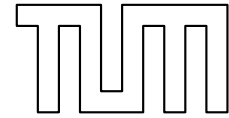




TECHNISCHE UNIVERSITÄT MÜNCHEN
Zentrum Mathematik



Feature Detection for Real Plane Algebraic Curves

Peter Michael Lebmeir

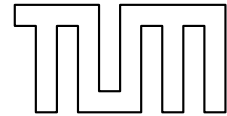
Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.



TECHNISCHE UNIVERSITÄT MÜNCHEN
Zentrum Mathematik



Feature Detection for Real Plane Algebraic Curves

Peter Michael Lebmeir

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Rupert Lasser
Prüfer der Dissertation: 1. Univ.-Prof. Dr. Dr. Jürgen Richter-Gebert
2. Univ.-Prof. Dr. Gerd Fischer
3. Prof. Dr. Tomás J. Recio Muñoz
Universidad de Cantabria / Spanien
(Schriftliche Beurteilung)

Die Dissertation wurde am 21. Januar 2009 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 20. Februar 2009 angenommen.

Danksagung

Ich danke Herrn Prof. Dr. Dr. Richter-Gebert und allen Mitarbeitern des Lehrstuhls für Geometrie und Visualisierung für Ihre Unterstützung bei verschiedenen Problemstellungen und für die vielen weitreichenden Einblicke auf didaktischer und fachlicher Seite. Ohne Sie wäre diese Arbeit nicht möglich gewesen. Mein besonderer Dank gilt auch den Gutachtern Prof. Dr. Gerd Fischer und Prof. Dr. Tomás Recio sowie dem Vorsitzenden des Prüfungsausschusses Prof. Dr. Rupert Lasser. Für TensorMania ein Danke an Michael Schmid.

Zusammenfassung

Diese Arbeit beschäftigt mit reellen ebenen algebraischen Kurven und ihren Eigenschaften. Dabei stellt sich heraus, dass reelle Kurven am besten durch komplexe Zahlen repräsentiert werden. Das Transformationsverhalten und Normalformen lassen sich in ihnen wesentlich einfacher beschreiben. So bewirken Rotationen der Ebene entsprechende Rotationen der Kurvenkoeffizienten. Eine translatorische Normalform kann darüber hinaus durch Auslöschung von Kurvenkoeffizienten erreicht werden. Eine derartige Form vereinfacht die Erkennung etwaiger Rotations- und Spiegelsymmetrien: Es sind nur Koeffizientenmuster aufzudecken und Winkeldiagramme zu analysieren. Für die Darstellung invarianter Eigenschaften werden Tensordiagramme herangezogen. Dabei genügt es geschlossene Diagramme aus den gewünschten Tensorknoten zu erstellen. Diagramm-Rechenregeln schaffen zusätzliche Vereinfachungen, um geometrische Eigenschaften der Kurven ansprechend darzustellen.

Abstract

This thesis concerns real plane algebraic curves and their attributes. It can be ascertained that real curves are best represented by complex numbers. In this case, normal forms and the behavior of curves under transformations of the plane can be described quite intuitively. Rotations of the plane effectuate rotations of the curve-coefficients. Even more: A translatorial normal form is accessible just by greedily making coefficients vanish. The presented form also permits an easy symmetry-detection. Potential rotational and reflectional symmetries can be uncovered by finding additional coefficient patterns and by analyzing angle diagrams. Invariant attributes of curves are visualized by closed tensor-diagrams with the desired tensors as nodes. Diagrammatical calculation rules cut complicated situations down and geometric properties of the curve have a simplified representation.

Contents

1. Introduction	1
1.1. Curves from dynamic geometry programs	1
1.2. Feature detection in general	2
1.3. Chapters and languages	3
2. Curves and Tensors in Algebra and Geometry	6
2.1. Polynomials	6
2.1.1. Attributes of polynomials	6
2.1.2. Resultant	7
2.1.3. Polynomials versus zero sets	8
2.1.4. Polynomials in the affine and projective plane	9
2.1.5. Polynomials as elements of a vector space	10
2.2. Curves	11
2.2.1. Curves as zero sets	11
2.2.2. Effects of transformations of the plane on curves	13
2.2.3. Rotating real curves around the origin	15
2.2.4. Excursus on tetrahedral coefficient structure	18
2.3. Complexification	22
2.3.1. Complexified curves	23
2.3.2. Structure of Complexification	25
2.3.3. Equivalence of real and complexified curves	26
2.3.4. The coefficient triangle	27
3. Curves and Tensor-Algebra	28
3.1. Connection between tensor algebra and curves via forms	28
3.1.1. Covariant and contravariant vectors	29
3.1.2. Tensors	30
3.1.3. Diagram notation of tensors	32
3.1.4. Tensoroperations	33
3.1.5. Curves as tensors	36
3.2. Invariants	39
3.2.1. The δ - and ε -tensor	39
3.2.2. Closed diagrams and projective invariants	41
3.2.3. Diagrams and (non)-Euclidean invariants	42
3.3. Diagram laws or how to calculate with diagrams	44
3.3.1. Complexification in tensor-notation	44
3.3.2. Juggling with \widehat{E} -, \widehat{A} - and one curve-tensor	45
3.3.3. Formal tensor-decomposition and the role of \widehat{E} and \widehat{A}	46
3.3.4. Excursus: The Pascal-Pyramid	48
3.3.5. Juggling with \widehat{E} -, \widehat{A} - and two curve tensors	53

3.3.6. More replacement rules	56
4. Curves and selected transformations of the plane	59
4.1. Rotating curves around the origin	59
4.1.1. Effects of the complexification	60
4.1.2. Invariants with respect to rotations around the origin	61
4.1.3. Reducing rotation invariants	62
4.1.4. Reconstruction of a curve by rotation invariants	66
4.1.5. Reconstruction of specific given curve	69
4.1.6. Normal form with respect to rotations around the origin	72
4.1.7. Invariants under rotations around the origin in tensor notation	73
4.2. Translation	76
4.2.1. Effects of translations on complexified curve coefficients	76
4.2.2. Structure of the coefficient transformation	77
4.2.3. Reference for translations	80
4.2.4. Minimizing coefficients by a translation	82
4.2.5. Minimizing coefficients by a translation of fixed direction	84
4.2.6. Translational normal form	87
4.2.7. Annihilation and normal form for complexified conics	89
4.2.8. Invariants under translation	92
4.3. Scaling	93
4.3.1. Effects of scaling on curve coefficients	93
4.3.2. Normal form and invariants	94
4.4. Projective transformations	95
4.4.1. Effects of projective transformations on curve coefficients	95
4.4.2. Projective invariant expressions	96
5. Symmetry detection	102
5.1. Rotational symmetry	102
5.1.1. Definition and restrictions	102
5.1.2. Rotational symmetry with respect to the origin as center	105
5.1.3. Rotational symmetry with respect to an arbitrary center	106
5.2. Reflective symmetry	109
5.2.1. Definition and restrictions	109
5.2.2. Reflective symmetry with axes passing through the origin	111
5.2.3. Reflective symmetry with arbitrary symmetry axes	113
5.3. Examples	115
6. Summary and Outlook	119
6.1. Summary	119
6.2. Outlook	120
A. Appendix	122
A.1. Rotation in real and monomial representation	122
A.2. Symmetric and skew-symmetric tensors	123
A.3. Basis and dual basis in the layers of the Pascal-Pyramid	124
B. Minimizing the norm of the coefficient vector	126

C. Coefficient Recognition	130
C.1. Introduction	130
C.2. Computationally constructed curves	132
C.3. Curve recognition	133
C.4. Type of constructed loci	136
C.5. Experimental results	137
C.5.1. Finding the degree	138
C.5.2. Shifting and preconditioning	142
C.5.3. Investigating curve invariants	143
C.6. Conclusions	143

1. Introduction

For searching and examining, for describing and understanding structures the appropriate language is always crucial: It must be powerful enough to encompass the circumstances, yet succinct enough to get to the point of things. The best way to describe structures is mathematics. But also within mathematics different things require different descriptions - have their own languages so to say. We follow this path of finding and analyzing real plane algebraic curves - or short: curves - in their *adequate languages*. We will see complex numbers in two- and three-dimensional coefficient tables representing curves, revealing deeper structures of transformations, closed diagrams *talk* about invariant properties, angle diagrams tell from symmetries and many more. In principle it does not matter where the curves came from, however nowadays they will most likely originate from dynamic geometry. All that matters is that in the end a curve coefficient vector can be provided for our analysis of features such as normal forms, invariant properties and symmetries.

1.1. Curves from dynamic geometry programs

In dynamic geometry programs algebraic curves mostly occur as loci. For example a ruler and compass construction is made by placing *free points* in a plane, restricting points to a line or a circle (*semi-free points*) and constructing one or more points, whose position is totally determined by the construction (*dependent points*). Already at this point the existing programs differ widely: Many programs do not incorporate points at infinity. Thus parallel lines do not have an intersection available for further constructions. Also complex intersections, such as intersection points of two unit circles, whose centers are more than two units apart, are not at all common.

A locus emerges when a distinct semi free point (*mover*) - a point with one degree of freedom like a point on a circle or a line - is moved while a distinct dependent point (*tracer*) is watched: To each position of the moved semi-free point, the dependent point describes a point on a curve (see Figure 1.1). If no more than ruler and compass constructions are used, the described curve is algebraic. A rigorous mathematical formalization by the use of straight-line programs can be found in [30] and [31].

In this way dynamic geometry programs are able to determine a large enough sample point set of a branch of a curve. These points are usually of high arithmetic precision. These points can be used to determine a curve coefficient vector of minimal degree. As a first entry-point on this topic my thesis-supervisor and I wrote a paper [35], which can be interpreted as preliminary work to this thesis (see also Appendix C). Alternatively we refer to [48], [49] and [50] for further work on coefficient-recognition.

Of course, the numerical quality of the coefficient vector depends on the selection of the points given by the dynamic geometry program. We would normally expect the points to be somehow “distributed well” over the “whole” curve, but certain difficulties impose restrictions. If the software does not make use of the projective extension of the

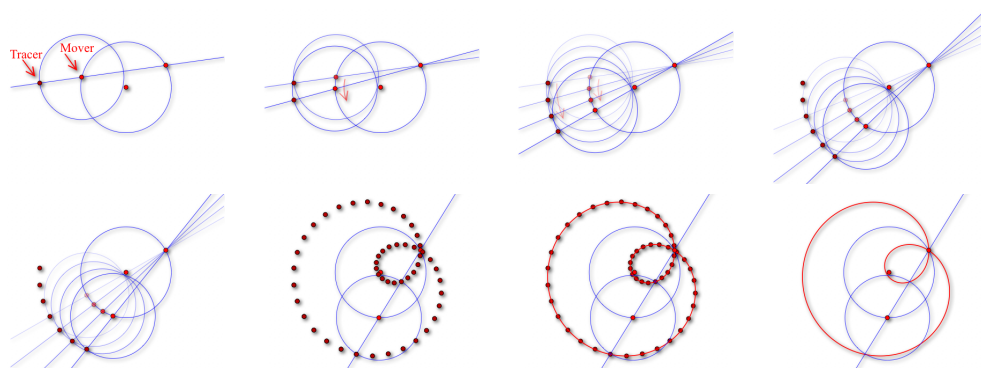


Figure 1.1.: Construction of a Limaçon with Cinderella

Euclidean plane or if it lacks complex numbers, only small parts of curve-branches might be represented by sample points. When the mover runs into a situation where points at infinity or (intermediately) complex points would be needed, no more sample points are available thereafter.

There is also a more subtle restriction by singularities: Singularities impose a serious boundary for the mover-movement in some dynamic geometry software. Only a complex detour can solve this problem (see [31]). Thus only dynamic geometry software incorporating complex projective calculation is capable of providing a reasonably distributed sample point set, in case the selection is not restricted by the user. We therefore use Cinderella (see [40]), not only for dealing with algebraic curves from dynamic geometry software, but also for our figures. However, a principle difficulty remains: The user himself may constructively restrict the sample-points to a tiny part of a curve. For example, an extreme case would be if the samples were taken from such a tiny part of a parabola that computationally it could not be told apart from a sample point set of a line.

1.2. Feature detection in general

In the framework of this thesis we refrain from all these interesting questions and begin our study with a curve coefficient vector to a real algebraic plane curve. It does not matter where it came from and which problems arose getting it. This thesis can extend existing coefficient-recognition-programs by a downstream feature recognition. Those may provide the user with additional information on his or her construction.

In this case we take a given coefficient-vector and the curve represented by it, as is. For example a Limaçon $\mathcal{C} = \mathcal{C}_g$ may be given via a zero set $V(g)$ of a polynomial g :

$$\mathcal{C}_g = V(g) = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0 \wedge g \in \mathbb{R}[x, y]\}$$

$$g(x, y) = -192 + 96\sqrt{2}y + 24x^2 - 96xy + 24y^2 + 24\sqrt{2}x^3 - 24\sqrt{2}x^2y + 24\sqrt{2}xy^2 - 24\sqrt{2}y^3 + 6x^4 + 12x^2y^2 + 6y^4$$

with curve coefficient vector

$$(-192 \ 0 \ 96\sqrt{2} \ 24 \ -96 \ 24 \ 24\sqrt{2} \ -24\sqrt{2} \ 24\sqrt{2} \ -24\sqrt{2} \ 6 \ 0 \ 12 \ 0 \ 6) .$$

This way the coefficients of the polynomial g become curve-coefficients. For some algorithms, however, it is necessary for the polynomial to represent an irreducible or at least a square-free curve. These terms are defined in detail in Chapter 2. Only so far: Superimposing the unit-circle with another unit-circle provides a new curve, which naturally looks like a unit-circle, but now is covered twice. This can generate problems.

To be able to apply the proposed techniques nonetheless, we feel free to restrict the given coefficient-vectors in that way. This naturally imposes restrictions on the coefficient-recognition-algorithm if one is used. However, this is more a desirable feature of such an algorithm than a restriction, since it enables fast and unique output.

Now starting with a given curve, we want to detect features. Features in our sense are characteristics with respect to a transformation group: Invariant properties and normal-forms, but also symmetries like rotational and reflectional symmetry. In general, the approach in this thesis is finding the right *languages* for feature detection followed by a detailed analysis of the expressions in that language.

1.3. Chapters and languages

In Chapter 2 we introduce many technical terms, but also *the* language of curves: Real algebraic plane curves are best represented by a complex coefficient vector (see Section 2.3 and also [46], [47], [52]). *Complexification* comes from the substitution of x and y by $\frac{z+\bar{z}}{2}$ and $\frac{z-\bar{z}}{2i}$, respectively. This is best complemented by *multinomial scaling*, altering the absolute values of the coefficients.

The Limaçon from the last section is represented by the following complexified coefficient vector (compare with left image in Figure 1.2 on page 5):

$$\left(-192 \quad 12\sqrt{2}e^{-i\frac{\pi}{2}} \quad 12\sqrt{2}e^{i\frac{\pi}{2}} \quad 4e^{i\frac{\pi}{2}} \quad 2 \quad 4e^{-i\frac{\pi}{2}} \quad 0 \quad 2e^{i\frac{\pi}{4}} \quad 2e^{-i\frac{\pi}{4}} \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \right)$$

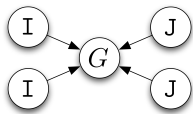
This leads to easy and beautiful structure-analyzations: Simple transformations of the plane, being complicated in real calculations, become simple again (see Section 2.2.3 and Section 4.1.1). However, we will not neglect profound, fascinating structures: Section 2.2.4 contains one of the first examinations on coefficient structures in this thesis. In it we speak in terms of two- and three-dimensional coefficient-tables, unveiling what impact transformations of the plane have on curve-coefficients. These structures confirm that the language of curves is complex (see Section 2.3.2).

Another very powerful language is *tensor-notation*. Chapter 3 introduces the basics on tensor-algebra in connection with curves and rapidly switches to *diagrams*. Curves of degree d are now represented via tensors G_{k_1, k_2, \dots, k_d} , whose components are taken from the curve coefficients. In diagrams they appear as nodes with d edges pointing inwards:



$\textcircled{G} \leftarrow \equiv$. We became aware of tensor-diagrams by a series written by Jim Blinn (cf. [9], [10]), but similar diagrams were even earlier in use (cf. [14], [15]). When it comes to invariant properties of curves, tensor diagrams are fantastic. Only closed diagrams with very few different nodes need to be drawn in order to obtain an invariant property (see Section 3.2.2 and Blinn's series). For projective ones only two types of tensors are needed: the curve-tensor and a totally skew-symmetric $3 \times 3 \times 3$ -tensor. For Euclidean properties two more tensors are added and further extensions of the used tensors lead

to yet further sub-geometries.



is a diagram resulting in a leading curve-coefficient of a curve of degree $d = 4$. In the case of our *Limaçon*, this is the rightmost non-vanishing coefficient in our complexified coefficient vector above. The nodes I and J represent the complex conjugate absolute circle points.

Diagrammatical calculations can also be carried out. We can juggle with tensor-nodes, replace the absolute circle-points I and J by corresponding conics and get another fascinating coefficient-study: Linear-combinations of generalized Pascal-Triangles build the connection and geometrically we may speak of a coefficient pyramid as the source of the scalars in these combinations (see Section 3.3.4).

Chapter 4 focuses on the actions of transformations of the plane on curves: rotations, translations, scalings and projective transformations are discussed. Thereby all the previously introduced languages are used and (partially) even new ones are created. Rotations of curves in complex representation are particularly simple, since they are just a rotation of the complexified coefficients (see Section 4.1.1). Invariant properties become visible and are easy to formulate in diagram-language (see Section 4.1.7). A complete set of rotational invariant expressions may be generated by multiplying the curve tensor with a series of I, J and the center of rotation. We propose a minimal system, permitting the reconstruction of any curve up to rotations. Also a normal-form with respect to rotations around the origin is proposed in Section 4.1.6. It is obtained by successively minimizing the angles $\varphi \in [0, 2\pi)$ of complex curve coefficients $c = re^{i\varphi}$. *Rotating our Limaçon by $\psi = \frac{-\pi}{4}$ we get*

$$\left(-192 \quad 12\sqrt{2}e^{-i\frac{\pi}{4}} \quad 12\sqrt{2}e^{i\frac{\pi}{4}} \quad -4 \quad 2 \quad -4 \quad 0 \quad 2i \quad -2i \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \right) .$$

Each new coefficient has the same absolute value as the previous one. The rotation of the coefficients is easily related to ψ as shown in Section 4.1.1. This curve is shown in the middle of Figure 1.2 on page 5.

Translations are a bit more complicated to handle. But again beautiful structures can be discovered, when curves are moved (cf. Section 4.2.2). It is astonishing to see that there are curves where certain coefficients are contained in a one-dimensional set, wherever the curve is translated (see Sections 4.2.4 and 4.2.5). We present a translatorial normal form, which can be established by coefficient analysis: Greedily making sub-leading, sub-sub-leading, ... coefficients vanish (or at least minimizing them) results in the desired form (see Section 4.2.6).

The translatorial normal form of our rotated Limaçon makes the “2i”-coefficient vanish:

$$\left(0 \quad 12 \quad 12 \quad 0 \quad -6 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \right) .$$

This is done by a translation with $t = -2i$ (see right image in Figure 1.2).

At first our normal form looks similar to the (inspiring) form, proposed in [46]. However, that form implies too many disadvantages and does not satisfy the “normal-form-condition” without additions. (For a detailed analysis see Appendix B.) Our normal form seems to be more natural to curves and also permits easier symmetry detection.

But extra care has to be taken for numerical stability in special cases. Scalings are the most simple transformations to handle. Together with rotations and translations they give us the features of curves with respect to orientation preserving Euclidean motions. The direct analysis of the effects of projective transformations on curve coefficients is almost useless (see Section 4.4.1). However, it never becomes any clearer than there that tensor-diagrams are the suitable language to describe features of curves. Analytical expressions of hundreds and more terms can be contained in simple and easy-to-handle diagrams. Bigger configurations can be cut down by diagrammatical calculation rules, which correspond in some way to Grassmann-Plücker relations. However, until now a complete set of invariant properties, permitting a reconstruction of any curve of any degree with respect to projective transformations, remains unknown to us. Now, with the observations made, it seems to be within reach.

Finally Chapter 5 focuses on symmetry-features. If a curve can be rotated non-trivially such that it is equal to the original, it is called rotational symmetric. Given a complexified curve, rotational symmetry with respect to the origin as center can be brought down to analyzing patterns in a triangular coefficient scheme. If the only non-vanishing coefficients of a curve are contained in equally spaced columns of that scheme, we have a rotational symmetry (see Section 5.1.2). The biggest benefit arises from the study of translations: Translating any curve in a way that it is in our proposed translatorial normal form, means it is rotational symmetric with respect to the origin or not at all (see Section 5.1.3). An analogous result holds for reflectional symmetries (see Section 5.2.3): Either the curve in translatorial normal form is not at all reflectionally symmetric or its axis is passing through the origin. The language of reflectional symmetries is the one of angle diagrams: The angle φ of each non-vanishing coefficient $c = re^{i\varphi}$ contains information on the angles of potential hyper-plane-mirrors. Taking all these angles together - for example in the form of angle diagrams - permits the detection of reflectional symmetry (see Section 5.2.2).

In the last (complexified) coefficient-vector of our Limaçon, all coefficients were real. The first coefficient being equal to twelve is named $c_{10} = r_{10}e^{i\varphi_{10}} = 12e^{0\pi i}$. It leads to the angle diagram



and thus to a detected reflectional symmetry with the horizontal coordinate axis as mirror.

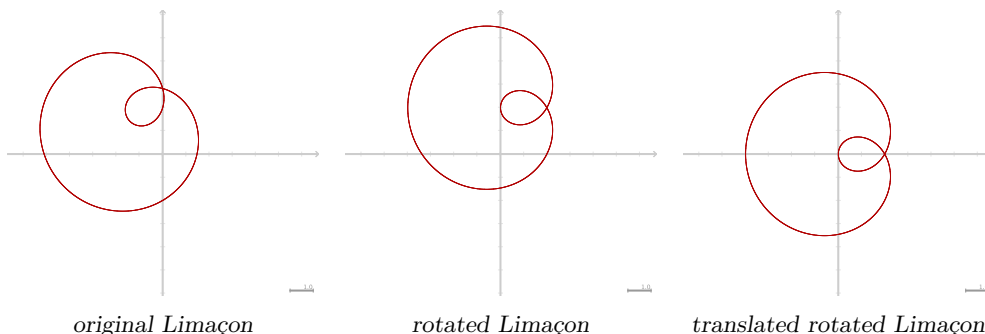


Figure 1.2.: A close-up of a Limaçon

2. Curves and Tensors in Algebra and Geometry

Real plane algebraic curves - or short: curves - can be represented in various ways. One of the most elementary representations is probably the curve as zero set of a polynomial. Therefore, we precede the examination of curves with some results on polynomials (Section 2.1), the concepts of which are then transferable to the zero sets and thus the curves. Polynomials, however, build a vector space. They can be represented with respect to infinitely many different bases. We will have a closer look at monomial basis (Section 2.1). Especially a change to a specific complex basis is discussed in 2.3. For us polynomials will be one possible ingredient to describe curves via zero sets (Section 2.2.1). If curves are subject to translations, we will make extensive use of polynomials with complex coefficients. To spice up this technical introduction, an excursus on a mathematically beautiful tetrahedral coefficient structure is also included (Section 2.2.4). We will encounter such coefficient structures more than once: for example, when curves are rotated.

2.1. Polynomials

We do not intend to cover the general theory of polynomials. The algebra behind these objects is only used in so far that it helps to work with curves later on. For a more in-depth algebraical study of polynomials we refer to algebra textbooks (cf. [21]). We also confine ourselves to polynomials over the field \mathbb{C} of complex numbers. Generalizations to arbitrary polynomial rings can be found in the cited algebra textbooks or in the context of curves in [11]. The benefit of complex polynomials in contrast to real polynomials is the algebraic closure of \mathbb{C} , which can be extensively exploited. In what follows, we will take the “detour” through \mathbb{C} and come back to real curves later on. For a direct in-depth study on real algebraic geometry we refer to [13].

2.1.1. Attributes of polynomials

Using homogenized and dehomogenized set-ups, we do not yet fix the number of arguments of polynomials. We also want to dwell on real and complex represented real polynomials as well as complex polynomials. Therefore, we also do not fix the field: $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Most concepts and terminology presented here will be transferred to curves in Section 2.2.1.

Definition 2.1

Let \mathbb{K} be the field \mathbb{C} or \mathbb{R} . Let $f \in \mathbb{K}[X_1, \dots, X_k]$ be a polynomial over \mathbb{K} . It can be written in the following form

$$f(x_1, \dots, x_k) = \sum_{\substack{r_1 + \dots + r_k = d \\ r_1, \dots, r_k \geq 0}} A_{r_1 \dots r_k} \cdot x_1^{r_1} \cdot \dots \cdot x_k^{r_k}$$

with coefficients $A_{r_1 \dots r_k} \in \mathbb{K}$. Without loss of generality we may assume that there are indices r_1, \dots, r_k with $\sum_{s=1}^k r_s = d$ and $A_{r_1 \dots r_k} \neq 0$. Then $d = \deg(f)$ is called the **degree** of f .

If $f(tx_1, \dots, tx_k) = t^d f(x_1, \dots, x_k)$ in $\mathbb{K}[X_1, \dots, X_k][T]$, then f is called **homogeneous** of degree d .

If the implication $a \in \mathbb{K}$ or $b \in \mathbb{K}$ holds for any pair of complex polynomials $a, b \in \mathbb{K}[X_1, \dots, X_k]$ satisfying $f = a \cdot b$, then f is called **irreducible**. f is named **prime**, in cases when f , dividing a polynomial $a \cdot b$, implies that f divides a or f divides b .

A point (x_1, \dots, x_k) is called a **zero** or **root** of f if $f(x_1, \dots, x_k) = 0$.

The set

$$V(f) = \left\{ (x_1, \dots, x_k) \in \mathbb{K}^k \mid f(x_1, \dots, x_k) = 0 \right\}$$

is called **zero set** of f or **variety** of f .

Examples of irreducible polynomials over \mathbb{C} are: $1, i, x+i, x_1+3ix_2$, etc. Irreducible polynomials over \mathbb{R} are for example: $1, x+2, x^2+1, \dots$. But the polynomial x^2+1 is not irreducible over \mathbb{C} , because it can be decomposed into two factors: $x^2+1 = (x+i)(x-i)$. From algebra we know that $\mathbb{C}[X_1, \dots, X_k]$ is factorial (see [11] or [21]). This means that any non-constant polynomial $f \in \mathbb{C}[X_1, \dots, X_k] \setminus \mathbb{C}$ is a product of finitely many irreducible elements and each irreducible element is prime. Now if f is a polynomial and $f = a \cdot b$ then $V(a) \subseteq V(f)$ and $V(b) \subseteq V(f)$. The conclusion from varieties back to polynomials is not that easy. Before we go into details we examine the resultant, which connects two polynomials by the use of their coefficients.

2.1.2. Resultant

Definition 2.2 Let $R = \mathbb{C}$ or $R = \mathbb{C}[X_1, \dots, X_k]$. Furthermore let $f, g \in R[X]$. Let $f(x) = \sum_{r=0}^m A_r x^r$ be of degree m and $g(x) = \sum_{r=0}^n B_r x^r$ be of degree n . Then

$$\det \begin{pmatrix} A_m & A_{m-1} & \dots & \dots & A_0 & & & & & & \\ & A_m & A_{m-1} & \dots & \dots & A_0 & & & & & \\ & & & \ddots & \ddots & & & & & & \ddots \\ & & & & A_m & A_{m-1} & \dots & \dots & A_0 & & \\ \hdashline & B_n & B_{n-1} & \dots & \dots & B_0 & & & & & \\ & & B_n & B_{n-1} & \dots & \dots & B_0 & & & & \\ & & & & \ddots & \ddots & & & & & \ddots \\ & & & & & B_n & B_{n-1} & \dots & \dots & & B_0 \end{pmatrix} = \text{res}(f, g)$$

is called the **resultant** of f and g . Thereby the coefficients of f are contained in the first n lines (above the dashed line). The coefficients of g appear in the remaining m lines. If g is the derivative f' of f , namely $g(x) = f'(x) = \sum_{r=1}^m r A_r x^{r-1}$, then $\text{res}(f, f')$ is called **discriminant** of f .

Resultants and discriminants have many interesting features. We will state a few of them here. For proofs and further treatment we refer to [21] and [24].

Theorem 2.1 Let $R = \mathbb{C}$ or $R = \mathbb{C}[X_1, \dots, X_k]$ and let $f, g \in R[X]$ be non-constant polynomials. Then the following statements are equivalent:

- $\text{res}(f, g) = 0$
- $\exists p, q \in R[X]$ with $\deg(p) < \deg(g)$ and $\deg(q) < \deg(f)$, such that $0 = pf + qg$
- $\deg(\text{gcd}(f, g)) \geq 1$, i.e. f and g have a non-constant common divisor.

Corollary 2.2 *The discriminant $\text{res}(f, f')$ of a polynomial f vanishes if and only if f has multiple components. If the discriminant of f is non-zero, f is called **square-free**.*

2.1.3. Polynomials versus zero sets

The following lemma is very important. It connects the zero sets and polynomials reasonably. It permits conclusions from a zero set to the underlying polynomial:

Theorem 2.3 (Study's Lemma)

Let $f, g \in \mathbb{C}[X_1, \dots, X_k]$ be two complex polynomials. If the polynomial f is irreducible and $V(f) \subseteq V(g)$, then f divides g .

PROOF This proof is very technical but straightforward algebra.

If $g = 0$, the above theorem is trivially true. Otherwise $g \neq 0$ and also $f \neq 0$ due to f being irreducible by assumption. Let f be non-constant because $f \in \mathbb{C}$ trivially divides g . Thus a non-constant term is contained in f . Without loss of generality we can say that this term involves x_k . More formally we have: Without loss of generality $f \notin \mathbb{C}[X_1, \dots, X_{k-1}]$. Thus f can be written as polynomial of degree $d \geq 1$ in x_k , namely

$f(x_1, \dots, x_k) = \sum_{n=0}^d A_n(x_1, \dots, x_{k-1})x_k^n$ with coefficients $A_n \in \mathbb{C}[X_1, \dots, X_{k-1}]$. For given (x_1, \dots, x_{k-1}) we abbreviate this polynomial by $f_{x_1, \dots, x_{k-1}}(x_k) = \tilde{f}(x_k)$.

We show first that this implies $g \notin \mathbb{C}[X_1, \dots, X_{k-1}]$, i.e. g also contains a non-vanishing term in x_k . Assume the contrary, then $A_d \cdot g \in \mathbb{C}[X_1, \dots, X_{k-1}]$. Due to $A_d \neq 0$, there must be a point (x_1, \dots, x_{k-1}) , which is not a root of $A_d \cdot g$. Consequently, g as element of $\mathbb{C}[X_1, \dots, X_k]$ satisfies $g(x_1, \dots, x_{k-1}, y) \neq 0$ for all $y \in \mathbb{C}$. But $\tilde{f}(y)$ must have a root due to $d \geq 1$ and the fundamental theorem of algebra. Therefore, $V(f) \not\subseteq V(g)$ in contradiction to the assumption.

The next step in this proof is to show that the resultant $\text{res}(\tilde{f}, \tilde{g}) = 0$ vanishes independently from the values x_1, \dots, x_{k-1} . Let us first look at the (c_1, \dots, c_{k-1}) , where the leading coefficient $A_d(c_1, \dots, c_{k-1}) \neq 0$ of \tilde{f} does not vanish: Due to the fundamental theorem of algebra, there must be a c_k such that $\tilde{f}(c_k) = 0$. By assumption $V(f) \subseteq V(g)$ and thus by Theorem 2.1 $\text{res}(\tilde{f}, \tilde{g})(c_1, \dots, c_{k-1}) = 0$. As a consequence $0 = A_n \cdot \text{res}(\tilde{f}, \tilde{g}) \in \mathbb{C}[X_1, \dots, X_{k-1}]$. Since A_n is not vanishing identically, we have the resultant $\text{res}(\tilde{f}, \tilde{g}) = 0$.

By Theorem 2.1 f and g have a non-constant common divisor. By the irreducibility of f , f divides g . □

The clue in the proof is that f and g are polynomials over an algebraically closed field. A corresponding theorem based on the field \mathbb{R} does not hold. Take $f(x) = 1 + x^2$, which is irreducible over \mathbb{R} . Then $V(f) = \emptyset \subseteq V(g)$ for any polynomial g . But clearly f does not divide any polynomial.

Study's Lemma connects the polynomials and the zero sets reasonably. If f and g with $V(f) = V(g)$ are irreducible, then f and g are equal up to a constant multiple. When it comes to feature detection of curves this will be very helpful. We will mainly use the

Corollary 2.4 *Let $f, g \in \mathbb{C}[X_1, \dots, X_k]$ be complex polynomials.*

If f and g are irreducible (and thus also square-free) with $V(f) = V(g)$, then f differs from g only by a constant multiple: $f = \lambda g$ for some $\lambda \in \mathbb{C}$.

If only f is irreducible (and thus also square-free) with $V(f) = V(g)$, then $g = \lambda f^u$ for some $\lambda \in \mathbb{C}$ and $u \in \mathbb{N}$.

2.1.4. Polynomials in the affine and projective plane

Geometry is largely the examination of objects and invariants under the action of transformation groups. Eventually, we will encounter the group of translations, rotations around the origin, Euclidean motions and projective transformations. Each framework has its optimal representation for objects such as points, polynomials or zero sets for visualizing results: Point-coordinates may be homogeneous in one case and inhomogeneous in the next or vice versa.

We briefly sketch the basic concepts and terminology of projective geometry. For more detailed and formal descriptions see [39]. Again let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then $\mathbb{K}\mathbb{P}^2$ denotes the projective plane over \mathbb{K} . The set of points (and due to the duality also the set of lines) is given by $\frac{\mathbb{K}^3 \setminus \{(0,0,0)\}}{\mathbb{K} \setminus \{0\}}$. Formally a point is an equivalence class $[p]$ ($p \neq (0,0,0)$) with $[p] = \{q \in \mathbb{K}^3 \mid q = \lambda p \wedge \lambda \in \mathbb{K} \setminus \{0\}\}$. In what follows, we will mostly omit the brackets $[]$ and write p instead of $[p]$. Non-zero-multiples λp will be identified with p and p is called a point in **homogeneous** point coordinates.

The canonical embedding of the affine plane \mathbb{K}^2 in $\mathbb{K}\mathbb{P}^2$ is given by

$$emb : \begin{cases} \mathbb{K}^2 & \longrightarrow \mathbb{K}\mathbb{P}^2 \\ (x, y) & \longmapsto (x, y, 1) \end{cases} . \quad (2.1)$$

The points at infinity are $\mathbb{K}\mathbb{P}^2 \setminus emb(\mathbb{K}^2) = \{(x, y, h) \in \mathbb{K}\mathbb{P}^2 \mid h = 0\}$. They are all contained in a line, the line at infinity l_∞ . In projective geometry any line can be singled out as the line at infinity.

In Euclidean geometry, we know that a point $(x, y) \in V(f)$ for some $f \in \mathbb{K}[X, Y]$ if $f(x, y) = 0$. We are now interested in the projective equivalent: Let

$$f = f_{(0)} + f_{(1)} + \dots + f_{(d)}$$

with homogeneous polynomials $f_{(n)}$ of degree n . The $f_{(n)}$ are so called **homogenous part of degree n** of f . Then $F \in \mathbb{K}[X, Y, H]$ with

$$F(x, y, h) = h^d f_{(0)}(x, y) + h^{d-1} f_{(1)}(x, y) + \dots + f_{(d)}(x, y)$$

is the homogenization of f . For $h \neq 0$ we have

$$F(x, y, h) = h \cdot f\left(\frac{x}{h}, \frac{y}{h}\right) \quad \text{and} \quad f(x, y) = F(x, y, 1) .$$

In this treatise we study plane curves. Therefore we use inhomogenous polynomials $f \in \mathbb{K}[X, Y]$ and their homogenized equivalents $f = f^{hom} \in \mathbb{K}[X, Y, H]$. In order to show that we are talking about the *pair* of inhomogenous polynomial and homogenized equivalent, we denote both by the same letter. The difference is in the number of arguments: $f(x, y)$ versus $f(x, y, h)$. (This notational technique is known as identifier-overload in computer science.)

2.1.5. Polynomials as elements of a vector space

Before we study the curves themselves, let us dwell on polynomials as elements of a vector space (compare with [22]): Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $d \in \mathbb{N}$. Then the set of polynomials $f \in \mathbb{K}[X, Y]$ up to degree d is denoted by $\mathfrak{P}_d(\mathbb{K})$. The set of all homogeneous polynomials $g \in \mathbb{K}[X, Y, H]$ of degree d is $\mathfrak{P}_d^{hom}(\mathbb{K})$. The polynomial $0 \in \mathbb{K}[X, Y, H]$ is homogeneous of any degree and thus contained in $\mathfrak{P}_d^{hom}(\mathbb{K})$. Now $\mathfrak{P}_d(\mathbb{K})$ and $\mathfrak{P}_d^{hom}(\mathbb{K})$ constitute a $\binom{d+2}{2}$ dimensional vector space. The standard (or monomial) basis in $\mathfrak{P}_d(\mathbb{K})$ is

$$1, \quad x, y, \quad x^2, xy, y^2, \quad \dots, \quad x^d, x^{d-1}y, \dots, y^d .$$

The monomial basis in $\mathfrak{P}_d^{hom}(\mathbb{K})$ can be obtained by multiplying the inhomogeneous basis elements from above by suitable powers of h such that they become homogeneous of degree d in x, y, h :

$$h^d, \quad xh^{d-1}, yh^{d-1}, \quad x^2h^{d-2}, xyh^{d-2}, y^2h^{d-2}, \quad \dots, \quad x^d, x^{d-1}y, \dots, y^d .$$

The number of basis elements determines the dimension. It is $\sum_{n=0}^d n + 1 = \frac{(d+2)(d+1)}{2}$, which equals $\binom{d+2}{2}$ as stated above. Thus, using the monomial basis, f and its (de-)homogenized equivalent can be represented by

$$f(x, y) = \sum_{k=0}^d \sum_{l=0}^{d-k} A_{kl} x^k y^l \quad \text{or} \quad f(x, y, h) = \sum_{k=0}^d \sum_{l=0}^{d-k} A_{kl} x^k y^l h^{d-k-l} .$$

In both cases, the coefficients can be stated by the use of two indices k and l . The polynomials can be written as

$$f = \sum_{k+l \leq d} A_{kl} b_{kl}$$

with a basis $b_{00}, b_{10}, b_{01}, \dots, b_{d0}, b_{d-1 0}, \dots, b_{0d}$ of the corresponding space of polynomials. In what follows we will always write our polynomials as a linear combination of basis vectors using coefficients with two indices. This allows a clearer view on the structure with interesting relations between transformations of the plane and corresponding transformations on the coefficients (see for example Section 4.1).

We will now introduce an order on the formal coefficients of polynomials, using both of their two indices.

Definition 2.3 Let $\Omega = \{\omega_{kl} \mid k, l \in I\}$ be a set of formal variables with two indices originating from an index set $I \subset \mathbb{N}$. Then \succ defines a total **order** on Ω by

- $\omega_{kl} \succ \omega_{pq}$ if $k + l > p + q$ and
- $\omega_{kl} \succ \omega_{pq}$, whenever $k + l = p + q$ and $k > p$

for all $\omega_{kl}, \omega_{pq} \in \Omega$. Thus equality holds only if $k = p$ and $l = q$.

Ω is separated into **blocks** by $block_n(\Omega) := \{\omega_{kl} \in \Omega \mid k + l = n\}$.

Two elements $\omega_{kl} \succ \omega_{pq}$ are called **succeeding** or **succeeding pair** if there is no $\omega_{ab} \in \Omega$ with $\omega_{kl} \succ \omega_{ab}$ and simultaneously $\omega_{ab} \succ \omega_{pq}$.

Now we are able collect all coefficients of f and put them into a vector, the **coefficient vector** of f . We do this in such a way that the elements are arranged according to \succ . Starting with the smallest and ending with the largest A_{kl} with respect to \succ , we get

$$A^T = (\underbrace{A_{00}}_{\text{block}_0(A)}, \underbrace{A_{10}, A_{01}}_{\text{block}_1(A)}, \underbrace{A_{20}, A_{11}, A_{01}}_{\text{block}_2(A)}, \dots, \underbrace{A_{d0}, A_{d-1\ 0}, \dots, A_{0,d}}_{\text{block}_d(A)}).$$

Thereby A^T denotes the transpose of the coefficient vector A . The order of the A_{kl} corresponds to the printed order from above:

$$\begin{aligned} f(x, y) &= A^T \cdot (1, x, y, x^2, xy, y^2, \dots, x^d, x^{d-1}y, \dots, y^d)^T =: A^T X \quad \text{and} \\ f(x, y, h) &= A^T \cdot (h^d, xh^{d-1}, yh^{d-1}, \dots, x^d, x^{d-1}y, \dots, y^d)^T =: A^T X^{\text{hom}}. \end{aligned}$$

This way we get an isomorphism $\iota : \mathfrak{P}_d(\mathbb{K}) \rightarrow \mathbb{K}^{\binom{d+2}{2}}$, which maps an $f(x, y) = A^T X$ onto A and $\iota^{\text{hom}} : \mathfrak{P}_d^{\text{hom}}(\mathbb{K}) \rightarrow \mathbb{K}^{\binom{d+2}{2}}$, which maps an $f(x, y, z) = A^T X^{\text{hom}}$ onto A . As a result we have the isomorphism $\mathfrak{P}_d(\mathbb{K}) \sim \mathfrak{P}_d^{\text{hom}}(\mathbb{K})$.

The block-structure of the coefficient vector A can be transferred to the formal vector X : $\text{block}_n(X) = \{x^k y^l \mid k+l=n\}$. The transfer to X^{hom} is analogous. We will use the block-structure to introduce auxiliary separating lines in vectors and matrices. This way the structure of transformations will be more clearly visible.

2.2. Curves

2.2.1. Curves as zero sets

Now, with the basics about polynomials and zero sets at hand, we can define curves with

Definition 2.4 Let \mathbb{K} be either \mathbb{R} or \mathbb{C} .

A subset $\mathcal{C} \subset \mathbb{K}^2$ is called a *real* respectively *complex affine algebraic plane curve* if there is a non-constant polynomial $f \in \mathbb{K}[X, Y] \setminus \mathbb{K}$ of degree $d \in \mathbb{N}$ such that $\mathcal{C} = V(f)$. A subset $\mathcal{C} \subset \mathbb{K}\mathbb{P}^2$ is called a *real* respectively *complex projective algebraic plane curve* if there is a non-constant polynomial $f \in \mathbb{K}[X, Y, H] \setminus \mathbb{K}$ of degree $d \in \mathbb{N}$ such that $\mathcal{C} = V(f)$.

Generally, we use affine algebraic curves by means of handling projective algebraic curves: When results are more clearly visible and easier to read without homogenizing components, we use the affine standpoint. Nevertheless, we do not forget the points at infinity, i.e. points with homogeneous coordinates $(x, y, 0) \neq (0, 0, 0)$, which lead us to the projective extension of the affine plane. Switching from affine to projective view is possible by using the *emb*-function from (2.1), i.e. by homogenization. Restricting the projective view to the affine is done by selecting a line to be l_∞ , e.g. $(0, 0, 1)$, and substituting x by $\frac{x}{h}$, y by $\frac{y}{h}$ and h by 1.

As curves are defined with the help of zero sets, we can transfer the terminology from the last section to curves by

Definition 2.5 A curve \mathcal{C} is *irreducible* if it is not a union of two different plane curves.

With Study's Lemma and our Corollary 2.4 we get

Lemma 2.5 Let \mathcal{C} be a curve with $\mathcal{C} = V(f)$ for some (homogeneous) polynomial $f \in \mathbb{C}[X_1, \dots, X_k]$. Let $f = f_1^{d_1} \cdot \dots \cdot f_r^{d_r}$ be a decomposition of f into irreducible (homogeneous) polynomials $f_1, \dots, f_r \in \mathbb{C}[X_1, \dots, X_k]$. Then

- \mathcal{C} is irreducible if and only if f is a power of an irreducible polynomial, i.e. if $r = 1$.
- \mathcal{C} has a unique decomposition into irreducible curves $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r$ with $\mathcal{C} = \bigcup_{n=1}^r \mathcal{C}_n$. With a suitable numeration we also get $\mathcal{C}_n = V(f_n)$. The \mathcal{C}_n determine the f_n up to a multiplicative constant. The \mathcal{C}_n are called **irreducible components** of \mathcal{C} .

For the proof we refer to [11] or [20].

What we get by this lemma is that curves may be identified with a certain class of polynomials: If $f = \prod_{n=1}^r f_n^{d_n}$ is a decomposition into irreducibles, then

$$\mathcal{C} = V(f) = V(\lambda \prod_{n=1}^r f_n)$$

with $\lambda \in \mathbb{K}$. Let $[f] := \{\lambda f \mid \lambda \in \mathbb{K}\}$. Then the set of all plane algebraic curves is equivalent to the set of all $[f]$, where f is a square-free product of irreducible polynomials.

Definition 2.6 Let \mathcal{C} be a curve and $\mathcal{C} = \bigcup_{n=1}^r \mathcal{C}_n$ its decomposition into irreducible curves. Further let the f_n be the up to a constant multiple uniquely determined irreducible polynomials with $\mathcal{C}_n = V(f_n)$. Then f with $f(x, y) = \prod_{n=1}^r f_n(x, y) = 0$ is called an **associated curve-equation** with respect to \mathcal{C} and f an **associated polynomial**. This definition is analogous for homogenous polynomials. $\deg(f)$ is called the **degree** of $\mathcal{C} = V(f)$.

If f is an associated polynomial of $\mathcal{C} = V(f)$, we will often speak of f as *curve* and identify all λf for all non-zero constants λ .

Now we are able to associate coefficient vectors of polynomials to curves and introduce curve coefficients: Let $\mathcal{C} = V(f)$. Then the polynomial f with $f(x, y) = A^T X$ has a coefficient vector A (see Section 2.1.5). With $[f] \sim \mathcal{C}$ we can associate the class $[A] = \{\lambda A \mid \lambda \in \mathbb{K}\}$ with \mathcal{C} , but we will not name A a curve coefficient vector: Curve coefficients shall be scaled coefficients of polynomials. The predefined scaling factors will be the so-called multinomial coefficients. This scaling will make calculations much easier. The connection between curves and tensors (see Section 3.1) will be much clearer in a much more natural way. Additionally, multinomial coefficients make interesting structural properties very easily accessible. The usage of these scaling factors could be exploited also in the field of pattern analysis. However, a direct identification of coefficients of polynomials and curves is common (for example see [26], [29]).

Definition 2.7

Let $p, q, r \in \mathbb{N}$. The numbers $m_{pqr} := \frac{(p+q+r)!}{p! \cdot q! \cdot r!}$ are called **multinomial coefficients**. Let $\mathcal{C} = V(f)$ be a curve of degree d for some (associated) polynomial f . Further let $f = \sum_{k+l \leq d} A_{kl} b_{kl}$ with the basis vectors b_{kl} . Then each coefficient A_{kl} can be associated with a multinomial coefficient $m_{kl} := m_{k \ l \ (d-k-l)} = \frac{d!}{k! \cdot l! \cdot (d-k-l)!}$, such that $A_{kl} = m_{kl} a_{kl}$. Thus

$$f(x, y) = A^T X = (ma)^T X ,$$

where $m = m_{(d)} = \text{diag}(m_{00}, m_{10}, m_{01}, \dots, m_{d0}, m_{d-10}, \dots, m_{0d})$ is a diagonal matrix with multinomial coefficients as entries. The corresponding vector a is called (associated) multinomially scaled **curve coefficient vector** of $\mathcal{C} = V(f)$.

$$[a] := \left\{ v \in \mathbb{K}^{\binom{d+2}{2}} \mid v = \lambda a \wedge \lambda \in \mathbb{K} \right\}$$

are the (associated) **coefficients** of $\mathcal{C} = V(f)$.

Through this definition we finally have

$$\mathcal{C} = V(f) \sim [f] \sim [A] = [m \cdot a] \sim [a] .$$

The diagonal matrix m is invertible by construction. Thus it constitutes an isomorphism between the space of all coefficient vectors of polynomials and the space of all multinomially scaled coefficient vectors of polynomials. As a consequence, the set of all curves is equivalent to the set of all $[a]$, with $[a]$ being the set of multinomially scaled coefficient vectors of square-free products of irreducible polynomials as stated above.

For example, take $\mathcal{C} = V(f)$ with $f(x, y) = (y - x^2)$. Thus \mathcal{C} is the standard parabola and $f(x, y) = A^T X = (0, 0, 1, -1, 0, 0) \cdot (1, x, y, x^2, xy, y^2)^T$. Introducing the multinomial coefficients, we get $f(x, y) = (0, 0, \frac{1}{2}, -1, 0, 0) \cdot m \cdot X$. Because $\mathcal{C} = V(\lambda f)$ for all $\lambda \neq 0$, we introduce equivalence classes and \mathcal{C} corresponds to $[f]$ and $[(0, 0, \frac{1}{2}, -1, 0, 0)]$. As already mentioned above, we mostly omit the brackets $[]$ during notation.

Summarizing our results, we have an implicit representation $f(x, y) = 0$ for a curve \mathcal{C} via the zero set $\mathcal{C} = V(f)$ of a polynomial f . One of the benefits of this representation is that the union $V(f) \cup V(g)$ of two curves $\mathcal{C}_1 = V(f)$ and $\mathcal{C}_2 = V(g)$ is just $V(f \cdot g)$.

2.2.2. Effects of transformations of the plane on curves

In geometry we examine the behavior of objects under the action of certain transformation groups. Here we want to clarify the principal effects transformations have on curves or on their associated polynomials.

Let $\mathbb{K}\mathbb{P}^2$ with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a projective plane containing a curve $\mathcal{C} = V(f)$. We use the projective viewpoint to have a general framework. For affine observations dehomogenization can be used. A point in the projective plane can be described by (x, y, h) . (We omit the brackets “[]” symbolizing equivalence classes.)

In case we want to transform the plane, we have to distinguish between the transformation of the plane itself and the corresponding coordinate transformation: The **transformation of the plane** determines the image of the used basis of the space. Identifying the basis elements with unit-vectors, this transformation can be identified with a matrix P . A **coordinate transformation** is an automorphism describable by its action on the point-coordinates:

$$\text{transformation} : \begin{cases} \mathbb{K}\mathbb{P}^2 & \longrightarrow \mathbb{K}\mathbb{P}^2 \\ (x, y, h)^T & \longmapsto Q \cdot (x, y, h)^T = (\tilde{x}, \tilde{y}, \tilde{z})^T \end{cases} \quad (2.2)$$

with an invertible matrix $Q = (q_{kl})_{1 \leq k, l \leq 3} \in \mathbb{K}^{3 \times 3}$ acting on the points of $\mathbb{K}\mathbb{P}^2$. The connection between P and Q is known from linear algebra (cf. [22]): $Q = (P^{-1})^T$. Due to the equivalence of (x, y, h) and $\lambda(x, y, h)$ for all $\lambda \in \mathbb{K} \setminus \{0\}$, the transformations Q

and μQ with $\mu \in \mathbb{K} \setminus \{0\}$ act in the same way. Thus we can confine our study to the actions of equivalence classes

$$[Q] := \{ N \in \mathbb{K}^{3 \times 3} \mid N = \mu Q \wedge \mu \in \mathbb{K} \setminus \{0\} \} .$$

Again we will omit the brackets and identify all elements of an equivalence class. A representation system of our transformations is, for example, the set of all three by three matrices with determinant one.

Let us look at how polynomials behave under transformations of the plane. Therefore let $f(x, y, h) = \sum_{k+l=0}^d A_{kl} x^k y^l h^{d-k-l} = A^T X$ be a polynomial of degree d in monomial basis-representation. If we want to transform the plane according to (2.2), we have to examine the monomial expressions $\tilde{x}^k \tilde{y}^l \tilde{z}^{d-k-l}$: Due to the linearity of (2.2) we get

$$\tilde{x}^k \tilde{y}^l \tilde{h}^{d-k-l} = (q_{11}x + q_{12}y + q_{13}h)^k (q_{21}x + q_{22}y + q_{23}h)^l (q_{31}x + q_{32}y + q_{33}h)^{d-k-l},$$

which in itself is a linear expression in the $x^r y^s h^{d-r-s}$ with $r, s \in \mathbb{N}$ and $r+s \leq d$. Consequently, a transformation of the plane, like the one in (2.2), induces a linear transformation of the vector X containing the monomials by $L = L_Q$: namely $X \mapsto LX$ for a $\binom{d+2}{2} \times \binom{d+2}{2}$ matrix L . Per assumption Q is invertible and thus also L .

Returning to our polynomial f we can see that

$$f(x, y, h) = A^T X = A^T L^{-1} LX = \underbrace{(L^{-T} A)^T}_{=\tilde{A}^T} \tilde{X} = \tilde{f}(\tilde{x}, \tilde{y}, \tilde{h}) \quad (2.3)$$

with L^{-T} being the transposed inverse of L . Thus the transformation of X and A is **contragredient**, i.e. L , acting on X , counteracts on A . If we transform the curve $\mathcal{C} = V(f)$ by some $Q' \in [Q]$, then there is a $\mu \neq 0$ with $Q' = \mu Q$. The monomials satisfy $\tilde{x}'^k \tilde{y}'^l \tilde{h}'^{d-k-l} = \lambda^d \tilde{x}^k \tilde{y}^l \tilde{h}^{d-k-l}$ and thus $L' = \lambda^d L$. This means that any transformation $Q' \in [Q]$ consistently maps $[a] \xrightarrow{Q'} [b]$ because $V(f) = V(\lambda f)$ for any non-zero λ . Thus it suffices to study a single curve coefficient vector a under transformations of the plane, rather than a whole set $[a]$ of vectors. According to Definition 2.7, the curve coefficients contained in a satisfy $A = ma$, where m contains multinomial coefficients. Then

$$f(x, y, h) = (ma)^T X = (ma)^T L^{-1} LX = (m \underbrace{(m^{-1} L^{-T} ma)}_{=\tilde{a}})^T \tilde{X} = \tilde{f}(\tilde{x}, \tilde{y}, \tilde{h}) . \quad (2.4)$$

and the curve coefficients are transformed by $m^{-1} L^{-T} m =: L_m^{-T}$. Thus L and L_m are similar matrices. We will say that Q and also L_m^{-T} transform a curve.

Summarizing this we can say that transformations T of the plane induce a change L in the basis $X \mapsto \tilde{X}$ of the associated polynomial corresponding to a contragredient change L_m^{-T} in the curve coefficients.

A further very important concept concerning curves and transformations is the one of normal forms. A normal form in our sense is defined by

Definition 2.8 *Let G be a transformation group of the plane, e.g. the group of all rotations or the group of all Euclidean mappings. Then G introduces an equivalence relation \sim on the set of all algebraic curves of a fixed degree d . For two curves f and g : $f \sim g$ if and only if there is a transformation $\tau \in G$ such that τ maps f onto g . Thus \sim subdivides the set of our curves into classes. A designated representation system with respect to this transformation group will be called a **normal form**. A curve is in normal form if it is given as one of the designated representatives.*

We present normal forms with respect to various transformation groups in the next chapter and show how a given curve must be transformed such that it is in normal form. When it comes to recognizing a curve, normal forms are very helpful: We can take a given curve and apply transformations such that the given curve is in normal form. Afterwards we can look for such a curve in a database and associate a corresponding name with it, such as conchoid, limaçon or something else. Given a zero set, one of the ultimate goals of a curve recognition algorithm would be an output like:

“The curve is a limaçon with parameters $r = 0.54$ and $s = 0.72$. A limaçon is given by the equation $(x^2 + y^2 - 2rx)^2 = s^2(x^2 + y^2)$.”

Optionally, the output is amended by the curve coefficients and a declaration of the relative position of the given curve with respect to the displayed equation. Potentially existing symmetries may also be of interest.

The term normal form has been defined such that it stays the same even if the respective transformation group acts on a curve. The normal form is invariant so to say. Another way of recognizing curves is to compare invariants of curves with database entries to attain an output like the one above. Invariants in this sense are defined by

Definition 2.9 Let f be a polynomial with coefficient vector A Let $\mathcal{C} = V(f)$ be a curve with coefficient vector $A = m \cdot a$ of f . Further let a transformation of the plane P with determinant $\Delta = \det(P)$ transform \mathcal{C} into $\tilde{\mathcal{C}} = V(\tilde{f})$ with coefficient vector $\tilde{A} = m \cdot \tilde{a}$ of \tilde{f} . Let \mathcal{I} be a function satisfying

$$\mathcal{I}(\tilde{a}) = \Delta^g \mathcal{I}(a) .$$

Then \mathcal{I} is called an **invariant** of \mathcal{C} . g is called the **weight** of the invariant. If $g = 0$, we speak of \mathcal{I} as **absolute invariant** and if $g \neq 0$, \mathcal{I} is called a **relative invariant**.

It can be shown (see [16], [25] or [55]) that the study of invariants can be reduced to examining homogenous integer rational algebraic \mathcal{I} 's. In addition a generalization in the direction of arbitrary factors λ instead of Δ^g is not necessary. When we restrict ourselves to transformations Q with $\det(Q) = 1$, then we are looking for invariants of the form $\mathcal{I}(\tilde{a}) = \mathcal{I}(a)$. This reduction is possible as stated above. This case resembles a common definition of invariants, used e.g. by Hurwitz in [27].

2.2.3. Rotating real curves around the origin

In this section we will exemplarily show how transformations act on curves. Therefore we take real plane curves, i.e. curves with associated polynomials f over \mathbb{R} . f shall be in monomial representation.

A rotation around the origin can be described very easily: Abbreviating $\cos(\varphi) = c$ and $\sin(\varphi) = s$, a point $p = (x, y, h)$ with homogenizing component h is rotated by applying the map

$$\begin{pmatrix} x \\ y \\ h \end{pmatrix} \mapsto \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ h \end{pmatrix} . \tag{2.5}$$

The transformation matrix $Q = Q(\varphi)$ has a determinant equal to one. The question now is, what happens to the polynomial $f(x, y, h)$ of the curve $\mathcal{C} = V(f)$, when the plane is rotated. As we have seen, a curve of degree d is given by a zero set of a polynomial equation:

$$f(x, y, h) = \sum_{k+l=0}^d m_{kl} a_{kl} \cdot x^k y^l h^{d-k-l} = (ma)^T X = 0 . \quad (2.6)$$

According to Equation (2.4) we have $f(x, y, h) = (m(L_m^{-T}a))^T \tilde{X} = (m\tilde{a})^T \tilde{X}$ with $\tilde{X} = LX$. This means that the rotated curve $\tilde{f}(x, y, h) = f_\varphi(x, y, h) = (m\tilde{a})^T X$. For a given curve coefficient vector a we are interested in the shape of $L_m^T = L_m^T(\varphi)$. We can extract L_m^T directly by observing that in the affine setup

$$\begin{aligned} \tilde{f}(x, y) &= \sum_{k+l=0}^d m_{kl} a_{kl} (cx + sy)^k (-sx + cy)^l \\ &= \sum_{k+l=0}^d m_{kl} \tilde{a}_{kl} x^k y^l . \end{aligned} \quad (2.7)$$

If we scaled the transformation matrix in (2.5) by λ , then we would have to scale all \tilde{a}_{kl} by λ^d .

Theorem 2.6 *If Equation (2.7) holds, then*

$$\tilde{a}_{kl} = \sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} (-1)^{k-p} \cdot c^{l+(p-q)} \cdot s^{k-(p-q)} \cdot a_{(p+q), ((k+l)-(p+q))} . \quad (2.8)$$

PROOF Expanding the product $(cx + sy)^k (-sx + cy)^l$, suitably changing summation indices a few times and recollecting the coefficients of monomial terms proves the theorem. A more detailed calculation can be found in the appendix (see Appendix A.1). \square

A first interesting property can be seen in Equation (2.8): A coefficient \tilde{a}_{kl} only depends on the coefficients a_{uv} of the same block, i.e. $u + v = k + l$. A second interesting property is contained in the coefficients: Using the generalized binomial formula we get

$$\begin{aligned} (c - s)^k (c + s)^l &= \left(\sum_{p=0}^k \binom{k}{p} c^p s^{k-p} (-1)^{k-p} \right) \cdot \left(\sum_{q=0}^l \binom{l}{q} s^q c^{l-q} \right) \\ &= \sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} (-1)^{k-p} c^{l+(p-q)} s^{k-(p-q)} . \end{aligned}$$

This corresponds directly to Equation (2.8).

We can rewrite Equation (2.8) in a much more descriptive way using matrix-vector notation. Doing so we can expect a block diagonal structure with entries corresponding linewise to the terms of the expanded version of the product $(c - s)^k (c + s)^l$. Indicating the block-affiliations by auxiliary lines, we have in fact

$$\left(a_{00} \mid a_{10} \ a_{01} \mid a_{20} \ a_{11} \ a_{02} \mid a_{30} \ a_{21} \ a_{12} \ a_{03} \mid \dots \right)^T \mapsto L_m^{-T} a = \quad (2.9)$$

$$\underbrace{\left(\begin{array}{c|c|c|c|c}
 1 & & & & \\
 \hline
 & c & -s & & \\
 & s & c & & \\
 \hline
 & & & c^2 & -2cs & s^2 \\
 & & & cs & c^2 - s^2 & -cs \\
 & & & s^2 & 2cs & c^2 \\
 \hline
 & & & & c^3 & -3c^2s & 3cs^2 & -s^3 \\
 & & & & c^2s & c^3 - 2cs^2 & s^3 - 2c^2s & cs^2 \\
 & & & & cs^2 & 2c^2s - s^3 & c^3 - 2cs^2 & -c^2s \\
 & & & & s^3 & 3cs^2 & 3c^2s & c^3 \\
 \hline
 & & & & & & & \ddots
 \end{array} \right) \cdot a} = L_m^{-T}$$

Having a closer look at this transformation we can see that

$$L_m^{-T} = m^{-1} L_m m . \quad (2.10)$$

This is due to the fact that L_m^{-T} is inverted by writing down the transformation matrix corresponding to a rotation by $-\varphi$, rather than φ . Thus inversion is equivalent to substituting s by $-s$, because $\cos(-\varphi) = \cos(\varphi)$ and $\sin(-\varphi) = -\sin(\varphi)$. Transposition of L_m^T sends the signs back to where they were in L_m^{-T} . The change in the coefficients of the matrix-components induced by transposition may be undone by multiplication by m and m^{-1} , respectively. Thus Equation (2.10) holds.

Let us have a short look at the coefficient vector A of the curve-describing polynomial: It incorporates the multinomial coefficients and thus $A = ma$. Under rotation of the plane around the origin, A changes according to $\tilde{A} = L^{-T} A$ and according to Equation (2.10) $L^{-T} = L_m$. This can also be shown analogously to Theorem 2.6.

Remark 2.1 *Let us briefly validate the third block in the diagonal of L_m^{-T} shown in (2.9) without the formalism of multiple sums from Appendix A.1: According to Equation 2.5 we have*

$$\begin{aligned}
 \tilde{x}^2 &= (cx - sy)^2 &= c^2x^2 - 2csxy + s^2y^2 \\
 \tilde{x}\tilde{y} &= (cx - sy)(sx + cy) &= csx^2 + (c^2 - s^2)xy - csy^2 \\
 \tilde{y}^2 &= (sx + cy)^2 &= s^2x^2 + 2csxy + c^2y^2
 \end{aligned}$$

and thus

$$\begin{pmatrix} \tilde{x}^2 \\ \tilde{x}\tilde{y} \\ \tilde{y}^2 \end{pmatrix} = \begin{pmatrix} c^2 & -2cs & s^2 \\ cs & c^2 - s^2 & -cs \\ s^2 & 2cs & c^2 \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix} =: L_{(2)} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix} .$$

The coefficients of $\text{block}_2(a)$, i.e. a_{20} , a_{11} and a_{02} , transform among themselves. According to (2.4):

$$\underbrace{(\tilde{a}_{20}, \tilde{a}_{11}, \tilde{a}_{02})}_{=: m_{(2)}} \begin{pmatrix} m_{20} & & \\ & m_{11} & \\ & & m_{02} \end{pmatrix} \begin{pmatrix} \tilde{x}^2 \\ \tilde{x}\tilde{y} \\ \tilde{y}^2 \end{pmatrix} = (\tilde{a}_{20}, \tilde{a}_{11}, \tilde{a}_{02}) m_{(2)} L_{(2)} m_{(2)}^{-1} m_{(2)} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix} ,$$

which gives us

$$\begin{pmatrix} \tilde{a}_{20} \\ \tilde{a}_{11} \\ \tilde{a}_{02} \end{pmatrix} = m_{(2)}^{-1} L_{(2)}^{-T} m_{(2)} \begin{pmatrix} a_{20} \\ a_{11} \\ a_{02} \end{pmatrix}$$

with

$$m_{(2)}^{-1} L_{(2)}^{-T} m_{(2)} = m_{(2)}^{-1} \begin{pmatrix} c^2 & -cs & s^2 \\ 2cs & c^2 - s^2 & -2cs \\ s^2 & cs & c^2 \end{pmatrix} m_{(2)} = \begin{pmatrix} c^2 & -\frac{m_{11}}{m_{20}} cs & \frac{m_{02}}{m_{20}} s^2 \\ \frac{2m_{20}}{m_{11}} cs & c^2 - s^2 & -\frac{2m_{02}}{m_{11}} cs \\ \frac{m_{20}}{m_{02}} s^2 & \frac{m_{11}}{m_{02}} cs & c^2 \end{pmatrix}$$

No matter what the degree of the curve is, $m_{20} = m_{02}$ and $\frac{m_{11}}{m_{20}} = \frac{m_{11}}{m_{02}} = 2$ holds in any case. Thus

$$\begin{pmatrix} \tilde{a}_{20} \\ \tilde{a}_{11} \\ \tilde{a}_{02} \end{pmatrix} = \begin{pmatrix} c^2 & -2cs & s^2 \\ cs & c^2 - s^2 & -cs \\ s^2 & 2cs & c^2 \end{pmatrix} \begin{pmatrix} a_{20} \\ a_{11} \\ a_{02} \end{pmatrix}$$

2.2.4. Excursus on tetrahedral coefficient structure

This section is dedicated to the beauty of mathematics. The transformation formulas (2.8) and (2.9) exhibit a lot of structure that we should dwell on for a while. Furthermore, the attributes shown will reappear in other transformation laws analogously (cf. Section 2.3).

line k of $block_3(L_m^{-T})$	coefficients (weighted)	sub-line weight	coefficients (not weighted)
0	1	1	1
	-3	-3	1
	3	3	1
	-1	-1	1
1	1 1	1	1 1
	-2 -2	-2	1 1
	1 1	1	1 1
2	1 2 1	1	1 2 1
	-1 -2 -1	-1	1 2 1
3	1 3 3 1	1	1 3 3 1

Table 2.1.: Tetrahedral coefficient structure of $block_3(L_m^{-T})$

Let $block_n(L_m^{-T})$ denote the n -th block in the diagonal of L_m^{-T} , starting with the zeroth block $block_0(L_m^{-T}) = (1)$. When it comes to line-numbering within blocks we will also start counting at zero and line $k = 0$ denotes the topmost line.

At first we can see that $block_n(L_m^{-T})$ is homogenous of degree n in c and s . The coefficients of the products of c and s exhibit a tetrahedral structure of binomially weighted binomial coefficients: For example take $block_3(L_m^{-T})$ from Expression (2.9). The coeffi-

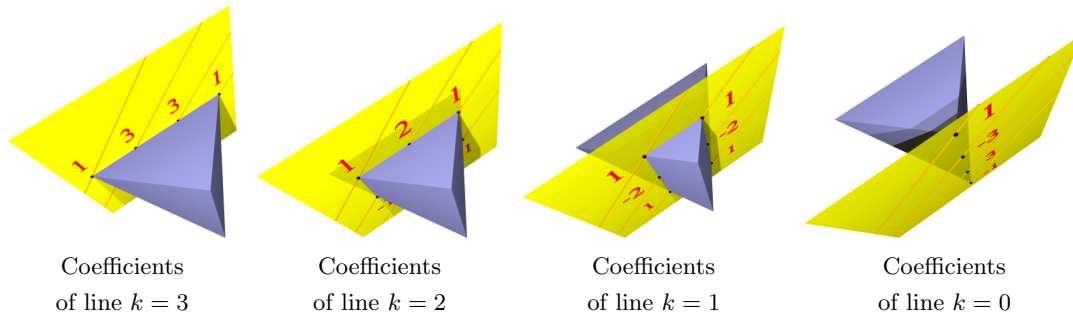


Figure 2.1.: Tetrahedron sliced four times

cient structure is

$$\begin{pmatrix} 1 & -3 & 3 & -1 \\ 1 & 1, -2 & 1, -2 & 1 \\ 1 & 2, -1 & 1, -2 & -1 \\ 1 & 3 & 3 & 1 \end{pmatrix} \text{ given by } \begin{pmatrix} \mathbf{1}c^3 & -\mathbf{3}c^2s & \mathbf{3}cs^2 & -\mathbf{1}s^3 \\ \mathbf{1}c^2s & \mathbf{1}c^3-\mathbf{2}cs^2 & \mathbf{1}s^3-\mathbf{2}c^2s & \mathbf{1}cs^2 \\ \mathbf{1}cs^2 & \mathbf{2}c^2s-\mathbf{1}s^3 & \mathbf{1}c^3-\mathbf{2}cs^2 & -\mathbf{1}c^2s \\ \mathbf{1}s^3 & \mathbf{3}cs^2 & \mathbf{3}c^2s & \mathbf{1}c^3 \end{pmatrix}. \quad (2.11)$$

Having a closer look at the rows of this coefficient structure we can detect binomial coefficients: In the bottom line we have the coefficient-structure as in $(a+b)^3 = \mathbf{1}a^3 + \mathbf{3}a^2b + \mathbf{3}ab^2 + \mathbf{1}b^3$. One line above we have twice $(1, 2, 1)$ but the second one is scaled by -1 and shifted by one position. We say that the two copies of $(1, 2, 1)$ form two sub-lines of that line $k = 2$. These coefficients are binomial coefficients corresponding to $(a+b)^2 = \mathbf{1}a^2 + \mathbf{2}ab + \mathbf{1}b^2$. This pattern may be continued. In the line $k = 1$, one line above, we have $(1, 1)$ three times, each shifted by one position. Additionally, the second copy, i.e. the second sub-line, is scaled by -2 . The total pattern is displayed in Table 2.1.

It is interesting to note that the weights of the sub-lines are themselves binomial coefficients: They belong to the coefficient-structure of $(a-b)^{d-k}$. Thus we have a binomially weighted binomial coefficient-structure as stated above. In general we can associate line k of any $block_n(L_m^{-T})$ with $d+1-k$ times the coefficients of $(a+b)^k$, weighted by the coefficients of $(a-b)^{d-k}$.

Let us once more look at the coefficient-pattern in Table 2.1. On the one hand we have a triangular structure of the binomial coefficients, corresponding to Pascal's Triangle. On the other hand we have certain lines of Pascal's Triangle more than once: In line $k = 0$ there is exactly the same number of sub-lines as we have coefficients in line $k = 3$. Now embedding the two-dimensional Pascal-Triangle in three-space, we can arrange the multiple copies of the corresponding line in the "surplus" dimension: The copies are positioned above each other. This way we get a three-dimensional *tetrahedral* coefficient pattern (compare with Table 2.1 and Figure 2.1).

To visualize the coefficients we slice our tetrahedron with the help of four parallel planes. Thereby the first and the last plane shall intersect the tetrahedron in two opposing edges and the coefficients of a line in $block_3(L_m^{-T})$ shall be contained in one plane. The result can be seen in Figure 2.1.

Summarizing the above, we got a tetrahedral structure by associating the lines in a block by powers of $(a+b)$ with weights according to powers of $(a-b)$. Likewise, we could have

line of $block_3(L_m^{-T})$	coefficients (weighted)	sub-line weight
0	1 -3 3 -1	1
1	1 -2 1	1
	1 -2 1	1
2	1 -1	1
	2 -2	2
	1 -1	1
3	1	1
	3	3
	3	3
	1	1

Table 2.2.: Tetrahedral coefficient structure of $block_3(L_m^{-T})$

done it the other way round arriving at a tetrahedron which is dual to the one above. Associating the lines of $block_3(L_m^{-T})$ by coefficients of powers of $(a - b)$ with weights according to coefficients of $(a + b)$ we get the pattern of table 2.2. The corresponding images of the tetrahedron will be omitted here. It can be obtained by rotating the tetrahedrons from Figure 2.1 by $\frac{\pi}{2}$. Thereby the rotation-axis is perpendicular to the displayed planes and passes through the center of the tetrahedrons.

Let us switch back to our first pattern in Table 2.1 and Figure 2.1. The affiliation of coefficients to a column in the Pattern 2.1 are shown by the auxiliary lines on the particular planes.

This brings us to the idea of scanning $block_3(L_m^{-T})$ in Expression (2.11) columnwise. In our tetrahedron this means that we scan it by four parallel planes aligned by two different opposing edges. Now each plane contains the coefficients of the corresponding column. In Figure 2.1 the intersections of these four planes with the ones from above are highlighted by orange lines. Unfortunately, the pattern is not as nice as the one before. This is due to Equation (2.10), saying $L_m^{-T} = m^{-1}L_m m$ (compare also with Remark 2.1). This means that we must introduce weighted weights, i.e. weights for a whole column, and line-weights in order to compensate the multiplication with the two diagonal matrices. The resulting pattern is shown in Table 2.3 and Figure 2.2.

Scanning the first column of the matrix in (2.11), we get $(1, 1, 1, 1)$, which corresponds to the fourth line in Pascal's Triangle $(1, 3, 3, 1)$, when the line weights are associated. The second column may be split into two sub-columns, the third into three and so on. Again, we will always encounter binomial coefficients corresponding to lines in Pascal's Triangle: for example in the column weights, in the sub-column weights for each column, in the coefficients of the columns and in the inverted line weights. Altogether we associated the first and the last column of $block_3(L_m^{-T})$ with two opposing edges in our tetrahedron. Now given a tetrahedron, we know that it has three different opposing edge-pairs. Two of them have already been identified with the coefficients of $block_3(L_m^{-T})$ shown in Expression (2.11). Now, if we scan through our tetrahedron, using the remaining opposing edge pair, we get again an interesting coefficient pattern (see Figure 2.3). It is in fact exactly the same one we arrived at while scanning columnwise. Thus the tetrahedral

column	0	1	2	3	line weights
coefficients (not weighted)	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{3}$
subcolumn weight	1	1	1	1	1
column weight	1	3	3	1	

a) Tabular with non-weighted coefficients

column	0	1	2	3	line weights
coefficients (weighted)	$\frac{1}{1}$	$\frac{-3}{1}$	$\frac{3}{-2}$	$\frac{-1}{1}$	$\frac{1}{3}$
subcolumn weight	1	1	1	1	1
column weight	1	3	3	1	

b) Tabular with weighted coefficients

Table 2.3.: Tetrahedral coefficient structure according to the columns and the diagonals of $block_3(L_m^{-T})$

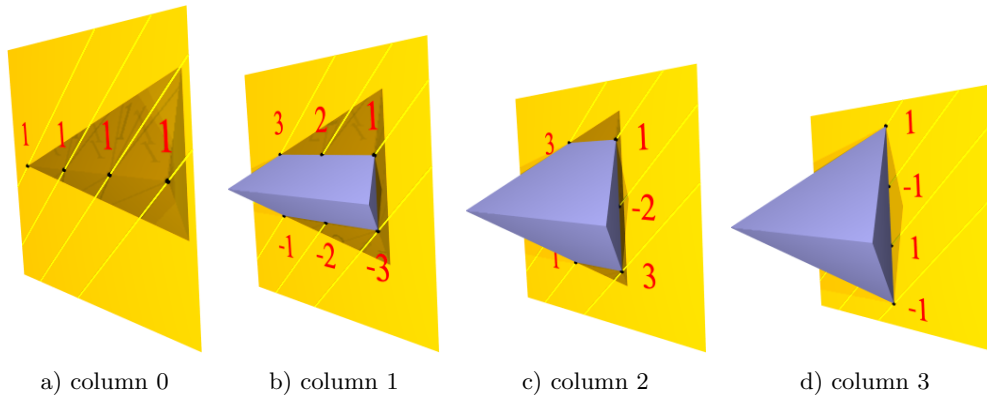


Figure 2.2.: Tetrahedron sliced four times corresponding to columns of (2.11)

structure is clear by the preceding examinations (compare with Table 2.3). What strikes most is that in a tetrahedron opposing edge pairs are equivalent due to the symmetry. However, our coefficient structure does not resemble this symmetry without exceptions: We got two times the structure shown in Table 2.3 and one time the one shown in Table 2.1. This corresponds to singling out one opposing edge pair in our tetrahedron. The corresponding coefficient pattern to this singled out edge pair is in fact simpler than the other patterns: No additional line and column-weights are needed.

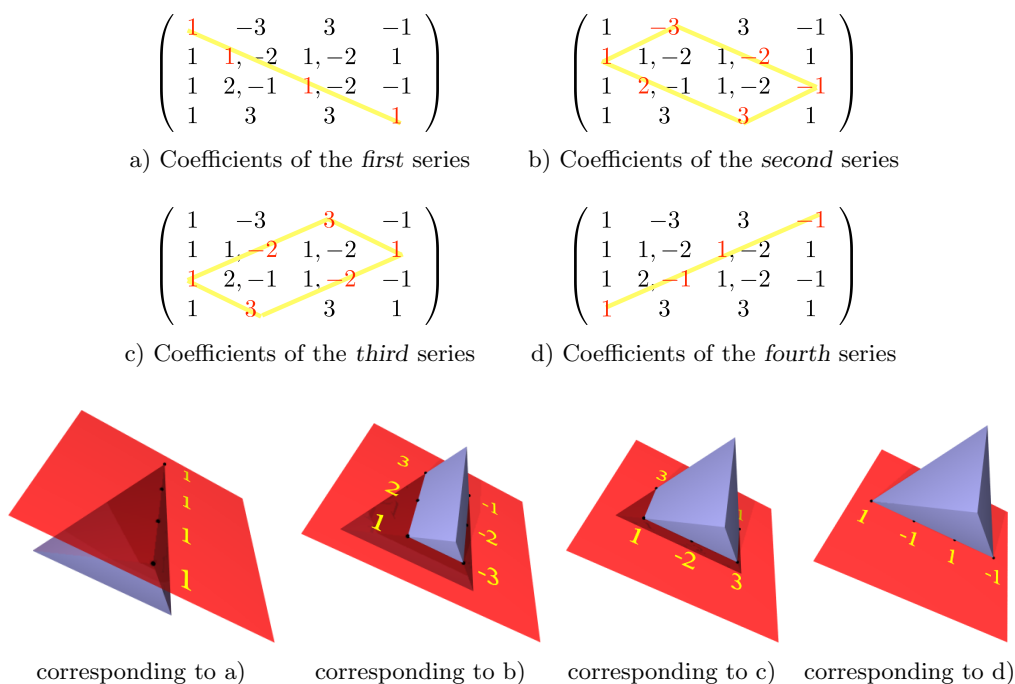


Figure 2.3.: Diagonal coefficient association in $block_3(L_m^{-T})$

The simpler pattern of Table 2.1 corresponds to the lines of $block_3(L_m^T)$. A single line in turn specifies the transformation of a curve coefficient under rotation of the plane. This connection of coefficient transformation and simple tetrahedral pattern is due to the introduction of multinomial coefficients.

In the last section and exemplarily in Remark 2.1 we saw that $m^{-1}L^{-T}m = L_m^{-T}$, which transforms a by $\tilde{a} = L_m^{-T}a$. Thus the introduction of multinomial coefficients leads to a rotation of the tetrahedron. According to the observations above and without multinomials we would have the lines of L^{-T} associated with the pattern of Table 2.3 (modulo signs), which would introduce unnecessary complexities.

2.3. Complexification

We have seen that Equation (2.8) and thus Equation (2.9) exhibit a lot of structure. But it is still unclear what this operation is doing geometrically with the curve coefficients. Can it be associated with a rotation of coefficients or is it something more complex? The answer is yes, but this operation is hidden because the examined polynomials were all in real monomial representation. In this section we will derive a better basis for our polynomials - better, in the sense of clearer structure and easier transformation matrices for curve coefficients under transformations of the plane. In Section 4.1 we will see that a rotation of the plane is in fact a rotation of the complexified curve coefficients.

Indeed complex representation was found useful for algebraic curves by two research groups independently and simultaneously (see [46] and [52]). We follow their approach of complexification and complement it by the use of multinomial coefficients. The previous examination of the tetrahedral coefficient structure may be understood as motivation for the subsequent transformation as complexification leads to similar structures. Thus

complexification will take care of the “complicated” parts that a rotation induces on the transformation of the curve coefficients. Consequently, rotations of the plane for complexified curves will be quite easy to describe (see Section 4.1).

2.3.1. Complexified curves

Looking back at the transformation in (2.5), it can be seen that it describes a rotation of the plane by

$$x \mapsto \cos(\varphi)x - \sin(\varphi)y \quad \text{and} \quad y \mapsto \sin(\varphi)x + \cos(\varphi)y .$$

This leads to the messy transformation (2.9) of the curve coefficient vector. From complex analysis we already know a shorter description for rotations: We may interpret the plane \mathbb{R}^2 as complex plane \mathbb{C} . The variables x and y can be transformed into complex variables z and \bar{z} by a (complex) linear transformation of the plane:

$$\begin{pmatrix} x \\ y \\ h \end{pmatrix} \mapsto \begin{pmatrix} 1 & i & 0 \\ 1 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ h \end{pmatrix} = \begin{pmatrix} x + iy \\ x - iy \\ h \end{pmatrix} = \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} . \quad (2.12)$$

In this framework a rotation by φ around the origin is simply

$$z \mapsto e^{i\varphi}z \quad \text{and consequently} \quad \bar{z} \mapsto e^{-i\varphi}\bar{z} .$$

Curves, former represented by $f(x, y, h) = 0$, are then represented by functions $F(z, \bar{z}, h)$ depending on a complex variable z and its conjugate \bar{z} . We will denote these curve-describing polynomials by

$$F(z, \bar{z}, h) = f\left(\frac{z + \bar{z}}{2}, \frac{z - \bar{z}}{2i}, h\right) = \sum_{k=0}^d \sum_{l=0}^{d-k} m_{kl} c_{kl} z^k \bar{z}^l h^{d-k-l} . \quad (2.13)$$

The dependence of F on z and \bar{z} is necessary for F to be an algebraic polynomial. Thereby d is the degree of f . The m_{kl} are multinomial coefficients and the c_{kl} will be called **(complex) curve coefficients**. $F(z, \bar{z}, h)$ is said to be the **complexification** of $f(x, y, h)$. The question now is, how the coefficients a_{kl} and the complex coefficients c_{kl} of a curve can be related. As complexification leaves the homogenizing component h invariant, we cover this relation in the affine set-up by

Theorem 2.7 *Let $f(x, y) = \sum_{k+l \leq d} m_{kl} a_{kl} x^k y^l = 0$ describe a curve $\mathcal{C} = V(f)$ of degree d . The introduction of $z = x + iy$ and $\bar{z} = x - iy$ permits a substitution of x by $\frac{z + \bar{z}}{2}$ and y by $\frac{z - \bar{z}}{2i}$ in f . The result is an algebraic polynomial $F(z, \bar{z}) = f\left(\frac{z + \bar{z}}{2}, \frac{z - \bar{z}}{2i}\right)$ in z and \bar{z} . Its homogenized version may be written in the form of Equation (2.13) with complex curve coefficients $c_{kl} \in \mathbb{C}$. The connection between the coefficients a_{kl} and c_{kl} is given by*

$$c_{kl} = \frac{1}{2^{k+l}} \sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} i^{(p-q)-(k-l)} a_{(p+q)((k+l)-(p+q))} . \quad (2.14)$$

PROOF The proof is completely analogous to the one of Theorem 2.6: We have

$$F(z, \bar{z}) = f\left(\frac{z + \bar{z}}{2}, \frac{z - \bar{z}}{2i}\right) = \sum_{k+l=0}^d m_{kl} a_{kl} \cdot \left(\frac{z + \bar{z}}{2}\right)^k \left(\frac{z - \bar{z}}{2i}\right)^l .$$

The right side of this equation may be expanded using the generalized binomial formula. Changing summation indices and collecting the coefficients of all terms containing a certain product $z^k \bar{z}^l$ results in the desired expression for c_{kl} . Carrying out these calculations is analogous to the detailed calculations shown in Appendix A.1. \square

The complex curve coefficients c_{kl} may be collected and written as coefficient vector c . The variable vector now containing products of powers of z , \bar{z} (and h) will be named Z . If we explicitly want to distinguish between the homogenous and inhomogenous variant, we use Z^{hom} and Z . Thus $F(z, \bar{z}) = (mc)^T Z$ and $F(z, \bar{z}, h) = (mc)^T Z^{hom}$. Thereby $Z = Z_d$ and $Z^{hom} = Z_d^{hom}$ depend on the degree d of F and analogous to Section 2.1.5:

$$Z^T = (1, z, \bar{z}, \dots, \bar{z}^d) \quad \text{and} \quad (Z^{hom})^T = (h^d, zh^{d-1}, \bar{z}h^{d-1}, \dots, \bar{z}^d) .$$

In Equation (2.14), we can see an interesting property analogous to the examinations after Theorem 2.6: The terms preceding $a_{(p+q)((n+l)-(p+q))}$ in Equation (2.14) can be viewed as an expansion of a generalized binomial formula:

$$\sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} i^{(p-q)-(k-l)} = (1+i)^k (1-i)^l .$$

Now, we may rewrite Equation (2.14) in a more descriptive way, using matrix-vector notation. According to the structure of the above equation, we may expect a block-diagonal structure with the coefficients of $(1+i)^k (1-i)^l$ scaled by $\frac{1}{2^{k+l}}$. For a better view on this structure we put these scaling factors $\frac{1}{2^{r-1}}$ afar from the transformation matrix and compare with

$$(1+i)^2 = 1 + 2i - 1; \quad (1+i)(1-i) = 1 + 0i + 1; \quad (1-i)^2 = 1 - 2i - 1 .$$

Altogether we get

$$\begin{aligned} & \left(a_{00} \mid a_{10} \quad a_{01} \mid a_{20} \quad a_{11} \quad a_{02} \mid a_{30} \quad a_{21} \quad a_{12} \quad a_{03} \mid \dots \right)^T \\ & \mapsto \left(c_{00} \mid c_{10} \quad c_{01} \mid c_{20} \quad c_{11} \quad c_{02} \mid c_{30} \quad c_{21} \quad c_{12} \quad c_{03} \mid \dots \right)^T = W^{-T} a \\ & = \left(\begin{array}{c|c|c|c|c} 1 & & & & \\ \hline & 1 & -i & & \\ & 1 & i & & \\ \hline & & & 1 & -2i & -1 \\ & & & 1 & 0 & 1 \\ & & & 1 & 2i & -1 \\ \hline & & & & & & 1 & -3i & -3 & i \\ & & & & & & 1 & -i & 1 & -i \\ & & & & & & 1 & i & 1 & i \\ & & & & & & 1 & 3i & -3 & -i \\ \hline & & & & & & & & & \ddots \end{array} \right) \begin{pmatrix} a_{00} \\ \frac{1}{2} \cdot a_{10} \\ \frac{1}{2} \cdot a_{01} \\ \frac{1}{4} \cdot a_{20} \\ \frac{1}{4} \cdot a_{11} \\ \frac{1}{4} \cdot a_{02} \\ \frac{1}{8} \cdot a_{30} \\ \frac{1}{8} \cdot a_{21} \\ \frac{1}{8} \cdot a_{12} \\ \frac{1}{8} \cdot a_{03} \\ \vdots \end{pmatrix} \quad (2.15) \end{aligned}$$

column	1	2	3	4	line weights
coefficients (not weighted)	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{3}$
	$\frac{3}{3}$	$\frac{2}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{3}{3}$
	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$
subcolumn weight	1	1 -1	1 -2	1	1 -3
column weight	1	3i	-3	-i	

Table 2.5.: Tetrahedral coefficient structure with respect to the columns of $block_3(W^{-T})$

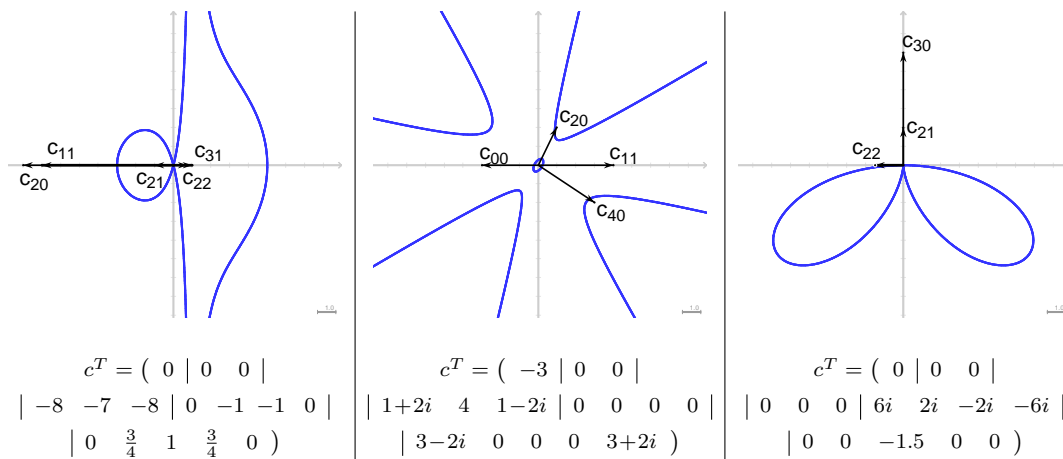


Figure 2.4.: Curves determined by a complex coefficient vector

2.3.3. Equivalence of real and complexified curves

The most important feature of the structure of (2.14) is that each $block_r(W^{-T})$ has $\lfloor \frac{r}{2} \rfloor$ complex conjugate line pairs. The first is conjugate to the last line, the second to the second-last line and so on:

$$c_{lk} = \frac{1}{2^{k+l}} \sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} \frac{1}{i^{(p-q)-(k-l)}} a_{(p+q)((k+l)-(p+q))} = \overline{c_{kl}} .$$

In every second block in the diagonal of W^{-T} , starting with $block_0(W^{-T})$, the entries of the line in the middle are real. Thus

$$c_{kl} = \overline{c_{lk}} \quad \text{and} \quad c_{kk} = \overline{c_{kk}} \in \mathbb{R} . \quad (2.16)$$

The relations in (2.16) are also exactly the conditions a curve $f(z, \bar{z}) = 0$ with complex coefficients must satisfy, such that it represents a real algebraic curve. One can easily prove this by counting the degrees of freedom of the curve describing polynomial.

The pairwise connection of the coefficients c_{kl} and c_{lk} is necessary to maintain the size, i.e. the dimension, of the coefficient-space of curves. The complexified representation of a curve, or better the complexified curve coefficient vector, contains redundant information. Figure 2.4 shows three curves. The curve coefficients are indicated as vectors below the curves. The vertical lines in the vectors indicate the separation into blocks.

2.3.4. The coefficient triangle

When speaking of a complexified curve, we normally have a coefficient vector c describing the curves totally. In many contexts it is beneficial not to use a linear notation for the vector components but a notation in form of a special triangle. Thereby we sort the coefficients according to their occurrence in the corresponding parts of the form $f(z, \bar{z})$: All coefficients which belong to the part being homogenous of a certain degree in z and \bar{z} are grouped in a line. This corresponds to Definition 2.3 and coefficients of a block are grouped linewise. The columns of our notation-scheme is ordered by the difference of indices in such a way that we obtain

$block_t(c)$	curve coefficients									
$t = 0$					c_{00}					
$t = 1$				c_{10}		c_{01}				
$t = 2$			c_{20}		c_{11}		c_{02}			
$t = 3$		c_{30}		c_{21}		c_{12}		c_{03}		
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots
index difference	\dots	-3	-2	-1	0	1	2	3	\dots	\dots

(2.17)

This notation-scheme will be called **coefficient triangle**. During the complexification-process, Equation (2.16) showed that $c_{kl} = \overline{c_{lk}}$ for any k and l . This may be interpreted as a kind of symmetry: Both c_{kl} and c_{lk} contain the *same information*. They only appear in complex conjugated versions of each other. In this context the coefficient triangle is symmetric with respect to the column exhibiting a vanishing index-difference. We will call this column the symmetry axis of the “reflectional symmetric” coefficient triangle. The axis is characterized by its real coefficients $c_{kk} \in \mathbb{R}$.

3. Curves and Tensor-Algebra

Tensors have been known about for approximately a century. Einstein showed the power of this kind of calculus in his disquisition of the relativity theory (cf. [18]) and brought tensors to popularity. Soon tensors found a wide application in mathematics, especially in the field of invariant theory. Tensors were extensively exploited for example by Weitzenböck in [55], by Gurevich in [25] and by many others.

However, tensor-notation has deficiencies despite its descriptive power: In this notation expressions get easily overcrowded with indices and indices of indices and even higher iterations of indices of indices. Things become very impractical and structural properties are literally buried under these indices. To avoid this we will simply replace the tensor-components by a node of a graph and attach in- and outgoing arrows to it, according to the co- and contravariant indices (see Section 3.1.3). For a first entry-point to this technique we recommend [41]. We happened to discover this technique in a series of articles written by Jim Blinn (cf. [9], [10]). He used this technique to treat projective geometry on a diagrammatic level and examined especially lines, conics and cubics and some of their invariants (see [3], [4], [5], [6], [7], [8]). But tensor-diagrams have already been used in the literature (see [17],[38]). First attempts can be traced back to Clifford and the so-called Clifford's Graphs (see [14], [15], [28], [42], [45]). There are many other examples in which tensor diagrams are useful tools. For example, in group theory (cf. [43]) or in quantum information theory (see [2]).

3.1. Connection between tensor algebra and curves via forms

We have seen that a plane algebraic curve $\mathcal{C} = V(f)$ of degree d can be represented by a zero-set of a polynomial

$$f(x) = \sum_{k=0}^d \sum_{l=0}^{d-k} m_{kl} c_{kl} \cdot x_1^k x_2^l x_3^{d-k-l} .$$

Here x may be $(x_1, x_2, x_3) = (x, y, h)$ or $(x_1, x_2, x_3) = (z, \bar{z}, h)$, depending on the working frame. A point $P = (p_1, p_2, p_3)$ in homogeneous coordinates is on a curve if $f(P) = 0$. In most cases this formula is very unhandy and structural results and especially invariants can be very hard to extract. As stated above, we will derive a very simple representation of \mathcal{C} , introducing tensor notation and visualizing \mathcal{C} using diagram techniques. In contrast to the triangular notation-scheme from Section 2.3.4, the coefficients are now ordered in form of squares, cubes and in general hyper-cubes.

We will start by recapitulating tensor-algebra basics from an operational point of view: We do not use tensors as abstract algebraic objects of a tensor-space. But tensors are viewed as multilinear objects defined via their transformation behavior. We begin with an examination of so-called co- and contravariant vectors. With a view to their transformation behavior general tensors are introduced. After examining tensor operations, such as multiplication, addition and contraction, curve-tensors are focused on.

3.1.1. Covariant and contravariant vectors

Working in the projective plane or in the projectively extended Euclidean plane, scalars, vectors with three elements, 3×3 matrices and in general $3 \times 3 \times \dots \times 3$ tensors are used. The simplest quantities besides scalars are vectors. Points $x = (x_1, x_2, x_3)^T$ as well as lines $l = (l_1, l_2, l_3)$ may be represented by vectors. The simplest operation between vectors is probably the scalar product $\langle l, x \rangle$. If for a point x and a line l the equation

$$\langle l, x \rangle = l_1x_1 + l_2x_2 + l_3x_3 = \sum_{i=1}^3 l_i x_i = 0 \quad (3.1)$$

is satisfied, the point x is on line l . Thus vectors represent two different but dual concepts: points and lines. In order to be able to distinguish between these two, we introduce a new notation for points by "moving" the index: A point will be represented by a vector whose components have an upper index. In this way Equation 3.1 changes to

$$\langle l, x \rangle = \sum_{i=1}^3 l_i x^i = 0 .$$

Dealing with higher order curves and more complex products, many sums show up. To shorten notation, we use the **Einstein Sum Convention** (ESC): Every time an index shows up as an upper and as a lower index, the whole expression is summed over all possible values of this index.

This notation is credited to Albert Einstein, who introduced it in [18]. It is widely used especially in relativistic theory (cf. [23]). With the ESC Equation (3.1) shortens to

$$\langle l, x \rangle = l_i x^i = 0 . \quad (3.2)$$

In Equation (3.2) the range of the index i must be clear. Here we deal with plane algebraic curves, utilizing homogeneous coordinates. The possible values for any index will always be 1, 2 and 3, i.e. $i \in \{1, 2, 3\}$, unless explicitly stated otherwise. Thus we will mostly omit the addendum of the range of our indices. Nevertheless, the concepts shown also apply for other finite dimensional spaces.

In Section 2.2.2 we introduced transformations P of the (projective) plane. Points (x^1, x^2, x^3) were mapped onto $Q(x^1, x^2, x^3)^T$, where $Q = P^{-T} \in \mathbb{K}^{3 \times 3}$ ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$). The result was a transformed point $(\tilde{x}^1, \tilde{x}^2, \tilde{x}^3)$. Let p_α^i with $1 \leq i, \alpha \leq 3$ be the components of P , then in an ESC-conform notation:

$$x^i = p_\alpha^i \tilde{x}^\alpha , \quad i, \alpha \in \{1, 2, 3\} . \quad (3.3)$$

In fact, we say that a geometric object consisting of three components (x^1, x^2, x^3) , which transforms according to (3.3) under a transformation of the plane, is called a **contravariant vector**.

Solving Equation (3.3) for the transformed coordinates \tilde{x}^α , we get

$$\tilde{x}^\alpha = q_i^\alpha x^i , \quad i, \alpha \in \{1, 2, 3\} , \quad (3.4)$$

where

$$p_\alpha^i q_i^\beta = \delta_\alpha^\beta = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases} \quad (3.5)$$

Let

$$\Delta = \det \left((p_{\alpha}^i)_{1 \leq \alpha, i \leq 3} \right)$$

be the determinant of the transformation matrix $P = (p_{\alpha}^i)_{1 \leq \alpha, i \leq 3}$. This way, q_i^{α} is the minor of Δ with respect to p_{α}^i divided by Δ .

Accordingly a **covariant vector** may be defined by its 3 components l_1, l_2, l_3 and its transformation behavior under linear transformations of space. It is:

$$\tilde{l}_{\alpha} = p_{\alpha}^i l_i, \quad i, \alpha \in \{1, 2, 3\} \quad (3.6)$$

or by utilizing Equation (3.5):

$$l_i = q_i^{\alpha} \tilde{l}_{\alpha}, \quad i, \alpha \in \{1, 2, 3\} .$$

With these definitions we ensure that an incidence of a point and a line remains unchanged under linear transformation of space. The invariance of equation (3.2) is shown by

$$l_i x^i = p_{\alpha}^i l_i \tilde{x}^{\alpha} = \tilde{l}_{\alpha} \tilde{x}^{\alpha} . \quad (3.7)$$

Remark 3.1 *Point-like vectors are named contravariant for the following reason: Points (vectors) as an element of a vector space are given by a linear combination of base vectors. Performing a linear transformation of space, the coordinate vectors of the points change. Let P be the transformation matrix representing the transformation of the base vectors. Then P^{-T} acts on the point-vectors, i.e. P counter-acts on the points (cf. [22]). Thus vectors which transform like base vectors are called covariant vectors and vectors transforming point-like are called contravariant.*

We may also have taken Equation (3.7) as a prerequisite for our treatment of tensors. This way a transformation of covariant vector would consequently act contragrediently (in a counter-acting manner) on contravariant vectors. Using matrix vector notation as in Remark 3.1 we get for example $Pl^T = \tilde{l}$. Then by 3.7

$$\langle l, x \rangle = lx = \underbrace{lP^T}_{=\tilde{l}} \underbrace{(P^T)^{-1} x}_{=Qx=\tilde{x}} = \tilde{l}\tilde{x} = \langle \tilde{l}, \tilde{x} \rangle .$$

Thus Equation (3.7) and contragredient transformations are closely related.

3.1.2. Tensors

We saw in Section 2.2.2 that a transformation of space may be interpreted as multiplication with a matrix P . Suppose we have a plane, which was already transformed by the use of another matrix R . Carrying P over to this plane means that we have to transform it by $\tilde{P} = R^{-1}PR$. This is clearly a transformation law for transformations: They can be regarded as geometric objects, which transform point- and line-like. We now generalize the concept of co- and contravariant vectors, which had been objects defined by components and a transformation law. This way we get co, contravariant and mixed tensors. A first example for mixed tensors are transformations.

Definition 3.1 (Tensor) A **tensor** $\left(t_{i_1, i_2, \dots, i_r}^{j_1, j_2, \dots, j_s}\right)$ of covariance r and contravariance s (or in short: a **r - s -tensor**) is a set of numbers $t_{i_1, i_2, \dots, i_r}^{j_1, j_2, \dots, j_s}$, which changes under linear transformations of space in accordance with the law

$$\tilde{t}_{\alpha_1, \alpha_2, \dots, \alpha_r}^{\beta_1, \beta_2, \dots, \beta_s} = p_{\alpha_1}^{i_1} p_{\alpha_2}^{i_2} \dots p_{\alpha_r}^{i_r} \cdot q_{j_1}^{\beta_1} q_{j_2}^{\beta_2} \dots q_{j_r}^{\beta_s} \cdot t_{i_1, i_2, \dots, i_r}^{j_1, j_2, \dots, j_s} . \quad (3.8)$$

Any index is within the range of $\{1, 2, \dots, n\}$ and $n = 3$ for the projective plane. (If it is clear, we often omit the brackets “()” around t , symbolizing that the tensor is meant rather than a component of that tensor.) The sum of the number of co- and contravariant indices $r + s$ of a tensor is called its **variance** or **order**.

Remark 3.2 There are alternative definitions for tensors in the literature: It is Equivalent to the above Definition 3.1 to speak of a tensor as a multilinearform. The components of which may be accessed by applying all combinations of the standard unit vectors (vectors of length one and with all but one component vanishing). For an introduction to this concept see [22]. A comprehensive study can be found in almost all books on tensor algebra (cf. [38], [55], ...). Here, we use the definition most common in physics. The focus lies on the transformation behavior which is most beneficial in our context of curve recognition using invariants.

Special cases of tensors are scalars, which are tensors of variance zero. Covariant vectors are 1-0-tensors and contravariant vectors are 0-1-tensors. As we have seen, transformations may be interpreted as 1-1-tensors. The number of components of a tensor depends on the range of its indices and its variance. If the index-range of all tensor-indices is n and if the variance of that tensor equals v , then it has n^v components. n^v is the number of possible combinations of indices attached to the tensor. Thus a $3 \times 3 \times 3$ tensor t_{ikl} has $3^3 = 27$ components, since $i, k, l \in \{1, 2, 3\}$.

In what follows we will encounter special tensors exhibiting symmetries.

Definition 3.2 Let $i_1, i_2, \dots, i_r \in I_n = \{1, 2, \dots, n\}$ be tensor-indices.

A covariant r -tensor $(t_{i_1, i_2, \dots, i_r})$ (or a contravariant r -tensor $(t^{i_1, i_2, \dots, i_r})$) is called (**partially**) **symmetric** with respect to the indices i_a and i_b if

$$t_{i_1, \dots, i_a, \dots, i_b, \dots, i_r} = t_{i_1, \dots, i_b, \dots, i_a, \dots, i_r} \quad (\text{or} \quad t^{i_1, \dots, i_a, \dots, i_b, \dots, i_r} = t^{i_1, \dots, i_b, \dots, i_a, \dots, i_r})$$

for all $i_a, i_b \in I_n$. A tensor is called (**totally**) **symmetric** if it is symmetric with respect to all index-pairs.

A covariant r -tensor $(t_{i_1, i_2, \dots, i_r})$ (or a contravariant r -tensor $(t^{i_1, i_2, \dots, i_r})$) is called **skew symmetric** with respect to the indices i_a and i_b if

$$t_{i_1, \dots, i_a, \dots, i_b, \dots, i_r} = -t_{i_1, \dots, i_b, \dots, i_a, \dots, i_r} \quad (\text{or} \quad t^{i_1, \dots, i_a, \dots, i_b, \dots, i_r} = -t^{i_1, \dots, i_b, \dots, i_a, \dots, i_r})$$

for all $i_a, i_b \in I_n$. A tensor is called (**totally**) **skew symmetric** if it is skew-symmetric with respect to all pairs (i_a, i_b) of different indices $i_a \neq i_b$.

Let us take the tensor (t_{ikl}) as an example. It is symmetric if $t_{ikl} = t_{ilk} = t_{kli} = t_{kil} = t_{lik} = t_{lki}$. It is skew-symmetric if $t_{ikl} = -t_{ilk} = t_{kli} = -t_{kil} = t_{lik} = -t_{lki}$.

This way the number of the effective components, i.e. the number of possibly different components, diminishes. For symmetric tensors $(t_{i_1, i_2, \dots, i_r})$ it is $\binom{n+r-1}{r}$ and for

skew-symmetric tensors it is $\binom{n}{r}$. This can be proven by analyzing the possible index-permutations (cf. [25]).

(Skew-)Symmetric tensors are important, because their symmetry is preserved by linear transformations of the plane (or linear transformations of higher dimensional spaces):

$$\begin{aligned} \tilde{t}_{\alpha_1, \dots, \alpha_a, \dots, \alpha_b, \dots, \alpha_r} &= p_{\alpha_1}^{i_1} p_{\alpha_2}^{i_2} \dots p_{\alpha_r}^{i_r} \cdot t_{i_1, \dots, i_a, \dots, i_b, \dots, i_r} \\ &= \pm p_{\alpha_1}^{i_1} p_{\alpha_2}^{i_2} \dots p_{\alpha_r}^{i_r} \cdot t_{i_1, \dots, i_b, \dots, i_a, \dots, i_r} \\ &= \pm \tilde{t}_{\alpha_1, \dots, \alpha_b, \dots, \alpha_a, \dots, \alpha_r} \cdot \end{aligned} \quad (3.9)$$

The space of the components of a tensor $(t_{i_1, i_2, \dots, i_d})$ is n^d -dimensional. A basis for this space is given by the objects $(t_{i_1, i_2, \dots, i_d})_{i_k \in \{1, 2, \dots, n\}, \forall k}$, where all but one component equal to one vanish. For our purposes another basis is more suitable. There is one, which gets along with only totally symmetric and skew-symmetric tensors with respect to two indices. No non-symmetric tensors are needed. For a proof we refer to Appendix A.2. For example, any 3×3 -tensor (t_{kl}) may be given by

$$(t_{kl}) = \begin{pmatrix} \lambda_1 & \lambda_4 & \lambda_5 \\ \lambda_4 & \lambda_2 & \lambda_6 \\ \lambda_5 & \lambda_6 & \lambda_3 \end{pmatrix} + \begin{pmatrix} 0 & \mu_1 & \mu_2 \\ -\mu_1 & 0 & \mu_3 \\ -\mu_2 & -\mu_3 & 0 \end{pmatrix} .$$

The nine degrees of freedom for a 3×3 -tensor are split into six degrees of freedom for symmetric matrices and three degrees of freedom for skew-symmetric matrices.

We will now define a partition on the set of all tensors $(t_{i_1, i_2, \dots, i_d})$ with i_1, i_2, \dots, i_d being elements of $\{1, 2, \dots, n\}$ into equivalence classes. This will help us in the context of curves. In our context usually $n = 3$. Therefore let $skew_{d,n}$ be the set d -0-tensors which are skew-symmetric with respect to at least two indices. Then

$$[t_{i_1, i_2, \dots, i_d}] := \left\{ \lambda (s_{i_1, i_2, \dots, i_d}) \mid \lambda \neq 0 \text{ and } (s_{i_1, i_2, \dots, i_d}) \text{ is a linear combination of elements of } skew_{d,n} \text{ and } (t_{i_1, i_2, \dots, i_d}) \right\} \quad (3.10)$$

Due to the above stated existence of a basis, consisting of totally symmetric and skew-symmetric elements, $[t_{i_1, i_2, \dots, i_d}]$ is an equivalence class. A representational system of the set of all such classes are the totally symmetric tensors. We can speak of an equivalence class of *tensors*, since linear transformations of the plane or space are consistent with the separation into such classes: This is ensured by the invariance of the symmetry attributes of tensors under such transformations (see Equation (3.9)).

3.1.3. Diagram notation of tensors

To get rid of the many indices, we represent a tensor by a node with in-going and out-going edges standing for co- and contravariant indices. We will name the in-going edges entries and the outgoing edges exits. If necessary, we will attach an index to the corresponding edge. If a tensor has more than one in- or outgoing arc, then these will be ordered counter-clockwise with respect to the order of sub- and superscripts (see Figure 3.1). Naturally, if a tensor is totally symmetric, this order is irrelevant. A scalar is just a node without dangling edges. A line-describing and therefore covariant vector is a node with one entry. A node with one exit represents a contravariant vector. The tensor t_{ij}^{lm} has three entries and two exits. These examples are displayed in Figure 3.1.

By Equation (3.6) a covariant vector transforms with $\tilde{l}_\alpha = p_\alpha^i l_i$. On the right hand side the ESC implies a summation of p_α^i and l_i over the joint index i . In diagram notation the

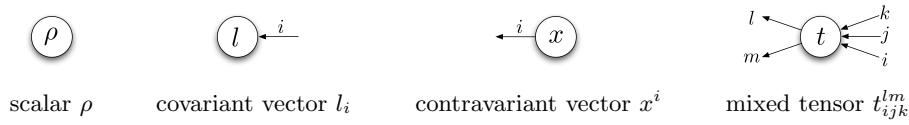


Figure 3.1.: Diagrams for simple tensors

summation is expressed by the connection of the nodes with an edge corresponding to index i . i is exit to x^i and an entry to p^i_α . The directedness of these edges corresponds to the ESC: Implicit summation takes place only with respect to an index appearing as *co* and as *contra*variant index. Figure 3.2 shows the transformation of tensors in accordance with Equations (3.3), (3.6) and (3.8).

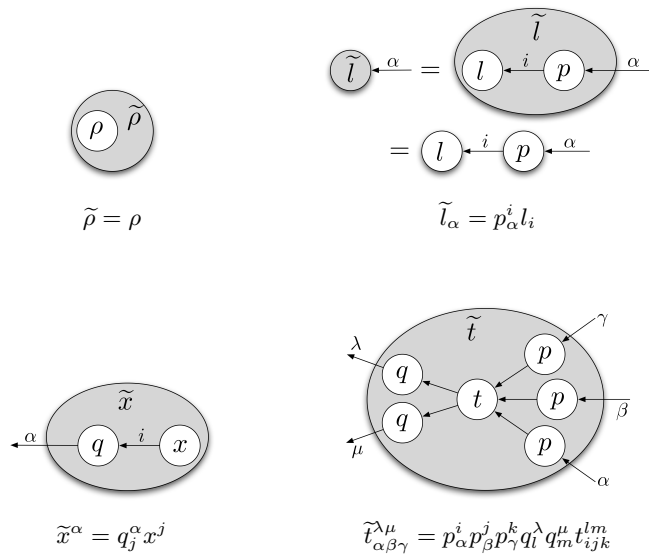


Figure 3.2.: Transformation of tensors in diagram notation

Connecting the dangling “ α -edges” of the second and third image of Figure 3.2 means building the scalar product of the two vectors l and x : The indices corresponding to the two dangling α -edges are identified - they are already named equally - and summation over that index takes place. Using Equation (3.5), saying that whenever p^i_α is connected with q^β_j , it can be replaced by an edge without nodes and we get our scalar product (cf. Figure 3.3).

3.1.4. Tensoroperations

Next we will discuss three operations on tensors: multiplication, addition and contraction (see also Figure 3.4). A multiplication of two tensors of any variance again equates to a tensor, where the *co*- and *contra*variances add up, respectively: We can write two tensor nodes next to each other and interpret the result as a new ”bigger“ node. The dangling edges of the two multiplied tensors become dangling edges of the result. Clearly, the new node represents a tensor because the edges, pointing inwards and outwards, define its behavior under transformations.

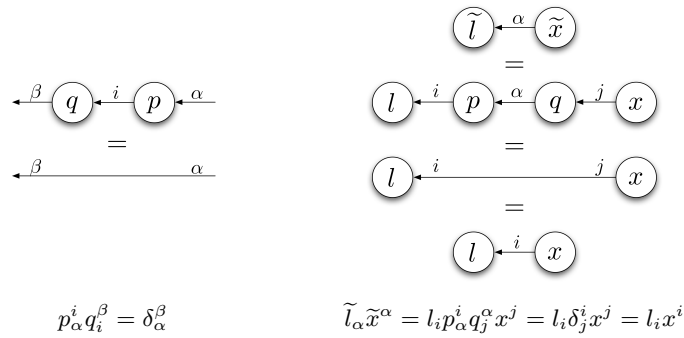


Figure 3.3.: Invariance of the algebraic equation of a line

Addition is a slightly more difficult. We can add up only tensors with the same number of co- and contravariant indices. If we add up two matrices A and B for example, the dimensions must match each other because addition is componentwise. For square matrices A and B , $A + B^T$ is also expressible with diagrams. Let A and B have components a_{ij} and b_{kl} , respectively. Let c_{rs} be the result of the addition. We get the result C of $A + B$ in tensor notation by identifying $i = k = r$ and $j = l = s$. $A + B^T$ is accessible by identifying $i = l = r$ and $j = k = s$ (see also the right image in Figure 3.4). We call every tuple of three identified indices in a tensor-summation a parallel triplet. In any case we need to specify these triplets in order to be able to perform tensor addition. Clearly, every index must be contained in exactly one parallel triplet. A co- and a contravariant index may of course never be parallel because the transformation laws do not match. Thus addition, parallel to covariant (contravariant) indices, results in a covariant (contravariant) structure. The summed expression has the same variance as the summands. The result of tensor addition is again a tensor inheriting the transformation law from the summands (see Figure 3.5).

Contraction is an operation we have already met, caring for transformations: We built a tensor by multiplying x^i by a transformation-tensor p_i^α . This way, we got a tensor $t_i^{i\alpha}$ with a loop: An exit was directly connected to an entry. Now we may simply draw a bigger node around our $t_i^{i\alpha}$, totally containing the “ i -loop” and name it u^α . This corresponds to tensor-contraction. The new tensor u^α behaves like a contravariant vector (see the right image in Figure 3.4), because $t_i^{i\alpha}$ transforms like u^α . In general, contraction of a tensor reduces its variance by two: The number of co- and contravariant indices reduce each by one.

Remark 3.3 *When we contract a tensor with respect to all its indices, we speak of total contraction. A totally contracted tensor is a scalar. It does not change under linear transformation of space. In diagram notation totally contractible tensors are closed diagrams. As a scalar they can be represented by a single node without dangling edges. Transformation of space replaces edges by edges with transformation-nodes. Thus we can also see from diagram notation that closed diagrams in total or single nodes without edges do not change.*

3. Curves and Tensor-Algebra

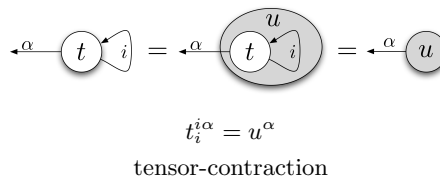
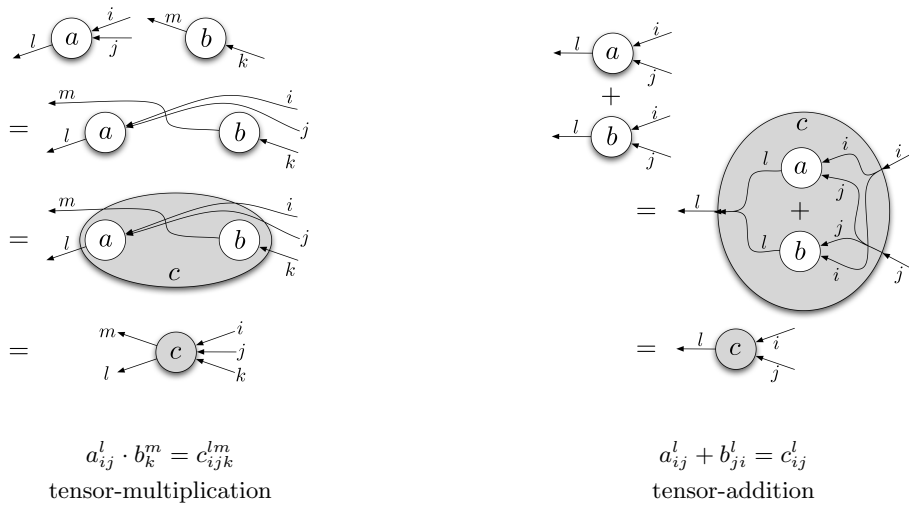
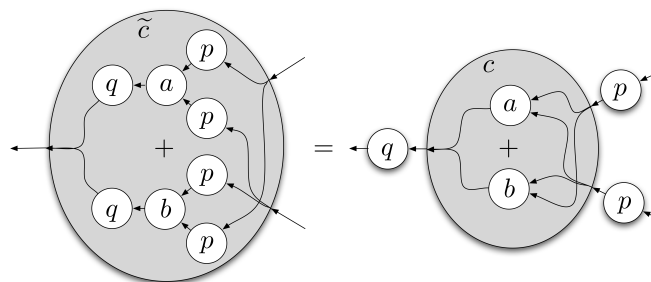


Figure 3.4.: Product, sum and contraction in diagram notation



The transformed tensor-sum has again tensor structure.

Figure 3.5.: Transformation law for the sum of two tensors

3.1.5. Curves as tensors

In Section 2.2.1 we introduced curves as zero-sets. A curve \mathcal{C} was defined by an implicit form-function f , namely $\mathcal{C} = V(f)$. Thereby we have a one to one correspondence between \mathcal{C} and $[f]$, when f is an associated polynomial of \mathcal{C} . Now we will show the connection between tensors and polynomials and plane algebraic curves in general. For this reason we set the index-range of all tensor indices to $\{1, 2, 3\}$. The curves and polynomials shall be of degree d for a predefined $d \in \mathbb{N} \setminus \{0\}$.

First let us examine $t_{i_1, i_2, \dots, i_d} x^{i_1} x^{i_2} \dots x^{i_d}$. Due to ESC we have to sum over all indices and

$$t_{i_1, i_2, \dots, i_d} \cdot x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_d} = \sum_{i_1=1}^3 \sum_{i_2=1}^3 \dots \sum_{i_d=1}^3 t_{i_1, i_2, \dots, i_d} \cdot x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_d} .$$

We may collect all terms according to the powers of x^1 , x^2 and x^3 and rewrite this expression as

$$t_{i_1, i_2, \dots, i_d} \cdot x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_d} = \sum_{k=0}^d \sum_{l=0}^{d-k} A_{kl} (x^1)^k (x^2)^l (x^3)^{d-k-l} . \quad (3.11)$$

Thereby $A_{kl} = \sum_{\pi \in S_d} t_{\pi(1, \dots, 1, 2, \dots, 2, 3, \dots, 3)}$ is the sum over all t_{i_1, i_2, \dots, i_d} with exactly k indices equal to 1 and exactly l indices equal to 2 in all permuted variations. Thus adding any skew-symmetric tensor to t_{i_1, \dots, i_d} leads to the same A_{kl} and thus to the same form. Furthermore, multiplication of this tensor by any constant λ leads to a multiplication of all A_{kl} by the same constant λ .

Theorem 3.1 *Let*

$$M_d = \left\{ [t_{i_1, \dots, i_d}] \mid \exists (j_1, \dots, j_d) \in \{1, 2\}^d : t_{j_1, \dots, j_d} \neq 0 \right\} .$$

The set of all plane algebraic curves \mathcal{C} of degree d is structurally isomorphic to M_d .

The restriction on the tensors in the above theorem ensures that $t_{i_1, i_2, \dots, i_d} x^{i_1} x^{i_2} \dots x^{i_d}$ contains a term of degree d in (x^1) and (x^2) .

For totally symmetric t_{i_1, i_2, \dots, i_d} we have

$$A_{kl} = m_{kl} t_{\underbrace{1, 1, \dots, 1}_{k \text{ times}} \underbrace{2, 2, \dots, 2}_{d-k-l \text{ times}} \underbrace{3, 3, \dots, 3}_{l \text{ times}}} =: m_{kl} a_{kl} \quad (3.12)$$

and Equation (3.11) evaluates to

$$t_{i_1, i_2, \dots, i_d} \cdot x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_d} = \sum_{k=0}^d \sum_{l=0}^{d-k} m_{kl} a_{kl} (x^1)^k (x^2)^l (x^3)^{d-k-l} . \quad (3.13)$$

Equations (3.12) and (3.13) give us a direct correspondence between totally symmetric tensors and curve describing polynomials. Thus with a totally symmetric t_{i_1, i_2, \dots, i_d} we can identify

$$\mathcal{C} \longleftrightarrow [f] \longleftrightarrow [t_{i_1, i_2, \dots, i_d}] \longleftrightarrow \{ \lambda t_{i_1, i_2, \dots, i_d} \mid \lambda \neq 0 \} .$$

f and t_{i_1, i_2, \dots, i_d} play an equivalent role in the definition of a curve.

Therefore we will often speak of totally symmetric tensors as **curve-tensors**. Simplifying notation we also speak of t_{i_1, i_2, \dots, i_d} as a curve. Two examples will provide a feeling for the connection between curves and tensors.

Example 3.1 Let us take an ellipsis given by $\frac{(x-1)^2}{a^2} + \frac{(y-1)^2}{b^2} - 1 = 0$ with major axis $a = 1$ and minor axis $b = \frac{1}{2}$. Homogenizing this equation and setting $x = x^1$, $y = x^2$ and the homogenizing component to x^3 , we get

$$f(x^1, x^2, x^3) = 4(x^3)^2 - 2(x^1)(x^3) - 8(x^2)(x^3) + (x^1)^2 + 4(x^2)^2 = 0 .$$

Thus taking care of the multinomial coefficients the curve-coefficient vector is

$$(a_{00} \ a_{10} \ a_{01} \ a_{20} \ a_{11} \ a_{02}) = (4 \ -1 \ -4 \ 1 \ 0 \ 4) .$$

The corresponding totally symmetric tensor is

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} = \begin{pmatrix} a_{20} & a_{11} & a_{10} \\ a_{11} & a_{02} & a_{01} \\ a_{10} & a_{01} & a_{00} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 4 & -4 \\ -1 & -4 & 4 \end{pmatrix}$$

and in fact

$$t_{kl}x^kx^l = (x^1 \ x^2 \ x^3) \begin{pmatrix} 1 & 0 & -1 \\ 0 & 4 & -4 \\ -1 & -4 & 4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = f(x^1, x^2, x^3) .$$

The curve $\mathcal{C} = V(f)$ can be identified with $[f] = \{ \lambda f \mid \lambda \neq 0 \}$ or with

$$[t_{kl}] = \left\{ \lambda \begin{pmatrix} 1 & \alpha & -1 + \beta \\ -\alpha & 4 & -4 + \gamma \\ -1 - \beta & -4 - \gamma & 4 \end{pmatrix} \mid \alpha, \beta, \gamma \in \mathbb{R} \ \lambda \neq 0 \right\} .$$

In diagram notation this looks like $\textcircled{x} \rightarrow \textcircled{T} \leftarrow \textcircled{x} = f(x) = f(x^1, x^2, x^3)$.

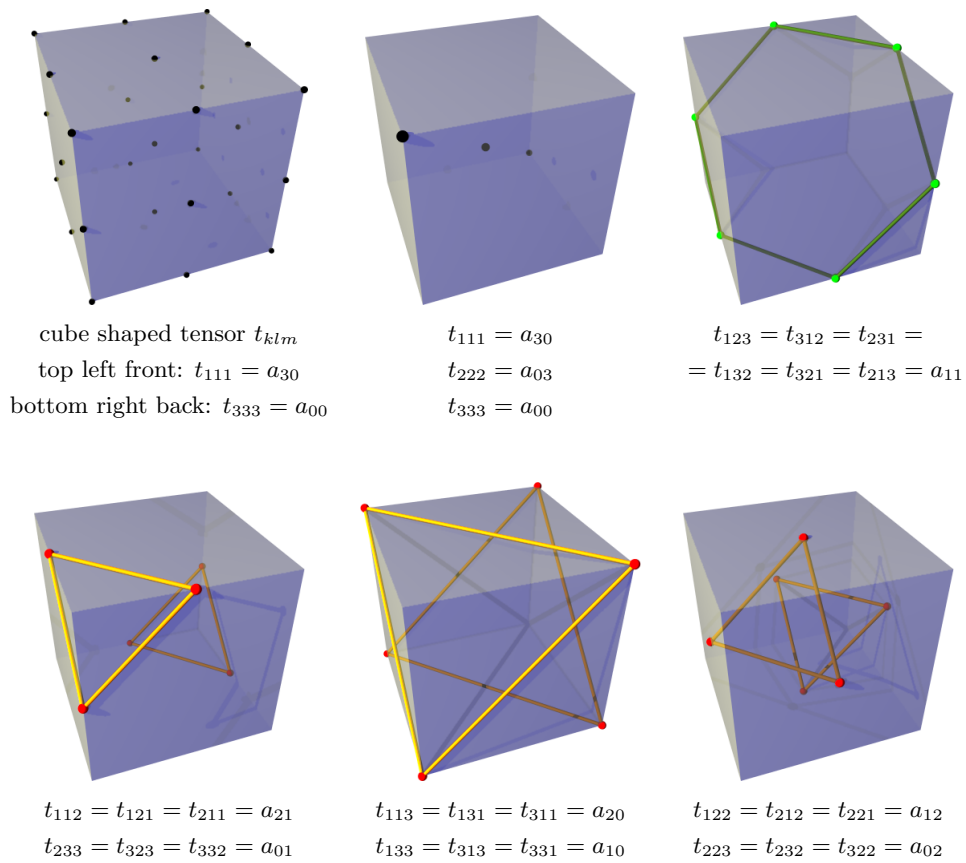
Example 3.2 Suppose we have a cubic given by

$$0 = f(x, y) = a_{00} + 3a_{10}x + 3a_{01}y + 3a_{20}x^2 + 6a_{11}xy + 3a_{02}y^2 + a_{30}x^3 + 3a_{21}x^2y + 3a_{12}xy^2 + a_{03}y^3 .$$

We can arrange the ten curve coefficients a_{kl} in a $3 \times 3 \times 3$ totally symmetric tensor t_{pqr} according to Equations (3.12) and (3.13). Thereby we can visualize t_{pqr} by a tensor, a cube-shaped “matrix”, (see Figure 3.6) with $3^3 = 27$ components. The number of effective components of the totally symmetric t_{klm} is $\binom{5}{3} = 10$, exactly matching the number of curve coefficients of our cubic. Each of the remaining 17 components of t_{pqr} are equal to one of those 10 components. Figure 3.6 shows this equality by highlighted orbits. The cardinality of tensor-components on each orbit corresponds to the associated multinomial coefficient. Visualizing the tensor as a cubical matrix, the curve describing polynomial $f(x, y)$ can be obtained by multiplying the cube by $(x, y, 1)$ from three directions (see Figure 3.7). We can write

$$\begin{array}{c} \textcircled{x} \\ \downarrow \\ \textcircled{T} \\ \uparrow \quad \uparrow \\ \textcircled{x} \quad \textcircled{x} \end{array} = f(x, y) = t_{klm}x^kx^lx^m = 0 .$$

Thereby due to the affine set-up we fix the homogenizing component to be $x^3 = 1$.



Components which are equal by virtue of the total symmetry of the tensor, are displayed within an orbit in each picture.

Figure 3.6.: Cubical and symmetry-structure of a totally symmetric tensor t_{klm} corresponding to a curve of degree three

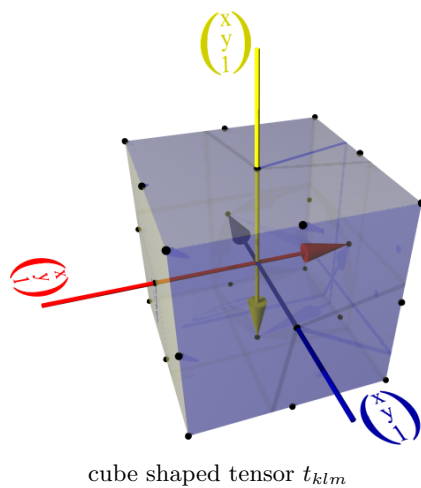


Figure 3.7.: Cubic polynomial via multiplication of a cubical tensor

3.2. Invariants

In this section we gather all ingredients we need to state invariants in an elegant tensor-diagram notation. Due to the concept of homogenous coordinates and the identification of objects with non-zero multiples, certain difficulties in stating the term of an invariant arise. We avoid these by referring to vector-configurations and therefore distinguish between multiples of the objects.

Definition 3.3 Let G be a transformation group and $T \in G$. Let O be a collection of tensors (representing geometric objects) and \tilde{O} the collection of the tensors of O transformed by T . A polynomial $\varphi(O)$ (or $\varphi(\text{tensor}_1, \text{tensor}_2, \dots)$) is called a **relative invariant** if

$$\varphi(\tilde{O}) = \det(T)^g \varphi(O)$$

for all choices of corresponding elements in O and for all $T \in G$. g is called the weight of the invariant. If $g = 0$, then φ is called an **absolute invariant**.

Projectively invariant properties may be expressed by the vanishing of a polynomial invariant φ . In those cases a multiplicative factor of $\det(T)^g$ changes nothing. Measuring projectively invariant functions such as the cross ratio, we have to consider special rational functions: a polynomial function divided by another, exhibiting exactly the same kind and the same number of objects. We will see examples later on.

In case of relative invariants φ has only geometric significance if tested against zero. In our case O will mainly consist of a single tensor representing an algebraic curve. We already encountered an example of an invariant with two distinct objects in Figure 3.3: The incidence between a point and a line: $\textcircled{l} \leftarrow \textcircled{x} = 0$. Another one is $p^i p^j Q_{ij} = 0$ with a point p^i and a conic Q_{ij} . $\textcircled{p} \rightarrow \textcircled{Q} \leftarrow \textcircled{p}$ is the corresponding diagram and the property of a point lying on a conic - the property of the diagram evaluating to zero - is projectively invariant.

3.2.1. The δ - and ε -tensor

So far we have built diagrams only by using “object-tensors” and the diagrams were quite linear in structure. This changes when we introduce the so called ε -tensor and the closely related δ -tensor. For further introduction regarding these tensors compare with [25] or [41].

Let us start with the seemingly unimposing δ -tensor. It is just the Kronecker-Delta δ_i^j , being $\delta_i^i = 1$ and $\delta_i^j = 0$ for $i \neq j$. The benefit of it is that it enables arrow-relabeling: For instance $p^i \delta_i^j = p^j$. In the diagrammatic point of view it is just an intermediate arrow:

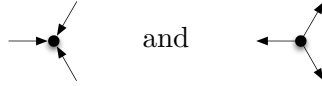
$$\textcircled{p} \xrightarrow{i} \xrightarrow{j} = \textcircled{p} \xrightarrow{j}$$

In total this additional arrow remains unnoticed because of the arbitrariness of summation indices. Its power in diagram simplification unfolds in connection with the ε -tensor (see Section 3.3.6): If the ambient space is \mathbb{RP}^2 as it is here, the ε -tensor is a threefold co- or a threefold contravariant tensor. The speciality is its total skew-symmetry: interchanging two indices flips the sign. Additionally fixing $\varepsilon_{123} = 1$ totally determines this

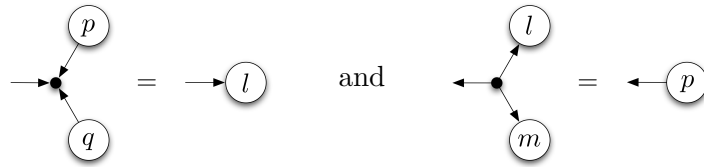
tensor:

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = -\varepsilon_{132} = -\varepsilon_{321} = -\varepsilon_{213} = 1 .$$

All other components vanish due to the skew-symmetry. The contravariant ε -tensor ε^{klm} is defined analogously. We represent the co- and contravariant ε -tensor in this order by



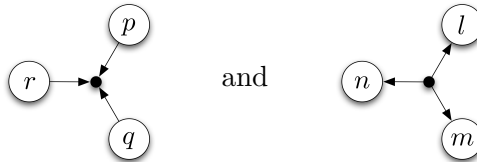
Let us state a few results concerning the ε -tensor. For further descriptive explanations we recommend [41]. If $l = p \times q$ results from the *join*-operation of two points and if $p = l \times m$ results from the *meet*-operation of two lines, we get:



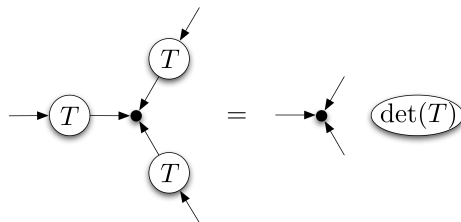
The first diagram represents an object with an incoming arrow. Thus it is contravariant, “line-like” so to say. The second diagram represents an object with an outgoing arrow, a “point-like” object. A simple calculation shows that, performing the necessary summations, we obtain exactly the stated vector products. Another remarkable identity is:

$$p_i q_j r_k \varepsilon^{ijk} = \det \begin{pmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_2 \end{pmatrix} .$$

The ε -tensor simply selects all six summands that are present in the evaluation of a determinant. In diagram notation determinants of three points or three lines read:



Now knowing about the effect of δ - and ε -tensor, we focus on their behavior under transformations. The δ -tensor has one co- and one contravariant index and thus transforms as we already have seen in the left of Figure 3.3. It does not change at all. The ε -tensor also demonstrates a remarkable transformation behavior: Applying a projective transformation T to an ε -tensor results in a multiple of the ε -tensor:



The covariant ε -tensor transforms analogously: When ε^{klm} is transformed as above, ε_{pqr} transforms by T^{-T} . The arrows are reverted and the factor becomes $\frac{1}{\det(-T)}$. This behavior gives rise to

Definition 3.4 Let A be a tensor with $A_{i_1, i_2, \dots}^{j_1, j_2, \dots}$ and T be the matrix representation of a transformation of space. Let $\tilde{A}_{i_1, i_2, \dots}^{j_1, j_2, \dots}$ be the transformed version of the tensor A . Then A is called **relative invariant tensor of weight g** if $A = \tilde{A} \cdot \det(T)^g$. In case $g = 0$, A is called an **absolute invariant tensor**.

As we have seen ε^{klm} is a relative invariant tensor of weight one. In projective geometry we do not distinguish between multiples. Doing so with the ε -tensor, it becomes an object which is not changed by any projective transformation. This property makes the ε -tensor a key ingredient in stating projective invariants.

3.2.2. Closed diagrams and projective invariants

The invariants we have met so far and the ones we will meet are all represented by a single or a linear combination of closed diagrams, i.e. there are no dangling arrows. In fact linear combinations of closed diagrams are invariants (cf. [41]) and the converse also holds due to the first fundamental theorem of invariant theory (see [25]). (Beware that the role of T and $T^* = T^{-T}$ is swapped in [41].)

Theorem 3.2 Let O be a collection of tensors representing geometric objects (like curve-tensors). Let \mathcal{D} be a closed diagram consisting of tensors in O and ε -tensors. Let there be k covariant and l contravariant ε -tensors in \mathcal{D} . If the space is transformed by the use of a matrix T , the diagram with the transformed tensors evaluates to $\det(T)^{l-k}$ times the diagram \mathcal{D} .

PROOF Scalar tensors are not affected by transformations. All other tensors get transformed according to the number and location of their indices. Lower covariant indices imply a transformation by T and upper contravariant indices imply transformations by T^{-T} . (We also introduce transformation tensors around the ε -tensors.) Speaking diagrammatically: Entries $\bigcirc \leftarrow$ get a T attached: $\bigcirc \leftarrow \textcircled{T} \leftarrow$ and exits $\bigcirc \rightarrow$ get a $T^* = T^{-T}$: $\bigcirc \rightarrow \textcircled{T^*} \rightarrow$. But in a closed diagram each arrow is an entry as well as an exit to corresponding nodes. Altogether an arrow is replaced by a T^{-T} - and a subsequent T -tensor and the diagrams are identical. All tensors in \mathcal{D} originating from O are thus surrounded by corresponding transformations. The entities of each such tensor with the corresponding transformations may be labeled as transformed tensors. The only remaining T and T^{-T} tensors surround the ε -tensors: The T^{-T} -tensors are grouped around ε^{klm} and the T -tensors around ε_{pqr} . Due to the transformation behavior of the ε -tensors the transformation results in a multiplication of the original diagram by the determinant of T to the power of $l - k$. Thereby k and l are the numbers of co- and contravariant ε -tensors used. Per construction no T or T^{-T} remains in our diagram. \square

For a detailed example on this theorem we refer to [41].

The next step is to reformulate the first fundamental theorem of invariant theory. It ensures the converse of the above theorem: Every polynomial invariant may be written in diagram notation using only our geometric objects and ε -tensors. For a detailed proof we refer to [25] and [39].

Theorem 3.3 *Let O be a collection of tensors, i.e. a collection of coordinate-collections with specified transformation behavior, representing geometric objects.*

Let $\varphi(\tilde{O}) = \det(T)^g \cdot \varphi(O)$ be a polynomial invariant of the objects in O with respect to projective transformations. Thereby $T \in GL(3)$ is arbitrary and \tilde{O} is the set of the by T transformed Objects in O . Then φ may be written as a linear combination of tensor-products of tensors in O and co- and contravariant ε -tensors.

According to this theorem, the set of all polynomial invariants of objects in O is equivalent to the set of all diagrams containing tensors from O and ε -tensors.

3.2.3. Diagrams and (non)-Euclidean invariants

Diagrams can also be used to represent invariants under other transformation groups than the projective group (cf. [41]). According to Felix Klein's Erlanger Program, Euclidean and non-Euclidean geometries can be considered as special cases of projective geometry, with a conic section - the fundamental object - attached to it (cf. [39]). Thus according to Felix Klein, we just have to add tensors to our diagram-ingredient-list, which represent these fundamental objects: Building diagrams consisting of object-tensors like a curve-tensor and ε -tensors together with the fundamental-form-tensors gives an invariant expression.

We consider points represented by homogeneous coordinate vectors (x, y, z) . In the standard embedding on the $z = 1$ plane Euclidean transformations are all those projective transformations that leave a certain pair of points $\{I, J\}$ invariant as a pair. These two points have the complex coordinates $I = (i, 1, 0)$ and $J = (-i, 1, 0)$. Thus they both lie on the line at infinity and both of them have complex coordinates. Rotations and translations leave each of the two points invariant, while reflections and glide reflections interchange them. As common in invariant theory, we will also consider similarities and similarities followed by (glide-)reflections as Euclidean transformations. They also leave the pair $\{I, J\}$ invariant. In the set-up of Cayley-Klein geometry the pair $\{I, J\}$ plays the role of a dual degenerate conic, the fundamental object. The corresponding primal conic is the (doubly covered) line at infinity. All tangents to this fundamental object will pass through either of the points I or J . A Euclidean (relative) invariant of a collection O of geometric objects is a polynomial in the coordinate entries of O such that $f(\tau(O)) = \det(T)^g f(O)$ for every Euclidean transformation τ . Here T is the corresponding 3×3 transformation matrix and g is a suitable integer.

If we want to represent Euclidean invariants on the level of tensors (or tensor diagrams), we could draw diagrams involving the tensors of O and I and J . Then we must take care of using I and J in a symmetric way in the diagram. However, we could also combine I and J immediately in a symmetric and an antisymmetric manner and introduce two (covariant!) quadratic forms E and A (compare [53] and [54]). In matrix/vector language the forms are defined as

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{IJ^T + JI^T}{2} \quad \text{and} \quad A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{IJ^T - JI^T}{2i} .$$

Diagrammatically we get:

$$\begin{aligned}
 \begin{array}{c} \textcircled{2} \textcircled{E} \\ \swarrow \searrow \end{array} &= \begin{array}{c} \textcircled{I} \textcircled{J} \\ \downarrow \downarrow \end{array} + \begin{array}{c} \textcircled{J} \textcircled{I} \\ \downarrow \downarrow \end{array} \quad \text{and} \\
 \begin{array}{c} \textcircled{2i} \textcircled{A} \\ \swarrow \searrow \end{array} &= \begin{array}{c} \textcircled{I} \textcircled{J} \\ \downarrow \downarrow \end{array} - \begin{array}{c} \textcircled{J} \textcircled{I} \\ \downarrow \downarrow \end{array} \quad .
 \end{aligned} \tag{3.14}$$

The tensors E and A are covariant (they have two outgoing arrows). Since I and J remain invariant (up to a factor) under rotations, the E - and A -tensor are also invariant (up to a certain factor). For rotations expressed by a projective 3×3 matrix T we have

$$\begin{aligned}
 \leftarrow \textcircled{T^*} \leftarrow \textcircled{E} \rightarrow \textcircled{T^*} \rightarrow &= \leftarrow \textcircled{E} \rightarrow \textcircled{\det(T)^{-2}} \quad \text{and} \\
 \leftarrow \textcircled{T^*} \leftarrow \textcircled{A} \rightarrow \textcircled{T^*} \rightarrow &= \leftarrow \textcircled{A} \rightarrow \textcircled{\det(T)^{-2}} \quad .
 \end{aligned}$$

For orientation reserving Euclidean transformations (they interchange I and J) the situation is slightly different. Here E becomes transformed as above. However, A will change its sign.

Thus having a tensor-diagram containing $\textcircled{I} \rightarrow$ and $\textcircled{J} \rightarrow$, we can substitute them by using $\leftarrow \textcircled{E} \rightarrow$ and $\leftarrow \textcircled{A} \rightarrow$:

$$\begin{aligned}
 \begin{array}{c} \textcircled{I} \textcircled{J} \\ \swarrow \searrow \\ \textcircled{X} \end{array} &= \begin{array}{c} \textcircled{I} \textcircled{J} \\ \swarrow \searrow \\ \frac{1}{2} \textcircled{X} \end{array} + \begin{array}{c} \textcircled{J} \textcircled{I} \\ \swarrow \searrow \\ \frac{1}{2} \textcircled{X} \end{array} + \begin{array}{c} \textcircled{I} \textcircled{J} \\ \swarrow \searrow \\ \frac{1}{2} \textcircled{X} \end{array} - \begin{array}{c} \textcircled{J} \textcircled{I} \\ \swarrow \searrow \\ \frac{1}{2} \textcircled{X} \end{array} \\
 &= \begin{array}{c} \textcircled{E} \\ \swarrow \searrow \\ \textcircled{X} \end{array} + i \begin{array}{c} \textcircled{A} \\ \swarrow \searrow \\ \textcircled{X} \end{array}
 \end{aligned}$$

Analogously we get

$$\begin{array}{c} \textcircled{J} \textcircled{I} \\ \swarrow \searrow \\ \textcircled{X} \end{array} = \begin{array}{c} \textcircled{E} \\ \swarrow \searrow \\ \textcircled{X} \end{array} - i \begin{array}{c} \textcircled{A} \\ \swarrow \searrow \\ \textcircled{X} \end{array}$$

This means that if X is symmetric with respect to the I - and J -tensor, $\textcircled{A} \textcircled{X}$ must vanish. This is also clear due to the skew-symmetry of A . Thus we have

Theorem 3.4 *If a diagram-term consists of I - and J -tensors in the symmetry of Equation (3.14), there is an equivalent diagram-expression without these tensors but with an E - and an A -tensor.*

If the number of I - and the number of J -tensors are equal, the I - and J -tensors may be removed completely by application of E - and A -tensors.

If the remaining tensor is symmetric with respect to its I and J pairs, the resulting expression without I- and J-tensors is a linear combination of diagrams, each containing no or an even number of A-tensors. Moreover: Diagrams with an odd number of A-tensors vanish.

In the sense of Cayley-Klein geometries we have

Theorem 3.5 *The set of all closed diagrams consisting of E-, A- and ε -tensors as well as of tensors from a collection O are linear combinations of Euclidean invariants of the objects whose tensor-representants are in O.*

3.3. Diagram laws or how to calculate with diagrams

We have seen how diagrams represent (complicated) analytical expressions. Now we focus on calculation formulas consisting of tensor diagrams. We especially examine diagrams containing curve-tensors. As we have seen: The “language” of real plane algebraic curves is also complex. Thus we almost always state our diagrams for complexified curves. The conversion from real to complex represented curves in tensor-notation will be examined before we focus on diagrammatical laws.

3.3.1. Complexification in tensor-notation

In Section 2.3 we got to know a special transformation for points, namely complexification: There

$$T^* = \begin{pmatrix} 1 & i & 0 \\ 1 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and $\det(T^*) = -2i$. Introducing complexification as transformation, our key ingredients for Euclidean invariants I and J transform to $2i$ times the first canonical unit vector $e_1 = (1, 0, 0)^T$ and to $-2i$ times the second canonical unit vector $e_2 = (0, 1, 0)^T$, respectively. Therefore, if we have a closed diagram containing copies of a single curve-tensor, ε -, I- and J-tensors, we can exchange the curve-tensor by its complexified version, I by e_1 and J by e_2 . In the end, up to a certain factor, we have essentially the same but complexified diagram.

For example, let us take the curve tensor Q_{kl} representing the conic

$$a_{11}(x^1)^2 + 2a_{21}(x^1)(x^2) + a_{22}(x^2)^2 + 2a_{13}(x^1)(x^3) + 2a_{23}(x^2)(x^3) + a_{33}(x^3)^2 = 0 .$$

Let c_{kl} be the corresponding complexified coefficients of this conic. Plugging I and J into Q we get the identity $\textcircled{I} \rightarrow \textcircled{Q} \leftarrow \textcircled{J} = 4 \cdot \textcircled{e_1} \rightarrow \textcircled{\hat{Q}} \leftarrow \textcircled{e_2} = 4 \cdot c_{11}$. Thereby \hat{Q} denotes the tensor with the complexified coefficients. It is useful to note that allowing asymmetric usage of I and J permits filtering leading coefficients:

$$\textcircled{I} \rightarrow \textcircled{Q} \leftarrow \textcircled{I} = -4c_{20} \quad \text{and} \quad \textcircled{J} \rightarrow \textcircled{Q} \leftarrow \textcircled{J} = -4c_{02} .$$

Taking a curve tensor, and using only I- and J-tensors, the number of I-tensors, say k , and the number of J-tensors, say l , essentially lead to c_{kl} .

In the same manner, we can complexify E -, A - and ε -tensors. The transformation of the latter is clear by Section 3.2.1: ε -tensors are relative invariants of weight ± 1 . Complexification replaces the covariant ε -tensor by itself multiplied by $\det(T) = \frac{1}{2i}$. The contravariant ε -tensor gets multiplied by $\det(T^*) = 2i$. For the remaining tensors we have

Theorem 3.6

$$\begin{aligned} \text{I} \rightarrow T^* &= \textcircled{2i} \textcircled{e_1} \rightarrow & \text{and} & \quad \text{J} \rightarrow T^* &= \textcircled{-2i} \textcircled{e_2} \rightarrow , \\ \leftarrow T^* \leftarrow E \rightarrow T^* &= \textcircled{2} \leftarrow \widehat{E} \rightarrow & \text{with} & \quad \widehat{E} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \\ \leftarrow T^* \leftarrow A \rightarrow T^* &= \textcircled{-2i} \leftarrow \widehat{A} \rightarrow & \text{with} & \quad \widehat{A} = A . \end{aligned}$$

For the complexified version of replacing I- and J- tensors by E - and A -tensors we get

$$\widehat{E} = \begin{array}{c} \textcircled{\widehat{E}} \\ \swarrow \quad \searrow \\ \textcircled{e_1} \quad \textcircled{e_2} \\ \downarrow \quad \downarrow \end{array} + \begin{array}{c} \textcircled{e_2} \quad \textcircled{e_1} \\ \downarrow \quad \downarrow \end{array} \quad \text{and} \quad \widehat{A} = \begin{array}{c} \textcircled{\widehat{A}} \\ \swarrow \quad \searrow \\ \textcircled{e_1} \quad \textcircled{e_2} \\ \downarrow \quad \downarrow \end{array} - \begin{array}{c} \textcircled{e_2} \quad \textcircled{e_1} \\ \downarrow \quad \downarrow \end{array}$$

3.3.2. Juggling with \widehat{E} -, \widehat{A} - and one curve-tensor


Now that we know what the complexified versions of the E -, A - and curve-tensors look like, we can stick them together and get various beautiful results. Of course they may be obtained analogously by putting E -, A - and (real) curve-tensors G together. In any case the complexified curve tensors will be denoted by \widehat{G} , the components of which are the complex coefficients c_{kl} (see Section 3.1.5).

In this and the next sections we will juggle with multiple tensors in an equal manner: Often we will have n multiple copies of one and the same tensor, say \widehat{E} , connected to the same symmetric tensors in the same way. Thus it will be convenient to introduce a diagrammatical abbreviation. We replace the multiple \widehat{E} -nodes by one \widehat{E}^n -node. Then n times the arrows from \widehat{E} are dangling from the new \widehat{E}^n -node. As an example, we have:

$$\begin{array}{c} \textcircled{\widehat{E}} \\ \swarrow \quad \searrow \\ \textcircled{\quad} \quad \textcircled{\quad} \\ \swarrow \quad \searrow \\ \textcircled{\widehat{E}} \end{array} = \textcircled{\widehat{E}^2}$$

Gluing only \widehat{E} -, \widehat{A} - and \widehat{G} -tensors, we will always get a corresponding analytical expression in the leading coefficients c_{kl} . This is due to the shape of \widehat{E} and \widehat{A} . The degree of the curve, given by the variance of \widehat{G} , is the number of arcs pointing to the node. By algebraic evaluation or by means of formal tensor composition it can be shown that:

$$\textcircled{\widehat{E}} \textcircled{\widehat{G}} = 2 \cdot c_{11} , \quad \textcircled{\widehat{E}^2} \textcircled{\widehat{G}} = 4 \cdot c_{22} \quad \text{and} \quad \textcircled{\widehat{E}^k} \textcircled{\widehat{G}} = 2k \cdot c_{kk} .$$

Due to the total symmetry of \widehat{G} and the skew-symmetry of $\widehat{A} = A$, the expression  must vanish, whatever tensors might also be connected to \widehat{G} .

3.3.3. Formal tensor-decomposition and the role of \widehat{E} and \widehat{A}

For a better understanding of the structure of diagrams we use yet a different notation. Let

$$a^n \sim \underbrace{\widehat{G}_{1,1,\dots,1,2,2,\dots,2}}_{n \text{ times}} = c_{n,d-n} . \quad (3.15)$$

The switch in the notation is a formal trick. This twist is similarly used, especially in older books on invariant theory (cf. [38] or [55]): Interpreting n as an index, a^n fully describes an entry of the \widehat{G} -tensor, provided d is known. Then a^n may have been defined to be equal to a tensor-component, and would consequently have an equality “=” in (3.15) instead of a similarity “ \sim ”. Interpreting n as an exponent, we get n copies of a meaningless formal component a . This a is only meaningful in an expanded and simplified expression, namely $a^n = \prod_{i=1}^n a$. This way we only have a similarity of a^n with the corresponding component of \widehat{G} . As such a^r , in a non-simplified term $a^r \cdot a^{n-r}$, can not be identified with a tensor-component. However, multiplying a term by a^r and subsequent simplification (gathering exponents) such that identification according to Equation (3.15) is possible, has the following effect: The complete term may be interpreted as an expression in components of \widehat{G} , which have at least r indices equal to one. Keeping all this in mind, we will ignore the difference between “=” and “ \sim ”. We simply write “=” instead of “ \sim ”.

Now, if we are juggling with multiple copies of a \widehat{G} -tensor, we need to distinguish the formal components. Those which originate from different copies must also be named differently. We do so by associating each copy of \widehat{G} with a different letter. Having two copies of a \widehat{G} -tensor, one of them gets associated with the formal component a and the other with b . Of course, in a term with gathered exponents we have an equivalence between these letters for any n :

$$a^n = \underbrace{\widehat{G}_{1,1,\dots,1,2,2,\dots,2}}_{n \text{ times}} = b^n \quad \text{but} \quad a \neq b .$$

This notational trick becomes clearer by looking at what an \widehat{E} -tensor is doing, when it is connected to \widehat{G} -tensor(s). First of all, \widehat{E} is creating a sum. In the first summand one index, say i , gets assigned to $i = 1$ and a second index, say j , gets assigned to $j = 2$. The second summand is equal to the first one but with switched roles of i and j . There $i = 2$ and $j = 1$. Consequently, an \widehat{E} -tensor connected to one and the same symmetric tensor, may be described by a multiplication with $(a + a)$. If \widehat{E} is connected to two tensors, we have a multiplication with $(a + b)$. Thus for the diagrams in the last section

$$\begin{aligned} \textcircled{\widehat{E}} \textcircled{\widehat{G}} &= \widehat{E}^{ij} \widehat{G}_{ij} = \widehat{G}_{12} + \widehat{G}_{21} = (a + a) , \\ \textcircled{\widehat{E}^2} \textcircled{\widehat{G}} &= \widehat{E}^{ij} \widehat{E}^{kl} \widehat{G}_{ijkl} = 4\widehat{G}_{1122} = (a + a)(a + a) , \\ \textcircled{\widehat{E}^k} \textcircled{\widehat{G}} &= 2^k \widehat{G}_{11\dots122\dots2} = (a + a)^k = 2^k a^k . \end{aligned}$$

In the second expression $(a+a)(a+a)$, a single factor $(a+a)$ may not be associated with tensor-components, let alone with $\widehat{G}_{12} + \widehat{G}_{12}$. This product first has to be expanded and then each term must be simplified: $(a+a)(a+a) = aa + aa + aa + aa = a^2 + a^2 + a^2 + a^2$. Now $a^2 = \widehat{G}_{1122}$ (see the diagram above).

With two copies of \widehat{G} involved, we get

$$\begin{aligned} \widehat{G} \leftarrow \widehat{E} \rightarrow \widehat{G} &= \widehat{E}^{ij} \widehat{G}_i \widehat{G}_j = \widehat{G}_1 \widehat{G}_2 + \widehat{G}_2 \widehat{G}_1 = (a+b) , \\ \widehat{G} \leftarrow \widehat{E}^2 \rightarrow \widehat{G} &= \widehat{E}^{ij} \widehat{E}^{kl} \widehat{G}_{ik} \widehat{G}_{jl} \\ &= \widehat{G}_{11} \widehat{G}_{22} + \widehat{G}_{12} \widehat{G}_{21} + \widehat{G}_{21} \widehat{G}_{12} + \widehat{G}_{22} \widehat{G}_{11} \\ &= a^2 b^0 + ab + ab + a^0 b^2 = (a+b)^2 , \\ \widehat{G} \leftarrow \widehat{E}^3 \rightarrow \widehat{G} &= \sum_{k=0}^d \binom{d}{p} a^k b^{d-k} = (a+b)^d . \end{aligned}$$

Taking an \widehat{A} -tensor connected to one or two \widehat{G} -tensors, we get a difference: In the minuend one index, say i , gets assigned to $i = 1$ and a second index, say j , gets assigned to $j = 2$. The subtrahend is equal to the minuend but with switched roles of i and j , where $i = 2$ and $j = 1$. Thus analogous to the \widehat{E} -tensor, the \widehat{A} -tensor may be described by a multiplication with $(a-a)$ or $(a-b)$, depending on whether \widehat{A} is connected to one or two different symmetric \widehat{G} -tensors. If \widehat{A} is connected to one and the same symmetric tensor, \widehat{A} produces two equal expressions to be subtracted from each other.

Thus $\widehat{A} \leftarrow \widehat{G} \rightarrow \widehat{G}$ vanishes as we have seen at the end of the last section. With two copies of a \widehat{G} -tensor we get:

$$\begin{aligned} \widehat{G} \leftarrow \widehat{A} \rightarrow \widehat{G} &= \widehat{G}_k \widehat{G}_l \widehat{A}^{kl} = \widehat{G}_1 \widehat{G}_2 - \widehat{G}_2 \widehat{G}_1 = 0 = (a-b)^1 , \\ \widehat{G} \leftarrow \widehat{A}^2 \rightarrow \widehat{G} &= \widehat{G}_{11} \widehat{G}_{22} - 2\widehat{G}_{12} \widehat{G}_{21} + \widehat{G}_{22} \widehat{G}_{11} = 2c_{20}c_{02} - 2(c_{11})^2 = (a-b)^2 , \\ \widehat{G} \leftarrow \widehat{A}^3 \rightarrow \widehat{G} &= \widehat{G}_{111} \widehat{G}_{222} - 3\widehat{G}_{112} \widehat{G}_{122} + 3\widehat{G}_{122} \widehat{G}_{211} - \widehat{G}_{222} \widehat{G}_{111} = 0 = (a-b)^3 , \\ \widehat{G} \leftarrow \widehat{A}^d \rightarrow \widehat{G} &= \sum_{k=0}^d \binom{d}{p} a^k b^{d-k} (-1)^{d-k} = (a-b)^d . \end{aligned}$$

It is interesting to note that by using exactly one or three \widehat{A} -tensors, the corresponding diagram evaluates to zero. This generalizes to

Theorem 3.7 For odd d

$$\widehat{G} \leftarrow \widehat{A}^d \rightarrow \widehat{G} = 0 .$$

PROOF The reason for this lies in the coefficient structure of $(a-b)^d$. For odd d it is skew-symmetric: The coefficient of $a^k b^{d-k}$ is $\binom{d}{k} (-1)^{d-k}$, which is exactly the negative of the coefficient $\binom{d}{d-k} (-1)^k$ of $a^{d-k} b^k$. Consequently summation yields

$$\widehat{G} \leftarrow \widehat{A}^d \rightarrow \widehat{G} = \sum_{k=0}^d \binom{d}{p} a^k b^{d-k} (-1)^{d-k} = \sum_{k=0}^{\lfloor \frac{d}{2} \rfloor} \binom{d}{p} (-1)^{d-k} (a^k b^{d-k} - a^{d-k} b^k) .$$

On the right hand side of this equation we only have expanded and simplified terms. Identification with tensor-components shows $a^k b^{d-k} = a^{d-k} b^k$ for any k . \square

Finally, mixing \widehat{E} - and \widehat{A} -tensors we get

Theorem 3.8

$$\begin{array}{c} \widehat{G} \\ \swarrow \quad \searrow \\ \widehat{E}^p \quad \widehat{A}^q \\ \swarrow \quad \searrow \\ \widehat{G} \end{array} = (a+b)^p (a-b)^q . \tag{3.16}$$

Summarizing the above, we can see by those two theorems what formal decomposition of \widehat{G} is capable of and what it can not do: Two \widehat{G} -tensors connected by several \widehat{E} - and \widehat{A} -tensors may be described by $(a+b)^p (a-b)^q$. This is just a linear-combination of all $a^k b^{d-k}$ ($0 \leq k \leq d$) with corresponding coefficients. The $a^k b^{d-k}$ may directly be associated with $c_{k,d-k} c_{d-k,k} = |c_{k,d-k}|^2$. Here we can see the redundancy contained in $(a+b)^p (a-b)^q$: Analogously to $c_{k,d-k} c_{d-k,k} = c_{d-k,k} c_{k,d-k}$ we have $a^k b^{d-k} = a^{d-k} b^k$. This redundancy was introduced by associating different names to different copies of one and the same tensor \widehat{G} . But the possibility to distinguish between the two tensors is necessary when calculating on the formal level. For example:

$$2c_{11} = \begin{array}{c} \widehat{E} \\ \swarrow \quad \searrow \\ \widehat{G} \end{array} = (a+a) \neq (a+b) = \begin{array}{c} \widehat{G} \\ \swarrow \quad \searrow \\ \widehat{E} \end{array} = 2c_{10}c_{01} .$$

By the use of Equations (3.15) and (3.16) we are able to easily transform diagrams with two copies of a curve tensor \widehat{G} into a corresponding linear combination of curve-coefficients.

3.3.4. Excursus: Generalizing Pascal-Triangles or the Pascal-Pyramid

We have seen the importance of the formulas $(a+b)^p (a-b)^q$ in Theorem 3.8. The coefficient structure of the corresponding expanded expressions exhibits a lot of structure. Now let us to have a closer look at the coefficient structure and examine its attributes. This will help us later on to determine linear-combinations of diagrams in \widehat{E} -, \widehat{A} - and \widehat{G} -tensors being equal to a single product and rotational type-2 expression $c_{kl}c_{lk}$. Again, after a little technical work, we will see a whole net of beautiful mathematical connections. We will see that for a fixed l the coefficients of $(a+b)^{l-q} (a-b)^q$ with $0 \leq q \leq l$ form a basis of \mathbb{R}^{l+1} respectively \mathbb{C}^{l+1} . Moreover, the dual basis and thus the coefficients of the linear-combinations we are looking for, will be easily accessible. By convention, the coefficients of $(a+b)^p (a-b)^q$ will be named $C_i^{(p,q)}$ with

$$(a+b)^p (a-b)^q = \sum_{i=0}^{p+q} C_i^{(p,q)} a^{p+q-i} b^i .$$

These coefficients $C_i^{(p,q)}$ are closely related to those of Pascal's Triangle. They may be nicely arranged in a pyramid, exhibiting recursive building-principles analogous to Pascal's Triangle. This is why the section-subtitle is Pascal-Pyramid.

Building principles of $(a+b)^p$ and $(a-b)^q$

Let us start with the (original) Pascal-Triangle: It is a geometric arrangement of binomial coefficients in a triangle (cf. [56] and Figure 3.8.a). In each row $l = p$ (starting with

$l = 0$) we have the coefficients $C_i^{(p,0)}$ of

$$(a + b)^p = \sum_{i=0}^p \binom{p}{i} a^{p-i} b^i = \sum_{i=0}^p C_i^{(p,0)} a^{p-i} b^i$$

and $C_i^{(p,0)} = \binom{p}{i}$ are binomial coefficients. The well-known building principle of Pascal's Triangle is that each coefficient $C_i^{(p,0)}$ with $p \geq 1$ is the sum of its two neighboring coefficients one line above. This recursion is due to the equivalence with a multiplication by $(a + b)$ and may be generalized:

Lemma 3.9 *If for a set of coefficients C_i and D_i*

$$(a + b) \cdot \sum_{i=0}^l C_i a^{l-i} b^i = \sum_{i=0}^{l+1} D_i a^{l+1-i} b^i ,$$

then with $C_{-1} = C_{l+1} = 0$ we get $D_i = C_{i-1} + C_i$ for $0 \leq i \leq l + 1$.

Consequently for the coefficients $C_i^{(p,q)}$ of $(a + b)^p (a - b)^q$ to arbitrary $p, q \in \mathbb{N}_0$:

$$C_i^{(p+1,q)} = C_{i-1}^{(p,q)} + C_i^{(p,q)} \quad \text{with} \quad C_{-1}^{(p,q)} = C_{l+1}^{(p,q)} = 0 . \quad (3.17)$$

PROOF With $C_{-1} = C_{l+1} = 0$:

$$\begin{aligned} (a + b) \cdot \sum_{i=0}^l C_i a^{l-i} b^i &= \sum_{i=0}^l C_i a^{l+1-i} b^i + \sum_{i=1}^{l+1} C_{i-1} a^{l+1-i} b^i \\ &= \sum_{i=0}^{l+1} (C_i + C_{i-1}) a^{l+1-i} b^i = \sum_{i=0}^{l+1} D_i a^{l+1-i} b^i . \quad \square \end{aligned}$$

We call Equation (3.17) our “+ **building-principle**”, because in combination with $C_0^{(0,0)} = 1$, the coefficient-triangle to $(a + b)^l$ is totally described.

Analogous to $(a+b)^p$ we also get a coefficient arrangement for $(a-b)^q$ and a corresponding “- **building-principle**” (see (3.18)):

Lemma 3.10 *If for a set of coefficients C_i and D_i*

$$(a - b) \cdot \sum_{i=0}^l C_i a^{l-i} b^i = \sum_{i=0}^{l+1} D_i a^{l+1-i} b^i ,$$

then with $C_{-1} = C_{l+1} = 0$ we get $D_i = -C_{i-1} + C_i$ for $0 \leq i \leq l + 1$.

Consequently for the coefficients $C_i^{(p,q)}$ of $(a + b)^p (a - b)^q$ to arbitrary $p, q \in \mathbb{N}_0$:

$$C_i^{(p,q+1)} = -C_{i-1}^{(p,q)} + C_i^{(p,q)} \quad \text{with} \quad C_{-1}^{(p,q)} = C_{l+1}^{(p,q)} = 0 . \quad (3.18)$$

PROOF With $C_{-1} = C_{l+1} = 0$:

$$\begin{aligned} (a - b) \cdot \sum_{i=0}^l C_i a^{l-i} b^i &= \sum_{i=0}^l C_i a^{l+1-i} b^i - \sum_{i=1}^{l+1} C_{i-1} a^{l+1-i} b^i \\ &= \sum_{i=0}^{l+1} (C_i - C_{i-1}) a^{l+1-i} b^i = \sum_{i=0}^{l+1} D_i a^{l+1-i} b^i . \quad \square \end{aligned}$$

The coefficient triangle to $(a - b)^q$ can be seen in Figure 3.8.f. It is a kind of a generalized Pascal-Triangle. Starting with the tip $C_0^{(0,0)} = 1$, we may generate this triangle just by using Equation (3.18).

Analyzing $(a + b)^p(a - b)^q$

Now we may generalize the Pascal-Triangle to obtain many coefficient triangles with some of the top lines cut off, coefficient trapezes so to say. We can do so in two different ways: The arrangement may correspond to $(a + b) \cdots (a - b)^q$ for a fixed q or to $(a + b)^p(a - b) \cdots$ for a fixed p and variable remaining exponent. In the first case we take the coefficients $C_i^{(0,q)}$ corresponding to $(a - b)^q$ and generate new coefficients of our arrangement by our “+ building principle” (3.17). This way the $C_i^{(0,q)}$ fill the topmost row in our triangle, which is not cut off. Figure 3.8.b-e show this for $q \in \{1, 2, 3, 4\}$. In the second case we take the coefficients $C_i^{(p,0)}$ corresponding to $(a + b)^p$ and our “- building principle” (3.18). Figure 3.8.g-j show this for $p \in \{1, 2, 3, 4\}$. In any case, we get an arrangement which is triangular in shape with the top q , respectively p , rows missing.

Looking at the coefficient pattern of $(a + b) \cdots (a - b)^q$ (Figure 3.8.a-e) we can see symmetric and skew-symmetric arrangements with the symmetry-axis passing through the tip of the triangle. In the patterns of $(a + b)^p(a - b) \cdots$ (Figure 3.8.f-j) we have alternating symmetric and skew-symmetric lines. In any case the elements on the symmetry-axis vanish in the skew-symmetric lines. Let us gather these observations in

Lemma 3.11 *The coefficient-pattern for $(a + b)^p(a - b)^q$ is skew-symmetric if q is odd and symmetric for even q . Thus*

$$C_{p+q-i}^{(p,q)} = (-1)^q C_i^{(p,q)} . \quad (3.19)$$

PROOF According to the “+ building principle” (3.17), the pattern for $(a + b) \cdots (a - b)^q$ inherits its symmetry from the symmetry of the line corresponding to $(a - b)^q$. We may think of the coefficient-arrangement corresponding to $(a + b) \cdots (a - b)^q$ as a line with coefficients from $(a - b)^q$. Attached to each coefficient there is a whole Pascal-Triangle, scaled by the coefficient the triangle is attached to. Summation over the overlap gives the desired coefficient pattern. (In Figure 3.8.c we have three overlapping triangles $\begin{matrix} 1 & & 1 \\ 1 & 2 & 1 \\ 1 & & 1 \end{matrix}$ scaled by 1, -2 and 1, respectively.) Algebraically (compare with Equation (A.1) in the appendix):

$$(a + b)^p(a - b)^q = \sum_{i=0}^{p+q} \underbrace{\sum_{r=\max(0,i-q)}^{\min(p,i)} \binom{p}{r} \binom{q}{i-r} (-1)^{i-r} a^{p+q-i} b^i}_{=C_i^{(p,q)}} ,$$

$$C_{p+q-i}^{(p,q)} = \sum_{r=\max(0,p-i)}^{\min(p,p+q-i)} \binom{p}{p-r} \binom{q}{-p+i+r} (-1)^{p+q-i+r}$$

$$\stackrel{s=p-r}{=} \sum_{s=\max(0,i-q)}^{\min(p,i)} \binom{p}{s} \binom{q}{i-s} (-1)^{q-i+s} = (-1)^q C_i^{(p,q)} . \quad \square$$

Analogously we may get a correlation between the coefficients of $(a + b)^p(a - b)^q$ and $(a + b)^q(a - b)^p$.

Lemma 3.12 *Corresponding coefficients of $(a + b)^p(a - b)^q$ and $(a + b)^q(a - b)^p$ differ at most by a factor of -1 :*

$$C_i^{(p,q)} = (-1)^i C_i^{(q,p)} \quad (3.20)$$

PROOF

$$\begin{aligned} C_i^{(q,p)} &= \sum_{r=\max(0,i-p)}^{\min(q,i)} \binom{q}{r} \binom{p}{i-r} (-1)^{i-r} \\ &\stackrel{s=i-r}{=} \sum_{s=\max(0,i-q)}^{\min(p,i)} \binom{q}{i-s} \binom{p}{s} (-1)^s = (-1)^i C_i^{(p,q)}. \quad \square \end{aligned}$$

Building the Pascal-Pyramid

Now that we have the coefficient structure of $(a + b)^p(a - b)^q$ for any given q we may stick these together and form a coefficient pyramid. Figure 3.8.p shows the pyramid viewed from the top. We call it the **Pascal-Pyramid**. The pyramidal structure is formed by successively superimposing the (tip-less) triangles for $(a + b)^x(a - b)^q$ with $x = 0, x = 1, \dots$ in a decent distance from the previous triangle, such that the planes containing the arrangements are parallel to each other. The same is true for the $(a + b)^p(a - b)^x$ -arrangements. This way the original Pascal-Triangle and the triangle to $(a - b)^q$ automatically are opposing triangular faces of the pyramid (see Figure 3.8).

In our pyramid we have a lot of coefficients, which may be generated by the “+ building-principle” as well as with the “- building-principle”. These are exactly those coefficients $C_i^{(p,q)}$ with $p \neq 0 \neq q$.

Slicing the pyramid into horizontal layers

Now we may slice our pyramid into layers. A layer l of the pyramid is formed by all coefficients of $(a + b)^p(a - b)^q$ with constant $l = p + q$. This corresponds to placing the square-face of the pyramid on the ground and slicing horizontally. Layer 0 consists of just the tip of our Pascal-Pyramid and layer l has an $(l + 1) \times (l + 1)$ -coefficient-arrangement (see Figure 3.8.k-o). Taking four neighboring coefficients, which form a square in a layer, these coefficients are related. This relation is due to the connection of these coefficients via the “+ and - building-principle” to one and the same coefficient one layer below:

Lemma 3.13 *For all $p, q \geq 1$ and $0 \leq i \leq p + q$ with $C_{-1}^{(\dots)} = C_{l+1}^{(l-r,r)} = 0$*

$$C_i^{(p,q-1)} = C_{i-1}^{(p-1,q)} + C_i^{(p-1,q)} + C_{i-1}^{(p,q-1)}. \quad (3.21)$$

PROOF Equation (3.17) and (3.18) prove the assumption. □

Thus graphically this lemma tells us: The upper left coefficient of four neighboring coefficients forming a square in a pyramid layer is the sum of the other three coefficients. Let us interpret the coefficients of a layer l as coefficients of a matrix $L^{(l)}$. The components of $L^{(l)}$ are $L_{qi}^{(l)} = C_i^{(l-q,q)}$ with $0 \leq q, i \leq l$. The “+ and - building-principle” can

also be seen in this layer-notation. For any $l \in \mathbb{N}_0$ we have

$$\begin{aligned}
 L^{(l+1)} &= \left(\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & L^{(l)} & & 0 \\ \hline \frac{C_0^{(0,l)}}{2} & \dots & \dots & \frac{C_l^{(0,l)}}{2} \end{array} \right) + \left(\begin{array}{c|ccc} 0 & & & \\ \vdots & & & \\ 0 & & L^{(l)} & \\ \hline \frac{C_0^{(0,l)}}{2} & \dots & \dots & \frac{C_l^{(0,l)}}{2} \end{array} \right) \\
 &= \left(\begin{array}{ccc|c} \frac{C_0^{(l,0)}}{2} & \dots & \dots & \frac{C_l^{(l,0)}}{2} \\ & & & 0 \\ & L^{(l)} & & \vdots \\ & & & 0 \end{array} \right) + \left(\begin{array}{c|ccc} \frac{C_0^{(l,0)}}{2} & \dots & \dots & \frac{C_l^{(l,0)}}{2} \\ 0 & & & \\ \vdots & & & \\ 0 & & -L^{(l)} & \end{array} \right)
 \end{aligned}$$

There are many other relations in the pyramid-layers: The $l - 2$ rows in the middle of layer l contain two copies of layer $l - 2$ or three binomially weighted copies of layer $l - 4$. However, we want to focus on a different property of the pyramid-layers: The linear independence of the row-vectors in a layer.

Theorem 3.14 *The row-vectors of each $L^{(l)}$ form a basis of \mathbb{R}^{l+1} respectively \mathbb{C}^{l+1} and the basis of the dual-space is formed by the column-vectors of $L^{(l)}$:*

$$L^{(l)} \cdot L^{(l)} = 2^l I ,$$

with the $(l + 1) \times (l + 1)$ -unit matrix I .

PROOF The proof of this theorem mainly makes use of Equations (3.17) to (3.21). As it is rather technical, we omit a proof here and refer to Appendix A.3. \square

3.3.5. Juggling with \widehat{E} -, \widehat{A} - and two curve tensors

Let us revisit Theorem 3.8. It said that the diagrams, consisting of two \widehat{G} -tensors connected by several \widehat{E} - and \widehat{A} -tensors, behave like $(a + b)^{d-q}(a - b)^q$:

$$\begin{array}{c} \widehat{G} \\ \widehat{E}^{d-q} \\ \widehat{G} \\ \widehat{A}^q \end{array} = (a + b)^{d-q}(a - b)^q .$$

Expanding this expression we get

$$\sum_{r=0}^d C_r^{(d-q,q)} a^{d-r} b^r = \sum_{r=0}^d C_r^{(d-q,q)} \underbrace{\widehat{G}_{1,\dots,1,2,\dots,2}}_{d-r \text{ times}} \underbrace{\widehat{G}_{1,\dots,1,2,\dots,2}}_{r \text{ times}} = \sum_{r=0}^d C_r^{(d-q,q)} c_{d-r,r} c_{r,d-r} .$$

Thus with the excursus on the Pascal-Pyramid we are now able to state the ‘‘converse’’: the linear-combination of diagrams evaluating to the products $c_{d-r,r} c_{r,d-r}$:

Theorem 3.15 *The product of leading coefficients $c_{d-r,r} c_{r,d-r}$ ($0 \leq r \leq d$) is a Euclidean invariant and*

$$c_{d-r,r} c_{r,d-r} = \begin{array}{c} \begin{array}{ccc} \begin{array}{c} \widehat{G} \\ \widehat{E}^{d-r} \\ \widehat{G} \end{array} \\ \begin{array}{c} e_1 \\ e_1 \end{array} \end{array} \begin{array}{c} \begin{array}{ccc} \widehat{G} \\ \widehat{A}^r \\ \widehat{G} \end{array} \\ \begin{array}{c} e_2 \\ e_2 \end{array} \end{array} = \frac{1}{2^d} \sum_{q=0}^d C_q^{(d-r,r)} \begin{array}{c} \widehat{G} \\ \widehat{E}^{d-q} \\ \widehat{G} \\ \widehat{A}^q \end{array}
 \end{array}$$

PROOF We use the terminology of our excursus on the Pascal-Pyramid. The expression on the right hand-side is a linear combination of $(a+b)^{d-q}(a-b)^q$ ($0 \leq q \leq d$). This corresponds to a linear-combination of lines in the Pascal-Pyramid layer $L^{(d)}$. The left hand-side corresponds to $a^{d-r}b^r$. Thus we are looking for a linear combination of layer-lines equating to the canonical unit-vector $e_r \in \mathbb{R}^{d+1}$. According to Theorem 3.14 the coefficients of the desired linear-combinations are the coefficients $C_q^{(d-r,r)}$ of $(a+b)^{d-r}(a-b)^r$. \square

In Theorem 3.7 we have seen that $\textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{A}^q} \rightarrow \textcircled{\widehat{G}} = 0$ for odd d . The proof used the antisymmetric coefficient structure from $(a-b)^d$. In Lemma 3.11 we have seen that also the coefficients of $(a+b)^{d-q}(a-b)^q$ with odd q have this property. Consequently

$$\textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{E}^{d-q}} \rightarrow \textcircled{\widehat{G}} = 0 \quad \text{for odd } q.$$

Thus the sum in Theorem 3.15 contains zero-summands if $d \geq 1$. This is no contradiction to Theorem 3.14. The summation over all q effectively is a summation over every second line having an even q . These lines have a symmetric structure, namely $C_i^{(d-q,q)} = (-1)^q C_{d-i}^{(d-q,q)} = C_{d-i}^{(d-q,q)}$ and thus all linear combinations have a symmetric structure: This symmetry corresponds exactly to the symmetry in the coefficient-products: $c_{d-q,q}c_{q,d-q} = a^{d-q}b^q = a^q b^{d-q} = c_{q,d-q}c_{d-q,q}$ (be aware of the indices).

Among the expressions $c_{kl}c_{lk}$ we have the interesting special cases, when $k = l$. Section 3.3.2 introduced c_{kk} as a \widehat{G} -node connected with k times an \widehat{E} -tensor. From Theorem 3.15 we get a different representation for the squared diagram. With $k = \frac{d}{2}$ we have

$$\left(\textcircled{\widehat{E}^k} \leftarrow \textcircled{\widehat{G}} \right)^2 = \sum_{q=0, q \text{ even}}^d C_q^{\left(\frac{d}{2}, \frac{d}{2}\right)} \textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{E}^{d-q}} \rightarrow \textcircled{\widehat{G}} \quad (3.22)$$

Especially:

$$\textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{E}} \rightarrow \textcircled{\widehat{E}} \leftarrow \textcircled{\widehat{G}} = \textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{E}^2} \rightarrow \textcircled{\widehat{G}} - \textcircled{\widehat{G}} \leftarrow \textcircled{\widehat{A}^2} \rightarrow \textcircled{\widehat{G}} \quad (3.23)$$

This is so to say the simplest version of (3.22). It encodes $4 \cdot ab = (a+b)^2 - (a-b)^2$. Equation (3.23) is not using any symmetry or other properties of \widehat{G} and is therefore independent of \widehat{G} . Thus

$$\begin{array}{c} \uparrow \\ \textcircled{\widehat{E}} \\ \downarrow \end{array} \quad \begin{array}{c} \uparrow \\ \textcircled{\widehat{E}} \\ \downarrow \end{array} = \begin{array}{cc} \leftarrow \textcircled{\widehat{E}} \rightarrow & \leftarrow \textcircled{\widehat{A}} \rightarrow \\ \leftarrow \textcircled{\widehat{E}} \rightarrow & \leftarrow \textcircled{\widehat{A}} \rightarrow \end{array} -$$

which corresponds to $\widehat{E}^{kl}\widehat{E}^{mn} = \widehat{E}^{km}\widehat{E}^{ln} - \widehat{A}^{km}\widehat{A}^{ln}$. In fact Equation (3.22) may be taken as an application of this rule. Consequently we get the

Theorem 3.16 *The Euclidean invariants $c_{kl}c_{lk}$ of a curve may be written as a linear-combination of diagrams containing only the curve tensors \widehat{G} and \widehat{E} -tensors:*




$$c_{kl}c_{lk} \in \text{span} \left\{ \left(\widehat{E}^r \right) \left(\widehat{G} \right) \left(\widehat{E}^s \right) \left(\widehat{G} \right) \left(\widehat{E}^r \right) \mid 2r + s = k + l \wedge r, s \in \mathbb{N}_0 \right\} .$$

The evaluation of the diagrams in the above theorem gives us:

$$(a + a)^r \cdot (a + b)^s \cdot (b + b)^r = 4^r a^r b^r (a + b)^s = 4^r \sum_{i=0}^s \binom{s}{i} c_{s+r-i, r+i} c_{r+i, s+r-i} .$$

Thus the coefficients are binomial coefficients. For a given variance of \widehat{G} , the relevant coefficient structures correspond to every second line in Pascal's Triangle. Due to the scaling by 4^r we have a coefficient sum of 2^{2r+s} . Thus this sum is equal for all such diagrams using two \widehat{G} -tensors of the same fixed variance.

Example 3.3 *Let \widehat{G} be a curve-tensor to a curve of degree $d = 4$. Then we have:*

diagrams and evaluation	coefficient structure
 $= 2c_{40}c_{04} + 8c_{31}c_{13} + 6c_{22}c_{22}$	1 4 6 4 1
 $= 8c_{31}c_{13} + 8c_{22}c_{22}$	4 · 1 4 · 2 4 · 1
 $= 16c_{22}c_{22}$	16 · 1

The first diagram corresponds to the fifth line in Pascal's Triangle: 1 4 6 4 1 and $1 \cdot c_{40}c_{04} + 4 \cdot c_{31}c_{13} + 6 \cdot c_{22}c_{22} + 4 \cdot c_{13}c_{31} + 1 \cdot c_{04}c_{40}$ is the diagram-evaluation. The next diagram corresponds to four times 1 2 1, etc. . The coefficient sum is always $2^4 = 16$.

Summarizing our results, we have seen that coefficient-products $c_{kl}c_{lk}$ of leading coefficients are Euclidean invariants. They may be written in simple diagrams using only the curve-tensor \widehat{G} and e_1 - and e_2 -tensors or G -, I - and J -tensors. We are also able to restate these invariants without e_1 - and e_2 -tensors (or I - and J -tensors). We can restrict ourselves to the invariant geometric objects of the corresponding Cayley-Klein-Geometry: the dual conics given by E and A or, in the complexified version, by \widehat{E} and \widehat{A} . In the connection between $c_{kl}c_{lk}$ and the diagrams with only \widehat{G} -, \widehat{E} - and \widehat{A} -tensors binomial coefficients, or generalized ones from the Pascal-Pyramid, can be found almost everywhere. Finally we are able to state these invariants as linear-combination of diagrams which dispense of \widehat{A} -tensors.

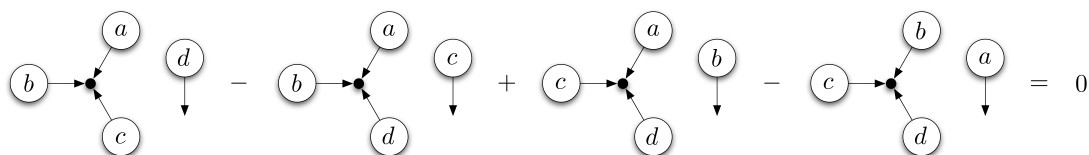
The linear-combinations of diagrams containing only \widehat{E} - and \widehat{G} -tensors evaluating to absolute values of leading coefficients are not far away. Theorem 3.15 in combination with Equation 3.22 would lead to corresponding expressions. We will omit this very special replacement-examination and focus on more general ones.

3.3.6. More replacement rules

The representation of invariants is by far not unique. We have already seen a replacement rule linking I-, J-, E- and A-tensors. But also projective invariants may be represented by several linear combinations of diagrams, which look completely different. This effect is common to invariant theory. For instance, in the classical projective invariant theory there exist certain syzygies. Grassmann-Plücker relations may, for example, be written as always vanishing expressions, independent of the value of the objects involved in the expression. Adding such a syzygy to an invariant does not change it as such but it changes its concrete algebraic appearance. On the level of tensor diagrams the structural components of a syzygy can be nicely separated from its coordinate entries. For instance the Grassmann-Plücker relation

$$\det(a, b, c)d - \det(a, b, d)c + \det(a, c, d)b - \det(b, c, d)a = 0$$

translates to the tensor diagram equation:



Here sub-diagrams in each summand have to be multiplied. The four letters a, b, c, d occur in each single summand. And each single summand has exactly one outgoing arrow. This is also the requirement on terms to enable tensor-addition (compare with Section 3.1.4). As a matter of fact, the specific points that are used in the summands can be neglected as long as the fact that the ends of the ε -tensors and the free arrows are connected to consistent parts of a surrounding diagram is ensured. If we rewrite the above equation in the tensor language, we get:

$$a^i b^j c^k d^l \varepsilon_{ijk} \delta_l^m - a^i b^j c^k d^l \varepsilon_{ijl} \delta_k^m + a^i b^j c^k d^l \varepsilon_{ikl} \delta_j^m - a^i b^j c^k d^l \varepsilon_{jkl} \delta_i^m = 0$$

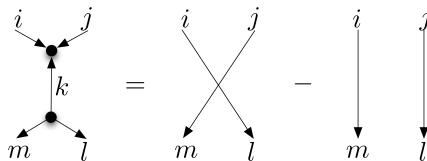
where m plays the role of the outgoing index. We can factor out the points and derive:

$$\varepsilon_{ijk} \delta_l^m - \varepsilon_{ijl} \delta_k^m + \varepsilon_{ikl} \delta_j^m - \varepsilon_{jkl} \delta_i^m = 0 .$$

This is the essential structural part of the Grassmann-Plücker relation. It tells us that for any diagram that involves an ε -tensor and a disjoint arrow, we get an equivalent diagram by summing three copies of the original diagram in which the ε -tensor and the arrow have been suitably rewired. One might wonder about the use of this replacement of a diagram by the sum of three diagrams with equivalent complexity. A first answer to this is that it imitates addition of a Grassmann-Plücker relation or a similar syzygy to an invariant as mentioned before. This is a powerful operation in classical projective geometry and invariant theory, for instance it is the basis of the straightening algorithm (see [44]). However, there is also a second answer: The above equation is not the only one of a similar type – and there are some of them that greatly help to cut down the complexity of diagrams. For the projective plane the following one is perhaps the most elementary and most useful. It is the so-called ε - δ -rule:

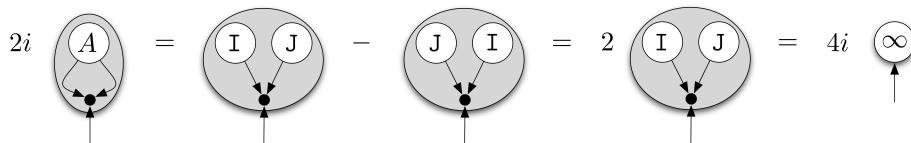
$$\varepsilon_{ijk} \varepsilon^{lmk} = \delta_i^l \delta_j^m - \delta_i^m \delta_j^l .$$

It can be proven by brute-force-expansion of the terms. Translated into diagram language this becomes:

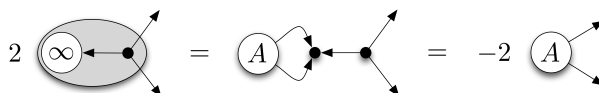


Thus the ε - δ -rule can take a diagram that contains two adjacent (co- and contravariant) ε -tensors and replace it by the difference of two diagrams, in which these ε -tensors have vanished. We will see later how useful this operation is. Perhaps one word of caution is required here: ε -tensors are antisymmetric tensors. This means that interchanging two indices corresponds to a multiplication by -1 . When dealing with diagrams one has to be careful not to interchange two arrows of an ε -tensor without changing the sign of the corresponding diagram. It is even possible to derive the Grassmann-Plücker relations by suitable application of the ε - δ -rule.

Another very important replacement rule concerns the E - and A -tensors. The tensor (A^{kl}) can also be generated in a different way. Let $\infty_m = (0, 0, 1)$ be a covariant tensor that represents the line at infinity. The line at infinity is spanned by I and J (in vector notation we have $I \times J = 2i \cdot \infty$) and we get:



Thus $A^{kl} \varepsilon_{klm} = 2\infty_m$. This can also be proven by expanding the tensors. Furthermore it is easy to prove that $\infty_m \varepsilon^{klm} = -A^{kl}$, as the following diagram shows:



Thus in the presence of ε -tensors, the A tensor and the line at infinity are essentially the same. Diagrams that only involve the A tensor, geometric objects and ε -tensors represent affine invariants.

In diagrams that contain both A tensors and E tensors we always may assume that there is at most one A tensor involved. This is due to the behavior of \widehat{E} and \widehat{A} as shown in the last section. We need only one A -tensor since we can replace any pair of A tensors by a suitable linear combination of a pair of E tensors, as the following theorem shows

Theorem 3.17

$$A^{ij} A^{kl} = E^{il} E^{jk} - E^{ik} E^{jl} .$$

This could again be easily proven either by expansion or by expanding the terms into linear combinations of I and J or by carrying over our results with \widehat{E} and \widehat{A} . Diagrams that only contain E -tensors are invariant under all Euclidean transformations. Diagrams that contain an (odd number of) A tensor(s) are only invariant under rotations. Nevertheless setting such a diagram to zero may result in an Euclidean invariant property.

At this point there is a wide range of possibilities for future work: The diagram notation and the replacement rules enable very elegant visualizations of geometric incidence theorems (for a first approach see [41]). It is also very interesting to examine substructures which make the whole diagram vanish. This way a diagram construction kit can be built up. We abstain from this general diagram examination and refocus our attention on curves, especially on the behavior of curves under transformations. Later we will visualize our results, mainly the invariants, by the use of tensor-diagrams.

4. Curves and selected transformations of the plane

In this chapter we focus on invariants and on features of real algebraic curves. We assume that we have a given coefficient vector. Furthermore we assume that the curve is complexified unless explicitly stated otherwise. A curve is, of course, entirely determined by a coefficient vector, which contains all information. Consequently, features of the curve, such as rotational symmetry, must find their correspondence in the algebraic structure of the coefficients. In this chapter we will investigate how geometric curve features can be extracted from the algebraic description of the curve. The complexified representation allows for a much simpler extraction process. Also normal forms and invariants are easily accessible.

Firstly, however, we will focus on rotations, more precisely on rotations around the origin. We will extract a set of invariants, which allows us to totally reconstruct the curve up to this kind of rotations. Also a normal form with respect to these transformations will be presented and diagrammatical results will be contained. In a second step we mainly study the same issue for translations. Together with rotations we get orientation preserving Euclidean motions. Reflection and scaling are studied afterwards as well as detection of rotational and reflectional symmetry.

4.1. Rotating curves around the origin

A rotation is a Euclidean motion with a fixpoint: the center of rotation. Any rotation with arbitrary center can be carried out in three steps: First apply a translation τ , which maps the center of rotation to the origin. Then perform the rotation by the desired angle φ and apply τ^{-1} afterwards. We will deal with translation later on (see Section 4.2). Thus, when speaking of rotations, we confine ourselves to rotations around the origin and an arbitrary rotation angle φ .

The description of the coefficient-transformation conceptually differs from [46] only in the use of multinomial coefficients. The presented complete set of invariants however is purely on the basis of the complexified curve-coefficients. A proof of completeness and a method for reconstruction of curves with only the invariants at hand is not contained in [46]. Even more: [46] as well as [47] lack the examination of potentially undefined invariants due to vanishing coefficients leading to undefined expressions or division by zero.

4.1.1. Effects of the complexification

The main benefit of complexification is that in this representation a rotation of the plane is very easy to describe. A rotation of the plane is a mapping where

$$\begin{aligned} z \mapsto \tilde{z} &= \tilde{x} + i\tilde{y} = (\cos(\varphi)x - \sin(\varphi)y) + i(\sin(\varphi)x + \cos(\varphi)y) \\ &= (\cos(\varphi) + i\sin(\varphi))(x + iy) = e^{i\varphi}z \end{aligned}$$

and consequently $\bar{z} \mapsto e^{-i\varphi}\bar{z}$.

This means that $z^k\bar{z}^l \mapsto e^{i(k-l)\varphi}z^k\bar{z}^l$, which in itself is also a rotation. Thus rotating a curve $F(z, \bar{z}, h) = 0$ by φ we get

$$\tilde{F}_\varphi(z, \bar{z}, h) = \tilde{F}(z, \bar{z}, h) = F(e^{-i\varphi}z, e^{i\varphi}\bar{z}, h) = \sum_{k+l=d} m_{kl}c_{kl} \cdot e^{i(-k+l)\varphi} \cdot z^k\bar{z}^l h^{d-k-l} = 0$$

and the transformation of the coefficients c_{kl} is particularly simple:

$$c_{kl} \mapsto e^{i(-k+l)\varphi}c_{kl} . \quad (4.1)$$

In matrix notation we have only entries of the form $e^{i\cdot n\varphi}$ with $n \in \mathbb{Z}$:

$$\left(\begin{array}{c|c|c|c|c|c|c|c|c|c|c|} c_{00} & c_{10} & c_{01} & c_{20} & c_{11} & c_{02} & c_{30} & c_{21} & c_{12} & c_{03} & \dots \end{array} \right)^T \mapsto M_\varphi^{-T} c = \underbrace{\left(\begin{array}{c|c|c|c|c|c|c|c|c|c|c|} 1 & & & & & & & & & & \\ \hline & e^{-i\cdot\varphi} & & & & & & & & & \\ \hline & & e^{i\cdot\varphi} & & & & & & & & \\ \hline & & & e^{-i\cdot 2\varphi} & & & & & & & \\ \hline & & & & 1 & & & & & & \\ \hline & & & & & e^{i\cdot 2\varphi} & & & & & \\ \hline & & & & & & e^{-i\cdot 3\varphi} & & & & \\ \hline & & & & & & & e^{-i\cdot\varphi} & & & \\ \hline & & & & & & & & e^{i\cdot\varphi} & & \\ \hline & & & & & & & & & e^{i\cdot 3\varphi} & \\ \hline & & & & & & & & & & \ddots \end{array} \right)} = M_\varphi^{-T} \begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \\ c_{30} \\ c_{21} \\ c_{12} \\ c_{03} \\ \vdots \end{pmatrix} \quad (4.2)$$

$M_\varphi^{-T} = W^{-T}L_m^{-T}W^T$ is the complexified version of L_m^{-T} from (2.9). Thereby multiplication by W^{-T} and W^T causes complexification (see (2.15)). The matrix M_φ^{-T} acts on our complexified curve coefficients c_{kl} and is of diagonal shape. We have:

Theorem 4.1 *Rotating the plane around the origin by an angle φ means rotating the complexified curve-coefficients according to Equation (4.2).*

If a curve is rotated by φ , the rotation of a c_{kl} is determined by the coefficients' indices: $c_{kl} \mapsto e^{i\varphi(-k+l)}c_{kl}$, i.e. c_{kl} is rotated by $\varphi \cdot (-k + l)$. Therefore we call $(-k + l)$ the (oriented) **rotation speed** of c_{kl} .

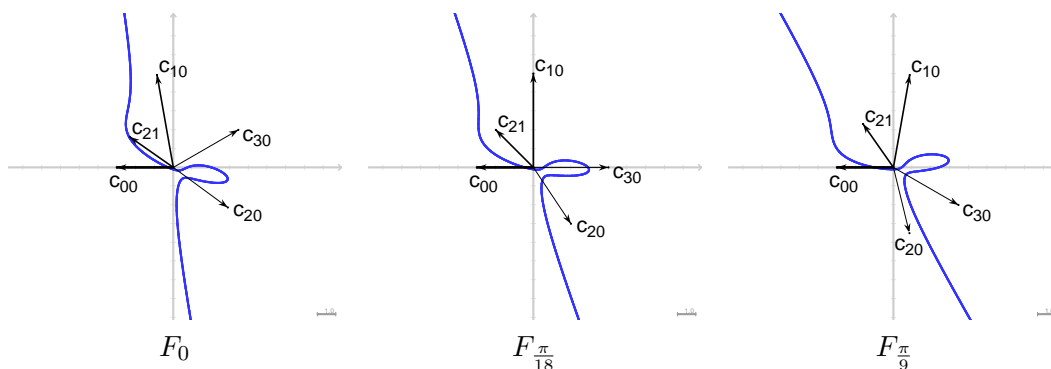


Figure 4.1.: Rotated curves: c_{00} is stationary, c_{10} and c_{21} rotate by $-\varphi$, c_{20} by -2φ and c_{30} by -3φ .

The coefficient triangle introduced in (2.17) on page 27 shows that the order of the columns corresponds to the rotation speeds of the coefficients:

$block_t(c)$	curve coefficients									
$t = 0$					c_{00}					
$t = 1$					c_{10}	c_{01}				
$t = 2$			c_{20}		c_{11}	c_{02}				
$t = 3$		c_{30}		c_{21}		c_{12}		c_{03}		
...
rotation speed	...	-3	-2	-1	0	1	2	3

The transformation matrix M_φ^{-T} from Equation (4.2) is consistent with our complexification process: Conjugate coefficient pairs remain conjugate after the transformation and the real c_{kk} stay real. Comparing (4.2) to the transformation for the coefficients a_{kl} of a curve in our real representation as shown in (2.9), we can directly see how powerful complexification is.

Figure 4.1 shows a curve F and by $\varphi \in \{\frac{\pi}{18}, \frac{\pi}{9}\}$ rotated copies F_φ of F . The coefficients behave according to (4.1) and (4.2) and are indicated by vectors.

4.1.2. Invariants with respect to rotations around the origin

When a curve is rotated around the origin, curve coefficients of the form c_{kk} remain unchanged - invariant so to say. But there are also other expressions in the curve coefficients, which remain unaltered. For example $c_{10} \cdot c_{01} \mapsto e^{-i\varphi} c_{10} \cdot e^{i\varphi} c_{01} = c_{10} \cdot c_{01}$. Any product of curve coefficients is multiplied by $e^{ir\varphi}$ (for suitable r), when the plane is rotated by φ around the origin. This is likewise a rotation. The rotation speed r of such a product is the sum of the rotation speeds of factors of the product. For example

$$c_{kl} \cdot c_{pq} \mapsto \exp(i \underbrace{[(-k + l) + (-p + q)]}_{=r} \varphi) \cdot c_{kl} \cdot c_{pq} = e^{ir\varphi} \cdot c_{kl} \cdot c_{pq} .$$

The rotation speed of c_{kl} is $-k + l$, the one of c_{pq} is $-p + q$ and the rotation speed of the product $c_{kl} \cdot c_{pq}$ is $r = (-k + l) + (-p + q)$.

The clue is now to find products of coefficients which have a vanishing rotation speed. These products must be invariant under rotations of the plane. To find these products, it is helpful to look at the coefficient triangle above, since the coefficients were ordered according to their rotation speed.

Now we have a wide range of invariant expressions. At a first glance we can directly see three types of invariant expressions:

1. form parameters c_{kk}
(rotation speed: $r = -k + k = 0$)
2. lengths $c_{kl}c_{lk} = |c_{kl}|^2$
(rotation speed: $r = -k + l - l + k = 0$)
3. angle relations contained in $(c_{kl})^{\frac{p-q}{m}}(c_{pq})^{\frac{-k+l}{m}}$ with positive $m = \gcd(|-p+q|, |-k+l|)$
(rotation speed: $r = \frac{(p-q)}{m}(-k+l) + \frac{(-k+l)}{m}(-p+q) = 0$)

These expressions are purely formal. Not all of them lead to necessarily well-defined expressions for a specific curve. For example, take a curve with $c_{10} = 0$, then type-3 invariant $c_{20}^1 c_{10}^{-2}$ is not defined. However, these expressions are already complete in the following sense: Given all values to the well-defined expressions for a specific curve \mathcal{C} , a curve-coefficient-vector may be reconstructed. This vector belongs to a curve \mathcal{C}_φ , differing from \mathcal{C} only by a rotation. We will see this in detail in Section 4.1.3.

The above expressions are complete but they are not well-defined for all curves nor are they free of redundancies. They have to be removed in order to get a minimal necessary invariant set.

4.1.3. Reducing rotation invariants

Rotation invariants, graphs and the folded coefficient triangle

Let us interpret the invariant expressions with the help of a graph whose nodes correspond to the entries of our coefficient-triangle. Type-1 invariants, for example, are represented by the coefficients (nodes) on the symmetry axis. All other expressions constitute the set of edges in a graph: An edge connects two coefficients which are the ingredients for one invariant expression. The invariant $c_{kl}c_{lk}$ corresponds to the edge connecting the nodes c_{kl} and c_{lk} . Interpreting the rotation invariants this way we get a graph whose nodes are the curve-coefficients and whose edges correspond to the type-2 and type-3 expressions. Moreover, some nodes are self-connected ($c_{kl} = c_{pq}$) and some edges occur more than once ($c_{lk} = c_{pq}$).

Removing the redundancy in the set of the above expressions translates to removing superfluous components of the graph, the self-connectedness for example. If in the type-3 expressions $c_{kl} = c_{pq}$, then $(c_{kl})^{\frac{k-l}{m}}(c_{kl})^{\frac{-k+l}{m}} = 1$. This is a scalar and thus naturally invariant under rotations. Such an invariant does not carry any information on the curve.

Talking about information contained in the invariants, we have to bear in mind that any coefficient c_{kl} carries the same information as $c_{lk} = \overline{c_{kl}}$. This is also the source of the symmetry of our coefficient triangle. Thus reflectionally symmetric edges in our graph carry the same information because they differ only by conjugation (see left image of Figure 4.2). One of these edges is superfluous. Instead of removing one edge in our graph

we switch to an easier representation, which does not contain the redundant symmetry: Imagine a folded version of the coefficient triangle, where the original symmetry axis is now triangle-border. This way conjugate coefficient-pairs get superimposed and c_{kk} -coefficients stay alone.

Using this folded triangle as basis for a graph, we get nodes of the form c_{kk} and of the form (c_{kl}, c_{lk}) with $k \neq l$. Former symmetric edges are now two edges connecting the same two nodes (see Figure 4.2). We also reintroduced self-connectedness: Type-2 invariants are the products $c_{kl}c_{lk}$, connecting (c_{kl}, c_{lk}) -nodes with itself.

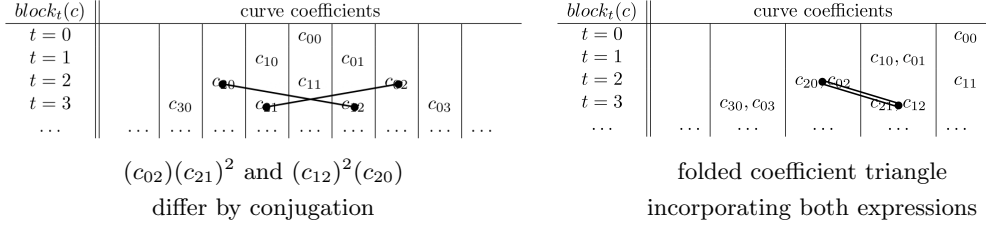


Figure 4.2.: Redundant rotation invariants

Removing redundant invariants

Focusing on the reduction of redundant expressions, we have already removed type-3 expressions with $c_{kl} = c_{pq}$. Also $c_{lk} = c_{pq}$ leads to $(c_{kl})^{\frac{l-k}{m}}(c_{lk})^{\frac{-k+l}{m}} = c_{kl}c_{lk}$, which is of type 2 and therefore excludable from type 3. Other trivialities are: type-2 expressions with $k = l$. These expressions may be obtained by squaring suitable type-1 invariants. Furthermore we may restrict $k > l$ to get rid of doubly listed type-2 expressions. Type-3 invariants containing a coefficient of the form c_{kk} may also be generated only by type-1 expressions: $(c_{kk})^{\frac{p-q}{m}}(c_{pq})^{\frac{-k+k}{m}} = (c_{kk})^{\pm 1}$ or $(c_{kl})^{\frac{p-p}{m}}(c_{pp})^{\frac{-k+l}{m}} = (c_{pp})^{\pm 1}$. Removing these invariants from type 3 isolates the c_{kk} -nodes in our graph. It also rids us of divisions by zero in case a c_{kk} vanishes for a given curve.

But there might be more division-by-zero problems. Let us focus on all invariants relating to the two nodes (c_{rs}, c_{sr}) and (c_{tu}, c_{ut}) . These are $c_{rs}c_{sr}$, $c_{tu}c_{ut}$ and altogether eight type-3 expressions

$$\begin{aligned} & \underbrace{(c_{rs})^{\frac{t-u}{m}}(c_{tu})^{\frac{-r+s}{m}}}_{= \alpha}, \quad \underbrace{(c_{tu})^{\frac{r-s}{m}}(c_{rs})^{\frac{-t+u}{m}}}_{= \frac{1}{\alpha}}, \quad \underbrace{(c_{sr})^{\frac{t-u}{m}}(c_{tu})^{\frac{-s+r}{m}}}_{= \beta}, \quad \underbrace{(c_{tu})^{\frac{s-r}{m}}(c_{sr})^{\frac{-t+u}{m}}}_{= \frac{1}{\beta}}, \\ & \underbrace{(c_{ut})^{\frac{s-r}{m}}(c_{sr})^{\frac{-u+t}{m}}}_{= \bar{\alpha}}, \quad \underbrace{(c_{sr})^{\frac{u-t}{m}}(c_{ut})^{\frac{-s+r}{m}}}_{= \frac{1}{\bar{\alpha}}}, \quad \underbrace{(c_{ut})^{\frac{r-s}{m}}(c_{rs})^{\frac{-u+t}{m}}}_{= \bar{\beta}}, \quad \underbrace{(c_{rs})^{\frac{u-t}{m}}(c_{ut})^{\frac{-r+s}{m}}}_{= \frac{1}{\bar{\beta}}}. \end{aligned}$$

The latter eight invariants correspond to the eight edges between two non-isolated nodes. They may be generated by taking one coefficient of each node and putting them into the context of a type-3 expression. Thereby the order of the coefficients is important due to the antisymmetric choice of the exponents.

Some relations between these expressions, i.e. conjugation or inversion, are already indicated below the brackets. But there are many more, when the type-2 expressions are

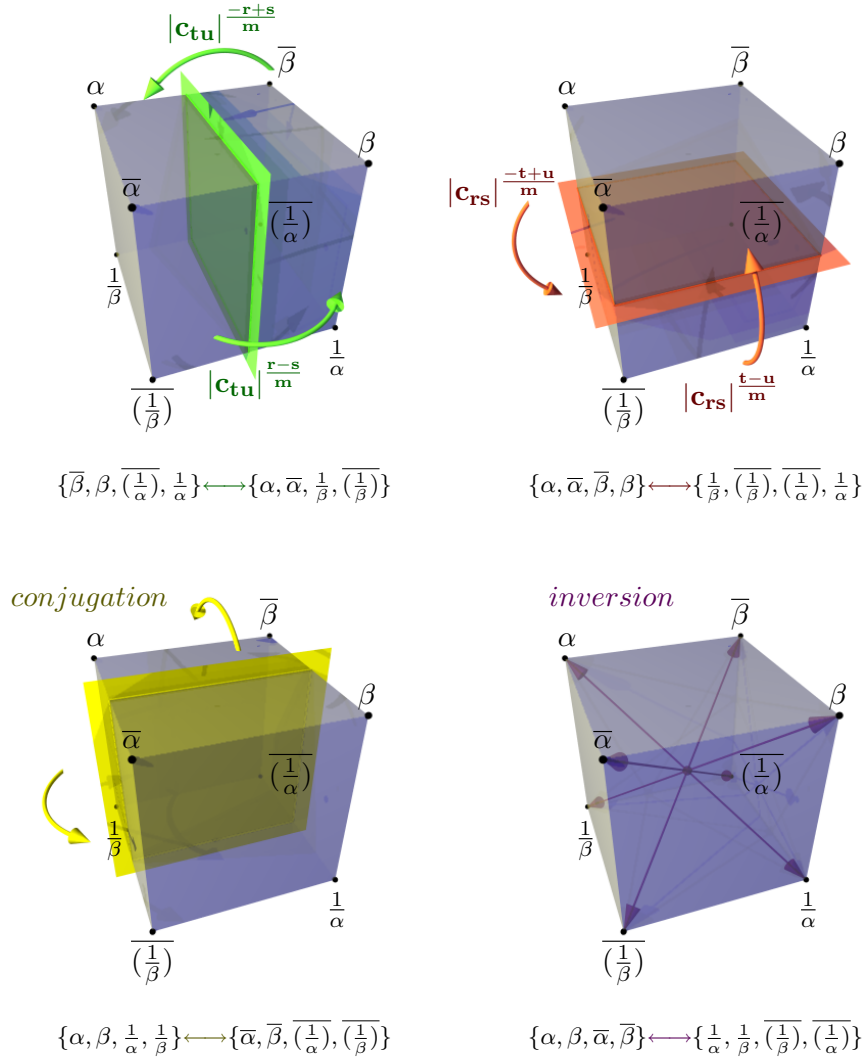


Figure 4.3.: Connections between type-3 expressions

utilized. For example $\alpha \cdot |c_{tu}|^{\frac{r-s}{m}} = \bar{\beta}$ or $\alpha \cdot |c_{rs}|^{\frac{-t+u}{m}} = \frac{1}{\beta}$. Figure 4.3 shows the rich connections between these expressions.

These connections make obvious that the two type-2 expressions $c_{rs}c_{sr}$ and $c_{tu}c_{ut}$ together with only one of the eight type-3 expressions contain all the information: With these invariants at hand, the seven missing expressions may formally be calculated. In our reduction process we are now free to choose one of the eight invariants. This is a great benefit because we may now ensure that we only use invariants with positive exponents. This way type-3 expressions never run into divisions-by-zero.

According to these observations and the equivalence of conjugate expressions with respect to their informational content, we restate our rotationally invariant expressions in reduced form:

reduced 1: expressions of the form c_{kk}

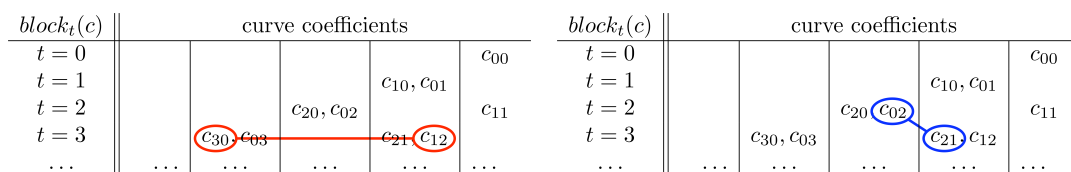
reduced 2: expressions of the form $c_{kl}c_{lk}$ with $k \neq l$

reduced 3: expressions of the form $(c_{kl})^{\frac{p-q}{m}} (c_{pq})^{\frac{-k+l}{m}}$ with $m = \gcd(p - q, -k + l)$

3a: and additionally if $k + l = p + q$, then $p - q > -k + l > 0$

3b: and additionally if $k + l < p + q$, then $p - q, -k + l > 0$

The exclusion of type-3 expressions with $k+l > p+q$ is due to informational conjugation-symmetry. In the same way we could have chosen to exclude $k + l < p + q$. The division into the two sub-cases 3a and 3b is necessary. 3a occupies itself only with coefficients contained in a single block $block_{k+l}(c)$. Therefore 3a-expressions will be called intra-block-invariants and 3b expressions inter-block-invariants. Figure 4.4 shows two expressions in the context of the folded coefficient triangle.



Left: type-3a expression is indicated by an edge between (c_{30}, c_{03}) and (c_{21}, c_{12}) . Due to $-k + l > 0$ and $p - q > 0$ follows: $(k, l) \in \{(0, 3), (1, 2)\}$ and $(p, q) \in \{(3, 0), (2, 1)\}$. With $p - q > -k + l$, the indicated invariant is $(c_{12})^3 (c_{30})^1$.

Right: type-3b expression is indicated by an edge between (c_{20}, c_{02}) and (c_{21}, c_{12}) . From $k + l < p + q$ follows: $(k, l) \in \{(2, 0), (0, 2)\}$. With $-k + l > 0$ and corresponding analysis for (p, q) , the indicated expression is exactly $(c_{02})^1 (c_{21})^2$.

Figure 4.4.: Expressions of type 3a and 3b

Returning to our graph, representing the invariant expressions, we now have a reduced one: The c_{kk} -nodes are isolated. All other nodes are self-connected due to invariant $c_{kl}c_{lk}$ expressions. Additionally any (c_{kl}, c_{lk}) -node is connected with any other (c_{pq}, c_{qp}) -node exactly once. Thus the non- c_{kk} -nodes build a complete graph. The edges of the complete sub-graph, together with the isolated nodes, resemble the set of the reduced expressions from above.

In the context of a given curve some nodes of the complete sub-graph might vanish. If this is the case, then all edges connected to this node represent vanishing type-3 expressions. However, the information whether a coefficient c_{kl} in the sub-graph vanishes (and thus also c_{lk}) is already contained in a corresponding type-2 expression $c_{kl}c_{lk}$. Consequently these type-3 invariants carry no additional information and are superfluous *in the context of this given curve*.

The number of invariant expressions listed in the reduced form may be calculated as follows: If we restrict ourselves to coefficients contained in $block_0(c)$ up to $block_d(c)$, then we have $n_1 = \lceil \frac{d+1}{2} \rceil$ type-1 expressions, $n_2 = \lceil \frac{d}{4}(d+2) \rceil$ type-2 expressions and $n_3 = \binom{n_2}{2}$ type-3 expressions.

Meaning of the invariants

The geometrical meaning of the invariants was already indicated when we listed the expressions for the first time. Now, with the reduced version, it becomes much clearer. In fact, only the reduced version allows such a strict association because the original type-

3 formally contains expressions of the other types. However, the general idea behind it justifies this approach.

Starting with type-1 expressions we have c_{kk} as form parameters or radii. c_{kk} appears in connection with $c_{kk}z^k\bar{z}^k$, which in real representation evaluates to $c_{kk}(x^2 + y^2)^k$. In $c_{11}z\bar{z} = c_{00}$ we have a circle with radius $\frac{c_{00}}{c_{11}}$.

Expressions of type 2 are lengths: $c_{kl}c_{lk} = |c_{kl}|^2$. The corresponding expressions with real curve coefficients are slightly more complicated but still lengths. For example, $4 \cdot |c_{10}|^2 = 4 \cdot c_{10}c_{01} = a_{10}^2 + a_{01}^2 = |(a_{10}, a_{01})|$ or $16 \cdot |c_{20}| = 16 \cdot c_{20}c_{02} = (a_{20} - a_{02})^2 + a_{11}^2 = |(a_{20} - a_{02}, a_{11})|$. We can rewrite our coefficients in polar form $c_{kl} = r_{kl} \cdot \exp(i\varphi_{kl})$ with non-negative r_{kl} . Then type-2 invariants fix r_{kl} because

$$r_{kl} = |c_{kl}| = \sqrt{c_{kl}c_{lk}} . \quad (4.3)$$

Thus in the coefficient-space of curves an invariant of type 2 fixes one degree of freedom. In the same way, we may rewrite invariants of (reduced) type 3 by

$$(c_{kl})^{\frac{p-q}{m}} (c_{pq})^{\frac{-k+l}{m}} = r_{klpq} \cdot \exp(i\varphi_{klpq}) .$$

According to our observations above, $r_{klpq} = (r_{kl})^{p-q}(r_{pq})^{-k+l}$ is already known by type-2 invariants.

So let us focus on the \exp -part and apply the logarithm. Consequently, with the \gcd of the corresponding rotation speeds $m = \gcd(|-p + q|, |-k + l|)$ we have

$$\frac{p-q}{m} \varphi_{kl} + \frac{-k+l}{m} \varphi_{pq} \equiv \varphi_{klpq} \pmod{2\pi} . \quad (4.4)$$

Thus a type-3 invariant contains an angle-relation between two coefficients. Invariants of type 3 essentially are weighted sums of angles. The corresponding expressions in real curve coefficients may be calculated by using trigonometric functions in the real coefficients a_{kl} . The type-3 products' main feature is that any coefficient may be combined with any other such that the result builds an invariant. Thus considering expressions with three or more different coefficients is superfluous. These can be generated by combining invariants listed above.

4.1.4. Reconstruction of a curve by rotation invariants

Suppose we have given values to all type-1 and type-2 invariants and (at least) all values to all non-vanishing expressions of type 3. By type 1 all c_{kk} -coefficients are determined and by type 2 all $r_{kl} \in \mathbb{R}_{\geq 0}$ with $c_{kl} = r_{kl}e^{i\varphi_{kl}}$ are known. If $r_{kl} = 0$ for all r_{kl} with $k \neq l$, all curve-coefficients are already determined and the invariants belong a curve consisting of one or several concentric circles. Otherwise, we are left with determining the angles φ_{kl} for coefficients $c_{kl} \neq 0$ with $k \neq l$.

According to the sole rotational degree of freedom, one of these angles will be left undetermined. For any pair of non-vanishing coefficients c_{kl} and c_{pq} ($k \neq l$ and $p \neq q$), we get a non-vanishing type-3a/3b invariant and thus a relation between the angles φ_{kl} and φ_{pq} (see Equation (4.4)). The principal problem of this Equation (4.4) is that with given φ_{klpq} and with one given angle, the other angle is in general not totally determined modulo 2π : We have either

$$\varphi_{kl} \equiv \frac{-k+l}{-p+q} \cdot \varphi_{pq} - \frac{m \cdot \varphi_{klpq}}{-p+q} \pmod{\frac{2\pi \cdot m}{-p+q}}$$

or

$$\varphi_{pq} \equiv \frac{-p+q}{-k+l} \cdot \varphi_{kl} + \frac{m \cdot \varphi_{klpq}}{-k+l} \pmod{\frac{2\pi \cdot m}{-k+l}}$$

with $m = \gcd(|-p+q|, |-k+l|)$. In any case, one given angle determines the other only up to certain complex roots of unity. Figure 4.5 illustrates this: We oppose the two related angles displayed as nodes with rays in the according angles.

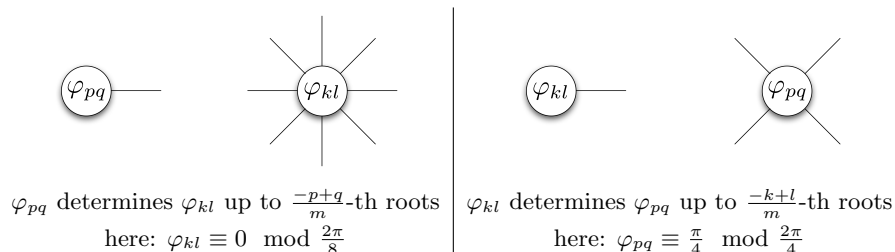


Figure 4.5.: Rotation invariants determine angles up to complex roots of unity

In addition to the ambiguity due to the rotational degree of freedom there is might be ambiguities due to the angle-relations. Such ambiguities are not always resolvable, even if we take all invariants into account. However, we will see that all coefficients satisfying the given relations lead to rotated copies of each other and reconstruction up to rotations is possible.

Now, let φ_{pq} be chosen arbitrarily. We show that all possible angles for φ_{kl} can be obtained by a rotation of the plane leaving φ_{pq} invariant. Rotating the plane around the origin by ψ , we get $\varphi_{kl} \mapsto \varphi_{kl} + (-k+l)\psi$ and $\varphi_{pq} \mapsto \varphi_{pq} + (-p+q)\psi$ (see Equation (4.1)). Thus rotations by $\psi = n \frac{2\pi}{-p+q}$ do not change φ_{pq} for any $n \in \mathbb{Z}$, but

$$\begin{aligned} \varphi_{kl} &\mapsto \varphi_{kl} + (-k+l)\psi = \varphi_{kl} + 2\pi n \frac{-k+l}{-p+q} \\ &= \varphi_{kl} + \frac{2\pi \cdot m}{-p+q} \cdot n \frac{-k+l}{m} \end{aligned}$$

With $m = \gcd(|-k+l|, |-p+q|)$ we have relatively prime $\frac{-p+q}{m}$ and $\frac{-k+l}{m}$. Thus

$$\left\{ \frac{2\pi \cdot m}{-p+q} \cdot n \frac{-k+l}{m} \pmod{2\pi} \mid n \in \mathbb{Z} \right\} = \left\{ \frac{2\pi \cdot m}{-p+q} \cdot r \pmod{2\pi} \mid r \in \mathbb{Z} \right\} .$$

Consequently, choosing the rotation properly we can switch between the choices (rays) for φ_{kl} . One brief note: Here it becomes clear why the type-3 invariants have their exponents divided by a corresponding gcd-value m . If we had stated our invariant expressions without m , then the choices for φ_{kl} would have included angles which could not be obtained by our special rotations of the plane.

Example 4.1 As an example take $c_{kl} = c_{02}$ and $c_{pq} = c_{40}$. To a chosen $\varphi_{40} = \widehat{\varphi}_{40}$ we obtain

$$\widehat{\varphi}_{02} \equiv -\frac{1}{2}\widehat{\varphi}_{40} + \frac{1}{4}\widehat{\varphi}_{0240} \pmod{\frac{\pi}{2}} \quad \text{versus} \quad \varphi_{02} \equiv -\frac{1}{2}\varphi_{40} + \frac{1}{2}\varphi_{0240} \pmod{\pi} .$$

In the first case, without the division by the gcd, we have twice as many angles as with the division. Now suppose $\widehat{\varphi}_{0240} = \varphi_{0240} = 0$ and $\widehat{\varphi}_{40} = \varphi_{40} = 0$.

Then $\widehat{\varphi}_{20} \in M = \{0, \pm\frac{\pi}{2}, \pi\}$ satisfies the above relation. Rotations leaving $\widehat{\varphi}_{40}$ invariant are rotations by $\psi \in \frac{\pi}{2}\mathbb{Z}$. But such rotations rotate $\widehat{\varphi}_{02}$ by multiples of π and M resolves into two rotationally inequivalent orbits.

Consequently, the invariants would not have determined the curve up to rotations! All observations from above have already shown

Theorem 4.2 *Let g be a complexified curve of known degree d . The values of all invariants of (reduced) type 1, 2, 3a and 3b with respect to g shall be known. Then it is possible to reconstruct g from the invariants up to a rotation of the plane around the origin: Expressions of type 1, 2 and 3 build a complete set of invariants, so to say.*

Let us dwell on this theorem to better understand the peculiarities of the reconstruction process. If we have a curve with exactly one non-vanishing coefficient of the form c_{kl} with $k \neq l$, then our graph-terminology shows only isolated nodes. Reconstruction is simple: All c_{kk} are given by type-1 invariants and $r_{kl} = |c_{kl}| = \sqrt{c_{kl}c_{\bar{k}\bar{l}}}$ is also clear. A single parameter, namely φ_{kl} , remains undetermined, resembling the one rotational degree of freedom. A rotation of the plane can arbitrarily transform φ_{kl} .

If we have exactly two different non-vanishing coefficients c_{kl} and c_{pq} with $k \neq l, p \neq q$ and $(k, l) \neq (q, p)$, we are in the situation which lead us to the theorem above. If there is a third coefficient c_{uv} or even more coefficients, we have more invariants. Arbitrarily fixing an angle, say φ_{kl} , determines the other angles up to angles of complex roots of unity. Within these angle-sets we may now determine subsets such that all the other invariant expressions are satisfied. We may do so by iteratively using Equation (4.4) or our graphical schemes representing the type 3 invariants. Finally we end up with a non-empty set of angles, which may be assigned to all the $\varphi_{kl}, \varphi_{pq}, \dots$. This set is non-empty because the original curve may be rotated such that $\tilde{\varphi}_{kl}$ is of the fixed value. Thus there is an angle set satisfying all invariant expressions. If we end up with multiple possibilities, we have shown that these are of one and the same (but rotated) curve. We exemplify reconstruction with three non-vanishing type-3a/3b expressions in the next section.

Theorem 4.3 *The set of type 1, 2, 3a and 3b expressions in their reduced form are minimal in the following sense: Any curve is reconstructible by given values to any of these expressions. If a single expression is left out, then there is a curve which is no longer reconstructible up to rotations.*

PROOF It is clear that no type-1 or type-2 expressions may be omitted. If a type-3a/3b expression is missing, then we would have problems with certain curves: There are curves, having exactly two non-vanishing coefficients (and their corresponding conjugate) of the form c_{kl}, c_{pq} with $k \neq l, p \neq q$. A suitable multiplication would lead to the missing invariant. Without the value for this invariant expression we only have the absolute value but no angle to these coefficients. Thus we get a two-dimensional pencil of curves satisfying all the remaining invariants. Consequently, the curve is not reconstructible up to rotations because two degrees of freedom can not be resolved. \square

The number of invariants which are necessary for reconstruction of a specific curve, depends a lot on the curve: All type-1 and all type-2 expressions are necessary. But if a coefficient c_{kl} with $k \neq l$ vanishes, then - speaking in our graph terminology - all

connections to and from this node are superfluous. Thus, if we have exactly r coefficients $c_{kl} \neq 0$ with $k \neq l$, then we need at most $\binom{r}{2}$ type-3a/3b expressions. This is only an upper bound, because in special cases some invariants may be implied by the remaining ones (see the example in the next section).

4.1.5. Reconstruction of specific given curve

Let us examine a curve g of degree $d = 8$ with the only non-vanishing coefficients $c_{22} = \frac{1}{2}$, $c_{20} = 5 \cdot e^{i\frac{\pi}{4}}$, $c_{80} = 4 \cdot e^{i\frac{3}{5}\pi}$ and $c_{71} = e^{-i\frac{\pi}{2}}$. The corresponding curve is displayed in Figure 4.6.a. The set of non-vanishing relevant rotational invariants to this curve is

type 1	type 2	type 3
$c_{22} = \frac{1}{2}$	$c_{20}c_{02} = 5^2$	$(c_{02})^4(c_{80}) = 5^4 \cdot 4 \cdot e^{i\frac{8}{5}\pi}$
	$c_{80}c_{08} = 4^2$	$(c_{17})^4(c_{80})^3 = 4^3 \cdot e^{i\frac{9}{5}\pi}$
	$c_{71}c_{17} = 1^2$	$(c_{02})^3(c_{71}) = 5^3 \cdot e^{i\frac{3}{4}\pi}$

Suppose we only have the above table of invariants and we try to reconstruct a curve \tilde{g} of degree eight from this data. The type-1 invariants directly give us $\tilde{c}_{22} = \frac{1}{2}$ and $\tilde{c}_{00} = \tilde{c}_{11} = \tilde{c}_{33} = \tilde{c}_{44} = 0$. The type-2 expressions provide us with the absolute values of all other coefficients and the only non-vanishing among them are $\tilde{r}_{20} = \tilde{r}_{02} = 5$, $\tilde{r}_{80} = \tilde{r}_{08} = 4$ and $\tilde{r}_{71} = \tilde{r}_{17} = 1$.

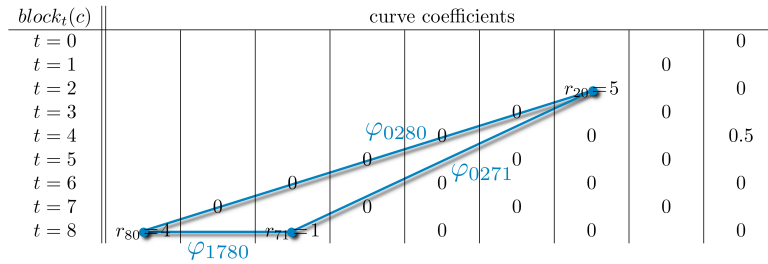
Using the notation $r_{klpq}e^{i\varphi_{klpq}} = (c_{kl})^{\frac{p-q}{m}}(c_{pq})^{\frac{-k+l}{m}}$ with $m = \gcd(|k-l|, |p-q|)$, we get the following invariants from the type-3 expressions in the tabular above:

$$\varphi_{0280} = \frac{8}{5}\pi, \quad \varphi_{1780} = \frac{9}{5}\pi \quad \text{and} \quad \varphi_{0271} = \frac{3}{4}\pi.$$

The absolute values r_{klpq} are not of interest since the type-2 expressions take care of consistency. According to Equation (4.4) and the fact that $\varphi_{kl} = -\varphi_{lk}$ we have the three equations

$$\begin{aligned} \varphi_{0280} = \frac{8}{5}\pi &\equiv -4\tilde{\varphi}_{20} + \tilde{\varphi}_{80} \pmod{2\pi}, & \varphi_{1780} = \frac{9}{5}\pi &\equiv -4\tilde{\varphi}_{71} + 3\tilde{\varphi}_{80} \pmod{2\pi} \\ \text{and } \varphi_{0271} = \frac{3}{4}\pi &\equiv -3\tilde{\varphi}_{20} + \tilde{\varphi}_{71} \pmod{2\pi}. \end{aligned} \quad (4.5)$$

These correspond to the (remaining) type-3 edges in our graph:



Now we are free to choose one of the angles. This freedom corresponds to the one dimensional degree of freedom when dealing with rotations around the origin. For example let $\tilde{\varphi}_{80} = 0$. Then $\tilde{\varphi}_{20}$ and $\tilde{\varphi}_{71}$ are restricted according to the invariants φ_{0280} and φ_{1780} . According to Equation (4.5):

$$\tilde{\varphi}_{20} \equiv \frac{\tilde{\varphi}_{80}}{4} - \frac{\varphi_{0280}}{4} \equiv \frac{\pi}{10} \pmod{\frac{\pi}{2}} \quad \text{and} \quad \tilde{\varphi}_{71} \equiv \frac{3\tilde{\varphi}_{80}}{4} - \frac{\varphi_{1780}}{4} \equiv \frac{\pi}{20} \pmod{\frac{\pi}{2}}.$$

The angle-combinations satisfying these two expressions are shown in Figure 4.6.b: The four choices of $\tilde{\varphi}_{20}$ and $\tilde{\varphi}_{71}$ are indicated by the two times four line segments, rotationally shifted according to $-\frac{\varphi_{0280}}{4}$ and $-\frac{\varphi_{1780}}{4}$.

Not all of these angle-combinations necessarily lead to a rotated copy of the original curve. We also have one constraint left: φ_{0271} relates $\tilde{\varphi}_{20}$ to $\tilde{\varphi}_{71}$. Let us choose $\tilde{\varphi}_{71}$, cycling through $\Phi = \{\frac{\pi}{20}, \frac{11\pi}{20}, \frac{21\pi}{20}, \frac{31\pi}{20}\}$ and check the equation with φ_{0271} in (4.5) to determine the possible angles for $\tilde{\varphi}_{20}$. Thereby Φ is the set of valid angles $\tilde{\varphi}_{71}$ according to our selection of $\tilde{\varphi}_{80}$ and φ_{1780} . In Figure 4.6.c-f the situations are illustrated. Whenever angles coincide with angles in Figure 4.6.b, the corresponding line segments are highlighted. The highlighted angle sets must satisfy all our given invariant expressions and thus may serve as angles for the reconstructed coefficients. In our case, we get four valid coefficient sets, satisfying all our invariant expressions. The corresponding curves are shown in Figure 4.6.g-j: The curves are all rotated copies of our original curve.

Instead of fixing $\tilde{\varphi}_{80}$, we likewise could have chosen to fix $\tilde{\varphi}_{71}$ or $\tilde{\varphi}_{20}$. Fixing $\tilde{\varphi}_{71}$ would have meant having a choice of three angles for $\tilde{\varphi}_{80}$ or $\tilde{\varphi}_{20}$ (see Figure 4.6.c) and in the end we would have got three rotated copies of our original curve. Fixing $\tilde{\varphi}_{20}$ as first angle would have been the fastest way for reconstruction: For example, setting $\tilde{\varphi}_{20} = \frac{\pi}{10}$ automatically determines the other angles modulo 2π . According to Equation (4.5) we have $\tilde{\varphi}_{80} \equiv \varphi_{0280} + 4\tilde{\varphi}_{20} \equiv 2\pi \equiv 0 \pmod{2\pi}$ and $\tilde{\varphi}_{71} \equiv \varphi_{0271} + 3\tilde{\varphi}_{20} \equiv \frac{21\pi}{20} \pmod{2\pi}$. The relation between $\tilde{\varphi}_{80}$ and $\tilde{\varphi}_{71}$ is automatically satisfied because of the equivalence $-4\varphi_{0271} + 3\varphi_{0280} \equiv -4\tilde{\varphi}_{71} + 3\tilde{\varphi}_{80} \equiv \varphi_{1780} \pmod{2\pi}$. As corresponding diagram we may take Figure 4.6.e but only with the highlighted segments.

In any case, we end up with one or more rotated copies of our original curve. As already mentioned, this would not be the case if the exponents in our invariant expressions $(c_{kl})^{\frac{p-q}{m}} (c_{pq})^{\frac{-k+l}{m}}$ had not been divided by the gcd-value $m = \gcd(|p-q|, |k-l|)$. Let us take our curve example again. Ignoring m we get

$$(c_{02})^8 (c_{80})^2 = 5^8 \cdot 4^2 \cdot e^{i\frac{6}{5}\pi}, \quad (c_{17})^8 (c_{80})^6 = 4^6 \cdot e^{i\frac{8}{5}\pi} \quad \text{and} \quad (c_{02})^6 (c_{71})^2 = 5^6 \cdot e^{i\frac{3}{2}\pi}.$$

These expressions are of course invariant but they do not permit a reconstruction up to rotations of the original curve:

$$\begin{aligned} \varphi'_{0280} &= \frac{6}{5}\pi \equiv -8\tilde{\varphi}'_{20} + 2\tilde{\varphi}'_{80} \pmod{2\pi}, & \varphi'_{1780} &= \frac{8}{5}\pi \equiv -8\tilde{\varphi}'_{71} + 6\tilde{\varphi}'_{80} \pmod{2\pi} \\ &\text{and} & \varphi'_{0271} &= \frac{3}{2}\pi \equiv -6\tilde{\varphi}'_{20} + 2\tilde{\varphi}'_{71} \pmod{2\pi}. \end{aligned}$$

Thus fixing $\tilde{\varphi}'_{80} = 0$ we would get the situation shown in Figure 4.6.k. A possible choice satisfying all above constraints, especially the one with φ'_{0271} , is highlighted. There $\tilde{\varphi}'_{71} = \frac{3\pi}{10}$ and $\tilde{\varphi}'_{20} = \frac{17\pi}{20}$. The corresponding curve is shown in Figure 4.6.l. Clearly it is not a rotated copy of our original curve.

By this example we already get a glimpse on rotational symmetry. The given curve has a twofold rotational symmetry. According to the above observations, the rotation speeds $\pm 8, \pm 6, \pm 2$ and 0 must be part of the cause: They all have a greatest common divisor of $\gcd(8, 6, 2, 0) = 2$. Consequently, a rotation of the plane by $\psi = \pi$ will not effect any of these coefficients. This finds its correspondence in our invariant expressions: In any type-3 invariant the gcd had been $m = 2$.

In fact, if we have a curve with a gcd of all rotation speeds being a number s , then the curve must have at least a s -fold rotational symmetry. In Section 5.1 we will see that there are more curves with a s -fold rotational symmetry. In fact, the criteria for

4. Curves and selected transformations of the plane

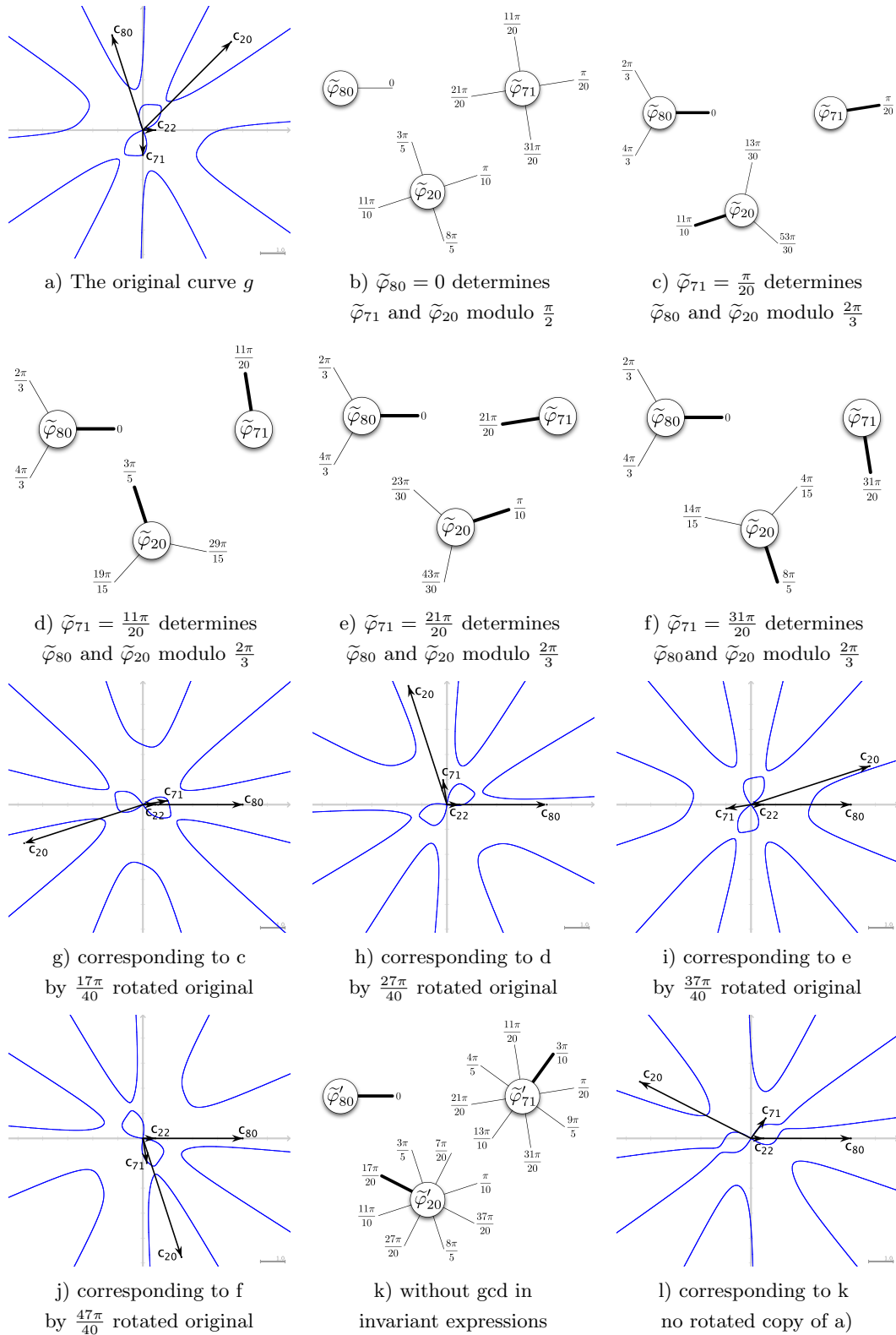


Figure 4.6.: Reconstructing a curve from invariants

such a symmetry is that s is the greatest common divisor of all possible differences of rotation speeds of non-vanishing coefficients (see Section 5.1). But the gcd of all our rotation speeds divides any sum and difference of the rotation speeds and thus also the corresponding gcd.

4.1.6. Normal form with respect to rotations around the origin

If we have a curve whose only non-vanishing coefficients are of the form c_{kk} , then the curve is not affected by rotations around the origin. Thus the whole curve is rotationally invariant. It must therefore consist of one or several circles concentric to the origin.

These curves are of the form $\sum_{k=0}^{\frac{d}{2}} m_{kk} c_{kk} z^k \bar{z}^k h^{d-2k} = 0$, which is and has to be a normal form with respect to rotations around the origin.

If there is a $c_{kl} \neq 0$ with $k \neq l$, the curve is affected by rotations. Let S be the set of all such coefficients for a given curve g . If a c_{kl} is contained in S , we can find angles by which the curve has to be rotated such that the transformed curve \tilde{g} has a positive real $\tilde{c}_{kl} \in \mathbb{R}_{>0}$. In general, there is a whole set of possible angles ψ with this property: The coefficient $c_{kl} = r_{kl} e^{i\varphi_{kl}}$ transforms under rotations according to $\tilde{c}_{kl} = c_{kl} e^{i\psi(-k+l)}$. For \tilde{c}_{kl} to be real, $\varphi_{kl} + \psi(-k+l) \equiv 0 \pmod{2\pi}$ must be satisfied. Thus $e^{i\psi}$ is determined up to the $(k-l)^{\text{th}}$ complex roots of unity and

$$\psi \equiv \frac{\varphi_{kl}}{k-l} \pmod{\frac{2\pi}{k-l}}.$$

The transformation group of rotations around the origin separates the set of curves into equivalence classes. The class of all rotationally equivalent curves to a given curve g will be denoted by $[g]_{rot}$. A representation system and thus a normal form with respect to rotations can be obtained by selecting a specific well-defined curve in each equivalence class. We do so by choosing the curve whose coefficients $c_{kl} = r_{kl} e^{i\varphi_{kl}}$ exhibit angles φ_{kl} satisfying a certain minimality-criterion. In what follows we describe this criterion and how the corresponding curve may be obtained.

Let g be an arbitrary curve and $S = S_g = \{c_{kl} \mid c_{kl} \neq 0 \wedge k \neq l\}$ the set of all non-vanishing non- c_{kk} -coefficients. Rotating g , the coefficients may change but S will always contain the same coefficients. Thus we have a rotationally invariant S . According to Definition 2.3, S is totally ordered by \succ . Let $|S| = n$ and $s = (s_1, s_2, \dots, s_n)$ be a vector, listing all elements of S in a by \succ ordered manner. Furthermore, let $s_k = r_k e^{i\varphi_k}$ with $r_k > 0$ for any k . The so called *angle-vector* to a vector s is denoted by φ with $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_n) \in [0, 2\pi)^n$. This vector contains exactly those values of a curve which are affected by rotations around the origin.

A total order on these (angle-) vectors in $[0, 2\pi)^n$ allows us to single out a unique vector and thus a rotational normal form. Therefore let $\varphi < \varphi'$, whenever there is an index $k < n$ with $\varphi_k < \varphi'_k$ and $\varphi_j = \varphi'_j$ for all $j < k$. Thus we have already proven

Theorem 4.4 *Among each class of rotationally equivalent curves there is a unique curve with a minimal angle-vector. The set of all these curves constitutes a rotational normal form.*

This theorem becomes even more obvious when we calculate the angle(s) ψ by which a given curve g has to be rotated such that it is in rotational normal form. At first all $\psi \in \Psi_0 = [0, 2\pi)$ are admissible. Now we successively minimize the components of the

angle-vector φ of g . In the r -th step we rotate g by any angle in Ψ_{r-1} and determine the resulting angle-vectors. Those rotation-angles leading to the smallest angle-vectors constitute Ψ_r . After $|S| = n$ steps we have the desired angle(s) contained in Φ_n . If in the meantime $|\Psi_r| = 1$, we may already stop, because then $\Phi_n = \Phi_r$.

In any case all angles $\psi \in \Psi_n$ lead to one and the same curve, independent from any rotations the curve may have previously undergone: For $|\Psi_n| = 1$ this is trivial and for $|\Psi_n| \neq 1$ we have minimized all coefficients in S . Thus we have several rotation angles but *exactly one* curve and this is crucial in the context of normal forms. If $|\Psi_n| \neq 1$, the given curve must have been rotationally symmetric. More on rotational symmetry will be covered in Section 5.1.

Example 4.2 As an example for transforming a curve into rotational normal form look at Figure 4.7. The coefficients of the original curve g are $c_{40} = 5 \cdot e^{i\frac{5}{6}\pi}$, $c_{31} = 4 \cdot e^{i\frac{1}{6}\pi}$, $c_{20} = 7 \cdot e^{i\frac{5}{4}\pi}$, $c_{22} = 3$ and $c_{00} = -3$. The remaining coefficients are either complex conjugates to the given values or zero. The (sorted) set S is given by $S = \{c_{40}, c_{31}, c_{20}\}$ and $|S| = 3$. We already left the complex conjugate out, because minimizing a φ_{kl} renders a minimization of φ_{lk} irrelevant. Ψ_1 must lead to a minimized angle of the largest coefficient in S with respect to \succ . This means that $\tilde{c}_{40} = 5$. Thus the elements $\psi \in \Psi_1$ satisfy $0 \equiv \varphi_{40} - 4\psi \pmod{2\pi}$ and consequently $\Psi_1 = \frac{5\pi}{24} + \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. Figure 4.7 shows the corresponding rotated copies of g . In a next step we can minimize $\tilde{\varphi}_{31}$ by using rotations and $\psi \in \Phi_1$. Rotations by $\psi \in \frac{5\pi}{24} + \{0, \pi\}$ lead to $\tilde{\varphi}_{31} \equiv \frac{7\pi}{4} \pmod{2\pi}$ and $\psi \in \frac{5\pi}{24} + \{\frac{\pi}{2}, \frac{3\pi}{2}\}$ lead to $\tilde{\varphi}_{31} \equiv \frac{3\pi}{4} \pmod{2\pi}$. Thus $\Phi_2 = \frac{5\pi}{24} + \{\frac{\pi}{2}, \frac{3\pi}{2}\}$. Any rotation by $\psi \in \Phi_2$ leads to the same $\tilde{\varphi}_{20} \equiv \frac{11\pi}{6} \pmod{2\pi}$ and consequently $\Phi_2 = \Phi_3$. Now, we end up with two angles which transform our given curve in one and the same curve shown in Figure 4.7.c. The curve is unique but the rotation-angle leading to this curve is not.

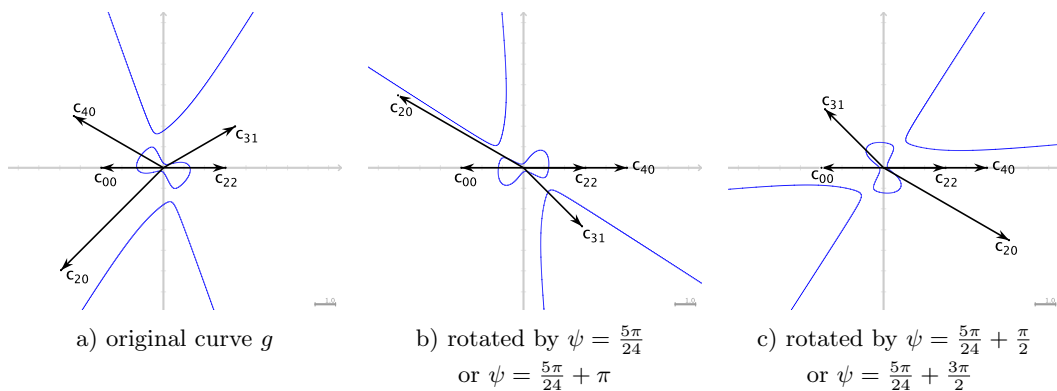


Figure 4.7.: Rotational normal form

4.1.7. Invariants under rotations around the origin in tensor notation

In Section 3.2.3 we saw that Euclidean invariants of curves are built up by the use of ε -tensors together with the E - and A -tensor. According to Theorem 3.4, we can also use I - and J -tensors in equal numbers, because the diagrams may be replaced by a linear combination of corresponding diagrams. Dealing with rotations around the origin, a new geometric object remains invariant: the origin itself as a point. Consequently, we may construct invariants of curves under rotations around the origin by building closed

diagrams consisting of curve-, ε -, E -, A - (or I - and J -) and 0 -tensors. Thereby \circlearrowleft denotes the contravariant 0 -tensor $(0, 0, 1)^T$. We will see that this additional tensor is enough to formally describe all our invariant expressions of type 1, 2 and 3 and thus all invariants necessary for reconstruction.

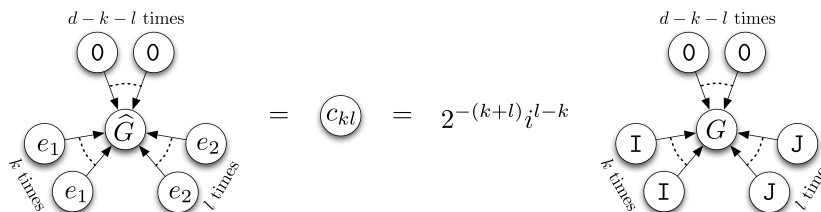
For example, if there is a curve-tensor for a curve \mathcal{C}_g of degree d , we have a node with d arrows pointing to the node. Connecting d times the 0 -tensor we get $a_{00} = c_{00}$ as a result. As we have seen, c_{00} is invariant under rotations around the origin because it is of type 1. Clearly the complexified version of 0 is $\widehat{0} = 0$ itself:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & i & 0 \\ 1 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 .$$

Connecting 0 to a curve-tensor \widehat{G} of variance d is like filtering a sub-tensor:

$$\widehat{G}_{k,i_2,i_3,\dots,i_d} 0^k = \widehat{G}_{3,i_2,i_3,\dots,i_d} .$$

The components of this sub-tensor could as well be the coefficients of a curve with one degree less, namely $d - 1$. Deleting the leading coefficients gives us the curve corresponding to the sub-tensor. In this sense we can deal with any 0 -tensor as a tensor virtually reducing the degree of the curve. We can address any curve-coefficient by using a curve-tensor connected to suitably many 0 -, I - and J -tensors. Theorem 3.6 provides:



In what follows we will confine ourselves again to the complexified notation (see Section 3.3.1). The curve-tensor in real representation G will be replaced by the complexified version \widehat{G} and the role of E , I and J is played by \widehat{E} , e_1 and e_2 , respectively.

Looking back at the construction of our invariants under rotations of the plane, we can see the following: All expressions satisfy that the rotation-speed of the expression is zero, i.e. the sum of all rotation speeds $-k + l$ of the participating coefficients c_{kl} vanishes. But this means that the sum of all first indices of the participating coefficients is exactly the sum of the second indices. However, stating our invariants as suitable products of the tensor-diagrams shown above, we see: The sum of the first indices indicates how many e_1 -tensors are used in total and the sum of the second indices indicates how many e_2 -tensors are used in total. Thus stating our invariants in diagram notation we have equally many e_1 - and e_2 -tensors. According to Theorem 3.4, we can now replace all e_1 - and e_2 -tensors and instead get linear combinations of diagrams with \widehat{E} - and \widehat{A} -tensors. Consequently, we have rotation-invariants because we only used curve-tensors together with \widehat{E} -, \widehat{A} - and 0 -tensors.

Type-1 invariants for a curve of degree d , the c_{kk} ($0 \leq k \leq \lfloor \frac{d}{2} \rfloor$), are particularly simple:

According to Section 3.3.2 and the behavior of the 0-tensor we have

$$\frac{1}{2^k} \widehat{E}^k \widehat{G} 0^{d-2k} = c_{kk}$$

According to Theorem 3.15 we get for the type-2 invariants $c_{n-r,r}c_{r,n-r}$ ($0 \leq r \leq n \leq d$):

$$c_{n-r,r}c_{r,n-r} = \frac{1}{2^n} \sum_{q=0}^n C_q^{(n-r,r)} 0^{d-n} \widehat{G} \widehat{E}^{n-q} \widehat{A}^q \widehat{G} 0^{d-n}$$

As we have seen, the summands for odd q vanish and an even number of \widehat{A} tensors may be substituted by an expression without \widehat{A} (compare with Theorems 3.7 and 3.17).

The type-3 invariants are more complex. Naturally, they may be written easily with e_1 - and e_2 -tensors. But in general it is difficult to write down the corresponding expression with \widehat{E} - and \widehat{A} -tensors. What we can do is replace pairs of one e_1 - and one e_2 -tensor connected to the same curve-tensor:

$$\begin{array}{c} e_1 \quad e_2 \\ \downarrow \quad \downarrow \\ \widehat{G} \\ \uparrow \quad \uparrow \\ \text{---} \quad \text{---} \end{array} = \frac{1}{2} \begin{array}{c} \widehat{E} \\ \downarrow \\ \widehat{G} \\ \uparrow \\ \text{---} \end{array}$$

Thus we can confine ourselves to diagrams with \widehat{G} -tensors which are only connected to (several) 0- and E -tensors as well as to *either* several e_1 - or e_2 -tensors. The remaining e_1 - and e_2 -tensors may then be combined according to Theorem 3.4 in Section 3.2.3. Concerning our type-3 invariants, we confine ourselves to the existence of a linear-combination of diagrams with \widehat{E} - and \widehat{A} - and without e_1 - and e_2 -tensors. The linear-combinations may be obtained similarly to the ones in Section 3.3.5 by applying a corresponding formal decomposition as described in Section 3.3.3. Anyway, the type-3 invariants may be written as

$$(c_{kl})^{\frac{p-q}{m}} (c_{pq})^{\frac{-k+l}{m}} = \left(\begin{array}{c} \text{\scriptsize } d-k-l \text{ times} \\ 0 \quad 0 \\ \downarrow \quad \downarrow \\ \widehat{G} \\ \uparrow \quad \uparrow \\ \text{\scriptsize } k \text{ times } e_1 \quad e_2 \text{\scriptsize } l \text{ times} \end{array} \right)^{\frac{p-q}{m}} \left(\begin{array}{c} \text{\scriptsize } d-p-q \text{ times} \\ 0 \quad 0 \\ \downarrow \quad \downarrow \\ \widehat{G} \\ \uparrow \quad \uparrow \\ \text{\scriptsize } p \text{ times } e_1 \quad e_2 \text{\scriptsize } q \text{ times} \end{array} \right)^{\frac{-k+l}{m}}$$

4.2. Translation

In Section 2.3 we complexified a curve by writing it as a zero set of a polynomial $g(z, \bar{z}, h) = 0$ with $z \in \mathbb{C}$. In addition to its great advantages when dealing with rotations, this notation will also be beneficial when curves are translated. As in the last chapter, we start by analyzing the effects of translations on curve coefficients. Getting to know special transformations making coefficients vanish, we arrive at a translatorial normal form. Curves having intersections with the line at infinity of multiplicity greater than one play a special role. A treatment of translation invariants of curves is also included in this chapter.

4.2.1. Effects of translations on complexified curve coefficients

When talking about translations, it is convenient to use dehomogenized coordinates or to scale the homogenizing component to $h = 1$. Thus a point p , given by $p = (z, \bar{z}, h)$ with $h = 1$, is translated by t when

$$\begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & \bar{t} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix}. \quad (4.6)$$

If $h \neq 1$, then t would be a scaled Euclidean translation parameter.

This transformation of the plane has its implications on the curve-coefficients. The transformation of the curve coefficients is covered in

Theorem 4.5 *Let g be a curve with coefficients c_{kl} and let \tilde{g} be the by t translated copy $\tilde{g}_t = \tilde{g}$ with coefficients \tilde{c}_{kl} . The translation induces the following map on the coefficients:*

$$c_{kl} \mapsto \tilde{c}_{kl} = \sum_{p=k}^{d-l} \sum_{q=l}^{d-p} (-1)^{(p+q)-(k+l)} \frac{(d-k-l)!}{(p-k)! (q-l)! (d-p-q)!} t^{p-k} \bar{t}^{q-l} c_{pq}. \quad (4.7)$$

PROOF The strategy of this proof is the same as in the proof of Theorem 2.7, which is carried out in detail in Appendix A.1. Translating g by t we get

$$0 = \tilde{g}(z, \bar{z}) = g(z-t, \bar{z}-\bar{t}) = \sum_{k+l=d} m_{kl} c_{kl} (z-t)^k (\bar{z}-\bar{t})^l.$$

Expanding the $(z-t)^k (\bar{z}-\bar{t})^l$, changing summation indices a few times and recollecting the coefficients of monomial terms proves the theorem. \square

Equation (4.7) completely describes the transformation of the coefficients. The fraction with the faculties may be restated by the use of binomial coefficients:

$$\frac{(d-k-l)!}{(p-k)! (q-l)! (d-p-q)!} = \binom{d-(k+l)}{d-(p+q)} \binom{(p+q)-(k+l)}{p-k}.$$

However, the matrix representation of this transformation is by far more instructive. To visualize the transformation using matrix notation, we choose $d = 3$.

$$\begin{aligned}
 & \left(c_{00} \mid c_{10} \ c_{01} \mid c_{20} \ c_{11} \ c_{02} \mid c_{30} \ c_{21} \ c_{12} \ c_{03} \right)^T \mapsto T_t c = \\
 & \underbrace{\begin{pmatrix} 1 & -3t & -3\bar{t} & 3t^2 & 6t\bar{t} & 3\bar{t}^2 & -t^3 & -3t^2\bar{t} & -3t\bar{t}^2 & -\bar{t}^3 \\ & 1 & & -2t & -2\bar{t} & & t^2 & 2t\bar{t} & \bar{t}^2 & \\ & & 1 & & -2t & -2\bar{t} & & t^2 & 2t\bar{t} & \bar{t}^2 \\ & & & 1 & & & -t & -\bar{t} & & \\ & & & & 1 & & & -t & -\bar{t} & \\ & & & & & 1 & & & -t & -\bar{t} \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & 1 \end{pmatrix}}_{= T_t} \begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \\ c_{30} \\ c_{21} \\ c_{12} \\ c_{03} \end{pmatrix} \quad (4.8)
 \end{aligned}$$

4.2.2. Structure of the coefficient transformation

The block-structure of the curve coefficients as defined in Definition 2.3 carries over to the transformation matrix: Auxiliary lines are shown in (4.8). In order to be able to uniquely refer to blocks, let

$$T_t = \begin{pmatrix} \text{block}_{0,0} & \text{block}_{0,1} & \text{block}_{0,2} & \text{block}_{0,3} \\ & \text{block}_{1,0} & \text{block}_{1,1} & \text{block}_{1,2} \\ & & \text{block}_{2,0} & \text{block}_{2,1} \\ & & & \text{block}_{3,0} \end{pmatrix}$$

with the block-borders corresponding to the auxiliary lines above. In general, the first index n of $\text{block}_{n,m}(T_t)$ specifies the ‘‘block-row’’: The transformation of coefficients \tilde{c}_{kl} with index-sum $k+l = n$ is described by these blocks. The second index b indicates that the block is contained in the b -th sup-diagonal. Thus $\text{block}_{n,m}(T_t)$ gets multiplied with $\text{block}_{n+m}(c)$.

Blocks below the diagonal are not needed due to the visually most dominant property of T_t : the upper triangular shape. The lowest possible values for the summation indices in Equation (4.7) are k and l , respectively. Thus \tilde{c}_{kl} depends only on coefficients c_{pq} with $c_{pq} \succeq c_{kl}$.

To analyze the structure within the blocks, it is best to reformulate Equation (4.7): It is convenient to rewrite \tilde{c}_{kl} by $\tilde{c}_{k,n-k}$ with n being the sum of the indices and the c_{pq} analogously. Changing summation accordingly we get

$$\tilde{c}_{k,n-k} = \sum_{m=0}^{d-n} \sum_{p=0}^m (-1)^m \binom{d-n}{m} \binom{m}{p} t^p \bar{t}^{m-p} c_{(p+k),(m+n)-(p+k)}. \quad (4.9)$$

Now the block-indexing is synchronized with the transformation formula: Analyzing a $\tilde{c}_{k,n-k}$ gives us the the number n of the *first* index to relevant blocks: $\text{block}_{n,*}(T_t)$. Then the first summation, the summation over m , gives us the second index: coefficients of $c_{(p+k),(m+n)-(p+k)}$ with the same m are contained in the same block, namely $\text{block}_{n,\mathbf{m}}(T_t)$.

Let us have a closer look at the coefficients of our curve-coefficients in Equation (4.9). Focusing on $block_{n,m}(T_t)$, we have

$$\sum_{p=0}^m (-1)^m \binom{d-n}{m} \binom{m}{p} t^p \bar{t}^{m-p} = (-1)^m \binom{d-n}{m} (t + \bar{t})^m .$$

Thus for a given n the whole coefficient structure for any $\tilde{c}_{k,n-k}$ is given by the summands of the corresponding expanded expression of

$$\sum_{m=0}^{d-n} (-1)^m \binom{d-n}{m} (t + \bar{t})^m = (1 - (t + \bar{t}))^{d-n} .$$

This proves the very nice and simple general structure we saw in the matrix-notation of Equation (4.8) for the case of $d = 3$: In each row within a block we have the same coefficients, only shifted. All entries in $block_{n,m}(T_t)$ have a sign of $(-1)^m$, i.e. odd sup-diagonals have minus signs. They depend on t and \bar{t} of order m . The coefficient structure of t and \bar{t} in the rows of the blocks are binomially scaled binomial coefficients. The overall scaling of $block_{n,m}(T_t)$ is $(-1)^m \binom{d-n}{m}$, corresponding to coefficients in a generalized Pascal-Triangle (compare with Figure 3.8.f on page 52). In our case for $d = 3$ we get the scaling pattern

$$\left(\begin{array}{c|c|c|c} 1 & -3 & 3 & -1 \\ \hline & 1 & -2 & 1 \\ \hline & & 1 & -1 \\ \hline & & & 1 \end{array} \right) .$$

It is not too surprising that tetrahedral coefficient structures may also be found in (4.8), or more generally in (4.7). Back in Section 2.2.4 we examined a tetrahedral coefficient structure of a coefficient transformation induced by rotations of the plane. The secret behind the tetrahedral structure was that we had, as here, binomially weighted binomial coefficients. Let us look at the columns of the transformation-matrix in (4.8), as an example, the rightmost column. We may arrange the coefficients on a tetrahedron in such a way that the coefficients of a block are all contained in a plane intersecting our tetrahedron. With parallel intersecting planes, we have the same structure as in Section 2.2.4. The signs may be integrated in the weights and the coefficients themselves are always binomial (see the subsequent Figure 4.8).

Another interesting structural property of the transformation (4.7) emerges when we translate it into terms of our coefficient triangle introduced by Equation (2.17): Highlighting all coefficients which are involved in the transformation of a specific coefficient c_{kl} , we can see that they form a sub-triangle Δ_{kl} in the coefficient triangle: The tip, i.e. row $m = 0$, of Δ_{kl} is c_{kl} itself (see Figure 4.9). The entries in row $m = 1$ of Δ_{kl} contain the coefficients depending linearly on t and \bar{t} in the transformation of c_{kl} :

$$\tilde{c}_{kl} = \dots + const_1 \cdot (tc_{k+1,l} + \bar{t}c_{k,l+1}) + \dots .$$

The coefficients in the row $m = 2$ evoke a quadratic change

$$\tilde{c}_{kl} = \dots + const_2 \cdot (t^2c_{k+2,l} + 2t\bar{t}c_{k+1,l+1} + \bar{t}^2c_{k,l+2}) + \dots$$

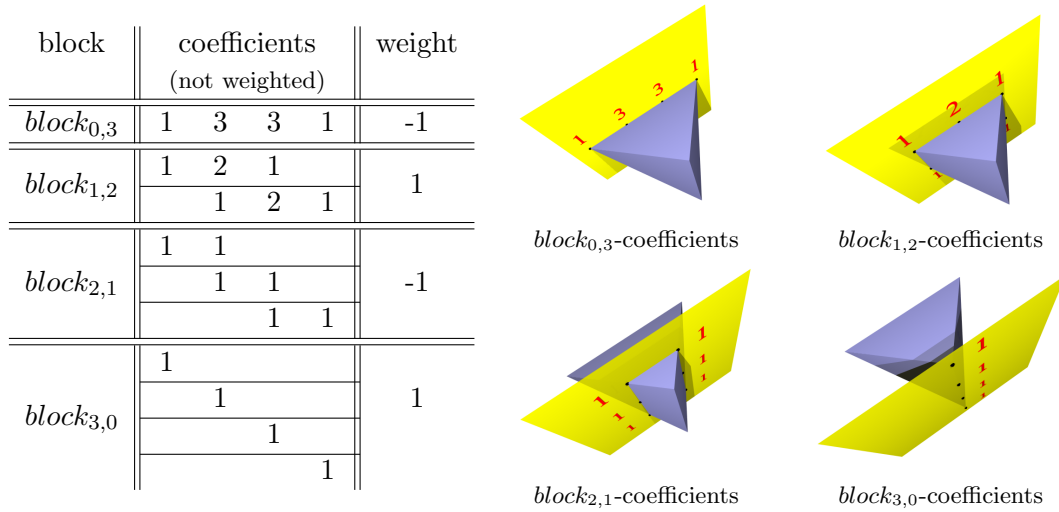


Figure 4.8.: Translation and tetrahedral coefficient structure

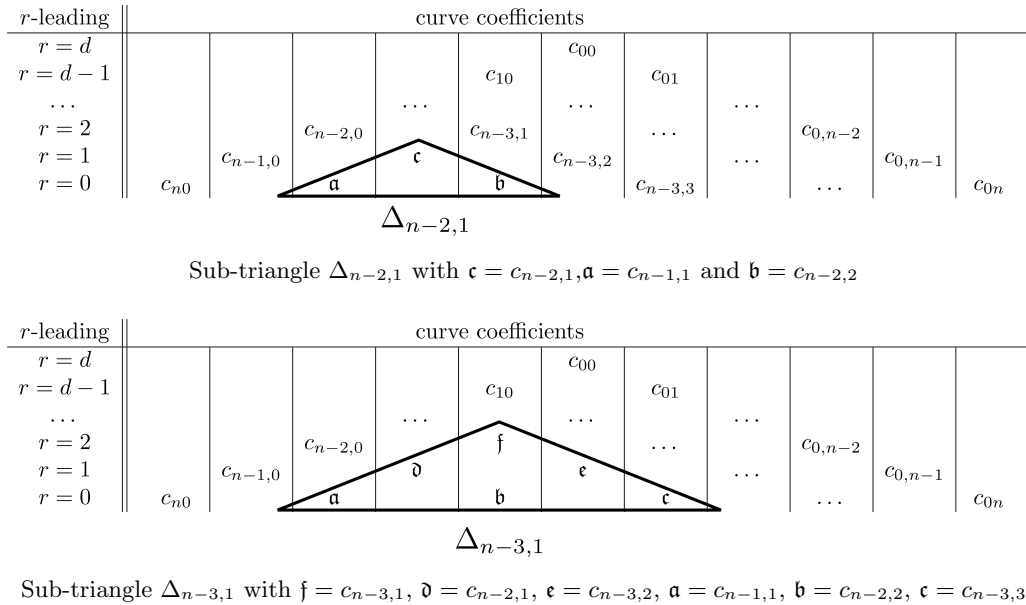


Figure 4.9.: Sub-triangles in the coefficient triangle

and so on. Thus the line in the sub-triangle determines the order of dependence on t and \bar{t} . The premultiplying constants may be determined according to their affiliation to blocks. These constants are exactly the overall block-scaling factors discussed above. Therefore $const_m = (-1)^m \binom{d-(k+l)}{m}$. Thus the integer factors, equipping the coefficients of a sub-triangle, form nice pattern. For $d=3$ and $c_{00} = j$ (or likewise for $d = 4$ and

abbreviate the coefficients in a specified sub-triangle by gothic letters. Unless explicitly stated otherwise, \mathfrak{a} will denote the greatest coefficient in Δ_{kl} with respect to \succ (see Definition 2.3), \mathfrak{b} the second largest and so on. Thus for a sub-leading c_{kl} the sub-triangle Δ_{kl} will be denoted by $\mathfrak{c} = c_{kl}$, $\mathfrak{a} = c_{k+1,l}$ and $\mathfrak{b} = c_{k,l+1}$ (see Figure 4.9 for a further example). The only sub-leading coefficient in this triangle is \mathfrak{c} . In many cases we make use of polar coordinates. There we will write $\mathfrak{a} = r_{\mathfrak{a}}e^{i\alpha}$, $\mathfrak{b} = r_{\mathfrak{b}}e^{i\beta}$, $\mathfrak{c} = r_{\mathfrak{c}}e^{i\gamma}$ and so on. The radii are supposed to be non-negative.

To be able to address the coefficients by the “distance” to the leading coefficients, we name coefficients c_{pq} **k-leading** if $p + q + k = d$, where d is the degree of the curve. Leading coefficients are 0-leading, sub-leading coefficients are 1-leading and so on.

Especially in the context of changing a sub-leading coefficient by translations of the plane, we come across certain angles very often. Therefore we will abbreviate these: Let Δ_{kl} be a sub-triangle in the coefficient triangle with tip \mathfrak{c} and base \mathfrak{a} and \mathfrak{b} (for an example see the top image in Figure 4.9). Then by convention:

$$\vartheta = \frac{-\alpha + \beta}{2} \quad , \quad \tau = \frac{\alpha + \beta}{2} \quad \text{and} \quad \sigma = \vartheta + \frac{\pi}{2} \quad .$$

The angle $2\vartheta = -\alpha + \beta$ will be called (oriented) **angle-difference** of the succeeding coefficient pair \mathfrak{a} and \mathfrak{b} .

These angles play an outstanding role and are needed to describe special lines, which in turn specify restrictions on transformed coefficients (see next section). For reference we name these auxiliary lines

$$L = \{ \lambda e^{i\tau} \mid \lambda \in \mathbb{R} \} \quad , \quad L_{\mathfrak{c}} = \{ \lambda e^{i\tau} + \mathfrak{c} \mid \lambda \in \mathbb{R} \} = \mathfrak{c} + L$$

$$\text{and} \quad L^{\perp} = \left\{ \lambda e^{i(\tau + \frac{\pi}{2})} \mid \lambda \in \mathbb{R} \right\} \quad .$$

Interpreting \mathbb{C} as a (real) two-dimensional space and L as a one-dimensional subset, we can see that $\mathbb{C}/L = \{ L_{\mathfrak{c}} \mid \mathfrak{c} \in \mathbb{C} \}$. A representation system can be obtained by intersecting all possible lines $L_{\mathfrak{c}}$ with a traversing line. Such a set of intersections may of course be identified with the traversing line itself. A special representation system is L^{\perp} (the line passing through the origin) being perpendicular to any $L_{\mathfrak{c}}$. The point of intersection p with $\{p\} = L^{\perp} \cap L_{\mathfrak{c}}$ is singled out by its absolute value: Among all points of $L_{\mathfrak{c}}$, p has minimal absolute value.

In the sections ahead we will have to examine translations by parameters t lying on lines which are closely related to the auxiliary lines above. To be able to reference these lines too, we introduce special meaningful points: The orthogonal projection of \mathfrak{c} onto L is given by

$$\mathfrak{c}_L = r_{\mathfrak{c}} \cos(\gamma - \tau) e^{i\tau} \quad .$$

In case when $r_{\mathfrak{a}} \neq 0$ we define

$$t_{\mathfrak{c}} = \frac{|\mathfrak{c}_L|}{2r_{\mathfrak{a}}} e^{i\vartheta} = \frac{r_{\mathfrak{c}} \cos(\gamma - \tau)}{2r_{\mathfrak{a}}} e^{i\vartheta} \quad .$$

If $r_{\mathfrak{a}} = r_{\mathfrak{b}}$, $t_{\mathfrak{c}}$ may be rewritten in a simpler form without trigonometric functions.

Expanding c_L , $\cos(\gamma - \tau)$ and the angle-abbreviations we get

$$\begin{aligned} t_{\mathbf{c}} &= \frac{r_{\mathbf{c}}}{4r_{\mathbf{a}}} \left(e^{i(\gamma - \frac{\alpha + \beta}{2})} + e^{i(-\gamma + \frac{\alpha + \beta}{2})} \right) e^{i(\frac{-\alpha + \beta}{2})} \\ &= \frac{r_{\mathbf{c}}}{4r_{\mathbf{a}}} \left(e^{i(\gamma - \alpha)} + e^{i(-\gamma + \beta)} \right) \\ &= \frac{\mathbf{c}\bar{\mathbf{a}} + \bar{\mathbf{c}}\mathbf{b}}{4r_{\mathbf{a}}^2} . \end{aligned}$$

Together with $t_{\mathbf{c}}$ we get further auxiliary lines with

$$\begin{aligned} T &= \{ \lambda e^{i\sigma} \mid \lambda \in \mathbb{R} \} \quad , \quad T_{\mathbf{c}} = \{ \lambda e^{i\sigma} + t_{\mathbf{c}} \mid \lambda \in \mathbb{R} \} = t_{\mathbf{c}} + T \\ \text{and} \quad T^{\perp} &= \{ \lambda e^{i\vartheta} \mid \lambda \in \mathbb{R} \} . \end{aligned}$$

Here we have analogous conditions for a representation-system compared to the lines $L_{\mathbf{c}}$. We have the set ${}^{\mathbb{C}}/T = \{ T_{\mathbf{c}} \mid \mathbf{c} \in \mathbb{C} \}$. As a special representational system we have T^{\perp} with $T_{\mathbf{c}} \perp T^{\perp}$ for all \mathbf{c} and $0 \in T^{\perp}$. Additionally, $t_{\mathbf{c}}$ with $T_{\mathbf{c}} \cap T^{\perp} = \{ t_{\mathbf{c}} \}$ has minimal absolute value amongst all points of $T_{\mathbf{c}}$.

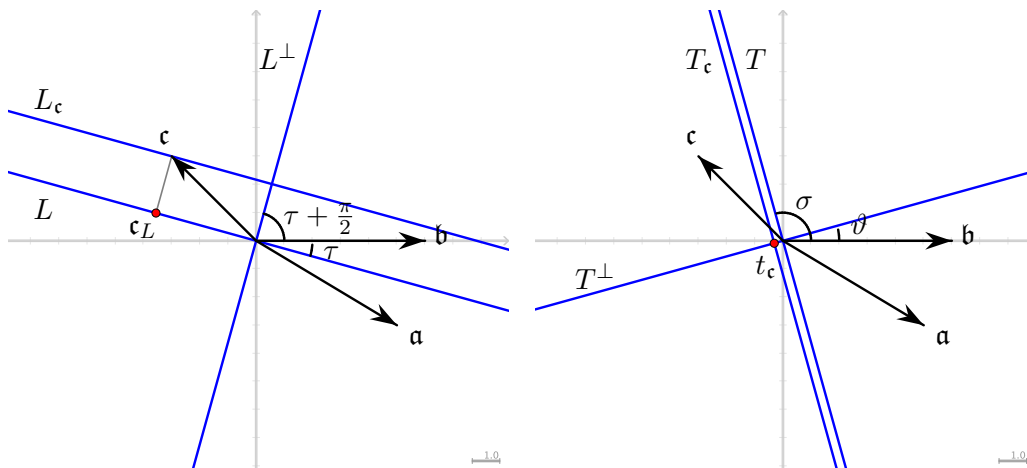


Figure 4.10.: Auxiliary lines for translations

4.2.4. Minimizing coefficients by a translation

With the abbreviations of the last section we can exploit the linear dependence of sub-leading coefficients on t and \bar{t} . We use it to find a translation such that the absolute value of a transformed sub-leading coefficient is minimized or annihilated. The process of translating a curve such that some $\tilde{c}_{kl} = 0$ with $k + l = d - 1$ will be called annihilation. If only a minimization of the absolute value is possible, we speak in short of minimization. We will study minimization and annihilation in detail because they are the key ingredients for our normal form for curves with respect to translations. Therefore an examination of the behavior of coefficients under arbitrary translations and under translations with a fixed direction takes place. By such translations of fixed direction we understand translations with translation parameter t restricted to a predefined line

passing through the origin. At first we will look at a Δ_{kl} to a sub-leading $c_{kl} = \mathbf{c}$. By Equation (4.7) we have

$$\tilde{\mathbf{c}} = \mathbf{c} - t\mathbf{a} - \bar{t}\mathbf{b} . \quad (4.11)$$

For almost all curves $r_{\mathbf{a}} \neq r_{\mathbf{b}}$ holds. In this case annihilation is possible: A translation parameter t may be found such that $\tilde{\mathbf{c}} = 0$. We state this in

Theorem 4.6 *Let $r_{\mathbf{a}} \neq r_{\mathbf{b}}$ and $\Delta = r_{\mathbf{a}}^2 - r_{\mathbf{b}}^2$. The by*

$$t = \frac{1}{\Delta} (\mathbf{c} \bar{\mathbf{a}} - \bar{\mathbf{c}} \mathbf{b}) \quad (4.12)$$

translated copy $\tilde{F} = \tilde{F}_t$ of F has a vanishing sub-leading coefficient $\tilde{\mathbf{c}} = 0$. Translations by a parameter different from t in Equation (4.12) imply $\tilde{\mathbf{c}} \neq 0$.

PROOF We determine t from the condition $\tilde{\mathbf{c}} = 0$ and the transformation law for sub-leading coefficients (4.11) $\tilde{\mathbf{c}} = \mathbf{c} - t\mathbf{a} - \bar{t}\mathbf{b}$. Splitting the participating coefficients into real and imaginary parts yields:

$$\mathbf{a} = x_{\mathbf{a}} + iy_{\mathbf{a}}, \quad \mathbf{b} = x_{\mathbf{b}} + iy_{\mathbf{b}} \quad \text{and} \quad \mathbf{c} = x_{\mathbf{c}} + iy_{\mathbf{c}} .$$

Together with $t = x_t + iy_t$ we get

$$x_{\mathbf{c}} + iy_{\mathbf{c}} = [(x_{\mathbf{a}} + x_{\mathbf{b}})x_t + (-y_{\mathbf{a}} + y_{\mathbf{b}})y_t] + i[(y_{\mathbf{a}} + y_{\mathbf{b}})x_t + (x_{\mathbf{a}} - x_{\mathbf{b}})y_t] .$$

Interpreting \mathbb{C} as two-dimensional real space gives us

$$\begin{pmatrix} x_{\mathbf{c}} \\ y_{\mathbf{c}} \end{pmatrix} = \begin{pmatrix} x_{\mathbf{a}} + x_{\mathbf{b}} & -y_{\mathbf{a}} + y_{\mathbf{b}} \\ y_{\mathbf{a}} + y_{\mathbf{b}} & x_{\mathbf{a}} - x_{\mathbf{b}} \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = M \begin{pmatrix} x_t \\ y_t \end{pmatrix} . \quad (4.13)$$

In order to be able to solve for t , $\det(M)$ needs to be non-zero. But

$$\det(M) = x_{\mathbf{a}}^2 - x_{\mathbf{b}}^2 + y_{\mathbf{a}}^2 - y_{\mathbf{b}}^2 = r_{\mathbf{a}}^2 - r_{\mathbf{b}}^2 = \Delta \neq 0 .$$

Thus, by the preconditions of our theorem, M is invertible and

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} x_{\mathbf{a}} - x_{\mathbf{b}} & y_{\mathbf{a}} - y_{\mathbf{b}} \\ -y_{\mathbf{a}} - y_{\mathbf{b}} & x_{\mathbf{a}} + x_{\mathbf{b}} \end{pmatrix} \begin{pmatrix} x_{\mathbf{c}} \\ y_{\mathbf{c}} \end{pmatrix} .$$

Consequently,

$$\begin{aligned} t &= \frac{1}{\Delta} [(x_{\mathbf{a}} - x_{\mathbf{b}})x_{\mathbf{c}} + (y_{\mathbf{a}} - y_{\mathbf{b}})y_{\mathbf{c}} + i(-y_{\mathbf{a}} - y_{\mathbf{b}})x_{\mathbf{c}} + i(x_{\mathbf{a}} + x_{\mathbf{b}})y_{\mathbf{c}}] \\ &= \frac{1}{\Delta} (\mathbf{c} \bar{\mathbf{a}} - \bar{\mathbf{c}} \mathbf{b}) . \end{aligned}$$

This asserts our theorem. □

An example for this and the subsequent theorems will be given in a separate section (Section 4.2.7) after we have arrived at our translatorial normal form. In the above theorem the translation, making a curve coefficient vanish, was unique. This was ensured by the precondition $r_{\mathbf{a}} \neq r_{\mathbf{b}}$. This means we use both translatorial degrees of freedom for our translation. In case of $r_{\mathbf{a}} = r_{\mathbf{b}}$, there might still exist a translation of the given curve, annihilating \mathbf{c} . But in general we can only minimize $|\tilde{\mathbf{c}}|$.

Theorem 4.7 Let $r_a = r_b \neq 0$. Then any translation by t moves \mathbf{c} to a point $\tilde{\mathbf{c}}$ on L_c and annihilation is only possible if $\mathbf{c} \in L$.

PROOF Translating the curve means calculating $\tilde{\mathbf{c}} = \mathbf{c} - t\mathbf{a} - t\bar{\mathbf{b}}$. Let $t = r_t e^{i\xi}$ be arbitrary and $z \in \mathbb{C}$ with $z = t\mathbf{a} + \bar{t}\mathbf{b}$. Then

$$\begin{aligned} z &= r_a r_t \left(e^{i(\alpha+\xi)} + e^{i(\beta-\xi)} \right) = r_a r_t \left(e^{i(\frac{\alpha-\beta}{2}+\xi)} + e^{i(\frac{-\alpha+\beta}{2}-\xi)} \right) e^{i\frac{\alpha+\beta}{2}} \\ &= r_a r_t \cdot 2 \cos(\xi - \vartheta) e^{i\tau} . \end{aligned} \quad (4.14)$$

With $\mathbf{c} \in L_c$ we have $\tilde{\mathbf{c}} \in L_c$. Thus annihilation is only possible if $0 \in L_c$, which means that $L_c = L_0 = L$. \square

According to the above theorem $\tilde{\mathbf{c}} = 0$ is only possible if already $\mathbf{c} = 0$ or $\gamma \in \tau + \pi\mathbb{Z}$. Therefore in the case of Figure 4.10 annihilating \mathbf{c} is not possible. But by the above proof it becomes clear that for suitable choices of t the transformed $\tilde{\mathbf{c}}$ may be moved to any position on L_c . Hence the absolute value of such a coefficient can be minimized.

Remark 4.1 It is interesting to note that in case of $r_a = r_b \neq 0$ a translation by any $t \in T$ does not change the sub-leading coefficient \mathbf{c} . If $t \in T$, then $\xi = \vartheta \pm \frac{\pi}{2}$ and thus $z = t\mathbf{a} + \bar{t}\mathbf{b} = 0$. Consequently, $|\mathbf{c}|$ is not minimized by a unique t . If t_{min} is one of those translation parameters, so is any $t \in t_{min} + T$.

Theorem 4.8 Let $r_a = r_b \neq 0$, then any $t \in T_c$ minimizes $|\tilde{\mathbf{c}}|$. Especially this is true for $t_c \in T_c$. ($t_c = \frac{r_c \cos(\gamma-\tau)}{2r_a} e^{i\vartheta}$ was defined in the reference section 4.2.3.)

PROOF According to Remark 4.1 it suffices to show that $T_c \cap T^\perp = \{t_c\}$ minimizes $|\tilde{\mathbf{c}}|$. With Equation (4.14) and $t = r_t e^{i\xi} = t_c = \frac{r_c \cos(\gamma-\tau)}{2r_a} e^{i\vartheta}$ we have

$$\begin{aligned} \tilde{\mathbf{c}} &= \mathbf{c} - t_c \mathbf{a} - \bar{t}_c \mathbf{b} \\ &= r_c e^{i\gamma} - r_c \cos(\gamma - \tau) e^{i\tau} \\ &= r_c \sin(\gamma - \tau) e^{i(\tau + \frac{\pi}{2})} \in L^\perp . \end{aligned}$$

$\tilde{\mathbf{c}} \in L^\perp$ means having a minimal $|\tilde{\mathbf{c}}|$ according to Section 4.2.3. \square

With Theorems 4.7 and 4.8 the auxiliary lines from Section 4.2.3 come alive. L , L_c and L^\perp are directly connected to the possible locations of $\tilde{\mathbf{c}}$. T , T_c and T^\perp are connected to the corresponding translation parameters evoking a movement of \mathbf{c} . What is most interesting is that in case of $r_a = r_b \neq 0$ we only need one translatorial degree of freedom to minimize (or annihilate) a sub-leading coefficient \mathbf{c} : Translations with any parameter $t \in T^\perp$ do the trick.

4.2.5. Minimizing coefficients by a translation of fixed direction

By the above observations it becomes clear why studying translations by parameters restricted to a line is so important: There are curves where minimization of a coefficient leaves us with one translatorial degree of freedom, i.e. with the freedom to choose the translation parameter from a line. To obtain a normal form this freedom must be used. But not all constellations are suitable:

Suppose the translation parameter t is restricted to a linear set $R = \{ \lambda e^{i\xi} \mid \lambda \in \mathbb{R} \}$ for a given ξ . Let Δ_{kl} be a sub-triangle to a sub-leading coefficient $\mathbf{c} = c_{kl}$. The two leading coefficients in this triangle \mathbf{a} and \mathbf{b} exhibit an angle-difference of 2ϑ . Now, if $\xi = \vartheta \pm \frac{\pi}{2}$, then $R = T$ and according to Remark 4.1 we have $\tilde{\mathbf{c}} = \mathbf{c}$ for any $t \in R$. Thus nothing can be done to \mathbf{c} , when t is restricted the way it is. The reason behind this is the special angle-difference of the leading coefficients, making $\tilde{\mathbf{c}}$ independent of \mathbf{a} and \mathbf{b} under these special translations. This independence may be generalized to arbitrary non-leading coefficients in the following way:

Theorem 4.9 *Let Δ_{kl} be a sub-triangle to an arbitrary non-leading coefficient c_{kl} . Furthermore, let Δ_{kl} contain a row of coefficients with identical absolute value and with constant angle-difference of 2ϑ for any succeeding coefficient pair of that row. Then these coefficients give no contribution to the calculation of \tilde{c}_{kl} if and only if $t \in R$ and $R = \left\{ \lambda e^{i(\vartheta + \frac{\pi}{2})} \mid \lambda \in \mathbb{R} \right\}$ holds for the translation parameter t .*

PROOF Let $\mathbf{c} = c_{kl}$. The first coefficient of the special line described above shall be \mathbf{a} . According to the given angle-difference the second coefficient of that line is $\mathbf{a}e^{i2\vartheta}$. The subsequent one is $\mathbf{a}e^{i4\vartheta}$ and so on (see Figure 4.11). Let there be $r+1$ coefficients in that specific line. The sum of the indices of coefficients in that row is $k+l+r$. According to Equation (4.9) a translation changes \mathbf{c} into $\tilde{\mathbf{c}}$, which depends on the coefficients contained in Δ_{kl} . This Equation states that the contribution of our coefficients in our special row is a constant multiple of

$$\sum_{p=0}^r \binom{r}{p} t^p \bar{t}^{r-p} \cdot \mathbf{a} e^{2(r-p)i\vartheta} = \mathbf{a} \left(t + e^{2i\vartheta} \bar{t} \right)^r .$$

But with $t \in R$ we get

$$\mathbf{a} \left(t + e^{i \cdot 2\vartheta} \bar{t} \right)^k = \mathbf{a} e^{ik\vartheta} \left(t e^{-i\vartheta} + e^{i\vartheta} \bar{t} \right)^k = \mathbf{a} e^{ik\vartheta} r_t \left(e^{i\frac{\pi}{2}} + e^{-i\frac{\pi}{2}} \right)^k = 0 .$$

Thus the coefficients of our special row do not contribute to the calculation of $\tilde{\mathbf{c}}$. □

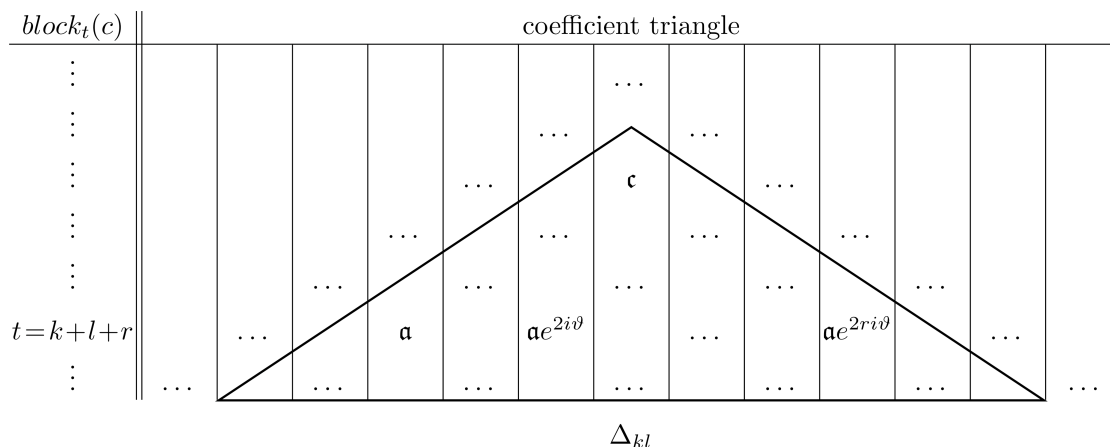


Figure 4.11.: Coefficient triangle with Δ_{kl} exhibiting a line of coefficients with identical absolute value and identical angle-differences

This theorem simplifies calculations immensely. For example, take a 2-leading coefficient $\mathbf{f} = c_{kl}$ as in Figure 4.9. Let \mathbf{a} , \mathbf{b} and \mathbf{c} in Δ_{kl} have the same absolute value and identical angle differences $-\alpha + \beta = -\beta + \gamma = 2\vartheta$. Translating the curve by a $t \in R$ with $R = \left\{ \lambda e^{i(\vartheta + \frac{\pi}{2})} \mid \lambda \in \mathbb{R} \right\}$ means that \mathbf{a} , \mathbf{b} and \mathbf{c} do not contribute and \mathbf{f} changes linearly, depending solely on \mathfrak{d} and \mathfrak{e} . (For an exemplary study of conics we refer to Section 4.2.7.)

Remark 4.2 *There is an interesting special case: Suppose all coefficients of each block have a common absolute value and all succeeding coefficient pairs have one and the same angle-difference 2ϑ . Then translations by any $t \in R = \left\{ \lambda e^{i(\vartheta + \frac{\pi}{2})} \mid \lambda \in \mathbb{R} \right\}$ have no effect on any curve coefficient and the curve is invariant under such translations. Thus the curve must be a collection of parallel lines, where some lines might coincide.*

At first sight, the above theorem and remark have nothing to do with minimizing coefficients for restricted translation parameters, but they give us the ingredients to continue our minimization process with k -leading coefficients for $k > 1$. If all blocks of leading up to $(k - 1)$ -leading coefficients have the same absolute and the same angle-difference for succeeding coefficient pairs, then these coefficients may be ignored. Thus minimizing k -leading coefficients with restricted translation parameters is still linear in t and \bar{t} . For the following cases, let t be restricted to $R = \left\{ \lambda e^{i\xi} \mid \lambda \in \mathbb{R} \right\}$ with a fixed ξ . Furthermore, let Δ_{kl} be an arbitrary coefficient-sub-triangle. If it consists of more than two rows, all coefficients below the second row shall be irrelevant in the sense of the above theorem. This means that for a curve of degree d the coefficient $\mathbf{c} = c_{kl}$ changes linearly by

$$\tilde{\mathbf{c}} = \mathbf{c} - (d - k - l)(t\mathbf{a} + \bar{t}\mathbf{b}) . \quad (4.15)$$

For sub-leading coefficients \mathbf{c} this degenerates to the well known Equation (4.11). It is instructive to compare the factor $(d - k - l)$ with the coefficients in the transformation formula (4.8), especially with those in blocks in the first sup-diagonal.

Now we have to study minimization in two cases, analogously to the last section: $r_{\mathbf{a}} \neq r_{\mathbf{b}}$ (Theorem 4.10) and $r_{\mathbf{a}} = r_{\mathbf{b}} \neq 0$ (Theorem 4.11).

Theorem 4.10 *Let $r_{\mathbf{a}} \neq r_{\mathbf{b}}$ and let the set of available translation parameters be restricted by $t \in \left\{ r_t e^{i\xi} \mid r_t \in \mathbb{R} \right\}$. Let $m = r_m e^{i\mu} = \mathbf{a}e^{i\xi} + \mathbf{b}e^{-i\xi}$. Then*

$$t = \frac{1}{2(d - k - l)} \left(\frac{\mathbf{c}}{m} + \frac{\bar{\mathbf{c}}}{\bar{m}} \right) e^{i\xi}$$

leads to a $\tilde{\mathbf{c}}$ of minimal absolute value with respect to all available translation parameters. The translation annihilates $\tilde{\mathbf{c}}$ if and only if $\gamma \equiv \mu \pmod{\pi}$.

PROOF Let $t = r_t e^{i\xi}$ for some $r_t \in \mathbb{R}$, $s = d - k - l$ and $m = r_m e^{i\mu}$. Then Equation (4.15) evaluates to

$$\tilde{\mathbf{c}} = \mathbf{c} - sr_t(\mathbf{a}e^{i\xi} + \mathbf{b}e^{-i\xi}) = \mathbf{c} - sr_t r_m e^{i\mu} .$$

Minimizing $|\tilde{\mathbf{c}}|$ in this context is equivalent to removing the $e^{i\mu}$ -component from \mathbf{c} . Consequently \mathbf{c} may be annihilated if it has only an $e^{i\mu}$ -component. This component is removed in case of $sr_t r_m e^{i\mu}$ being the orthogonal projection of \mathbf{c} onto the line $\left\{ \lambda e^{i\mu} \mid \lambda \in \mathbb{R} \right\}$. This is the case if

$$sr_t r_m = r_{\mathbf{c}} \cos(\gamma - \mu) \quad \text{or} \quad r_t = \frac{r_{\mathbf{c}} \cos(\gamma - \mu)}{sr_m} = \frac{1}{2s} \left(\frac{\mathbf{c}}{m} + \frac{\bar{\mathbf{c}}}{\bar{m}} \right) . \quad \square$$

Theorem 4.11 *Let $r_a = r_b \neq 0$ and let the set of available translation parameters be restricted to $t \in R = \{r_t e^{i\xi} \mid r_t \in \mathbb{R}\}$. Furthermore, let $\frac{-\alpha+\beta+\pi}{2} \notin \xi + \pi\mathbb{Z}$. Then*

$$t = \frac{2t_c}{d-k-l} \cdot \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}}$$

leads to a \tilde{c} of minimal absolute value with respect to all available translation parameters. The translation annihilates \tilde{c} if and only if $c \in L$.

PROOF The precondition $\frac{-\alpha+\beta+\pi}{2} \notin \xi + \pi\mathbb{Z}$ assures that a translation by a $t \in R$ may have an effect on c at all. T and R are not parallel.

In Theorem 4.8, we stated that the set of all translation parameters leading to a minimal $|\tilde{c}|$ is $T_c = t_c + T$. The situation here differs from that former one by the factor $(d-k-l)$ in the transformation law (compare Equations (4.11) and (4.15)). Consequently, here any $t \in \frac{t_c}{d-k-l} + T$ would minimize $|\tilde{c}|$.

However, t is restricted to $R \not\parallel T$ by our preconditions. Thus the desired translation parameter is contained in the intersection of those two sets. Bearing in mind that $t_c \in T^\perp$ we get:

$$\begin{aligned} t = r_t \cdot e^{i\xi} &= \frac{\left| \frac{t_c}{d-k-l} \right|}{\cos(\xi - \gamma)} \cdot e^{i\xi} = \frac{t_c}{d-k-l} \cdot \frac{e^{-i\vartheta}}{\cos(\xi - \vartheta)} \cdot e^{i\xi} \\ &= \frac{2t_c}{d-k-l} \cdot \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}} \end{aligned}$$

This proves our theorem. □

With these ingredients we are now able to construct translatorial normal forms.

4.2.6. Translatorial normal form

To establish a translatorial normal form for curves, we must make use of both translatorial degrees of freedom and move each curve in a predefined manner. The result must be independent of any previous translations. We can achieve this by selectively annihilating or minimizing coefficients making use of the last two sections.

Building a normal form for a curve of degree d , we proceed block-wise, starting with the sub-leading coefficients (leading coefficients are invariant under translations) and going upwards in our coefficient triangle. Given a block we can greedily annihilate coefficients: If more than one coefficient can be annihilated, we start with the largest coefficient with respect to \succ . If annihilation is not possible, we try minimization within this block, again starting with the largest coefficient with respect to \succ . Either annihilation or minimization of a sub-leading coefficient is possible because either Theorem 4.6 or Theorem 4.8 applies. If a translatorial degree of freedom is left, we try to annihilate (or if impossible minimize) another sub-leading coefficient. Thereby larger coefficients with respect to \succ are preferred. If none is possible, we neglect the leading coefficients, because all succeeding pairs must exhibit the same angle-difference and we can pretend that the degree of the curve is $d-1$. Then we start over from the beginning and greedily annihilate (or minimize) coefficients. Eventually we have used both translatorial degrees of freedom or otherwise the curve consists of parallel lines. Anyway we end up in a predefined curve. Summarizing this we have

Theorem 4.12 *Block-wise greedily annihilating - and if impossible minimizing - coefficients translates given curves in a predefined form, a normal form.*

PROOF Application of Theorems 4.6, 4.8, 4.10, 4.11 and Remark 4.2 proves the assumption. \square

Remark 4.3 *We give a detailed step-by-step algorithm for establishing the above normal form for a given curve. Thereby we will use the necessary translatorial degrees of freedom. Additionally we refer to the translation parameters, which must be used and refer to results from the last two sections. If a case does not apply in the given situation, we step over to the next.*

Whenever a process of annihilation or minimization is applicable to multiple coefficients it shall be applied to the largest one with respect to \succ . The used expressions always refer to the corresponding Δ_{kl} . Whenever normal form is reached “ \diamond ” is displayed.

In this algorithm we can see the freedom of swapping steps 2b and 2c and/or steps 2d and 2e. However, for a normal form the order must be specified in advance. Changing the order leads to different normal forms. With all the ingredients we may of course think of a whole lot more normal forms: For example one, where annihilation is ignored and only minimization is greedily performed. However, most of these different normal forms have an unfavorable drawback, when it comes to symmetry-detection.

Translatorial normal-form-algorithm:

- If there is a sub-leading coefficient \mathbf{c} with $r_{\mathbf{a}} \neq r_{\mathbf{b}}$:
translate by $t = \frac{1}{r_{\mathbf{a}}^2 - r_{\mathbf{b}}^2} (\mathbf{c} \bar{\mathbf{a}} - \bar{\mathbf{c}} \mathbf{b})$ (Theorem 4.6). \diamond

- Otherwise all leading coefficients have identical absolute value ($\neq 0$) and thus two translations are necessary to establish normal form.

1. If for any sub-leading coefficient $\mathbf{c} \in L$: annihilate \mathbf{c} by $t = t_{\mathbf{c}}$. Otherwise minimize the first sub-leading coefficient by $t = t_{\mathbf{c}}$ (Theorem. 4.8).

In any case one translatorial degree of freedom is used and further parameters t are restricted to $R = \{r_t e^{i\xi} \mid r_t \in \mathbb{R}\}$ with ξ being equal to σ from this first translation.

2. a) There is any sub-leading $\mathbf{c} \in L$ and $2\vartheta + \pi \not\equiv 2\xi \pmod{2\pi}$: annihilate \mathbf{c} by $t = 2t_{\mathbf{c}} \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}}$. Otherwise if the same holds for any sub-leading $\mathbf{c} \notin L$: minimize \mathbf{c} by $t = 2t_{\mathbf{c}} \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}}$ (Theorem 4.11). \diamond

** Now all succeeding leading coefficient-pairs have the same angle-difference and are irrelevant for translations by $t \in R$ (Theorem 4.9). Reduce the degree of the curve by one and neglect the leading coefficients. (Former sub-leading coefficients are now leading.) Let $m = r_m e^{i\mu} = \mathbf{a} e^{i\xi} + \mathbf{b} e^{-i\xi}$ with \mathbf{a} and \mathbf{b} corresponding to the cases.*

- b) There is any sub-leading coefficient $\mathbf{c} = c_{kl}$ with $r_{\mathbf{a}} \neq r_{\mathbf{b}}$ and $e^{i\mu} = \pm e^{i\gamma}$: annihilate \mathbf{c} by $t = \frac{1}{2(d-k-l)} \left(\frac{\mathbf{c}}{m} + \frac{\bar{\mathbf{c}}}{\bar{m}} \right) e^{i\xi}$ (Theorem 4.10). \diamond

- c) There is any sub-leading coefficient $\mathbf{c} = c_{kl}$ with $r_{\mathbf{a}} = r_{\mathbf{b}} \neq 0$, $\mathbf{c} \in L$ and $2\vartheta + \pi \not\equiv 2\xi \pmod{2\pi}$: annihilate \mathbf{c} by $t = \frac{2t_{\mathbf{c}}}{d-k-l} \cdot \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}}$ (Theorem 4.11). \diamond

d) There is any sub-leading coefficient $\mathbf{c} = c_{kl}$ with $r_{\mathbf{a}} \neq r_{\mathbf{b}}$: minimize \mathbf{c} by $t = \frac{1}{2(d-k-l)} \left(\frac{\mathbf{c}}{m} + \frac{\bar{\mathbf{c}}}{\bar{m}} \right) e^{i\xi}$ (Theorem 4.10). \diamond

e) There is any sub-leading coefficient $\mathbf{c} = c_{kl}$ with $r_{\mathbf{a}} = r_{\mathbf{b}} \neq 0$ and $2\vartheta + \pi \not\equiv 2\xi \pmod{2\pi}$: minimize \mathbf{c} by $t = \frac{2t_{\mathbf{c}}}{d-k-l} \cdot \frac{e^{2i\xi}}{e^{2i\xi} + e^{2i\vartheta}}$ (Theorem 4.11). \diamond

If there is still one translatorial degree of freedom left, we stop calculations if c_{00} is sub-leading. In this case the curve is invariant under any translation by $t \in \mathbb{R}$ (Remark 4.2). Otherwise we start over with $*$. \diamond

Assume that a curve g is in translatorial normal form. Rotating this curve preserves the absolute value of each coefficient, changing only the corresponding angle by a multiplication with $e^{ik\varphi}$ for some angle $k\varphi$. Thus transforming g into rotational normal form preserves the translatorial normal form. Taking the two normal forms together we have a normal form with respect to arbitrary rotations and translations: a normal form with respect to orientation preserving Euclidean transformations.

Naturally there are much more possible translatorial normal forms. The benefit of this one is that it easily permits feature extraction (see Chapter 5). In Appendix B we discuss a “normal form” which is proposed for shape-recognition purposes.

4.2.7. Annihilation and normal form for complexified conics

We exemplify the process of annihilation with conics. Knowledge from the theory of conics is not needed but may help to understand. The treatment of curves of higher order is in principle not very different from the following. Throughout this section we will use the notation from above and rid ourselves of the double-indices by setting

$$\mathbf{a} = r_{\mathbf{a}}e^{i\alpha} = c_{20}, \quad \mathbf{b} = r_{\mathbf{b}}e^{i\beta} = c_{11}, \quad \mathbf{c} = r_{\mathbf{c}}e^{i\gamma} = c_{10} \quad \text{and} \quad \mathfrak{d} = r_{\mathfrak{d}}e^{i\delta} = c_{00} .$$

The radii $r_{\mathbf{a}}$, $r_{\mathbf{b}}$, $r_{\mathbf{c}}$ and $r_{\mathfrak{d}}$ shall be real and non-negative unless explicitly stated otherwise. Thus a conic is given by

$$q(z, \bar{z}) = \mathfrak{d} + 2\mathbf{c}z + 2\bar{\mathbf{c}}\bar{z} + \mathbf{a}z^2 + 2\mathbf{b}z\bar{z} + \bar{\mathbf{a}}\bar{z}^2 = 0 .$$

The only sub-leading coefficients of $q(z, \bar{z})$ are \mathbf{c} and $\bar{\mathbf{c}}$. Because of $\mathbf{c} = \overline{(\bar{\mathbf{c}})}$, annihilating \mathbf{c} means automatically annihilating $\bar{\mathbf{c}}$ and vice versa. Thus, without loss of generality, we focus on \mathbf{c} alone.

According to Theorem 4.6 we can perform annihilation when $r_{\mathbf{a}} \neq r_{\mathbf{b}}$. Figure 4.12.a illustrates a quadric curve satisfying this condition. The uniquely determined annihilating translation vector

$$t = \frac{\mathbf{c}\bar{\mathbf{a}} - \bar{\mathbf{c}}\mathbf{b}}{r_{\mathbf{a}}^2 - r_{\mathbf{b}}^2}$$

according to Theorem 4.6 is also indicated. A translation by this parameter t uses both translatorial degrees of freedom. This concludes our study of this kind of curves.

Now let $r_{\mathbf{a}} = r_{\mathbf{b}}$. In this case our auxiliary lines come to life. Figures 4.12.b+c illustrate those lines, given already by the leading coefficients. By $\mathbf{b} = c_{11} \in \mathbb{R}$ we have $e^{i\beta} = \text{sign}(\mathbf{b})$ and thus

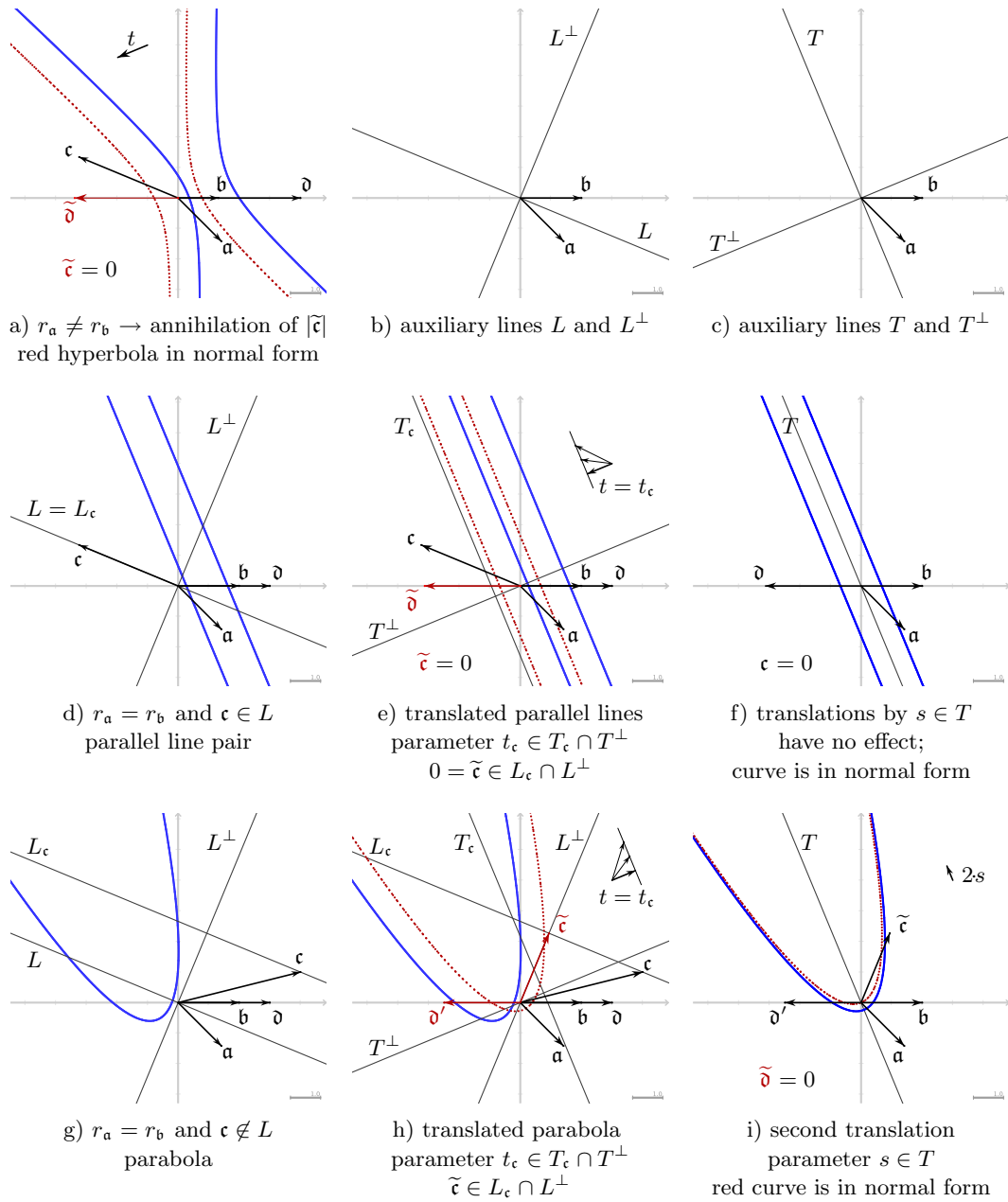


Figure 4.12.: Translation of a parabola and a pair of parallel lines

$$\begin{aligned}
 0 = q(z, \bar{z}) &= \vartheta + 2cz + 2\bar{c}\bar{z} + r_a (e^{i\alpha}z + 2 \cdot \text{sign}(\mathbf{b})z\bar{z} + e^{-i\alpha}\bar{z}) \\
 &= \vartheta + 2cz + 2\bar{c}\bar{z} + r_a \left(e^{i\frac{\alpha}{2}}z + \text{sign}(\mathbf{b})e^{-i\frac{\alpha}{2}}\bar{z} \right)^2 \\
 &= \vartheta + 2cz + 2\bar{c}\bar{z} + r_a \left(e^{i\frac{\alpha}{2}}z + e^{i\beta}e^{i\frac{-\alpha}{2}}\bar{z} \right)^2 \\
 &= \vartheta + 2cz + 2\bar{c}\bar{z} + r_a e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}}z + e^{i\frac{-\alpha+\beta}{2}}\bar{z} \right)^2 .
 \end{aligned}$$

The condition $r_a = r_b$ affects only the leading terms of q . Let us have a short look at the projective extension of the complex plane - to be more precise: at the restriction of q to the line at infinity, namely $r_a e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right)^2$. There we can see a double root for any $z \in T$ with

$$T = \left\{ \lambda e^{i\frac{-\alpha+\beta+\pi}{2}} \mid \lambda \in \mathbb{R} \right\} .$$

From the theory of quadric equations, we know that double roots at infinity are characteristic for parabolas or pairs of parallel lines. In Figure 4.12 both cases are illustrated. Looking back to the case where $r_a \neq r_b$, we must have had hyperbolas, ellipses, circles or intersecting line pairs. There the linear (sub-leading) part of the equation was extinguishable by translations, which is also confirmed by the theory of quadratic equations. Now with $r_a = r_b$ we have two cases: either $\mathbf{c} \in L$ (Figure 4.12.d) or $\mathbf{c} \notin L$ (Figure 4.12.e) with $L = \left\{ \lambda e^{i\tau} \mid \lambda \in \mathbb{R} \right\}$ and $\tau = \frac{\alpha+\beta}{2}$. In the first case annihilation is possible by Theorems 4.7 and 4.8. With $\mathbf{c} = r_c e^{i\tau}$ for a $r_c \in \mathbb{R}$ we get

$$\begin{aligned} q(z, \bar{z}) &= \mathfrak{d} + 2r_c \left(e^{i\frac{\alpha+\beta}{2}} z + e^{i\frac{-\alpha-\beta}{2}} \bar{z} \right) + r_a e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right)^2 \\ &= \mathfrak{d} + 2r_c e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha-3\beta}{2}} \bar{z} \right) + r_a e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right)^2 . \end{aligned}$$

With $\beta \in \pi\mathbb{Z}$ we have $e^{i\frac{-3\beta}{2}} = e^{i\frac{\beta}{2}}$ and we can employ Vieta's theorem to get

$$\begin{aligned} q(z, \bar{z}) &= \mathfrak{d} + 2r_c e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right) + r_a e^{i\beta} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right)^2 \\ &= \left(\xi + \sqrt{r_a} e^{i\frac{\beta}{2}} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right) \right) \cdot \left(\zeta + \sqrt{r_a} e^{i\frac{\beta}{2}} \left(e^{i\frac{\alpha-\beta}{2}} z + e^{i\frac{-\alpha+\beta}{2}} \bar{z} \right) \right) \end{aligned}$$

for suitable $\xi, \zeta \in \mathbb{R}$, satisfying $\xi \cdot \zeta = c_{00}$ and $\xi + \zeta = \frac{2r_c}{\sqrt{r_a}} e^{i\frac{\beta}{2}}$. This means that q degenerates in two parallel lines: q is a product of two terms which are linear in z and \bar{z} exhibiting equal leading coefficients. This again corresponds nicely to the theory of conics.

Now taking our two parallel lines, we may translate them such that the origin is equally far away from both lines. Consequently $\xi = -\zeta$ holds for the translated curve. Thus $\xi + \zeta = 0 = \frac{2r_c}{\sqrt{r_a}}$, which is only satisfied if $r_c = 0$. This means that $\mathbf{c} = 0$ after this translation (see Figure 4.12.e). In fact, annihilating \mathbf{c} in this case we need Theorem 4.8 and a $t \in T_c$. An additional second translation with translation parameter s restricted to $s \in T$ is not needed: Any such translation would leave the parallel lines invariant. Anyway, there is no further sub-leading coefficient to minimize and $c_{00} = \mathfrak{d}$ may not be changed by the use of an $s \in T$. As a result our parallel lines are in normal form (see Figure 4.12.f).

Now only the examination of parabolas is missing, i.e. when $r_a = r_b$ and $\mathbf{c} \notin L$ (Figure 4.12.g). In a first translation according to Theorem 4.8, $|\tilde{\mathbf{c}}|$ may be minimized by any $t \in T_c$ (Figure 4.12.h). In a second translation by parameter $s \in T$, we may only change the constant term $c_{00} = \mathfrak{d}$. But the leading coefficients have no influence on it and $\tilde{\mathfrak{d}}$ depends only on the minimized \mathbf{c} and $\bar{\mathbf{c}}$. These two coefficients have the same absolute value and are conjugate. Thus automatically $\tilde{\mathfrak{d}}$ can only be moved on the real axis, which is fine because \tilde{c}_{00} must be real anyway. Consequently $\tilde{\mathfrak{d}}$ is annihilable and the resulting parabola is in normal form and must contain the origin (Figure 4.12.i).

Retrospectively, we can easily observe what the single translations did with the original parabola: The second translation was in the direction of the axis of the parabola, sending

the apex to the origin. Thus the first translation moves the original axis such that it passes through the origin.

4.2.8. Invariants under translation

We will confine ourselves to curves where the absolute values $|c_{kl}|$ of all leading coefficients c_{kl} are not equal. By Theorem 4.6 there is at least one sub-leading coefficient c_{pq} which can be annihilated by a unique translation: Let c_{pq} be the one with greatest p . A normal form for a given curve with respect to translation can be achieved by annihilating the coefficient c_{pq} .

For curves with leading coefficients of the same absolute value the translation process, transforming a curve into translatorial normal-form, is uniquely determined: $|T_{\min}| = 1$. Thus the coefficients of the translated curve are invariants of the curve describing equation with respect to translations. The leading coefficients are not affected by translations and are thus also invariant. The other coefficients \tilde{c}_{kl} with $k + l < d$ can be calculated using Expression (4.7) and have a more complicated structure.

Take a quadric $q(z, \bar{z}) = 0$ for example. Annihilating a sub-leading coefficient by translation means $\tilde{c}_{10} = 0$. The translated quadric will be

$$\tilde{q}(z, \bar{z}) = \tilde{c}_{00} + \tilde{c}_{20}z^2 + \tilde{c}_{11}z\bar{z} + \tilde{c}_{02}\bar{z}^2 .$$

The translation parameter t is $t = \frac{c_{10}\overline{c_{20}} - \overline{c_{10}}c_{11}}{|c_{20}|^2 - |c_{11}|^2}$ according to Theorem 4.6. Thus by the transformation law (4.7) we get (with $\Delta = c_{20}c_{02} - c_{11}^2$):

$$\begin{aligned} \tilde{c}_{00} &= q(t, \bar{t}) = c_{00} + \frac{1}{\Delta} (2c_{10}c_{01}c_{11} - c_{10}^2c_{02} - c_{01}^2c_{20}) \\ &= \frac{1}{\Delta} \det \begin{pmatrix} c_{20} & c_{11} & c_{10} \\ c_{11} & c_{02} & c_{01} \\ c_{10} & c_{01} & c_{00} \end{pmatrix} = \frac{\det(M)}{\Delta} , \end{aligned}$$

$$\tilde{c}_{10} = 0, \quad \tilde{c}_{01} = 0,$$

$$\tilde{c}_{20} = c_{20}, \quad \tilde{c}_{11} = c_{11} \quad \text{and} \quad \tilde{c}_{02} = c_{02} .$$

We have the following invariants of the quadric equation: \tilde{c}_{00} , $\tilde{c}_{20} = \overline{\tilde{c}_{02}}$ and \tilde{c}_{11} . Therefore $\Delta = |c_{20}|^2 - c_{11}^2$ is invariant. Thus $\Delta \cdot \tilde{c}_{00} = \det(M)$ is also invariant. $\det(M)$ is even a projective invariant of the quadric equation.

Invariants of the conic as curve can be obtained by normalizing the coefficients. This can be done by dividing the coefficients by the Frobenius norm $\|c\|_F$ of the coefficient vector c . Thus $\frac{\det(M)}{\|c\|_F^3}$ and $\frac{\Delta}{\|c\|_F^2}$ are invariants of the curve. If $\det(M) \neq 0$, the same is true for the combination $\frac{\Delta^3}{\det(M)^2}$.

It is instructive to analyze what this specific invariant means. To do so, consider the very simple quadric given by $ax^2 + by^2 - 1 = 0$. If both a and b are positive, this quadric is an ellipse, whose axes of symmetry are aligned to the coordinate system. In this case, the complex representation turns out to be $\frac{a-b}{4}z^2 + \frac{a+b}{2}z\bar{z} + \frac{a-b}{4}\bar{z}^2 - 1$. Thus we have $c_{20} = c_{02} = \frac{a-b}{4}$, $c_{11} = \frac{a+b}{4}$ (remember we have multinomial pre-scaling), $c_{10} = 0$ and $c_{00} = -1$. We want to study what happens if we insert these coefficients into $\frac{\Delta^3}{\det(M)^2}$. First observe that the numerator as well as the denominator are of homogeneous degree

6 in the c_{ij} . Thus we may multiply all coefficients by 4 without altering the value of the expression. So we get:

$$\frac{\det^3 \begin{pmatrix} a-b & a+b \\ a+b & a-b \end{pmatrix}}{\det^2 \begin{pmatrix} a-b & a+b & 0 \\ a+b & a-b & 0 \\ 0 & 0 & -4 \end{pmatrix}} = -\frac{ab}{4}.$$

This amazingly simple expression is nothing else but $-(\Omega \cdot \frac{1}{2\pi})^2$, where $\Omega = \pi\sqrt{ab}$ is the area of our ellipse. Thus we have a Euclidean explanation for our expression being invariant: It is just a function of the area!

More translatorial invariants can be obtained by establishing translatorial normal form and taking the resulting curve coefficients. However, they will not be necessary in most cases because we want to establish a normal form anyway and therefore do not need to reconstruct the curve up to translations. The translatorial normal form allows an easy symmetry detection, which would be quite difficult for arbitrary curves (see Chapter 5).

4.3. Scaling

Another elementary operation is scaling. In the last chapter we already dealt with translations and so we confine ourselves here to scalings with respect to the origin as center. Together with rotations and translations, we get the orientation preserving Euclidean motions. In this chapter we will analyze the effects of scaling on the curve coefficients and take a look at a normal form and invariants with respect to scalings.

4.3.1. Effects of scaling on curve coefficients

Let us start with the analysis of what scaling does to points contained in our plane. With the homogenizing component h we have

$$\begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix}. \quad (4.16)$$

Thereby $\lambda \in \mathbb{R}_{>0}$ is the scaling parameter. $\lambda = 0$ is excluded because everything would shrink to a point. Negative scaling parameters may be ignored in the presence of rotations: Rotating a curve around the origin by an angle $\psi = \pi$ is equivalent to scaling by $\lambda = -1$. Consequently, a scaling by a negative λ may be split up to a rotation by π with subsequent scaling by $|\lambda|$. The effect of scaling on the curve coefficients is given by

Theorem 4.13 *Let g be a curve with coefficients c_{kl} and let \tilde{g} be the by $\lambda \in \mathbb{R}_{>0}$ scaled copy $\tilde{g} = \tilde{g}_\lambda$ of g with coefficients \tilde{c}_{kl} . Then the mapping on the coefficients, induced by the scaling, is*

$$c_{kl} \mapsto \tilde{c}_{kl} = \lambda^{-(k+l)} c_{kl}. \quad (4.17)$$

PROOF Equation (4.16) directly proves (4.17). □

4.3.2. Normal form and invariants

Theorem 4.13 shows that a (uniform) scaling of the plane results in a (non-uniform) scaling of the curve coefficients. All coefficients c_{kl} are scaled except c_{00} , which consequently is an invariant under these transformations. But there are whole curves, too, which are invariant under scaling by an arbitrary parameter: In the Euclidean plane these are exactly the curves consisting of one or more lines passing through the origin. In the projective plane this set is complemented by the points on the line at infinity. Talking about normal forms with respect to scalings, these curves are already in normal form. All other curves may be transformed to a sheer wealth of different normal forms: We can normalize any collection $S \setminus \{c_{00}\}$ of coefficients $c_{kl} \neq c_{00}$, containing at least one non-vanishing coefficient, with respect to any norm. In the following we list a few of them:

- We can scale a curve such that the largest non-vanishing coefficient $|c_{kl}|$ with respect to \succ has an absolute value of one: Simply choose λ to be the unique positive real $(k + l)$ -th root of $|c_{kl}|$.
- We can make use of the block-structure of the curve-coefficient-vector as defined in Definition 2.3. Via scaling we may achieve that the sum of the absolute values of coefficients in a non-vanishing $block_t(c)$ ($t \neq 0$) equals one: Set λ to the unique positive t -th real root of $\|block_t(c)\|_1 = \sum_{l=0}^t |c_{t-l,l}|$. Generically we only know that the leading block $block_d(c)$ contains at least one non-zero coefficient.
- A normal form using *all* coefficients except c_{00} may also be calculated: Let c^p be the projection of the curve coefficient vector into a vector space of one dimension less such that $c_{kl}^p = c_{kl}$ except for $k = l = 0$. Selecting λ to be the unique positive root of

$$0 = 1 - \sum_{t=1}^d \lambda^t \underbrace{\sum_{l=0}^t |c_{t-l,l}|}_{\geq 0}$$

leads to $\|c^p\|_1 = 1$. The root is unique due to the non-negative coefficients of the polynomial in λ , which is subtracted from unity. Thus this polynomial is monotonically increasing for positive λ and zero for $\lambda = 0$ because there must be at least one non-vanishing leading coefficient.

With all these normal forms, care has to be taken about the overall scaling of the coefficient vector because of the identity $V(f) = V(\lambda f)$ for all $\lambda \neq 0$. In any case these normal forms do not interfere with a rotational or translational normal form potentially established earlier: The angles φ_{kl} of coefficients $c_{kl} = r_{kl}e^{i\varphi_{kl}}$ are preserved under scalings by positive λ and the property of being a vanishing or non-vanishing coefficient is also invariant under scalings. Thus taking translatorial, rotational and one of the scaling normal forms together, we obtain a normal form with respect to orientation preserving Euclidean mappings of the plane.

As invariant expressions with respect to scalings, we get all quotients of products of coefficients with equal index-sum per product: Let Q as well as D be a family of coefficients. A single coefficient c_{kl} may be contained in Q or D more than once. Its index-sum is $k + l$. The sum over all index-sums of all coefficients in Q shall be equal to the one of

D . Index-sums to coefficients which are contained more than once have to be counted according to their multiplicity. Then

$$\frac{\prod_{c_{kl} \in Q} c_{kl}}{\prod_{c_{pq} \in D} c_{pq}}$$

is an invariant with respect to scalings. This property directly drops out of Equation (4.17) from the end of the last section.

4.4. Projective transformations

In the previous sections we studied special Euclidean transformations. Now we want to focus on projective ones. In general, we can proceed analogously to the rotations, translations or scalings. But we will see very early the impracticability of this ansatz to derive a normal form or invariants. In the literature, a complete projective classification to higher order curves can not be found. “Higher order curves” already means curves of order five (compare with [11]). To overcome most of these problems, we propose using diagram techniques. We will follow this path for a while and refer to Section 3.2. Further observation on diagrammatical projective invariants are covered by [1].

4.4.1. Effects of projective transformations on curve coefficients

Projective transformations of the plane may be considered as mere matrix multiplications by matrices with a non-vanishing determinant. After complexification and simplification, a real projective transformation of our plane curve is given by

$$\begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} m & n & o \\ \bar{n} & \bar{m} & \bar{o} \\ p & \bar{p} & \kappa \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \quad (4.18)$$

with $m, n, o, p \in \mathbb{C}$ and $\kappa \in \mathbb{R}$. This transformation has nine real degrees of freedom. We might restrict the transformation-matrix to exhibit a determinant equal to unity: This is due to the fact that in projective geometry $(z, \bar{z}, h)^T$ is not at all different from $\det(M) \cdot (z, \bar{z}, h)^T$. This would reduce the degrees of freedom by one.

However, extracting structures and thus normal forms or invariants from this point of view is too complicated. Already the transformation of the curve-coefficients of a conic, induced by the above transformation, is too messy:

$$\begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \end{pmatrix} \mapsto M^{-1} \begin{pmatrix} \kappa^2 & 2p\kappa & 2\kappa\bar{p} & p^2 & 2\bar{p}p & \bar{p}^2 \\ \kappa o & p o + \kappa m & \bar{p} o + \kappa n & p m & \bar{p} m + p n & \bar{p} n \\ \kappa \bar{o} & p \bar{o} + \kappa \bar{n} & \bar{p} \bar{o} + \kappa \bar{m} & p \bar{n} & \bar{p} \bar{m} + p \bar{n} & \bar{p} \bar{m} \\ o^2 & 2m o & 2n o & m^2 & 2n m & n^2 \\ o \bar{o} & m \bar{o} + o \bar{n} & n \bar{o} + o \bar{m} & m \bar{n} & n \bar{n} + m \bar{m} & n \bar{m} \\ \bar{o}^2 & 2\bar{n} o & 2\bar{m} o & \bar{n}^2 & 2\bar{n} \bar{m} & \bar{m}^2 \end{pmatrix}^{-1} M \begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \end{pmatrix}$$

Thereby M denotes the diagonal-matrix containing the multinomial coefficients. The expansion, especially the inversion of the matrix in the middle, results in a very in-transparent expression. Extracting structures and invariants must be done by using a different approach.

4.4.2. Projective invariant expressions

A promising ansatz is the use of tensor-diagrams. According to Theorem 3.3 we can draw closed diagrams using only the co- and contravariant ε -tensors together with the curve tensor and we will obtain an invariant. Thus it is highly interesting which closed diagrams are generically non-vanishing. We call these diagrams non-trivial. Having all these at hand, we are able to form *any* invariant according to Theorems 3.2 and 3.3.

The ultimate goal would be to determine a (minimal) system of diagrams, which permits the generation of any projective invariant. If we had such a system for any degree of a curve, we would be able to reconstruct any curve from its invariants. According to a famous basis theorem of Hilbert (see [25] or [51]) the corresponding minimal systems are finite.

The question now is how to find such basis-sets for curves of a fixed degree. We do not solve this problem here, but we propose a few steps in the direction to the solution which seem promising to us.

According to Theorem 3.2 and 3.3, projective invariants of a curve are represented by the set \mathcal{I} of closed diagrams consisting of the curve tensor together with co- and contravariant ε -tensors. Due to the ε - δ -rule from Section 3.3.6 we know that two directly connected ε -tensors may be replaced. Applying this rule to the diagrams in \mathcal{I} we are left with \mathcal{I}' . These diagrams consist of curve-tensors and not directly connected co- and contravariant ε -tensors. However, the curve-tensor is per definition covariant and thus not directly connectable to a covariant ε -tensor. As the diagrams in \mathcal{I}' are closed, they do not contain covariant ε -tensors.

Theorem 4.14 *The set of all projective invariants of a curve may be described by the set of closed diagrams containing (multiple copies of) the corresponding curve-tensor and contravariant ε -tensors.*

If we have a non-trivial invariant, then no ε -tensor is connected twice or even more often to one and the same curve-tensor. This is because

$$\begin{array}{c} \textcircled{\widehat{G}} \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \nearrow \\ \searrow \end{array} \rightarrow \\ \textcircled{\widehat{G}} \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \nearrow \\ \searrow \end{array} \rightarrow \end{array} = \begin{array}{c} \textcircled{\widehat{G}} \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \nearrow \\ \searrow \end{array} \rightarrow \\ \textcircled{\widehat{G}} \begin{array}{l} \nearrow \\ \searrow \end{array} \begin{array}{l} \nearrow \\ \searrow \end{array} \rightarrow \end{array} = 0 . \quad (4.19)$$

\widehat{G} is totally symmetric and thus we may interchange the incoming edges. However, if we swap the outgoing edges at the ε -tensor, we have to change the sign. Thus the original expression must have been identical to its negative and therefore it is vanishing.

In order to build a diagrammatical invariant at all, we are not allowed to have dangling arcs. But this has implications on the degree of a non-trivial invariant in the curve coefficients. This degree is given by the number of copies of the involved curve-tensor. It is constrained by the fact that the variance $var(\widehat{G})$ of a curve-tensor \widehat{G} is given by the degree d of the corresponding curve and the variance $var(\varepsilon) = 3$ of the ε -tensor. Consequently, there is no non-trivial invariant or non-trivial closed diagram of degree one or two in the curve-coefficients.

Theorem 4.15 *If d is the degree of a curve and k the degree of a non-trivial invariant of this curve, then $k \cdot d \equiv 0 \pmod{3}$.*

Now the search for a diagrammatical basis of invariants may start. At first we may fix the degree d of curves under observation. Then we start to list a few diagrams and try

to extract building blocks with interesting attributes that may help listing or ignoring further diagrams: For example diagram-parts which vanish identically reduce the amount of diagrams that have to be considered. One such rule is already known: If we have a diagram where one curve-tensor is directly connected to an ε -tensor more than once, then it must vanish (see Equation (4.19)). We are also interested in (more complicated) diagrammatical calculation rules to be able to cut complex and big diagrams down to small ones. These rules are nothing more than adapted Grassmann-Plücker-Relations. In the end, we must reach the point where we listed a basis of invariants due to its secured finiteness.

For an exemplarily study, let $d = 2$. From the theory of quadratic equations we know that essentially the sole invariant is the determinant of the corresponding curve-tensor. But for now let us ignore this knowledge and let us check whether we can confirm this from a diagrammatical point of view. By Theorem 4.14 we only need two kinds of nodes in our diagrams. We therefore replace our curve-tensor-node by a small circle:

$$\rightarrow \textcircled{\widehat{G}} \leftarrow = \rightarrow \circ \leftarrow .$$

The smallest possible non-trivial diagram must contain $k = 3$ \circ -nodes. In fact

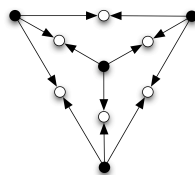
$$\begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \swarrow \quad \searrow \\ \circ \quad \circ \end{array} = \begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \bullet \\ \swarrow \quad \searrow \\ \circ \quad \circ \end{array} = 6 \cdot \det(\widehat{G}) . \tag{4.20}$$

This is also the only non-trivial diagram containing three curve-tensors. One interesting building block, taken from this diagram, is the following (beware of the sign-flip!):

$$\begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \swarrow \quad \searrow \\ \circ \quad \circ \end{array} \rightarrow = -2 \cdot \det(\widehat{G}) \rightarrow \tag{4.21}$$

This means that whenever two curve-tensors share the same two ε -tensors we can find the sub-graph shown above. Replacing this subgraph by $-2 \cdot \det(\widehat{G})$ times the identity-transformation (δ_i^j), we get a diagram with three \circ -nodes less.

Now, by Theorem 4.15, the next smallest (non-trivial) diagram(s) must already contain six copies of the curve tensor. If two of them share the same two ε -tensors, we can apply Equation (4.21) and get a diagram with three nodes. This is either trivial or a multiple of $\det(\widehat{G})$. Thus the whole diagram was either trivial or a multiple of $(\det(\widehat{G}))^2$. Consequently, we can discard such diagrams because we can generate something equivalent by Equation (4.20). The only remaining diagram is of tetrahedral shape with the ε -tensors at the vertices:



Resolving this one is a bit more complicated and we need to revisit Grassmann-Plücker relations:

$$[abc][def] = [abd][cef] - [abe][cdf] + [abf][cde]$$

This formula uses two ε -tensors sharing a curve tensor. This leads to a vanishing last diagram according to Equation (4.19). Now, let us apply this rule to our tetrahedral diagram: The first summand is just the negative of the original diagram because we may swap the blue nodes, changing the connection-order of one ε -tensor. Thus with Equation (4.21) (blue sub-graph):

$$2 \cdot \text{tetrahedral diagram} = \text{blue sub-graph} = -2 \cdot \det(\widehat{G}) \cdot \text{tetrahedral diagram} = 12(\det(\widehat{G}))^2$$

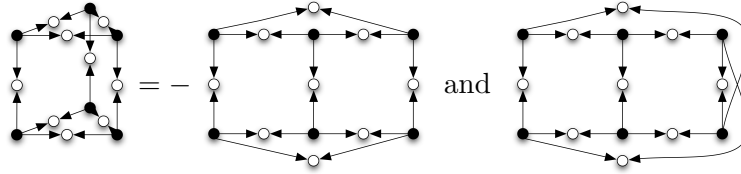
Consequently, all non-trivial diagrams with six curve-tensors are equivalent to $(\det(\widehat{G}))^2$. In essence, we used the Grassmann-Plücker relation above and exploited the fact that the entry-points c and d as well as b and e were connected to an ε -tensor. Thus the diagrammatical rule is

$$2 \cdot \text{diamond diagram} = \text{two diamond diagrams} = 4 \cdot (\det(\widehat{G}))^2$$

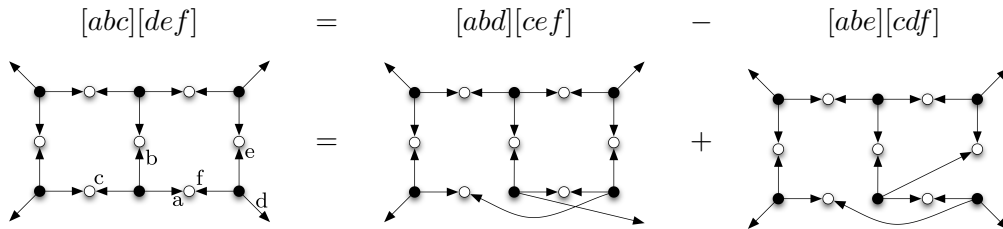
Comparing this rule with Equation (4.21), we can see certain cells, bounded by nodes and arcs. On the left hand-side in Equation (4.21), we have one cell with two ε -nodes and in Equation (4.22) two neighboring cells with three ε -nodes. Having six \circ - and four \bullet -nodes, it is not possible to have a cell with four black ε -nodes without shortcuts and without configurations from Equation (4.21) (try it!).

Analogously the largest cells without shortcuts are bounded in size for all diagrams. For example in diagrams with nine curve-tensors, there is no cell with five or more ε -nodes, which does not have any short-cuts: Just suppose the contrary. Then we have a ring with at least five ε -tensors and equally many curve-tensors. The curve tensors in the ring are already fully connected but the ε -tensors have one dangling arc each. But now we have only four (or less) free curve-tensors left, to which these arcs may connect. Consequently, one of the free curve-tensors must be connected to two ε -tensors in the ring, building a short-cut.

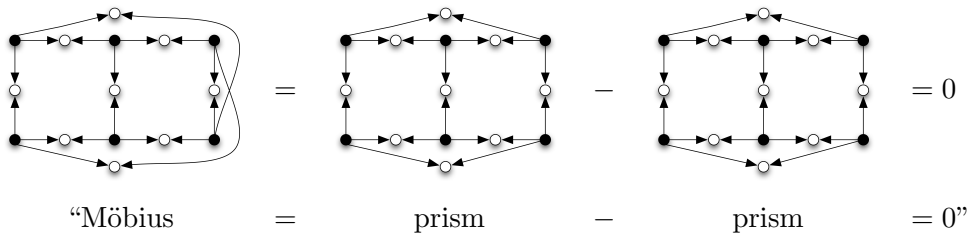
In fact (by enumeration) we get only two possible diagrams with six ε -tensors: a prism and a Möbius strip:



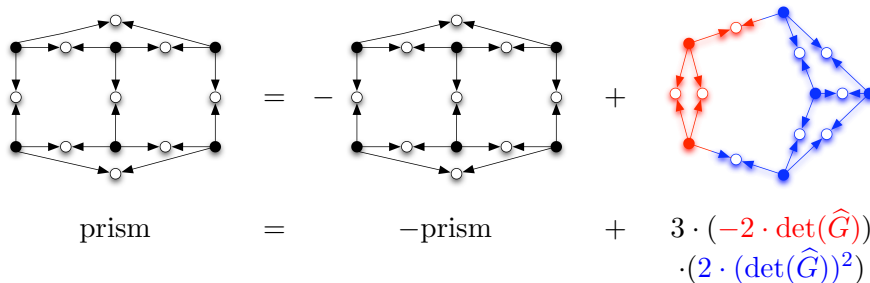
Thus with nine curve-tensors, we must either have a known sub-configuration, reducing the diagram, or there are two neighboring cells of size four. The neighboring cells must either have two or three ε -tensors in common. Applying the Grassmann-Plücker relation, we get (beware of the signs and the vanishing term)



In any of these summands we can see a cell with five ε -nodes. But in case of the prism or Möbius strip, the bottom left node gets connected to one of the rightmost nodes via an intermediate curve-tensor. Thus a careful application of this rule on our Möbius strip shows



and the application on the prism gives us



Consequently, the prism equates to $-6 \cdot (\det(\hat{G}))^3$. Now we are able continue to cut one connection and formulate diagrammatic calculation rules which can be used for higher order invariants. We will omit this here. However, we want to advert to the

vanishing Möbius strip. It is interesting that the configuration exhibits none of the former structures and vanishes anyway. All cells in this Möbius strip are of size four, i.e. contain four ε -tensors. It does not matter where the Grassmann-Plücker relation with the two cells of size four is applied: The result is always “prism – prism” and thus zero. As such, the Möbius strip is the second non-trivially vanishing diagram.

Let us close the topic of conics and switch to cubics. Thus $d = 3$ and Theorem 4.15 impose no restrictions on the degree of the invariant: Any number of curve-tensors is conceivable. However, it is clear that the number of curve-tensors must match the number of ε -tensors in order to get a closed diagram. At first we get the pendant to Equation (4.19), which is of course extendable to curves of arbitrary degree:

$$\rightarrow \widehat{G} \rightarrow = \rightarrow \widehat{G} \rightarrow = 0 . \quad (4.23)$$

Thus the smallest diagram is

$$\begin{array}{c} \circ \\ \uparrow \\ \circ \text{---} 3 \text{---} \circ \\ \downarrow \\ \circ \end{array} = 0 . \quad (4.24)$$

In contrast to conics, this diagram vanishes and in it we can discover symmetries: Reflecting the diagram at an axis through the ε -tensors and one curve-tensor, we obtain the same diagram. On such an axis, we have an odd number of ε -tensors. Finally, we can say that the following sub-configuration leads to a vanishing tensor:

$$\rightarrow \circ \text{---} 2 \text{---} \begin{array}{c} \circ \\ \uparrow \\ \circ \text{---} \circ \\ \downarrow \\ \circ \end{array} \rightarrow = 0 \rightarrow \quad (4.25)$$

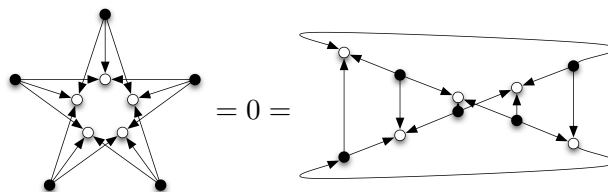
The next smallest configuration must use four nodes of each type. The only possibility with no sub-configuration (4.25) is a cube:

$$\begin{array}{c} \circ \text{---} \circ \\ \uparrow \quad \downarrow \\ \circ \text{---} \circ \\ \uparrow \quad \downarrow \\ \circ \text{---} \circ \\ \uparrow \quad \downarrow \\ \circ \text{---} \circ \end{array} = \sum_{\substack{k+m+o+q=4 \\ l+n+p+r=4}} \alpha_{klmnopqr} \cdot c_{kl}c_{mn}c_{op}c_{qr} .$$

This diagram also showed up in [9], but additionally we can state the following: Expanding this expression we get a sum of 25 terms. The terms themselves are products of four coefficients $c_{kl}c_{mn}c_{op}c_{qr}$, where the sum of the first indices $k + m + o + q$ and the sum of the second indices $l + n + p + r$ equals four. In fact, this diagram is a weighted sum over all such possible expressions. This is our “smallest” non-trivial invariant for cubics.

In the next step we have to examine diagrams with five curve- and five ε -tensors. Discarding diagrams with vanishing sub-configurations, we get as sole possibility a pentagram

or if reconfigured, a Möbius strip



This configuration also vanishes, which can be proven by evaluating this diagram algebraically. Again we have reflectional symmetries with an odd number of ε -tensors on the axes. We have verified this property for a lot of diagrams and arrive at the

Conjecture 4.16 *For cubics (or any other curve of odd degree), a closed tensor-diagram vanishes if it can be arranged such that there is a reflectional symmetry-axis with an odd number of ε -tensors on it.*

Studying the diagrams with six curve- and six ε -tensors, we get an interesting diagram in the shape of a hexagonal cylinder:

$$= \sum_{\substack{k+m+o+q+s+u=6 \\ l+n+p+r+t+v=6}} \alpha_{klmnopqrstuv} \cdot C_{kl}C_{mn}C_{op}C_{qr}C_{st}C_{uv} \cdot$$

In fact, all other diagrams of this size are either vanishing or equivalent to this one. Again if this diagram is evaluated algebraically, one can see that only summands of degree six in the curve-coefficients occur. There are in total 103 of them. Thereby the sum of the first and the sum of the second indices always equate to six. [9] shows how this diagram, in connection with the cube above, is related to the discriminant of a cubic. Dividing the cube-invariant to the power of three by the squared hexagonal cylinder, we get an expression which is independent from the homogeneous scaling of a curve equation: The cube-invariant to the power of three and the squared hexagonal cylinder are polynomials of homogenous degree twelve in the curve coefficients. Dividing these two polynomials by each other, it does not matter whether the curve coefficients belonged to $V(g)$ or to a $V(\lambda g)$ for arbitrary $\lambda \neq 0$. Thus we have an invariant which must be equivalent to the Weierstrass invariant for cubics. Consequently, we already have all basic diagrams for cubics and other non-vanishing diagrams must be reducible.

At this status, proving and calculating a minimal or at least a generating set of invariants for curves of arbitrary degree would involve a detailed observation of invariants. We omit this here and refer to [1], [8] and [9] for additional information that might help finding the desired bases.

5. Symmetry detection

5.1. Rotational symmetry

5.1.1. Definition and restrictions

An interesting feature of a curve is its rotational symmetry. A curve is rotationally symmetric if it can be rotated around a center by an angle ψ such that the rotated curve is identical to the original one. Such an angle ψ is called symmetry angle. In Figure 5.1 we can see a few rotationally symmetric curves.

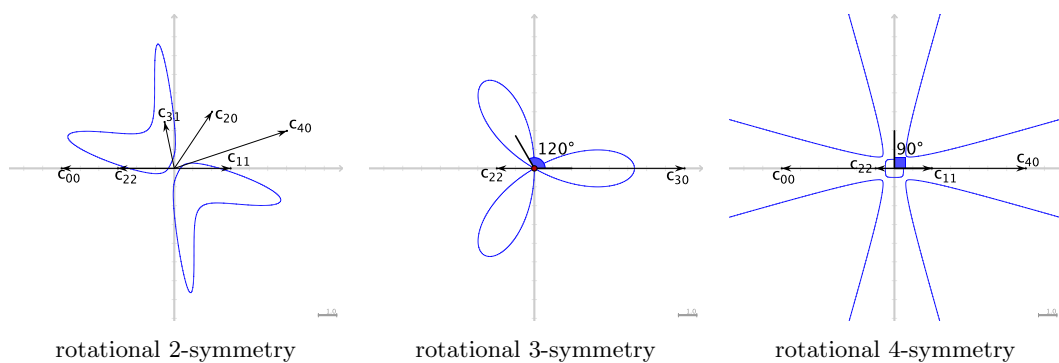


Figure 5.1.: Rotationally symmetric curves

Definition 5.1 Let g be a curve given by the zero set $\mathcal{C}_g = \{z \in \mathbb{C} \mid g(z, \bar{z}) = 0\}$ and let $n \in \mathbb{N} \setminus \{0, 1\}$. g is said to be **n -fold rotationally symmetric** with respect to the origin as center of rotation if \mathcal{C}_g and

$$\mathcal{C}_{g,n} = \tilde{\mathcal{C}}_g = \left\{ z \in \mathbb{C} \mid g\left(e^{\frac{2\pi i}{n}} z, e^{-\frac{2\pi i}{n}} \bar{z}\right) = 0 \right\} \quad (5.1)$$

coincide. If $n \in \mathbb{N} \setminus \{0, 1\}$ is maximal, then g is said to be **rotationally n -symmetric** with respect to the origin as center of rotation.

g is rotationally symmetric with respect to a center $t \in \mathbb{C}$ if the by $-t$ translated copy of g is rotationally symmetric with respect to the origin.

Rotationally invariant curves consist of one or several concentric circles and are **rotationally ∞ -symmetric**.

An angle is called **symmetry angle** if a curve can be rotated around an arbitrary center by this angle such that the rotated copy and the original curve coincide.

The curves from Figure 5.1 are all rotationally symmetric with respect to the origin as center of the rotation. From left to right we have a rotational 2-symmetry, 3-symmetry and 4-symmetry. Naturally, the right curve is also 2-fold rotationally symmetric. The

symmetry angles of the displayed curves are integral multiples of π , $\frac{2\pi}{3}$ and $\frac{\pi}{2}$, respectively.

In general, the symmetry-angles are directly readable from Equation 5.1: The set $\tilde{\mathcal{C}}_g$ in Definition 5.1 is exactly the curve g rotated by $\frac{-2\pi}{n}$. Consequently, if $\tilde{\mathcal{C}}_g = \mathcal{C}_g$, integral multiples of $\frac{2\pi}{n}$ are symmetry-angles. We will now revisit some results from [36] to clarify the connections between symmetry angles and rotationally symmetric curves. The following two lemmata show that the above definition of a symmetry angle is consistent to the one of rotational n -symmetry.

Lemma 5.1 *Let g be a rotationally symmetric algebraic curve with respect to a given center. Then any symmetry angle ψ is a rational multiple of 2π or g is an arrangement of concentric circles.*

PROOF We interpret the origin as a circle with zero radius. If the curve is a circle, then any ψ is a symmetry angle. Assume that ψ is an irrational multiple of 2π and g is not a circle. Then there is a point z on the curve which is not the origin. The images of z under iterated rotations by ψ build a set of points which is dense on a circle c_z through z and centered at the origin. By assumption $g \neq c_z$, and consequently g must intersect c_z in infinitely many points. But g was algebraic and therefore of finite degree d . Thus the number of intersections of g and c_z is at most $2 \cdot d$ by Bezout's theorem (cf. [11]). \square

Lemma 5.2 *If g is rotationally symmetric with respect to a given center and not a circle, then the smallest positive symmetry angle ψ may be written in the form $\psi = \frac{2\pi}{n}$ with $n \in \mathbb{N} \setminus \{0, 1\}$ and the curve is rotationally n -symmetric.*

PROOF Due to the rotational symmetry and by Lemma 5.1, there is a symmetry angle $\psi = \frac{p}{q} \cdot 2\pi$ with $p, q \in \mathbb{N}$. Without loss of generality, we may assume $\gcd(p, q) = 1$. Thus there are $v, w \in \mathbb{Z}$ with $vp + wq = 1 = \gcd(p, q)$. If ψ is a symmetry angle, then for any $l \in \mathbb{Z}$ the angles $v \cdot \psi + l \cdot 2\pi$ have this property, too. Let $l = w$. Then

$$v \cdot \frac{p}{q} \cdot 2\pi + w \cdot \frac{q}{q} \cdot 2\pi = \frac{vp + wq}{q} \cdot 2\pi = \frac{1}{q} \cdot 2\pi < \psi .$$

Having more than one symmetry angle $\psi_i = \frac{p_i}{q_i} \cdot 2\pi$, we know that all angles $\frac{1}{q_i} \cdot 2\pi$ are symmetry angles. n equal to the least common multiple of all the q_i proves this lemma. \square

The n -fold rotational symmetry is always connected to a rotation center. The question now is how many of these centers a single curve can have. The answer is given in

Theorem 5.3 *Let g be an algebraic curve which is not a single line or an arrangement of parallel lines. Then g has at most one center of rotational symmetry.*

PROOF Suppose g is rotationally m -symmetric with respect to the center a and rotationally n -symmetric with respect to the center $b \neq a$. We will show that if these conditions hold, g is invariant under a special translation by $t \neq 0$. Consequently, all iterations of that translations map the curve onto itself. Thus curve-intersecting lines which are parallel to the translation direction intersect the curve in infinitely many points. Now Bezout's Theorem (see [11]) implies that the line itself must be part of the curve. Therefore, the original curve must have been an arrangement of parallel lines contradicting the assumptions.

To retrieve the non-trivial translation, let $\alpha = \frac{2\pi}{m}$ and $\beta = \frac{2\pi}{n}$. The rotation around a by α will be denoted by ρ_a and the one around b by β will be denoted by ρ_b . Let us decompose each of ρ_a and ρ_b into two concurrent hyperplane-reflections with one common reflection: $\rho_a = \sigma_2 \circ \sigma_1$ and $\rho_b = \sigma_3 \circ \sigma_2$. (For more details on geometry with mirrors see [12].) Thereby the fixed line of σ_2 contains a and b (see Figure 5.2). The angles between the fixed lines of the reflections are half the rotation angles and thus at most $\frac{\pi}{2}$. If both $\alpha = \beta = \frac{\pi}{2}$, then a $\rho_2 \circ \rho_1$ is clearly a translation. Figure 5.2 shows this case with half the translation vector. This means that the curve must be invariant under this translation.

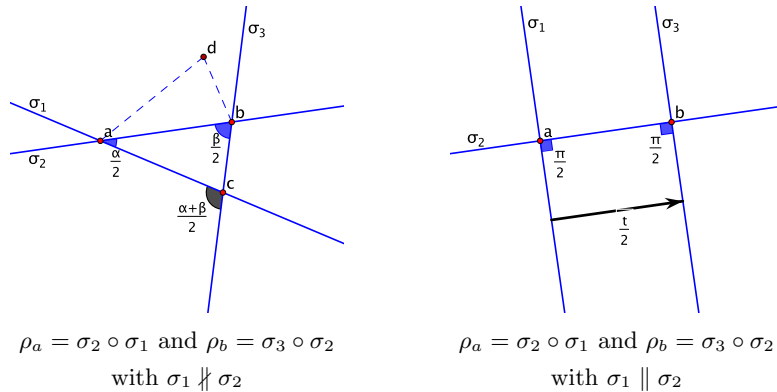


Figure 5.2.: Multiple rotation centers

If at least one of α and β is different from $\frac{\pi}{2}$, then the fixed lines of σ_1 and σ_3 intersect in a point. Let us call it c . We get $\rho_b \circ \rho_a = \sigma_3 \circ \sigma_2 \circ \sigma_2 \circ \sigma_1 = \sigma_3 \circ \sigma_1$, which is a rotation around c by the angle $\alpha + \beta$. Thus g is also rotationally symmetric with respect to c and $\alpha + \beta$ is an associated symmetry angle. We now play the same game analogously with the inverse rotations ρ_a^{-1} and ρ_b^{-1} . Thus we get another rotation symmetry center $d \neq c$ with an associated symmetry angle $-\alpha - \beta$. Consequently $\rho_b^{-1} \circ \rho_a^{-1} \circ \rho_b \circ \rho_a = (\sigma_2 \circ \sigma_3 \circ \sigma_1) \circ (\sigma_2 \circ \sigma_3 \circ \sigma_1)$ is a squared glide-reflection and thus a translation. The fixed lines of σ_1 , σ_2 and σ_3 constitute a non-degenerate triangle and are not concurrent. Thus the translation is non-trivial and g must be invariant under this translation. \square

Let us continue with the detection of a potential rotational symmetry of a curve g . To detect one, we have to locate the center and determine the symmetry angles. For the latter it suffices to calculate a maximal $n \in \mathbb{N} \setminus \{0, 1\}$ such that $\frac{2\pi}{n}$ is a symmetry angle. To keep it simple, we will first focus on determining a potential rotational symmetry with respect to the origin as center of rotation (see the following Section 5.1.2). Rotational symmetry with an arbitrary center will be covered afterwards.

In general, we will pull everything down to Definition 5.1: We check whether there is a n such that $\tilde{\mathcal{C}}_g = \mathcal{C}_g$. Comparing sets is much more difficult than comparing polynomials. Thus we make use of Hilbert's Nullstellensatz or alternatively of Study's Lemma (see Theorem 2.3 and Corollary 2.4). They both grant us the equality of $g(z, \bar{z})$ and $g(e^{\frac{2\pi i}{n}} z, e^{-\frac{2\pi i}{n}} \bar{z})$ up to a scalar factor for arbitrary $z \in \mathbb{C}$, whenever $\tilde{\mathcal{C}}_g = \mathcal{C}_g$. But these theorems impose restrictions: We have to confine ourselves to irreducible curves or at least to square-free curves.

5.1.2. Rotational symmetry with respect to the origin as center

According to Definition 5.1, we have to determine whether there is an n such that $\mathcal{C}_g = \tilde{\mathcal{C}}_g$. Let g be an irreducible curve (or at least a square-free curve). Then by Corollary 2.4 there is a $\lambda \in \mathbb{C}$ with

$$g(z, \bar{z}) = \lambda g\left(e^{\frac{2\pi i}{n}} z, e^{-\frac{2\pi i}{n}} \bar{z}\right) \quad \forall z \in \mathbb{C} . \quad (5.2)$$

The importance of this statement is its validity for all z . This equation holds even for those z where $g(z, \bar{z}) \neq 0$. This means that the equality of the two polynomials in (5.2) implies the equality of all their coefficients: Consequently, there must be a $\lambda \in \mathbb{C}$ such that for all coefficients c_{pq} of g

$$c_{pq} = \lambda \cdot c_{pq} \cdot e^{\frac{2\pi i}{n}(p-q)} . \quad (5.3)$$

The aim is now to determine the largest $n \in \mathbb{N}$ satisfying Equation (5.3) by using the uniqueness of λ . Vanishing coefficients obviously impose no restriction because Equation (5.3) is satisfied for all λ and all n . If $c_{pq} \neq 0$, then $\lambda = e^{\frac{2\pi i}{n}(q-p)}$. In the case of $p = q$ we have $\lambda = 1$.

Let us examine the restrictions on n , imposed by two non-vanishing coefficients $c_{kl} \neq 0$ and $c_{pq} \neq 0$. We get

$$\begin{aligned} \lambda &= \exp\left(\frac{2\pi i}{n}(l-k)\right) = \exp\left(\frac{2\pi i}{n}(q-p)\right) \\ &\Leftrightarrow \frac{2\pi}{n}(l-k) \equiv \frac{2\pi}{n}(q-p) \pmod{2\pi} \\ &\Leftrightarrow (l-k) - (q-p) \in n\mathbb{Z} \\ &\Leftrightarrow n \mid (l-k) - (q-p) . \end{aligned}$$

This means that the differences $(l-k) - (q-p)$ of the rotation speeds $(q-p)$ and $(l-k)$ of any pair of non-vanishing coefficients c_{kl} and c_{pq} determine the rotational symmetry: Let R be the set of all such differences $(l-k) - (q-p)$, then n must divide any element of R . Because of $c_{kl} = \overline{c_{lk}}$ we have $R = \{0\}$ exactly while the only non vanishing coefficients are of the form c_{kk} . Thus $R = \{0\}$ only for circles and unions of circles where n , dividing the elements of R , is arbitrary. Otherwise, if $R \neq \{0\}$, the maximal n is $\gcd(R)$ and g is rotationally $\gcd(R)$ -symmetric. This proves

Theorem 5.4 *Let R be the set of all differences in the rotation speed of all pairs of non-vanishing coefficients for a curve g . Then g is n -fold rotationally symmetric if*

$$n \mid r \quad \forall r \in R .$$

To determine $\gcd(R)$ and thus a maximal n , we can look at our coefficient triangle in (2.17). We highlight a column whenever it contains a non-zero entry (see Figure 5.3). If two neighboring columns are highlighted, no rotational symmetry exists because the difference between two special rotation speeds equals one. If we have a pattern with equally spaced highlighted columns, the distance in between specifies the maximal n . Figure 5.3 shows the pattern on the coefficient triangle for rotational symmetry of the first two curves of Figure 5.1: The non-vanishing coefficients are highlighted. The distance between them specifies the type of rotational symmetry.

5. Symmetry detection

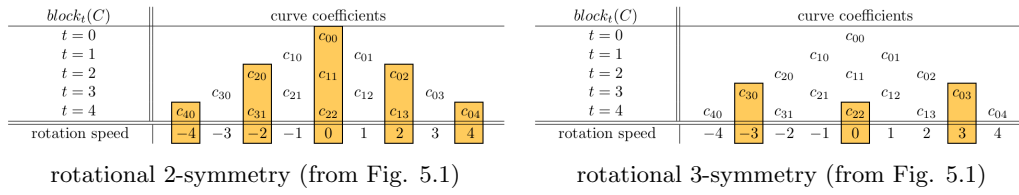


Figure 5.3.: Pattern on the coefficient triangle for rotationally symmetric curves

5.1.3. Rotational symmetry with respect to an arbitrary center

Knowing how to detect rotational symmetry with respect to the origin as center of rotation, we switch to arbitrary centers. It would be nice if we had a translation which moves a center of rotational symmetry onto the origin. In fact, we already have such a translation: It is our translatorial normal form, which focuses on the annihilation of coefficients.

Before we can conclude this, let us first gather some results:

Lemma 5.5 *Let g be a rotationally symmetric curve with respect to the origin as center of rotation. Furthermore, let $c_{pq} \neq 0$ be a non-vanishing coefficient of g . Then $c_{p-1,q} = c_{p,q-1} = 0$ vanish if they are coefficients of g .*

PROOF Let $c_{pq} \neq 0$ be a coefficient of g . Its rotation speed is $q - p$. If $c_{p-1,q}$ or $c_{p,q-1}$ were non-zero, then the corresponding rotation speed would differ by one from $q - p$. As a consequence, g would not be rotationally symmetric (see Theorem 5.4). If $p = q = 0$ then this lemma introduces no coefficient-restrictions because $c_{-1,0}$ and $c_{0,-1}$ are no valid curve coefficients. \square

Speaking in terms of our coefficient triangle, this lemma says: A curve can not be rotationally symmetric if there are non-zero coefficients to the upper left or to the upper right of a non-vanishing coefficient. Of course, we can extend this lemma to all coefficients in the neighboring columns of a non-zero coefficient. The proof is of course analogous:

Lemma 5.6 *Let the preliminaries be the same as in Lemma 5.5. Then any coefficient c_{kl} must vanish if one of the following conditions holds:*

- $\exists \alpha \in \mathbb{Z} : c_{p-1,q} = c_{k+\alpha,l+\alpha}$
- $\exists \alpha \in \mathbb{Z} : c_{p,q-1} = c_{k+\alpha,l+\alpha}$

Now we are already able to prove

Theorem 5.7 *Let g be a curve whose leading coefficients are not of the same absolute value. By Theorem 4.6 there is a translation such that the translated copy of g has a vanishing sub-leading coefficient. Let \tilde{g} be a copy of g , which had been translated by such a transformation. Then g is rotationally n -symmetric with respect to an arbitrary center if and only if \tilde{g} is rotationally n -symmetric with respect to the origin as center of rotation.*

PROOF If g is not rotationally symmetric, then any translated copy of g is also not rotationally symmetric. But if g is in fact rotationally symmetric, then there naturally

exists a translation such that the center of the rotational symmetry is mapped onto the origin. Let the translated curve be g' . It must satisfy the above lemmata.

Let $c_{k+1,l}$ and $c_{k,l+1}$ be an arbitrary pair of succeeding leading coefficients with different absolute values. Then at least one of them is non-zero. Of course, the same is true for one of the coefficients $c'_{k+1,l}$ and $c'_{k,l+1}$ of g' . (Translations do not change leading coefficients.) Thus by Lemma 5.5 the coefficient c'_{kl} must vanish. But by Theorem 4.6 there is exactly one translation annihilating c_{kl} . Consequently, annihilation of an arbitrary sub-leading coefficient sends the center of rotational symmetry onto the origin if g is rotationally symmetric. \square

The clue of the above proof is that the transformation annihilating a coefficient, which had to be zero, was unique. But for curves with leading coefficients of the same absolute value this argument does not hold. There the transformations minimizing a sub-leading coefficient were not unique. For our normal forms we introduced a second translation to overcome this deficiency. This will also save us here: The combination of the two translations was a unique transformation minimizing two coefficients. But before we deal with these special curves, let us examine the behavior of coefficients in the rows of our coefficient triangle (2.17).

We remember from previous chapters what blocks of coefficients had been: There $block_b(c)$ was the collection of all coefficients c_{pq} with index-sum $p + q = b$. We will see that under certain circumstances a whole family of blocks must vanish, i.e. all coefficients of these blocks are zero.

Lemma 5.8 *Let g be of degree d and a rotationally n -symmetric curve with respect to the origin as center of rotation. Further let $n \in 2\mathbb{N}$ and let $c_{pq} \neq 0$ be a leading coefficient of g with $p \neq q$. Then all coefficients of $block_b(c)$ with $b \in \{d - 1, d - 3, d - 5, \dots\} \subset \mathbb{N}$ vanish.*

PROOF Suppose $c_{kl} \neq 0$ is a non-vanishing coefficient of g and $k + l = d - \alpha$ with $\alpha \in 2\mathbb{N} + 1$. Thus $p + q = d = k + l + \alpha$. Then by Theorem 5.4:

$$n \mid r \quad \text{with} \quad r = (q - p) - (l - k) = 2(k - p) + \alpha$$

Consequently, $r \in 2\mathbb{Z} + 1$ and $n \notin 2\mathbb{N}$, which contradicts our assumption. \square

Thus by the above lemma, at least every second line vanishes if we have an even rotational symmetry and a leading coefficient $c_{pq} \neq 0$ with $p \neq q$. Conversely, if every second line in the coefficient triangle vanishes, then the curve is rotationally n -symmetric with even n or it is rotationally ∞ -symmetric. Figure 5.3 shows two patterns: One with even rotational symmetry and every second line vanishing and one with odd rotational symmetry.

Lemma 5.9 *Let g be of degree $d \geq 1$, rotationally symmetric with the origin as center of rotation and a curve where Theorem 4.6 is not applicable. Thus all leading coefficients of g are of the same absolute value. Then all $block_b$ with $b \in \{d - 1, d - 3, d - 5, \dots\} \subset \mathbb{N}$ have only vanishing coefficients.*

PROOF By assumption every leading coefficient is non-vanishing. Thus the curve is rotationally 2-symmetric. Lemma 5.8 proves this lemma. \square

With these ingredients at hand we are finally able to examine rotational symmetry with arbitrary centers for any curve. This theorem is one main result, finalizing our observations on rotational symmetry.

Theorem 5.10 *Let g be an arbitrary curve and let \tilde{g} be the corresponding curve in translatorial normal form, where coefficients were greedily annihilated (see Theorem 4.12). Then g is rotationally symmetric with respect to an arbitrary center or rotation if and only if \tilde{g} is rotationally symmetric with respect to the origin as center of rotation.*

PROOF If g has leading coefficients of different absolute values, Theorem 5.7 proves the assumption.

Now let g have leading coefficients of the same absolute value. If g is not rotationally symmetric, then an arbitrary translation does not change this. The resulting curve can never be rotationally symmetric with respect to the origin.

However, if g is an arrangement of parallel lines, then it has either no center of rotational symmetry or infinitely many. In the latter case these must be contained in a line, which passes through the origin if g is in translatorial normal form.

Let g be rotationally symmetric with respect to an arbitrary center and let g be no arrangement of parallel lines. Naturally, there is a translation moving the center of rotation to the origin. Let this translated curve be g' . By Lemma 5.9 we get vanishing $block_b(c')$ with $b \in \{d-1, d-3, d-5, \dots\} \subset \mathbb{N}$ and with d as the degree of g . We remember from the discussions about our translatorial normal form: There were two translations necessary to send g to g' . The composition of these translations is uniquely determined if and only if g is not a line or an arrangement of parallel lines.

Establishing translatorial normal form, the first sub-leading coefficient of g can be minimized. Due to the existing rotational symmetry of g , this must be an annihilation because $c_{d-1,0} \in block_{d-1}$, which has to vanish. This transformation leaves us with a one parameter family of translations. Possibly a further sub-leading coefficient c_{kl} can be minimized. This is the case when the angle differences of the succeeding leading coefficient pairs are not always the same. Then $c_{kl} \in block_{d-1}$ must also be annihilated. The translation annihilating $c_{d-1,0}$ and c_{kl} is unique and therefore the rotational symmetry center must be at the origin.

Now let the angle differences of the leading coefficient pairs be always the same. Then minimizing $c_{d-1,0}$ is still an annihilation and we have a one parameter family \mathcal{F} of translations left. But \mathcal{F} is exactly the set of all translations preserving the annihilated $c_{d-1,0}$. Even more, any $f \in \mathcal{F}$ does not change any sub-leading coefficient. Consequently, the annihilation of $c_{d-1,0}$ annihilates all sub-leading coefficients along the way. By Theorem 4.9 we know that all curve coefficients change independently from the leading coefficients under any $f \in \mathcal{F}$. Thus we can treat the curve with annihilated $c_{d-1,0}$ as a curve of degree d' . Thereby c_{kl} is its first non-vanishing non-leading coefficient and $\hat{d} = k+l$. We have $d' \notin \{d-1, d-3, d-5, \dots\}$. Otherwise, the g would not be rotationally symmetric. Now we are in the same situation as before: There might be a coefficient in the block below the one of c_{kl} which can be minimized. This minimization would necessarily be an annihilation and the grand total translation would be uniquely determined. If none of these coefficients can be changed, then they already must all be zero due to the annihilation of $c_{d-1,0}$. Thus we effectively have a curve of even lower degree $d'' < d'$ and the game begins anew. The case where no coefficient besides $c_{d-1,0}$ can be changed is the case when g is a line or an arrangement of parallel lines. This case was discussed above. Thus the theorem is proven. \square

5.2. Reflectional symmetry

Detecting reflectional symmetry with respect to a line is not so different from detecting rotational symmetry. We will take the zero set \mathcal{C}_g of a curve g and compare it to a $\tilde{\mathcal{C}}_g$ which denotes the reflected curve with curve describing polynomial \tilde{g} . The aim is to detect whether there is such a non-trivial equality $\mathcal{C}_g = \tilde{\mathcal{C}}_g$ as well as if and where there is a mirror and how many mirrors there are.

5.2.1. Definition and restrictions

Definition 5.2 Let g be a curve given by the zero set $\mathcal{C}_g = \{z \in \mathbb{C} \mid g(z, \bar{z}) = 0\}$. Let $l : p + \mu e^{i\alpha}$ be an arbitrary line ($\mu \in \mathbb{R}$). g is said to be (**reflectionally**) **symmetric** with respect to l if and only if \mathcal{C}_g and

$$\mathcal{C}_{\tilde{g}} = \tilde{\mathcal{C}}_g = \left\{ z \in \mathbb{C} \mid \tilde{g}(z, \bar{z}) = g\left(\overline{(z-p)e^{2\alpha i}} + p, (z-p)e^{-2\alpha i} + \bar{p}\right) = 0 \right\} \quad (5.4)$$

coincide. Thereby $\tilde{\mathcal{C}}_g$ is the reflection of \mathcal{C}_g on the **mirror** l . The line l is also called **symmetry axis**. The angle α of l is called **mirror-angle**.

Figure 5.4 shows reflectionally symmetric curves. We can already see that intersecting symmetry axes imply a rotational symmetry. However, point-symmetric curves will be ignored here, because we may interpret such curves as rotationally 2-symmetric. Thus we restrict ourselves to hyperplane-reflections.

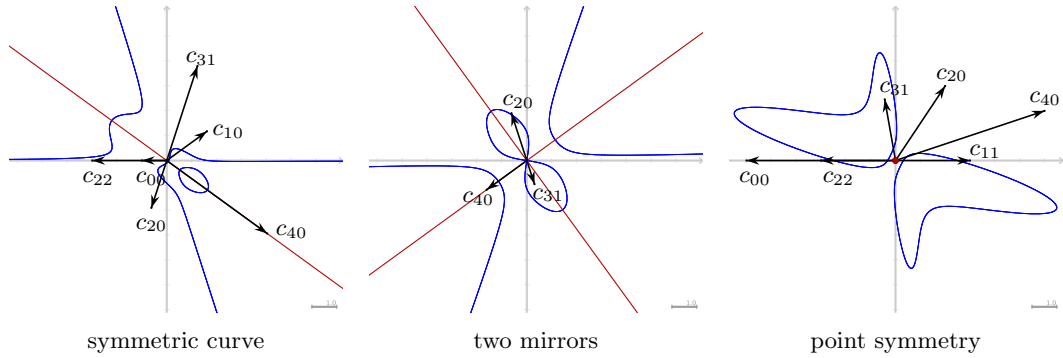


Figure 5.4.: Reflectionally symmetric curves

Let us now have a closer look at the special shape of $\tilde{\mathcal{C}}_g$ in Definition 5.2. If g is symmetric with respect to the real axis, then clearly

$$\mathcal{C}_g = \tilde{\mathcal{C}}_g = \{z \in \mathbb{C} \mid g(\bar{z}, z) = 0\} .$$

Reflections on mirrors l passing through the origin in an angle α to the real axis are easily described. We can perform such a reflection by the following steps: First rotate the plane by $-\alpha$ around the origin, then reflect at the real axis and third rotate back. This means that $z \xrightarrow{\text{rotate}} e^{-i\alpha} z \xrightarrow{\text{reflect}} \overline{e^{-i\alpha} z} = e^{i\alpha} \bar{z} \xrightarrow{\text{rotate}} e^{2i\alpha} \bar{z}$ and thus

$$\mathcal{C}_g = \tilde{\mathcal{C}}_g = \{z \in \mathbb{C} \mid g(e^{2i\alpha} \bar{z}, e^{-2i\alpha} z) = 0\} .$$

If l does not contain the origin but an arbitrary point p , we can prepend and append a corresponding translation by $-p$ and p and get \tilde{C}_g from Definition 5.2.

With respect to reflectional symmetry there are special curves: A circle or concentric circles are of course reflectionally symmetric. All lines passing through the circle center are symmetry axes. Given an arbitrary symmetry angle there is always a corresponding mirror. Thus these curves have infinitely many mirrors, which build a pencil. Lines or arrangement of parallel lines also have infinitely many symmetry axes, namely their perpendiculars. For the converse we can use

Lemma 5.11 *If an algebraic curve g has two different but parallel symmetry axes, then g is a line or an arrangement of parallel lines.*

PROOF The reason for this is that two parallel mirrors l_1 and l_2 imply a translation τ by two times their distance. g must be invariant under this translation. Consequently, if z is a point on g , so is $\tau^k(z)$ for any $k \in \mathbb{Z}$. Let the line passing through all these points be h . Then g must either intersect h in all $\tau^k(z)$ or contain the line. But g is algebraic and therefore intersects h in finitely many points or contains it (see [11] for Bezout's Theorem). \square

Here we have used the equivalence of translations and two reflections with parallel mirrors. But we can also use the equivalence of two intersecting mirrors and rotations. (For more details on the geometry with mirrors see [12].)

Lemma 5.12 *Let g be an algebraic curve with two different intersecting symmetry axes $l : p + \lambda e^{i\alpha}$ and $m : q + \mu e^{i\beta}$ ($\lambda, \mu \in \mathbb{R}$). The point of intersection shall be w . Then g is rotationally symmetric with respect to w as center of rotation.*

PROOF Different intersecting symmetry axes mean $\alpha \neq \beta$. The reflections on l and m imply a rotation ρ of the plane by $2(\alpha - \beta)$ around w . The first case is that $2(\alpha - \beta)$ is a rational multiple of 2π , i.e. $2(\alpha - \beta) = \frac{s}{r} \cdot 2\pi$ with $s, r \in \mathbb{Z}$ and $\gcd(r, s) = 1$. Thus if z coincides with g , so does $ze^{i\frac{s \cdot 2\pi i}{r}}$. Due to $\gcd(r, s) = 1$ our curve g is r -fold rotationally symmetric. In the second case $2(\alpha - \beta)$ is no rational multiple of 2π . Let $g(z) = 0$ and $z \neq w$. Then the set of all iteratively rotated copies of z are contained in a circle c_z . This set has infinitely many elements. g can only intersect this circle in finite many points or contain it due to Bezout's Theorem (see [11]). Therefore g is a circle or a union of concentric circles. \square

Example 5.1 *For example take a curve g which is reflectionally symmetric with respect to the real axis and with respect to the line given by $\mu e^{i\beta}$ ($\mu \in \mathbb{R}$). Then g is invariant under the transformations*

$$\begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} 0 & e^{2i\beta} & 0 \\ e^{-2i\beta} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix}.$$

As a consequence, g is invariant under combination of these reflections, especially under

$$\begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & e^{2i\beta} & 0 \\ e^{-2i\beta} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix} = \begin{pmatrix} e^{-2i\beta} & 0 & 0 \\ 0 & e^{2i\beta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ h \end{pmatrix}.$$

But this transformation is just a rotation of g by -2β around the origin.

Using the above insights and our knowledge from our observations on rotational symmetry, we get

Theorem 5.13 *Let g be an algebraic curve which is not a line and not an arrangement of parallel lines. Then all symmetry axes are concurrent.*

PROOF By Lemma 5.11 g has no parallel symmetry axes. Suppose there are at least three different non-concurrent mirrors. They would have three different points of intersection. By Lemma 5.12 each of these intersections must be a center of a rotational symmetry. But Theorem 5.3 states that we can only have one if g is algebraic. \square

Now we are able to start with the detection of reflectional symmetry. We proceed analogous to the detection of rotational symmetry: We pin everything down to the Definition 5.2. Again we use Corollary 2.4 to get an equality of polynomials rather than an equality of zero-sets. Therefore we restrict ourselves to irreducible or at least square-free curves.

5.2.2. Reflectional symmetry with axes passing through the origin

In accordance with Definition 5.2 and the observations thereafter, we have to determine whether there is a p and α such that $\mathcal{C}_g = \tilde{\mathcal{C}}_g$. In this section we focus on the detection of reflectional symmetry with respect to axes passing through the origin. Thus w.l.o.g. p from Definition 5.2 is the origin $p = 0$. Let g be an irreducible curve. If g is a line, then all symmetry axes are clear: All perpendiculars to g and g itself are mirrors. Let g be of degree $d > 1$ and reflectionally symmetric with respect to the mirror $\mu e^{i\alpha}$ ($\mu \in \mathbb{R}$). Then by Corollary 2.4 there is a $\lambda \in \mathbb{C}$ such that for all $z \in \mathbb{C}$

$$g(z, \bar{z}) = \lambda g(e^{2\alpha i} \bar{z}, e^{-2\alpha i} z) .$$

The irreducibility of g and $d > 1$ secure the existence of a point $q = \mu_q e^{i\alpha}$ on our mirror, which is not a point on our curve g . Thus $g(q, \bar{q}) \neq 0$ and

$$g(q, \bar{q}) = g(\mu_q e^{i\alpha}, \mu_q e^{-i\alpha}) = \lambda g(\mu_q e^{i\alpha}, \mu_q e^{-i\alpha}) = \lambda g(e^{2\alpha i} \bar{q}, e^{-2\alpha i} q)$$

and consequently $\lambda = 1$. This means that $g(z, \bar{z}) = g(e^{2\alpha i} \bar{z}, e^{-2\alpha i} z) = \tilde{g}(z, \bar{z})$ for all $z \in \mathbb{C}$. Thus g and \tilde{g} must have equal polynomial coefficients:

$$c_{pq} = c_{qp} e^{2\alpha i(-p+q)} = \tilde{c}_{pq} . \tag{5.5}$$

Again the rotation speed of c_{pq} , namely $-p + q$, comes in. If c_{pq} vanishes or if $p = q$, then this equation is trivially satisfied.

Now if we want to detect a potential reflectional symmetry of a given curve g , we have to find all angles α such that Equation (5.5) holds for all coefficients c_{pq} of g . As we have seen, vanishing coefficients and coefficients of the form c_{pp} impose no restrictions on α . This fits nicely: If we have a curve with vanishing coefficients c_{pq} when $p \neq q$, then we have (concentric) circle(s). Thus we have infinitely many mirrors and arbitrary α . If a $c_{pq} \neq 0$ with $p \neq q$, we get a constraint on α from Equation (5.5). Let us write the coefficients in exponential form: $c_{pq} = r_{pq} \exp(i\varphi_{pq})$ with positive $r_{pq} \in \mathbb{R}_{>0}$. Thus Equation (5.5) is equivalent to

$$\varphi_{pq} \equiv -\varphi_{pq} + 2\alpha(q - p) \pmod{2\pi} .$$

Consequently,

$$\varphi_{pq} \equiv \alpha(q - p) \pmod{\pi} .$$

This means that there is a $k \in \mathbb{Z}$ such that

$$\varphi_{pq} = \alpha(q - p) + k\pi ,$$

which is equivalent to

$$\alpha = \frac{\varphi_{pq} - k\pi}{q - p} \tag{5.6}$$

for some $k \in \mathbb{Z}$. Thus a single non-vanishing coefficient c_{pq} with $p \neq q$ constrains α to the form of (5.6). If α is a mirror-angle, so is $\alpha + \pi$, because both angles lead to the same mirror. One may live with this ambiguity, or as we do, one can restrict α to $\alpha \in [0, \pi)$. Consequently, α must be contained in the finite set

$$\mathfrak{A}_{pq} = \left\{ \frac{\varphi_{pq} - k\pi}{q - p} \pmod{\pi} \mid k \in \mathbb{Z} \right\} .$$

The cardinality of \mathfrak{A}_{pq} is $|-p + q|$, the absolute value of the rotation speed of the corresponding non-vanishing coefficient c_{pq} .

The above observation already proves

Theorem 5.14 *Let S be the set of all coefficients $c_{pq} \neq 0$ with $p \neq q$ for a given curve g . Let \mathfrak{A} be the set of all mirror-angles $\alpha \in [0, \pi)$ of g with respect to symmetry axes passing through the origin. Then*

$$\mathfrak{A} = \bigcap_{c_{pq} \in S} \mathfrak{A}_{pq} .$$

Graphically one can take an angle φ_{pq} of a non-vanishing coefficient c_{pq} with $p \neq q$, divide it by its rotation speed $-p + q$ and add the $2(-p + q)$ -th complex roots of unity. Figure 5.5 shows \mathfrak{A}_{20} for $\varphi_{20} = -\frac{\pi}{5}$. The mirror-angles in $[\pi, 2\pi)$ are displayed in grey and lead to the same mirrors as the angles in \mathfrak{A}_{20} . Generating such diagrams to each non-vanishing coefficient c_{pq} with $p \neq q$ leads to potentially many graphics. If all the diagrams have one or more lines, or respectively angles, in common, then this line (angle) denotes a reflection-symmetry with the corresponding mirror-angle. In this way we are able to detect reflectional symmetries with respect to axes passing through the origin.

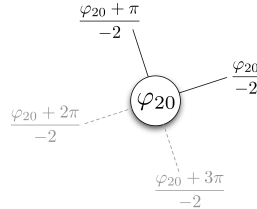


Figure 5.5.: Graphical version of \mathfrak{A}_{20} with additional (grey) angles

Before we switch to reflectional symmetry with respect to arbitrary positioned axes, let us once more focus on the reflection-operation: Reflecting a curve g on a line passing through the origin with angle α to the real axis gives us a reflected copy \tilde{g} of g . Equation

(5.5) gives us the transformation law for the curve coefficients. As with rotations and translations it is instructive to look at this transformation, using matrix-vector notation: We have

$$\left(\begin{array}{c|c|c|c|c|c|c|c|c|c|c} c_{00} & c_{10} & c_{01} & c_{20} & c_{11} & c_{02} & c_{30} & c_{21} & c_{12} & c_{03} & \dots \end{array} \right)^T \mapsto S_\alpha c = \underbrace{\left(\begin{array}{c|c|c|c|c|c|c|c|c|c|c} 1 & & & & & & & & & & \\ & e^{-2i\alpha} & & & & & & & & & \\ & e^{2i\alpha} & & & & & & & & & \\ & & & e^{-4i\alpha} & & & & & & & \\ & & e^{4i\alpha} & 1 & & & & & & & \\ & & & & & & & & & e^{-6i\alpha} & \\ & & & & & e^{2i\alpha} & & & & e^{-2i\alpha} & \\ & & & & e^{6i\alpha} & & & & & & \\ & & & & & & & & & & \ddots \end{array} \right)} = S_\alpha \begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \\ c_{30} \\ c_{21} \\ c_{12} \\ c_{03} \\ \vdots \end{pmatrix} \quad (5.7)$$

5.2.3. Reflectional symmetry with arbitrary symmetry axes

Let us look at reflections from an elementary geometric point of view. Therefore let p be a point on the mirror (symmetry axes) and α be the angle between the mirror and the (positive) real (half-)axis. The reflection operation $S_{(p,\alpha)}$ on this mirror may be decomposed into the following operations as indicated in Section 5.2.1:

1. Translation by $-p$, which sends the axis to a line passing through the origin. The operation on the coefficient-vector may be described by matrix-multiplication with T_{-p} (compare with Equation (4.8)).
2. Rotation by $-\alpha$, which sends the axis to the real coordinate axis. The corresponding transformation on the coefficient-vector is encoded in $R_{-\alpha}$ (compare with Equation (4.2)).
3. Reflection on the real axis, which may be described by S_0 (compare with Equation (5.7)).
4. Rotation by α , representable by R_α and
5. Translation by p , with the corresponding translation-matrix T_p .

Thus

$$S_{(p,\alpha)} = T_p \cdot R_\alpha \cdot S_0 \cdot R_{-\alpha} \cdot T_{-p} .$$

By expansion or by elementary geometric means (see [12]) one can show the identities $R_\alpha S_0 R_{-\alpha} = S_\alpha = R_{2\alpha} S_0 = S_0 R_{-2\alpha}$ and $T_p S_0 = S_0 T_{\bar{p}}$. Additionally, $T_{e^{i\alpha} p} R_\alpha = R_\alpha T_p$. Consequently,

$$S_{(p,\alpha)} = T_p R_{2\alpha} T_{-\bar{p}} S_0 = R_{2\alpha} T_{e^{-2i\alpha} p - \bar{p}} S_0 . \quad (5.8)$$

Thus the decomposition into five steps may be reduced to three steps: the reflection at the real axis, a translation by $e^{-2i\alpha} p - \bar{p}$ and a rotation around the origin by 2α . Equation

(5.8) describes the shape and structure of $S_{(p,\alpha)}$: Basically, it is like the translation matrix $T_{e^{-2i\alpha}p-\bar{p}}$. The difference is that the lines of the translation matrix have to be scaled by $e^{2i\alpha\cdot}$ according to the entries of $R_{2\alpha}$ and the order of the columns in each block is inverted due to the multiplication with S_0 .

Analogously we can think of $S_{(p,\alpha)}$ as $S_{(p,\alpha)} = R_{2\alpha}S_0T_{e^{2i\alpha}\bar{p}-p}$, a translation-matrix with block-wise inverted line-order and scaled lines. To better see this it may be instructive to write the corresponding matrices down. We will omit this here.

Having a reflectionally symmetric curve, we know that for the curve coefficient-vector c the Equation $\tilde{c} = S_{(p,\alpha)}c = c$ must hold. The curve is reflectionally symmetric if and only if the coefficient vector is invariant under multiplication with $S_{(p,\alpha)}$. One of our main results in studying reflectional symmetry is that we have a translation, sending potentially unknown mirrors to lines passing through the origin. This is already clear for curves with two or more different concurrent mirrors: Lemma 5.12 shows that the point of intersection is a center of rotational symmetry and Theorem 5.10 states that this center is the origin if the curve is in translatorial normal form. But in general we have the following theorem:

Theorem 5.15 *A reflectionally symmetric curve in translatorial normal form with greedy coefficient annihilation (see Theorem 4.12) has a mirror passing through the origin.*

PROOF We prove this theorem in three steps: First we look at reflectionally symmetric curves whose mirrors are the real axis. In a second step the mirrors are lines passing through the origin and, lastly, the mirrors are arbitrary.

Let g be a curve which is reflectionally symmetric with respect to the real axis. Then $c_{kl} = c_{lk} \in \mathbb{R}$ and all coefficients are real. The translatorial normal form \tilde{g} of g can be established according to the algorithm presented in Remark 4.3. Any of the described translations use a real $t \in \mathbb{R}$: $\frac{c\bar{a}-\bar{c}b}{r_a^2-r_b^2} \in \mathbb{R}$, $t_c \in \mathbb{R}$, $\frac{e^{2i\xi}}{e^{2i\xi}+e^{2i\vartheta}} \in \mathbb{R}$ (because $e^{2i\vartheta} = \frac{b}{a}$ and $2\xi \in \{0, 2\pi\}$) and $(\frac{c}{m} + \frac{\bar{c}}{\bar{m}}) e^{i\xi} \in \mathbb{R}$ (because $m = a e^{i\xi} + b e^{-i\xi}$ with $\xi \in \{0, \pm\frac{\pi}{2}, \pi\}$ and thus either $m \in \mathbb{R}$ or $\bar{m} = -m$). Consequently, the translatorial normal form \tilde{g} is also reflectionally symmetric with respect to the real axis.

If g is reflectionally symmetric with respect to an arbitrary axis passing through the origin, then so is \tilde{g} : A rotation of \tilde{g} in translatorial normal form does not destroy the normal-form-property because the coefficients are just multiplied with a factor having an absolute value of one. Thus rotations do not change the coefficients' absolute values. The translatorial normal form, however, is minimizing certain absolute values. Now let $R_{-\alpha}$ be the rotation which transforms g into a curve g' which is reflectionally symmetric with respect to the real axis. By the above observations \tilde{g}' , the translatorial normal form of g' , has also this property. Finally, the by R_α rotated \tilde{g} must be the normal form of g , exhibiting the same mirror as g .

Now let g be a curve with arbitrary reflectional symmetry. Let p be a point on the mirror. Then the by $t = -p$ translated copy g' of g is a curve which is reflectionally symmetric with respect to a mirror passing through the origin. Both curves g and g' differ only by a translation and therefore they must have the same translatorial normal form \tilde{g} . But g' has a normal form which is reflectionally symmetric with a mirror passing through the origin. Thus the same is true for g . \square

By this theorem and our preliminary observation we know that a reflectionally symmetric curve in translatorial normal form, which is neither a line nor an arrangement of parallel

lines, has only mirrors containing the origin. Detecting these symmetries was the subject of our last section and thus we know how to detect arbitrary reflectional symmetries.

A nice special relationship for mirror-angles is contained in

Theorem 5.16 *A curve g with leading coefficients of the same absolute value, which is neither a line nor an arrangement of parallel lines, has at most two different symmetry axes. If there are two different mirrors, then these are perpendicular to each other.*

PROOF Let c_{kl} and $c_{k-1,l+1}$ be leading coefficients of g . If g is reflectionally symmetric, the curve-coefficient-vector of g is invariant under multiplication with $S_\alpha T_{e^{2\alpha i} \bar{p}-p}$. Thereby p is a point on the corresponding mirror with mirror-angle α . Thus these leading coefficients imply

$$\tilde{c}_{kl} = c_{kl} = e^{2\alpha(-k+l)i} c_{lk} \quad \text{and} \quad \tilde{c}_{k-1,l+1} = c_{k-1,l+1} = e^{2\alpha(-k+l+2)i} c_{l+1,k-1} .$$

After substituting $c_{kl} = r_{kl} e^{i\varphi_{kl}}$ and $c_{k-1,l+1} = r_{kl} e^{i(\varphi_{kl}+2\vartheta)}$, 2ϑ denotes the angle-difference between our two succeeding leading coefficients. Bearing this in mind, we get the equivalent expressions

$$2\varphi_{kl} \equiv 2\alpha(-k+l) \pmod{2\pi} \quad \text{and} \quad 2\varphi_{kl} + 4\vartheta \equiv 2\alpha(-k+l+2) \pmod{2\pi} .$$

Therefore the angle-difference fixes our possible mirror-angles to

$$\alpha \equiv \vartheta \pmod{\frac{\pi}{2}} .$$

In combination with Theorem 5.13 the assumption is proven. □

This proof gives rise to the following Corollary:

Corollary 5.17 *Let g be neither a line nor an arrangement of parallel lines and let g have a non-vanishing succeeding coefficient-pair in any of its blocks. Then g has at most two different symmetry axes. If there are two different mirrors, then these are perpendicular to each other.*

5.3. Examples

Let us review the preceding sections by looking at two curves of degree four. Multinomial coefficients are denoted by $m_{kl} = \frac{4!}{k! l! (4-k-l)!}$.

- \mathcal{C}_{f_1} : $f_1(x, y) = \sum_{k+l \leq 4} A_{kl}^{(1)} x^k y^l = \sum_{k+l \leq 4} m_{kl} a_{kl}^{(1)} x^k y^l$ with the coefficient vector $a^{(1)} = (a_{00}^{(1)}, a_{10}^{(1)}, \dots, a_{04}^{(1)})$ given by

$$(11906 \mid -1958 \quad -806 \mid 350 \quad 76 \quad 154 \mid -62 \quad -14 \quad -2 \quad -50 \mid 10 \quad 4 \quad -2 \quad 4 \quad 10)$$

- \mathcal{C}_{f_2} : $f_2(x, y) = \sum_{k+l \leq 4} A_{kl}^{(2)} x^k y^l = \sum_{k+l \leq 4} m_{kl} a_{kl}^{(2)} x^k y^l$ with the coefficient vector $a^{(2)} = (a_{00}^{(2)}, a_{10}^{(2)}, \dots, a_{04}^{(2)})$ given by

$$(158 \mid -168 \quad -36 \mid 60 \quad 24 \quad 20 \mid -6 \quad -12 \quad -2 \quad -12 \mid 6 \quad 0 \quad 2 \quad 0 \quad 6)$$

5. Symmetry detection

These curves can be seen in Figure 5.6 together with transformed versions. After complexification we get two complex coefficient vectors $c^{(1)}$ for f_1 and $c^{(2)}$ for f_2 . We describe them with the help of coefficient triangles:

$block_t(c^{(1)})$	curve coefficients								
$t = 0$					11906				
$t = 1$				-979		-979			
				+403i		-403i			
$t = 2$			49 - 38i		126		49 + 38i		
$t = 3$		-7 - i		-8 + 8i		-8 - 8i		-7 + i	
$t = 4$	2		-i		1		i		2
rotation speed	-4	-3	-2	-1	0	1	2	3	4

$block_t(c^{(2)})$	curve coefficients								
$t = 0$					158				
$t = 1$				-84 + 18i		-84 - 18i			
$t = 2$			10 - 12i		20		10 + 12i		
$t = 3$		3i		-1 + 3i		-1 - 3i		-3i	
$t = 4$	0		0		1		0		0
rotation speed	-4	-3	-2	-1	0	1	2	3	4

Calculating the translatorial normal form for f_1 means annihilating the coefficient $c_{30}^{(1)}$ by a translation with $t_1 = -5 - 3i$. The translatorial normal form for f_2 is obtained by a translation of $t_2 = -1 - 3i$, annihilating $c_{21}^{(2)}$. We denote the translated curve \tilde{f}_1 and \tilde{f}_2 . The corresponding coefficient vectors are given by the triangles

$block_t(\tilde{c}^{(1)})$	curve coefficients								
$t = 0$					10				
$t = 1$				0		0			
$t = 2$			1		-2		1		
$t = 3$		0		0		0		0	
$t = 4$	2		-i		1		i		2
rot. speed	-4	-3	-2	-1	0	1	2	3	4

and

$block_t(\tilde{c}^{(2)})$	curve coefficients								
$t = 0$					-10				
$t = 1$				0		0			
$t = 2$			0		0		0		
$t = 3$		3i		0		0		-3i	
$t = 4$	0		0		1		0		0
rot. speed	-4	-3	-2	-1	0	1	2	3	4

Looking at the structure of the non-vanishing coefficients in the corresponding coefficient triangles, we can directly see the rotational symmetry: \tilde{f}_1 and thus f_1 is rotationally 2-symmetric and f_2 is rotationally 3-symmetric.

The rotational normal form of \tilde{f}_1 can be obtained in the following way: $\tilde{c}_{40}^{(1)}$ is already contained in $\mathbb{R}_{>0}$. Rotations by $\psi \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{4}\}$ preserve this property. But $\tilde{c}_{31}^{(1)}$ may be changed. If $\psi \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$, then the coefficient $\tilde{c}_{31}^{(1)}$ of the rotated copy \hat{f}_1 of \tilde{f}_1 changes its sign: $\hat{c}_{31}^{(1)} = \tilde{c}_{31}^{(1)} e^{-2\psi i} = -\tilde{c}_{31}^{(1)}$. If $\psi \in \{0, \pi\}$, then $\hat{c}_{31}^{(1)} = \tilde{c}_{31}^{(1)}$. Consequently $\hat{c}_{31}^{(1)} \in \{\pm i\}$. Writing $\hat{c}_{31}^{(1)}$ in form of $re^{i\varphi}$ with $\varphi \in [0, 2\pi)$, we can see that $\varphi \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$ and $\frac{\pi}{2} < \frac{3\pi}{2}$. This means that we have to rotate the curve by an angle φ , which implies $\varphi = \frac{\pi}{2}$. Consequently $\psi \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$. ψ is not uniquely determined. So we continue with the next relevant non-vanishing coefficient: $\tilde{c}_{20}^{(1)}$. It has the same angular velocity as $c_{31}^{(1)}$ and therefore no new restriction on ψ can be implied. Thus we end up with two possible rotations to transform \tilde{f}_1 into rotational normal form. This confirms again the 2-fold rotational symmetry. Both $\psi \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$ give us the same coefficient vector of \hat{f}_1 :

$$\hat{c}^{(1)} = (10 \mid 0 \ 0 \mid -1 \ -2 \ -1 \mid 0 \ 0 \ 0 \ 0 \mid 2 \ i \ 1 \ -i \ 2) .$$

Similar calculations for \tilde{f}_2 lead to rotations by $\varphi \in \{\frac{\pi}{6}, \frac{5\pi}{6}, \frac{3\pi}{2}\}$, resulting all in a positive and real $c_{30}^{(2)}$. The final coefficient vector for \hat{f}_2 is

$$\hat{c}^{(2)} = (-10 \mid 0 \ 0 \mid 0 \ 0 \ 0 \mid 3 \ 0 \ 0 \ 3 \mid 0 \ 0 \ 1 \ 0 \ 0) .$$

Thus, finally, \hat{f}_1 and \hat{f}_2 are in rotational and translational normal form and hence in normal form with respect to orientation preserving euclidean motions of the plane.

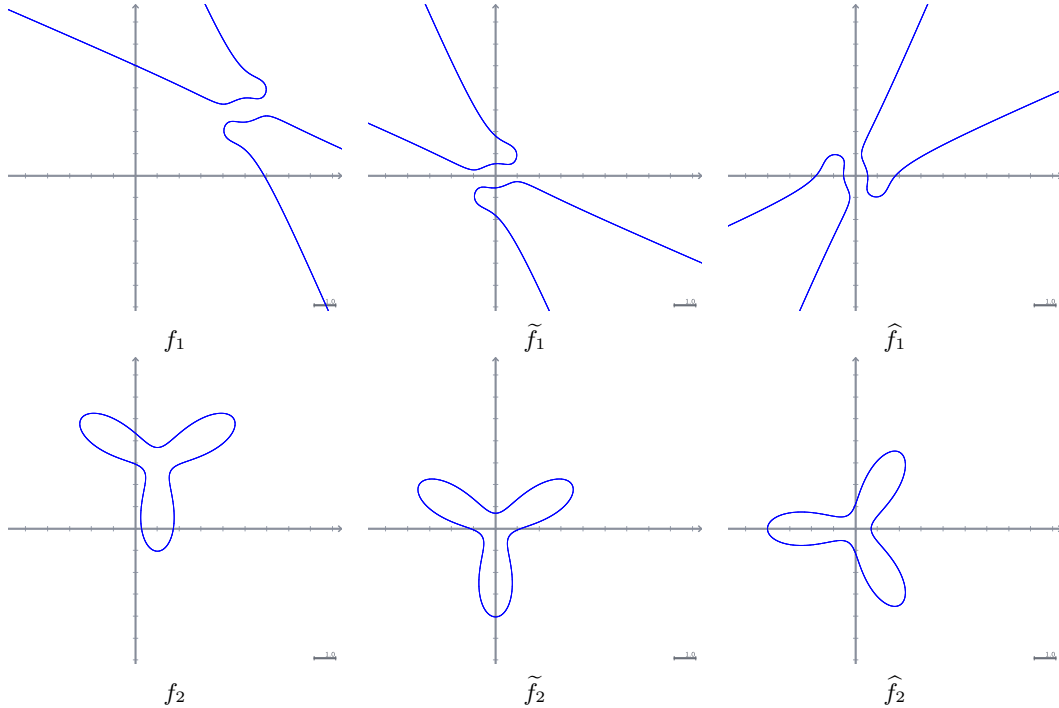


Figure 5.6.: Two curves under observation

Let us continue to examine $g = \hat{f}_2$ and detect its reflectional symmetries. The curve is in translational normal form and thus potentially existing mirrors pass through the origin.

To detect these we need the set S of all non-vanishing coefficients c_{kl} with $k \neq l$. Here $S = \{c_{30}\}$. Thus Theorem 5.14 gives us the set of mirror angles:

$$\mathfrak{A} = \bigcap_{c_{kl} \in S} \mathfrak{A}_{kl} = \mathfrak{A}_{kl} = \left\{ \frac{\varphi_{30} - n\pi}{-3} \pmod{\pi} \mid n \in \mathbb{Z} \right\} = \left\{ 0, \frac{\pi}{3}, \frac{2\pi}{3} \right\},$$

which means that we have three different concurrent mirrors.

A normal form with respect to scalings remains to be established: Choosing to normalize the largest non-vanishing coefficient, we can see that \widehat{f}_2 is already in normal form. However, \widehat{f}_2 must be scaled such that $c'_{40} = 1$. Thus the scaling parameter must be $\lambda = 16$ and the resulting coefficient vector is

$$c'^{(1)} = (10 \mid 0 \ 0 \mid -4 \ -8 \ -4 \mid 0 \ 0 \ 0 \ 0 \mid 32 \ 16i \ 16 \ -16i \ 32) .$$

6. Summary and Outlook

6.1. Summary

Summarizing the preceding chapters, we have seen that the language of curves is complex. Complex numbers in connection with multinomial scaling allow simple translation descriptions and short formulations. The effects of transformations of the plane on curve coefficients demonstrate best why complex numbers are *the* right language for describing curves. When it comes to invariant properties, tensor-diagram language is best suited. Many invariant expressions would involve an enormous amount of summands, which can be reduced to a simple diagram. The clearest examples are given in Section 4.4.2 by the projective invariants of cubics: An expression consisting of 103 terms can be reduced to a diagram of a hexagonal cylinder.

With these ingredients, it becomes evident how rotations, translations, scalings and projective transformations can be exploited. We established normal forms with respect to these transformations. Integrating the knowledge from symmetry detection, arbitrary Euclidean motions are covered. The presented normal forms and curve describing invariant sets may be used for further examination. In a searchable database the desired normal form can be looked up and additional information such as the name of the curve can be retrieved: Given an arbitrary curve, it can be translated such that it is in translatorial and scaling normal form. By using the rotational normal form or rotational invariants and optionally reflections, one may identify a curve in such a database and inform the user:

“You constructed a Limaçon with parameters $r = 0.54$ and $s = 0.72$. A Limaçon is given by the equation $(x^2 + y^2 - 2rx)^2 = s^2(x^2 + y^2)$. Your construction will generically generate a Limaçon.”

The statement concerning the generical construction may be accessible by producing randomized input curves of the same construction.

The main results include also the normal forms, especially the translatorial one. Greedily annihilating (or at least greedily minimizing) coefficients is already a normal form algorithm. This normal form may be used as a basis for further examination: Features such as rotational and reflectional symmetry may be easily extracted because the symmetry centers and reflectional axes contain the origin. Besides these theoretical results, many beautiful cross-references to interesting coefficient structures are presented. Two-dimensional as well as three-dimensional coefficient tables show up during transformations: For example tetrahedral structures in rotations and complexification, generalized Pascal Triangles or the Pascal Pyramid in linear-combinations of tensor-diagrams, angle-diagrams in connection with rotations and reflectional symmetries and, last but not least, patterns in coefficient-triangles in connection with rotational symmetry.

With all these ingredients, features of curves and even curves itself become easily accessible. You only have to speak the right language.

6.2. Outlook

By far this work does not close the topic on feature detection, it rather opens up an even greater range of new lines of action. One of the biggest and probably most challenging one is the practical realization in the framework of dynamic geometry.

As indicated in the introduction, the first task is to extract a curve coefficient vector from a sample point set. We also dealt with this problem in [35], reprinted in Appendix C. There, and also in the field of pattern recognition, a real coefficient vector representing a real algebraic curve was extracted by means of minimizing a kind of distance from a curve to the sample-point set. In this thesis, we have directly seen that a recognition of a complexified coefficient vector is more desirable. Otherwise, we would have to translate into the *right language*. Furthermore, there is the big question of the distances from the sample-point-set to curves. Usually an approximated Euclidean distance to the samples is used (see also [35], [49], [50]). We also apply such an approximation, which is quite good for curves of low degree, since we have highly accurate samples originating from constructions. In most cases an eigenvalue or singular-value analysis is sufficient.

[48] proposes other algorithms on the basis of ridge regression to overcome some numerical deficiencies for higher order curves. The Euclidean distance approximation is, however, still objected. At the latest when curve recognition with respect to projective transformations are of interest, it is not at all clear why the Euclidean distance shall be used for coefficient extraction. In fact, all algorithms known to us approximate the sample-point set very well with virtually all sample-points contained in the recognized curve. But higher order curves often have several branches and the recognition algorithms result in curves with *artifacts* visually destroying features like symmetries. Often an intelligent preprocessing helps to some extent. For us, the future lies not in such approximations but rather in other measures. With the best language not yet known to us, polynomial interpolated measures might offer some solutions.

A further entry-point to recognition is also an examination of what curve can be constructed at all. This means that not always all plane algebraic curves have to be considered. For example, rational curves are sufficient in many cases. Thus the range of possible curve coefficient vectors is restricted, which may be exploited during calculations. Another promising option is to extract coefficients and “almost” features and use these as restrictions for a subsequent new recognition: For example, a recognized coefficient vector belonging to an almost reflectionally symmetric curve might be given. Then we could try to extract a curve-coefficient vector from the same given sample points, which is bound to belong to a reflectionally symmetric curve. Comparing the two results might provide interesting clues and further information can be expected.

As well as specialization, generalization is also of great interest. Here the questions are: How can the proposed algorithms be adapted to transcendental curves? What about complex curves? ... or curves in space? Most preconditions we have used prohibit an extension to transcendental curves. But some of them permit complex algebraic (spacial) curves: complexification, for example. However, it is still not clear whether this is the right language in this case. Additionally, we might extend the class of examined features and turn our interest to singularities, the number of disconnected branches or something else.

To be able to apply the results of this thesis in practice, additional numerical examinations are necessary. Take, for example, the translatorial normal form. There are curves with leading coefficients $\mathbf{a} = c_{d0}$ and $\mathbf{b} = c_{d-1,1}$, where $r_{\mathbf{a}} = r_{\mathbf{b}} + \varepsilon$. In this case the curve

will be translated by $t = \frac{\bar{c}a - \bar{c}b}{r_a^2 - r_b^2}$. As a consequence, the denominator is an expression of order $\mathcal{O}(\varepsilon)$. If $c \neq 0$ and $0 < \varepsilon \ll 1$, then $|t|$ downright explodes. If nothing cancels out, the change in the curve's constant term \tilde{c}_{00} is of order d in t and \bar{t} . Thus the explosion of t is magnified and the numerical information originally contained in c_{00} is lost. For many applications this might be acceptable, but for higher order curves more coefficients may suffer. However, we also gain some information: In such cases we have curves of degree d , which have almost identical absolute values for all leading coefficients. This means that the observed curves are near a curve which has an intersection with the line at infinity of multiplicity d . Following the idea from above, we can restart the recognition process and force a curve with an intersection with the line at infinity of multiplicity d . In the context of curve-recognition in dynamic geometry, we may also use the technique of randomization: Usually a whole bunch of sample point sets may be generated, belonging to different instances of one and the same construction. Thus not only statements about the whole construction are possible. Several *similar* generated curves may also help improving the stability of (numerical) recognition and feature detection because all these curves may be compared.

Comparisons are also necessary when additional information is desired: An indexed and searchable data-base can help, but some mathematical work still has to be done: Due to capacity and handling-reasons, the data-base should only contain classes of curves. For example, conchoids are curves with coefficient vectors

$$(0 \mid 0 \ 0 \mid -2a^2 \ 4b^2 - 2a^2 \ -2a^2 \mid 0 \ -4b \ -4b \ 0 \mid 0 \ 3 \ 4 \ 3 \ 0)$$

to arbitrary $a, b \in \mathbb{R}$. Thus, given a conchoid, we still find corresponding parameters a and b such that the given coefficient vector matches the one above. A direct comparison in general is not possible.

Switching to diagrams, we can see many possible future “playgrounds”. The diagrams are incredibly powerful and we have only scratched the surface. With diagrams whole geometric incidence theorems may be formulated (see [41]). These theorems may be used to build and extend our diagram construction kit. With such ingredients more complicated non-trivial and non-vanishing diagrams can be found and geometrically interpreted. Finding a basis of projective invariants by means of diagrammatical rules is probably one of the most interesting future tasks. But not only projective and Euclidean invariants are of interest: With the diagrams one should - or let us say must - be able to carry out feature detection also in elliptic, hyperbolic or any other Cayley-Klein-Geometry. Even for higher order curves, where the algebraic expressions have an exploding number of summands, tensor diagrams stay handy and interesting.

By all we have seen and by all we can expect to come we can finally say:

This thesis on feature detection
is a glimpse on the power and the beauty of mathematical language.

A. Appendix

A.1. Rotation in real and monomial representation

In this section we give a detailed proof of Theorem 2.6. Thereby we use the abbreviations $c = \cos(\varphi)$ and $s = \sin \varphi$.

$$\begin{aligned}\tilde{f}(x, y) &= \sum_{k+l=0}^d m_{kl} a_{kl} (cx + sy)^k (-sx + cy)^l \\ &= \sum_{r=0}^d \sum_{k=0}^r m_{k(r-k)} a_{k(r-k)} \left(\sum_{u=0}^k \binom{k}{u} c^u x^u s^{k-u} y^{k-u} \right) \\ &\quad \cdot \left(\sum_{v=0}^{r-k} \binom{r-k}{v} (-1)^v s^v x^v c^{r-k-v} y^{r-k-v} \right)\end{aligned}$$

Now using the formula for Cauchy-products of finite sums

$$\left(\sum_{u=0}^k \alpha_u \right) \left(\sum_{v=0}^{r-k} \beta_v \right) = \sum_{n=0}^r \sum_{q=\max(0, n-r+k)}^{\min(k, n)} \alpha_q \beta_{n-q} , \quad (\text{A.1})$$

we continue with

$$\begin{aligned}\tilde{f}(x, y) &= \sum_{r=0}^d \sum_{k=0}^r m_{k(r-k)} a_{k(r-k)} \cdot \\ &\quad \cdot \sum_{n=0}^r \sum_{q=\max(0, n-r+k)}^{\min(k, n)} \binom{k}{q} \binom{r-k}{n-q} (-1)^{n-q} c^{2q-k-n+r} s^{-2q+k+n} x^n y^{r-n} \\ &= \sum_{n=0}^d \sum_{r=n}^d \sum_{k=0}^r \sum_{q=\max(0, n-r+k)}^{\min(k, n)} \dots \\ &= \sum_{n=0}^d \sum_{l=0}^{d-n} \sum_{k=0}^{l+n} \sum_{q=\max(0, k-l)}^{\min(k, n)} m_{k(n+l-k)} a_{k(n+l-k)} \binom{k}{q} \binom{n+l-k}{n-q} (-1)^{n-q} \\ &\quad \cdot c^{2q-k+l} s^{-2q+k+n} x^n y^l\end{aligned}$$

Using Equation (A.1) again with respect to the last two sums and taking care of the terms with mixed indices, we get $k \rightarrow p + q$ and thus

$$= \sum_{n=0}^d \sum_{l=0}^{d-n} x^n y^l \sum_{p=0}^n \sum_{q=0}^l m_{(p+q)((n+l)-(p+q))} a_{(p+q)((n+l)-(p+q))} \cdot \binom{p+q}{p} \binom{(n+l)-(p+q)}{n-p} (-1)^{n-p} c^{l+(p-q)} s^{n-(p-q)}$$

Now we make use of

$$\begin{aligned} & \frac{m_{(p+q)((n+l)-(p+q))}}{m_{nl}} \binom{p+q}{p} \binom{(n+l)-(p+q)}{n-p} \\ &= \frac{d!}{(p+q)!(l+n-p-q)!(d-n-l)!} \frac{n!!(d-n-l)!}{d!} \frac{(p+q)!}{p!q!} \frac{(l+n-p-q)!}{(n-p)!(l-q)!} \\ &= \frac{n!!}{p!(n-p)!q!(l-q)!} = \binom{n}{p} \binom{l}{q} \end{aligned}$$

Thus relabeling $n \rightarrow k$, we get

$$\tilde{f}(x, y) = \sum_{k=0}^d \sum_{l=0}^{d-k} m_{kl} x^k y^l \cdot \tilde{a}_{kl}$$

with

$$\tilde{a}_{kl} = \sum_{p=0}^k \sum_{q=0}^l \binom{k}{p} \binom{l}{q} (-1)^{k-p} \cdot c^{l+(p-q)} \cdot s^{k-(p-q)} \cdot a_{(p+q)((k+l)-(p+q))} ,$$

which proves Theorem 2.6.

A.2. Tensors contained in the linear hull of totally symmetric and skew-symmetric tensors

In Section 3.1.2 we wanted to build up a contra- or covariant tensor of variance d - or rather its coefficient-space - by a linear combination of a totally symmetric and (several) skew-symmetric tensors. We can do so in the following way:

Without loss of generality, we confine ourselves to d -0-tensors. For the sake of simplicity we introduce the multi-indices i and j by $j = (j_1, j_2, \dots, j_d)$ and $i = (i_1, i_2, \dots, i_d)$. Thus $i, j \in \{1, 2, \dots, n\}^d$.

For any $j \in \{1, 2, \dots, n\}^d$ let $b_i^{(j)}$ with $i \in \{1, 2, \dots, n\}$ be defined by

$$b_i^{(j)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} . \end{cases}$$

The $b_i^{(j)}$ are d -0-tensors with tensor indices i_1, i_2, \dots, i_d . They clearly build a basis of our coefficient space.

Let us linearly combine $b_i^{(j)}$ for any given j by one totally symmetric and further (partially) skew-symmetric tensors. Therefore let $\Pi_{(j)}$ be the set of all permuted versions

of the multi-index j . If j contains r_1 times the number 1, r_2 times 2, ... and r_n times n , then the cardinality $|\Pi_{(j)}|$ of $\Pi_{(j)}$ is $\frac{d!}{r_1! r_2! \dots r_n!}$.

We use totally symmetric d -0-tensors of the form

$$s_i^{(j)} = \begin{cases} 1 & \text{if } i \in \Pi_{(j)} \\ 0 & \text{otherwise .} \end{cases}$$

Example A.1 Figure 3.6 on page 38 gives an intuition on how the ten cube-shaped tensors with $d = n = 3$ of the above form look like. There are exactly three tensors with $j_1 = j_2 = j_3$ having only one non-vanishing component (compare with the middle image in the top row of that figure). The other tensors are indicated by the subsequent images and each displayed orbit belongs to a different tensor $s_i^{(j)}$.

Let j and k be of such a kind that $j \neq k$ and that there is a permutation $\pi \in S_d$ with $k = \pi(j)$. We define some d -0-tensors by

$$a_i^{(j,k)} = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

With the definition of $s_i^{(j)}$ and $a_i^{(j,k)}$ we get for arbitrary j :

$$b_i^{(j)} = \frac{1}{|\Pi_{(j)}|} \left(s_i^{(j)} + \sum_{k \in \Pi_{(j)} \setminus \{id\}} a_i^{(j,k)} \right) . \quad (\text{A.2})$$

The $a_i^{(j,k)}$ in Equation (A.2) are not necessarily skew-symmetric. However, we can replace those by skew-symmetric ones in the following way: Let π be associated to a non-skew-symmetric $a_i^{(j,k)}$. As permutation, π can be rewritten by a composition of transpositions. Each such transposition can be associated with a skew-symmetric tensor and $a_i^{(j,k)}$ is the sum of these tensors.

Consequently, there is a basis of the coordinate space of a contra- or covariant tensor of variance d consisting only of totally symmetric and (partially) skew-symmetric tensors.

A.3. Basis and dual basis in the layers of the Pascal-Pyramid

In Section 3.3.4 we encountered layers in the Pascal-Pyramid (see Figure 3.8 on page 52). A layer l was interpreted as a matrix $L^{(l)}$ with components $L_{qi}^{(l)} = C_i^{(l-q,q)}$. Thereby $C_i^{(l-q,q)}$ denoted the i -th coefficient in $(a+b)^{l-q}(a-b)^q$ (starting with $i = 0$). The relations between these coefficients were mainly covered in Equations (3.17) to (3.21). Here we will prove Theorem 3.14, stating that $L^{(l)} \cdot L^{(l)} = 2^l \cdot I$ with I as identity-matrix. Let $N^{(l)} = L^{(l)} \cdot L^{(l)}$. Then for $q \neq l$ we have:

$$\begin{aligned} N_{qr}^{(l)} &= \sum_{p=0}^l L_{qp}^{(l)} L_{pr}^{(l)} = \sum_{p=0}^l C_p^{(l-q,q)} C_r^{(l-p,p)} \\ &\stackrel{(3.17)}{=} \underbrace{\sum_{p=0}^{l-1} \left(C_{p-1}^{(l-1-q,q)} + C_p^{(l-1-q,q)} \right) \left(C_{r-1}^{(l-1-p,p)} + C_r^{(l-1-p,p)} \right)}_{=: \zeta} + \underbrace{C_l^{(l-q,q)} C_r^{(0,l)}}_{=: \xi} \end{aligned}$$

$$\begin{aligned}
 \zeta \quad C_{-1}^{(\dots)}=0 &= \sum_{p=1}^{l-1} C_p^{(l-1-q,q)} C_{r-1}^{(l-1-p,p)} + \sum_{p=1}^{l-1} C_p^{(l-1-q,q)} C_r^{(l-1-p,p)} \\
 &+ \sum_{p=0}^{l-1} C_p^{(l-1-q,q)} C_{r-1}^{(l-1-p,p)} + \sum_{p=0}^{l-1} C_p^{(l-1-q,q)} C_r^{(l-1-p,p)} \\
 &= \sum_{p=0}^{l-2} C_p^{(l-1-q,q)} \left(C_{r-1}^{(l-2-p,p+1)} + C_r^{(l-2-p,p+1)} + C_{r-1}^{(l-1-p,p)} + C_r^{(l-1-p,p)} \right) \\
 &+ C_{l-1}^{(l-1-q,q)} C_{r-1}^{(0,l-1)} + C_{l-1}^{(l-1-q,q)} C_r^{(0,l-1)} \\
 \stackrel{(3.21)}{=} &2 \left(\sum_{p=0}^{l-2} C_p^{(l-1-q,q)} C_r^{(l-1-p,p)} \right) + C_{l-1}^{(l-1-q,q)} C_{r-1}^{(0,l-1)} + C_{l-1}^{(l-1-q,q)} C_r^{(0,l-1)} \\
 &\quad \zeta \stackrel{(3.18)}{=} C_{l-1}^{(l-1-q,q)} \left(-C_{r-1}^{(0,l-1)} + C_r^{(0,l-1)} \right)
 \end{aligned}$$

Thus for $q \neq l$ and $r < l$:

$$N_{qr}^{(l)} = \zeta + \xi = 2 \sum_{p=0}^{l-1} C_p^{(l-1-q,q)} C_r^{(l-1-p,p)} = 2N_{qr}^{(l-1)}$$

and for $q \neq l$ and $r = l$ we get $N_{ql}^{(l)} = 0$. Additionally, we have for any q and any r :

$$N_{qr}^{(l)} = \sum_{p=0}^l C_p^{(l-q,q)} C_r^{(l-p,p)} \stackrel{(3.19)}{=} \sum_{p=0}^l C_p^{(q,l-q)} C_{l-r}^{(l-p,p)} \stackrel{(3.20)}{=} N_{l-q,l-r}^{(l)} .$$

Therefore, $N_{lr}^{(l)} = N_{0,l-r}^{(l)}$ and

$$N^{(l)} = 2 \left(\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & N_{00}^{(l-1)} \end{array} \right)$$

With $N^{(0)} = L^{(0)}L^{(0)} = (1)$ or $N^{(1)} = L^{(1)}L^{(1)} = 2 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ it follows by induction that

$$N^{(l)} = 2^l \cdot I .$$

B. Minimizing the norm of the coefficient vector and translatorial normal forms

In Section 4.2.6 we presented a translatorial normal form. In the field of shape recognition and especially in pose estimation, translations are also of special interest. There, curves are translated such that they are in (translatorial) normal form, in order to be able to compare shape-representing curves (see [46], [47], [52]). In principle, the task of finding a translatorial normal form is common with ours. However, there translatorial “normal forms” are obtained by minimizing not a single sub-leading coefficient, but some equivalent to the sum of the squared absolute values of *all* subleading coefficients (see [46]):

$$\min_{t \in \mathbb{C}} \sum_{k=0}^{d-1} |c_{d-1-k,k} - c_{d-k,k}t - c_{d-1-k,k+1}\bar{t}|^2 . \quad (\text{B.1})$$

In matrix-vector notation this is

$$\min_{t \in \mathbb{C}} \left\| \begin{pmatrix} c_{d-1,0} \\ \vdots \\ c_{0,d-1} \end{pmatrix} - \begin{pmatrix} c_{d,0} & c_{d-1,1} \\ \vdots & \vdots \\ c_{1,d-1} & c_{0,d} \end{pmatrix} \begin{pmatrix} t \\ \bar{t} \end{pmatrix} \right\|_2^2 = \min_{t \in \mathbb{C}} \left\| c_{sub} - M \begin{pmatrix} t \\ \bar{t} \end{pmatrix} \right\|_2^2 \quad (\text{B.2})$$

Thereby c_{sub} is a vector containing all sub-leading coefficients and M is a matrix consisting of the corresponding entries of the sub-triangles to the sub-leading coefficients. Let the set of all $t \in \mathbb{C}$ minimizing this expression be T_{\min} . [46] and [47] claim that $|T_{\min}| = 1$, which is not true for all curves. In fact, translations by $t \in T_{\min}$ do not always use both translatorial degrees of freedom. Consequently, the by a $t \in T_{\min}$ translated curve is not necessarily independent of prior translations. Thus Equation (B.1) can not be the sole criterion for a normal form.

We will show this in three steps. Let $t = x_t + iy_t$ be the decomposition of t into real and imaginary part. At first we interpret

$$q(x_t, y_t) := \left\| c_{sub} - M \begin{pmatrix} x_t + iy_t \\ x_t - iy_t \end{pmatrix} \right\|_2^2 = h \quad (\text{B.3})$$

as a conic in three-space, exhibiting the level-sets T_h to levels $h \in \mathbb{R}$. In a second step, we will show that T_{\min} , interpreted as level-set, may contain more than one translation parameter. And last we will see that there are curves g with $t_1, t_2 \in T_{\min}$ such that the corresponding translated copies $V(g_1)$ and $V(g_2)$ are different.

Lemma B.1 *There are real $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta \in \mathbb{R}$ such that*

$$q(x_t, y_t) = \alpha + 2\beta x_t + 2\gamma y_t + \delta x_t^2 + 2\varepsilon x_t y_t + \zeta y_t^2 = h . \quad (\text{B.4})$$

PROOF Let $q(x_t, y_t) = \|v(x_t, y_t)\|_2^2$, then

$$q(x_t, y_t) = \|\Re(v(x_t, y_t))\|_2^2 + \|\Im(v(x_t, y_t))\|_2^2$$

with corresponding real and imaginary parts of $v(x_t, y_t)$.

Let $c_{sub} = c_R + ic_I$ and $M = M_R + iM_I$ be the decomposition into real and imaginary parts of c_{sub} and M . Expanding the equation above we directly get

$$\begin{array}{c|c} \alpha = c_R^T c_R + c_I^T c_I & \beta = (-c_R^T M_R - c_I^T M_I) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \hline \gamma = (c_R^T M_I - c_I^T M_R) \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \delta = (1 \ 1) (M_R^T M_R + M_I^T M_I) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \hline \varepsilon = (1 \ 1) (M_I^T M_R - M_R^T M_I) \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \zeta = (1 \ -1) (M_R^T M_R + M_I^T M_I) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{array}$$

This proves the assumption. \square

By construction $q(x_t, y_t)$ is strictly non-negative and for variable $h \in \mathbb{R}$ we get a conic $q(x_t, y_t) - h = 0$ in \mathbb{R}^3 . Due to the non-negativity and the linearity of h , this conic may only be either an elliptic paraboloid or a parabolic cylinder. In case of an elliptic paraboloid, the minimum is unique and $|T_{\min}| = 1$. But for a parabolic cylinder T_{\min} contains a whole line.

Lemma B.2 *Let g be a curve of degree d with the following attribute: All leading coefficients have the same absolute value. All succeeding leading coefficient-pairs have the same angle difference φ . Then $q(x_t, y_t) - h = 0$ is a parabolic cylinder and $|T_{\min}| > 1$.*

PROOF W.l.o.g. we may assume that $c_{d0} = 1$. Otherwise, divide $g(z, \bar{z})$ by c_{d0} and take the resulting polynomial, representing the same curve ($V(g) = V(\frac{g}{c_{d0}}$). Thus

$$\underbrace{\begin{pmatrix} 1 & e^{i\varphi} \\ e^{i\varphi} & e^{i2\varphi} \\ \vdots & \vdots \\ e^{i(d-1)\varphi} & e^{id\varphi} \end{pmatrix}}_{= M} = \underbrace{\begin{pmatrix} \cos(0\varphi) & \cos(1\varphi) \\ \cos(1\varphi) & \cos(2\varphi) \\ \vdots & \vdots \\ \cos((d-1)\varphi) & \cos(d\varphi) \end{pmatrix}}_{= M_R} + i \underbrace{\begin{pmatrix} \sin(0\varphi) & \sin(1\varphi) \\ \sin(1\varphi) & \sin(2\varphi) \\ \vdots & \vdots \\ \sin((d-1)\varphi) & \sin(d\varphi) \end{pmatrix}}_{= iM_I}$$

By Lemma B.1

$$q(x_t, y_t) = (x_t \ y_t) \begin{pmatrix} \delta & \varepsilon \\ \varepsilon & \zeta \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} + 2(\beta \ \gamma) \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \alpha$$

with

$$\delta = 2d(1 + \cos(\varphi)), \quad \varepsilon = 2d \sin(\varphi), \quad \zeta = 2d(1 - \cos(\varphi)) \quad \text{and} \quad \det \begin{pmatrix} \delta & \varepsilon \\ \varepsilon & \zeta \end{pmatrix} = 0.$$

Consequently, q is degenerated and therefore a parabolic cylinder. The eigenvectors of the above 2×2 matrix to the eigenvalue zero constitute the elements of T_{\min} . \square

Taking these two lemmas together we arrive at

A drawback of this normal form is that, in general, it does not preserve vanishing sub-leading coefficients. Vanishing coefficients will be needed for symmetry detection (see Chapter 5).

Example B.2 Let g be a curve with coefficient vector

$$c = (-100, 1, 1, 0, 0, 0, 1, 0, 0, 1) = (c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}, c_{30}, c_{21}, c_{12}, c_{03}) .$$

Clearly all sub-leading coefficients vanish and Expression (B.1) is already minimal because $\|c_{sub}\|_2^2 = 0$. Additionally, we have $\|M_0c\|_2^2 = \|c\|_2^2 = 10004$. But

$$\begin{aligned} \|M_t c\|_2^2 &= \|(-100 - 3t - 3\bar{t}, 1, 1, -t, 0, -\bar{t}, 1, 0, 0, 1)\|_2^2 \\ &= \|c\|_2^2 + 1200\Re(t) + 38\Re(t)^2 + 2\Im(t) . \end{aligned}$$

Thus $\|M_t c\|_2^2$ is minimal for a certain $t \in \mathbb{R}_{<0}$. But then $\tilde{c}_{20} \neq 0$.

Remark B.1 [47] proposes to calculate a translation parameter t , which minimizes Expression B.1 in a first step. The by this t translated copy \tilde{g} of the original curve g has the coefficients \tilde{c} . As we have seen in Theorem B.3 and Example B.1, \tilde{g} is not independent of previous translations of the curve in contrast to the claims in [47]. In a second step Expression (B.5) shall be minimized a few times to get a good approximation of the curve \tilde{g}^* with minimal $\|\tilde{c}^*\|_2$. By Theorem B.4 we have a translatorial normal form and therefore an independence of \tilde{c}^* with respect to prior translations. Thus step one is at least theoretically useless. In the above example, we have seen that step two may and in general will, even destroy the property, established by step one: $\|\tilde{c}_{sub}^*\|_2^2 \neq 0$.

C. Coefficient Recognition

This is a joint paper by Prof. Dr. Dr. Jürgen Richter-Gebert and myself. It is a preliminary study to this thesis concerning the recognition of a curve coefficient vector. Given a curve coefficient vector feature detection can take place. The subsequent paper emerged from studies preceding the ADG 2006 - conference and appeared in Lecture Notes in Computer Science/Artificial Intelligence [35]. Some of the “open questione”, stated at the end of this paper, are already answered in this thesis.

Abstract Curve recognition is one of the possible applications this thesis. The application of the discussed normal forms and the calculation of the invariants give the possibility to take any curve and compare the corresponding values with database-entries. The result may thus be to display the curves name to the user, who provides the curve-coefficients. But mostly the curve-coefficients are not directly accessible and must be calculated in beforehand.

In this context [35] was written: The focus lies on the extraction of curve-coefficients, when only a high-precision sample point set is given. We reprint this study here with only minor changes and adaptations. At the end, we include a short outlook on further possibilities to achieve a better preconditioning and a numerically more stable recognition-procedure for curves of higher degree and lower quality samples.

C.1. Introduction

The generation of loci is one of the central applications in today’s computer geometry programs. In abstract terms, a *locus* consists of all locations that a specific point of a geometric configuration can take, while one parameter of the configuration may vary. For instance, the locus of all points that are at a fixed distance from a fixed point is simply a circle. Typically the locus data generated by a geometry program does not consist of a symbolic description of the locus. Rather than that a more or less dense collection of sample points on the locus is generated. In many cases, the user of such a program can identify the underlying curve visually by simply looking at it or by a pre-knowledge of the underlying construction or by a combination of both. However, for several applications (such as automatic assistance, geometric expert systems, etc.), it is highly desirable to recognize these curves automatically. After a geometric construction for a locus is specified, such a recognition procedure can be carried out on two different levels. First, one is interested in which locus is generated for a *concrete* instance (i.e. position of free elements) of the construction. Second, one is interested in invariants of *all* the loci that can be generated by a specific construction. We refer to this set as the *type* of the locus. Very often in the literature *types of loci* (which are algebraic curves in our setup) are closely related to *names* of the corresponding curves like *cardioide*, *limaçon*, *lemniscate*, *Watt curve*, *conic*, etc. One of the ultimate goals of a locus recognition algorithm would be an output like:

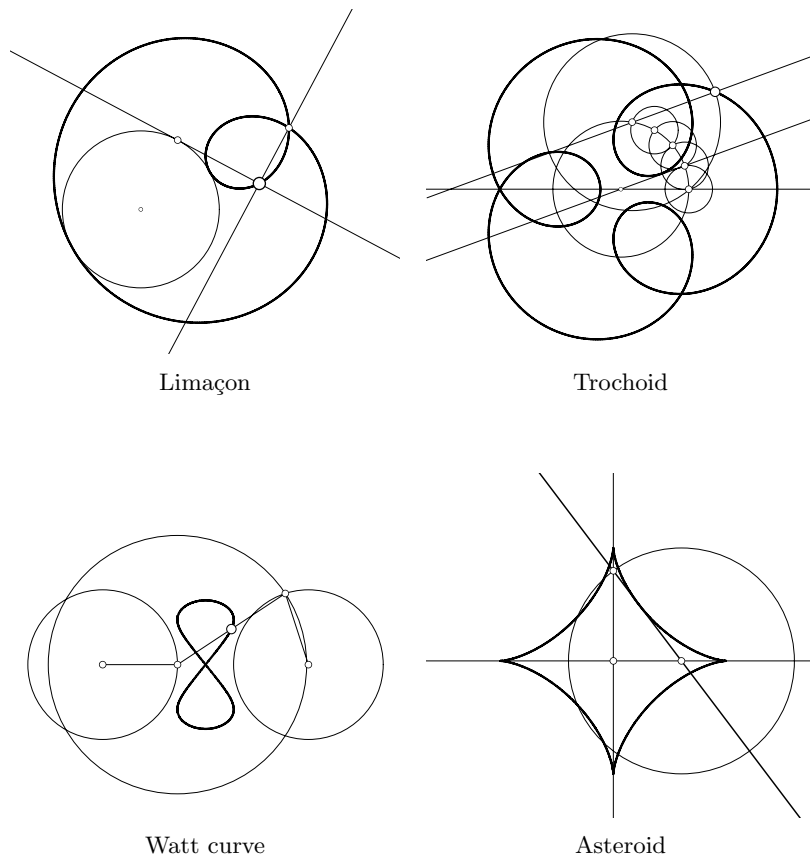


Figure C.1.: Geometric constructions of loci; each curve can be interpreted as zero set of an algebraic equation.

“You constructed a limaçon with parameters $r = 0.54$ and $s = 0.72$. A limaçon is given by the equation $(x^2 + y^2 - 2rx)^2 = s^2(x^2 + y^2)$. Your construction will generically generate a limaçon.”

If one treats the geometric construction as a black box which takes positions of the free elements of a construction as input and produces sample points of the generated locus as output, the above demands imply the necessity of a curve recognition algorithm based on discrete sets of sample points.

In our constructions we only allow primitive operations coming from ruler and compass operations. In this case, a locus turns out to be contained in a zero set of a polynomial equation for any specific instance of a construction. Fixing the free construction elements in a way that only one degree of freedom is left, a dependent point is restricted to an algebraic curve. The locus data is given by a discrete set of sample points \mathcal{P} on this curve, constructed for example via a dynamic geometry program. In general, these programs trace a single point and thus return a set \mathcal{P} contained in a single real branch of the zero set of an algebraic equation. Therefore we presume \mathcal{P} to be of that kind.

In our case, curve recognition for a specific instance of a construction means extracting a single algebraic *equation* of minimal degree from a discrete point set \mathcal{P} and determining

its degree. In a second step we focus on the set \mathcal{B} of curves obtainable by *all* instances of a specific construction – the *type* of the curve. We presume that all coordinates of dependent elements in a construction are analytic functions in the free parameters of the free elements of a construction. Under this assumption, an examination of the degree of the curves in \mathcal{B} and geometric invariant extraction is possible. In this way we can, at least for simple curves, associate a *name* with \mathcal{B} , characterizing the type of the contained curves. In particular, the minimal degree itself turns out to be an invariant property. Recognizing implicit multivariate polynomials has been investigated also by other research groups, most often in different contexts and with different side constraints. A genuine feature of the setup treated in this article is that the sample points usually come with a high arithmetic precision and that it is known in advance that they belong to an algebraic curve with an a priori upper degree bound (usually the real degree will be much lower than this bound). This is in contrast to related research work in computer vision or pattern recognition where one is interested to approximate camera pictures of geometric shapes by algebraic curves in order to recognize these shapes. There the shape data does not a priori belong to an algebraic curve and one is only interested in relatively roughly approximating curves of fixed degree (compare for instance [50],[37],[48]). Another related setup was treated and implemented in the dynamic geometry program Cabri Géomètre [33]. There a first locus recognition algorithm was implemented. Unfortunately, no algorithmic details are available and practical experiments show that the algorithm used there is relatively instable in particular under rigid transformations. (We will deal with this particular issue later in Section 5.2).

In this article we introduce a concept of locus recognition that deals with the construction on two different levels. On both levels the construction itself is treated as a black box that generates sample points on the locus for specific parameters of the free elements of the construction. In the first step we propose an algorithm that takes the sample points of one locus as sole input data and reconstructs the parameters and degree of the underlying algebraic curve for this specific instance. In a second step we allow the free construction elements to vary and thus produce a whole collection of sample point sets $\mathcal{P}_1, \mathcal{P}_2, \dots$ each of which represents locus data of the same type. Randomization techniques are then used to finally extract common features of these loci and determine the type of the locus (with high probability).

C.2. Computationally constructed curves

We use ruler and compass constructions as starting point for our investigations. They provide us with locus data of high arithmetic precision. In principle, it does not matter where discrete sample point sets come from, as long as they are very close to an algebraic curve.

Dynamic geometry programs facilitate ruler and compass constructions in a plane: Elementary construction steps consist of placing free points in the plane, joining two points by a line, constructing circles by midpoint and perimeter and intersecting lines with lines, lines with circles or circles with circles. Furthermore, points can be restricted to already constructed elements like a line or a circle. We will call such points *semi-free* as they have only one degree of freedom. Lines (circles) are zero sets of linear (quadratic) equations and can be computationally represented by the parameters of these equations. Semi-free points can be described by a one-parametric solution manifold of an equation.

All-in-all (for details see [32]), any instance of a geometric construction corresponds to the solution set of a finite set of polynomial equations in the parameters generated by the free and semi-free elements. Allowing only finitely many ruler and compass construction steps, every position of a dependent point can be described by algebraic equations depending on the position of the free points.

While moving a free point in a dynamic geometry program, one can watch the dependent points move consistently with the construction. Fixing all (semi-) free elements except for one semi-free point in a construction, a single dependent point p gets restricted to a zero set of an algebraic equation. Tracing p with a dynamic geometry program under the movement of a single semi-free point (a *mover*), a single branch B of an algebraic curve is revealed. The algebraic curve can be described by a polynomial equation $b(p) = 0$ of minimal degree. B is the set of locations of the traced point p . Unless stated otherwise, all points and curves are assumed to be given in homogeneous coordinates. Under these assumptions a constructed branch B is contained in the zero set $Z(b)$ of a homogeneous polynomial function $b : \mathbb{RP}^2 \rightarrow \mathbb{R}$ of minimal degree d_b :

$$Z(b) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{RP}^2 \mid b(x_1, x_2, x_3) = \sum_{i+j+k=d_b} \beta_{i,j,k} \cdot x_1^i x_2^j x_3^k = 0 \right\} \quad (\text{C.1})$$

From a construction in a dynamic geometry program, we get a discrete set \mathcal{P} containing points lying almost exactly on a branch B of some $Z(b)$. Using algebraic methods, we could, in principle, determine a corresponding curve of very high degree by the following approach: In a first step each elementary construction step is described as a polynomial relation in the coordinates of geometric elements. In a second step elimination techniques (for instance based on Gröbner Bases or Ritt’s algebraic decomposition method) are used to find an implicit representation of the possible position of the locus generating point. This type of approach to curve recognition is purely symbolic but it significantly suffers from combinatorial explosion of the algebraic structures. Even for small constructions one could, in general, not hope for results in reasonable computation time. In contrast, we use the construction as a “black box” and consider only the sample point data of the computationally constructed branch B . Our aim is to derive its describing algebraic equation b of minimal degree. Now our driving questions are:

1. Which plane algebraic curve of minimal degree “reasonably” fits the numerical locus data? (Which b fits \mathcal{P} “reasonably” and what is d_b ?)
2. Regarding a ruler and compass construction, which type of plane algebraic curves are constructed? (What is the degree d_b of almost all curves B in \mathcal{B} and what name characterizes \mathcal{B} ?)

Clearly the problem needs some mathematic modeling which preserves quality criteria. The next section will deal with this issue. Additionally, the first part of our work, i.e. the parameter extraction, can be applied to any curve or even surface recognition problem given sufficiently many points of sufficiently high precision.

C.3. Curve recognition

Given a finite set of data points $\mathcal{P} = \{p_1, p_2, \dots, p_m\}$, e.g. a set of samples in homogeneous coordinates on a locus from a ruler and compass construction. The problem of fitting

an algebraic curve B to the data set \mathcal{P} is usually cast as minimizing the mean square distance

$$\frac{1}{m} \sum_{i=1}^m \text{dist}(p_i, Z(b))^2 \quad (\text{C.2})$$

from the data points to the curve (see [49]). This is a function in the coefficients β of the polynomial b , where $\text{dist}(p, Z(b))$ denotes the Euclidean distance from a point p to the zero set $Z(b)$.

Unfortunately, there is in general no closed form for dist . In principle, dist could be approximated by iterative calculations. However, they turn out to be expensive and very clumsy for the required optimization procedure. Therefore we will use other approximation. Without any numerical noise $b(p) = 0$ for all $p \in Z(b)$ the first choice to replace (C.2) is

$$\frac{1}{m} \sum_{i=1}^m b(p_i)^2. \quad (\text{C.3})$$

We will use this formula as central ansatz for modeling the distance of curve and sample points. This is adequate since our sample points are given with high arithmetic precision. We now will describe how the coefficients of b that minimize (3) can be calculated. For this let $T_1(x), T_2(x), \dots, T_k(x)$ be a basis of the linear space of all homogeneous polynomials of degree d in three variables. If $d = 2$ one could take, for instance, the monomial basis

$$T_1(x) = x_1^2, T_2(x) = x_1x_2, T_3(x) = x_1x_3, T_4(x) = x_2^2, T_5(x) = x_2x_3, T_6(x) = x_3^2.$$

In general we have $k = \frac{1}{2}(d+1)(d+2)$ basis polynomials. Furthermore let

$$\tau_d = (T_1, T_2, \dots, T_k) : \mathbb{RP}^2 \rightarrow \mathbb{R}^k$$

be the function that associates a point $p \in \mathbb{RP}^2$ with its evaluation $\tau_d(p) \in \mathbb{R}^k$ of all the basis functions of degree d .

Assuming we have a discrete set of n points $\mathcal{P} = \{p_1, \dots, p_n\}$ from an underlying construction and assuming we know d_b , the degree of the computationally constructed branch B , we can fix τ_d for polynomials of degree d_b . Furthermore, we can form an $n \times k$ matrix $P := P_{d,\mathcal{P}}$ with row vectors $\tau_d(p_i)$ of τ_d -transformed points $p_i \in \mathcal{P}$. Thus minimizing (C.3) is equivalent to minimizing $\|P\beta\|_2$, with the unknown parameter vector β . So far the minimum can easily be obtained by setting $\beta = (0, \dots, 0)$. This comes from the homogeneity of the problem setup and the fact that β and $\lambda \cdot \beta$ for $\lambda \in \mathbb{R} \setminus \{0\}$ define the same curve. We overcome this problem by forcing additional side-constraints on β and require $\|\beta\|_2 = 1$. In relation to our ansatz (C.3), this side constraint has two advantages: First of all, it is geometrically reasonable, since it ensures a bound on each coefficient of the polynomial. Second, it leads to a nicely solvable minimization problem. All in all curve recognition in our case is stated as

$$\min_{\|\beta\|_2=1} \|P\beta\|_2 \quad (\text{C.4})$$

The above minimization problem leads to a nicely structured Eigenvector-Eigenvalue analysis as the following considerations show. We have

$$\min_{\|\beta\|_2=1} \|P\beta\|_2 = \min_{\|\beta\|_2=1} \sqrt{\beta^T P^T P \beta},$$

which is minimized by an eigenvector of $P^T P$ corresponding to an eigenvalue λ of minimal absolute value. The minimum itself is this very eigenvalue λ . In terms of a singular value decomposition of P , λ is no more than a singular value of P of minimal absolute value and (C.4) is minimized by a corresponding right singular vector of P .

PROOF $P^T P$ is symmetric and features an orthogonal diagonalization employing eigenvectors. Thus $\exists Q \in \mathcal{O}(k) : Q^T P^T P Q = \Lambda$ with a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ exhibiting the eigenvalues λ_i ($1 \leq i \leq k$) of $P^T P$. The j -th column of Q is the eigenvector corresponding to the j -th diagonal entry of Λ . Abbreviating $Q^T \beta$ by γ , we get:

$$\min_{\|\beta\|_2=1} \|P\beta\|_2 = \min_{\|\gamma\|_2=1} \sqrt{\gamma^T \Lambda \gamma}.$$

Let λ_j be an eigenvalue of minimal absolute value $|\lambda_j| = \min\{|\lambda_j| \mid 1 \leq j \leq k\}$ and let $e_j^T = (0, \dots, 0, 1, 0, \dots, 0)$ be the j -th canonical basis vector. Then

$$\min_{\|\gamma\|_2=1} \sqrt{\gamma^T \Lambda \gamma} = \sqrt{e_j^T \Lambda e_j} = \sqrt{\lambda_j}.$$

By undoing the transformation from β to $\gamma = e_j$, we get $\beta^T = (q_{1,j}, q_{2,j}, \dots, q_{k,j})$, the j -th eigenvector of $P^T P$. If the minimal absolute eigenvalue is a single eigenvalue, then we get a unique eigenvector minimizing eigenvector β . If not, then β may be any vector in the span of all eigenvectors to the eigenvalues with an absolute value of λ_j . Since a singular value of P is the root of an eigenvalue of $P^T P$, the right singular vector corresponds to the eigenvectors from above. \square

In practice, with large \mathcal{P} the minimal absolute eigenvalue of $P^T P$ will always be unique. In cases where \mathcal{P} does not specify a single curve and a whole bunch of curves could have \mathcal{P} as sample point set, we can not expect an algorithm to always return the correct curve.

Let us assume having the sample points on a branch B of a yet unknown degree d_b in a no-numerical-noise environment. If we choose an arbitrary testing degree $d \in \mathbb{N} \setminus \{0\}$ and examine the behavior of P and (C.4), we have the following cases:

$d = d_b$: As we investigate in homogeneous polynomials, the length of the parameter vector β is $k_d = \frac{1}{2}(d+1)(d+2)$, but the corresponding curve is determined by $k_d - 1$ points in general position. Per assumption, \mathcal{P} contains this many points in general position and thus $\text{rank}(P) = k_d - 1 = \text{rank}(P^T P)$ and $P^T P$ has a single vanishing eigenvalue $\lambda = 0$. The corresponding eigenvector is the desired parameter vector of the curve.

$d < d_b$: An analysis of the corresponding P yields: P is of maximal rank and $P^T P$ has no vanishing eigenvalue.

$d > d_b$: Let $d = d_b + r$ for some $r \in \mathbb{N}$. An analogous observation shows that $P^T P$ has $k_r = \frac{1}{2}(r+1)(r+2)$ vanishing eigenvalues. This is a reasonable behavior because it shows that the resulting curves are degenerate: One component is the curve of degree d_b fixed by \mathcal{P} and the other component is a more or less random curve of degree r .

This shows that for a given \mathcal{P} , the product $P^T P$ has a vanishing eigenvalue $\lambda = 0$ if and only if the testing degree is greater than or equals to d_b . Thus for constructed curves, we can compute the eigenvalues λ of $P^T P$ of minimal absolute value for testing degrees $d = 1, d = 2, \dots$, successively. This leads to a point where λ is zero. At that point we reached the desired degree.

In practice, where the sample points are disturbed by a small numerical noise, we will not observe vanishing eigenvalues but a significant drop (usually a few orders of magnitude) of the absolute minimal eigenvalues (see section 5). Thus we are able to say:

“You constructed a curve of degree d_b .”

C.4. Type of constructed loci

Let \mathcal{B} be the whole continuum of constructed branches B and let $d_{\mathcal{B}}$ denote the maximum of all degrees d_b of constructed curves B contained in \mathcal{B} . Since we deal with finite constructions, $d_{\mathcal{B}}$ is bounded. \mathcal{P} is a discrete set of points originating from discrete positions of a semi-free mover of a construction. In the last section we saw that $P^T P$ becoming (nearly) singular indicates that the degree of a curve is equal to or below a certain value. Due to the analyticity (omitting the details here (see [32] for an in-depth analysis), almost any curve contained in \mathcal{B} is of degree $d_{\mathcal{B}}$. Therefore we can select a generic parameter set in the parameter space of a specific construction and get a generic curve B . The degree d_b of B equals $d_{\mathcal{B}}$ with probability one. This means, we can detect $d_{\mathcal{B}}$ and prompt the user of a dynamic geometry program:

“You constructed a curve of degree d_b . Your construction will generically generate a curve of degree $d_{\mathcal{B}}$.”

To further specify \mathcal{B} with associated degree $d_{\mathcal{B}}$, we can calculate invariants with respect to either projective, affine or Euclidean transformations. If all curves that belong to a given construction show the same invariants stored in a database, a name like circles, conchoids or limaçon can be associated with \mathcal{B} .

One possibility to achieve this is to examine the orthogonal space \mathcal{B}^{\perp} . This is the space of all $k_{d_{\mathcal{B}}} = \frac{1}{2}(d_{\mathcal{B}} + 2)(d_{\mathcal{B}} + 1)$ dimensional vectors orthogonal to any coefficient vector of any curve in \mathcal{B} . Therefore \mathcal{B}^{\perp} is an invariant of the underlying construction. Generating a matrix V containing at least $k_{d_{\mathcal{B}}}$ parameter vectors to generic curves of \mathcal{B} , \mathcal{B}^{\perp} is the null space of V . It can be determined by an eigenvalue/eigenvector analysis: \mathcal{B}^{\perp} is spanned by the eigenvectors of $V^T V$ to eigenvalues equal to zero.

The remaining task for a graduation of \mathcal{B} would be to correlate the calculated orthogonal space with a database and retrieve a corresponding name. Then we can say something like:

“Your construction will generically generate a circle.”

The idea of how to avoid costly calculations of \mathcal{B}^{\perp} is to simply look up some orthogonal spaces for curve classes of degree $d_{\mathcal{B}}$ in a database and perform a multiplication with V . Getting a zero result, we know \mathcal{B} to be a subset of all curves with the orthogonal space taken.

As a simple example, let $d_{\mathcal{B}} = 2$ and let \mathcal{B} be a set of circles. We assume that all coefficient calculations are performed with the monomial basis for quadrics (see example after equation (C.3)). Looking in a database for orthogonal spaces to curve

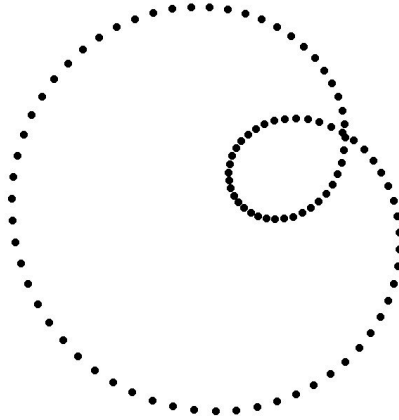


Figure C.2.: Raw data of a degree-four-curve

types of degree two, we will find *circles* with an associated orthogonal space $\mathcal{B}^\perp = \text{span}((1, 0, 0, -1, 0, 0), (0, 1, 0, 0, 0, 0)) =: \text{span}(o_1, o_2)$, because circles are characterized by the following properties

1. the coefficient of x^2 equals the coefficient of y^2 :
 $(1, 0, 0, -1, 0, 0) \cdot \tau_2(x, y, z) = x^2 - y^2 = 0, \forall(x, y, z)$ on a circle.
2. the coefficient of xy equals zero:
 $(0, 1, 0, 0, 0, 0) \cdot \tau_2(x, y, z) = xy = 0, \forall(x, y, z)$ on a circle.

If in fact all members of \mathcal{B} are circles, $V \cdot o_1$ and $V \cdot o_2$ are zero or almost zero due to numerical effects.

To specify a more complex \mathcal{B} , a linear test like matrix multiplication may not suffice. It has to be tested whether potentially non-linear expressions hold for $k_{d_{\mathcal{B}}}$ generic curves out of \mathcal{B} . Additional difficulties arise when \mathcal{B} is tested to be a subset of a parametrized class, like the class of conchoids (A conchoid symmetric to the x-axis and with the singularity in the Euclidean origin can be written as $b(x, y, z) = (x - \sigma)^2(x^2 + y^2) - \rho^2 x^2 = 0$ for some parameters σ, ρ). Here, research is in progress to unify and simplify the tests of type affiliations. Focusing on Euclidean graduation, a promising approach seems to be to calculate an intrinsic rotation invariant center of a curve using all curve parameters as introduced in [47]. Then mapping \mathcal{B} to a class of curves in a database by transformations is reduced to comparing invariants under rotations and scaling.

C.5. Experimental results

Constructing a curve C in a dynamic geometry program and applying a curve recognition algorithm based on minimizing (C.4) yields some curve parameters. This corresponding curve should fit the sample data if the corresponding eigenvalue is significantly small. We present some data of numerical experiments with loci whose sample points have been calculated with the program Cinderella [40].

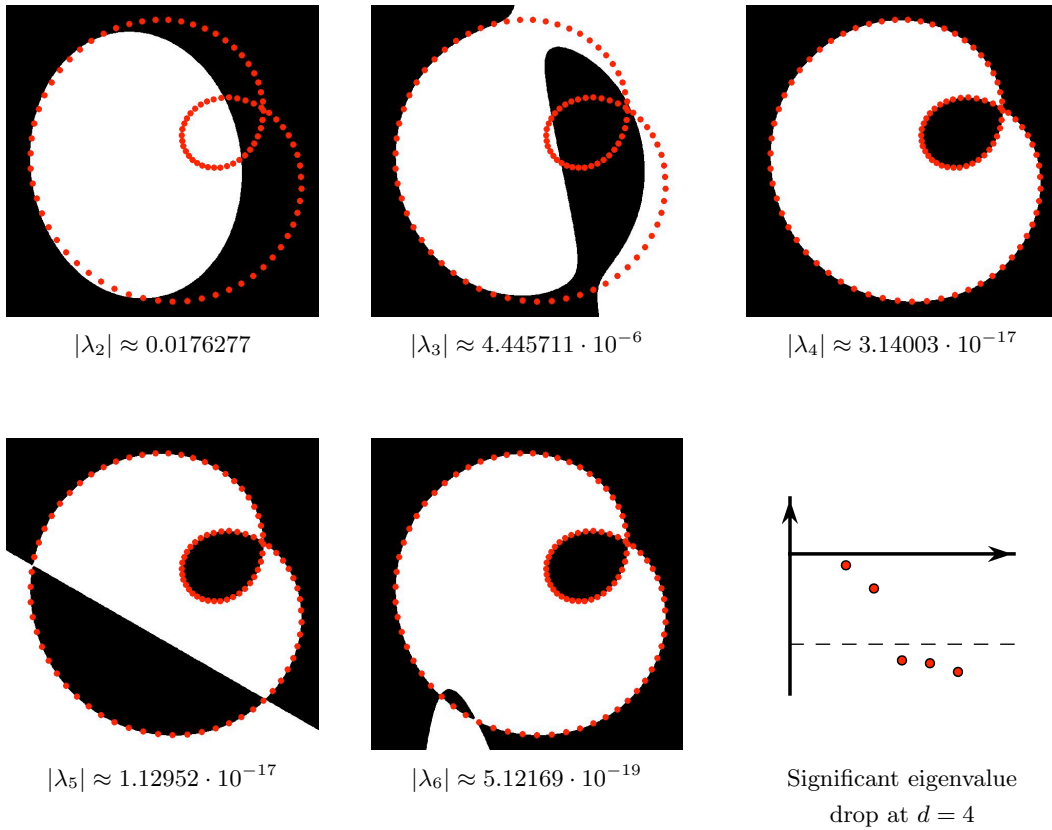


Figure C.3.: Estimated curves for degrees 2, 3, 4, 5 and 6 with logarithmic plot of the minimal eigenvalues against the tested degrees and threshold-line corresponding to $\lambda = 10^{-14}$

C.5.1. Finding the degree

We start with the example of a limaçon corresponding to the first picture in Figure 1. The corresponding sample data \mathcal{P} of the locus corresponds to the cloud of points given in Figure 2. The numerical precision of the data is approximately 14 digits. The Euclidean coordinates of points $p \in \mathcal{P}$ range from -10 to 10 . Figure C.3. shows a sequence of plots for the estimated curves of degree $d = 2, \dots, 6$. The smallest absolute values of eigenvalues λ_d in these five situations are given below the pictures. The sample points are given for reference. One observes a significant drop of the eigenvalues from $d = 3$ to $d = 4$. The picture in the bottom right shows the logarithm of the minimal eigenvalues with a threshold-line at $\ln(\lambda) = \ln(10^{-14})$ marked vertically against the tested degree $d \in \{2, 3, 4, 5, 6\}$. Moreover, for $d = 5$ we obtain (as expected) three absolute eigenvalues below $\lambda = 10^{-14}$ and for $d = 6$ we obtain six absolute eigenvalues below 10^{-14} . The next larger absolute eigenvalues are around 10^{-6} . This suggests that the curve under investigation is of degree $d = 4$. For the degree five approximation the plot unveils an additional component of degree one. And for degree six an additional component of degree two shows up.

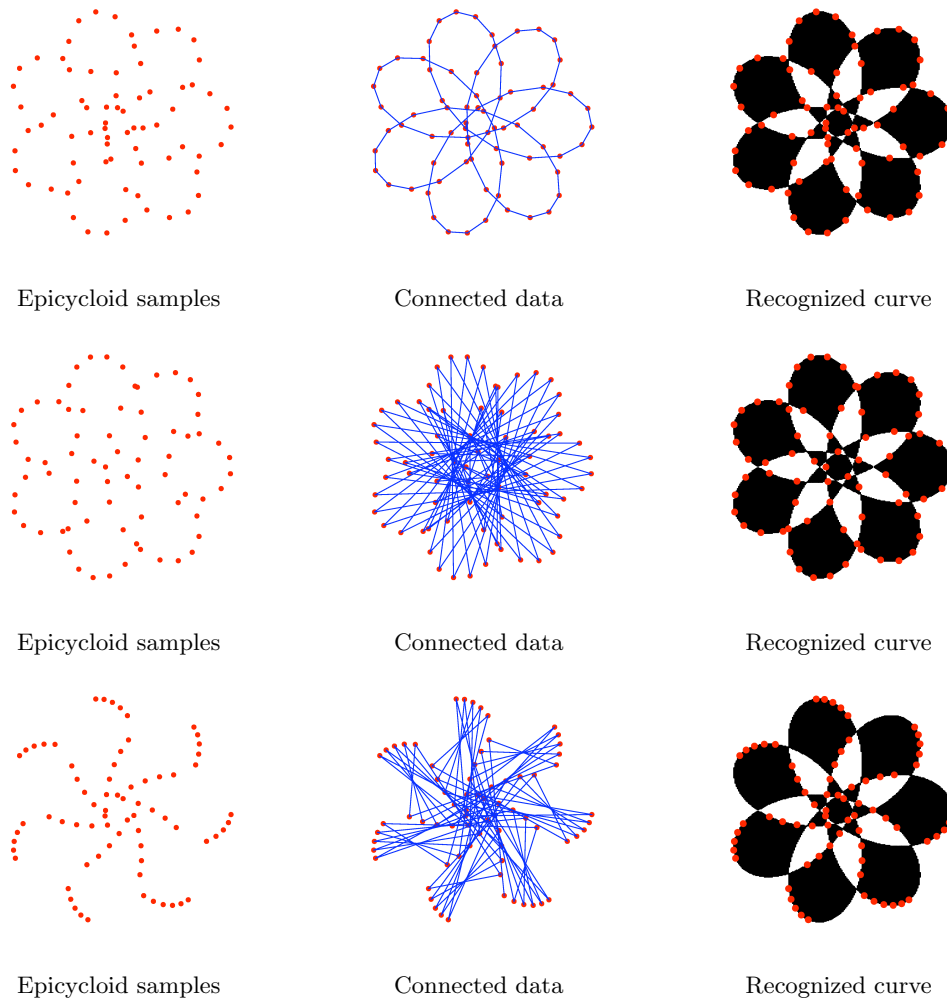


Figure C.4.: Recognition of Epicycloids (see Equation (C.5))

As a second example we examine a parametrizable Epicycloid

$$f(t) = \begin{pmatrix} r \cdot \cos(k \cdot t) - s \cdot \cos(l \cdot t) \\ -r \cdot \sin(k \cdot t) - s \cdot \sin(l \cdot t) \end{pmatrix} . \quad (\text{C.5})$$

If we choose $r = 2.4$, $s = 3$, $k = 2$ and $l = 5$, we get samples depending on the range and sampling rate for t . Three instances of sample point sets are shown in Figure C.4. In the top row $t \in [0; 2\pi)$, in the middle row $t \in [0; 40\pi)$ and in the bottom row $t \in [0; 10\pi)$. In any case 72 equidistant values are used. Thus we have an ordered sequence of sample points given by our parametrization. In dynamic geometry programs, ordered samples can be calculated even if the constructed curve is not rationally parametrizable. (More precisely a curve can be parametrized by rational functions if it comes from a construction that uses exclusively ruler *and* compass may lead to more general but still algebraic curves.) A good idea of what the curve belonging to the data looks like can be provided by connecting the ordered samples linearly. The middle column of Figure C.4 shows these line segments. In the case of $t \in [0; 40\pi)$ in the

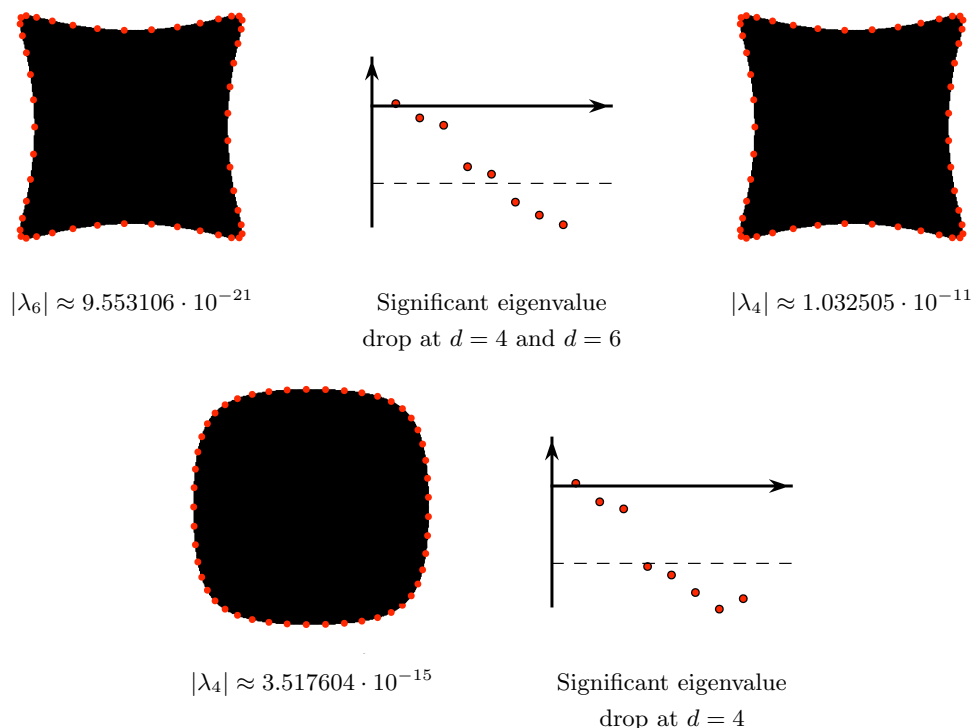


Figure C.5.: Recognition of Epicycloids (see Equation (C.5))

second row, the sampling rate is too low to show the curve itself but the picture gives a good impression of the contour. In cases, where the data is not scattered along the curve but locally concentrated, more intuition is needed when looking at the connected data. In either case, our algorithm detects the correct curve of degree ten (ignoring roundoff errors).

As a further example, we take a Epicycloid of degree six with parameters $r = 1$, $s = 6$, $k = 12$ and $l = 4$. The top left image of Figure C.5 shows it, correctly recognized by our algorithm. The corresponding minimal eigenvalue is sufficiently small, i.e. below the dashed threshold. The logarithmic plot of the minimal eigenvalues against the tested degree in the subsequent middle picture reveals another significant eigenvalue-drop. It suggests that the curve looks quite like a curve of degree four. There is, in fact, such a curve of degree $d = 4$, approximating the sample points very well. It can be seen in the right picture in the top row of Figure C.5. We could have guessed this only by looking at the parameter vector of the correctly recognized curve: All parameters corresponding to the $x_1^i x_2^j x_3^k$ -terms with $i + j \in \{5, 6\}$ are very small or vanish totally (x_3 is the coordinate for homogenization). This curve of degree four can be discarded because of a (second) significant drop in the minimal eigenvalues when switching to a testing degree of six. An acceptable threshold, e.g. 10^{-14} , may be chosen to tell these curves apart. By altering the Epicycloid's parameter r to $r = 0.4$, (C.5) still provides us with a Epicycloid of degree six. In this case, there is only one significant drop in the minimal eigenvalues. λ_4 is already in the range of roundoff errors. Thus our presented algorithm falsely assumes that the data belongs to a curve of degree four instead of six.

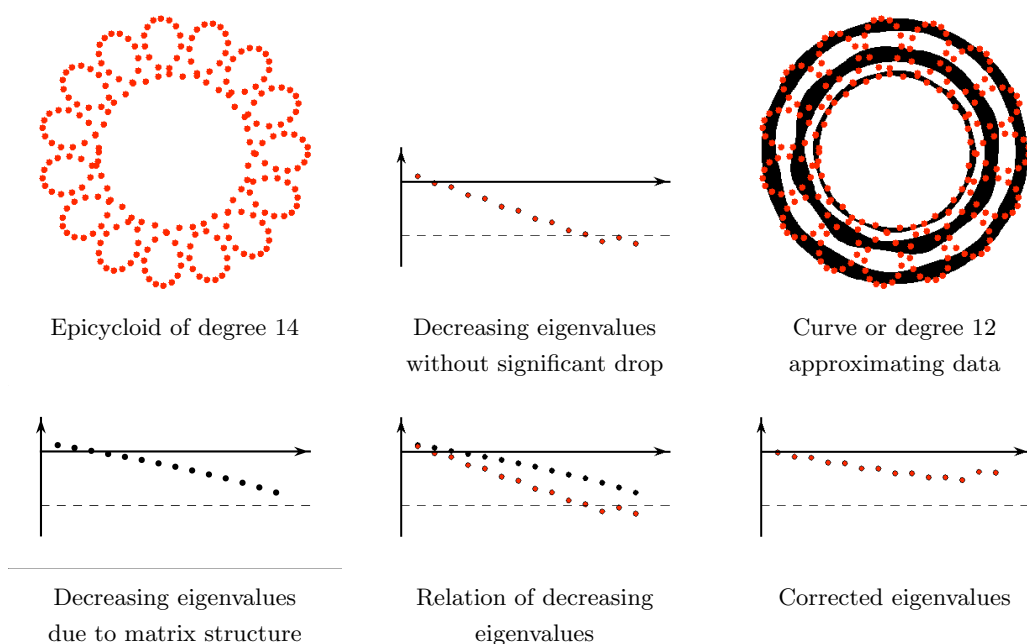


Figure C.6.: Problematic recognition of an Epicycloid of degree 14 and structural decreasing eigenvalues

Our algorithm makes a false degree guess more often for curves of relatively high degree. The may be recognized as curves of lower degree. Usually in most of these cases the approximation by the low degree curve is so good that it visually fits the sample data extremely well. The situation is qualitatively the same as with the Epicycloid in Figure C.5. For the example in the top row, the sample data is already very nicely approximated by an algebraic curve of degree 4.

In general, one can observe that the absolute value of the smaller eigenvalue becomes smaller the higher the degree gets. Let us take for example the Epicycloid with a degree of 28 from Figure C.6. It has the parameters $r = 1.4$, $s = 5$, $k = 14$ and $l = 1$ and it comes up with a decreasing minimal eigenvalue the higher the selected tested degree is. This is not only the case with Epicycloids but with any curve. The minimal eigenvalues stop decreasing once they reach the region of roundoff errors. A reason for this behavior is in the structure of the matrix P . P is used to determine the degree in the minimization (C.4). It is built up row-by-row by $\tau_d(p)$ with p being a sample point. Thus the columns of P consist of values $x^i y^j z^k$, where x , y and z are the coordinates of a sample point. Since the number of free parameters grows quadratically with the degree d , curves of higher degree can simply approximate a set of sample points much better. We did not analyze this effect qualitatively but experimental results exploit a roughly exponential behavior. For this we we took a sample set of 300 random points within the range of -4 and 4 for x and y . With probability one, these points will not be contained in a curve of degree 20 or lower. We can do this a hundred times and thereby calculate the mean minimal eigenvalues for our testing degrees. The resulting diminishing eigenvalues, due to our matrix structure, are shown in a logarithmic plot in the bottom left of Figure C.6. In this logarithmic scale it is almost linear. The next picture (middle of lower

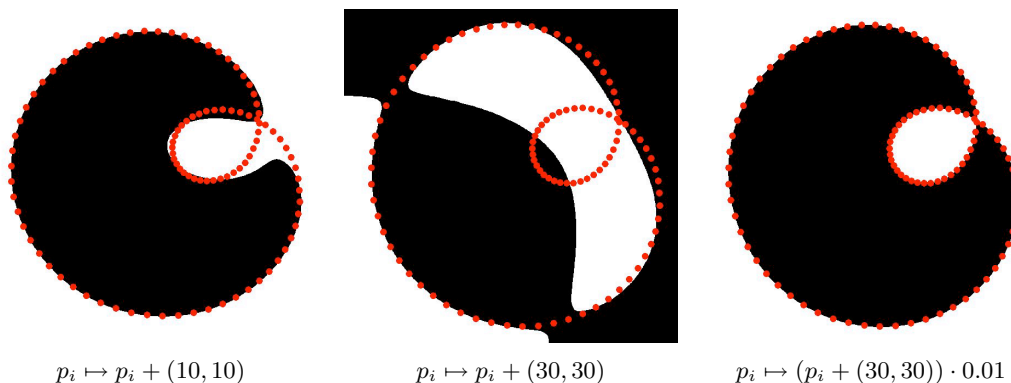


Figure C.7.: Experiments with shifted sample data for shifts $(10, 10)$, $(30, 30)$ and a shift $(30, 30)$ followed by scaling

row) compares this generic effect with the behavior for the Epicycloid. We see that at least part of the systematic eigenvalue fall can be explained by this generic effect. Knowing the structural diminution approximately, we can introduce a correction term in our calculations. The bottom right plot of Figure C.6 reflects the outcome. This means that in case of our Epicycloid we have small eigenvalues for high testing degrees but not a single significant drop in the eigenvalues. The fact that λ_{12} in itself is lower than our chosen threshold is no indicator that a curve and the correct degree was found. If we accepted the curve at the testing degree $d = 12$, we would get the curve shown in the right of Figure C.6. It is obviously not a good approximation to the sample points. In relation to the structural diminution, the minimal eigenvalues of the Epicycloid decrease even more. We interpret this as follows: The higher the testing degree gets, the better the data may be approximated. The increasing values in the corrected plot are due to roundoff errors.

C.5.2. Shifting and preconditioning

Our method as presented here has one significant drawback: It is very sensitive to the location of the sample points. This is due to the following effect. If, for instance, in our example we shift the Euclidean sample points by a vector $(100, 100)$, then the corresponding parameters in $\tau_d(p_i)$ exhibit very high parameter values since the shifts are amplified by the large exponents in the polynomial bases. This results in the fact that our matrix P becomes more and more numerically ill-conditioned when the sample points are far away from the origin. The first two pictures of Figure 4 show how the estimated curve becomes more and more inaccurate the further away all sample points are from the origin.

So far, we do not have a unified method to attack these numerical problems, however we have several reasonable heuristics that work well in practice. As a first step one could a priori investigate the data and translate it so that it is not too far away from the origin. One could also scale the data. After calculating the parameter vector this vector has to be transformed correctly to match the original data points again. We call

this process preconditioning. The third picture of Figure C.7 demonstrates the result of this method. There the data points have first been translated by a vector $(30, 30)$. This would normally result in a very badly conditioned matrix and the estimated curve would be far off the sample points (see middle picture). However, now the whole data set is scaled by a factor of 0.01, which again moves the data points close to the origin. The last picture shows that after this preconditioning the estimated curve nicely matches the sample data again.

A more subtle method to attack numerical instabilities can be achieved by using topological curve invariants. In our example the middle picture could never stem from a locus generated by a dynamic geometry program: The curve obviously breaks up in (at least) two branches and \mathcal{P} contains points on different branches. This is also true for the wrongly recognized Epicycloid of Figure C.6. Using, for example, a Bernstein Basis (with respect to the appropriate bounded domain in each case) for representing and dealing with the involved polynomials could help to improve the conditioning problem, too.

C.5.3. Investigating curve invariants

The parameter vector of the curve is the eigenvector that corresponds to the smallest absolute eigenvalue. A search for curve invariants as proposed in Section 4 would require the generation of many instances of similar curves and the calculation of the orthogonal space. The orthogonal space would hint to specific dependencies on the curve's parameters. However, very often such dependencies also are indicated if one investigates only in one such parameter vector. In our example of Figure 2 after multiplying by a suitable factor, the eigenvector of the degree four approximation takes the following form (four digits after the decimal point are shown):

$$\beta = (1.0, -9.4034 \cdot 10^{-8}, 2.0, -8.8127 \cdot 10^{-8}, 1.0, -28.08, -0.55999, -28.08, \\ -0.55999, 237.3311, 7.8623, 40.2880, -423.9761, 66.5167, -1396.3979)$$

Looking at these values one may suspect that the curve has the characteristic properties

$$\beta_2 = \beta_4 = 0, \quad 2\beta_1 = \beta_3 = 2\beta_5, \quad \beta_6 = \beta_8, \quad \beta_7 = \beta_9.$$

In fact, comparing these conjectures with the coefficients of a generic limaçon generated by a computer algebra system shows that the above relations indeed hold for the general case. Using techniques like the PSLQ algorithm [19] that is able to “guess” integral relations between real numbers, one could also use a single eigenvector to derive many more reasonable conjectures about the underlying curve type. Further research in this direction is in progress.

C.6. Conclusions

With a given set of points \mathcal{P} representing an algebraic curve approximately, we can determine the curve's coefficients by computing eigenvalues and eigenvectors only. This article presented the first steps along this road. Still there are many open questions and problems to tackle. The main research problems currently are:

- What are good ways of preconditioning?
- How to deal with curve types that depend on parameters?
- How to derive geometric transformations that map a curve to a kind of standard representation?
- How can one derive a reliable measure for the quality of the result?
- To what extent can randomization techniques be used to speed up the calculations?
- How to use topological curve characteristics to restrict the search space of potential curves?
- Can similar approaches be used within the more special class of rational algebraic functions?

Bibliography

- [1] M. BERTRAM: *Generating Sets of Invariants in Tensor-Diagram-Notation*, Diploma-
Thesis at Technical University of Munich, tbp
- [2] R. BISCHOF *Geometry of Qubits*, Diploma-Thesis in Computer-Science, ETH Zurich,
2003
- [3] J.F. BLINN: *How many different cubic curves are there?*
Computer Graphics and Applications, IEEE
Part 1: Volume 9, Issue 3 (May 1989) 78-83
Cubic curve update: Volume 9, Issue 6 (Nov. 1989) 70-72
- [4] J.F. BLINN: *How many rational parametric cubic curves are there?*
Computer Graphics and Applications, IEEE
Part 1: Inflection points, Volume 19, Issue 4 (July-Aug. 1999) 84-87
Part 2: The “same” game, Volume 19, Issue 6 (Nov.-Dec. 1999) 88-92
Part 3: The catalog, Volume 20, Issue 2 (March-April 2000) 85-88
- [5] J.F. BLINN: *How to Solve a Cubic Equation*
Computer Graphics and Applications, IEEE
Part 1: The shape of the discriminant, Volume 26, Issue 3 (May-June 2006) 84-93
Part 2: The 11 Case, Volume 26, Issue 4 (July-Aug. 2006) 90-100
*Part 3: General Depression and a New Covariant, Volume 26, Issue 6 (Nov.-Dec.
2006) 92-102*
Part 4: The 111 Case, Volume 27, Issue 1 (Jan.-Feb. 2007) 100-103
Part 5: Back to Numerics, Volume 27, Issue 3 (May-June 2007) 78-89
- [6] J.F. BLINN: *How to solve a quadratic equation?*
Computer Graphics and Applications, IEEE
Part 1: Volume 25, Issue 6 (Nov.-Dec. 2005) 76-79
Part 2: Volume 26, Issue 2 (March-April 2006) 82-87
- [7] J.F. BLINN: *Lines in space*
Computer Graphics and Applications, IEEE
Part 1: The 4D cross product, Volume 23, Issue 2 (March-April 2003) 84-91
Part 2: The line formulation, Volume 23, Issue 3 (May-June 2003) 72-79
Part 3: The two matrices, Volume 23, Issue 4 (July-Aug. 2003) 96-101
Part 4: Back to the diagrams, Volume 23, Issue 5 (Sept.-Oct. 2003) 84-93
Part 5: A tale of two lines, Volume 23, Issue 6 (Nov.-Dec. 2003) 84-97
*Part 6: Our Friend the Hyperbolic Paraboloid, Volume 24, Issue 3 (May-Jun 2004)
92-100*
Part 7: The Algebra of Tinkertoys, Volume 24, Issue 4 (July-Aug. 2004) 96-102
Part 8: Line(s) through four lines, Volume 24, Issue 5 (Sept.-Oct. 2004) 100-106

- [8] J.F. BLINN: *Polynomial discriminants*,
Computer Graphics and Applications, IEEE
Part 1: Matrix magic, Volume 20, Issue 6 (Nov.-Dec. 2000) 94-98
Part 2: Tensor diagrams, Volume 21, Issue 1 (Jan-Feb 2001) 86-92
- [9] J.F. BLINN: *Quartic discriminants and tensor invariants*, Computer Graphics and Applications, IEEE Volume 22, Issue 2 (March-April 2002) 86-91
- [10] J.F. BLINN: *Uppers and downers*
Computer Graphics and Applications, IEEE
Part 1: Volume 12, Issue 2 (March 1992) 85-91
Part 2: Volume 12, Issue 3 (May 1992) 80-85
- [11] E. BRIESKORN E. AND H. KNÖRRER: *Plane Algebraic Curves*, Birkhäuser Verlag Basel, 1986
- [12] F. BACHMANN: *Aufbau der Geometrie aus dem Spiegelungsbegriff*, Springer Verlag Berlin Heidelberg New York, 1973
- [13] J. BOCHNAK, M. COSTE AND M-F. ROY: *Géométrie algébrique réelle*, Springer Verlag Berlin Heidelberg, 1987
- [14] A. BUCHHEIM: *On Clifford's Theory of Graphs*, Proc. London Math. Soc., Vol. s1-17 (1885) 80-106
- [15] W. CLIFFORD: *Extract of a Letter to Mr. Sylvester from Prof. Clifford of University College*, Am. J. Math., Vol. 1 (1878) 126-128.
- [16] C.M. CRAMLET: *The Derivation of Algebraic Invariants by Tensor Algebra*, Bull. Amer. Math. Soc. Volume 34, Number 3 (1928), 334-342
- [17] V. DOLOTIN AND A. MOROZOV: *Introduction to Non-Linear Algebra*, World Scientific Publishing (2007)
- [18] A. EINSTEIN: *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik, Vol. 354, Issue 7 (1916), 769-822
- [19] H.R.P. FERGUSON, D.H. BAILEY AND S. ARNO: *Analysis of PSLQ, An Integer Relation Finding Algorithm*, Mathematics of Computation, Vol 68, No. 225 (Jan 1999) 351-369.
- [20] G. FISCHER: *Ebene algebraische Kurven*, Vieweg Verlag Braunschweig / Wiesbaden, 1994
- [21] G. FISCHER: *Lehrbuch der Algebra*, Vieweg Verlag Wiesbaden, 2008
- [22] G. FISCHER: *Lineare Algebra*, Vieweg Verlag Wiesbaden, 2005
- [23] T. FLIESSBACH: *Allgemeine Relativitätstheorie*, Spektrum Akademischer Verlag Heidelberg/Berlin, 2003
- [24] I.M. GELFAND, M.M. KAPRANOV AND A.V. ZELEVINSKY: *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser Verlag Bosten Baden Berlin, 1994

- [25] G.B. GUREVICH *Foundations of the Theory of Algebraic Invariants*, Noordhoff, Groningen (1964)
- [26] A. HELZER, M. BARZOHAR AND D. MALAH: *Stable Fitting of 2D Curves and 3D Surfaces by Implicit Polynomials*, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 No. 10 (Oct 2004) 1283-1294
- [27] A. HURWITZ: *Zur Invariantentheorie*, Mathematische Annalen, Springer Berlin / Heidelberg, Vol. 45 (1894) 381-404
- [28] A.B. KEMPE: *On the Application of Clifford's Graphs to Ordinary Binary Quantics*, Proc. London Math. Soc., Vol. s1-17 (1885) 107-123
- [29] D. KEREN, D. COOPER AND J. SUBRAHMONIA: *Describing complicated Objects by Implicit Polynomials*, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 No. 1 (Jan 1994) 38-52
- [30] U. KORTENKAMP: *Foundations of Dynamic Geometry*, PhD thesis, ETH Zurich, Switzerland (1999)
- [31] U. KORTENKAMP AND J. RICHTER-GEBERT: *Complexity issues in dynamic geometry* In *Festschrift in the honor of Stephen Smale's 70th birthday*, M. Rojas, F. Cucker (eds.), World Scientific, pp 355-404, (2002).
- [32] U. KORTENKAMP AND J. RICHTER-GEBERT: *Grundlagen dynamischer Geometrie, Zeichnung - Figur - Zugfigur* (2001) 123-144
- [33] J.-M. LABORDE AND F. BELLEMAIN: *Cabri-Geometry II*, Texas Instruments, 1993-1998.
- [34] P. LEBMEIR AND J. RICHTER-GEBERT: *Recognition of Computationally Constructed Loci*, Conference proceedings of 31. Süddeutsches Kolloquium über Differentialgeometrie, (2006) 27-37
- [35] P. LEBMEIR AND J. RICHTER-GEBERT: *Recognition of Computationally Constructed Loci*, Lecture Notes in Computer Science/Artificial Intelligence, Vol. 4869 (2007) 52-68
- [36] P. LEBMEIR AND J. RICHTER-GEBERT: *Rotations, Translations and Symmetry Detection for Complexified Curves*, Computer Aided Geometric Design, Vol. 25 (2008) 707-719
- [37] Z. LEI AND T. TASDIZEN AND D.B. COOPER: *PIMs and Invariant Parts for Shape Recognition*, Proceedings of Sixth International Conference on Computer Vision (ICCV'98) (Jan 1998) 827-832.
- [38] P.J. OLVER: *Classical Invariant Theory*, Cambridge University Press, Cambridge New York Melbourne (1999)
- [39] J. RICHTER-GEBERT: *Hands on Projective Geometry*, tbp
- [40] J. RICHTER-GEBERT AND U. KORTENKAMP: *Cinderella - The interactive geometry software*, Springer 1999; see also <http://www.cinderella.de>, link dated Dec. 2008

- [41] J. RICHTER-GEBERT AND P. LEBMEIR: *Diagrams, Tensors and Geometric Reasoning*, DCG, in review
- [42] W. SPOTTISWOODE: *On Clifford's Graphs*, Proc. London Math. Soc., Vol. s1-10 (1878) 204-214
- [43] G.E. STEDMAN: *Diagram techniques in group theory*, Cambridge University Press, New York (1990)
- [44] B. STURMFELS: *Algorithms in Invariant Theory*, Springer-Verlag Wien New York (1993)
- [45] J.J. SYLVESTER: *On an Application of the New Atomic Theory to the Graphical Representation of the Invariants and Covariants of Binary Quantics, with Three Appendices*, Am. J. Math., Vol. 1 (1878) 64-125.
- [46] J.-P. TAREL AND D.B. COOPER: *A New Complex Basis for Implicit Polynomial Curves and its Simple Exploitation for Pose Estimation and Invariant Recognition*, Conference on Computer Vision and Pattern Recognition (CVPR'98) (June 1998) 111-117
- [47] J.-P. TAREL AND D.B. COOPER: *The Complex Representation of Algebraic Curves and its Simple Exploitation for Pose Estimation and Invariant Recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (2000) 663-674.
- [48] T. TASDIZEN, J.-P. TAREL AND D.B.COOPER: *Algebraic Curves That Work Better*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99) (June 1999) 35-41.
- [49] G. TAUBIN: *An improved algorithm for algebraic curve and surface fitting*, Proc. ICCV'93, Berlin, Germany (may 1993) 658-665.
- [50] G. TAUBIN: *Estimation of Planar Curves, Surfaces and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (1991) 1115-1138.
- [51] N.M. THIÉRY: *Computing Minimal Generating Sets of Invariant Rings of Permutation Groups with SAGBI-Gröbner Basis* Discrete Mathematics and Theoretical Computer Science Proceedings AA (DM-CCG) (2001) 315-328
- [52] M. UNEL AND W.A. WOLOVICH *Complex Representations of Algebraic Curves*, International Conference on Image Processing ICIP (Oct. 1998), 272-276
- [53] T.L. WADE: *Euclidean Concomitants of the Triangle* National Mathematics Magazine, vol. 17, no. 8 (May 1943) 335-340
- [54] T.L. WADE: *Euclidean concomitants of the ternary cubic*, Bull. Amer. Math. Soc., vol. 48, no. 8 (1942) 589-601
- [55] R. WEITZENBÖCK: *Invariantentheorie*, Noordhoff, Groningen (1923)
- [56] http://en.wikipedia.org/wiki/Pascal's_triangle *Pascal's triangle*, link dated: Dec. 2008

