Because it is a conformal map, it can be used to describe the airflow around an airfoil by describing the airflow around the corresponding cylinder and transforming the results into the airflow around the airfoil.

12 Summary

In the preceding chapters we have discussed linear and rational families of circles. We started from the very beginning with the Euclidean definition of a circle and refined it step by step. Finally we reached the projective space $\mathbb{P}(\mathfrak{Circ})$ of circles where we could give a unified description of circles and straight lines. After that we discussed the inversion about these generalized circles and its action on the space of circles.

The most important result at that point was that in the projective space of circles inversions can be represented as linear maps. This fact made it interesting to investigate linear families of circles. This is a classical subject and we treated only some of its aspects. We changed our focus to families of circles that are represented by a conic section in $\mathbb{P}(\mathfrak{Circ})$ and called them rational families of circles. We used some of their properties to learn something about bicircular quartic curves.

Why are we interested in bicircular quartics? First of all, they arise in a natural way in many different situations. The cardioid is a striking example for this. The same holds for the lemniscate of Bernoulli. Special bicircular quartics and circular cubics have been studied since ancient times, for example to solve the Delian problem and the problem of angle trisection. Another important motivation is their universality. The equation defining a bicicular quartic has a shape that can be easy generalized to fit any degree of circularity and any dimension.

Bicircular quartics have many interesting geometrical properties. We have discussed only a few of them. Many of these properties are found in the treatise of John Casey on this subject, others are covered in the works listed in the bibliography. There are many directions in which we could continue. We may for example ask what conditions can be changed without rendering a certain theorem invalid. We already saw that it is possible to replace circles by general quadrics with two fixed points. We could examine the metric invariants of bicircular quartics, a task that has been performed in [42]. Other fields of interest are the algebraic properties of one-dimensional bispherical quartics and bicircular quartic curves over finite fields.

We have seen that bicircular quartics and also their cousins in a different degree or a different dimension are fruitful subjects of study. This is true for purely geometric questions, but also for technical applications. This text was written to narrow the gap between the knowledge available at the end of the 19^{th} century and the mathematics of today. In some aspects the mathematical notation has changed considerably and the proofs in those days are very short and make certain assumptions on the knowlegde of the reader, which makes them hard to read now. This work should make the reception of older sources easier. Even though projective geometry was already well established around 1870, when bicircular quartics were intensively studied, the original texts are written mainly in the language of affine geometry. This text provides the reader with a description that was rewritten in the language of projective geometry by homogenizing as many parts of the involved equations as possible.

Many topics have to remain untreated here.

- The calculation of the envelope of a family of circles can be done by a geometrical construction in the space of circles. We can take a conic section in $\mathbb{P}(\mathfrak{Circ})$ that represents a rational family of circles. The representant $\mathcal{O}^\circ$ of its orthogonal circle and the conic section then define a cone with vertex $\mathcal{O}^\circ$ containing the conic section. This cone intersects the paraboloid Π of nullcircles in a spatial curve. The projection of this curve into $\mathbb{P}_2(\mathbb{C})$ by identifying the representants of nullcircles on Π with the corresponding points in projective space is the bicicircular quartic that is the envelope of the given rational family of circles.

- The four circles of inversion of a bicircular quartic are mutually orthogonal. They have a lot of other interesting properties, among others that their radii r_i fulfil the equation
$$\sum_{i=1}^{4} \frac{1}{{r_i}^2} = 0.$$

- Bicircular quartics can be generated by mechanical devices. A lot of research has been done on this field at the end of the 19$^{\text{th}}$ century. One example for such a mechanical generation is the construction of the cardioid as an epicycloid.

- Cyclides can be used for the interpolation and approximation of surfaces. This is an interesting application, because cyclides are more flexible than quadric surfaces. It could be useful to study families of circles represented by a cubic Bezier curve in $\mathbb{P}(\mathfrak{Circ})$.

- In the upper half plane model of hyperbolic geometry circles that are orthogonal to the real axis (including straight lines with this property) are the geodesics of this space. The transformation $z \mapsto z^2$ squeezes this space into the first quadrant. The geodesics in this new space are bicircular curves that have the real and the imaginary axis as circles of inversion.

- Interesting results have been found concerning the rectification of bicircular quartics. It turns out that elliptic functions play an important role here.

- Bicircular quartics (and all n-circular $2n$-tics in general) allow the definition a power function on $\mathbb{P}(\mathbb{C})$ in analogy to the power of a point with respect to a circle. This raises questions, for instance, about the shape of the locus of points that have the same power with respect to two or more given bicircular quartics.

This is just a short selection from the things that are not treated here.

The author of this text would greatly appreciate a comprehensive work covering the subject of circles and bicircular quartics, giving an overview over the history of these curves as well as discussing them under the aspect of algebraic geometry, covering practical applications of them and their relatives in higher dimensions and providing a rich collection of references to other works on this or similar topics.

Bibliography

[1] Arnold V. I., Gusein-Zade, S. M., Varchenko, A. N.: *Singularities of Differential Maps*, Volume I: The Classification of Critical Points, Caustics and Wave Fronts, translated by Ian Porteous, based on a previous translation by Mark Reynolds, in *Monographs in Mathematics*, Vol. 82, Birkhäuser, Boston - Basel - Stuttgart, 1985

[2] Ameseder, Adolf: *Über Fusspunktcurven der Kegelschnitte*, in *Grunert Arch.*, LXIV (1879), p. 143-144

[3] Ameseder, Adolf: *Zur Theorie der Fusspunktcurven der Kegelschnitte*, in *Grunert Arch.*, LXIV (1879), p. 145-164

[4] Anderson James, W.: *Hyperbolic geometry*, second edition, Springer, London, 2005

[5] Bartl, Josef Ludwig: *Analytische Geometrie*, 4 volumes, unpublished manuscript, posthumously xerocopied by the University of Hannover, 1979

[6] Barth, Wolf: *Kreise*, lecture notes, Universität Erlangen-Nürnberg, 1997

[7] Basset, Alfred Barnard: *An elementary treatise on cubic and quartic curves*, Cambridge, Deighton Bell and Co., 1901

[8] Bauer, Thomas: *Über eben algebraische Kurven vom Grad vier*, Zulassungsarbeit, Universität Erlangen-Nürnberg, 1990

[9] Blaschke, Wilhelm: *Projektive Geometrie*, Notdruck in *Bücher der Mathematik und Naturwissenschaften*, Wolfenbütteler Verlagsanstalt, Wolfenbüttel-Hannover, 1947

[10] Blaschke, Wilhelm: *Griechische und anschauliche Geometrie*, Verlag von R. Oldenbourg, München, 1953

[11] Bloch, Otto: *Die Ortskurven der graphischen Wechselstromtechnik nach einheitlicher Methode behandelt*, Verlag von Rascher & Co., Zürich, 1917

[12] Boehm, Wolfgang: *On cyclides in geometric modeling*, in *Conputer Aided Geometric Design*, Volume 7 (1989), p. 243-255

[13] Booth, J.: *Sur la rectification de quelques courbes*, in *Brioschi Ann., second series (Annali di Matematica pura ed applicata)*, Vol. II (1868), p. 81-88

[14] Brieskorn, Egbert; Knörrer, Horst: *Ebene algebraische Kurven*, Birkhäuser, Basel, 1981

[15] Brieskorn, Egbert: *Plane algebraic curves*, Birkhäuser, Basel, 1986

[16] Camerer, J. W.: *Apollonii de tactionibus quae supersunt ac maxime lemmata Pappi in hos libros graece nunc primum edita e codicibus mscptis, cum Vietae librorum Apollonii restitutione, adjectis observationibus, computationibus ac problematis Apolloniani historia*, Karl Wilhelm Ettinger, Gotha; J. St. van Esveldt Holtrop & Co., Amsterdam, 1795

[17] Casey, John: *On Bicircular Quartics*, in *Transactions of the Royal Irish Academy*, Volume 24 (1871), p. 457-569

[18] Casey, John: *On Cyclides and Sphero-Quartics*, in *Philosophical Transactions of the Royal Society of London*, Volume 161 (1871), p. 585-721

[19] Casey, John: *On a new Form of Tangential Equation*, in *Philosophical Transactions of the Royal Society of London*, Volume 167 (1877), p. 367-460

[20] Casey, John: *The First Six Books of the Elements of Euclid and Propositions I-XXI of Book XI*, Hodges, Figgis, & Co., Dublin; Longmans, Green, & Co., London, 1885

[21] Casey, John: *A Sequel to the First Six Books of the Elements of Euclid*, fifth edition, Hodges, Figgis, & Co., Dublin; Longmans, Green, & Co., London, 1885

[22] Cayley, Arthur: *On the Mechanical Description of a Nodal Bicircular Quartic*, in *Proceedings of the London Mathematical Society*, Vol. III (1871), p. 100-106

[23] Cayley, Arthur: *On certain Constructions for Bicircular Quartics*, in *Proceedings of the London Mathematical Society*, Vol. V (1874), p. 29-31

[24] Cox, David; Little, John; O'Shea, Donal: *Ideal, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, New York, 1998

[25] Coolidge, Julian Lowell: *A Treatise on the Circle and the Sphere*, At the Clarendon Press, Oxford, 1916

[26] Crofton, M. W.: *On Various Properties of Bicircular Quartics*, in *Proceedings of the London Mathematical Society*, Vol. II (1869), p. 33-45

[27] Dutta, Debasish; Martin, Ralph R.; Pratt, Michael J.: *Cyclides in Surface and Solid Modeling*, in *IEEE Computer Graphics and its Applications*, 13 (1993), p. 53-59

[28] Emch, Arnold: *Illustration of the elliptic integral of the first kind by a certain link-work*, in *Annals of Mathematics*, second series, Vol. 1 (1900), p. 81-92

[29] Emch, Arnold: *An application of elliptic functions to Peaucellier's link-work (inversor)*, in *Annals of Mathematics*, second series, Vol. 2 (1901), p. 60-63

[30] Euklides; Thaer, Clemens: *Die Elemente*, published as *Ostwalds Klassiker der exakten Wissenschaften 235, 236, 240, 241, 243*, Akademische Verlagsgesellschaft, Leipzig, 1975

[31] Faucette, William Mark: *A Geometric Interpretation of the Solution of the General Quartic Polynomial*, in *The American Mathematic Monthly*, 103 (1996), p. 51-57

[32] Fischer, Gerd: *Ebene algebraische Kurven*, Vieweg, Braunschweig, 1994

[33] Fladt, Kuno: *Analytische Geometrie spezieller ebener Kurven*, Akademische Verlagsgesellschaft, Frankfurt am Main, 1962

[34] Ganguli, Surendramohan: *Lectures on the theory of plane curves*, Calcutta University Press, Calcutta, 1919

[35] Hart, Harry: *On the Focal Conics of a Bicircular Quartic*, in *Proceedings of the London Mathematical Society*, Vol. XI (1880), p. 143-151

[36] Hart, Harry: *On the linear vectorial equation of the central of a conic*, in *Messenger of mathematics*, Vol. 12 (1882), p. 33

[37] Hart, Harry: *On the Focal Conics of a Bicircular Quartic*, in *Proceedings of the London Mathematical Society*, Vol. XIV (1883), p. 199-202

[38] Hart, Harry: *On the focal quadrics of a cyclide*, in *Messenger of mathematics*, Vol. 14 (1884), p. 1-8

[39] Hirst, T. A.: *On the Quadric Inversion of Plane Curves*, in *Proceedings of the Royal Society of London*, Volume 14 (1865), p. 91-106

[40] Hirsch, Anton: *Extension of the 'Villarceau-Section' to Surfaces of Revolution with a generating conic*, in *Journal for Geometry and Graphics*, Volume 6 (2002), No. 2, p. 121-132

[41] Loria, Gino: *Spezielle algebraische und transcendente ebene Kurven – Theorie und Geschichte*, german edition by Fritz Schütte, B. G. Teubner, Leipzig, 1902

[42] Lüders, Otto: *Über orthogonale Invarianten der bizirkularen Kurven vierter Ordnung*, Inaugural-Dissertation zur Erlangung der Doktorwürde der Hohen philosophischen Fakultät der Vereinigten Friedrichs-Universität Halle-Wittenberg, Buchdruckerei des Waisenhauses, Halle an der Saale, 1910

[43] Michael, W.: *Die Konstruktion des singulären Punktes der bizirkularen Quartik und der durch ihn gehenden Tangentialkreise*, in *Archiv für Elektrotechnik*, 30 (1936), p. 199-206

[44] Pedoe, Daniel: *Circles, A Mathematical View*, The Mathematical Association of America, Washington, DC, 1995

[45] Pišl, Milan: *Cirkulární kubiky a bicirkulární kvartiky*, in *Pokroky matematiky, fyziky a astronomie*, 3 (1958), p. 32-41

[46] Ramamurti, B.: *A contribution to the inversion geometry of the bicircular quartic*, in *Journal of the Indian Mathematical Society*, Vol. 19 (1932), p. 177-181

[47] Reye, Theodor: *Synthetische Geometrie der Kugeln und linearen Kugelsysteme*, B. G. Teubner, Leipzig, 1879

[48] Roberts, Samuel: *On the Ovals of Descartes*, in *Proceedings of the London Mathematical Society*, Vol. III (1871), p. 106-126

[49] Salmon, George: *A treatise on the higher plane curves – intended as a sequel to the treatise on conic sections*, first edition, Hodges and Smith, Dublin, 1852

[50] Salmon, George: *A treatise on the higher plane curves – intended as a sequel to the treatise on conic sections*, second edition, Hodges and Foster, Dublin, 1873

[51] Salmon, George: *A treatise on conic sections*, 6th ed., reprinted, AMS Chelsea Pub, Providence, RI, 2005, originally published: Chelsea Pub., New York, 1954

[52] Silberhorn, Christine: *Klassifikation der Hüllkurven symmetrischer elliptischer Kreisscharen*, Zulassungsarbeit, Universität Erlangen-Nürnberg, 1997

[53] Simpson, Harold: *A note on two-circuited circular cubics and bicircular quartics*, in *Journal of the London Mathematical Society*, Vol. 17 (1942), p. 31-33

[54] Waerden, Bartel Leendert van der: *Erwachende Wissenschaft – Ägyptische, babylonische und griechische Mathematik*, Birkhäuser Verlag, Basel; Stuttgart, 1966

[55] Villarceau, Yvon; François, Antoine-Joseph: *Théorème sur le tore*, in *Nouvelles Annales de Mathématiques*, Série 1, Tome 7, Gauthier-Villars, Paris, 1848, p. 345-347

[56] Volenec, Vladimir: *Concyclic or collinear points on a rational circular cubic or on a rational bicircular quartic*, in *Rad Hrvatske Akademije Znanosti i Umjetnosti, Razreda Matematicke, Fizicke, Kemij i Tehnicke Znanosti*, Vol. 11 (1994), p. 77-83

[57] Wieleitner, Heinrich: *Spezielle ebene Kurven*, Göschen, Leipzig, 1908

[58] Wieleitner, Heinrich: *Algebraische Kurven*, Bd.1: Gestaltliche Verhältnisse, Bd.2: Allgemeine Eigenschaften, reprint by de Gruyter, Berlin, 1943

[59] Wilson, R.: *Power Coordinates and the Bicircular Quartic*, in *Proceedings of the Mathematical Society of Edinburgh*, 43 (1925), p. 26-34