

MATHEMATICS AND VISUALIZATION

Helmut Pottmann
Johannes Wallner

Computational Line Geometry

LuraTech

www.luratech.com

 Springer

Mathematics and Visualization

Series Editors

Gerald Farin
Hans-Christian Hege
David Hoffman
Christopher R. Johnson
Konrad Polthier
Martin Rumpf



LuraTech

www.luratech.com

Helmut Pottmann
Johannes Wallner

Computational Line Geometry

With 264 Figures, 17 in Color

LuraTech

www.luratech.com

 Springer

Helmut Pottmann
Geometric Modeling and Industrial Geometry
Vienna University of Technology
Wiedner Hauptstr. 8-10/104
1040 Wien
Austria
pottmann@geometrie.tuwien.ac.at

and

Geometric Modeling and Scientific Visualization Research Center
King Abdullah University of Science and Technology
Thuwal 23955-6900
Saudi Arabia
helmut.pottmann@kaust.edu.sa

Johannes Wallner
Institute of Geometry
Graz University of Technology
Kopernikusgasse 24
8010 Graz
Austria
j.wallner@tugraz.at

ISSN 1612-3786
ISBN 978-3-540-42058-3 (hardcover)
ISBN 978-3-642-04017-7 (softcover)
DOI 10.1007/978-3-642-04018-4
Springer Heidelberg Dordrecht London New York

e-ISBN 978-3-642-04018-4

Library of Congress Control Number: 2009942430

Mathematics Subject Classification (2000): 51M30, 53A25, 51J15, 70B10, 65Y25, 68U05, 14Qxx

© Springer-Verlag Berlin Heidelberg 2001, First softcover printing 2010

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: deblik, Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Line geometry in its original sense is concerned with the set of lines of three-dimensional space. This set of lines may be considered from different angles: there is the incidence geometry, i.e., projective geometry, of lines and linear line manifolds. Line space is a smooth manifold, whose simplest elements are ruled surfaces. There is the algebraic geometry of lines. Computing with lines requires coordinates and geometric models of line space. There are relations to mechanics and spatial kinematics, and therefore applications of line geometry in mechanism design and robotics.

When studying geometric problems of applied mathematics, in Computing and Visualization, line geometry often occurs naturally and leads to the most elegant and efficient solution. Our interest and research in this area showed us that line-geometric methods play an important role — not only do results of old geometers contribute to new problems, but new challenges typically inspire the theory behind the applications as well. Thus there is a very fruitful interplay between academic research and practical applications in this active area. It is one of the aims of this book to show these connections. For this purpose we have interpreted ‘line geometry’ in a broad sense.

This book is written for those graduate students and researchers in science and engineering, who want to learn more about geometry, geometric computing, and its applications. We assume that the reader is familiar with linear algebra, calculus, and elementary differential geometry. Material which requires more background and is not essential for the rest of the book (some remarks and one section) is marked with an asterisk.

In the first chapter we introduce projective geometry in some detail, and briefly consider projective differential geometry, some basic concepts concerning algebraic curves and surfaces, and elementary facts about Bézier and B-spline methods in Geometric Design. The core of this book is an exposition of line geometry both from the abstract point of view and the viewpoint of scientific computing. The presentation of the classical material is influenced by the work of Heinrich Brauner and Gunter Weiss and their consequent use of the Klein model.

The selection of the material has been governed by the applications we are aiming at: Computer Aided Design, Geometric Modeling, Scientific Visualization, Computer Aided Manufacturing and Robotics. We do not treat data structures and

do not perform analysis of algorithms. There is not much in the literature about the line-geometric side of these topics — the interested reader should consult [24].

We have included some recent and so far unpublished results from a research project on computational line geometry. We gratefully acknowledge the funding of this project by grant No. P13648-MAT of the Austrian Science Fund.

In the past years we have been encouraged by several people to write a book on ‘classical’ geometry and its applications in geometric computing. This task seemed too ambitious, so we have narrowed down the area a bit. However, as line geometry has so many connections with other areas which are addressed briefly as well, we hope that we can at least partially fulfill the wishes of our friends in the scientific community.

We should point out that the idea to write a book on ‘computational line geometry’ is not ours, but came from Bahram Ravani at the University of California in Davis, USA. We want to express our thanks to Bahram Ravani, especially for his continuing support of this project. He made it possible that the first author could enjoy the stimulating academic atmosphere at UC Davis for several times. Many discussions with Bahram Ravani, Rida Farouki, Kenneth Joy and Bernd Hamann had an important impact on the progress of the present book.

We have been lucky to find many people who have been willing to help us in various ways. Michael Hofer, Stefan Leopoldseder and Heidrun Mühlthaler have put much time and effort in reading the text carefully. They have eliminated numerous errors and inaccuracies, and their excellent suggestions have greatly improved the readability of the text. Any remaining errors are, of course, the authors’ sole responsibility.

Special emphasis has been laid on visualization, which is also expressed by the fact that this book has 264 figures. There would have been much less illustrations and much less elaborate ones, if we had not had help from Gershon Elber, Georg Glaeser, Bert Jüttler, Hannes Kaufmann, Stefan Leopoldseder, Martin Peternell, Norbert Pfeifer, Norbert Pomaroli, Hellmuth Stachel, Tamás Várády, Michael Wagner, and Tony Wills. Most of the color plates have been created by Georg Glaeser and Hans-Peter Schröcker from the University of Applied Arts in Vienna — the software package ‘Open Geometry’ [64] turned out to be an excellent tool. Last, but by no means least, most of our figures and many examples have been created by Boris Odehnal, whose work was supported by grant No. P13648-MAT of the Austrian Science Fund. Many thanks again to all our friends who helped us with the present book.

Finally we are very grateful to Martin Peters and Ruth Allewelt from Springer-Verlag, Heidelberg. They supported this book project in the best possible way, and provided conditions where it has been a pleasure to work.

Table of Contents

| | |
|--|------------|
| Preface | v |
| Table of Contents | vii |
| 1. Fundamentals | 1 |
| 1.1 Real Projective Geometry | 1 |
| 1.1.1 The Real Projective Plane | 1 |
| 1.1.2 n -dimensional Projective Space | 6 |
| 1.1.3 Projective Mappings | 11 |
| 1.1.4 Projectivities, Cross Ratio and Harmonic Position | 20 |
| 1.1.5 Polarities and Quadrics | 28 |
| 1.1.6 Complex Extension and the Way from Projective to Euclidean Geometry | 54 |
| 1.2 Basic Projective Differential Geometry | 68 |
| 1.2.1 Curves | 68 |
| 1.2.2 Surfaces | 78 |
| 1.2.3 Duality | 82 |
| 1.3 Elementary Concepts of Algebraic Geometry | 86 |
| 1.3.1 Definitions and Algorithms | 86 |
| 1.3.2 Geometric Properties of Varieties in Projective Space | 99 |
| 1.3.3 Duality | 104 |
| 1.4 Rational Curves and Surfaces in Geometric Design | 105 |
| 1.4.1 Rational Bézier Curves | 105 |
| 1.4.2 Dual Bézier Curves | 121 |
| 1.4.3 Rational Bézier Surfaces | 126 |
| 2. Models of Line Space | 133 |
| 2.1 The Klein Model | 133 |
| 2.1.1 Plücker Coordinates | 133 |
| 2.1.2 Computing with Plücker Coordinates | 137 |
| 2.1.3 The Klein Quadric | 141 |
| 2.2 The Grassmann Algebra | 144 |
| 2.3 The Study Sphere | 154 |

| | |
|---|-----|
| 3. Linear Complexes | 159 |
| 3.1 The Structure of a Linear Complex | 159 |
| 3.1.1 Linear Complexes and Null Polarities in Projective Space . . | 159 |
| 3.1.2 Linear Complexes and Helical Motions in Euclidean Space . | 163 |
| 3.1.3 Linear Complexes in the Klein Model | 168 |
| 3.2 Linear Manifolds of Complexes | 171 |
| 3.2.1 Pencils of Linear Line Complexes | 172 |
| 3.2.2 Euclidean Properties of Pencils of Linear Complexes | 178 |
| 3.3 Reguli and Bundles of Linear Complexes | 181 |
| 3.4 Applications | 185 |
| 3.4.1 Spatial Kinematics | 185 |
| 3.4.2 Statics and Screw Theory | 191 |
| 4. Approximation in Line Space | 195 |
| 4.1 Fitting Linear Complexes | 195 |
| 4.2 Kinematic Surfaces | 202 |
| 4.3 Approximation via Local Mappings into Euclidean 4-Space | 211 |
| 4.4 Approximation in the Set of Line Segments | 221 |
| 5. Ruled Surfaces | 223 |
| 5.1 Projective Differential Geometry of Ruled Surfaces | 223 |
| 5.1.1 Infinitesimal Properties of First Order | 225 |
| 5.1.2 Infinitesimal Properties of Higher Order | 234 |
| 5.2 Algebraic Ruled Surfaces | 238 |
| 5.2.1 Rational Ruled Surfaces | 242 |
| 5.2.2 The Bézier Representation of Rational Ruled Surfaces . . . | 247 |
| 5.2.3 Skew Cubic Surfaces | 252 |
| 5.3 Euclidean Geometry of Ruled Surfaces | 261 |
| 5.3.1 First Order Properties | 263 |
| 5.3.2 A Complete System of Euclidean Invariants | 270 |
| 5.4 Numerical Geometry of Ruled Surfaces | 282 |
| 5.4.1 Discrete Models and Difference Geometry | 282 |
| 5.4.2 Interpolation and Approximation Algorithms | 291 |
| 5.4.3 Variational Design | 296 |
| 5.4.4 Offset Surfaces and their Applications | 303 |
| 5.4.5 Intersection of Ruled Surfaces | 309 |
| Color Plates | 311 |
| 6. Developable Surfaces | 327 |
| 6.1 Differential Geometry of Developable Surfaces | 327 |
| 6.2 Dual Representation | 334 |
| 6.2.1 Differential Geometry of the Dual Surface | 334 |
| 6.2.2 Developable Bézier and B-Spline Surfaces | 343 |

| | | |
|-----------|---|------------|
| 6.2.3 | Interpolation and Approximation Algorithms with Developable Surfaces | 352 |
| 6.3 | Developable Surfaces of Constant Slope and Applications | 358 |
| 6.3.1 | Basics | 359 |
| 6.3.2 | The Cyclographic Mapping and its Applications | 366 |
| 6.3.3 | Rational Developable Surfaces of Constant Slope and Rational Pythagorean-Hodograph Curves | 383 |
| 6.4 | Connecting Developables and Applications | 396 |
| 6.4.1 | Basics | 396 |
| 6.4.2 | Convex Hulls and Binder Surfaces | 400 |
| 6.4.3 | Geometric Tolerancing | 405 |
| 6.4.4 | Two-Dimensional Normed Spaces and Minkowski Offsets | 410 |
| 6.5 | Developable Surfaces with Creases | 416 |
| 7. | Line Congruences and Line Complexes | 423 |
| 7.1 | Line Congruences | 423 |
| 7.1.1 | Projective Differential Geometry of Congruences | 423 |
| 7.1.2 | Rational Congruences and Trivariate Bézier Representations | 428 |
| 7.1.3 | Euclidean Differential Geometry of Line Congruences | 434 |
| 7.1.4 | Normal Congruences and Geometrical Optics | 446 |
| 7.1.5 | Singularities of Motions Constrained by Contacting Sur- faces and Applications in Sculptured Surface Machining | 452 |
| 7.1.6 | Numerical Geometry of Line Congruences | 465 |
| 7.1.7 | Projection via Line Congruences | 469 |
| 7.2 | Line Complexes | 474 |
| 7.2.1 | Differential Geometry of Line Complexes | 474 |
| 7.2.2 | Algebraic Complexes and Congruences | 480 |
| 7.2.3 | Special Quadratic Complexes | 487 |
| 8. | Linear Line Mappings — Computational Kinematics | 497 |
| 8.1 | Linear Line Mappings and Visualization of the Klein Model | 497 |
| 8.1.1 | Linear Line Mappings into P^2 | 498 |
| 8.1.2 | Linear Line Mappings into P^3 | 511 |
| 8.1.3 | Visualization of the Klein Image | 519 |
| 8.2 | Kinematic Mappings | 522 |
| 8.2.1 | Quaternions | 523 |
| 8.2.2 | The Spherical Kinematic Mapping | 527 |
| 8.2.3 | Other Kinematic Mappings | 535 |
| 8.3 | Motion Design | 538 |
| | References | 547 |
| | List of Symbols | 556 |
| | Index | 557 |

1. Fundamentals

“All geometry is projective geometry.” (Arthur Cayley, 1821–1895)

The first chapter of this book is an introduction into projective geometry. Although projective geometry is, from the abstract viewpoint, nothing but linear algebra in disguise, it is a geometric language which allows a unified approach to such different things as Euclidean geometry of points, the differential geometry of ruled surfaces, or spherical kinematics.

Sec. 1.1 introduces projective space and coordinates for points and hyperplanes, studies linear subspaces and quadratic varieties, and shows the relations between some of the classical geometries. In Sec. 1.2, we consider curves and surfaces from the viewpoint of projective geometry. Afterwards we describe some basic facts about algebraic varieties in affine and projective space. Finally, Sec. 1.4 deals with polynomial and rational curves and surfaces in Computer-Aided Geometric Design.

1.1 Real Projective Geometry

1.1.1 The Real Projective Plane

History and Definition of Ideal Points

Projective geometry has its origin in the development of the geometric rules of perspective drawing. A central projection maps parallel lines that are not parallel to the image plane to *intersecting* lines. This example indicates that a geometry used to study central projections will have to regard parallelity as a form of intersection. This is realized by the concept of *projective extension* of Euclidean space.

We first discuss the projective extension of the Euclidean plane. We add an *ideal point* or *point at infinity* to each line L , such that parallel lines share the same ideal point. Thus, two lines are parallel if and only if they intersect at infinity — all other lines intersect in a proper point. The *ideal line*, denoted by ω , consists of all ideal points. In figures we often indicate the ideal point of a line L by an arrow labelled with L_ω or L_u . It is part of the definition of the point at infinity that a line contains exactly *one* ideal point. It can be approached by traveling along the line in either direction.

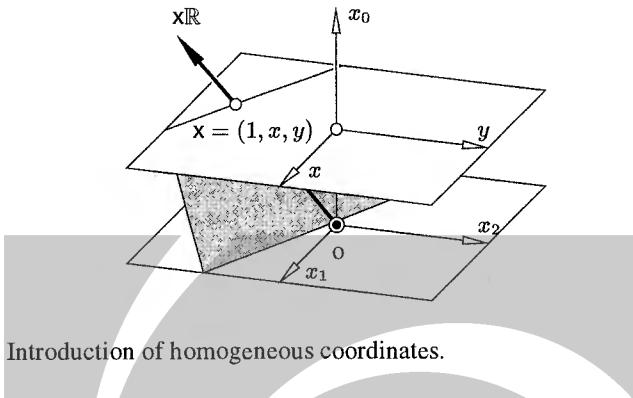


Fig. 1.1. Introduction of homogeneous coordinates.

Homogeneous Point Coordinates

We introduce coordinates which describe proper points as well as ideal points: We choose a Cartesian coordinate system in the Euclidean plane E^2 . Its points are represented by coordinate vectors $\mathbf{x} = (x, y) \in \mathbb{R}^2$. The linear space \mathbb{R}^3 is assumed to have coordinates x_0, x_1, x_2 . We embed the Euclidean plane into \mathbb{R}^3 by mapping $(x, y) \mapsto (1, x, y)$. Thus E^2 is identified with the plane $x_0 = 1$ of \mathbb{R}^3 (see Fig. 1.1).

The one-dimensional linear subspace spanned by $\mathbf{x} = (1, x, y)$ consists of the vectors

$$\mathbf{x}\mathbb{R} = (\lambda, \lambda x, \lambda y), \quad \lambda \in \mathbb{R}.$$

If $\lambda \neq 0$, then such a coordinate triple uniquely defines the subspace $\mathbf{x}\mathbb{R}$ and the corresponding proper point of E^2 . Thus, to a point with two coordinates in E^2 we assign three, which are unique only up to a common nonzero factor, and are called *homogeneous Cartesian coordinates*. The symbol $\mathbf{x}\mathbb{R}$ denotes the point as well as the symbol \mathbf{x} , if $\mathbf{x} = \lambda(1, x, y)$ and $\mathbf{x} = (x, y)$. To emphasize the homogeneity, we use the notation $(x_0 : x_1 : x_2)$ for $(x_0, x_1, x_2)\mathbb{R}$. The ‘ordinary’ coordinates are recovered from the homogeneous ones by

$$\mathbf{x}\mathbb{R} \cong (x_0 : x_1 : x_2) \cong (x_1/x_0, x_2/x_0) \cong \mathbf{x}, \quad (x_0 \neq 0). \quad (1.1)$$

When using matrix notation we regard \mathbf{x} (or \mathbf{x}) as a column vector, even if we write $\mathbf{x} = (x, y)$ for simplicity.

Representation of Euclidean Lines and Ideal Points

Consider a line L parallel to the vector (l_1, l_2) in the Euclidean plane. Its points \mathbf{x} are represented by one-dimensional subspaces $\mathbf{x}\mathbb{R}$ of \mathbb{R}^3 . They are contained in a two-dimensional linear subspace of \mathbb{R}^3 . One one-dimensional subspace, however, is missing: It is spanned by $(0, l_1, l_2)$, is parallel to L , and is contained in the plane $x_0 = 0$. If L' is parallel to L , then the one-dimensional subspace parallel to L' obviously is the same as it was for L .

Conversely, all one-dimensional subspaces of \mathbb{R}^3 , except those in $x_0 = 0$, represent a proper point of E^2 . A one-dimensional subspace in $x_0 = 0$ is contained in two-dimensional subspaces which correspond to a family of parallel lines. This shows how to represent ideal points: The ‘missing’ one-dimensional subspaces of the previous paragraph are the obvious candidates.

Now we have established a one-to-one correspondence between one-dimensional linear subspaces of \mathbb{R}^3 and points (both proper and ideal) of the projectively extended plane. This extended plane is called the *real projective plane*, and is denoted by P^2 .

Lines of P^2 are represented by two-dimensional linear subspaces of \mathbb{R}^3 . All except the subspace $x_0 = 0$ determine a Euclidean line, and $x_0 = 0$ corresponds to the ideal line ω that contains all ideal points.

Homogeneous Line Coordinates and the Incidence Relation

If a point $p\mathbb{R}$ is contained in a line L , we call these two objects *incident*. The pair $(p\mathbb{R}, L)$ then is a member of the *incidence relation*. A line L spanned by the points $a\mathbb{R}$ and $b\mathbb{R}$ can be parametrized by

$$x\mathbb{R} = (\lambda_0 a + \lambda_1 b)\mathbb{R}, \quad \lambda_0, \lambda_1 \in \mathbb{R}. \quad (1.2)$$

We may also write this two-dimensional subspace as the solution of a homogeneous linear equation

$$u_0 x_0 + u_1 x_1 + u_2 x_2 = 0. \quad (1.3)$$

The coefficients $(u_0, u_1, u_2) = u$ are unique up to a scalar factor, and are called *homogeneous line coordinates* of L . In order to distinguish line coordinates from point coordinates, we write $L = \mathbb{R}(u_0, u_1, u_2) = \mathbb{R}u$ with the symbol \mathbb{R} at the left hand side of the coordinate vector.

We can give a Euclidean interpretation to Equ. (1.3): If we introduce the canonical scalar product in \mathbb{R}^3 , then Equ. (1.3) means that the vectors (u_0, u_1, u_2) and (x_0, x_1, x_2) are orthogonal. We write

$$u \cdot x = 0. \quad (1.4)$$

This is the condition of *incidence* of a line $\mathbb{R}u$ and a point $x\mathbb{R}$. To compute the line $\mathbb{R}u$ incident with two points $a\mathbb{R}$, $b\mathbb{R}$, we therefore have to determine u such that it is orthogonal to both $a\mathbb{R}$ and $b\mathbb{R}$. This can be done by letting

$$u = a \times b. \quad (1.5)$$

Conversely, the point $a\mathbb{R}$ incident with lines $\mathbb{R}u$, $\mathbb{R}v$ is computed by

$$a = u \times v. \quad (1.6)$$

The Incidence Structure of the Projective Plane

The incidence structure of the projective plane is very simple: Two different points $A_1 = \mathbf{a}_1\mathbb{R}$, $A_2 = \mathbf{a}_2\mathbb{R}$ are incident with exactly one line $L = \mathbb{R}\mathbf{u}$ (cf. Equ. (1.5)). We write

$$L = A_1 \vee A_2.$$

Two different lines L_1 , L_2 are incident with exactly one point P . We write

$$P = L_1 \cap L_2.$$

This is in contrast to the situation in the Euclidean plane, where there are lines which do not intersect (parallel lines). The incidence of a point P and the line L could be denoted by the symbol $L \sqcap P$, in order to emphasize the ‘symmetry’ of this relation. But we prefer to write $P \in L$.

We call points incident with the same line *collinear*, and lines incident with the same point *concurrent*. If $\mathbf{a}_1\mathbb{R}$, $\mathbf{a}_2\mathbb{R}$, $\mathbf{a}_3\mathbb{R}$ are incident with the line $\mathbb{R}\mathbf{u}$, then $\mathbf{a}_1 \cdot \mathbf{u} = \mathbf{a}_2 \cdot \mathbf{u} = \mathbf{a}_3 \cdot \mathbf{u} = 0$, and $\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = 0$. Conversely, if this determinant is zero, the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are contained in a two-dimensional subspace and so $\mathbf{a}_1\mathbb{R}$, $\mathbf{a}_2\mathbb{R}$, $\mathbf{a}_3\mathbb{R}$ are collinear points. We have shown:

$$\mathbf{a}_1\mathbb{R}, \mathbf{a}_2\mathbb{R}, \mathbf{a}_3\mathbb{R} \text{ collinear} \iff \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = 0.$$

Duality in the Projective Plane

The incidence condition (1.4) is completely symmetric: If we interchange point and line coordinates, it remains unchanged. This has an important consequence.

Definition. *The geometric objects ‘line’ and ‘point’ are dual to each other in the projective plane. If in a statement about points and lines and their incidences every occurrence is replaced by its dual, this new statement is called the dual of the original one.*

We have defined points to be *collinear* if they are contained in the same line, lines to be *concurrent*, if they contain the same point. If we want to convert a statement which contains the words ‘collinear’ and ‘concurrent’ into its dual version, we have to replace ‘collinear’ with ‘concurrent’ and vice versa.

Theorem 1.1.1. *(The principle of duality in the projective plane) If a true statement about geometric objects in the projective plane employs only ‘points’, ‘lines’ and ‘incidence’, its dual is also true, and vice versa.*

Proof. The proof of the dual statement is obtained by interchanging the words ‘point’ and ‘line’ in the original proof. Equations (1.4), (1.5), and (1.6) show that in that way the dual statement is proved. \square

Remark 1.1.1. Th. 1.1.1 is actually a meta-theorem: a theorem about other theorems. Its proof is a sketch of a meta-proof. It is also necessary to specify further what a ‘statement’ is. As this is no course in logic, we just mention that it should not contain free variables, i.e., all variable names should be preceded by a ‘for all’ or ‘exists’ quantifier. \diamond

Usually the dual statement differs from its original. If not, the statement is called *self-dual*. Geometric objects can be dualized by dualizing their definition, which is a statement.

Example 1.1.1. A *range of points*, i.e., the points incident with a line, is dual to a *pencil of lines*, i.e., the lines concurrent in a point. The line spanned by two points is dual to the intersection point of two lines. A *line element*, i.e., a line with a point in it, is a self-dual geometric object. \diamond

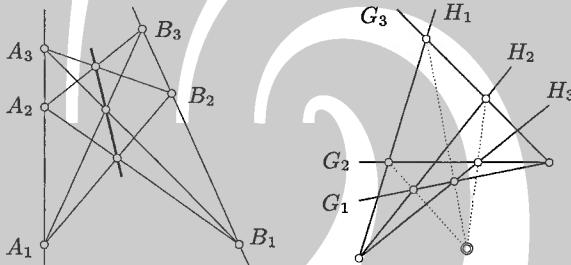


Fig. 1.2. Left: Pappos’ theorem. Right: its dual (Brianchon’s theorem).

Example 1.1.2. We give a statement that can be dualized: Suppose A_1, A_2, A_3 are collinear, and likewise B_1, B_2, B_3 . Then the three points $(A_1 \vee B_2) \cap (A_2 \vee B_1)$, $(A_2 \vee B_3) \cap (A_3 \vee B_2)$, $(A_3 \vee B_1) \cap (A_1 \vee B_3)$ are collinear (see Fig. 1.2, left).

The dual statement is the following: Suppose G_1, G_2, G_3 are concurrent, and likewise H_1, H_2, H_3 . Then the three lines $(G_1 \cap H_2) \vee (G_2 \cap H_1)$, $(G_2 \cap H_3) \vee (G_3 \cap H_2)$, $(G_3 \cap H_1) \vee (G_1 \cap H_3)$ are concurrent (see Fig. 1.2, right). We will give a proof in Ex. 1.1.9. \diamond

Models of the Projective Plane

The use of homogeneous coordinates motivates the so-called *bundle model* of the real projective plane. All lines of Euclidean 3-space concurrent in a point are called a *bundle* of lines. There is an obvious one-to-one correspondence between the lines of the bundle and the one-dimensional linear subspaces of \mathbb{R}^3 . Recall that we have established a one-to-one correspondence between the one-dimensional linear subspaces of \mathbb{R}^3 and the points of the projective plane. Thus we can say that the bundle is (a model of) the projective plane P^2 .

A pencil of lines in the bundle corresponds to the range of points contained in a line of P^2 .

Another model is the following: All one-dimensional linear subspaces of \mathbb{R}^3 intersect the unit sphere S^2 in two points. Thus the unit sphere with antipodal points identified is a model of P^2 . Yet another model is the northern hemisphere, with antipodal points of the equator identified.

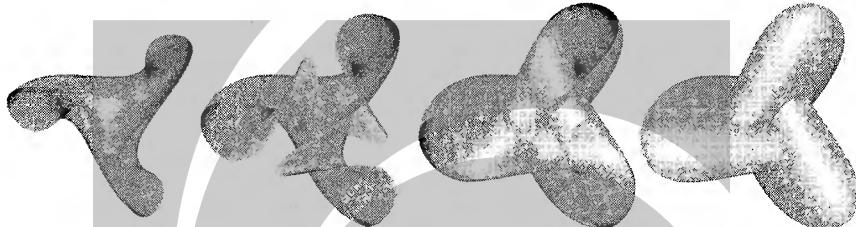


Fig. 1.3. Boy's surface and its planar sections.



Remark 1.1.2. We saw that the projective plane can be constructed by identifying antipodal points of the unit sphere, which makes the projective plane a closed surface. Instead of identifying all points with their antipodes we could restrict ourselves to the upper hemisphere together with the equator, where all pairs of antipodal points are represented by at least one point, and precisely the points of the equator by two points. This shows that we can make a projective plane from an upper hemisphere by gluing the equator to itself such that antipodal points come together. This construction is called a *crosscap* attached to a circular hole in the sphere (in this case, the lower hemisphere is the hole).

Yet another equivalent method is to take a Möbius band and a disk, and to glue them together along their boundaries. The resulting surface is non-orientable and cannot be embedded into Euclidean three-space without self-intersections. But it has been known for a long time that the projective plane admits an immersion in \mathbb{R}^3 , i.e., it can be mapped to a closed surface in without local singularities. The first to actually draw a picture of an immersion was W. Boy in 1901. Ever since, any similarly shaped surface which is an immersion of the projective plane is called a Boy's surface (cf. [4]). Fig. 1.3 shows an example. \diamond

1.1.2 n -dimensional Projective Space

The construction of n -dimensional real projective space P^n is completely analogous to that of P^2 . We augment Euclidean n -space by ideal points, one for each class of parallel lines, and collect the ideal points in the *ideal hyperplane*, denoted by ω . We define a one-to-one correspondence between one-dimensional subspaces of \mathbb{R}^{n+1} and the points of P^n :

Definition. Assume that the linear space \mathbb{R}^{n+1} is equipped with coordinates x_0, x_1, \dots, x_n . The set of its one-dimensional subspaces is called *n-dimensional real projective space* P^n . We embed *n-dimensional Euclidean space* E^n into P^n by $(x_1, \dots, x_n) \mapsto (1, x_1, \dots, x_n)\mathbb{R}$. If $x_0 = 0$, the point $(x_0, \dots, x_n)\mathbb{R}$ is called an *ideal point*.

We denote the point $(x_0, \dots, x_n)\mathbb{R} = x\mathbb{R}$ also by the symbol $(x_0 : \dots : x_n)$. If $x\mathbb{R}$ is a proper point (i.e., $x_0 \neq 0$), its coordinates in E^n are recovered by

$$(x_0, \dots, x_n)\mathbb{R} \in P^n \iff (x_1/x_0, \dots, x_n/x_0) \in E^n \quad (x_0 \neq 0). \quad (1.7)$$

If a line of E^n is parallel to the vector (l_1, \dots, l_n) , its ideal point is $(0, l_1, \dots, l_n)\mathbb{R}$. The ideal points are contained in the *ideal hyperplane*, whose equation is $x_0 = 0$. The coordinate vector x of the point $x\mathbb{R}$ is called its *homogeneous coordinate vector*. The correspondence between one-dimensional subspaces and points represents one of the most important connections between algebra and geometry.

If we want to distinguish this model of P^n from others, we refer to it as to ‘the real projective space of \mathbb{R}^{n+1} ’.

If we repeat the construction of *n*-dimensional projective space and homogeneous coordinates in it for *complex* numbers instead of real ones, we get so-called *complex projective space* $\mathbb{C}P^n$. It will be discussed in more detail in Sec. 1.1.6, but we will occasionally refer to it before. As a general rule, everything which can be computed linearly holds true for both projective spaces, real and complex. Objects defined by nonlinear equations behave in different ways. This is mainly because polynomials always have zeros over the complex number field, but not necessarily over the field of real numbers.

Projective Subspaces

Definition. A $(k+1)$ -dimensional linear subspace U of \mathbb{R}^{n+1} defines a k -dimensional projective subspace, which consists of all points $x\mathbb{R}$ with $x \in U$.

The zero subspace of \mathbb{R}^{n+1} in this way corresponds to the empty set (there is no point contained in the zero subspace). It is assigned the dimension $k = -1$. Points $x\mathbb{R}$ are 0-dimensional projective subspaces. If $k = 1$, the projective subspace is called a *line*, if $k = 2$, a *plane*, if $k = n-1$, a *hyperplane*. The only *n*-dimensional projective subspace is P^n itself.

Definition. Points $P_0 = p_0\mathbb{R}, \dots, P_k = p_k\mathbb{R}$ of P^n are *projectively independent*, if and only if the vectors p_0, \dots, p_k are *linearly independent*.

Any k -dimensional projective subspace is a k -dimensional projective space itself. If U is the corresponding linear subspace, we write $\Pi(U)$ to indicate the projective subspace. We will often omit the symbol Π and write U for the projective subspace as well.

Projective Span

Assume $k + 1$ linearly independent vectors p_0, \dots, p_k which span a $(k + 1)$ -dimensional linear subspace U . We denote the projective subspace corresponding to this linear subspace by the symbol

$$U = [p_0, \dots, p_k] = P_0 \vee \dots \vee P_k, \text{ where } p_i \mathbb{R} = P_i.$$

It is called the *projective span* of P_0, \dots, P_k . Two projectively independent points span a line and three projectively independent points span a plane. U can be parametrized by

$$x\mathbb{R} = (\lambda_0 p_0 + \dots + \lambda_k p_k)\mathbb{R} \quad (1.8)$$

with $k + 1$ homogeneous parameters $\lambda_0, \dots, \lambda_k \in \mathbb{R}$.

Hyperplanes can be represented alternatively by *one linear equation*

$$u_0 x_0 + \dots + u_n x_n = u \cdot x = 0. \quad (1.9)$$

The homogeneous $(n + 1)$ -tuple $(u_0 : \dots : u_n)$ is called *homogeneous hyperplane coordinate*, and we denote the hyperplane with coordinate vector u by $\mathbb{R}u$. The incidence relation is described by (1.9): A point $x\mathbb{R}$ and a hyperplane $\mathbb{R}u$ are incident if and only if $u \cdot x = 0$.

Example 1.1.3. We compute the hyperplane coordinates of the plane spanned by the points $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ of Euclidean space E^3 : After the embedding into projective space P^3 these three points have coordinates $x_1\mathbb{R} = (1, 1, 0, 0)\mathbb{R}$, $x_2\mathbb{R} = (1, 0, 1, 0)\mathbb{R}$, and $x_3\mathbb{R} = (1, 0, 0, 1)\mathbb{R}$. Their span $\mathbb{R}u$ must fulfill $u \cdot x_1 = u \cdot x_2 = u \cdot x_3 = 0$. These are three homogeneous linear equations for the four unknown components of u . In this case we see even without computing that the solution is $\mathbb{R}(-1, 1, 1, 1)$. \diamond

Projective Span and Intersection of Subspaces

Suppose U and V are linear subspaces of \mathbb{R}^{n+1} . They define projective subspaces, also denoted by U and V . The symbol $U \vee V$ denotes the smallest (linear or projective) subspace which contains both U, V . By $U \cap V$ we denote their intersection, which again is a (linear or projective) subspace.

The following theorem shows a relation between the dimensions of these subspaces:

Theorem 1.1.2. Assume that G, H are projective subspaces of n -dimensional projective space P^n . Then

$$\dim G + \dim H = \dim(G \cap H) + \dim(G \vee H). \quad (1.10)$$

Proof. We consider the linear subspaces corresponding to the projective ones, and denote their dimension with \dim_L . Now it is well known that $\dim_L G + \dim_L H = \dim_L(G \cap H) + \dim_L(G \vee H)$. Because $\dim_L(\cdot) = \dim(\cdot) + 1$, the result follows. \square

If $G \cap H$ is empty and $G \vee H = P^n$, the projective subspaces G, H are called *complementary*. (1.10) implies that in this case, $\dim G + \dim H = n - 1$.

Example 1.1.4. There are only two possibilities for the dimensions of complementary subspaces in the projective plane: The trivial possibility is the empty subspace and entire space, which are complementary. The nontrivial possibility is a point and a non-incident line.

In projective three-space we have, apart from the trivial case, the cases of a plane plus a non-incident point, and of two non-intersecting lines. Such lines are called *skew*. \diamond

If U_i is an indexed family of subspaces, where the index ranges in some set I , we use the symbols

$$V = \bigcap_{i \in I} U_i, \quad W = \bigvee_{i \in I} U_i \quad (1.11)$$

to denote the intersection V and the span W of the subspaces U_i . The former consists of the points which are contained in all U_i , and is a projective subspace again. The latter is the smallest subspace which contains all U_i .

The Annulator Mapping

The incidence condition (1.9) shows that a point $x\mathbb{R}$ is contained in a hyperplane $\mathbb{R}u$, if and only if the vectors x and u are *orthogonal* with respect to the canonical scalar product in \mathbb{R}^{n+1} .

A $(k+1)$ -dimensional linear subspace U of \mathbb{R}^{n+1} has an $(n-k)$ -dimensional orthogonal complement U° . We consider a second copy of P^n , denoted by P^{n*} . U defines a k -dimensional projective subspace in P^n , and U° defines an $(n-k-1)$ -dimensional projective subspace in P^{n*} .

Definition. *The space P^{n*} is called dual projective space. The mapping $U \mapsto U^\circ$ is called the annulator mapping.*

Proposition 1.1.3. *Assume that G and H are projective subspaces. The annulator mapping has the following properties:*

$$(G \cap H)^\circ = G^\circ \vee H^\circ, \quad (G \vee H)^\circ = G^\circ \cap H^\circ, \quad G \subset H \iff G^\circ \supset H^\circ.$$

If U_i is an indexed family of subspaces, then

$$(\bigcap U_i)^\circ = \bigvee U_i^\circ, \quad (\bigvee U_i)^\circ = \bigcap U_i^\circ.$$

Proof. This follows from the respective properties of the orthogonal complement in \mathbb{R}^{n+1} . \square

Remark 1.1.3. We emphasize that from the ‘clean’ point of view, something like orthogonality should not be intermixed with the annulator mapping, which is a natural anti-morphism of the lattice of subspaces. The orthogonality defined by the canonical coordinate system in \mathbb{R}^{n+1} appears here because we expect the reader to be familiar with this concept. \diamond

Duality in n -dimensional Projective Space

Definition. If a statement involving k -dimensional subspaces, projective span, intersection, and inclusion of subspaces in P^n is modified by replacing these items by $(n - k - 1)$ -dimensional subspace, intersection, projective span, and reverse inclusion of subspaces, then this new statement is called dual to the original one.

Obviously, the iterated dual gives the original statement again. A configuration of subspaces, which is also called a ‘geometric figure’, is assigned a dual configuration, by dualizing its definition.

Theorem 1.1.4. A statement which fulfills the criteria of the previous definition is true if and only if its dual is.

Proof. The annihilator mapping transforms a configuration of subspaces into a dual configuration. This shows that the truth of a statement in P^n is equivalent to the truth of the dual statement in P^{n*} . As P^{n*} is just a second copy of P^n , the theorem follows. \square

Remark 1.1.4. The points of dual projective space can be identified with the hyperplanes of P^n . A projective subspace of P^{n*} is then a set of hyperplanes. If U is a projective subspace of P^n , then U° consists of all hyperplanes which contain U . \diamond

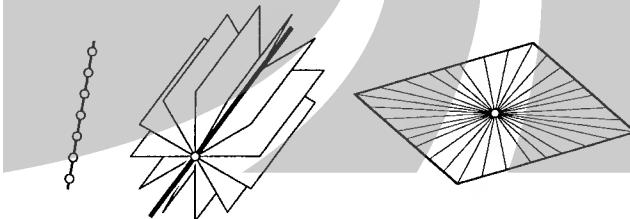


Fig. 1.4. A range of points, a pencil of planes, and a pencil of lines in a plane (from left to right).

Example 1.1.5. We illustrate duality in P^3 : points are dual to planes, and lines are dual to lines. The points of a line — a *range of points* — is dual to the planes which contain a line — a *pencil of planes*.

The *field of points* of a plane is dual to the *bundle of planes* incident with a point. If a point P is contained in the plane ε , then the lines which contain P and are contained in ε are called a *pencil of lines* (cf. Ex. 1.1.1). Its dual figure consists of those lines which are contained in a plane P° , and contain a point ε° , where $\varepsilon^\circ \in P^\circ$. This is again a pencil of lines, so ‘pencil of lines’ is a self-dual object in P^3 . Likewise ‘concurrent lines’ are the same as two ‘co-planar lines’, so this notion is self-dual (cf. Fig. 1.4). \diamond

Remark 1.1.5. Note that there is no such thing as duality in Euclidean space E^n . It is true that E^n is contained in P^n , but what is missing is not self-dual: There are many ideal points, but only one ideal hyperplane. \diamond

Models of Projective Space

Some of the models of the projective plane can be immediately generalized to projective space: a bundle of lines in Euclidean space E^{n+1} is an n -dimensional projective space itself, as well as the unit sphere S^n with antipodal points identified.

Remark 1.1.6. The construction of P^n by identifying antipodal points makes it an n -dimensional manifold whose double cover is the sphere. It is orientable if n is odd, and its fundamental group is \mathbb{Z}_2 if $n \geq 2$. One-dimensional projective space (a projective line) is homeomorphic to a circle. \star

The question of embedding topological spaces into Euclidean spaces as surfaces is very interesting: There is no embedding of P^n into E^{n+1} , and the projective spaces are among those manifolds which are most resistant against embedding into \mathbb{R}^n 's. \diamond

1.1.3 Projective Mappings

Projective Automorphisms

‘Geometry’ in the sense of F. Klein’s Erlangen program (1872) means the theory of objects invariant with respect to a given group of transformations, which are the automorphisms of the geometry. Thus Euclidean geometry is the theory of objects invariant with respect to Euclidean congruence transformations.

The group of automorphisms of n -dimensional projective space are induced by the linear automorphisms of \mathbb{R}^{n+1} :

Definition. If $A \in \mathbb{R}^{(n+1) \times (n+1)}$ is a regular matrix, the linear automorphism $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$

$$f : \mathbf{x} \mapsto \mathbf{x}' = A \cdot \mathbf{x} \quad (1.12)$$

induces the projective mapping $\phi : P^n \rightarrow P^n$ by

$$\phi : \mathbf{x} \mathbb{R} \mapsto (A \cdot \mathbf{x}) \mathbb{R}. \quad (1.13)$$

ϕ is called projective automorphism, projective collineation, or regular projective map. The group of projective automorphisms of P^n is denoted by PGL_n , and is called the projective linear group.

Definition. A projective isomorphism ϕ from one n -dimensional projective space onto another is defined by the same formulae, but with the source coordinate system in the first, and the target coordinate system in the second projective space.

A projective isomorphism is also called projective collineation or regular projective map from the first projective space onto the second.

ϕ maps one-dimensional subspaces $x\mathbb{R}$ of \mathbb{R}^{n+1} to one-dimensional subspaces $x'\mathbb{R}$ of \mathbb{R}^{n+1} , i.e., points of P^n to points of P^n . f is one-to-one and onto, and so is ϕ .

Remark 1.1.7. Two linear automorphisms f_1, f_2 of \mathbb{R}^{n+1} define the same projective automorphism if and only if $f_1 = \lambda f_2$, with $\lambda \neq 0$ (the proof of this fact is left as an exercise to the reader — cf. also the proofs of Prop. 1.1.8 and Th. 2.2.6).

If we denote the group of linear automorphisms of \mathbb{R}^{n+1} by GL_{n+1} and the group of *homotheties* of the form $x \mapsto \lambda x$ with $\mathbb{R} \cdot \mathrm{id}$, then obviously

$$\mathrm{PGL}_n = \mathrm{GL}_{n+1} / (\mathbb{R} \cdot \mathrm{id}).$$

◇

We will show later that a one-to-one mapping of P^n onto itself, which preserves collinearity of points, is necessarily a projective automorphism.

Remark 1.1.8. To indicate that the mapping ϕ is applied to an object P we will use the notation $\phi(P)$ or $P\phi$. The symbol $\phi\psi$ means that we apply ϕ first and then ψ . This is the same as what is usually denoted by $\psi \circ \phi$. ◇

The Action of Projective Automorphisms on Hyperplanes

A linear automorphism f of \mathbb{R}^{n+1} preserves the \vee , \cap , and \subset relations between linear subspaces, and therefore so does the projective automorphism ϕ which is induced by f .

In particular, a hyperplane $\mathbb{R}u$ is mapped to a hyperplane $\mathbb{R}u'$. We want to compute u' from u and from the matrix A describing f .

Proposition 1.1.5. *The projective automorphism $x\mathbb{R} \mapsto x'\mathbb{R} = (A \cdot x)\mathbb{R}$ transforms hyperplanes according to*

$$u' = (A^{-T}) \cdot u, \quad (1.14)$$

where A^{-T} denotes the transpose inverse of A .

Proof. We have to show that $u' \cdot x' = 0$ is equivalent to $u \cdot x = 0$: We compute $u' \cdot x' = u'^T x' = (A^{-T} u)^T (Ax) = u^T A^{-1} Ax = u \cdot x$. The proposition is proved. ◻

Example 1.1.6. Consider the projective automorphism $\kappa : x\mathbb{R} \mapsto (A \cdot x)\mathbb{R}$ of the real projective plane with

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad A^{-T} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & -1 & -1 \end{bmatrix}.$$

The points $P = p\mathbb{R} = (2, 0, -1)\mathbb{R}$ and $Q = q\mathbb{R} = (0, 1, 0)\mathbb{R}$ are contained in the line $L = \mathbb{R}u = \mathbb{R}(1, 0, 2)$ (cf. Equ. (1.4)). Their images $P\kappa = (A \cdot p)\mathbb{R} = (1, 1, 3)\mathbb{R}$ and $Q\kappa = (-1, 1, 0)\mathbb{R}$ are contained in the line $L\kappa = \mathbb{R}(A^{-T} \cdot u) = \mathbb{R}(3, 3, -2)$. ◻

Fundamental Sets

A linear automorphism $f \in \mathrm{GL}_{n+1}$ is uniquely defined by its values on a basis $\{\mathbf{b}_0, \dots, \mathbf{b}_n\}$. Basis vectors, however, are no geometric objects in projective geometry. A vector \mathbf{b}_i defines a projective point, but not conversely.

It turns out that a projective automorphism is uniquely determined by its values on $n + 2$ points ‘in general position’:

Definition. A set of $n + 2$ points of P^n is called a fundamental set if every subset of $n + 1$ points is projectively independent.

Proposition 1.1.6. If P_0, \dots, P_{n+1} and P'_0, \dots, P'_{n+1} are fundamental sets of P^n , there exists a projective automorphism ϕ which maps P_i to P'_i for $i = 0, \dots, n + 1$.

The same is true for projective isomorphisms, where the points P_i and the points P'_j are contained in different projective spaces.

Proof. We let $P_i = \mathbf{b}_i \mathbb{R}$, and $P'_i = \mathbf{b}'_i \mathbb{R}$ ($i = 0, \dots, n + 1$). Then $\{\mathbf{b}_0, \dots, \mathbf{b}_n\}$ and $\{\mathbf{b}'_0, \dots, \mathbf{b}'_n\}$ are bases of \mathbb{R}^{n+1} , and we can write $\mathbf{b}_{n+1} = \sum_{i=0}^n \lambda_i \mathbf{b}_i$, and $\mathbf{b}'_{n+1} = \sum_{i=0}^n \lambda'_i \mathbf{b}'_i$. All coefficients λ_i (and λ'_i) are nonzero, because otherwise $n + 1$ of the vectors \mathbf{b}_i (or \mathbf{b}'_i) would be linearly dependent.

We let $\mathbf{c}_i = \lambda_i \mathbf{b}_i$ and $\mathbf{c}'_i = \lambda'_i \mathbf{b}'_i$. Then $\{\mathbf{c}_0, \dots, \mathbf{c}_n\}$ and $\{\mathbf{c}'_0, \dots, \mathbf{c}'_n\}$ are bases of \mathbb{R}^{n+1} . The linear mapping f which maps $\mathbf{c}_i \mapsto \mathbf{c}'_i$ ($i = 0, \dots, n$) induces a projective automorphism ϕ which maps $P_i \rightarrow P'_i$ for $i = 0, \dots, n$.

But ϕ also maps $P_{n+1} \mapsto P'_{n+1}$, because $f(\mathbf{b}_{n+1}) = f(\mathbf{c}_0 + \dots + \mathbf{c}_n) = \mathbf{c}'_0 + \dots + \mathbf{c}'_n = \mathbf{b}'_{n+1}$. \square

Example 1.1.7. We want to show that for any two hyperplanes H, H' there exists a projective automorphism ϕ with $H\phi = H'$. It is well known that any independent set of vectors in \mathbb{R}^{n+1} can be completed to a basis. So we can choose fundamental sets (P_0, \dots, P_{n+1}) and (P'_0, \dots, P'_{n+1}) such that P_0, \dots, P_n span H , and P'_0, \dots, P'_n span H' . Then a projective automorphism with $P_i\phi = P'_i$ ($i = 0, \dots, n + 1$) maps H to H' . \diamond

Example 1.1.8. Assume that $\mathbf{b}_0 = (1, 0, 1)$, $\mathbf{b}_1 = (0, 1, -1)$, $\mathbf{b}_2 = (1, 0, 0)$, $\mathbf{b}_3 = (1, 1, -2)$, and $\mathbf{b}'_0 = (2, -1, 0)$, $\mathbf{b}'_1 = (-2, 2, 1)$, $\mathbf{b}'_2 = (1, 0, 1)$, $\mathbf{b}'_3 = (-2, 3, 3)$. We look for a projective automorphism κ of the projective plane such that $\kappa(\mathbf{b}_i \mathbb{R}) = \mathbf{b}'_i \mathbb{R}$ for $i = 0, \dots, 3$. We use the notation of the proof of Prop. 1.1.6. First we determine λ_i such that $\sum_{i=0}^2 \lambda_i \mathbf{b}_i = \mathbf{b}_3$ and similarly λ'_i for \mathbf{b}'_i . This gives $\lambda_0 = -1$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda'_0 = -1$, $\lambda'_1 = 1$, $\lambda'_2 = 2$. We let $\mathbf{c}_i = \lambda_i \mathbf{b}_i$ and $\mathbf{c}'_i = \lambda'_i \mathbf{b}'_i$ for $i = 0, 1, 2$. Then the matrix $A = (\mathbf{c}'_1, \mathbf{c}'_2, \mathbf{c}'_3) \cdot (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)^{-1}$ is the coordinate matrix of κ . It turns out that the matrix A equals the matrix A used in Ex. 1.1.6. \diamond

Projective Coordinate Systems

A fundamental set also defines a *projective coordinate system* or *projective frame* in the following way: The proof of Prop. 1.1.6 shows that a fundamental set $(P_0, \dots, P_n; E)$ can always be written in the form

$$(\mathbf{b}_0\mathbb{R}, \dots, \mathbf{b}_n\mathbb{R}; \mathbf{e}\mathbb{R} = (\mathbf{b}_0 + \dots + \mathbf{b}_n)\mathbb{R}).$$

The points $\mathbf{b}_i\mathbb{R}$ are called *fundamental points* and $\mathbf{e}\mathbb{R}$ is called *unit point*. All vectors $\mathbf{p} \in \mathbb{R}^{n+1}$ can be written in the form

$$\mathbf{p} = x_0\mathbf{b}_0 + \dots + x_n\mathbf{b}_n. \quad (1.15)$$

The coefficients (x_0, \dots, x_n) are called *projective coordinates* of the point $\mathbf{p}\mathbb{R}$ with respect to the projective frame $(P_0, \dots, P_n; E)$. These coordinates are homogeneous, since $\lambda\mathbf{p}$ leads to coefficients $(\lambda x_0, \dots, \lambda x_n)$. Thus, we say that a point has, with respect to the given projective frame, the coordinates $(x_0 : \dots : x_n)$ or $(x_0, \dots, x_n)\mathbb{R}$.

The coordinates of the fundamental points and the unit point are

$$(1 : 0 : \dots : 0), \dots, (0 : \dots : 0 : 1); \quad (1 : 1 : \dots : 1).$$

The following is obvious:

Theorem 1.1.7. *Assume that $(P_0, \dots, P_n; E)$ is a projective frame, and that ϕ is a projective automorphism. If P has coordinates $(x_0 : \dots : x_n)$ with respect to this frame, then $P\phi$ has the same coordinates with respect to $(P_0\phi, \dots, P_n\phi; E\phi)$.*

A projective frame induces coordinates for hyperplanes in a natural way: The coordinates of a hyperplane are the coefficients of its linear equation.

If A is the matrix of the projective automorphism which transforms one projective frame to another ($\mathbf{b}'_i = A \cdot \mathbf{b}_i$), then point coordinates transform with A^{-1} : $(x'_0, \dots, x'_n) = A^{-1} \cdot (x_0, \dots, x_n)$. Hyperplane coordinates transform with A^T : $(u'_0, \dots, u'_n) = A^T \cdot (u_0, \dots, u_n)$.

Example 1.1.9. We prove Pappos' theorem (see Fig. 1.2 and Ex. 1.1.2): Let $S = [A_1 A_2 A_3] \cap [B_1 B_2 B_3]$ and consider the projective frame $(S, A_1, B_1; (A_1 \vee B_2) \cap (A_2 \vee B_1))$. Then the points $A_1, A_2, A_3, B_1, B_2, B_3$, have the homogeneous coordinates $(0 : 1 : 0), (1 : 1 : 0), (a : 1 : 0), (0 : 0 : 1), (1 : 0 : 1), (b : 0 : 1)$ (in that order) with real numbers a, b , and $(A_1 \vee B_2) \cap (A_2 \vee B_1) = (1 : 1 : 1)$, $(A_2 \vee B_3) \cap (A_3 \vee B_2) = (ab - 1 : b - 1 : a - 1)$, $(A_3 \vee B_1) \cap (A_1 \vee B_3) = (ab : b : a)$. These three points are immediately seen to be collinear. \diamond

Homogeneous Cartesian Coordinates

A Cartesian coordinate system with its origin $(0, \dots, 0)$, unit point $(1, \dots, 1)$ and axes defines a projective coordinate system $(P_0, P_1, \dots, P_n; E)$ in the following way: P_0 is the origin, P_1, \dots, P_n are the ideal points of the coordinate axes, and E is the unit point. If (x_1, \dots, x_n) are the Cartesian coordinates of a point P , its homogeneous coordinates with respect to the projective frame $(P_0, \dots, P_n; E)$ are $(1 : x_1 : \dots : x_n)$. Such a projective frame is called *homogeneous Cartesian coordinate system*.

Projective Mappings in E^n

We consider P^n as the projective extension of E^n , where ideal points are characterized by the condition $x_0 = 0$. A projective automorphism ϕ need not map ideal points to ideal points, so its restriction to E^n perhaps does not map all points. The inverse image of the ideal hyperplane is called the *vanishing hyperplane*.

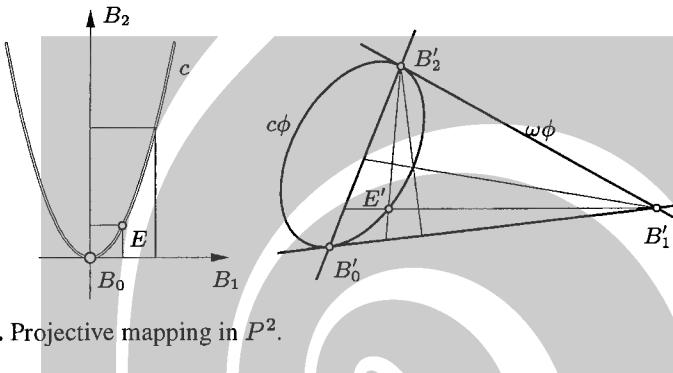


Fig. 1.5. Projective mapping in P^2 .

Example 1.1.10. Fig. 1.5 illustrates some properties of a projective mapping in the plane P^2 . The homogeneous Cartesian coordinate system $(B_0, B_1, B_2; E)$ is mapped onto a projective coordinate system $(B'_0, B'_1, B'_2; E')$. The image of the ideal line ω is the line $\omega\phi = B'_1 B'_2$. The metric scales on the coordinate axes B_0B_1 and B_0B_2 are mapped onto *projective scales* on $B'_0B'_1$ and $B'_0B'_2$.

Consider the parabola with $c : y = x^2$. In homogeneous coordinates (see (1.1)) c 's equation is

$$c : x_1^2 - x_0 x_2 = 0. \quad (1.16)$$

The image curve $c\phi$ possesses the same equation with respect to the frame $(B'_0, B'_1, B'_2; E')$. It is a conic whose tangent at B'_2 is the line $\omega\phi$. This shows that the curve c touches the ideal line in the ideal point B_2 (for the definition of tangency, see Sec. 1.1.5). \diamond

Fixed Points and Perspective Collineations

A projective map ϕ may have *fixed points* $X = X\phi$. The homogeneous coordinate vector \mathbf{x} of a fixed point is characterized by

$$A \cdot \mathbf{x} = \lambda \mathbf{x},$$

where A is ϕ 's coordinate matrix. This shows that \mathbf{x} is an eigenvector of the matrix A .

Proposition 1.1.8. *A projective automorphism κ which fixes a fundamental set is the identity. There is exactly one projective isomorphism which transforms a given fundamental set into another one.*

Proof. A regular linear mapping with $n + 2$ eigenvectors of which any $n + 1$ are linearly independent, must be a scalar multiple of the identity. This shows that κ is the identity mapping.

The existence of a projective isomorphism which maps a given fundamental set onto another one has already been shown by Prop. 1.1.6. It is unique, because for any two such isomorphisms, say κ_1, κ_2 , the mapping $\kappa_1\kappa_2^{-1}$ fixes a fundamental set and is therefore the identity mapping. This shows $\kappa_1 = \kappa_2$. \square

A nontrivial projective map can therefore have at most $n + 1$ projectively independent fixed points. An important special case is where all points of a hyperplane A are fixed. Then, we speak of a *perspective collineation* with *axis hyperplane* A .

Theorem 1.1.9. *A perspective collineation ϕ possesses a point C , called its center, with the property that $X, X\phi, C$ are collinear for all points $X \in P^n$.*

Proof. To show this, we assume that ϕ is not the identity mapping (the theorem is true in this case) and choose a projective coordinate system such that the axis hyperplane A is given by the equation $x_0 = 0$. If all vectors of the hyperplane $x_0 = 0$ are eigenvectors of a linear mapping, then this linear mapping must have the following coordinate representation:

$$\begin{aligned} x'_0 &= a_{00}x_0, \\ x'_1 &= a_{10}x_0 + ax_1, \\ &\dots \\ x'_n &= a_{n0}x_0 + ax_n. \end{aligned} \tag{1.17}$$

We introduce the vector $c := (a_{00} - a, a_{10}, \dots, a_{n0})$ and see that (1.17) can be written in the form

$$x' = x_0 c + a x,$$

which shows collinearity of the points $X = x\mathbb{R}$, $X\phi = x'\mathbb{R}$ and $C = c\mathbb{R}$. \square

Obviously the center C of a perspective collineation is a fixed point. If $a_{00} = a$, the C is incident with the axis A and we speak of an *elation* ϕ . Otherwise, the center is not contained in the axis and ϕ is called a *homology*.

Perspective Collineations of Euclidean Space

Example 1.1.11. Equ. (1.17) contains special cases of perspective collineations in the projective extension of Euclidean space. Assume that the axis A is the ideal hyperplane ω . Then the coordinate system used in (1.17) is a homogeneous Cartesian coordinate system. We see that an elation whose axis hyperplane is ω is a *translation*, and that a homology with this property is a *central similarity*. This is illustrated in Fig. 1.6. \diamond

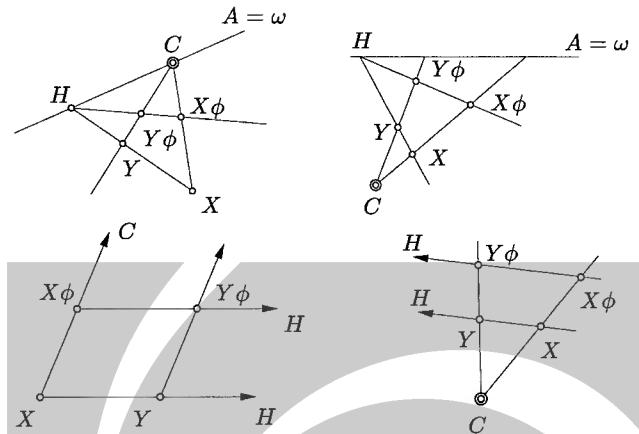


Fig. 1.6. Translation and central similarity as special perspective collineations.

Example 1.1.12. A perspective collineation ϕ is uniquely defined by its center C , its axis A , and an admissible pair of points $X, X\phi$. Admissible means that both X and $X\phi$ are not in A , are different from C , and that $X, X\phi, C$ are collinear. A geometric construction of the images of further points Y, Z is shown by Fig. 1.7. Both image $\omega\phi$ and preimage $\omega\phi^{-1}$ of the ideal line ω are constructed as well. Note that the axis' ideal point is fixed and thus $\omega\phi$ and $\omega\phi^{-1}$ must be parallel to A .

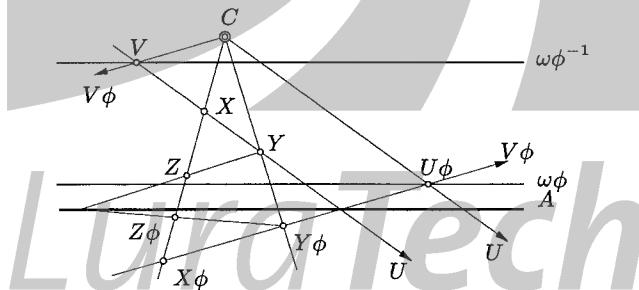


Fig. 1.7. Perspective collineation in P^2 .

◇

Projective Mappings of Subspaces

Definition. The restriction of a projective automorphism ϕ to a k -dimensional projective subspace Q^k is called a projective mapping of Q^k onto $Q^k\phi$.

Clearly $Q^k\phi$ is of the same dimension. We introduce a projective coordinate system $b_0\mathbb{R}, \dots, b_n\mathbb{R}, e\mathbb{R}$ in P^n such that $e = b_0 + \dots + b_n$ and $Q^k = b_0\mathbb{R} \vee \dots \vee b_k\mathbb{R}$,

and let $\bar{e}\mathbb{R} = (b_0 + \dots + b_k)\mathbb{R}$. Then $b_0\mathbb{R}, \dots, b_k\mathbb{R}, \bar{e}\mathbb{R}$ is a projective coordinate system in Q^k .

Its ϕ -image is a projective coordinate system of $Q^k\phi$ and ϕ maps points of Q^k to points of $Q^k\phi$ via equal coordinates with respect to these systems.

Lemma 1.1.10. *For any two fundamental sets $b_0\mathbb{R}, \dots, b_k\mathbb{R}, \bar{e}\mathbb{R}$, and $b'_0\mathbb{R}, \dots, b'_k\mathbb{R}, \bar{e}'\mathbb{R}$ of projective subspaces $Q^k, Q^{k'}$ of P^n there is exactly one projective mapping of Q^k onto $Q^{k'}$ which transforms the points of the first set onto the respective points of the second.*

Proof. Without loss of generality we can assume that $\bar{e} = b_0 + \dots + b_k$. Choose b_{k+1}, \dots, b_n such that b_0, \dots, b_n is a basis of \mathbb{R}^{n+1} . Do the same for the vectors $\bar{e}', b'_0, \dots, b'_k$. Then the projective automorphism which maps $b_0\mathbb{R}, \dots, b_n\mathbb{R}, (\sum b_i)\mathbb{R}$ to $b'_0\mathbb{R}, \dots, b'_n\mathbb{R}, (\sum b'_i)\mathbb{R}$ gives the desired projective mapping. Uniqueness is shown as in the case of P^n (cf. Prop. 1.1.8). \square

Example 1.1.13. Consider two planes $Q^2, Q^{2'}$ of projective three-space. In order to set up a collineation of Q^2 onto $Q^{2'}$ we need the images of a fundamental set. In this case it consists of four points with not three of them being collinear. Such a set is called a *quadrilateral*.

We assume that the projective automorphism ϕ which induces the projective mapping $Q^2 \rightarrow Q^{2'}$ is a perspective collineation with center C . Then $C, P, P\phi$ are collinear for all $P \in Q^2$, and the points of $Q^2 \cap Q^{2'}$ are fixed points of ϕ .

Conversely, a projective mapping $Q^2 \rightarrow Q^{2'}$ which has these two properties (it is then called a *perspectivity*) can always be extended to a perspective collineation of P^3 . \diamond

Example 1.1.14. The situation of Ex. 1.1.13 has the following interpretation: A planar figure and its image under a central projection are connected by a projective mapping.

This geometric situation occurs when we take a photograph. This is literally true only in the case of the ancient *camera obscura*, which does not have a lens but only a very small hole for the light rays to pass through. Modern cameras give good approximations, especially those used in surveying and photogrammetry.

Developing and printing the film is a Euclidean similarity, which is a projective mapping also. Thus a planar figure and a photograph are connected via a projective mapping of subspaces, and Lemma 1.1.10 shows that in order to reconstruct a planar figure we must know the position of four points in general position ('fitting points'). This fact is of course important in photogrammetry and computer vision (cf. Fig. 1.8). \diamond

Characterization of Projective Automorphisms

It is interesting to ask which properties of a bijection of projective space imply that this bijection is a projective automorphism, i.e., is described by a linear mapping of homogeneous coordinates. There is the following theorem:

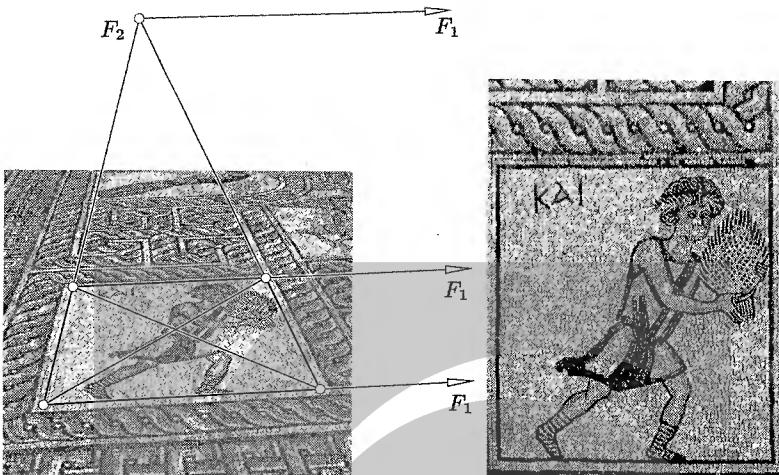


Fig. 1.8. Rectifying the photograph of a planar figure (courtesy of Institute of Photogrammetry and Remote Sensing, Vienna University of Technology).

Theorem 1.1.11. Assume that κ is a bijective mapping of P^n onto itself. If κ preserves collinearity of points, and there is a line L such that the restriction $\kappa|L$ is a projective transformation, then κ is a projective automorphism of P^n .

Proof. We fix a fundamental set (P_0, \dots, P_{n+1}) of P^n such that $L = P_0 \vee P_1$, and consider its image $(\kappa(P_0), \dots, \kappa(P_{n+1}))$, which is again a fundamental set. There is a projective automorphism λ with $\lambda(P_i) = \kappa(P_i)$. The mapping $\kappa\lambda^{-1}$ preserves collinearity and fixes the fundamental set (P_0, \dots, P_{n+1}) . If we can show that it is the identity mapping, then $\kappa = \lambda$ and we are done.

This is done by induction: If $n = 1$, then necessarily $P^n = L$ and Prop. 1.1.8 shows that $\kappa\lambda^{-1}$ is the identity. Now assume that the theorem has been proved for all projective spaces of dimension less than n .

If $U = P_0 \vee \dots \vee P_{n-1}$, then $\dim(U) = n - 1$. $\kappa\lambda^{-1}$ fixes $\{P_0, \dots, P_{n-1}\}$ and therefore leaves U invariant. Further $P'_n = (P_n \vee P_{n+1}) \cap U$ is a fixed point of $\kappa\lambda^{-1}$, and $(P_0, \dots, P_{n-1}, P'_n)$ is a fundamental set of U . Thus $\kappa\lambda^{-1}|U$ is the identity, by our assumption.

Now consider a point $X \notin (P_n \vee P_{n+1})$. The lines $X \vee P_n$ and $X \vee P_{n+1}$ both intersect U , and are therefore invariant under $\kappa\lambda^{-1}$, which shows that $\kappa(X) = X$. The only remaining points (those of $P_n \vee P_{n+1}$) are now easily shown to be fixed also, because all lines incident with these points contain at least two fixed points. \square

Remark 1.1.9. In real projective space P^n with $n \geq 2$ the last assumption in Th. 1.1.11 is actually not required. ★

We give a sketch of the proof: As the proof of Th. 1.1.11 shows, it is enough to show that a bijective mapping κ of P^2 onto itself, which preserves collinearity of points, and which fixes a fundamental set O, P_1, P_2, E , fixes all points of $O \vee$

P_1 . We use $(O, P_1, P_2; E)$ as a projective coordinate system. With respect to the corresponding affine coordinate system, we have the coordinates $O = (0, 0)$, $E = (1, 1)$, $O \vee P_1$ and $O \vee P_2$ are x_1 - and x_2 -axis.

We define the mapping ϕ by $\kappa(x, 0) = (\phi(x), 0)$. Obviously $\phi(0) = 0$ and $\phi(1) = 1$. $\kappa(P_1) = P_1$ shows that ϕ maps \mathbb{R} to \mathbb{R} . It is an elementary exercise to construct, from the points $(x_1, 0)$ and $(x_2, 0)$, the points $(x_1 + x_2, 0)$ and $(x_1 x_2, 0)$, using only the operations \vee and \cap applied to these two points and the points O, E, P_1, P_2 .

This construction is transformed by κ such that the base points are fixed, which shows that $\phi(x_1 + x_2) = \phi(x_1) + \phi(x_2)$ and $\phi(x_1 x_2) = \phi(x_1)\phi(x_2)$. It is well known that such a mapping is the identity mapping, which is a special property of the real number field. \diamond

1.1.4 Projectivities, Cross Ratio and Harmonic Position

Projectivities

The restriction of a projective mapping in P^n to a line L is called a *projectivity* of L onto $L' = L\phi$. This is a special case of a projective mapping of subspaces. By Lemma 1.1.10 it is uniquely defined by the images of three distinct points of L .

If we write these points as $b_0\mathbb{R}, b_1\mathbb{R}, e\mathbb{R} = (b_0 + b_1)\mathbb{R}$, and their images as $b'_0\mathbb{R}, b'_1\mathbb{R}, e'\mathbb{R} = (b'_0 + b'_1)\mathbb{R}$, then ϕ acts via equal coordinates with respect to these two projective frames:

$$\phi : x\mathbb{R} = (x_0 b_0 + x_1 b_1)\mathbb{R} \mapsto x'\mathbb{R} = (x_0 b'_0 + x_1 b'_1)\mathbb{R}.$$

(x_0, x_1) are projective coordinates on L . The ratio $x_1 : x_0$ sets up a *projective scale* on L . Points $B_0 = b_0\mathbb{R}, B_1 = b_1\mathbb{R}, E = e\mathbb{R}$ have, in this order, the scale values $\infty (= 1 : 0)$, 0 and 1, and are called *end point*, *origin* and *unit point* of the projective scale. This shows that a projectivity of a range of points onto another may also be seen as a transformation via equal values in projective scales on two lines L, L' .

The Cross Ratio of Four Points

It is possible to assign a numerical projective invariant to *four* collinear points A, \dots, D : Consider the projective scale with A as end point, B as origin and C as unit point; the scale value of D then is called the *cross ratio* $\text{cr}(A, B, C, D)$ of the four points. If

$$A = a\mathbb{R}, B = b\mathbb{R}, C = (a + b)\mathbb{R}, D = (x_1 a + x_0 b)\mathbb{R},$$

then $\text{cr}(A, B, C, D) = \frac{x_1}{x_0}.$ (1.18)

Clearly, projective maps preserve the cross ratio of collinear points.

Proposition 1.1.12. *If the four points A, B, C, D of the line L have coordinates $(a_0 : a_1), \dots, (d_0 : d_1)$ in some projective coordinate system of L , their cross ratio is given by*

$$\text{cr}(A, B, C, D) = \frac{\begin{vmatrix} a_0 & c_0 \\ a_1 & c_1 \end{vmatrix} \cdot \begin{vmatrix} b_0 & d_0 \\ b_1 & d_1 \end{vmatrix}}{\begin{vmatrix} b_0 & c_0 \\ b_1 & c_1 \end{vmatrix} \cdot \begin{vmatrix} a_0 & d_0 \\ a_1 & d_1 \end{vmatrix}}. \quad (1.19)$$

Proof. A change of coordinate system in L is expressed by multiplication of homogeneous coordinate vectors with some regular 2×2 matrix M . Because of $\det(M \cdot N) = \det(M) \cdot \det(N)$ for all matrices M, N , the numerical value defined by Equ. (1.19) is not affected by coordinate changes. If we choose the projective coordinate system A, B, C , we immediately see that both Equ. (1.18) and Equ. (1.19) give the same result. \square

From this it is easy to see how certain permutations of the arguments change the value of the cross ratio:

Corollary 1.1.13. *If A, B, C, D are collinear points, then*

$$\begin{aligned} \text{cr}(A, B, C, D) &= \text{cr}(B, A, D, C) = \text{cr}(C, D, A, B) = \text{cr}(D, C, B, A) \\ &= 1 / \text{cr}(A, B, D, C) = 1 - \text{cr}(A, C, B, D). \end{aligned} \quad (1.20)$$

If the projective frame $(b_0\mathbb{R}, b_1\mathbb{R}; e\mathbb{R})$ on L is such that $b_1\mathbb{R}$ is an ideal point, then the projective scale becomes a *metric scale* with b_0 as origin and $e\mathbb{R}$ as unit point. The value a in this metric scale corresponds to coordinates $a_0 = 1, a_1 = a$ in the projective coordinate system given by $(b_0\mathbb{R}, b_1\mathbb{R}; e\mathbb{R})$. The cross ratio becomes

$$\text{cr}(A, B, C, D) = \frac{\begin{vmatrix} 1 & 1 \\ a & c \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 \\ b & d \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ b & c \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 \\ a & d \end{vmatrix}} = \frac{c-a}{c-b} \cdot \frac{d-b}{d-a} = \frac{\text{ratio}(A, B, C)}{\text{ratio}(A, B, D)}, \quad (1.21)$$

where the symbol ‘ $\text{ratio}(A, B, C)$ ’ means the quotient $(c-a)/(c-b)$. This quotient of ratios explains the name *cross ratio*. There is another relation between the affine ratio and the projective cross ratio: Consider (A, \dots, D_ω) with D_ω at infinity. Now, $\text{cr}(A, B, C, D_\omega) = \text{cr}(D_\omega, C, B, A)$ is the value of A in a projective scale with D_ω as end point, C as origin and B as unit point. This is a metric scale and clearly the scale value of A equals $\text{ratio}(A, B, C)$. Thus, for D_ω at infinity,

$$\text{cr}(A, B, C, D_\omega) = \text{ratio}(A, B, C). \quad (1.22)$$

Let us return to the concept of projectivities. We consider two lines L, L' in the projective plane, and define a special projectivity of L onto L' by

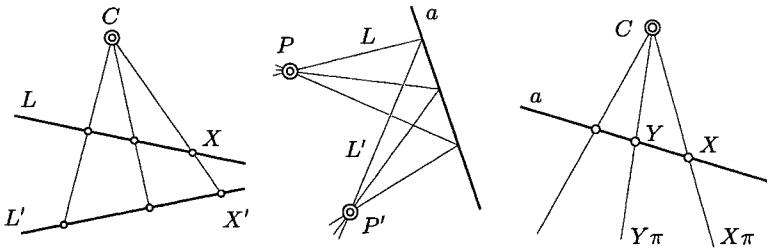


Fig. 1.9. Perspectivities of ranges of points and of line pencils.

$$X \in L \mapsto (C \vee X) \cap L',$$

where the point C is not in L, L' . This mapping is called a *perspectivity* and it equals the restriction of an appropriate perspective collineation to L . (cf. Ex. 1.1.13 and Fig. 1.9, left). The point C is called the *center* of the perspectivity. We say that the ranges of points L, L' are ‘perspective’, which means that we have defined a perspectivity $L \rightarrow L'$.

There is also a dual version of this: Consider two pencils of lines with vertices P, P' . A perspectivity of the first pencil onto the second is defined by

$$L \mapsto (L \cap a) \vee P',$$

where L is contained in the first pencil, i.e., contains P , and a is the *axis* of the perspectivity (see Fig. 1.9, center).

In line coordinates, a perspectivity of line pencils is expressed by a linear mapping, because a perspectivity of ranges of points is.

The third type of mapping called a perspectivity is the following: A point X contained in a line a is mapped to the line $X\pi = X \vee C$. An appropriate projective coordinate system reveals that this mapping in homogeneous coordinates is again given by a linear mapping. Its inverse is called a perspectivity as well (see Fig. 1.9, right).

The following theorem is useful:

Theorem 1.1.14. *All projectivities can be decomposed into a finite product of perspectivities.*

Proof. We have to show the theorem only for projectivities of ranges of points. We already know that such a mapping is uniquely determined by the images of three distinct points. So if we can find a product of perspectivities which maps three points in exactly the same way as the given projectivity, we are done. Such a construction can be seen in Fig. 1.10. \square

This allows the following

Definition. *Any composition of perspectivities is called a projectivity.*

Because of Th. 1.1.14, this is consistent with the previous use of ‘projectivity’. It actually extends the previous definition, which was restricted to the case that domain and range was of the same type.

Example 1.1.15. A projectivity of the line \bar{L} onto the line L' , which fixes the intersection point $\bar{L} \cap L'$, is always a perspectivity (see Fig. 1.10). The proof of this fact is left to the reader as an exercise. \diamond

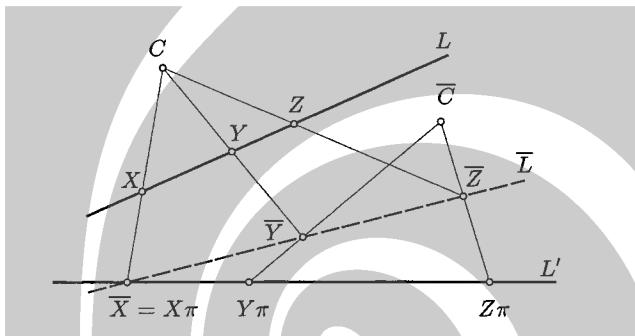


Fig. 1.10. Decomposing a projectivity into two perspectivities.

Projective Automorphisms of a Line

Let us now focus on the case $L = L'$. A projectivity π of L onto itself, i.e., a projective automorphism of L , is written as

$$\begin{bmatrix} x'_0 \\ x'_1 \end{bmatrix} = A \cdot \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}, \quad (1.23)$$

where x_0, x_1 are homogeneous coordinates with respect to a projective frame of L .

Lemma 1.1.15. A projective automorphism ϕ of a line L has 0, 1, or 2 fixed points, if it is not the identity mapping.

Proof. A point $x\mathbb{R}$ is a fixed point of ϕ if and only if x is an eigenvector of the matrix A of Eqn. (1.23). If ϕ is not the identity mapping (i.e., A is not a scalar multiple of the unit matrix), then there are 0, 1 or 2 one-dimensional eigenspaces. \square

Definition. A projective automorphism of a line is called elliptic, parabolic, or hyperbolic, if it has 0, 1 or 2 fixed points, respectively.

In the complex projective plane there are no elliptic projectivities, because there is always an eigenvector of A . If the matrix A is real and has no real eigenvectors, then there are two conjugate complex ones. We may say that there are two conjugate complex fixed points of A . This will be discussed in Sec. 1.1.6.

Involutions and Harmonic Points

An involutory projective automorphism of a range of points, or of a pencil of lines, is called an *involution*. This means that it is its own inverse, but it does not equal the identity mapping.

Lemma 1.1.16. *A non-identical projective automorphism is an involution if and only if its coordinate matrix satisfies*

$$\text{trace}(A) = a_{00} + a_{11} = 0. \quad (1.24)$$

Proof. The matrix A is no multiple of the unit matrix I , but its square is. This condition is easily seen to be equivalent to Equ. (1.24). \square

Clearly an involution π is uniquely determined by the images $A\pi, B\pi$ of *two* points A, B (if $A\pi \neq B$), because automatically $A\pi \mapsto A$ serves as a third pair of points. On the other hand, with $B\pi \mapsto B$ we already have the images of four points, and the involution may seem over-determined. A closer look at (1.24) reveals that actually for all admissible $A, B, A\pi, B\pi$ there is a unique involution π .

Furthermore, (1.24) implies that there is no parabolic involution, so we either have two real fixed points $E = E\pi, F = F\pi$ or a pair of conjugate imaginary fixed points.

Example 1.1.16. Consider a line L in the projective extension of the Euclidean plane, and a hyperbolic involution π of L . Assume that the ideal point F_ω of L is fixed: $F_\omega\pi = F_\omega$. The other fixed point E then is a proper Euclidean point. The Euclidean reflection in E obviously is an involutory projectivity, and therefore coincides with π . Further it is easily verified that for all X we have

$$-1 = \text{ratio}(X, X\pi, E) = \text{cr}(X, X\pi, E, F_\omega),$$

where the latter equality is given by (1.22). \diamond

Theorem 1.1.17. *If π is an involution with fixed points E, F , then for all points $X \in E \vee F$ we have*

$$\text{cr}(X, X\pi, E, F) = -1. \quad (1.25)$$

Proof. We transform the entire situation by a suitable collineation κ (or simply a perspectivity like the one in Fig. 1.11) such that F is transformed into an ideal point $F\kappa = F_\omega$. Cross ratios are not affected by κ , and by Ex. 1.1.16, we have $\text{cr}(X, X\pi, E, F) = \text{cr}(X\kappa, X\pi\kappa, E\kappa, F\kappa) = -1$. \square

We say that a pair $X, X\pi$ with property (1.25) is in *harmonic position* with respect to the points E, F . This is a generalization of Euclidean symmetry to projective geometry. When we use the complex extension, (1.25) is valid also for an elliptic involution.

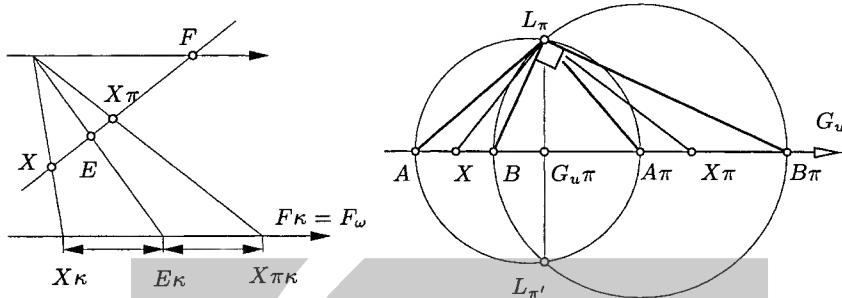


Fig. 1.11. Mapping involutions to Euclidean special cases.

Definition. Four collinear points A, \dots, D are in harmonic position if and only if their cross ratio equals -1 :

$$H(A, B; C, D) \iff \text{cr}(A, B, C, D) = -1. \quad (1.26)$$

We use the notation $H(A, B; C, D)$ to indicate harmonic position.

If A, B, C, D are in harmonic position, then also the following cross ratios have the value -1 :

$$\text{cr}(A, B, C, D) = \text{cr}(B, A, C, D) = \text{cr}(A, B, D, C) = \text{cr}(C, D, A, B) = -1.$$

The equality of these various cross ratios follows immediately from (1.20), and shows that the property of being harmonic remains invariant if we interchange A and B or C and D , or the pairs A, B and C, D . Thus, harmonic position is a symmetric relation between unordered pairs of collinear points. Note that harmonic pairs always *interlace* in the closed projective line.

The object dual to a range of points in P^n is a pencil of hyperplanes, which therefore has involutions as well. The following two examples illustrate the case $n = 2$:

Example 1.1.17. We obtain a Euclidean special case of a hyperbolic involution π in a line pencil, if the two fixed lines E, F are orthogonal. Then, π is the *reflection* in E (or F), since this reflection is an involutory projective automorphism, and acts on E, F in the same way as π does.

E, F are the bisector lines of all pairs $L, L\pi$. This shows that harmonic position of two pairs of concurrent lines, one of which is orthogonal, is characterized by the orthogonal pair being the angular bisector of the other. \diamond

Example 1.1.18. Consider the elliptic involution ι in a pencil of lines which maps a line L to the line $L\iota$ orthogonal to it. It is easily verified that this mapping is indeed an involution. It is called the *right angle involution*. It can be used to handle elliptic involutions π in a range g of points constructively. (see Fig. 1.11):

Take two pairs $A, A\pi$ and $B, B\pi$. The Thales circles with diameters $A, A\pi$ and $B, B\pi$ intersect in two points L_π, L'_π (the *Laguerre points* of π).

Now we ‘project π from L_π ’, which means that we define a new involution which maps a line x incident with L_π to $(x \cap g)\pi \vee L_\pi$. This mapping coincides with the right angle involution in the pencil L_π , because (i) it is an involution, and (ii) it maps two different lines like the right angle involution.

This can be used to find the image of further points by a simple geometric construction. In particular, we can construct the image $G_u\pi$ of g ’s ideal point, which is called the *central point* of π . \diamond

Projective Reflections

Suppose κ is an *involutory perspective collineation* in P^n , i.e., κ^2 is the identity mapping, but κ is not (as an example you may think of the reflection in a hyperplane). Its restriction to any line containing the collineation’s center C is a projectivity $X \mapsto X\kappa$. The fixed points of this projectivity are the center C and the line’s intersection D with the axis hyperplane. κ is involutory, so this projectivity must be an involution, with fixed points C, D , which shows that $X, X\kappa, C, D$ are in harmonic position for all X collinear with C, D . The mapping κ is called a *harmonic homology* or *projective reflection* (its being a homology, and not an elation, is clear because $C \neq D$ implies that the center is not incident with the axis hyperplane).

Lemma 1.1.18. *If the center C and the axis hyperplane A of a projective reflection κ have coordinates $(1, 0, \dots, 0)\mathbb{R}$ and $\mathbb{R}(1, 0, \dots, 0)$, respectively, then $\kappa : x\mathbb{R} \mapsto x'\mathbb{R}$ has, with respect to this coordinate system, the equation*

$$x'_0 = x_0, x'_1 = -x_1, \dots, x'_n = -x_n. \quad (1.27)$$

Proof. We let $a_{10} = \dots = a_{n0} = 0$ and $a = -a_{00}$ in Equ. (1.17). Because of $a \neq 0$ and homogeneity we may without loss of generality let $a_{00} = 1$. \square

In the projective extension of Euclidean space there are the following special cases of harmonic homologies: If the axis is the hyperplane at infinity, Equ. (1.27) shows that κ is the reflection in its center.

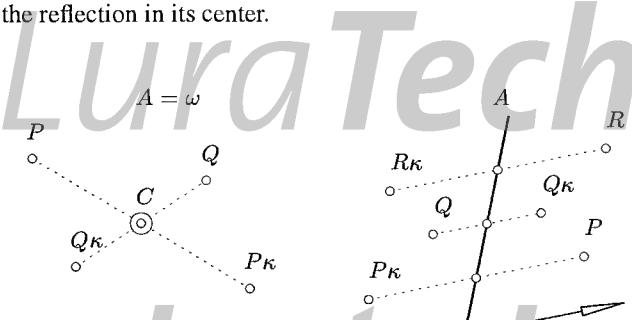


Fig. 1.12. Projective reflections whose axes or centers are ideal.

If C is an ideal point, κ is, in general, an *affine reflection* in A , a special case of which is the Euclidean reflection in A . It is remarkable that reflection in a point and reflection in a hyperplane are the same from the projective point of view: both are harmonic homologies, with either the center or the axis at infinity. That is why we call this type of projective automorphism a projective reflection (see Fig. 1.12).

Remark 1.1.10. In P^2 , harmonic homologies are the only involutory collineations. In P^3 this is no longer true: Consider the projective extension of E^3 and a reflection in a line A , which is the same as the rotation by an angle π about the same line. The projective extension κ of this mapping is an involutory collineation. The set of its fixed points comprises two lines, namely A and the line at infinity \bar{A} orthogonal to A .

If we apply any projective automorphism to this situation, the reflection in the line is transformed to an involutory collineation which fixes the points of two skew (i.e., non-intersecting) lines L, \bar{L} .

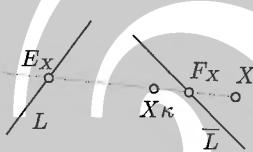


Fig. 1.13. Skew involution.

For all points X not in L, \bar{L} , there is a unique line L_X which intersects both L, \bar{L} . Let $E_X = L \cap L_X, F_X = \bar{L} \cap L_X$. Then $L_X = E_X \vee F_X, E_X\kappa = E_X, F_X\kappa = F_X$, so L_X is left invariant by κ . It is clear that the restriction of κ to $E_X \vee F_X$ is an involution, whose fixed points are precisely E_X, F_X . We therefore have $H(X, X\kappa; E_X, F_X)$. This type of involutory collineation is called *skew involution* (see Fig. 1.13). \diamond

Singular Projective Mappings

We have defined projective mappings by Equ. (1.12), where we used a regular matrix A . It is often useful to consider singular linear mappings, whose domain is a projective space of dimension n and whose image space has a different dimension.

Definition. A mapping λ from an n -dimensional projective space onto an m -dimensional projective space is called *linear*, if it is described by the equation $(x\mathbb{R})\lambda = (A \cdot x)\mathbb{R}$, with an $(m+1) \times (n+1)$ -matrix A .

If the matrix A is quadratic and regular, a linear mapping becomes a projective isomorphism. *Singular projective mapping* means a linear mapping which is no projective isomorphism. It turns out that such mappings are generalizations of the familiar concept of central projection from projective three-space onto a plane. We first define a *central projection* from P^n onto a subspace U via a *center* Z by

$$P\pi = (Z \vee P) \cap U. \quad (1.28)$$

We require that Z and U are complementary subspaces. Then Th. 1.1.2 shows that $P\pi$ is actually a well-defined point if $P \notin Z$, and $P\pi$ is the only point of U such that P , $P\pi$, and some point of Z are collinear. Precisely the points $Q \in P \vee Z$ have the property that $P\pi = Q\pi$. If $P \in Z$, then $P\pi$ is undefined or void, depending on interpretation. Formally, if Z is the empty set and $U = P^n$, then the projection is the identity mapping.

There is the following lemma which establishes a connection between central projections and linear mappings:

Lemma 1.1.19. *For all linear mappings $\lambda : P^n \rightarrow P^m$ there is a central projection π from P^n onto a subspace U and a projective isomorphism α of U onto P^m such that $\lambda = \pi\alpha$. All such mappings are singular linear mappings.*

Proof. We assume that $(x\mathbb{R})\lambda = (A \cdot x)\mathbb{R}$. We let Z equal the kernel of the linear mapping $x \mapsto A \cdot x$ and choose a linear subspace U of \mathbb{R}^{n+1} which is complementary to Z . Obviously all vectors $x \in \mathbb{R}^{n+1}$ can be written in the form $x = x' + x''$ with $x' \in Z$ and $x'' \in U$. If $x \notin Z$, then $x'' \neq 0$. As $x\mathbb{R}$, $x'\mathbb{R}$, and $x''\mathbb{R}$ are collinear points, the mapping $x\mathbb{R} \mapsto x''\mathbb{R}$ indeed equals the central projection π with center Z and image subspace U .

The linear mapping $x'' \mapsto A \cdot x''$ from U to \mathbb{R}^{m+1} is regular, since its kernel $U \cap Z$ is the zero subspace. Furthermore, $A \cdot x = A \cdot x''$, so the corresponding mapping $\alpha : x''\mathbb{R} \mapsto (A \cdot x'')\mathbb{R}$ of projective points is a projective isomorphism, and $\lambda = \pi\alpha$.

To show the converse, we have to show that a central projection π with center Z and image subspace U is a singular projective mapping (α is even a regular one). We introduce a projective coordinate system $(B_0, \dots, B_n; E)$ such that B_0, \dots, B_m span U and the remaining points span Z . We use $(B_0, \dots, B_m, \underbrace{1, \dots, 1}_{m+1}, 0, \dots,$

$0)\mathbb{R}$ as a projective coordinate system in U . Obviously the linear mapping $\pi' : (x_0, \dots, x_n)\mathbb{R} \mapsto (x_0, \dots, x_m)\mathbb{R}$ equals π : If $X = (x_0, \dots, x_n)\mathbb{R}$, the line $X \vee X\pi'$ meets Z in the point $(0, \dots, 0, x_{m+1}, \dots, x_n)\mathbb{R}$. \square

It is an immediate consequence of Lemma 1.1.19 that a linear mapping has a *kernel* (or *center* or *exceptional subspace*) Z , whose points have no image, and which is independent of the decomposition of λ into a projection π and a projective isomorphism. Precisely the points $Q \in P \vee Z$ have the property that $P\pi = Q\pi$.

1.1.5 Polarities and Quadrics

Having studied linear objects in projective space, we are now going to have a closer look at *quadratic* ones. This is done by first considering some further projective mappings.

Correlations

We have already defined projective mappings of one projective space onto another — such mappings are defined by a linear transformation of homogeneous coordinates. In this section we consider projective mappings of P^n onto its dual P^{n*} :

Definition. A projective mapping κ of P^n onto its dual P^{n*} is called a projective correlation. For $P \in P^n$, the points contained in the hyperplane $P\kappa$ are called conjugate to P .

Thus a projective correlation maps points to hyperplanes. In homogeneous coordinates, a projective correlation reads

$$(x\mathbb{R})\kappa = \mathbb{R}(C \cdot x), \quad \text{with } C \in \mathbb{R}^{(n+1) \times (n+1)}, \quad \det C \neq 0. \quad (1.29)$$

The point $y\mathbb{R}$ is conjugate to $x\mathbb{R}$ if and only if

$$y^T \cdot C \cdot x = 0. \quad (1.30)$$

The Duality Associated with a Correlation

Consider a projective correlation κ and the image $Q\kappa$ of a projective subspace Q . $Q\kappa$ is a set of hyperplanes. More precisely, it is a projective subspace of P^{n*} , i.e., a linear manifold of hyperplanes. If Q is defined by linear equations

$$Q : a_0^T \cdot x = \dots = a_k^T \cdot x = 0,$$

and κ is given by Equ. (1.29), then $Q\kappa$ is defined by the equations

$$Q\kappa : (a_0^T \cdot C^{-1}) \cdot u = \dots = (a_k^T \cdot C^{-1}) \cdot u = 0.$$

The intersection of all hyperplanes of $Q\kappa$ is a subspace of P^n . We call it $Q\delta$. Obviously its annihilator $Q\delta^\circ$ coincides with the set $Q\kappa$ of hyperplanes. The points of $Q\delta$ are conjugate to all points of Q .

Definition. If Q is a subspace and κ a correlation, there is a subspace $Q\delta$ whose points are conjugate to the points of Q . The mapping $Q \mapsto Q\delta$ is called the duality associated with κ .

Proposition 1.1.20. The duality associated with a correlation maps k -dimensional projective subspaces to $(n - k - 1)$ -dimensional ones, and has the following properties:

$$(G \cap H)\delta = G\delta \vee H\delta, \quad (G \vee H)\delta = G\delta \cap H\delta, \quad G \subset H \iff G\delta \supset H\delta.$$

If U_i is an indexed family of subspaces, then

$$(\bigcap U_i)\delta = \bigvee U_i\delta, \quad (\bigvee U_i)\delta = \bigcap U_i\delta.$$

Proof. This follows immediately from Prop. 1.1.3 and the relation $Q\delta^\circ = Q\kappa$. \square

We see that a duality δ is an explicit realization of the duality principle within the same projective space: Subspaces are mapped to subspaces of complementary dimension, and inclusion and span are interchanged.

The Adjoint Correlation

In particular a duality maps hyperplanes to points. Assume a hyperplane has coordinates $\mathbb{R}\mathbf{v}$. Its points $\mathbf{x}\mathbb{R}$ are those which satisfy $\mathbf{v}^T \mathbf{x} = 0$. If the correlation κ is given by Equ. (1.29), then the image hyperplanes $\mathbb{R}\mathbf{u}$ of $\mathbf{x}\mathbb{R}$ satisfy $\mathbf{v}^T C^{-1} \mathbf{u} = 0$, and all of them contain the point $\mathbf{y}\mathbb{R}$, where $\mathbf{y} = (\mathbf{v}^T \cdot C^{-1})^T = C^{-T} \cdot \mathbf{v}$. This shows the following

Lemma 1.1.21. *The duality associated with the correlation (1.29) maps hyperplanes to points according to $\mathbb{R}\mathbf{v} \mapsto (C^{-T} \cdot \mathbf{v})\mathbb{R}$.*

Definition. *If κ is the correlation given by (1.29), then the correlation $\widehat{\kappa}$ defined by*

$$\widehat{\kappa} : \mathbf{y}\mathbb{R} \mapsto \mathbb{R}(C^T \cdot \mathbf{y}) \quad (1.31)$$

is called adjoint to κ .

The relation between a correlation κ and its adjoint correlation $\widehat{\kappa}$ is, according to Lemma 1.1.21, the following: κ maps a point P to a hyperplane $P\kappa$ of points conjugate to P . But what is the set of points X such that P is conjugate to X ? It is the hyperplane $P\widehat{\kappa}$.

Particularly interesting is the case of a *self-adjoint correlation* $\kappa = \widehat{\kappa}$. Clearly a correlation is self-adjoint if and only if conjugacy with respect to this correlation is a *symmetric relation*, which means that $P \in Q\kappa$ is the same as $Q \in P\kappa$, for all $P, Q \in P^n$. The associated duality then has the property that $\delta\delta = \text{id}$.

Lemma 1.1.22. *The matrix C of Equ. (1.29) belongs to a self-adjoint correlation if $C = \pm C^T$, and vice versa.*

Proof. Equations (1.29) and (1.31) show that the correlation κ equals its adjoint $\widehat{\kappa}$ if and only if $C = \lambda C^T$. This implies $C = \lambda^2 C$. Now C is not the zero matrix and the only solutions of $1 = \lambda^2$ are $\lambda = 1, -1$. \square

Null Polarities

Let us discuss the case of a skew-symmetric matrix C in Equ. (1.29). Because $\det C = \det(C^T) = \det(-C) = (-1)^{n+1} \det C$, an even n implies $\det C = 0$, and such a matrix does not correspond to a correlation.

Therefore we study the skew-symmetric case in projective spaces of odd dimension only. If C is a regular skew-symmetric matrix, then for all \mathbf{x} ,

$$\mathbf{x}^T \cdot C \cdot \mathbf{x} = \sum_{i,k=0}^n c_{ik} x_i x_k = \sum_{i \leq k} (c_{ik} - c_{ki}) x_i x_k = 0. \quad (1.32)$$

This means that $\mathbf{x}\mathbb{R}$ is contained in the hyperplane $\mathbb{R}(C \cdot \mathbf{x})$, and is therefore conjugate to itself, for all points $\mathbf{x}\mathbb{R}$. In this case the correlation κ defined by the matrix C is called a *null polarity*. Null polarities of P^3 play a fundamental role in line geometry and will be discussed later in more detail (see Sec. 3.1.1).

Polarities and Quadrics

If C of Equ. (1.29) is a symmetric matrix, then the correlation κ defined by C is called *polarity*. A point P and its image $P\kappa$ are called *pole* and *polar hyperplane* of each other. $P\kappa$ contains all points conjugate to P .

We extend the definition of conjugacy to hyperplanes: Hyperplanes $P\kappa, Q\kappa$ are conjugate if and only if P, Q are. This is equivalent to $P\kappa$'s containing $Q\kappa$'s pole and vice versa.

Definition. If κ is a polarity, the set of self-conjugate points is called a *quadric*, if it is not empty. κ is then called a *hyperbolic* polarity, otherwise it is called *elliptic*. A quadric of the projective plane is called a *conic*.

The points $x\mathbb{R}$ of the quadric Φ defined by κ are those which satisfy (by Equ. (1.30))

$$x^T \cdot C \cdot x = 0, \quad (1.33)$$

where C is the symmetric coordinate matrix of κ . This equation need not have a real solution. To ‘see’ a quadric also in the case of an elliptic polarity we have to use the complex extension of real projective space, where all quadratic equations have a solution.

Remark 1.1.11. We have defined a quadric by its polarity. Conversely, we may start with a quadric as a set Φ and ask for a polarity. Clearly, if Φ has an equation of the form (1.33) with a symmetric matrix C , then C defines a polarity whose quadric coincides with Φ . It can be shown that the quadric as a point set determines the associated quadratic equation uniquely and therefore also the matrix C , up to a scalar factor. We therefore speak of poles and polar hyperplanes with respect to quadrics as well. ◇

Symmetric Bilinear Forms and Quadratic Forms

The previous discussion can be re-formulated with a slightly different vocabulary: Consider a symmetric bilinear form

$$(x, y) \mapsto \Omega(x, y) = x^T \cdot C \cdot y, \quad (1.34)$$

which is defined by a symmetric matrix C . The corresponding *quadratic form*

$$x \mapsto \Omega_q(x) = x^T \cdot C \cdot x \quad (1.35)$$

has the property that $\Omega_q(x) = \Omega(x, x)$. Clearly the equation $\Omega_q(x) = 0$ defines a quadric, and the polar hyperplane of a point $p\mathbb{R}$ has the equation $\Omega(p, x) = 0$. Conjugacy of points $x\mathbb{R}, y\mathbb{R}$ is expressed by the relation $\Omega(x, y) = 0$.

Projective Reflections Leaving a Quadric Invariant

In projective geometry it is important to know the projective automorphisms of objects. We are going to describe a family of projective reflections which leave a quadric invariant:

Theorem 1.1.23. *Assume that κ is a polarity and Φ is its quadric. Then the projective reflection σ with center P and axis hyperplane $P\kappa$ leaves Φ invariant for all points P not in Φ . σ commutes with κ , which means that $\kappa\sigma = \sigma\kappa$.*

Proof. We introduce a projective frame such that $P = (1 : 0 : \dots : 0)$ and $P\kappa$ has the equation $x_0 = 0$. Computing $P\kappa$ by multiplying $C \cdot (1, 0, \dots, 0)^T$ shows that $(c_{00}, \dots, c_{0n}) = \lambda(1, 0, \dots, 0)$. By symmetry, also $(c_{00}, c_{10}, \dots, c_{n0}) = (\lambda, 0, \dots, 0)$.

The projective reflection σ has the coordinate matrix $A = \text{diag}(1, -1, \dots, -1)$, according to Lemma 1.1.18. We compute $(x\mathbb{R})\kappa\sigma = (\mathbb{R}C \cdot x)\sigma = \mathbb{R}(A^{-T}Cx)$ and $(x\mathbb{R})\sigma\kappa = \mathbb{R}(CA \cdot x)$, which is the same. Φ is invariant because $X \in \Phi$, which means $X \in X\kappa$, implies $X\sigma \in X\kappa\sigma = (X\sigma)\kappa$, so $X\sigma \in \Phi$ as well. \square

We may say that any polar pair (i.e., pole and polar hyperplane) defines a *projective symmetry* of Φ . A projective reflection σ with center P has the property that $X, X\sigma, P$ are collinear. So σ maps all points $X \in \Phi$ to the other intersection point of $X \vee P$ with Φ , which is unique (see Fig. 1.14).

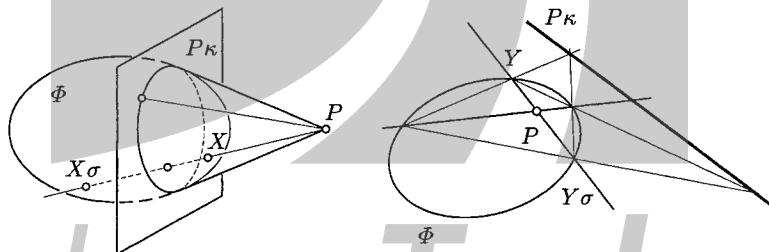


Fig. 1.14. Pole and polar plane with respect to a quadric in P^3 (left) and in P^2 (right).

Polarities Induced in Hyperplanes

We want to compute the intersection of a quadric with a hyperplane.

Proposition 1.1.24. *If P is not contained in the quadric Φ , the intersection $\Phi \cap P\kappa$ is empty or a quadric of $P\kappa$.*

Proof. We use the coordinate system used in the proof of Th. 1.1.23. The intersection of Φ with $x_0 = 0$ has the equation

$$\Phi_0 : \bar{x}^T \cdot C_0 \cdot \bar{x} = x_0 = 0, \quad \bar{x} = (x_1, \dots, x_n) \quad (1.36)$$

where C_0 is the $(n \times n)$ -submatrix of C , which is obtained from C by deleting the first row and column. Its rank is n , because $\text{rk}(C) = n + 1$, and all c_{0i} and c_{i0} are zero except c_{00} . \square

Definition. Consider a polarity κ , its quadric Φ and a hyperplane H whose pole is not contained in Φ . The polarity induced in H by Φ maps $Q \in H$ to $Q\kappa \cap H$.

Proposition 1.1.25. Under the assumptions of the previous definition, the polarity induced in a hyperplane H indeed is a polarity, its quadric is the intersection $\Phi \cap H$, and its coordinate matrix is the matrix C_0 of Equ. (1.36), if H has equation $x_0 = 0$.

Proof. We assume a coordinate system such that H has equation $x_0 = 0$. Without loss of generality we may even assume the coordinate system used in the proof of Th. 1.1.23. The matrix C_0 of (1.36) defines a polarity κ_H in H , because $\text{rk}(C_0) = n$. In order to show that $Q\kappa_H = Q\kappa \cap H$ for all $Q \in H$ we have to show that the points which are κ_H -conjugate to Q are also κ -conjugate to Q . This is clear from the definition of C_0 . The rest follows immediately from Prop. 1.1.24. \square

Quadratic Cones

If Φ is a quadric in an $(n - 1)$ -dimensional projective subspace U and V is a point not contained in U , then the union Φ_V of the lines $X \vee V$ where X runs in Φ is called the *quadratic cone* with vertex V and base quadric Φ (see Fig. 1.15).

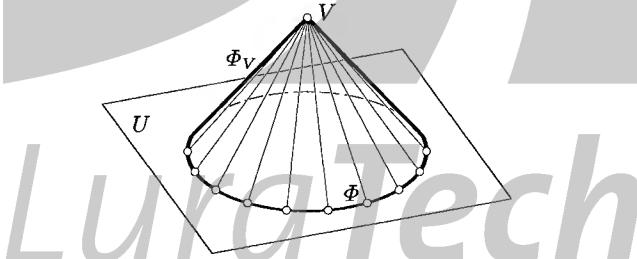


Fig. 1.15. Quadratic cone Φ_V with base quadric $\Phi \subset U$ and vertex V .

More generally, if V is a k -dimensional subspace which does not intersect an $(n - k - 1)$ -dimensional subspace U , then the union Φ_V of all $(k + 1)$ -spaces $X \vee V$ where X runs in a quadric $\Phi \subset U$ is called the quadratic cone with vertex space V and base quadric Φ .

Assume a projective coordinate system $(b_0\mathbb{R}, \dots, b_n\mathbb{R}; e\mathbb{R})$ such that the vertex space $V = [b_0, \dots, b_k]$ and $U = [b_{k+1}, \dots, b_n]$. Assume that the base quadric has, within U , the equation

$$\bar{x}^T \cdot C_0 \cdot \bar{x} = 0, \quad \bar{x} = (x_{k+1}, \dots, x_n). \quad (1.37)$$

Then the equation of Φ_V with respect to the original coordinate system obviously is just (1.37).

Tangents

We call a line *tangent* to a quadric Φ if it has exactly one point in common with Φ or is entirely contained in Φ . This is motivated by considering a tangent as limit $P \vee Q$ with both $P, Q \in \Phi$ converging towards each other, and the following

Proposition 1.1.26. *Assume that P is a point of the quadric Φ , which is defined by the polarity κ . All lines L incident with P , which are not contained in $P\kappa$, intersect Φ in exactly one further point. The lines of $P\kappa$ incident with P are precisely Φ 's tangents at P .*

Proof. Assume $L = P \vee Q$ with $P = p\mathbb{R}$, $Q = q\mathbb{R}$, and Φ is given by Equ. (1.33). We know $p^T \cdot C \cdot p = 0$. L is contained in $P\kappa$ if and only if Q is, which is expressed by $q^T \cdot C \cdot p = 0$. To compute further points of $L \cap \Phi$, we solve $(tp + q)^T \cdot C \cdot (tp + q) = 0$. This equation simplifies to $2tq^T \cdot C \cdot p + q^T \cdot C \cdot q = 0$, which shows the result. \square

Thus we call the hyperplane $P\kappa$ the *tangent hyperplane* of Φ at P , if $P \in \Phi$.

The Tangent Cone

We return to Equ. (1.36) and consider the intersection $\Phi_0 = \Phi \cap H$. In view of Prop. 1.1.26, this can be called a *non-tangential intersection* of Φ , if $H = P\kappa$ with $P \notin \Phi$.

Proposition 1.1.27. *If $\Phi_0 = \Phi \cap P\kappa$ is a non-tangential intersection of the quadric Φ , then the cone Φ_P with base quadric Φ_0 and vertex P consists of the tangents of Φ in those points of Φ_0 , which contain P .*

Proof. If $X \in \Phi_0$, then $X \in X\kappa$. Further $X \in P\kappa$ implies $P \in X\kappa$, and so $X \vee P \in X\kappa$, and $X \vee P$ is a tangent. \square

The cone Φ_P is called the *tangent cone* associated with P (cf. Fig. 1.14, left). If Φ_0 is given by Equ. (1.36), then the cone's equation is given by (1.36) as well, without the restriction $x_0 = 0$.

Example 1.1.19. In the projective extension of E^3 , we consider the quadric

$$\Phi : x_0^2 - x_1^2 - x_2^2 + x_3^2 = 0.$$

The inhomogeneous equation describing the affine part of Φ is obtained by the substitution $x = x_1/x_0$, $y = x_2/x_0$, $z = x_3/x_0$ and reads

$$x^2 + y^2 - z^2 = 1.$$

This is a one-sheeted hyperboloid of revolution. It is centered at the origin and has rotational symmetry about the z -axis. Its center $M = (0, 0, 0)$ is the pole of the ideal hyperplane $\omega : x_0 = 0$. The projective reflection with center M and axis hyperplane ω leaves Φ invariant.

Consider an ideal point $P_\omega = (0 : p_1 : p_2 : p_3)$. P 's polar plane $P_\omega\kappa = \mathbb{R}u$ is given by

$$u = \text{diag}(1, -1, -1, 1) \cdot (0, p_1, p_2, p_3)^T = (0, -p_1, -p_2, p_3)^T.$$

Its inhomogeneous equation is $-p_1x - p_2y + p_3z = 0$. It contains M and the projective reflection with center P_ω and axis plane $P_\omega\kappa$ is an affine one. If $p_1 = p_2 = 0$ or $p_3 = 0$, then $P_\omega\kappa$'s normal vector $(0, -p_1, -p_2, p_3)$ is parallel to $(0, p_1, p_2, p_3)$ and so this affine reflection is the Euclidean reflection in $P_\omega\kappa$. This shows that Φ possesses Euclidean symmetry with respect to the plane $z = 0$, and with respect to any plane $(0 : p_1 : p_2 : 0)$ which contains the z -axis. \diamond

Example 1.1.20. (Continuation of Ex. 1.1.19) The intersection $\Phi \cap \omega$ is given by

$$\Phi_0 : x_1^2 + x_2^2 - x_3^2 = x_0 = 0.$$

This is a quadric in the ideal plane, which is called the *ideal conic* of Φ . The tangent cone Φ_M has the homogeneous equation $x_1^2 + x_2^2 - x_3^2 = 0$. In inhomogeneous coordinates its equation reads

$$x^2 + y^2 - z^2 = 0.$$

Φ_M 's generator lines are tangent to Φ at Φ 's ideal points. Any two generators G_1, G_2 of Φ_M span a plane which intersects Φ in a conic with tangents G_1, G_2 . From the Euclidean point of view this conic is a hyperbola with asymptotes G_1, G_2 . Φ_M is called *asymptotic cone*. It is a cone of revolution and contains the asymptotes of all meridian hyperbolae of Φ , which are defined as intersections of Φ with planes containing the z -axis.

Let us emphasize that we use this special Euclidean situation just for illustration and for interpretation of well-known Euclidean properties of quadrics in the more general setting of projective geometry. The notions of ideal plane, center, and asymptotes, are not invariant under projective transformations, and therefore none of them is a notion of projective geometry. \diamond

Euclidean Properties of Quadrics

Consider the projective extension of Euclidean space E^n . Ex. 1.1.19 and Ex. 1.1.20 showed already that some of the projective symmetries of a quadric Φ are Euclidean symmetries. As in the previous text κ and δ denote the polarity and duality associated with Φ .

If the ideal plane ω is not self-conjugate, its pole $M = \omega\delta$ is a proper Euclidean point, and Φ is symmetric with respect to M . M is called *center* of Φ . If $P_\omega \in \omega$, then $P_\omega\kappa$ contains $M = \omega\delta$. Conversely, the pole $H\delta$ of a hyperplane H which contains M is an ideal point. The projective reflection defined by an ideal point P_ω and its polar hyperplane $P_\omega\kappa$ is an *affine* reflection.

If we choose a coordinate system such that ω has equation $x_0 = 0$ and $M = (1, 0, \dots, 0)\mathbb{R}$, then κ 's coordinate matrix C must be as described in the proof of Th. 1.1.23, and the equation of Φ becomes

$$\Phi : x_0^2 + \sum_{i,k=1}^n c_{ik}x_i x_k = 0. \quad (1.38)$$

Theorem 1.1.28. Assume a quadric Φ with equation (1.38). There is a Cartesian coordinate system such that the equation of Φ becomes

$$a_1 X_1^2 + \dots + a_n X_n^2 = 1. \quad (1.39)$$

The Euclidean reflections in the coordinate hyperplanes are projective automorphisms of Φ .

Proof. Consider the matrix C_0 of Equ. (1.36). The spectral theorem of linear algebra says that there is an orthonormal basis $\mathbf{p}_1, \dots, \mathbf{p}_n$ of eigenvectors of C_0 such that $C_0 \cdot \mathbf{p}_i = a_i \mathbf{p}_i$. Consider the Cartesian coordinate system with origin \mathbf{o} and unit vectors $\mathbf{p}_1, \dots, \mathbf{p}_n$. Because of $\mathbf{p}_i^T \cdot C_0 \cdot \mathbf{p}_j = \mathbf{p}_i^T \cdot a_j \mathbf{p}_j = a_j \delta_{ij}$ the equation of Φ is of the form $x_0^2 + \sum a_j x_j^2 = 0$. Its inhomogeneous version is just (1.39).

The ideal points $U_j = (0, \mathbf{p}_j)$ are pairwise conjugate, so $U_j\kappa$ is the hyperplane spanned by the U_k 's with $k \neq j$. Thus the projective symmetry of Φ associated with U_j is the Euclidean reflection in this hyperplane. \square

The coordinate axes of the Cartesian coordinate system referred to in Th. 1.1.28 are called the *axes* of Φ . They need not be unique. The intersections of Φ with its axes, if existent, are called the *vertices* of the quadric. Its coordinates are

$$(0, \dots, 0, 1/\sqrt{a_i}, 0, \dots, 0).$$

Projective Transformations Applied to Quadrics

Let us apply the projective transformation $\mathbf{x}\mathbb{R} \mapsto (A \cdot \mathbf{x})\mathbb{R}$ to the quadric (1.33). With $B := A^{-1}$, we can write the image quadric as

$$\Phi' : \mathbf{y}^T \cdot B^T \cdot C \cdot B \cdot \mathbf{y} = 0. \quad (1.40)$$

Clearly this again is a quadric. A change of coordinate system is expressed by the same equation, if old coordinates \mathbf{x} and new ones \mathbf{x}' are connected by $\mathbf{x}' = A \cdot \mathbf{x}$.

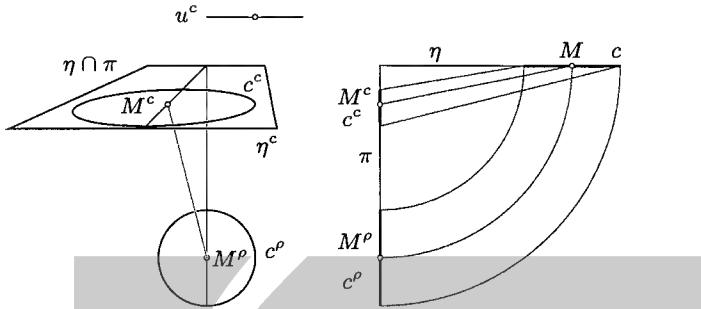


Fig. 1.16. Perspective view c^c of a circle c together with rotated version c^ρ (see text).

Example 1.1.21. Consider a perspectivity which maps a plane η to an image plane π , which is identified with the sheet of paper Fig. 1.16 is printed at. This mapping is denoted by the superscript “ c ”.

A circle c of η (i.e., a special type of quadric) is mapped to the quadric c^c (Fig. 1.16 shows an ellipse).

Denote the polarity and duality associated with c by κ and δ , respectively, and the same for c^c with κ' , δ' . The ideal line of η is denoted by u . If η is not parallel to π , then u^c is a proper line. In the context of perspective drawing, is it called the *horizon*.

The center $M = u\delta$ of the circle c is mapped to the pole $M^c = (u^c)\delta'$ of u^c with respect to c^c . The reader may test his or her knowledge of polarities by finding the point of η which is mapped to c^c 's center. \diamond

Auto-polar Simplices

Th. 1.1.28 shows that there is always a special Cartesian coordinate system where a quadric Φ has a simple equation, provided the ideal plane is not self-conjugate, i.e., the quadric has a center.

Here we want to find a suitable *projective* coordinate system such that the equation of any quadric becomes simple. Alternatively we ask for projective automorphisms which transform Φ to a quadric with a simple equation.

Definition. If Φ is a quadric with associated polarity κ , and P_0, \dots, P_n are projectively independent points such that $P_i\kappa$ is the hyperplane spanned by all P_j with $i \neq j$, then P_0, \dots, P_n is called an auto-polar simplex of Φ (resp., of κ).

The equation of Φ becomes very simple in a projective coordinate system $(P_0, \dots, P_n; E)$ such that (P_0, \dots, P_n) is an auto-polar simplex: If C is the coordinate matrix of κ with respect to such a coordinate system, then $e_i^T \cdot C \cdot e_j$ equals 0 if $i \neq j$. This shows that C is a diagonal matrix and Φ has the equation

$$a_0x_0^2 + \dots + a_nx_n^2 = 0, \quad a_i \neq 0. \quad (1.41)$$

Conversely, if the equation of a quadric Φ is given by (1.41), then the fundamental points of the underlying projective coordinate system are an auto-polar simplex of Φ .

Proposition 1.1.29. *All polarities κ , resp., quadrics Φ , possess auto-polar simplices.*

Proof. It is well known that we can transform a symmetric matrix C into a diagonal matrix by a suitable sequence of the following operations on columns, which are followed by the analogous operations on rows: (i) Add a multiple of the i -th column to the j -th column; (ii) interchange the i -th column with the j -th column; (iii) multiply the i -th column with a nonzero factor.

These column and row operations have the same effect as multiplying C with an appropriate matrix B from the right and the multiplying with B^T from the left afterwards. Equ. (1.40) shows that in such a way we achieve a coordinate transform such that the equation of Φ has diagonal form. \square

A special auto-polar simplex is given by the center and the ideal points of the axes of a quadric Φ , provided the ideal plane is not self-conjugate.

Remark 1.1.12. A more geometric ‘construction’ of an auto-polar simplex of a polarity κ , resp., its quadric Φ , is the following:

Start with any point $B_0 \notin \Phi$. In $B_0\kappa$ choose $B_1 \notin \Phi$. In $B_0\kappa \cap B_1\kappa$, choose $B_2 \notin \Phi$. Assume that we have already constructed B_0, \dots, B_k : In $B_0\kappa \cap \dots \cap B_k\kappa$ choose $B_{k+1} \notin \Phi$. The last point B_n of this sequence equals the only point of the zero-dimensional subspace $B_0\kappa \cap \dots \cap B_{n-1}\kappa$. Fig. 1.17 shows the case $n = 2$. \diamond

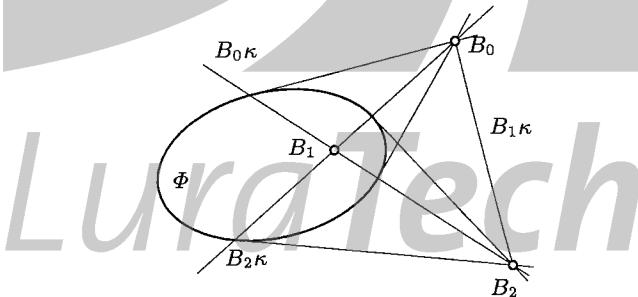


Fig. 1.17. Auto-polar triangle B_0, B_1, B_2 of a conic Φ .

Projective Normal Form of a Quadric

Prop. 1.1.29 shows that in a suitable coordinate system a quadric has the simple equation (1.41). We want to know if we can even achieve that in some coordinate system all coefficients a_i have special values.

A second question is the *projective classification* of quadrics, which means an enumeration of all classes of projectively non-equivalent quadrics. We will see that both the real and complex number fields allow a nice classification.

Theorem 1.1.30. *For all quadrics Φ there is a projective coordinate system such that Φ has equation*

$$\varepsilon_0 x_0^2 + \cdots + \varepsilon_n x_n^2 = 0, \quad \varepsilon_i = \pm 1. \quad (1.42)$$

Proof. Prop. 1.1.29 shows that we may assume a coordinate system such that the equation of Φ is given by (1.41). The coordinate transformation $x'_i = x_i / \sqrt{|a_i|}$ transforms Φ 's equation into (1.42). \square

Assume that all numbers ε_i in (1.42) have the same sign. Then the quadric obviously has no real points. Assume that there are p positive signs and q negative signs ($p + q = n + 1$). After multiplication of the equation with ± 1 we can assume that $q \leq p$ and rewrite (1.42) in the form

$$x_0^2 + \cdots + x_{p-1}^2 - x_p^2 - \cdots - x_{p+q-1}^2 = 0 \quad \text{with } q \leq p. \quad (1.43)$$

Lemma 1.1.31. *The quadric given by Equ. (1.43) contains the subspace*

$$x_0 - x_p = x_1 - x_{p+1} = \cdots = x_{q-1} - x_{p+q-1} = 0, \quad x_q = \cdots = x_{p-1} = 0, \quad (1.44)$$

which is of dimension $n - p = q - 1$.

Proof. Equ. (1.43) is equivalent to $(x_0 - x_p)(x_0 + x_p) + \cdots + (x_{q-1} - x_{p+q-1})(x_{q-1} + x_{p+q-1}) + x_q^2 + \cdots + x_{p-1}^2 = 0$. \square

The number $q - 1$ is called the *index* of the quadric.

Theorem 1.1.32. *(the inertia theorem of J. J. Sylvester) Two quadrics are projectively equivalent if and only if they have the same index. The index is the maximum dimension of a projective subspace contained in a quadric Φ .*

Proof. Consider Equ. (1.43). We define the (linear) subspace U by the equations $x_p = \cdots = x_n = 0$. It has linear dimension p . The quadratic form (1.43) is positive definite when restricted to U .

If a subspace Q is contained in Φ , the restriction of the quadratic form (1.42) to Q is zero. If we consider U, Q as linear subspaces, this shows $U \cap Q = 0$ and therefore the linear dimension of Q is less or equal $(n + 1) - \dim(U) = q$.

On the other hand the subspace defined by Equ. (1.44) is contained in Φ , and its dimension equals the index. Thus we have established the second assertion of the theorem. The first follows immediately, because the maximal dimension of a subspace contained in Φ is a projective invariant. \square

Remark 1.1.13. It is easily verified that n mutually orthogonal eigenvectors of C (which exist, because C is a symmetric matrix) determine an auto-polar simplex of Φ , and therefore the index of Φ may be found by the signs of the eigenvalues of C . This however is not the easiest way to determine the index. \diamond

Example 1.1.22. We compute an auto-polar simplex of the conic $\Phi : 2x_0x_2 - x_1^2 = 0$ by transforming its coordinate matrix into diagonal form by row and column operations as described in the proof of Prop. 1.1.29, and applying the column operations to the unit matrix (below):

$$\begin{array}{ccc|ccc|ccc} & 0 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 0 \\ C := & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -\frac{1}{2} \end{array} = B^T C B$$

$$\begin{array}{ccc|ccc|ccc} & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & -\frac{1}{2} \\ & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & \frac{1}{2} \end{array} =: B$$

Now $C' = B^T C B = \text{diag}(2, -1, -1/2)$ is a diagonal matrix. The coordinate transform $Bx' = x$ achieves transformation of Φ 's equation to diagonal form. The base points of the new coordinate system are the points of the auto-polar simplex. Their coordinates in the old system are the columns of B^{-1} . This means that the points $(1, 0, -2)\mathbb{R}, (0, 1, 0)\mathbb{R}, (1, 0, 2)\mathbb{R}$ are an auto-polar simplex of Φ .

A last sequence of row and column operations (multiplication of columns and rows one, two, and three with factors $1/\sqrt{2}, 1$, and $\sqrt{2}$, respectively) shows that Φ has, with respect to a certain coordinate system, the equation $x_0^2 - x_1^2 - x_2^2 = 0$. This is projective normal form, up to a permutation of coordinates and a negligible scalar factor. The index of Φ obviously equals zero. \diamond

Projective Automorphisms of Quadrics

There are infinitely many projective transformations which transform a quadric Φ in a projectively equivalent quadric Φ' . Two such transformations α, α' differ from each other by a projective automorphism $\beta = \alpha'^{-1} \circ \alpha$ of Φ : The matrix C in (1.33) which defines the quadric Φ is unique up to a nonzero factor. Thus, by (1.40), the matrix B of β satisfies

$$B^T \cdot C \cdot B = \lambda C. \quad (1.45)$$

We say that B is an orthogonal matrix with respect to C . An interpretation will be given later when we discuss metrics.

Clearly the projective reflections defined by the point P and its polar hyperplane are projective automorphisms of Φ . If we are given two points P_1, P_2 of Φ , and we choose P on the line $P_1 \vee P_2$, then the corresponding projective reflection interchanges P_1 and P_2 . This shows the following

Theorem 1.1.33. *For any two points P_1, P_2 of a quadric Φ , there is a projective automorphism of Φ which transforms P_1 to P_2 .*

Example 1.1.23. A quadric in the real projective plane P^2 is a conic. By (1.43), its projective normal form is given by

$$x_0^2 + x_1^2 - x_2^2 = 0. \quad (1.46)$$

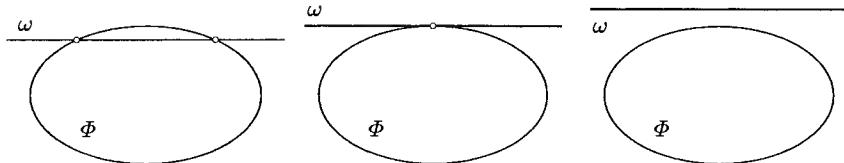


Fig. 1.18. Hyperbola, parabola, and ellipse from the projective point of view.

Thus in real projective geometry there is only one conic, which means that all conics are projectively equivalent. The usual classification of conics is not a projective classification, but an affine or Euclidean one. An *ellipse* has no ideal points, a *parabola* exactly one, and a *hyperbola* has two (cf. Fig. 1.18).

A projective frame such that a given conic Φ has equation (1.46) is in general no homogeneous Cartesian coordinate system. If it is, then (1.46) is the equation of the hyperbola $y^2 - x^2 = 1$. We apply the projective transformation $x'_0 = x_2$, $x'_1 = x_1$, $x'_2 = x_0$ to it. Then the image conic has the equation $-x_0^2 + x_1^2 + x_2^2 = 0$, which is nothing but the unit circle $x^2 + y^2 = 1$. The collineation $x_0 = x'_0 + x'_2$, $x'_1 = x_1$, $x_2 = x'_0 - x'_2$ maps Φ to the parabola with equation $x_1^2 + 4x_0x_2 = 0$ or $x^2 + 4y = 0$. \diamond

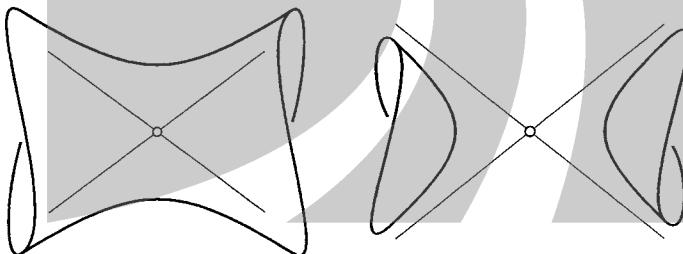


Fig. 1.19. One-sheeted and two-sheeted hyperboloids.

Example 1.1.24. In P^3 , quadrics can have two essentially different normal forms: one with signs $(+++ -)$ and one with signs $(++ - -)$ in (1.42). The *oval quadrics* have projective normal form

$$x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0 \quad (1.47)$$

and carry no straight lines. Its *affine types*, which means classification with respect to their ideal points in projectively extended E^3 , are *ellipsoid*, *two-sheeted hyperboloid* (see Fig. 1.19) and *elliptic paraboloid*.

The *ruled quadrics* have projective normal form

$$x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0. \quad (1.48)$$

Its affine subtypes are the *one-sheeted hyperboloid* (the ideal plane is no tangent plane, see Ex. 1.1.19 and Fig. 1.19) and the *hyperbolic paraboloid* (the ideal plane is a tangent plane, see Fig. 1.23). \diamond

Proof of Projective Properties by Special Cases

The ruled quadrics (Ex. 1.1.24) carry two families of lines. One such family, also called a *regulus*, is seen most easily in a hyperboloid of revolution. Rotation of one such line about the axis generates a regulus. Reflection in a plane which contains the axis gives the second regulus. We also see that two lines of the same regulus are skew and lines of different reguli intersect, possibly in an ideal point.

These properties, easily proved in the special case of a hyperboloid of revolution, involve only elements of projective geometry, such as ‘line’ and ‘intersection’. Therefore all ruled quadrics share these properties. This shows another advantage of embedding Euclidean geometry in projective geometry: projective properties may be proved by means of a Euclidean special case without loss of any generality. More generally, if we find a projective transformation which maps a given situation into one which can be studied better, we can work there.

Example 1.1.25. Let us apply this idea to prove the following result: *Given three pairwise skew lines E_1, E_2, E_3 in P^3 , the set of lines intersecting all three lines is a regulus*, i.e., one of the two families of lines contained in a ruled quadric of projective normal form (1.48).

We consider three skew lines E'_1, E'_2, E'_3 contained in a one-sheeted hyperboloid Φ' , which belong to one of its two reguli. Clearly, for all points $P \in E'_1$, there exists exactly one line L concurrent with E'_1 in P , and in addition meeting E'_2 and E'_3 . It is obtained as the intersection $(P \vee E'_2) \cap (P \vee E'_3)$. We choose a fundamental set B_0, \dots, B_3, E adapted to the three given lines as in Fig. 1.20. We do the same with E'_1, E'_2, E'_3 . These two fundamental sets define a projective transformation which obviously maps E'_i to E_i , maps the other regulus of Φ' (the one which does not contain E'_i , but all lines L) onto a regulus, which is formed by all lines which meet E_1, E_2, E_3 . \diamond

The Polarity Induced in a Subspace

We already discussed the polarity in a hyperplane, induced by a polarity defined in entire projective space. More generally, we define the following: Assume a polarity κ and its associated quadric Φ , and consider a projective subspace G^k of dimension k which is complementary to its $(n - k - 1)$ -dimensional polar space $G^k\delta$. The construction of auto-polar simplices shows that there exists such an auto-polar simplex B_0, \dots, B_n with $G^k = B_0 \vee \dots \vee B_k$. This automatically implies $G^k\delta = B_{k+1} \vee \dots \vee B_n$. Then the polarity *induced in G^k* is defined geometrically by

$$\kappa_k : P \in G^k \mapsto P\kappa \cap G^k. \quad (1.49)$$

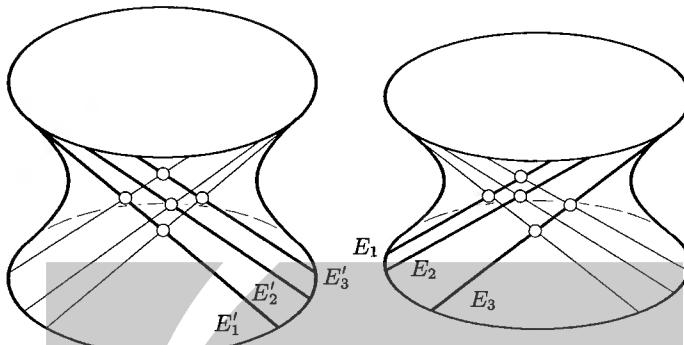


Fig. 1.20. Setting up a projective isomorphism of two ruled quadrics.

We introduce a projective frame such that the polarity has normal form. By (1.42), if $P = (p_0 : \dots : p_n)$, then its polar hyperplane equals

$$P\kappa = \mathbb{R}(\varepsilon_0 p_0, \dots, \varepsilon_n p_n).$$

Thus κ_k is given by

$$\kappa_k : (p_0, \dots, p_k) \mathbb{R} \mapsto \mathbb{R}(\varepsilon_0 p_0, \dots, \varepsilon_k p_k). \quad (1.50)$$

This equation shows that κ_k is indeed a polarity defined in G^k . Its ellipticity or hyperbolicity depends on the $\varepsilon_0, \dots, \varepsilon_k$. If it is hyperbolic, it defines a quadric G^k , which obviously coincides with the intersection $\Phi \cap G^k$. Also the converse is true: If Φ intersects G^k , then the induced polarity is hyperbolic.

The Involution of Conjugate Points

It is important to elaborate somewhat more the special case of a polarity induced in a one-dimensional subspace, i.e., in a line G^1 . Equ. (1.50) takes the form

$$(p_0, p_1) \mathbb{R} \mapsto \mathbb{R}(\varepsilon_0 p_0, \varepsilon_1 p_1).$$

It is a bijective mapping of the line's points to its hyperplanes, which coincide with its points. Note that $\mathbb{R}(\varepsilon_0 p_0, \varepsilon_1 p_1)$ are *hyperplane coordinates*, which means that the corresponding point coordinates $(x_0 : x_1)$ solve the equation $\varepsilon_0 p_0 x_0 + \varepsilon_1 p_1 x_1 = 0$. Thus, the induced polarity is given in point coordinates by

$$(p_0, p_1) \mathbb{R} \mapsto (-\varepsilon_1 p_1, \varepsilon_0 p_0). \quad (1.51)$$

This is an involution, called the *involution of conjugate points*. For $\varepsilon_0 \varepsilon_1 = 1$ it is elliptic, which means that G^1 does not intersect Φ . For $\varepsilon_0 \varepsilon_1 = -1$, the quadric defined by the induced polarity, i.e., the fixed points of the involution, are the intersection points of G^1 and Φ .

Its geometric construction in P^2 is shown by Fig. 1.21, where also the dual counterpart, the *involution of conjugate lines* in P , is illustrated.

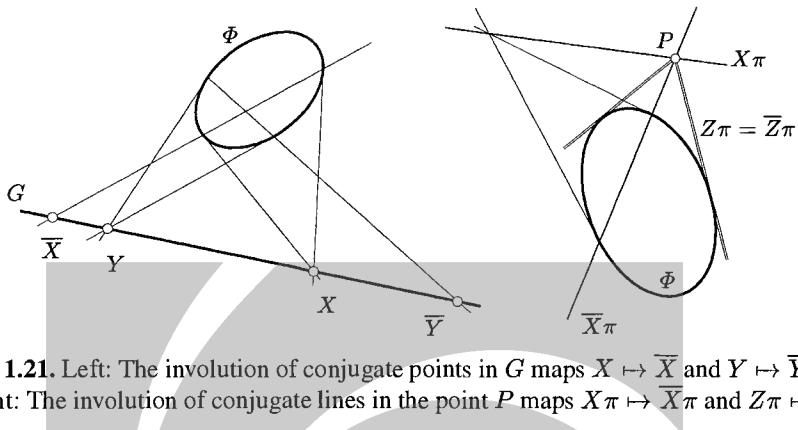


Fig. 1.21. Left: The involution of conjugate points in G maps $X \mapsto \bar{X}$ and $Y \mapsto \bar{Y}$. Right: The involution of conjugate lines in the point P maps $X\pi \mapsto \bar{X}\pi$ and $Z\pi \mapsto \bar{Z}\pi$.

Example 1.1.26. In projectively extended E^2 , consider a conic Φ with center M (which is the pole of the ideal line), and denote the duality defined by this conic with δ . We study the involution κ_1 of conjugate lines at M , which maps a line L containing M (a diameter line) to another diameter line. The pole $L\delta$ of L is an ideal point (because M is the pole of the ideal line). The point $L\delta$ is contained in the diameter $L' = L\kappa_1$ conjugate to L .

If L intersects Φ in two points E, F , the conic tangents $E\delta, F\delta$ in E and F are parallel to the diameter conjugate to L , because $E \in L \implies E\delta \ni L\delta$.

The *axes* of Φ are, by definition, an *orthogonal* pair of conjugate diameter lines. Such a pair exists as has been shown before, and is unique if Φ is not a circle.

The involution κ_1 is illustrated in Figure 1.22 by means of an ellipse and a hyperbola. In the latter case, the so-called conjugate hyperbola and the parallelograms formed by the tangents at opposite points are shown. The *asymptotes* $U\delta, V\delta$ of Φ are defined as the conic tangents at the conic's ideal points U, V . They are concurrent in M (because $U, V \in M\delta$), they are fixed lines of κ_1 (because $(U\delta)\kappa_1$ must contain $(U\delta)\delta = U$), and they are diagonals of the parallelograms shown by Fig. 1.22. \diamond

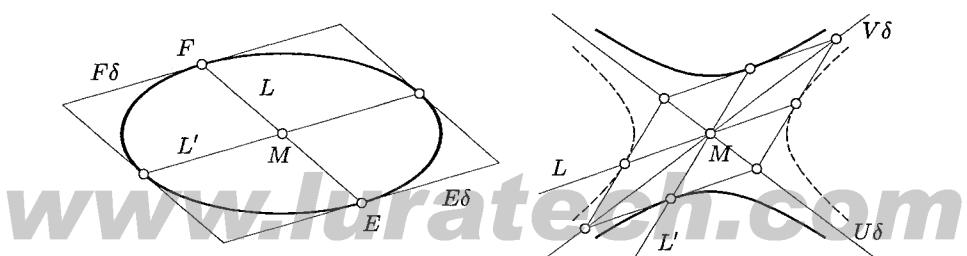


Fig. 1.22. Involution of conjugate diameter lines.

The Set of Tangent Hyperplanes

Consider the polar hyperplane $P\kappa$ of a point $P \in \Phi$, i.e., $P \in P\kappa$. Assume a coordinate system such that $P = (1, 0, \dots, 0)\mathbb{R}$ and $P\kappa = \mathbb{R}(0, \dots, 0, 1)$. Then the matrix C of κ must satisfy $c_{00} = \dots = c_{0,n-1} = 0, c_{0n} \neq 0$. Thus, Φ has in this coordinate system an equation of the form

$$x_0 x_n + \sum_{i,k=1}^n c_{ik} x_i x_k = 0. \quad (1.52)$$

If this coordinate system is a homogeneous Cartesian coordinate system in E^n , then in inhomogeneous coordinates (X_1, \dots, X_n) , (1.52) reads as

$$X_n + \sum_{i,k=1}^n c_{ik} X_i X_k = 0. \quad (1.53)$$

P becomes $P = (0, \dots, 0)$ and $P\kappa$ has equation $X_n = 0$. The left hand side of (1.53) clearly shows that $X_n = 0$ is a first order approximation of Φ in a neighbourhood of the point P , which shows that the polar hyperplane at P coincides with the tangent hyperplane in the sense of differential geometry. Projective coordinate transforms are differentiable, and so this is true for all proper points of a quadric.

All lines of $P\kappa$ concurrent in P are *tangents* of Φ at P . More generally, any subspace $G^k \subset P\kappa$ which contains the ‘point of tangency’ P is called *tangent subspace*. Its polar space $G^k\delta$ consequently contains P , and therefore intersects G^k . Conversely, assume that a subspace G^k intersects its polar space $G^k\delta$. Since any point of G^k is conjugate to any point of $G^k\delta$, all points of the intersection S are self-conjugate and thus contained in Φ . If $P \in S$, then $P\delta$ contains G^k , since all points of G^k are conjugate to P . This shows the following theorem, the proof of whose second statement is completely analogous:

Theorem 1.1.34. *A subspace G^k is tangent to a quadric if and only if it intersects its polar subspace $G^k\delta$. A subspace G^k is contained in a quadric Φ if and only if it is contained in its polar space $G^k\delta$.*

The duality associated with a polarity maps points of a quadric to its tangent hyperplanes and vice versa. This is formulated in the following

Theorem 1.1.35. *The set of tangent hyperplanes of a quadric Φ is a quadric in dual projective space. This duality is explicitly realized by the duality defined by Φ . If we apply it twice, we get the original quadric again.*

Singular Quadratic Varieties

Assume that $P = (1 : 0 : \dots : 0)$ is contained in the quadric Φ and that its tangent hyperplane $P\kappa$ is given by $x_n = 0$. Then the *tangential intersection* $P\kappa \cap \Phi$ of Φ quadric has the equation

$$\sum_{i,k=1}^{n-1} c_{ik} x_i x_k = 0, \quad x_n = 0. \quad (1.54)$$

which is the zero set of a quadratic form in x_0, \dots, x_{n-1} . Its matrix annihilates the vector $(1, 0, \dots, 0)$ by construction, it therefore has rank $\leq n - 1$. Because the original matrix $(c_{ik})_{i,k=0}^n$ had rank $n + 1$, its rank equals exactly $n - 1$. The tangential intersection contains P and has the form of a *quadratic cone* in the $(n - 1)$ -dimensional projective subspace $P^{n-1} |_{x_n = 0}$. For any two points $P, P' \in \Phi$, there is always a projective automorphism of Φ which maps $P \mapsto P'$, and which clearly also maps the tangent plane of P onto $P' \kappa$. Therefore, any two tangential intersections of Φ are projectively equivalent.

In general, a *singular quadratic variety* Φ in P^n , given by the equation

$$\mathbf{x}^T \cdot C \cdot \mathbf{x} = 0, \quad (\det C = 0, \quad C = C^T), \quad (1.55)$$

possesses a *vertex space* V . As a projective subspace it corresponds to a linear subspace. This linear subspace equals the null space $\{\mathbf{x} \mid C \cdot \mathbf{x} = 0\}$ of C , which has dimension $n + 1 - \text{rank } C$. Therefore V has the projective dimension $d = n - \text{rank } C$. In a projective coordinate system $(B_0, \dots, B_n; E)$ such that B_0, \dots, B_d span V , Equ. (1.55) reads

$$\sum_{i,k=d+1}^n c_{ik} x_i x_k = 0. \quad (1.56)$$

If $d = n - 1$, this equation has the form $x_n^2 = 0$ and therefore it describes the hyperplane $x_n = 0$. We say that the quadratic variety consists of the hyperplane, with algebraic multiplicity 2. Otherwise, (1.56) is the zero set of a quadratic form, which only employs the variables x_{d+1}, \dots, x_n , and the matrix $(c_{ik})_{i,k=d+1}^n$ is regular.

There are two possibilities: Either Equ. (1.37) has no solutions within the subspace $U = B_{d+1} \vee \dots \vee B_n$, which has the equation $x_0 = \dots = x_d = 0$, or it defines a quadric Φ' in this subspace.

In the first case, Φ consists of the vertex subspace alone. In the second case, the definition of a quadratic cone and Equ. (1.37) show that Φ is a quadratic cone with base quadric Φ' and vertex space V . Summing up, we have

Theorem 1.1.36. *A singular quadratic variety Φ in P^n either coincides with its vertex space V of dimension d , or there is a regular quadric Φ' in a subspace G^{n-d-1} complementary to V , such that Φ equals the quadratic cone Φ'_V .*

Example 1.1.27. Two examples: The vertex of a quadratic cone in P^2 is a point. A subspace complementary to this point is a non-incident line. A quadric in this line is a pair of points and thus the cone is a *pair of lines*.

In P^3 , there are the following possibilities for a singular quadratic variety: (i) a quadratic cone which consists of all lines concurrent in the vertex V ($d = 0$) and intersecting a conic whose carrier plane does not contain V , or (ii) a pair of planes which intersect in the vertex line ($d = 1$), or (iii) a plane of multiplicity two ($d = 2$). \diamond

The definition of ‘tangent’ which we used for quadrics is valid for quadratic cones as well. It turns out that quadratic cones have tangent hyperplanes, which can be described analogously to Prop. 1.1.26.

Proposition 1.1.37. *Assume that a quadratic cone Φ has the equation (1.55). If $P = p\mathbb{R}$ is not contained in the cone’s vertex, then Φ ’s tangents in P are the lines of the hyperplane $\mathbb{R}(C \cdot p)$ which are incident with P .*

Proof. Because P is not in the vertex space, $C \cdot p \neq 0$, and $\mathbb{R}(C \cdot p)$ is actually a hyperplane. Then the proof is the same as for Prop. 1.1.26. \square

We know that cones $\Phi = \Phi'_V$ in three-space which connect a vertex V with a conic Φ' have the property that the tangent plane of P is also tangent plane in all points of $P \vee V$. This result generalizes to:

Proposition 1.1.38. *We use the notation of Th. 1.1.36. If $P \in \Phi$, $P \notin V$, then P ’s tangent hyperplane is also the tangent hyperplane of all points of $P \vee V$. All tangent hyperplanes contain the vertex subspace V .*

Proof. Prop. 1.1.37 shows how to compute tangent hyperplanes. We use the coordinate system of the discussion preceding Th. 1.1.36, such that the equation of Φ is given by (1.56). If $P \in \Phi$, $P = p\mathbb{R}$, $Q = q\mathbb{R}$, then $Q \in P \vee V$ if and only if $(q_{d+1}, \dots, q_n) = \lambda(p_{d+1}, \dots, p_n)$. Equ. (1.56) does not involve the variables x_0, \dots, x_d , so $\mathbb{R}(C \cdot q) = \mathbb{R}(C \cdot p)$ for all $q\mathbb{R} \in P \vee V$.

The second statement is clear: If $v\mathbb{R} \in V$, then $C \cdot v = 0$, so $(C \cdot p)^T \cdot v = 0$. \square

Finally a result about the dual of a singular quadratic variety:

Proposition 1.1.39. *We use the notation of Th. 1.1.36. The set of Φ ’s tangent hyperplanes is a quadric contained in an $(n - d - 1)$ -dimensional subspace of the dual space P^{n*} .*

Proof. If Φ is given by (1.55), we have to show that the set of all planes $\mathbb{R}(C \cdot p)$ for $p\mathbb{R} \in \Phi$, $p\mathbb{R} \notin V$, is a quadric. We use a coordinate system such that Φ is given by (1.56). If we disregard the variables x_0, \dots, x_d , (1.56) describes a regular quadric. We let $C' = (c_{ik})_{i,k=d+1}^n$ and $p' = (p_{d+1}, \dots, p_n)$. By Th. 1.1.35, the set of hyperplanes $\mathbb{R}(C' \cdot p')$ is a quadric. This shows the statement, because $C \cdot p = (0, \dots, 0, C' \cdot p')$, where the number of initial zeros equals d . \square

Example 1.1.28. Consider the quadratic cone $\Phi : x^2 + y^2 = z^2$ in Euclidean three-space. Its projective extension has the equation $x_1^2 + x_2^2 - x_3^2 = 0$. It is a singular quadratic variety. If we write its equation in the form $x^T \cdot C \cdot x = 0$, then $C = \text{diag}(0, 1, 1, -1)$. The vertex space V equals the point $(1, 0, 0, 0)\mathbb{R}$.

The set Φ^* of tangent planes, i.e., the set of planes $\mathbb{R}(C \cdot p)$ with $p\mathbb{R} \in \Phi \setminus V$ is the set of planes $\{\mathbb{R}(0, p_1, p_2, -p_3)\}$ with $p_1^2 + p_2^2 = p_3^2$. It is therefore defined by the equations

$$\Phi^* = \{\mathbb{R}u \mid u_1^2 + u_2^2 - (-u_3)^2 = 0, u_0 = 0\}.$$

Clearly Φ^* is a conic contained in the subspace $u_0 = 0$. \diamond

Intersections of Quadrics with Subspaces

The description of (possibly singular) quadratic varieties enables us to describe the intersections of quadrics with projective subspaces. Apart from empty intersections we have the following cases: If G^k is not tangent to Φ , the intersection is a quadric in G^k . If G^k is tangent to Φ , the intersection is a singular quadratic variety.

Example 1.1.29. Consider the quadric M_2^4 of P^5 which is given by the equation

$$M_2^4 : x_0x_3 + x_1x_4 + x_2x_5 = 0. \quad (1.57)$$

It plays a fundamental role in line geometry and therefore we have a closer look at it now. Its projective normal form is

$$M_2^4 : x_0^2 + x_1^2 + x_2^2 - x_3^2 - x_4^2 - x_5^2 = 0, \quad (1.58)$$

which shows that M_2^4 is of index 2, which is the maximum index for a quadric in P^5 . This can already be seen from equation (1.57), since the two-dimensional subspaces $x_0 = x_1 = x_2 = 0$ or $x_3 = x_4 = x_5 = 0$ are contained in M_2^4 .

Intersections with non-tangent hyperplanes H are quadrics of H . Their index equals one, because the planes contained in M_2^4 intersect H in lines.

In order to investigate the tangential intersections, it is sufficient to consider one of them. We use the hyperplane $T^4 : x_5 = 0$. It contains its pole $P = (0 : 0 : 1 : 0 : 0 : 0)$ and thus indeed is a tangent hyperplane. Its intersection with M_2^4 has equation $x_0x_3 + x_1x_4 = x_5 = 0$. This is a quadratic cone with vertex P and base quadric $\Phi' : x_0x_3 + x_1x_4 = x_2 = x_5 = 0$, which is a ruled quadric in a three-dimensional projective subspace.

Intersections of M_2^4 with three-dimensional projective subspaces G^3 may be pairs of different planes, quadratic cones with a point as vertex, as well as oval or ruled quadrics. The intersection can never be empty, since a 3-space G^3 intersects each of the planes contained in M_2^4 , which is a consequence of the projective dimension theorem. These are all possible cases, and all of them actually occur: one can find a three-space for each particular case. \diamond

Paraboloids

We have already mentioned some quadrics which are called *paraboloids*. In general, this term is reserved for quadrics which do not have a center. We call a quadric Φ a *paraboloid*, if the ideal hyperplane ω is tangent to Φ . The point of tangency A_ω is found as the pole $A_\omega = \omega\delta$ of the ideal hyperplane. It can be shown that for each paraboloid there is a Cartesian coordinate system such that the equation of the proper part of the paraboloid reads

$$X_n = a_1 X_1^2 + \cdots + a_{n-1} X_{n-1}^2. \quad (1.59)$$

Example 1.1.30. The paraboloids of E^3 belong to two different projective types of quadrics: An *elliptic paraboloid* Φ is an oval quadric (i.e., its index is zero) and may be written in the inhomogeneous normal form

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad (1.60)$$

where $x = x_1/x_0$, $y = x_2/x_0$, $z = x_3/x_0$. Its intersection with the ideal plane $x_0 = 0$ has the equation $x_0 = b^2x_1^2 + a^2x_2^2 = 0$. The only real solution of this equation is the point $A_\omega = (0 : 0 : 0 : 1)$.

This shows that all planar sections of Φ have at most one point at infinity, and are therefore ellipses, parabolae, or possibly empty, or consist of just one point. This explains the name *elliptic*. If $a = b$, then Φ is a *paraboloid of revolution*, which is generated by rotating a parabola about its axis. \diamond

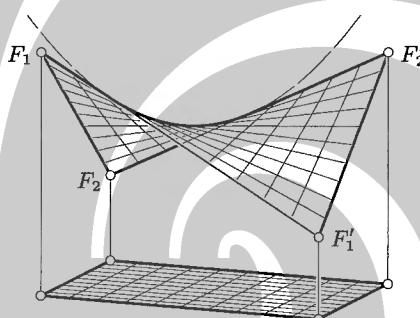


Fig. 1.23. Hyperbolic paraboloid.

Example 1.1.31. A paraboloid Φ of E^3 may be a ruled quadric. It is then called a *hyperbolic paraboloid*. Its inhomogeneous normal form is given by

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}. \quad (1.61)$$

The intersection with the ideal plane $x_0 = 0$ is given by $x_0 = (bx_1 + ax_2)(bx_1 - ax_2) = 0$, and thus consists of two lines $E_\omega : x_0 = bx_1 + ax_2 = 0$ and $F_\omega : x_0 = bx_1 - ax_2 = 0$, one from each regulus of the quadric. Therefore the intersection of Φ with a plane ε either has one or two ideal points, depending on the number of intersection points of ε with E_ω and F_ω . Planar intersections of Φ are thus parabolae, hyperbolae, or pairs of lines (if the intersection is tangential).

All lines of one regulus intersect the ideal line of the other. This means that all lines of one regulus are parallel to all planes containing the ideal line of the other, which has equation $bx \pm ay = c$, with an arbitrary c .

Such a plane is called *director plane*. A parallel projection in direction $(0, 0, 1)$ maps all director planes onto lines, and so both reguli are mapped to pencils of parallel lines (see Fig. 1.23). \diamond

Examples of Quadrics Generated by Projectivities and Steiner's Theorem

We may ask the following elementary question: Consider a line in the plane or in space. Move this line in arbitrary manner. Then, for all points of the line, connect the original with the new position by a straight line. How does this set of lines look like?

We will discuss this and some related questions. It is fortunate that the hyperbolic paraboloid (see Ex. 1.1.31) carries two families of lines which project to parallel pencils under an appropriate parallel projection p as shown by Fig. 1.23. We will exploit this fact.

Proposition 1.1.40. *Three lines of one regulus of a hyperbolic paraboloid intersect any two lines of the other regulus in point triples of equal ratio.*

Proof. This is true because ratios are not affected by parallel projections, and after application of our special projection p the statement is trivial. \square

The converse is also true:

Proposition 1.1.41. *Assume a correspondence between two skew lines in space which preserves ratios of triples of points. Connecting corresponding points gives one of the two reguli of a hyperbolic paraboloid.*

Proof. Such a correspondence (which is called a *similarity*) is necessarily one-to-one and onto, and is uniquely determined by the images \mathbf{a}', \mathbf{b}' of two points \mathbf{a}, \mathbf{b} . The similarity then is given by

$$(1 - \lambda)\mathbf{a} + \lambda\mathbf{b} \mapsto (1 - \lambda)\mathbf{a}' + \lambda\mathbf{b}'$$

for $\lambda \in \mathbb{R}$. In an appropriate coordinate system it is now easy to show that the surface of all lines which connect corresponding points is indeed a hyperbolic paraboloid. \square

Prop. 1.1.41 is true in higher dimensions as well because the projective span of two skew lines is three-dimensional.

Consider a smooth surface Φ in P^3 and a central projection with center C and image plane π . Then the *contour outline* c of Φ is the set of all points whose tangent plane passes through C . It projects to a curve $c' \subset \pi$, which is called the *silhouette* of Φ . If a one-parameter family of curves contained in Φ traverses c , their projections possess c' as an envelope, because the tangent plane of points of c is mapped to a line.

Proposition 1.1.42. *The silhouette c' of a hyperbolic paraboloid with respect a parallel projection is a parabola, if the center is not contained in the surface. (cf. Fig. 1.23).*

Proof. If Φ is a quadric and $C \notin \Phi$, then the contour outline c is the intersection of Φ with C 's polar hyperplane, and so c' is a conic (if C is an exterior point of Φ) or empty (if C is an interior point of Φ). A hyperbolic paraboloid does not have interior points.

A *parallel projection* is a projection whose center is an ideal point. A *general parallel projection* in this context means that the center is not contained in the hyperbolic paraboloid Φ , which means $C \notin E_\omega$ and $C \notin F_\omega$. Then its polar plane $C\kappa$ contains A_ω and the contour outline c touches the ideal plane at A_ω . The curve c projects to a conic which contains only one ideal point (the only ideal point whose tangent plane contains the center is A_ω). Therefore it projects to a conic c' with one ideal point, i.e., a parabola. \square

Proposition 1.1.43. *Assume a similarity between two different lines in the plane. The lines l which connect corresponding points are tangent to some parabola, if they are not concurrent.*

Proof. In order to prove this, we find a hyperbolic paraboloid and lines in it which, under an appropriate parallel projection, are mapped to the lines l in question. A regulus of a hyperbolic paraboloid is generated by two skew ranges F_1, F_2 of points connected by a similarity. They project to similar ranges F'_1, F'_2 , and the lines of the regulus project to the lines l . This situation is again seen in Fig. 1.23. Now the envelope of the lines l coincides with the silhouette of Φ , which is a parabola, if the center of projection is not contained in the surface.

If it is, there is a line L of the other regulus incident with the center, and the images of all lines of the first regulus intersect the image of L , which is a point. \square

Proposition 1.1.44. *Assume two skew lines E_1, E_2 and a projectivity $E_1 \rightarrow E_2$. The lines connecting corresponding points consist precisely of one regulus of a ruled quadric. Conversely, the lines of one regulus of a ruled quadric intersect any two lines of the other regulus in point pairs of a projectivity between these two lines.*

Proof. This follows immediately from the previous discussion of hyperbolic paraboloids: The theorem to prove is a theorem of projective geometry. All ruled quadrics are projectively equivalent. The similarities discussed above are special cases of projectivities, and any projectivity can be transformed into a similarity by an appropriate projective automorphism of P^3 (just map a pair of corresponding points such that both become ideal points.) \square

Proposition 1.1.45. *Assume two lines E_1, E_2 in the projective plane and a projectivity $E_1 \rightarrow E_2$. Either, the envelope of the lines which connect corresponding points is a conic, or all lines are concurrent (see Fig. 1.24).*

Proof. This has been proved for the special case of a similarity; the general case follows from the fact that a similarity is a special case of a projectivity, and that a projectivity may be transformed into a similarity by an appropriate projective automorphism of the plane. \square

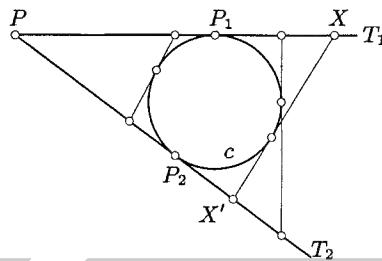


Fig. 1.24. A projectivity between lines generates the tangents of a conic, and vice versa.

Proposition 1.1.46. Fix two tangents T_1, T_2 of a conic c , which touch c in points P_1, P_2 . A further tangent of c intersect T_1, T_2 in points X, X' . The mapping of T_1 onto T_2 which maps $X \mapsto X'$ is part of a projectivity $T_1 \rightarrow T_2$ (cf. Fig. 1.24).

Proof. This follows from the previous proposition and the fact that if a point is not contained in c but in one of c 's tangents, there is exactly one other tangent passing through this point. We had to say ‘part of’ a projectivity, because P_1, P_2 and $P = T_1 \cap T_2$ have no partner yet. By continuity, the pairs (P_1, P) and (P, P_2) complete the projectivity. \square

Theorem 1.1.47. (of J. Steiner) Assume two points $P_1 \neq P_2$ in the projective plane and a projectivity π which maps the pencil of lines concurrent in P_1 to the line pencil P_2 , such that $P_1 \vee P_2$ is not mapped to itself. Then for all lines l incident with P_1 consider the point $l \cap l\pi$. These points are exactly the points of a certain conic (see Fig. 1.25).

Proof. To prove this theorem, we have nothing to do: We already know that the set of tangents of a conic is a conic in the dual plane. So this theorem is nothing but the dual version of Prop. 1.1.45. \square

Proposition 1.1.48. Assume a conic c and two points $P_1 \neq P_2$ in it. For all P in c , consider the lines $P \vee P_1$ and $P \vee P_2$. This defines a mapping of the line pencil P_1 to the line pencil P_2 , which is a projectivity.

Proof. Again the dual statement has been shown in Prop. 1.1.46, so there is nothing to prove. \square

There are of course other proofs of Steiner's theorem which do not need the entire theory of quadrics. To become more familiar with the concepts of projective geometry, let us show one of them:

Proof. Consider two pencils of lines with vertices $A \neq B$ and a projectivity π from the first pencil to the second. By our assumptions, the line $L = A \vee B$ has a preimage $T_A = L\pi^{-1}$ and an image $T_B = L\pi$, which are different from L (Fig. 1.25). Let

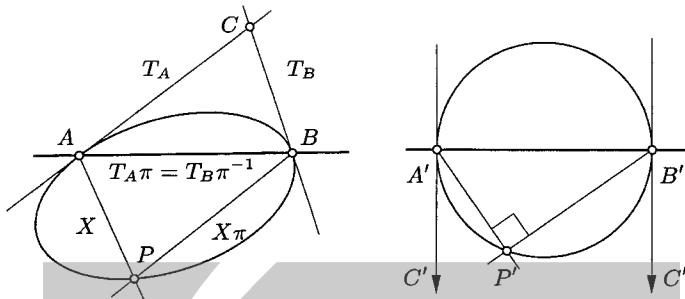


Fig. 1.25. Left: A projective isomorphism π of line pencils with vertices A and B generates a conic by Steiner's theorem. Right: Thales' theorem is a Euclidean special case.

$X, X\pi$ be another pair of corresponding lines. With respect to the projective frame $(T_A \cap TB, A, B; X \cap X\pi)$ the projectivity π is given by

$$\mathbb{R}(a, b, 0) \mapsto \mathbb{R}(b, 0, a), \quad (a, b) \in \mathbb{R}^2 \setminus \{(0, 0)\},$$

because this mapping is linear and transforms the points T_A, L , and X in the desired way. The intersection $\mathbb{R}(a, b, 0) \cap \mathbb{R}(b, 0, a)$ is given by the point $(ab, -a^2, -b^2)\mathbb{R}$, which is a parametrization of the conic $x_0^2 - x_1 x_2 = 0$. \square

Yet another proof is the following:

Proof. We use a projective automorphism of the plane which transforms the given projectivity into a special one, where all lines are orthogonal to their image. Fig. 1.25 shows how to set up this collineation by corresponding fundamental sets $ABCD$ and $A'B'C'D'$. By Thales' theorem, these orthogonal pencils generate a circle. Hence, general projective pencils generate a projective image of a circle, which is a conic. We also see that the vertices A, B of the two pencils lie on the conic; preimage and image T_A, T_B of the common line L are the tangents there. \square

Steiner's theorem and the fact that projectivities are determined by the images of three elements can be used to prove the following theorem on the number of points and/or tangents which uniquely determine a conic:

Theorem 1.1.49. *A conic in the projective plane is determined uniquely by (i) five points, (ii) four points and the tangent at one of them, (iii) three points and the tangents at two of them. (iv) five tangents, (v) four tangents and the contact point of one of them, (vi) three tangents and the contact points of two of them (if these elements are in admissible position).*

Admissible position means that no three points may be collinear, a tangent in a point may not contain a further point, and no three tangents may be concurrent.

1.1.6 Complex Extension and the Way from Projective to Euclidean Geometry

The Complex Extension of Real Projective Space

The *complex extension* of real projective n -space P^n is very convenient in many places, especially when we deal with polynomial equations. It is denoted by $\mathbb{C}P^n$ and is constructed as follows: We embed the linear space \mathbb{R}^{n+1} in \mathbb{C}^{n+1} and consider the complex projective space over it. This means that a point of $\mathbb{C}P^n$ consists of all complex multiples $z\mathbb{C}$ of a vector $z \in \mathbb{C}^{n+1}$. If two *real* points have no common real multiple, then they neither have a common complex multiple, so the embedding

$$P^n \rightarrow \mathbb{C}P^n, \quad x\mathbb{R} \mapsto x\mathbb{C}$$

is well defined. In coordinates the embedding is simple: We just allow points whose coordinates are complex numbers. Of course the coordinate vector of a real point becomes complex when multiplied with a non-real number (e.g., the points $(1, 1, 1)\mathbb{C}$ and $(i, i, i)\mathbb{C}$ are the same point). It is nevertheless easy to recognize the real points $z\mathbb{C}$: They are precisely those points $(z_0 : \dots : z_n)\mathbb{C}$ where all ratios $z_i : z_j$ are real.

If $z = x + iy \in \mathbb{C}$, then the *conjugate* of z is given by $\bar{z} = x - iy$. If $z = (z_0, \dots, z_n)$, then $\bar{z} = (\bar{z}_0, \dots, \bar{z}_n)$ is the complex conjugate of z . So for all points $z\mathbb{C}$ the point $\bar{z}\mathbb{C}$ is called the *complex conjugate* of $z\mathbb{C}$. It is easily verified that $z\mathbb{C} = \bar{z}\mathbb{C}$ if and only if $z\mathbb{C}$ is a real point: If $x \in \mathbb{R}^{n+1}$ and $z = \lambda x$ with $\lambda \in \mathbb{C}$, then $\bar{z} = \bar{\lambda}\bar{x} = \bar{\lambda}x$, so $\bar{z}\mathbb{C} = z\mathbb{C}$. On the other hand, if $\bar{z}\mathbb{C} = z\mathbb{C}$, then there is a $\lambda \in \mathbb{C}$ with $\bar{z} = \lambda z$, which implies that the real vector $x = z + \bar{z}$ equals $(1 + \lambda)z$, and so $z\mathbb{C} = x\mathbb{C}$.

The mapping $z\mathbb{C} \mapsto \bar{z}\mathbb{C}$ preserves incidences and subspaces, but it is *not induced by a linear mapping*, because the cross ratio of $(\bar{a}\mathbb{C}, \bar{b}\mathbb{C}, \bar{c}\mathbb{C}, \bar{d}\mathbb{C})$ is the complex conjugate of the cross ratio of $(a\mathbb{C}, b\mathbb{C}, c\mathbb{C}, d\mathbb{C})$. So this mapping is an example of a mapping which preserves subspaces and incidence, but is *no projective automorphism*. Such a mapping does not exist in real projective space (this follows from Remark 1.1.9).

The point conjugate to a given point is independent of the choice of a coordinate system, if its fundamental points are real.

Real and Complex Subspaces

We call a subspace of complex projective space *real*, if it is spanned by real points. A real line spanned by points $a\mathbb{R}, b\mathbb{R}$ of course contains all points $(\lambda a + \mu b)\mathbb{C}$ with $\lambda, \mu \in \mathbb{R}$, which are real points, but also all points $(\lambda a + \mu b)\mathbb{C}$ with $\lambda, \mu \in \mathbb{C}$, not all of which are real. It is easily verified that a projective subspace is real if and only if it equals its complex conjugate (i.e., the projective subspace consisting of the conjugates of its points). Thus the invariant subspaces of the mapping $\iota : z\mathbb{C} \mapsto \bar{z}\mathbb{C}$ are precisely the real subspaces.

The complex conjugate of the hyperplane C_u is given by $C_{\bar{u}}$, because $u \cdot z = 0 \iff \bar{u} \cdot \bar{z} = 0$.

If two real subspaces G_1, G_2 intersect, then they intersect in a real subspace, which follows from $\overline{G_1 \cap G_2} = \overline{G_1} \cap \overline{G_2}$.

Proposition 1.1.50. *An imaginary point $z\mathbb{C}$ is contained in exactly one real line, which is spanned by $z\mathbb{C}$ and $\bar{z}\mathbb{C}$.*

Proof. The line $z\mathbb{C} \vee \bar{z}\mathbb{C}$ is real, since it is spanned by the real points $(z + \bar{z})\mathbb{C}$ and $i(z - \bar{z})\mathbb{C}$. If $z\mathbb{C}$ is contained in another real line, then it is the intersection of two real lines, and therefore it is a real point. \square

The dual version is:

Proposition 1.1.51. *A non-real hyperplane $\mathbb{C}U$ contains exactly one real $(n - 2)$ -dimensional projective subspace, namely its intersection with $\mathbb{C}\bar{U}$.*

Example 1.1.32. There are two types of *imaginary lines* in $\mathbb{C}P^3$. A non-real line L does not coincide with its complex conjugate \bar{L} . It may intersect \bar{L} in one point, or L, \bar{L} may be skew.

In the latter case, L does not contain any real point P (because then \bar{P} is contained in \bar{L} , but $\bar{P} = P$). Analogously it is easy to show that L is contained in no real plane. Such lines are called *imaginary lines of the second kind*.

If L and \bar{L} intersect, then the point $P = L \cap \bar{L}$ is real, because then $\bar{P} = \bar{L} \cap \bar{\bar{L}} = \bar{L} \cap L = P$. Such a line is called *imaginary line of the first kind*. The plane $\varepsilon = L \vee \bar{L}$ is real, because its complex conjugate $\bar{\varepsilon} = \bar{L} \vee \bar{\bar{L}}$ obviously equals ε .

Examples of such lines are found in tangential intersections of oval quadrics, which intersect in the point of tangency. Thus we may say that an oval quadric, when embedded in $\mathbb{C}P^n$, becomes a ruled quadric, both of its reguli being imaginary lines of the first kind. \diamond

Real Collineations and Polarities

We call a collineation or a polarity in complex projective space *real* if in some real projective frame it is expressed by a real linear mapping. Nevertheless we also apply these mappings to complex points, and we may ask for complex points contained in quadrics.

The complex number field is algebraically closed and thus all polynomial equations have a solution. Thus in complex projective space all projectivities and collineations have fixed points, and all involutions and polarities are hyperbolic. There is only one quadric, and it has normal form

$$x_0^2 + \cdots + x_n^2 = 0. \quad (1.62)$$

A real elliptic polarity which has no real points, now defines a quadric. So there may be a quadric without real points whose associated polarity is real. An important example of such a quadric will occur when we embed Euclidean geometry in projective geometry.

Affine Space as a Subset of Projective Geometry

'Geometry' in the sense of Felix Klein consists of some 'space' and some group of transformations acting on it. It happens that the most familiar geometries, the Euclidean and the affine geometries, can be embedded in projective geometry, as we have already seen in some places. The statement that 'all geometry is projective geometry' is true insofar as we restrict ourselves to linear automorphisms acting on subsets of linear spaces, and to projective automorphisms acting on subsets of projective spaces. This class of geometries actually is a very rich one.

Affine geometry is constructed from projective geometry as follows: We choose a hyperplane ω . If we use a projective frame $(B_0, \dots, B_n; E)$ such that B_1, \dots, B_n span ω , then ω has the equation $x_0 = 0$. Points in *affine space* A^n are characterized by $x_0 \neq 0$. They can be described by homogeneous coordinates of the form $(1, x_1, \dots, x_n)\mathbb{R}$, or by the *affine coordinate vector* $\mathbf{x} = (x_1, \dots, x_n)$. We see that A^n is isomorphic to the well known affine space \mathbb{R}^n . The projective coordinate frame is called an *affine frame*. The affine coordinates are also uniquely determined by the point $B_0 = \mathbf{o}$ and the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, such that $\mathbf{o} + \mathbf{e}_i$ have affine coordinates $(0, \dots, 0, 1, 0, \dots, 0)$ (see Fig. 1.26). Linear independence is the only condition on the vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$; attributes like unit length and orthogonality do not belong to affine geometry. Cartesian coordinate systems are special cases of affine frames. They can be defined only if more than just one hyperplane has been distinguished in a projective space.

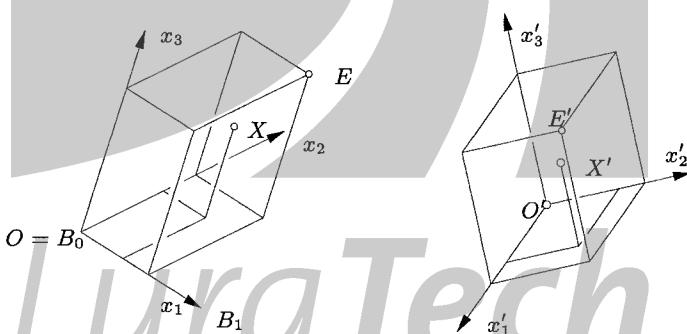


Fig. 1.26. Affine map as transformation via equal coordinates with respect to two affine frames.

What we have just done is the operation reverse to the projective extension of \mathbb{R}^n , which was performed by adding an ideal plane to an affine space. Because all hyperplanes of projective space are equal with respect to the group of projective transformations, it is justified that we can take away any hyperplane and call the rest an affine space.

If there is an affine space embedded in a projective space, we may speak of ideal points, which are those not in the affine space. An ideal point A_ω may however be

'seen' from the affine space as it is uniquely determined by all lines concurrent there (which are parallel from the affine viewpoint). This also shows that projectively extending an affine space, which was previously obtained by deleting a hyperplane of a projective space, recovers the original projective space.

Affine Transformations

Even more is true: We show that not only the point sets of affine and projective space correspond in the way described above, but also the affine and projective transformation groups are related: In an n -dimensional linear space such as \mathbb{R}^n , there is the group of affine transformations of the form

$$\mathbf{x}' = \mathbf{a} + A \cdot \mathbf{x}, \quad \mathbf{a} \in \mathbb{R}^n, \quad A \in \mathbb{R}^{n \times n}, \det A \neq 0, \quad (1.63)$$

i.e., a composition of a linear automorphism and a translation. Clearly these transformations give rise to unique projective automorphisms of the projective extension of this affine space, because we can write (1.63) in the form

$$(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n)^T \mapsto \begin{bmatrix} 1 & \mathbf{o}^T \\ \mathbf{a} & A \end{bmatrix} \cdot (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n), \quad (1.64)$$

if $x_0 = 1$. Such a transformation maps proper points to proper points and ideal points to ideal points, so the ideal hyperplane is fixed. It is also easily verified that any projective transformation which maps ideal points to ideal points has a matrix of the form

$$\begin{bmatrix} 1 & \mathbf{o}^T \\ \mathbf{a} & A \end{bmatrix}.$$

Thus the group of affine transformations consists precisely of those projective transformations which leave the ideal hyperplane invariant. Because all hyperplanes of P^n are equal with respect to the projective transformation group, we can choose an arbitrary plane ω , and the affine space constructed above, endowed with the group of projective transformations which leave ω fixed, is isomorphic to the affine space with the affine transformation group (1.63). Such a transformation has, in an affine frame, the coordinate representation (1.63).

The group of affine transformations is denoted by GA_n , and algebraically, it is a semidirect product of the general linear group GL_n and the additive group \mathbb{R}^n of translations.

Let us apply an affine map (1.63) to the underlying affine frame: the origin \mathbf{o} is mapped to a point \mathbf{a} and the unit vectors \mathbf{e}_i are mapped to $\mathbf{e}'_i = A \cdot \mathbf{e}_i$. By Eqn. (1.63),

$$\mathbf{x}' = \mathbf{a} + A \cdot \mathbf{x} = \mathbf{a} + A \cdot (x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n) = \mathbf{a} + x_1 \mathbf{e}'_1 + \dots + x_n \mathbf{e}'_n,$$

i.e., the coordinates (x_1, \dots, x_n) of a point with respect to the original frame agree with the coordinates of the image point \mathbf{x}' with respect to the image frame $(\mathbf{a}; \mathbf{e}'_1, \dots, \mathbf{e}'_n)$. Hence, an affine map can be defined as *transformation via equal coordinates* with respect to two affine frames (see Fig. 1.26).

Affine Invariants

Properties of geometric objects which are invariant with respect to the action of the affine transformation group are called *affine invariants* or affine properties of those objects. Parallelity of lines is an affine invariant, orthogonality is not. Let us give some examples and look at them from projective space:

Parallel lines are those whose intersection point is an ideal point. The set of ideal points is invariant with respect to the affine transformation group GA_n . The notion of ‘line’ of course is, because it is even invariant with respect to PGL_n of which GA_n is a subgroup. Thus parallelity of lines is an affine property.

An *affine subspace* is a projective subspace not contained in the ideal hyperplane ω . This property is not affected by GA_n , so being an affine subspace is an affine property. Note that an affine subspace can mean two things: As a subset of projective space it may contain ideal points (it certainly does if its dimension is greater than zero). As a subset of affine space (its ‘affine part’) it of course does not. But the ideal points of an affine subspace are, when performing the projective extension, easily recovered as the ideal points of its lines, so there is no harm if we allow this imprecise way of speech. This also shows that the *dimension* of an affine subspace is an affine property.

Affine subspaces of dimensions n_1, n_2 , whose intersection consists of ideal points only, are called totally parallel. This is again an affine property. If their intersection consists of a k_1 -dimensional projective subspace, a k_2 -dimensional subspace of which is contained in ω , then their *degree of parallelity* is defined by $d = (k_2 + 1) / \min(n_1, n_2)$. The numbers k_1, k_2, d are again affine invariants.

The *ratio* of three collinear points is an affine invariant because of Equ. (1.22).

Further examples of affine invariants are the center of a quadric, the property of being a diameter line, and the classification of conics as ellipses, hyperbolae, and parabolae. In a projective plane there is only one conic, which means that all conics are projectively equivalent. Likewise we can show that in an affine plane there are only three conics (the ellipse, the hyperbola, and the parabola) if we identify all conics which are just affine transforms of each other: Given two ellipses c, c' it is easy to set up an affine transformation mapping c onto c' by choosing affine frames whose origins are the centers of c and c' , and whose unit vectors define conjugate diameters of c and c' . A similar argument shows that any two hyperbolae and any two parabolae are affinely equivalent.

Equi-affine Geometry

In \mathbb{R}^n there is the familiar notion of *volume* of a polyhedron, which is called *area* if $n = 2$ and *length* if $n = 1$. In general the affine transformation given by (1.63) changes this volume: The volume of the affine transform of a polyhedron is the original volume multiplied by $\det(A)$ (more precisely, Lebesgue measure of \mathbb{R}^n is transformed by multiplication with $\det(A)$).

Affine maps whose matrices A have determinant 1 form a subgroup of GA_n , which is called the *equi-affine group* or *group of volume preserving affine trans-*

formations and is denoted by the symbol SA_n . By its definition, the volume of a Lebesgue measurable subset of affine space is an *equi-affine invariant*.

The usual first and second order differential invariants of Euclidean geometry, such as arc length or curvature, are invariant with respect to Euclidean transformations. It turns out that it is possible to define differential invariants of higher order which are equi-affine invariants. They are the subject of *affine differential geometry* which should be called *equiaffine differential geometry*.

Note that the ratio of volumes is already a general affine invariant: If measurable sets M_1, M_2 have volume ratio $V(M_1)/V(M_2)$, their images possess the volume ratio $\det(A)V(M_1)/\det(A)V(M_2) = V(M_1)/V(M_2)$, so especially the property of two subsets having equal volume is an affine invariant.

Example 1.1.33. Consider an ellipse, a pair of conjugate diameters and the circumscribed parallelogram which touches the ellipse in their endpoints (see Fig. 1.22). If the ellipse is a circle, all these parallelograms are quadrangles of equal size and therefore of equal area. All ellipses are affinely equivalent, especially they are affinely equivalent to some fixed circle. Thus we have shown that all parallelograms circumscribed in this way have equal area.

This is also true for a hyperbola and can be shown by considering the special hyperbola with equation $y = 1/x$, where this property is easily verified by direct computation. The general case again follows from the fact that all hyperbolae are affinely equivalent (see Fig. 1.22). \diamond

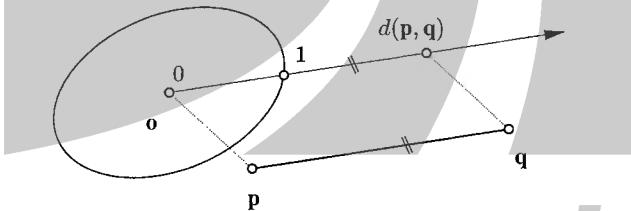


Fig. 1.27. Measuring Euclidean distance.

Euclidean Geometry Embedded in Affine Geometry

In order to specialize affine geometry further to *Euclidean geometry*, we define an ellipsoid Σ in affine space A^n and call it *unit sphere*. We use the unit sphere to define distances between points. If \mathbf{o} is the center of Σ , and \mathbf{p} is one of its points, then the distance of \mathbf{o} and \mathbf{p} is set to 1. If $\mathbf{q} = \mathbf{o} + \lambda(\mathbf{p} - \mathbf{o})$ then the center and \mathbf{q} have distance $|\lambda|$. The distance of all other pairs \mathbf{p}, \mathbf{q} of points is measured after translating them such that \mathbf{p} is transformed to \mathbf{o} (see Fig. 1.27).

In an affine coordinate system whose origin is \mathbf{o} , Σ has the equation

$$\mathbf{x}^T \cdot Q \cdot \mathbf{x} = 1, \quad (1.65)$$

where Q is a positive definite symmetric matrix. This matrix defines an inner product by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \cdot Q \cdot \mathbf{y}, \quad (1.66)$$

and induces a vector norm, which is used to compute the distance $d(\mathbf{a}, \mathbf{b})$ of two points \mathbf{a}, \mathbf{b} as

$$d^2(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|^2 = \langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle, \quad (1.67)$$

where the symbol $\|\mathbf{x}\|$ stands for the length of the vector \mathbf{x} . The inner product also serves for the introduction of *angles*. The angle ϕ between two vectors \mathbf{g}, \mathbf{h} is given by

$$\cos \phi = \frac{\langle \mathbf{g}, \mathbf{h} \rangle}{\sqrt{\langle \mathbf{g}, \mathbf{g} \rangle} \sqrt{\langle \mathbf{h}, \mathbf{h} \rangle}}. \quad (1.68)$$

The angle between lines is given by the angle of their direction vectors. In particular *orthogonality* is characterized by

$$\langle \mathbf{g}, \mathbf{h} \rangle = 0. \quad (1.69)$$

An affine map which preserves Euclidean distances is called a *Euclidean congruence transformation*. Properties invariant with respect to the Euclidean transformation group, called *Euclidean invariants*, are the contents of *Euclidean geometry*. The matrix A of a Euclidean congruence must preserve the inner product of two vectors and therefore satisfy the orthogonality condition

$$A^T \cdot Q \cdot A = Q. \quad (1.70)$$

If the matrix Q is the unit matrix, then the coordinate system is called *Cartesian coordinate system*. Such coordinate systems always exist and can be found by Gram-Schmid orthonormalization.

If we use a Cartesian coordinate system, the inner product coincides with the canonical inner product $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \cdot \mathbf{y}$, which will be written in the form $\mathbf{x} \cdot \mathbf{y}$. We obtain the familiar formulae for distance and angle. The matrix A of a Euclidean congruence transformation $\mathbf{x}' = \mathbf{a} + A \cdot \mathbf{c}$ has the simple characterization

$$A^T \cdot A = I, \quad (1.71)$$

which means that A is an orthogonal matrix. Orthogonal matrices form the *orthogonal group* O_n , an important subgroup of which is characterized by positive determinant ($\det A = 1$) and called *special orthogonal group* SO_n . If we use a Cartesian coordinate system, then the Euclidean congruence transformations are precisely those given by (1.63) with a matrix $A \in O_n$. This shows that the group OA_n of Euclidean congruence transformations is a semidirect product of O_n and the additive group \mathbb{R}^n of translations. If $A \in SO_n$, then the transformation is called *Euclidean motion* or *Euclidean displacement*. This group is denoted by SOA_n .

Finally, we mention that all Euclidean geometries which possibly can be constructed by the choice of different ellipsoids are isomorphic, because all ellipsoids are affinely equivalent. This justifies our speaking of *the* Euclidean geometry of n dimensions.

Euclidean Geometry Embedded in Projective Geometry

Assume that we have a Euclidean geometry embedded in an affine geometry, which in turn is embedded in a projective geometry. The unit sphere Σ , which is, from the projective point of view, just an oval quadric which has no intersection with the ideal plane, defines a polarity κ . The affine frame $(\mathbf{o}, \mathbf{e}_1, \dots, \mathbf{e}_n)$ where Σ has equation (1.65) defines a projective frame $(B_0, \dots, B_n; E)$ with $B_0 = \mathbf{o}$, $E = \mathbf{o} + \mathbf{e}_1 + \dots + \mathbf{e}_n$, and B_1, \dots, B_n are the ideal points of the lines $\mathbf{o} + [\mathbf{e}_1], \dots, \mathbf{o} + [\mathbf{e}_n]$. In this projective frame, Σ has the equation

$$\mathbf{x}^T \cdot C \cdot \mathbf{x} = 0 \quad \text{with} \quad C = \begin{bmatrix} -1 & \mathbf{o}^T \\ \mathbf{o} & Q \end{bmatrix}.$$

Assume that $A_\omega = (0, \mathbf{a})\mathbb{R}$ is an ideal point and the line L contains A_ω . The polar hyperplane $A_\omega\kappa$ is a diameter hyperplane of Σ and has the *affine* equation

$$\mathbf{x}^T \cdot Q \cdot \mathbf{a} = 0. \quad (1.72)$$

We see that this hyperplane is orthogonal to \mathbf{a} in the Euclidean geometry defined by Σ . It intersects ω in ideal points which belong to lines orthogonal to L . Furthermore, the matrix Q can be used to define a polarity in the ideal plane $\omega : x_0 = 0$ by

$$\mathbf{x}\mathbb{R} \mapsto Q \cdot \mathbf{x}\mathbb{R}, \quad (1.73)$$

where the homogeneous coordinate vector $\mathbf{x}\mathbb{R} = (x_1, \dots, x_n)\mathbb{R}$ corresponds to the ideal point $(0, x_1, \dots, x_n)\mathbb{R}$. Clearly this is just the polarity induced by κ in the ideal plane, and it does not define a quadric because Σ does not intersect ω . It is called the *absolute polarity* defined by Euclidean geometry, and there is the following theorem, which is not very deep at all, and only expresses a change of viewpoint. It should motivate the reader to think in a more ‘projective’ way instead of in an ‘affine’ one.

Theorem 1.1.52. *If Euclidean geometry is embedded in projective geometry with ideal hyperplane ω , there is an absolute polarity in ω , which is elliptic, and has the property that the orthogonality relation between affine subspaces G, H is expressed by the conjugacy of $G \cap \omega, H \cap \omega$ with respect to the absolute polarity.*

Proof. Equ. (1.72) shows that vectors \mathbf{a}, \mathbf{x} are orthogonal if and only if the corresponding ideal points $(0, \mathbf{a})\mathbb{R}$ and $(0, \mathbf{x})\mathbb{R}$ are conjugate with respect to the absolute polarity. Orthogonality of affine subspaces G, H means that all lines contained in G are orthogonal to all lines contained in H , which is nothing but the fact that all ideal points of G are conjugate to all ideal points of H . \square

In the complex extension of projective space there are no elliptic polarities, so also the absolute polarity defines a quadric. It is called the *absolute quadric* i_ω and coincides with the intersection of the unit sphere with ω . This leads to the following fact which is almost trivial but perhaps surprising if one realizes it for the first time:

All translates of the unit sphere, and also all scaled versions of it, intersect ω in the absolute quadric i_ω , because translations and homotheties leave ω point-wise fixed. The equation of the absolute quadric is

$$x_0 = \mathbf{x}^T \cdot Q \cdot \mathbf{x} = 0 \quad \text{with } \mathbf{x} = (x_1, \dots, x_n). \quad (1.74)$$

An imaginary line whose ideal point $(0, \mathbf{i})\mathbb{C}$ lies in i_ω is orthogonal to itself, and $\langle \mathbf{i}, \mathbf{i} \rangle = 0$. Such a line is called *isotropic* or *minimal line*. If we use the Euclidean distance formula (1.67) for complex points also, it turns out that points contained in the same isotropic line have zero distance.

A *homogeneous Cartesian coordinate system* can be defined as a projective frame $(B_0, \dots, B_n; E)$ such that $B_0 = \mathbf{o}$ and B_1, \dots, B_n are pairwise orthogonal ideal points. This means that any two B_i, B_j with $1 \leq i < j \leq n$ are conjugate with respect to the absolute polarity, and B_1, \dots, B_n is its auto-polar simplex. In addition, B_0, B_1, \dots, B_n is an auto-polar simplex of Σ .

The equation of i_ω is very simple in a homogeneous Cartesian coordinate system: It is given by

$$i_\omega : x_0 = x_1^2 + \dots + x_n^2 = 0. \quad (1.75)$$

Example 1.1.34. In order to introduce a Euclidean metric in the affine plane, we choose an ellipse Σ (the *unit circle*), as in Fig. 1.28. Then, conjugate diameters of the ellipse become orthogonal and can be used to define a Cartesian coordinate system, where Σ is given by the inhomogeneous equation $x^2 + y^2 = 1$.

Σ 's ideal points are complex conjugates of each other, and are denoted by J, \bar{J} . They comprise the absolute quadric i_ω , which consists of two points only (because the dimension of the ideal hyperplane ω equals one), and are called *absolute points*.

The homogeneous representation of Σ is $x_1^2 + x_2^2 - x_0^2 = 0$ and thus i_ω is given by $x_0 = x_1^2 + x_2^2 = 0$. The only solutions of this equation are the points $J = (0, 1, i)\mathbb{C}$ and $\bar{J} = (0, 1, -i)\mathbb{C}$.

All other circles $x^2 + y^2 + ax + by + c = 0$ intersect ω precisely in the set $\{J, \bar{J}\}$. The converse is also true: All real conics whose complex extension contains J, \bar{J} , are circles.

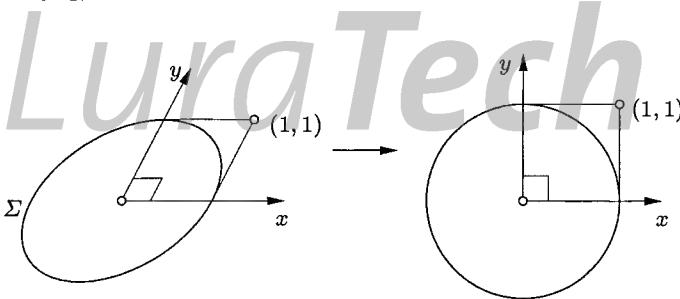


Fig. 1.28. Euclidean plane and its standard version.

The absolute polarity is an elliptic involution, whose complex extension has the fixed points J, \bar{J} . It maps an ideal point to its orthogonal ideal point. The minimal

lines concurrent in \mathbf{o} have equation $x \pm iy = 0$, so their union is given by the equation

$$(x - iy)(x + iy) = x^2 + y^2 = 0$$

Thus this pair of lines may be viewed as a circle of zero radius. These lines are also the fixed lines of the involution of conjugate diameters in \mathbf{o} (with respect to the polarity defined by Σ), and they are Σ 's tangents in J, \bar{J} . \diamond

Example 1.1.35. Perhaps the reader wonders why we introduce Euclidean geometry in such a complicated way. One reason is that it is often desirable to introduce more than just one metric, each defined by a different unit sphere. We give an example here, later we will show others.

Consider a set of data points (x_i, y_i) , $i = 1, \dots, N$, in \mathbb{R}^2 and associated values f_i . Then, the problem of *scattered data interpolation* is to find a smooth function f defined on \mathbb{R}^2 (or a subset containing the data points) which interpolates the given function values at the given data points: $f(x_i, y_i) = f_i$, $(i = 1, \dots, N)$. Depending on the application, it might not be clear which scales to use on x - and y -axis. However, there are many methods, such as the use of radial basis functions, or methods based on triangulations, which give different results if we scale the x - and y -coordinates by different factors (these methods are not affinely invariant), but do not change if we rotate or translate the data points.

To obtain affine invariance, G. Nielson [131] proposed the following approach: He constructs (by principal component analysis) an ellipse Σ which is connected in an affinely invariant way to the data set. This ellipse is used to define a Euclidean metric in exactly the way we have just described. Since the metric is tied in an affinely invariant way to the data set, we can use any Euclidean construction of an interpolant, which then will automatically depend on the given data in an affinely invariant way.

Equivalent to the introduction of a new Euclidean metric is to apply an affine transformation to the entire situation such that Σ becomes the ordinary unit circle, and then to apply procedures which are invariant with respect to Euclidean transformations. Fig. 1.29 shows an example of a Dirichlet triangulation constructed in an adapted Euclidean metric. \diamond

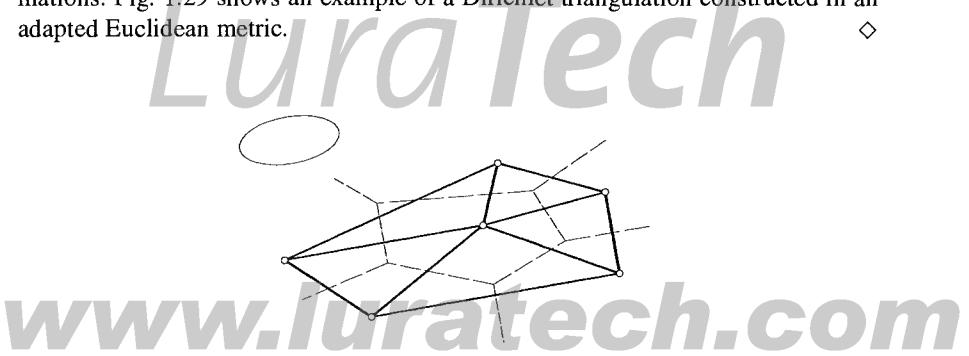


Fig. 1.29. Euclidean construction with respect to a modified unit circle (Dirichlet triangulation).

Example 1.1.36. In the projective extension of Euclidean three-space, the Euclidean metric defines an *absolute conic*. In a homogeneous Cartesian coordinate system it is given by

$$i_\omega : x_0 = x_1^2 + x_2^2 + x_3^2 = 0.$$

The situation is completely analogous to that of the Euclidean plane, which was discussed in Ex. 1.1.35. Quadrics which contain i_ω are the Euclidean spheres.

In Euclidean geometry, a planar section of a sphere is a circle. If the plane is tangent to the sphere, the intersection are two conjugate complex lines, i.e., a circle of zero radius as described in Ex. 1.1.34. The complex extension of a sphere is a ruled quadric, and the two conjugate complex lines mentioned above belong to the two reguli. \diamond

Equiformal Geometry

A *Euclidean similarity* σ is an affine map which preserves Euclidean orthogonality. This means that the matrix A of σ is a scalar multiple of an orthogonal matrix. These mappings are therefore compositions of a Euclidean congruence transformation and a homothetic transformation (central similarity) $\mathbf{x} \mapsto t\mathbf{x}$. From a projective point of view, a *similarity* is a projective map which maps the absolute quadric i_ω as a whole onto itself. It is not difficult to show the converse: If a projective automorphism maps i_ω onto itself, then it is a Euclidean similarity transformation. In coordinates, such a mapping has the form

$$\mathbf{x}' = \mathbf{a} + t \cdot A \cdot \mathbf{x}, \quad (1.76)$$

where A is the matrix of a Euclidean congruence. The group of Euclidean similarities is a subgroup of the affine transformation group. Properties which are invariant under Euclidean similarities are called *equiformal invariants* and form the content of *equiformal geometry*. Examples are: the angle between two lines, or the vertices of a quadric (as those points where the tangent plane is orthogonal to the diameter).

In fact, many properties and results which are often considered as typically Euclidean do not involve distance, but only the orthogonality relation, and should be classified as *equiformal*. Examples are many of the remarkable points in a triangle, such as the circumcenter, incenter, orthocenter, and others. Note that the barycenter is even an affine invariant.

Remark 1.1.14. Consider two lines G, H concurrent in a point P . The absolute points in the plane $G \vee H$ are denoted by J, \bar{J} . The lines $j = P \vee J$ and $\bar{j} = P \vee \bar{J}$ are the two minimal lines concurrent in P . The cross ratio of these four lines is connected with the angle $\phi = \sphericalangle(G, H)$ via the *Laguerre formula*

$$\phi = \frac{i}{2} \ln(\text{cr}(G, H, j, \bar{j})). \quad (1.77)$$

The verification is left to the reader as an exercise (use Equ. (1.19)). This result is not at all surprising when we know that a Euclidean similarity of the plane maps the set of absolute points onto itself, and thus maps minimal lines to minimal lines.

Now the angle is invariant under similarities, and the cross ratio is invariant under projective transformations, which motivates that there is some connection. \diamond

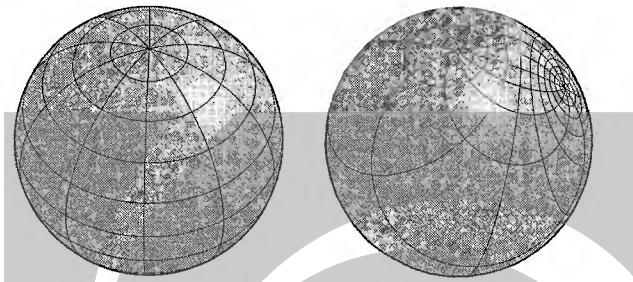


Fig. 1.30. Conformal mapping of a sphere determined by a projective automorphism.

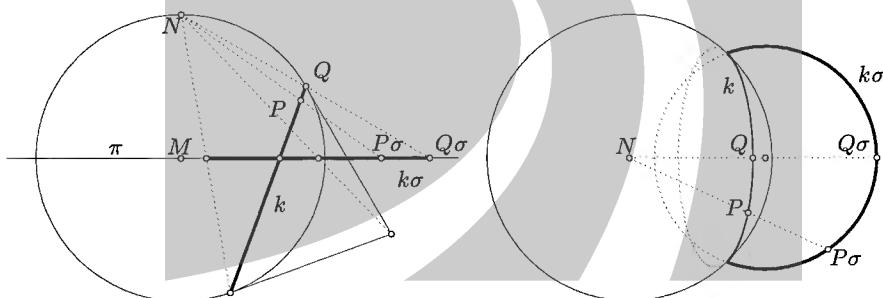


Fig. 1.31. Stereographic projection σ from the north pole N onto the equator plane π . The image of the circle k is a circle $k\sigma$. Left: front view. Right: top view onto π .

Example 1.1.37. As an exercise, the reader may prove the following result: a projective automorphism of a sphere, i.e., a projective automorphism of projective three-space, which maps a given sphere onto itself, is *conformal*, which means that the angle between curves on the sphere remains invariant. (see Fig. 1.30).

As a second exercise, consider the *stereographic projection*. It is the restriction of a central projection to the sphere — the center of the projection being the ‘north pole’ and the image plane being the ‘equatorial plane’ (see Fig. 1.31). Show that this projection is also conformal.

Hint: In both cases consider the complex extension and show that minimal lines are mapped to minimal lines, then use the Laguerre formula.

Both the projective automorphisms of the sphere and the stereographic projection have an additional property: They map circles to circles or possibly lines. For the former this is easily shown by the fact that planes are transformed into planes, and so planar sections of the sphere into planar sections. The reader may try to prove this fact for the stereographic projection as an exercise (or see the paragraph preceding Equ. (8.21)).

If we map a projective automorphism of the sphere to the plane by a stereographic projection from the north pole, this gives a transformation of the Euclidean plane where exactly one point has no image (the corresponding sphere point is mapped to the north pole), and circles/lines are mapped to circles/lines. Such a transformation is called *Möbius transformation*. \diamond

Cayley-Klein Geometries

We have seen that affine and equiformal transformations are precisely those projective transformations which fix some ‘absolute figure’. In the former case it is the ideal hyperplane, in the latter the absolute quadric.

There are transformation groups (such as the groups of Euclidean or equi-affine transformations), which cannot be characterized in this way.

Nevertheless, there is a wide variety of remarkable geometries (the so-called *Cayley-Klein geometries*) which are defined as those subgroups of PGL_n which fix some ‘absolute figure’ (cf. [61]). We give a few prominent examples:

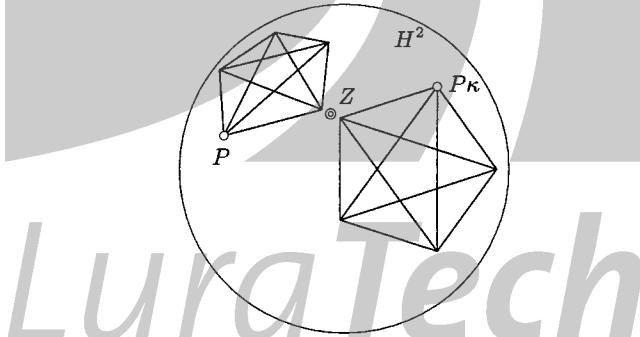


Fig. 1.32. Projective model of the hyperbolic plane H^2 and hyperbolic rotation κ about the center Z .

Example 1.1.38. Consider the interior of a circle c in the Euclidean plane E^2 . It is called *hyperbolic plane* H^2 . The group of hyperbolic transformations is defined as the group of projective automorphisms of the circle c . All of them map the interior of the circle onto itself in a one-to-one way, because an interior point is distinguished from an exterior one by the non-existence of a circle tangent incident with it. Such transformations are called *hyperbolic congruence transformations* (see Fig. 1.32).

The theory of invariants of this group acting on H^2 is *planar hyperbolic geometry*. More precisely, we have described the so-called projective model of the hyperbolic plane. This works in higher dimensions as well — we may use any oval quadric as an absolute figure. \diamond

Remark 1.1.15. The planar Euclidean and hyperbolic geometries share many properties which have been important in the history of mathematics. Starting with Euclid, mathematicians have tried to characterize *Euclidean geometry* in a unique way by ‘axioms’ (those of Euclid are, according to modern standards, no axioms at all). One starts by describing how ‘points’ and ‘lines’ should behave, and what a ‘congruence transformation’ should be. It turns out that there is something called *absolute geometry* which has several different models, but is rich enough that interesting theorems can be proved, and it turns out that both Euclidean and hyperbolic geometry are instances of absolute geometry.

It is possible to formulate a set of axioms such that by adding just one axiom (the *parallel axiom*) or its negation, we get either Euclidean or hyperbolic geometry [172]. This answers the ancient question whether Euclid’s parallel postulate is a consequence of his other axioms or not. This was first shown by C.F. Gauss (1777–1855), who did not publish his results, and independently by J. Bolyai (1802–1860) and N.I. Lobachevskij (1792–1856). The Euclidean parallel axiom says: Given a line L and a point $P \notin L$, there exists exactly one line containing P , which does not intersect L . This is not true in hyperbolic geometry: There is always an infinite number of straight lines which do not meet L , as is clearly seen in the projective model discussed above, where a hyperbolic line appears as a straight line segment bounded by the absolute circle. \diamond

Example 1.1.39. Another example of a Cayley-Klein geometry is *elliptic geometry*: We construct a projective space P^n whose points are the one-dimensional subspaces of a real vector space, which is equipped with a positive definite scalar product and norm

$$\langle x, y \rangle = x^T \cdot C \cdot y, \quad \|x\| = \sqrt{\langle x, x \rangle}. \quad (1.78)$$

If two vectors a, b are orthogonal, i.e., $\langle a, b \rangle = 0$, we say $a\mathbb{R}, b\mathbb{R}$ are *orthogonal points*. The equation $\langle a, x \rangle = 0$ is the equation of the hyperplane orthogonal to the point $a\mathbb{R}$. The elliptic *distance* of points $a\mathbb{R}, b\mathbb{R}$ is defined as the *angle* of the corresponding one-dimensional subspaces:

$$\cos d(a\mathbb{R}, b\mathbb{R}) = \left| \frac{\langle a, b \rangle}{\|a\| \|b\|} \right|. \quad (1.79)$$

The distance assumes values in the interval $[0, \pi/2]$. Elliptic congruence transformations are mappings which leave the elliptic distance invariant: It turns out that all of them are projective automorphisms $\kappa : x\mathbb{R} \mapsto (A \cdot x)\mathbb{R}$, where the matrix A has the property $A^T C A = \lambda C$.

The same definitions except (1.79) apply to points of the complex extension of P^n as well. As the scalar product is no longer positive definite, if we allow complex

arguments, there are now points $x\mathbb{R}$ with $\langle x, x \rangle = 0$. This condition defines a quadric whose polarity is governed by the matrix C .

An elliptic congruence transformation must leave this quadric invariant, because $x^T \cdot C \cdot x = 0$ implies that $(A \cdot x)^T \cdot C \cdot (A \cdot x) = x^T \cdot A^T \cdot C \cdot A \cdot x = \lambda x^T \cdot C \cdot x = 0$. Conversely, if a projective automorphism $x\mathbb{R} \mapsto (A \cdot x)\mathbb{R}$ leaves the quadric $\langle x, x \rangle = 0$ invariant, it is not difficult to show that necessarily $A^T \cdot C \cdot A = \lambda \cdot C$, with $\lambda \in \mathbb{C}$.

This shows that elliptic geometry is a Cayley-Klein geometry, where the ‘absolute figure’ is the quadric $\langle x, x \rangle = 0$ and the geometric transformations are those which leave the absolute figure invariant.

A familiar instance of an elliptic geometry is the Euclidean geometry of a bundle of lines in E^n . \diamond

Example 1.1.40. An example of an elliptic geometry we already know is the hyperplane ω at infinity of the projective extension P^n of a Euclidean space E^n .

The proper points of P^n have the form $x\mathbb{R} = (1, x)\mathbb{R}$ with $x \in \mathbb{R}^n$, and the ideal points have coordinates $x\mathbb{R} = (0, x)\mathbb{R}$. A Euclidean congruence transformation has the form $(1, x)\mathbb{R} \mapsto (1, a + A \cdot x)\mathbb{R}$, where A is an orthogonal matrix ($A^T \cdot A$ equals the identity matrix, cf. Equ. (1.71)). Analogously a Euclidean similarity transformation has the form $(1, x)\mathbb{R} \mapsto (1, a + \lambda A \cdot x)\mathbb{R}$, according to Equ. (1.76). We extend this definition to all points of P^n by letting

$$x\mathbb{R} \mapsto \left(\begin{bmatrix} 1 & 0 \\ a & \lambda A \end{bmatrix} \cdot x \right) \mathbb{R}, \quad (1.80)$$

where $\lambda \neq 0$, $a \in \mathbb{R}^n$, and $A^T \cdot A$ equals the identity matrix. In particular, ideal points are mapped according to

$$(0, x)\mathbb{R} \mapsto (0, \lambda A \cdot x)\mathbb{R} = (0, A \cdot x)\mathbb{R}.$$

This shows that the ideal hyperplane, which is itself a projective space of $n - 1$ dimensions, is equipped with an elliptic geometry according to Ex. 1.1.39, where the matrix C of Equ. (1.78) is the identity matrix. The ‘absolute figure’ of this elliptic geometry is the set of points $(0, x)\mathbb{R}$ of ω which satisfy $x^T \cdot x = 0$, i.e., coincides with the absolute quadric i_ω . \diamond

1.2 Basic Projective Differential Geometry

1.2.1 Curves

We are going to study curves and surfaces in projective space. We know what a curve/surface in the linear space \mathbb{R}^{n+1} is, and we also know that all nonzero vectors of \mathbb{R}^{n+1} correspond to projective points. Thus it seems natural to define a *curve* of P^n as a mapping $c : I \rightarrow P^n$, where I is some parameter interval, and

$$c(t) = c(t)\mathbb{R}, \quad \text{with} \quad c(t) = (c_0(t), c_1(t), \dots, c_n(t)) \neq (0, \dots, 0) \quad (1.81)$$

for all $t \in I$. If all coordinate functions $c_i(t)$ are C^r (which we use short for r times continuously differentiable) we say that $c(t)$ is at least r times differentiable. *Closed* curves can be defined either as smooth mappings of the unit circle into projective space, or as *periodic* mappings of the real line. Here we only study local properties of curves, so we will only use curves defined in an interval.

Normalization

Clearly multiplication of $c(t)$ with a scalar function $\rho(t)$ does not change $c(t)$, so there are a lot of different homogeneous coordinate functions for the same curve. Many of them are not differentiable, even not continuous, and nevertheless describe the same curve c .

We study the curve c in the neighbourhood of some curve point $c(t_0) = (c_0(t_0), \dots, c_n(t_0))\mathbb{R}$. Some $c_i(t_0)$ is nonzero, and without loss of generality we assume that $c_0(t_0)$ is nonzero. If the c_i are continuous, there is a neighbourhood of t_0 such that $c_0(t)$ is nonzero, and in this neighbourhood we can write

$$c(t) = (1, c_1(t)/c_0(t), \dots, c_n(t)/c_0(t))\mathbb{R}.$$

If we ignore the first coordinate, this is just a curve $\bar{c}(t)$ in the affine space with equation $x_0 \neq 0$ embedded in P^n . As such it is a well known object. We define the differentiability class of c at t_0 as the differentiability class of the curve \bar{c} at t_0 . It is easily verified that the differentiability class does not change if we use a different coordinate for normalization instead of x_0 .

Remark 1.2.1. The reader who is familiar with the concept of differentiable manifolds perhaps thinks that before defining curves we should have introduced a differentiable manifold structure on a projective space. Indeed, the normalization $(c_0(t), \dots, c_n(t))\mathbb{R} \rightarrow (1, c_1(t)/c_0(t), \dots, c_n(t)/c_0(t))$ does just this; there are $n+1$ different affine charts defined in this way. ◇

A property of the homogeneous coordinate functions, which is a property of the curve, must be invariant with respect to re-normalizations, which means multiplication with a scalar function $\rho(t)$ as described above. Calculations with derivatives have to be handled with care, because $\rho(t)$ may be ill behaved. We may visualize a renormalization as follows: Let us view P^n as bundle in \mathbb{R}^{n+1} , i.e., each point is seen as line through the origin of \mathbb{R}^{n+1} . Then, a curve in P^n is represented by a cone with vertex at the origin. If $c(t) = c(t)\mathbb{R} = \bar{c}(t)\mathbb{R}$, then both $c(t)$ and $\bar{c}(t)$ are curves contained in the same cone (see Fig. 1.33).

Invariants of Projective Differential Geometry

Projective differential geometry deals with properties of curves which are invariant with respect to regular re-parametrizations, and projective transformations. Invariance with respect to a projective transformation κ can mean different things: If we assign a subset of projective space to a curve c , such as its tangent T , then invariance means the following: The curve $\kappa \circ c$ has the tangent $T\kappa$.

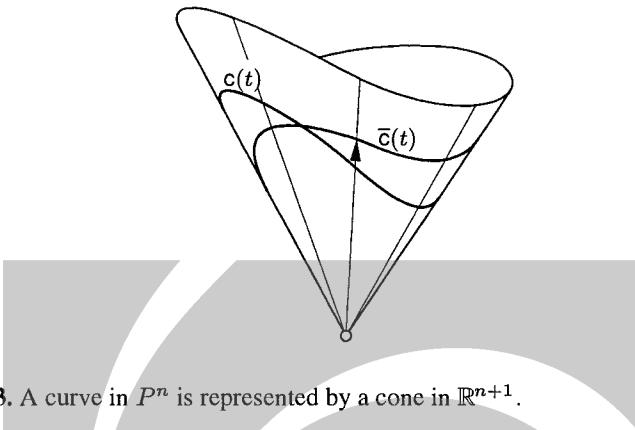


Fig. 1.33. A curve in P^n is represented by a cone in \mathbb{R}^{n+1} .

If we assign to a curve a numerical invariant, such as the dimension of its first osculating subspace, then invariance means that the dimension assumes the *same* value for the curve $\kappa \circ c$.

If we assign to a curve an invariant which is contained in some space which is *naturally* defined by projective space, and the group PGL_n of projective transformations induces a group G of transformations there in a *natural* way, then invariance means that the invariant must transform with κ 's natural image in G .

An example of this is dual projective space: A projective transformation naturally induces a transformation in the set of hyperplanes.

Curve Tangents and Singularities

Assume that $c(t) = (c_0(t), \dots, c_n(t))$ is differentiable, and that its first derivative $\dot{c}(t)$ is nonzero for all t . This means that $c(t)$ is a *regular curve* in \mathbb{R}^{n+1} .

If $c_0(t_0) \neq 0$, then we can look at the affine curve $\bar{c} = (1, c_1/c_0, \dots, c_n/c_0)$, whose derivative is given by

$$\dot{\bar{c}} = -\frac{\dot{c}_0}{c_0^2}c + \frac{1}{c_0}\dot{c},$$

which is zero if and only if $\dot{c}_0c = c_0\dot{c}$, or, equivalently, if and only if c, \dot{c} are linearly dependent. This means that either $\dot{c} = 0$ or that $c\mathbb{R}, \dot{c}\mathbb{R}$ are the same point. We summarize these facts in the following definition:

Definition. If $c : I \rightarrow P^n$ is a curve in projective space with $c(t) = c(t)\mathbb{R}$, we denote the point $\dot{c}(t)\mathbb{R}$ its first derivative point $c^1(t)$ at the parameter value t , if it exists, i.e., $\dot{c}(t) \neq 0$.

The curve c is called *regular* at t , if $c^1(t) \neq c(t)$. Then the curve tangent at $c(t)$ is the line $c(t) \vee c^1(t)$. Otherwise it is said to have a *singularity* at t .

Remark 1.2.2. If $\dot{c}(t) = 0$, we let $c^1(t) = \{\}$, $c^1(t)$ in that case being a projective subspace of dimension -1 , not of dimension 0. The curve tangent then obviously is

just 0-dimensional, and no line. It nevertheless equals the projective subspace which corresponds to the linear span of $c(t), \dot{c}(t)$. \diamond

Theorem 1.2.1. *The property of being a regular point and the curve tangent are projective differential invariants.*

Proof. The derivative of $\rho(t)c(t)$ is given by $\dot{\rho}(t)c(t) + \rho(t)\dot{c}(t)$, so the linear span of $c(t), \dot{c}(t)$ is not altered by re-normalization.

If ϕ is a regular differentiable mapping $I \rightarrow I'$ used for re-parametrization, we calculate $\frac{d}{dt}(c(\phi(t))) = \dot{c}(\phi(t))\dot{\phi}(t)$ and see that the linear span of c, \dot{c} is not altered by re-parametrization.

If κ is the projective transformation given by $x\mathbb{R} \mapsto (A \cdot x)\mathbb{R}$, then the curve $c\kappa$ is parametrized by $t \mapsto (A \cdot c)\mathbb{R}$. This implies that $(c\kappa)^1 = (c^1)\kappa$.

Thus we have shown that both the regularity of a curve at a certain parameter value, and the curve tangent there are projective differential invariants. \square

Remark 1.2.3. The reader familiar with differentiable manifolds knows how to define tangent vectors of curves, which are elements of the projective space's tangent bundle. We are content with defining the curve's tangent, thus rather employing the linear incidence structure of projective space than its differentiable structure. \diamond

Remark 1.2.4. A line bundle in Euclidean \mathbb{R}^{n+1} is a model of projective n -space. A curve in the line bundle is a *cone*. If the vertex of the line bundle is the origin, we may describe the cone by a projective curve $c(t) = c(t)\mathbb{R}$. The line $c(t)\mathbb{R}$ is a generator line of the cone. The tangent plane of the cone at this generator, which is, per definition, the projective curve tangent, is spanned by the generator $c(t)$ and its first derivative generator $c^1(t)$. A generator $c(t)$ is regular if and only if $c(t), \dot{c}(t)$ are linearly independent. It is easily shown that a non-regular generator indeed corresponds to a surface singularity in the usual sense (see Fig. 1.34) \diamond

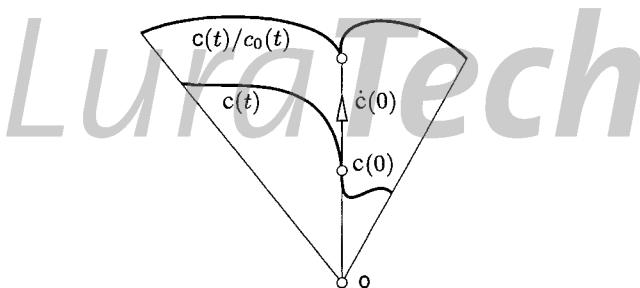


Fig. 1.34. Regular curve in \mathbb{R}^{n+1} representing a curve with singularity in P^n .

Convergence in Projective Space

The tangent is the limit of a line connecting neighbouring points: In order to state this familiar statement precisely, we first have to define *convergence* in the set of lines. This is done easily after introduction of a ‘point model’ for the set of lines, which is one of the main topics of this book. Meanwhile, we show that for points $c(t), c(t + \Delta t)$ there is a point on the line $c(t) \vee c(t + \Delta t)$ which converges to $c^1(t)$.

Convergence of points P_1, P_2, \dots towards a point P means that there are vectors p_1, p_2, \dots and p such that $P_i = p_i\mathbb{R}$ and $P = p\mathbb{R}$ and the series p_i converges to p . The same definition is used for a continuous family $c(t)$ of points.

Now we have

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} (c(t + \Delta t) - c(t)) = \dot{c}(t),$$

so the point $(c(t + \Delta t) - c(t))\mathbb{R}$ converges to the point $\dot{c}(t)\mathbb{R} = c^1(t)$, if it exists.

Remark 1.2.5. The definition of convergence in a projective space is in accordance with the definition of a *topology* on P^n : There is the mapping $\pi : p \rightarrow p\mathbb{R}$ which assigns to a vector the corresponding projective point. Now a subset of P^n is called *open* if and only if its π -preimage is open as a subset of $\mathbb{R}^{n+1} \setminus \{0\}$. \diamond

Remark 1.2.6. After embedding affine space \mathbb{R}^n in projective space P^n by sending (x_1, \dots, x_n) to $(1, x_1, \dots, x_n)\mathbb{R}$, we can try to compute the projective curve tangent of an affine curve $c : I \rightarrow \mathbb{R}^n$, which gives rise to a projective curve $c(t) = (1, c(t))$: We obtain $c^1(t) = (0, \dot{c}(t))\mathbb{R}$ and see that the projective curve tangent agrees with the usual tangent, as it is spanned by $c(t)$ and the ideal point $c^1(t)$ which indicates the direction $\dot{c}(t)$. \diamond

Osculating Spaces of a Curve

Assume a curve $c : I \rightarrow P^n$ with $c(t) = c(t)\mathbb{R}$. If $c(t)$ is sufficiently differentiable, the flag of linear subspaces spanned by

$$[c(t)], \quad [c(t), \dot{c}(t)], \quad [c(t), \dot{c}(t), \ddot{c}(t)], \quad \dots$$

defines projective subspaces

$$c(t), \quad c(t) \vee c^1(t), \quad c(t) \vee c^1(t) \vee c^2(t), \quad \dots$$

$$\text{where } c(t) = c(t)\mathbb{R}, \quad c^1(t) = \dot{c}(t)\mathbb{R}, \quad c^2(t) = \ddot{c}(t)\mathbb{R}, \quad \dots$$

The projective subspace $c(t) \vee \dots \vee c^k(t)$ is called the *osculating subspace of order k*. The dimensions of the osculating subspaces, which clearly are a nondecreasing sequence starting with zero, cannot exceed the dimension of entire space.

Theorem 1.2.2. *The projective osculating subspaces of order k are projective differential invariants.*

Proof. We have to show that the linear span of the first k derivative vectors of $c(t)$ is not affected by re-normalization and re-parametrization. Further we have to show that when applying a linear transformation to $c(t)$, they transform accordingly.

The former is easily done and completely analogous to the case $k = 1$, the latter is trivial. \square

The theorem shows that the sequence of dimensions of the osculating subspaces is a projective differential invariant. There are special names for curve points where this sequence behaves in certain ways:

Definition. *There are the following names for curve points, depending on the sequence of dimensions of osculating subspaces:*

| | |
|--------------------|---|
| 0, 1 | <i>regular point</i> |
| 0, 1, ..., $n - 1$ | <i>point of main type</i> |
| 0, 1, 1 | <i>inflection point (ordinary inflection if 0, 1, 1, 2)</i> |
| 0, 1, 1, 1 | <i>flat point</i> |
| 0, 1, 2, 2 | <i>point with stationary osculating plane (handle point).</i> |
| 0, 0 | <i>cusp (ordinary cusp if 0, 0, 1)</i> |

In order to illustrate the connections between the curve $c(t)$ in \mathbb{R}^{n+1} and the curve $c(t)\mathbb{R}$ in P^n , we study the case of an inflection point in more detail: Such a point is characterized by

$$\text{rank}(c(t_0), \dot{c}(t)) = \text{rank}(c(t_0), \dot{c}(t_0), \ddot{c}(t_0)) = 2. \quad (1.82)$$

This means that either the plane $c(t_0) + [\dot{c}(t_0), \ddot{c}(t_0)]$, which is the osculating plane of elementary differential geometry, contains the origin o , or that it is not a plane at all, i.e., $\dot{c}(t_0)$ and $\ddot{c}(t_0)$ are linearly dependent, so $c(t_0)$ is an inflection point in the sense of elementary differential geometry. For the case of the projective plane, this is illustrated in Fig. 1.35, where the affine part of P^2 is visualized as the plane $x_0 = 1$ in \mathbb{R}^3 .

This figure looks very much like a central projection from P^3 onto a plane, and motivates the following:

Proposition 1.2.3. *Consider a curve in P^3 and a point p which lies in the osculating plane at $c(t_0)$, but not in the curve tangent. Then projecting the curve from the center p onto a plane π yields a curve c' with an inflection point at $c'(t_0)$.*

Proof. We choose a projective coordinate system such that $p = (1, 0, 0, 0)\mathbb{R}$, π has the equation $x_0 = x_3$, and $c(t_0) = (1, 0, 0, 1)\mathbb{R}$. Then the projection has the equation $(1, x, y, z)\mathbb{R} \mapsto (1, x/z, y/z, 1)\mathbb{R}$, or, equivalently $(x_0, x_1, x_2, x_3)\mathbb{R} \mapsto (x_3, x_1, x_2, x_3)\mathbb{R}$.

If $c(t) = c(t)\mathbb{R} = (c_0(t), \dots, c_3(t))\mathbb{R}$, then the projection equals $c'(t) = c'(t)\mathbb{R}$ with $c' = (c_3, c_1, c_2, c_3)\mathbb{R}$. The center of the projection is not contained in c 's tangent at $t = t_0$, so the vectors $(1, 0, 0, 0)$, $c(t_0)$, $\dot{c}(t_0)$ are linearly independent. But it is contained in c 's osculating plane, so $(1, 0, 0, 0)$, $c(t_0)$, $\dot{c}(t_0)$, $\ddot{c}(t_0)$ are linearly dependent.

Elementary linear algebra shows (i) that $\{c'(t_0), \dot{c}'(t_0)\}$ is linearly independent, and (ii) that $\{c'(t_0), \dot{c}'(t_0), \ddot{c}'(t_0)\}$ is linearly dependent. This implies (i) that c' is regular at $t = t_0$ and (ii) that it has an inflection point there. \square

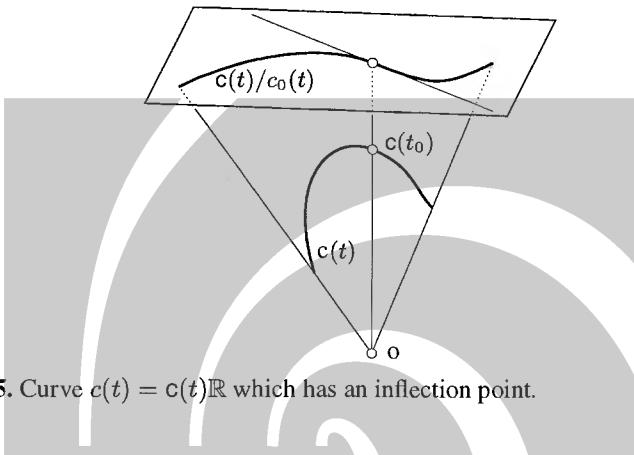


Fig. 1.35. Curve $c(t) = c(t)\mathbb{R}$ which has an inflection point.

Example 1.2.1. In the Euclidean plane there is the semicubic parabola with equation

$$y^2 = x^3.$$

It has the parametrization (t^2, t^3) . We re-parametrize this curve by $u = 1/t$. This is a regular parameter transformation only if $t \neq 0$. The curve is now described by $u \mapsto (1/u^2, 1/u^3)$ ($u \neq 0$). After embedding the Euclidean plane in the projective plane we can write the curve in the form $u \mapsto (1, 1/u^2, 1/u^3)\mathbb{R} = (u^3, u, 1)\mathbb{R}$ ($u \neq 0$). It is continuously extended by the curve $u \mapsto c(u) = (u^3, u, 1)\mathbb{R}$ ($u \in \mathbb{R}$). The point $c(0) = (0, 0, 1)\mathbb{R}$ is called the ideal point of the original curve. It is an inflection point because $c^1(u) = (3u^2, 1, 0)\mathbb{R}$, $c^2(u) = (6u, 0, 0)\mathbb{R}$, and so $\dim(c(0) \vee c^1(0)) = \dim(c(0) \vee c^1(0) \vee c^2(0)) = 1$.

Note that the (projective extension of the) semicubic parabola is projectively equivalent to the (projective extension of the) cubic parabola $y = x^3$. This is seen from the projective mapping κ defined by $(x_0, x_1, x_2)\mathbb{R} \mapsto (x_2, x_1, x_0)\mathbb{R}$. In the future we skip ‘projective extension of’, if the curves in question can be extended uniquely by continuity.

κ interchanges the point $(0, 0)$ with the ideal point $(0, 0, 1)\mathbb{R}$. This shows that both the cubic parabola and the semicubic parabola possess a cusp and an inflection point, which are located at these two points. A further curve projectively equivalent to both of them is shown in Fig. 1.36. \diamond

Remark 1.2.7. The reader perhaps has noticed that we used the term ‘curve’ for an object more general than just a mapping $c : I \rightarrow P^n$. We have actually considered

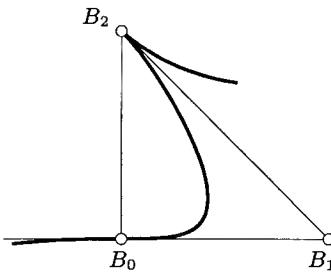


Fig. 1.36. Planar cubic with cusp and inflection point.

closed curves which consist of a curve $c : \mathbb{R} \rightarrow P^n$ together with an additional point which becomes ‘visible’ after the parameter transform $t \mapsto 1/t$, and extending the curve continuously. We could say that the domain of the mapping c is the projective line, which immediately leads us to such things as submanifolds of differentiable manifolds on the one hand (we will not pursue this) and rational mappings on the other hand (see Sec. 1.3.1). \diamond

Contact Order of Curves

A central concept of projective differential geometry is that of contact order:

Definition. Two curves $c(t)$, $d(u)$ are said to have contact of order k at parameter values t_0 , u_0 , if there exists a regular parameter transform $u = u(t)$ with $u_0 = u(t_0)$ and curves $c(t)$, $d(u)$ in \mathbb{R}^{n+1} such that $c(t) = c(t)\mathbb{R}$, $d(u) = d(u)\mathbb{R}$, and

$$\frac{d^l}{dt^l}c(t_0) = \frac{d^l}{dt^l}d(u(t_0)) \quad \text{for } l = 0, \dots, k. \quad (1.83)$$

The order of contact of two curves is the maximum k such that the curves have contact of order k there.

This definition of contact of order k obviously is projectively invariant (more generally, it can be shown that contact of order k is invariant with respect to C^k diffeomorphisms).

Contact of order k between curves is clearly a reflexive and symmetric relation. It is also *transitive* in the following sense: If curves c_1 and c_2 as well as curves c_2 and c_3 have contact of order k in the same point p , then also the curves c_1 and c_3 have contact of order k at p .

It is not difficult to show how contact of order k is expressed in an affine normalization of projective curves, e.g. $c(t) = (1, c_1(t), \dots, c_n(t))\mathbb{R}$, $d(u) = (1, d_1(u), \dots, d_n(u))$: c, d have contact of order k if and only if there is a regular parameter transform $u = u(t)$ such that the affine curves $(c_1(t), \dots, c_n(t))$ and $(d_1(u(t)), \dots, d_n(u(t)))$ share the same derivatives up to order k .

Example 1.2.2. We study the behaviour of the graph of the monomial

$$y = x^n, \quad n \geq 3,$$

at the origin $(0, 0)$. The curve tangent at the origin has the inhomogeneous parameterization $t \mapsto (t, 0)$.

The inhomogeneous parameter representation $\mathbf{c}(t) = (t, t^n)$ immediately shows that it has contact of order $n - 1$ with this tangent.

The point $\mathbf{c}(0)$ is also an inflection point, which follows from $\frac{d^2y}{dx^2} = 0$. The local behaviour of the curve at $(0, 0)$ needs not be the one we expect from an ordinary inflection point, namely that the curve changes the side of the tangent. This is true only if n is odd (see Fig. 1.37, where contact of order n with the tangent is visualized as limit of $n + 1$ common points of curve and tangent). Curves with a Taylor

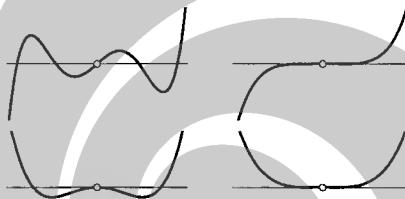


Fig. 1.37. Local behaviour at higher inflection points.

expansion of the form

$$y = a_n x^n + a_{n+1} x^{n+1} + \dots$$

show the same behaviour: Locally they change from one side of the tangent to the other if and only if n is odd. \diamond

Example 1.2.3. Consider a curve $\mathbf{c}(t)$ in Euclidean n -space E^n .

The *Frenet frame* at a point of main type $\mathbf{c}(s)$ is an ortho-normal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$, such that $\mathbf{e}_1, \dots, \mathbf{e}_j$ span the j -th osculating subspace ($j < n$). We temporarily let $v = \|\dot{\mathbf{c}}\|$. It is well known that then

$$\begin{aligned} v\dot{\mathbf{e}}_1 &= \kappa_1 \mathbf{e}_2, & v\dot{\mathbf{e}}_2 &= -\kappa_1 \mathbf{e}_1 + \kappa_2 \mathbf{e}_3, \dots \\ v\dot{\mathbf{e}}_{n-1} &= -\kappa_{n-2} \mathbf{e}_{n-2} + \kappa_{n-1} \mathbf{e}_n, & v\dot{\mathbf{e}}_n &= -\kappa_{n-1} \mathbf{e}_{n-1}. \end{aligned}$$

These equations are called the *Frenet equations*. The Frenet frame is uniquely determined by the requirements $\mathbf{e}_1 = \dot{\mathbf{c}}/\|\dot{\mathbf{c}}\|$, $\kappa_1, \dots, \kappa_{n-2} > 0$ if $n > 2$, and $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) > 0$.

If $v = 1$, i.e., the curve is parametrized with its arc length, we indicate differentiation with a prime. In this case the $n - 1$ functions $\kappa_1, \dots, \kappa_{n-1}$ determine the curve uniquely up to a Euclidean congruence transformation. The Frenet equations for E^2 and E^3 read as follows:

$$\begin{aligned} \mathbf{e}'_1 &= \kappa \mathbf{e}_2, & \mathbf{e}'_2 &= -\kappa \mathbf{e}_1 \quad (n = 2), \\ \mathbf{e}'_1 &= \kappa \mathbf{e}_2, & \mathbf{e}'_2 &= -\kappa \mathbf{e}_1 + \tau \mathbf{e}_3, & \mathbf{e}'_3 &= -\tau \mathbf{e}_2 \quad (n = 3). \end{aligned}$$

κ and τ are *curvature* and *torsion* of the curve. $\mathbf{e}_1, \mathbf{e}_2$ are called unit tangent vector and *principal normal vector*. In E^3 , $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$ is called the *binormal vector*.

Assume that \mathbf{c} is parametrized with arc length and consider a second curve \mathbf{d} which has contact of order 2 with \mathbf{c} at $\mathbf{c}(s_0)$. Then, there exists a parametrization $\mathbf{d}(t)$ with

$$\mathbf{d}(t_0) = \mathbf{c}(s_0), \quad \dot{\mathbf{d}}(t_0) = \mathbf{e}_1(s_0), \quad \ddot{\mathbf{d}}(t_0) = \kappa \mathbf{e}_2(s_0).$$

This shows that necessarily also $\dot{\mathbf{d}}(t_0)$ is a unit vector, and that $\ddot{\mathbf{d}}(t_0)$ is orthogonal to $\dot{\mathbf{d}}(t_0)$. All parametrizations which achieve this, agree with the arc length parametrization up to differentiation order 2.

Further we see that that second order contact implies equality of the curvature and the first two Frenet vectors at the contact point. In particular, a curve \mathbf{c} has second order contact with its osculating circle in all points which are not inflection points (in which case the curve has second order contact with the curve tangent).

The osculating circle may be defined as the limit of the circumcircle of points $\mathbf{c}(t), \mathbf{c}(t + \Delta t_1), \mathbf{c}(t + \Delta t_2)$ where $\Delta t_1, \Delta t_2$ converge to zero independently. This motivates the expression that curves having second order contact have three neighbouring points in common. \diamond

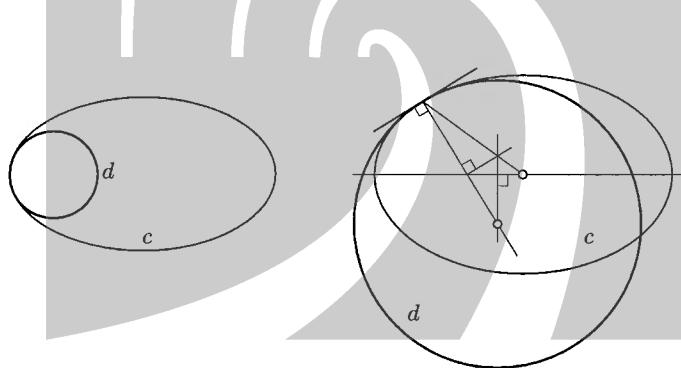


Fig. 1.38. Osculating circle d of an ellipse c in a vertex (left) and in a general point (right).

Remark 1.2.8. We do not recommend to perform proofs by using just the visualization of k -th order contact as limit case of $k+1$ neighbouring intersection points. For qualitative considerations it is however very useful. Let us give an example. What is the locally best fitting conic to a *planar* curve at a point $c(t)$? A conic is determined by 5 points. Thus we might expect the existence of a conic ('passing through 5 neighbouring points') which has fourth order contact with a curve c at $c(t)$.

A conic has no inflection points and thus this can only happen if $c(t)$ is no inflection point of c . The reader may try to prove that for all curves with sufficient differentiability there actually exists such a conic, which is a projective differential

invariant. It is called the *projective osculating conic* and plays an important role in projective differential geometry of planar curves.

Note that planarity of c is essential. Otherwise already four neighbouring points would not be planar and could not be interpolated by a conic. Thus, a twisted curve can in general have at most second order contact with a conic. The osculating circle of Euclidean differential geometry is an example of such a conic (see Fig. 1.38). \diamond

Remark 1.2.9. So far we have not defined *numerical* projective differential invariants (except the dimension of osculating spaces). Such invariants are well known in Euclidean differential geometry: There is the curvature and torsion of a curve. In Euclidean geometry there is also an invariant differential form of first order, which is usually called arc length differential.

A systematic treatment of these topics involves the Lie theory of local transformation groups and their actions on differential forms, and is beyond the scope of this book. As the group of projective automorphisms is rather large, the projectively invariant differential form of lowest order has order 3. There are numerical invariants, one of which is called projective curvature. For these and other topics see [17]. \diamond

1.2.2 Surfaces

An m -surface in P^n is a mapping $s : D \rightarrow P^n$, where D is an open domain in \mathbb{R}^m , and

$$s(u) = s(u)\mathbb{R} = (s_0(u), \dots, s_n(u))\mathbb{R}, \quad u = (u^1, \dots, u^m) \in D. \quad (1.84)$$

Projective differential geometry deals with properties of surfaces, which are not affected by re-parametrizations, and when the surface undergoes a projective transformation, this property has to transform accordingly (e.g., if it is a subspace, it has to undergo the same transformation, if it is a numerical invariant, it must not change). A property assigned to a surface means that it must of course also be independent of re-normalizations.

The Tangent Space

If $s : D \rightarrow P^n$ is a surface with $D \subset \mathbb{R}^m$, $s(u^1, \dots, u^m) = s(u^1, \dots, u^m)\mathbb{R}$, we consider the partial derivatives $\partial s / \partial u^i$. These vectors are the homogeneous coordinate vectors of points, and both the vector and the corresponding point will be denoted by subscripts which indicate the differentiation:

$$s_{,i} = s_{,u^i} = \frac{\partial s}{\partial u^i}, \quad s_{,i} = s_{,u^i} = s_{,i}\mathbb{R}. \quad (1.85)$$

If the surface point depends on two real parameters u and v , we write $s_{,1} = s_{,u}$ and $s_{,2} = s_{,v}$. Iterated derivatives are denoted by $s_{,uu}$, $s_{,uvv}$, etc. (but note that we sometimes write simply u for (u^1, \dots, u^m)).

Definition. Assume a surface $s : D \rightarrow P^n$ with $D \subset \mathbb{R}^m$ and $s(u) = s(u)\mathbb{R} = s(u^1, \dots, u^m)\mathbb{R}$. The projective subspace

$$s(u) \vee \frac{\partial s}{\partial u^1}\mathbb{R} \vee \dots \vee \frac{\partial s}{\partial u^m}\mathbb{R} = s(u) \vee s_{,1}(u) \vee \dots \vee s_{,m}(u)$$

spanned by $s(u)$ and its partial derivative points is called tangent space of s at u . The surface is called regular at u , if the dimension of the tangent space equals m .

Proposition 1.2.4. The tangent space and regularity of a surface point are properties of projective differential geometry.

Proof. We have to study the behaviour of the linear span of $s(u), s_{,1}(u), \dots, s_{,m}(u)$ under re-normalizations, re-parametrizations, and projective transformations. We leave the details to the reader. \square

A curve obviously is a 1-surface, and there are no regular k -surfaces if $k > n$. A curve $c : I \rightarrow D$ gives rise to a curve $s(c(t))$ which is contained in the surface $s(D)$. Its tangent is contained in the surface's tangent space. Conversely, it is easy to see that the tangent space is the union of all possible curve tangents which can arise in this way.

Contact Order of Surfaces

Definition. Two m -surfaces $s_1(u), s_2(v)$ have contact of order k in a common point, if there exist homogeneous parametrizations s_1, s_2 and a locally regular parameter transform $v = v(u)$ such that the derivatives of $s_1(u)$ and $s_2(v(u))$ in this point are equal up to order k .

An m -surface $s_1 : D_1 \rightarrow P^n$ and an l -surface $s_2 : D_2 \rightarrow P^n$ with $m < l$ have contact of order k , if there exists a smooth mapping $s_3 : D_1 \rightarrow D_2$ such that $s_1(u)$ and $s_2 \circ s_3(u)$ have contact of order k .

Example 1.2.4. A curve $c \subset E^n$ and its osculating plane at some point $c(t)$ have contact of order two, since the osculating circle lies in the osculating plane and has second order contact with the curve. Of course, this holds in P^n as well, because for all curve points there is an appropriate affine normalization where the curve becomes a Euclidean curve. \diamond

Example 1.2.5. Contact of order k is preserved if both surfaces undergo the same differentiable mapping. An example is a projection:

Project a curve in P^3 from a point in its osculating plane. Choose a conic in this plane which osculates the curve and does not contain the projection center. The conic projects to the tangent of the projection of the curve. Thus this projection has second order contact with its tangent, and therefore has an inflection point.

This argument can be completed to a proof of Prop. 1.2.3. \diamond

Osculating Quadrics of 2-Surfaces in P^3

Consider a 2-surface Φ and a regular surface point P in it. We may use this point as origin and its tangent plane as plane $z = 0$ in a locally adapted coordinate system (x, y, z) . Then it is known that Φ can be parametrized locally as graph of the function $z = f(x, y)$: $s(x, y) = (x, y, f(x, y))$. It has second order contact with the surface Ψ obtained as graph $(x, y, \bar{s}(x, y))$ of its second order Taylor approximant

$$\bar{s} : (x, y) \mapsto (x, y, f_{xx}x^2 + 2f_{xy}xy + f_{yy}y^2). \quad (1.86)$$

Here subscripts after a comma indicate partial differentiation. If all second order partial derivatives of f are zero, the tangent plane $z = 0$ has second order contact with Φ at P . Then, P is called *flat point*. Otherwise, Ψ is either an elliptic paraboloid, hyperbolic paraboloid or a parabolic cylinder; accordingly, P is called an *elliptic, hyperbolic or parabolic surface point*. Ψ is one example of an osculating quadric, i.e. a quadric having second order contact with Φ at P ; there are many other osculating quadratic varieties, which are singular in the case of a parabolic surface point. However, in general there is no quadric which has third order contact with s .

Proposition 1.2.5. *The property of being an elliptic, hyperbolic, parabolic, or flat point is a property of projective differential geometry.*

Proof. We only sketch the proof: The second order Taylor approximant used in (1.86) is a quadratic function, whose zeros can be: (i) only the point $(0, 0)$, (ii) two lines, (iii) one line, (iv) the entire plane. It is easily seen that these cases correspond to elliptic, hyperbolic, parabolic and flat points. Assume a line $t \cdot (a, b)$ in the xy -plane and consider the curves $s(ta, tb)$ and $(ta, tb, 0)$. It is not difficult to show that this line has second order contact if and only if (a, b) is a zero of the Taylor approximant, and conversely, no other line of the form $(ta, tb, 0)$ has second order contact with s .

This shows that the four classes of points are characterized by the number of lines in the tangent plane which have second order contact with the surface: none, 2, 1, or all lines, respectively. The proposition follows because this characterization has already been established as projectively invariant. \square

The Involution of Conjugate Tangents for 2-Surfaces in P^3

When we discuss properties which involve only the first and second derivative, we can replace Φ by any osculating quadric Ψ . We give some examples which belong to projective differential geometry.

Consider a quadric $\Psi \subset P^3$ and a point $P \in \Psi$. The quadric defines a duality denoted by δ . In the tangent plane $P\delta$ of P we pick a point $C \neq P$ and consider the contour outline c of Ψ with respect to a central projection with center C (see Fig. 1.39). The contour c is the intersection of $\Psi \cap C\delta$. Its tangent at P equals $T_C = P\delta \cap C\delta = (P \vee C)\delta$, so $P \vee C$ and the contour's tangent are *conjugate*.

On the other hand we can calculate the contour outline of a surface and see that its tangent can be computed using only first and second derivatives.

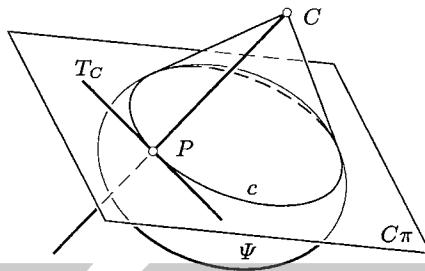


Fig. 1.39. Conjugate surface tangents.

This shows that the involution of conjugate surface tangents is the same for all quadrics which have contact of order two, and that the following definition assigns a projective differential invariant to an elliptic or hyperbolic surface point:

Definition. *In an elliptic or hyperbolic surface point we define the involution of conjugate surface tangents as the involution of conjugate surface tangents of any osculating quadric.*

It is clear that the correspondence between projection center and contour outline, which has been established for quadrics, holds for all surfaces: If the projection center C is contained in the tangent plane of P , then P is part of the contour outline, whose tangent in P is conjugate to the line $P \vee C$.

The involution of conjugate surface tangents is most easily computed if the surface has the form of a graph of a function, because the second order Taylor approximant immediately gives the polarity of the osculating quadric. This shows that the lines mentioned in the proof of Prop. 1.2.5, which are found as zero set of the Taylor approximant, and are exactly the lines which have second order contact with the surface, are also the lines fixed by the involution of conjugate tangents. They are called *asymptotic tangents*. Surface curves all of whose tangents are asymptotic tangents, are called *asymptotic curves* in the surface.

Remark 1.2.10. If a surface point is parabolic or flat, then the Taylor approximant defines no osculating quadric, but a singular osculating quadratic variety. There is no involution of conjugate tangents. In the parabolic case we define the unique asymptotic tangent conjugate to all other tangents, and in a flat point all tangents are called conjugate to each other. ◇

It can be shown that a non-tangential planar intersection of a surface has an inflection point if the intersecting plane contains an asymptotic tangent. This is a consequence of the fact that these tangents have contact of order 2 with the surface.

Remark 1.2.11. Asymptotic tangents are also invariants of Euclidean differential geometry, because the Euclidean transformation group is a subgroup of the projective one. So there must be a Euclidean characterization of them as well: It turns out that they are precisely the tangents of vanishing normal curvature.

The involution of conjugate tangents is easily described in terms of the second fundamental form of Euclidean surface theory, which is calculated for a Taylor approximant surface: It turns out that the second fundamental form has coefficients $h_{11} = f_{,xx}$, $h_{12} = f_{,xy}$, $h_{22} = f_{,yy}$. Because ‘everything depends only on the first and second derivatives’ and the invariance of the properties in question, we have shown that conjugate tangents defined by direction vectors (s_1, s_2) and (t_1, t_2) in the parameter domain are characterized by

$$h_{11}s_1t_1 + h_{12}(s_1t_2 + s_2t_1) + h_{22}s_2t_2 = 0. \quad (1.87)$$

◇

Hypersurfaces

A *regular hypersurface* is a regular $(n - 1)$ -surface in P^n . Without going into details, we mention that the second order Taylor approximant defines an osculating quadratic variety, which is possibly singular. If it is regular, we call it an osculating quadric and use it to define a conjugacy relation between surface tangents. The self-conjugate tangents are precisely the asymptotic tangents which have second order contact with the surface. We can classify the regular surface points according to the index of an osculating quadric, which is a projective differential invariant. If the osculating quadrics are oval, the surface point is called elliptic again.

1.2.3 Duality

For a hypersurface $s : D \rightarrow P^n$ all of whose points are regular we can define a *dual surface* $s^* : D \rightarrow P^{n*}$ which assigns to a parameter value $u \in D$ the tangent hyperplane at u , which is attached to the point $s(u)$. The most simple case is that of $n = 2$, where a curve defines its dual curve, consisting of all tangents.

Remark 1.2.12. The hyperplane coordinates $\mathbb{R}\mathbf{u}$ of the projective span of n independent points $\mathbf{a}_1\mathbb{R}, \dots, \mathbf{a}_n\mathbb{R}$ are found as coefficients of the linear equation

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{x}) = u_0x_0 + \dots + u_nx_n = 0.$$

Obviously then $\mathbb{R}\mathbf{u} = \mathbb{R}(u_0, \dots, u_n)$. The values u_i can be computed by applying Laplace’s theorem to the last column of the determinant expression: u_i equals $(-1)^{n+1-i}$ times the determinant of the matrix $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ minus the i -th row. The vector \mathbf{u} is called the *vector product* of $\mathbf{a}_1, \dots, \mathbf{a}_n$ and is denoted by

$$\mathbf{u} = \mathbf{a}_1 \times \dots \times \mathbf{a}_n.$$

If we apply this procedure to a surface point $s(u) = \mathbf{s}(u)\mathbb{R}$ and the partial derivative points $(\partial\mathbf{s}/\partial u^i)\mathbb{R}$ we see that the dual surface is indeed an $(n - 1)$ -surface whose differentiability class is C^{r-1} , if s was C^r . ◇

Planar Curves

If the surface is C^2 , we may construct the dual of the dual. It turns out that we get the original surface again. For planar curves, this is easily shown: Assume that $c(t) = (c_0(t), c_1(t), c_2(t))\mathbb{R} = c(t)\mathbb{R}$. The dual curve is given by $c^*(t) = \mathbb{R}c^*(t) = \mathbb{R}(c(t) \times \dot{c}(t))$. The iterated dual c^{**} is calculated as $c^{**} = (c^* \times \dot{c}^*)\mathbb{R} = (c \times \dot{c}) \times (\dot{c} \times \ddot{c} + c \times \ddot{c})\mathbb{R} = \det(c, \dot{c}, \ddot{c})c\mathbb{R} = c\mathbb{R} = c$.

The fact that iterated duality gives the curve again (if the curve is well behaved) leads to the notion of *curve* as a self-dual object, which can be described either as its family of points or the family of lines (i.e., tangents).

Example 1.2.6. The curve

$$c(t) = c(t)\mathbb{R} = (1, t^m, t^n)\mathbb{R}, \quad (m < n)$$

has the line representation (which means its dual curve)

$$c^*(t) = \mathbb{R}u(t) = \mathbb{R}(c(t) \times \dot{c}(t)) = \mathbb{R}((n-m)t^{m+n-1}, -nt^{n-1}, mt^{m-1}).$$

Renormalization yields the equivalent form

$$c^*(t) = \mathbb{R}((n-m)t^n, -nt^{n-m}, m).$$

In particular, $(m, n) = (1, 3)$ gives a curve c with an inflection at $t = 0$. The dual representation is given by $\mathbb{R}(2t^3, -3t^2, 1)$ and has a cusp at $t = 0$.

It is not difficult to see that the duals of curves with ordinary inflection points have ordinary cusps at the appropriate parameter values, and vice versa. This shows that an inflection point gives a singularity of the dual curve, whereas an ordinary cusp, which is not a regular point, yields a dual curve which is regular there.

This duality between inflections and cusps has been used in algorithms for detecting inflections of curves by means of cusps of their duals (see [77, 80]).

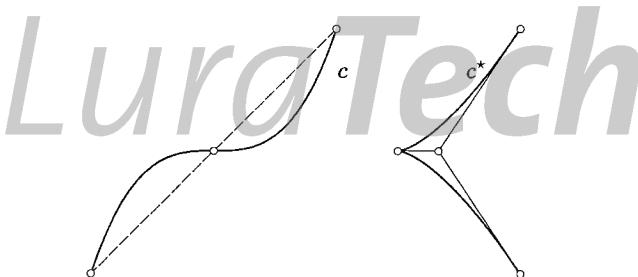


Fig. 1.40. Inflection and cusp are dual to each other. Left: a line intersects in three points. Right: A point is incident with three tangents.



Example 1.2.7. Another interesting special case is $(m, n) = (1, 4)$, which means that $c(0)$ is a flat point, and c has third order contact with its tangent there (see Fig. 1.40, left). The dual curve is given by $\mathbb{R}(3t^4, -4t^3, 1)$, which is a singularity whose osculating spaces have dimensions $0, 0, 0, 1, 2, \dots$ (cf. the definition of a singularity of a curve). Such a singularity is sometimes called *extraordinary point*. The next example shows a situation where such points occur naturally. \diamond

Example 1.2.8. *Offsets* of planar curves are constructed as follows: We start with a curve c and compute its dual c^* . There are two dual curves $c_{r,+}^*, c_{r,-}^*$ whose lines are parallel to those of c^* and have constant distance r to them. They correspond to curves $c_{r,+}$ and $c_{r,-}$ which are called the *offsets* of c at distance r .

Fig. 1.41 shows an example where the offset curve has a singularity of the type described in the previous example, which occurs as limit of two cusps and a double point. \diamond

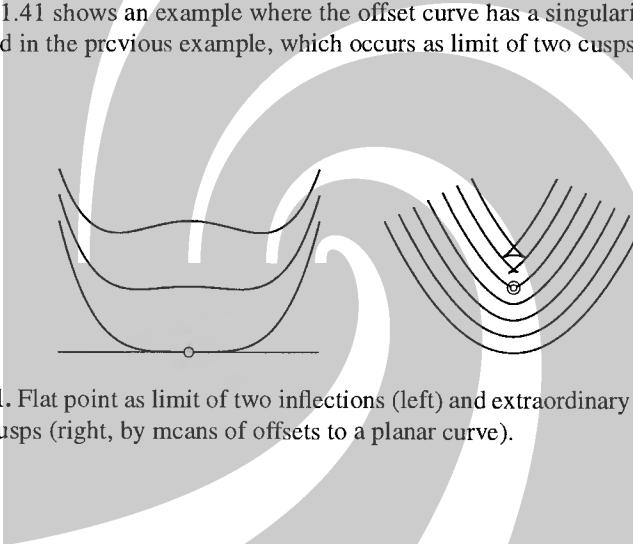


Fig. 1.41. Flat point as limit of two inflections (left) and extraordinary point as limit of two cusps (right, by means of offsets to a planar curve).

Contact Order and Duality

If two curves have contact of order k in a point, it is clear that their duals have at least contact of order $k - 1$ (the loss of one order comes from the differentiation when computing the dual). We are going to show that in the case of *regular* curve and dual curve the duals actually have contact of order k .

Lemma 1.2.6. *Suppose that two planar curves $c(t), \bar{c}(t)$ have contact of order k at $t = t_0$ ($k \geq 1$), and that both c and its dual c^* are regular at $t = t_0$. Then also the duals $c^*(t_0)$ and $\bar{c}^*(t)$ have contact of order k at $t = t_0$.*

Proof. Without loss of generality we assume $t_0 = 0$. We choose a coordinate system such that $c(0)$ is the origin, and the x_1 -axis is the curve tangent. We re-parametrize the curves c and \bar{c} such that $c(t) = c(t)\mathbb{R} = (1, t, y(t))\mathbb{R}$ and $\bar{c}(t) = \bar{c}(t)\mathbb{R} = (1, t, y(t) + t^{k+1}d(t))\mathbb{R}$. Because the x_1 -axis is c 's tangent, and c^* is supposed to be regular, $\dot{y}(0) = 0$ and $\ddot{y}(0) \neq 0$.

We compute $\mathbf{c}^* = \mathbf{c} \times \dot{\mathbf{c}} = (1, t, y) \times (0, 1, \dot{y}) = (t\dot{y} - y, -\dot{y}, 1)$. Then $c^*(t) = \mathbf{c}^*(t)\mathbb{R}$. Analogously, we compute $\bar{\mathbf{c}}^*(t)$: We abbreviate the first derivative of $y(t) + t^{k+1}d(t)$ with $\lambda(t)$ and get the result $\bar{\mathbf{c}}^*(t) = \bar{\mathbf{c}}^*(t)\mathbb{R} = (t\lambda(t) - (y + t^{k+1}d), -\lambda(t), 1)\mathbb{R}$.

We construct a local parameter transform $t = t(u)$ such that $\dot{y}(u) = \lambda(t(u))$. As $\dot{\lambda}(0) = \ddot{y}(0) \neq 0$, the inverse function theorem shows that $t = t(u)$ exists and is regular. Further, the chain rule shows that $t(u) = u + u^k r(u)$, i.e., $t(u)$ is the identity mapping up to derivatives of order $k - 1$.

Now $\bar{\mathbf{c}}^*(t(u)) = (t(u)\dot{y}(u) - y(t(u)) - t(u)^{k+1}d(t(u)), -\dot{y}(u), 1)$. It remains to show that the first coordinate function has the same derivatives up to order k with the first coordinate function of \mathbf{c}^* , i.e., with $u\dot{y}(u) - y(u)$. This is easily done by comparing Taylor polynomials up to order k . \square

Remark 1.2.13. Lemma 1.2.6 is not true for singular curves: Consider the curves $c_1(t) = (1, t^3, t^4)\mathbb{R}$ and $c_2(t) = (1, t^3, t^5)\mathbb{R}$ which have contact order three at $t = 0$. Their duals $c_1^*(t) = \mathbb{R}(t^4, -4t, 3)$ and $c_2^*(t) = \mathbb{R}(2t^5, -5t^2, 3)$ don't have even contact of order one, because the former is regular and the latter is not. Furthermore, Lemma 1.2.6 is not true for contact order zero. \diamond

Space Curves and Tangent Developables

The tangent space of a C^2 curve $c : I \rightarrow P^3$ has dimension at most 1. If we want to assign a plane to each curve point, we have to use the osculating subspace of dimension two. If c has no inflection points, we define the dual curve

$$c^* : t \mapsto c(t) \vee c^1(t) \vee c^2(t),$$

which is a one-parameter family of planes. We will later see that the envelope of these planes is the surface of tangents of c . $c^*(t)$ is the tangent plane in all points of the tangent at $c(t)$. We will study these surfaces in detail in Chap. 6.

Surfaces in P^3

We discuss surfaces in three-dimensional space in more detail: We have already seen that the family of tangent planes of a surface may be just a curve in dual projective space. On the other hand, in the neighbourhood of an elliptic or hyperbolic surface point, the correspondence between surface points and their tangent planes is one-to-one. If $s : D \rightarrow P^3$, $(u, v) \mapsto s(u, v)\mathbb{R}$ is a surface, its dual is given by

$$s^*(u, v) = \mathbb{R}s^*(u, v) \quad \text{with} \quad s^*(u, v) = s(u, v) \times \frac{\partial}{\partial u} s(u, v) \times \frac{\partial}{\partial v} s(u, v).$$

If s^* is a surface $D \rightarrow P^{3*}$, then $s = s^{**}$ is similarly found with s^* instead of s .

If (u, v) is a parabolic or flat point of s , it is no regular point of s^* . If it is hyperbolic or elliptic, it is a regular point of s^* .

1.3 Elementary Concepts of Algebraic Geometry

We will briefly describe some basic concepts of algebraic geometry, including Gröbner bases, which are a main algorithmic tool. We give proofs only if they are simple. This section is not an introduction into algebraic geometry, i.e., the geometry of algebraic varieties, the regular functions defined on them, and their morphisms, but into the projective geometry of algebraic curves and surfaces.

1.3.1 Definitions and Algorithms

We consider the real and complex affine spaces \mathbb{R}^n and \mathbb{C}^n , embedded in their respective projective extensions P^n and $\mathbb{C}P^n$. The embedding is given by $(x_1, \dots, x_n) \mapsto (1 : x_1 : \dots : x_n)$. As in the previous chapters we will write $x = (x_1, \dots, x_n)$ for an affine point, and $x = (x_0, \dots, x_n)$ for the homogeneous coordinate vector of a projective point. We will also think of \mathbb{R}^n and P^n as embedded in \mathbb{C}^n and $\mathbb{C}P^n$.

We write k for any of the two fields \mathbb{R}, \mathbb{C} . The reader should be aware of the fact that most of the things which are true for $k = \mathbb{R}$ are true for all fields, and most of the things true for $k = \mathbb{C}$ are true for all algebraically closed fields. In order to keep the presentation simple, we restrict ourselves to $k = \mathbb{R}$ and $k = \mathbb{C}$.

The ring of polynomials in n indeterminates x_1, \dots, x_n with coefficients from k will be denoted by $k[x_1, \dots, x_n]$ or $k[\mathbf{x}]$. The ring of polynomials in $n + 1$ indeterminates x_0, \dots, x_n is denoted by $k[x_0, \dots, x_n]$ or $k[\mathbf{x}]$. (In the future we will not repeat definitions employing \mathbf{x} for x).

Monomial Orderings

A *monomial* is a polynomial of the form $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and we abbreviate this by \mathbf{x}^α , where $\alpha = (\alpha_1, \dots, \alpha_n)$. Its *total degree* is the number $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

We define an ordering in the set of monomials by $\mathbf{x}^\alpha > \mathbf{x}^\beta$ if and only if \mathbf{x}^α precedes \mathbf{x}^β in the alphabetical order when writing both in their ‘long’ form without exponents, and the blank is the last letter in the alphabet. This will be made clear in the examples. We should also mention that different monomial orderings are used for different purposes in computational algebraic geometry. In order to keep the presentation simple, we restrict ourselves to the pure lexicographic ordering here.

Definition. Let $g \in k[\mathbf{x}]$. The leading monomial $\text{LM}(g)$ is the greatest monomial of g . The leading term $\text{LT}(g)$ is the term belonging to this monomial, i.e., this monomial together with its coefficient in g .

Example 1.3.1. $x_1 > x_2$ because ‘ x_1 ’ precedes ‘ x_2 ’ in the alphabet. For the same reason, $x > y$, if our variables bear the names x and y . Further examples are $x_1 x_2^2 > x_1$ and $x_1^2 > x_1 x_2^3$, because ‘ $x_1 x_2 x_2$ ’ precedes ‘ $x_1 \square \square$ ’, and ‘ $x_1 x_1 \square \square$ ’ precedes ‘ $x_1 x_2 x_2 x_2$ ’. If $g = 2x_1 x_2^2 + 2x_2 - 3x_1^2$, then $\text{LM}(g) = x_2^2$ and $\text{LT}(g) = -3x_1^2$. ◇

Ideals

Definition. An ideal I of $k[\mathbf{x}]$ is a set of polynomials with the following property: If $g, h \in I$, then $g + h \in I$, and if $g \in I$, then $g \cdot h \in I$ for all $h \in k[\mathbf{x}]$. We write $I \trianglelefteq k[\mathbf{x}]$.

An ideal I is said to be generated by polynomials g_i , if it is the smallest ideal which contains all g_i . We write $I = \langle g_1, \dots, g_m \rangle$ if the index i ranges from 1 to m . I coincides with the set of all polynomial combinations $g_{i_1} \cdot h_{i_1} + \dots + g_{i_r} \cdot h_{i_r}$, with $h_{i_k} \in k[\mathbf{x}]$. The ideal generated by a single polynomial g , which is just the set of its multiples $g \cdot h$, $h \in k[\mathbf{x}]$, is called the *principal ideal* generated by g .

Monomial Ideals

The notion of a monomial ideal is important for understanding the concept of Gröbner bases. The reader not interested in the details may skip this paragraph, he or she will be able to apply Gröbner bases for elimination without knowing what a monomial ideal is.

Definition. An ideal generated by monomials is called a *monomial ideal*.

The ideal membership problem is easily solved for monomials and monomial ideals:

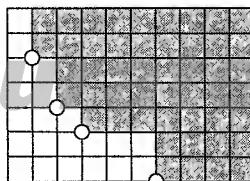
Proposition 1.3.1. Suppose I is a monomial ideal generated by monomials \mathbf{x}^α , where α ranges in some set A . Then \mathbf{x}^β is in I if and only if it is a multiple of one of the \mathbf{x}^α .

Proof. Assume that $\mathbf{x}^\beta \in I$, which means that $\mathbf{x}^\beta = \sum h_i \mathbf{x}^{\alpha_i}$. Every term of the right hand side of this equation is a multiple of some α_i . All but one of them cancel, and this one is precisely \mathbf{x}^β . \square

We will eventually show that all ideals are generated by a *finite* number of polynomials. A first step is to show this for monomial ideals:

Proposition 1.3.2. Every monomial ideal I is generated by a finite number of monomials.

Proof. We consider the case of two indeterminates first and visualize a possible distribution of monomials belonging to I . If a grid point (i, j) , symbolizing the monomial $x_1^i x_2^j$ in the following diagram belongs to I , the previous proposition shows that all grid points (k, l) with $k \geq i$, $l \geq j$ also belong to I . Clearly there is a finite number of grid points which generate I in this way.



The case of more than two indeterminates is completely analogous, but not so easy to visualize. A ‘rigorous’ proof is easily derived from this visualization. \square

Gröbner Bases

Definition. If I is an ideal, then the leading monomials $\text{LM}(g)$ of all $g \in I$ generate an ideal, called $\text{LM}(I)$. A finite collection of polynomials $g_1, \dots, g_r \in I$ is called Gröbner basis of I , if the leading monomials $\text{LM}(g_1), \dots, \text{LM}(g_r)$ generate $\text{LM}(I)$.

The name ‘basis’ is justified by the following

Proposition 1.3.3. All ideals I have Gröbner bases, and a Gröbner basis is a finite set of polynomials which generates I .

Proof. The ideal $\text{LM}(I)$ is generated by the monomials $\text{LM}(g)$ with $g \in I$. Prop. 1.3.2 shows that finitely many of them already generate $\text{LM}(I)$. This shows the existence of a Gröbner basis g_1, \dots, g_r .

The ideal I' generated by g_1, \dots, g_r is contained in I . Assume now that there are polynomials $g \in I$, but not in I' . Choose one of them with minimal possible leading monomial, call it g . Because $\text{LM}(g_1), \dots, \text{LM}(g_r)$ generate $\text{LM}(I)$, there are polynomials h_1, \dots, h_r with $\text{LM}(g) = \text{LM}(g_1) \cdot h_1 + \dots + \text{LM}(g_r) \cdot h_r$. Now $g' = g - \sum g_i h_i$ is in I , but not in I' (because otherwise g would be in I'). Its leading monomial is less than $\text{LM}(g)$ by construction. This contradicts the minimal choice of g . \square

Example 1.3.2. We have not yet an algorithm for computation of Gröbner bases, so we can only show a very simple example: Assume that I is generated by the polynomials $x^2 - y$ and y . Clearly it is also generated by x^2 and y , so it is a monomial ideal and both x^2, y and $x^2 - y, y$ are its Gröbner bases. \diamond

The existence of Gröbner bases immediately implies the following theorem of David Hilbert:

Theorem 1.3.4. (Hilbert basis theorem) For all ideals $I \trianglelefteq k[\mathbf{x}]$ there is a finite number of polynomials $g_1, \dots, g_r \in I$ which generate this ideal.

Ideal Chains

When considering algebraic varieties contained in each other, we will need the following theorem:

Theorem 1.3.5. There is no infinite chain of ideals $I_1 \subset I_2 \subset I_3 \subset \dots$ where all inclusions are proper, i.e., $I_i \neq I_{i+1}$.

Proof. Consider the union of all I_i , which is an ideal. It has a finite generator set g_1, \dots, g_r . Every g_i is an element of some $I_{k(i)}$. If $k = \max_i(k(i))$, then all g_i are contained in I_k , which implies $I_k = I_{k+1} = I_{k+2} = \dots = I$. \square

Affine and Projective Algebraic Varieties

We consider subsets of an affine space (\mathbb{R}^n or \mathbb{C}^n) which are defined by polynomial equations. It is no exaggeration to say that they occur in every branch of mathematics, and their various properties are exploited in numerous ways. Not all of these however are easy to show and ‘algebraic geometry’ is a very large field. In the last decades computational methods have become increasingly important for applications.

Definition. *An affine algebraic variety M in k^n ($k = \mathbb{R}, \mathbb{C}$) is the set of common zeros of a set F of polynomials $g_i \in k[\mathbf{x}]$.*

The definition of a *projective* algebraic variety is similar, but we have to take into account that a point of P^n is described by the set of multiples $\lambda \cdot \mathbf{x}$ of a nonzero coordinate vector $\mathbf{x} \in k^{n+1}$. So it is necessary to restrict ourselves to *homogeneous polynomials* which consist of monomials of the same total degree r , which is then called the degree of the homogeneous polynomial. If p is a homogeneous polynomial of degree r , then $p(\lambda\mathbf{x}) = \lambda^r p(\mathbf{x})$, and the set of zeros of p consists of a union of one-dimensional subspaces. This enables the following

Definition. *A set F of homogeneous polynomials of $k[\mathbf{x}]$ defines an affine algebraic variety $V(F)$ in k^{n+1} . A point \mathbf{x} of n -dimensional projective space is contained in the projective algebraic variety defined by F if and only if its homogeneous coordinate vector \mathbf{x} is in $V(F)$.*

Because of this correspondence between affine algebraic varieties in k^{n+1} defined by *homogeneous* polynomials on the one hand, and projective algebraic varieties on the other hand, we formulate some theorems only for affine algebraic varieties.

Example 1.3.3. Consider \mathbb{R}^2 and the polynomial $g(x_1, x_2) = x_1^2 + x_2^2 - 1$. The one-element set $G = \{g\}$ defines the affine algebraic variety $V(G)$, which is nothing but the unit circle.

After embedding affine \mathbb{R}^2 in P^2 by letting $(x_1, x_2) \mapsto (1, x_1, x_2)\mathbb{R}$, the same set is the projective algebraic variety defined by the homogeneous polynomial $x_1^2 + x_2^2 - x_0^2$. None of the ideal points not in affine \mathbb{R}^2 satisfies this equation.

The line $x_1 = 0$ of affine \mathbb{R}^2 is an affine algebraic variety, likewise defined by only one polynomial equation. The same equation in P^2 defines this line plus its ideal point. It is easy to show that the *affine* line is no projective variety, and that the smallest projective variety containing the affine line is just the projectively extended line. ◇

Assume affine space k^n embedded into projective space by $(x_1, \dots, x_n) \mapsto (1 : x_1 : \dots : x_n)$. The affine part of a projective algebraic variety $V(F)$ is an affine algebraic variety $V(F')$, where F' is found by substituting 1 for x_0 in all elements of F .

Conversely, any affine algebraic variety M is contained in a projective algebraic variety, which is found by making the defining equations M homogeneous by

adding appropriate powers of x_0 . This need *not* be the smallest projective variety which contains the affine one.

Example 1.3.4. Consider the variety $x_1^2 - x_2 = 0$ in \mathbb{R}^2 , which is a parabola. After embedding affine \mathbb{R}^2 in the projective plane, this set is defined by $x_1^2 - x_2x_0 = 0$, $x_0 \neq 0$. Its unique projective extension has the equation $x_1^2 - x_2x_0 = 0$. The difference between the affine and the projective variety in this case is just the point $(0, 0, 1)\mathbb{R}$.

The polynomial cubic $(x, y, z) = (t, t^2, t^3)$ in \mathbb{R}^3 is an affine variety, because it is the intersection of the surfaces $x^2 = y$, $x^3 = z$. The projective extensions of these two surfaces have, in homogeneous coordinates $x = x_1/x_0$, $y = x_2/x_0$, $z = x_3/x_0$, the equations $x_1^2 = x_2x_0$, $x_1^3 = x_3x_0^2$.

The ‘affine part’ $x_0 \neq 0$ of their intersection is the original cubic, and the ‘ideal part’ $x_0 = 0$ obviously consists of the line $x_0 = x_1 = 0$. There is, however, a smaller projective variety which contains the cubic. It is defined by $x_1^2 = x_2x_0$, $x_1^3 = x_3x_0^2$, $x_2^3 = x_3^2x_0$, and its only ideal point is $(0, 0, 0, 1)\mathbb{R}$. ◇

The notion of affine (projective) algebraic variety is an affine (projective) invariant, because affine and projective mappings transform polynomials to polynomials.

The Correspondence Ideal—Variety

Assume we have defined an algebraic variety M in k^n ($k = \mathbb{R}$ or $k = \mathbb{C}$) by a set F of polynomials $g_i \in k[\mathbf{x}]$, and now consider the set of *all* polynomials which vanish for all points of M . These obviously include the original polynomials g_i , and form an ideal, denoted by $I(M)$, because if g, h vanish at all $\mathbf{x} \in M$, then $g + h$ vanishes likewise, and if g vanishes at all $\mathbf{x} \in M$, then so does the product $g \cdot h$ for all $h \in k[\mathbf{x}]$.

More generally we may assume that M is any subset of k^n . Again the set of polynomials which vanish at all points of M is an ideal, called $I(M)$. The algebraic variety $V(I(M))$ defined by the polynomials of $I(M)$ contains the original set M .

The following properties of the operations V and I are immediately verified:

- If $M = V(F)$ where F is a set of polynomials, then $M = V(I)$ where I is the ideal generated by F . This is clear because if polynomials vanish at \mathbf{x} , then also polynomial combinations of these polynomials vanish there.
- If $M_1 \subseteq M_2$, then $I(M_1) \supseteq I(M_2)$, and if $F_1 \subseteq F_2$, then $V(F_1) \supseteq V(F_2)$. This is an immediate consequence of the definition: If more polynomial equations are to satisfy, the number of solutions becomes smaller, and vice versa.
- For all sets F of polynomials we have $F \subseteq I(V(F))$ and likewise for all sets M of points we have $M \subseteq V(I(M))$. This is also an immediate consequence of the definition.
- We have $V(I(V(F))) = V(F)$ and $I(V(I(M))) = I(M)$, which follows from the two previous inclusions.

The last of these statements shows that algebraic varieties $V(F)$ on the one hand and the ideals of the form $I(M)$ on the other hand are *closed* under the operations $V(I(\cdot))$ and $I(V(\cdot))$, respectively, and they correspond to each other under the operations V and I .

Proposition 1.3.6. *An algebraic variety M is always defined by a finite number of polynomial equations.*

Proof. Consider the ideal $I(M)$. By the Hilbert basis theorem it has a *finite* number g_1, \dots, g_r of generators. Obviously $V(g_1, \dots, g_r)$ equals $V(I)$. \square

Example 1.3.5. Consider affine \mathbb{R}^2 and the polynomials $g_1(\mathbf{x}) = x_1, g_2(\mathbf{x}) = x_2$. Then $V(\{g_1\})$ is the x_2 -axis, and $V(\{g_1, g_2\})$ is only the point $(0, 0)$.

Consider the x_1 -axis without the point $(0, 0)$ and denote this set by M . All polynomials vanishing along the x_1 -axis must also vanish in $(0, 0)$, because the zero set of a continuous function is closed. On the other hand the x_1 -axis, denoted by M' , is an algebraic variety, which implies $V(I(M)) = M'$. \diamond

Union and Intersection of Algebraic Varieties

If $M_1 = V(F_1), M_2 = V(F_2)$ are algebraic varieties, their intersection is the algebraic variety $V(F_1 \cup F_2)$. More generally, if $M_i = V(F_i)$ is an arbitrary collection of algebraic varieties, their intersection is the algebraic variety defined by the union of all F_i .

The union of M_1, M_2 is given by $V(S)$ where S is the set of products $g_1 \cdot g_2$, with $g_1 \in F_1, g_2 \in F_2$. So also the union of two algebraic varieties is an algebraic variety.

The proofs of these statements are easy and left as an exercise for the reader.

Irreducibility of Algebraic Varieties

In contrast to the family of *subsets* of an affine space, the family of *algebraic* varieties is very inflexible. This is expressed in the following definition and theorems following it:

Definition. *An algebraic variety M is called irreducible if it is not the union of algebraic varieties M_1, M_2 unless one of M_1, M_2 is contained in the other.*

Proposition 1.3.7. *There is no infinite chain $M_1 \supset M_2 \supset \dots$ of algebraic varieties where all inclusions are proper.*

Proof. Consider the ideals $I(M_1) \subset I(M_2) \subset \dots$ If all inclusions $M_i \supset M_{i+1}$ are proper, then also all inclusions $I_i \subset I_{i+1}$ are proper, which contradicts Th. 1.3.5. \square

Theorem 1.3.8. *An algebraic variety can be uniquely decomposed into a finite number of irreducible varieties, called its irreducible components: $M = M_1 \cup M_2 \cup \dots \cup M_r$.*

Proof. Call all varieties which have such a decomposition *good*, and the others *bad*. We have to show that there are no bad varieties. Assume that M is bad. Then M is not irreducible. So $M = M_1 \cup M_2$, where M_i are algebraic varieties, M_1 is not contained in M_2 and vice versa, and at least one of them, call it M' , must be bad. If we repeat this argument we get an infinite chain $M \supset M' \supset M'' \supset \dots$ of algebraic varieties where all inclusions are proper. This contradicts Prop. 1.3.7.

The uniqueness is also shown easily: If there are two decompositions $M = M_1 \cup \dots \cup M_r = M'_1 \cup \dots \cup M'_{r'}$, such that one of the M'_j does not occur among the M_i , then we can write $M'_j = (M_1 \cap M'_j) \cup \dots \cup (M_r \cap M'_j)$, which shows that M'_j is not irreducible. \square

Example 1.3.6. Consider the variety V defined by the equation $f(x_0, x_1, x_2) = -x_0^2 + x_1^2 + x_2^2$ in the complex projective plane. It is the complex and projective extension of the unit circle.

Suppose that $V = V_1 \cup V_2$ with neither $V_1 \subseteq V_2$ nor $V_2 \subseteq V_1$. Choose a polynomial $g_1 \in I(V_1) \setminus I(V_2)$ and a polynomial $g_2 \in I(V_2) \setminus I(V_1)$. Obviously the product $g_1 g_2$ is contained in $I(V)$. Then Prop. 1.3.10 (below) implies that some power $(g_1 g_2)^r$ is a multiple of f . As f has no nontrivial factors, this is not possible, and V is indeed irreducible. \diamond

The Hilbert Nullstellensatz

We define the *radical* of an ideal I as the set of the polynomials a power of which is contained in I . The radical of I is denoted by \sqrt{I} :

$$\sqrt{I} = \{g \in k[\mathbf{x}] \mid \exists r : g^r \in I\}$$

and is an ideal itself.

The following theorem is called the Hilbert Nullstellensatz (German for ‘theorem of zeros’, see e.g. [182]):

Theorem 1.3.9. *If k is algebraically closed (e.g., $k = \mathbb{C}$), and $I \trianglelefteq k[\mathbf{x}]$, then $I(V(I)) = \sqrt{I}$.*

The Hilbert Nullstellensatz gives precise information on the solvability of polynomial equations: If the equations $g_1(\mathbf{x}) = \dots = g_r(\mathbf{x}) = 0$ have no solution, then $V(I)$ is void where $I = \langle g_1, \dots, g_r \rangle$, and the constant polynomial 1 is in $I(V(I))$. The Hilbert Nullstellensatz then ensures that there is polynomial combination of the g_i which gives the constant polynomial 1. This means that by appropriately manipulating the equations $g_1(\mathbf{x}) = 0, \dots, g_r(\mathbf{x}) = 0$ we can reach the contradictory equation $1 = 0$.

If the underlying number field is \mathbb{R} , then there are other reasons why an equation has no solution: An example is the equation $x^2 + y^2 + 1 = 0$, which does not lead to an algebraic contradiction. A contradiction can be reached only if we use *inequalities* such as $x^2 \geq 0$.

Another proposition is an immediate consequence of the Hilbert Nullstellensatz:

Proposition 1.3.10. Assume that M is an algebraic variety defined by one polynomial $g(\mathbf{x})$. If a polynomial h vanishes at all $\mathbf{x} \in M$, an appropriate power of h is a multiple of g . If g is irreducible, then h is a multiple of g .

Proof. Consider the ideal I of all multiples of g . Clearly h belongs to $I(V(I))$ and the Hilbert Nullstellensatz shows the result. The statement about an irreducible g follows from the unique prime factor decomposition in the polynomial ring. \square

Hypersurfaces

Algebraic varieties defined by exactly *one* polynomial equation are called *hypersurfaces*. If the dimension of entire space is two, then they are called *curves*. This definition is valid for both affine and projective algebraic varieties. The total degree of this polynomial is called the *degree* of the algebraic hypersurface. We will sometimes use *order* as a synonym for degree.

Proposition 1.3.11. Assume that the algebraic hypersurface M is defined by the equation $g(\mathbf{x}) = 0$. If $g = g_1^{\alpha_1} \cdots g_r^{\alpha_r}$ is the factorization of g , then M is likewise defined by the polynomial $g_1 \cdots g_r = 0$. If $k = \mathbb{C}$, the irreducible components of M are the algebraic varieties $M_i : g_i(\mathbf{x}) = 0$.

Proof. Clearly M is the union of the M_i . If a polynomial h vanishes for all $\mathbf{x} \in M_i$, then Prop. 1.3.10 says that it is a multiple of g_i . This shows that the decomposition $M = M_1 \cup \dots \cup M_r$ is a proper one, and the M_i are M 's irreducible components. \square

We are now going to describe the computation of Gröbner bases, which are a useful tool for elimination of variables, which is the basic ingredient of many algorithms in computational algebraic geometry.

Polynomial Division

There is a procedure called *division with remainder* of a polynomial f by polynomials g_1, \dots, g_m , whose result is not unique, and which is defined as follows:

Consider the greatest term m of f which is a multiple of one of the leading monomials $\text{LM}(g_i)$. Then $f = m / \text{LT}(g_i) \cdot g_i + r_1$, where r_1 is called the *remainder* of this division. We repeat this procedure with r_1 instead of f , and get r_2 , and so on.

The procedure stops if the remainder r_{i+1} becomes zero, or if we cannot find a monomial of r_i which is divisible by one of the leading monomials $\text{LM}(g_i)$. In the former case we say that the remainder of division of f by g_1, \dots, g_m is zero; in the latter case we say that r_i is the remainder.

Example 1.3.7. Let $f = x^2y - x^2 - 2xy^2 + 1$, $g_1 = xy - x$, $g_2 = y + 1$. Then

$$\begin{aligned} f - \frac{x^2y}{xy}g_1 &= -2xy^2 + 1 = r_1 \\ r_1 - \frac{-2xy^2}{xy}g_1 &= -2xy + 1 = r_2 \\ r_2 - \frac{-2xy}{xy}g_1 &= -2x + 1 = r_3 \end{aligned}$$

Here the procedure terminates with remainder $-2x + 1$. \diamond

Note that the result of this procedure is not unique if it is possible to choose more than one $\text{LM}(g_i)$ for division, so we better speak of *a* remainder instead of *the* remainder.

An Algorithm for Computing Gröbner Bases

The following algorithm of B. Buchberger is presented without proof:

Consider an ideal generated by polynomials g_1, \dots, g_r . The greatest common divisor of $\text{LM}(g_i)$ and $\text{LM}(g_j)$ is denoted by t_{ij} . If $t_{ij} = 1$, let $h_{ij} = 0$, otherwise compute

$$m_{ij} = \text{LT}(g_i)/t_{ij} \cdot g_j - \text{LT}(g_j)/t_{ij} \cdot g_i.$$

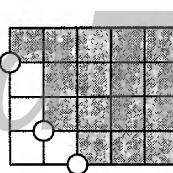
Then h_{ij} is the remainder after division of m_{ij} by g_1, \dots, g_r (it is possible to show that in this special case the remainder is unique).

If all h_{ij} are zero, then g_1, \dots, g_r is a Gröbner basis. Otherwise, we add all nonzero h_{ij} to the generating set g_1, \dots, g_r , and repeat the entire procedure (but we can restrict ourselves to those pairs i, j which have not already been considered in a previous step).

Example 1.3.8. Let $I = \langle g_1, g_2 \rangle$ with $g_1 = x^2$, $g_2 = xy + y^2$. The leading monomials are $\text{LM}(g_1) = x^2$, $\text{LM}(g_2) = xy$. We compute $t_{12} = x$, $m_{12} = (x^2/x)g_2 - (xy/x)g_1 = xy^2$, and division results in $h_{12} = y^3$ (to be exact, $-y^3$, but multiplication by an invertible element, i.e., a real or complex number, makes no difference).

So we let $g_3 = y^3$ and repeat the same for g_1, g_2, g_3 . We compute $t_{13} = 1$ and $t_{23} = y$. This means that $h_{13} = 0$. In order to compute h_{23} we write $m_{23} = (y^3/y)g_2 - (xy/y)g_3 = y^4$, and division results in $y^4 = y \cdot g_3$, so the remainder h_{23} equals zero and we are done.

This shows that $(x^2, xy + y^2, y^3)$ is a Gröbner basis of I , and the leading monomials of all elements of g can be described by the following diagram:



\diamond

Remark 1.3.1. Computer algebra software is often capable of computing Gröbner bases. One example is MAPLE™

After loading the ‘grobner’ package with

```
> with(grobner);
```

we can repeat Ex. 1.3.8 by typing

```
> gbasis([x^2, x*y+y^2], [x, y], plex);
```

\diamond

Remark 1.3.2. A Gröbner basis is not unique. An example is the following: If $(x^3 + xy^2, x^2 + 1, y^2 - 1)$ is a Gröbner basis, then also $(x^2 + 1, y^2 - 1)$ is one, because the leading monomial of $x^3 + xy^2$ is a multiple of $\text{LM}(x^2 + 1)$, and the monomials x^3, x^2, y^2 obviously generate the same ideal as the monomials x^2, y^2 . \diamond

Elimination of Variables

Suppose the algebraic variety M is defined by a set $F = \{g_1, \dots, g_r\}$ of polynomial equations, and we want to find the ‘projection’ of M onto the subspace spanned by x_1, \dots, x_r .

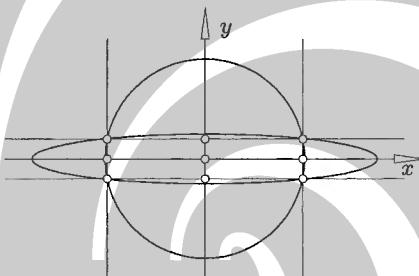


Fig. 1.42. Projecting the intersection of varieties onto the coordinate axes.

Example 1.3.9. The curve $x^2 + y^2 - 1 = 0, x + y + z - 2 = 0$ is an ellipse in the plane $x + y + z = 2$. Its ‘top view’, i.e., its projection onto the xy -plane, is given by the circle $x^2 + y^2 = 1$.

The algebraic variety defined by the equations $g_1(x, y) = x^2 + y^2 - 1 = 0$, $g_2(x, y) = x^2/3 + 4y^2 - 1 = 0$ consists of the intersection points of the two curves $g_1 = 0$ and $g_2 = 0$. Its projection onto the x -axis gives the x -coordinates of the intersection points, and its projection onto the y -axis gives their y -coordinates (cf. Fig. 1.42).

This example shows that the ‘projection’ mentioned above contains also the computation of intersection points. \diamond

The following results show that Gröbner bases can be effectively used for elimination:

Proposition 1.3.12. We consider polynomials in $n + m$ indeterminates which are denoted by $x_1, \dots, x_n, y_1, \dots, y_m$. They are ordered by $x_1 > x_2 > \dots > x_m > y_1 > \dots > y_n$.

Assume that I' is an ideal in $k[x_1, \dots, y_m] = k[\mathbf{x}, \mathbf{y}]$. Then the ideal $I = I' \cap k[\mathbf{y}]$ is spanned by those elements of a Gröbner basis g_1, \dots, g_r of I' which do not contain any x_i .

Proof. Assume that $g \in I$, i.e., g does not contain any of the indeterminates x_i . Its leading term $\text{LT}(g)$ is a multiple of some $\text{LM}(g_i)$. Now also g_i must be in I , because otherwise its leading monomial would contain one of the x_i .

This shows that the leading monomials of those g_i which do not contain any x_i generate $\text{LM}(I)$, and therefore they are a Gröbner basis of I . \square

Proposition 1.3.13. *Assume that $M = V(I')$ is an algebraic variety in k^{n+m} , where $I' \subseteq k[x_1, \dots, x_n, y_1, \dots, y_m]$. The smallest algebraic variety containing M 's projection onto $[y_1, \dots, y_m]$ equals $V(I)$ with $I = I' \cap k[y]$.*

Proof. Clearly M 's projection is contained in $V(I)$. We have to show that $V(I)$ is indeed the *smallest* algebraic variety containing M 's projection: All polynomials which vanish for all y in M 's projection must, if seen as polynomials of $k[\mathbf{x}, \mathbf{y}]$ vanish for all points of M also, and are therefore contained in I' . \square

Example 1.3.10. The projection of an algebraic variety need not be an algebraic variety: The projection of the hyperbola $xy - 1 = 0$ onto the x -axis is the x -axis without the point 0. The smallest algebraic variety containing this set is the x -axis itself.

We compute the intersection points of the curves $x^2 + y^2 = 1$ and $x^2/4 + 3y^2 = 1$: A Gröbner basis of the ideal $\langle x^2 + y^2 - 1, x^2/4 + 3y^2 - 1 \rangle$ is given by $\langle 11x^2 - 8, 11y^2 - 3 \rangle$, which shows that the projection onto the y -axis has y -coordinates $\pm\sqrt{3/11}$, and similarly for the x -axis. \diamond

Applications to Implicitization

If we have a parametrized curve or surface in k^n ($k = \mathbb{R}, \mathbb{C}$) we would like to know the smallest algebraic variety containing it.

This is easily done if the parametrization is polynomial or rational (which is nothing but polynomial in homogeneous coordinates)

Proposition 1.3.14. *Consider affine spaces k^n and k^m , and m polynomials $g_1, \dots, g_m \in k[x_1, \dots, x_n]$. Compute a Gröbner basis of the ideal which is generated by the polynomials $y_1 - g_1(\mathbf{x}), \dots, y_m - g_m(\mathbf{x})$ with respect to the ordering $x_1 > \dots > x_n > y_1 > \dots > y_m$.*

The smallest affine algebraic variety M in k^m which contains all points $(g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))$, $\mathbf{x} \in k^n$, is determined by those elements of the Gröbner basis which do not contain any x_i .

Proof. Clearly the set of points $(g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))$ is the projection of the algebraic variety defined by equations $y_1 - g_1(\mathbf{x}) = \dots = y_m - g_m(\mathbf{x}) = 0$ in k^{n+m} onto the linear subspace spanned by the first n coordinates. \square

Example 1.3.11. Consider the curve $x = t^2 - 1, y = 2 - t$. A Gröbner basis of the ideal $\langle x - t^2 + 1, y - 2 + t \rangle$ with respect to the ordering $t > x > y$ is given by $\langle -y + t - 2, x - 3 - y^2 - 4y \rangle$, so the implicit equation of the curve is given by $x - 3 = y^2 + 4y$. \diamond

Example 1.3.12. Consider the curve $t \mapsto (\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2})$. After embedding affine \mathbb{R}^2 in P^2 we can write this curve as $t \mapsto (1+t^2, 2t, 1-t^2)\mathbb{R}$. The set of all homogeneous coordinate vectors in \mathbb{R}^3 which belongs to these projective points is parametrized by two variables t, u in the form

$$(t, u) \mapsto ((1+t^2)u, 2tu, (1-t^2)u).$$

Now we compute a Gröbner basis of the ideal

$$\langle x_0 - (1+t^2)u, x_1 - 2tu, x_2 - (1-t^2)u \rangle$$

with respect to the ordering $t > u > x_0 > x_1 > x_2$ and get

$$\langle 2u - x_0 - x_2, tx_0 + tx_2 - x_1, tx_1 - x_0 + x_2, x_0^2 - x_1^2 - x_2^2 \rangle.$$

This shows that the implicit equation of the curve is given by $x_0^2 - x_1^2 - x_2^2 = 0$. \diamond

Applications to Polynomial and Rational Mappings

Similar to rational parametrization of curves and surfaces is the problem of polynomial or rational mapping of algebraic varieties: We are given an algebraic variety M , defined by polynomials f_1, \dots, f_r . We apply a mapping $\mathbf{x} \mapsto (g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))$ and search for the smallest algebraic variety containing the image.

This is again done by computing a Gröbner basis of the ideal generated by polynomials $y_1 - g_1(\mathbf{x}), \dots, y_m - g_m(\mathbf{x}), f_1, \dots, f_r$ with respect to the ordering $x_1 > \dots > x_n > y_1 > \dots > y_m$, and disregarding all basis elements which contain any of the variables x_i .

Remark 1.3.3. The image of an algebraic variety need not be an algebraic variety. However it is possible to show that the image of an algebraic variety is in some sense ‘dense’ in the smallest algebraic variety containing it. We will not go into details because they immediately lead into deeper waters. \diamond

Examples where the rational mapping of algebraic varieties is useful will be given later in connection with the computation of *dual* surfaces.

Products of Projective Spaces — Algebraic Relations

Consider projective spaces P^n and P^m over the ground field k ($k = \mathbb{R}$ or $k = \mathbb{C}$). Their product $P^n \times P^m$ is the set of pairs (x, y) with $x \in P^n, y \in P^m$. We write $x = (x_0 : \dots : x_n)$ and $y = (y_0 : \dots : y_m)$. The mapping $(x, y) \mapsto (x_0y_0 : x_0y_1 : \dots : x_0y_m : \dots : x_ny_m)$ of $P^n \times P^m \rightarrow P^{(n+1)(m+1)-1}$ is well-defined. It is called the *Segre embedding*. We denote the coordinates in $P^{(n+1)(m+1)-1}$ by z_{00}, \dots, z_{nm} .

Lemma 1.3.15. *The Segre embedding is one-to-one. The image of $P^n \times P^m$ under the Segre embedding is the Segre variety, which is the zero set of the polynomials*

$$z_{ij}z_{kl} = z_{il}z_{kj}$$

for all possible $i, k = 0, \dots, n$ and $j, l = 0, \dots, m$.

The proof is not difficult and left to the reader. This result shows that we can identify the product $P^n \times P^m$ with a projective algebraic variety contained in a higher-dimensional projective space. When we speak of a projective algebraic variety in $P^n \times P^m$, we always mean a projective variety contained in the Segre manifold.

An interpretation of a projective variety $M \in P^n \times P^m$ is the following: We define a relation ' \sim ' by $x \sim y$ if and only if $(x, y) \in M$. Such a relation is called *algebraic*.

The mapping $\pi_1 : P^n \times P^m \rightarrow P^n$ which assigns the point x to the pair (x, y) is called the projection onto the first factor. Analogously we define the projection π_2 onto the second factor. The image of π_1 is the set of points $x \in P^n$ which are in relation to some $y \in P^m$. The following result will be used in several places, but is here presented without proof (cf. [182], Vol. I, Sec. 5.3.).

Proposition 1.3.16. *We use homogeneous coordinates (x_0, \dots, x_n) in $\mathbb{C}P^n$ and (y_0, \dots, y_m) in $\mathbb{C}P^m$. A subset of the product $\mathbb{C}P^n \times \mathbb{C}P^m$ which is defined by a finite number of polynomials, which are separately homogeneous in the first and second group of variables, is a projective algebraic subvariety (after the Segre embedding), and vice versa.*

If $M \subset \mathbb{C}P^n \times \mathbb{C}P^m$ is a projective algebraic variety, then so is its projection onto the first or second factor.

Example 1.3.13. We define a relation between points in the Euclidean plane by $x \sim y$ if and only if p is a point of the unit circle and y is in x 's tangent. This relation is easily extended to the projective plane and to complex points.

If we use homogeneous coordinates $x = (x_0 : x_1 : x_2)$ and $y = (y_0 : y_1 : y_2)$, we have $x \sim y$ if and only if $x_0^2 = x_1^2 + x_2^2$ (i.e., x is contained in the circle), and $x_0y_0 - x_1y_1 - x_2y_2 = 0$ (i.e., x and y are conjugate with respect to the circle's polarity).

So \sim is an algebraic relation. Its projection onto the first factor is the circle $x_0^2 = x_1^2 + x_2^2$. Its projection onto the second factor is the entire plane (all points of $\mathbb{C}P^2$ are incident with some tangent of the unit circle).

This shows that Prop. 1.3.16 is not valid for real projective spaces, because then the projection onto the second factor would be precisely the outside of the unit circle, which is no algebraic variety. \diamond

Suppose that an algebraic relation is the zero set of polynomials g_1, \dots, g_k as described in Prop. 1.3.16. We want to compute the projection onto the first (or second) factor. It would seem logical just to eliminate the variables y_0, \dots, y_m , as described by Prop. 1.3.12 and Prop. 1.3.13. This won't work, because projections onto subspaces of projective spaces operate with points $(x_0 : \dots : x_n)$ which are well-defined if and only if not all x_i are zero, whereas projections from products of projective spaces operate with pairs $(x_0 : \dots : x_n, y_0 : \dots : y_m)$ of points which are well-defined if and only if neither all x_i nor all y_i are zero.

Example 1.3.14. Consider the algebraic relation in $\mathbb{C}P^3 \times \mathbb{C}P^3$ defined by $x \sim y$ if and only if $x_1 + x_2 = x_2 + x_3 = 0$, $y_1 - y_2 = y_2 - y_3 = 0$, $x_3y_0 = y_3x_0$, $x_2y_0 = y_2x_0$.

It is easy to verify that the solution set of these equations consists of the pairs $(1 : 0 : 0 : 0, 1 : 0 : 0 : 0)$ and $(0 : 1 : -1 : 1, 0 : 1 : 1 : 1)$. Thus the projection π_1 maps the relation onto two points, and the same is true for π_2 .

Elimination of x_0, \dots, x_3 by computing a Gröbner basis of the above set of polynomials with respect to the ordering $x_0 < \dots < x_3$ gives the result $y_1 - y_3 = y_2 - y_3 = 0$, i.e., a line. This is clear because $x_0 = \dots = x_3 = 0$ causes the last two polynomials to evaluate to zero.

To compute the projection, we first restrict ourselves to $x_0 \neq 0$. By homogeneity, we may let $x_0 = 1$. Then elimination of x_1, \dots, x_3 gives the result $y_1 = y_2 = y_3 = 0$, which has the nonzero solution $(1 : 0 : 0 : 0)$.

Second we look for solutions with $x_0 = 0, x_1 = 1$: We get the result $y_0 = y_1 - y_3 = y_2 - y_3 = 0$, which has the nonzero solution $(0 : 1 : 1 : 1)$.

Third, we look for solutions with $x_0 = x_1 = 0, x_2 = 1$, and finally we check whether $x_0 = x_1 = x_2 = 0, x_3 = 1$ leads to a nonzero solution. In both cases we find that there are none. \diamond

1.3.2 Geometric Properties of Varieties in Projective Space

Dimension

We would like to assign a dimension to an algebraic variety. We would expect that an algebraic variety defined by one polynomial equation has dimension $n - 1$, if the dimension of entire space equals n . The following is a ‘natural’ definition of dimension:

Definition. An (affine or projective) algebraic variety M is of dimension m , if there is a chain of irreducible varieties

$$M_0 \subsetneq M_1 \subsetneq \dots \subsetneq M_m$$

with $M_m \subseteq M$, and no longer chain exists.

This definition makes it difficult to compute the dimension of concrete varieties, and it is not clear that the dimension is finite. Further it is not clear that the dimension of affine or projective subspaces (which are algebraic varieties, defined by linear equations) equals their previously defined dimension. Indeed the dimension defined in this way has all the properties we expect, but we will not prove this.

A variety all of whose components have equal dimension is called *pure*. An algebraic variety whose dimension equals 1 is called a *curve*.

Intersection Multiplicities and Degree of a Variety

It is easy to see that a line l , which is not contained in an algebraic hypersurface $M : g(\mathbf{x}) = 0$, intersects M in a finite number of points: The restriction of g to the line l is again a polynomial, which is either the zero polynomial (then $l \subset M$) or has finitely many zeros (the intersection points). Each intersection point has a *multiplicity* which is defined as the multiplicity of the zero of the polynomial.

Proposition 1.3.17. *An algebraic hypersurface $M : g(\mathbf{x}) = 0$ has at most $\deg(g)$ intersection points with a line not contained in M . A projective algebraic hypersurface $M : g(\mathbf{x}) = 0$ has exactly $\deg(g)$ intersection points if $k = \mathbb{C}$ and $g \notin M$ (each intersection point is counted according to its multiplicity).*

Proof. Clearly the restriction of g to l has degree at most $d = \deg(g)$. Assume now that M is a projective algebraic hypersurface, $k = \mathbb{C}$ and $l \not\subset M$. Choose a point $p \in l$ but $p \notin M$, and choose a coordinate system such that the line consists of the points $(u, v, 0, \dots, 0)\mathbb{C}$, and $p = (0, 1, 0, \dots, 0)\mathbb{C}$. $l \not\subset M$ and $p \notin M$ means that the coefficient of the monomial x_1^d is nonzero; so substituting $x_0 = 1, x_1 = t, x_2 = \dots = x_n = 0$ gives a polynomial of degree d with exactly d zeros. \square

There is the following generalization of this proposition:

Theorem 1.3.18. *Almost all lines intersect a complex algebraic hypersurface M in the same number of points (not counting multiplicities). Assume that M is an affine (projective) algebraic variety of dimension d in complex n -dimensional affine (projective) space. Then almost all affine (projective) subspaces of dimension $n - d$ intersect M in the same number of points.*

All $(n - d)$ -dimensional subspaces having more than this number of points in common with M are either contained in M (if $n - d \leq d$) or contain an irreducible component of M (if $n - d > d$).

This so called *generic number of intersection points* is called the *degree* of an algebraic variety V , and is denoted by the symbol $\deg(V)$. We will sometimes use *order* as a synonym for the degree of an algebraic variety. The proposition above shows that the degree of a hypersurface equals the degree of its defining polynomial, if this polynomial is irreducible.

The Hilbert Polynomial

If I is a monomial ideal, we can count the number of monomials *not* in I , which are of total degree s , and the number of monomials not in I of total degree $\leq s$. It turns out that both functions are polynomials for all s greater than some s_0 . This is easy to see from the diagram used in the proof of Prop. 1.3.2: If s is large enough such that we are beyond the circles indicating generators of I , the number of monomials not in I of total degree s (all of which are situated on a line) remains constant, and the number of monomials $\leq s$ not in I is a linear function in s . The general case is also easy to show.

Definition. *Assume that $M = V(I)$ is a projective algebraic variety ($I \trianglelefteq \mathbb{C}[\mathbf{x}]$ is generated by homogeneous polynomials). The Hilbert polynomial $H_M(t)$ is defined as the polynomial function that eventually counts the number of monomials of total degree t not in $\text{LM}(I)$.*

If $M = V(I)$ is an affine algebraic variety, then its Hilbert polynomial is defined as the polynomial function which eventually counts the number of monomials which are of total degree less or equal t , and are not contained in $\text{LM}(I)$.

Example 1.3.15. Consider the twisted cubic $t \mapsto (t, t^2, t^3)$ in \mathbb{R}^3 . Any projective extension must contain the point $(0, 0, 0, 1)\mathbb{R}$ by continuity. Thus the projective extension is defined by the polynomials

$$g_1(x) = x_0x_2 - x_1^2, \quad g_2(x) = x_0x_3 - x_2x_1, \quad g_3(x) = x_1x_3 - x_2^2,$$

because the only ideal point in their zero set is $(0, 0, 0, 1)\mathbb{R}$, and for $x_0 = 1$ this gives the original cubic. Consider the ideal $I = \langle g_1, g_2, g_3 \rangle$. It is easily verified that $\{g_1, g_2, g_3\}$ is a Gröbner basis of I . Then $\text{LM}(I)$ is generated by

$$x_0x_2 = x^{(1,0,1,0)}, \quad x_0x_3 = x^{(1,0,0,1)}, \quad x_1x_3 = x^{(0,1,0,1)}.$$

All monomials x^α with $\alpha_0, \alpha_2 \geq 1$ or $\alpha_0, \alpha_3 \geq 1$ or $\alpha_1, \alpha_3 \geq 1$ are in $\text{LM}(I)$. Thus the monomials not in $\text{LM}(I)$ are the monomials of the form

$$x_0^{\alpha_0}x_1^{\alpha_1}, \quad x_1^{\alpha_1}x_2^{\alpha_2}, \quad x_2^{\alpha_2}x_3^{\alpha_3}.$$

The number of monomials of total degree s among them is given by $3s + 1$. The number of such monomials of total degree less or equal s is given by $(3s^2 + 5s + 2)/2$. \diamond

Theorem 1.3.19. *The degree of the Hilbert polynomial $H_M(t)$ of an algebraic variety equals its dimension $\dim(M)$; the leading coefficient is of the form $d/\dim(M)!$ where d is the degree of M .*

Example 1.3.16. With Ex. 1.3.15 and Th. 1.3.19 we are able to compute the dimension and the degree of the twisted cubic (more precisely, its projective extension M) in complex projective space. From $H_M(t) = 3t + 1$ we conclude $\dim(M) = 1$, $\deg(M) = 3$.

Note that the twisted cubic is defined by three equations, but has dimension one. It can be shown that it is impossible to define the projective extension of the twisted cubic by only two equations, so this curve is not the intersection of two algebraic surfaces. \diamond

The following proposition sums up some properties of the dimension and the degree.

Proposition 1.3.20. *An affine subspace of dimension d also has dimension d as an affine algebraic variety. Likewise the projective and algebraic dimensions agree for projective subspaces. Their degree equals 1.*

Algebraic hypersurfaces have dimension $n - 1$, and all algebraic varieties of dimension $n - 1$ are hypersurfaces. If an algebraic hypersurface is defined by a polynomial of the form $g_1 \cdots g_r$ with g_i irreducible and $g_i \neq g_j$ for $i \neq j$, then its degree equals $\deg(g_1) + \cdots + \deg(g_r)$.

Lüroth Parametrizations

The following lemma, which is presented without proof, is very useful:

Lemma 1.3.21. *If c is a rational curve in $\mathbb{C}P^n$, there is a parametrization $u \in \mathbb{C} \mapsto c(u)\mathbb{R}$, such that (i) for all but a finite number of points p of c there is an $u \in \mathbb{C}$ with $c(u)\mathbb{R} = p$, and (ii) for all but a finite number of points this parameter value u is unique.*

This is a consequence of Lüroth's theorem (cf. [182], p.9, and [197]), and such a parametrization is called a *Lüroth parametrization*.

Remark 1.3.4. The parameter u used in a Lüroth parametrization is an inhomogeneous one. We could make it homogeneous by substituting $u = u_1/u_0$. Because $\mathbb{C}P^1 = \mathbb{C} \cup \infty$, it makes no difference. Lemma 1.3.21 uses arguments from the ordinary complex plane or from the projective complex line. \diamond

Proposition 1.3.22. *If $u \in \mathbb{C} \mapsto c(u)\mathbb{R}$ is a Lüroth parametrization of degree d of the algebraic curve c , then $\deg(c) = d$.*

Proof. To compute $\deg(c)$, we have to count intersection points with hyperplanes. Such a hyperplane has an equation of the form $h \cdot x = 0$, and the equation $h \cdot c(u) = 0$ has d solutions, if we count multiplicities.

If $c(u)$ is a Lüroth parametrization, almost all hyperplanes intersect c in points p_1, \dots, p_r such that there are unique parameter values u_1, \dots, u_r with $c(u_i)\mathbb{R} = p_i$. This shows $\deg(c) = d$. \square

Bézout's Theorem

Assume we have algebraic varieties M_1, M_2 and consider their intersection $M = M_1 \cap M_2$, which is again an algebraic variety. We would like to know something about the degree of M from the degree of M_1 and M_2 . This problem has an easy solution if both M_1, M_2 are planar curves in the complex projective plane. Two curves of degree n_1, n_2 intersect in general in $n_1 \cdot n_2$ points; in any case it is possible to assign multiplicities to intersection points such that the sum of multiplicities equals $n_1 \cdot n_2$. This is Bézout's theorem, which is presented here together with another tool useful for eliminating variables, the *resultant*.

Definition. *If R is a commutative ring and $f, g \in R[x]$ with $f = \sum_{i=0}^m a_i x^i$, $g = \sum_{i=0}^n b_i x^i$, we define $\rho(f, g) = \det(c_{ij})$, where the $(n+m) \times (n+m)$ -matrix c_{ij} is given by $c_{ij} = a_j - i$ for $i \leq n$ and $c_{ij} = b_j - i + n$ for $i > n$. $\rho(f, g)$ is called the resultant of f, g .*

Here the symbol $R[x]$ means the ring of polynomials with coefficients in R and the indeterminate x . All undefined values c_{ij} are set to zero.

Theorem 1.3.23. *(of E. Bézout): Two algebraic curves c_1, c_2 defined by homogeneous polynomials $p_1, p_2 \in \mathbb{C}[x_0, x_1][x_2]$ of degree n_1, n_2 , respectively, without common components have finitely many intersection points s_1, \dots, s_r . Assume without loss of generality that $(0 : 0 : 1) \notin c_1, c_2$. Then*

$$\rho(p_1, p_2) = C \cdot \prod_{\rho=1}^r (x_0 s_{\rho,1} - x_1 s_{\rho,0})^{k(\rho)},$$

i.e., the s_ρ can be calculated in part by a prime factor decomposition of the resultant. The points s_ρ themselves are then found by intersecting the lines $x_1 : x_0 = s_{\rho,1} : s_{\rho,0}$ with both curves. If such a line contains exactly one intersection point, then the exponent $k(\rho)$ is called multiplicity of s_ρ , and the sum of multiplicities equals $m \cdot n$.

The curves c_1, c_2 intersect transversely, if and only if $k_\rho = 1$.

The proof is a rather lengthy, but not very difficult, investigation of various properties of the resultant. Transverse intersection means that the common point of the two curves is regular for both of them and the curve tangents are different (this will be discussed later).

Example 1.3.17. We want to compute the intersection points of the circles $c_1 : x^2 + y^2 = 1$ and $c_2 : x^2 + (y - 1)^2 = 1$, including the points at infinity which their projective and complex extensions may have in common.

First we embed \mathbb{R}^2 into \mathbb{C}^2 and further into the complex projective plane via $x = x_1/x_0$ and $y = x_2/x_0$. The projective extensions of c_1 and c_2 have equations

$$(-x_0^2 + x_1^2)x_2^0 + 0 \cdot x_2^1 + 1 \cdot x_2^2 = 0, \quad x_1^2 \cdot x_2^0 - 2x_0 \cdot x_2^1 + 1 \cdot x_2^2 = 0.$$

We have written the equations in this form because Th. 1.3.23 requires polynomials in the indeterminate x_2 . The resultant of these polynomials then equals

$$\begin{vmatrix} x_1^2 - x_0^2 & 0 & 1 & 0 \\ 0 & x_1^2 - x_0^2 & 0 & 1 \\ x_1^2 & -2x_0 & 1 & 0 \\ 0 & x_1^2 & -2x_0 & 1 \end{vmatrix} = x_0^2(2x_1 - \sqrt{3}x_0)(2x_1 + \sqrt{3}x_0).$$

This shows that the intersection points are contained in the line at infinity $x_0 = 0$ (with multiplicity 2) and in the lines $x_1 : x_0 = x = \pm\sqrt{3}/2$. Intersecting these lines with both curves gives the two conjugate complex points $(0, 1, \pm i)\mathbb{C}$ at infinity and $(1, \pm\sqrt{3}/2, 1/2)\mathbb{C}$, which stem from the two real and finite intersections $(\pm\sqrt{3}/2, 1/2)$. \diamond

Regular and Singular Points

We want to investigate the behaviour of an algebraic variety M in the neighbourhood of a point p . For simplicity, we assume that M is irreducible.

Assume that M is an affine variety given by the equations $g_1(\mathbf{x}) = \dots = g_r(\mathbf{x}) = 0$, and $\mathbf{p} \in M$. We consider all lines with parametrization $\mathbf{p} + \lambda \mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. If we substitute $\mathbf{x} = \mathbf{p} + \lambda \mathbf{v}$ in g_1, \dots, g_r we get r polynomials which have a zero at $\lambda = 0$. The minimum multiplicity of these zeros minus 1 is called the *contact order* of the line with M .

If M is a projective algebraic variety, we calculate the contact order in an affine part of projective space. A line which has contact order 0 is called *transverse* to M , the others are called *tangent* to M .

It is easily seen that the line $\mathbf{p} + \lambda\mathbf{v}$ is tangent to the affine variety M if and only if $\mathbf{v} \cdot \text{grad}_{g_1}(\mathbf{p}) = \dots = \mathbf{v} \cdot \text{grad}_{g_r}(\mathbf{p}) = 0$; and if M is a projective variety, the points $x = x\mathbb{R}$ or $x = x\mathbb{C}$ of the tangents fulfill $x \cdot \text{grad}_{g_1}(\mathbf{p}) = \dots = x \cdot \text{grad}_{g_r}(\mathbf{p}) = 0$. This shows that the union of all tangents is a projective subspace, which is called *tangent space*. If its dimension equals $\dim(M)$, the corresponding point is called a *regular point*, otherwise the point is called *singular*. The classification of singularities then uses tangents whose contact order is greater than the minimal contact order of lines in the point in question. The reader nevertheless should note that the presence of lines which enjoy a high (or even infinite) contact order does not imply a singularity.

Example 1.3.18. Consider the planar curve $t \mapsto (t^2, t^3)$. Its projective extension M , which consists of all points $(1, t^2, t^3)\mathbb{C}$ together with the point $(0, 0, 1)\mathbb{C}$, has the equation $g(x) = x_1^3 - x_2^2x_0 = 0$.

The gradient of g is given by $\text{grad}_g(x) = (-x_2^2, 3x_1^2, -2x_0x_2)$. Singular points therefore fulfill the equations $x_2^2 = 3x_1^2 = -2x_0x_2 = 0$. The only solution is the point $(1, 0, 0)\mathbb{C}$. All other points are regular points.

The affine part of M has the equation $x_1^3 - x_2^2 = 0$. Substituting $x_1 = \lambda u, x_2 = \lambda v$ gives the polynomial $\lambda^2(\lambda u^3 - v^2) = 0$. This shows that all lines incident with $(1, 0, 0)\mathbb{C}$ except the x_1 -axis have contact order one, and the x_1 -axis has contact order two. Such a singularity is called a *cusp* (see Fig. 1.40, right). ◇

We have defined a tangent in several ways: A curve or surface has a tangent in the sense of differential geometry. An algebraic curve or surface also has tangents as an algebraic variety. It is however easy to show that if the image of a curve in the sense of differential geometry is contained in an algebraic variety M , its tangent is contained in M 's tangent space. In regular points p of M the tangents of all regular curves passing through p span the tangent space.

1.3.3 Duality

A projective algebraic hypersurface M , given by the equation $g(x) = 0$, has in its regular points a tangent hyperplane, whose equation is given by $\text{grad}_g(p) \cdot x = 0$. Thus the hyperplane coordinates of the tangent plane at $p\mathbb{R}$ (or $p\mathbb{C}$) are given by $\mathbb{R} \cdot \text{grad}_g(p)$ (or $\mathbb{C} \cdot \text{grad}_g(p)$). The smallest projective variety in dual projective space which contains all tangent hyperplanes is the image of M under the mapping

$$x \mapsto \text{grad}_g(x).$$

It can be computed e.g. using Gröbner bases.

Example 1.3.19. Consider the curve $M : g(x) = x_0x_1^2 - x_2^3 = 0$ in the real projective plane. Then $\text{grad}_f(x) = (x_1^2, 2x_0x_1, -3x_2^2)$. The dual curve is the image of M under the mapping $(x_0, x_1, x_2) \mapsto (x_1^2, 2x_0x_1, -3x_2^2)$. A Gröbner basis of the ideal

$$\langle y_0 - x_1^2, y_1 - 2x_0x_1, y_2 + 3x_2^2, x_0x_1^2 - x_2^3 \rangle$$

with respect to the ordering $x_0 > x_1 > x_2 > y_0 > y_1 > y_2$ contains one polynomial which does not involve any x_i , namely

$$27y_0y_1^2 + 4y_2^3,$$

which defines the dual curve M^* . \diamond

Example 1.3.20. The quadratic cone $x_1^2 + x_2^2 - x_0^2 = 0$ of P^3 is mapped by

$$(x_0, x_1, x_2, x_3) \mapsto (-2x_0, 2x_1, 2x_2, 0)$$

to the variety defined by the equations

$$y_1^2 + y_2^2 - y_0^2 = 0, \quad y_3 = 0.$$

This shows that the dual M^* need not be a hypersurface. In this case it is of dimension 1. \diamond

Definition. If M is a projective algebraic variety, the degree of its dual variety M^* is called the class of M .

1.4 Rational Curves and Surfaces in Geometric Design

A rational curve is a mapping $c : \mathbb{R} \rightarrow P^n$ which has a polynomial representation when using homogeneous coordinates. A rational d -surface is likewise a mapping $s : \mathbb{R}^d \rightarrow P^n$. We will apply the term rational curve (or surface) also to the smallest projective algebraic variety which contains the image of c (or s). If it is sometimes not clear which concept we refer to, this is because it makes no difference anyway (e.g., when we speak of the curve's tangent in a regular point).

1.4.1 Rational Bézier Curves

Polynomial Bézier Curves

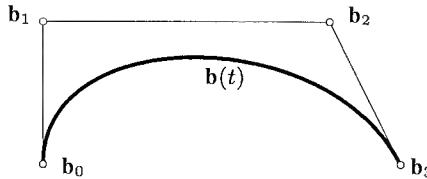
We consider the Bernstein polynomials of degree n , which are given by

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i} \quad (i = 0, \dots, n). \quad (1.88)$$

A polynomial function $\mathbb{R} \rightarrow \mathbb{R}^m$ which is a linear combination of Bernstein polynomials

$$\mathbf{b}(t) = \sum_{i=0}^n B_i^n(t) \mathbf{b}_i \quad (1.89)$$

is called *Bézier curve*, and the coefficients \mathbf{b}_i are called *control points* (see Fig. 1.43). The sequence $\mathbf{b}_0, \mathbf{b}_1, \dots$ is called the *control polygon* of the Bézier curve.

**Fig. 1.43.** Cubic Bézier curve.

Lemma 1.4.1. A Bézier curve $\mathbf{b}(t)$ defined by control points $\mathbf{b}_0, \dots, \mathbf{b}_n \in \mathbb{R}^m$ has the following properties:

1. The first derivative of $\mathbf{b}(t)$ is given by

$$\dot{\mathbf{b}}(t) = n \cdot \sum_{i=0}^{n-1} B_i^{n-1}(t) \Delta \mathbf{b}_i, \quad (1.90)$$

where $\Delta \mathbf{b}_i = \mathbf{b}_{i+1} - \mathbf{b}_i$, and is therefore itself a Bézier curve.

2. The sum $\sum_{i=0}^n B_i^n(t) = 1$, and the curve point $\mathbf{b}(t)$ depends on the control points in an affinely invariant way.
3. The curve point $\mathbf{b}(t)$ is contained in the convex hull of the control points if $0 \leq t \leq 1$. This is called the convex hull property.

Proof.

1. This follows directly by differentiating $B_i^n(t)$.
2. We compute $1 = (t + (1-t))^n = \sum_{i=0}^n \binom{n}{i} t^i (1-t)^{n-i} = \sum B_i^n(t)$. Therefore $\mathbf{b}(t)$ is an *affine* combination of the points \mathbf{b}_i , which means a linear combination whose coefficients sum up to 1.
3. If $0 \leq t \leq 1$, then $B_i^n(t) \geq 0$ by definition. Because $\sum B_i^n(t) = 1$ it follows that $B_i^n(t) \leq 1$. This shows that $\mathbf{b}(t)$ is even a *convex* combination of the points \mathbf{b}_i .

□

Iteration of (1.90) shows that all derivatives of a Bézier curve are Bézier curves, and furthermore it shows that the r -th derivative at $t = 0$ is a linear combination of the vectors

$$\mathbf{b}_1 - \mathbf{b}_0, \dots, \mathbf{b}_r - \mathbf{b}_0,$$

and the coefficient of $\mathbf{b}_r - \mathbf{b}_0$ is nonzero. Likewise the r -th derivative at $t = 1$ is a linear combination of the vectors

$$\mathbf{b}_{n-1} - \mathbf{b}_n, \dots, \mathbf{b}_{n-r} - \mathbf{b}_n,$$

such that the coefficient of $\mathbf{b}_{n-r} - \mathbf{b}_n$ is nonzero. It follows that the points $\mathbf{b}_{r+1}, \dots, \mathbf{b}_n$ have no influence on the first r derivatives of $\mathbf{b}(t)$ in $t = 0$.

The Algorithm of de Casteljau

We assume that points $\mathbf{b}_0, \dots, \mathbf{b}_n$ and a real number t are given, and construct a point $\mathbf{b}(t)$ by the following recursion, which is called the *algorithm of de Casteljau* (see Fig. 1.44).

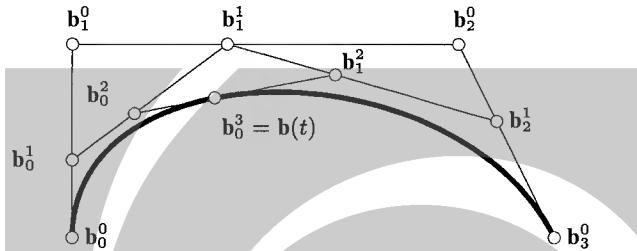


Fig. 1.44. Algorithm of de Casteljau.

1. $\mathbf{b}_0^0 = \mathbf{b}_0, \dots, \mathbf{b}_n^0 = \mathbf{b}_n$.
2. $\mathbf{b}_i^r = (1-t)\mathbf{b}_i^{r-1} + t\mathbf{b}_{i+1}^{r-1}$ ($r = 1, \dots, n$, $i = 0, \dots, n-r$).
3. $\mathbf{b}(t) = \mathbf{b}_n^n$.

Lemma 1.4.2. *The algorithm of de Casteljau applied to points $\mathbf{b}_0, \dots, \mathbf{b}_n$ and the real number t evaluates the polynomial Bézier curve with these control points at the parameter value t .*

Proof. We first show that evaluating the Bézier curve with control points $\mathbf{b}_i, \dots, \mathbf{b}_{i+r}$ at the parameter value t equals the point \mathbf{b}_i^r of the algorithm. This is trivial for $r = 0$ and easily shown by induction for $r > 0$, because

$$(1-t) \sum_{j=0}^{r-1} B_j^{r-1}(t) \mathbf{b}_{i+j} + t \sum_{j=0}^{r-1} B_j^{r-1}(t) \mathbf{b}_{i+1+j} = \sum_{j=0}^r B_j^r(t) \mathbf{b}_{i+j}.$$

The statement of the lemma is now the special case $i = 0$ and $r = n$. □

Multi-affine Polar Forms

Clearly all Bézier curves are polynomial curves. It is even possible to show that all polynomial curves are Bézier curves. This is done in the following way: Assume that we are given a polynomial function $\mathbf{c} : \mathbb{R} \rightarrow \mathbb{R}^m$ of degree $\leq n$, and we want to find $\mathbf{b}_0, \dots, \mathbf{b}_n$ such that $\mathbf{c}(t) = \sum_i B_i^n(t) \mathbf{b}_i$. It is sufficient to do this component-wise, so we restrict ourselves to the case of dimension one. It is also sufficient to do this for all monomials which take part in c , because everything is linear.

Consider the monomial t^r . For all $n \geq r$ we define an n -variate function P by

$$P(t_1, \dots, t_n) = \binom{n}{r}^{-1} \sum_{\substack{I \subset \{1, \dots, n\} \\ \#I=r}} \prod_{i \in I} t_i.$$

The meaning of this definition is the following: Take an r -element subset I of the set $\{1, \dots, n\}$, and multiply the variables t_i for all $i \in I$. Do this for all r -element subsets and sum up the products. Divide by $\binom{n}{r}$, which is the number of all possible I 's. This function (apart from the factor) is the r -th elementary symmetric function of n arguments.

Example 1.4.1. The following list shows some examples for low values of n and r :

$$\begin{aligned} n=2, r=0: \quad P(t_1, t_2) &= 1 \\ n=2, r=1: \quad P(t_1, t_2) &= (t_1 + t_2)/2 \\ n=2, r=2: \quad P(t_1, t_2) &= t_1 t_2 \\ n=3, r=0: \quad P(t_1, t_2, t_3) &= 1 \\ n=3, r=1: \quad P(t_1, t_2, t_3) &= (t_1 + t_2 + t_3)/3 \\ n=3, r=2: \quad P(t_1, t_2, t_3) &= (t_1 t_2 + t_1 t_3 + t_2 t_3)/3 \\ n=3, r=3: \quad P(t_1, t_2, t_3) &= t_1 t_2 t_3 \end{aligned}$$

◇

The function P has the property $P(t, \dots, t) = t^r$, and its value does not change when we permute the arguments, which is clear from the definition. A further important property of P will be stated after the following definition.

Definition. A mapping $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called *affine*, if $f(t\mathbf{a} + (1-t)\mathbf{b}) = tf(\mathbf{a}) + (1-t)f(\mathbf{b})$ for all arguments \mathbf{a}, \mathbf{b} and real values t .

We see that being an *affine* mapping is a bit less than being a linear one. It is easily verified that P is affine in each argument, i.e.,

$$P(\dots, ta + (1-t)b, \dots) = tP(\dots, a, \dots) + (1-t)P(\dots, b, \dots), \quad (1.91)$$

for all $a, b, t \in \mathbb{R}$. The dots stand for arguments which are the same in all three instances of P .

Therefore the mapping P is called the *multi-affine polar form* of the corresponding monomial which was used for the definition of P . If a polynomial is a linear combination of monomials, the appropriate linear combination of the monomials' polar forms is called the polar form of the polynomial. The polar form of an \mathbb{R}^m -valued polynomial is defined component-wise (for an example, see Ex. 1.4.2).

Theorem 1.4.3. If $\mathbf{c}(t) = \sum \mathbf{c}_i t^i$ is a real-valued polynomial function of degree $\leq n$, then it can be written in the form $c(t) = \sum_i B_i^n(t) \mathbf{b}_i$, where

$$\mathbf{b}_i = P(\underbrace{0, \dots, 0}_{n-i}, \underbrace{1, \dots, 1}_i), \quad (1.92)$$

and P is the n -variate polar form of c .

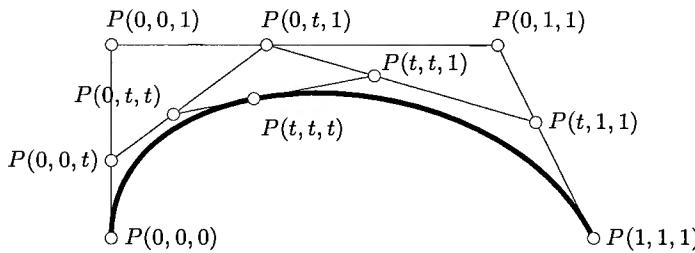


Fig. 1.45. Bézier curve and selected values of its polar form.

Proof. The multi-affinity of P ensures that P satisfies the equation $(1 - t_n) P(t_1, \dots, t_{n-1}, 0) + t_n P(t_1, \dots, t_{n-1}, 1) = P(t_1, \dots, t_n)$. Thus the points $\mathbf{p}_i^r := P(\underbrace{0, \dots, 0}_{n-r-i}, \underbrace{1, \dots, 1}_i, \underbrace{t, \dots, t}_r)$ satisfy the same recurrence relation as the points \mathbf{b}_i^r used in the algorithm of de Casteljau. This shows that $\mathbf{p}_i^r = \mathbf{b}_i^r$, and especially that $\mathbf{p}_0^n = \mathbf{b}_0^n$. Thus Lemma 1.4.2 implies the statement of the theorem (see Fig. 1.45). \square

Example 1.4.2. Consider the planar polynomial curve $c(t) = (t^2 - 1, 2t)$. We want to write this curve as Bézier curve of degree 2. The 2-variate polar forms of the monomials $1, t, t^2$, are given by $1, \frac{1}{2}(t_1 + t_2), t_1 t_2$, respectively. Thus the polar form of c is given by $P(t_1, t_2) = (t_1 t_2 - 1, t_1 + t_2)$, and we can compute the control points as $\mathbf{b}_0 = P(0,0) = (-1,0)$, $\mathbf{b}_1 = P(0,1) = (-1,1)$, and $\mathbf{b}_2 = P(1,1) = (0,2)$. \diamond

The polar form of a polynomial is uniquely determined by the properties listed above:

Lemma 1.4.4. *If $p(u)$ is a polynomial function of degree $\leq n$, and $Q(u_1, \dots, u_n)$ is a symmetric multi-affine n -variate real-valued function with $Q(u, \dots, u) = p(u)$, then Q coincides with the n -variate polar form P of p .*

Proof. Q is, by multi-affinity, uniquely determined by its values $Q(\underbrace{0, \dots, 0}_{n-k}, 1, \dots, 1)$. By Th. 1.4.3, these values are uniquely determined by the polynomial p , because the proof of Th. 1.4.3 (which employs P instead of Q) uses only such properties of P which are shared by Q . \square

Another property of the multi-affine polar form is the following:

Lemma 1.4.5. *Two polynomial curves \mathbf{f}, \mathbf{g} with multi-affine polar forms S, T share the same derivatives at $u = u_0$ up to order k , if and only if*

$$S(u_1, \dots, u_k, \underbrace{u_0, \dots, u_0}_{n-k}) = T(u_1, \dots, u_k, \underbrace{u_0, \dots, u_0}_{n-k})$$

for all u_1, \dots, u_k .

Proof. It is sufficient to show this for $u_0 = 0$. Iteration of Equ. (1.90) shows that the first k derivatives of \mathbf{f} , \mathbf{g} at $u = 0$ determine the first $k + 1$ Bézier points $\mathbf{b}_0, \dots, \mathbf{b}_k$ of \mathbf{f} (or \mathbf{g} , which is the same), and vice versa. By Th. 1.4.3, these are the points $S(1, \dots, 1, \underbrace{0, \dots, 0}_{\geq n-k})$. This shows the ‘if’ part of the lemma. The ‘only if’ part follows if we note that these special values actually determine *all* values $(u_1, \dots, u_k, \underbrace{0, \dots, 0}_{n-k})$. \square

Segments of Bézier Curves — Subdivision

Any segment of a Bézier curve may be re-parametrized so that it is a polynomial curve segment whose parameter ranges in the interval $[0, 1]$. Thus any Bézier curve segment is a Bézier curve in its own right. We describe how to find the control points of such a segment:

Corollary 1.4.6. *If $\mathbf{c}(t)$ is an \mathbb{R}^m -valued polynomial curve of degree $\leq n$, and P is its n -variate polar form, then the control points of the curve segment $\mathbf{c}([a, b])$ are the points \mathbf{b}_i with*

$$\mathbf{b}_i = P(\underbrace{a, \dots, a}_{n-i}, \underbrace{b, \dots, b}_i). \quad (1.93)$$

Proof. We consider the curve $\mathbf{d}(t) = \mathbf{c}((t-a)/(b-a))$. Then $\mathbf{d}([0, 1]) = \mathbf{c}([a, b])$, and it is easily verified that the polar form Q of \mathbf{d} likewise fulfills the equation $Q(t_1, \dots, t_n) = P((t_1-a)/(b-a), \dots, (t_n-a)/(b-a))$. Th. 1.4.3 now immediately implies the statement we wanted to show. \square

Example 1.4.3. Consider the Bézier curve of Fig. 1.44 and Fig. 1.45. The point $\mathbf{b}_0^3 = \mathbf{b}(t)$ divides the curve into two segments, namely $\mathbf{b}([0, t])$ and $\mathbf{b}([t, 1])$. By Cor. 1.4.6, the control points of the former are $\mathbf{b}_0^0 = P(0, 0, 0)$, $\mathbf{b}_0^1 = P(0, 0, t)$, $\mathbf{b}_0^2 = P(0, t, t)$, $\mathbf{b}_0^3 = P(t, t, t)$, and the control points of the latter are $\mathbf{b}_0^0 = P(t, t, t)$, $\mathbf{b}_1^1 = P(t, t, 1)$, $\mathbf{b}_2^2 = P(t, 1, 1)$, $\mathbf{b}_3^3 = P(1, 1, 1)$. We see that the control points of the two curve segments are points which already appear in the algorithm of de Casteljau. \diamond

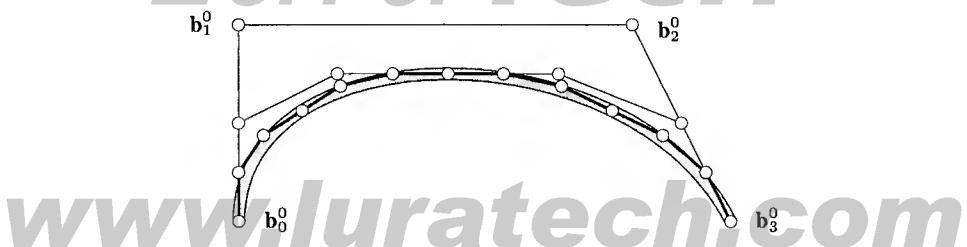


Fig. 1.46. Subdivision of Bézier curves with control polygons.

Remark 1.4.1. If we iterate the process described by Ex. 1.4.3 we get a sequence of $2, 4, 8, \dots$ curve segments together with their control polygons. It is possible to show that the sequence of control polygons converges towards the Bézier curve rapidly, in the sense that points and lines of the polygon approximate points and tangents of the Bézier curve (see Fig. 1.46). \diamond

The following corollary shows how the algorithm of de Casteljau computes not only the curve point, but also its tangent. This result will have several applications, e.g., when estimating the degree of a dual curve.

Corollary 1.4.7. *If \mathbf{b} is a Bézier curve with control points $\mathbf{b}_0, \dots, \mathbf{b}_n$, then the curve tangent in the point $\mathbf{b}(t)$ is spanned by the points \mathbf{b}_0^{n-1} and \mathbf{b}_1^{n-1} which are computed with the algorithm of de Casteljau.*

Proof. We consider the curve segment $\mathbf{b}([0, t])$. The proof of Th. 1.4.3 and Cor. 1.4.6 show that its control points are $\mathbf{b}_0^0, \dots, \mathbf{b}_0^n$ (cf. also Ex. 1.4.3). Equ. (1.90) implies that the end tangent of a Bézier curve is spanned by the last two control points. Thus \mathbf{b}_0^{n-1} is contained in the curve tangent at $\mathbf{b}(t)$. To show that also \mathbf{b}_1^{n-1} is contained in this tangent, we consider the segment $\mathbf{b}([t, 1])$ and repeat the same argument. \square

Variation Diminishing Property and Convex Bézier Curves

There is the following theorem concerning the number of intersections of a Bézier curve and a hyperplane:

Theorem 1.4.8. *A hyperplane H intersects a polynomial Bézier curve (not contained in H) in at most as many points as it intersects the control polygon of this curve.*

This is called the *variation diminishing property* of Bézier curves.

Proof. (Sketch) We consider the process of dividing a Bézier curve into $2, 4, 8, \dots$ segments as described by Remark 1.4.1. We consider the polygons P_i which consist of all 2^i control polygons of the i -th iteration step for $i \geq 1$, and P_0 is the control polygon of the original curve. Obviously the polygon P_1 is constructed from P_0 by ‘cutting vertices’. P_2 is constructed from P_1 likewise, and so on. This procedure does not increase the essential number of intersection points with a given hyperplane, and the convergence mentioned in Remark 1.4.1 shows the statement of the theorem. \square

Corollary 1.4.9. *A Bézier curve is convex if its control polygon is convex.*

Proof. Convexity means that the number of intersections with any line does not exceed two. Thus, Th. 1.4.8 immediately implies the result. \square

Examples of convex control polygons are shown by most illustrations in this section.

Rational Bézier Curves and their Osculating Subspaces

The Bézier curve $c : \mathbb{R} \rightarrow P^m$ with

$$c(t) = \mathbf{c}(t)\mathbb{R} = \left(\sum_{i=0}^n B_i^n(t) \mathbf{b}_i \right) \mathbb{R}, \quad (\mathbf{b}_i \in \mathbb{R}^{m+1}) \quad (1.94)$$

is called a *rational* Bézier curve. It follows from the previous discussion that all rational curves can be written as Bézier curves. The Bézier representation is advantageous from the viewpoint of projective differential geometry, because from the control points we immediately know osculating subspaces at two different values of t . If we use the monomial basis, we know them only at $t = 0$.

Theorem 1.4.10. *The k -th projective osculating subspace of a rational Bézier curve $c(t) = (\sum_i B_i^n(t) \mathbf{b}_i) \mathbb{R}$ at $t = 0$ equals $\mathbf{b}_0 \mathbb{R} \vee \dots \vee \mathbf{b}_k \mathbb{R}$. Likewise the k -th osculating subspace at $t = 1$ is given by $\mathbf{b}_n \mathbb{R} \vee \dots \vee \mathbf{b}_{n-k} \mathbb{R}$.*

Proof. The k -th osculating subspace at $t = 0$ is spanned by the points $\mathbf{c}(0)\mathbb{R}$, $\dot{\mathbf{c}}(0)\mathbb{R}, \dots, \mathbf{c}^{(k)}(0)\mathbb{R}$. We already know that $\mathbf{c}^{(r)}(0)$ is a linear combination of the vectors $\mathbf{b}_1 - \mathbf{b}_0, \dots, \mathbf{b}_r - \mathbf{b}_0$, with the coefficient of $\mathbf{b}_r - \mathbf{b}_0$ being nonzero. This shows that $\mathbf{b}_0 \mathbb{R} \vee \dots \vee \mathbf{b}_k \mathbb{R}$ actually equals the k -th osculating subspace at $t = 0$. The situation at $t = 1$ is completely analogous. \square

Example 1.4.4. Consider the semicubic parabola with equation $y^2 = x^3$ (cf. Fig. 1.48, left, which shows control points different from those computed here). It has the parametrization $t \mapsto (t^2, t^3)$. After the parameter transform $t = u/(u-1)$ we have the parametrization $u \mapsto ((u/(u-1))^2, (u/(u-1))^3)$, which in homogeneous coordinates can be written as $u \mapsto ((u-1)^3, u^2(u-1), u^3)\mathbb{R} = (u^3 - 3u^2 + 3u - 1, u^3 - u^2, u^3)\mathbb{R}$. The trivariate polar form of this polynomial is given by $P(u_1, u_2, u_3) = (u_1 u_2 u_3 - (u_1 u_2 + u_1 u_3 + u_2 u_3) + (u_1 + u_2 + u_3) - 1, u_1 u_2 u_3 - (u_1 u_2 + u_1 u_3 + u_2 u_3)/3, u_1 u_2 u_3)$. Now the control points $\mathbf{b}_i \mathbb{R}$ can be computed by $\mathbf{b}_0 = P(0, 0, 0) = (-1, 0, 0)$, $\mathbf{b}_1 = P(0, 0, 1) = (0, 0, 0)$ (which does not represent a point), $\mathbf{b}_2 = P(0, 1, 1) = (0, -1/3, 0)$, $\mathbf{b}_3 = P(1, 1, 1) = (0, 0, 1)$. This shows that the sequence of dimensions of the osculating spaces is given by $0, 0, 1, 2, \dots$ at $u = 0$, and by $0, 1, 1, 2, \dots$ at $u = 1$, which means that the curve has an ordinary cusp at $u = 0$, and an inflection point at $u = 1$ (at infinity). \diamond

Rational Normal Curves

It is sometimes useful to remember that a rational curve of degree k in P^n ($k > n$) can be generated from a standard polynomial curve in P^k by applying a central projection plus a projective isomorphism. We write this down in the form of a lemma:

Lemma 1.4.11. *For a rational curve $c(u) = \mathbf{c}(u)\mathbb{R}$ of degree k in P^n ($k > n$) there is a singular projective mapping $\pi : P^k \rightarrow P^n$ such that $c(u) = ((1, u, \dots, u^k)\mathbb{R})\pi$.*

Proof. We assume that $\mathbf{c}(u) = (c_0(u), \dots, c_n(u))$ with $c_j(u) = \sum_{i=0}^k a_{ij} u^i$. We collect the coefficients a_{ij} in a matrix A . Obviously the linear mapping $\pi : \mathbb{X}\mathbb{R} \mapsto (A \cdot \mathbf{x})\mathbb{R}$ has the required properties. \square

A curve which has, with respect to a certain projective coordinate system, the parametrization $c(u) = (1, u, u^2, \dots, u^k)\mathbb{R}$ is called a *rational normal curve* of projective k -space. To make it complete, we have to add the point $c(\infty) = (0, \dots, 0, 1)\mathbb{R}$. Obviously all of them are projectively equivalent.

A normal curve has no ‘characteristic points’ — for all u and v there is a projective automorphism of P^k which leaves the curve invariant and takes $c(u)$ to $c(v)$. A rational curve in general can have a certain number of singular points, inflection points, and so on, which are distinguished among the other points of the curve by the dimension of their osculating subspaces. If this curve is represented, according to Lemma 1.4.11, as linear image of a normal curve, the ‘shape information’ must be contained in the center of this linear mapping, as linear mappings with the same center yield projectively equivalent images. A discussion of shapes of rational curves based on the location of the projection center can be found in [148].

Geometric Control Polygons

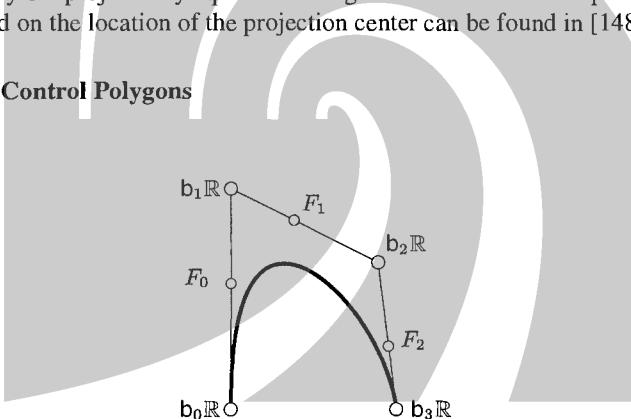


Fig. 1.47. Geometric control polygon of a rational Bézier curve.

We want to describe a rational Bézier curve $c(t) = (\sum_i B_i^n(t)b_i)\mathbb{R}$ of P^n not by a sequence b_i of control vectors of \mathbb{R}^{n+1} , but by points of projective space P^n itself. This can be done as follows: If we know $b_0 \in \mathbb{R}^{n+1}$, and the projective points $(b_0 + b_1)\mathbb{R}$ and $b_1\mathbb{R}$, then we can reconstruct b_1 :

Suppose $\mathbf{c} = \lambda(b_0 + b_1)$, $\mathbf{d} = \mu b_1$, and only $b_0, \mathbf{c}, \mathbf{d}$ are known. These vectors are linearly dependent, so we can write $\mathbf{c} = \alpha b_0 + \beta \mathbf{d}$, or $\frac{1}{\alpha} \mathbf{c} = b_0 + \frac{\beta}{\alpha} \mathbf{d}$, which shows $b_1 = \frac{\beta}{\alpha} \mathbf{d}$.

Clearly multiplying *all* control points b_i with the same scalar factor does not change the curve $c(t)$; so the points $b_i\mathbb{R}$ together with the points $(b_i + b_{i+1})\mathbb{R}$ completely determine the control vectors b_i , after we have chosen b_0 arbitrarily, and so the curve $c(t)$ is uniquely defined. An example is shown in Fig. 1.47.

The points $B_i = b_i\mathbb{R}$ of projective space are again called *control points*. The points $F_i = (b_i + b_{i+1})\mathbb{R}$ are called *frame points* and have been introduced by G. Farin.

This approach fails if successive control vectors b_i belong to the same projective point, or do not belong to a projective point at all. In that case it is necessary to assign *weights* to the points b_i , which are nothing but b_i 's first coordinate. Fig. 1.48 shows an example of this.

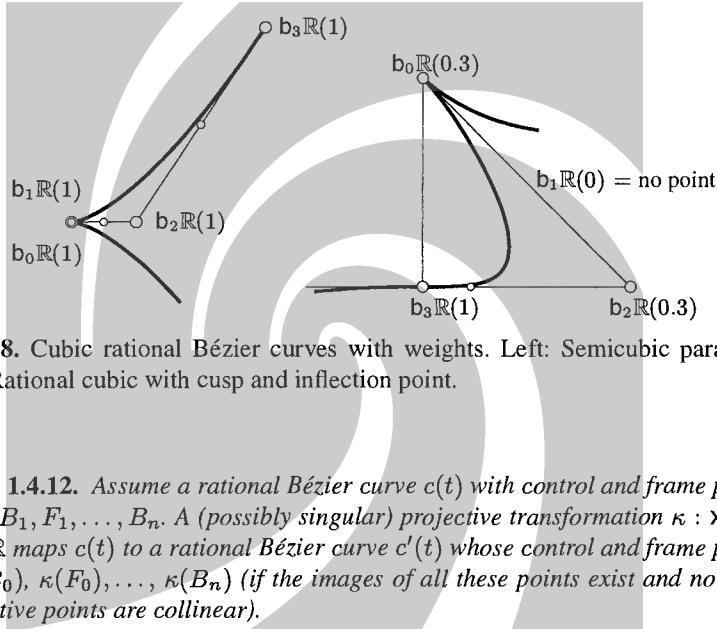


Fig. 1.48. Cubic rational Bézier curves with weights. Left: Semicubic parabola; Right: Rational cubic with cusp and inflection point.

Lemma 1.4.12. Assume a rational Bézier curve $c(t)$ with control and frame points $B_0, F_0, B_1, F_1, \dots, B_n$. A (possibly singular) projective transformation $\kappa : \mathbf{x}\mathbb{R} \mapsto (A \cdot \mathbf{x})\mathbb{R}$ maps $c(t)$ to a rational Bézier curve $c'(t)$ whose control and frame points are $\kappa(B_0), \kappa(F_0), \dots, \kappa(B_n)$ (if the images of all these points exist and no three consecutive points are collinear).

Proof. Assume $c(t) = c(t)\mathbb{R}$ with $c(t) = \sum_i B_i^n(t)b_i$, $b_i\mathbb{R} = B_i$, $(b_i + b_{i+1})\mathbb{R} = F_i$. The polynomial curve $c'(t) = A \cdot c(t)$ has control points $A \cdot b_i$, and $c'(t) = c'(t)\mathbb{R}$ has control points $(A \cdot b_i)\mathbb{R}$ and frame points $(A \cdot (b_i + b_{i+1}))\mathbb{R}$. \square

Lemma 1.4.12 is the reason why we call the sequence $B_0, F_0, B_1, F_1, \dots, B_n$ the *geometric control polygon* of the curve $c(t)$.

Projective Variation Diminishing Property

The geometric control polygon of a rational Bézier curve allows us to find an upper bound for the number of intersections of a line with a rational Bézier curve in the same manner as for polynomial Bézier curves. We have defined the geometric control polygon as a sequence of points, three of which are collinear. If we mark two points of a line in projective space, this determines two connected components of the line, and a third point selects one of them. Thus the point F_i determines a projective line segment beginning in B_i , ending in B_{i+1} , and containing F_i . If we speak of

the intersection of a line with the geometric control polygon, we mean intersection with the union of the line segments determined by the sequence B_0, F_0, \dots, B_n .

Corollary 1.4.13. *A hyperplane H intersects a rational Bézier curve $c(t)$ (not contained in H) in at most as many points as it intersects the geometric control polygon of this curve, if no vertex has zero weight.*

Proof. Suppose that $c(t) = \mathbf{c}(t)\mathbb{R}$ with $\mathbf{c}(t) = \sum B_i^n(t)\mathbf{b}_i$, $\mathbf{b}_i \in \mathbb{R}^{m+1}$. The set of homogeneous coordinate vectors of the points of H is an m -dimensional linear subspace \bar{H} of \mathbb{R}^{m+1} . Each point of the polynomial Bézier curve $\mathbf{c}(t)$ which is in \bar{H} corresponds to a point of c which is in H , and vice versa. So the result follows from Th. 1.4.8. \square

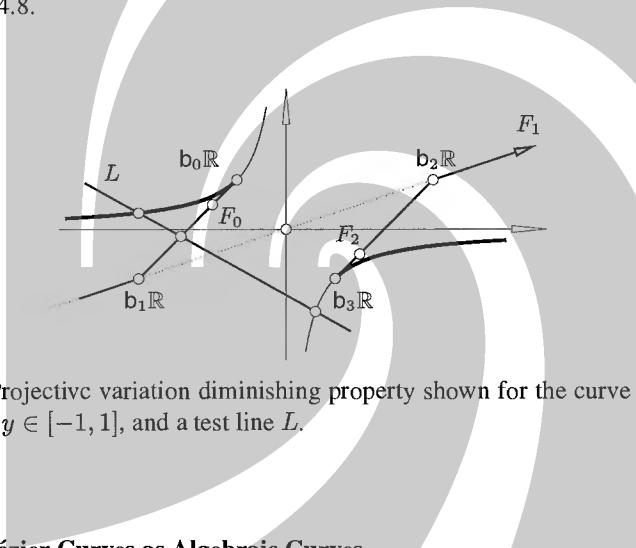


Fig. 1.49. Projective variation diminishing property shown for the curve $x = 1/y$, the interval $y \in [-1, 1]$, and a test line L .

Rational Bézier Curves as Algebraic Curves

The smallest projective algebraic variety which contains a rational Bézier curve $c(t) = \mathbf{c}(t)\mathbb{R}$ in P^m of degree n is either a point (if the curve is completely degenerate and constant) or an algebraic curve. We want to estimate the degree of this curve: We use the geometric definition of ‘degree’ and intersect the curve with a test hyperplane. This amounts to intersect the curve $c(t)$ in \mathbb{R}^{m+1} with a linear subspace of dimension m . If the curve is not contained in the hyperplane, then this gives a polynomial equation which has ‘in general’ n solutions. If $c(t)$ passes through the origin of \mathbb{R}^{m+1} k times, then these k solutions do not belong to projective points. We can therefore expect the curve to be of degree $n - k$.

Example 1.4.5. We consider the rational Bézier curve with control points $B_0 = (-1, 0) = (1, -1, 0)\mathbb{R}$, $B_1 = (0, 0, 1)\mathbb{R}$ (an ideal point), $B_2 = (1, 0) = (1, 1, 0)\mathbb{R}$; and frame points $F_0 = (-1, 1)$ and $F_1 = (1, 1)$. This shows that the control vectors \mathbf{b}_i are (up to a scalar factor) given by $\mathbf{b}_0 = (1, -1, 0)$, $\mathbf{b}_1 = (0, 0, 1)$, and $\mathbf{b}_2 = (1, 1, 0)$.

With the Bernstein polynomials $B_0^2(t) = (1-t)^2$, $B_1^2(t) = 2t(1-t)$, and $B_2^2(t) = t^2$, we can compute the curve $c(t) = (\sum_i B_i^2(t)\mathbf{b}_i)\mathbb{R}$, and get $c(t) = (1-2t+2t^2, -1+2t, 2t(1-t))\mathbb{R}$. Implicitization shows that the smallest projective algebraic variety which contains this curve is the unit circle $x_1^2 + x_2^2 = x_0^2$. \diamond

Polynomial B-spline Curves

It would lead too far to give here an overview of the theory of B-splines. We confine ourselves to a geometric definition, and to evaluation and decomposition of B-spline curves into Bézier curves. For alternative definitions (which are better suited to show other properties) and the approximation theory of B-splines, see [33, 176].

Consider an integer (the *degree*) n , and an ordered list T of real numbers $t_0 \leq t_1 \leq \dots$ with $t_i < t_{i+n+1}$ (the *knot vector*). We define a function $N_i^n(u)$ by the recursion

$$\begin{aligned} N_i^0(u) &:= \begin{cases} 1 & \text{for } t_i \leq u < t_{i+1} \\ 0 & \text{else} \end{cases} \\ N_i^r(u) &:= \frac{u - t_i}{t_{i+r} - t_i} N_i^{r-1}(u) + \frac{t_{i+r+1} - u}{t_{i+r+1} - t_{i+1}} N_{i+1}^{r-1}(u) \end{aligned} \quad (1.95)$$

for $1 \leq r \leq n$, which is called the *i-th B-spline basis function* of degree n corresponding to the *knot vector* T (as knot vectors are hardly ever added and multiplied by scalar factors, the term ‘vector’ is misleading, but it is usually called so). We also define functions $P_{i,j}^n(u_1, \dots, u_n)$ by the following recursion ($1 \leq r \leq n$):

$$\begin{aligned} P_{i,j}^0(\) &= \delta_{i,j} \\ P_{i,j}^r(u_1, \dots, u_r) &= \frac{u_r - t_i}{t_{i+r} - t_i} P_{i,j}^{r-1}(u_1, \dots, u_{r-1}) \\ &+ \frac{t_{i+r+1} - u_r}{t_{i+r+1} - t_{i+1}} P_{i+1,j}^{r-1}(u_1, \dots, u_{r-1}) \end{aligned} \quad (1.96)$$

Definition. If n is a positive integer, $\mathbf{d}_0, \dots, \mathbf{d}_m$ are points of \mathbb{R}^d , $T = (t_0 \leq \dots \leq t_{m+n+1})$ is a knot vector, then the B-spline curve $s(u)$ of degree n with control points $\mathbf{d}_0, \dots, \mathbf{d}_m$ and knot vector T , and its associated polar forms S_j are defined by

$$s(u) = \sum_{i=0}^m N_i^n(u) \mathbf{d}_i, \quad S_j(u_1, \dots, u_n) = \sum_{i=0}^m P_{i,j}^n(u_1, \dots, u_n) \mathbf{d}_i. \quad (1.97)$$

Remark 1.4.2. In this book we never use the equation ‘order = degree + 1’ which is familiar in spline theory. ‘Order’ is always a synonym for ‘degree’. \diamond

Fig. 1.50 shows an example of a B-spline curve of degree three over the knot vector $T = (0, 0, 0, 0, 2, 4, 6, 8, 8, 8, 8)$. Properties of B-spline curves are summarized in the following

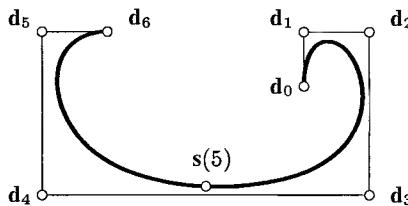


Fig. 1.50. Cubic B-spline curve over the knot vector $(0, 0, 0, 0, 2, 4, 6, 8, 8, 8, 8)$.

Theorem 1.4.14. Fix a degree n , an admissible knot vector $T = (t_0 \leq \dots \leq t_{n+m+1})$, and control points $\mathbf{d}_0, \dots, \mathbf{d}_m$. Then the B-spline basis functions $N_i^n(u)$ and the B-spline curve $\mathbf{s}(u)$ defined by these data have the following properties:

1. $N_i^n(u) \geq 0$ and $\sum_{i=1}^m N_i^n(u) = 1$. $N_i^n(u) = 0$ if $u \notin [t_i, t_{i+n+1}]$.
2. The curve point $\mathbf{s}(u)$ only depends on $\mathbf{d}_{i-n}, \dots, \mathbf{d}_i$, if $u \in [t_i, t_{i+1}]$. This dependency is affinely invariant.
3. The curve point $\mathbf{s}(u)$ is contained in the convex hull of $\mathbf{d}_{i-n}, \dots, \mathbf{d}_i$, if $u \in [t_i, t_{i+1}]$.
4. The restriction of the B-spline basis functions N_i^n to any interval $[t_j, t_{j+1}]$ is polynomial with multi-affine polar form $P_{i,j}^n(u_1, \dots, u_n)$.
5. The restriction of the curve \mathbf{s} to any interval $[t_j, t_{j+1}]$ is polynomial with multi-affine polar form $S_j(u_1, \dots, u_n)$.
6. If $t_j < t_{j+1}$ and $j - n \leq l \leq j$, then $P_{i,j}^n(t_{l+1}, \dots, t_{l+n}) = \delta_{i,l}$.
7. Under the same assumptions, $\mathbf{d}_l = S_j(t_{l+1}, \dots, t_{l+n})$ (note that all polar forms S_l, \dots, S_{l+n} have the same value, namely \mathbf{d}_l).
8. The curve $\mathbf{s}(u)$ is $n - \mu$ times differentiable at a μ -fold knot t_j .
9. The point $S_j(u_1, \dots, u_n)$ is contained in the osculating $(n - 1)$ -spaces of the curve \mathbf{s} at parameter values u_1, \dots, u_n . If any of the u_i occurs μ times, this point is contained in the osculating $n - \mu$ -space of the curve $\mathbf{s}(u)$ at u_i .
10. The B-spline curve intersects a hyperplane in at most as many points as its control polygon.

Proof. Part of the proof consists of lengthy calculations, which are left to the reader.

1. This is easily proved by induction directly from the definition, 2. and 3. follow directly.
4. The polynomiality is clear from the definition. To show that $P_{i,j}^n$ actually equals the polar form, we use Lemma 1.4.4. We have to show that $P_{i,j}^n$ is multi-affine (which is shown by induction), that $P_{i,j}^n(u, \dots, u) = N_i^n(u)$ (which is clear from the definition) and that $P_{i,j}^n(u_1, \dots, u_n)$ does not depend on the order of the arguments u_1, \dots, u_n (shown by induction).
5. This can be shown component-wise, which is nothing but 4.
6. Follows directly from the definition by induction.
7. By 6., $S_j(t_{l+1}, \dots, t_{l+n}) = \sum P_{i,j}^n(t_{l+1}, \dots, t_{l+r}) \mathbf{d}_i = \sum \delta_{i,l} \mathbf{d}_i$.

8. We assume that $t_{j+1} = \dots = t_{j+\mu}$ is a μ -fold knot. It is sufficient to show the statement for the basis functions N_i^n . We first compute $P_{i,j}^\mu(t_{j+1}, \dots, t_{j+1}) = P_{i,j}^\mu(t_{j+1}, \dots, t_{j+\mu}) = \delta_{ij} = P_{i,j+\mu}^\mu(t_{j+1}, \dots, t_{j+\mu}) = P_{i,j+\mu}^\mu(t_{j+1}, \dots, t_{j+1})$.

Thus the polar forms $P_{i,j}^n$ and $P_{i,j+\mu}^n$ of $N_i^n|[t_j, t_{j+1}]$ and $N_i^n|[t_{j+\mu}, t_{j+\mu+1}]$ coincide whenever the argument t_{j+1} occurs μ times. By Lemma 1.4.5, this means that the first $n - \mu$ derivatives of these polynomials are equal.

9. Th. 1.4.10 and Cor. 1.4.6 show that the osculating $(n - \mu)$ -space of a polynomial curve $s(u)$ at $u = a$ is spanned by the points $S(\underbrace{a, \dots, a}_{\geq \mu}, b, \dots, b)$, where $a \neq b$.

This shows that any point of the form $S(\underbrace{a, \dots, a}_{\geq \mu}, u_{\mu+1}, \dots, u_n)$ is contained in this space.

10. The proof is the same as for Bézier curves.

□

This theorem shows (i) how to evaluate a B-spline curve geometrically, i.e., how to compute $S_j(u, \dots, u)$ from the control points, (ii) how to decompose a B-spline curve into Bézier curve segments, i.e., how to compute the control points $S_j(t_j, \dots, t_j, t_{j+1}, \dots, t_{j+1})$ of the j -th segment, (iii) how the curve behaves when control points are changed. A B-spline curve exhibits the convex hull property, the variation diminishing property, and *local* control. This is illustrated in Fig. 1.51.

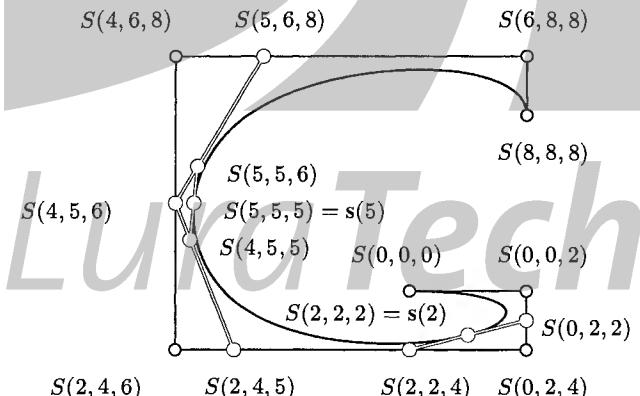


Fig. 1.51. Polar form of a cubic B-spline curve over the knot vector $(0, 0, 0, 0, 2, 4, 6, 8, 8, 8, 8)$.

It is important to note that *all* piecewise polynomial curves are B-spline curves over some knot vector. This is shown by

Theorem 1.4.15. If $p(u)$ is a piecewise polynomial curve of degree $\leq n$ which is only $n - \mu_i$ times differentiable at $u = u_i$, then there is a knot vector T such that

$$\begin{aligned} T &= (\underbrace{u_1, \dots, u_1}_{\mu_1}, \underbrace{u_2, \dots, u_2}_{\mu_2}, \dots) \text{ with } u_1 < u_2 < \dots, \\ p(u) &= \sum_{i=0}^m N_i^n(u) \mathbf{d}_i, \quad \text{with} \\ \mathbf{d}_l &= S_j(t_{l+1}, \dots, t_{l+n}), \quad (t_j < t_{j+1}, j-n \leq l \leq j), \end{aligned}$$

where S_j is the polar form of the polynomial $p|[t_j, t_{j+1}]$.

Proof. We first have to show that \mathbf{d}_l is well-defined, because according to the theorem, it may be found by evaluating the different polar forms S_l, \dots, S_{l+n} . But Lemma 1.4.5 shows that all these polar forms evaluate to the same point.

By Th. 1.4.14, the so defined B-spline curve coincides with $p(u)$ in the interval $p|[t_j, t_{j+1}]$. \square

Rational B-spline Curves (NURBS Curves)

A polynomial B-spline curve $\mathbf{s}(u)$ of degree n , defined by a knot vector T and control points $\mathbf{d}_i \in \mathbb{R}^{d+1}$, gives rise to a rational B-spline curve $s(u)$ in P^d via

$$s(u) = \mathbf{s}(u)\mathbb{R} = \left(\sum N_i^n(u) \mathbf{d}_i \right) \mathbb{R}. \quad (1.98)$$

If no \mathbf{d}_i is the zero vector, we can encode the vectors \mathbf{d}_i (up to an irrelevant common scalar factor) by the sequence of points

$$\mathbf{d}_0\mathbb{R}, \mathbf{f}_0\mathbb{R} = (\mathbf{d}_0 + \mathbf{d}_1)\mathbb{R}, \mathbf{d}_1\mathbb{R}, \dots, \mathbf{d}_m\mathbb{R},$$

and get control points $\mathbf{d}_i\mathbb{R}$ and frame points $\mathbf{f}_i\mathbb{R}$, just like in the case of Bézier curves. We call the polygon with vertices $\mathbf{d}_i\mathbb{R}$ and edges containing $\mathbf{d}_i\mathbb{R}, \mathbf{f}_i\mathbb{R}, \mathbf{d}_{i+1}\mathbb{R}$ the *geometric control polygon* of the curve.

The name NURBS for such curves stands for ‘non-uniform rational B-spline’, where non-uniformity means that the knots need not necessarily be evenly distributed on the real number line.

To show some properties of NURBS curves, we consider properties which are enjoyed by polynomial Bézier and B-spline curves alike, and then try to generalize them to the rational case, thereby imitating the case of Bézier curves.

The following theorem gives a summary:

Theorem 1.4.16. A NURBS curve $s(u)$ of degree n in d -dimensional projective space, defined by the knot vector $T = (t_0 \leq \dots \leq t_{n+m+1})$ and its geometric control polygon $\mathbf{d}_0\mathbb{R}, \mathbf{f}_0\mathbb{R}, \mathbf{d}_1\mathbb{R}, \dots, \mathbf{d}_m\mathbb{R}$, have the following properties:

1. The restriction of $s(u)$ to a nonempty interval $[t_j, t_{j+1}]$ is a rational Bézier curve.

2. A hyperplane H intersects $s(u)$ (not contained in H) in at most as many points as it intersects the geometric control polygon. This is the projective variation diminishing property.
3. The control point $d_l\mathbb{R}$ is contained in the $(n - 1)$ -dimensional osculating subspace of $s(u)$ at the parameter values t_{l+1}, \dots, t_{l+n} .
4. Any piecewise rational curve $s(u)$ of degree $\leq n$ is a NURBS curve of degree n .

Proof. The proofs are the same as for rational Bézier curves or follow more or less directly from the respective properties of polynomial B-spline curves. \square

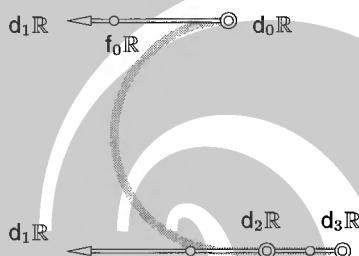


Fig. 1.52. NURBS curve of Ex. 1.4.6.

Example 1.4.6. Consider the half circle $x^2 + y^2 = 1, x \leq 0$ and the line segment which joins $(0, -1)$ with $(1, -1)$. We want to parametrize this curve as a NURBS curve (cf. Fig. 1.52).

First we let $c(u) = c(u)\mathbb{R} = (1+u^2, u^2-1, -2u)\mathbb{R}$. Then the arc is $c([-1, 1])$. We see that $\dot{c}(1) = (2, 2, -2)$. If we parametrize the line $y = -1$ by $d(u) = (1, u-1, -1)\mathbb{R} = 2u \cdot (1, u-1, -1)\mathbb{R} = d(u)\mathbb{R}$, we see that $d(1) = (2, 2, -2)$, so that the curve $c([-1, 1] \cup d([1, 2]))$ is a piecewise polynomial curve of degree 2. It is discontinuous at $u = -1$ and $u = 2$, and C^1 at $u = 1$. A suitable knot vector is therefore $(-1, -1, -1, 1, 2, 2, 2)$.

The polar form S_3 of c is given by $S_3(u_1, u_2) = (u_1 u_2 + 1, u_1 u_2 - 1, -u_1 - u_2)$, and the polar form S_4 of d is given by $S_4(u_1, u_2) = (u_1 + u_2, 2u_1 u_2 - u_1 - u_2, -u_1 - u_2)$. The control points $d_0\mathbb{R}, \dots, d_3\mathbb{R}$ are found by evaluating $d_0 = S_3(-1, -1) = (2, 0, 2)$, $d_1 = S_3(-1, 1) = (0, -2, 0) = S_4(-1, 1)$, $d_2 = S_3(1, 2) = (3, 1, -3) = S_4(1, 2)$, $d_3 = S_4(2, 2) = (4, 4, -4)$.

Thus the geometric control polygon $d_0\mathbb{R}, (d_0 + d_1)\mathbb{R}, d_1\mathbb{R}, \dots, d_3\mathbb{R}$ of this NURBS curve consists of the line segments defined by the points $(0, 1), (-1, 1), (0, -1, 0)\mathbb{R}$ (at infinity), $(-1/3, -1), (1/3, -1), (5/7, -1), (1, -1)$. \diamond

1.4.2 Dual Bézier Curves

The Bézier representation of a rational curve expresses the polynomial homogeneous parametrization $\mathbf{c}(t)\mathbb{R}$ in terms of the Bernstein polynomials. Then the coefficients have the remarkable geometric meaning of control points with a variety of important and practically useful properties which have been discussed above.

The *tangent* of a planar rational curve $c(t) = \mathbf{c}(t)\mathbb{R}$ at $t = t_0$ is computed as the line which connects $c(t_0)$ with its first derivative point $c^1(t_0)$ and has the homogeneous coordinate vectors $\mathbb{R}\mathbf{u}(t_0) = \mathbb{R}(\mathbf{c}(t_0) \times \dot{\mathbf{c}}(t_0))$. Thus the family of tangents has again a polynomial parametrization, which can be expressed in the Bernstein basis. This leads to a *dual Bézier curve*

$$U(t) = \mathbb{R}\mathbf{u}(t) = \mathbb{R} \sum_{i=0}^m B_i^m(t) \mathbf{u}_i, \quad (1.99)$$

which can be seen as a family of lines in the (ordinary) projective plane, or as a family of points of its dual plane.

Theorem 1.4.17. *The family of tangents of a rational Bézier curve is a dual Bézier curve, and vice versa.*

Proof. We have seen above that the family of tangents of a rational Bézier curve has a polynomial representation in homogeneous coordinates. This proves the first statement of the lemma. The second follows from applying duality twice (cf. Sec. 1.2.3). \square

Remark 1.4.3. All *projective* properties of rational Bézier curves are shared by its dual, because it is again a rational Bézier curve (of the dual plane). We thus may convert statements about rational families of points (i.e., curves) into statements about the families of their tangents. \diamond

Remark 1.4.4. When speaking of a Bézier curve we often mean a *curve segment*. In the form we have written the Bernstein polynomials, the curve segment is parametrized over the interval $[0, 1]$. For any $t \in [0, 1]$, Equ. (1.99) yields a line $U(t) = \mathbb{R}\mathbf{u}(t)$. The original curve segment is the envelope of the lines $U(t)$, where t ranges in $[0, 1]$. \diamond

Dual Control Structure

As an example of Remark 1.4.3, let us discuss the dual control structure of a Bézier curve c (see Fig. 1.53). It is defined as the control structure of c 's dual, which is a Bézier curve of the dual plane. Its control *points* are therefore *lines*, when viewed from the original projective plane.

There are the *Bézier lines* $U_i = \mathbb{R}\mathbf{u}_i$, $i = 0, \dots, m$, and the *frame lines* F_i , whose line coordinate vectors are given by

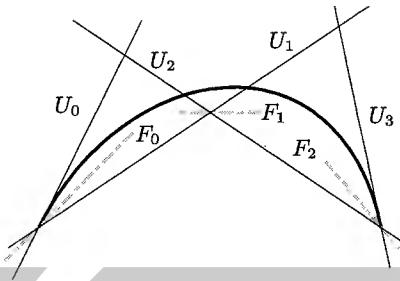


Fig. 1.53. Dual Bézier curve with dual control structure.

$$f_i = u_i + u_{i+1}, \quad i = 0, \dots, m-1. \quad (1.100)$$

From (1.100) we see that the frame line F_i is concurrent with the Bézier lines U_i and U_{i+1} . This is dual to the collinearity of a frame point with its two adjacent Bézier points.

We could also use weights instead of frame lines, just as we could have used weights instead of frame points. Because weights are no projective invariants, we prefer frame lines and frame points. An invariant statement of theorems is also important for their dualization.

The complete geometric input, consisting of the $m+1$ Bézier lines and m frame lines, defines a dual Bézier curve as follows: Each line has a one-dimensional subspace of homogeneous coordinate vectors with respect to a given coordinate system. It is possible to choose the Bézier line coordinate vectors u_i such that Equ. (1.100) holds. This choice is unique up to a common factor of the vectors u_0, \dots, u_m , which does not change the curve. Now (1.99) uniquely defines the corresponding curve segment.

For a Bézier curve, the control points b_0 and b_m are the end *points* of the curve segment and the line $b_0 \vee b_1$ and $b_{m-1} \vee b_m$ are the *tangents* there. Dual to this, the end *tangents* of a dual Bézier curve are U_0 and U_m , and their *points of contact* are given by $U_0 \cap U_1$ and $U_{m-1} \cap U_m$, respectively.

The control points b_i, b_{i+1} divide the range of points $b_i \vee b_{i+1}$ into two subsets. The one which contains the frame point is the edge of the *control polygon*. We dualize this construction: The line pencil spanned by lines U_i and U_{i+1} is divided into two subsets, bounded by U_i and U_{i+1} . The one which contains the frame line is part of the *complete dual control structure* (see Fig. 1.54).

Dual Variation Diminishing Property and Convexity

Dual to the variation diminishing property of a rational Bézier curve with respect to its projective control polygon we can state the following result.

Theorem 1.4.18. *If c is a planar rational Bézier curve, the number of c 's tangents incident with a given point P does not exceed the number of lines of the complete dual control structure which are incident with P (if no control line has zero weight).*

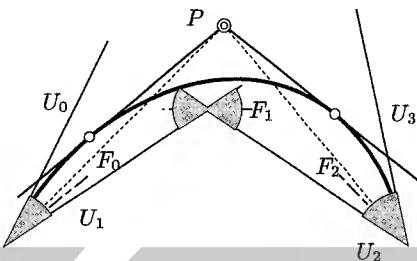


Fig. 1.54. Complete dual control structure and variation diminishing property.

Proof. This is the dual version of Cor. 1.4.13. \square

This result easily implies the following sufficient condition for *convexity* of a dual Bézier curve (see Fig. 1.54).

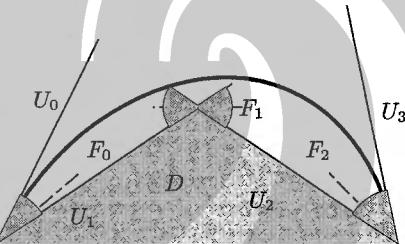


Fig. 1.55. Convex dual control polygon.

By a *convex* curve we understand part of the boundary of a convex domain. A support line L of a convex domain D is a line through a point of the boundary of D such that D lies entirely on one side of L .

Corollary 1.4.19. *If the Bézier lines U_i and the frame lines F_i of a dual Bézier curve c are among the edges and support lines of a convex domain D , and the points $U_i \cap U_{i+1}$ are among D 's vertices, then c is convex and lies completely outside D (cf. Fig. 1.55).*

Proof. The proof is left to the reader as an exercise. The main ingredient is the shape of the dual control polygon. It ensures that there are at most two curve tangents incident with any point of the plane, and the outside of a smooth convex domain is precisely the set of points where two such tangents exist. \square

Remark 1.4.5. A planar rational curve segment, or more precisely, a rational parametrization of it, possesses two Bézier representations: the usual, point-based

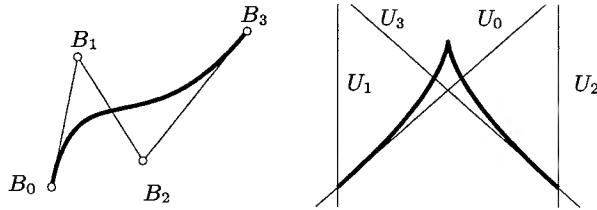


Fig. 1.56. Standard control structure tends to generate inflections, whereas the dual control structure tends to introduce cusps.

form, and the dual line-based representation, which was introduced by J. Hoschek [76]. However, their behaviour when used for design purposes is different. By using the standard representation it is difficult to design cusps but quite easy to achieve inflections of the curve segment. In the dual representation, very special conditions on the control structure must be met to design an inflection, but it is easy to get cusps. This is illustrated by Fig. 1.56. For many applications, cusps are not desirable and therefore Cor. 1.4.19 plays an important role. To achieve inflections, it is best to locate them at end points of the curve segments (see [145]). \diamond

Evaluation and Conversion of Dual Curves

To evaluate the polynomial (1.99), we can use the algorithm of de Casteljau in the same way as for ‘ordinary’ polynomial Bézier curves:

1. we let $u_i^0 = u_i$, ($i = 0, \dots, m$),
2. then recursively define lines $U_i^k(t) = \mathbb{R}u_i^k(t)$ with

$$u_i^k(t) = (1-t)u_i^{k-1}(t) + tu_{i+1}^{k-1}(t), \quad (1.101)$$

3. and get $U(t) = \mathbb{R}u_0^m(t)$.

This algorithm yields the line $U(t)$ for all real parameter values t . For many purposes however we want to compute the ‘point curve’ c which corresponds to the family of lines $U(t)$. We know that iterated duality of planar curves gives the original curve again (cf. the computation at p. 83). Thus we have to compute a tangent to the dual curve at $U(t)$ in order to get the point $c(t)$.

The dual version of Cor. 1.4.7 shows that the point is the intersection of the lines U_0^{m-1} and U_1^{m-1} , and

$$c(t) = u_0^{m-1}(t) \times u_1^{m-1}(t). \quad (1.102)$$

Note that computation of the point $c(t)$ via the dual algorithm of de Casteljau and the previous formula is as efficient as applying the algorithm of de Casteljau to $c(t)$ itself. We summarize these computations in the following

Corollary 1.4.20. *If $U(t) = \mathbb{R} \sum_{i=0}^m B_i^m(t) u_i$ is a dual Bézier curve, then the corresponding point curve $c(t)$ is parametrized by*

$$\begin{aligned}
c(t) &= \mathbf{c}(t)\mathbb{R} = \left(\sum_{k=0}^{2m-2} B_k^{2m-2}(t) \mathbf{b}_k \right) \mathbb{R}, \\
\mathbf{b}_k &= \binom{2m-2}{k}^{-1} \sum_{i+j=k} \binom{m-1}{i} \binom{m-1}{j} (\mathbf{u}_i \times \mathbf{u}_{j+1}).
\end{aligned} \tag{1.103}$$

Proof. We use the notation of the dual algorithm of de Casteljau (see Equ. (1.101)). We have to compute $\mathbf{c}(t) = \mathbf{u}_0^{m-1}(t) \times \mathbf{u}_1^{m-1}(t)$. Lemma 1.4.2 shows that

$$\mathbf{u}_0^{m-1}(t) = \sum_{i=0}^{m-1} B_i^{m-1}(t) \mathbf{u}_i, \text{ and } \mathbf{u}_1^{m-1}(t) = \sum_{j=0}^{m-1} B_j^{m-1}(t) \mathbf{u}_{j+1},$$

and Equ. (1.103) follows by expanding the cross product. \square

We see that $c(t)$ possesses the degree $2m - 2$ as a Bézier curve.

Degree and Class of Rational Bézier Curves

If (1.99) is a Lüroth parametrization of a dual Bézier curve of degree m , the corresponding point curve c is a rational curve of *class* m (cf. Prop. 1.3.22). In general, it has the degree $n = 2m - 2$ as an algebraic variety.

This degree decreases if inflection points and certain higher order singularities of the dual curve occur — they are characterized by linear dependence of $\mathbf{u}_0^{m-1}(t_0)$ and $\mathbf{u}_1^{m-1}(t_0)$. This means that $\mathbf{c}(t_0) = (0, 0, 0)$ and $\mathbf{c}(t) = (t - t_0)\bar{\mathbf{c}}(t)$, where $\deg(\bar{\mathbf{c}}) = \deg(\mathbf{c}) - 1$.

By duality (and interchanging line and point coordinates), (1.103) also describes the change from the standard point representation to the dual form. Thus we can conclude that a Bézier curve with a Lüroth parametrization of polynomial degree n , has degree n as an algebraic curve, and class less or equal $2n - 2$.

Note that for a rational curve c , whose class m is less than its degree n , the dual form is preferable from the viewpoint of computational efficiency.

Dual B-spline Curves

The generalization of the previous paragraphs to *dual NURBS curves* is mostly straightforward. Clearly the dual of a planar piecewise rational curve is rational, which means that the dual of a NURBS curve is a NURBS curve of the dual projective plane.

However the formula which expresses the control points of the dual curve is more complicated. This is because for a piecewise rational curve $\mathbf{c}(u)\mathbb{R}$ of degree $\leq n$ which is $C^{n-\mu}$ at $u = u_0$ the dual curve $\mathbf{c}(u) \times \dot{\mathbf{c}}(u)\mathbb{R}$ is of degree $2n - 2$, but only $n - \mu - 1$ times differentiable. We therefore have to choose a different knot vector for the dual curve.

One solution of this problem is based on the product formulae for B-spline basis functions (Mørken [126]). As an alternative, one may use the piecewise Bézier representation in an intermediate step.

Remark 1.4.6. Assume that a NURBS parametrization $\mathbf{c}(u)\mathbb{R}$ of a planar curve $c(u)$ has degree n and is $n - \mu$ times differentiable at $u = u_0$. Therefore the knot vector contains u_0 with multiplicity μ . Its dual, constructed with an analogue of formula (1.103), has degree $2n - 2$ and is $n - \mu - 1$ times differentiable at $u = u_0$. Lemma 1.2.6 however implies that after an appropriate regular parameter transform the curve is actually $n - \mu$ times differentiable, if some regularity conditions are met. \diamond

1.4.3 Rational Bézier Surfaces

We will discuss two different types of Bézier *surfaces*. One (tensor product surfaces) is a very simple generalization of the notion of Bézier curve, but not all polynomial surface parametrizations are tensor product parametrizations. The other type (triangular surfaces) is a bit more complicated, but this class of surfaces indeed coincides with the class of rational surfaces.

Tensor Product Surfaces

Consider the polynomial 2-surface

$$s(u, v) = \sum_{i=0}^n \sum_{j=0}^m B_i^n(u) B_j^m(v) \mathbf{b}_{ij}, \quad (1.104)$$

which is a linear combination of control points \mathbf{b}_{ij} with Bernstein polynomials. Such a surface is called tensor product (TP) surface of degree (n, m) . The lines $u = \text{const}$ and the lines $v = \text{const}$ are Bézier curves of degree m and n , respectively.

The obvious generalization of a TP 2-surface to d variables is given by

$$s(u_1, \dots, u_d) = \sum_{i_1=0}^{m_1} \cdots \sum_{i_d=0}^{m_d} B_{i_1}^{m_1}(u_1) \cdots B_{i_d}^{m_d}(u_d) \mathbf{b}_{i_1 \dots i_d}. \quad (1.105)$$

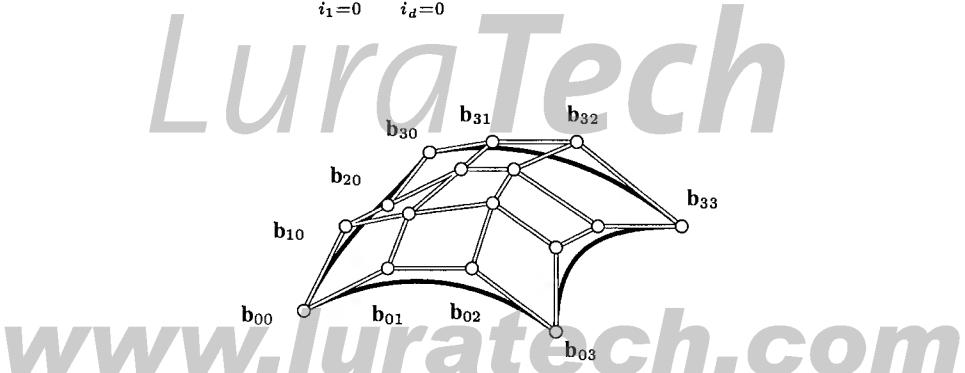


Fig. 1.57. Tensor product Bézier surface of degree $(3, 3)$.

Fig. 1.57 shows an example of a tensor product Bézier surface. Such surfaces have the property that parameter curves are Bézier curves, and exhibit some of their properties, e.g., the convex hull property, which is shown in exactly the same way as for curves. There is no easy generalization of the variation diminishing property, however.

The construction of ‘tensor product’ surfaces is not restricted to Bézier surfaces. We may replace the Bernstein polynomials in Equ. (1.104) by other basis functions: The *tensor product B-spline surfaces* have the form

$$s(u, v) = \sum_{i=0}^{m_1} \sum_{j=0}^{m_2} N_i^{n_1}(u) N_j^{n_2}(v) \mathbf{d}_{ij},$$

where $N_i^{n_1}(t)$ are the B-spline basis functions defined by a certain knot vector t_0, \dots , and $N_j^{n_2}(t)$ are B-spline basis functions defined over a different knot vector.

Ruled Surfaces as TP Surfaces

A TP Bézier surface of degree $(n, 1)$ is of the form

$$s(u, v) = (1 - v) \sum_{i=0}^n B_i^n(u) \mathbf{b}_{i0} + v \sum_{i=0}^n B_i^n(u) \mathbf{b}_{i1}, \quad (1.106)$$

because the Bernstein polynomials of first order are given by $B_0^1(t) = (1 - t)$, and $B_1^1(t) = t$. The lines $u = \text{const}$ are straight lines, which shows that s is a *ruled surface*. Ruled surfaces will be discussed in detail in Chap. 5.

Example 1.4.7. A TP surface of degree $(1, 1)$ has the representation

$$s(u, v) = (1 - u)(1 - v) \mathbf{b}_{00} + u(1 - v) \mathbf{b}_{10} + (1 - u)v \mathbf{b}_{01} + uv \mathbf{b}_{11}.$$

By Prop. 1.1.41, this surface is a hyperbolic paraboloid (see Fig. 1.23). \diamond

Barycentric Coordinates

We consider the points 0 and 1 of the real number line as *base points* $s_0 = 0, s_1 = 1$, and write all real numbers in the form

$$t = t_0 s_0 + t_1 s_1 \quad \text{with } t_0 + t_1 = 1. \quad (1.107)$$

This is uniquely done by letting $t_0 = (1 - t), t_1 = t$. These ‘coordinates’ of t with respect to the base points s_0, s_1 are called *barycentric coordinates*. The Bernstein polynomials attain a more symmetric form when written in barycentric coordinates: Instead of $B_i^n(t)$ we write $B_{i_0 i_1}(t_0, t_1)$, where $i = i_0$, and $n - i = i_1$:

$$B_i^n(t) = \frac{n!}{i!(n-i)!} (1 - t)^i t^{n-i} = B_{i_0 i_1}(t_0, t_1) = \frac{n!}{i_0! i_1!} t_0^{i_0} t_1^{i_1}. \quad (1.108)$$

The definition is not restricted to $s_0 = 0$ and $s_1 = 1$; we can choose any two different real numbers as base points. This new complication leads to the ‘right’ generalization of Bézier *curves* to Bézier *surfaces* which are capable of representing all polynomial or rational d -surfaces:

Choose a base simplex $\mathbf{s}_0, \dots, \mathbf{s}_d$ (base triangle $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2$ in the case of a bivariate surface) in \mathbb{R}^d . We can write all points $\mathbf{u} = (u_1, \dots, u_d)$ of \mathbb{R}^d in the form

$$\mathbf{u} = \sum t_i \mathbf{s}_i \quad \text{with } \sum t_i = 1. \quad (1.109)$$

The t_i are called *barycentric coordinates* of \mathbf{u} . Then, a Bernstein polynomial of degree n is any of the following polynomials given by

$$B_{i_0 \dots i_d}(\mathbf{u}) = \frac{n!}{i_0! \dots i_d!} t_0^{i_0} \dots t_d^{i_d}, \quad (1.110)$$

such that $i_0 + \dots + i_d = n$, $i_0 \geq 0, \dots, i_d \geq 0$. The restriction that all indices have to be nonnegative will not be mentioned explicitly in the sequel.

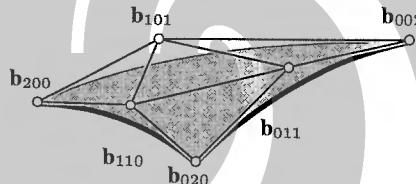


Fig. 1.58. Triangular Bézier surface of degree two.

Triangular Bézier Surfaces

A *triangular Bézier d-surface* of degree n is a surface of the form

$$s(\mathbf{u}) = \sum_{i_0 + \dots + i_d = n} B_{i_0 \dots i_d}(\mathbf{u}) \mathbf{b}_{i_0 \dots i_d}. \quad (1.111)$$

The points $\mathbf{b}_{i_0 \dots i_d}$ are called *control points*.

Such a triangular Bézier d -surface has the following properties: The sum of all coefficient functions $B_{i_0 \dots i_d}(\mathbf{u})$ of degree n equals 1. This means that if the control points undergo an affine transformation, the entire surface transforms accordingly. If \mathbf{u} is in the interior of the base simplex $\mathbf{s}_0, \dots, \mathbf{s}_d$, then the coefficient functions assume values between 0 and 1. This shows that $s(\mathbf{u})$ lies in the convex hull of the control points.

The restriction of the surface to an affine r -dimensional subspace is a triangular Bézier r -surface. Especially the restriction to a line in its domain is a Bézier curve.

For arguments $\mathbf{u} = \mathbf{s}_0, \mathbf{u} = \mathbf{s}_1, \dots, \mathbf{u} = \mathbf{s}_d$ the surface s assumes the values $\mathbf{b}_{n,0,\dots,0}, \mathbf{b}_{0,n,0,\dots,0}, \dots, \mathbf{b}_{0,\dots,0,n}$.

The Multi-affine Polar Form

A triangular Bézier d -surface is polynomial. Like in the curve case it is possible to show that in fact all polynomial surfaces can be written as triangular Bézier surfaces. Suppose a polynomial d -surface of degree n in \mathbb{R}^m is given as a linear combination

$$\mathbf{f}(x_1, \dots, x_d) = \sum \mathbf{a}_\alpha \mathbf{x}^\alpha$$

of monomials with coefficients $\mathbf{a}_\alpha \in \mathbb{R}^m$ (for the notation, see Sec. 1.3.1). Of course we can restrict ourselves to describe the procedure of conversion from monomial form to Bézier form for $m = 1$, because the general case can be done component-wise. By linearity, we can restrict ourselves to the case of one monomial \mathbf{x}^α , where $\alpha = (\alpha_1, \dots, \alpha_d)$. The degree of the Bézier representation will be called n , it must be greater or equal the total degree $|\alpha| = \alpha_1 + \dots + \alpha_d$.

We determine α_0 such that $\alpha_0 + \dots + \alpha_d = n$. Then we define the multi-index $\mathbf{i} = (i_1, \dots, i_n)$ such that

$$\mathbf{i} = (\underbrace{0, \dots, 0}_{\alpha_0}, \underbrace{1, \dots, 1}_{\alpha_1}, \dots, \underbrace{d, \dots, d}_{\alpha_d}).$$

The symbol $\bar{\mathbf{x}}$ abbreviates a $(d+1)$ -tuple of variables x_0, \dots, x_d , whereas the symbol \mathbf{x} abbreviates the d -tuple of variables (x_1, \dots, x_d) . If we use *several* $\bar{\mathbf{x}}$'s, we assign *upper* indices to them. So an n -tuple of $\bar{\mathbf{x}}$'s is written in the form $\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n$. The same notation is used for \mathbf{x} instead of $\bar{\mathbf{x}}$. The j -th component of the i -th vector $\bar{\mathbf{x}}$ then is x_j^i .

The symbol $\bar{\mathbf{x}}_{\bullet i}$ with $\mathbf{i} = (i_1, \dots, i_n)$ means $x_{i_1}^1 \cdots x_{i_n}^n$ (here $0 \leq i_j \leq d$). We consider the multilinear mapping

$$\bar{P}(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n) = \frac{1}{\#\sigma} \sum_{\sigma} \bar{\mathbf{x}}_{\bullet \sigma(\mathbf{i})}, \quad (1.112)$$

where σ ranges over all possible permutations of the multi-index \mathbf{i} , and the symbol $\#\sigma$ denotes the number of these permutations. Finally the mapping P is defined by letting

$$P(\mathbf{x}^1, \dots, \mathbf{x}^n) = \bar{P}(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n) \quad \text{with} \quad x_0^1 = \dots = x_0^n = 1.$$

It is called the *multi-affine polar form* of the monomial \mathbf{x}^α . The construction of \bar{P} and P indeed is a bit complicated, but it looks simple enough if we use this multi-index notation. The following theorem describes the relation between polar forms and triangular d -surfaces:

Theorem 1.4.21. Assume $\mathbf{f}(x_1, \dots, x_d) = \sum_\alpha \mathbf{a}_\alpha \mathbf{x}^\alpha$ is a polynomial d -surface in \mathbb{R}^m . If $n > |\alpha|$ for all α which occur in the definition of \mathbf{f} , determine the multi-affine polar form $P(\mathbf{x}^1, \dots, \mathbf{x}^n)$ off component-wise and by linear combination from the multi-affine polar forms of monomials as described above. Then \mathbf{f} is a triangular Bézier d -surface over the base simplex $\mathbf{s}_0, \dots, \mathbf{s}_d$ with control points

$$\mathbf{b}_{i_0, \dots, i_d} = P\left(\underbrace{\mathbf{s}_0, \dots, \mathbf{s}_0}_{i_0}, \dots, \underbrace{\mathbf{s}_d, \dots, \mathbf{s}_d}_{i_d}\right).$$

Example 1.4.8. Consider the 2-surface

$$(x_1, x_2) \mapsto \mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 \\ x_2 \\ x_1 x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{x}^{(1,0)} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mathbf{x}^{(0,1)} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{x}^{(1,1)},$$

which is the graph of the function $x_1 \cdot x_2$. We want to write this surface as a triangular Bézier surface of degree $n = 2$ over the base triangle $\mathbf{s}_0 = (0, 0)$, $\mathbf{s}_1 = (1, 0)$, $\mathbf{s}_2 = (0, 1)$.

This means we have to compute the polar forms P' , P'' , P''' for the monomials $\mathbf{x}^{(1,0)}$, $\mathbf{x}^{(0,1)}$, and $\mathbf{x}^{(1,1)}$. As to the first monomial, we have $\alpha_1 = 1$ and $\alpha_2 = 0$, which implies $\alpha_0 = 1$ and $\mathbf{i} = (0, 1)$.

To define the polar form \bar{P}' , we have to consider all permutations of \mathbf{i} . In this case there are only two permutations, namely $(0, 1)$ itself and $(1, 0)$. Thus

$$\bar{P}'(\bar{\mathbf{x}}^1, \bar{\mathbf{x}}^2) = \frac{1}{2}(\bar{\mathbf{x}}_{\bullet 0,1} + \bar{\mathbf{x}}_{\bullet 1,0}) = \frac{1}{2}(x_0^1 x_1^2 + x_1^1 x_0^2).$$

Letting $x_0^1 = x_0^2 = 1$ gives the multi-affine polar form

$$P'(\mathbf{x}^1, \mathbf{x}^2) = \frac{1}{2}(x_1^2 + x_1^1).$$

A similar procedure yields $\mathbf{i} = (0, 2)$ and $P''(\mathbf{x}^1, \mathbf{x}^2) = (x_2^2 + x_2^1)/2$ for the monomial $\mathbf{x}^{(0,1)}$, and further, $\mathbf{i} = (1, 2)$ and $P'''(\mathbf{x}^1, \mathbf{x}^2) = (x_1^1 x_2^2 + x_2^1 x_1^2)/2$ for the monomial $\mathbf{x}^{(1,1)}$. The polar form of \mathbf{f} then is given by the linear combination

$$P(\mathbf{x}^1, \mathbf{x}^2) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} P' + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} P'' + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} P''' = \frac{1}{2} \begin{bmatrix} x_1^1 + x_1^2 \\ x_2^1 + x_2^2 \\ x_1^1 x_2^2 + x_2^1 x_1^2 \end{bmatrix}.$$

Inserting the values $\mathbf{s}_0 = (0, 0)$, $\mathbf{s}_1 = (1, 0)$, $\mathbf{s}_2 = (0, 1)$ according to Th. 1.4.21 results in

$$\begin{aligned} \mathbf{b}_{200} &= P((0, 0), (0, 0)) = (0, 0, 0), & \mathbf{b}_{020} &= P((1, 0), (1, 0)) = (1, 0, 0), \\ \mathbf{b}_{002} &= P((0, 1), (0, 1)) = (0, 1, 0), & \mathbf{b}_{110} &= P((0, 0), (1, 0)) = (\frac{1}{2}, 0, 0), \\ \mathbf{b}_{101} &= P((0, 0), (0, 1)) = (0, \frac{1}{2}, 0), & \mathbf{b}_{011} &= P((1, 0), (0, 1)) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}). \end{aligned}$$

We see that the surface $\mathbf{f}(\mathbf{x})$ can be written in the form

$$\mathbf{f}(\mathbf{x}) = \sum_{i+j+k=2} B_{ijk}(u_0, u_1, u_2) \cdot \mathbf{b}_{ijk},$$

where u_0, u_1, u_2 are the barycentric coordinates of the point \mathbf{x} with respect to the triangle $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2$, as defined in Equ. (1.109). \diamond

Rational Surfaces

A rational Bézier surface is a surface of the form

$$f(x_1, \dots, x_d) = f(x_1, \dots, x_d)\mathbb{R}, \quad (1.113)$$

where f is an ordinary polynomial Bézier surface (a triangular one or a TP surface). A rational B-spline surface has the same form, with f being a polynomial B-spline surface. Rational tensor product B-spline surfaces are also called NURBS surfaces.

The result that all polynomial d -surfaces are triangular Bézier surfaces generalizes immediately to the rational case, so all rational surfaces have rational Bézier representations. Another fact, which is valid in the rational case as well as for polynomial surfaces, is that if f is a TP surface of degree $(1, n)$, then $f = f\mathbb{R}$ is a ruled surface unless the rulings of f are incident with the origin of the coordinate system.

Example 1.4.9. The sphere is a rational surface (any quadric is), and we want to represent part of it in Bézier form. The stereographic projection with center $(0, 0, 1)$ (cf. Ex. 1.1.37) projects the point \mathbf{p} onto the point $(x, y, 0)$, if

$$\mathbf{p} = \frac{1}{1 + x^2 + y^2} (2x, 2y, x^2 + y^2 - 1).$$

Thus $f(x_1, x_2) = f(x_1, x_2)\mathbb{R} = (1 + x_1^2 + x_2^2, 2x_1, 2x_2, x_1^2 + x_2^2 - 1)\mathbb{R}$ is a quadratic rational parametrization of the sphere without the north pole $(0, 0, 1)$. The sphere is therefore a triangular Bézier surface of degree n for all $n \geq 2$. Here we let $n = 2$. The polar form of the polynomial f can be computed by the algorithm preceding Th. 1.4.21. We need the polar forms P' , P'' of the polynomials $f(x_1, x_2) = x_1$ and $f(x_1, x_2) = x_2$, respectively, which had been computed in Ex. 1.4.8, and the polar forms P'''' , P''''' of $f(x_1, x_2) = x_1^2$, and $f(x_1, x_2) = x_2^2$, respectively. They are given by $P''''(\mathbf{x}^1, \mathbf{x}^2) = x_1^1 x_1^2$ and $P'''''(\mathbf{x}^1, \mathbf{x}^2) = x_2^1 x_2^2$. The polar form of the constant polynomial 1 is the constant function 1 again. Then the polar form of f is given by

$$P(\mathbf{x}^1, \mathbf{x}^2) = \begin{bmatrix} 1 + P'''' + P''''' \\ 2P' \\ 2P'' \\ P'''' + P''''' - 1 \end{bmatrix} = \begin{bmatrix} x_1^1 x_1^2 + x_2^1 x_2^2 + 1 \\ x_1^1 + x_1^2 \\ x_2^1 + x_2^2 \\ x_1^1 x_1^2 + x_2^1 x_2^2 - 1 \end{bmatrix}.$$

We choose a base triangle $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2$. The control vectors are computed by $\mathbf{b}_{200} = P(\mathbf{s}_0, \mathbf{s}_0)\mathbb{R}$, $\mathbf{b}_{110} = P(\mathbf{s}_0, \mathbf{s}_1)\mathbb{R}$, and so on. Fig. 1.59 shows two examples. \diamond

Example 1.4.10. We consider the special case of a rational TP surface of degree $(1, n)$. One family of parameter lines consists of lines, therefore such a surface is a ruled surface. We will see later (cf. Equ. (5.19)) that all rational ruled surfaces are rational TP surfaces of degree $(1, n)$ for an appropriate n .

The special case $n = 1$ is a surface which carries two families of lines, i.e., a ruled quadric. Conversely, all ruled quadrics Φ are rational TP surfaces of degree $(1, 1)$: Choose three skew generators G_1, G_2, G_3 from one regulus and three

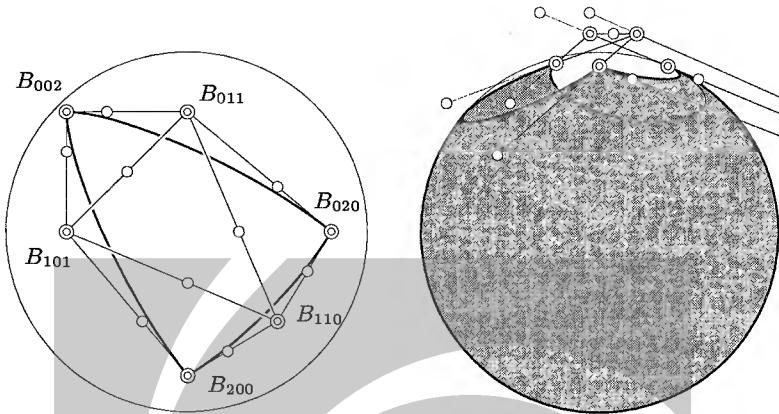


Fig. 1.59. Left: Bézier points and frame points of a rational triangular Bézier surface contained in the unit sphere (top view). The vertices of the base triangle are $s_0 = (0, 1/2)$, $s_1 = (1/2, 0)$, $s_2 = (-1/2, -1/2)$. Right: Bézier points and frame points of a second rational surface, covering almost the entire sphere, and defined by the base triangle $(5, 0), (0, 5), (-4, -4)$ (front view).

skew generators H_1, H_2, H_3 from the other. Then without loss of generality we can assume that $G_1 = b_{00}\mathbb{R} \vee b_{10}\mathbb{R}$, $G_3 = b_{01}\mathbb{R} \vee b_{11}\mathbb{R}$, $H_1 = b_{00}\mathbb{R} \vee b_{01}\mathbb{R}$, $H_3 = b_{10}\mathbb{R} \vee b_{11}\mathbb{R}$, and even that $G_2 = (b_{00} + b_{01})\mathbb{R} \vee (b_{10} + b_{11})\mathbb{R}$, $H_2 = (b_{00} + b_{10})\mathbb{R} \vee (b_{01} + b_{11})\mathbb{R}$. Then the rational TP surface with control vectors $b_{00}, b_{01}, b_{10}, b_{11}$ coincides with the original ruled quadric Φ . \diamond

2. Models of Line Space

The set of straight lines of projective three-space is a four-dimensional manifold with a geometric structure induced by the underlying geometry of the projective space P^3 . This projective geometry of lines is easily understood in the *Klein model*, where the lines of P^3 are identified with the points of a certain quadric in projective 5-space P^5 . This Klein quadric is a special case of a *Grassmann manifold*, a concept which is studied in Sec. 2.2.

The lines of Euclidean space with their metric properties lead to a different model, the *Study sphere*. Here oriented lines of E^3 are identified with the points of a unit sphere constructed with dual numbers instead of real ones.

These models are closely tied to the problem of introducing suitable coordinates for lines: The Klein quadric corresponds to Plücker coordinates of lines, which are a special case of Grassmann coordinates of subspaces. The Study sphere leads to dual coordinate vectors for oriented lines.

Working in a geometric point model enables better understanding and a simple interpretation of various objects of line space. The design of efficient algorithms involving lines is greatly simplified if it is based on the right geometric model.

2.1 The Klein Model

2.1.1 Plücker Coordinates

It is possible to introduce coordinates in the space of lines, such that span and intersection of points, lines, and planes are easily computed. This is the special three-dimensional case of the general theory of computing with subspaces in the Grassmann algebra, which will be presented in Sec. 2.2. Many results of this section are special cases of more general theorems, but we will also give independent proofs in most cases, or leave them to the reader as an exercise.

Following H. Grassmann (1844) and J. Plücker (1865), we introduce coordinates in the set \mathcal{L} of straight lines of real projective three-dimensional space P^3 in the following way: We choose a projective coordinate system in P^3 , which is equivalent to choosing a basis e_0, \dots, e_3 of \mathbb{R}^4 . All lines L can be written in the form $L = X \vee Y$, where X, Y are two projective points with homogeneous coordinate vectors x and y , respectively.

We introduce the concept of *exterior product* of vectors. This operation, denoted by the symbol ‘ \wedge ’, assigns to two vectors x and y of \mathbb{R}^4 a new vector $x \wedge y$, which is contained in a new vector space, denoted by $\Lambda^2\mathbb{R}^4$. The operation shall be *bilinear* in both arguments, i.e.,

$$x \wedge (\lambda y + \mu z) = \lambda x \wedge y + \mu x \wedge z, \quad (\lambda x + \mu y) \wedge z = \lambda x \wedge z + \mu y \wedge z;$$

and it shall be *anti-commutative*, i.e.,

$$x \wedge y = -y \wedge x.$$

If $x = \sum x_i e_i$ and $y = \sum y_i e_i$, the bilinearity of the exterior product (whose existence we have not yet shown) implies that

$$x \wedge y = \sum_{i,j=0}^3 x_i y_j e_i \wedge e_j.$$

The products of the basis vectors (there are 16 of them) are not linearly independent because of the anti-commutativity relation. Obviously it is necessary that

$$e_i \wedge e_j = -e_j \wedge e_i,$$

for all $i, j = 0, \dots, 3$. This also implies that the product of a basis vector with itself is zero. So there remain six products which can possibly be linearly independent, namely

$$e_0 \wedge e_1, e_0 \wedge e_2, e_0 \wedge e_3, e_2 \wedge e_3, e_3 \wedge e_1, e_1 \wedge e_2. \quad (2.1)$$

The product of two vectors is a combination of these six vectors, and we can write

$$x \wedge y = \sum_{(i,j) \in I} (x_i y_j - x_j y_i) e_i \wedge e_j, \quad (2.2)$$

with $I = \{(0,1), (0,2), (0,3), (2,3), (3,1), (1,2)\}$.

Having studied the consequences, we are now ready to *define* the exterior product:

Definition. Let the vector space $\Lambda^2\mathbb{R}^4$ be six-dimensional and have the basis (2.1). The exterior product of the vectors $x = \sum x_i e_i$ and $y = \sum y_i e_i$ is an element of $\Lambda^2\mathbb{R}^4$, and is given by Equ. (2.2).

Lemma 2.1.1. The exterior product is bilinear and anti-commutative. If the vector space of products is to be six-dimensional, it is uniquely determined by these two requirements.

Proof. It is an elementary exercise to verify that the exterior product as defined by Equ. (2.2) indeed is bilinear and anti-commutative. The uniqueness is clear because we were able to deduce the defining equation (2.2) from the assumptions of bilinearity and anti-commutativity. \square

We will eventually show (see Lemma 2.1.2 below) that the exterior product of two vectors essentially depends only on the line spanned by the points $x\mathbb{R}$ and $y\mathbb{R}$. Thus we call the coordinates of $x \wedge y$ with respect to the basis (2.1) the *Plücker coordinates* of this line. We write

$$(x_0, \dots, x_3) \wedge (y_0, \dots, y_3) = (l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12}), \quad l_{ij} = x_i y_j - x_j y_i.$$

The number l_{ij} is the determinant of columns i and j of the matrix

$$\begin{pmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \end{pmatrix}.$$

Remark 2.1.1. The *Grassmann coordinates* of a line are the numbers $l_{01}, l_{02}, l_{03}, l_{23}, l_{13} = -l_{31}, l_{12}$. They are coordinates with respect to the basis $\mathbf{e}_i \wedge \mathbf{e}_j$ ($0 \leq i < j \leq 3$). \diamond

It turns out that not all elements of $\Lambda^2 \mathbb{R}^4$ occur as exterior products of vectors of \mathbb{R}^4 . This and other properties are summarized in the following:

Lemma 2.1.2. *If we replace the points $x\mathbb{R}, y\mathbb{R}$ by other points of their span, the Plücker coordinate vector $x \wedge y$ remains the same up to a scalar factor.*

A nonzero six-tuple $L = (l_{01}, \dots, l_{12})$ of real numbers is the result of an exterior product if and only if it fulfills the relation

$$\Omega_q(L) = l_{01}l_{23} + l_{02}l_{31} + l_{03}l_{12} = 0. \quad (2.3)$$

Proof. First we compute $(a_1x + b_1y) \wedge (a_2x + b_2y) = a_1a_2x \wedge x + a_1b_2x \wedge y + b_1a_2y \wedge x + b_1b_2y \wedge y = (a_1a_2 - b_1b_2)x \wedge y$, which shows that $(x \wedge y)\mathbb{R}$ does not change if we replace x and y by two linearly independent linear combinations of x, y .

As to the second statement, it is an elementary calculation to verify that if $l_{ij} = x_i y_j - x_j y_i$ then L satisfies (2.3). To show the converse, we consider the four vectors

$$\begin{aligned} s_0 &= (0, l_{01}, l_{02}, l_{03}), \\ s_1 &= (-l_{01}, 0, l_{12}, -l_{31}), \\ s_2 &= (-l_{02}, -l_{12}, 0, l_{23}), \\ s_3 &= (-l_{03}, l_{31}, -l_{23}, 0). \end{aligned} \quad (2.4)$$

The exterior product of any two of them is a scalar multiple of L , and not all six possible exterior products are zero unless L is zero itself. The result will also follow from Th. 2.2.4 and Ex. 2.2.5. \square

The four points $s_0\mathbb{R}, \dots, s_3\mathbb{R}$ used in the previous proof are the points where the line with Plücker coordinate vector L intersects the coordinate planes $x_0 = 0, \dots, x_3 = 0$.

Equ. (2.3) is called the *Plücker identity*. We call elements of $\Lambda^2 \mathbb{R}^4$ which can be written as an exterior product of two vectors *simple*. As Lemma 2.1.2 shows, an element is simple if and only if it fulfills the Plücker identity.

Example 2.1.1. We compute the Plücker coordinates of the line spanned by the points with affine coordinates $(1, 0, -1)$ and $(2, -2, 0)$. The homogeneous coordinate vectors of these points are $\mathbf{a} = (1, 1, 0, -1)$ and $\mathbf{b} = (1, 2, -2, 0)$. The (2×2) -subdeterminants of the matrix

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & -2 & 0 \end{bmatrix}$$

defined by the pairs $(0, 1), (0, 2), \dots, (1, 2)$ of columns are given by

$$(l_{01}, \dots, l_{12}) = \left(\left| \begin{array}{cc} 1 & 1 \\ 1 & 2 \end{array} \right|, \dots, \left| \begin{array}{cc} 1 & 0 \\ 2 & -2 \end{array} \right| \right) = (1, -2, 1, -2, -2, -2).$$

Equ. (2.4) shows that the points $(0, 1, -2, 1)\mathbb{R}$, $(-1, 0, -2, 2)\mathbb{R}$, $(2, 2, 0, -2)\mathbb{R}$, and $(-1, -2, 2, 0)\mathbb{R}$ are contained in L . \diamond

The Range of Points Incident with a Line

The following question arises: If the Plücker coordinates of a line L are given, how do we find the points of L ?

First we extend exterior multiplication in an associative and bilinear way by $(\mathbf{e}_i \wedge \mathbf{e}_j) \wedge \mathbf{e}_k = \mathbf{e}_i \wedge (\mathbf{e}_j \wedge \mathbf{e}_k) = \mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k$, and by letting

$$\mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k = \mathbf{e}_j \wedge \mathbf{e}_k \wedge \mathbf{e}_i = -\mathbf{e}_k \wedge \mathbf{e}_j \wedge \mathbf{e}_i. \quad (2.5)$$

The product of three vectors of \mathbb{R}^4 is an element of $\Lambda^3 \mathbb{R}^4$, which contains the vectors $\mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k$. Because of the relations (2.5), any linear combination of the 64 possible products of basis vectors is actually a linear combination of the vectors

$$\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3, \quad \mathbf{e}_0 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3, \quad \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_3, \quad \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2. \quad (2.6)$$

We therefore define $\Lambda^3 \mathbb{R}^4$ to be a four-dimensional vector space with the basis (2.6).

Lemma 2.1.3. A point $\mathbf{z}\mathbb{R}$ is contained in the line $L = \mathbf{x}\mathbb{R} \vee \mathbf{y}\mathbb{R}$ if and only if $\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} = \mathbf{o}$.

Proof. We use the same notation as in Lemma 2.1.2 and let $\mathbf{L} = \mathbf{x} \wedge \mathbf{y}$. First it is easy to verify that the points $\mathbf{s}_1\mathbb{R}, \dots, \mathbf{s}_4\mathbb{R}$ of Equ. (2.4) fulfill the equation $\mathbf{s}_i \wedge \mathbf{L} = \mathbf{o}$. Two of these points, $\mathbf{s}_1\mathbb{R}$ and $\mathbf{s}_2\mathbb{R}$ say, span the line L , and any point $\mathbf{z} \in L$ can be written in the form $\mathbf{z} = \lambda_1 \mathbf{s}_1 + \lambda_2 \mathbf{s}_2$, which shows $\mathbf{z} \wedge \mathbf{L} = \mathbf{o}$.

The equation $\mathbf{z} \wedge \mathbf{L} = \mathbf{o}$ is a linear system of equations. We leave it to the reader to verify that its rank is two, so the solution space is two-dimensional, and therefore coincides with the set of homogeneous coordinate vectors of points in L . The result will follow also from Th. 2.2.2. \square

Axis Coordinates — The Pencil of Planes Incident with a Line

A line L does not only carry a pencil of points, but by duality it is the carrier of a pencil of planes. This pencil is spanned by any two of these planes, say \mathbf{Ra}^* , \mathbf{Rb}^* . The duality principle ensures that lines are described by six-tuples of *dual* Plücker coordinates, derived from plane coordinates, as well: They are computed by

$$(l_{01}^*, l_{02}^*, l_{03}^*, l_{23}^*, l_{31}^*, l_{12}^*) = \mathbf{a}^* \wedge \mathbf{b}^*, \quad (2.7)$$

and are also called *axis coordinates*.

Lemma 2.1.4. *If $(l_{01} : l_{02} : l_{03} : l_{23} : l_{31} : l_{12})$ are Plücker coordinates of a line, then*

$$(l_{01}^* : l_{02}^* : l_{03}^* : l_{23}^* : l_{31}^* : l_{12}^*) = (l_{23} : l_{31} : l_{12} : l_{01} : l_{02} : l_{03}) \quad (2.8)$$

are its dual Plücker coordinates.

Proof. The result follows directly from Ex. 2.2.7 and Th. 2.2.6, which will be shown later. An elementary proof is the following: We know that the four points of Equ. (2.4) are contained in the line L . All of them are incident with any of the planes $\mathbb{R}f_i$ whose homogeneous plane coordinate vectors are given by

$$\begin{aligned} f_0 &= (0, l_{23}, l_{31}, l_{12}), \\ f_1 &= (-l_{23}, 0, l_{03}, -l_{02}), \\ f_2 &= (-l_{31}, -l_{03}, 0, l_{01}), \\ f_3 &= (-l_{12}, l_{02}, -l_{01}, 0). \end{aligned} \quad (2.9)$$

It may happen that some of f_0, \dots, f_3 are zero, but never more than two. Therefore at least two of f_0, \dots, f_3 are homogeneous coordinate vectors of planes which contain L . Now we can compute the dual Plücker coordinates from any of the expressions $f_0 \wedge f_1, f_0 \wedge f_2, \dots$, which shows the result. \square

2.1.2 Computing with Plücker Coordinates

Abbreviations for Coordinate n -Tuples

Often real projective three-space P^3 occurs as the projective extension of real affine three-space E^3 . Then we use homogeneous Cartesian coordinates (x_0, \dots, x_3) , with proper points being characterized by $x_0 \neq 0$ and ideal points by $x_0 = 0$. We write the homogeneous coordinate vector \mathbf{x} of a point as

$$(x_0 : x_1 : x_2 : x_3) = \mathbf{x}\mathbb{R} = (x_0, \mathbf{x})\mathbb{R} \quad \text{with} \quad \mathbf{x} = (x_1, x_2, x_3). \quad (2.10)$$

In this way the point $(1, \mathbf{x})\mathbb{R}$ has the affine coordinates \mathbf{x} . The ideal point $(0, \mathbf{x})\mathbb{R}$ is contained in all lines parallel to the vector \mathbf{x} . Analogously, for the plane with equation $u_0x_0 + \dots + u_3x_3 = 0$ and homogeneous plane coordinate vector $\mathbf{u} = (u_0, \dots, u_3)$ we write

$$(u_0 : u_1 : u_2 : u_3) = \mathbb{R}\mathbf{u} = \mathbb{R}(u_0, \mathbf{u}) \quad \text{with } \mathbf{u} = (u_1, u_2, u_3). \quad (2.11)$$

The plane at infinity has coordinates (u_0, \mathbf{o}) with $u_0 \neq 0$. All other planes $\mathbb{R}(u_0, \mathbf{u})$ have the normal vector \mathbf{u} . The Plücker coordinate vector of a line G will be written in the form

$$\begin{aligned} (l_{01} : l_{02} : l_{03} : l_{23} : l_{31} : l_{12}) &= (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R} \quad \text{with} \\ \mathbf{l} &= (l_{01}, l_{02}, l_{03}), \quad \bar{\mathbf{l}} = (l_{23}, l_{31}, l_{12}). \end{aligned} \quad (2.12)$$

If $\mathbf{l} \neq \mathbf{o}$, then the ideal point

$$G_u = (0, \mathbf{l})\mathbb{R} = (0, l_{01}, l_{02}, l_{03})\mathbb{R}$$

is contained in G , which follows from Lemma 2.1.3.

Computing Span and Intersection

The line L spanned by points $(x_0, \mathbf{x})\mathbb{R}$ and $(y_0, \mathbf{y})\mathbb{R}$ is given by

$$L = (x_0, \mathbf{x}) \wedge (y_0, \mathbf{y}) = (x_0\mathbf{y} - y_0\mathbf{x}, \mathbf{x} \times \mathbf{y})\mathbb{R}, \quad (2.13)$$

where \times denotes the vector product in \mathbb{R}^3 . This formula is easy to verify directly. Dual to Equ. (2.13), the dual Plücker coordinates of the line L contained in the two planes $\mathbb{R}(u_0, \mathbf{u})$ and $\mathbb{R}(v_0, \mathbf{v})$ are

$$(\mathbf{l}^*, \bar{\mathbf{l}}^*) = (u_0\mathbf{v} - v_0\mathbf{u}, \mathbf{u} \times \mathbf{v}). \quad (2.14)$$

By Lemma 2.1.4, its Plücker coordinates are given by

$$(\mathbf{l}, \bar{\mathbf{l}}) = (\mathbf{u} \times \mathbf{v}, u_0\mathbf{v} - v_0\mathbf{u}). \quad (2.15)$$

The following formulae are verified either by expanding and comparing with the formulae for span and intersection discussed in Sec. 2.2 for the general case, or directly, which is a list of elementary exercises. We therefore do not prove them here. If a line $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ and a point $(x_0, \mathbf{x})\mathbb{R}$ are not incident, there is a unique plane $\mathbb{R}(u_0, \mathbf{u})$ spanned by them. It can be computed by the formula

$$(u_0, \mathbf{u}) = (\mathbf{x} \cdot \bar{\mathbf{l}}, -x_0\bar{\mathbf{l}} + \mathbf{x} \times \mathbf{l}), \quad (2.16)$$

The dual formula computes the intersection point $(p_0, \mathbf{p})\mathbb{R}$ of a line $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ and a non-incident plane $\mathbb{R}(u_0, \mathbf{u})$:

$$(p_0, \mathbf{p}) = (\mathbf{u} \cdot \mathbf{l}, -u_0\mathbf{l} + \mathbf{u} \times \bar{\mathbf{l}}). \quad (2.17)$$

If two lines $G = (\mathbf{g}, \bar{\mathbf{g}})\mathbb{R}$ and $H = (\mathbf{h}, \bar{\mathbf{h}})\mathbb{R}$ intersect, their span $G \vee H = \mathbb{R}(u_0, \mathbf{u})$ can be computed as one of the spans $G \vee (0, \mathbf{h})\mathbb{R}$ or $H \vee (0, \mathbf{g})\mathbb{R}$, which leads to the formula

$$\mathbb{R}(u_0, \mathbf{u}) = \mathbb{R}(\bar{\mathbf{g}} \cdot \mathbf{h}, -\mathbf{g} \times \mathbf{h}) = \mathbb{R}(\mathbf{g} \cdot \bar{\mathbf{h}}, \mathbf{g} \times \mathbf{h}). \quad (2.18)$$

This formula fails if the ideal points $(0, \mathbf{g})\mathbb{R}$ of G and $(0, \mathbf{h})\mathbb{R}$ of H coincide.

The dual version of this formula computes the intersection point $(p_0, \mathbf{p})\mathbb{R} = G \cap H$ by

$$(p_0, \mathbf{p})\mathbb{R} = (\mathbf{g} \cdot \bar{\mathbf{h}}, -\bar{\mathbf{g}} \times \bar{\mathbf{h}})\mathbb{R} = (\bar{\mathbf{g}} \cdot \mathbf{h}, \bar{\mathbf{g}} \times \bar{\mathbf{h}})\mathbb{R}. \quad (2.19)$$

This fails if $\{\bar{\mathbf{g}}, \bar{\mathbf{h}}\}$ is linearly dependent.

These formulae are valid in general projective coordinate frames, because all such frames define an affine part with equation $x_0 \neq 0$ and an ideal plane with equation $x_0 = 0$ in P^3 .

Changes of Coordinates

The choice of a new coordinate system in P^3 changes the homogeneous coordinate vectors of points according to the rule

$$x'_j = \sum_k c_{jk} x_k,$$

where c_{jk} is a regular matrix. We denote the homogeneous coordinate vectors of a point X with respect to the old and new coordinate systems by \mathbf{x} and \mathbf{x}' , respectively. We now compute the Plücker coordinate vector of a line spanned by two points with respect to the new coordinate system:

$$\begin{aligned} l'_{ij} &= x'_i y'_j - x'_j y'_i \\ &= (\sum_k c_{ik} x_k)(\sum_l c_{jl} y_l) - (\sum_l c_{jl} x_l)(\sum_k c_{ik} y_k) \\ &= \sum_{k,l} c_{ik} c_{jl} (x_k y_l - x_l y_k) \\ &= \sum_{k,l} c_{ik} c_{jl} l_{kl}, \end{aligned} \quad (2.20)$$

where l_{kl} is an abbreviation for the expression $x_k y_l - x_l y_k$. This coincides with the previous definition of l_{kl} if (k, l) is one of $(0, 1), (0, 2), (0, 3), (2, 3), (3, 1)$, or $(1, 2)$. Obviously $l_{kk} = 0$ and $l_{kl} = -l_{lk}$, so l'_{ij} is a linear combination of the coefficients of the original Plücker coordinate vector (l_{01}, \dots, l_{12}) . This shows that Plücker coordinates transform in the same way as homogeneous point coordinates: We have

$$(l'_{01}, \dots, l'_{12})^T = \tilde{C} \cdot (l_{01}, \dots, l_{12}), \quad (2.21)$$

where the entries of the matrix \tilde{C} are given in detail by Equ. (2.20). The matrix \tilde{C} is regular, because all possible Plücker coordinates (with respect to the new coordinate system) are linear combinations of its columns, and it is easy to find six linear independent Plücker coordinate vectors.

As \mathbf{L}' is the Plücker coordinate vector of a line if and only if $\Omega_q(\mathbf{L}') = 0$, this transformation must have the property that $\Omega_q(\mathbf{L}) = 0$ if and only if $\Omega_q(\mathbf{L}') = 0$.

The Incidence Relation

The point $(x_0, \mathbf{x})\mathbb{R}$ is incident with the plane $\mathbb{R}(u_0, \mathbf{u})$ if and only if

$$u_0 x_0 + \mathbf{u} \cdot \mathbf{x} = 0, \quad (2.22)$$

where the dot indicates the canonical scalar product in \mathbb{R}^3 . A point $\mathbb{R}(x_0, \mathbf{x})$ is contained in the line $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ if and only if in the exterior product $(x_0, \mathbf{x}) \wedge (\mathbf{l}, \bar{\mathbf{l}}) = 0$. This is expressed by

$$\mathbf{x} \cdot \bar{\mathbf{l}} = 0, \quad -x_0 \bar{\mathbf{l}} + \mathbf{x} \times \mathbf{l} = \mathbf{o}. \quad (2.23)$$

There is also a dual version of this equation: A line $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ and plane $\mathbb{R}(u_0, \mathbf{u})$ are incident if and only if

$$\mathbf{u} \cdot \mathbf{l} = 0, \quad -u_0 \mathbf{l} + \mathbf{u} \times \bar{\mathbf{l}} = \mathbf{o}. \quad (2.24)$$

We define the bilinear form Ω by

$$\begin{aligned} \Omega(\mathbf{G}, \mathbf{H}) &= \Omega((\mathbf{g}, \bar{\mathbf{g}}), (\mathbf{h}, \bar{\mathbf{h}})) = \mathbf{g} \cdot \bar{\mathbf{h}} + \bar{\mathbf{g}} \cdot \mathbf{h} \\ &= g_{01}h_{23} + g_{02}h_{31} + g_{03}h_{12} + g_{23}h_{01} + g_{31}h_{02} + g_{12}h_{03}. \end{aligned} \quad (2.25)$$

This bilinear form corresponds to the quadratic form $\Omega_q(\mathbf{G})$ as defined by Equ. (2.3): $\Omega(\mathbf{G}, \mathbf{G}) = 2\Omega_q(\mathbf{G})$, and $\Omega(\mathbf{G}, \mathbf{H}) = \Omega_q(\mathbf{G} + \mathbf{H}) - \Omega_q(\mathbf{G}) - \Omega_q(\mathbf{H})$.

Lemma 2.1.5. *Two lines G, H with Plücker coordinate vectors $\mathbf{G} = (\mathbf{g}, \bar{\mathbf{g}})$ and $\mathbf{H} = (\mathbf{h}, \bar{\mathbf{h}})$ intersect if and only if*

$$\mathbf{g} \cdot \bar{\mathbf{h}} + \bar{\mathbf{g}} \cdot \mathbf{h} = \Omega((\mathbf{g}, \bar{\mathbf{g}}), (\mathbf{h}, \bar{\mathbf{h}})) = 0. \quad (2.26)$$

Proof. This is a special case of Th. 2.2.2, but we will also show an independent proof. First the statement is true if $G = H$, so can leave this case aside. As $\Omega_q(\mathbf{G}) = \Omega_q(\mathbf{H}) = 0$, we have $\Omega(\mathbf{G}, \mathbf{H}) = \Omega_q(\mathbf{G} + \mathbf{H})$. So we have to show that $\Omega_q(\mathbf{G} + \mathbf{H}) = 0$ if and only if G and H intersect.

A change of coordinates does not affect the vanishing of Ω_q , so we can choose any suitable projective coordinate system $(B_0, B_1, B_2, B_3; E)$ and compute $\Omega(\mathbf{G} + \mathbf{H})$: If G and H intersect, we choose $B_0 = G \cap H$, $B_1 \in G$ and $B_2 \in H$, which implies that $\mathbf{G} = \mathbf{e}_0 \wedge \mathbf{e}_1$, $\mathbf{H} = \mathbf{e}_0 \wedge \mathbf{e}_2$. Computation shows that $\Omega_q(\mathbf{G} + \mathbf{H}) = 0$. If G and H are skew, we choose $B_0, B_1 \in G$ and $B_2, B_3 \in H$. Then $\mathbf{G} = \mathbf{e}_0 \wedge \mathbf{e}_1$, $\mathbf{H} = \mathbf{e}_2 \wedge \mathbf{e}_3$, and $\Omega_q(\mathbf{G} + \mathbf{H}) \neq 0$. \square

Direction and Moment Vector of a Line

We want to investigate the geometric meaning of the two vectors $\mathbf{l}, \bar{\mathbf{l}}$ which occur if we write the Plücker coordinates of a line L in the form given by Equ. (2.12), and if the underlying coordinate system is a homogeneous Cartesian coordinate system.

We already know that $(0, \mathbf{l})\mathbb{R}$ is the line's ideal point if the line is not itself contained in the ideal plane — the vector \mathbf{l} indicates the *direction* of L . A non-ideal line is spanned by the point $(0, \mathbf{l})\mathbb{R}$ and a proper point $\mathbf{x} = (1, \mathbf{x})\mathbb{R}$. Then, the Plücker coordinates of L are computed with Equ. (2.13) as

$$(\mathbf{l}, \bar{\mathbf{l}}) = (\mathbf{l}, \mathbf{x} \times \mathbf{l}). \quad (2.27)$$

Thus the vector $\bar{\mathbf{l}}$ is the *moment vector* of a line-bound force of magnitude $\|\mathbf{l}\|$ with respect to the origin. The Plücker relation (2.3) expresses the orthogonality of \mathbf{l} and $\bar{\mathbf{l}}$. This interpretation will be elaborated later (see Sec. 2.3 and Sec. 3.4.2).

2.1.3 The Klein Quadric

The homogeneity of Plücker coordinates suggests to view the coordinates of a line as homogeneous coordinates of points in five-dimensional projective space P^5 . This projective space is modeled over the linear space \mathbb{R}^6 , which we identify with $\Lambda^2 \mathbb{R}^4$, and which has the basis $e_0 \wedge e_1, e_0 \wedge e_2, e_0 \wedge e_3, e_2 \wedge e_3, e_3 \wedge e_1, e_1 \wedge e_2$. Homogeneous coordinates in P^5 are written in the form $(l_{01} : l_{02} : l_{03} : l_{23} : l_{31} : l_{12})$.

Definition. *The Klein mapping $\gamma : \mathcal{L} \rightarrow P^5$ assigns to a line L of P^3 the point $(l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12})\mathbb{R}$ of P^5 , where (l_{01}, \dots, l_{12}) are the line's Plücker coordinates (F. Klein, 1868).*

Lemma 2.1.2 then immediately implies

Theorem 2.1.6. *The lines of projective three-space are, via the Klein mapping, in one-to-one correspondence with the points of the quadric $\Omega_q = 0$ of real projective 5-space (cf. Equ. (2.3)).*

Definition. *The quadric of P^5 defined by Equ. (2.3) is called the Klein quadric, and is denoted by the symbol M_2^4 .*

We split a coordinate six-tuple \mathbf{L} into its parts $(\mathbf{l}, \bar{\mathbf{l}})$ as described above. The equation $\Omega_q(\mathbf{L}) = \mathbf{l} \cdot \bar{\mathbf{l}} = 0$ actually defines a quadric, because the corresponding bilinear form Ω defined by Equ. (2.25) is regular: It has the coordinate matrix

$$K = \begin{bmatrix} 0 & \text{diag}(1, 1, 1) \\ \text{diag}(1, 1, 1) & 0 \end{bmatrix} \quad (2.28)$$

with respect to the basis $e_0 \wedge e_1, \dots, e_1 \wedge e_2$ — the equation $\mathbf{l} \cdot \bar{\mathbf{l}} = 0$ of M_2^4 can be written in the form $(\mathbf{l}, \bar{\mathbf{l}})^T \cdot K \cdot (\mathbf{l}, \bar{\mathbf{l}}) = 0$.

Remark 2.1.2. An algebraic variety of dimension d and degree k will sometimes be denoted by a capital letter with subscript k and superscript d . The Klein quadric M_2^4 is an example. If the variety has degree 1 (i.e., it is a subspace), then we omit the subscript 1, so notations like P^3, P^5 also fit into the scheme. ◇

One-dimensional Subspaces Contained in the Klein Quadric

A point $A = a\mathbb{R}$ of P^3 and a non-incident line G define a pencil of lines, which consists of all lines that contain A and intersect G . If $G = X \vee Y$ with $X = x\mathbb{R}$ and $Y = y\mathbb{R}$, then the line G is parametrized by $p(\lambda, \mu) = \lambda x + \mu y$ ($\lambda, \mu \in \mathbb{R}$, not both zero), and the Plücker coordinates of the pencil's lines are given by

$$L(\lambda, \mu) = p(\lambda, \mu) \wedge a = \lambda(x \wedge a) + \mu(y \wedge a) = \lambda X + \mu Y, \quad (2.29)$$

which is a linear combination of the Plücker coordinate vectors X and Y , which belong to the lines $A \vee X$ and $A \vee Y$, respectively (see also Fig. 2.1). This shows the first part of the following

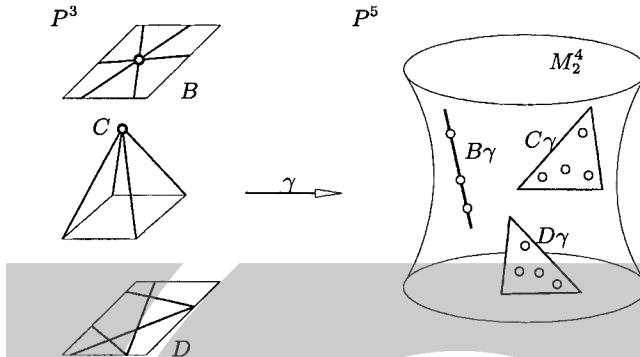


Fig. 2.1. One-dimensional and two-dimensional subspaces contained in the Klein quadric

Lemma 2.1.7. *The Klein mapping takes a pencil of lines to a straight line contained in M_2^4 . Vice versa, two points $X\mathbb{R}$, $Y\mathbb{R}$ of M_2^4 correspond to intersecting lines if and only if their span is contained in M_2^4 .*

Proof. We still have to show the second statement. Suppose that X and Y are Plücker coordinate vectors of lines G and H , and consider a linear combination $Z = \lambda X + \mu Y$ with $\lambda, \mu \neq 0$. The computation

$$\Omega(Z, Z) = \lambda^2 \Omega(X, X) + 2\lambda\mu \Omega(X, Y) + \mu^2 \Omega(Y, Y) = 2\lambda\mu \Omega(X, Y)$$

shows that $\Omega(X, Y) = 0$ is equivalent to $\Omega(Z, Z) = 0$. Therefore, Z is the Plücker coordinate vector of a line if and only if the lines G and H intersect (by Equ. (2.26)). \square

Recall the discussion of bilinear forms and quadratic forms on p. 31. By definition, the quadratic form Ω_q defined by Equ. (2.3) is a quadratic form whose vanishing is the defining equation of the Klein quadric. The corresponding bilinear form, whose vanishing expresses conjugacy of points, is $\Omega/2$. As scalar factors have no influence on whether an expression is zero or not, we have the following result:

Corollary 2.1.8. *The Klein mapping maps intersecting lines to conjugate points of the Klein quadric, and vice versa.*

2-dimensional Subspaces Contained in the Klein Quadric

The set of lines of P^3 concurrent in a point $P = p\mathbb{R}$ are called a bundle of lines. In order to find a parametrization of the lines of a bundle, we choose a plane which is not incident with $p\mathbb{R}$. It is spanned by three points $X = x\mathbb{R}$, $Y = y\mathbb{R}$, and $Z = z\mathbb{R}$. Then the Plücker coordinate vector of a line L of the bundle is given by

$$L = p \wedge (\lambda X + \mu Y + \nu Z) = \lambda X + \mu Y + \nu Z. \quad (2.30)$$

We see that the Klein mapping transforms line bundles to two-dimensional subspaces contained in the Klein quadric.

Remark 2.1.3. As a bundle of lines is itself a projective plane (cf. p. 5), we can ask whether the Klein mapping, restricted to a bundle, is a projective mapping. This is clear from Equ. (2.30), and also clear from Lemma 2.1.7: The ‘lines’ of the bundle are line pencils, and the Klein mapping transforms these ‘lines’ to lines in P^5 . ◇

The object dual to a *bundle* of lines is a *field of lines*, which consists of all lines incident with a given plane. By the duality principle the Klein mapping transforms a field of lines to the points of a two-dimensional subspace contained in M_2^4 . This is illustrated in Fig. 2.1.

Lemma 2.1.9. *The Klein quadric M_2^4 carries two 3-parameter families Π_b, Π_f of two-dimensional subspaces (planes), which correspond to line bundles and line fields, respectively. There are no other two-dimensional subspaces contained in M_2^4 . Any two different planes of the same family intersect in a single point, whereas the intersection of planes of different families is either empty or a line.*

Proof. We have to show that a plane which is entirely contained in the Klein quadric, is the Klein image of a bundle or a field. Take three points of P^5 which span this plane: Lemma 2.1.7 shows that they are images of pairwise intersecting lines. Three such lines are either concurrent or co-planar, which shows the first part of the lemma.

The statements about the intersection of the two types of planes follow immediately from the fact that two bundles have exactly one line in common, two fields have exactly one line in common, and a bundle and a field either have no line in common (if the vertex of the bundle is not contained in the carrier plane of the field) or they share a line pencil. ◻

Remark 2.1.4. By Th. 1.1.32, a regular quadric in P^5 cannot contain 3-dimensional subspaces. Thus Lemma 2.1.9 shows that the Klein quadric has maximal index, namely 2. The behaviour of the planes contained in M_2^4 is in some respects similar to the behaviour of the subspaces contained in a ruled quadric in P^3 (two one-parameter families of one-dimensional subspaces, i.e., reguli). ◇

Remark 2.1.5. We discussed Plücker coordinates and the Klein quadric only for the case of a real number field, but everything is valid for complex projective space as well. We get a complex Klein quadric $M_2^4(\mathbb{C})$ contained in $\mathbb{C}P^5$, defined by the same equation. Without going into details, we mention that the existence of planes in $M_2^4(\mathbb{C})$ which do not intersect, shows that it is not possible to find an ‘algebraic model’ for the lines of projective space which itself is a projective space, because every two algebraic sub-varieties of complementary dimension of a complex projective space must intersect. ◇

Remark 2.1.6. There is even no one-to-one *continuous* mapping of the space of lines \mathcal{L} onto a projective space, where the topology in \mathcal{L} is that of the Klein model. ◇

Projective Automorphisms of the Klein Quadric

A projective automorphism $\kappa : P^3 \rightarrow P^3$ maps lines to lines in a bijective way, and therefore induces, via the Klein mapping γ , a bijection of the Klein quadric, which is denoted by $\tilde{\kappa}$:

$$\begin{array}{ccc} M_2^4 & \xrightarrow{\tilde{\kappa}} & M_2^4 \\ \gamma \uparrow & & \uparrow \gamma \\ \mathcal{L} & \xrightarrow{\kappa} & \mathcal{L} \end{array} \quad (2.31)$$

We can write κ in the form $\kappa(x\mathbb{R}) = x'\mathbb{R}$ with $x'_j = \sum_k c_{jk}x_k$. If a line L is spanned by points $x\mathbb{R}, y\mathbb{R}$, its image is spanned by $x'\mathbb{R}, y'\mathbb{R}$. As a projective mapping transforms homogeneous coordinates in the same way as a change of coordinates, we can use Equ. (2.20) and the discussion following it to deduce that the mapping $\tilde{\kappa}$ is described by a linear mapping of Plücker coordinates, i.e., is itself a projective collineation of P^5 .

Because κ transforms bundles to bundles and fields to fields, $\tilde{\kappa}$ does not transform a plane of Π_b to a plane of Π_f ; these two families of planes are only permuted within themselves.

There are, however, projective collineations of P^3 into its dual — the correlations of P^3 — which map lines to lines, but bundles to fields and fields to bundles. They also induce projective automorphisms of M_2^4 , which is seen from a calculation completely analogous to the one which lead to Equ. (2.20). This motivates the following theorem:

Theorem 2.1.10. *Projective collineations and correlations of P^3 induce projective automorphisms of the Klein quadric, and the Klein quadric does not admit other projective automorphisms.*

Proof. A projective automorphism α cannot change the dimension of an intersection or span. This shows that the two families Π_b and Π_f of planes in M_2^4 are either (i) interchanged or (ii) transformed within themselves; mixing is not possible because the dimension of the intersection of two planes decides whether these planes belong to the same family or to different families.

In case (ii) bundles are mapped to bundles, and their vertices are mapped by a bijective point-to-point transformation $\kappa : P^3 \rightarrow P^3$. Now collinear points can be characterized by vertices of bundles which share the *same* line. They are therefore mapped to collinear points, and Th. 1.1.11 shows that κ is a projective automorphism of P^3 . Case (i) is similar. \square



2.2 The Grassmann Algebra

Subspaces of linear spaces or, which is essentially the same, of projective spaces, can be efficiently coordinatized by alternating multilinear forms. We will briefly describe this concept in general and for dimension three in particular, which will again

show the results of Sec. 2.1. The proofs are intended to be more or less complete, but brief. Their reading will require more ‘mathematical maturity’ than other parts of this book, and is not essential for what follows.

Exterior Multiplication

Consider a linear space V of dimension n , and assume it has the basis $\mathbf{e}_1, \dots, \mathbf{e}_n$. The underlying number field will be \mathbb{R} unless explicitly stated. It should be emphasized that everything in this section also holds for the complex number field. For all $k = 1, \dots, n$, we call the symbols

$$\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} = \mathbf{e}_{\wedge i_1 \dots i_k} \quad (2.32)$$

the exterior product of the basis vectors $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k}$. The vector space $\Lambda^k V$, which is called the k -fold exterior product of V with itself, has the basis

$$\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}, \quad (i_1 < \dots < i_k). \quad (2.33)$$

It is therefore of dimension $\binom{n}{k}$. The basis given in Equ. (2.33) will be referred to as the *canonical* basis in $\Lambda^k V$. The vector space $\Lambda^0 V$ is identified with the number field. All vector spaces $\Lambda^k V$ with $k > n$ are identified with the zero vector space. The vector space

$$\Lambda V = \Lambda^0 V \oplus \Lambda^1 V \oplus \dots \oplus \Lambda^n V \quad (2.34)$$

is called the *exterior algebra* over V . Its dimension as a real vector space is $\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = 2^n$ (multiplication in this algebra will be defined below).

Example 2.2.1. Assume that $V = \mathbb{R}^4$, and the basis vectors are numbered from 0 to 3. Then $\Lambda^0 V = \mathbb{R}$, $\Lambda^1 V = V$, $\Lambda^2 V$ has the basis $\mathbf{e}_0 \wedge \mathbf{e}_1, \mathbf{e}_0 \wedge \mathbf{e}_2, \mathbf{e}_0 \wedge \mathbf{e}_3, \mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3$, and is therefore six-dimensional. $\Lambda^3 V$ has the basis $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3, \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2$, and is four-dimensional like V itself. $\Lambda^4 V$ has the basis vector $\mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$, and is one-dimensional. The entire exterior algebra $\Lambda \mathbb{R}^4$ therefore has dimension $1 + 4 + 6 + 4 + 1 = 16$. ◇

We define *exterior multiplication* in ΛV , which will be written as $\mathbf{a} \wedge \mathbf{b}$, as follows: It is linear in both arguments, and the products of the basis vectors are given by the obvious definition

$$(\mathbf{e}_{r_1} \wedge \dots \wedge \mathbf{e}_{r_k}) \wedge (\mathbf{e}_{s_1} \wedge \dots \wedge \mathbf{e}_{s_l}) = \mathbf{e}_{r_1} \wedge \dots \wedge \mathbf{e}_{r_k} \wedge \mathbf{e}_{s_1} \wedge \dots \wedge \mathbf{e}_{s_l},$$

where we identify

$$\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} = \pm \mathbf{e}_{j_1} \wedge \dots \wedge \mathbf{e}_{j_k}$$

with a plus sign if (j_1, \dots, j_k) is an *even* permutation of (i_1, \dots, i_k) , and with a minus sign if (j_1, \dots, j_k) is an *odd* permutation of (i_1, \dots, i_k) ; and by defining all products with two equal factors as zero.

Example 2.2.2. With $V = \mathbb{R}^4$, we have $\mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_2 \wedge \mathbf{e}_1$, and $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 = \mathbf{e}_3 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_3 \wedge \mathbf{e}_2 \wedge \mathbf{e}_1$.

We compute a product: $(3\mathbf{e}_1 + \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2) \wedge (\mathbf{e}_1 \wedge \mathbf{e}_0 - 1) = 3\mathbf{e}_1 \wedge \mathbf{e}_1 \wedge \mathbf{e}_0 - 3\mathbf{e}_1 + \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_1 \wedge \mathbf{e}_0 - \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 = -3\mathbf{e}_1 - \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2$. \diamond

Theorem 2.2.1. *The exterior product has the following properties:*

1. *The exterior product actually is consistently defined.*
2. *For all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \Lambda V$, we have $(\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c} = \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c})$, i.e., the exterior product is associative.*
3. *For all $\mathbf{a}_1, \dots, \mathbf{a}_k \in V$, and for all i, j with $1 \leq i < j \leq k$ we have*

$$\dots \wedge \mathbf{a}_i \wedge \dots \wedge \mathbf{a}_j \wedge \dots = - \dots \wedge \mathbf{a}_j \wedge \dots \wedge \mathbf{a}_i \wedge \dots$$

i.e., the exterior product is anti-commutative with respect to elements of V .

4. *If two factors are equal in an exterior product, then this product equals zero.*
5. *Assume that $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, and $\mathbf{a}_1, \dots, \mathbf{a}_n \in V$. Then $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_n = \det(\mathbf{a}_1, \dots, \mathbf{a}_n) \cdot \mathbf{e}_{\wedge 1\dots n}$.*
6. *An exterior product of vectors $\mathbf{a}_1, \dots, \mathbf{a}_k$ is zero if and only if $\mathbf{a}_1, \dots, \mathbf{a}_k$ are linearly dependent.*
7. *An element of $\Lambda^k V$ can be canonically identified with an alternating k -linear form $V^* \times \dots \times V^* \rightarrow \mathbb{R}$, where V^* is the vector space dual to V , consisting of all linear mappings $V \rightarrow \mathbb{R}$. Thereby $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ maps a k -tuple of linear forms according to*

$$\langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, \mathbf{b}_1^*, \dots, \mathbf{b}_k^* \rangle = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \mathbf{b}_{\sigma(1)}^*(\mathbf{a}_1) \cdots \mathbf{b}_{\sigma(k)}^*(\mathbf{a}_k),$$

where $\sigma \in S_k$ means that σ runs through all permutations of k elements.

8. *A linear mapping $L : V \rightarrow V$ induces a linear mapping $L_k : \Lambda^k V \rightarrow \Lambda^k V$, which commutes with the exterior product: $L_k(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k) = L(\mathbf{a}_1) \wedge \dots \wedge L(\mathbf{a}_k)$.*
9. *If $\mathbf{b}_1, \dots, \mathbf{b}_n$ is a basis of V , then $\mathbf{b}_{i_1} \wedge \dots \wedge \mathbf{b}_{i_k}$ for $1 \leq i_1 < \dots < i_k \leq n$ is a basis of $\Lambda^k V$.*

- Proof.*
1. If $\mathbf{i} = (i_1, \dots, i_r), \mathbf{j} = (j_1, \dots, j_s)$ are lists of indices, and $\sigma(\mathbf{i}), \tau(\mathbf{j})$ are permutations, we have to show that $\mathbf{e}_{\wedge \sigma(\mathbf{i})} \wedge \mathbf{e}_{\wedge \tau(\mathbf{j})} = \text{sgn}(\sigma) \text{sgn}(\tau) \mathbf{e}_{\wedge \mathbf{i}} \wedge \mathbf{e}_{\wedge \mathbf{j}}$. This is true if some i_k equals some j_l because then both expressions are zero. If \mathbf{i}, \mathbf{j} are distinct, this is true because the left hand side expands to $\mathbf{e}_{\wedge \sigma(\mathbf{i}), \sigma(\mathbf{j})}$, and the sign of the permutation $(\sigma(\mathbf{i})\sigma(\mathbf{j}))$ equals $\text{sgn}(\sigma) \text{sgn}(\tau)$.
 2. This holds for the basis by definition and is by linearity valid for all arguments.
 3. This holds if the \mathbf{a}_i are basis vectors. The general case follows from linearity.
 4. This is an immediate consequence of the previous assertion.
 5. This follows directly from the definition of the determinant: If $\mathbf{a}_i = (a_{i1}, \dots, a_{in})$, we have $\sum a_{1i_1} \mathbf{e}_{i_1} \wedge \dots \wedge \sum a_{ni_n} \mathbf{e}_{i_n} = \sum_{i_1, \dots, i_n} a_{1i_1} \cdots a_{ni_n} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_n} = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \text{sgn}(\sigma) \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n = \det(\mathbf{a}_1, \dots, \mathbf{a}_n) \cdot \mathbf{e}_{\wedge 1, \dots, n}$.

6. If one of the vectors \mathbf{a}_i is expressible as a linear combination of the others, their exterior product is zero by 4. If they are linearly independent, there are $\mathbf{a}_{k+1}, \dots, \mathbf{a}_n$ such that $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq 0$. Then 5. shows that $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_n$ is nonzero, which means necessarily that also $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ is nonzero.
7. This mapping is obviously linear in each argument. That it is alternating means nothing but the properties of the exterior product listed above, and is not important for what follows. We have to show that the mapping only depends on $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$. By linearity, it is sufficient to show this for the basis forms \mathbf{e}_j^* which map \mathbf{e}_j to 1 and the other basis vectors to zero. If $\mathbf{b}_i^* = \mathbf{e}_{j_i}^*, \mathbf{a}_i = \sum a_{ij} \mathbf{e}_j$, then $\langle \mathbf{e}_{j_1}^*, \dots, \mathbf{e}_{j_k}^* \rangle \mapsto \sum_{\sigma} \text{sgn}(\sigma) a_{1\sigma(j_1)} \cdots a_{k\sigma(j_k)}$, where σ ranges in the set of permutations of $\{j_1, \dots, j_k\}$. On the other hand, $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ expands to $\sum_{i_1, \dots, i_k} a_{1i_1} \cdots a_{ki_k} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} = \sum_{i_1 < \dots < i_k} \sum_{\sigma} a_{1\sigma(i_1)} \cdots a_{k\sigma(i_k)} \text{sgn}(\sigma) \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}$. This shows that $\langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, \mathbf{b}_1^*, \dots, \mathbf{b}_k^* \rangle$ can be computed using only the coefficients of $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ with respect to the basis $\{\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} \mid i_1 < \dots < i_k\}$, and therefore only depends on $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$.
8. We ask: If $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k = \mathbf{a}'_1 \wedge \dots \wedge \mathbf{a}'_k$, does $L(\mathbf{a}_1) \wedge \dots \wedge L(\mathbf{a}_k) = L(\mathbf{a}'_1) \wedge \dots \wedge L(\mathbf{a}'_k)$? A linear mapping L induces a linear mapping $L^* : V^* \rightarrow V^*$, defined by $(L^* \mathbf{b}^*) \mathbf{a} = \mathbf{b}^*(L\mathbf{a})$. Then for all \mathbf{b}_i^* , 7. implies that $\langle L\mathbf{a}_1 \wedge \dots \wedge L\mathbf{a}_k, \mathbf{b}_1^*, \dots, \mathbf{b}_k^* \rangle = \sum_{\sigma} \text{sgn } \sigma \mathbf{b}_{\sigma(1)}^*(L\mathbf{a}_1) \cdots \mathbf{b}_{\sigma(k)}^*(L\mathbf{a}_k) = \sum_{\sigma} \text{sgn } \sigma (L^* \mathbf{b}_{\sigma(1)}^*) \mathbf{a}_1 \cdots (L^* \mathbf{b}_{\sigma(k)}^*) \mathbf{a}_k = \langle \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k, (L^* \mathbf{b}_1^*, \dots, L^* \mathbf{b}_k^*) \rangle$, and the same for \mathbf{a}'_i instead of \mathbf{a}_i . This shows that $L(\mathbf{a}_1) \wedge \dots \wedge L(\mathbf{a}_k)$ only depends on $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$, and the answer is affirmative.
9. There is a linear mapping L defined by $L(\mathbf{e}_i) = \mathbf{b}_i$, which has an inverse L^{-1} . Both L and L^{-1} induce linear endomorphisms $\tilde{L}, \widetilde{L^{-1}}$ of $\Lambda^k V$. Because $L^{-1} L$ induces the identity in $\Lambda^k V$, we have $\widetilde{L^{-1}} = \widetilde{L^{-1}}$, so \widetilde{L} is invertible, and a \widetilde{L} -image of a basis is a basis. We observe that $\widetilde{L}(\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}) = \mathbf{b}_{i_1} \wedge \dots \wedge \mathbf{b}_{i_k}$, and we are done. \square

Simple Vectors

The reason why we consider the Grassmann algebra is that we can elegantly describe linear subspaces by exterior products. This is summarized in the following definition and theorem:

Definition. If $\mathbf{b} = \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ for $\mathbf{b} \in \Lambda^k V$, $\mathbf{a}_1, \dots, \mathbf{a}_k \in V$, then \mathbf{b} is called simple.

Theorem 2.2.2. For $\mathbf{a} \in \Lambda^k V$ the set of all $\mathbf{b} \in V$ such that $\mathbf{a} \wedge \mathbf{b} = 0$ is a linear subspace $L(\mathbf{a})$. If \mathbf{a} is nonzero and simple, i.e., we have $\mathbf{a} = \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$, then $L(\mathbf{a})$ equals the span $[\mathbf{a}_1, \dots, \mathbf{a}_k]$, which has dimension k . For simple \mathbf{a}, \mathbf{b} , the spaces $L(\mathbf{a}), L(\mathbf{b})$ are equal if and only if \mathbf{a} is a scalar multiple of \mathbf{b} .

Proof. The fact that $L(\mathbf{a})$ is a linear subspace follows from linearity. Assume that $\mathbf{a} = \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ is simple and nonzero. Then $\mathbf{a}_1, \dots, \mathbf{a}_k$ are linearly independent and span a k -dimensional linear subspace W of V . From the fact that $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k \wedge \mathbf{b} = 0$

if and only if the set $\{\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}\}$ is linearly dependent, immediately follows that $L(\mathbf{a}) = W$.

Clearly $L(\mathbf{a}) = L(\mathbf{b})$ if \mathbf{b} is a multiple of \mathbf{a} . Assume that $\mathbf{b}_1, \dots, \mathbf{b}_k$ span the same k -dimensional subspace as $\mathbf{a}_1, \dots, \mathbf{a}_k$. Then we can write $\mathbf{b}_j = \sum_i \lambda_{ij} \mathbf{a}_i$. We compute $\mathbf{b} = \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k = \sum_{i_1, \dots, i_k} \lambda_{i_1} \dots \lambda_{i_k} \mathbf{a}_{i_1} \wedge \dots \wedge \mathbf{a}_{i_k}$. Each of the summands is zero if (i_1, \dots, i_k) is not a permutation of $(1, \dots, k)$. If it is, then the summand equals a multiple of $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$, which implies that \mathbf{b} is a multiple of \mathbf{a} . \square

This gives a one-to-one correspondence between the k -dimensional linear subspaces of V and those one-dimensional subspaces of $\Lambda^k V$, which are spanned by simple elements. This can be used to assign *coordinates* to a subspace:

Grassmann Coordinates

Definition. For a k -dimensional linear subspace $W = [\mathbf{a}_1, \dots, \mathbf{a}_k]$ compute $\mathbf{a} = \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}$. The coefficients $a_{i_1 \dots i_k}$ are called Grassmann coordinates of W . The same definition applies to projective subspaces of a projective space.

Obviously the Grassmann coordinates are defined only up to a constant factor. Note that not all elements of $\Lambda^k V$ are decomposable, i.e., are exterior products of vectors, so it is to be expected that the Grassmann coordinates of subspaces fulfill some relations. They will be investigated later.

Example 2.2.3. (cf. Ex. 2.1.1 and Remark 2.1.1) Consider $V = \mathbb{R}^4$ and the subspace $W = [\mathbf{a}_1, \mathbf{a}_2]$ spanned by $\mathbf{a}_1 = (1, 1, 0, -1)$, $\mathbf{a}_2 = (1, 2, -2, 0)$. We compute its Grassmann coordinates, thereby numbering the basis vectors from 0 to 3: $\mathbf{a}_1 \wedge \mathbf{a}_2 = (\mathbf{e}_0 + \mathbf{e}_1 - \mathbf{e}_3) \wedge (\mathbf{e}_0 + 2\mathbf{e}_1 - 2\mathbf{e}_2) = \mathbf{e}_0 \wedge \mathbf{e}_1 - 2\mathbf{e}_0 \wedge \mathbf{e}_2 + \mathbf{e}_0 \wedge \mathbf{e}_3 - 2\mathbf{e}_1 \wedge \mathbf{e}_2 + 2\mathbf{e}_1 \wedge \mathbf{e}_3 - 2\mathbf{e}_2 \wedge \mathbf{e}_3$. The list of coordinate values is $(1, -2, 1, -2, 2, -2)$. Considered as a projective subspace of P^3 , W is a line. \diamond

The Grassmann Variety $G_{n,k}$

Theorem 2.2.3. The set of simple elements of $\Lambda^k V$ is an algebraic variety, and the one-dimensional subspaces spanned by simple elements form a projective variety in the projective space over $\Lambda^k V$.

Proof. We first prove that $\dim L(\mathbf{a}) < k$, if $\mathbf{a} \in \Lambda^k V$ is not simple: Assume that $\mathbf{x}_1, \dots, \mathbf{x}_l$ is a basis of $L(\mathbf{a})$ and complete this set to a basis $\mathbf{x}_1, \dots, \mathbf{x}_n$ of V . The element \mathbf{a} can be written in the form

$$\mathbf{a} = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \mathbf{x}_{i_1} \wedge \dots \wedge \mathbf{x}_{i_k}. \quad (2.35)$$

By construction, $\mathbf{a} \wedge \mathbf{x}_i = 0$ for $i = 1, \dots, l$. This equation is possible only if the coefficient $a_{i_1 \dots i_k}$ is zero if one of $1, \dots, l$ is not one of i_1, \dots, i_k . This is always

the case if $l > k$, and if $k = l$ this happens always unless $i_1, \dots, i_k = 1, \dots, l$. In the former case \mathbf{a} is zero and $L(\mathbf{a}) = V$. In the latter only one summand is nonzero and \mathbf{a} is simple.

So we have shown that a nonzero element \mathbf{a} is simple if and only if the kernel of the linear mapping $\mathbf{x} \mapsto \mathbf{a} \wedge \mathbf{x}$ has dimension $\geq k$ (in which case its dimension actually equals k), or equivalently, if its rank is $\leq n - k$ (in which case it actually equals $n - k$).

The matrix of this linear mapping depends on \mathbf{a} in a polynomial manner, and the fact that the rank is less or equal $n - k$ is expressed by the vanishing of all its $(n - k + 1)$ -minors, which are polynomial equations. \square

This projective variety is called the *Grassmann variety* $G_{n,k}$. We write $G_{n,k}(\mathbb{C})$ to indicate that the ground field is \mathbb{C} . Its dimension equals $k(n - k)$ and it is irreducible.

Example 2.2.4. Consider the special case $n = 4, k = 2$: We choose the coordinate systems $\mathbf{e}_0, \dots, \mathbf{e}_3$ in V and $\mathbf{e}_{\wedge 0,1,2}, \mathbf{e}_{\wedge 0,1,3}, \mathbf{e}_{\wedge 0,2,3}, \mathbf{e}_{\wedge 1,2,3}$ in $\Lambda^3 V$. If $\mathbf{a} = \sum_{0 \leq i < j \leq 3} a_{ij} \mathbf{e}_i \wedge \mathbf{e}_j$, then the mapping $\mathbf{x} \mapsto \mathbf{a} \wedge \mathbf{x}$ has the coordinate matrix

$$\begin{bmatrix} a_{12} & -a_{02} & a_{01} & \\ a_{13} & -a_{03} & & a_{01} \\ a_{23} & & -a_{03} & a_{02} \\ & a_{23} & -a_{13} & a_{12} \end{bmatrix},$$

which is found by computing $\mathbf{a} \wedge \mathbf{e}_0, \dots, \mathbf{a} \wedge \mathbf{e}_3$. The condition that it has rank 2 is expressed by the vanishing of all of its (3×3) -subdeterminants. This yields the equations

$$\begin{aligned} a_{01}(a_{12}a_{03} + a_{01}a_{23} - a_{02}a_{13}) &= 0, \\ a_{02}(\quad \quad \quad) &= 0, \dots \end{aligned}$$

where the expression in brackets is always the same. It is easy to convince oneself that these 16 equations are equivalent to $a_{12}a_{03} - a_{02}a_{13} + a_{01}a_{23} = 0$. \diamond

The Plücker Relations

It is a classical result of algebraic geometry that the Grassmann variety is defined by the equations described by Th. 2.2.4.

Theorem 2.2.4. Assume that V is an n -dimensional linear space, equipped with the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$, and that elements $\mathbf{a} \in \Lambda^k V$ are written in the form $\mathbf{a} = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}$. If \mathbf{a} is simple, then for all possible indices $i_0 < \dots < i_k, j_1 < \dots < j_{k-1}$,

$$\sum_{s=0}^k (-1)^s a_{i_0 \dots \hat{i}_s \dots i_k} a_{j_1 \dots i_s \dots j_{k-1}} = 0. \quad (2.36)$$

The hat means that i_s is deleted from the first list — it is then inserted into the second list at the appropriate place. All undefined Grassmann coordinates are set to

zero. This set of equations characterizes simple elements, i.e., defines the Grassmann variety.

Example 2.2.5. We derive the Plücker relations for $n = 4$ and $k = 2$, which must, of course, give the result of Ex. 2.2.4 again. If we let $i_0, i_1, i_2 = 0, 1, 2$ and $j_1 = 3$, then Th. 2.2.4 gives

$$a_{12}a_{03} - a_{02}a_{13} + a_{01}a_{23} = 0. \quad (2.37)$$

All other combinations either give the same relation or the zero relation (if j_1 is one of i_0, i_1, i_2). Thus the Grassmann variety $G_{4,2}$ of two-dimensional linear subspaces of \mathbb{R}^4 , or of lines in projective three-space, is defined by *one quadratic equation*.

If $n = 4, k = 1$, or $n = 4, k = 3$, there are no relations at all. This confirms the fact that the set of one-dimensional linear subspaces (or points of projective three-space) as well as the set of three-dimensional linear subspaces (planes of projective three-space) are projective spaces themselves: The former is trivial, and the latter follows from the duality principle. \diamond

Duality

Consider an n -dimensional linear space V with the basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ and the $(n-1)$ -dimensional projective space whose points are V 's one-dimensional subspaces.

A k -dimensional linear subspace W is at the same time a $(k-1)$ -dimensional projective subspace, and by duality it is an $(n-k-1)$ -dimensional projective subspace of the dual projective space, when we regard it as a set of hyperplanes.

Now we try to find *dual* Grassmann coordinates of subspaces, which generalize hyperplane coordinates. This is done as follows:

The \star Operator

We define the \star operator as a linear mapping $\Lambda^k V \rightarrow \Lambda^{n-k} V$ by its action on the basis: Consider an element $\mathbf{e}_{\wedge i_1 \dots i_k}$ of the canonical basis of $\Lambda^k V$. Denote the multi-index which complements (i_1, \dots, i_k) by (j_1, \dots, j_{n-k}) . Then $(i_1 \dots i_k, j_1 \dots j_{n-k})$ is a permutation of $(1 \dots n)$, which is either even or odd. We let

$$\star \mathbf{e}_{\wedge i_1 \dots i_k} = \text{sgn}(i_1 \dots i_k, j_1 \dots j_{n-k}) \cdot \mathbf{e}_{\wedge j_1 \dots j_{n-k}}. \quad (2.38)$$

This definition can be given a different meaning: We introduce an orthogonality relation in ΛV by defining the canonical bases of $\Lambda^k V$ to be orthonormal, and by letting $\mathbf{a} \cdot \mathbf{b} = 0$ if $\mathbf{a} \in \Lambda^k V$, $\mathbf{b} \in \Lambda^l V$ with $k \neq l$. Then there is the following theorem:

Theorem 2.2.5. For all $\mathbf{a} \in \Lambda^k V$, $\mathbf{x} \in \Lambda^{n-k} V$, the \star operator fulfills the equation

$$(\star \mathbf{a} \cdot \mathbf{x}) \cdot \mathbf{e}_{\wedge 1 \dots n} = \mathbf{a} \wedge \mathbf{x}. \quad (2.39)$$

Proof. $\Lambda^n V$ is one-dimensional and a basis is given by the exterior product $\mathbf{e}_{\wedge 1 \dots n}$, so if $\mathbf{a} \in \Lambda^k V$, then the mapping

$$\phi : \Lambda^{n-k} V \rightarrow \Lambda^n V : \mathbf{x} \mapsto \mathbf{a} \wedge \mathbf{x}$$

maps \mathbf{x} in a linear way to $\phi(\mathbf{x}) \cdot \mathbf{e}_{\wedge 1 \dots n}$ (we identify $\Lambda^n V$ and \mathbb{R}). The mapping $\mathbf{x} \mapsto \phi(\mathbf{x})$ is a linear form, and therefore there is $\bar{\mathbf{x}}\mathbf{a} \in \Lambda^{n-k} V$ which fulfills

$$(\bar{\mathbf{x}}\mathbf{a}) \cdot \mathbf{x} = \phi(\mathbf{x}) \quad (2.40)$$

for all $\mathbf{x} \in \Lambda^k V$. Clearly the mapping $\mathbf{a} \mapsto \bar{\mathbf{x}}\mathbf{a}$ is linear.

We compute $\bar{\mathbf{x}}(\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k})$ for $i_1 < \dots < i_k$. Assume that j_1, \dots, j_{n-k} with $j_1 < \dots < j_{n-k}$ are precisely those indices of $1, \dots, n$ which complement the set i_1, \dots, i_k . Then

$$\mathbf{e}_{\wedge i_1 \dots i_k} \wedge \mathbf{e}_{\wedge j_1 \dots j_{n-k}} = \text{sgn}(i_1 \dots i_k, j_1 \dots j_{n-k}) \cdot \mathbf{e}_{\wedge 1, \dots, n}.$$

This shows that $\bar{\mathbf{x}} = \star$. □

Example 2.2.6. Consider $V = \mathbb{R}^4$ with basis vectors $\mathbf{e}_0, \dots, \mathbf{e}_3$, and choose $k = 2$. We compute $\star(\mathbf{e}_0 \wedge \mathbf{e}_2)$: The indices 1, 3 complete the set $(0, 1, 2, 3)$, but the permutation $(0, 2, 1, 3)$ is odd. So $\star(\mathbf{e}_0 \wedge \mathbf{e}_2) = -\mathbf{e}_1 \wedge \mathbf{e}_3$.

To compute $\star \star \mathbf{e}_1$ we first notice that $\star \mathbf{e}_1 = -\mathbf{e}_0 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$ because $(1, 0, 2, 3)$ is an odd permutation, and that $\star(-\mathbf{e}_{\wedge 0,2,3}) = -\mathbf{e}_1$, because $0, 2, 3, 1$ is an even permutation. ◇

It is easy to show that $\star \star \mathbf{a} = (-1)^{k(n-k)} \mathbf{a}$, if \mathbf{a} is one of the vectors of the canonical basis of $\Lambda^k V$. From linearity follows that this equation holds for all $\mathbf{a} \in \Lambda^k V$. Ex. 2.2.6 and Ex. 2.2.7 confirm this result for the case $n = 4$.

Dual Grassmann Coordinates

A linear subspace W of V defines its annulator, which consists of all linear forms vanishing on W . The annulator is a subspace of the dual space V^* . From the projective viewpoint, a projective subspace defines its annulator, consisting of all hyperplanes which contain W . The linear and the projective annulator are, of course, closely related: The hyperplanes of the projective annulator clearly are the zero sets of the linear forms of the linear annulator.

If V is equipped with coordinates x_0, \dots, x_n , corresponding to the canonical basis $\mathbf{e}_0, \dots, \mathbf{e}_n$, then its dual basis $\mathbf{e}_0^*, \dots, \mathbf{e}_n^*$ are the linear forms $\mathbf{x} \mapsto x_0, \dots, \mathbf{x} \mapsto x_n$. The dual vector space V^* gives rise to an exterior algebra ΛV^* , and again all k -dimensional subspaces of V^* are represented by the simple elements of $\Lambda^k V^*$.

Definition. The dual Grassmann coordinates of a linear or projective subspace are the Grassmann coordinates of its annulator space.

It is fortunate that the dual Grassmann coordinates are easily computed from the Grassmann coordinates: It turns out that after identification of ΛV and ΛV^* by means of their canonical bases the passage to the annihilator space is given by the $*$ operator.

Theorem 2.2.6. *If $\mathbf{a} \in \Lambda^k V$ represents a linear or projective subspace, then $*\mathbf{a} \in \Lambda^{n-k} V$ represents its annihilator, if we identify V and V^* by means of the bases $\mathbf{e}_0, \dots, \mathbf{e}_n$ and $\mathbf{e}_0^*, \dots, \mathbf{e}_n^*$.*

Proof. (Sketch) The scalar product in ΛV has the property that $(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k) \cdot (\mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k) = \det(\mathbf{a}_i \cdot \mathbf{b}_j)_{i,j=1}^k$. This follows from the fact that the previous formula defines a symmetric bilinear form which agrees with the scalar product on the canonical basis.

If we apply an orthogonal linear mapping ϕ to the vectors $\mathbf{a}_i, \mathbf{b}_j$, then obviously the scalar product does not change. Also $\phi(\mathbf{e}_1) \wedge \dots \wedge \phi(\mathbf{e}_n) = \det(\phi) \cdot \mathbf{e}_{\wedge 1, \dots, n} = \mathbf{e}_{\wedge 1, \dots, n}$.

Furthermore, ϕ induces linear mappings ϕ_k in all $\Lambda^k V$. Then, the previous considerations show that $\phi_{n-k}(*\mathbf{a}) = *(\phi_k(\mathbf{a}))$ if $\mathbf{a} \in \Lambda^k V$.

Now the proof of the theorem is easy: Given a subspace U , we select an orthonormal basis $\mathbf{a}_1, \dots, \mathbf{a}_n$ of V such that $U = [\mathbf{a}_1, \dots, \mathbf{a}_k]$. There is an orthogonal linear transformation ϕ with $\phi(\mathbf{a}_i) = \mathbf{e}_i$. We have to show that $*(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k)$ represents U 's annihilator. It is enough to show that $*(\phi(\mathbf{a}_1) \wedge \dots \wedge \phi(\mathbf{a}_k))$ represents $\phi(U)$'s annihilator, but this follows directly from the definition. \square

Example 2.2.7. We illustrate duality in three-dimensional projective space. Its $(k-1)$ -dimensional subspaces are described by the simple elements of $\Lambda^k V$ with $V = \mathbb{R}^4$. We assume the canonical basis $\mathbf{e}_0, \dots, \mathbf{e}_3$. A basis of $\Lambda \mathbb{R}^4$ is given by $1, \mathbf{e}_0, \dots, \mathbf{e}_3, \mathbf{e}_{\wedge 01}, \dots, \mathbf{e}_{\wedge 23}, \mathbf{e}_{\wedge 012}, \dots, \mathbf{e}_{\wedge 123}, \mathbf{e}_{\wedge 0123}$. The $*$ operator is linear and acts on this basis in the following way:

$$\begin{array}{llll} *1 = \mathbf{e}_{\wedge 0123} & *\mathbf{e}_0 = \mathbf{e}_{\wedge 123} & *\mathbf{e}_{\wedge 01} = \mathbf{e}_{\wedge 23} & *\mathbf{e}_{\wedge 123} = -\mathbf{e}_0 \\ * \mathbf{e}_{\wedge 0123} = 1 & *\mathbf{e}_1 = -\mathbf{e}_{\wedge 023} & *\mathbf{e}_{\wedge 02} = -\mathbf{e}_{\wedge 13} & *\mathbf{e}_{\wedge 023} = \mathbf{e}_1 \\ & *\mathbf{e}_2 = \mathbf{e}_{\wedge 013} & *\mathbf{e}_{\wedge 03} = \mathbf{e}_{\wedge 12} & *\mathbf{e}_{\wedge 013} = -\mathbf{e}_2 \\ & *\mathbf{e}_3 = -\mathbf{e}_{\wedge 012} & *\mathbf{e}_{\wedge 12} = \mathbf{e}_{\wedge 03} & *\mathbf{e}_{\wedge 012} = \mathbf{e}_3 \\ & & *\mathbf{e}_{\wedge 13} = -\mathbf{e}_{\wedge 02} & \\ & & *\mathbf{e}_{\wedge 23} = \mathbf{e}_{\wedge 01} & \end{array}$$

We compute the dual Grassmann coordinates of the line spanned by the points with affine coordinates $(1, 0, -1)$ and $(2, -2, 0)$: Its Grassmann coordinates are found as $\mathbf{a} = (1, 1, 0, -1) \wedge (1, 2, -2, 0) = \mathbf{e}_0 \wedge \mathbf{e}_1 - 2\mathbf{e}_0 \wedge \mathbf{e}_2 + \mathbf{e}_0 \wedge \mathbf{e}_3 - 2\mathbf{e}_1 \wedge \mathbf{e}_2 + 2\mathbf{e}_1 \wedge \mathbf{e}_3 - 2\mathbf{e}_2 \wedge \mathbf{e}_3$ (see Ex. 2.2.3). The dual Grassmann coordinates (the Grassmann coordinates of the annihilator, which is a line in dual projective space) are therefore given by $*\mathbf{a}$. To indicate the fact we are in dual space, we add a star to the basis vectors. So $*\mathbf{a} = *(\mathbf{e}_0 \wedge \mathbf{e}_1 + \dots) = \mathbf{e}_2^* \wedge \mathbf{e}_3^* + 2\mathbf{e}_1^* \wedge \mathbf{e}_3^* + \mathbf{e}_1^* \wedge \mathbf{e}_2^* - 2\mathbf{e}_0^* \wedge \mathbf{e}_3^* - 2\mathbf{e}_0^* \wedge \mathbf{e}_2^* - 2\mathbf{e}_0^* \wedge \mathbf{e}_1^*$. \diamond

Example 2.2.8. (Continuation of Ex. 2.2.7) Here we identify the well known cross product of three vectors with our computations with Grassmann coordinates.

For three points $a \mathbb{R}$, $b \mathbb{R}$, $c \mathbb{R}$ of projective three-space, the plane spanned by these points has the Grassmann coordinates $(a \wedge b \wedge c) \mathbb{R}$. The plane coordinates of this plane are given by $\mathbb{R}u$ with $u = \star(a \wedge b \wedge c)$. An explicit computation reveals that the vector (u_0, u_1, u_2, u_3) equals

$$\left(\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, - \begin{vmatrix} a_0 & b_0 & c_0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}, - \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} \right)^T.$$

We write $u = a \times b \times c$. This is the same vector product as defined by Remark 1.2.12. \diamond

Span and Intersection of Subspaces

If a subspace U is given by $U = [a_1, \dots, a_r]$ with linearly independent a_i , then it is described by the exterior product $u = a_1 \wedge \dots \wedge a_r$. If the a_i are linearly dependent, then their exterior product is zero. If another subspace W is given by $W = [b_1, \dots, b_s]$, then their span is given by $U \vee W = [a_1, \dots, a_r, b_1, \dots, b_s]$. The union of the vectors a_i and b_j can be linearly dependent. If $w = b_1 \wedge \dots \wedge b_s$, then this set is independent if and only if $u \wedge w \neq 0$, in which case $U \vee W$ is described by the exterior product $u \wedge w$. If $u \wedge w = 0$, we can find the span $U \vee W$ by successively computing

$$u \wedge b_{i_1} \neq 0, \quad u \wedge b_{i_1} \wedge b_{i_2} \neq 0, \quad \dots$$

until we arrive at a maximal sequence

$$x = u \wedge b_{i_1} \wedge \dots \wedge b_{i_l}$$

which has the property that $x \wedge b_j$ is zero for all j . Then $U \vee W$ is described by $x \in \Lambda^{r+l} V$.

Example 2.2.9. Consider projective three-space and the two lines $g = (1, 0, 0) \vee (1, 2, 2)$ and $h = (0, 0, 0) \vee (2, 2, 2)$, where the points are written in affine coordinates. Then $a_1 = (1, 1, 0, 0)$, $a_2 = (1, 1, 2, 2)$, $b_1 = (1, 0, 0, 0)$ and $b_2 = (1, 2, 2, 2)$. The line g is described by $g = a_1 \wedge a_2 \in \Lambda^2 V$: $g = (e_0 + e_1) \wedge (e_0 + e_1 + 2e_2 + 2e_3) = 2e_{\wedge 02} + 2e_{\wedge 03} + 2e_{\wedge 12} + 2e_{\wedge 13}$. We compute $g \wedge b_1 = g \wedge e_0 = 2e_{\wedge 012} + 2e_{\wedge 013}$, and further $g \wedge b_1 \wedge b_2 = 0$. This shows that $g \vee h$ is a plane, whose Grassmann coordinates are given by $g \wedge e_0$. Its dual Grassmann coordinates are, according to Ex. 2.2.7, given by $2e_3^* - 2e_2^*$, which implies that the plane has homogeneous plane coordinates $\mathbb{R}(0, 0, -2, 2)$. \diamond

If U, W are subspaces and the subspace $T = U \cap W$ is their intersection, then their duals fulfill the relation $T^* = U^* \vee W^*$. Thus we can compute intersections of subspaces by computing the span of their annihilators, using dual Grassmann coordinates.

Example 2.2.10. We want to compute the intersection of the line g spanned by $(0, 0, 0)$ and $(1, 1, 1)$ with the plane $\varepsilon : x_1 = 2$ (affine coordinates). We first compute the Grassmann coordinates of the line g by $\mathbf{g} = \mathbf{e}_0 \wedge (\mathbf{e}_0 + \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = \mathbf{e}_{\wedge 01} + \mathbf{e}_{\wedge 02} + \mathbf{e}_{\wedge 03}$. Its dual Grassmann coordinates are, according to Ex. 2.2.7, given by $\mathbf{g}^* = \mathbf{e}_{\wedge 23}^* - \mathbf{e}_{\wedge 13}^* + \mathbf{e}_{\wedge 12}^*$. The plane coordinates of ε are $\mathbb{R}(2, -1, 0, 0)$, so its is described by the element $2\mathbf{e}_0^* - \mathbf{e}_1^*$ of $\Lambda^1(\mathbb{R}^{4*})$. Now $\mathbf{g}^* \wedge (2\mathbf{e}_0^* - \mathbf{e}_1^*) = 2\mathbf{e}_{\wedge 230}^* - 2\mathbf{e}_{\wedge 130}^* + 2\mathbf{e}_{\wedge 120}^* - \mathbf{e}_{\wedge 231}^*$, which shows that the intersection point $g \cap \varepsilon$ is given by $\mathbf{e}_0 + 2\mathbf{e}_1 + 2\mathbf{e}_2 + 2\mathbf{e}_3$, or, in affine coordinates, by the point $(2, 2, 2)$. \diamond

2.3 The Study Sphere

Dual Numbers

The construction of *dual numbers* from the reals is similar to the construction of complex numbers: a dual number is a formal sum

$$a + \varepsilon b, \quad a, b \in \mathbb{R} \quad (2.41)$$

and addition and multiplication rules are given by

$$\begin{aligned} (a_1 + \varepsilon b_1) + (a_2 + \varepsilon b_2) &= (a_1 + a_2) + \varepsilon(b_1 + b_2) \\ (a_1 + \varepsilon b_1) \cdot (a_2 + \varepsilon b_2) &= (a_1 a_2) + \varepsilon(a_1 b_2 + a_2 b_1). \end{aligned} \quad (2.42)$$

The set of dual numbers is denoted by \mathbb{D} . The addition is just the vector addition of the linear space \mathbb{R}^2 , and the multiplication is commutative, associative, distributive, and $1 + \varepsilon \cdot 0$ serves as a unit. This means that \mathbb{D} is a commutative ring with unit. There is, however, not always a multiplicative inverse. Especially the dual unit ε fulfills the relation

$$\varepsilon^2 = 0 \quad (2.43)$$

and is therefore a divisor of zero.

Example 2.3.1. We compute $(a + \varepsilon b)^k$: The binomial formula gives $a^k + \varepsilon k a^{k-1} b + \varepsilon^2(\dots)$, so the result is $a^k + \varepsilon k a^{k-1} b$. \diamond

Analytic Functions of Dual Arguments

If f is a real analytic function, i.e., it is represented by a power series $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ which converges in some interval, then we can extend its definition to the ring of dual numbers by letting

$$\begin{aligned} f(x + \varepsilon y) &= \sum_{k=0}^{\infty} a_k (x + \varepsilon y - x_0)^k \\ &= \sum_{k=0}^{\infty} a_k (x - x_0)^k + \varepsilon \sum_{k=0}^{\infty} k a_k (x - x_0)^{k-1} y \\ &= f(x) + \varepsilon y f'(x). \end{aligned} \quad (2.44)$$

Example 2.3.2. We compute $\sin(x + \varepsilon y) = \sin(x) + \varepsilon y \cos(x)$, $\cos(x + \varepsilon y) = \cos(x) - \varepsilon y \sin(x)$, $\exp(x + \varepsilon y) = \exp(x) - \varepsilon y \exp(x)$. Also the result of the previous example is found again: $(a + \varepsilon b)^k = a^k + \varepsilon k a^{k-1} b$. \diamond

Oriented Lines

We consider Euclidean 3-space and use Cartesian coordinates.

Definition. An oriented line is a line of Euclidean space together with a unit vector parallel to it. An oriented line is also called a spear.

There are exactly two possibilities for this unit vector: Each line defines two oriented lines. Oriented lines will be denoted by the symbol \vec{L} .

If a line G is defined by a point $p \in G$ and a unit direction vector g , its Plücker coordinates are

$$(g, \bar{g}) = (g, p \times g). \quad (2.45)$$

On the other hand, if $(g, \bar{g})\mathbb{R}$ are the Plücker coordinates of G , then

$$\frac{\pm 1}{\|g\|} (g, \bar{g}) \quad (2.46)$$

are the two possible Plücker coordinate vectors of G such that the first part is a unit vector. They are called *normalized Plücker coordinates* (see Fig. 2.2). They can be used to describe lines of Euclidean space only.

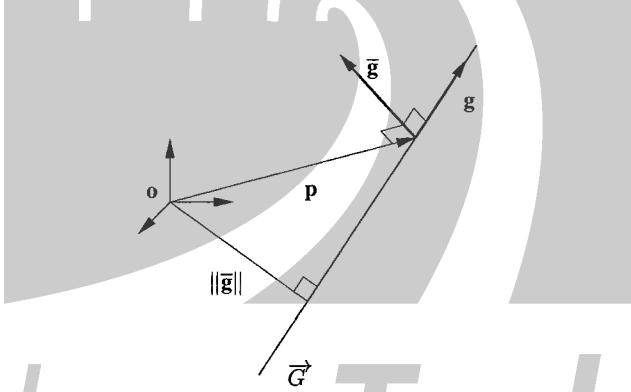


Fig. 2.2. Normalized Plücker coordinates.

The vector \bar{g} is the *moment vector* of a unit force on L with respect to the origin O . The equation $\bar{g} = p \times g$ (see Eq. (2.45)) gives a simple geometric interpretation of the moment vector: It is orthogonal to the plane $OG\vec{G}$. Its length equals the distance of \vec{G} to O , and the frame p, g, \bar{g} is positively oriented.

The Study Sphere

E. Study [192] first combined the two parts of the normalized Plücker coordinate vector of a line into a *dual vector* by letting

$$\hat{g} = g + \varepsilon \bar{g}. \quad (2.47)$$

This vector is an element of $\mathbb{D}^3 = \mathbb{R}^3 + \varepsilon\mathbb{R}^3$. We define a \mathbb{D} -valued bilinear form in \mathbb{D}^3 by letting

$$\widehat{\mathbf{x}} \cdot \widehat{\mathbf{y}} = \mathbf{x} \cdot \mathbf{y} + \varepsilon(\mathbf{x} \cdot \bar{\mathbf{y}} + \bar{\mathbf{x}} \cdot \mathbf{y}). \quad (2.48)$$

In \mathbb{D}^3 , which has dimension 6 as a linear space over the reals, this bilinear form defines a kind of degenerate scalar product of rank 3. It induces a ‘norm’, which will be denoted by $\|\cdot\|$.

We compute the norm of the vector $\widehat{\mathbf{g}}$ of Equ. (2.47) and see that

$$\|\widehat{\mathbf{g}}\|^2 = \widehat{\mathbf{g}} \cdot \widehat{\mathbf{g}} = \|\mathbf{g}\|^2 + 2\varepsilon \mathbf{g} \cdot \bar{\mathbf{g}} = 1, \quad (2.49)$$

because $\|\mathbf{g}\| = 1$ and because of Equ. (2.3).

Definition. *The mapping which assigns to an oriented line of Euclidean space the dual vector $\widehat{\mathbf{g}} = \mathbf{g} + \varepsilon\bar{\mathbf{g}}$, where $(\mathbf{g}, \bar{\mathbf{g}})$ are its normalized Plücker coordinates, is called the Study mapping. Its image is called the Study model of oriented lines of E^3 .*

Because of Equ. (2.49), the image of the Study mapping is also called the *Study sphere*.

The Dual Angle of Spears

The scalar product of two dual unit vectors $\widehat{\mathbf{g}}, \widehat{\mathbf{h}}$ has a simple geometric meaning in terms of the spears they represent:

We define the *distance* $d(G, H)$ between two lines G, H in E^3 as the smallest distance between points $\mathbf{a} \in G$, and $\mathbf{b} \in H$. The minimum value is attained if \mathbf{a}, \mathbf{b} are the points where the common perpendicular of G, H meets G and H , respectively (cf. Fig. 2.3).

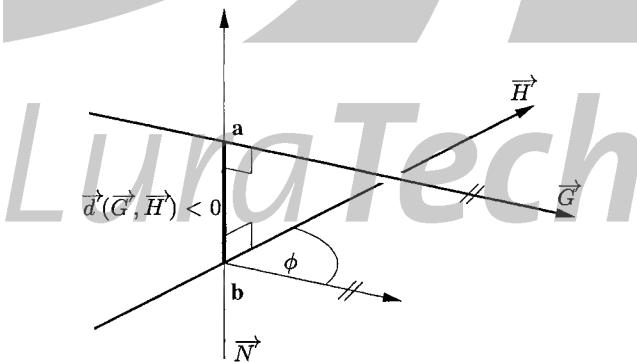


Fig. 2.3. Common perpendicular of two spears.

The common perpendicular of an ordered pair $(\overrightarrow{G}, \overrightarrow{H})$ of spears can be given an orientation: If $G = (\mathbf{g}, \bar{\mathbf{g}})$ and $H = (\mathbf{h}, \bar{\mathbf{h}})$ are normalized Plücker coordinates, the common perpendicular N is given an orientation by the vector $\mathbf{g} \times \mathbf{h}$.

Definition. Consider the oriented perpendicular \vec{N} of spears \vec{G}, \vec{H} . The distance $d(G, H)$ is attained at points $\mathbf{a} \in G$ and $\mathbf{b} \in H$. Then, the oriented distance is defined by

$$\vec{d}(\vec{G}, \vec{H}) = \sigma \cdot d(G, H), \quad (2.50)$$

where $\sigma = 1$ if $\mathbf{b} - \mathbf{a}$ has the same orientation as \vec{N} , and $\sigma = -1$ otherwise (cf. Fig. 2.3).

Definition. The dual angle of two spears \vec{G}, \vec{H} is defined by

$$\hat{\alpha}(\vec{G}, \vec{H}) = \alpha(\vec{G}, \vec{H}) + \varepsilon \vec{d}(\vec{G}, \vec{H}). \quad (2.51)$$

Lemma 2.3.1. The scalar product (2.48) in \mathbb{D}^3 is a Euclidean invariant. If \vec{G}, \vec{H} are two lines whose Study images are $\hat{\mathbf{g}}, \hat{\mathbf{h}}$, then there is the equation

$$\hat{\mathbf{g}} \cdot \hat{\mathbf{h}} = \cos \hat{\phi} = \cos \phi - \varepsilon d \sin \phi, \quad (2.52)$$

where $\phi = \alpha(\vec{G}, \vec{H})$, $\hat{\phi} = \hat{\alpha}(\vec{G}, \vec{H})$, and $d = \vec{d}(\vec{G}, \vec{H})$.

Proof. Assume that \mathbf{a}, \mathbf{b} are the points where their common perpendicular meets the lines G, H . We compute

$$\hat{\mathbf{g}} \cdot \hat{\mathbf{h}} = \mathbf{g} \cdot \mathbf{h} + \varepsilon(\mathbf{g} \cdot \bar{\mathbf{h}} + \bar{\mathbf{g}} \cdot \mathbf{h}) = \cos \alpha(\mathbf{g}, \mathbf{h}) + \varepsilon \Omega((\mathbf{g}, \bar{\mathbf{g}}), (\mathbf{h}, \bar{\mathbf{h}})). \quad (2.53)$$

and further, using Equ. (2.45) we deduce that

$$\begin{aligned} \mathbf{g} \cdot \bar{\mathbf{h}} + \bar{\mathbf{g}} \cdot \mathbf{h} &= \mathbf{g} \cdot (\mathbf{b} \times \mathbf{h}) + (\mathbf{a} \times \mathbf{g}) \cdot \mathbf{h} \\ &= -\det(\mathbf{b}, \mathbf{g}, \mathbf{h}) + \det(\mathbf{a}, \mathbf{g}, \mathbf{h}) = -\det(\mathbf{b} - \mathbf{a}, \mathbf{g}, \mathbf{h}) \\ &= -d \sin \phi. \end{aligned}$$

Equations (2.48) and (2.53) now show the result. \square

This shows again that two lines of Euclidean space with Plücker coordinates $(\mathbf{g}, \bar{\mathbf{g}})$ and $(\mathbf{h}, \bar{\mathbf{h}})$ intersect if and only if $\mathbf{g} \cdot \bar{\mathbf{h}} + \mathbf{h} \cdot \bar{\mathbf{g}} = 0$.

Lemma 2.3.1 has a remarkable consequence: We call the group of bijective mappings of \mathbb{D}^3 onto itself which keep the dual scalar product (2.48) invariant, the group of *dual spherical motions*.

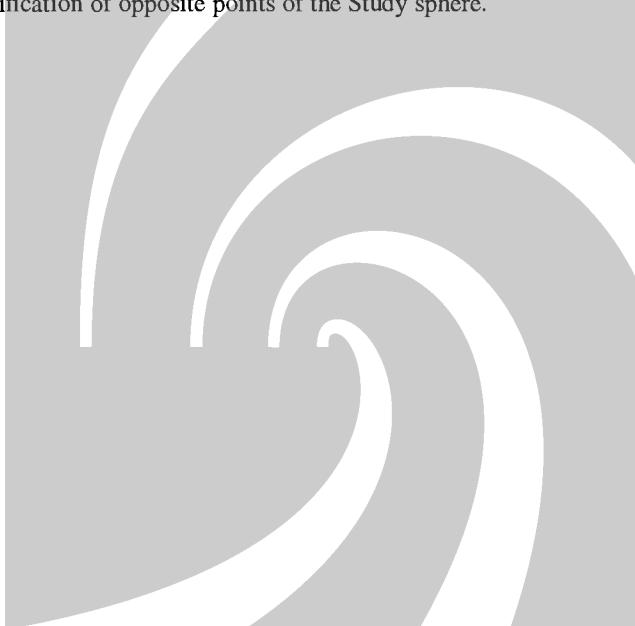
Theorem 2.3.2. The Study mapping provides an isomorphism between the group of motions of E^3 and the group of dual spherical motions.

Proof. Lemma 2.3.1 shows that a motion of E^3 defines, via the Study mapping, a dual spherical motion. The converse is also clear: Invariance of the dual scalar product means invariance of distance and angle, and such a bijection in the set of lines of Euclidean space must be induced from a Euclidean motion. \square

Remark 2.3.1. In this sense, kinematics of Euclidean line space is equivalent to dual spherical kinematics, a fact that can be effectively used to build up the theory of spatial kinematics and to perform practical calculations for spatial mechanisms.

This correspondence between a motion of E^3 and a dual spherical motion is a 'kinematic mapping' in the sense defined in Chap. 8. \diamond

Any line G carries two spears \vec{G}_1, \vec{G}_2 , which are mapped points $\hat{\mathbf{g}}_1 = -\hat{\mathbf{g}}_2$ of D^3 . Thus a Study model of the unoriented lines of Euclidean space can be constructed by identification of opposite points of the Study sphere.



LuraTech

www.luratech.com

3. Linear Complexes

Having identified the set of lines with the set of points of a certain quadric contained in projective five-space (the Klein quadric), we are able to apply concepts of projective geometry to the set of lines. The basic elements in a projective space are its subspaces, ‘linear manifolds’ of points, defined by linear equations in the points’ homogeneous coordinates. Accordingly in this chapter we are going to study ‘linear manifolds’ of lines. These correspond to intersections of the Klein quadric with subspaces, and they are defined by linear equations in the lines’ Plücker coordinates.

The basic object will be the so-called linear line complex, a three-dimensional linear manifold of lines defined by one linear equation. It is on the one hand connected to null polarities of projective three-space, and has on the other hand several nice interpretations in Euclidean geometry, kinematics and statics.

Other linear manifolds of lower dimension are intersections of complexes. Unlike projective subspaces of the same dimension, all of which ‘look the same’, i.e., are projectively equivalent, linear manifolds of lines exhibit a slightly more complicated behaviour.

3.1 The Structure of a Linear Complex

We continue the discussion started in Sec. 1.1.5, where we considered correlations, which are projective mappings of projective space onto its dual. A correlation π of the form (1.29) in P^3 maps a point P to a plane $P\pi$. The points of $P\pi$ are called conjugate to P . If the conjugacy relation is *symmetric* (i.e., $Q \in P\pi$ implies $P \in Q\pi$) then π is called self-adjoint. In Sec. 1.1.5 we showed that in this case the matrix C of Equ. (1.29) must be either symmetric or skew-symmetric. In Sec. 1.1.5 we discussed the symmetric case in greater detail. Now we resume the skew-symmetric case.

3.1.1 Linear Complexes and Null Polarities in Projective Space

According to Equ. (1.29), a null polarity π has the form $x\mathbb{R} \mapsto \mathbb{R}(C \cdot x)$ with a skew-symmetric matrix C . We already know (cf. Equ. (1.32)) that all points are self-conjugate, i.e., are contained in their respective image plane $P\pi$. The duality δ associated with this polarity then maps $P\pi$ to P again.

For simplicity, we will use the symbol π also for the duality δ associated with π . This means we write $L\pi$ and $\varepsilon\pi$ instead of $L\delta$ and $\varepsilon\delta$ for a line L and a plane ε .

Definition. If π is a null polarity, then $P\pi$ is called the null plane of P , and P is called the null point of $P\pi$. An invariant line $L = L\pi$ is called null line.

We are going to study the set of all null lines of a null polarity π :

Lemma 3.1.1. If P, Q are conjugate points, then $P \vee Q$ is a null line, and all null lines can be represented in this way.

Proof. We collect incidence relations: $P \in P\pi, Q \in Q\pi$ because π is a null polarity. $P \in Q\pi, Q \in P\pi$, because P and Q are conjugate. $(P \vee Q)\pi = P\pi \cap Q\pi$, which holds true for all correlations. Thus $(P \vee Q)\pi$ contains both P, Q and $(P \vee Q)\pi = P \vee Q$.

Conversely, assume that L is a null line, and choose $P, Q \in L$. Then $L = L\pi, L \subset P\pi$, and $L \subset Q\pi$ show that P, Q are conjugate. \square

There is a simple characterization of null lines in terms of their Plücker coordinates:

Lemma 3.1.2. Consider the null polarity π given by Equ. (1.29), denote the entries in the matrix C by c_{ij} , and let

$$\bar{\mathbf{c}} = (c_{01}, c_{02}, c_{03}), \quad \mathbf{c} = (c_{23}, c_{31}, c_{12}). \quad (3.1)$$

Assume the line L has Plücker coordinates $L\mathbb{R} = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$. Then

$$L = L\pi \iff \Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{l}, \bar{\mathbf{l}})) = \bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0. \quad (3.2)$$

Proof. We assume that L is spanned by points $x\mathbb{R}$ and $y\mathbb{R}$, and that $(\mathbf{l}, \bar{\mathbf{l}}) = x \wedge y$. Conjugacy of $x\mathbb{R}$ and $y\mathbb{R}$ is expressed by

$$\mathbf{x}^T \cdot C \cdot \mathbf{y} = \sum_{i < k} c_{ik} (x_i y_k - x_k y_i) = 0.$$

By the definition of exterior multiplication, this is equivalent to $\bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0$. \square

If a linear complex \mathcal{C} is defined by Equ. (3.2), then we call $\mathbf{C} = (\mathbf{c}, \bar{\mathbf{c}})$ the *homogeneous coordinate vector* of \mathcal{C} .

Remark 3.1.1. Note that we did *not* define $\bar{\mathbf{c}} = (c_{23}, c_{31}, c_{12})$, but $\bar{\mathbf{c}} = (c_{01}, c_{02}, c_{03})$. \diamond

Lemma 3.1.2 shows that the null lines are characterized by a *linear equation* in their Plücker coordinates. We give a name to such a set of lines:

Definition. A set \mathcal{C} of lines defined by a linear homogeneous equation in Plücker coordinates is called a *linear line complex*, or short, *linear complex*.

We temporarily denote the coefficient matrix of the bilinear form Ω by K (see Equ. (2.28)). Any linear homogeneous equation of the form $U^T \cdot X = 0$ in \mathbb{R}^6 is equivalent to $(K^{-T} \cdot U)^T \cdot K \cdot X = \Omega(K^{-T} \cdot U, X) = 0$, so it is no restriction to write it in the form $\Omega(C, X) = 0$.

Definition. A linear line complex C defined by the equation $\Omega(C, X) = 0$ is called singular, if C represents a line, i.e., $\Omega_q(C) = 0$; otherwise it is called regular.

The linear complexes may be characterized as follows.

Theorem 3.1.3. A regular linear complex consists of the null lines of some null polarity. A singular linear complex consists of all lines intersecting a fixed line (called the axis of the singular complex).

Proof. If (c, \bar{c}) represents a line L , Equ. (2.26) shows that C consists of all lines which meet L .

If (c, \bar{c}) represents no line (i.e., $c \cdot \bar{c} \neq 0$), we define a skew-symmetric matrix $C = (c_{ik})$ by letting $(c_{01}, c_{02}, c_{03}) = \bar{c}$ and $(c_{23}, c_{31}, c_{12}) = c$. Its determinant equals $(c_{01}c_{23} + c_{02}c_{31} + c_{03}c_{12})^2 = (c \cdot \bar{c})^2 \neq 0$, so the matrix C actually defines a null polarity, and Lemma 3.1.2 shows the result. \square

The Complex Cone and Planar Complex Curve

We will briefly give the definitions of some objects which will be studied in greater detail in Chap. 7. The set of lines defined by a 3-dimensional algebraic variety contained in the Klein quadric is called an *algebraic line complex*. The set of lines defined by a 3-surface contained in M_2^4 is called a *smooth line complex*. A linear line complex is of dimension three as an algebraic variety (see also Sec. 3.1.3).

If C is a line complex, the lines of C concurrent in a point P are called the *complex cone* of P , and the lines contained in a plane ε define, by planar duality, a curve in ε , which is called the *planar complex curve*. In the cases where the dimensions of these objects are different from 1, we call them *degenerate*. If the complex curve is not degenerate, its lines may be the envelope of an ‘ordinary’ curve in ε .

In the case of a regular linear complex, both the complex cone and the planar complex curve are line pencils: The pencil incident with a point and its null plane.

If C is a singular line complex consisting of all lines incident with the axis L , there are two possibilities: If the point P is not incident with L , then the situation is like in a regular linear complex. If $P \in L$, the degenerate complex cone is a bundle. Analogously, if a plane ε does not contain L , the planar complex curve is a pencil. If L is incident with ε , it equals the entire field of lines and is therefore degenerate.

Computations in Homogeneous Coordinates

Consider a regular linear complex, consisting of the null lines of the null polarity π , which is characterized, uniquely up to a scalar factor, either by its coordinate matrix C , or the vector $(c, \bar{c}) \in \mathbb{R}^6$ (cf. Lemma 3.1.2).

If we use homogeneous Cartesian coordinates and the notation introduced in Sec. 2.1.2, then Equ. (1.29) assumes the form

$$((x_0, \mathbf{x})\mathbb{R})\pi = \mathbb{R}(\mathbf{x} \cdot \bar{\mathbf{c}}, -x_0\bar{\mathbf{c}} + \mathbf{x} \times \mathbf{c}). \quad (3.3)$$

The mapping of planes is described, in plane coordinates, by the inverse transpose of the matrix C . As C is skew-symmetric, C^{-T} is given by

$$C^{-T} = \frac{1}{\mathbf{c} \cdot \bar{\mathbf{c}}} \begin{bmatrix} c_{23} & c_{31} & c_{12} \\ -c_{23} & c_{03} & -c_{02} \\ -c_{31} & -c_{03} & c_{01} \\ -c_{12} & c_{02} & -c_{01} \end{bmatrix},$$

and apart from a scalar factor, the passage from C to its transpose inverse is the same as the passage from Plücker coordinates to axis coordinates. Thus the null point of the plane $\mathbb{R}(u_0, \mathbf{u})$ is found by

$$(\mathbb{R}(u_0, \mathbf{u}))\pi = (\mathbf{u} \cdot \mathbf{c}, -u_0\mathbf{c} + \mathbf{u} \times \bar{\mathbf{c}})\mathbb{R}. \quad (3.4)$$

Remark 3.1.2. For the singular linear complex consisting of all lines incident with the axis L , the plane $P \vee L$ plays the role of the null plane of P ; and the point $\varepsilon \cap L$ plays the role of the null point of ε . If $L = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$, then they are computed with formulae (2.16) and (2.17), which are completely analogous to formulae (3.3) and (3.4), respectively. \diamond

Lemma 3.1.4. *The image of a line $L = \mathbb{L}\mathbb{R} = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ under the null polarity $(\mathbf{c}, \bar{\mathbf{c}})$ is given by*

$$L\pi = (\mathbf{l}', \bar{\mathbf{l}'})\mathbb{R} = (\mathbf{c} \cdot \bar{\mathbf{c}})(\mathbf{l}, \bar{\mathbf{l}}) - (\mathbf{c} \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}} \cdot \mathbf{l})(\mathbf{c}, \bar{\mathbf{c}}) = \Omega_q(C)L - \Omega(C, L)C. \quad (3.5)$$

The pair $L, L\pi$ of lines is called a *reciprocal pair*.

Proof. The mapping $(\mathbf{l}, \bar{\mathbf{l}}) \mapsto (\mathbf{l}', \bar{\mathbf{l}'})$ is a regular linear mapping and therefore induces a projective automorphism of P^5 . By Prop. 1.1.6, it coincides with the projective automorphism $L\gamma \mapsto L\pi\gamma$ induced by π , if it coincides for a fundamental set (which has seven points). This condition is fulfilled, as all of π 's null lines are obviously fixed by both mappings. \square

Projective Equivalence of Linear Complexes

We want to know how many different regular line complexes exist in P^3 , where ‘different’ means nonequivalent with respect to the group of projective automorphisms. The answer turns out to be very simple:

Theorem 3.1.5. *All regular linear line complexes in P^3 are projectively equivalent.*

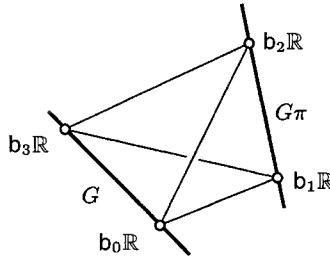


Fig. 3.1. Projective coordinate system for normal form of linear line complex.

Proof. We are going to find, for all linear complexes \mathcal{C} , a projective coordinate system, such that \mathcal{C} has the equation $l_{03} + l_{12} = 0$ in the Plücker coordinates defined by this coordinate system. This will show the result, because \mathcal{C} 's equation is then independent of \mathcal{C} .

Consider the null polarity π attached to \mathcal{C} (see Fig. 3.1). Choose a line $G \notin \mathcal{C}$ and $b_0R, b_3R \in G$. It is left to the reader as an exercise to show that $G, G\pi$ must be skew. So we can choose $b_1R, b_2R \in G\pi$. The unit point is chosen arbitrarily.

In any such coordinate system the line $b_0R \vee b_1R$ has the Plücker coordinates $(1 : 0 : 0 : 0 : 0 : 0)$. It is a null line by Lemma 3.1.1. The coordinate matrix C of π must fulfill Lemma 3.1.2, which shows that $c_{01} = 0$. Analogously we derive $c_{02} = c_{31} = c_{23} = 0$. Then $\Omega(C, L) = c_{03}l_{03} + c_{12}l_{12}$ with $\Omega_q(C) = c_{03}c_{12} \neq 0$. With $p = c_{03}/c_{12}$, the equation of \mathcal{C} is given by

$$p \cdot l_{03} + l_{12} = 0.$$

If we change the unit point such that its new position has, in the previous coordinate system, coordinates $(p : 1 : 1 : 1)$, then coordinates transform according to $x_0 = px'_0$, $x_1 = x'_1$, $x_2 = x'_2$, $x_3 = x'_3$, and Plücker coordinates change according to $l_{03} = pl'_{03}$, $l_{12} = l'_{12}$. This shows that in the new coordinate system the complex has the equation $l_{03} + l_{12} = 0$. \square

3.1.2 Linear Complexes and Helical Motions in Euclidean Space

In this section we study linear complexes in Euclidean space E^3 . More precisely, we consider the projective extension P^3 of E^3 and look for special Euclidean properties of linear complexes.

Definition. *The family of Euclidean congruence transformations, which is parameterized by the real parameter t , and which, in a suitable Cartesian coordinate system, maps points according to*

$$\mathbf{x} \mapsto \mathbf{x}(t) = \begin{bmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ p \cdot t \end{bmatrix}, \quad (3.6)$$

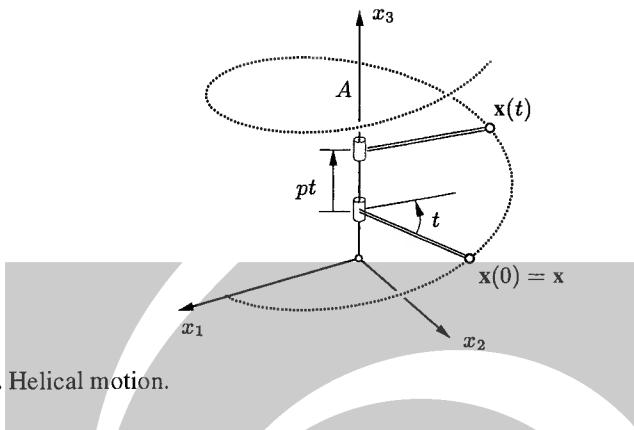


Fig. 3.2. Helical motion.

is called a one-parameter group of helical motions (or a uniform helical motion) if $p \neq 0$. If $p = 0$, it is called a one-parameter group of rotations, or a uniform rotation. A uniform helical motion is the superposition of a uniform rotation and a translation by the vector $(0, 0, pt)$. The line $x_1 = x_2 = 0$ is called its axis, and p is called its pitch (see Fig. 3.2).

Assume that the pitch p is nonzero. If the point \mathbf{x} is situated on the axis, its path coincides with the axis. The trajectories of the other points are *helices*. A helical motion is called *right-handed*, if $p > 0$, and *left-handed*, if $p < 0$.

The Path Normals of a Helical Motion

The field of velocity vectors of a uniform helical motion is independent of t in the sense that any time a point has traveled to a position $\mathbf{x}(t) = \mathbf{x}_0$, its velocity vector \mathbf{v} is the same: If $\mathbf{x}(t) = (x, y, z)$, then

$$\mathbf{v} = \frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} -y \\ x \\ p \end{bmatrix}. \quad (3.7)$$

All lines incident with \mathbf{x}_0 and orthogonal to \mathbf{v} are path normals (see Fig. 3.3). The following theorem shows a connection between a linear complex and the path normals of a helical motion:

Theorem 3.1.6. *The path normals of a helical motion of E^3 are precisely the non-ideal lines of a regular linear complex \mathcal{C} of P^3 , and vice versa.*

Proof. We compute the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ of a path normal from Equ. (3.6) and (3.7): This normal is spanned by $(1, \mathbf{x})$ and $(0, \mathbf{n})$ with \mathbf{n} orthogonal to \mathbf{v} , so $(\mathbf{l}, \bar{\mathbf{l}}) = (\mathbf{n}, \mathbf{x} \times \mathbf{n})$, and $\mathbf{v} \cdot \mathbf{l} = -yl_{01} + xl_{02} + pl_{03} = 0$. Further we evaluate $l_{12} = -yl_{01} + xl_{02}$ from $\bar{\mathbf{l}} = \mathbf{x} \times \mathbf{l}$. This shows that $pl_{03} + l_{12} = 0$, so all path normals are contained in the linear complex \mathcal{C} with equation

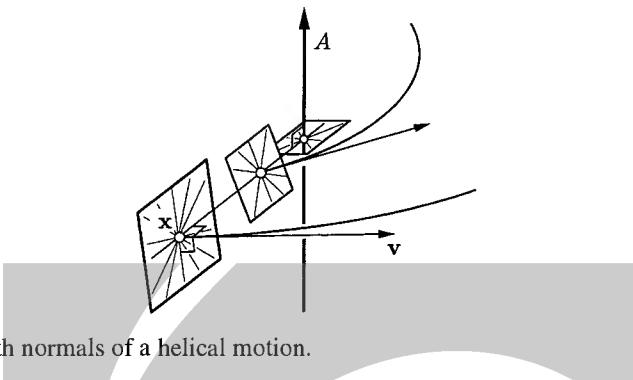


Fig. 3.3. Path normals of a helical motion.

$$pl_{03} + l_{12} = 0. \quad (3.8)$$

All points of E^3 are incident with a pencil of path normals, as well as with a pencil of non-ideal lines of \mathcal{C} , so these two sets actually coincide.

The proof of the converse statement follows the proof of Th. 3.1.5. We consider a linear line complex \mathcal{C} and its associated null polarity π . The ideal plane is denoted by ω . Let $A_u = \omega\pi$, and consider the ideal line L_u orthogonal to the point A_u (for orthogonality of points in ω , see Ex. 1.1.40). Then let $A = L_u\pi$, and choose a Cartesian coordinate system with origin $O \in A$, A as z -axis. The associated projective coordinate system $b_0\mathbb{R}, \dots, b_3\mathbb{R}, e\mathbb{R}$ then has the property $b_0\mathbb{R} \in A$, $b_3\mathbb{R} \in A$, $b_1\mathbb{R} \in L_u = A\pi$, $b_2\mathbb{R} \in L_u = A\pi$. According to Th. 3.1.5, the complex \mathcal{C} has the equation $pl_{03} + l_{12} = 0$ in this coordinate system. The first part of this proof shows that Equ. (3.6) defines a helical motion whose path normals agree with the non-ideal lines of \mathcal{C} . \square

Remark 3.1.3. In order to avoid complicated formulations, also the non-ideal part of a linear complex \mathcal{C} is called a linear complex. The missing lines are the pencil incident with the ideal plane's null point $\omega\pi$, where π is the null polarity associated with \mathcal{C} . \diamond

Remark 3.1.4. If the points of Euclidean space undergo a helical motion, each has its pencil of path normals. So we could expect that there is a four-parameter family of path normals. Th. 3.1.6 however shows that the set of path normals is a line complex, which can be identified with a three-dimensional planar section of the Klein quadric (see Sec. 3.1.3). It should be pointed out that other one-parameter transformation groups do not have this property. An example are uniform spiraloid motions (the family of similarity transformations which is automorphic for a snail's shell), where every line is a path normal for some point of space. \diamond

Remark 3.1.5. A linear complex is an entity of projective geometry: A projective automorphism transforms a linear complex into an object of the same type. It is therefore rather surprising that it has a Euclidean interpretation as given by Th. 3.1.6, which also shows that a projective automorphism transforms the set of path normals

of a helical motion to the set of path normals of another helical motion — but neither the axes nor the point paths are mapped onto each other. \diamond

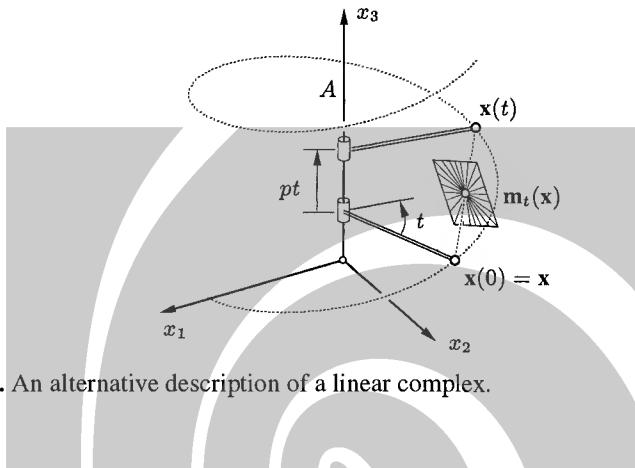


Fig. 3.4. An alternative description of a linear complex.

Remark 3.1.6. Consider the helical motion $\mathbf{x} \mapsto \mathbf{x}(t)$ of Equ. (3.6). Fix a time t , and for all \mathbf{x} consider the midpoint $\mathbf{m}_t(\mathbf{x}) = \frac{1}{2}(\mathbf{x} + \mathbf{x}(t))$, and the lines incident with $\mathbf{m}_t(\mathbf{x})$ and orthogonal to the vector $\mathbf{x}(t) - \mathbf{x}$ (see Fig. 3.4).

It is left as an exercise to the reader to show that this set of lines is a linear complex. \diamond

Null Polarities and Helical Motions

Th. 3.1.6 allows a nice visualization of a null polarity:

Corollary 3.1.7. *To every null polarity π in the projective extension of E^3 there exists exactly one helical motion with axis A such that π maps points in E^3 to their path normal planes, and the ideal point of A to the ideal plane.*

Proof. By Th. 3.1.3, the set of null lines of π is a regular linear complex \mathcal{C} . The result now follows from Th. 3.1.6. \square

Singular Linear Complexes in Euclidean Space

In projective space, there is only one type of singular linear complex: It consists of all lines which intersect a given axis L . In E^3 , this axis may be a proper Euclidean line or it may be an ideal line. In the former case, the complex consists of the path normals of a uniform *rotation*, parametrized by Equ. (3.6) with $p = 0$. This is shown in the same way as the case $p \neq 0$.

In the latter case all lines of the complex \mathcal{C} intersect an ideal line L_u , which means that they are parallel to a fixed plane ε (which then contains L_u). Alternatively we can say that the lines of \mathcal{C} are orthogonal to a line A (which is then orthogonal to ε). They are therefore path normals of a uniform *translation*. The pitch is said to be infinite.

Equation (3.8) of a regular linear complex in E^3 is a linear combination of the singular linear complex $l_{12} = 0$ associated with the rotation about the helical axis A and the singular linear complex $l_{03} = 0$ associated with the translation along A . The pitch p appears as quotient of coefficients in this combination. An illustration of the path normals in these singular cases can be seen in Fig. 3.5.

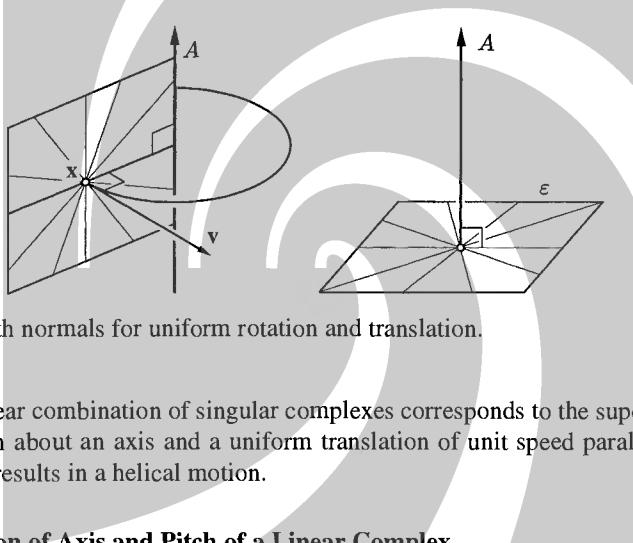


Fig. 3.5. Path normals for uniform rotation and translation.

This linear combination of singular complexes corresponds to the superposition of a rotation about an axis and a uniform translation of unit speed parallel to this axis which results in a helical motion.

Computation of Axis and Pitch of a Linear Complex

We want to compute the axis and the pitch of the helical motion which corresponds to a linear complex \mathcal{C} according to Th. 3.1.6. We call these the axis and the pitch of the linear complex, because they are uniquely defined by \mathcal{C} . We first show a lemma which exhibits a *Euclidean invariant* of a linear complex.

Lemma 3.1.8. *Assume that a linear complex \mathcal{C} in E^3 is defined by the equation $\bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0$. Then the value of $\mathbf{c} \cdot \bar{\mathbf{c}} / \mathbf{c}^2$ does not change if we apply Euclidean congruence transformations.*

Proof. We have to show that this value does not change if we apply a rotation about the origin, a reflection, and a translation. In a Cartesian coordinate system the first two cases can be written in the form $\alpha : \mathbf{x} \mapsto A \cdot \mathbf{x}$, with an orthogonal matrix A .

Assume a line $L = (0, \mathbf{g})\mathbb{R} \vee (1, \mathbf{x})\mathbb{R}$. It has Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}}) = (\mathbf{g}, \mathbf{x} \times \mathbf{g})$. Its image $\alpha(L)$ has Plücker coordinates $(A\mathbf{g}, (A\mathbf{x}) \times (A\mathbf{g})) = (A\mathbf{l}, A\bar{\mathbf{l}})$. The α -images of lines in the complex obviously fulfill the equation $(A\bar{\mathbf{c}}) \cdot (A\mathbf{l}) + (A\mathbf{c}) \cdot (A\bar{\mathbf{l}}) = 0$.

$(A\bar{l}) = 0$, the image of the complex has coordinates $(A \cdot \mathbf{c}, A \cdot \bar{\mathbf{c}})$. Further, $A\mathbf{c} \cdot A\bar{\mathbf{c}}/(A\mathbf{c})^2 = \mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2$.

A translation $\mathbf{x} \mapsto \mathbf{x} + \mathbf{u}$ transforms L in the following way: $(l, \bar{l}) \mapsto (g, (\mathbf{x} + \mathbf{u}) \times g) = (l, \bar{l} + \mathbf{u} \times l)$. The α -images of lines in the complex fulfill the equation $(\bar{\mathbf{c}} - \mathbf{c} \times \mathbf{u}) \cdot l + \mathbf{c} \cdot (\bar{l} + \mathbf{u} \times l) = 0$, because $(\mathbf{c} \times \mathbf{u}) \cdot l = \det(l, \mathbf{c}, \mathbf{u}) = \mathbf{c} \cdot (\mathbf{u} \times l)$, and $(\bar{\mathbf{c}} - \mathbf{c} \times \mathbf{u}) \cdot \mathbf{c}/\mathbf{c}^2 = \mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2$. Thus the lemma is proved. \square

Theorem 3.1.9. *Pitch p and Plücker coordinates $(\mathbf{a}, \bar{\mathbf{a}})\mathbb{R}$ of the axis A of a regular linear complex C with equation $\bar{\mathbf{c}} \cdot l + \mathbf{c} \cdot \bar{l} = 0$ in E^3 are given by*

$$p = \mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2, \quad (\mathbf{a}, \bar{\mathbf{a}}) = (\mathbf{c}, \bar{\mathbf{c}} - p\mathbf{c}). \quad (3.9)$$

Proof. We denote the null polarity associated with C by π . The ideal plane $\mathbb{R}(1, \mathbf{o})$ is denoted by ω . By (3.4), $A_u = \omega\pi = (0, \mathbf{c})\mathbb{R}$, which shows that A is parallel to \mathbf{c} . Consider the plane $\varepsilon = \mathbb{R}(0, \mathbf{c})$ which contains the origin and is orthogonal to A . Then, again by (3.4), $\varepsilon\pi = (\mathbf{c}^2, \mathbf{c} \times \bar{\mathbf{c}})\mathbb{R}$. According to the proof of Th. 3.1.6, the ideal line L_u orthogonal to A_u equals $\varepsilon \cap \omega$, so $A = L_u\pi = \varepsilon\pi \vee \omega\pi$. We compute its Plücker coordinates with Equ. (2.13):

$$(\mathbf{a}, \bar{\mathbf{a}})\mathbb{R} = (-\mathbf{c}^2\mathbf{c}, \mathbf{c} \times (\mathbf{c} \times \bar{\mathbf{c}}))\mathbb{R} = (-\mathbf{c}^2\mathbf{c}, (\mathbf{c} \cdot \bar{\mathbf{c}})\mathbf{c} - \mathbf{c}^2\bar{\mathbf{c}})\mathbb{R} = (\mathbf{c}, \bar{\mathbf{c}} - (\mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2)\mathbf{c})\mathbb{R}.$$

To show that $p = \mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2$, we first note that this is true in a Cartesian coordinate system where C is given by (3.8), because then $\mathbf{c} = (0, 0, 1)$, $\bar{\mathbf{c}} = (0, 0, p)$. By Lemma 3.1.8, $\mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2$ is a Euclidean invariant, and can therefore be computed in any Cartesian coordinate system. \square

Remark 3.1.7. The pitch is ‘the only’ numerical Euclidean invariant of a regular linear complex, because all complexes which have the same pitch can be transformed into each other by Euclidean congruence transformations. \diamond

Example 3.1.1. We consider the null polarity $\pi : (x_0, x_1, x_2, x_3)\mathbb{R} \mapsto \mathbb{R}(x_3, x_2, -x_1, -x_0)$. The nonzero entries in its matrix are $c_{03} = -c_{30} = 1$, $c_{12} = -c_{21} = 1$. Thus the equation of the linear complex of null lines has the equation $\bar{\mathbf{c}} \cdot l + \mathbf{c} \cdot \bar{l} = 0$, if we let $\mathbf{c} = \bar{\mathbf{c}} = (0, 0, 1)$, according to Lemma 3.1.2. The pitch of the helical motion associated with π equals $\mathbf{c} \cdot \bar{\mathbf{c}}/\mathbf{c}^2 = 1$, and its axis has Plücker coordinates $(\mathbf{c}, \bar{\mathbf{c}} - p\mathbf{c})\mathbb{R} = ((0, 0, 1), (0, 0, 0))\mathbb{R}$. \diamond

3.1.3 Linear Complexes in the Klein Model

We have defined a linear complex to be a set of lines whose Plücker coordinates fulfill a linear homogeneous equation. In the Klein Model, where the Plücker coordinates serve as point coordinates of points, and the lines of P^3 correspond to the points of the Klein quadric M_2^4 , this linear equation defines a hyperplane. In this way a linear complex is a hyperplanar section of the Klein quadric.

The Extended Klein Mapping

Recall the Klein mapping γ , which identifies the set of lines in projective three-space with the Klein quadric in P^5 . We still lack a geometric interpretation of the remaining points of P^5 . It turns out that the points of P^5 can be identified with linear complexes, such that the singular ones correspond to points of the Klein quadric. We define:

Definition. *The extended Klein mapping, denoted by γ^* , transforms a linear complex defined by the equation (3.2) to the point $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$*

Proposition 3.1.10. *If a linear complex C is singular with axis L , then $C\gamma^* = L\gamma$. If it is regular, then the set $C\gamma$ consists of all points of the Klein quadric conjugate to $C\gamma^*$.*

Proof. We have already established (see the proof of Th. 3.1.3) that the singular complex C consisting of all lines which meet L is mapped to $L\gamma$.

If C is regular, its equation can be written in the form $\Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{l}, \bar{\mathbf{l}})) = 0$, where $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} = C\gamma^*$ (see Equ. (3.2)). This equation expresses conjugacy of the points $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ and $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ with respect to M_2^4 (see the paragraph following Equ. (1.34)). \square

Corollary 3.1.11. *The extended Klein mapping is a bijection between linear complexes and points of P^5 , where the singular complexes correspond to the points of M_2^4 .*

Remark 3.1.8. Note that $C\gamma$ denotes the (ordinary) Klein image of a linear complex as a set of lines and is a subset of the Klein quadric. But the symbol $C\gamma^*$ means the point of P^5 which corresponds to C . \diamond

In the following, the symbol μ_2^4 shall denote the polarity of M_2^4 , and the duality associated with it. It maps a projective subspace of dimension k to its polar subspace of dimension $4 - k$. Prop. 3.1.10 can be expressed as

$$C\gamma = C\gamma^* \mu_2^4 \cap M_2^4. \quad (3.10)$$

A Regular Linear Complex as a Quadratic Variety

The Klein image $C\gamma$ of a regular linear complex C consists, by Prop. 3.1.10, of those points of M_2^4 which are contained in the polar hyperplane of $C\gamma^*$. This is expressed by Equ. (3.10). Thus $C\gamma$ is a quadric, because $C\gamma^* \not\subset M_2^4$. As a projective algebraic variety, it is of dimension three and degree 2, which is seen from the fact that it is defined by one quadratic equation in a four-dimensional hyperplane of P^5 .

$C\gamma$ is a quadric of the maximal possible index 1, because M_2^4 contains planes: $C\gamma$ contains a three-parameter family of lines which correspond to the three-parameter family of line pencils in C .

A null polarity π which fixes the lines of C induces an automorphic collineation $\tilde{\pi}$ of M_2^4 , as described by Equ. (2.31) and Th. 2.1.10:

Proposition 3.1.12. *The null polarity associated with a linear complex \mathcal{C} induces a projective reflection $\tilde{\pi}$ of M_2^4 which fixes $\mathcal{C}\gamma$ and $\mathcal{C}\gamma^*$.*

Proof. Since $\tilde{\pi}$ fixes the points of $\mathcal{C}\gamma$ and thus fixes its carrier hyperplane $[\mathcal{C}\gamma]$ pointwise, it must also (because it leaves M_2^4 invariant) fix the μ_2^4 -image of $[\mathcal{C}\gamma]$, i.e., the point $\mathcal{C}\gamma^*$. It is therefore an automorphic projective reflection of M_2^4 according to Th. 1.1.23. \square

Remark 3.1.9. We could directly verify Prop. 3.1.12 from formula (3.5): The Klein images $L\mathbb{R}, L'\mathbb{R}$ of a reciprocal pair L, L' fulfill $L' = \Omega_q(C)L - \Omega(C, L)C$, which shows that $L\mathbb{R}, L'\mathbb{R}, C\mathbb{R}$ are collinear. \diamond

A Singular Linear Complex as a Quadratic Variety

A *singular* linear complex consists of all points of M_2^4 conjugate to the point $\mathcal{C}\gamma^* \in M_2^4$, and is therefore a *tangential intersection* of M_2^4 . All singular linear complexes are projectively equivalent, so we consider the case $\mathcal{C}\gamma^* = C\mathbb{R} = (0, 0, 0, 1, 0, 0)\mathbb{R}$: The equations $\mathcal{C}\gamma : \Omega(C, L) = \Omega_q(L) = 0$ then read

$$l_{01} = 0, \quad l_{02}l_{31} + l_{03}l_{12} = 0. \quad (3.11)$$

Thus $\mathcal{C}\gamma$ contains the ruled quadric

$$l_{01} = l_{23} = 0, \quad l_{02}l_{31} + l_{03}l_{12} = 0. \quad (3.12)$$

If L' fulfills (3.12), then any linear combination

$$L = \lambda C + \mu L'$$

satisfies (3.11). This shows that $\mathcal{C}\gamma$ is a quadratic cone with vertex $\mathcal{C}\gamma^* = C\mathbb{R}$ and base quadric given by (3.12).

The base quadric contains two one-parameter families of lines, which shows that $\mathcal{C}\gamma$ contains two one-parameter families of planes. They correspond to bundles whose vertex is incident with the axis of the complex, and to fields whose carrier planes are incident with the axis.

As a projective algebraic variety, $\mathcal{C}\gamma$ is of dimension three and of degree two, as is seen from Equ. (3.11). The point $C\mathbb{R}$ is a singular point, all other points are regular.

Projective Automorphisms of the Klein Quadric

Any projective automorphism of P^3 or projective correlation onto its dual induces a projective automorphism of M_2^4 (see Th. 2.1.10). All lines L, L' of P^3 are projectively equivalent: This means that for any two points $L\gamma, L'\gamma$ of M_2^4 there is a projective automorphism of M_2^4 which maps $L\gamma$ to $L'\gamma$. Even more is true:

Proposition 3.1.13. *Any two admissible triples $L_1\gamma, L_2\gamma, L_3\gamma$ and $L'_1\gamma, L'_2\gamma, L'_3\gamma$ of points of M_2^4 can be mapped onto each other by a projective automorphism of M_2^4 , where ‘admissible triple’ means that each triple consists of pairwise skew lines.*

Proof. In Ex. 1.1.25 we showed that for two triples L_1, L_2, L_3 and L'_1, L'_2, L'_3 of pairwise skew lines there is a projective automorphism of P^3 which transforms L_i to L'_i . \square

Th. 3.1.5 shows that the regular linear complexes are projectively equivalent, which implies the following

Corollary 3.1.14. *Any two points \mathbb{CR} , $\mathbb{C}'\mathbb{R}$ of P^5 and not in M_2^4 can be mapped onto each other by a projective automorphism of M_2^4 .*

3.2 Linear Manifolds of Complexes

How many lines L_i of a linear complex C are necessary in order to determine it uniquely? The Klein image immediately tells us the answer: The span of the points $L_i\gamma$ must be a hyperplane. This means that *five independent lines* uniquely determine a linear complex. Independence of lines L_i means independence of the points $L_i\gamma$ in the sense of projective geometry, or linear independence of Plücker coordinates, which is the same.

Another question is the following: ‘How many’ linear complexes contain two given lines, three given lines, four given lines? This leads to the following definition:

Definition. *A set G of linear complexes is called a k -dimensional linear manifold of complexes, if $G\gamma^*$ is a k -dimensional projective subspace of P^5 .*

Definition. *The carrier $C(G)$ of a linear manifold G of linear complexes is the set of lines contained in all complexes $C \in G$. The axes of the singular complexes in G are its singular set $S(G)$.*

Lemma 3.2.1. *Assume that two linear manifolds G and G' of linear complexes are related by $G\gamma^* = G'\gamma^*\mu_2^4$, i.e., the polarity of the Klein quadric transforms their respective Klein images into each other. Then*

$$\begin{aligned} S(G) &= C(G'), & C(G) &= S(G'), \quad \text{and} \\ S(G)\gamma &= M_2^4 \cap G\gamma^*, & S(G')\gamma &= M_2^4 \cap G'\gamma^*, \\ C(G)\gamma &= M_2^4 \cap G'\gamma^*, & C(G')\gamma &= M_2^4 \cap G\gamma^*. \end{aligned}$$

Proof. As $G\gamma^*\mu_2^4 = G'\gamma^*\mu_2^4\mu_2^4 = G'\gamma^*$, it is clear that the left half of the bottom group of equations implies the right half. Further it is obvious that these four equations imply the two former ones.

$C\gamma^* \in M_2^4$ means that C is a singular complex. In this case its axis A satisfies the relation $A\gamma = C\gamma^*$, by definition of γ^* . This shows the equation $S(G)\gamma = M_2^4 \cap G\gamma^*$. By Prop. 1.1.20,

$$\bigcap_{C \in G} C\gamma^*\mu_2^4 = (\bigvee_{C \in G} C\gamma^*)\mu_2^4 = G\gamma^*\mu_2^4 = G'\gamma^*.$$

We use Equ. (3.10) to compute the Klein image of the carrier $C(G)$:

$$C(G)\gamma = \bigcap_{c \in G} C\gamma = \bigcap_{c \in G} (M_2^4 \cap C\gamma^* \mu_2^4) = M_2^4 \cap \bigcap_{c \in G} C\gamma^* \mu_2^4 = M_2^4 \cap G'\gamma^*. \quad \square$$

We will study the cases $\dim(G) = 1$ and $\dim(G) = 2$ in greater detail in Sec. 3.2.1 and Sec. 3.3, respectively. At first sight, this material might seem rather abstract. It has, however, a variety of applications in various fields as kinematics, classical mechanics, and computational geometry, and it will be important in our further development of line geometry.

3.2.1 Pencils of Linear Line Complexes

A one-dimensional linear manifold of linear complexes is also called a *pencil* of linear complexes. Its γ^* -image is a line in P^5 which can be written in the form

$$CR = (\lambda A + \mu B)R.$$

Its intersection with the Klein quadric is computed by solving

$$0 = \Omega(C, C) = \lambda^2 \Omega(A, A) + 2\lambda\mu \Omega(A, B) + \mu^2 \Omega(B, B). \quad (3.13)$$

The number of solutions depends on whether this homogeneous polynomial of λ, μ is the zero polynomial or not, and on the sign of the discriminant

$$\Delta := \Omega^2(A, B) - \Omega(A, A)\Omega(B, B) = \Omega^2(A, B) - 4\Omega_q(A)\Omega_q(B). \quad (3.14)$$

This shows that, from the projective viewpoint, there are at least four possible types of pencils: (a) the polynomial (3.13) is the zero polynomial, (b) the polynomial is not the zero polynomial, and (b1) $\Delta < 0$, (b2) $\Delta = 0$, and (b3) $\Delta > 0$. The number of intersection points of $G\gamma^*$ with M_2^4 is $\infty, 0, 1$, and 2 , respectively.

The following lemma is useful:

Lemma 3.2.2. *If the pencil G of linear complexes is spanned by C_1 and C_2 , then $C(G) = C_1 \cap C_2$.*

Proof. The inclusion $C_1 \cap C_2 \supset C(G)$ is trivial. We show $C_1 \cap C_2 \subset C(G)$: Assume that C_i has the coordinate vector $(\mathbf{c}_i, \bar{\mathbf{c}}_i)$. If a line L with Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ is contained in both C_1 and C_2 , then $\mathbf{c}_1 \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}}_1 \cdot \mathbf{l} = 0$ and $\mathbf{c}_2 \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}}_2 \cdot \mathbf{l} = 0$ imply that $(\lambda\mathbf{c}_1 + \mu\mathbf{c}_2) \cdot \bar{\mathbf{l}} + (\lambda\bar{\mathbf{c}}_1 + \mu\bar{\mathbf{c}}_2) \cdot \mathbf{l} = 0$, so $L \in C(G)$. \square

Pencils of Singular Linear Complexes

Here G consists of singular complexes, i.e., $G\gamma^*$ is entirely contained in M_2^4 . This is only possible if all coefficients in Equ. (3.13) are zero:

$$\Omega_q(A) = \Omega_q(B) = \Omega(A, B) = 0.$$

The Klein image $S(G)\gamma$ of G 's singular set coincides with $G\gamma^*$, which shows that $S(G)$ is a pencil of lines.

The carrier $C(G)$, i.e., the lines which belong to *all* complexes of G , are those which intersect *all* lines in a pencil. So $C(G)$ consists of the bundle of lines concurrent in the pencil's vertex, and the field of lines incident with the pencil's carrier plane.

Lemma 3.2.1 shows that $\gamma(C(G)) = M_2^4 \cap \mu_2^4 \circ \gamma(S(G))$. Because $\gamma(S(G)) = G\gamma^*$ is one-dimensional, the Klein image of $C(G)$ is contained in a three-space. It consists of two planes in M_2^4 , which intersect in $\gamma(S(G))$.

Hyperbolic Pencils of Linear Complexes and Hyperbolic Linear Congruences

Here we study the case $\Delta > 0$, where Δ is defined in Equ. (3.14). This means that G contains exactly two singular complexes. Such a pencil is called *hyperbolic*.

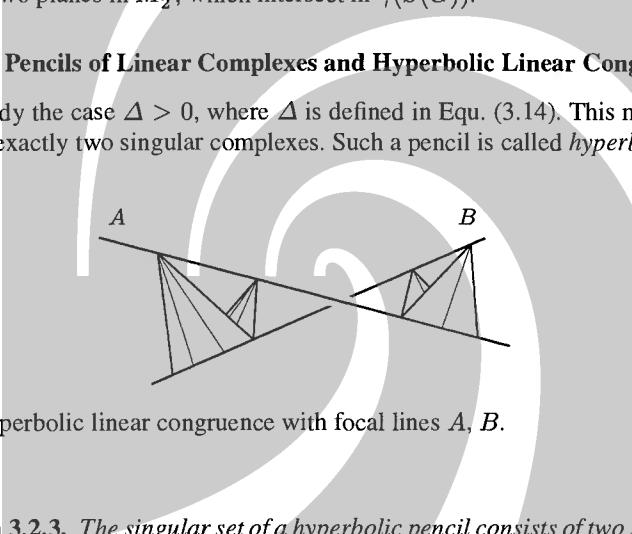


Fig. 3.6. Hyperbolic linear congruence with focal lines A, B .

Proposition 3.2.3. *The singular set of a hyperbolic pencil consists of two skew lines A, B . Its carrier consists of all lines which intersect A, B (see Fig. 3.6).*

Proof. As $\Delta > 0$, Equ. (3.13) has two solutions, and G contains two singular complexes \mathcal{A}, \mathcal{B} . \mathcal{A} consists of the lines which meet a line A and analogously for \mathcal{B} and B . Thus $S(G) = \{A, B\}$ and $G\gamma^* = \mathcal{A}\gamma^* \vee \mathcal{B}\gamma^* = A\gamma \vee B\gamma$.

A, B are skew because the line $A\gamma \vee B\gamma = G\gamma^*$ is not contained in M_2^4 (see Lemma 2.1.7). By Lemma 3.2.2, the carrier $C(G)$ equals $\mathcal{A} \cap \mathcal{B}$, i.e., consists of the lines which meet both A and B . \square

Definition. *The set \mathcal{N} of lines which intersect two given skew lines A, B is called the hyperbolic linear congruence with focal lines A, B (see Fig. 3.6). These are also called the focal lines of \mathcal{N} .*

A set of lines whose Klein image is two-dimensional, is called a *line congruence*. If this Klein image is algebraic, or smooth, the congruence is called algebraic, or smooth, respectively. Line congruences in general are studied in Sec. 7.1.1.

Theorem 3.2.4. *The carrier of a hyperbolic pencil G of linear complexes is a hyperbolic linear congruence \mathcal{N} . Its Klein image $\mathcal{N}\gamma$ is a ruled quadric contained in M_2^4 .*

Proof. By Lemma 3.2.1, $\mathcal{N}\gamma$ is the intersection of M_2^4 with $G\gamma^*$'s polar subspace, which is three-dimensional. This intersection is a regular quadric, because $G\gamma^*$ is non-tangential. \mathcal{N} contains pencils of lines, so $\mathcal{N}\gamma$ is a *ruled* quadric. \square

Pencils of Complexes in Complex Projective Space

A linear complex \mathcal{C} in *complex* projective space \mathbb{CP}^3 is defined in the same way as in real projective space — its lines satisfy a linear equation in Plücker coordinates. Such a complex is either singular, or associated with a null polarity π , according to Lemma 3.1.2. Its coordinate vector $(\mathbf{c}, \bar{\mathbf{c}})$ is composed from its linear equation like in the real case, and we let $\mathcal{C}\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{C}$. Th. 3.1.3 holds true for complex complexes as well. Two linear complexes are said to be *conjugate complex*, if they have conjugate complex coordinate vectors. This definition is independent of the coordinate system.

If \mathcal{C} is a linear complex in real projective space, and has the equation $\mathbf{c} \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}} \cdot \mathbf{l} = 0$, then this equation has also complex solutions $(\mathbf{l}, \bar{\mathbf{l}}) \in \mathbb{C}^6$, so it defines a linear complex in \mathbb{CP}^3 , denoted by $\mathcal{C}_{\mathbb{C}}$. \mathcal{C} equals the set of real lines in $\mathcal{C}_{\mathbb{C}}$. Obviously $\mathcal{C}\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ and $\mathcal{C}_{\mathbb{C}}\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{C}$.

The definition of a pencil G of linear complexes carries over to \mathbb{CP}^3 — $G\gamma^*$ is a line in \mathbb{CP}^5 . Depending on the number of intersection points of $G\gamma^*$ with the Klein quadric $M_2^4(\mathbb{C})$, the pencil is called singular, parabolic, or hyperbolic. Elliptic complexes do not exist in complex projective space, because Equ. (3.13) always has a solution.

A pencil G , which is spanned by complexes $\mathcal{C}_1, \mathcal{C}_2$ in real projective space, defines a pencil $G_{\mathbb{C}}$ spanned by $(\mathcal{C}_1)_{\mathbb{C}}$ and $(\mathcal{C}_2)_{\mathbb{C}}$. We call a pencil in complex projective space a *real* pencil, if it is spanned by two real complexes.

Lemma 3.2.5. *If G is a pencil of linear complexes in P^3 and $G_{\mathbb{C}}$ the corresponding pencil in \mathbb{CP}^3 , then the carrier $C(G)$ consists of the real lines in $C(G_{\mathbb{C}})$.*

Proof. This follows from Lemma 3.2.2 and the fact that the real lines of $C(G_{\mathbb{C}}) = (\mathcal{C}_1)_{\mathbb{C}} \cap (\mathcal{C}_2)_{\mathbb{C}}$ comprise the set $\mathcal{C}_1 \cap \mathcal{C}_2$. \square

Remark 3.2.1. Beware of the two meanings of ‘complex’: first the noun, which denotes a set of lines, and second, the adjective, which means an object associated with, or consisting of, complex numbers.

When speaking about conjugate complex complexes, another difficulty arises: Usually complex conjugation is denoted by a bar, but the symbol $\bar{\mathbf{c}}$ here simply denotes the second part of a Plücker coordinate vector. Thus we cannot use the bar to indicate complex conjugation without comment. \diamond

Elliptic Pencils of Linear Complexes and Elliptic Linear Congruences

If $\Delta < 0$ in Equ. (3.14), the pencil G has no singular complexes at all and is called *elliptic*. It is useful to employ the complex extension of P^3 and to consider the pencil $G_{\mathbb{C}}$. As (3.14) has two conjugate complex solutions over the complex number field, the pencil $G_{\mathbb{C}}$ is hyperbolic and is spanned by two conjugate complex complexes \mathcal{A}, \mathcal{B} . \mathcal{A} consists of all lines which meet a line A , and the same for \mathcal{B} and a line B . The lines A and B are complex conjugate, i.e., possess complex conjugate Plücker coordinates.

Definition. *The carrier of an elliptic pencil of linear complexes is called an elliptic linear congruence.*

By Lemma 3.2.5, this elliptic linear congruence, denoted by \mathcal{N} , is the set of real lines which intersect A and B . By Prop. 3.2.3, A and B are skew. The Klein image of \mathcal{N} is a quadric, because it equals $M_2^4 \cap G\gamma^*\mu_2^4$.

Proposition 3.2.6. *If \mathcal{N} is the elliptic linear congruence corresponding to a pencil G of linear complexes, then for all points $P \in P^3$ there is exactly one line of \mathcal{N} incident with P , and likewise for all planes of P^3 .*

Proof. \mathcal{N} consists of the real lines which intersect, in the complex extension, the focal lines A, B of $G_{\mathbb{C}}$. As P is contained neither in A nor in B , there is exactly one line $L_{\mathbb{C}}$ of $\mathbb{C}P^3$ incident with P , which meets both A and B . Complex conjugation in $\mathbb{C}P^3$ interchanges A and B , and fixes P , so it fixes $L_{\mathbb{C}}$. Thus $L_{\mathbb{C}}$ is actually a real line and is contained in \mathcal{N} . The statement for planes follows by duality. \square

Theorem 3.2.7. *An elliptic linear congruence \mathcal{N} consists of all real lines which intersect a skew pair of complex conjugate focal lines, and vice versa. Its Klein image is an oval quadric contained in M_2^4 .*

Proof. We have already shown that \mathcal{N} consists of the real lines which intersect a pair of conjugate complex skew lines. To show the converse, we consider the pencil spanned by singular complexes \mathcal{A}, \mathcal{B} , such that \mathcal{A} has coordinates $(\mathbf{a}, \bar{\mathbf{a}})$ and \mathcal{B} has coordinates $(\mathbf{b}, \bar{\mathbf{b}})$ conjugate complex to $(\mathbf{a}, \bar{\mathbf{a}})$.

This pencil is also spanned by the complexes with coordinates $(\mathbf{a} + \mathbf{b}, \bar{\mathbf{a}} + \bar{\mathbf{b}})$ and $i(\mathbf{a} - \mathbf{b}, \bar{\mathbf{a}} - \bar{\mathbf{b}})$, where i is the imaginary unit. These are real coordinates, so the pencil is the complexification of a real elliptic pencil G . By the proof of Prop. 3.2.6, the carrier of G consists of the real lines which intersect A and B .

To show that $\mathcal{N}\gamma$ is an oval quadric, we note that by Prop. 3.2.6, \mathcal{N} contains no pencil, so $\mathcal{N}\gamma$ contains no lines. \square

Example 3.2.1. Let us look at the following special linear congruence \mathcal{N} in the projective extension of Euclidean space. We will use the complex extension and consider first a linear congruence $\mathcal{N}_{\mathbb{C}}$ in complex projective space: Let $I = (1, 0, 0, ip)\mathbb{C}$, $\bar{I} = (1, 0, 0, -ip)\mathbb{C}$, $J = (0, 1, i, 0)\mathbb{C}$, $\bar{J} = (0, 1, -i, 0)\mathbb{C}$, where i is the

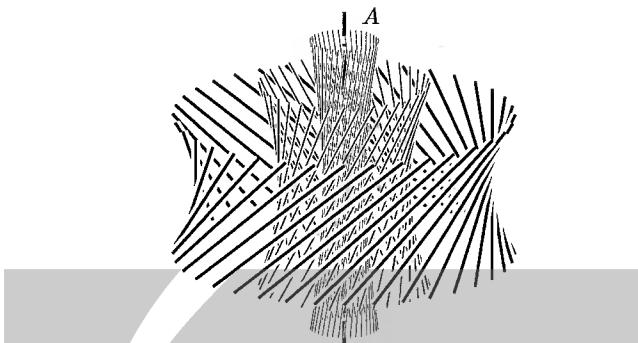


Fig. 3.7. Elliptic linear congruence with rotational symmetry about the line A .

imaginary unit, and p is a nonzero real number. $\mathcal{N}_{\mathbb{C}}$ is the hyperbolic linear congruence with focal lines $I \vee J$ and $\bar{I} \vee \bar{J}$. Their Plücker coordinates are $(1, i, 0, p, ip, 0)$ and $(1, -i, 0, p, -ip, 0)$.

We are looking for the pencil $G_{\mathbb{C}}$ of linear complexes which corresponds to this linear congruence. As the focal lines of $\mathcal{N}_{\mathbb{C}}$, i.e., the set $S(G_{\mathbb{C}})$, is given, we use the formula $S(G_{\mathbb{C}})\gamma = G_{\mathbb{C}}\gamma^* \cap M_2^4$ of Lemma 3.2.1: $G_{\mathbb{C}}\gamma^*$ is a line, and $S(G_{\mathbb{C}})\gamma$ consists of two points, so the former is spanned by the latter:

$$G_{\mathbb{C}}\gamma^* = (1, i, 0, p, ip, 0)\mathbb{C} \vee (1, -i, 0, p, -ip, 0)\mathbb{C}.$$

By taking sum and difference of these Plücker coordinates, we see that

$$G_{\mathbb{C}}\gamma^* = (1, 0, 0, p, 0, 0)\mathbb{C} \vee (0, 1, 0, 0, p, 0)\mathbb{C},$$

These are real coordinates, so $G_{\mathbb{C}}\gamma^*$ is actually the complexification of a real line $G\gamma^*$:

$$G\gamma^* = (1, 0, 0, p, 0, 0)\mathbb{R} \vee (0, 1, 0, 0, p, 0)\mathbb{R}.$$

G is a pencil of linear complexes in P^3 , and the linear congruence $\mathcal{N} = C(G)$ defined by the pencil G consists of the real lines in $\mathcal{N}_{\mathbb{C}}$. \mathcal{N} is an elliptic linear congruence, because it has no focal lines. The above complexes which span G have the equations

$$pl_{01} + l_{23} = 0, \quad pl_{02} + l_{31} = 0. \quad (3.15)$$

In order to visualize the congruence \mathcal{N} , we consider the helical motion (3.6). Equ. (3.7) shows that the path tangents of points $(x, y, 0)$ have the Plücker coordinates

$$(1, \bar{1}) = (-y, x, p, py, -px, x^2 + y^2).$$

This is exactly the set of non-ideal solutions of Equ. (3.15). The only other solution is the horizontal line at infinity. Obviously this elliptic linear congruence has rotational symmetry (see Fig. 3.7). \diamond

Proposition 3.2.8. Any elliptic linear congruence in E^3 is the image of the special congruence with rotational symmetry given by Equ. (3.15), under the mapping $x \mapsto k \cdot x$, $y \mapsto y$, $z \mapsto z$.

Proof. It is not hard to find a homogeneous Cartesian coordinate system such that the focal lines are given in the form $(1, 0, 0, ip)\mathbb{C}\vee(0, k, i, 0)\mathbb{C}$, and $(1, 0, 0, -ip)\mathbb{C}\vee(0, k, -i, 0)\mathbb{C}$. This is left to the reader as an exercise. \square

Figure 3.7 also shows the decomposition of an elliptic linear congruence in a one-parameter family of coaxial and concentric reguli which belong to one-sheeted hyperboloids with common vertices $(0, 0, \pm ip)$.

Parabolic Pencils of Linear Complexes and Parabolic Linear Congruences.

If $\Delta = 0$ in Equ. (3.14), but the pencil G is not singular, then $G\gamma^*$ is tangent to M_2^4 and G contains exactly one singular complex. The pencil is called *parabolic*.

Definition. The carrier of a parabolic pencil of linear complexes is called a *parabolic linear congruence*.

The unique singular complex \mathcal{A} in the pencil G determines a line A , which alone comprises the singular set $S(G)$. The pencil G is spanned by \mathcal{A} and any other regular complex $\mathcal{B} \in G$. If $\mathcal{A}\gamma^* = A$, $\mathcal{B}\gamma^* = B$, then G is parametrized by

$$\mathbf{C} = \lambda\mathbf{A} + \mu\mathbf{B}, \quad \text{with} \quad \Omega_q(\mathbf{A}) = 0, \quad \Delta = 0 \implies \Omega(\mathbf{A}, \mathbf{B}) = 0. \quad (3.16)$$

The complex \mathcal{B} can be chosen arbitrarily from the pencil G , but different from \mathcal{A} . This shows that $A\mathbb{R}$ is conjugate to all possible $B\mathbb{R}$ with respect to the polarity of M_2^4 , and that the line A belongs to all complexes of the pencil. Especially it is contained in the carrier $\mathcal{N} = C(G)$. It is called the *focal line* or *axis* of \mathcal{N} .

Theorem 3.2.9. A parabolic linear congruence \mathcal{N} with focal line A is the set of all lines of a regular linear complex \mathcal{B} which intersect a line $A \in \mathcal{B}$. Its Klein image is a quadratic cone. The congruence consists of a one-parameter family of line pencils $\mathcal{P}(u)$, whose vertices $a(u)$ are in A , and whose planes $\varepsilon(u)$ contain A . The mapping $a(u) \mapsto \varepsilon(u)$ is a projective mapping.

Proof. We assume that G is spanned by \mathcal{A} and \mathcal{B} with \mathcal{A} singular and \mathcal{B} regular, and the singular complex \mathcal{A} consists of the lines which meet A . Then $C(G) = \mathcal{A} \cap \mathcal{B}$ shows the first statement. To find the lines of \mathcal{N} incident with a point of A , we use the null polarity π associated with \mathcal{B} : The lines of \mathcal{B} incident with a point $a(u) \in A$ comprise the pencil with vertex $a(u)$ and carrier plane $a(u)\pi = \varepsilon(u)$. π is a correlation, so $a(u) \mapsto a(u)\pi$ is a projective mapping.

Lemma 3.2.1 shows that $\mathcal{N}\gamma$ is a tangential intersection of M_2^4 with a three-dimensional subspace. This implies that $\mathcal{N}\gamma$ is a quadratic cone; its vertex is $A\gamma$. \square

Remark 3.2.2. The parabolic linear congruence \mathcal{N} contains a one-parameter family of pencils: The Klein mapping translates this statement as follows: The quadratic cone $\mathcal{N}\gamma$ carries a one-parameter family of lines.

Note also that the focal line of \mathcal{N} belongs to \mathcal{N} ; hyperbolic and elliptic linear congruences do not contain their focal lines. \diamond

Definition. *The projective mapping referred to in Th. 3.2.9 is called the fundamental projectivity of the parabolic linear congruence.*

3.2.2 Euclidean Properties of Pencils of Linear Complexes

We have seen that in projective space there are three different nonsingular linear congruences. With respect to Euclidean geometry, there are different ‘sizes’ and ‘shapes’ within each projective class.

Parabolic Linear Congruences in Euclidean Space

First we show how the fundamental projectivity of a parabolic linear congruence works in Euclidean space. This will be of importance later, when we study ruled surfaces.

Consider the parabolic pencil (3.16) of linear complexes. Assume a homogeneous Cartesian coordinate system such that $A = A\mathbb{R} = (\mathbf{a}, \bar{\mathbf{a}})\mathbb{R}$ is no ideal line ($\mathbf{a} \neq \mathbf{0}$). Among the regular complexes with coordinates $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ there is exactly one whose axis is orthogonal to A : It is directly verified that the choice of

$$\lambda = -(\mathbf{a} \cdot \mathbf{b}), \quad \mu = \mathbf{a}^2, \quad (3.17)$$

gives $\mathbf{c} \cdot \mathbf{a} = 0$. We use the axis A_0 of this special complex \mathcal{C}_0 as z -axis, and A as x -axis of a Cartesian coordinate system. This is illustrated in Fig. 3.8. The line A_0 is the axis of the helical motion which corresponds to \mathcal{C}_0 , and the lines of \mathcal{N} are the path normals of A ’s points.

We introduce some notations for parabolic linear congruences in E^3 : The fundamental projectivity is denoted by π , and the ideal plane by ω . The xy -plane is called *central plane* and denoted by π_c , the origin is called *central point*, and denoted by S . A_0 is called *central normal*. We let $A_u = A \cap \omega$, denote the xz -plane with α , and call it *asymptotic plane*.

The fundamental projectivity π maps $S \mapsto \pi_c$ and $A_u \mapsto \alpha$. This is easily verified and also clear from the interpretation of null lines as path normals.

If p is the pitch of \mathcal{C}_0 , then $d := -p$ is called the *distribution parameter* of \mathcal{N} . We compute the distribution parameter with Equ. (3.9), using $(\mathbf{c}, \bar{\mathbf{c}}) = \lambda(\mathbf{a}, \bar{\mathbf{a}}) + \mu(\mathbf{b}, \bar{\mathbf{b}})$ with λ, μ from Equ. (3.17). As $\mathbf{a} \cdot \bar{\mathbf{b}} + \bar{\mathbf{a}} \cdot \mathbf{b} = 0$, this expression simplifies to

$$d = \frac{\mathbf{a}^2(\mathbf{b} \cdot \bar{\mathbf{b}})}{(\mathbf{a} \cdot \mathbf{b})^2 - \mathbf{a}^2\mathbf{b}^2} = -\frac{\mathbf{a}^2(\mathbf{b} \cdot \bar{\mathbf{b}})}{(\mathbf{a} \times \mathbf{b})^2}. \quad (3.18)$$

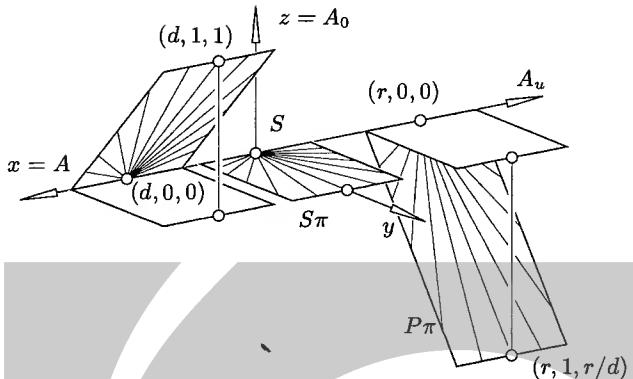


Fig. 3.8. Parabolic linear congruence \mathcal{N} in Euclidean space. The helical motion about the z -axis such that \mathcal{N} 's lines are path normals is left-handed, so $d > 0$.

According to Equ. (3.7), the vector $(0, r, p)$ is orthogonal to the null plane of the point $P = (r, 0, 0)$. This shows that

$$\overline{SP} = d \tan \sphericalangle(P\pi, S\pi). \quad (3.19)$$

In particular, if $\overline{SP} = d$, then $\sphericalangle(S\pi, P\pi) = \pi/4$.

The Axes of Complexes in a Pencil

We use a Cartesian coordinate system. Assume that a pencil G of linear complexes is spanned by the complexes \mathcal{A}, \mathcal{B} with coordinate vectors $\mathcal{A}\gamma^* = \mathbb{A}\mathbb{R} = (\mathbf{a}, \bar{\mathbf{a}})\mathbb{R}$, $\mathcal{B}\gamma^* = \mathbb{B}\mathbb{R} = (\mathbf{b}, \bar{\mathbf{b}})\mathbb{R}$, respectively. We consider only the case of linearly independent \mathbf{a}, \mathbf{b} here. The case of linear dependence is briefly addressed in Remark 3.2.3.

The axis of the complex \mathcal{C} with $\mathcal{C}\gamma^* = \mathcal{C} = \lambda\mathcal{A} + \mu\mathcal{B}$ is parallel to $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$. This shows that we can assume, without loss of generality, that G is spanned by complexes \mathcal{A}, \mathcal{B} with *orthogonal* axes $A_{\mathcal{A}}, A_{\mathcal{B}}$, and that we can use a coordinate system as shown in Fig. 3.9.

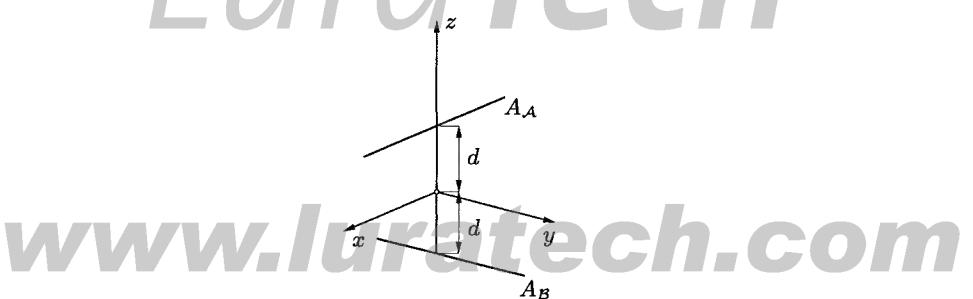


Fig. 3.9. Coordinate system for a pencil of linear complexes in Euclidean space.

In this coordinate system, the Plücker coordinates of the axes are given by $\mathbf{A}_\mathcal{A} = (1, 0, 0, 0, d, 0)$ and $\mathbf{A}_\mathcal{B} = (0, 1, 0, d, 0, 0)$. Equ. (3.9) shows that \mathbf{A}, \mathbf{B} must be of the form

$$\mathbf{A} = (1, 0, 0, p_\mathcal{A}, d, 0), \quad \mathbf{B} = (0, 1, 0, d, p_\mathcal{B}, 0).$$

The pitches of \mathcal{A}, \mathcal{B} then are $p_\mathcal{A}, p_\mathcal{B}$, respectively.

We normalize the coefficients λ, μ such that $\lambda = \cos \phi, \mu = \sin \phi$, so we have $\mathbf{C} = \mathbf{A} \cos \phi + \mathbf{B} \sin \phi$, which expands to

$$\mathbf{C} = (\mathbf{c}, \bar{\mathbf{c}}) = (\cos \phi, \sin \phi, 0, p_\mathcal{A} \cos \phi + d \sin \phi, d \cos \phi + p_\mathcal{B} \sin \phi, 0).$$

The Plücker coordinates $\mathbf{A}_\mathcal{C}$ of \mathcal{C} 's axis $A_\mathcal{C}$ are given by

$$\begin{aligned} \mathbf{A}_\mathcal{C} &= (\cos \phi, \sin \phi, 0, -h(\phi) \sin \phi, h(\phi) \cos \phi, 0), \quad \text{with} \\ h(\phi) &= (p_\mathcal{B} - p_\mathcal{A}) \sin \phi \cos \phi + d - 2d \sin^2 \phi. \end{aligned}$$

This shows that $A_\mathcal{C}$ intersects the z -axis orthogonally in the point $(0, 0, h(\phi))$. We are looking for complexes of G with *orthogonally intersecting axes*. They belong to values of ϕ which solve $h(\phi) = h(\phi + \pi/2)$. This equation is equivalent to

$$(p_\mathcal{B} - p_\mathcal{A}) \sin 2\phi + 2d \cos 2\phi = 0,$$

which always has solutions. Now without loss of generality we can assume that \mathcal{A}, \mathcal{B} are two such complexes. Then automatically $d = 0$ and the formulae for $\mathbf{A}_\mathcal{C}$ simplify to

$$\begin{aligned} \mathbf{A}_\mathcal{C} &= (\cos \phi, \sin \phi, 0, -h(\phi) \sin \phi, h(\phi) \cos \phi, 0) \quad (3.20) \\ \text{with} \quad h(\phi) &= \frac{1}{2}(p_\mathcal{B} - p_\mathcal{A}) \sin 2\phi. \end{aligned}$$

The pitch of \mathcal{C} has the value

$$p_\mathcal{C} = p_\mathcal{A} \cos^2 \phi + p_\mathcal{B} \sin^2 \phi. \quad (3.21)$$

If $p_\mathcal{A} = p_\mathcal{B}$, the set of axes is a line pencil and the pitch is independent of ϕ : G has rotational symmetry about the z -axis. If $p_\mathcal{A} = p_\mathcal{B} \neq 0$, the carrier $C(G)$ is an elliptic linear congruence with rotational symmetry. If $p_\mathcal{A} = p_\mathcal{B} = 0$, the pencil G is singular, and its carrier is the field of lines incident with the xy -plane together with the bundle of lines concurrent in the origin.

The Cylindroid

We continue the discussion above. If $p_\mathcal{A} \neq p_\mathcal{B}$, the set of axes is called *cylindroid* or *Plücker conoid* (see Fig. 3.10).

Equ. (3.20) shows that the cylindroid is generated as follows: A line L moves such that during its motion it intersects the z -axis orthogonally. The motion is a superposition of a rotation about the z -axis with a harmonic translation of double frequency in z -direction.

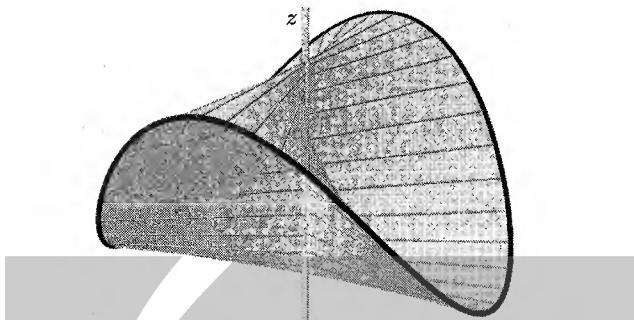


Fig. 3.10. The Plücker conoid.

The pitch assumes its extremal values p_A, p_B at the complexes whose axes intersect orthogonally. Singular complexes are characterized by zero pitch:

$$p = p_A \cos^2 \phi + p_B \sin^2 \phi = 0. \quad (3.22)$$

Their axes are the focal lines of the carrier $C(G) = \mathcal{N}$, which is a hyperbolic linear congruence if $p_A p_B < 0$, a parabolic one if exactly one of p_A, p_B is zero, and an elliptic one for $p_A p_B > 0$.

Remark 3.2.3. If the vectors \mathbf{a}, \mathbf{b} are linearly dependent, the situation is different. There are the following possibilities: $\mathbf{a} = \mathbf{b} = \mathbf{o}$ means that all complexes are singular, their axes are a line pencil in the ideal plane.

If not both \mathbf{a}, \mathbf{b} are zero, we may assume, without loss of generality, that neither is zero, and we easily deduce that the set of axes is a pencil of parallel lines. Either the pencil of complexes is singular, or its carrier is a parabolic or hyperbolic linear congruence with one focal line at infinity. ◇

3.3 Reguli and Bundles of Linear Complexes

A two-dimensional linear manifold G of linear complexes is called a *bundle of linear complexes*. The set $G\gamma^*$ is a plane in P^5 which is parametrized by $\mathbb{CR} = (\lambda_0 \mathbf{A}_0 + \lambda_1 \mathbf{A}_1 + \lambda_2 \mathbf{A}_2)\mathbb{R}$. The intersection $G\gamma^* \cap M_2^4$ is computed by solving $\Omega_q(C) = 0$, which expands to

$$(\lambda_0, \lambda_1, \lambda_2) \cdot \Delta \cdot (\lambda_0, \lambda_1, \lambda_2)^T = 0 \quad \text{with} \quad \Delta := (\Omega(\mathbf{A}_i, \mathbf{A}_j))_{i,j=0,1,2}. \quad (3.23)$$

The projective type of the intersection depends on the rank of the matrix Δ and on the existence of real solutions of a quadratic equation. The classification obviously is equivalent to the classification of quadratic varieties in P^2 (cf. also Ex. 1.1.27). $(\lambda_0, \lambda_1, \lambda_2)$ are homogeneous coordinates in $G\gamma^*$ with respect to the projective coordinate system $(\mathbf{A}_0\mathbb{R}, \mathbf{A}_1\mathbb{R}, \mathbf{A}_2\mathbb{R}; (\mathbf{A}_0 + \mathbf{A}_1 + \mathbf{A}_2)\mathbb{R})$. If $\text{rk}(\Delta) < 3$, there are the following possibilities:

1. $\text{rk}(\Delta) = 0$: $G\gamma^*$ is contained in M_2^4 : All $\mathcal{C} \in G$ are singular.
2. $\text{rk}(\Delta) = 1$: The equation $\Omega_q(\mathcal{C}) = 0$ defines a line in P^5 , so G contains one pencil of singular complexes.
3. $\text{rk}(\Delta) = 2$: The equation $\Omega_q(\mathcal{C}) = 0$ defines two lines in $\mathbb{C}P^5$. As both lines are contained in the plane $G\gamma^*$, they intersect.
 - 3a. These two lines are real: G contains two pencils of singular complexes.
 - 3b. These two lines are conjugate complex. Their intersection point is real, so G contains exactly one singular complex. Its complex extensions contains two singular pencils.

Remark 3.3.1. From $\dim(G\gamma^*) = 2$ we conclude that $\dim(G\gamma^*\mu_2^4) = 2$. This shows (cf. Lemma 3.2.1) that both $C(G)\gamma$ and $S(G)\gamma$ are planar sections of the Klein quadric. \diamond

If $\text{rk}(\Delta) = 3$, the intersection $G\gamma^* \cap M_2^4$ is a conic or void. In the latter case, the complex extension of G intersects M_2^4 in a conic without real points. In both cases we speak of a *regular* bundle of linear complexes.

Reguli

In Ex. 1.1.25 we discussed ruled quadrics which carry two families $\mathcal{R}_1, \mathcal{R}_2$ of lines. If we take three lines L_1, L_2, L_3 of \mathcal{R}_1 , then \mathcal{R}_2 consists of all lines which intersect L_1, L_2, L_3 , and vice versa. Such a family of lines was called a *regulus*.

We repeat the definition of a regulus, and show some further properties. Pictures of reguli and their carrier quadrics can be found in Fig. 1.20 and Fig. 1.23.

Definition. If L_1, L_2, L_3 are pairwise skew lines, the set \mathcal{R} of all lines which intersect L_1, L_2, L_3 is called a *regulus*.

Proposition 3.3.1. The Klein image $\mathcal{R}\gamma$ of a regulus \mathcal{R} is a non-tangential planar section $V \cap M_2^4$ of the Klein quadric, and vice versa.

Any three lines of \mathcal{R} define another regulus \mathcal{R}' , whose Klein image is given by $\mathcal{R}'\gamma = V\mu_2^4 \cap M_2^4$. \mathcal{R} and \mathcal{R}' are the two reguli of the same ruled carrier quadric.

Proof. Assume that \mathcal{R} is the set of lines L which meet the three lines L_1, L_2, L_3 (these lines are skew). By Cor. 2.1.8, $\mathcal{R}\gamma$ is the set of points $L\gamma \in M_2^4$ which are conjugate to the three points $L_1\gamma, L_2\gamma, L_3\gamma$ (these points are pairwise non-conjugate).

It follows that such an $L\gamma$ is conjugate to all points of the plane $U = L_1\gamma \vee L_2\gamma \vee L_3\gamma$, so $\mathcal{R}\gamma$ coincides with the set $V \cap M_2^4$ with $V = U\mu_2^4$. V is not tangent to the Klein quadric, because this would imply that also U has this property and at least two of the points $L_i\gamma$ are conjugate to each other.

If we repeat this construction with three points $L'_1\gamma, L'_2\gamma, L'_3\gamma$ from $\mathcal{R}\gamma$, their span is the plane V and the regulus \mathcal{R}' defined by the lines L'_1, L'_2, L'_3 has the property that $\mathcal{R}'\gamma = M_2^4 \cap V\mu_2^4$.

Every line of \mathcal{R} intersects every line of \mathcal{R}' , so the carrier quadrics of \mathcal{R} and \mathcal{R}' coincide. The proposition is proved. \square

Corollary 3.3.2. *The Klein image of a regulus is a conic which is not contained in a generator plane of the Klein quadric, and vice versa.*

Proof. These conics are those planar sections $U \cap M_2^4$ of the Klein quadric such that U is not tangent to M_2^4 . Therefore the statement follows immediately from Prop. 3.3.1. \square

Definition. *Reguli $\mathcal{R}, \mathcal{R}'$ which have the same carrier quadric are called complementary to each other.*

Clearly, if a regulus \mathcal{R} is defined by lines L_1, L_2, L_3 , its complementary regulus \mathcal{R}' contains L_1, L_2, L_3 . This shows that a regulus is completely determined by three pairwise skew lines contained in it.

Theorem 3.3.3. *Singular set $S(G)$ and carrier $C(G)$ of a regular bundle of linear complexes are two complementary reguli.*

Proof. This follows directly from Lemma 3.2.1 and Prop. 3.3.1. \square

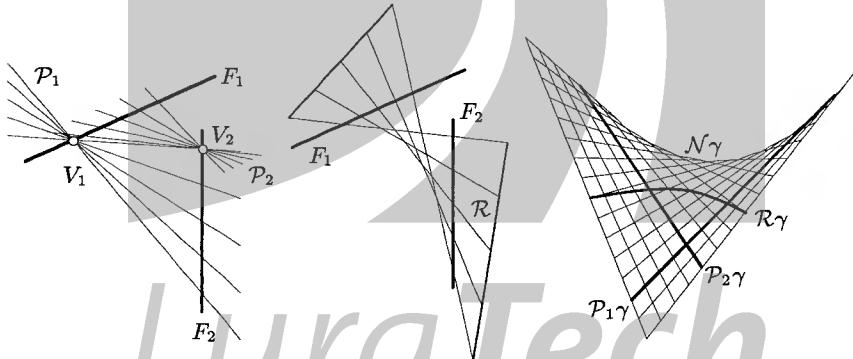


Fig. 3.11. Left: Line pencils $\mathcal{P}_1, \mathcal{P}_2$ in a hyperbolic linear congruence \mathcal{N} with focal lines F_1, F_2 . Center: Regulus \mathcal{R} in \mathcal{N} . Right: Klein image.

Example 3.3.1. The Klein image $\mathcal{N}\gamma$ of a hyperbolic linear congruence \mathcal{N} is, by Th. 3.2.4, a ruled quadric contained in a three-dimensional subspace G^3 of P^5 . We may apply any projective isomorphism $G^3 \rightarrow P^3$ to visualize the Klein image $\mathcal{N}\gamma$.

Fig. 3.11 shows an example: We choose two pencils $\mathcal{P}_1, \mathcal{P}_2$, and a regulus \mathcal{R} contained in \mathcal{N} . The Klein images $\mathcal{P}_{i\gamma}$ are straight lines, and the Klein image $\mathcal{R}\gamma$ is a conic. All three are contained in the ruled quadric $\mathcal{N}\gamma$.

Visualization of line space in general will be discussed in Sec. 8.1.3. \diamond

Linear Congruences Tangent to Reguli

Consider a regulus \mathcal{R} and a line $L \in \mathcal{R}$. Then $\mathcal{R}\gamma$ is a conic, and $L\gamma$ one of its points. This conic's tangent in the point $L\gamma$ is a line in P^5 , denoted by $G\gamma^*$, which means that G is a pencil of complexes. The carrier of G is a parabolic linear congruence $\mathcal{N} = C(G)$ with focal line L . By Lemma 3.2.1, the Klein image $\mathcal{N}\gamma = C(G)\gamma$ equals $G\gamma^* \mu_2^4 \cap M_2^4$.

Do the lines of \mathcal{N} have a geometric meaning for \mathcal{R} ? The answer will be given by Lemma 3.3.4, whose proof requires to do Ex. 3.3.2 first.

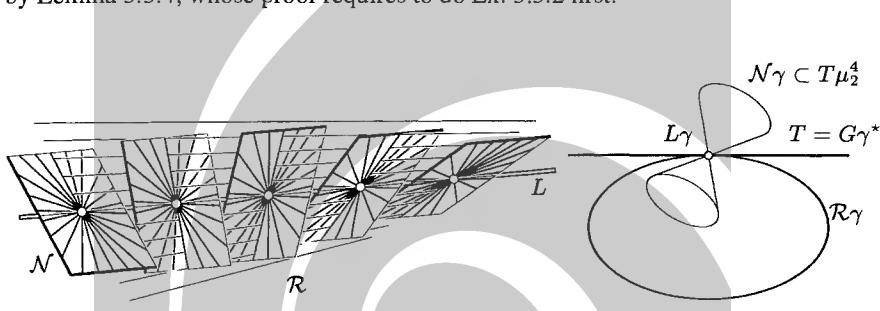


Fig. 3.12. Left: The parabolic linear congruence $\mathcal{N} = C(G)$ tangent to a regulus \mathcal{R} (see Lemma 3.3.4). Right: Klein image.

Lemma 3.3.4. Assume a regulus \mathcal{R} and a line $L \in \mathcal{R}$. Consider the quadric Φ which carries \mathcal{R} . Then the set of Φ 's surface tangents in points of L equals the parabolic linear congruence $\mathcal{N} = C(G)$ where the line $G\gamma^*$ is tangent to the conic $\mathcal{R}\gamma$ in the point $L\gamma$ (cf. Fig. 3.12).

We first show that Lemma 3.3.4 is true for one particular ruling of one particular regulus:

Example 3.3.2. Consider the lines $L_1 = (1, 0, 0, 0)\mathbb{R} \vee (0, 1, 0, 0)\mathbb{R}$, $L_2 = (1, 0, 1, 0)\mathbb{R} \vee (0, 1, 0, 1)\mathbb{R}$, $L_3 = (0, 0, 1, 0)\mathbb{R} \vee (0, 0, 0, 1)\mathbb{R}$, and the regulus \mathcal{R} which contains L_1, L_2, L_3 . The quadric Φ which carries \mathcal{R} has the equation $x_0x_3 - x_1x_2 = 0$. This is verified by testing that L_1, L_2, L_3 are contained in the quadric.

The Plücker coordinates of L_1, L_2, L_3 are $L_1 = (1, 0, 0, 0, 0, 0)$, $L_2 = (1, 0, 1, 0, -1, 0)$, $L_3 = (0, 0, 0, 1, 0, 0)$. The span $[L_1, L_2, L_3]$ has the equation $l_{02} = l_{31} = l_{03} + l_{12} = 0$. We parametrize it by $(\lambda_0, 0, \lambda_1, \lambda_2, 0, -\lambda_1)$. Its intersection with the Klein quadric is the conic given by $\lambda_0\lambda_2 = \lambda_1^2$ and, by Prop. 3.3.1, coincides with $\mathcal{R}\gamma$.

The tangent of the curve $\mathcal{R}\gamma$ in the point $L_1\mathbb{R}$ has the equation $\lambda_2 = 0$ and therefore equals the line $L_1\mathbb{R} \vee (0, 0, 1, 0, 0, -1)\mathbb{R}$. This line equals the set $G\gamma^*$, for some well-defined parabolic pencil G of linear complexes. If $C \in G$, then by construction $C\gamma^* = (\lambda, 0, \mu, 0, 0, -\mu)\mathbb{R}$.

\mathcal{C} is regular if $\mu \neq 0$ and has an associated null polarity π . We choose a point $P = (\sigma, \tau, 0, 0)\mathbb{R}$ in L_1 and compute its null plane $P\pi$: Equ. (3.3) gives $P\pi =$

$\mathbb{R}\mu(0, 0, -\tau, \sigma)$. Further, we compute Φ 's tangent plane $P\pi_\Phi$ at P , which is done by computing $\mathbb{R}A \cdot (\sigma, \tau, 0, 0)$, where A is the coordinate matrix of the quadric Φ . This results in $\mathbb{R}(0, 0, -\tau, \sigma)$ again. \diamond

Proof. (of Lemma 3.3.4) It has been shown in Ex. 1.1.25 that for ruled quadrics Φ, Φ' and lines $L \subset \Phi, L' \subset \Phi'$ there is a projective automorphism which transforms Φ to Φ' , and L to L' . Therefore it is sufficient to consider one ruled quadric and one of its lines, which has been done in Ex. 3.3.2. \square

Remark 3.3.2. Consider a regulus \mathcal{R} and a line $L_0 \in \mathcal{R}$. Assume a smooth family $L(t)$ of lines such that $L(t) \in \mathcal{R}$ and $\lim_{t \rightarrow 0} L(t) = L_0$. The lines which intersect both L_0 and $L(t)$ are contained in a hyperbolic linear congruence $\mathcal{N}(t)$. All of them intersect \mathcal{R} 's carrier quadric in two points. As $L(t)$ converges to L_0 , the linear congruence $\mathcal{N}(t)$ converges to the set of lines *tangent* to Φ in a point of L_0 , which is the parabolic linear congruence discussed in Ex. 1.1.25 and Lemma 3.3.4.

This shows that a parabolic linear congruence occurs as limit of hyperbolic linear congruences if their axes converge to the same line. \diamond

3.4 Applications

We describe two fields of application of the concept of linear complex and linear manifold of linear complexes. One is kinematics of Euclidean three-space, where path normals comprise linear complexes. The other one is statics of forces, which has many analogues with spatial kinematics.

3.4.1 Spatial Kinematics

Consider a rigid body moving in Euclidean three-space E^3 . We think of two copies of E^3 : One copy associated with the moving body and called *moving space* or *moving system* Σ^0 , and one copy called the *fixed space* or *fixed system* Σ . We use Cartesian coordinates and denote points of the moving system Σ^0 by x^0, y^0, \dots , and points of the fixed system by x, y , and so on.

One-parameter Motions

A *one-parameter motion* Σ^0/Σ is a smooth family of Euclidean congruence transformations depending on a parameter t which can be thought of as time. A point x^0 of Σ^0 is, at time t , mapped to the point

$$\mathbf{x}(t) = A(t) \cdot \mathbf{x}^0 + \mathbf{a}(t) \quad (3.24)$$

of Σ , where $A(t) \in \text{SO}_3$ and $\mathbf{a}(t) \in \mathbb{R}^3$. In this way all points of Σ^0 have a *path curve* or *trajectory* $\mathbf{x}(t)$ in Σ . The trajectory of the origin is $\mathbf{a}(t)$. $A(t)$ is called the *linear part* of the motion. We have $A^T = A^{-1}$ and $\det(A) = 1$.

The first derivative

$$\dot{\mathbf{x}}(t) = \dot{A}(t) \cdot \mathbf{x}^0 + \dot{\mathbf{a}}(t)$$

of the path of \mathbf{x}^0 is called its *velocity vector* at time t . We write $\mathbf{v}(\mathbf{x})$ for the vector field of vectors $\dot{\mathbf{x}}(t)$ attached to the points $\mathbf{x}(t)$.

Proposition 3.4.1. *The velocity vector of \mathbf{x}^0 at time t can be written in the form*

$$\mathbf{v}(\mathbf{x}) = \dot{\mathbf{x}}(t) = \bar{\mathbf{c}} + \mathbf{c} \times \mathbf{x}, \quad \text{where} \quad (3.25)$$

$$\dot{A}(t)A^T(t) = \begin{bmatrix} & -c_3 & c_2 \\ c_3 & & -c_1 \\ -c_2 & c_1 & \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}, \quad \bar{\mathbf{c}} = \dot{\mathbf{a}}(t) - \dot{A}(t)A^T(t)\mathbf{a}(t).$$

Proof. By Equ. (3.24), $\mathbf{x}^0 = A^T \cdot (\mathbf{x} - \mathbf{a})$. We insert this into the expression for $\dot{\mathbf{x}}(t)$ and get

$$\mathbf{v}(\mathbf{x}) = \dot{A}A^T\mathbf{x} + \dot{\mathbf{a}} - \dot{A}A^T\mathbf{a} =: C \cdot \mathbf{x} + \bar{\mathbf{c}}.$$

Differentiating $AA^T = E_3$ shows that $\dot{A}A^T + A\dot{A}^T = C + C^T = 0$, so the matrix $C = \dot{A}A^T$ is skew-symmetric. It therefore has the form described above, and it is elementary to verify that $(c_1, c_2, c_3) \times \mathbf{x} = C \cdot \mathbf{x}$ for all \mathbf{x} . \square

The vector \mathbf{c} in Equ. (3.25) is called the *Darboux vector*. We see that at a given time t , all velocity vectors are zero if and only if $\mathbf{c} = \bar{\mathbf{c}} = \mathbf{0}$, and that otherwise there are always points with nonzero velocity. The former case is called a *stationary instant* of the motion.

Theorem 3.4.2. *In every non-stationary instant t of a smooth one-parameter motion, the velocities of points agree with the velocities of a uniform helical motion, a uniform rotation, or a uniform translation. The path normals $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ are the lines of the linear complex $\mathcal{C} : \bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0$, where $\mathbf{c}, \bar{\mathbf{c}}$ are given by Equ. (3.25).*

Proof. The path normals $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ at the point \mathbf{x} satisfy

$$0 = \mathbf{v} \cdot \mathbf{l} = \bar{\mathbf{c}} \cdot \mathbf{l} + (\mathbf{c} \times \mathbf{x}) \cdot \mathbf{l} = \bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot (\mathbf{x} \times \mathbf{l}) = \bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}}.$$

This is the equation of a linear complex \mathcal{C} . The rest follows from Sec. 3.1.2. \square

The three cases mentioned in Th. 3.4.2 are distinguished as follows: If $\mathbf{c} = \mathbf{0}$, the path normals agree with those of a translation parallel to $\bar{\mathbf{c}}$. Otherwise the path normals are those of a rotation or helical motion, whose pitch can be computed with Equ. (3.9). Zero pitch ($\mathbf{c} \cdot \bar{\mathbf{c}} = 0$) characterizes rotations. The velocity vector field $\mathbf{v}(\mathbf{x}) = \bar{\mathbf{c}} + \mathbf{c} \times \mathbf{x}$ of a smooth motion is called an *infinitesimal motion*. It is determined by the pair $(\mathbf{c}, \bar{\mathbf{c}})$. The previous paragraph shows how to distinguish different cases of infinitesimal motions: Infinitesimal translations with $\mathbf{c} = \mathbf{0}$, infinitesimal rotations with $\mathbf{c} \cdot \bar{\mathbf{c}} = 0$, and infinitesimal helical motions with $\mathbf{c} \cdot \bar{\mathbf{c}} \neq 0$. If the velocity vector field at time t is an infinitesimal rotation, then the line with Plücker coordinates $(\mathbf{c}, \bar{\mathbf{c}})$ is called the *instantaneous axis* at time t .

Note that a path normal is orthogonal to trajectories in all of its points. This is the reason why we consider path normal *lines* instead of path normal *vectors* attached to points.

Remark 3.4.1. Chasles' theorem [18] states that every rigid body motion is a helical motion. Th. 3.4.2 is an infinitesimal version of this theorem: It states that any smooth family of rigid body motions agrees with helical motions in its first derivatives. \diamond

One-parameter Subgroups of Motions

The one-parameter motions with constant velocity vector field play an important role. We begin with a definition which also applies in a more general setting:

Definition. A smooth one-parameter family $\alpha(t)$ of Euclidean motions which fulfill $\alpha(t) \circ \alpha(s) = \alpha(t + s)$ is called a one-parameter group of Euclidean motions.

The following theorem characterizes one-parameter groups among one-parameter motions:

Theorem 3.4.3. A one-parameter motion $\alpha(t)$ is a one-parameter group if and only if its velocity vector field is constant, which means that $v(x) = \bar{c} + c \times x$ does not depend on t . In that case, $\alpha(t)$ is either the translation $x \mapsto x + tv$, or a uniform rotation about an axis with constant angular velocity, or a uniform helical motion as described by Equ. (3.6).

Proof. Assume that $\alpha(t)$ maps x^0 to $A(t) \cdot x^0 + a(t)$. The group property implies $\alpha(t+0) = \alpha(t)\alpha(0)$, so $\alpha(0)$ is the identity mapping. Further it shows that $A(t+s)x^0 + a(t+s) = A(t)(A(s)x^0 + a(s)) + a(t)$ for all s, t, x^0 . Differentiating this identity with respect to t and evaluating at $t = 0$ yields

$$\dot{A}(s)x^0 + \dot{a}(s) = \dot{A}(0)(A(s)x^0 + a(s)) + \dot{a}(0). \quad (3.26)$$

If we let $x = A(s)x^0 + a(s)$, this shows that $v(x)$ is the same at time 0 and at time s . On the one hand, the differential equation (3.26) has a unique solution $A(s), a(s)$ for given initial values. On the other hand, we have already found one-parameter motions which fulfill Equ. (3.26): the one-parameter motions described in this theorem. \square

k -Parameter Motions

If an orthogonal matrix $A \in SO_3$ and a vector $a \in \mathbb{R}^3$ depend smoothly on k parameters, they determine a k -parameter motion. The path or trajectory surface of a point x^0 is then parametrized by

$$x(u) = A(u) \cdot x^0 + a(u), \quad u = (u_1, \dots, u_k) \in D \subset \mathbb{R}^k. \quad (3.27)$$

Fig. C.2 shows an example of a two-parameter motion with some path surfaces. Since the group of Euclidean displacements is six-dimensional, the cases $k = 1, \dots, 5$ are of main interest from a geometric point of view. A k -parameter motion is here denoted by the symbol M^k .

The restriction of M^k to a smooth curve $\mathbf{u}(t)$ in its parameter domain is a one-parameter motion. For all $\mathbf{u}_0 \in D$, there are k special one-parameter motions M_j^1 ($j = 1, \dots, k$) defined by

$$\mathbf{u}_j(t) = \mathbf{u}_0 + t \cdot \mathbf{e}_j,$$

where \mathbf{e}_j denotes the j -th canonical basis vector of \mathbb{R}^k . According to Prop. 3.4.1, the velocity vector fields $\mathbf{v}_j(\mathbf{x})$ are of the form

$$\mathbf{v}_j(\mathbf{x}) = \frac{\partial A}{\partial u_j} \mathbf{x}^0 + \frac{\partial \mathbf{a}}{\partial u_j} = \bar{\mathbf{c}}_j(\mathbf{u}) + \mathbf{c}_j(\mathbf{u}) \times \mathbf{x}. \quad (3.28)$$

The infinitesimal motions $\mathbf{v}_j(\mathbf{x})$ have path normal complexes with equations

$$\mathcal{C}_j : \bar{\mathbf{c}}_j \cdot \mathbf{l} + \mathbf{c}_j \cdot \bar{\mathbf{l}} = 0.$$

A general one-parameter motion, defined by the curve $\mathbf{u}(t)$, has the velocity vector field

$$\dot{\mathbf{x}}(t) = \frac{\partial \mathbf{x}}{\partial u_1} \dot{u}_1(t) + \dots + \frac{\partial \mathbf{x}}{\partial u_k} \dot{u}_k(t) = \mathbf{v}(\mathbf{x}) = \sum_{j=1}^k \dot{u}_j \mathbf{v}_j(\mathbf{x}). \quad (3.29)$$

The infinitesimal motion $\mathbf{v}(\mathbf{x})$ is determined by $\dot{\mathbf{u}} = (\dot{u}_1, \dots, \dot{u}_k)$. Its path normal complex \mathcal{C} (with coordinate vector $(\mathbf{c}, \bar{\mathbf{c}})$) is, by linearity, related to the path normal complexes \mathcal{C}_j by

$$(\mathbf{c}, \bar{\mathbf{c}}) = \sum_{j=1}^k \dot{u}_j (\mathbf{c}_j, \bar{\mathbf{c}}_j). \quad (3.30)$$

Lemma 3.4.4. *For all instants \mathbf{u} , the set of possible path normals complexes \mathcal{C} of one-parameter motions contained in a k -parameter motion M^k comprise a linear manifold of complexes spanned by $\mathcal{C}_1, \dots, \mathcal{C}_k$.*

Proof. Equ. (3.30) shows that the set of possible coordinate vectors equals the linear span of the coordinate vectors $(\mathbf{c}_j, \bar{\mathbf{c}}_j)$. \square

The dimension of this linear manifold of path normal complexes has a geometric meaning:

Definition. *The infinitesimal degree of freedom of a k -parameter motion M^k at an instant \mathbf{u} is the dimension of the linear space of infinitesimal one-parameter motions at this instant. M^k is called regular at \mathbf{u} if this number equals k , otherwise singular.*

Equations (3.28), (3.27), and (3.30) show that the infinitesimal degree of freedom of M^k equals the dimension of the linear span of the coordinate vectors $[(\mathbf{c}_1, \bar{\mathbf{c}}_1), \dots, (\mathbf{c}_k, \bar{\mathbf{c}}_k)]$. The projective dimension of the linear manifold $\mathcal{C}_1 \gamma^* \vee \dots \vee \mathcal{C}_k \gamma^*$ of path normal complexes equals the infinitesimal degree of freedom minus 1.

Remark 3.4.2. A k -parameter motion is regular if and only if it is regular as a smooth mapping of D into the Euclidean motion group. \diamond

2-Parameter Motions

Consider a two-parameter motion M^2 . The paths of points $\mathbf{x}^0 \in \Sigma^0$ are 2-surfaces of Σ . We want to describe its normals — the *path normals* — in line-geometric terms. If the path is a regular surface, it has a unique surface normal. In a singular point it may have a pencil or even a bundle of normals.

Corollary 3.4.5. *If \mathbf{u} is a regular instant of a two-parameter motion there is a pencil G of linear complexes such that its carrier $C(G)$ coincides with the set of path normals.*

Proof. A line is path normal for the path surface, if it is path normal for all paths of one-parameter motions contained in this path surface. Now Lemma 3.4.4 now shows the result. \square

This allows a classification of regular instants of two-parameter motions according to their derivatives. The pencil G of complexes assigned to a regular instant may be singular, parabolic, hyperbolic, or elliptic. This means that there are, up to a scalar factor, ∞ , 1, 2 or 0 infinitesimal one-parameter motions which are infinitesimal rotations or translations, i.e., their path normal complex is singular.

If the pencil G is nonsingular, its carrier is a hyperbolic, parabolic or elliptic linear congruence \mathcal{N} , whose focal lines are the axes of the infinitesimal rotations contained in M^2 .

Remark 3.4.3. The fact that the set of path normals is a linear congruence \mathcal{N} has several consequences. One is the following: For all points of Σ^0 , except those in the focal lines of \mathcal{N} , there is exactly one path normal, and this line is orthogonal to the path surfaces in all of its points. \diamond

3-Parameter Motions

If M^3 is a three-parameter motion $\mathbf{x}(\mathbf{u}) = A(\mathbf{u}) \cdot \mathbf{x}^0 + \mathbf{a}(\mathbf{u})$, the paths of points are expected to fill out a three-dimensional subset of Euclidean space. If the linear space of velocity vectors $\mathbf{v}(\mathbf{x})$ of some fixed point \mathbf{x} has dimension three at some instant \mathbf{u}_0 , then the inverse function theorem says that the mapping $\mathbf{u} \mapsto \mathbf{x}(\mathbf{u})$ is a diffeomorphism in a neighbourhood of \mathbf{u}_0 . It is therefore important to know for which points this dimension drops. Such points have path normals and are called *singular*.

Corollary 3.4.6. *For all regular instants \mathbf{u} there is, in general, a ruled quadric Φ (possibly non-real) of singular points. One regulus of Φ comprises the set of path normals, and the lines of the other one are the axes of infinitesimal rotations contained in M^3 .*

The meaning of ‘in general’ will be explained in the proof.

Proof. The point \mathbf{x}^0 is singular at the instant \mathbf{u} if and only if it has a path normal. This path normal must be contained in the path normal complexes for all one-parameter motions contained in M^3 . By Lemma 3.4.4, we have to determine the carrier of the bundle G spanned by C_1, C_2, C_3 . If this carrier is a regulus \mathcal{R} (this is the ‘general case’) (cf. the discussion at the beginning of Sec. 3.3), then the set of points in the carrier is a ruled quadric Φ . A line L of \mathcal{R} is a path normal for all points in L .

The singular complexes in G belong to infinitesimal rotations or translations. Th. 3.3.3 shows that its set of axes $S(G)$ equals \mathcal{R} ’s complementary regulus. \square

Remark 3.4.4. The path (orbit) of a point under the action of a three-parameter motion may have boundary points. Boundary points either stem from boundary points \mathbf{u} of the parameter domain D of the motion or they are singular points at the appropriate instant. Points whose path is a smooth surface are singular at all instants. \diamond

The classification of regular instants of a three-parameter motion with respect to its first derivatives, i.e., its path normals, coincides with the classification of bundles of linear complexes, as has already been indicated in the proof of Cor. 3.4.6. Each of the cases listed in Sec. 3.3 can occur. There are four singular cases and two regular ones. In one case there are no singular points: The plane $C_1\gamma^* \vee C_2\gamma^* \vee C_3\gamma^*$ need not intersect the Klein quadric, so the carrier of the bundle G can be empty. We should also remark that this is only a projective classification of the singular set. An affine or Euclidean classification would have much more different cases.

Relative Motions

We consider three copies $\Sigma^0, \Sigma^1, \Sigma^2$ of Euclidean space. Σ^0 performs a one-parameter motion with respect to Σ^1 and Σ^1 does the same with respect to Σ^2 . These motions are denoted by

$$\mathbf{x}^1(t) = A_{01}(t) \cdot \mathbf{x}^0 + \mathbf{a}_{01}(t), \quad \mathbf{x}^2(t) = A_{12}(t) \cdot \mathbf{x}^1 + \mathbf{a}_{12}(t). \quad (3.31)$$

The *composite motion* of Σ^0 with respect to Σ^2 is then defined as

$$\mathbf{x}^2(t) = A_{12}(t)A_{01}(t) \cdot \mathbf{x}^0 + A_{12}(t)\mathbf{a}_{01}(t) + \mathbf{a}_{12}(t) = A_{02}(t)\mathbf{x}^0 + \mathbf{a}_{02}(t). \quad (3.32)$$

We want to derive a relation between the velocity vector fields $\mathbf{v}_{01}(\mathbf{x}^1), \mathbf{v}_{12}(\mathbf{x}^2)$, and $\mathbf{v}_{02}(\mathbf{x}^2)$, where the superscript of the argument indicates its domain. It is easier to consider \mathbf{v}_{01} as a function of \mathbf{x}^2 as well, where the connection between \mathbf{x}^1 and \mathbf{x}^2 is given by Equ. (3.31). Then $\mathbf{v}_{01}(\mathbf{x}^2) = A_{12} \cdot \mathbf{v}_{01}(\mathbf{x}^1)$.

Corollary 3.4.7. *The velocity vector fields defined by the one-parameter motions of Equ. (3.31) and (3.32) are related by*

$$\mathbf{v}_{02}(\mathbf{x}^2) = \mathbf{v}_{01}(\mathbf{x}^2) + \mathbf{v}_{12}(\mathbf{x}^2). \quad (3.33)$$

The three path normal complexes are contained in a pencil.

Proof. Differentiating Equ. (3.32) shows that $\dot{\mathbf{x}}^2$ equals the sum $[A_{12}(A_{01} \cdot \mathbf{x}^0 + \mathbf{a}_{01}) + \dot{\mathbf{a}}_{12}] + [A_{12}(A_{01} \cdot \mathbf{x}^0 + \dot{\mathbf{a}}_{01})]$. The first summand equals $\mathbf{v}_{12}(\mathbf{x}^2)$, and the latter $\mathbf{v}_{01}(\mathbf{x}^2)$. \square

3.4.2 Statics and Screw Theory

Linear complexes occur not only in spatial kinematics, but also in statics and the corresponding ‘screw theory’, whose name already indicates this connection. The statics of forces acting on a rigid body is founded on the following principles:

- I. Forces acting on the same particle are equivalent if their vector sums agree.
(The principle of addition of forces).
- II. A force acting on a rigid body may be shifted along its line of action so as to act on any particle of that line (Principle of transmissibility of a force).
- III. A rigid body is in static equilibrium if and only if the complete set of forces acting on it can be reduced to zero by principles I and II.

According to principle II, we may view a force as a line bound vector \mathbf{f} and represent it by \mathbf{f} and the moment vector $\bar{\mathbf{f}}$ of \mathbf{f} about the origin: We say the force has coordinates $\mathbf{F} = (\mathbf{f}, \bar{\mathbf{f}})$. The homogeneous Plücker coordinates of its line of action are then given by $(\mathbf{f}, \bar{\mathbf{f}})\mathbb{R}$.

Systems of Forces

If two forces $(\mathbf{p}, \bar{\mathbf{p}})$ and $(\mathbf{q}, \bar{\mathbf{q}})$ act on a particle with vector \mathbf{x} , then

$$(\mathbf{p}, \mathbf{x} \times \mathbf{p}) = (\mathbf{p}, \bar{\mathbf{p}}), \quad (\mathbf{q}, \mathbf{x} \times \mathbf{q}) = (\mathbf{q}, \bar{\mathbf{q}}).$$

By principle I, both forces may be replaced by the resultant force $(\mathbf{f}, \bar{\mathbf{f}}) = (\mathbf{p} + \mathbf{q}, \mathbf{x} \times (\mathbf{p} + \mathbf{q}))$.

Analogously, two parallel forces $(\mathbf{p}, \bar{\mathbf{p}})$, $(\mathbf{q}, \bar{\mathbf{q}})$ possess a resultant force which equals their sum, provided that $\mathbf{p} + \mathbf{q} \neq \mathbf{o}$. To see this, we introduce, by principle I, a pair of opposite forces $\pm(\mathbf{f}, \bar{\mathbf{f}})$ acting on a line which intersects the lines of $(\mathbf{p}, \bar{\mathbf{p}})$ and $(\mathbf{q}, \bar{\mathbf{q}})$. The resultant forces $(\mathbf{p}, \bar{\mathbf{p}}) + (\mathbf{f}, \bar{\mathbf{f}})$ and $(\mathbf{q}, \bar{\mathbf{q}}) - (\mathbf{f}, \bar{\mathbf{f}})$ act on intersecting lines, because they are not parallel:

$$(\mathbf{p} + \mathbf{f}) \times (\mathbf{q} - \mathbf{f}) = \mathbf{f} \times (\mathbf{p} + \mathbf{q}) \neq \mathbf{o}.$$

Their resultant force is therefore given by $(\mathbf{p}, \bar{\mathbf{p}}) + (\mathbf{f}, \bar{\mathbf{f}}) + (\mathbf{q}, \bar{\mathbf{q}}) - (\mathbf{f}, \bar{\mathbf{f}}) = (\mathbf{p}, \bar{\mathbf{p}}) + (\mathbf{q}, \bar{\mathbf{q}})$.

If $\mathbf{p} + \mathbf{q} = \mathbf{o}$ and $\bar{\mathbf{p}} + \bar{\mathbf{q}} \neq \mathbf{o}$, the parallel forces $(\mathbf{p}, \bar{\mathbf{p}})$, $(\mathbf{q}, \bar{\mathbf{q}})$ are equal in magnitude and opposite in direction. They are said to form a *couple* of forces, which has the moment $\bar{\mathbf{p}} + \bar{\mathbf{q}}$. Its line of action may be defined as the ideal line with Plücker coordinates $(\mathbf{o}, \bar{\mathbf{p}} + \bar{\mathbf{q}})$.

If we transform a system of forces $(\mathbf{f}_i, \bar{\mathbf{f}}_i)$ into an equivalent system $(\mathbf{g}_i, \bar{\mathbf{g}}_i)$ by principles I and II, both the force and moment sums remain unchanged,

$$\sum \mathbf{f}_i = \sum \mathbf{g}_i, \quad \sum \bar{\mathbf{f}}_i = \sum \bar{\mathbf{g}}_i. \quad (3.34)$$

Conversely, it can be shown that (3.34) characterizes equivalent force systems. It makes therefore sense to represent the system of forces by

$$(\mathbf{m}, \bar{\mathbf{m}}) = (\sum \mathbf{f}_i, \sum \bar{\mathbf{f}}_i), \quad (3.35)$$

which is sometimes called a *motor*.

Reduction of Force Systems

The following cases can be distinguished:

- A. $\mathbf{m} \neq \mathbf{o}$, $\Omega_q((\mathbf{m}, \bar{\mathbf{m}})) = \mathbf{m} \cdot \bar{\mathbf{m}} \neq 0$: This is the generic case. The system of forces cannot be reduced to a single force. The motor $(\mathbf{m}, \bar{\mathbf{m}})$ is also called a *screw*. There is a regular linear complex \mathcal{M} with $\mathcal{M}\gamma^* = (\mathbf{m}, \bar{\mathbf{m}})\mathbb{R} \in P^5$.
- B. $\mathbf{m} \neq \mathbf{o}$, $\mathbf{m} \cdot \bar{\mathbf{m}} = 0$: The system of forces is equivalent to the single force $(\mathbf{m}, \bar{\mathbf{m}})$ acting along the line with Plücker coordinates $(\mathbf{m}, \bar{\mathbf{m}})$.
- C. $\mathbf{m} = \mathbf{o}$, $\bar{\mathbf{m}} \neq \mathbf{o}$: This force system is equivalent to the couple $(\mathbf{o}, \bar{\mathbf{m}})$ with moment vector $\bar{\mathbf{m}}$.
- D. $\mathbf{m} = \bar{\mathbf{m}} = \mathbf{o}$: The system reduces to zero and according to principle III the rigid body is in *equilibrium*.

Proposition 3.4.8. *A generic system $(\mathbf{m}, \bar{\mathbf{m}})$ of forces is in infinitely many ways equivalent to a system of two forces $(\mathbf{p}, \bar{\mathbf{p}})$ and $(\mathbf{q}, \bar{\mathbf{q}})$, acting on skew lines P and Q . If π is the null polarity associated with $(\mathbf{m}, \bar{\mathbf{m}})$, then $P\pi = Q$.*

Proof. Denote the linear complex with coordinate vector $(\mathbf{m}, \bar{\mathbf{m}})$ with \mathcal{M} . Choose a line $G \notin \mathcal{M}$ such that $G\pi$ is no ideal line. We assume that it has Plücker coordinates $(\mathbf{g}, \bar{\mathbf{g}})$, and we temporarily let $\alpha = \Omega((\mathbf{m}, \bar{\mathbf{m}}), (\mathbf{g}, \bar{\mathbf{g}}))$, $\beta = \Omega_q(\mathbf{m}, \bar{\mathbf{m}})$. The condition $G \notin \mathcal{M}$ means that $\alpha \neq 0$, and $\beta \neq 0$ because the system of forces is ‘generic’. By Equ. (3.5), $H = G\pi$ has the Plücker coordinates

$$(\mathbf{h}, \bar{\mathbf{h}}) = \beta(\mathbf{g}, \bar{\mathbf{g}}) - \alpha(\mathbf{m}, \bar{\mathbf{m}}).$$

We let $(\mathbf{p}, \bar{\mathbf{p}}) = (\beta/\alpha)(\mathbf{g}, \bar{\mathbf{g}})$ and $(\mathbf{q}, \bar{\mathbf{q}}) = (-1/\alpha)(\mathbf{h}, \bar{\mathbf{h}})$. Then $(\mathbf{m}, \bar{\mathbf{m}}) = (\mathbf{p}, \bar{\mathbf{p}}) + (\mathbf{q}, \bar{\mathbf{q}})$ and we are done. \square

Proposition 3.4.9. (Poinsot's theorem) *A generic system of forces is equivalent to a single force and a couple whose moment vector is parallel to this force.*

Proof. In the proof of Prop. 3.4.8 choose G as the axis of \mathcal{M} . Then H is at infinity, and represents a couple of forces whose moment is parallel to G . \square

The Complex of Vanishing Moments

The *moment* \mathbf{f}_p of a force $(\mathbf{f}, \bar{\mathbf{f}})$ about a point \mathbf{p} is defined by

$$\mathbf{f}_p = (\mathbf{x} - \mathbf{p}) \times \mathbf{f} = \bar{\mathbf{f}} - \mathbf{p} \times \mathbf{f},$$

where \mathbf{x} is an arbitrary point on the line of action. Recall that $\bar{\mathbf{f}}$ is the moment with respect to the origin. The moment $m((\mathbf{f}, \bar{\mathbf{f}}), \vec{A})$ of the force *about an oriented line* \vec{A} with normalized Plücker coordinates $(\mathbf{a}, \bar{\mathbf{a}})$ ($\|\mathbf{a}\| = 1$) is defined by

$$m((\mathbf{f}, \bar{\mathbf{f}}), \vec{A}) = \mathbf{a} \cdot \mathbf{f}_p,$$

where \mathbf{p} is an arbitrary point of \overrightarrow{A} . If we select a different point, the result is the same because

$$m((\mathbf{f}, \bar{\mathbf{f}}), \overrightarrow{A}) = \mathbf{a} \cdot (\bar{\mathbf{f}} - \mathbf{p} \times \mathbf{f}) = \mathbf{a} \cdot \bar{\mathbf{f}} + \bar{\mathbf{a}} \cdot \mathbf{f}.$$

If a system of forces $(\mathbf{f}_i, \bar{\mathbf{f}}_i)$ is equivalent to $(\mathbf{m}, \bar{\mathbf{m}})$ then its moment about the oriented line \overrightarrow{A} equals per definition

$$m((\mathbf{m}, \bar{\mathbf{m}}), \overrightarrow{A}) = \sum m((\mathbf{f}_i, \bar{\mathbf{f}}_i), \overrightarrow{A}) = \mathbf{a} \cdot \bar{\mathbf{m}} + \bar{\mathbf{a}} \cdot \mathbf{m}. \quad (3.36)$$

This leads to a physical interpretation of the linear complex associated with $(\mathbf{m}, \bar{\mathbf{m}})$:

Corollary 3.4.10. *If the moment of a system $(\mathbf{m}, \bar{\mathbf{m}})$ of forces about an oriented line \overrightarrow{A} vanishes, then this line is contained in the linear complex with coordinate vector $(\mathbf{m}, \bar{\mathbf{m}})$.*

Virtual Work and Screw Theory

Both in kinematics and in statics, ‘motors’ appear as a basic entity. A theory which unifies both concepts is R.S. Ball’s *screw theory* [8]. It is based on the concept of *reciprocal screws*: Consider a rigid body and apply a force system to it, which is represented by a screw $(\mathbf{a}, \bar{\mathbf{a}})$. At the same time the body performs an infinitesimal one-parameter motion represented by a screw $(\mathbf{b}, \bar{\mathbf{b}})$.

Definition. *Assume a force $(\mathbf{f}, \bar{\mathbf{f}})$ acting on a point \mathbf{x} which moves with velocity \mathbf{v} . The virtual work done by this force is defined as the product $\mathbf{f} \cdot \mathbf{v}$. The virtual work of a collection of forces is the sum of the single virtual works.*

Lemma 3.4.11. *Let forces $(\mathbf{f}_i, \bar{\mathbf{f}}_i)$ act on points \mathbf{x}_i , which have velocity vectors $\mathbf{v}_i = \bar{\mathbf{c}} + \mathbf{c} \times \mathbf{x}_i$, and let $(\mathbf{m}, \bar{\mathbf{m}}) = \sum (\mathbf{f}_i, \bar{\mathbf{f}}_i)$. Then the sum of virtual works of these forces equals*

$$\Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{m}, \bar{\mathbf{m}})) = \mathbf{c} \cdot \bar{\mathbf{m}} + \bar{\mathbf{c}} \cdot \mathbf{m}. \quad (3.37)$$

Proof. Summing the single virtual works yields $\sum \mathbf{v}(\mathbf{x}_i) \cdot \mathbf{f}_i = \sum (\mathbf{c} \times \mathbf{x}_i + \bar{\mathbf{c}}) \cdot \mathbf{f}_i = \sum (\mathbf{c} \cdot (\mathbf{x}_i \times \mathbf{f}_i) + \bar{\mathbf{c}} \cdot \mathbf{f}_i) = \mathbf{c} \cdot \sum \bar{\mathbf{f}}_i + \bar{\mathbf{c}} \cdot \sum \mathbf{f}_i$. \square

A force system $(\mathbf{m}, \bar{\mathbf{m}})$ defines a point $(\mathbf{m}, \bar{\mathbf{m}})\mathbb{R}$ in P^5 . Clearly this point does not uniquely determine the force system. Analogously an infinitesimal motion $(\mathbf{c}, \bar{\mathbf{c}})$ defines a point of P^5 . Although information about the magnitude of forces and velocities is lost, there is the following theorem, which follows directly from Lemma 3.4.11.

Theorem 3.4.12. *The virtual work done under the circumstances of Lemma 3.4.11 vanishes if and only if the points $(\mathbf{m}, \bar{\mathbf{m}})\mathbb{R}$ and $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ are conjugate with respect to the Klein quadric.*

A pair of such screws with zero virtual work are called *reciprocal*.

Virtual Work and k -Parameter Motions

All the screws belonging to an instant of a k -parameter motion are called a *screw system of order k* or a k -system of screws. By Lemma 3.4.4, the associated points in P^5 form a $(k - 1)$ -dimensional subspace which is the Klein image of a $(k - 1)$ -dimensional linear manifold of linear complexes.

The screws reciprocal to all elements of a k -system are called its *reciprocal system*. Reciprocal systems are clearly related by the polarity of the Klein quadric, which transforms a $(k - 1)$ -dimensional projective subspace of P^5 into a $(5 - k)$ -dimensional one. The reciprocal system is therefore of order $6 - k$.

The cases $k = 1, 2, 3$ correspond via this polarity to the cases $k = 5, 4, 3$, so for a projective classification it is sufficient to consider the cases $k = 1, 2, 3$ only. Clearly the classification of screw systems from the projective point of view is nothing but the classification of linear manifolds of line complexes, which has been done in Sec. 3.2 and Sec. 3.3.

Remark 3.4.5. The *affine* classification of linear manifolds of line complexes (and of screw systems) is straightforward. If a projective subspace U is assigned to such a linear manifold we have to distinguish several ‘cases’, depending on whether U is an ideal subspace, or depending on the dimension of the intersection of U with the ideal plane. A Euclidean classification can be found in [84]. \diamond

Remark 3.4.6. An extension of screw theory has been developed by Ohwovorile and Roth (cf. [134]). They consider two rigid bodies A and B which are in contact at a point. Let A be acted upon by a force system $(\mathbf{b}, \bar{\mathbf{b}})$ due to contact with the body B , while undergoing an infinitesimal motion $(\mathbf{a}, \bar{\mathbf{a}})$. Under appropriate assumptions, the sign of the virtual work (3.37) has a physical meaning. If this sign is positive, A moves infinitesimally away from B . The corresponding screw pair is called *repelling*. For negative sign, A would move into B (which cannot happen for rigid bodies) and they speak of a *contrary* screw pair.

In this extension of the theory one considers ‘oriented’ screws, which can be considered as points of the 5-sphere, the double covering of P^5 [190]. \diamond

LuraTech

www.luratech.com

4. Approximation in Line Space

4.1 Fitting Linear Complexes

The applications of the concept of linear line complex discussed in Sec. 3.4 show that it is important to study the problem of approximation of and with linear complexes. We consider the following questions: which linear complex fits a given set of data lines best? What is an appropriate definition of ‘best’ for various applications? For this we have to define distance functions for lines and linear complexes, which make the problem computationally tractable. It turns out that most approximation problems can be accessed by least-square methods.

The Moment of a Line with respect to a Linear Complex

Consider Euclidean three-dimensional space E^3 , a line L and a linear complex \mathcal{C} . Assume that their homogeneous Plücker coordinates are $(\mathbf{l}, \bar{\mathbf{l}})$ and $(\mathbf{c}, \bar{\mathbf{c}})$, respectively, with $\mathbf{c} \neq \mathbf{o}$, $\mathbf{l} \neq \mathbf{o}$ (i.e., L is not an ideal line, and \mathcal{C} is not the path normal complex of a translation). The following definition was given by F. Klein:

Definition. *Under the assumptions above, the moment of the line L with respect to the linear complex \mathcal{C} is defined by*

$$m(L, \mathcal{C}) = \frac{1}{\|\mathbf{c}\| \|\mathbf{l}\|} |\bar{\mathbf{c}} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}}| = \frac{1}{\|\mathbf{c}\| \|\mathbf{l}\|} |\Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{l}, \bar{\mathbf{l}}))|. \quad (4.1)$$

If we use *normalized* Plücker coordinates for L and \mathcal{C} , which means $\|\mathbf{l}\| = \|\mathbf{c}\| = 1$, then (4.1) becomes $m(L, \mathcal{C}) = |\Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{l}, \bar{\mathbf{l}}))|$.

Remark 4.1.1. As we did with lines, we can assign two *oriented* linear complexes to a linear complex \mathcal{C} . If \mathcal{C} has Plücker coordinates $(\mathbf{c}, \bar{\mathbf{c}})$ with $\mathbf{c} \neq \mathbf{o}$, there are two ways to normalize the vector \mathbf{c} . Any of the coordinate six-tuples $\pm(1/\|\mathbf{c}\|)(\mathbf{c}, \bar{\mathbf{c}})$ is called an oriented linear complex $\vec{\mathcal{C}}$ belonging to \mathcal{C} .

If $\vec{\mathcal{C}} = (\mathbf{c}, \bar{\mathbf{c}})$ with $\|\mathbf{c}\| = 1$, and the line L has the normalized Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$, we let $m(\vec{L}, \vec{\mathcal{C}}) = \Omega((\mathbf{c}, \bar{\mathbf{c}}), (\mathbf{l}, \bar{\mathbf{l}}))$, and call this value the moment of the oriented line \vec{L} with respect to $\vec{\mathcal{C}}$.

Lemma 3.4.11 shows that $m(\vec{L}, \vec{\mathcal{C}})$ has an interpretation as a virtual work. We will not need the moments of oriented lines and complexes, and we will always use Equ. (4.1), which defines $m(L, \mathcal{C})$ to be nonnegative. \diamond

We are interested in the set of lines whose moment with respect to the linear complex \mathcal{C} equals a certain constant. This set has the following geometric description:

Lemma 4.1.1. *Assume that L is a line and \mathcal{C} is a linear complex of pitch p such that $m(L, \mathcal{C})$ is defined. If $P \in L$, r is P 's distance to the axis A of \mathcal{C} and α is the minimum of angles enclosed by L and a line of \mathcal{C} incident with P , then*

$$m(L, \mathcal{C}) = \sqrt{r^2 + p^2} \sin \alpha. \quad (4.2)$$

Proof. The lines of \mathcal{C} are path normals of the unique helical motion with pitch p and axis A . If T denotes the path tangent of P , then obviously $\pi/2 - \alpha = \sphericalangle(L, T)$. Without loss of generality, the helical motion has the form (3.6) and \mathcal{C} therefore the equation (3.8). Now (4.2) follows by a simple computation. \square

We see that the moment $m(L, \mathcal{C})$ vanishes if and only if $L \in \mathcal{C}$. The set \mathcal{K} of lines such that $\sqrt{r^2 + p^2} \sin \alpha$ is constant (with the symbols used in Lemma 4.1.1) is called a *cyclic quadratic complex*. Its complex cones are right circular cones, and its complex curves are circles. We will discuss it in Sec. 7.2.3.

Deviation of a Linear Complex from a Set of Lines

We assume that lines L_1, \dots, L_k are given by their normalized Plücker coordinates $(\mathbf{l}_i, \bar{\mathbf{l}}_i)$ ($\|\mathbf{l}_i\| = 1$). It is fortunate that the moment of a line L_i with respect to a linear complex \mathcal{X} is a bilinear form if we restrict ourselves to normalized Plücker coordinates. The sum of moments equals a quadratic function:

$$\sum_{i=1}^k m(L_i, \mathcal{X})^2 = F(\mathbf{x}, \bar{\mathbf{x}}) = \sum_{i=1}^k (\bar{\mathbf{x}} \cdot \mathbf{l}_i + \mathbf{x} \cdot \bar{\mathbf{l}}_i)^2. \quad (4.3)$$

We use the expression (4.3) as a measure of the deviation of a linear complex \mathcal{X} from the set $\{L_1, \dots, L_k\}$ of lines. The function F is a positive semidefinite quadratic form defined in \mathbb{R}^6 . To compute its coordinate matrix, we observe that the product $(\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d})$ is in matrix notation written as $\mathbf{b}^T \cdot (\mathbf{a} \cdot \mathbf{c}^T) \cdot \mathbf{d}$. Thus we let

$$M = \sum_{i=1}^k \begin{pmatrix} \bar{\mathbf{l}}_i \cdot \bar{\mathbf{l}}_i^T & | & \mathbf{l}_i \cdot \bar{\mathbf{l}}_i^T \\ \hline \bar{\mathbf{l}}_i \cdot \mathbf{l}_i^T & | & \mathbf{l}_i \cdot \mathbf{l}_i^T \end{pmatrix}, \quad (4.4)$$

and we see that F can be expressed in the form

$$F(\mathbf{x}, \bar{\mathbf{x}}) = (\mathbf{x}, \bar{\mathbf{x}})^T \cdot M \cdot (\mathbf{x}, \bar{\mathbf{x}}). \quad (4.5)$$

Computation of an Approximant Complex in the Regular Case

In order to compute a linear complex \mathcal{X} which approximates the given set of lines L_1, \dots, L_k , we look for normalized Plücker coordinates $(\mathbf{x}, \bar{\mathbf{x}})$ which minimize

$F(\mathbf{x}, \bar{\mathbf{x}})$. This approximation problem makes sense only for $k \geq 5$, because otherwise there is at least a one-parameter family of linear complexes which contains all given lines. There is a unique linear complex which contains five lines in general position, so we always assume that $k > 5$.

Lemma 4.1.2. *In the set of Plücker coordinate vectors $(\mathbf{x}, \bar{\mathbf{x}})$ with $\|\mathbf{x}\| = 1$, the function F of Equ. (4.3) attains its minimum λ precisely for those $(\mathbf{x}, \bar{\mathbf{x}})$ which fulfill*

$$(M - \lambda D) \cdot (\mathbf{x}, \bar{\mathbf{x}}) = \mathbf{o}, \quad \|\mathbf{x}\| = 1, \quad (4.6)$$

where $D = \text{diag}(1, 1, 1, 0, 0, 0)$, M is defined by (4.4), and λ is the smallest solution of the equation

$$\det(M - \lambda D) = 0. \quad (4.7)$$

Proof. We have to minimize $F(\mathbf{x}, \bar{\mathbf{x}})$ subject to the side condition $G(\mathbf{x}, \bar{\mathbf{x}}) = \|\mathbf{x}\| = 1$. This side condition is expressed in matrix notation by $(\mathbf{x}, \bar{\mathbf{x}})^T \cdot D \cdot (\mathbf{x}, \bar{\mathbf{x}}) = 1$. We introduce the Lagrangian multiplier λ and solve $\text{grad}_F(\mathbf{x}, \bar{\mathbf{x}}) - \lambda \text{grad}_G(\mathbf{x}, \bar{\mathbf{x}}) = \mathbf{o}$, which simplifies to Equ. (4.6). A nonzero solution obviously is possible if and only if $M - \lambda D$ is singular.

Then $M \cdot (\mathbf{x}, \bar{\mathbf{x}}) = \lambda \cdot D \cdot (\mathbf{x}, \bar{\mathbf{x}})$, and Equ. (4.5) shows that $F(\mathbf{x}, \bar{\mathbf{x}}) = \lambda$. This shows that in order to minimize F we have to choose the smallest possible λ . \square

Remark 4.1.2. The statistical standard deviation of the approximating linear complex \mathcal{X} from the given lines equals

$$\sigma = \sqrt{\lambda/(k-5)}.$$

The smaller λ is, the better the solution complex fits the original data L_1, \dots, L_k . ‘Small’ however has only a relative meaning. If the input data are scaled by a factor α , the value of λ scales with α^2 . So σ , which scales linearly with the input data, and whose definition accounts for the number of input lines, should be compared with the size of objects under discussion. \diamond

Complexes of Zero Pitch

If the pitch p of the approximant complex \mathcal{C} turns out to be very small, one might be interested in fitting the input data by a singular linear complex \mathcal{C}_0 of zero pitch. By Th. 3.1.3, such a complex consists of all lines which intersect a line A_0 . Thus we have to determine A_0 such that it ‘almost’ intersects the input lines L_1, \dots, L_k .

The simplest choice is to take the axis of the approximant complex, computed by Equ. (3.9). We achieve a better result if we minimize within the set of singular complexes:

Lemma 4.1.3. *In the set of normalized Plücker coordinates $(\mathbf{x}, \bar{\mathbf{x}})$ of singular linear complexes (i.e., $\|\mathbf{x}\| = 1, \mathbf{x} \cdot \bar{\mathbf{x}} = 0$), the function F of Equ. (4.3) attains its minimum λ for a pair $(\mathbf{x}, \bar{\mathbf{x}})$ which fulfills*

$$(M - \lambda D - \mu K) \cdot (\mathbf{x}, \bar{\mathbf{x}}) = \mathbf{o}, \quad \|\mathbf{x}\| = 1, \quad \mathbf{x} \cdot \bar{\mathbf{x}} = 0, \quad (4.8)$$

where D is as in Lemma 4.1.2, M is given by Equ. (4.4), K is from Equ. (2.28), and λ, μ satisfy

$$\det(M - \lambda D - \mu K) = 0. \quad (4.9)$$

Proof. With the matrices M , D , and K the minimization problem $F(\mathbf{x}, \bar{\mathbf{x}}) \rightarrow \min$, $\mathbf{x}^2 = 1$, $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$ is transformed into $F(\mathbf{x}, \bar{\mathbf{x}}) = (\mathbf{x}, \bar{\mathbf{x}})^T \cdot M \cdot (\mathbf{x}, \bar{\mathbf{x}}) \rightarrow \min$, $G(\mathbf{x}, \bar{\mathbf{x}}) = (\mathbf{x}, \bar{\mathbf{x}})^T \cdot D \cdot (\mathbf{x}, \bar{\mathbf{x}}) = 1$, $H(\mathbf{x}, \bar{\mathbf{x}}) = (\mathbf{x}, \bar{\mathbf{x}})^T \cdot K \cdot (\mathbf{x}, \bar{\mathbf{x}}) = 0$.

We introduce two Lagrangian multipliers λ, μ and solve $(\text{grad}_F - \lambda \text{grad}_G - \mu \text{grad}_H)(\mathbf{x}, \bar{\mathbf{x}}) = 0$, $G(\mathbf{x}, \bar{\mathbf{x}}) = 1$, $H(\mathbf{x}, \bar{\mathbf{x}}) = 0$. This simplifies to Equ. (4.8). The nonzero vector $(\mathbf{x}, \bar{\mathbf{x}})$ is contained in the kernel of $M - \lambda D - \mu K$, which implies Equ. (4.9). Equ. (4.8) immediately shows that $F(\mathbf{x}, \bar{\mathbf{x}}) = \lambda$. \square

The solution of this problem is not as straightforward as that of the previous one. Equ. (4.9) defines an algebraic curve in the (λ, μ) -plane. The set of solutions of (4.9) together with the equations $(M - \lambda D - \mu K) \cdot (\mathbf{x}, \bar{\mathbf{x}}) = \mathbf{o}$, $\|\mathbf{x}\| = 1$ in the variables $\lambda, \mu, \mathbf{x}, \bar{\mathbf{x}}$ is an algebraic variety in \mathbb{R}^8 . Its projection onto the coordinate space \mathbb{R}^6 of variables $\mathbf{x}, \bar{\mathbf{x}}$ is, in the generic case, an algebraic curve Φ , because then for all pairs (λ, μ) satisfying (4.9) there are exactly two solutions of the other two equations.

The solutions of the minimization problem are contained in the intersections of the curve Φ with the Klein quadric, which corresponds to the equation $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$.

Because of the high degree of the problem, we do not pursue further its algebraic aspects. The solution is best computed numerically. A good starting point for an iterative algorithm would be the approximant complex computed by minimizing (4.3).

Complexes of Infinite Pitch

Linear complexes \mathcal{C} of infinite pitch consist of lines which intersect a certain line at infinity. In Euclidean space, the proper lines of such a complex comprise the set of lines orthogonal to a certain vector $\bar{\mathbf{c}}$. A line L with Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ is in \mathcal{C} if and only if $\mathbf{l} \cdot \bar{\mathbf{c}} = 0$.

Such complexes have been excluded from the discussion so far, because we worked with moments and they are defined only for complexes of finite pitch.

We define the deviation of a line L of Euclidean \mathbb{R}^3 from \mathcal{C} by $|\cos \sphericalangle(L, \bar{\mathbf{c}})|$. The deviation of a complex \mathcal{X} with Plücker coordinates $(\mathbf{o}, \bar{\mathbf{x}})$ with $\|\bar{\mathbf{x}}\| = 1$ from a given finite set L_1, \dots, L_k of lines with normalized Plücker coordinates $(\mathbf{l}_i, \bar{\mathbf{l}}_i)$ is defined by

$$F_\infty(\bar{\mathbf{x}}) = \sum_{i=1}^k (\bar{\mathbf{x}} \cdot \mathbf{l}_i)^2. \quad (4.10)$$

The minimization of $F_\infty(\bar{\mathbf{x}})$ subject to the side condition $\bar{\mathbf{x}}^2 = 1$ leads to an ordinary eigenvalue problem in \mathbb{R}^3 .

Note that one might not know in advance whether an approximation of the given data with such a special case of a singular complex makes sense. Input data which

are actually contained in a complex of infinite pitch cause all coefficients in (4.7) to vanish (the vector $(0, \bar{\mathbf{c}})$ obviously is contained in the kernel of both M and D). Thus data which can be well approximated by such a complex can be detected by the magnitude of these coefficients.

Remark 4.1.3. If small coefficients in Equ. (4.7) cause numerical difficulties then of course one remedy is to approximate with a complex of infinite pitch. Another possibility is to use the normalization $\mathbf{x}^2 + \bar{\mathbf{x}}^2 = 1$ instead of $\mathbf{x}^2 = 1$. This is equivalent to letting $D = E_6$ and leads to an equation of degree six instead of the cubic equation (4.7). \diamond

Pencils of Minimizing Complexes

It is possible that Equ. (4.7) of Lemma 4.1.2 has two small solutions λ_1, λ_2 . This means that there are two solution complexes $\mathcal{C}_1, \mathcal{C}_2$ of Equ. (4.6) corresponding to λ_1 and λ_2 , respectively, which fit the input data equally well. In fact, all complexes of the pencil spanned by $\mathcal{C}_1, \mathcal{C}_2$ are close to the input lines:

Lemma 4.1.4. *If \mathcal{C} is a complex of the pencil spanned by $\mathcal{C}_1, \mathcal{C}_2$ with Plücker coordinates $\mathcal{C}_i\gamma^* = (\mathbf{c}_i, \bar{\mathbf{c}}_i)\mathbb{R}$ ($i = 1, 2$) and $\alpha = \sphericalangle(\mathbf{c}_1, \mathbf{c}_2)$, then for all lines L*

$$m(L, \mathcal{C}) \leq 1/|\sin \alpha| \cdot (m(L, \mathcal{C}_1) + m(L, \mathcal{C}_2)). \quad (4.11)$$

Proof. Assume that $L\gamma = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$ with $\|\mathbf{l}\| = 1$, that $\|\mathbf{c}_i\| = 1$ and that $\mathcal{C}\gamma = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$, $\|\mathbf{c}\| = 1$ with $(\mathbf{c}, \bar{\mathbf{c}}) = \mu_1(\mathbf{c}_1, \bar{\mathbf{c}}_1) + \mu_2(\mathbf{c}_2, \bar{\mathbf{c}}_2)$. Then $\mu_1, \mu_2 \leq 1/|\sin \alpha|$ and we have $m(L, \mathcal{C}) = |\mu_1(\mathbf{c}_1 \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}}_1 \cdot \mathbf{l}) + \mu_2(\mathbf{c}_2 \cdot \bar{\mathbf{l}} + \bar{\mathbf{c}}_2 \cdot \mathbf{l})| \leq |\mu_1| \cdot m(L, \mathcal{C}_1) + |\mu_2| \cdot m(L, \mathcal{C}_2)$, so the lemma is proved. \square

Lemma 4.1.4 shows that the input data in the case of two small solutions λ_1, λ_2 of (4.7) are close to all complexes of a pencil of complexes. In the generic case the carrier of this pencil (cf. Sec. 3.2) is a linear line congruence. We could also ask whether a line which is close to all complexes of a pencil G is also close to the lines of the carrier $C(G)$. The answer is affirmative, but we will not give precise estimates.

Remark 4.1.4. Assume that a pencil G of linear complexes is spanned by \mathcal{C}_1 and \mathcal{C}_2 . Consider the set of lines L with $m(L, \mathcal{C}) < \varepsilon$ for all complexes $\mathcal{C} \in G$. The Klein image of this set of lines is bounded by a certain non-Euclidean distance surface. \diamond

Bundles of Minimizing Complexes

It is possible that Equ. (4.7) has three small solutions $\lambda_1, \lambda_2, \lambda_3$. This means that there are three projectively independent solution complexes \mathcal{C}_i of Equ. (4.6), which fit the input data equally well, and all complexes of the *bundle* G spanned by $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ are close to the input data. This is shown by applying Lemma 4.1.4 twice: We have an inequality of the form

$$m(L, \mathcal{C}) \leq k \cdot \sum m(L, \mathcal{C}_i)$$

if \mathcal{C} is a complex of G , with a constant k depending on the angles between the vectors \mathbf{c}_i , where $\mathcal{C}_i \gamma^* = (\mathbf{c}_i, \bar{\mathbf{c}}_i) \mathbb{R}$.

Again it is possible to show that the lines close to *all* complexes of a bundle G are close to lines of the carrier $C(G)$, which is a regulus in the generic case. Approximation by reguli is discussed later (see p. 217). We first describe two degenerate cases.

Fitting Bundles to Lines

To fit lines L_i with a proper *line bundle*, i.e., one with a proper vertex, we choose two orthogonal planes $\sigma_i : \mathbf{m}_i \cdot \mathbf{x} - f_i = 0$ and $\tau_i : \mathbf{n}_i \cdot \mathbf{x} - g_i = 0$ which contain L_i . Without loss of generality we choose \mathbf{m}_i and \mathbf{n}_i as orthogonal unit vectors. The distance of a point \mathbf{x} to the planes σ_i and τ_i equals $\mathbf{m}_i \cdot \mathbf{x} - f_i$ and $\mathbf{m}_i \cdot \mathbf{c} - g_i$, respectively, so the distance of \mathbf{x} to the line L_i is computed by

$$d(\mathbf{x}, L_i)^2 = (\mathbf{m}_i \cdot \mathbf{x} - f_i)^2 + (\mathbf{n}_i \cdot \mathbf{x} - g_i)^2.$$

The vertex \mathbf{v} of the approximating bundle is therefore found as minimizer of

$$\sum_{i=1}^k d(\mathbf{x}, L_i)^2 = \sum_{i=1}^k [(\mathbf{m}_i \cdot \mathbf{x} - f_i)^2 + (\mathbf{n}_i \cdot \mathbf{x} - g_i)^2]. \quad (4.12)$$

A bundle of parallel lines is fitted to the lines L_1, \dots, L_k in the following way: We may assume that there is a vector \mathbf{w}_0 and vectors \mathbf{l}_i parallel to L_i such that $\mathbf{w}_0 \cdot \mathbf{l}_i > 0$ for all i . Then let

$$\mathbf{w} = \frac{1}{k} \sum \mathbf{l}_i.$$

The bundle of lines parallel to \mathbf{w} then is the bundle which minimizes the sum of squared Euclidean distances (in \mathbb{R}^3) of \mathbf{l}_i to \mathbf{w} .

Fitting Fields to Lines

This problem is dual to the previous one, so we could solve it by applying some duality to the input data, fitting a bundle to it, and applying the inverse duality to this bundle. But this has the disadvantage that the resulting field is no longer minimizing any distances defined by Euclidean geometry.

On the other hand imagine a plane ε and a line L nearly parallel to it. There is a line segment in L whose points are all close to ε , but the distances of L 's points to ε may become arbitrarily large. That is why it is better to solve a different problem: The fitting of a field to *line segments*, which is discussed below.

Fitting Complexes to Line Segments

In applications it often makes sense to define the deviation of a line L from some unbounded set M by intersecting this line with a certain domain of interest D and considering the deviation of $L \cap D$ from M . A simple special case is to choose a segment in L . We will show how to modify the algorithms for fitting linear complexes to lines such that they become algorithms for fitting linear complexes to line segments.

We need an appropriate definition of the distance of a line segment \overline{ab} to a linear complex C with equation $\bar{c} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}} = 0$ and corresponding null polarity π : By Equ. (3.3), the lines of C incident with a are contained in the plane $a\pi$ with normal vector $\mathbf{n}_a = \bar{c} + \mathbf{c} \times \mathbf{a}$. The distance of b to this plane is given by

$$d(\mathbf{b}, a\pi) = \frac{|\mathbf{n}_a \cdot (\mathbf{b} - \mathbf{a})|}{\|\mathbf{n}_a\|}.$$

Assume that the line $L = \mathbf{a} \vee \mathbf{b}$ has Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ with $\mathbf{l} = \mathbf{b} - \mathbf{a}$. Then by insertion of \mathbf{n}_a

$$d(\mathbf{b}, a\pi) = \frac{\bar{c} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}}}{\|\bar{c} + \mathbf{c} \times \mathbf{a}\|}. \quad (4.13)$$

Obviously interchanging \mathbf{a} and \mathbf{b} gives an analogous expression for the distance $d(\mathbf{a}, b\pi)$. This motivates the following

Definition. With the notation of the previous paragraph, the distance $d(\overline{ab}, C)$ between the line segment \overline{ab} and the complex C is defined by

$$d(\overline{ab}, C)^2 = d(\mathbf{b}, a\pi)^2 + d(\mathbf{a}, b\pi)^2 = (\bar{c} \cdot \mathbf{l} + \mathbf{c} \cdot \bar{\mathbf{l}})^2 \left(\frac{1}{v_a^2} + \frac{1}{v_b^2} \right),$$

where $\mathbf{l} = \mathbf{b} - \mathbf{a}$, $v_a = \|\bar{c} + \mathbf{c} \times \mathbf{a}\|$, $v_b = \|\bar{c} + \mathbf{c} \times \mathbf{b}\|$. (4.14)

If the line spanned by the points \mathbf{a} and \mathbf{b} is contained in C , then $d(\overline{ab}, C) = 0$.

Remark 4.1.5. The complex C defines a uniform helical motion (cf. Th. 3.1.6). Then v_a is the norm of the velocity vector \mathbf{n}_a of the point \mathbf{a} . If p is the pitch of C and r_a is the distance of \mathbf{a} from C 's axis, then

$$v_a = \sqrt{r_a^2 + p^2}. \quad \diamond$$

The linear complex \mathcal{X} which fits the line segments $\overline{a_1b_1}, \dots, \overline{a_kb_k}$ best is defined to be the minimizer of

$$\sum_{i=1}^k d(\overline{a_i b_i}, \mathcal{X})^2 = \sum_{i=1}^k \left(\frac{1}{v_{a_i}^2} + \frac{1}{v_{b_i}^2} \right) (\bar{x} \cdot \mathbf{l}_i + \mathbf{x} \cdot \bar{\mathbf{l}}_i)^2, \quad (4.15)$$

where $\mathbf{l}_i = \mathbf{b}_i - \mathbf{a}_i$, $\bar{\mathbf{l}}_i = \mathbf{a}_i \times \mathbf{l}_i = \mathbf{b}_i \times \mathbf{l}_i$, and v_{a_i}, v_{b_i} are the velocities of $\mathbf{a}_i, \mathbf{b}_i$ as defined by Equ. (4.14). The solution may be computed using *weight iteration*: We minimize the weighted sum

$$F(\mathbf{x}, \bar{\mathbf{x}}, w_1, \dots, w_k) = \sum_{i=1}^k w_i (\bar{\mathbf{x}} \cdot \mathbf{l}_i + \mathbf{x} \cdot \bar{\mathbf{l}}_i)^2, \quad (4.16)$$

subject to an appropriate side condition which expresses a normalization of $(\mathbf{x}, \bar{\mathbf{x}})$: Typically this would be $\|\mathbf{x}\| = 1$, but if a solution complex of very large pitch is expected it is better to use $\|\bar{\mathbf{x}}\| = 1$. The solution is analogous to the one described by Lemma 4.1.2. In the beginning we let $w_1 = \dots = w_k = 1$. After the first step we use

$$w_i = 1/v_{\mathbf{a}_i}^2 + 1/v_{\mathbf{b}_i}^2, \quad i = 1, \dots, k,$$

where $v_{\mathbf{a}_i}, v_{\mathbf{b}_i}$ are defined by Equ. (4.14), and compute a new minimizer of (4.16). This is iterated until the change of weights from step to step is less than some threshold value.

To complete the discussion of singular cases in the previous section, we mention how to fit a field of lines to given line segments $\mathbf{a}_i \mathbf{b}_i$. The simplest solution of this problem is to find a plane approximating the points $\mathbf{a}_i, \mathbf{b}_i$ in the usual least-squares sense.

This procedure can be used to find a field of lines which fits given lines L_1, \dots, L_k within a bounded region of interest $D \subset E^3$. We simply use the smallest line segments $\overline{\mathbf{a}_i \mathbf{b}_i}$ which contain the intersection $L_i \cap D$.

4.2 Kinematic Surfaces

As an application of the theoretical problem of fitting linear complexes to given lines or line segments, we look at a problem which arises in the context of *reverse engineering* of geometric models [196]: Given are scattered data points which are expected to fit to a simple surface, where the meaning of ‘simple’ is explained later in this section. We want to reconstruct the surface or the geometric data which uniquely determine it.

A particular instance of this problem is the automated reconstruction of parts from laser scanner data and the decomposition of their boundary surface in its planar, cylindrical, etc., pieces [118, 155, 158, 195].

This section is organized as follows: First we determine surfaces which are invariant with respect to uniform motions and show that their surface normals are contained in certain linear manifolds of lines. The next topic is the estimation of these linear manifolds from scattered surface normals, and how to compute the surfaces from them. Last we discuss the reconstruction of surfaces which are ‘piecewise almost’ invariant.

Consider a one-parameter subgroup $M(t)$ of Euclidean motions (a uniform helical, rotational or translational motion, cf. Th. 3.4.3) and a curve $c : I \rightarrow E^3$ in Euclidean three-space. The symbol

$$g(u, v) = M(u)(c(v)), \quad u \in \mathbb{R}, v \in I. \quad (4.17)$$

means that we apply the Euclidean motion $M(u)$ to the curve point $c(v)$. Then g parametrizes a surface in Euclidean space. It is a *cylindrical surface* or *cylinder* if $M(t)$ is a group of translations, a *surface of revolution*, if $M(t)$ is a group of rotations about an axis, and a *helical surface* otherwise. We refer to these types of surface as *kinematic surface*.

Remark 4.2.1. A subset of Euclidean space may be a kinematic surface in several ways: The cylinder of revolution is an example of a surface which is at the same time a cylindrical surface, a surface of revolution, and a helical surface (see Fig. 4.1, right). \diamond

Invariant Surfaces

Clearly the surfaces described by Equ. (4.17) are *invariant* with respect to the one-parameter group of motions which generates them: If p is a point of the surface, then so is $M(t)(p)$ for all $t \in \mathbb{R}$ (see Fig. 4.1). We ask for all surfaces which are invariant under the one-parameter group $M(t)$:

Theorem 4.2.1. *A subset Φ of Euclidean space E^3 which is both a two-dimensional C^1 submanifold and invariant with respect to a one-parameter group of motions $M(t)$ is a kinematic surface as described by Equ. (4.17).*

Proof. (Sketch) It is sufficient to prove this locally. Choose a point $p \in \Phi$ whose velocity vector $v_p = d/dt|_{t=0} M(t)(p)$ is nonzero and consider a curve $c : I \rightarrow E^3$ with $c(0) = p$ and with $\{c'(0), v_p\}$ linearly independent, such that the surface $g(u, v) = M(u)(c(v))$ is locally regular. The point $g(u, v)$ is contained in Φ by invariance of Φ . Thus g locally is a diffeomorphism of an open neighbourhood of $(0, 0)$ in the u, v parameter domain to an open subset of Φ , and Φ is locally parameterized as a kinematic surface.

If $v_p = 0$, there is nothing to show, because $M(t)(p) = p$ for all t . \square

There is a simple characterization of invariant surfaces, respectively, kinematic surfaces, in terms of their normals:

Lemma 4.2.2. *The surface normals of a regular C^1 surface $g : U \subset \mathbb{R}^2 \rightarrow E^3$ are contained in a linear complex if and only if the surface is contained in a kinematic surface as defined by Equ. (4.17).*

Proof. The surface normals of a kinematic surface are path normals of the curves $M(t)(p)$, and so the ‘if’ part of the theorem follows from Th. 3.1.6.

To prove the converse, assume that all of g ’s surface normals are contained in a linear complex \mathcal{C} , which, by Th. 3.1.6, is the path normal complex of a one-parameter group $M(t)$ of motions. We consider a point $p = g(u_0, v_0)$ (cf. Fig. 4.1, left). The image $g(V)$ of a small neighbourhood V of (u_0, v_0) is a two-dimensional C^1 submanifold of E^3 . The integral curves of M ’s velocity vector field $v : p \mapsto v(p)$ are tangent to $g(U)$ by our assumption, which shows that especially the integral

curve $M(t)(p)$ starting in p is contained in $g(V)$. Fig. 4.1, left, shows a sequence of such integral curves, which are helices.

Now we can choose a curve $c : I \rightarrow g(V)$ such that $c(0) = p$ and $\{\dot{c}(0), v(p)\}$ is linearly independent. The kinematic surface $h(t, s) = M(t)(c(s))$ locally parametrizes $g(V)$, and the proof is complete. \square

Fig. 4.1, left shows a helical surface generated as envelope of a sphere which undergoes a uniform helical motion. This surface is both an invariant surface and a pipe surface, which means envelope of a smooth family of spheres.

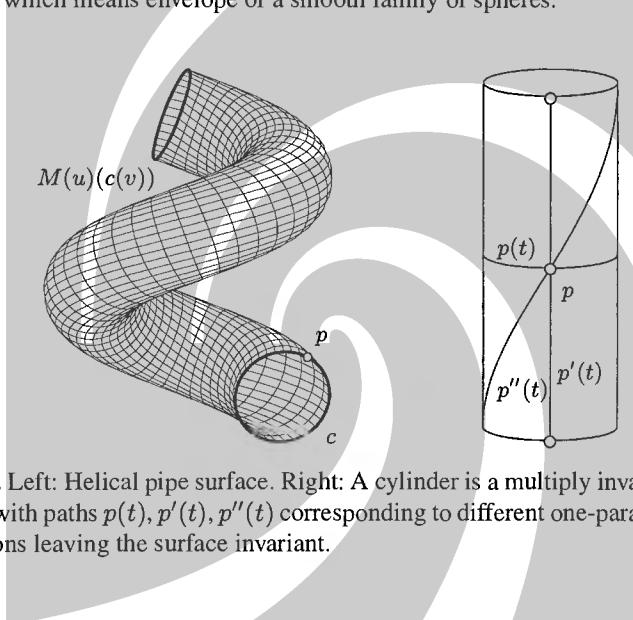


Fig. 4.1. Left: Helical pipe surface. Right: A cylinder is a multiply invariant surface: point p with paths $p(t), p'(t), p''(t)$ corresponding to different one-parameter groups of motions leaving the surface invariant.

Multiply Invariant Surfaces

It is important to know which surfaces are invariant with respect to more than just one one-parameter group. Before answering that question, we prove the following

Lemma 4.2.3. *If C_1, C_2 are two different linear complexes, and $M_1(t), M_2(t)$ are corresponding one-parameter groups of motions, and a linear line congruence is invariant with respect to both $M_1(t), M_2(t)$, then this congruence is a bundle of lines, or a field of lines, or consists of all lines which intersect a proper line orthogonally.*

Proof. Clearly the possibilities mentioned have this property. The converse is also shown easily: A hyperbolic congruence is invariant only if both its axes are invariant, but non-intersecting and non-parallel lines invariant under rotations and helical motions are only the axis and the ideal line orthogonal to it.

A parabolic congruence is not invariant with respect to any Euclidean motion, which follows from the detailed description in Sec. 3.2.2.

An elliptic linear congruence can have rotational symmetry, like the one described by Equ. (3.15). It is left to the reader as an exercise to show that no elliptic congruence can be multiply invariant (this follows from Prop. 3.2.8). \square

The following lemma enumerates all surfaces which are multiply invariant. The result is intuitively clear anyway (cf. Fig. 4.1, right).

Lemma 4.2.4. *A connected regular C^1 surface $g : U \rightarrow E^3$, all of whose surface normals are contained in two different linear complexes C_1, C_2 , is contained in a plane, or sphere, or cylinder of revolution.*

Proof. Denote the pencil of linear complexes spanned by C_1 and C_2 by G . If N is a surface normal, then by Lemma 3.2.2, $N \in C_1, N \in C_2$ is equivalent to $N \in C(G)$.

If $C \in G$ and $M(t)$ is the uniform motion whose path normal complex equals C , then Lemma 4.2.2 shows that the surface $g(U)$ locally is a kinematic surface with respect to $M(t)$. As $C(G)\gamma$ is a two-dimensional quadric, it is determined by the surface normals of arbitrarily small open subsets of $g(U)$. This implies that $C(G)$ is invariant by such an $M(t)$, and so $C(G)$ must be as described by Lemma 4.2.3.

We now choose special uniform motions $M(t)$, namely: (i) if $C(G)$ is degenerate with a proper vertex O , consider all rotations about O ; if (ii) $C(G)$ is degenerate with an ideal vertex A_u , choose all translations orthogonal to A_u ; if (iii) $C(G)$ is hyperbolic with proper axis A , choose a rotation about A and a translation parallel to A . The last remaining case, a field of lines, cannot occur. The statement we want to prove now follows locally from Lemma 4.2.2 and globally from connectedness of $g(U)$. Cases (i), (ii), and (iii) correspond to the sphere, the plane, and the cylinder of revolution. \square

Corollary 4.2.5. *The only connected surfaces which are invariant with respect to two independent one-parameter groups of Euclidean motions are spheres, planes, and cylinders of revolution.*

Fitting Complexes to Scattered Surface Normals

To approximate scattered data points p_1, \dots, p_k by a kinematic surface we proceed as follows: First we estimate the surface normals N_1, \dots, N_k at the data points. There exist solutions of this problem which will not be discussed here (see e.g. [78]). We assume that the input data are evenly distributed. If not, we have to apply data reduction algorithms first (see [122]). Lemma 4.2.2 shows that in order to find the one-parameter subgroup which generates the approximating kinematic surface, we have to fit a linear complex C to the surface normals. This is done by using the methods described in this section. With Equ. (3.9) we compute axis A and pitch p of the generating motion. Of course there are several different cases according to dimensionality of the solution and magnitude of the pitch.

1. If p is small compared to the diameter of the input point cloud we may want to fit a complex of zero pitch to the input normals. This corresponds to input data with rotational symmetry.

2. If p is very large, we could fit a linear complex of infinite pitch. This leads to a cylindrical surface which approximates the input data.
3. If a pencil of complexes fits the input normals, Cor. 4.2.5 shows that they are close to the surface normals of spheres, planes, or cylinders.

Remark 4.2.2. Small portions of input data will often lead to a pencil of nearly equally good solutions and it depends on further information what one can do in this case: If the input data are not expected to be multiply invariant, we may be able to gather more sample points and run the approximation algorithm again. If a kinematic surface fits the input data well in small regions, we have to paste together pieces of kinematic surfaces. This is discussed below. \diamond

Remark 4.2.3. In reverse engineering applications the input data can have large measurement errors or possibly include data points which belong to another part of the object which does not fit the same kinematic surface.

Thus an approximation method must be able to cope with outliers, which is not the case for the least squares method in the form presented above. Therefore, one may use a robust regression method to compute an initial estimate (e.g., an estimate of a least median of squares solution) and then refine it by either rejecting outliers or by down-weighting their influence on the final approximant.

This kind of noise filtering based on so-called *M-estimators* has been investigated in detail both in statistics and in computer vision. Note that the formulae presented here are nicely compatible with M-estimation, since we just have to reformulate the various functions F to be minimized as a weighted sum of squares of moments, using one of the weighting schemes suggested in the literature [171]. \diamond

Remark 4.2.4. The minimization of the function F of Equ. (4.3) is motivated by the following approach: Consider the one-parameter group $M(t)$ of helical motions associated with a linear complex C with $C\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$, $\|\mathbf{c}\| = 1$. The velocity $\mathbf{v}(\mathbf{x})$ of the point \mathbf{x} is given by Equ. (3.25). Assume that we have found estimates N_i of surface normals at the data points \mathbf{x}_i , which have normalized Plücker coordinates $(\mathbf{n}_i, \bar{\mathbf{n}}_i)$. Then

$$\cos \gamma_i = \cos \sphericalangle(\mathbf{v}(\mathbf{x}_i), \mathbf{n}_i) = \frac{\mathbf{n}_i \cdot \mathbf{v}(\mathbf{x}_i)}{\|\mathbf{v}(\mathbf{x}_i)\|} = \frac{\bar{\mathbf{c}} \cdot \mathbf{n}_i + \mathbf{c} \cdot \bar{\mathbf{n}}_i}{\|\mathbf{v}(\mathbf{x}_i)\|}.$$

Minimizing the function

$$G(\mathbf{c}, \bar{\mathbf{c}}) = \sum_{i=1}^k \cos^2 \gamma_i$$

with the side condition $\|\mathbf{c}\| = 1$ is a nonlinear problem. We did minimize the function F of Equ. (4.3). Equ. (4.2) shows that

$$F(\mathbf{c}, \bar{\mathbf{c}}) = \sum_{i=1}^k (r_i^2 + p^2) \cos^2 \gamma_i,$$

where r_i is the distance of data point \mathbf{d}_i to the axis of \mathcal{C} . This is a reasonable and geometrically meaningful simplification of $G \rightarrow \min$. In fact by using a weight iteration as in Equ. (4.16) it is even possible to minimize G . \diamond

Fitting Kinematic Surfaces to Scattered Data Points

After we have found a one-parameter subgroup $M(t)$ of Euclidean motions which fits the estimates of surface normals, we try to find the approximating surface itself. This is done by ‘projecting’ the input data points \mathbf{d}_i , $i = 1, \dots, k$ into an appropriate plane ε , where the image of the point \mathbf{d}_i is found by following its trajectory $M(t)(\mathbf{d}_i)$ until it intersects ε . As the trajectories of points may intersect planes more often than once (indeed, even infinitely many times), this definition has to be more precise in order to make the projection well defined:

1. If $M(t)$ is subgroup of translations, the only restriction is that ε must not be parallel to the trajectories of points. It makes sense to choose the plane orthogonal to the vector of the translation.
2. If $M(t)$ is a subgroup of rotations about a fixed axis A , we choose ε such that it contains A . Then all trajectories except those of A ’s points will intersect ε in two points (see Fig. 4.2). The mapping becomes well defined if we choose one of the two closed half-planes defined by A in ε and intersect all trajectories with this half-plane.
3. If $M(t)$ is a uniform helical motion of pitch p and axis A , choose ε orthogonal to A . Then all trajectories intersect ε once. It turns out that for actual computations this choice is not always the best, especially if p is small. A choice which works well in practice is to choose ε as the path normal plane of one of the data points. All trajectories intersect this plane in a finite number of points. If the points are close together the intersection points cluster in a finite number of well separated subsets.

After performing the intersection, the points \mathbf{p}_i should lie close to a certain curve, if the original points lie close to a kinematic surface. This curve can be fitted to the points \mathbf{p}_i (cf. [106, 158] and Fig. 4.2). The surface generated by this curve under the action of $M(t)$ is the approximant we have been looking for.

Remark 4.2.5. ‘Projecting’ data points into a reference plane should keep artificial distortions to a minimum. In cases 1 and 2 of the above list, the distance of a data point to the eventual solution surface is the same as the distance of the projected point to the curve which is fitted to these projected points.

In case 3, however, the first method, which always works, does not have this property — the smaller p and the farther the data points are from the axis, the more distances increase. The second method of projection is less trivial in its implementation, but avoids these distortions to a certain extent. \diamond

Fitting Special Surfaces

In computer-aided design, simple surfaces whose normals form a well known subset of line space occur very often. These include planes, spheres, cylinders or cones of revolution, and tori, some of which are multiply invariant. It is useful to specialize the general approximation algorithms for these surfaces:

1. A plane is easily fitted to scattered data points, so if the input data are known to be contained in a plane, the whole machinery described above is actually not necessary.
2. A line is easily fitted to scattered data points in a plane. This situation occurs when the original data points belong to a part of a half-cone of revolution, where half-cone means one of the two halves of a cone which are separated by the vertex. After reconstruction of the axis A and projecting the input data into a half-plane as described above, a line has to be fitted to the projections of points.
3. A sphere with center \mathbf{x} is fitted to scattered points \mathbf{p}_i by minimizing

$$F(\mathbf{x}) = \sum_{i,j} \frac{((2\mathbf{x} - \mathbf{p}_i - \mathbf{p}_j) \cdot (\mathbf{p}_i - \mathbf{p}_j))^2}{\|\mathbf{p}_i - \mathbf{p}_j\|^2}, \quad (4.18)$$

where summation is over all index pairs i, j such that $\|\mathbf{p}_i - \mathbf{p}_j\|$ is not too small compared with the extension of the point cloud \mathbf{p}_i .

The motivation for this is the following: The single terms in this sum are the squared distances of the point \mathbf{x} to the bisector plane of \mathbf{p}_i and \mathbf{p}_j , which is numerically ill-defined if $\|\mathbf{p}_i - \mathbf{p}_j\|$ is small.

4. A circle is fitted to points in a plane by minimizing the same expression, with the only difference that the variables (4.18) are vectors of \mathbb{R}^2 . This occurs if the input data belong to a torus: After projection of the data points to a half-plane which contains the axis we have to fit a circle to these points.

For solutions of the problem of fitting special surfaces based on their representation as algebraic varieties, we refer to the literature [118, 163].

Example 4.2.1. This example concerns scattered data (e.g. obtained by a laser scanner) from an object whose boundary is a surface of revolution. The surface normals at the data points are estimated (see Fig. 4.2, left). The pitch in this case is nearly zero, which shows that the original data come from a surface of revolution. We let $p = 0$ and project the input data into a half-plane which contains the axis (Fig. 4.2, center). The curve which fits these points was found by a moving least squares method according to [106]. \diamond

Example 4.2.2. We again consider reconstruction of surfaces from scattered data. The difference to Ex. 4.2.1 is that the data do not come from a surface of revolution, but from a *pipe surface*, which is defined as the envelope of spheres of equal radius whose center runs in the *spine curve*.

A pipe surface is locally well approximated by a surface of revolution: If we replace the spine curve by its osculating circle, we get a torus which is in second

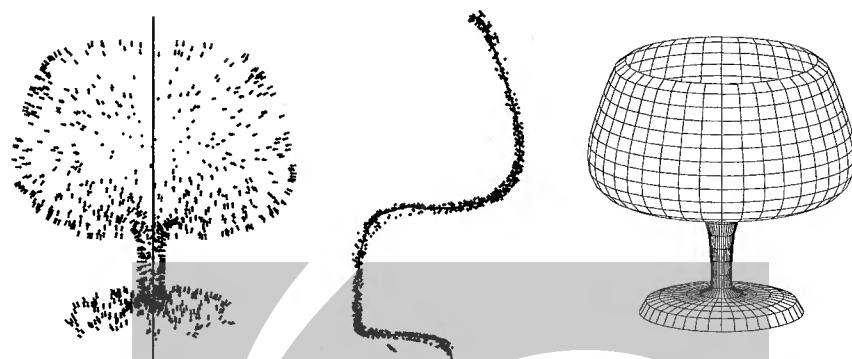


Fig. 4.2. Reconstruction of a surface of revolution: Left: data points, estimates of normal vectors, and axis computed from this estimation. Center: points projected onto a plane and a curve approximating this point set, Right: final surface of revolution.

order contact with the pipe surface in all points of a common circle. If we consider only small parts of the given point cloud, it is easy to determine such approximating tori, whose spine curves then locally approximate the spine curve of the pipe surface. The result of such an approximation is shown in Fig. 4.3. An application of this is the recovery of constant radius rolling ball blends in reverse engineering [99]. Using locally approximating surfaces of revolution, we can also reconstruct so-called moulding surfaces [108]. These are generated by a moving planar curve, whose carrier plane rolls on a cylinder surface. With local fits by right circular cones or cylinders, the reconstruction of developable surfaces may be performed [25]. ◇

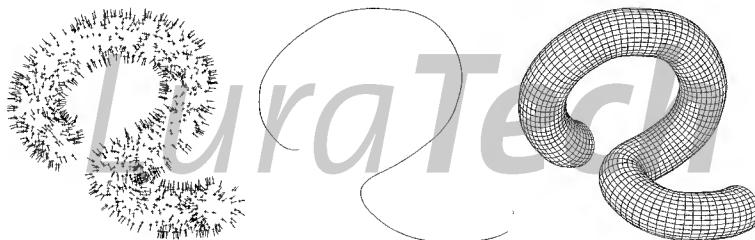


Fig. 4.3. Pipe surface: Left: data points and estimates of normal vectors, Center: approximate spine curve, Right: reconstruction of pipe surface.

www.luratech.com

Example 4.2.3. Fig. C.3 illustrates reverse engineering of an actual object.

This object is one of the so-called Darmstadt benchmarks suggested by J. Hoschek; the pictures were generated by the BSolid prototype Reverse Engineering

system, developed by the Computer and Automation Research Institute and Cadmus Consulting and Development Ltd., Budapest. \diamond

Approximation by Kinematic Surfaces Using Additional Information

Reverse engineering of CAD models might also make use of further geometric information such as parallelity, concentricity, or orthogonality, if available.

1. If the direction \mathbf{x} of the axis A of a surface of revolution is known in advance, minimization of the function $F(\mathbf{x}, \bar{\mathbf{x}})$ of Equ. (4.3) is much easier, because the constraint $\|\mathbf{x}\| = 1$ can be fulfilled automatically, and $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$ is now a linear side-condition.
2. Similarly, the problem simplifies if we reconstruct a surface of revolution and know a point \mathbf{b} of the axis A . If A 's normalized Plücker coordinates are $(\mathbf{x}, \bar{\mathbf{x}})$, then $\bar{\mathbf{x}} = \mathbf{b} \times \mathbf{x}$ and $F(\mathbf{x}, \bar{\mathbf{x}})$ of Equ. (4.3) is a function of \mathbf{x} alone. The constraint $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$ is fulfilled automatically.
Thus minimizing F is an ordinary eigenvalue problem.
3. An analogous situation occurs when fitting a surface of revolution, whose axis is constrained in a plane. Suppose that this plane is defined by a point \mathbf{b} and two independent basis vectors $\mathbf{v}_1, \mathbf{v}_2$. We let $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2$ and have the linear relations

$$\mathbf{x} = \rho_1 \mathbf{v}_1 + \rho_2 \mathbf{v}_2, \quad \bar{\mathbf{x}} = (\mathbf{b} + \tau_1 \mathbf{v}_1 + \tau_2 \mathbf{v}_2) \times \mathbf{x} = \mathbf{b} \times \mathbf{x} + \tau \mathbf{n}.$$

Thus the function F may be expressed in terms of ρ_1, ρ_2, τ .

4. Finally, if the axis has to intersect a given line L with Plücker coordinates $\mathbf{l}, \bar{\mathbf{l}}$, we have to make use of the linear intersection condition

$$\bar{\mathbf{x}} \cdot \mathbf{l} + \mathbf{x} \cdot \bar{\mathbf{l}} = 0.$$

The minimization of F subject to this condition together with the normalization $\|\mathbf{x}\| = 1$ is similar to minimization of F subject to the condition that $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$.

Remark 4.2.6. The first three cases in the list above lead to the minimization of a quadratic function subject to a normalization condition. This common property can also be seen in the Klein image: In line-geometric terms, we are looking for an approximant complex whose Klein image must lie in a certain projective subspace U of P^5 :

For helical surface reconstruction without additional constraints (cf. Lemma 4.1.2), we can define $U = P^3$, which is actually no restriction. For reconstruction of surfaces of revolution in cases 1–3 of the above list, U is a plane contained in the Klein quadric:

In case 1, this plane is the Klein image of all lines parallel to a given vector. In case 2, this plane is the Klein image of all lines incident with the point \mathbf{b} . In case 3, this plane is the Klein image of a field. The fact that U is a subset of the Klein quadric explains why the quadratic side condition $\mathbf{x} \cdot \bar{\mathbf{x}} = 0$ is fulfilled automatically in cases 1–3.

Case 4 of the list above is different, as is the problem of finding a minimizing singular complex (cf. Lemma 4.1.3): Here the solution complexes are restricted to a quadratic variety V in P^5 . In the case of singular complexes, V is the Klein quadric. In case 4 of the above list V is a tangential intersection of the Klein quadric. \diamond

4.3 Approximation via Local Mappings into Euclidean 4-Space

In this section we discuss approximation methods in line space which are based on local mappings into Euclidean 4-space. ‘Local’ means that these mappings are defined in open subsets of line space. It does not necessarily mean that these subsets are small.

Stereographic Projection

It is sometimes useful to identify a part of a quadric with an affine space. One familiar example is a map projection which identifies part of the globe with (part of) a sheet of paper.

Definition. Assume that Φ is a quadric in P^n , the point Z is in Φ , and Q is a hyperplane with $Z \notin Q$. Then the projection

$$\sigma : \Phi \rightarrow Q, \quad X \mapsto X\sigma = (X \vee Z) \cap Q \quad (4.19)$$

is called the stereographic projection of Φ to Q with center Z .

An example of a stereographic projection has been given in Ex. 1.1.37.

Lemma 4.3.1. We use the notation of the definition above. If T is Φ ’s tangent plane at the center Z , then the stereographic projection is a one-to-one correspondence between $\Phi \setminus T$ and $Q \setminus T$. Hyperplanar sections Φ' of Φ are mapped onto $(n - 2)$ -dimensional projective subspaces of Q if and only if Φ' contains the projection center.

Proof. We know that all lines not tangent to Φ intersect Φ in exactly two points or not at all. This shows that especially all non-tangential lines $Z \vee X$ with $X \in Q$ intersect the quadric Φ in exactly one further point besides Z . Thus $\sigma : \Phi \setminus T \rightarrow Q \setminus T$ is one-to-one and onto.

If $\Phi' = \Phi \cap H$, where H is a hyperplane, then Φ' contains Z if and only if H contains Z , so $Z \vee H$ is $(n - 1)$ -dimensional (otherwise, if Φ' does not contain Z , it has dimension n), and the statement follows. \square

The meaning of Lemma 4.3.1 is the following: If we disregard all points of Φ which are also contained in the center’s tangent plane, the stereographic projection gives a one-to-one correspondence between the quadric and the *affine space* $Q \setminus T$.

Example 4.3.1. (cf. Ex. 1.1.37) We consider the projective extension of Euclidean space E^3 . If Φ is the unit sphere, Z is its north pole, and Q is the equator plane, then besides Z there are no points of Φ which are contained in the north poles' tangent plane T . The line $T \cap Q$ is at infinity. Lemma 4.3.1 says that the stereographic projection is a one-to-one correspondence between the points of the unit sphere different from the north pole, and the points of the equator plane. \diamond

Stereographic Projection of the Klein Quadric

The set \mathcal{L} of the lines of three-dimensional projective space can be identified, via the Klein mapping, with the Klein quadric $M_2^4 \subset P^5$. The set \mathcal{L}^o of proper lines of Euclidean space can be identified with the Klein quadric without a plane, which is the Klein image of the ideal field of lines (cf. Sec. 2.1.3). Unfortunately \mathcal{L}^o (or its Klein image) does not have the structure of an affine space. Therefore we try to find an appropriate stereographic projection which identifies a certain subset of \mathcal{L} with an affine space.

If L_Z is a line, then $Z := L_Z\gamma \in M_2^4$. The tangent plane of M_2^4 at Z contains the γ -images of all lines which intersect L_Z . We introduce a Cartesian coordinate system in E^3 and let Z equal the horizontal line at infinity, which is contained in all planes $z = \text{const}$. Then $L_Z\gamma = Z = (0, 0, 0, 0, 1)$.

A line L of Euclidean space, which does not intersect L_Z , i.e., is not horizontal, intersects both planes $\pi_- : z = 0$ and $\pi_+ : z = 1$ in points $\mathbf{x}_- = (x_1, x_2, 0)$ and $\mathbf{x}_+ = (x_3, x_4, 1)$. The Plücker coordinates of L are computed by

$$\begin{aligned} L\gamma &= (1, x_1, x_2, 0) \wedge (1, x_3, x_4, 1)\mathbb{R} \\ &= (x_3 - x_1, x_4 - x_2, 1, x_2, -x_1, x_1 x_4 - x_2 x_3)\mathbb{R}. \end{aligned} \quad (4.20)$$

P^5 is equipped with homogeneous coordinates $(x_{01} : \dots : x_{12})$. The hyperplane $Q : x_{12} = 0$ does not contain the center Z , and projection of $L\gamma$ onto Q from the center Z gives the point

$$(Z \vee L\gamma) \cap Q = (x_3 - x_1, x_4 - x_2, 1, x_2, -x_1, 0)\mathbb{R}.$$

Thus x_1, x_2, x_3, x_4 are affine coordinates in the hyperplane Q , and the mapping

$$\sigma : L \in \mathcal{L}^o \mapsto (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \quad (4.21)$$

is, apart from a linear coordinate transformation, nothing but the stereographic projection.

Distances between Lines

We have to define a distance function between lines. If we restrict ourselves to non-horizontal lines L , we can use the stereographic projection σ defined by Equ. (4.21): If q is a positive definite quadratic form in \mathbb{R}^4 , then $\sqrt{q(G\sigma - H\sigma)}$ serves as a *distance* between lines G, H .

We give a more detailed construction of such a distance function: We use the intersection points of lines with the planes $\pi_- : z = 0$ and $\pi_+ : z = 1$. If G, H are two non-horizontal lines, denote the four intersection points by $\mathbf{g}_- = (g_1, g_2, 0)$, $\mathbf{g}_+ = (g_3, g_4, 1)$, $\mathbf{h}_- = (h_1, h_2, 0)$, $\mathbf{h}_+ = (h_3, h_4, 1)$. With the notation of Equ. (4.21), $G\sigma = (g_1, g_2, g_3, g_4)$, $H\sigma = (h_1, h_2, h_3, h_4)$.

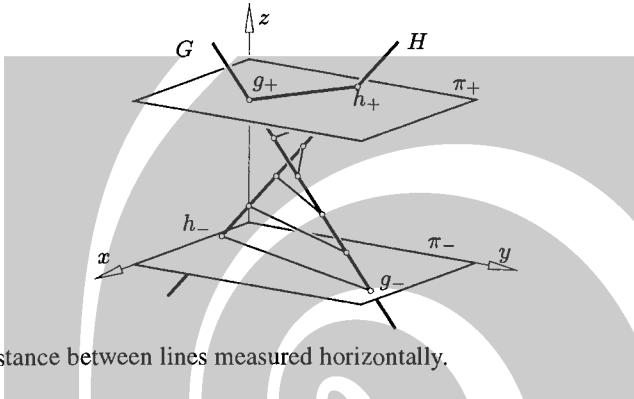


Fig. 4.4. Distance between lines measured horizontally.

We consider the correspondence

$$(1 - \lambda)\mathbf{g}_- + \lambda\mathbf{g}_+ \mapsto (1 - \lambda)\mathbf{h}_- + \lambda\mathbf{h}_+, \quad \lambda \in [0, 1],$$

(cf. Fig. 4.4) between the line segments $[\mathbf{g}_-, \mathbf{g}_+]$ and $[\mathbf{h}_-, \mathbf{h}_+]$. The lines which connect associated points are all horizontal and are contained in a hyperbolic paraboloid (by Prop. 1.1.40), or in a plane (if G, H are coplanar). An average distance of associated points is then given by

$$\begin{aligned} d(G, H)^2 &= 3 \int_0^1 [(1 - \lambda)(\mathbf{g}_- - \mathbf{h}_-) + \lambda(\mathbf{g}_+ - \mathbf{h}_+)]^2 d\lambda \\ &= (\mathbf{g}_- - \mathbf{h}_-)^2 + (\mathbf{g}_+ - \mathbf{h}_+)^2 + (\mathbf{g}_- - \mathbf{h}_-) \cdot (\mathbf{g}_+ - \mathbf{h}_+). \end{aligned} \quad (4.22)$$

Because of

$$d(G, H)^2 = \sum_{i=1}^4 (g_i - h_i)^2 + (g_1 - h_1)(g_3 - h_3) + (g_2 - h_2)(g_4 - h_4), \quad (4.23)$$

this is the distance of points $G\sigma, H\sigma$ defined by the quadratic form

$$q(x_1, x_2, x_3, x_4) = x_1^2 + \cdots + x_4^2 + x_1x_3 + x_2x_4, \quad (4.24)$$

which is positive definite, as is clearly seen from its definition. If $\mathbf{x} = (x_1, \dots, x_4)$, then

$$q(\mathbf{x}) = \mathbf{x}^T \cdot G \cdot \mathbf{x}, \quad \text{with} \quad G = \begin{bmatrix} 1 & & 1/2 & 1/2 \\ & 1 & & \\ 1/2 & & 1 & \\ & 1/2 & & 1 \end{bmatrix} \quad (4.25)$$

We define the q -norm and the q -scalar product by

$$\|\mathbf{x}\|_q = q(\mathbf{x})^{1/2}, \quad \langle \mathbf{x}, \mathbf{y} \rangle_q = \mathbf{x}^T \cdot G \cdot \mathbf{y}. \quad (4.26)$$

Remark 4.3.1. Consider a point $\mathbf{p} = (x_p, y_p, z_p)$ of the line segment $[\mathbf{g}_-, \mathbf{g}_+]$. In Equ. (4.22) we integrated its distance to the line H , which is measured within the horizontal plane $z = z_p$. The shortest distance of \mathbf{p} to H differs from that distance by a factor λ with $\cos \phi \leq \lambda \leq 1$, ϕ being the angle enclosed by H and the z -axis.

If lines have the property that the angle enclosed with the z -axis does not exceed a certain value, then the distance function defined here is bounded by the Euclidean distance between lines times a certain factor.

This is in accordance with the well known property of stereographic map projections that $d(\sigma(P), \sigma(Q))/d(P, Q)$ increases if P and Q converge towards the center of the projection. \diamond

Stereographic Projection of Complexes

We apply the stereographic projection σ of Equ. (4.21) to sets of lines, such as reguli \mathcal{R} , linear congruences \mathcal{N} , or linear complexes \mathcal{C} . Their subsets of non-horizontal lines are denoted by the symbols \mathcal{R}° , \mathcal{N}° , \mathcal{C}° , respectively. But we will still speak of a regulus, a linear congruence, and a linear complex, even if the horizontal lines are missing.

Lemma 4.3.2. Consider a linear complex \mathcal{C} with Plücker coordinates $\mathcal{C}\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} = (c_{01} : \dots : c_{12})$. If \mathcal{C} is regular and $c_{03} \neq 0$, then $\mathcal{C}^\circ\sigma$ is a quadric. If $c_{03} = 0$, then $\mathcal{C}^\circ\sigma$ is a hyperplane. If $c_{03} \neq 0$ and \mathcal{C} is a singular complex, then $\mathcal{C}^\circ\sigma$ is a quadratic cone. All image quadrics are homothetic, and so are all image cones.

Proof. Equ. (4.20) shows that σ maps \mathcal{C}° to the algebraic variety with equation

$$\begin{aligned} \mathcal{C}^\circ\sigma : \quad & c_{03}(x_1x_4 - x_2x_3) - (c_{02} + c_{23})x_1 + (c_{01} - c_{31})x_2 \\ & + c_{23}x_3 + c_{31}x_4 + c_{12} = 0. \end{aligned} \quad (4.27)$$

This is a hyperplane if and only if $c_{03} = 0$, and a quadratic variety otherwise. Up to the scalar factor c_{03} , all possible image quadrics share the quadratic term $x_1x_4 - x_2x_3$. Two quadrics with the same quadratic terms in their equations are *homothetic*, i.e., differ from each other only by a central similarity and a translation, and so do singular quadratic varieties.

A linear complex is singular if and only if its Klein image is a quadratic cone (and therefore contains pencils of concurrent lines). This property is preserved by a central projection. This shows that regular complexes project to regular quadrics, and singular complexes to singular quadratic varieties. \square

Remark 4.3.2. We projectively extend Euclidean four-space and consider the projective quadric Φ which contains $\mathcal{C}^\circ\sigma$. Its affine part is given by Equ. (4.27), and its equation in homogeneous coordinates $(x_0 : \dots : x_4)$ is $c_{03}(x_1x_4 - x_2x_3) +$

$(-(c_{02} + c_{23})x_1 + (c_{01} - c_{31})x_2 + c_{23}x_3 + c_{31}x_4 + c_{12}x_0)x_0 = 0$. Its intersection with the ideal hyperplane $x_0 = 0$ is the quadric

$$\Psi : x_0 = x_1x_4 - x_2x_3 = 0,$$

which is independent of \mathcal{C} . If we look at the stereographic projection, we see why this must be so: The horizontal lines of $\mathcal{C} \setminus \mathcal{C}^o$ are those whose Klein image is contained in the tangent hyperplane of the Klein quadric at the projection center Z . They project to the quadric Ψ . Every surface tangent at Z contained in the Klein quadric appears as a projection ray, and so the ideal part of the stereographic image is the same for all linear complexes. \diamond

Remark 4.3.3. The linear complexes \mathcal{C} with $c_{03} = 0$ contain the horizontal line at infinity. Their Klein image contains the projection center, so the stereographic image is the hyperplane $\mathcal{C}^o\sigma : -(c_{02} + c_{23})x_1 + (c_{01} - c_{31})x_2 + c_{23}x_3 + c_{31}x_4 + c_{12} = 0$. (cf. Equ. (4.27)). \diamond

Remark 4.3.4. The geometry of circles and lines in the Euclidean plane is called *Möbius geometry*. Circles may be defined as those real conics whose intersection with the ideal line has the equation $x_0 = x_1^2 + x_2^2 = 0$. It is well known (see Ex. 1.1.37 or the paragraph preceding Equ. (8.21)) that circles and lines are precisely the stereographic images of planar sections of the unit sphere in \mathbb{R}^3 .

Analogously, those quadrics in P^4 whose intersection with the ideal hyperplane has the equation $x_0 = x_1x_4 - x_2x_3 = 0$, together with the hyperplanes of P^3 , are the elements of a generalized Möbius geometry, and they are precisely the stereographic images of planar sections of the Klein quadric. \diamond

Fitting a Linear Complex to Data Lines

Consider a finite number of lines L_i in Euclidean three-space and assume a Cartesian coordinate system such that no line is horizontal. Then the angles enclosed by the lines L_i and the z -axis are less or equal some value $\phi_0 < \pi/2$. A smaller ϕ_0 is better for what follows than a larger one.

We apply the transformation σ of Equ. (4.21) to L_i and get points $X_i = L_i\sigma$ with coordinate vectors $\mathbf{x}_i \in \mathbb{R}^4$. We have to fit a surface of the form (4.27) to them. One possibility to solve this problem is the following (cf. [193]): Consider the function

$$F_{\mathbf{a}}(\mathbf{x}) = a_0(x_1x_4 - x_2x_3) + a_1x_1 + \cdots + a_4x_4 + a_5,$$

where $\mathbf{a} = (a_0, \dots, a_5)$, $\mathbf{x} = (x_1, x_2, x_3, x_4)$. We could define a ‘distance’ of the point $\mathbf{x}_0 \in \mathbb{R}^4$ to the surface $F_{\mathbf{a}} = 0$ by the value $|F_{\mathbf{a}}(\mathbf{x}_0)|$. However, what we actually want is to measure distances between points with the quadratic form q defined by Equ. (4.24). In order to relate these two distance functions in a neighbourhood of the surface $F_{\mathbf{a}}(\mathbf{x}) = 0$, we first consider ordinary Euclidean distance in \mathbb{R}^4 , which has no geometric meaning in line space. The function $F_{\mathbf{a}}$ has the Taylor expansion

$$F_{\mathbf{a}}(\mathbf{x} + \mathbf{h}) = F_{\mathbf{a}}(\mathbf{x}) + \langle \text{grad}_{F_{\mathbf{a}}}(\mathbf{x}), \mathbf{h} \rangle + o(\|\mathbf{h}\|). \quad (4.28)$$

The symbol $\langle \cdot, \cdot \rangle$ denotes the ordinary Euclidean scalar product, and the symbol $o(f(\mathbf{h}))$ means a remainder term with the property that $o(f(\mathbf{h}))/\|\mathbf{h}\| \rightarrow 0$ as \mathbf{h} converges towards zero.

Further, $F_{\mathbf{a}}(\mathbf{x})/\|\text{grad}_{F_{\mathbf{a}}}(\mathbf{x})\|$ is a first order approximant to the distance of a point \mathbf{x} to the surface $F_{\mathbf{a}} = 0$, as \mathbf{x} tends towards this surface.

Let G equal the coordinate matrix of the bilinear form q (see Equ. (4.25)). We consider the q -gradient of $F_{\mathbf{a}}$, which is defined by

$$\text{grad}_{F_{\mathbf{a}}}^{(q)}(\mathbf{x}) = G^{-1} \cdot \text{grad}_{F_{\mathbf{a}}}(\mathbf{x}). \quad (4.29)$$

Then the Taylor expansion (4.28) is valid if we replace the scalar product, norm, and gradient, by their q -variants, as defined by (4.26). This is clear from definition of the q -gradient. Moreover, the function

$$\tilde{F}_{\mathbf{a}}(\mathbf{x}) = F_{\mathbf{a}}(\mathbf{x})/\|\text{grad}_{F_{\mathbf{a}}}^{(q)}(\mathbf{x})\|_q$$

is a first order approximant of the signed q -distance of the point \mathbf{x} to the surface $F_{\mathbf{a}} = 0$. We therefore look for \mathbf{a} which minimizes

$$\sum_{i=1}^k (\tilde{F}_{\mathbf{a}}(\mathbf{x}_i))^2. \quad (4.30)$$

Then the surface $F_{\mathbf{a}}(\mathbf{x}) = 0$ will be a reasonable least squares fit for the data points X_i . Because the coefficients of \mathbf{a} enter F in a linear way and because of the low polynomial degree of the problem this nonlinear least-squares problem is computationally tractable.

Remark 4.3.5. If the surface $F_{\mathbf{a}}(\mathbf{x}) = 0$ has a singular point \mathbf{s} (this happens if the coefficient vector \mathbf{a} belongs to a singular linear complex), then the vector $(\nabla_{\mathbf{x}} F_{\mathbf{a}})(\mathbf{s})$ is zero, and so data points near \mathbf{s} will cause problems (such data points belong to lines near the axis of the singular complex).

The surfaces $F_{\lambda\mathbf{a}}(\mathbf{x}) = 0$ and $F_{\mathbf{a}}(\mathbf{x}) = 0$ are the same. This shows that in order to minimize the expression in (4.30) it is necessary to normalize the vector \mathbf{a} in some way. The methods to solve this minimization problem are similar to those of Sec. 4.1. \diamond

Fitting a Linear Congruence to Data Lines

A linear line congruence \mathcal{N} is the carrier of a pencil of linear complexes. If this pencil is spanned by linear complexes \mathcal{C} and \mathcal{C}' , then $\mathcal{N}\sigma = \mathcal{C}\sigma \cap \mathcal{C}'\sigma$. The two equations $\mathcal{C}^o\sigma : a'(x_1x_4 + x_2x_3) + \dots$ and $\mathcal{C}'^o : a''(x_1x_4 + x_2x_3) + \dots$ of the form (4.21) which define $\mathcal{N}^o\sigma$ have the same quadratic terms, so an appropriate linear combination is linear. We see that $\mathcal{N}\sigma$ is defined by one linear and one quadratic equation, so it is a quadric or possibly a singular quadratic variety contained

in a three-dimensional hyperplane of \mathbb{R}^4 . Because σ preserves the linear incidence structure, it is clear that oval and ruled quadrics $\mathcal{N}\sigma$ correspond to elliptic and hyperbolic linear congruences \mathcal{N} , respectively. A parabolic congruence \mathcal{N} is mapped to a quadratic cone, and a bundle or field of lines is mapped to a plane.

If L_1, L_2, \dots is a finite set of lines, we consider the stereographic image points $\mathbf{x}_i = L_i\sigma$. In order to fit a linear congruence to the lines L_i , we first look for a hyperplane H which contains, as best as possible, all points \mathbf{x}_i , where ‘best’ is in the sense of the metric define by the quadratic form q of Equ. (4.24). This procedure is easy and consists of an obvious modification of the same procedure in Euclidean space: Assume that H has the equation $\langle \mathbf{a}, \mathbf{x} \rangle_q + a_0 = 0$, where the q -scalar product is defined by Equ. (4.26). We minimize $\sum_i (\langle \mathbf{a}, \mathbf{x}_i \rangle_q + a_0)^2$ under the side condition that $\|\mathbf{a}\|_q = 1$. This leads to an ordinary eigenvalue problem.

Having found H , we project the points \mathbf{x}_i q -orthogonally onto H . Within H , we have to approximate the points \mathbf{x}_i by a quadric Q which appears as intersection of one of the quadrics of Equ. (4.27) with H . This is similar to the problem of fitting a complex to data points — the quadratic coefficients of Q ’s equations are already known up to a common factor.

If the best fitting quadric degenerates into a plane within H , then the original hyperplane fitting problem would have had a one-parameter family of solutions. This case can be detected by the appearance of two small solutions of the characteristic equation analogous to (4.7).

Fitting a Regulus to Data Lines

The Klein image of a linear congruence is a (possibly degenerate) quadratic variety contained in a three-dimensional subspace of P^5 . A regulus is a quadratic contained in a subspace which is two-dimensional. So the problem of fitting a regulus to data lines is transformed into the problem of fitting a plane to the data points, and afterwards fitting a conic (or possibly a line) to a planar set of points. Like in the previous case, the quadratic coefficients in the conic’s equation are already known up to a scalar factor. The metric used for a least squares fit is based on the quadratic form q of Equ. (4.24).

Interpolating and Approximating Real-Valued Functions of Lines

We consider the problem of *scattered data interpolation* and *approximation* for functions defined in line space. Assume that we are given a finite sequence of data lines L_i and real numbers r_i . We also define a region of interest D within line space, where the interpolant is to be defined.

We want to construct a function F which is defined on all lines L within some domain of interest, and either exactly or approximately assumes the values r_i at the lines L_i .

One possible way to solve this problem is the following [140]: Assume that a finite number of open spherical caps Γ_i cover the unit sphere, and that this covering

is symmetric with respect to the origin. A antipodal pair $+\Gamma_i, -\Gamma_i$ of caps is determined by its axis A_i of rotational symmetry and its spherical radius ρ_i . For each axis A_i , consider the set $\mathcal{L}_i \subset D$ of lines with $\sphericalangle(A_i, L) < \rho_i$.

We choose a Cartesian coordinate system such that A_i is parallel to its z -axis and consider the mapping σ of Equ. (4.21). Thus for all i we get image points $\sigma_i(L_j) \in \mathbb{R}^4$ for those data lines L_j which are contained in \mathcal{L}_i . It is well known how to perform scattered data interpolation and approximation in Euclidean \mathbb{R}^4 (cf. [78]), so we may assume that we have constructed real-valued functions F_i whose domain is \mathbb{R}^4 and which assume the values r_j at the points $\sigma_i(L_j)$. At last we have to glue the functions F_i together, using weight functions w_i defined on the unit sphere. Denote one of the unit vectors parallel to the line L by $\mathbf{l}(L)$. Then

$$F(L) = \sum_i w_i(\mathbf{l}(L)) F_i(\sigma_i(L))$$

is a solution, if the weight functions w_i have the following properties: (i) $w_i(\mathbf{n}) = w_i(-\mathbf{n})$ for all i, \mathbf{n} (the weights are symmetric with respect to the origin), (ii) $\sum_i w_i(\mathbf{n}) = 1$ for all \mathbf{n} (the weights form a partition of unity), and (iii) w_i is zero outside $\Gamma_i \cup (-\Gamma_i)$.

Remark 4.3.6. Examples of such function are the so-called *Franke-Little weights* (cf. e.g. [9], p.112), which are defined as follows: assume that \mathbf{a}_i is a unit vector parallel to the axis A_i and choose an integer $m \geq 1$. Then let

$$\begin{aligned}\tilde{w}_i(\mathbf{n}) &= \max(|\mathbf{n} \cdot \mathbf{a}_i| / \cos \rho_i, 0)^m \\ w(\mathbf{n}) &= \sum_i \tilde{w}_i(\mathbf{n}) \\ w_i(\mathbf{n}) &= \tilde{w}_i(\mathbf{n}) / w(\mathbf{n}).\end{aligned}$$

These weights are m times continuously differentiable, their support is $+\Gamma_i \cup -\Gamma_i$ by definition, and they sum up to one, also by definition. \diamond

Parallel Robots

Consider a *six-legged parallel manipulator* (or parallel robot, cf. Fig. 4.5). Here a rigid body Σ (the ‘moving system’) is connected with another one (the ‘fixed system’ Σ_0) by six bars of variable length (‘legs’), which have spherical links at both ends. A motion of Σ is understood as smooth family of positions $\Sigma(t)$, caused by a change in the legs’ lengths.

Geometrically, we may think of the legs L_i , $i = 1, \dots, 6$ as of lines. Special cases of such manipulators appear for example in flight simulators. They have, within bounds, the maximum degree of freedom possible, and the direct connection between the moving and the fixed system makes it possible to move heavy loads in a fast and precise manner. Research on parallel manipulators has been quite active in the past decade [123, 86]: We will show how the investigation of the mechanism’s stability benefits from line geometry and from approximation methods in line space [155].

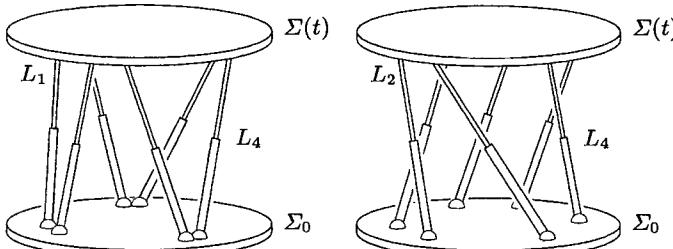


Fig. 4.5. Six-legged parallel manipulator. Left: Mechanism which is stable for all leg lengths. Right: Unstable position of another mechanism. The infinitesimal degree of freedom equals two.

Static Stability and Kinematic Stability

We consider ‘mechanical structures’ which consist of rigid bodies joined together by links. In statics there is the notion of *stability* of such a structure: A force acting on some point causes forces in the various joints of the structure, and if the structure is considered linearly elastic, also reversible deformations which depend linearly on these forces. An *unstable* or *singular* structure responds to exterior forces by ‘infinite’ induced forces — we can imagine that arbitrarily small exterior forces cause finite displacements.

Obviously if the structure admits a flexion (i.e., it is a *mechanism*), then it is unstable in this sense. A flexion consists of motions of the individual parts of the mechanical structure which are compatible with the joints.

For structures whose parts are joined by spherical links (*frameworks*), this notion of static stability is known to be equivalent to the following ‘kinematic’ stability.

Definition. Assume a framework F consisting of rigid parts K_0, \dots, K_n , where K_0 is assumed fixed. An infinitesimal flexion of F is an assignment of velocity vectors $\mathbf{v}(\mathbf{x})$ to all $\mathbf{x} \in F$ such that \mathbf{v} is not identically zero, but is zero in K_0 and coincides with an infinitesimal motion of Euclidean space if restricted to any K_i . If F admits an infinitesimal flexion, then F is called *unstable* or *singular*; otherwise it is called *stable*.

If the framework admits a smooth flexion which leaves K_0 fixed, then clearly the velocity vector field of this flexion serves as an infinitesimal flexion. It is even possible to show that a framework which admits any continuous flexion has an infinitesimal flexion, so that all flexible frameworks are singular. The following lemma is an immediate consequence of Th. 3.4.2:

Lemma 4.3.3. A six-legged parallel manipulator is singular as a framework if and only if the legs L_1, \dots, L_6 are contained in a linear line complex.

Proof. A velocity vector \mathbf{v}_i assigned to a L_i ’s ‘moving’ endpoint \mathbf{x}_i fits into an infinitesimal rigid body motion of L_i leaving the other endpoint fixed if and only if

$v_i \perp L_i$, and so L_i is a path normal for any infinitesimal motion of Σ . As the set of path normals is a linear complex for all infinitesimal motions, the lemma follows. \square

A singular position of the parallel manipulator is one where small forces cause large displacements, which makes it e.g. unfit for use in a flight simulator. In practice it is also important to keep at a certain distance from singular positions, because the above mentioned ratio of displacement to force is not actually infinite, but too large in a whole neighbourhood of a singular position.

Snapping between Neighbouring Positions

It may happen that the parts of two different frameworks are congruent, without the frameworks themselves being congruent. We can think of a framework taken apart and rebuilt in a different way. Two such positions may be very close to each other. This situation is to be avoided in structures actually built, because it can easily happen that the framework snaps from one position to the other.

It is easy to see that a parallel manipulator which has two neighbouring positions F_1, F_2 is very close to a singular structure: Consider the endpoints x_1, \dots, x_6 of the legs and their two position x_i^+ and x_i^- . Obviously there is a rigid body motion which transforms x_i^- to x_i^+ . According to Remark 3.1.6 the lines incident with the midpoints $m_i = \frac{1}{2}(x_i^+ + x_i^-)$ and orthogonal to the vector $x_i^+ - x_i^-$ belong to a linear complex. This shows that the closer the points x_i^+, x_i^- are, the closer is the manipulator to a singular one, and the better the legs can be fitted by a linear complex.

This can be checked with the tools shown in Sec. 4.1. Because of unavoidable imperfections of an actual structure these methods are also needed when testing for singular positions.

Degree of Singularity

If the legs are contained in a linear manifold of complexes, then its dimension plus one is called the *degree* of singularity. Degree 1 means that there is only one linear complex which the legs belong to. The methods presented in Sec. 4.1 can be used to detect such cases.

Example 4.3.2. As an example, we consider a so-called *Stewart-Gough platform* (see Fig. 4.5). Here the spherical links are arranged in two planes π of the moving and π^0 of the fixed system. We assume that the moving system moves in a translational manner such that the plane π remains constant. The legs have to vary their lengths and we get a two-parameter family of structures which is e.g. parametrized by the position of a point (u, v) within π .

Fig. 4.6 shows the deviation of the legs which belong to parameter values (u, v) to the nearest linear complex (the *stability function*). This deviation is defined by Equ. (4.3) and the nearest linear complex is found according to Lemma 4.1.2.

Computation of the stability function requires the fitting of many linear complexes to lines. This is done numerically and is reasonably fast if we have good

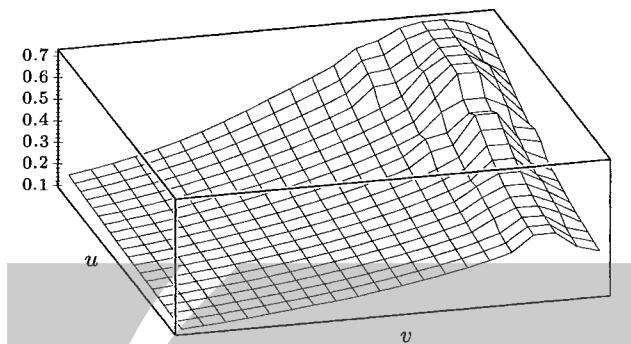


Fig. 4.6. The stability function of a six-legged parallel manipulator.

initial values for iterative algorithms — in this particular case initial values are provided by the solution complexes for neighbouring positions. \diamond

Remark 4.3.7. Line geometry also plays a role in the study of singular positions of *serial* robots. For example, if the mapping of the configuration space of a $6R$ robot to the Euclidean motion group is singular, then the six rotation axes are contained in a linear complex. \diamond

4.4 Approximation in the Set of Line Segments

We have already studied line segments in various places. Here we investigate briefly the set of *oriented line segments* of Euclidean 3-space E^3 and show how to solve approximation problems in it.

An *oriented line segment* \overrightarrow{pq} is identified with a member of $\mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$ in the obvious way:

$$(x_1, \dots, x_6) = (p_1, p_2, p_3, q_1, q_2, q_3). \quad (4.31)$$

Oriented line segments \overrightarrow{pp} of zero length are identified with a *point* of Euclidean space, and are mapped to the three-dimensional diagonal subspace Δ with equation $x_1 - x_4 = x_2 - x_5 = x_3 - x_6 = 0$.

The mapping $\overrightarrow{pq} \mapsto \overrightarrow{qp}$ is expressed in \mathbb{R}^6 by $(x_1, \dots, x_6) \mapsto (x_4, x_5, x_6, x_1, x_2, x_3)$.

Distances between Line Segments

In order to define a *distance* between two line segments $\overrightarrow{p_i q_i}$ ($i = 1, 2$), we do as in (4.22) and consider a similarity transformation which maps $p_1 \mapsto p_2$, $q_1 \mapsto q_2$:

$$(1 - \lambda)p_1 + \lambda q_1 \mapsto (1 - \lambda)p_2 + \lambda q_2 \quad (\lambda \in [0, 1]).$$

The distance is then defined as

$$\begin{aligned} d(\overrightarrow{\mathbf{p}_1 \mathbf{q}_1}, \overrightarrow{\mathbf{p}_2 \mathbf{q}_2})^2 &= 3 \int_0^1 ((1-\lambda)(\mathbf{p}_1 - \mathbf{p}_2) + \lambda(\mathbf{q}_1 - \mathbf{q}_2))^2 d\lambda \quad (4.32) \\ &= (\mathbf{p}_1 - \mathbf{p}_2)^2 + (\mathbf{q}_1 - \mathbf{q}_2)^2 + (\mathbf{p}_1 - \mathbf{p}_2) \cdot (\mathbf{q}_1 - \mathbf{q}_2). \end{aligned}$$

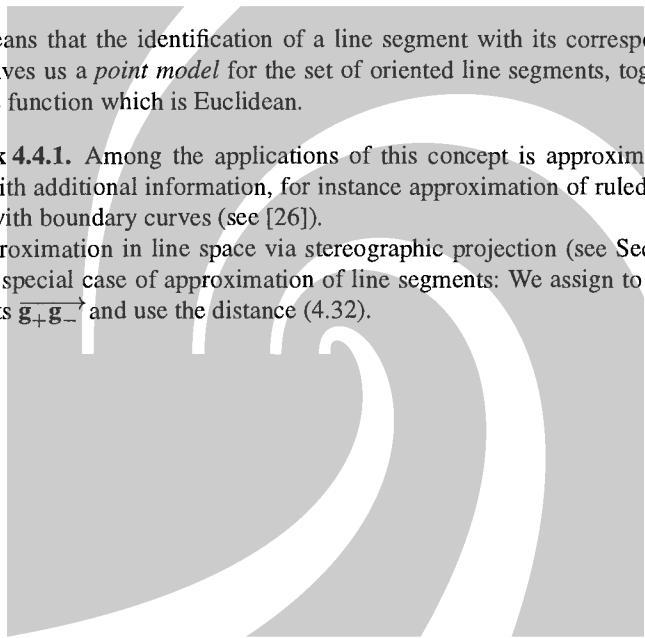
Obviously this distance is simply the distance between the two points in \mathbb{R}^6 which correspond to the segments $\overrightarrow{\mathbf{p}_i \mathbf{q}_i}$ with respect to the quadratic form

$$q(\mathbf{x}) = x_1^2 + \cdots + x_6^2 + x_1x_4 + x_2x_5 + x_3x_6. \quad (4.33)$$

This means that the identification of a line segment with its corresponding point in \mathbb{R}^6 gives us a *point model* for the set of oriented line segments, together with a distance function which is Euclidean.

Remark 4.4.1. Among the applications of this concept is approximation in line space with additional information, for instance approximation of ruled surfaces together with boundary curves (see [26]).

Approximation in line space via stereographic projection (see Sec. 4.3) is actually a special case of approximation of line segments: We assign to a line G the segments $\overrightarrow{\mathbf{g}_+ \mathbf{g}_-}$ and use the distance (4.32). \diamond



LuraTech

www.luratech.com

5. Ruled Surfaces

This chapter is devoted to ruled surfaces. A one-parameter smooth manifold of lines — we could say we take a line and move it around — traces out a smooth surface in space. We first study the subject from the viewpoint of projective differential geometry by discussing the distribution of tangent planes. It turns out that there are the classes of skew and torsal surfaces which exhibit a totally different tangent behaviour. The latter will be treated in a chapter of their own (see Chap. 6).

In Sec. 5.2, we look at algebraic ruled surfaces and especially at rational ones. An enumeration and projective classification of low degree surfaces on the one hand leads to interesting specimens, and on the other hand shows some kinds of typical behaviour by means of simple examples.

The third part of this chapter deals with the Euclidean differential geometry of ruled surfaces. This is in some aspects similar to the classical theory of space curves, which is only to be expected from an inherently one-dimensional object. Finally we show how to deal with ruled surfaces in numeric computations, and how to solve approximation and interpolation problems concerned with them.

5.1 Projective Differential Geometry of Ruled Surfaces

Ruled surfaces are one-parameter families of lines in projective three-space. We use the Klein mapping γ to give a definition of a ruled surface which is most suitable for our purposes:

Definition. A family $\mathcal{R} = R(u)$ of lines in P^3 is called a ruled surface, if its Klein image $\mathcal{R}\gamma$ is a curve $R\gamma(u)$ in the Klein quadric. Its differentiability class is defined to be the differentiability class of $R\gamma$. The lines $R(u)$ are the generator lines (generators, rulings) of the surface.

The meaning of this definition is that a C^r ruled surface has a C^r parametrization in Plücker coordinates:

$$c : I \rightarrow M_2^4, \quad u \mapsto (\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R}. \quad (5.1)$$

This situation is depicted in Fig. 5.1.

Note that the definition of ‘ruled surface’ depends on the definition of ‘curve’. Usually curves are either defined over an open interval, or over the unit circle, if they are closed curves.

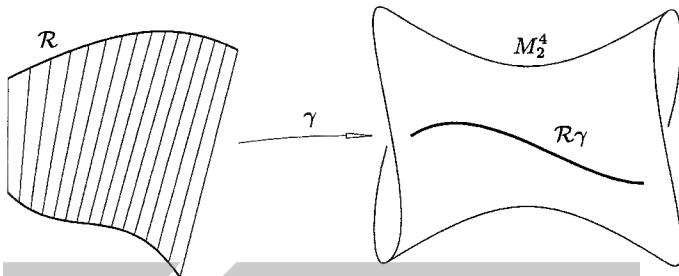


Fig. 5.1. Ruled surface \mathcal{R} and Klein image $\mathcal{R}\gamma$.

We will also define curves whose domain is the real projective line, which is nothing but the unit circle in disguise: The real projective line is the set of subspaces of a two-dimensional real vector space, such that the linear subspace spanned by (x_0, x_1) corresponds to the point $(x_0 : x_1)$. The unit circle in \mathbb{R}^2 intersects each such linear space in two points $\pm(x_0, x_1)$ with $x_0^2 + x_1^2 = 1$. This shows that the projective line is a closed curve. When following the unit circle, we come back to the same point already after an angle of 180 degrees.

However, when investigating infinitesimal or local properties of curves we will always consider curves defined in intervals.

Ruled Surface Parametrizations whose Domain is the Projective Line

Example 5.1.1. The Klein image of a *pencil of lines* is a straight line contained in the Klein quadric and can be parametrized by $(\lambda_0(\mathbf{p}_1, \bar{\mathbf{p}}_1) + \lambda_1(\mathbf{p}_2, \bar{\mathbf{p}}_2))\mathbb{R}$, where $\mathbf{p}_1 \cdot \bar{\mathbf{p}}_1 = \mathbf{p}_2 \cdot \bar{\mathbf{p}}_2 = \mathbf{p}_1 \cdot \bar{\mathbf{p}}_2 + \mathbf{p}_2 \cdot \bar{\mathbf{p}}_1 = 0$. These conditions mean that $(\mathbf{p}_1, \bar{\mathbf{p}}_1)$ and $(\mathbf{p}_2, \bar{\mathbf{p}}_2)$ are the Plücker coordinate vectors of two intersecting lines. If $(\lambda_0 : \lambda_1)$ is identified with a point of the projective line, this is a parametrization of a curve $c : P^1 \rightarrow M_2^4$, which shows that a pencil of lines corresponds to a *closed curve* in the Klein quadric. \diamond

Example 5.1.2. The Klein image of a *regulus* is a conic contained in the Klein quadric. There are several ways to write a conic as a rational curve defined over the projective line. Ex. 1.1.23 shows that any conic has, in a suitable coordinate system, the equation $x_0^2 + x_1^2 = x_2^2$, and Ex. 1.4.5 gives a parametrization of this conic as a rational Bézier curve.

Another way is the following: Assume that $a\mathbb{R}, b\mathbb{R}$ are points of the conic in P^n , and their tangents meet in $c\mathbb{R}$. If $d\mathbb{R}$ is another point of the conic, we may without loss of generality assume that $d = a + b + c$, because we can scale all vectors by arbitrary nonzero factors. It has been shown in Sec. 1.1 that a conic is uniquely determined by three points and the tangents in two of them. The closed rational curve

$$c : P^1 \rightarrow P^n, \quad (\lambda, \mu)\mathbb{R} = (\lambda : \mu) \mapsto (\lambda^2 a + \lambda\mu c + \mu^2 b)\mathbb{R}$$

coincides with the conic, if we can show (i) that $c(1 : 0) = \mathbf{a}\mathbb{R}$, $c(1 : 1) = \mathbf{d}\mathbb{R}$, $c(0 : 1) = \mathbf{b}\mathbb{R}$, and (ii) that c 's tangents in $\mathbf{a}\mathbb{R}$ and $\mathbf{b}\mathbb{R}$ are the lines $\mathbf{a}\mathbb{R} \wedge \mathbf{c}\mathbb{R}$ and $\mathbf{b}\mathbb{R} \wedge \mathbf{c}\mathbb{R}$, respectively. This is simply done by evaluation and differentiation. \diamond

Ruled Surfaces as Sets of Points

If \mathcal{R} is a ruled surface, the union of its lines has a parametrization as a two-surface of points in P^3 : Assume that the lines of \mathcal{R} are parametrized in the form $R(u)$, with $u \in D$ (where D is an interval or the projective line). If we choose curves $a(u) = \mathbf{a}(u)\mathbb{R}$, $b(u) = \mathbf{b}(u)\mathbb{R}$ such that $a(u) \vee b(u) = R(u)$, then

$$s : D \times \mathbb{R} \rightarrow P^3, \quad s(u, v) = ((1 - v)a(u) + vb(u))\mathbb{R}$$

is a parametrization of part of this surface, and

$$\begin{aligned} s : D \times P^1 &\rightarrow P^3, \\ s(u, \lambda_0 : \lambda_1) &= x(u, \lambda_0 : \lambda_1)\mathbb{R} = (\lambda_0 a(u) + \lambda_1 b(u))\mathbb{R} \end{aligned} \tag{5.2}$$

parametrizes the entire ruled surface.

Locally such curves $a(u)$ and $b(u)$ can be obtained by $a(u) = R(u) \cap \varepsilon_1$, $b(u) = R(u) \cap \varepsilon_2$, where ε_i are two planes. Any pair of curves $a(u)$, $b(u)$, with $a(u) \vee b(u) = R(u)$ is called a pair of *director curves*, or *directrices*.

If $R(u)$ is r times continuously differentiable, (2.17) shows that so are $a(u)$, $b(u)$, and $s(u, v)$. Conversely, if $a(u)$, $b(u)$ are C^r , then so is $R(u)$ and we see that the two representations (5.1) and (5.2) of a ruled surface are equivalent as far as local properties are concerned.

5.1.1 Infinitesimal Properties of First Order

A generator line $R(u)$ of a smooth ruled surface is called *regular* if $R(u)\gamma$ is a regular point of the curve $R\gamma(u)$ in the Klein quadric. If $R(u) = (\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R}$, then this means that the vectors $(\mathbf{r}, \bar{\mathbf{r}})$ and $(\dot{\mathbf{r}}, \dot{\bar{\mathbf{r}}})$ are linearly independent, or that the points $R\gamma$, $(R\gamma)^1 = (\dot{\mathbf{r}}, \dot{\bar{\mathbf{r}}})\mathbb{R}$ are distinct (for the notation, see p. 70). Their span $R\gamma(u) \vee (R\gamma)^1(u)$ is the tangent of the curve $R\gamma$ at the parameter value u .

A ruled surface with only regular generators is called *regular*.

If we are given directrices a , b as in (5.2), then $(\mathbf{r}, \bar{\mathbf{r}})\mathbb{R} = (\mathbf{a} \wedge \mathbf{b})\mathbb{R}$, and without loss of generality we can assume that $(\mathbf{r}, \bar{\mathbf{r}}) = \mathbf{a} \wedge \mathbf{b}$. By differentiation we get

$$(\dot{\mathbf{r}}, \dot{\bar{\mathbf{r}}}) = \dot{\mathbf{a}} \wedge \mathbf{b} + \mathbf{a} \wedge \dot{\mathbf{b}}. \tag{5.3}$$

The fact that the product rule for differentiation is also valid for exterior multiplication follows from its expression in coordinates. The point $\dot{\mathbf{a}}(u)\mathbb{R}$ is the tangent point $\mathbf{a}^1(u)$ of the director curve $a(u)$, and analogously $\mathbf{b}(u)\mathbb{R} = \mathbf{b}^1(u)$.

Lemma 5.1.1. *The ruled surface given by director curves $a(u)$, $b(u)$ has a singular generator at parameter value u if and only if the four points $a(u)$, $b(u)$, $a^1(u)$, $b^1(u)$ are collinear (cf. Fig. 5.2).*

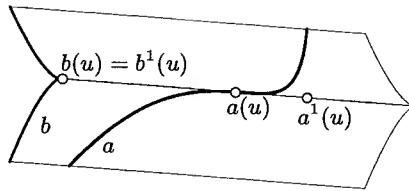


Fig. 5.2. Singular generator $a(u) \vee b(u)$ of a ruled surface with derivative points of directrices.

Proof. Clearly if these four points are collinear, then $\dot{a} = \lambda_1 a + \mu_1 b$, $\dot{b} = \lambda_2 a + \mu_2 b$, and substitution in (5.3) gives $\dot{a} \wedge b + a \wedge \dot{b} = (\lambda_1 + \mu_2)a \wedge b$, which means that the tangent point of the curve $R\gamma$ coincides with the point $R\gamma$ itself, and the surface is singular at this parameter value.

Conversely, assume that $\dot{a} \wedge b + a \wedge \dot{b} = \lambda a \wedge b$. Exterior multiplication with a gives $a \wedge \dot{a} \wedge b + a \wedge a \wedge \dot{b} = \lambda a \wedge a \wedge b$, which simplifies to

$$a \wedge \dot{a} \wedge b = 0, \text{ or } \dim[a, \dot{a}, b] \leq 2.$$

Analogously we get $\dim[b, \dot{b}, a] \leq 2$, which shows $\dim[a, \dot{a}, b, \dot{b}] = 2$. \square

Obviously the notion of ‘singular ruling’ is a *geometric* property of projective differential geometry of ruled surfaces: It is unchanged if we apply re-normalizations and regular parameter transforms, and if we apply projective transformations. This follows from the respective properties of the regularity of a curve, and from the fact that a projective automorphism of P^3 induces a projective automorphism of the Klein quadric.

Singular Surface Points

Even if a ruled surface has only regular rulings, the parametrization (5.2) of its point set may have singular points.

Lemma 5.1.2. *A regular ruling $R(u)$ carries at most one singular surface point. A singular ruling consists entirely of singular surface points.*

Proof. We consider surface curves $c(t)$ with $c(0) = p$ and look for their tangent points $c^1(t)$. If the set of all possible tangent points is a plane, the point p is a regular point of the surface.

A curve $(u(t), \lambda_0(t) : \lambda_1(t))$ in the parameter domain $I \times P^1$ defines a surface curve, if we substitute this expression in (5.2):

$$c(t) = c(t)\mathbb{R} = x(u(t), \lambda_0(t) : \lambda_1(t))\mathbb{R}.$$

We indicate differentiation with respect to t with a prime and compute the first derivative point $c^1(t) = c'\mathbb{R}$ of c :

$$\mathbf{c}' = \lambda'_0 \mathbf{a} + \lambda_0 u' \dot{\mathbf{a}} + \lambda'_1 \mathbf{b} + \lambda_1 u' \dot{\mathbf{b}} = \lambda'_0 \mathbf{a} + \lambda'_1 \mathbf{b} + u' (\lambda_0 \dot{\mathbf{a}} + \lambda_1 \dot{\mathbf{b}}), \quad (5.4)$$

which depends linearly on $\lambda'_0, \lambda'_1, u'$. If $\mathbf{a}, \mathbf{b}, \lambda_0 \dot{\mathbf{a}} + \lambda_1 \dot{\mathbf{b}}$ are linearly independent, then clearly there is a linear three-space of possible vectors \mathbf{c}' , therefore a projective two-space (a tangent plane) of possible c^1 's, and the surface is regular.

If the ruling $R(u)$ is singular, then by Lemma 5.1.1 these three vectors are always linearly dependent, so all points of $R(u)$ are singular.

If $R(u)$ is regular, they can be linearly dependent for at most one ratio $\lambda_0 : \lambda_1$ — this depends on the intersection of the planes $[\mathbf{a}, \mathbf{b}]$ and $[\dot{\mathbf{a}}, \dot{\mathbf{b}}]$ in \mathbb{R}^4 . The lemma is proved. \square

Remark 5.1.1. If $R(u)$ is a *singular* ruling, the preceding proof shows that the derivative point c^1 of any surface curve is contained in the ruling itself, which means that all surface curves touch the ruling (this includes surface curves which are singular there). \diamond

First Order Contact of Ruled Surfaces

We define first order contact of ruled surfaces via their Klein image:

Definition. Two ruled surfaces \mathcal{R} and \mathcal{S} , parametrized by $R(u)$ and $S(u)$, are said to be in first order contact at the parameter value u_0 , if the curves $R\gamma(u) = (\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R}$, $S\gamma(u) = (\mathbf{s}(u), \bar{\mathbf{s}}(u))\mathbb{R}$ are in first order contact at u_0 .

Thus first order contact means that after a regular parameter transform and appropriate re-normalization the curves $(\mathbf{r}, \bar{\mathbf{r}})$ and $(\mathbf{s}, \bar{\mathbf{s}})$ agree in their value and first derivative at the parameter value u_0 (cf. Sec. 1.2.1). Regular curves are in first order contact if and only if they have the same tangent in a common point.

Theorem 5.1.3. If two ruled surfaces \mathcal{R} , \mathcal{S} , parametrized by $R(u)$ and $S(u)$, are in first order contact at a common ruling, then their tangent planes in all points of this ruling coincide.

Proof. We can choose planes $\varepsilon_1, \varepsilon_2$ such that the curves $R(u) \cap \varepsilon_i$ and $S(u) \cap \varepsilon_i$ can be used as directrices ($i = 1, 2$). As we can compute the derivatives of the directrices from the derivatives of $R(u)\gamma$ and $S(u)\gamma$ (which are equal), and the derivatives of any surface curve from the derivatives of the directrices (by Equ. (5.4)), the tangent planes of \mathcal{R} and \mathcal{S} are the same in all points of the common ruling. \square

Example 5.1.3. Fig. 5.3 shows a ruled surface \mathcal{R} and a regulus \mathcal{S} which are in first order contact in all points of their common generator R_0 . The Klein image curves $\mathcal{R}\gamma$ and $\mathcal{S}\gamma$ touch each other in the point $R_0\gamma$. \diamond

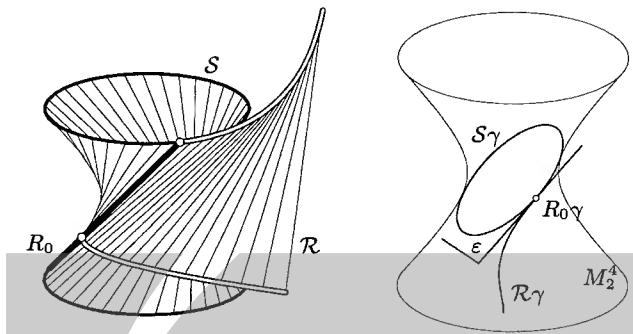


Fig. 5.3. Left: Ruled surface \mathcal{R} and regulus \mathcal{S} which are in first order contact at R_0 . Right: Klein image.

Non-torsal Generators

The proof of Lemma 5.1.2 showed that there are two types of regular rulings of a ruled surface: Rulings which carry a singular surface point, and rulings which do not. We discuss these two cases in more detail:

Definition. A regular generator $R(u)$ of a smooth ruled surface is called *torsal*, if the tangent of the curve $R\gamma(u)$ is contained in the Klein quadric, and *non-torsal* otherwise.

A tangent T of a quadric Φ is contained in Φ if and only if $T \cap \Phi$ has more than one point. The tangent of the curve $R\gamma(u)$ is clearly tangent to the Klein quadric, and it is entirely contained in M_2^4 if and only if one further point (e.g., the derivative point $(R\gamma)^1(u)$) is contained in M_2^4 .

It will turn out that among regular generators precisely the torsal ones carry a singular surface point. The points of a non-torsal generator are regular. Note that this definition does not involve singular generators. We will see later how to distinguish between classes of singular generators.

Theorem 5.1.4. All points of a non-torsal ruling $R(u_0)$ of a ruled surface are regular. The mapping of points of $R(u_0)$ to their tangent planes is a projective mapping. The set of surface tangents is a parabolic linear congruence \mathcal{N} , which fulfills the relation $\mathcal{N} = C(R\gamma(u_0) \vee (R\gamma)^1(u_0))$.

(For the notation, cf. Sec. 3.2).

Proof. We assume two director curves $a(u), b(u)$ as in (5.2). The ruling $R(u_0)$ is torsal if and only if the derivative point $(a \wedge b)^1 \mathbb{R} = (\dot{a} \wedge b + a \wedge \dot{b}) \mathbb{R}$ is contained in the Klein quadric. Equivalent to this is that the line $(\dot{a} \wedge b) \mathbb{R} \vee (a \wedge \dot{b}) \mathbb{R}$ is entirely contained in M_2^4 (because then three of its points are). Lines in M_2^4 correspond to line pencils of P^3 , so the ruling $R(u_0)$ is torsal if and only if the lines $a(u_0) \vee b^1(u_0)$ and $b(u_0) \vee a^1(u_0)$ are contained in a pencil, i.e., intersect.

For a non-torsal ruling $R(u_0)$ the four points $a(u_0), b(u_0), a^1(u_0), b^1(u_0)$ therefore are a tetrahedron and the points $a\mathbb{R}, b\mathbb{R}, (\lambda_0 a + \lambda_1 b)\mathbb{R}$ are never collinear. As the proof of Lemma 5.1.2 shows, in that case there is no singular surface point on $R(u_0)$.

Consider a plane which contains the tangent to the curve $R\gamma(u)$ at $R\gamma(u_0)$, but is not tangent to the Klein quadric. It intersects M_2^4 in a conic $S\gamma(u)$ which is in first order contact with $R\gamma(u)$. This shows that the regulus $S(u)$ is in first order contact with the surface $R(u)$, and so $S(u)$ and $R(u)$ have the same surface tangents along the common ruling $R(u_0) = S(u_0)$.

Now Th. 5.1.3, Lemma 3.3.4, and Th. 3.2.9, together with Ex. 3.3.2 and the discussion preceding it show the statement about the surface tangents. \square

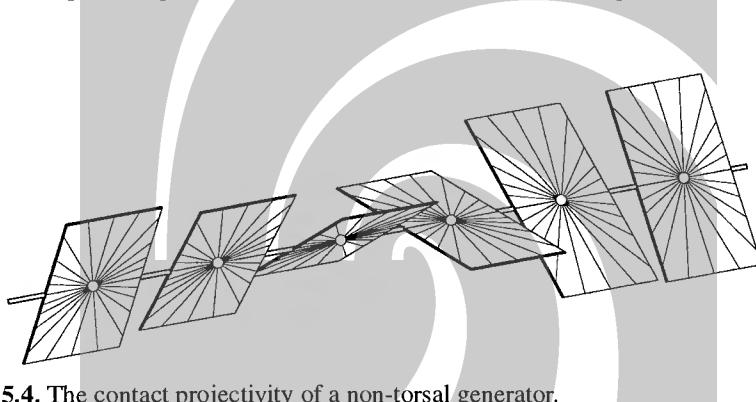


Fig. 5.4. The contact projectivity of a non-torsal generator.

Definition. *The projective mapping of points of a non-torsal ruling to their tangent planes is called the contact projectivity of this ruling.*

Fig. 5.4 illustrates the contact projectivity for a non-torsal generator, and also shows lines of the parabolic linear congruence \mathcal{N} mentioned in Th. 5.1.4.

Remark 5.1.2. We have proved that, within projective differential geometry, at a non-torsal ruling any ruled surface behaves like a regulus with respect to first order of differentiation. In fact there are a lot of reguli which are in this sense ‘tangent’ to the ruled surface, because we can choose many different conics tangent to the surface’s Klein image.

This can be compared to the situation on a quadric in three-space: For any curve on the quadric, there are a lot of conics which touch the curve there (if the curve tangent is not contained in the tangent space). \diamond

Being a non-torsal generator is a geometric property. This follows from the description of surface tangents, which is done using only objects of projective differential geometry.

Torsal Generators

Let us now consider torsal generators. The following theorem is an analogue to Th. 5.1.4 and describes the behaviour of tangent planes along a torsal ruling.

Theorem 5.1.5. *A regular torsal generator $R(u_0)$ of a ruled surface \mathcal{R} , parametrized by $R(u)$, carries exactly one singular surface point. All regular points possess the same tangent plane.*

The tangent $R\gamma(u_0) \vee (R\gamma)^1(u_0)$ of the Klein image curve lies in M_2^4 and is the Klein image of a line pencil, which is in first order contact with the ruled surface.

Definition. *The singular surface point of a torsal generator is called its cuspidal point, and the tangent plane in its other points is called torsal plane.*

Proof. (of Th. 5.1.5) We consider the line $S\gamma = R\gamma(u_0) \vee (R\gamma)^1(u_0)$. This line is contained in the Klein quadric. We parametrize it as a ruled surface $S\gamma(u)$, such that $S\gamma(u_0) = R\gamma(u_0)$ and $(S\gamma)^1(u_0) = (R\gamma)^1(u_0)$. The lines $S(u)$ are the lines of a pencil. $R(u)$ and $S(u)$ are in first order contact as ruled surfaces, therefore so are their respective point sets.

The surface traced out by $S(u)$ is a plane, but if we choose one director curve constant (in the vertex), and another one e.g. as a line, we see that the vertex is a singular point, whereas in all other points there is the same tangent plane. As surfaces which are in first order contact have the same surface tangents, this shows the statement of the theorem. \square

The proofs of Lemma 5.1.2 and Th. 5.1.4 show that singular, torsal, and non-torsal generator lines are distinguished in the following way:

$$\begin{aligned} \text{rk}(a(u_0), b(u_0), \dot{a}(u_0), \dot{b}(u_0)) = 2 &\iff R(u_0) \text{ is singular} \\ 3 &\iff R(u_0) \text{ is torsal} \\ 4 &\iff R(u_0) \text{ is non-torsal} \end{aligned} \quad (5.5)$$

which is equivalent to $\dim(a(u_0) \vee b(u_0) \vee a^1(u_0) \vee b^1(u_0))$ being equal to one, two, or three, respectively. This is expressed by

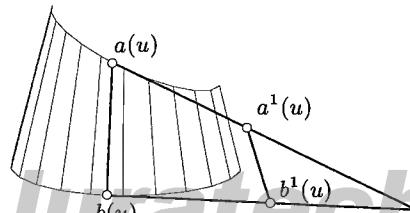


Fig. 5.5. Torsal generator $a(u) \vee b(u)$ with derivative points $a^1(u), b^1(u)$.

Corollary 5.1.6. Assume that the ruled surface \mathcal{R} is parametrized by $R(u) = a(u) \vee b(u)$, where a and b are directrices. If the tangents $a(u_0) \vee a^1(u_0)$ and $b(u_0) \vee b^1(u_0)$ are contained in the same line, the ruling $R(u_0)$ is singular. If they span a plane, then $R(u_0)$ is regular and torsal (cf. Fig. 5.5). If they are skew and span entire P^3 , this ruling is non-torsal

If the surface \mathcal{R} is parametrized in Plücker coordinates by $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$, we can compute the cuspidal point $(c_0(u), \mathbf{c}(u))\mathbb{R}$ and the torsal plane $\mathbb{R}(p_0(u), \mathbf{p}(u))$ as the vertex and the carrier of the pencil spanned by $R(u_0)$ and its derivative line whose Plücker coordinates are $(\dot{\mathbf{r}}(u), \ddot{\mathbf{r}}(u))$ (cf. the proof of 5.2). With Equ. (2.19) and (2.18) we get

$$(c_0, \mathbf{c})(u_0) = (\bar{\mathbf{r}} \cdot \dot{\mathbf{r}}, \bar{\mathbf{r}} \times \dot{\mathbf{r}})(u_0), \quad (p_0, \mathbf{p})(u_0) = (\mathbf{r} \cdot \dot{\mathbf{r}}, \mathbf{r} \times \dot{\mathbf{r}})(u_0). \quad (5.6)$$

Remark 5.1.3. Recall that (2.18) and (2.19) do not work in all cases. If they fail, we can use Equ. (2.4) to find three independent points contained in $R(u_0)$ or its derivative line. The span of these points yields the torsal plane.

Analogously, we can use Equ. (2.9) to find three independent planes which contain either $R(u_0)$ or its derivative line. Their intersection is the cuspidal point. \diamond

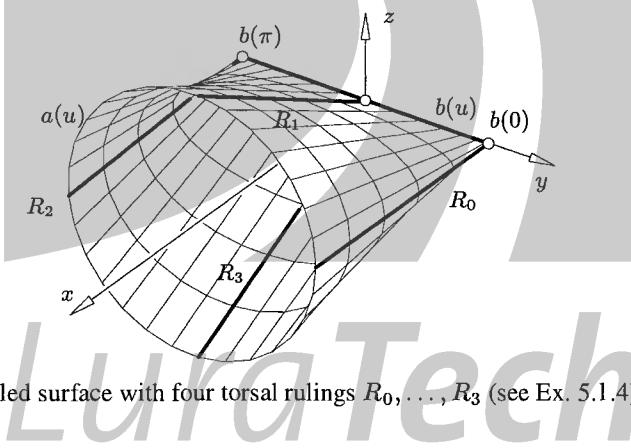


Fig. 5.6. Ruled surface with four torsal rulings R_0, \dots, R_3 (see Ex. 5.1.4).

Example 5.1.4. We consider a ruled surface defined by directrices $a(u) = \mathbf{a}(u)\mathbb{R}$, $b(u) = \mathbf{b}(u)\mathbb{R}$ in projectively extended Euclidean space, which is equipped with a Cartesian coordinate system. We let

$$\mathbf{a}(u) = (1, d, \cos u, \sin u), \quad \mathbf{b}(u) = (1, 0, \cos u, 0).$$

Obviously the rulings $R(u) = a(u) \vee b(u)$ intersect the y -axis under a right angle (see Fig. 5.6). We compute the rank of $\{\mathbf{a}, \mathbf{b}, \dot{\mathbf{a}}, \dot{\mathbf{b}}\}$ and find that it equals four if u is no integer multiple of $\pi/2$, where it equals three. The surface therefore has

precisely four regular torsal rulings, contained in its symmetry planes. Note that a ruling contained in a symmetry plane must be torsal unless the surface has a self-intersection there, because the contact projectivity never is symmetric.

Plücker coordinate representations of $R\gamma = (\mathbf{r}, \bar{\mathbf{r}})\mathbb{R} = (\mathbf{a} \wedge \mathbf{b})\mathbb{R}$ and its derivative point $(R\gamma)^1$ are given by

$$\begin{aligned} (\mathbf{r}, \bar{\mathbf{r}}) &= (d, 0, \sin u, \sin u \cos u, 0, -d \cos u), \\ (\dot{\mathbf{r}}, \dot{\bar{\mathbf{r}}}) &= (0, 0, \cos u, \cos 2u, 0, d \sin u). \end{aligned}$$

Obviously $(R\gamma)^1 \in M_2^4$ and therefore represents a line if and only if u is an integer multiple of $\pi/2$. This again shows the location of the torsal generators.

We compute the cuspidal point of $R(u)$ as the intersection of the line and its derivative line. For $u = 0$ and $u = \pi$ we see from $\mathbf{b}(u) = 0$ that $b(0)$ and $b(\pi)$ are the cuspidal points. Equ. (5.6) fails for $u = \pi/2$ and $u = 3\pi/2$, but obviously in both cases $\dot{\mathbf{r}}$ is zero, the derivative line is therefore an ideal line, and the cuspidal point is at infinity. The torsal planes are orthogonal to the symmetry planes for all four torsal generators. \diamond

Torsal Surfaces

Definition. A ruled surface all of whose rulings are torsal, is called a *torsal ruled surface*. If a ruled surface is not torsal, it shall be called a *skew ruled surface*.

Torsal ruled surfaces play an important role, and Chap. 6 will be entirely devoted to them. We first show some examples and then give a local classification theorem.

Example 5.1.5. A *cone* is a ruled surface with a *constant* director curve, which means a curve $c(u)$ with $c(u) = \mathbf{a}\mathbb{R} = \text{const}$. The point $\mathbf{a}\mathbb{R}$ is called the *vertex* of the cone. It can be parametrized by

$$\mathbf{x}(u, \lambda_0 : \lambda_1) = \lambda_0 \mathbf{a} + \lambda_1 \mathbf{b}(u), \quad \text{where } \mathbf{a} = \text{const.} \quad (5.7)$$

The rulings of this surface are regular if no tangent of the curve $\mathbf{b}(u) = \mathbf{b}\mathbb{R}(u)$ degenerates or contains the point $a = \mathbf{a}\mathbb{R}$. As $\dot{\mathbf{a}}(u) = 0$, we have $\text{rk}(\mathbf{a}, \dot{\mathbf{a}}, \mathbf{b}, \dot{\mathbf{b}}) = 3$ in regular points, so all rulings are torsal.

In the projective extension of Euclidean space, the cones with ideal vertex are called *cylindrical surfaces* or cylinders. A special case of a cone is a pencil of lines which had already been encountered in the proof of Th. 5.1.5. \diamond

Example 5.1.6. Other examples of torsal ruled surfaces are the *tangent surfaces* of regular C^2 curves $a(u) = \mathbf{a}\mathbb{R}(u)$ in projective three-space. Its family of tangents $R(u) = a(u) \vee a^1(u)$ has the parametrization

$$\mathbf{x}(u, \lambda_0 : \lambda_1) = (\lambda_0 a(u) + \lambda_1 \dot{a}(u))\mathbb{R}. \quad (5.8)$$

The curves $a(u)$ and $a^1(u)$ are the surface's director curves, and $a(u)$ is called its *curve of regression* or *edge of regression*. If it has no inflection points, then $\text{rk}(\mathbf{a}, \dot{\mathbf{a}}, \ddot{\mathbf{a}}) = 3$ and (5.5) shows that then $R(u)$ has only regular torsal rulings.

The cuspidal point of the ruling $R(u_0)$ is $a(u_0)$, and the osculating plane $a(u_0) \vee a^1(u_0) \vee a^2(u_0)$ of the curve $a(u)$ is the torsal plane. \diamond

We show that any torsal ruled surface is locally one of these examples, and globally a (maybe complicated) composition of them. This is described by the following theorem.

Theorem 5.1.7. *Assume that $R(u)$ ($u \in I$) parametrizes a C^2 torsal ruled surface. All intervals $J \subset I$ contain an interval J' where the surface either is a cone or a tangent surface.*

Proof. The main idea of the proof is the following: If a curve is singular in an entire interval, it is constant, and the set of parameter values for which a curve is singular, is closed. This argument is used in several places.

If the curve $R\gamma$ is singular in J , then it is constant in J . This degenerate case was forbidden by the definition of ‘torsal generator’.

Otherwise there is a regular generator $R(u_0)$ with $u_0 \in J$, and $R(u)$ is regular in a neighbourhood of u_0 . Consider the curve $c(u)$ of cuspidal points. If $c(u)$ is singular in an interval J' , it is constant in J' , and the surface $R(u)$ is a cone for $u \in J'$.

Otherwise there is $u'_0 \in J$ such that $c(u'_0)$ is nonsingular. Then $c(u'_0)$ is nonsingular in an interval J' containing u'_0 . We assume that $R(u)$ has Plücker coordinates $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$. The torsality condition $(R\gamma)^1 \in M_2^4$ then reads $\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} = 0$. We compute the derivative point $c^1 = \dot{c}\mathbb{R}$ of the curve $c(u) = c\mathbb{R}(u)$ by differentiating (5.6):

$$\dot{c} = (\dot{\bar{\mathbf{r}}} \cdot \dot{\mathbf{r}} + \bar{\mathbf{r}} \cdot \ddot{\mathbf{r}}, \dot{\bar{\mathbf{r}}} \times \dot{\mathbf{r}} + \bar{\mathbf{r}} \times \ddot{\mathbf{r}}) = (\bar{\mathbf{r}} \cdot \ddot{\mathbf{r}}, \bar{\mathbf{r}} \times \ddot{\mathbf{r}}),$$

Equations (2.23) and (5.9) show that $c^1(u)$ is contained in the ruling $R(u)$, so we have $R(u) = c(u) \vee c^1(u)$ and the surface $R(u)$ is indeed the tangent surface of $c(u)$ for $u \in J'$. \square

A picture of a tangent surface can be seen in Fig. 5.7.

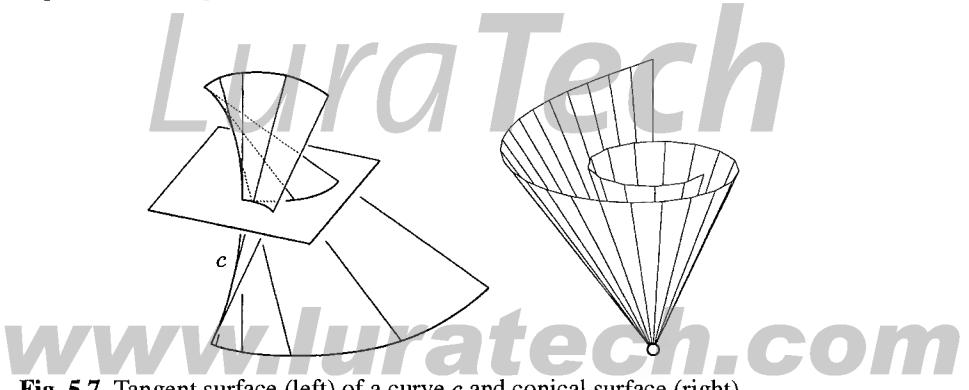


Fig. 5.7. Tangent surface (left) of a curve c and conical surface (right).

Remark 5.1.4. Note that the singular generators and the generators where the cuspidal curve is singular without being constant can have accumulation points without the surface being constant or conical. These phenomena will not be important for ‘practical’ applications. \diamond

Remark 5.1.5. The Klein image of a torsal ruled surface $R(u)$ is a curve all of whose tangents are contained in the Klein quadric. We know that an *asymptotic tangent* of a surface is one which is in second order contact with this surface. This shows that the tangents of $\mathcal{R}\gamma$ are asymptotic tangents.

A curve all of whose tangents are asymptotic tangents, is an asymptotic curve, by definition. This shows that asymptotic curves $\mathcal{R}\gamma$ correspond to torsal ruled surfaces.

The Klein image of a cone is contained in a plane, because all of its rulings are incident with the vertex, and therefore are contained in a bundle (cf. Lemma 2.1.9). \diamond

5.1.2 Infinitesimal Properties of Higher Order

Contact of Order k

A concept very important for the investigation of higher order infinitesimal properties is the order of contact. Recall that two curves or surfaces are said to have *contact of order k* at some parameter value u , if after a regular parameter transform they agree in derivatives of order $0, 1, \dots, k$ at this parameter value. For fixed point of contact, this is an equivalence relation for curves and surfaces which are k times continuously differentiable.

Two ruled surfaces now have two different contact orders: Their Klein images have some order of contact as curves, and their point sets, parametrized by Equ. (5.2), have a contact order as surfaces. The following theorem asserts that contact of curves ensures contact of surfaces. The proof is analogous to that of Th. 5.1.3.

Theorem 5.1.8. Assume that $R_1(u)$, $R_2(u)$ are C^k parametrizations of ruled surfaces \mathcal{R}_1 , \mathcal{R}_2 . If the curves $R_i\gamma$ have contact of order k at u_0 as curves in the Klein quadric, then the surfaces have contact of order k in all points of the ruling $R_1(u_0) = R_2(u_0)$.

Proof. If the Klein images $R_i\gamma = (\mathbf{r}_i, \bar{\mathbf{r}}_i)\mathbb{R}$ agree in their first k derivatives, then so do planar sections of the surfaces \mathcal{R}_i , which are computed by 2.17. This shows that the surface parametrizations given by (5.2) agree in their first k derivatives. \square

In the terminology of Computer Aided Geometric Design, this result says that G^r joins of curves on the Klein quadric imply G^r joins of ruled surfaces along generators.

Th. 5.1.8 is one of the basic tools when studying infinitesimal properties of ruled surfaces, because it allows to use projective differential geometry of curves. We will discuss second order properties later.

Remark 5.1.6. It is possible that ruled surfaces have second order contact in all regular points of a torsal generator, but that their Klein images do not have second order contact. An example will be given by the osculating cone of a torsal ruled surface (Th. 6.1.4).

On the other hand, we will see that second order contact in all points of a non-torsal generator line implies second order contact of the Klein images (see Prop. 5.1.11) A consequence of this behaviour of torsal generators is that the Klein model is useful for the design of non-torsal ruled surfaces, but perhaps less so for the design of torsal ones. \diamond

Singular Generators

If a ruled surface \mathcal{R} , parametrized by $R(u)$, has a singular generator $R(u_0)$, then it is in first order contact with the degenerate ruled surface \mathcal{S} , parametrized by $S(u) = R(u_0) = \text{const}$. We consider the osculating subspaces

$$R\gamma, \quad R\gamma \vee (R\gamma)^1, \quad R\gamma \vee (R\gamma)^1 \vee (R\gamma)^2, \dots$$

and the sequence of their dimensions for $u = u_0$ (cf. Th. 1.2.2). Because $R(u_0)$ is a singular generator, this sequence starts with 0, 0. If the $(k+1)$ -th osculating subspace $R\gamma \vee \dots \vee (R\gamma)^{k+1}$ is the first to have dimension one, then k is called the *order of singularity* of the generator $R(u_0)$. In this case the surface $R(u)$ is in k -th order contact with the constant ruled surface $S(u)$. If no finite order exists, we call $R(u_0)$ a singular generator of infinite order.

If the order is finite and equals k , the *limit tangent* of $R\gamma$ at $u = u_0$ is spanned by $R\gamma(u_0)$ and $(R\gamma)^{k+1}(u_0)$. We distinguish between *singular torsal* and *singular non-torsal* generators, depending on whether the limit tangent is contained in the Klein quadric or not.

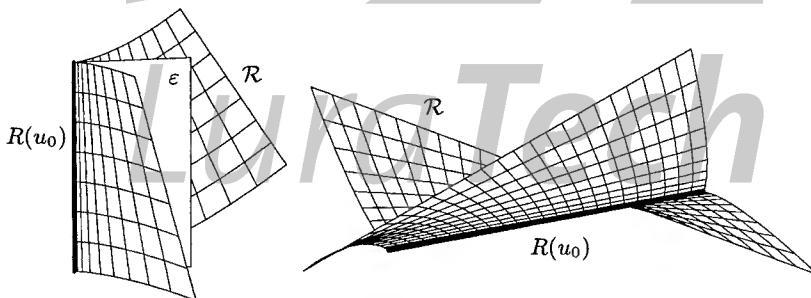


Fig. 5.8. Left: Singular torsal generator $R(u_0)$ with limit tangent plane ε . Right: Singular non-torsal generator $R(u_0)$.

Example 5.1.7. Fig. 5.8, left, shows a cone whose director curve is a semicubic parabola (see Fig. 1.40, right). The singular torsal ruling $R(u_0)$ is incident with the

singular point of the directrix. The limit tangent $R\gamma(u_0) \vee (R\gamma)^2(u_0)$ is the Klein image of a line pencil, whose vertex is the cone's vertex, and whose carrier plane ε can be seen as limit tangent plane of \mathcal{R} in the points of the singular generator. \diamond

Example 5.1.8. An example of a singular non-torsal ruling $R(u_0)$ is shown by Fig. 5.8, right. There is no common limit tangent plane for all points of $R(u_0)$ — the distribution of limit tangent planes is the same as for a regular non-torsal generator.

This ruled surface is contained in an elliptic linear congruence with rotational symmetry, and the planar sections shown in Fig. 5.8 are semicubic parabolae. \diamond

Second Order Behaviour at Non-torsal Generators

The Plücker coordinates $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$ of a ruled surface fulfill the Plücker identity $\mathbf{r}(u) \cdot \bar{\mathbf{r}}(u) = 0$. Repeated differentiation of this identity with respect to u results in the sequence

$$0 = \dot{\mathbf{r}} \cdot \bar{\mathbf{r}} + \mathbf{r} \cdot \dot{\bar{\mathbf{r}}} = \ddot{\mathbf{r}} \cdot \bar{\mathbf{r}} + 2\dot{\mathbf{r}} \cdot \dot{\bar{\mathbf{r}}} + \mathbf{r} \cdot \ddot{\bar{\mathbf{r}}} = \dots \quad (5.9)$$

These identities are useful when studying higher order infinitesimal properties of ruled surfaces. The first identity can be used to show (5.5) and just says that the derivative point $(R\gamma)^1$ is contained in the tangent plane of the Klein quadric's at $R\gamma$.

Theorem 5.1.9. For all non-torsal generators $R(u_0)$ of a C^2 ruled surface \mathcal{R} , there is a regulus which is in second order contact with \mathcal{R} in all points of $R(u_0)$.

Definition. The regulus mentioned in Th. 5.1.9 belongs to a quadric, which is called the Lie quadric of the generator $R(u_0)$.

Proof. (of Th. 5.1.9) We assume that \mathcal{R} is parametrized by $R(u)$, and $R(u)$ has the Plücker coordinates $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$. Since $R(u_0)$ is non-torsal, we have $\dot{\mathbf{r}} \cdot \bar{\mathbf{r}} \neq 0$. The second identity of (5.9) says that also $\ddot{\mathbf{r}} \cdot \bar{\mathbf{r}} + \mathbf{r} \cdot \ddot{\bar{\mathbf{r}}} = \Omega((\mathbf{r}, \bar{\mathbf{r}}), (\ddot{\mathbf{r}}, \ddot{\bar{\mathbf{r}}})) \neq 0$ locally. This means that the second derivative point $(R\gamma)^2(u_0)$ is not conjugate to $R\gamma$, or that the osculating plane $R\gamma \vee (R\gamma)^1 \vee (R\gamma)^2$ is not tangent to the Klein quadric.

Therefore the osculating plane intersects the Klein quadric M_2^4 in a conic, which is in second order contact with the curve $R\gamma$. As any conic in M_2^4 is the Klein image of a regulus, the theorem now follows from Th. 5.1.8. \square

Example 5.1.9. Consider the ruled surface \mathcal{R} with Plücker coordinate parametrization $R(u)\gamma = (\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R}$

$$(\mathbf{r}, \bar{\mathbf{r}})(u) = (1 + u - u^2, -1 + u + u^2, 1, 1 - u^2, u, -1 + u^2 - u^4).$$

The osculating plane $O^2 = R\gamma \vee (R\gamma)^1 \vee (R\gamma)^2$ is parametrized by

$$\lambda_0(\mathbf{r}, \bar{\mathbf{r}}) + \lambda_1(\dot{\mathbf{r}}, \dot{\bar{\mathbf{r}}}) + \lambda_2(\ddot{\mathbf{r}}, \ddot{\bar{\mathbf{r}}}).$$

Its intersection with the Klein quadric equals the Klein image of the Lie quadric. It is computed by solving $(\lambda_0\mathbf{r} + \lambda_1\dot{\mathbf{r}} + \lambda_2\ddot{\mathbf{r}}) \cdot (\lambda_0\bar{\mathbf{r}} + \lambda_1\dot{\bar{\mathbf{r}}} + \lambda_2\ddot{\bar{\mathbf{r}}}) = 0$. If we let $u = -1/2$, we get the equation

$$-4\lambda_0\lambda_2 + 2\lambda_1^2 - 4\lambda_1\lambda_2 + 4\lambda_2^2 = 0.$$

A picture of the surface \mathcal{R} and its Lie quadric \mathcal{S} for $u = -1/2$ is shown in Fig. 8.15, together with a central projection of the Klein image curves $\mathcal{R}\gamma$ and $\mathcal{S}\gamma$ (cf. Ex. 8.1.6). \diamond

A color picture of a Lie quadric is shown by Fig. C.4.

Corollary 5.1.10. *If $R(u_0)$ is a non-torsal generator of a ruled surface, then the asymptotic tangents in the points of $R(u_0)$ which are different from $R(u_0)$, comprise the other regulus of the Lie quadric.*

Proof. By Th. 5.1.9, the Lie quadric and the ruled surface in question are in second order contact. Therefore they share the asymptotic tangents, which, for a ruled quadric, comprise one of its reguli. \square

Fig. C.4 shows asymptotic tangents of a ruled surface \mathcal{R} . In this case the intersection of \mathcal{R} and its Lie quadric \mathcal{S} contains an asymptotic tangent. We have already mentioned that second order contact of ruled surfaces in all regular points of a torsal generator need not imply that their Klein image curves have second order contact (Remark 5.1.6). For non-torsal generators, this is different:

Proposition 5.1.11. *If two ruled surfaces $\mathcal{R}_1, \mathcal{R}_2$ are in second order contact in all points of a non-torsal generator, then their Klein image curves $\mathcal{R}_1\gamma, \mathcal{R}_2\gamma$ are in second order contact.*

Proof. This follows from the fact that $\mathcal{R}_1, \mathcal{R}_2$ share the same Lie quadric, whose Klein image is the intersection of M_2^4 with the osculating planes of both $\mathcal{R}_1\gamma$ and $\mathcal{R}_2\gamma$ (see the proof of Th. 5.1.9). These two osculating planes are equal and it is well known that two surface curves with a common non-tangent osculating plane are in second order contact. \square

This result together with Remark 5.1.6 shows to what extend the two different notions of second order contact of ruled surfaces coincide.

Second Order Behaviour at Torsal Generators

We continue the discussion of second order properties, which becomes slightly more complicated for regular torsal generators. The principle of analysis is the same as for non-torsal generators. Depending on the intersection of $R\gamma$'s osculating plane with the Klein quadric, there are the possibilities listed below.

We denote the curve tangent $R\gamma \vee (R\gamma)^1$ with T^1 , and the subspace $R\gamma \vee (R\gamma)^1 \vee (R\gamma)^2 = T^1 \vee (R\gamma)^2$ with the symbol O^2 (for *osculating plane*, although it need not be two-dimensional).

(i) If O^2 is not contained in M_2^4 , it intersects M_2^4 in T and a further line L , which may coincide with T . The line L is the Klein image of a line pencil — its vertex is called *osculation point*, and its plane *osculating plane*.

These names are perhaps misleading because the osculation plane does not have second order contact with the ruled surface. The geometric meaning of the osculation plane is the following: A planar section of the ruled surface which contains the cuspidal point, but does not contain the generator, has a cusp in the cuspidal point, with a limit tangent contained in the osculation plane.

(ii) If $\dim(O^2) = 2$ and O^2 is contained in M_2^4 , we choose a conic in O^2 which has second order contact with the curve $R\gamma$. There are two cases to be distinguished:

(iia) O^2 is the Klein image of a bundle: Then the conic is the Klein image of a quadratic cone, which is then in second order contact with the ruled surface.

(iib) O^2 is the Klein image of a field of lines: Then the conic is the Klein image of the family of tangents to a planar conic, which is again in second order contact with $R(u)$.

(iii) If $\dim(O^2) = 1$ (and therefore $O^2 = T^1$), the curve $R\gamma$ has an inflection point. By Th. 5.1.8, the surface $R(u)$ is in second order contact with the pencil $(T^1)\gamma^{-1}$.

Example 5.1.10. We consider again the surface $R(u)$ of Ex. 5.1.4 and especially the torsal ruling $R(0)$. The subspace $O^2 = R\gamma \vee (R\gamma)^1 \vee (R\gamma)^2$ has the parametrization $(\mathbf{r}, \bar{\mathbf{r}}) + \lambda(\dot{\mathbf{r}}, \ddot{\mathbf{r}}) + \mu(\ddot{\mathbf{r}}, \ddot{\mathbf{r}})$. For $u = 0$, we compute $O^2 \cap M_2^4$ and get $\lambda\mu = 0$. This shows that we have case (i). The line pencil whose Klein image corresponds to the line $\mu = 0$, is spanned by $R(0)$ and the second derivative line R^2 . The carrier plane of this pencil (the osculation plane of the torsal ruling $R(0)$) is spanned by $b(0) = (1, 0, 1, 0)\mathbb{R}$ and R^2 , which has the Plücker coordinates $(\mathbf{o}, \mathbf{e}_3)$ — with Equ. (2.16) we get the result $\mathbb{R}(0, 0, 0, 1)$. \diamond

5.2 Algebraic Ruled Surfaces

In this section we use the *complex extensions* $\mathbb{C}P^3$ and $\mathbb{C}P^5$ of real projective three-space P^3 and five-space P^5 whenever necessary, with the obvious embeddings of the real spaces into the complex ones, as well as the obvious embedding of the real Klein quadric in the complex Klein quadric, which is defined by the same equation.

Definition. An algebraic ruled surface \mathcal{R} in P^3 is the Klein preimage of a real algebraic curve $\mathcal{R}\gamma$ in the Klein quadric $M_2^4 \subset P^5$ (cf. Fig. 5.1).

Properties of algebraic varieties, such as irreducibility or the degree always refer to the image $\mathcal{R}\gamma$ of \mathcal{R} . However, we will show that no ambiguities can arise, because the point set associated with an algebraic ruled surface is again algebraic, and irreducibility and degree are the same as for $\mathcal{R}\gamma$. First we show a simple lemma:

Lemma 5.2.1. We use homogeneous coordinates x_0, \dots, x_n in P^n . Consider the affine spaces $A_i \subset P^n$ whose equations are $x_i \neq 0$. Then a subset $M \subset P^n$ is a projective algebraic variety, if $M \cap A_i$ is an affine algebraic variety for all A_i .

Proof. Let $M_i = M \cap A_i$. If we can find a projective algebraic variety M'_i such that $M'_i \cap A_i = M_i$ and $M \subset M'_i$ for all i , then obviously $M = M'_0 \cap \dots \cap M'_n$.

The sets M'_i are defined as follows: We have $M_i = V(F_1, \dots, F_k)$, with polynomials F_1, \dots, F_k not involving x_i . We make homogeneous polynomials $\tilde{F}_1, \dots, \tilde{F}_k$ from them by multiplying all monomials with an appropriate power of x_i . Then we let $M'_i = V(x_i \tilde{F}_1, \dots, x_i \tilde{F}_k)$. By construction $M'_i \cap A_i = M_i \cap A_i$ and $M'_i \setminus A_i = P^n \setminus A_i$, so M'_0, \dots, M'_n have the required properties. \square

Theorem 5.2.2. *The union Φ of generators of a complex irreducible algebraic ruled surface \mathcal{R} is an irreducible hypersurface in $\mathbb{C}P^3$.*

The proof is broken into two lemmas, which are interesting in their own right:

Lemma 5.2.3. *If \mathcal{R} is a set of lines, and $\mathcal{R}\gamma$ is a projective algebraic subvariety of the Klein quadric, then the union $\Phi \subset \mathbb{C}P^3$ of all lines in \mathcal{R} is a projective algebraic variety.*

Proof. We consider the relation in $\mathbb{C}P^3 \times \mathbb{C}P^5$ defined by $p \sim q$ if and only if $q \in \mathcal{R}\gamma$ and p is incident with the line L whose Klein image is q . The incidence condition (2.23) shows that ‘ \sim ’ is an algebraic relation. Clearly $p \in \Phi$ if and only if there is a q with $p \sim q$, and by Prop. 1.3.16, Φ is a projective algebraic variety. \square

Lemma 5.2.4. *If \mathcal{R} is the set of lines contained in a hypersurface $\Phi : F(\mathbf{x}) = 0$ in $\mathbb{C}P^3$, then $\mathcal{R}\gamma$ is a projective algebraic variety in $\mathbb{C}P^5$.*

Proof. We use Plücker coordinates (l_{01}, \dots, l_{12}) for lines, and consider lines with $l_{01} \neq 0$. From Equ. (2.4) we choose the points $s_0 \mathbb{C}, s_1 \mathbb{C} \in L$, which have coordinates $s_0 = (0, l_{01}, l_{02}, l_{03})$ and $s_1 = (-l_{01}, 0, l_{12}, -l_{31})$. Because $l_{01} \neq 0$, these two points are well-defined and not equal.

If $\deg(F) = n$, the line L is contained in Φ if and only if

$$\left. \frac{d^k}{dt^k} \right|_{t=0} F(s_0 + ts_1) = 0, \quad k = 0, \dots, n. \quad (5.10)$$

This condition is polynomial in the indeterminates l_{01}, \dots, l_{12} , and so is the Plücker relation. Thus $\mathcal{R}\gamma$ intersects the five-dimensional affine space $l_{01} \neq 0$ in an affine algebraic variety.

We repeat this for lines with $l_{02} \neq 0, \dots, l_{12} \neq 0$, and apply Lemma 5.2.1 to conclude that $\mathcal{R}\gamma$ is a projective algebraic variety. \square

Proof. (of Th. 5.2.2) Lemma 5.2.3 shows that Φ is an algebraic variety. We are going to show that Φ is irreducible.

Assume that $\Phi = \Phi' \cup \Phi''$, and consider the sets S', S'' of lines contained in Φ' , Φ'' , respectively, and let $\mathcal{R}' = \mathcal{R} \cap S'$, $\mathcal{R}'' = \mathcal{R} \cap S''$. Obviously $\mathcal{R} = \mathcal{R}' \cup \mathcal{R}''$. Lemma 5.2.4 shows that $\mathcal{R}'\gamma, \mathcal{R}''\gamma$ are algebraic. $\mathcal{R}\gamma$ is irreducible, so $\mathcal{R}' = \mathcal{R}$ or $\mathcal{R}'' = \mathcal{R}$. In the first case all points of Φ are contained in some line of \mathcal{R}' , i.e., in a

line contained in Φ' , so $\Phi = \Phi'$. The second case corresponds to $\Phi = \Phi''$, so Φ is irreducible.

To show that Φ is a hypersurface we would have to use properties of ‘dimension’ which have not been proved in this book, e.g., the fact that dimensions of algebraic varieties equal their dimensions as differentiable manifolds in regular points. \square

Remark 5.2.1. If Φ is irreducible, this does not imply that $\mathcal{R}\gamma$ is. A counterexample is provided by a quadric Φ carrying two reguli. We may choose $\mathcal{R}\gamma$ as the union of two conics. \diamond

Remark 5.2.2. Th. 5.2.2 is not true for real algebraic ruled surfaces, only for complex ones. An example is provided by the tangent surface of a planar curve, e.g., an ellipse. The union of all generators is the outside of the ellipse, but the smallest algebraic variety which contains this set is the entire plane. \diamond

For algebraic ruled surfaces, the property of being skew is much stronger than for non-algebraic ones.

Lemma 5.2.5. *An algebraic ruled surface has, apart from a finite number of singular generators, either only torsal generators, or only a finite number of them.*

Proof. The singularity condition is polynomial, and is therefore satisfied everywhere or only in a finite number of points. If it is satisfied everywhere, the ruled surface consists of a single ruling, its Klein image is only a point, and this has been excluded by the definition of ‘algebraic ruled surface’.

The torsality condition is also polynomial and is therefore satisfied always or only in a finite number of points. \square

Thus it makes sense to distinguish between *skew* and *torsal* algebraic ruled surfaces.

The following facts are given without proof, because they are very intuitive anyway and we do not want to overburden the reader with algebraic geometry.

Proposition 5.2.6. *If \mathcal{R} is an irreducible algebraic torsal ruled surface, the dual of its point set Φ is a point in the case that Φ is planar, and is a curve otherwise; the dual of a non-torsal algebraic ruled surface is again a surface.*

The proof depends on the fact that for all generators of a torsal ruled surface there is exactly one tangent plane. The mapping of the curve $\mathcal{R}\gamma$ to $(P^3)^*$, which assigns to a generator its tangent plane, cannot increase the dimension.

Proposition 5.2.7. *Assume that \mathcal{R} is an algebraic ruled surface and Φ its point set. If Φ is no quadric and no plane, then all points of Φ which are incident with more than one ruling, or are a cuspidal point of a ruling, are singular points of the surface (in the sense of algebraic geometry). These singular points are contained in an algebraic subvariety of dimension ≤ 1 .*

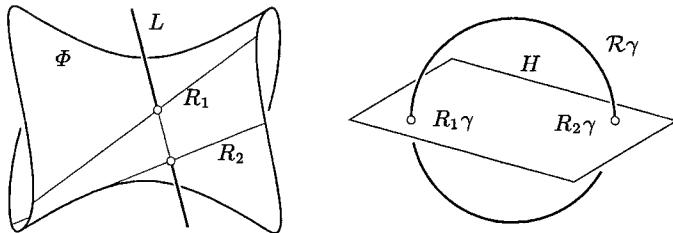


Fig. 5.9. A line L intersects an algebraic ruled surface \mathcal{R} with point set Φ .

The Degree of a Ruled Surface

An algebraic ruled surface has a degree as a curve in the Plücker quadric and also as a point set. The following theorem shows that these two degrees are the same.

Theorem 5.2.8. *Assume that \mathcal{R} is an irreducible algebraic ruled surface, whose point set Φ is not planar. Then $\deg(\mathcal{R}\gamma) = \deg(\Phi)$. If \mathcal{R} is not torsal, also the class of Φ (the degree of its dual surface) equals $\deg(\mathcal{R}\gamma)$.*

Proof. The theorem is true for reguli (since then $\deg(\mathcal{R}\gamma) = \deg(\Phi) = 2$) and false for planar Φ if $\deg(\mathcal{R}\gamma) > 1$. Assume now that Φ is no plane and no quadric.

A generic line L intersects Φ transversely in $\deg(\Phi)$ points, and does not meet a singular point of Φ . Prop. 5.2.7 shows that it therefore intersects $\deg(\Phi)$ lines of \mathcal{R} (see Fig. 5.9).

If the set of lines which meet L is denoted by $\mathcal{C}(L)$, this means that $\deg(\Phi) = \#(\mathcal{R} \cap \mathcal{C}(L))$. $\mathcal{C}(L)$ is a singular line complex with $\mathcal{C}(L)\gamma = H \cap M_2^4$, where H is a hyperplane tangent to the Klein quadric, and $\deg(\Phi) = \#(H \cap \mathcal{R}\gamma)$ (see Fig. 5.9). Obviously $\mathcal{R}\gamma \not\subset H$, and it is not difficult to show that the intersection is transverse, so $\deg(\mathcal{R}\gamma) = \#(H \cap \mathcal{R}\gamma) = \#(\Phi \cap L) = \deg(\Phi)$.

Assume now that \mathcal{R} is not torsal, so its dual is indeed a surface (cf. Prop. 5.2.6). In order to compute the generic number of tangent planes incident with a given line, we note that by Th. 5.1.4 at non-torsal generators R all planes which contain R occur as tangent plane of exactly one point of R . This shows that the number of intersection points of L with Φ equals the number of Φ 's tangent planes incident with L , and so $\deg(\Phi^*) = \deg(\Phi) = \deg(\mathcal{R})$. \square

The class of a *torsal* ruled surface is to be defined in a different way:

Definition. *The class m of a torsal ruled surface is defined as the generic number of tangent planes incident with a point.*

If \mathcal{R} is a cone, its class equals the number of tangent planes which contain a generic line incident with the vertex. We apply this to prove the following

Proposition 5.2.9. *The apparent contour of an algebraic non-torsal ruled surface \mathcal{R} with respect to a central projection is an algebraic curve of class $\deg(\mathcal{R})$, if the projection center C is not contained in \mathcal{R} 's point set Φ .*

Proof. We consider the tangent cone Ψ of Φ with vertex in C . Its intersection with the image plane is the apparent contour c , and the class of c equals the class of Ψ .

To compute the class of Ψ , we dualize the entire situation: Let Φ^* equal the set of tangent planes of Φ , and let C^* equal the bundle of planes incident with C . C^* is a plane in P^{3*} . The tangent planes of Φ which are incident with C , comprise the set $\Psi^* = \Phi^* \cap C^*$.

The condition that $C \not\subset \Phi$ ensures that C^* and Φ^* intersect transversely, so the intersection curve $C^* \cap \Phi^*$ has the same degree as Φ^* . This shows that the class of c equals the class of Ψ , the class of Φ , and the degree of \mathcal{R} . \square

5.2.1 Rational Ruled Surfaces

Definition. A rational ruled surface \mathcal{R} is a ruled surface such that $\mathcal{R}\gamma$ is a rational curve.

A parametrization $(\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R}$ with rational functions $\mathbf{r}(u) = (r_{01}(u), r_{02}(u), r_{03}(u))$, $\bar{\mathbf{r}}(u) = (r_{23}(u), r_{31}(u), r_{12}(u))$, can always be transformed into a polynomial one by multiplying with a common denominator. By substituting $u = u_1/u_0$ and by multiplying with an appropriate power of u_0 we can make all coordinate functions homogeneous polynomials in the indeterminates u_0, u_1 :

$$(R\gamma)(u) = (\mathbf{r}(u), \bar{\mathbf{r}}(u))\mathbb{R} = \left(\sum_{i=0}^d (\mathbf{c}_i, \bar{\mathbf{c}}_i) u_1^i u_0^{d-i} \right) \mathbb{R}. \quad u = u_0 : u_1 \in P^1. \quad (5.11)$$

So $R(u)$ is a ruled surface whose parameter domain is the projective line. By the definition of ‘rational ruled surface’, all such surfaces have a representation in the form (5.11).

Remark 5.2.3. There may be parameter values u where $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$ is zero and defines no projective point. This is easily avoided: If α is a common zero of all polynomials $r_{ij}(u)$, all of them are divisible by $u - \alpha$. By successive division we can therefore eliminate all common zeros without altering the rest of the curve, and make it defined for all arguments. \diamond

Proposition 5.2.10. We use the notation of Equ. (5.11): If $(\mathbf{r}, \bar{\mathbf{r}})$ has no zeros, then the set of points $(\mathbf{r}, \bar{\mathbf{r}})(u_0 : u_1)\mathbb{C}$ for $(u_0 : u_1) \in \mathbb{C}P^1$ is an algebraic curve, and Equ. (5.11) parametrizes an algebraic ruled surface.

Proof. (Sketch) By Th. 2.2.1.6, $(\mathbf{p}, \bar{\mathbf{p}})\mathbb{C} = (\mathbf{r}(u_0 : u_1), \bar{\mathbf{r}}(u_0 : u_1))\mathbb{C}$, if and only if the exterior product $(\mathbf{p}, \bar{\mathbf{p}}) \wedge (\mathbf{r}, \bar{\mathbf{r}})$ in $\Lambda^2 \mathbb{C}^6$ is zero, provided both factors are nonzero. As $(\mathbf{r}, \bar{\mathbf{r}})$ has no zeros, this condition is fulfilled, and the equality above is an algebraic relation between points $(\mathbf{p}, \bar{\mathbf{p}})\mathbb{C}$ in $\mathbb{C}P^5$ and parameter values $(u_0 : u_1) \in \mathbb{C}P^1$. Prop. 1.3.16 shows that the image of $(\mathbf{r}, \bar{\mathbf{r}})$ is an algebraic variety. The statement about the dimension is not shown here. \square

Remark 5.2.4. Prop. 5.2.10 is neither true for real parametrizations, nor if we allow zeros of $(\mathbf{r}, \bar{\mathbf{r}})$. In these cases the point set of a rational ruled surface is only *part* of an algebraic surface. We give two simple examples:

Consider two intersecting lines A, B with Plücker coordinates $(\mathbf{a}, \bar{\mathbf{a}})$ and $(\mathbf{b}, \bar{\mathbf{b}})$. The ruled surface \mathcal{R} parametrized by $R\gamma(u_0 : u_1) = (u_0^2(\mathbf{a}, \bar{\mathbf{a}}) + u_1^2(\mathbf{b}, \bar{\mathbf{b}}))\mathbb{R}$, where $(u_0 : u_1) \in P^1$, is defined for all $u_0 : u_1$, and parametrizes only part of the pencil spanned by A, B , whereas the smallest algebraic ruled surface which contains \mathcal{R} is just this pencil.

An example for the complex number field with zeros of $(\mathbf{r}, \bar{\mathbf{r}})$ is the following: $R\gamma(u_0 : u_1) = (u_0 u_1 (\mathbf{a}, \bar{\mathbf{a}}) + u_1^2 (\mathbf{b}, \bar{\mathbf{b}}))\mathbb{R}$ with $u_0 : u_1 \in \mathbb{C}P^1$ parametrizes the pencil spanned by A and B without the line A . \diamond

The following is an immediate consequence of Lemma 1.3.21 and Prop. 1.3.22:

Corollary 5.2.11. *If (5.11) is a Lüroth parametrization of the rational ruled surface \mathcal{R} , then this surface has degree d as an algebraic ruled surface.*

The Point Set of a Rational Ruled Surface

Two planes $\varepsilon_0, \varepsilon_1$ intersect the ruled surface \mathcal{R} given by (5.11) in two rational curves $a_0(u) = \mathbf{a}_0(u)\mathbb{R}$, $a_1(u) = \mathbf{a}_1(u)\mathbb{R}$, as is obvious from the intersection formula (2.17).

We assume that the rational parametrizations of a_0 and a_1 are of order d_0 and d_1 , respectively. Equ. (5.2), which describes the point set Φ of a ruled surface, takes the form

$$\begin{aligned} s(u_0 : u_1, v_0 : v_1) &= v_0 a_0(u_0 : u_1) + v_1 a_1(u_0 : u_1) = \\ &= v_0 \sum_{i=0}^{d_0} \mathbf{a}_{i,0} u_1^i u_0^{d_0-i} + v_1 \sum_{j=0}^{d_1} \mathbf{a}_{j,1} u_1^j u_0^{d_1-j}, \quad \text{with } (u_0 : u_1), (v_0 : v_1) \in P^1. \end{aligned} \quad (5.12)$$

This parametrizes the point set Φ of \mathcal{R} , except for those generators which intersect both a_0 and a_1 in the same point (and obviously cannot be recovered from these two points, which are equal). Equ. (5.12) is not restricted to *planar* curves a_0 and a_1 — it is valid for any directrices.

Conversely, any parametrization of the form (5.12) defines a rational ruled surface via $(\mathbf{r}, \bar{\mathbf{r}})(u_0 : u_1) = \mathbf{a}_0(u_0 : u_1) \wedge \mathbf{a}_1(u_0 : u_1)$. The following theorem shows the existence of a ‘minimal’ representation of the form (5.12):

Theorem 5.2.12. *All rational ruled surfaces \mathcal{R} of degree d in complex projective space P^3 have a representation as a linear blend of the form (5.12) of rational director curves a_0, a_1 of degrees d_0, d_1 with $d_1 = d - d_0 \geq d_0$, and such that $a_0(u_0 : u_1) \neq a_1(u_0 : u_1)$ for all $(u_0 : u_1) \in P^1$.*

Proof. We assume a Lüroth parametrization $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$ of $\mathcal{R}\gamma$, which determines Lüroth parametrizations of generic planar intersections $a_0(u) = \mathbf{a}_0(u)\mathbb{R}$, $a_1(u) =$

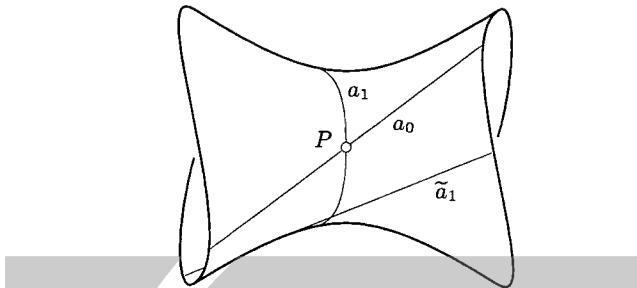


Fig. 5.10. Directrices $a_0(u)$, $a_1(u)$ with a common value $P = a_0(u_0) = a_1(u_0)$. New directrix \tilde{a}_1 .

$\mathbf{a}_1(u)\mathbb{R}$. Let $\deg(a_0) = d_0$, $\deg(a_1) = d_1$. We choose a plane $\varepsilon : \mathbf{h} \cdot \mathbf{x} = 0$ and compute its intersection c with the ruled surface, which is parametrized by

$$\mathbf{c}(u) = -(\mathbf{h} \cdot \mathbf{a}_1(u))\mathbf{a}_0(u) + (\mathbf{h} \cdot \mathbf{a}_0(u))\mathbf{a}_1(u).$$

The degree of $\mathbf{c}(u)$ is obviously less than or equal to $d_0 + d_1$. From the definition of degree of an algebraic variety it is clear that for almost all planes ε we have $\deg(c) = \deg(\Phi)$. We are going to show that

$$d \leq d_0 + d_1, \quad (5.13)$$

with equality if and only if $a_0(u) \neq a_1(u)$ for all u . This is done as follows: The component functions of \mathbf{c} have a common factor $(u - \alpha)$ (which means $d < d_0 + d_1$) if and only if $\mathbf{c}(\alpha) = 0$, i.e.,

$$\mathbf{c}(\alpha) = -(\mathbf{h} \cdot \mathbf{a}_1(\alpha))\mathbf{a}_0(\alpha) + (\mathbf{h} \cdot \mathbf{a}_0(\alpha))\mathbf{a}_1(\alpha) = 0.$$

This equation implies $a_0(\alpha) = a_1(\alpha)$, which is shown as follows: If $a_0(\alpha) \neq a_1(\alpha)$, then $\mathbf{a}_0(\alpha)$ and $\mathbf{a}_1(\alpha)$ are linearly independent. Not both $\mathbf{h} \cdot \mathbf{a}_0(\alpha)$ and $\mathbf{h} \cdot \mathbf{a}_1(\alpha)$ vanish, as ε was supposed to contain no generator line. This contradicts $\mathbf{c}(\alpha) = 0$.

Division by $u - \alpha$ diminishes the degree, and we have shown (5.13). We now show that it is possible to choose director curves a_0 , a_1 without common points: Assume that $\mathbf{a}_0(\alpha) = \lambda \mathbf{a}_1(\alpha)$. Then we consider the new director curve

$$\tilde{\mathbf{a}}_1 : \tilde{\mathbf{a}}_1(u) = -\mathbf{a}_0(u) + \lambda \mathbf{a}_1(u).$$

$\tilde{\mathbf{a}}_1$ has at least one zero for $u = \alpha$, and division by an appropriate power of $u - \alpha$ gives a new director curve which is defined for all arguments, and can be used instead of a_1 (see Fig. 5.10). Obviously its degree is less than $\max(d_0, d_1)$ and for $u = \alpha$ it assumes a value different from $a_0(\alpha)$.

If there are still common values of the two directrices, we repeat this procedure. At last (5.13) shows the statement of the theorem. \square

Linear Families of Director Curves

Proposition 5.2.13. *If the director curves whose existence is shown by Th. 5.2.12 have degrees d_0, d_1 with $d_0 < d_1$, then there is only one director curve which has degree d_0 . In any case there is a linear $(d_1 - d_0 + 1)$ parameter family of director curves of degree d_1 .*

Proof. If there are two director curves of degree d_0 , we could use them to define the ruled surface, which would then have degree $\leq 2d_0 < d$.

If $d_0 \leq d_1$, consider a curve of the form

$$\mathbf{c}(u) = v_0(u)\mathbf{a}_0(u) + v_1\mathbf{a}_1(u), \quad (5.14)$$

with a polynomial $v_0(u)$ of degree $d_1 - d_0$ and a constant v_1 . It is a director curve of degree d_1 . Therefore there is a linear $(d_1 - d_0 + 1)$ parameter family of director curves of degree d_1 . \square

If $d_0 = d_1$, there is a one-parameter linear family of director curves of degree d_1 , given by Equ. (5.14) with constants v_0, v_1 .

Example 5.2.1. Rational cones of degree d are obtained if one director curve is constant ($d_0 = 0$). There is a linear $(d_1 + 1)$ -parameter family of director curves of degree d_1 . \diamond

Example 5.2.2. If $\deg(\mathcal{R}) = 2$, Th. 5.2.12 gives us directrices a_0, a_1 of degree $d_0 = 0, d_1 = 2$, or $d_0 = d_1 = 1$. The case of a directrix of zero degree leads to a quadratic cone. If $d_0 = d_1 = 1$, then the correspondence $a_0(u) \longleftrightarrow a_1(u)$ is a projective mapping, and Prop. 1.1.44 shows that the surface is a regulus. The director curves are two lines of its complementary regulus. \diamond

Example 5.2.3. If $\deg(\mathcal{R}) = 4$, Th. 5.2.12 ensures the existence of director curves of degrees $(2, 2)$ or of degrees $(1, 3)$, (or possibly $(0, 4)$, which leads to a cone). As has been shown above, the first kind of surface carries a one-parameter family of director curves of degree two, i.e., conics. Fig. 5.11, right shows an example. The parametrizations used in (5.12) obviously define a projective mapping between these conics. Among this type of ruled surfaces there are torsal ruled surfaces, namely the tangent surfaces of twisted cubics (cf. Ex. 1.3.15).

The second kind of ruled surface has a unique director line and carries a three parameter linear family of cubic curves. If such a surface is torsal, then it must be planar because it has a director line. Fig. 5.11, left shows an example of a surface of the second kind. \diamond

Example 5.2.4. We show how to compute directrices of minimal degree according to Th. 5.2.12. First we consider the ruled surface parametrized by

$$\begin{aligned} (\mathbf{r}(u), \bar{\mathbf{r}}(u)) = & (2u(1+u^2), (1+u^2)(u^2-2u+3), \\ & -(1+u^2)(3+u^2), -(3+u^2)^2, 2u(3+u^2), -4u^2). \end{aligned}$$

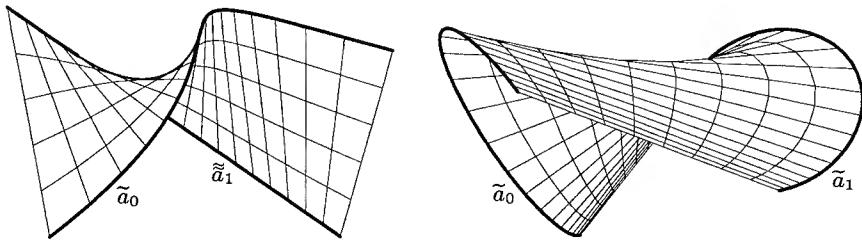


Fig. 5.11. Left: Quartic ruled surface with a linear and a cubic directrix. Right: Quartic ruled surface with two quadratic directrices (courtesy M. Peternell).

We use Equ. (2.17) to compute the intersection curves $\mathbf{a}_0(u)\mathbb{R}$ and $\mathbf{a}_1(u)\mathbb{R}$ with the planes $\mathbb{R}(-1, 0, 0, 1)$ and $\mathbb{R}(1, 0, 0, 1)$. Common points of these curves are $\mathbf{a}_0(\pm i)\mathbb{R} = \mathbf{a}_1(\pm i)$ and $\mathbf{a}_0(\pm i\sqrt{3})\mathbb{R} = -\mathbf{a}_1(\pm i\sqrt{3})$. If we let $\tilde{\mathbf{a}}_0 = (\mathbf{a}_0 - \mathbf{a}_1)/(1 + u^2)$ and $\tilde{\mathbf{a}}_1 = (\mathbf{a}_0 + \mathbf{a}_1)/(3 + u^2)$, then $\tilde{\mathbf{a}}_i(u) = \tilde{\mathbf{a}}_i(u)\mathbb{R}$ are directrices without common points. The result of this division is

$$\tilde{\mathbf{a}}_0(u) = (0, 2u, u^2 - 2u + 3, -(3 + u^2)), \quad \tilde{\mathbf{a}}_1(u) = (1 + u^2, 2u, 3 + u^2, 0).$$

Fig. 5.11, right, shows rulings and representatives of a linear family of conics of this surface. Another example is the homogeneous ruled surface parametrization

$$(\mathbf{r}(u_0 : u_1), \bar{\mathbf{r}}(u_0 : u_1)) = (u_0^3(u_0 - u_1), -u_1 u_0^2(u_0 + u_1), -u_1 u_0(u_1^2 - u_0^2), u_1^3(u_0 + u_1), u_1^2 u_0(u_1 - u_0), -2u_0^2 u_1^2).$$

We use Equ. (2.17) to compute the intersection curves $\mathbf{a}_0(u_0 : u_1)\mathbb{R}$ and $\mathbf{a}_1(u_0 : u_1)\mathbb{R}$ with the planes $\mathbb{R}(0, -1, 1, 0)$ and $\mathbb{R}(0, 1, 1, 0)$. We get $\mathbf{a}_0(1 : 0) = -\mathbf{a}_1(1 : 0)$ and $\mathbf{a}_0(0 : 1) = \mathbf{a}_1(0 : 1)$. The result of the divisions $\tilde{\mathbf{a}}_0 = (\mathbf{a}_0 + \mathbf{a}_1)/u_1$ and $\tilde{\mathbf{a}}_1 = (\mathbf{a}_0 - \mathbf{a}_1)/u_0$ is

$$\begin{aligned}\tilde{\mathbf{a}}_0(u) &= (u_0^2(u_0 + u_1), 2u_0^2 u_1, 0, u_1^2(u_0 + u_1)), \\ \tilde{\mathbf{a}}_1(u) &= (u_0^2(u_1 - u_0), 0, -2u_0 u_1^2, -u_1^2(u_1 - u_0)).\end{aligned}$$

The curves $\tilde{\mathbf{a}}_0\mathbb{R}$ and $\tilde{\mathbf{a}}_1\mathbb{R}$ still have points in common, so we can repeat the procedure. We see that $\tilde{\mathbf{a}}_1 = (\tilde{\mathbf{a}}_0 + \tilde{\mathbf{a}}_1)/u_0 u_1$ is linear, and that $\tilde{\mathbf{a}}_0 = \tilde{\mathbf{a}}_0\mathbb{R}$ and $\tilde{\mathbf{a}}_1 = \tilde{\mathbf{a}}_1\mathbb{R}$ have no common values. Thus these two curves can be used as minimal directrices (see Fig. 5.11, left). \diamond

Example 5.2.5. We again consider the surface \mathcal{R} of Ex. 5.1.4 (see Fig. 5.6). The change of parameters $\cos u = (1 - t^2)/(1 + t^2)$ and $\sin u = 2t/(1 + t^2)$ yields the surface parametrization

$$\mathbf{x}(t, v)\mathbb{R} = (1 + t^2, dv(1 + t^2), 1 - t^2, 2vt)\mathbb{R}.$$

Thus \mathcal{R} may be parametrized as a rational tensor product Bézier surface of bidegree $(2, 1)$. The Plücker coordinates of the rulings $R(u)$ are

$$(\mathbf{r}, \bar{\mathbf{r}}) = (d(1+t^2)^2, 0, 2t(1+t^2), 2t(1-t^2), 0, -d(1-t^4)),$$

so \mathcal{R} is of degree four. The implicit inhomogeneous equation $x^2y^2 + d^2z^2 - x^2 = 0$ of its point set Φ is found by implicitization. Obviously the planes $x = c = \text{const}$ ($c \neq 0$) intersect Φ in ellipses $y^2 + (dz/c)^2 = 1$. For $c = \pm d$ this ellipse is a circle.

The intersection with both the yz -plane and the ideal plane yield quadratically parametrized line segments, so \mathcal{R} is an example of a quartic ruled surface with a one-parameter family of quadratic directrices (type (2,2), cf. Ex. 5.2.3). \diamond

Example 5.2.6. The ruled surface parametrized by $(\mathbf{r}, \bar{\mathbf{r}}) = (0, t+t^3, 0, t^2+t^4, 0, 0)$ has the directrices $a_0(t) = (1, 0, 0, t)\mathbb{R}$ and $a_1(t) = (0, 0, t+t^3, 0)\mathbb{R}$, both of which parametrize lines. The surface is therefore of degree four, and of type (1, 3). It is also easy to see that it is a polynomial cubic tensor product surface and has the equation $x_0^3x_2 = x_1(x_3x_0^2 + x_3^3)$. Fig C.4 shows an affine image of this surface together with the Lie quadric for $t = 0$. The intersection of the surface and its Lie quadric consists of the y -axis of the coordinate system (i.e., the ruling $R(0)$, with multiplicity three), the z -axis, and the horizontal line at infinity (with multiplicity four). \diamond

5.2.2 The Bézier Representation of Rational Ruled Surfaces

To write polynomials as linear combinations of Bernstein polynomials instead of monomials has many advantages. In geometric design the most famous reason is the numerical stability of the Bernstein basis in the interval $[0, 1]$ and the subdivision algorithm of P. de Casteljau. Another reason is shown by Th. 1.4.10 and Ex. 1.4.4: we are able to determine osculating subspaces at two parameter values by taking the projective span of control points.

In Sec. 1.4.1 we discussed how to write a polynomial $p(u)$ of degree $\leq d$ as a linear combination of Bernstein polynomials $B_i^d(u)$ of degree d . If we substitute $u = u_1/u_0$ in $B_i^d(u)$ of Equ. (1.88) and multiply by u_0^d , we get the *homogeneous Bernstein polynomials*

$$B_i^d(u_0, u_1) = \binom{d}{i} u_1^i (u_0 - u_1)^{d-i} \quad (5.15)$$

Remark 5.2.5. Equ. (1.108) defines a different kind of homogeneous Bernstein polynomial $B_{i,d-i}(u_0, u_1)$ whose arguments are barycentric coordinates with respect to the base points 0 and 1. The equations

$$B_i^d(u) = B_{i,d-i}(1-u, u) = B_i^d(1, u)$$

show how the three types of Bernstein polynomials are related to each other. \diamond

Obviously the curve defined by Equ. (5.11) can be written in the form

$$(\mathcal{R}\gamma)(u_0 : u_1) = (\mathbf{r}(u_0 : u_1), \bar{\mathbf{r}}(u_0 : u_1))\mathbb{R} = \left(\sum_{i=0}^d (\mathbf{b}_i, \bar{\mathbf{b}}_i) B_i^d(u_0, u_1) \right) \mathbb{R} \quad (5.16)$$

with $(u_0 : u_1) \in P^1$.

The vectors $(\mathbf{b}_i, \bar{\mathbf{b}}_i)$ are Bézier coefficient vectors. Th. 1.4.3 and Ex. 1.4.2 show how to compute them from the coefficient vectors $(\mathbf{c}_i, \bar{\mathbf{c}}_i)$ of Equ. (5.11).

Control Null Polarities

The $d + 1$ points $(\mathbf{b}_i, \bar{\mathbf{b}}_i)\mathbb{R}$ are called the projective *Bézier points* of the curve $\mathcal{R}\gamma$. As has been discussed in Sec. 1.4.1, the control *points* alone are not sufficient to determine the curve — we need the *vectors* $(\mathbf{b}_i, \bar{\mathbf{b}}_i)$, at least up to a common scalar factor.

One possibility to encode the vectors in a geometric way has been described at p. 113: We use the following *frame points*

$$(\mathbf{f}_i, \bar{\mathbf{f}}_i)\mathbb{R} = ((\mathbf{b}_i, \bar{\mathbf{b}}_i) + (\mathbf{b}_{i+1}, \bar{\mathbf{b}}_{i+1}))\mathbb{R}, \quad i = 0, \dots, d - 1.$$

The knowledge of the points $(\mathbf{b}_i, \bar{\mathbf{b}}_i)\mathbb{R}$, $(\mathbf{f}_i, \bar{\mathbf{f}}_i)\mathbb{R}$, and $(\mathbf{b}_{i+1}, \bar{\mathbf{b}}_{i+1})\mathbb{R}$ determines the vectors $(\mathbf{b}_i, \bar{\mathbf{b}}_i)$, $(\mathbf{f}_i, \bar{\mathbf{f}}_i)$, and $(\mathbf{b}_{i+1}, \bar{\mathbf{b}}_{i+1})$ up to a scalar factor.

Apart from the first and last point, the control points are in general not contained in the Klein quadric, even if the curve $\mathcal{R}\gamma$ is. They are Klein images of linear complexes, which define a null polarity. Thus we speak of *Bézier null polarities* and *frame null polarities*. Together they are the set of *control null polarities*. The first and last of them is singular, the others need not be.

Algorithms with Control Null Polarities

There are some geometric processing algorithms in which the line geometric representation and the associated control null polarities turn out to be very useful. A simple example is the following

Example 5.2.7. We want to compute the intersection of a rational ruled surface \mathcal{R} with a line L , where $L\gamma = (\mathbf{l}, \bar{\mathbf{l}})\mathbb{R}$. We use the intersection condition $\Omega((\mathbf{l}, \bar{\mathbf{l}}), (\mathbf{r}(u_0 : u_1), \bar{\mathbf{r}}(u_0 : u_1))) = 0$, which is a homogeneous polynomial equation of degree d in the homogeneous indeterminates $u_0 : u_1$. Its solutions give the generators of \mathcal{R} which intersect L .

Often algorithms for curve/surface intersections are based on an implicit equation for the surface. Obviously it is not necessary to implicitize the surface \mathcal{R} to compute its intersection with a line. For implicitization of rational ruled surfaces we refer to [179], Prop. 1.3.14, and Ex. 1.3.11. \diamond

Rational Ruled Surfaces Contained in Linear Manifolds of Lines

Another application of the line geometric control structure concerns ruled surfaces \mathcal{R} contained in a linear complex \mathcal{C} . Obviously this happens if and only if $\mathcal{R}\gamma$ is contained in $\mathcal{C}\gamma$. As $\mathcal{C}\gamma$ is a hyperplanar section of the Klein quadric, this is true if and only if the control points $(\mathbf{b}_i, \bar{\mathbf{b}}_i)\mathbb{R}$ are contained in this hyperplane. A slight generalization is the following

Lemma 5.2.14. *A rational ruled surface \mathcal{R} lies in a linear complex (in a linear congruence, respectively), if and only if its control null polarities $(\mathbf{b}_i, \bar{\mathbf{b}}_i)\mathbb{R}$ of Equ. (5.16) are contained in a four-dimensional (three-dimensional, respectively) projective subspace of P^5 .*

Proof. The case of a linear complex has been discussed above, and the case of a linear congruence is completely similar. \square

Remark 5.2.6. Of course, Lemma 5.2.14 is valid for ‘regulus, bundle, plane’ and ‘two-dimensional’ instead of ‘complex’ and ‘four-dimensional’ as well. But this case is not as interesting as the higher dimensional cases. \diamond

The following lemma is useful when studying rational curves of low degree:

Lemma 5.2.15. *An algebraic curve c in P^n of degree d is contained in a projective subspace of dimension d .*

Proof. Any hyperplane which contains more than $\deg(c)$ points of c , must contain c , by the properties of the degree of an algebraic variety. If there are $d+1$ projectively independent points P_0, \dots, P_d of c , then c is contained in all hyperplanes which contain P_0, \dots, P_d , i.e., c is contained in the d -space $P_0 \vee \dots \vee P_d$. \square

Example 5.2.8. Examples of algebraic ruled surfaces which are contained in a linear complex are ruled surfaces of degree four. Their Klein images are algebraic curves of degree four, which must lie in a subspace of dimension ≤ 4 , by Lemma 5.2.15. The same argument shows that ruled surfaces of degree three are contained in a linear congruence.

The Klein image of a rational ruled surface \mathcal{R} is a rational curve. If its Bézier control points span a projective subspace of dimension d with $d \leq 4$, then \mathcal{R} is contained in a linear complex. If $d \leq 3$, \mathcal{R} is contained in a linear congruence. This condition is always fulfilled, if there are at most $d+1$ control points. \diamond

Control Points of Planar Sections

Let us intersect the ruled surface \mathcal{R} of Equ. (5.16) with the plane $\mathbb{R}(v_0, \mathbf{v})$. Equ. (2.17) shows that a parametrization $(p_0(u_0 : u_1), \mathbf{p}(u_0 : u_1))$ of the intersection curve is given by

$$(p_0, \mathbf{p})\mathbb{R} = (\mathbf{v} \cdot \mathbf{r}, -v_0 \mathbf{r} + \mathbf{v} \times \bar{\mathbf{r}}) = \sum_{i=0}^d B_i^d \mathbf{d}_i, \quad (5.17)$$

with $\mathbf{d}_i = (\mathbf{v} \cdot \mathbf{b}_i, -v_0 \mathbf{b}_i + \mathbf{v} \times \bar{\mathbf{b}}_i).$

Equ. (3.4) shows that $\mathbf{d}_i\mathbb{R}$ is the null point of the plane $\mathbb{R}(v_0, \mathbf{v})$ with respect to the Bézier null polarity $(\mathbf{b}_i, \bar{\mathbf{b}}_i)$. By linearity, the points $(\mathbf{d}_i + \mathbf{d}_{i+1})\mathbb{R}$ are the null points of the same plane with respect to the frame null polarities $(\mathbf{b}_i + \mathbf{b}_{i+1}, \bar{\mathbf{b}}_i + \bar{\mathbf{b}}_{i+1})$. These facts are summarized in the following theorem:

Theorem 5.2.16. *An irreducible intersection of a rational ruled surface \mathcal{R} of degree d with a plane ε is a rational Bézier curve of degree d . Its Bézier points and frame points are exactly the null points of ε with respect to the Bézier and frame null polarities of \mathcal{R} .*

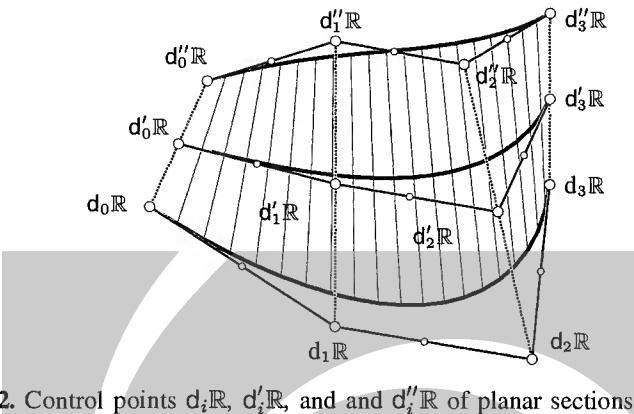


Fig. 5.12. Control points d_iR , d'_iR , and d''_iR of planar sections of a rational ruled surface.

Example 5.2.9. Consider a ruled surface \mathcal{R} and planar sections with planes $\varepsilon_1, \dots, \varepsilon_k$. Th. 5.2.16 shows how to compute the control structure of the Bézier curve $\mathcal{R} \cap \varepsilon_j$: The i -th control point is computed by applying the i -th control polarity to ε_j .

The mapping of the planes $\varepsilon_1, \dots, \varepsilon_k$ to the k different i -th control points is a duality (it is even a null polarity). For instance it maps the planes of a pencil to points of a pencil. This is illustrated in Fig. 5.12: We intersect a ruled surface which three parallel planes. The i -th control points are collinear. \diamond

Remark 5.2.7. The parametrization (5.17) is undefined for parameter values $(\alpha_0 : \alpha_1)$ where $d(\alpha_0 : \alpha_1) = 0$. In that case we may divide by the factor $(u_0 : u_1 - \alpha_0 : \alpha_1)$, which is the same as dividing by the factor $(\alpha_1 u_0 - \alpha_0 u_1)$, and obtain a parametrization of the intersection curve whose degree is lower than d .

The parametrization (5.17) is still useful if this does not occur within the parameter interval of interest. \diamond

Remark 5.2.8. Recall that the set of tangent planes of a surface which are incident with a fixed point Z is called the surface's tangent cone Δ with vertex Z . Assume that the skew ruled surface \mathcal{R} has a rational Plücker coordinate parametrization $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$ of degree d . By Equ. (2.7), its dual is again a rational ruled surface of degree d . As tangent cones (as sets of planes) are dual to planar sections, we can use the dual of Th. 5.2.16 to compute them:

According to Th. 5.2.16, Δ is rational of class d (i.e., a rational planar Bézier curve of dual projective space), if it is irreducible. Its Bézier planes and frame planes are exactly the null planes of Z with respect to the Bézier and frame null polarities of \mathcal{R} .

The intersection of Δ with a plane π is the contour outline of \mathcal{R} with respect to the central projection with center Z onto the plane π . To compute its dual control structure, we intersect the dual control structure of Δ with ε . A proof of this will be shown later (see Cor. 6.2.11). \diamond

Tensor Product Representation of Rational Ruled Surfaces

Planar sections of rational ruled surfaces can be used to convert the parametrization (5.16) of a rational ruled surface \mathcal{R} into a parametrization as a rational tensor product surface of degree $(n, 1)$. If $\mathbf{a}_0(u)\mathbb{R}$ and $\mathbf{a}_1(u)\mathbb{R}$ are two planar sections of \mathcal{R} (cf. Equ. (5.17)) with Bézier representation

$$\mathbf{a}_j(u_0 : u_1) = \sum_{i=0}^d B_i^d(u_0, u_1) \mathbf{b}_{i,j}, \quad j = 0, 1,$$

then the point set of \mathcal{R} may be parametrized by

$$\begin{aligned} \mathbf{x}(u_0 : u_1, v_0 : v_1) &= v_0 \mathbf{a}_0(u_0 : u_1) + v_1 \mathbf{a}_1(u_0 : u_1) \\ &= \sum_{i=0}^d \sum_{j=0,1} B_i^d(u_0, u_1) B_j^1(v_0, v_1) \mathbf{b}_{i,j}. \end{aligned} \quad (5.18)$$

Note that $B_0^1(v_0, v_1) = v_0 - v_1$ and $B_1^1(v_0, v_1) = v_1$. Conversion into an inhomogeneous parameter $u = u_1/u_0$ and $v = v_1/v_0$ yields

$$\mathbf{x}(u, v) = (1 - v) \mathbf{a}_0(u) + v \mathbf{a}_1(u) = \sum_{i=0}^d \sum_{j=0,1} B_i^d(u) B_j^1(v) \mathbf{b}_{i,j}. \quad (5.19)$$

Remark 5.2.9. The curves $\mathbf{a}_0\mathbb{R}$, $\mathbf{a}_1\mathbb{R}$ are contained in planes ε_0 and ε_1 , respectively. After performing a complex extension, $\varepsilon_0 \cap \varepsilon_1$ has generically d points in common with the ruled surface. Thus $\mathbf{a}_0\mathbb{R}$ and $\mathbf{a}_1\mathbb{R}$ have d points in common. We see that the proof of Th. 5.2.12 gives the right degree $\deg(\mathcal{R}) = d + d - d$ again. ◇

The rational tensor product surface

$$\mathbf{x}(u, v) = \sum_{i=0}^n \sum_{j=0,1} B_i^n(u) B_j^1(v) \mathbf{b}_{i,j}, \quad (5.20)$$

can be converted into a line coordinate representation of degree $2n$ by $(\mathbf{r}(u), \bar{\mathbf{r}}(u)) = \mathbf{x}(u, 0) \wedge \mathbf{x}(u, 1)$. More generally, if $\mathbf{a}_0(u)$ and $\mathbf{a}_1(u)$ are rational Bézier curves of degree m and n , respectively, with Bézier control vectors $\mathbf{b}_{i,0}$ and $\mathbf{b}_{i,1}$, respectively, the ruled surface $R(u) = \mathbf{a}_0(u)\mathbb{R} \vee \mathbf{a}_1(u)\mathbb{R}$ has the Bézier representation

$$(\mathbf{r}(u), \bar{\mathbf{r}}(u)) = \sum_{k=0}^{n+m} B_k^{n+m}(u) \binom{m+n}{k}^{-1} \sum_{i+j=k} \binom{m}{i} \binom{n}{j} \mathbf{b}_{i,0} \wedge \mathbf{b}_{j,1}. \quad (5.21)$$

Fig. 5.13 shows a polynomial TP Bézier surface of degree $(1, 3)$. A simple example of a rational tensor product surface is a ruled quadric, which has been shown in Ex. 1.4.10.

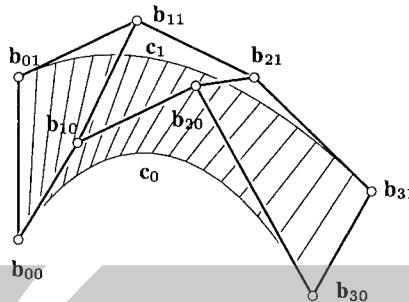


Fig. 5.13. Ruled polynomial tensor product Bézier surface.

Remark 5.2.10. Another way of controlling a rational ruled surface is closely related to its tensor product representation (5.20). By connecting associated Bézier points $b_{i,0}\mathbb{R}$, $b_{i,1}\mathbb{R}$ and also the corresponding frame points, we get a sequence of lines

$$L_i = b_{i,0}\mathbb{R} \vee b_{i,1}\mathbb{R}, \quad M_i = (b_{i,0} + b_{i+1,0})\mathbb{R} \vee (b_{i,1} + b_{i+1,1})\mathbb{R}.$$

The sequence L_0, M_0, \dots determines the ruled surface, but not the boundary curves which had been used for its definition: We pick two different points $\bar{b}_{0,0}\mathbb{R}, \bar{b}_{0,1}\mathbb{R} \in L_0$ and construct vectors $\bar{b}_{i,j}$ such that $\bar{b}_{i,j}\mathbb{R} \in L_i, (\bar{b}_{i,j} + \bar{b}_{i+1,j})\mathbb{R} \in M_i$.

If the lines L_i and M_i are skew, the construction is unique for all possible choices of $\bar{b}_{0,0}$ and $\bar{b}_{0,1}$, and the ruled surface defined by the new control vectors is the same as the original one. Except for the skewness condition, the sequences L_0, \dots, L_n and M_0, \dots, M_{n-1} can be chosen arbitrarily and define a rational ruled surface of degree $\leq 2n$. \diamond

5.2.3 Skew Cubic Surfaces

The simplest skew ruled algebraic surfaces after the reguli are the skew cubic surfaces. We will enumerate the possible projective cases and show affine and Euclidean special cases. We will frequently omit the attribute ‘ruled’ when speaking of ‘skew cubic’ surfaces, as ‘skew’ has no meaning for non-ruled surfaces anyway.

It is fortunate that all cubic skew ruled surfaces turn out to be rational, so we can apply the whole theory of rational surfaces, especially Th. 5.2.12. First we need a lemma concerning cubic curves in three-space:

Lemma 5.2.17. *A non-planar curve c of degree three in real or complex projective three-space is rational.*

Proof. (Sketch) Choose two points $A, B \in c$ and parametrize the pencil of planes incident with $A \vee B$ in the form $\varepsilon(u_0 : u_1) = \mathbb{R}(u_0 u_0 + u_1 u_1)$. Computing the intersection $\varepsilon(u_0 : u_1) \cap c$ leads to a polynomial equation of degree three, whose coefficients are polynomials in u_0 and u_1 . Two solutions (corresponding to points

A and B) are already known. The third solution can therefore be computed by polynomial division, the result of which is a rational parametrization of c . \square

Lemma 5.2.18. *A real or complex cubic skew surface \mathcal{R} is rational, if it is irreducible. Such a surface is contained in exactly one linear line congruence, which is either hyperbolic or parabolic.*

Proof. The Klein image $\mathcal{R}\gamma$ of the surface is an irreducible cubic curve. If $\mathcal{R}\gamma$ is planar and its carrier plane ε is not contained in the Klein quadric, then $\mathcal{R}\gamma$ would be a conic (and therefore not cubic). If $\varepsilon \subset M_2^4$, then ε is the Klein image of a bundle (then \mathcal{R} is a cone, therefore not skew) or of a field (then \mathcal{R} is planar, therefore not skew). This shows that $\mathcal{R}\gamma$ is not planar.

By Lemma 5.2.15, $\mathcal{R}\gamma$ spans a three-space $G^3 \subset P^5$, and by Lemma 5.2.17 it is rational. Because $\mathcal{R}\gamma$ is not planar, it is not contained in any other three-space, which shows that \mathcal{R} is contained in exactly one linear congruence \mathcal{N} with $\mathcal{N}\gamma = G^3 \cap M_2^4$.

A real rational cubic intersects all hyperplanes in at least one real point, as real polynomials of degree three have at least one real zero. Therefore $\mathcal{R}\gamma$ cannot be contained in an oval quadric. Th. 3.2.7 shows that \mathcal{N} must be hyperbolic or parabolic. \square

Directrices of Skew Cubic Surfaces

We extend the definition of projective isomorphism, or projective mapping, to the case of a mappings between lines and conics: If $a(u_0 : u_1) = (u_0 a_0 + u_1 a_1)\mathbb{R}$ is a parametrization of a line in homogeneous coordinates, and $b(u_0 : u_1) = (u_0^2 b_0 + u_0 u_1 b_1 + u_1^2 b_2)\mathbb{R}$ is a parametrization of a conic, then the mapping $a(u_0 : u_1) \mapsto b(u_0 : u_1)$ is said to be a projective isomorphism. It is left to the reader to show that this leads to a well defined class of projective mappings between lines, pencils, and conics.

Proposition 5.2.19. *All real or complex skew cubic surfaces \mathcal{R} possess a director line a_0 and a two-parameter linear family of quadratic directrices a_1 , one of which is possibly degenerate. There is a projective mapping $\kappa : a_0 \rightarrow a_1$ and \mathcal{R} consists of the lines $X \vee X\kappa$ for $X \in a_0$ (see Fig. 5.14, left).*

Proof. By Th. 5.2.12, \mathcal{R} has directrices of degrees one and two. By Prop. 5.2.13, the director line is unique and there is a linear two-parameter family of quadratic director curves, which are conics or quadratic parametrizations of lines.

If there are two such degenerate quadratic curves, then all generators intersect three lines, and \mathcal{R} is planar or a regulus, i.e., either not skew or not cubic.

The linear and an arbitrary non-degenerate quadratic directrix have parametrizations $(a'_0 u_0 + a''_0 u_1)\mathbb{R}$ and $(a'_1 u_0^2 + a''_1 u_0 u_1 + a'''_1 u_1^2)\mathbb{R}$, respectively. Both are one-to-one and onto, and the correspondence $a_0(u_0 : u_1) \leftrightarrow a_1(u_0 : u_1)$ is a projective isomorphism κ , by definition. The ruled surface \mathcal{R} obviously consists of the lines $X \vee X\kappa$, where X ranges in the director line. \square

Torsal Generator Lines of Skew Cubic Surfaces

A skew algebraic ruled surface can only have a finite number of torsal generators. We will show that cubic surfaces can have at most two.

Lemma 5.2.20. *We use the notation of Prop. 5.2.19. Consider the intersection point P of a_0 with the carrier plane $[a_1]$. The line $a_0(u) \vee a_1(u)$ is a torsal generator if and only if a_1 's tangent in $a_1(u)$ is incident with P . (cf. Fig. 5.14, left and center).*

Proof. If R is a torsal generator, the surface's tangent plane is the same for all points of R . Clearly the line a_0 is contained in all tangent planes of points of a_0 . This implies that the tangent plane τ_R which belongs to R must contain a_0 and the conic's tangent in $R \cap a_1$. If P is inside the conic a_1 , none of a_1 's tangents are incident with P , if $P \in a_1$, there is exactly one, and if P is outside, there are two. This shows the statement about the number of torsal generators. In the case of one torsal generator, we know that $P\kappa^{-1} \vee P$ is a generator and both $P, P\kappa^{-1} \in a_0$, so a_0 is this torsal generator. \square

This proof immediately implies

Corollary 5.2.21. *The number of torsal generators of a skew cubic surface equals zero, one, or two.*

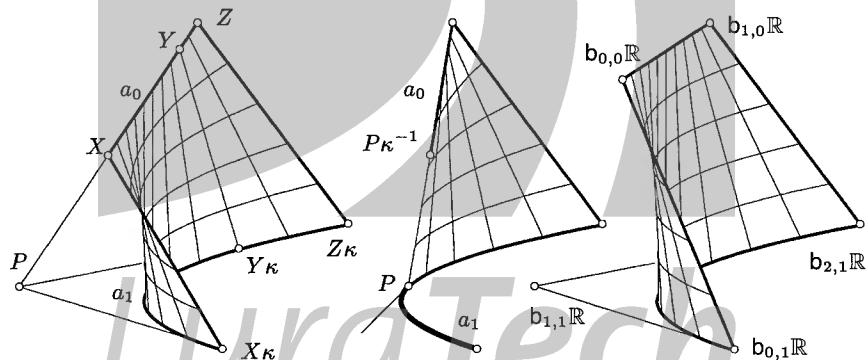


Fig. 5.14. Left: Cubic skew surface with linear directrix a_0 , quadratic directrix a_1 , and projective mapping $X \mapsto X\kappa$. The generators $X \vee X\kappa$ and $Z \vee Z\kappa$ are torsal. Center: Cubic skew surface with only one torsal generator. Right: Bézier points of a skew cubic surface.

Tangential Planar Sections of Skew Cubic Surfaces

This paragraph is based on the elementary fact that a bivariate polynomial p (or a trivariate homogeneous polynomial) of degree three, which is divisible by a polynomial p_0 of degree d , is a product $p = p_0 p_1$, with a polynomial p_1 of degree $3 - d$.

A tangential planar section of a skew cubic surface is an algebraic curve of degree three. As all tangent planes contain a generator, this curve contains a line. The remaining quadratic component may be a conic, or the union of two lines, or one ‘double’ line.

Conversely, consider a conic c contained in the point set Φ of a skew cubic surface. We intersect its carrier plane ε with Φ . This gives a cubic curve, which contains a conic, and therefore also a line l . We use the complex extensions $c_{\mathbb{C}}$ and $l_{\mathbb{C}}$ of the curves c and l to show that ε is a tangent plane: If $c_{\mathbb{C}}$ and $l_{\mathbb{C}}$ intersect in two points (this means that $\#(l \cap c)$ equals 0 or 2), then Φ ’s tangent plane in the points of $l_{\mathbb{C}} \cap c_{\mathbb{C}}$ is spanned by tangents to $c_{\mathbb{C}}$ and $l_{\mathbb{C}}$ there, and consequently coincides with ε . If c and l touch each other in a point p , then Φ has a surface singularity in p , and all planes incident with l may be considered tangent to Φ in p .

The fact that the conics contained in skew cubic surfaces are precisely those in tangential intersections is also verified by discussion of the three different skew cubic surfaces which exist in projective three-space (see below).

Models of the Projective Plane

We consider the linear family of quadratic directrices which exists according to Prop. 5.2.13, and which is parametrized by Equ. (5.14). We use a homogeneous parameter $u = u_1/u_0$ and write both the linear directrix a_0 and the quadratic directrix a_1 in homogeneous Bézier form (cf. Fig. 5.14, right). Then a quadratic directrix is parametrized by

$$\begin{aligned} x_{v_0, v_1}(u_0 : u_1) \mathbb{R} &= (v_0(u_0, u_1)((u_0 - u_1)b_{0,0} + u_1 b_{1,0}) \\ &\quad + v_1((u_0 - u_1)^2 b_{0,1} + 2u_1(u_0 - u_1)b_{1,1} + u_1^2 b_{2,1})) \mathbb{R}, \end{aligned} \quad (5.22)$$

where v_1 is constant and v_0 is a linear function of u_0 and u_1 .

Equ. (5.22) parametrizes a quadratic directrix, which is a conic except for one particular choice of v_0 and v_1 , where it parametrizes a line. A different interpretation of the same equation yields a parametrization of the point set of the ruled surface, whose domain is the projective plane:

$$\begin{aligned} x(u_0 : u_1 : v_0) \mathbb{R} &= (v_0((u_0 - u_1)b_{0,0} + u_1 b_{1,0}) \\ &\quad + (u_0 - u_1)^2 b_{0,1} + 2u_1(u_0 - u_1)b_{1,1} + u_1^2 b_{2,1}) \mathbb{R}. \end{aligned} \quad (5.23)$$

A line in P^2 consists of all points $(u_0 : u_1 : v_0)$ which satisfy a linear relation of the form $\alpha u_0 + \beta u_1 + \gamma v_0 = 0$. The image of such a line is a curve of degree less or equal two.

We see that (5.23) is undefined for $(u_0 : u_1 : v_0) = (0 : 0 : 1)$. The straight lines of P^2 incident with this point are mapped to the generator lines contained in Φ .

Proposition 5.2.22. *Any cubic skew ruled surface \mathcal{R} is a model of the projective plane in the sense that there is a mapping of P^2 to the point set of \mathcal{R} such that straight lines are mapped to lines or conics.*

Remark 5.2.11. A surface $\mathbf{x}(u_0 : u_1 : u_2)$ whose domain is P^2 and whose parametrization is quadratic in the homogeneous indeterminates $(u_0, u_1, u_2) = \mathbf{u}$ is called a Steiner's surface Φ . A famous example is Steiner's 'Roman surface', which is parametrized by $\mathbf{x}(u_0 : u_1 : u_2) = (u_0^2 + u_1^2 + u_2^2, 2u_0u_2, 2u_0u_1, 2u_1u_2)$ and has the implicit equation $x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2 = 2x_0x_1x_2x_3$ (see Fig. C.1).

It follows from Prop. 1.3.16 that for the complex number field, Φ is algebraic. To estimate its degree, we count the number of intersection points with a line $\mathbf{h}_1 \cdot \mathbf{x} = \mathbf{h}_2 \cdot \mathbf{x} = 0$: The condition $\mathbf{h}_i \cdot \mathbf{x}(u_0 : u_1 : u_2) = 0$ is quadratic, so there are, in the generic case, four intersection points. This shows $\deg(\Phi) \leq 4$.

It is not difficult to show that $\deg(\Phi) = 3$ if and only if the parametrization has a 'base point', i.e., $\mathbf{x}(\mathbf{u}_0) = \mathbf{o}$ for at least one choice \mathbf{u}_0 of \mathbf{u} . In that case the curve $\mathbf{x}(\mathbf{u}_0 + \lambda\mathbf{u}')$ is a line for all \mathbf{u}' . Thus a cubic Steiner's surface is a ruled surface. On the other hand Equ. (5.23) shows that a skew cubic surface is a Steiner's surface. \diamond

Skew Cubic Surfaces Contained in a Hyperbolic Linear Congruence

Consider a skew cubic surface \mathcal{R} with point set Φ . By Lemma 5.2.18, \mathcal{R} is contained in exactly one linear congruence \mathcal{N} . We assume that \mathcal{N} is hyperbolic. Its focal lines are denoted by L_0 and L_1 . All generators intersect L_0, L_1 . $L_i \cap \Phi$ does not consist of discrete points (in that case \mathcal{R} would consist of planar parts), and so both L_0, L_1 are contained in Φ .

Lemma 5.2.23. *One of L_0, L_1 is a linear directrix, and the other carries a degenerate quadratic directrix, which has 0 or 2 singular points. The torsal generators are incident with these singular points.*

Proof. We choose planes ε_i which contain L_i . The intersection of the rulings with ε_i are quadratic or linear curves, if \mathcal{R} is parametrized according to Th. 5.2.12. This shows that one of L_0, L_1 is the director line $a_0(t)$ of \mathcal{R} , and the other one carries the degenerate quadratic director curve $a_1(t)$. All generators $a_0(t) \vee a_1(t)$ are regular, because a_0 is regular, and clearly they are torsal if $a_1(t)$ is a singular point.

The curve $a_1(t)$, which parametrizes part of a line can always be seen as projection of a regular quadratic parametrization of a conic c onto a line (see Fig. 5.15).

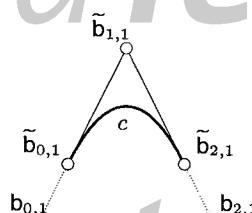


Fig. 5.15. Projecting a Bézier curve onto a singular one.

If the projection center is inside c , there are two conjugate complex singular points. If it is outside there are two real ones. Without loss of generality we can assume that the parameter values 0 and 1 are precisely those singularities. Then we have the situation depicted in Fig. 5.15. Th. 1.4.10 shows that the vector $b_{1,1}$ must be zero, which is also clear from the projection.

The lines L_0, L_1 are skew, which shows that all generators corresponding to regular points of a_1 are not torsal. \square

A skew cubic surface contained in \mathcal{N} is therefore obtained by parametrizing one focal line linearly and the other one as a degenerate quadratic curve, and then blending these two curves in a linear way (cf. Equ. (5.23)).

The singular quadratic directrix is parametrized by

$$a_1(u_0 : u_1) = ((u_0 - u_1)^2 b_{0,1} + u_1^2 b_{2,1})\mathbb{R}. \quad (5.24)$$

The two vectors $b_{0,1}, b_{2,1}$ can be recovered from the points $b_{0,1}\mathbb{R}, b_{2,1}\mathbb{R}, (b_{0,1} + b_{2,1})\mathbb{R}$, up to a scalar factor, so the curve a_1 is determined by three points.

If there are no real singularities, this is also true, but with complex points $b_{0,1}$ and $b_{2,1}$.

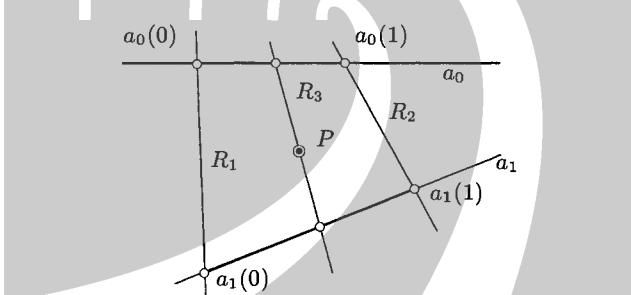


Fig. 5.16. Torsal rulings R_1, R_2 of a skew cubic surface \mathcal{R} contained in a hyperbolic linear congruence with focal lines F_1, F_2 (see proof of Prop. 5.2.24).

Proposition 5.2.24. *A skew cubic surface contained in a hyperbolic linear congruence \mathcal{N} which has two torsal rulings R_1, R_2 , is determined by R_1, R_2 , a further point P of the surface, and by the knowledge which of the two focal lines L_1, L_2 of \mathcal{N} is the director line a_0 . All such surfaces are projectively equivalent.*

Proof. The generators which belong to the singularities of the degenerate quadratic directrix a_1 are torsal (see the proof of Lemma 5.2.23 or Lemma 5.2.20). We assume that a_0 parametrizes L_1 , and that a_1 is parametrized as in Equ. (5.24). Then $a_0(0) = R_1 \cap L_1, a_0(1) = R_2 \cap L_1, a_1(0) = R_1 \cap L_2, a_1(1) = R_2 \cap L_2$ (see Fig. 5.16). A third generator line is the unique line R_3 incident with P, L_1, L_2 . We choose an admissible parametrization of a_0 and have $L_1 \cap R_3 = a_0(u_0 : u_1)$ with $(u_0 : u_1) \in P^1$. Then we know $a_1(u_0 : u_1) = L_2 \cap R_3$ and a_1 is uniquely determined.

A projective mapping is uniquely determined by the image of the five points $a_0(0)$, $a_0(1)$, $a_1(0)$, $a_1(1)$, P , which are in general position. This shows that all such surfaces are projectively equivalent. \square

Proposition 5.2.25. *All skew cubic surfaces without real torsal generators are projectively equivalent.*

Proof. The proof is the same as the proof of Prop. 5.2.24, but with complex singularities instead of real ones. \square

The Plücker Conoid

We want to derive projective normal forms for skew cubic surfaces. First we consider surfaces contained in hyperbolic linear congruences.

Proposition 5.2.26. *The surfaces of Prop. 5.2.24 are projectively equivalent to the surface with inhomogeneous equation $z(x^2 + y^2) - 2xy = 0$.*

Such a surface is called a *Plücker conoid* or a *cylindroid* (see Fig. 3.10 and Fig. 5.17).

Proof. We choose \mathcal{N} such that its focal lines are the vertical z -axis of a Cartesian coordinate system and the horizontal line at infinity. We let

$$\begin{aligned} a_0(u_0 : u_1) &= (0, u_0, u_1, 0)\mathbb{R} \quad ((u_0 : u_1) \in P^1) \\ &= (0, \cos \phi, \sin \phi, 0)\mathbb{R}, \quad (\tan \phi = u_1/u_0, \phi \in [0, \pi)). \end{aligned} \quad (5.25)$$

The singularities of a_1 are chosen at $z = \pm 1$, and so

$$a_1(u_0 : u_1) = (u_0^2 + u_1^2, 0, 0, 2u_0u_1)\mathbb{R} = (1, 0, 0, \sin 2\phi)\mathbb{R}. \quad (5.26)$$

We have used the angle parameter ϕ to show that the surface discussed here is the same which has already been discussed at p. 180. The line $a_0 \vee a_1$ has Plücker coordinates

$$(\mathbf{r}, \bar{\mathbf{r}})(u_0 : u_1) = (u_0(u_0^2 + u_1^2), u_1(u_0^2 + u_1^2), 0, -2u_0u_1^2, 2u_0^2u_1, 0). \quad (5.27)$$

We let $u = u_1 : u_0$ and parametrize the ruling $a_0(u) \vee a_1(u)$ in the form $v(0, 1, u, 0) + (1 + u^2, 0, 0, 2u)$. This yields the inhomogeneous parametrization

$$x(u, v) = (1 + u^2, v, uv, 2u)\mathbb{R} \quad (5.28)$$

of the surface. Implicitization shows that the surface is defined by $z(x^2 + y^2) = 2xy$. \square

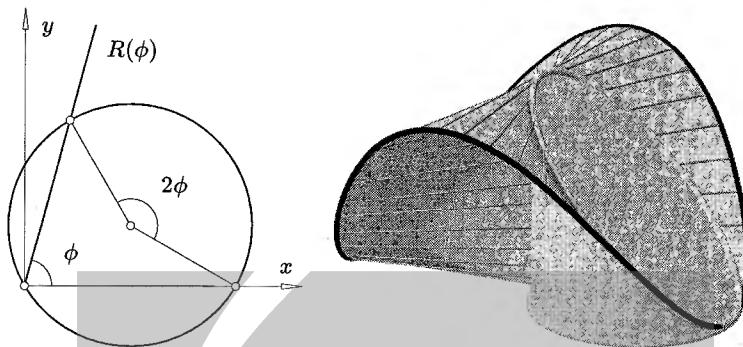


Fig. 5.17. Intersection of a right circular cylinder with the Plücker conoid (cf. Remark 5.2.12). Left: top view. Right: perspective view.

Remark 5.2.12. Equations (5.25) and (5.26) show that the Plücker conoid consists of ‘horizontal’ generators $R(\phi)$ parallel to vectors $(\cos \phi, \sin \phi, 0)$, whose z -coordinate varies with $\sin 2\phi$. We intersect $R(\phi)$ with the right circular cylinder

$$(x - r \cos \phi_0)^2 + (y - r \sin \phi_0)^2 = r^2,$$

which contains the z -axis and obtain the point $(0, 0, \sin 2\phi)$ together with another point $(x(\phi), y(\phi), \sin 2\phi)$. Its ‘top view’ $(x(\phi), y(\phi))$ rotates about the center with coordinates $(r \cos \phi_0, r \sin \phi_0)$ with angular velocity 2 (this follows from the theorem of the peripheral angle, cf. Fig. 5.17). Hence, the intersection curve is planar, and therefore an ellipse.

This shows that the Plücker conoid carries a two-parameter family of ellipses whose top views are circles. It is left as an exercise for the reader to verify that this family coincides with the two-parameter family of quadratic directrices which exist by Prop. 5.2.13. \diamond

Corollary 5.2.27. Consider a non-circular ellipse c contained in a right circular cylinder. Choose a generator a of this cylinder. Then the family of lines orthogonal to a which intersect c is a Plücker conoid.

Proof. This follows immediately from the discussion in Remark 5.2.12. \square

The Zindler Conoid

By Prop. 5.2.25, all skew cubic surfaces which do not have real torsal generators are projectively equivalent. It is therefore sufficient to describe *one* such surface. An example is shown by Fig. 5.18. We use the following quadratic director curve:

$$a_1(u_0 : u_1) = (u_0^2 - u_1^2, 0, 0, 2u_0u_1)\mathbb{R} = (1, 0, 0, \tan 2\phi)\mathbb{R}.$$

and the linear director curve of Equ. (5.25). The family $R(u)$ of generators then has Plücker coordinates

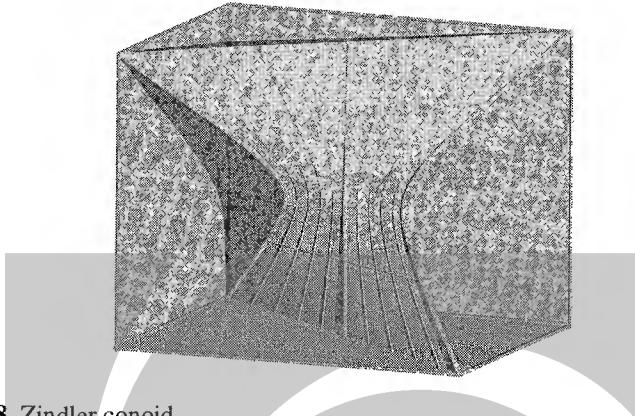


Fig. 5.18. Zindler conoid.

$$(\mathbf{r}, \bar{\mathbf{r}})(u_0 : u_1) = (u_0(u_0^2 - u_1^2), u_1(u_0^2 - u_1^2), 0, -2u_0u_1^2, 2u_0^2u_1, 0), \quad (5.29)$$

and the point set of this surface has the parametrization in inhomogeneous parameters

$$x(u, v) = (1 - u^2, v, uv, 2u)\mathbb{R}. \quad (5.30)$$

Implicitization of its affine part gives the equation

$$z(x^2 - y^2) - 2xy = 0. \quad (5.31)$$

This surface is known as the *Zindler conoid* (see Fig. 5.18).

Remark 5.2.13. The Zindler conoid has many properties which are similar to those of the Plücker conoid. This comes from the fact that the complex extensions of the two surfaces are the same (up to a coordinate transform).

An example of such a property is: The director conics of the Zindler conoid are hyperbolae, whose projections onto the xy plane have the equation $x^2 - y^2 + ax + by = 0$, therefore contain the origin and have asymptotes parallel to the lines $y = \pm x$. \diamond

The Cayley Surface

We have classified the skew cubic surfaces which are contained in hyperbolic linear congruences. We will come to the case of a parabolic congruence now. By Lemma 5.2.20, such a surface has a director line a_0 , a director conic a_1 which intersects a_0 , and the rulings $R(u)$ are of the form $R(u) = a_0(u) \vee a_1(u)$, where always $a_0(u) \neq a_1(u)$. The correspondence $\kappa : a_0(u) \longleftrightarrow a_1(u)$ is a projective mapping, as has been discussed in Prop. 5.2.19. Such a surface is called a *Cayley surface* (see Fig. 5.19).

We use the notation of Fig. 5.14, center. Assume that $P = a_0(u') = a_1(u'')$. Then $P\kappa^{-1} = a_0(u'')$ and the generator $R(u'') = a_0(u'') \vee a_1(u'')$ coincides with

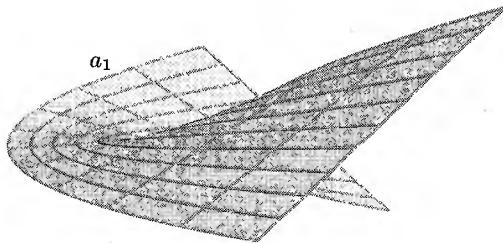


Fig. 5.19. Cayley surface with linear directrix a_0 at infinity and quadratic directrix a_1 .

the line a_0 . It is the only torsal generator. Assume further that u''' is different from u' , u'' . If we know the three points $a_0(u')$, $a_0(u'')$, $a_0(u''')$ and the corresponding points $a_1(u')$, $a_1(u'')$, $a_1(u''')$, the projective mapping $a_0 \rightarrow a_1$ is determined uniquely. It is left as an exercise for the reader to verify that all such input configurations (line, conic, and three pairs of points) are projectively equivalent. This shows that all Cayley surfaces are projectively equivalent. It is therefore sufficient to consider one of them: Let

$$a_0(u_0 : u_1) = (u_1, 0, 0, u_0)\mathbb{R}, \quad a_1(u_0 : u_1) = (u_0^2, u_0 u_1, u_1^2, 0)\mathbb{R}.$$

The generators $R(u_0 : u_1)$ have the Plücker coordinates

$$(\mathbf{r}, \bar{\mathbf{r}})(u_0 : u_1) = (u_0 u_1^2, u_1^3, -u_0^3, -u_0 u_1^2, u_0^2 u_1, 0). \quad (5.32)$$

A parametrization of the point set of the Cayley surface and an implicit inhomogeneous equation are given by

$$\mathbf{x}(u, v) = (v + u^2, u, 1, uv), \quad x^3 + zy^2 - xy = 0. \quad (5.33)$$

5.3 Euclidean Geometry of Ruled Surfaces

Euclidean Differential Geometry of Surfaces

Here we sum up very briefly some elementary facts concerning the Euclidean differential geometry of surfaces. Some aspects of the Euclidean differential geometry of curves (the Frenet equations) have been mentioned in Ex. 1.2.3.

We consider a regular surface $\mathbf{x}(u^1, u^2)$, where $\mathbf{x} : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$. The partial derivatives $\mathbf{x}_{,1}$, $\mathbf{x}_{,2}$ span the *tangent space*. The *unit normal vector field* \mathbf{n} is found by normalizing $\mathbf{x}_{,1} \times \mathbf{x}_{,2}$. We use the *Einstein convention* on summation: If an index appears both as a subscript and as a superscript in a product, this means summation for $i = 1, 2$. The real numbers

$$g_{ij} = \mathbf{x}_{,i} \cdot \mathbf{x}_{,j} \quad (5.34)$$

are the coefficients of the Euclidean scalar product (the *first or metric fundamental form*) with respect to the basis $\mathbf{x}_{,1}, \mathbf{x}_{,2}$: If $\mathbf{v} = v^i \mathbf{x}_{,i}$ and $\mathbf{w} = w^j \mathbf{x}_{,j}$, then $\mathbf{v} \cdot \mathbf{w} = v^i w^j g_{ij}$.

A curve $(u^1(t), u^2(t)) = u(t)$ in \mathbf{x} 's parameter domain defines a *surface curve* $\mathbf{c}(t) = \mathbf{x}(u(t))$. Since $\dot{\mathbf{c}} = \dot{u}^i \mathbf{x}_{,i}$, we have $\|\dot{\mathbf{c}}\|^2 = \dot{u}^i \dot{u}^j g_{ij}$. Two surfaces are *isometric*, if there exists a bijective *isometric* mapping (an *isometry*) of the first onto the second such that arc lengths of surface curves are preserved. If surfaces \mathbf{x} and $\bar{\mathbf{x}}$ happen to possess equal coefficient functions $g_{jk}(u^1, u^2)$ and $\bar{g}_{jk}(u^1, u^2)$, obviously the mapping $\mathbf{x}(u^1, u^2) \mapsto \bar{\mathbf{x}}(u^1, u^2)$ is locally an isometry.

As to the curvature of a surface curve $\mathbf{c}(t) = \mathbf{x}(u(t))$, we let $\|\ddot{\mathbf{c}}(t)\| = v(t)$, $\mathbf{e}_1(t) = \dot{\mathbf{c}}(t)/v(t)$, and $\mathbf{m}(t) = \mathbf{e}_1(t) \times \mathbf{n}(u(t))$. We define κ_g (the *geodesic curvature*) and κ_n (the *normal curvature*) of \mathbf{c} by

$$\dot{\mathbf{e}}_1 = v(\kappa_n \mathbf{n} + \kappa_g \mathbf{m}), \quad \ddot{\mathbf{c}} = \dot{v} \mathbf{e}_1 + v^2 (\kappa_n \mathbf{n} + \kappa_g \mathbf{m}). \quad (5.35)$$

Obviously then $\kappa_g = \det(\mathbf{e}_1, \mathbf{n}, \ddot{\mathbf{c}})/v^2$, and $\kappa^2 = \kappa_n^2 + \kappa_g^2$. We define the coefficients

$$h_{jk} = -\mathbf{x}_{,j} \cdot \mathbf{n}_{,k} = \mathbf{n} \cdot \mathbf{x}_{,jk}. \quad (5.36)$$

If \mathbf{v} and \mathbf{w} are as above, the *second fundamental form* $h(\mathbf{v}, \mathbf{w}) = v^i w^j h_{jk}$. The normal curvature of \mathbf{c} is computed via $h(\dot{\mathbf{c}}, \dot{\mathbf{c}})/(\dot{\mathbf{c}} \cdot \dot{\mathbf{c}}) = (v^i v^j h_{jk})/(v^i v^j g_{jk})$. This shows that κ_n depends only on the *first derivative* of \mathbf{c} . If two surfaces touch each other, then they are even in second order contact if and only if all tangent vectors in the contact point have the same normal curvature. Further, $h(\mathbf{v}, \mathbf{w}) = 0$ characterizes *conjugate tangents*.

In all surface points there are two orthogonal tangent vectors $\mathbf{v}_i = v_i^j \mathbf{x}_{,j}$ ($i = 1, 2$), which indicate the *principal directions*, if uniquely defined, and which have extremal normal curvatures $\kappa_i = \kappa_n(\mathbf{v}_i)$. These are called the *principal curvatures*. The column vectors (v_1^1, v_2^1) are eigenvectors of the matrix $(g_{jk})^{-1} \cdot (h_{jk})$ with eigenvalues κ_i . Surface curves tangent to either \mathbf{v}_1 or \mathbf{v}_2 in each point are called *principal curvature lines*. A surface point is called an *umbilic*, if $\kappa_1 = \kappa_2$. It is possible to show that away from umbilic points there is a local parametrization of the surface whose parameter lines are principal curvature lines.

Two equivalent definitions of the *Gaussian curvature* are

$$K = \kappa_1 \kappa_2 = \det(h_{jk}) / \det(g_{jk}). \quad (5.37)$$

It turns out that a surface point is *elliptic*, *parabolic/flat*, or *hyperbolic*, if $K > 0$, $K = 0$, or $K < 0$, respectively. A surface point is flat if and only if $\kappa_1 = \kappa_2 = 0$.

The Gaussian curvature and the geodesic curvature of curves are *intrinsic properties*, i.e., invariant with respect to isometric mappings. Likewise intrinsic is the property of being a *geodesic parallel vector field* along a surface curve: Assume that $\mathbf{v}(t)$ is a vector attached to $\mathbf{c}(t) = \mathbf{x}(u(t))$. It is a *parallel field* if $\dot{\mathbf{v}}$ is orthogonal to the surface in the point $\mathbf{x}(u(t))$. For given $\mathbf{v}(t_0)$, there is a unique parallel field $\mathbf{v}(t)$ along \mathbf{c} which contains $\mathbf{v}(t_0)$. For all t_1 , the mapping $\mathbf{v}(t_0) \mapsto \mathbf{v}(t_1)$ is called *parallel transport* and is a Euclidean displacement of the surface's tangent space in $\mathbf{x}(u(t_0))$ onto the tangent space on $\mathbf{x}(u(t_1))$.

If the unit tangent vector field $\mathbf{e}_1(t)$ of a surface curve is a parallel field, or equivalently, $\kappa_g = 0$, then \mathbf{c} is called a *geodesic line* or *geodesic*. Further equivalent characterizations of geodesics are that the osculating plane contains the surface normal, or that they are locally the shortest paths between surface points.

5.3.1 First Order Properties

We consider a ruled surface $\mathcal{R}(u)$ in real Euclidean 3-space E^3 . We use a Cartesian coordinate system and assume that $\mathcal{R}(u)$ is defined by two directrices a_0 and a_1 , which are, in Cartesian coordinates, parametrized by $\mathbf{a}_0(u)$, $\mathbf{a}_1(u)$, with $u \in I \subset \mathbb{R}$. Then the point set of the ruled surface \mathcal{R} is parametrized by

$$\mathbf{x}(u, v) = (1 - v)\mathbf{a}_0(u) + v\mathbf{a}_1(u), \quad u \in I, v \in \mathbb{R}. \quad (5.38)$$

The vectors $\mathbf{r}(u) = \mathbf{a}_1(u) - \mathbf{a}_0(u)$ parallel to the generator $R(u)$ define a *cone* whose vertex is the origin \mathbf{o} . It is parametrized as a surface by $\mathbf{c}(u, v) = v\mathbf{r}(u)$ and is called *director cone* of \mathcal{R} .

If we let $\mathbf{a} = \mathbf{a}_0$, the parametrization of Equ. (5.38) is equivalent to

$$\mathbf{x}(u, v) = \mathbf{a}(u) + v\mathbf{r}(u), \quad u \in I, v \in \mathbb{R}. \quad (5.39)$$

Remark 5.3.1. After embedding E^3 into projective space and adding to the point set of \mathcal{R} the ideal points of the generators, we can intersect \mathcal{R} with the ideal plane, which gives the curve $a_\infty(u) = (0, \mathbf{r}(u))\mathbb{R}$. Equ. (5.39) then appears as a special case of Equ. (5.2) with $\lambda_0 = 1$, $\lambda_1 = v$: The ruled surface \mathcal{R} is defined by the directrices a_0 and a_∞ . \diamond

Equ. (5.39) defines a ruled surface only if $\mathbf{r}(u) \neq \mathbf{o}$ for all u . The surface's tangent plane is spanned by the partial derivatives $\mathbf{x}_u = \dot{\mathbf{a}}(u) + v\dot{\mathbf{r}}(u)$ and $\mathbf{x}_v = \mathbf{r}$. Therefore a surface normal vector is given by

$$\mathbf{n}(u, v) = \mathbf{x}_u \times \mathbf{x}_v = \dot{\mathbf{a}}(u) \times \mathbf{r}(u) + v(\dot{\mathbf{r}}(u) \times \mathbf{r}(u)). \quad (5.40)$$

The surface is regular at $\mathbf{x}(u, v)$, if $\mathbf{n}(u, v) \neq \mathbf{o}$.

The Distribution of Surface Normals

A regular generator $R(u)$ has at most one singular surface point. If $R(u)$ is torsal, the tangent plane is the same for all of its regular points. We want to determine the set of surface normals in the points of a *non-torsal generator*:

Lemma 5.3.1. *The set of surface normals in the points of a non-torsal generator of a ruled surface is (the affine part of) a regulus belonging to a hyperbolic paraboloid.*

Proof. Equ. (5.5) shows that the regular generator $R(u)$ of the surface defined by Equ. (5.39) is torsal if and only if

$$\det(\dot{\mathbf{a}}(u), \mathbf{r}(u), \dot{\mathbf{r}}(u)) = 0. \quad (5.41)$$

We let $\mathbf{n}_0 = \dot{\mathbf{a}} \times \mathbf{r}$ and $\mathbf{n}_1 = \dot{\mathbf{r}} \times \mathbf{r}$. Equ. (5.41) implies that for a non-torsal generator, $\mathbf{n}_0, \mathbf{n}_1$ are linearly independent. By (5.40), the set of surface normals in the points $\mathbf{a}(u) + v\mathbf{r}(u)$, u fixed, is a ruled surface whose point set is parametrized by

$$\mathbf{y}(v, w) = \mathbf{a}(u_0) + v\mathbf{r}(u_0) + w(\mathbf{n}_0 + v\mathbf{n}_1).$$

This is an affine bilinear parametrization which is not planar (otherwise $\mathbf{n}_0, \mathbf{n}_1$ would be linearly dependent), and so the proof is complete (cf. Ex. 1.4.7). \square

The Striction Point

By Th. 5.1.4, the set of surface tangents in the points of a non-torsal generator $R(u)$ of a ruled surface \mathcal{R} in projective space is a parabolic linear congruence. The set of surface tangents in Euclidean three-space is this congruence minus the pencil of surface tangents in the ideal point of $R(u)$.

The tangent plane in the ruling's point at infinity is spanned by the three points $a_0(u) = (1, \mathbf{a}(u))\mathbb{R}$, $a_\infty(u) = (0, \mathbf{r}(u))\mathbb{R}$, $a_\infty^1(u) = (0, \dot{\mathbf{r}}(u))\mathbb{R}$. From the Euclidean viewpoint, it is spanned by the point $\mathbf{a}(u)$, and the two vectors $\mathbf{r}(u)$, $\dot{\mathbf{r}}(u)$. It is called *asymptotic plane*. Recall that the director cone of the ruled surface is parametrized by $\mathbf{x}(u, v) = v \cdot \mathbf{r}(u)$. The computation of a_∞^1 immediatly shows the following

Lemma 5.3.2. *The director cone's tangent plane and the asymptotic plane of \mathcal{R} are parallel in corresponding rulings.*

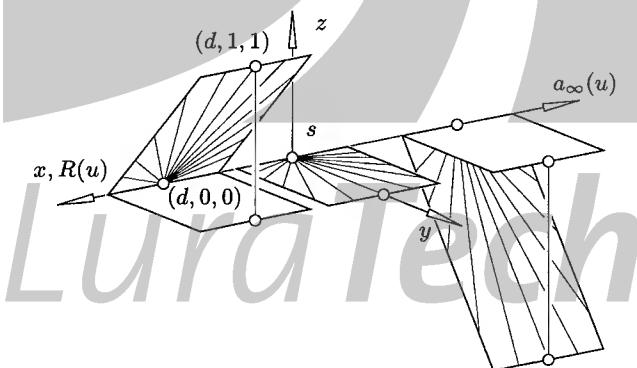


Fig. 5.20. Distribution of tangent planes along a generator $R(u)$. The origin, x -axis, y -axis, and z -axis of the coordinate system are striction point, generator line, central tangent, and central normal. The distribution parameter $\delta = -d$ is negative.

The plane which contains the generator $R(u)$ and is orthogonal to the asymptotic plane is called the *central plane*. Its normal vector is given by $\mathbf{r} \times (\dot{\mathbf{r}} \times \mathbf{r})$. The point of contact of the central plane is the *central point* or *striction point* of the generator.

The surface tangent at the striction point which is orthogonal to the ruling is called the *central tangent* (see Fig. 5.20). The *central normal* is the surface normal in the striction point.

We want to compute the striction point of the generator $R(u)$. This means that for given u we want to find the value $v_s(u)$ such that $\mathbf{x}(u, v_s(u))$ is the striction point. This is done as follows: The surface normal $\mathbf{n}(u, v)$ at the striction point is, by definition, orthogonal to $\dot{\mathbf{r}}(u) \times \mathbf{r}(u)$. We use (5.40) to write this condition in the form

$$(\dot{\mathbf{a}} \times \mathbf{r} + v_s(\dot{\mathbf{r}} \times \mathbf{r})) \cdot (\dot{\mathbf{r}} \times \mathbf{r}) = 0.$$

Therefore, the value v_s is given by

$$v_s = -\frac{(\dot{\mathbf{a}} \times \mathbf{r}) \cdot (\dot{\mathbf{r}} \times \mathbf{r})}{(\dot{\mathbf{r}} \times \mathbf{r})^2} = -\frac{(\dot{\mathbf{a}} \cdot \dot{\mathbf{r}})\mathbf{r}^2 - (\dot{\mathbf{a}} \cdot \mathbf{r})(\dot{\mathbf{r}} \cdot \mathbf{r})}{\dot{\mathbf{r}}^2\mathbf{r}^2 - (\dot{\mathbf{r}} \cdot \mathbf{r})^2}. \quad (5.42)$$

If $\|\mathbf{r}(u)\| = 1$ for all u , this expression assumes a simpler form. In this case we write $\mathbf{r} = \mathbf{e}_1$ and note that $\mathbf{e}_1 \cdot \dot{\mathbf{e}}_1 = 0$. Equ. (5.42) simplifies to

$$v_s = -\dot{\mathbf{a}} \cdot \dot{\mathbf{e}}_1 / \dot{\mathbf{e}}_1^2. \quad (5.43)$$

The Distribution Parameter

We consider again the surface of Equ. (5.39), but with $\|\mathbf{r}(u)\| = 1$, and we let $\mathbf{r} = \mathbf{e}_1$. The Plücker coordinates of the ruling $R(u)$ are given by $R(u)\gamma = (\mathbf{r}, \bar{\mathbf{r}})\mathbb{R} = (\mathbf{e}_1, \mathbf{a} \times \mathbf{e}_1)\mathbb{R}$. The derivative point $(R\gamma)^1$ has coordinates $(\dot{\mathbf{e}}_1, \dot{\mathbf{a}} \times \mathbf{e}_1 + \mathbf{a} \times \dot{\mathbf{e}}_1)\mathbb{R}$.

The Klein image $\mathcal{R}\gamma$ of the ruled surface \mathcal{R} is a curve contained in the Klein quadric, whose tangent is spanned by the points $R\gamma(u)$ and $(R\gamma)^1(u)$. According to Th. 5.1.4, this curve tangent completely determines the parabolic linear congruence \mathcal{N} of surface tangents along $R(u)$, and Equ. (3.18) shows how to compute the distribution parameter d of \mathcal{N} from $R\gamma(u)$, $(R\gamma)^1(u)$:

$$-d = \frac{\mathbf{r}^2(\dot{\mathbf{r}} \cdot \bar{\mathbf{r}})}{(\dot{\mathbf{r}} \times \mathbf{r})^2}. \quad (5.44)$$

The value $\delta(u) = -d(u)$ is referred to as *distribution parameter of the ruled surface* at the generator $R(u)$. If we insert the definition of \mathbf{r} and $\bar{\mathbf{r}}$ into (5.44), we get

$$\delta = \det(\dot{\mathbf{a}}, \mathbf{e}_1, \dot{\mathbf{e}}_1) / \dot{\mathbf{e}}_1^2. \quad (5.45)$$

Remark 5.3.2. In Sec. 3.2.2 we discussed a parabolic linear congruence from the Euclidean point of view. There we showed that the tangent planes whose contact points are at distance d from the striction point are the bisectors of the central plane and the asymptotic plane (see Fig. 3.8 and Fig. 5.20).

We see that a small absolute value of the distribution parameter means that near the striction point the tangent planes are winding faster around the generator. ◇

Remark 5.3.3. There is a way to obtain the striction point and the distribution parameter in a difference-geometric limit process: Consider two neighbouring generators $R(u)$ and $R(u + h)$, oriented by the vectors $\mathbf{r}(u)$, $\mathbf{r}(u + h)$. The smallest distance $D(u, h)$ between $R(u)$ and $R(u + h)$ is measured between the footpoints $\mathbf{f}_1(u, h)$ and $\mathbf{f}_2(u, h)$ of the common perpendicular. If $\phi(u, h)$ is the angle between these two generators, then Prop. 5.4.1 will show that the distribution parameter and the striction point are found by

$$\delta(u) = \lim_{h \rightarrow 0} \frac{D(u, h)}{\phi(u, h)}, \quad s(u) = \lim_{h \rightarrow 0} \mathbf{f}_1(u, h) = \lim_{h \rightarrow 0} \mathbf{f}_2(u, h).$$

◇

Remark 5.3.4. The distribution parameter δ is the pitch of a helical motion whose axis is the central tangent, and whose velocity vectors at points of the generator are tangent to the ruled surface. If the pitch is positive, the winding of the tangent planes around the generators is negative. That is why $\delta = -d$. Fig. 5.20 shows path normals for negative pitch, and Fig. 5.21 shows path tangents for positive pitch. ◇

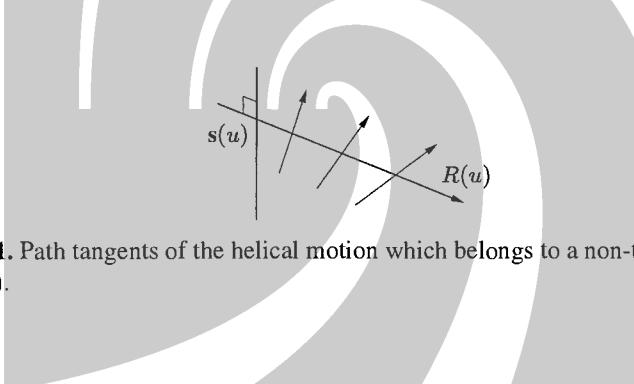


Fig. 5.21. Path tangents of the helical motion which belongs to a non-torsal generator $R(u)$.

Torsal Generators

A regular generator $R(u)$ of the ruled surface \mathcal{R} is torsal if and only if

$$\det(\dot{\mathbf{a}}(u_0), \mathbf{r}(u_0), \dot{\mathbf{r}}(u_0)) = 0. \quad (5.46)$$

This condition is also fulfilled if the generator is singular. By Equ. (5.5), $R(u)$ is regular and torsal if and only if

$$\text{rk}(\dot{\mathbf{a}}(u), \mathbf{r}(u), \dot{\mathbf{r}}(u)) = 2. \quad (5.47)$$

We distinguish two cases:

Definition. The regular torsal generator $R(u)$ is called cylindrical, if $\text{rk}(\mathbf{r}(u), \dot{\mathbf{r}}(u)) = 1$.

The condition of $R(u)$'s being cylindrical obviously is equivalent to

$$\mathbf{r}(u) \times \dot{\mathbf{r}}(u) = \mathbf{0}. \quad (5.48)$$

The cuspidal point of a cylindrical generator is at infinity. Note that all generator lines of a cylindrical surface are cylindrical. The value of the distribution parameter, which is undefined, is set to ∞ .

We compute the cuspidal point of a non-cylindrical generator $R(u)$. It can be written in the form $\mathbf{a}(u) + v_c(u)\mathbf{r}(u)$ with a certain value v_c . Among the other points of $R(u)$ it is characterized by the linear dependence of $\{\dot{\mathbf{a}} + v_c\dot{\mathbf{r}}, \mathbf{r}\}$, i.e., $(\dot{\mathbf{a}} + v_c\dot{\mathbf{r}}) \times \mathbf{r} = \mathbf{0}$. Thus v_c can be computed by

$$v_c = -\frac{(\dot{\mathbf{a}} \times \mathbf{r}) \cdot (\dot{\mathbf{r}} \times \mathbf{r})}{(\dot{\mathbf{r}} \times \mathbf{r})^2}. \quad (5.49)$$

Remark 5.3.5. We see that Equ. (5.49) is the same as (5.42), which computes the v parameter value of the striction point. Thus it is justified to define the striction point also in the case of a torsal generator, and let it equal the cuspidal point.

If the cuspidal point is not at infinity, it also makes sense to identify the generator's tangent plane with the asymptotic plane, and to define the central plane (orthogonal to the asymptotic plane), the central tangent and the central normal (both incident with the cuspidal point). The value of the distribution parameter is zero. \diamond

The Central and Asymptotic Developables

We consider a ruled surface \mathcal{R} and the families of its central and asymptotic planes. The envelopes of these two families are called the *central developable* and the *asymptotic developable* (see Th. 6.2.1). It is easy to see that the central developable touches the point set of the ruled surface in the points of the striction curve.

The Striction Curve

We consider the ruled surface \mathcal{R} parametrized by Equ. (5.39), and the striction points $\mathbf{s}(u)$ of the generators $R(u)$. We assume that $\|\mathbf{r}(u)\| = 1$ and let $\mathbf{e}_1 = \mathbf{r}$. The curve of all striction points — the *striction curve* of \mathcal{R} — has, according to (5.43), the parametrization

$$\mathbf{s}(u) = \mathbf{a}(u) + v_s(u)\mathbf{r}(u) = \mathbf{a} - \frac{\dot{\mathbf{a}} \cdot \dot{\mathbf{e}}_1}{\dot{\mathbf{e}}_1^2} \mathbf{e}_1. \quad (5.50)$$

Theorem 5.3.3. *If the ruled surface \mathcal{R} has no cylindrical generators, the family $R(u)$ of generators is geodesically parallel along the striction curve, and the striction curve is the unique directrix which has this property.*

Proof. We assume a parametrization like in (5.39) with $\mathbf{e}_1 = \mathbf{r}$, $\|\mathbf{e}_1\| = 1$. Differentiation shows that $\mathbf{e}_1 \cdot \dot{\mathbf{e}}_1 = 0$, and so the vector $\dot{\mathbf{e}}_1$, which is by definition parallel to the asymptotic plane, indicates the direction of the central normal. Thus the tangential component of $\dot{\mathbf{e}}_1$ vanishes. This is the definition of ‘geodesic parallel’.

To show the converse, we assume a directrix $\mathbf{a}(u)$. Geodesic parallelity of $\mathbf{e}_1(u)$ along $\mathbf{a}(u)$ means that $\dot{\mathbf{e}}_1 \cdot \mathbf{e}_1 = \dot{\mathbf{e}}_1 \cdot \dot{\mathbf{a}} = 0$. The former of these equations is fulfilled automatically, and the latter together with (5.43) shows that $v_s = 0$ and therefore \mathbf{a} is the striction curve. \square

Proposition 5.3.4. If \mathcal{R} is a ruled surface without cylindrical generators, any two of the following three statements about a curve $\mathbf{c}(u) = \mathbf{a}(u) + v(u)\mathbf{r}(u)$ imply the third one: (i) \mathbf{c} is the striction curve; (ii) \mathbf{c} is a geodesic; (iii) \mathbf{c} intersects all generators under constant angle.

Proof. Parallel transport along a curve is a Euclidean displacement, and so the angle between vectors of two parallel fields is constant. If the striction curve is a geodesic, its unit tangent vector field is a parallel field, and therefore the angle between the striction curve and the generators $R(u)$ is constant. This shows that (i) and (ii) imply (iii). Conversely, (i) and (iii) show that the unit tangent vector field of the striction curve is a parallel field which means that it is a geodesic, so (i) and (iii) imply (ii). Finally, if (ii) holds then (iii) shows that the family of generators is geodesically parallel along the curve \mathbf{c} , and (i) follows from Th. 5.3.3. \square

Examples: Conoidal Surfaces and Surfaces of Constant Slope

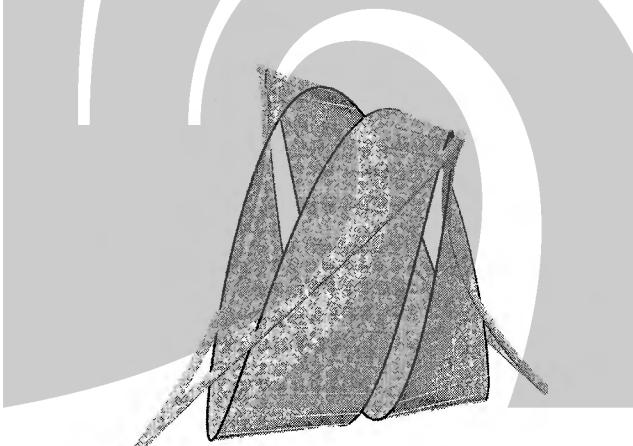


Fig. 5.22. Conoidal surface with striction curve.

Example 5.3.1. A surface \mathcal{R} whose director cone is a plane is called *conoidal*. This means that all generators are parallel to a fixed plane ε , which will be assumed horizontal (see Fig. 5.22). This implies that also all asymptotic planes are parallel to ε , and the asymptotic developable degenerates.

All central planes are orthogonal to ε , and so the central developable is a vertical cylinder. Its generators are vertical tangents of the ruled surface, and the striction curve is the contour with respect to vertical projection. \diamond

Example 5.3.2. (Continuation of Ex. 5.3.1.) If a conoidal ruled surface \mathcal{R} has a director line, it is called a *conoid*. An example of a conoid is a regulus contained in a hyperbolic paraboloid.

If the director line is orthogonal to the plane ε , we have a *right conoid*. Examples are given by the circle conoid of Ex. 5.1.4 and Ex. 5.2.5 (Fig. 5.6), by a right hyperbolic paraboloid, the Plücker conoid and the Zindler conoid (cf. Sec. 5.2.3). The striction curve of a right conoid is contained in the director line. \diamond

Example 5.3.3. If a ruled surface \mathcal{R} with generators $R(u)$ has the property that there is a line L , which we imagine as vertical, such that $\angle(L, R(u)) = \alpha = \text{const.}$, then \mathcal{R} is called a ruled surface of *constant slope* (see Fig. 5.23).

If we assume a parametrization of the ruled surface of the form (5.39), the director cone is parametrized by $\mathbf{x}(u, v) = v \cdot \mathbf{r}(u)$. If the angle enclosed by $\mathbf{r}(u)$ and L is constant, also the angle enclosed by the tangent planes of the director cone and the line L is constant.

This shows that the angle enclosed by \mathcal{R} 's asymptotic planes and L is constant, so the asymptotic developable is a developable surface of constant slope (with respect to L). Such surfaces will be discussed in Sec. 6.3.

The central planes are parallel to L , and therefore the central developable is a vertical cylinder, or possibly degenerate. Like in Ex. 5.3.1, the striction curve is the contour line with respect to vertical projection.

An example is a regulus with rotational symmetry about a vertical axis. Its striction curve is its smallest parallel circle. This is also an example of a geodesic striction curve as described by Prop. 5.3.4. \diamond

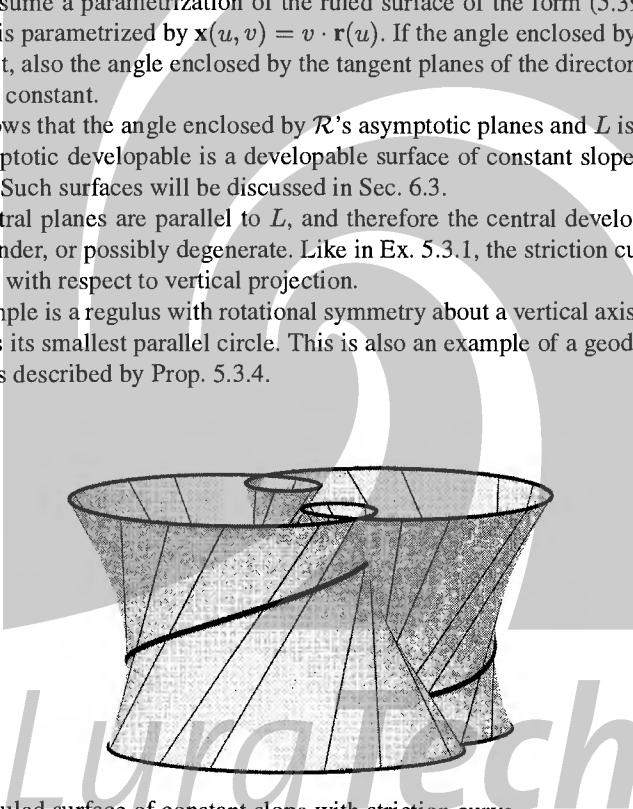


Fig. 5.23. Ruled surface of constant slope with striction curve.

The Striction Curve of Rational Ruled Surfaces

For a rational ruled surface \mathcal{R} , parametrized by (5.39), we may assume that both $\mathbf{a}(u)$ and $\mathbf{r}(u)$ are rational functions. Equ. (5.42) and (5.50) show that the striction curve is rational, and so are the central and asymptotic developables.

Without proof we mention that the degree of the striction curve as an algebraic curve is less than or, in general, equal to $4d - 4$, if d is the degree of \mathcal{R} . For example, a

regulus which has no rotational symmetry possesses a striction curve of degree four (see Fig. C.5).

This may be motivated by the following formula for the striction curve whose proof is left to the reader as an exercise: If $(\mathbf{r}(u), \bar{\mathbf{r}}(u))$ is a parametrization of \mathcal{R} in Plücker coordinates, then the striction curve $s(u)$ is given by

$$s(u) = ((\mathbf{r} \times \dot{\mathbf{r}})^2, \det(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}})\mathbf{r} + \det(\mathbf{r}, \bar{\mathbf{r}}, \dot{\mathbf{r}})\dot{\mathbf{r}} + (\bar{\mathbf{r}} \cdot \dot{\mathbf{r}})(\mathbf{r} \times \dot{\mathbf{r}})) \mathbb{R}. \quad (5.51)$$

The Striction Curve of Torsal Ruled Surfaces

The striction curve of a *cylinder* is not well defined. The generators are geodesically parallel along any surface curve. The striction curve of a *cone* degenerates to its vertex. For the *tangent surface* of a curve \mathbf{s} (cf. Fig. 5.7), which is parametrized by

$$\mathbf{x}(u, v) = \mathbf{s}(u) + v\dot{\mathbf{s}}(u),$$

the curve \mathbf{s} itself agrees with the striction curve, and consists of the set of cuspidal points.

The tangent plane along a generator $R(u)$ is parallel to $\dot{\mathbf{s}}(u), \ddot{\mathbf{s}}(u)$, and thus it is the curve's osculating plane at $\mathbf{s}(u)$.

Remark 5.3.6. The analogue to the central plane for torsal surfaces (cf. Remark 5.3.5) is orthogonal to the osculating plane and contains the curve tangent. It therefore coincides with the curve's rectifying plane at $\mathbf{s}(u)$.

Generator, central normal and central tangent of \mathcal{R} coincide with tangent, principal normal and bi-normal of the curve \mathbf{s} , which suggests to extend Euclidean curve theory to ruled surfaces. This will be done in Sec. 5.3.2. \diamond

5.3.2 A Complete System of Euclidean Invariants

Regularity of the Striction Curve

We consider ruled surfaces which possess a regular striction curve. This excludes e.g., cylindrical generators, or conical surfaces whose striction curve degenerates to a point, but includes tangent surfaces of curves, which are torsal. We enumerate some conditions which ensure that the striction curve is regular.

Lemma 5.3.5. *The striction curve \mathbf{s} of a C^2 ruled surface \mathcal{R} is regular in its point $\mathbf{s}(u)$, if one of the following is satisfied:*

1. *The generator $R(u)$ is not torsal.*
2. *\mathcal{R} is C^3 , $R(u)$ is torsal, and the curve $(R\gamma)^1(u)$ intersects the Klein quadric transversely (i.e., $(R\gamma)^2(u)$ is not contained in M_2^4).*
3. *The ruled surface \mathcal{R} has contact of second or higher order with the tangent surface of a regular curve.*

Proof. We are going to show that a singularity of s forces both $(R\gamma)^1(u)$ and $(R\gamma)^2(u)$ to be contained in the Klein quadric. This will prove 1. and 2.

We use a parametrization like in (5.39) and assume that $\mathbf{r} = \mathbf{e}_1$ with $\|\mathbf{e}_1\| = 1$. We may use s as directrix and let $\mathbf{a} = s$ locally. Then (5.43) shows that $\dot{s} \cdot \dot{\mathbf{e}}_1 = 0$ locally, and also $\ddot{s} \cdot \dot{\mathbf{e}}_1 + \dot{s} \cdot \ddot{\mathbf{e}}_1 = 0$. Singularity of s means $\dot{s}(u) = \mathbf{o}$, and therefore $\ddot{s} \cdot \dot{\mathbf{e}}_1(u) = 0$.

We have $\dot{\mathbf{e}}_1 \neq \mathbf{o}$ (the generators of \mathcal{R} are non-cylindrical), and so we may locally re-parametrize such that $\|\dot{\mathbf{e}}_1\| = 1$, which implies $\dot{\mathbf{e}}_1 \cdot \ddot{\mathbf{e}}_1 = 0$. We compute $R\gamma = (\mathbf{e}_1, s \times \mathbf{e}_1)\mathbb{R}$, $(R\gamma)^1 = (\dot{\mathbf{e}}_1, \dot{s} \times \mathbf{e}_1 + s \times \dot{\mathbf{e}}_1)\mathbb{R}$, $(R\gamma)^2 = (\ddot{\mathbf{e}}_1, \ddot{s} \times \mathbf{e}_1 + 2\dot{s} \times \dot{\mathbf{e}}_1 + s \times \ddot{\mathbf{e}}_1)\mathbb{R}$.

Obviously $\dot{s}(u) = \mathbf{o}$ implies that $(R\gamma)^1(u)$ is contained in the Klein quadric. The point $(R\gamma)^2(u)$ is contained in M_2^4 , if and only if $\ddot{\mathbf{e}}_1 \cdot (\ddot{s} \times \mathbf{e}_1 + s \times \ddot{\mathbf{e}}_1) = 0$. This condition reduces to $\det(\ddot{\mathbf{e}}_1, \ddot{s}, \mathbf{e}_1) = 0$. If $\dot{s} = 0$, then the three vectors $\ddot{\mathbf{e}}_1$, \ddot{s} , \mathbf{e}_1 , are orthogonal to $\dot{\mathbf{e}}_1$, so we have indeed $(R\gamma)^2(u) = 0$. This shows 1. and 2.

To show 3., we use the following argument: If surfaces are in second order contact, their striction curves are in first order contact. But the striction curve of a tangent surface of a regular curve is just this regular curve. This completes the proof. \square

If the striction curve s is regular, we can assume a parametrization of the ruled surface such that $\|\dot{s}\| = 1$. We will assume such a parametrization and indicate differentiation with respect to such a *natural* parameter with primes instead of dots. If we use a parametrization like in (5.39) then (5.43) shows that

$$\mathbf{s}' \cdot \mathbf{e}_1' = 0. \quad (5.52)$$

The Sannia Frame

We use the notations of the previous paragraph. The *Sannia frame* is the orthonormal frame which is attached to the striction point $s(u)$, and consists of the three unit vectors \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 , where

$$\mathbf{e}_2 = \frac{\mathbf{e}_1'}{\|\mathbf{e}_1'\|}, \quad \mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2 = \frac{\mathbf{e}_1 \times \mathbf{e}_1'}{\|\mathbf{e}_1'\|}. \quad (5.53)$$

Obviously \mathbf{e}_2 and \mathbf{e}_3 indicate the directions of the central normal and central tangent, respectively.

Lemma 5.3.6. *The Sannia frame of a C^2 ruled surface fulfills the equations*

$$\begin{aligned} \mathbf{e}_1' &= \kappa \mathbf{e}_2, \\ \mathbf{e}_2' &= -\kappa \mathbf{e}_1 + \tau \mathbf{e}_3, \\ \mathbf{e}_3' &= -\tau \mathbf{e}_2, \end{aligned} \quad (5.54)$$

$$\text{with } \kappa = \|\mathbf{e}_1'\|, \quad \tau = \det(\mathbf{e}_1, \mathbf{e}_1', \mathbf{e}_1'')/\|\mathbf{e}_1'\|^2. \quad (5.55)$$

Proof. The proof is analogous to the construction of the Frenet frame of space curves: We let $\mathbf{e}_i' = \sum a_{ik} \mathbf{e}_k$ with unknown coefficients a_{ik} . Equ. (5.53) shows that $a_{12} = \|\mathbf{e}_1'\|$ and $a_{13} = 0$.

Differentiating the identities $\mathbf{e}_i \cdot \mathbf{e}_l = \text{const.}$ for all i, l shows $a_{jk} = -a_{kj}$, for all j, k , and especially $a_{jj} = 0$. Finally we let $a_{12} = \kappa$ and $a_{23} = \tau$. To compute $\tau = -\mathbf{e}'_3 \cdot \mathbf{e}_2$, we first differentiate \mathbf{e}_3

$$\mathbf{e}'_3 = \frac{1}{\|\mathbf{e}'_1\|^3} (\mathbf{e}'^2(\mathbf{e}_1 \times \mathbf{e}'_1) - (\mathbf{e}'_1 \cdot \mathbf{e}'_1)(\mathbf{e}_1 \times \mathbf{e}'_1)),$$

and then use (5.53) to get (5.55). \square

By construction, the coefficients $\kappa(u)$ and $\tau(u)$ are Euclidean invariants of the ruled surface. They are referred to as *curvature* and *torsion* at the generator $R(u)$.

Remark 5.3.7. The curvature κ is always nonnegative. Zero curvature means that $\|\mathbf{e}'_1\| = 0$, but this can only happen at cylindrical generators which had been excluded anyway (they do not possess a proper striction point).

If $\tau(u)$ is zero for all u , the central tangent vector $\mathbf{e}_3(u)$ is constant, and all generators are orthogonal to \mathbf{e}_3 . This characterizes conoidal surfaces. \diamond

Spherical Kinematics of the Sannia Frame

We attach the three vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of the Sannia frame to the origin. If we move along the striction curve, this orthonormal frame performs some spherical motion. The three path curves $\mathbf{e}_1(u), \mathbf{e}_2(u)$, and $\mathbf{e}_3(u)$ are called *spherical generator image*, *spherical central normal image*, and *spherical central tangent image*, respectively.

If we parametrize the motion of this frame by the unit length of the striction curve of the underlying ruled surface \mathcal{R} , the velocities $\|\mathbf{e}'_1\|, \|\mathbf{e}'_3\|$ of the points $\mathbf{e}_1, \mathbf{e}_3$ are κ and $|\tau|$, respectively. If we denote the arc length differentials of the curves $\mathbf{e}_1(u), \mathbf{e}_3(u)$, by ds_1, ds_3 , respectively, then

$$\kappa = \frac{ds_1}{du}, \quad \tau = \frac{ds_3}{du}. \quad (5.56)$$

By Prop. 3.4.1, the spherical motion of the Sannia frame assigns to each point \mathbf{x} of Euclidean space a velocity vector $\mathbf{v}(\mathbf{x})$ which can always be expressed in the form $\mathbf{v}(\mathbf{x}) = \mathbf{d} \times \mathbf{x}$. Especially we have

$$\mathbf{v}(\mathbf{e}_1) = \mathbf{e}'_1 = \mathbf{d} \times \mathbf{e}_1, \quad \mathbf{v}(\mathbf{e}_2) = \mathbf{e}'_2 = \mathbf{d} \times \mathbf{e}_2, \quad \mathbf{v}(\mathbf{e}_3) = \mathbf{e}'_3 = \mathbf{d} \times \mathbf{e}_3, \quad (5.57)$$

with the Darboux vector \mathbf{d} . The Sannia equations (5.54) show that in our case

$$\mathbf{d} = \tau \mathbf{e}_1 + \kappa \mathbf{e}_3. \quad (5.58)$$

Remark 5.3.8. The infinitesimal spherical motion of the Sannia frame is an infinitesimal rotation about the Darboux vector. Equ. (5.58) shows that the axis is always contained in the plane spanned by \mathbf{e}_1 and \mathbf{e}_3 .

We consider the union of Darboux vectors in the fixed frame and in the moving frame: We get two cones \mathcal{P} and \mathcal{P}_0 , whose vertex is the origin, and whose generators are spanned by the Darboux vectors: With respect to the fixed frame, the set of

Darboux vectors is parametrized by $P_0(u) = [\mathbf{d}(u)] = [\tau(u)\mathbf{e}_1(u) + \kappa(u)\mathbf{e}_3(u)]$. With respect to the moving frame $\mathbf{e}_1(u), \mathbf{e}_2(u), \mathbf{e}_3(u)$, the set of Darboux vectors is simply parametrized by $P(u) = [(\tau(u), 0, -\kappa(u))]$. Obviously \mathcal{P} is contained in the plane $x_2 = 0$.

We will show (see Prop. 8.2.15) that the spherical motion of the Sannia frame is generated by the *rolling* of \mathcal{P} ('moving polhode') on \mathcal{P}_0 ('fixed polhode'). This is illustrated in Fig. 5.24. Thus this motion is a spherical involute motion, and the paths of \mathbf{e}_1 and \mathbf{e}_3 are spherical involutes of the curve $\mathbf{d}(u)/\|\mathbf{d}(u)\|$. \diamond

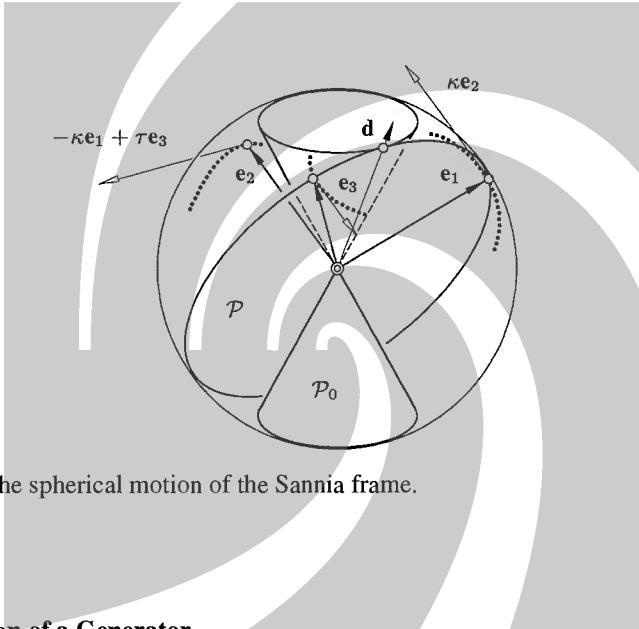


Fig. 5.24. The spherical motion of the Sannia frame.

The Striction of a Generator

So far the curvature theory of ruled surfaces was essentially the same as for space curves. But it is not true that a ruled surface is uniquely determined by its curvature and torsion — we need one further invariant. We use the angle enclosed by the striction curve and the generator lines:

Definition. If \mathcal{R} is a ruled surface in Euclidean space with striction curve s , which is regular at $s(u)$, then the striction $\sigma(u)$ of the generator $R(u)$ is the angle between the striction curve and $R(u)$:

$$\mathbf{s}'(u) = \mathbf{e}_1 \cos \sigma(u) + \mathbf{e}_3 \sin \sigma(u). \quad (0 \leq \sigma(u) < 2\pi) \quad (5.59)$$

Recall that the surface's tangent plane in a striction point is spanned by the vectors \mathbf{e}_1 and \mathbf{e}_3 . Th. 5.3.8 will show that the four Euclidean invariants curvature, torsion, striction and distribution parameter cannot be unrelated. The relation expressed by Lemma 5.3.7 is a very simple one:

Lemma 5.3.7. *If \mathcal{R} is a C^2 ruled surface and $\delta(u)$, $\kappa(u)$, $\sigma(u)$ are the distribution parameter, the curvature, and the striction of the generator $R(u)$, respectively, then*

$$\delta = \sin \sigma / \kappa. \quad (5.60)$$

Proof. Equations (5.59), (5.55), and (5.45) show that

$$\sin \sigma = s' \cdot e_3 = s' \cdot \frac{(e_1 \times e'_1)}{\|e'_1\|} = \frac{\det(s', e_1, e'_1)}{\|e'_1\|} = \kappa \delta. \quad \square$$

Behaviour of κ, τ, σ under Reversion of Orientations

Of course a ruled surface \mathcal{R} with generators $R(u)$ determines the vectors $e_1(u)$ uniquely only up to multiplication with ± 1 . Assume that we have chosen a continuous family $e_1(u)$ of unit vectors, and have already computed $\kappa(u)$, $\tau(u)$, and $\sigma(u)$. If we choose $-e_1$ instead, we also get $-e_2 = (-e'_1)/\|-e'_1\|$ instead of e_2 , but $e_3 = (-e_1) \times (-e_2)$ remains unchanged. Equ. (5.55) shows that κ remains unchanged. Further, (5.55), (5.54), and (5.59) show that τ and $\cos \sigma$ reverse their sign, but $\sin \sigma$ remains unchanged.

There are two possible ways to parametrize the striction curve with its arc length. If we reverse this parametrization, but keep the vector e_1 unchanged, first derivatives reverse their sign, which shows $s' \rightarrow -s'$, $e'_1 \rightarrow -e'_1$. The other two vectors of the Sannia frame therefore change according to $e_2 \rightarrow -e_2$ and $e_3 \rightarrow e_1 \times (-e_2) = -e_3$. Because of $e'_3 \rightarrow -(-e'_3) = e'_3$, $\sin \sigma$ does not change, whereas τ and $\cos \sigma$ reverse their sign.

Since the definition of the distribution parameter did not involve any arbitrary choice of an invariant parametrization or a unit vector, it does not change under either change of orientation, and we could use (5.60) to show the invariance of $\sin \sigma$. We also see that the quotient

$$\delta_c = \cos \sigma / \tau, \quad (5.61)$$

is not affected by changes of orientation.

Remark 5.3.9. It is easy to see that δ_c is the distribution parameter of the central tangent surface, whose directrix is the striction curve s of the original surface, and whose generators are its central tangents. It has striction $\pi/2 - \sigma$, if σ is the striction of the original surface, and it has the same central normals. This follows from Th. 5.3.3. \diamond

The Fundamental Theorem of Euclidean Differential Geometry of Ruled Surfaces

Before formulating the theorem, a remark concerning tangent surfaces: Consider the tangent surface \mathcal{T} of a curve s . Its Sannia frame is the Frenet frame of s . Thus curvature and torsion of the tangent surface are equal to curvature and torsion of the striction curve s , and the striction is zero. If u is an arc length parameter for

s , then the functions $\kappa(u)$ and $\tau(u)$ determine the curve up to motions in E^3 . The following is a generalization of the fundamental theorem of Euclidean curve theory, which states that a curve is uniquely determined by its curvature and torsion.

Theorem 5.3.8. (*Fundamental theorem of Euclidean ruled surface theory by G. Sannia, 1925*). Assume that $\kappa(u)$, $\tau(u)$, $\sigma(u)$ are real-valued functions defined in an interval I with κ positive, κ and σ continuously differentiable, and τ continuous. Then there exists a twice continuously differentiable ruled surface, parametrized with the arc length u of its striction curve, with curvature $\kappa(u)$, torsion $\tau(u)$, and whose sine and cosine of the striction equal $\sin \sigma(u)$, $\cos \sigma(u)$. This surface is unique up to Euclidean motions.

Proof. We choose $u_0 \in I$ and the orthonormal frame $\mathbf{e}_1(u_0)$, $\mathbf{e}_2(u_0)$, $\mathbf{e}_3(u_0)$. The differential equation (5.54) for vectors $\mathbf{e}_1(u)$, $\mathbf{e}_2(u)$, $\mathbf{e}_3(u)$ has a unique solution in I . The skew symmetry of the coefficient matrix ensures that $(\mathbf{e}_i \cdot \mathbf{e}_j)' = 0$. This shows that $\mathbf{e}_1(u)$, $\mathbf{e}_2(u)$, $\mathbf{e}_3(u)$ is an orthonormal frame for all $u \in I$. Finally we choose $s_0 \in E^3$ and integrate (5.59) to define the curve $s(u)$ via

$$s(u) = s_0 + \int_{u_0}^u (\mathbf{e}_1(u) \cos \sigma(u) + \mathbf{e}_3(u) \sin \sigma(u)) du. \quad (5.62)$$

It is an easy exercise to verify that the ruled surface whose directrix is s and whose generators $R(u)$ are parallel to $\mathbf{e}_1(u)$ fulfills the requirements of the theorem.

It is clear that any solution surface must fulfill both (5.54) and (5.59), so we have also shown that any possible solution must arise in this way, and is therefore uniquely determined by the choice of $\mathbf{e}_1(u_0)$, $\mathbf{e}_2(u_0)$, $\mathbf{e}_3(u_0)$, and s_0 . For any two such choices there is a Euclidean motion which transforms the first into the second, which shows the uniqueness statement of the theorem. \square

Remark 5.3.10. In order to define the striction uniquely we have to restrict ourselves to $0 \leq \sigma < 2\pi$. If the striction grows and reaches 2π , it jumps to zero. This discontinuity can be avoided by allowing arbitrary values of σ and considering all values whose difference is an integer multiple of 2π as equivalent. This is the reason why we formulated Th. 5.3.8 with $\sin \sigma$ and $\cos \sigma$ instead of σ itself. \diamond

Ruled Surfaces with Constant κ , τ , and σ

Consider a ruled surface \mathcal{R} generated by (i) a helical motion or (ii) a rotation. In roughly decreasing order of generality, there are the following possibilities: (ia) \mathcal{R} is a skew helical ruled surface (ib) \mathcal{R} is the tangent surface of a helix, (iia) \mathcal{R} is a regulus with rotational symmetry, or (iib1) the tangent surface of a circle, or (iib2) a cone of revolution, or (iic) a cylinder of revolution, or (iid) a pencil of lines. Cases (ia) to (iib1) are admissible surfaces for our present theory. Fig. 5.25 illustrates these surfaces.

Proposition 5.3.9. Cases (ia) to (iib1) of the list contained in the previous paragraph are exactly the ruled surfaces with constant invariants κ , τ , and σ .

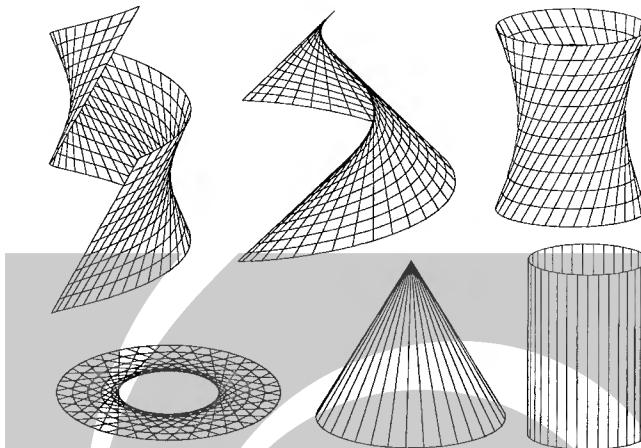


Fig. 5.25. Cases (ia), (ib), (iia), (iib1), (iib2), (iic) of ruled surfaces generated by uniform rotations or helical motions.

Proof. Because these surfaces admit a one-parameter subgroup of automorphic Euclidean motions, their invariants are constant. To show the converse, we do the following: Choose $u_0 \in \mathbb{R}$ and the initial values of \mathbf{e}_i and s as described in the proof of Th. 5.3.8, and construct a solution surface \mathcal{R}_0 . Repeat the procedure for $u_1 \in \mathbb{R}$, but with initial values $\mathbf{e}_i(u_1)$, $s(u_1)$, which are found by evaluating the respective functions of \mathcal{R}_0 at a second parameter value u_1 . This gives a surface \mathcal{R}_1 .

Denote the Euclidean motion which maps $\mathbf{e}_i(u_0)$ to $\mathbf{e}_i(u_1)$ and $s(u_0)$ to $s(u_1)$ with α_{u_1} . Then $\alpha_{u_1}(\mathcal{R}_0)$ and \mathcal{R}_1 share the values of \mathbf{e}_i and s at u_0 , and their curvature, torsion, and striction agree for all $u \in \mathbb{R}$. Th. 5.3.8 shows that $\alpha_{u_1}(\mathcal{R}_0) = \mathcal{R}_1$, and \mathcal{R}_0 admits an automorphic Euclidean motion.

The same argument shows $\alpha_{s+t}(\mathcal{R}_0) = \alpha_s(\alpha_t(\mathcal{R}_0))$, and so the family of all such automorphic motions is a one-parameter group α_t , parametrized with $t \in \mathbb{R}$. \square

The Intrinsic Metric of Ruled Surfaces

Our aim here is to express some properties of the Euclidean differential geometry of surfaces in terms of curvature, torsion, and striction.

We parametrize a ruled surface by $\mathbf{x}(u, v) = \mathbf{s}(u) + v\mathbf{e}_1(u)$ with $\|\mathbf{e}_1\| = 1$, and compute the partial derivatives

$$\mathbf{x}_u = s' + v\mathbf{e}'_1 = \mathbf{e}_1 \cos \sigma + \mathbf{e}_3 \sin \sigma + v\kappa \mathbf{e}_2, \quad \mathbf{x}_v = \mathbf{e}_1. \quad (5.63)$$

This shows that the coefficients g_{ij} of the metric fundamental form take the form

$$g_{11} = \mathbf{x}_u \cdot \mathbf{x}_u = 1 + v^2\kappa^2, \quad g_{12} = \mathbf{x}_u \cdot \mathbf{x}_v = \cos \sigma, \quad g_{22} = \mathbf{x}_v \cdot \mathbf{x}_v = 1. \quad (5.64)$$

In the following text the notion ‘local isometry’ of surfaces $\mathbf{x}(u, v)$ means that all (u, v) have a neighbourhood U such that the mapping from $\mathbf{x}(U)$ is well defined and an isometry. It is necessary to formulate the statement of the theorem in this way because the point set of a ruled surface may have self-intersections and there is no well-defined mapping of surface points, but only of parameter values.

Theorem 5.3.10. Assume that \mathcal{R} and $\bar{\mathcal{R}}$ are two ruled surfaces parametrized by $\mathbf{x}(u, v) = \mathbf{s}(u) + v\mathbf{e}_1(u)$ and $\bar{\mathbf{x}}(u, v) = \bar{\mathbf{s}}(u) + v\bar{\mathbf{e}}_1(u)$ with curvatures κ , $\bar{\kappa}$, and strictions σ , $\bar{\sigma}$. The mapping $\mathbf{x}(u, v) \mapsto \bar{\mathbf{x}}(u, v)$ is a local isometry if $\kappa(u) = \bar{\kappa}(u)$ and $\sigma(u) = \bar{\sigma}(u)$.

Proof. The metric fundamental form depends on κ and σ only. This shows that the mapping $\mathbf{x}(u, v) \mapsto \bar{\mathbf{x}}(u, v)$ is a local isometry. \square

Obviously the generators $R(u)$, $\bar{R}(u)$, and the striction points $s(u)$, $\bar{s}(u)$ correspond in this isometry.

Remark 5.3.11. If two ruled surfaces \mathcal{R} , $\bar{\mathcal{R}}$ fulfill the assumptions of Th. 5.3.10, we consider the family \mathcal{R}_h ($0 \leq h \leq 1$) of ruled surfaces, which are uniquely determined, up to a Euclidean motion, by their invariants

$$\kappa_h(u) = \kappa(u), \quad \sigma_h(u) = \sigma(u), \quad \tau_h(u) = (1 - h)\tau(u) + h\bar{\tau}(u).$$

This family of ruled surfaces, all of which are isometric, yields a *continuous bending* according to F. Minding [124], which connects \mathcal{R} and $\bar{\mathcal{R}}$. \diamond

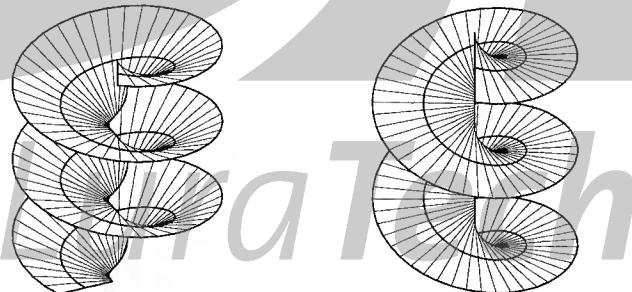


Fig. 5.26. Left: Ruled helical surface. Right: isometric conoidal surface.

Corollary 5.3.11. All C^2 ruled surfaces with regular striction curve are locally isometric to a conoidal surface.

Proof. Conoidal surfaces are characterized by vanishing torsion, so the statement follows from Th. 5.3.10. \square

Fig. 5.26 shows an example of a ruled surface and a conoidal surface isometric to it, which has helical symmetry because the original surface has. In this case the result is actually a right conoid whose generator line is the helical axis.

Remark 5.3.12. If we apply the isometry described in Cor. 5.3.11 to the tangent surface of a curve, the resulting conoid is planar, because the striction of a tangent surface is zero. Thus this isometry is a generalization of the development of tangent surfaces. \diamond

Gaussian Curvature

We consider the ruled surface \mathcal{R} whose point set is parametrized by $\mathbf{x}(u, v) = \mathbf{s}(u) + v\mathbf{e}_1(u)$, with \mathbf{s} as striction curve and $\|\mathbf{e}_1\| = 1$. We compute the unit normal vector field $\mathbf{n}(u, v)$ by normalizing $\mathbf{x}_u \times \mathbf{x}_v$. By Equ. (5.63), we have

$$\begin{aligned}\mathbf{n}(u, v) &= \frac{1}{G(u, v)} (\sin \sigma(u) \mathbf{e}_2(u) - v\kappa(u) \mathbf{e}_3(u)), \quad \text{with} \\ (G(u, v))^2 &= g_{11}(u, v)g_{22}(u, v) - g_{12}(u, v)^2 = \sin^2 \sigma(u) + v^2 \kappa^2(u).\end{aligned}$$

The coefficients h_{22} and h_{12} of the second fundamental form are computed by

$$h_{22} = \mathbf{x}_{vv} \cdot \mathbf{n} = 0, \quad h_{12} = \mathbf{x}_{uv} \cdot \mathbf{n} = \frac{\kappa \sin \sigma}{G}.$$

This shows that the Gaussian curvature equals

$$K(u, v) = -\frac{h_{12}(u, v)^2}{G(u, v)^2} = -\frac{\kappa^2(u) \sin^2 \sigma(u)}{(\sin^2 \sigma(u) + v^2 \kappa^2(u))^2}. \quad (5.65)$$

Equations (5.65) and (5.60) imply

Corollary 5.3.12. (Formula of E. Larmarle, 1858) The Gaussian curvature $K(u, v)$ of the ruled surface $\mathbf{x}(u, v) = \mathbf{s}(u) + v\mathbf{e}_1(u)$, with \mathbf{s} as striction curve, and $\|\mathbf{e}_1\| = 1$, equals

$$K(u, v) = -\frac{\delta(u)^2}{(\delta(u)^2 + v^2)^2}. \quad (5.66)$$

δ is the distribution parameter of the ruled surface.

Example 5.3.4. Fig. 5.27 shows contour lines of the Gaussian curvature function on the hyperboloid with equation $x^2 - 2y^2 + z^2/2 = 1$, for values in the interval $[-1/2, -1/30]$. It follows from Cor. 5.3.12 that the point of a generator with maximal Gaussian curvature is the striction point. Thus the contour lines must be tangent to the generators in their striction points. \diamond

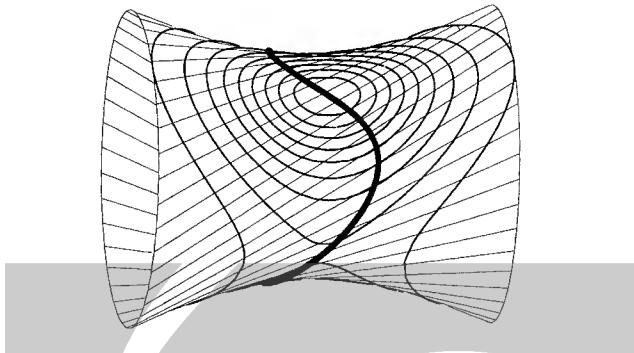


Fig. 5.27. Regulus with striction curve (thick) and contour lines of the Gaussian curvature function (cf. Ex. 5.3.4).

Remark 5.3.13. The fact that the torsion of a ruled surface does not enter the formula for the Gaussian curvature $K(u, v)$ illustrates the fact that $K(u, v)$ is actually defined by the intrinsic metric.

The distribution parameter vanishes for a non-cylindrical torsal generator $R(u)$. This shows that the Gaussian curvature is zero for all regular points of $R(u)$.

The points of a non-torsal generator are hyperbolic ($K < 0$). We had already seen this from the existence of two asymptotic tangents (Cor. 5.1.10).

The absolute value of Gaussian curvature assumes its maximum $1/\delta^2$ in the striction point, which is a direct consequence of (5.66). This is another characterization of the striction point in terms of the inner geometry of the surface. Thus isometric mappings of ruled surfaces onto ruled surfaces must map striction curves to striction curves. We also see that there are no ruled surfaces of constant negative Gaussian curvature $K < 0$. \diamond

Euclidean Differential Geometry of Cones and Cylinders

So far we considered surfaces which have a regular striction curve. A generator $R(u)$ of a ruled surface \mathcal{R} which has no striction point or where the striction curve is not regular, has one of the following three properties: (i) the surface is a cylinder in a neighbourhood of $R(u)$, (ii) the surface is a cone in a neighbourhood of $R(u)$, or (iii) every neighbourhood of $R(u)$ contains a piece of ruled surface with regular striction curve. The last case can be arbitrarily complicated and it makes no sense to attempt to classify surfaces with ‘piecewise’ regular striction curve. The first case, the cylinder surfaces, are completely characterized by a planar section orthogonal to the generators. Their theory is equivalent to the theory of planar curves. The second case, the cones, will be considered here in more detail.

A cone with parametrization $\mathbf{x}(u, v) = v\mathbf{e}_1(u)$ is defined by the curve $\mathbf{e}_1(u)$ contained in the unit sphere. A natural parametrization of the cone is induced by an arc length parametrization of the curve \mathbf{e}_1 , which means

$$\|\mathbf{e}'_1\| = 1. \quad (5.67)$$

Here again the prime denotes differentiation with respect to a natural parameter. A canonical orthonormal frame associated with the cone is $\mathbf{e}_1, \mathbf{e}_2 = \mathbf{e}'_1, \mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$. This frame is the well known *Darboux frame* of the curve $\mathbf{e}_1(u)$ with respect to the unit sphere Σ . The derivatives of the frame vectors satisfy the equations

$$\begin{aligned}\mathbf{e}'_1 &= & \mathbf{e}_2, \\ \mathbf{e}'_2 &= & -\mathbf{e}_1 & + k\mathbf{e}_3, \\ \mathbf{e}'_3 &= & -k\mathbf{e}_2,\end{aligned}\tag{5.68}$$

which follows in exactly the same way as Lemma 5.3.6. It is well known that k is the *geodesic curvature* of the curve \mathbf{e}_1 , when regarded as a surface curve of the unit sphere (the normal curvature being equal to 1 and the geodesic torsion vanishing). The value $k(u)$ is called *conical curvature* of the generator $R(u)$.

Theorem 5.3.13. (*The fundamental theorem of Euclidean differential geometry of cones*): *If $k(u)$ is a continuous function defined in an open interval I , then there exists a C^2 cone, parametrized canonically with the arc length u of its spherical image, such that $k(u)$ is its conical curvature. This cone is unique up to Euclidean motions.*

Proof. The existence part of the proof is completely analogous to the first part of the proof of Th. 5.3.8. The uniqueness is also seen analogously. \square

The conical curvature can be computed via

$$k = \det(\mathbf{e}_1, \mathbf{e}'_1, \mathbf{e}''_1),\tag{5.69}$$

which follows directly from $\mathbf{e}''_1 = \mathbf{e}'_2$ and the second line of (5.68).

Remark 5.3.14. If we choose to describe the cone \mathcal{R} by the curve $-\mathbf{e}_1$ instead of \mathbf{e}_1 , this changes the sign of the conical curvature k . (5.69) shows that $k > 0$ means that \mathbf{e}''_1 points to the left, if we sit on the sphere and look in the direction indicated by \mathbf{e}'_1 , so $k > 0$ is equivalent to a ‘left bend’ if \mathbf{e}_1 is a road in the unit globe. \diamond

Remark 5.3.15. The vanishing of k characterizes a point $\mathbf{e}_1(s_0)$ where $\mathbf{e}_1(s_0), \mathbf{e}'_1(s_0), \mathbf{e}''_1(s_0)$ are linearly dependent. In this case the cone’s tangent plane $T(s_0) = [\mathbf{e}_1(s_0), \mathbf{e}'_1(s_0)]$ has second order contact with the curve \mathbf{e}_1 . The generator $R(s_0)$ is called an *inflection generator*.

If the zero at $s = s_0$ of the function $k(s)$ has a finite order and this order is odd, then the positive part ($v > 0$) of the cone locally is contained in two different half-spaces bounded by $T(s)$ for $s < s_0$ and $s > s_0$. \diamond

Remark 5.3.16. A simple geometric interpretation of $|k|$ for $k \neq 0$ may be found as follows. Consider the osculating circle $c(s_0)$ of the curve $\mathbf{e}_1(s)$ at $s = s_0$ and project it from the origin \mathbf{o} by a cone of revolution Γ . Γ has aperture angle 2ω (see Fig. 5.28). The radius $\rho_c = \sin \omega$ of $c(s_0)$ equals $1/\kappa_c$, where

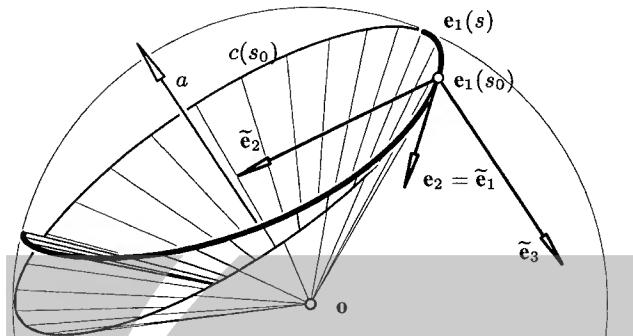


Fig. 5.28. Geometric interpretation of the conical curvature (cf. Remark 5.3.16).

$$|k| = \kappa_c \cos \omega = \frac{\cos \omega}{\rho_c} = \cot \omega. \quad (5.70)$$

By construction, the Darboux vector $\mathbf{d} = \tau \mathbf{e}_1 + \kappa \mathbf{e}_3$ spans Γ 's axis a .

The cone Γ is in second order contact with the original cone, because the director curves c and e_1 are in second order contact, and they share a common constant director curve (the origin). Γ is called *osculating cone of revolution*. It is a counterpart of the osculating circle of a curve. \diamond

The Conical Curvature of a Ruled Surface

Let us now return to a C^2 ruled surface with regular striction curve. We compute the conical curvature of its director cone. The Darboux frame of the director cone obviously equals the Sannia frame of the ruled surface, but they have different canonical parametrizations: The former is parametrized with the arc length of \mathbf{e}_1 , and the latter with the arc length of the striction curve.

Lemma 5.3.14. *If \mathcal{R} is a ruled surface with regular striction curve, with curvature κ and torsion τ , and k is the conical curvature of its director cone, then $k = \tau/\kappa$.*

Proof. We denote the arc length differential of \mathbf{e}_1 by ds_1 and use (5.68), (5.54), and (5.56) to compute

$$-k\mathbf{e}_2 = \frac{d\mathbf{e}_3}{ds_1} = \frac{d\mathbf{e}_3}{du} \frac{du}{ds_1} = -\tau \mathbf{e}_2 \frac{1}{\kappa}. \quad \square$$

Thus the quotient τ/κ is also called the *conical curvature* of the ruled surface \mathcal{R} .

Remark 5.3.17. Consider the osculating cone $\Gamma(u)$ at parameter value u of the director cone of \mathcal{R} . The spherical motion of the Sannia frame has, by (5.58), the instantaneous axis $\mathbf{d}(u) = \tau(u)\mathbf{e}_1(u) + \kappa(u)\mathbf{e}_3(u)$. It coincides with the axis of $\Gamma(u)$. \diamond

Example 5.3.5. The director cone of a ruled surface of *constant conical curvature* $k \neq 0$ is a cone of revolution. Therefore this ruled surface is a *surface of constant slope* (cf. Ex. 5.3.3).

Constant zero conical curvature means that the director cone is a plane, and the surface is conoidal (cf. Ex. 5.3.1). \diamond

Example 5.3.6. Consider the tangent surface \mathcal{T} whose point set is parametrized by $\mathbf{x}(u, v) = \mathbf{s}(u) + vs'(\mathbf{u})$. We will show later (see Th. 6.1.4) that, in general, for a generator $T(u)$ there is an *osculating cone of revolution* with vertex $\mathbf{s}(u)$, denoted by Γ_u , which is in second order contact with \mathcal{T} in the regular points of $T(u)$. It is parallel to the osculating cone of revolution of the director cone (cf. Fig. 5.28). Its axis is parallel to the Darboux vector $\mathbf{d}(u)$ (see Equ. (5.58)). Γ_u becomes planar if $k(u) = 0$, which happens if $\mathbf{s}(u)$ is a point with stationary osculating plane.

We sketch an idea how to prove that Γ_u is indeed in second order contact with \mathcal{T} . Consider the rectifying developable \mathcal{R} of the curve \mathbf{s} . The rectifying plane at $\mathbf{s}(u)$ has the equation

$$(\mathbf{x} - \mathbf{s}(u)) \cdot \mathbf{e}_2(u) = 0.$$

To compute the envelope of the rectifying planes, we compute the equation of the derivative plane:

$$(\mathbf{x} - \mathbf{s}(u)) \cdot (-\kappa \mathbf{e}_1 + \tau \mathbf{e}_3) = 0.$$

These two planes intersect in the generator $R(u)$ of \mathcal{R} . We find that the generator $R(u)$ coincides with the axis of Γ_u .

It is well known that \mathcal{R} carries the finite principal curvature centers of \mathcal{T} — this is also clear from the intuitive definition of a principal curvature as an ‘intersection point of nearby surface normals’, all of which are contained in the rectifying plane. Thus the principal curvature centers of \mathcal{T} coincide with those of Γ_u , which means second order contact. \diamond

5.4 Numerical Geometry of Ruled Surfaces

5.4.1 Discrete Models and Difference Geometry

In geometric computing, discrete models of curves and surfaces play an increasingly important role. In part this is due to the easy availability of storage capacity: it is no longer necessary to store a surface via the control nets of patches. It is possible to store and process a rather dense polyhedral approximation. Particularly in connection with subdivision algorithms and multiresolution representations, this is an attractive alternative to the freeform surface approach, especially for applications in computer graphics and scientific visualization. This is our motivation for briefly discussing *discrete models of ruled surfaces*.

Such a discrete model consists of a sequence of lines R_0, R_1, \dots , which algorithms of geometric processing are applied to. Invariants of the Euclidean differential geometry of a ruled surface can be approximated by appropriate invariants

of the discrete model. This is the subject of so-called *difference geometry* [173] — this area of classical geometry deserves further development in view of the many applications in computational geometry.

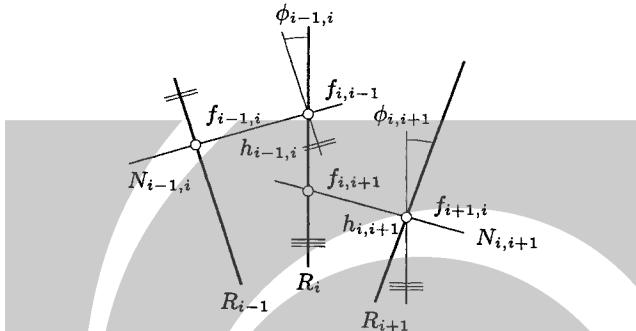


Fig. 5.29. Discrete ruled surface model: Notations.

Discrete Differential Invariants

In order to describe some basics of difference geometry of ruled surfaces we introduce the following notations (see Fig. 5.29). Consecutive lines R_i , R_{i+1} enclose the angle $\phi_{i,i+1}$ and possess the common perpendicular $N_{i,i+1}$. The distance $h_{i,i+1}$ of the lines R_i , R_{i+1} equals the distance of the two footpoints $f_{i,i+1} \in R_i$ and $f_{i+1,i} \in R_{i+1}$. The line R_i is incident with the footpoints $f_{i-1,i}$ and $f_{i,i+1}$.

We further define the discrete distribution parameter

$$\delta_{i,i+1} := \frac{h_{i,i+1}}{\phi_{i,i+1}}. \quad (5.71)$$

We assume that there is a smooth ruled surface with generators $R(t)$ such that $R_i = R(u_i)$ with $u_0 < u_1 < u_2 < \dots$, and study the behaviour of $f_{i,i+1}$, $f_{i+1,i}$, and $h_{i,i+1}$ as u_{i+1} converges to u_i .

Proposition 5.4.1. *Assume that $R(t)$ is the family of rulings of a twice continuously differentiable ruled surface \mathcal{R} , and let $R_i = R(u)$, $R_{i+1} = R(u+h)$. Then*

$$\delta(u) = \lim_{h \rightarrow 0} \delta_{i,i+1}, \quad (5.72)$$

$$s(u) = \lim_{h \rightarrow 0} f_{i,i+1} = \lim_{h \rightarrow 0} f_{i-1,i}, \quad (5.73)$$

$$s(u) + [e_3(u)] = \lim_{h \rightarrow 0} N_{i,i+1}. \quad (5.74)$$

Here $\delta(u)$ is the distribution parameter, $s(u)$ is the striction point, and $s(u) + [e_3(u)]$ is the central tangent (cf. Sec. 5.3).

Proof. We assume that \mathcal{R} is parametrized by $\mathbf{x}(u, v) = \mathbf{s}(u) + v\mathbf{e}(u)$, where $\mathbf{s}(u)$ is the striction line, and that $\|\dot{\mathbf{s}}(u)\| = \|\mathbf{e}(u)\| = 1$. Without loss of generality we let $u = 0$ and let $\mathbf{s}(h) = \mathbf{s}(0) + h\dot{\mathbf{s}}(0) + (h^2/2)\ddot{\mathbf{s}}(0) + o(h^2)$, and $\mathbf{e}(h) = \mathbf{e}(0) + h\dot{\mathbf{e}}(0) + o(h)$. It is a matter of elementary linear algebra to compute $\delta_{i,i+1}$ and $f_{i,i+1}$. We have

$$\begin{aligned} h_{i,i+1} &= \det(\mathbf{s}(h) - \mathbf{s}(0), \mathbf{e}(0)/\|\mathbf{e}(h) \times \mathbf{e}(0)\|, \mathbf{e}(h)), \\ \text{and } \phi_{i,i+1} &= \sphericalangle(\mathbf{e}(h), \mathbf{e}(0)) = h\|\dot{\mathbf{e}}(0)\| + o(h), \end{aligned}$$

which shows that

$$\delta_{i,i+1} = \det(\dot{\mathbf{s}}(0), \mathbf{e}(0), \dot{\mathbf{e}}(0)) + o(1). \quad (5.75)$$

Analogously, we compute $f_{i,i+1}$. If $f_{i,i+1} = \mathbf{s}(0) + f(h)\mathbf{e}(0)$, then

$$f(h) = -(h/2)\mathbf{e}(0)\dot{\mathbf{s}}(0) + o(h). \quad (5.76)$$

Equ. (5.75) and (5.45) show (5.72). Equ. (5.76) shows that $\lim_{h \rightarrow 0} f(h) = 0$, which implies (5.73). Equ. (5.74) is a consequence of (5.73), because the line $N_{i,i+1}$ must now converge to a surface tangent orthogonal to the ruling $R(u)$. \square

Remark 5.4.1. To show Equ. (5.72) we could also use a geometric argument: The ratio $h_{i,i+1}/\phi_{i,i+1}$ is the parameter of the helical motion with axis $N_{i,i+1}$, which takes R_i to R_{i+1} . In the limit we obtain the parameter of a helical motion whose axis is the central tangent and whose velocity vectors in points of R_i are tangent to \mathcal{R} . By Remark 5.3.4, this is the distribution parameter, which shows (5.72) (cf. also Remark 5.3.3). \diamond

Proposition 5.4.2. *We use the notation of Prop. 5.4.1. The convergence of the following limits is quadratic:*

$$\mathbf{s}(u) = \lim_{h \rightarrow 0} \frac{1}{2}(f_{i,i+1} + f_{i,i-1}), \quad \delta(u) = \lim_{h \rightarrow 0} \frac{1}{2}(\delta_{i-1,i} + \delta_{i,i+1}). \quad (5.77)$$

Proof. We use the notations of the proof of Prop. 5.4.1. Equations (5.75) and (5.76) show that both $\delta_{i,i+1}$ and $f_{i,i+1}$ are smoothly dependent on h . We write $\delta(h)$ for $\delta_{i,i+1}$. Then $\delta_{i,i+1} = \delta(h) = \delta(0) + h\dot{\delta}(0) + o(h)$, and $\delta_{i-1,i} = \delta(-h) = \delta(0) - h\dot{\delta}(0) + o(h)$. This shows that

$$\frac{1}{2}(\delta_{i-1,i} + \delta_{i,i+1}) = \delta(0) + o(h).$$

which means that the convergence of the left hand side of this equation towards $\delta(0)$ is quadratic in h . The same argument applies for $f(h)$, which completes the proof. \square

Because of the quadratic convergence of the midpoint of the two footpoints $f_{i,i+1}$ and $f_{i,i-1}$ we use the sequence of these midpoints as a discrete model for the striction curve.

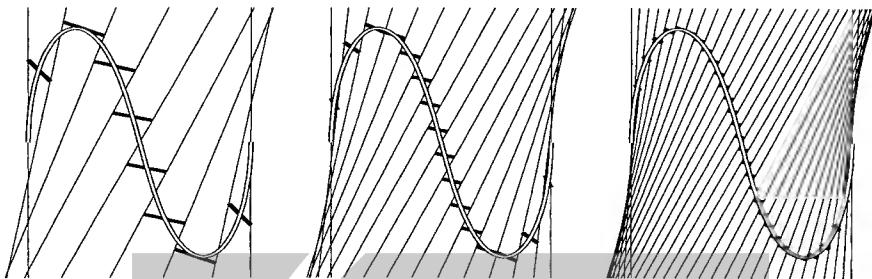


Fig. 5.30. Discretized ruled surface with striction curve and footpoints $f_{i,i-1}$ and $f_{i,i+1}$.

Example 5.4.1. Fig. 5.30 shows three discrete ruled surfaces. We choose one of the two reguli contained in the hyperboloid

$$4x^2 + y^2 - z^2 = 4,$$

discretize it and show the line segments $f_{i,i+1}, f_{i+1,i}$. The fineness of the discretization increases from left to right. The first order convergence to the striction curve (also shown) is clearly visible, and obviously the polygon of midpoints $(f_{i,i+1} + f_{i,i-1})/2$ is a far better approximation of the striction curve than either footpoint polygon. ◇

Other invariants of ruled surfaces can also be expressed by means of the discrete model; the interested reader may find this in the monograph on difference geometry by R. Sauer [173].

Isometric Mapping of Ruled Surfaces

The difference geometric approach also nicely visualizes *isometric mappings* of ruled surfaces. Consider a discrete ruled surface R_0, \dots and a polygon p_0, \dots with vertices $p_i \in R_i$, which should be a discrete curve contained in the ruled surface.

If a sequence of discrete models has been obtained by discretizing an actual smooth ruled surface $R(t)$ and a rectifiable curve $c(t)$ in it, i.e., $R_i = R(t_i)$, $p_i = c(t_i)$, then the arc length of the polygon p_0, p_1, \dots converges to the arc length of c , if $t_{i+1} - t_i$ tends to zero uniformly.

Mappings of one discrete model onto another, which preserve the edge lengths of all polygons are therefore discrete versions of isometric mappings of ruled surfaces. Such transformations are constructed as follows: We consider a pair R_i, R_{i+1} of consecutive rulings as a rigid body $S_{i,i+1}$, and the common perpendicular $N_{i,i+1}$ is assumed to be part of this rigid body. We think of the ruling R_i as a hinge which connects $S_{i-1,i}$ and $S_{i,i+1}$. In this way these two adjacent bodies can rotate freely about the line R_i (Fig. 5.31).

We obtain a kinematic chain (i.e., a mechanism) whose internal degree of freedom equals the number of hinges, i.e., the number of rulings R_i minus the first and the last one. All transformations of discrete models which are composed of such a sequence of rotations about the lines R_i preserve the arc length of all discrete polygons p_0, p_1, \dots with $p_i \in R_i$. We therefore use these mappings as discrete models of the Minding isometries of ruled surfaces which have been mentioned in Remark 5.3.11.

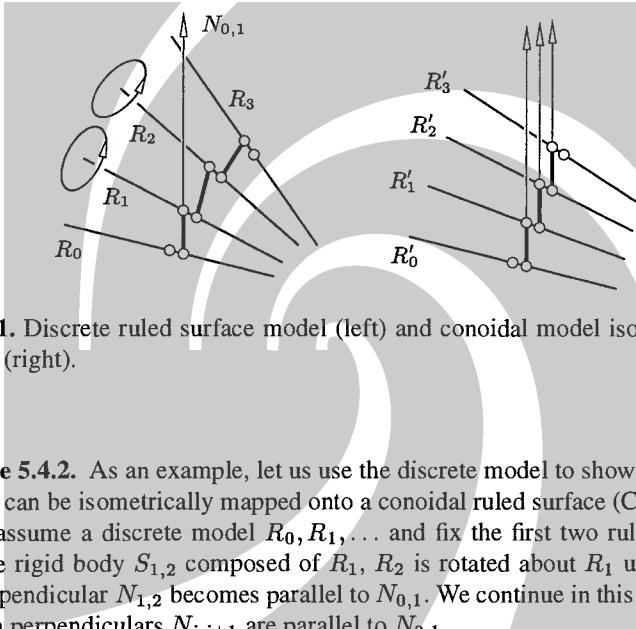


Fig. 5.31. Discrete ruled surface model (left) and conoidal model isometric to the first one (right).

Example 5.4.2. As an example, let us use the discrete model to show that all ruled surfaces can be isometrically mapped onto a conoidal ruled surface (Cor. 5.3.11).

We assume a discrete model R_0, R_1, \dots and fix the first two rulings R_0, R_1 . Then the rigid body $S_{1,2}$ composed of R_1, R_2 is rotated about R_1 until the common perpendicular $N_{1,2}$ becomes parallel to $N_{0,1}$. We continue in this way, until all common perpendiculars $N_{i,i+1}$ are parallel to $N_{0,1}$.

The new discrete model $R'_0 = R_0, R'_1 = R_1, R'_2, \dots$ has the property that all its rulings R'_i are orthogonal to $N_{0,1}$, and therefore parallel to R_0 and R_1 . This means that R'_0, R'_1, \dots is a discrete model of a conoidal surface (see Fig. 5.31).

It is not difficult to show that if the original rulings R_i come from an actual smooth ruled surface \mathcal{R} , and we use finer and finer discrete models, this procedure actually converges to a conoidal surface which is isometric to \mathcal{R} .

Fig. 5.32, left, shows an image of an actual kinematic chain which is a discrete model of a regulus with rotational symmetry. The right hand part of this figure shows an ‘isometric’ position of this kinematic chain (which is no conoidal surface), but which is a discrete ruled surface of constant slope and straight striction polygon. ◇

Remark 5.4.2. If we apply the procedure described in Ex. 5.4.2 to a developable surface, we get its development.

Thus we may compute an ‘approximate development’ of a skew ruled surface, which is desirable in certain applications as follows: First compute the isometric

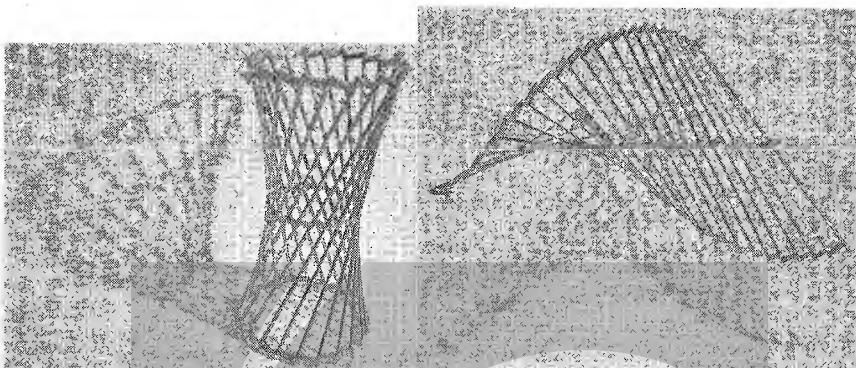


Fig. 5.32. Left: Discrete regulus with rotational symmetry. Right: Isometric discrete surface with straight striction polygon.

conoidal surface (e.g., by using a discrete model) and then project orthogonally onto a plane parallel to the rulings.

This mapping is isometric if restricted to any of the rulings. Using difference geometry, many variants and improvements of such an approximate development are possible (for a discussion of the approximate development of ruled surface strips near their striction curve, see G. Aumann [5]). \diamond

Subdivision Algorithms

A subdivision algorithm computes a refined discrete model from a coarser one. We are going to study subdivision algorithms which converge to smooth ruled surfaces.

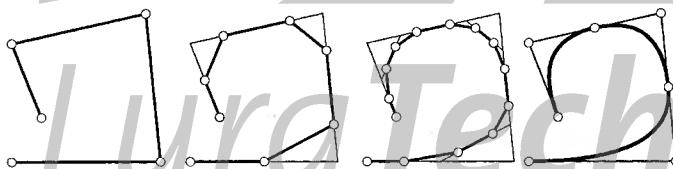


Fig. 5.33. Chaikin's algorithm.

Example 5.4.3. Let us illustrate this general principle by means of a simple subdivision algorithm for polygons: *Chaikin's algorithm* [22], which is illustrated in Fig. 5.33, produces a sequence of polygons which converge to a smooth curve: For a polygon P_0, \dots, P_n , we construct the polygon

$$\begin{aligned}
 \mathbf{p}'_0 &= \mathbf{p}_0, & \mathbf{p}'_1 &= \frac{1}{2}(\mathbf{p}_0 + \mathbf{p}_1), \\
 \mathbf{p}'_{2i} &= \frac{1}{4}(3\mathbf{p}_i + \mathbf{p}_{i+1}), & \mathbf{p}'_{2i+1} &= \frac{1}{4}(\mathbf{p}_i + 3\mathbf{p}_{i+1}), \quad (1 < i < n-1) \\
 \mathbf{p}'_{2n-2} &= \frac{1}{2}(\mathbf{p}_{n-1} + \mathbf{p}_n), & \mathbf{p}'_{2n-1} &= \mathbf{p}_n.
 \end{aligned} \tag{5.78}$$

Iteration of this subdivision procedure yields a sequence of polygons which converges to a quadratic B-spline curve with uniform knot vector and control points $\mathbf{p}_0, \dots, \mathbf{p}_n$ (Fig. 5.33). Note that the limit curve touches the interior edges of the control polygon in their midpoints.

The algorithms of de Casteljau and de Boor are subdivision algorithms as well. Their repeated application yields a sequence of polygons which converges towards a Bézier curve or B-spline curve, respectively (cf. Remark 1.4.1 and [78]). \diamond

Example 5.4.4. Another subdivision scheme has been found by Dyn, Levin and Gregory [38]. We fix a real number w which acts as a shape parameter. From four consecutive points $\mathbf{p}_{i-1}, \dots, \mathbf{p}_{i+2}$ of a polygon we compute the point $\mathbf{p}_{i+1/2}$ according to

$$\mathbf{p}_{i+1/2} = \left(\frac{1}{2} + w\right)(\mathbf{p}_i + \mathbf{p}_{i+1}) - w(\mathbf{p}_{i-1} + \mathbf{p}_{i+2}). \tag{5.79}$$

The refined polygon then has vertices $\dots, \mathbf{p}_i, \mathbf{p}_{i+1/2}, \mathbf{p}_{i+1}, \dots$

Dyn et al. have shown that the subdivision scheme produces polygons which converge to a C^1 curve for $0 < w < \frac{1}{8}$. An example is shown in Fig. 5.34. By its definition, the limit curve interpolates the vertices of all intermediate polygons obtained via subdivision and hence such schemes are referred to as *interpolatory subdivision schemes*. \diamond

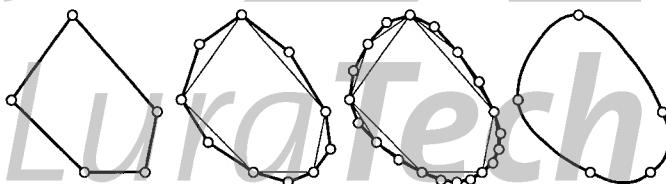


Fig. 5.34. Interpolatory subdivision according to Dyn, Levin and Gregory ($w = 1/13$).

Polygonal subdivision schemes can be used to generate discrete models of ruled surfaces: Given a sequence of line segments $\mathbf{p}_i \mathbf{q}_i$, ($i = 0, \dots, n$), we apply the same subdivision scheme to both boundary polygons $\mathbf{p}_0, \dots, \mathbf{p}_n$ and $\mathbf{q}_0, \dots, \mathbf{q}_n$. Connecting corresponding points of the refined polygon gives a refined sequence of line segments. Smoothness of the limit surface is guaranteed if the two limit polygons are smooth.

If not line segments but entire lines are given, we can apply polygonal subdivision algorithms by selecting appropriate segments, e.g., segments of constant length which lie symmetric to the footpoints $f_{i,i-1}$ and $f_{i,i+1}$.

Variational Interpolatory Subdivision of Ruled Surfaces

Reducing the computation of a surface to the computation of two curves contained in it without looking at the surface itself may not be the best solution. Hence, we propose the following *variational interpolatory subdivision for ruled surface strips* (cf. [132]).

Suppose that R_1, \dots, R_N is a sequence of generators of a ruled surface \mathcal{R} , and that $R_i = \mathbf{a}_i \vee \mathbf{b}_i$. We want to insert a line $R_{i,i+1} = \mathbf{a}'_i \vee \mathbf{b}'_i$ between R_i and R_{i+1} . Suppose we have already defined \mathbf{a}'_i and \mathbf{b}'_i . Then we do the following (see Fig. 5.35, left):

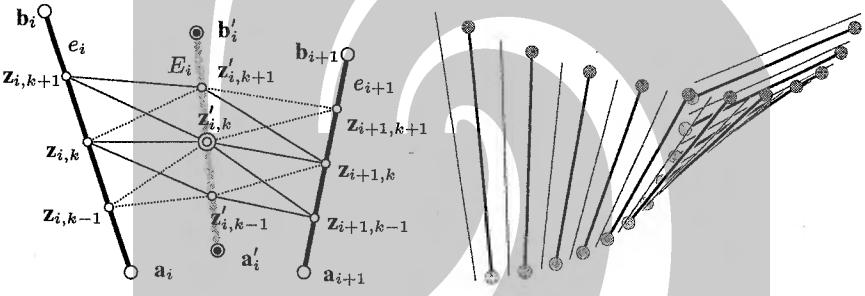


Fig. 5.35. Subdividing a discrete ruled surface. Left: Notation. Right: Result of subdivision. The light grey spheres indicate the points a'_i , the dark grey ones the points b'_i .

Choose an integer $n > 1$ and for $0 \leq k \leq n$ let $\mathbf{z}_{i,k} = (n-k)/n \cdot \mathbf{a}_i + n/k \cdot \mathbf{b}_i$, $\mathbf{z}'_{i,k} = (n-k)/n \cdot \mathbf{a}'_i + n/k \cdot \mathbf{b}'_i$. These points define two different triangulations of the discrete ruled surface. If $\mathbf{z}'_{i,k}$ is no boundary point (i.e., $k \neq 0, n$), it has the six neighbours $\mathbf{z}'_{i,k-1}$, $\mathbf{z}'_{i,k+1}$, $\mathbf{z}_{i,k}$, $\mathbf{z}_{i,k-1}$, $\mathbf{z}_{i+1,k+1}$, $\mathbf{z}_{i+1,k}$ with respect to the first triangulation and $\mathbf{z}'_{i,k-1}$, $\mathbf{z}'_{i,k+1}$, $\mathbf{z}_{i,k}$, $\mathbf{z}_{i,k+1}$, $\mathbf{z}_{i+1,k-1}$, $\mathbf{z}_{i+1,k}$ with respect to the second.

We compute the so-called *umbrella vectors* $\mathbf{u}_{i,k}$ and $\tilde{\mathbf{u}}_{i,k}$ with respect to the first and second triangulation by

$$\begin{aligned}\mathbf{u}_{i,k} &= \mathbf{z}'_{i,k} - \frac{1}{6} (\mathbf{z}'_{i,k-1} + \mathbf{z}'_{i,k+1} + \mathbf{z}_{i,k} + \mathbf{z}_{i,k-1} + \mathbf{z}_{i+1,k+1} + \mathbf{z}_{i+1,k}) \\ \tilde{\mathbf{u}}_{i,k} &= \mathbf{z}'_{i,k} - \frac{1}{6} (\mathbf{z}'_{i,k-1} + \mathbf{z}'_{i,k+1} + \mathbf{z}_{i,k} + \mathbf{z}_{i,k+1} + \mathbf{z}_{i+1,k-1} + \mathbf{z}_{i+1,k}).\end{aligned}$$

These umbrella vectors have been introduced by L. Kobbelt [98]. The norms $\|\mathbf{u}_{i,k}\|$ and $\|\tilde{\mathbf{u}}_{i,k}\|$ are a discrete analogue of the absolute value of the mean curvature in the point $\mathbf{z}'_{i,k}$.

Finally we choose the points \mathbf{a}'_i and \mathbf{b}'_i such that the discrete energy function

$$F = \sum_{i,k} (\mathbf{u}_{i,k}^2 + \tilde{\mathbf{u}}_{i,k}^2). \quad (5.80)$$

is minimized. As F is quadratic, this amounts to the solution of a linear system of equations.

Example 5.4.5. We take twelve generators from the quartic ruled surface $R(u, v) = (1 + vu^2, u^2/2 + uv, u - v)$ and let $n = 12$. The result of one step of variational subdivision is shown by Fig. 5.35, right. \diamond

Remark 5.4.3. We did not prove the convergence of this interpolatory subdivision scheme. This however seems to be irrelevant for practical applications, since after a few iterations employing this variational scheme we can continue with any scheme which is known to be convergent. \diamond

Quasilinear Interpolation of Lines

Certain subdivision algorithms are based on repeated linear interpolation, such as the algorithm of de Casteljau. Extensions of such schemes to lines (not line segments) are possible, if for two lines L_0 and L_1 we define an appropriate family $L(t)$ of ‘linearly interpolating’ lines with $L(0) = L_0$ and $L(1) = L_1$. Two possibilities have been discussed in the literature:

Example 5.4.6. In [60, 166], two lines L_0 and L_1 with Plücker coordinate vectors $(\mathbf{l}_0, \bar{\mathbf{l}}_0)$ and $(\mathbf{l}_1, \bar{\mathbf{l}}_1)$ are interpolated by

$$(\mathbf{l}(t), \bar{\mathbf{l}}(t)) = (1 - t)(\mathbf{l}_0, \bar{\mathbf{l}}_0) + t(\mathbf{l}_1, \bar{\mathbf{l}}_1),$$

which does not define lines unless L_0, L_1 intersect. Besides, this formula is not geometric in the sense that the interpolant does not depend on the given lines, but on their Plücker coordinates in an ambiguous way.

The first problem is overcome by defining $L(t)$ to be the axis of the linear complex \mathcal{C} with equation $\bar{\mathbf{l}}(t) \cdot \mathbf{x} + \mathbf{l}(t) \cdot \bar{\mathbf{x}} = 0$ (i.e., $\mathcal{C}\gamma^* = (\mathbf{l}(t), \bar{\mathbf{l}}(t))\mathbb{R}$) — for the computation of the axis see Th. 3.1.9.

The ‘interpolating lines’ $L(t)$ are therefore the axes of a pencil of linear complexes and thus form the rulings of a Plücker conoid (see p. 180, Fig. 3.10, and the discussion beginning at p. 258).

The second problem can be solved by using normalized Plücker coordinates, or by assigning weights to the given lines. For details, we refer the reader to [60]. \diamond

Example 5.4.7. Another idea is the following: Consider the helical motion α which has the smallest possible angle, whose axis is the common perpendicular of L_0 and L_1 , and which maps L_0 to L_1 .

The corresponding one-parameter subgroup $\alpha(t)$ of helical motions with $\alpha(1) = \alpha$ generates a helical surface with rulings $L(t) = \alpha(t)(L_0)$. This surface joins L_0 and L_1 . This nonlinear interpolation method has been used in [187]. \diamond

Remark 5.4.4. Convergence analysis of subdivision schemes based on nonlinear interpolation is difficult.

Spherical interpolation by the arc length of great circles is a prominent and as yet unsolved example of this. It is also part of the scheme described in Ex. 5.4.7. Its extension to the three-sphere of \mathbb{R}^4 has been used for motion design and computer animation (cf. Chap. 8). \diamond

5.4.2 Interpolation and Approximation Algorithms

Interpolating Line Segments

Consider a sequence of oriented line segments $\overrightarrow{\mathbf{p}_1 \mathbf{q}_1}, \dots, \overrightarrow{\mathbf{p}_N \mathbf{q}_N}$. We want to interpolate or approximate this sequence by a ruled surface strip of the form

$$\mathbf{x}(u, v) = (1 - v)\mathbf{p}(u) + v\mathbf{q}(u), \quad v \in [0, 1], u \in [a, b]. \quad (5.81)$$

Interpolation means that

$$\mathbf{p}(u_i) = \mathbf{p}_i, \mathbf{q}(u_i) = \mathbf{q}_i, \quad i = 1, \dots, N.$$

We identify the oriented line segment $\overrightarrow{\mathbf{p} \mathbf{q}}$ with the element $(p_1, p_2, p_3, q_1, q_2, q_3) \in \mathbb{R}^6$ (cf. Equ. (4.31)). Then the input line segments are points $(\mathbf{p}_i, \mathbf{q}_i) \in \mathbb{R}^6$, the ruled surface strip (5.81) is identified with a curve $(\mathbf{p}(u), \mathbf{q}(u))$, and the interpolation problem is nothing but an ordinary curve interpolation problem in \mathbb{R}^6 , which can be solved with standard methods. For example, we could compute a C^2 cubic spline interpolant using a chordal parametrization based on the distance function (4.32):

$$u_{i+1} = u_i + d(\overrightarrow{\mathbf{p}_i \mathbf{q}_i}, \overrightarrow{\mathbf{p}_{i+1} \mathbf{q}_{i+1}}). \quad (5.82)$$

For unevenly distributed data, the so-called centripetal parametrization of E.T.Y. Lee [105],

$$u_{i+1} = u_i + \sqrt{d(\overrightarrow{\mathbf{p}_i \mathbf{q}_i}, \overrightarrow{\mathbf{p}_{i+1} \mathbf{q}_{i+1}})}. \quad (5.83)$$

should give better results.

Remark 5.4.5. If we use a linear interpolation method, such as cubic spline interpolation, we could apply it to any subset of the six coordinates in \mathbb{R}^6 separately. Especially we could interpolate the points \mathbf{p}_i by a curve $\mathbf{p}(u)$, and the points \mathbf{q}_i by a curve $\mathbf{q}(u)$. The interpretation of a line segment as a point in \mathbb{R}^6 is necessary only for the choice of the parameters u_i . \diamond

Approximating Line Segments

To approximate a sequence of given line segments $(\mathbf{p}_1, \mathbf{q}_1), \dots, (\mathbf{p}_N, \mathbf{q}_N)$, we basically do the same as for interpolation: After identification of a line segment with a point of \mathbb{R}^6 , the present approximation problem reduces to a curve approximation problem in \mathbb{R}^6 . Fitting a B-spline curve of degree k in 6-space to the given points

$(\mathbf{p}_i, \mathbf{q}_i)$ corresponds to fitting a B-spline tensor product patch of degree $(1, k)$ to the given line segments. Standard curve approximation techniques including parameter correction can be applied (see e.g., [78]), but distances and angles should preferably be measured according to the inner product (4.33).

Piecewise Quadratic Ruled Surfaces

The previous algorithms for fitting a ruled surface to given line segments leads to surfaces whose degree is higher than necessary. Let us look at the simplest example: For a G^1 curve approximant we would use quadratic splines. They correspond — under the interpretation of a line segment as a point of \mathbb{R}^6 , which is described above — to surfaces which are piecewise of degree four, and whose differentiability is also merely G^1 .

It turns out that the class of *quadratic* rational ruled surfaces is flexible enough to allow G^1 spline approximation. We will show how the Klein image can be used to reduce the problem to well known algorithms.

A *quadratic* ruled surface is either a regulus, or a quadratic cone, or the tangent surface of a conic. Its Klein image is a conic contained in the Klein quadric. We first consider skew ruled surfaces, since there the line geometric approach is the most elegant one. If two reguli touch each other in the points of a common ruling R_0 , they possess the same contact projectivity there, and their Klein images are conics which touch each other in the point $R_0\gamma$.

We see that the design of piecewise quadratic skew G^1 ruled surfaces in three-space is equivalent to the design of quadratic G^1 spline curves contained in the Klein quadric M_2^4 . The carrier planes of the quadratic segments shall not be contained in M_2^4 (i.e., the quadratic segments are actually Klein images of reguli).

Other quadratic spline curves in M_2^4 correspond to surfaces which consist of pieces of quadratic cones and pieces of tangent surfaces of conics. Note that smooth developable surfaces which consist of pieces of quadratic cones need not have a smooth Klein image (see [110]), and the converse of Th. 5.1.8 is not true. For the design of such surfaces, see Chap. 6.

Hermite Interpolation with Biarcs in the Klein Quadric

Originally biarcs are defined as smooth curves which consist of two circular arcs. If we confine ourselves to circular arcs contained in a fixed sphere, we may say that a biarc is a smooth curve which consists of two arcs of planar sections of this sphere. For us a *biarc in a quadric* is a smooth curve which consists of two arcs of planar sections.

A method which allows to approximate a ruled surface \mathcal{R} with rulings $R(u)$ by a piecewise quadratic G^1 ruled surface is the following. We select a number of non-torsal rulings $R_0 = R(u_0), \dots, R_N = R(u_N)$ and compute the tangent behaviour (contact projectivity) there. All pairs R_i, R_{i+1} of consecutive rulings will be interpolated by a pair Q_{2i}, Q_{2i+1} of quadratic ruled surfaces, such that the tangent projectivities of Q_{2i} and \mathcal{R} coincide at R_i , the tangent projectivities of Q_{2i+1} and

\mathcal{R} coincide at R_{i+1} , and such that $\mathcal{Q}_{2i}, \mathcal{Q}_{2i+1}$ are in first order contact somewhere in between.

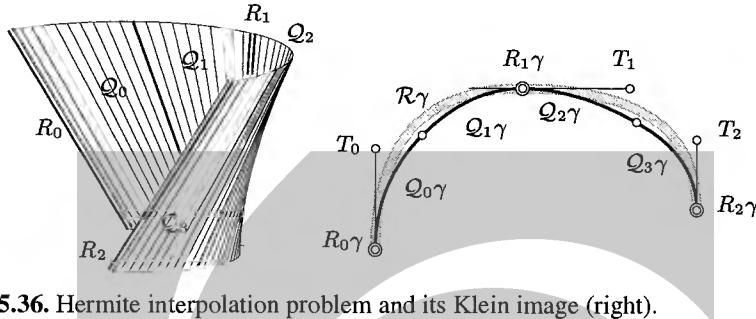


Fig. 5.36. Hermite interpolation problem and its Klein image (right).

In order to achieve this, we look at the Klein image. The curve $\mathcal{R}\gamma$ contains the points $R_i\gamma$, and we assume that the curve tangents in $R_i\gamma$ are spanned by $R_i\gamma$ and a further point T_i .

If the ruled surface \mathcal{R} has director curves $a(u)$ and $b(u)$, we may use the homogeneous coordinate representations $a(u)\mathbb{R}$ and $b(u)\mathbb{R}$ of these director curves to compute the points T_i : From $R(u) = a(u) \vee b(u)$ we have $R(u)\gamma = a(u)\mathbb{R} \wedge b(u)\mathbb{R}$. By differentiation we get the point

$$T_i = (\mathcal{R}\gamma)^1(u_i) = (\dot{a}(u_i) \wedge b(u_i) + a(u_i) \wedge \dot{b}(u_i))\mathbb{R}.$$

We have to join $R_i\gamma$ and $R_{i+1}\gamma$ by a pair $\mathcal{Q}_{2i}\gamma, \mathcal{Q}_{2i+1}\gamma$ of conic segments, such that

1. both segments are contained in the Klein quadric;
2. the line $R_i\gamma \vee T_i$ is tangent to $\mathcal{Q}_{2i}\gamma$ in the point $R_i\gamma$;
3. the line $R_{i+1}\gamma \vee T_{i+1}$ is tangent to $\mathcal{Q}_{2i+1}\gamma$ in the point $R_{i+1}\gamma$; and
4. the conics $\mathcal{Q}_{2i}\gamma$ and $\mathcal{Q}_{2i+1}\gamma$ have a point plus tangent in common.

In the general case, the four points $R_i\gamma, R_{i+1}\gamma, T_i, T_{i+1}$ (the input data) span a three-space $Q^3 \subset P^5$. The intersection $\Phi = Q^3 \cap M_2^4$ is a quadric which must contain both $\mathcal{Q}_{2i}\gamma$ and $\mathcal{Q}_{2i+1}\gamma$, because conics are planar.

In the language of projective three-space, both quadratic ruled surfaces \mathcal{Q}_{2i} and \mathcal{Q}_{2i+1} are contained in a linear line congruence spanned by the input data.

We have now reduced our problem of constructing a biarc in the Plücker quadric to the problem of finding a biarc in the quadric Φ , which is contained in a projective three-space. We show briefly how to solve this problem.

Construction of Biarcs in Quadrics of P^3

The construction of a biarc which joins two pairs of point plus tangent, the segments of which are planar sections of a given quadric Φ , is an elementary one (see [204]).

The input data are the points R_0, R_1 , together with the tangents $R_0 \vee T_0, R_1 \vee T_1$. We use homogeneous coordinates with respect to some projective coordinate system. Then $R_i = r_i\mathbb{R}$, $T_i = t_i\mathbb{R}$. If the quadric Φ has the equation $\mathbf{x}^T \cdot A \cdot \mathbf{x} = 0$, then its polarity κ has the coordinate matrix A and by Prop. 1.1.26, we have

$$\mathbf{r}_i^T \cdot A \cdot \mathbf{r}_i = 0, \quad \mathbf{r}_i^T \cdot A \cdot \mathbf{t}_i = 0, \quad i = 0, 1. \quad (5.84)$$

The conic segments which solve the problem are quadratic rational curves, so we

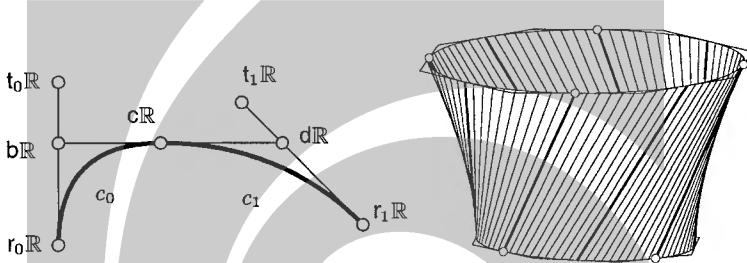


Fig. 5.37. Left: Bézier points for a biarc of conic segments c_0, c_1 . Right: Hermite interpolation with quadratic ruled surfaces (courtesy M. Peternell).

can use a Bézier representation for them. By Th. 1.4.10, the Bézier points of the two segments are $r_0\mathbb{R}, b\mathbb{R}, c\mathbb{R}, d\mathbb{R}, r_1\mathbb{R}$, with yet undetermined weights. They must be such that the point $c\mathbb{R}$ is contained in the quadric Φ :

$$c^T \cdot A \cdot c = 0. \quad (5.85)$$

Further, the inner Bézier points of the two segments have to be contained in the given tangents:

$$b = r_0 + \lambda t_0, \quad d = r_1 + \mu t_1. \quad (5.86)$$

Third, $b\mathbb{R}, c\mathbb{R}, d\mathbb{R}$ should be collinear:

$$c = \alpha b + \beta d,$$

and finally their span is tangent to Φ in the point $c\mathbb{R}$ (see Fig. 5.37, left). This tangency condition is expressed by the fact that the line $b\mathbb{R} \vee d\mathbb{R}$ has exactly one point in common with the quadric Φ : that is, Equ. (5.85), which expands to the quadratic equation $(\alpha b + \beta d)^T \cdot A \cdot (\alpha b + \beta d) = 0$, has a double root. This is equivalent to

$$(b^T \cdot A \cdot d)^2 - (b^T \cdot A \cdot b)(d^T \cdot A \cdot d) = 0.$$

If we insert (5.84) and (5.86), we see that this condition factors into the product of two bilinear equations involving λ and μ :

$$r_0^T \cdot A \cdot r_1 + \lambda t_0^T \cdot A \cdot r_1 + \mu r_0^T \cdot A \cdot t_1 + \lambda \mu (t_0^T \cdot A \cdot t_1 \pm c) = 0, \quad (5.87)$$

with

$$c^2 = (\mathbf{t}_0^T \cdot A \cdot \mathbf{t}_0)(\mathbf{t}_1^T \cdot A \cdot \mathbf{t}_1). \quad (5.88)$$

If we can solve (5.87) for λ and μ , we can determine $\alpha : \beta$ from (5.85).

The solvability of (5.87) depends on the existence of the number c , i.e., depends on whether the right hand side of (5.88) is nonnegative. If we disregard the case that one of $\mathbf{t}_0, \mathbf{t}_1$ is contained in Φ , this means that either both T_0 and T_1 belong to the set $\mathbf{x}^T \cdot A \cdot \mathbf{x} > 0$, or both belong to the set $\mathbf{x}^T \cdot A \cdot \mathbf{x} < 0$.

Remark 5.4.6. In the case of an oval quadric $\Phi : \mathbf{x}^T \cdot A \cdot \mathbf{x} = 0$ the two sets $\mathbf{x}^T \cdot A \cdot \mathbf{x} > 0$ and $\mathbf{x}^T \cdot A \cdot \mathbf{x} < 0$ are the *interior* and *exterior* of the quadric Φ . All tangents of Φ are contained in the exterior, except for the single point which they have in common with the quadric.

If Φ is not oval, these two sets are still the connected components of Φ 's complement, but are indistinguishable by the projective transformation group.

So if Φ happens to be an oval quadric, the right hand side of (5.88) is positive, because T_0 and T_1 were supposed to be contained in tangents of Φ — the biarc problem is always solvable. If Φ is not oval, this need not be the case. \diamond

Remark 5.4.7. If the biarc problem is solvable, it obviously does not have a unique solution. To investigate the manifold of solutions, we first note that we can choose one of the two possible signs in Equ. (5.87). It turns out that this sign determines the ‘sign’ of the first Bézier segment’s initial tangent vector.

If we fix one sign, we see that for all $\lambda \in \mathbb{R} \cup \{\infty\}$ there is a unique $\mu \in \mathbb{R} \cup \{\infty\}$, such that (5.87) is fulfilled, and the bilinearity of this equation means that the mapping $\lambda \mapsto \mu$ (and also the mapping $b\mathbb{R} \mapsto d\mathbb{R}$) is a *projective mapping*.

Therefore, for each choice of a sign, there is a one-parameter family of solution biarcs, and by Prop. 1.1.44, the one-parameter family of contact tangents $b\mathbb{R} \vee d\mathbb{R}$ is a regulus.

By its definition, this regulus is tangent to the quadric Φ . It is well known that the set of points of tangency is a conic itself, which contains the points R_0, R_1 . \diamond

By translating Remark 5.4.7 back into the language of the Hermite interpolation problem for ruled surfaces, we have the following

Corollary 5.4.3. *The problem of Hermite interpolation with reguli Q_{2i}, Q_{2i+1} described on p. 292 has two one-parameter families of solutions, if it is solvable at all. The family of common rulings of Q_{2i}, Q_{2i+1} is a regulus itself, which contains both R_i and R_{i+1} .*

Example 5.4.8. Fig. 5.37 shows the result of a Hermite interpolation by pairs of quadratic ruled surfaces based on biarcs. In this example, $\lambda = \mu$. \diamond

Remark 5.4.8. Approximation of a given curve in a quadric by interpolating discrete points and tangents by biarc segments is, according to Remark 5.4.6, possible if and only if all tangents T_i belong to the same connected component of Φ 's complement. The curve tangent can change from one component to the other, but must be an asymptotic tangent somewhere in between, by continuity. If $\mathcal{R}\gamma$ is the Klein

image of a ruled surface, and Φ is the Klein quadric, asymptotic tangents belong to torsal rulings. So the biarc interpolating problem is solvable for ruled surface segments without torsal rulings. \diamond

5.4.3 Variational Design

Often the solutions of interpolation or approximation problems have to be ‘well behaved’, the meaning of which of course depends on the application one has in mind. In many cases it means the solution should ‘oscillate’ as little as possible, it should be ‘pleasing to the eye’.

The Linearized Thin Plate Functional

These requirements lead to combined problems, where we look for an interpolant/approximant which minimizes a certain regularization functional. For ruled surfaces (5.81), a functional which is not geometric, but is usually a good choice, is the simplified thin plate spline functional

$$F(\mathbf{x}) = \int_a^b \int_0^1 (\mathbf{x}_{uu}^2 + 2\mathbf{x}_{uv}^2 + \mathbf{x}_{vv}^2) du dv. \quad (5.89)$$

It is easy to find surface parametrizations which minimize F . Before we state the result, we make a change of parameters and use the curve $\mathbf{c}(u) := (\mathbf{p}(u) + \mathbf{q}(u))/2$ as directrix. We let $\mathbf{g}(u) = (\mathbf{p}(u) - \mathbf{q}(u))/2$, and re-parametrize (5.81) in the form

$$\mathbf{x}(u, v) = \mathbf{c}(u) + v\mathbf{g}(u), \quad v \in [-1, 1]. \quad (5.90)$$

Proposition 5.4.4. *If a ruled surface strip parametrized by (5.90) interpolates given line segments and minimizes the simplified thin plate spline functional (5.89), then the central curve $\mathbf{c}(u)$ is a cubic C^2 spline curve, and $\mathbf{g}(u)$ is an exponential C^2 spline in tension, whose segments have parametrizations of the form*

$$\mathbf{s}(u) = \mathbf{a}_0 + u\mathbf{a}_1 + \exp(u\sqrt{6})\mathbf{a}_2 + \exp(-u\sqrt{6})\mathbf{a}_3. \quad (5.91)$$

Proof. The thin plate spline functional simplifies to

$$\begin{aligned} F(\mathbf{x}) &= \int_a^b \int_{-1}^1 (\mathbf{x}_{uu}^2 + 2\mathbf{x}_{uv}^2 + \mathbf{x}_{vv}^2) du dv \\ &= 2 \int_a^b \ddot{\mathbf{c}}^2 du + \frac{2}{3} \int_a^b (\dot{\mathbf{g}}^2 + 6\ddot{\mathbf{g}}^2) du = 2\|\ddot{\mathbf{c}}\|_2^2 + \frac{2}{3}T(\mathbf{g}). \end{aligned}$$

The first term in this sum — the squared L^2 norm of $\ddot{\mathbf{c}}$ — is a linearized version of \mathbf{c} 's bending energy. Its minimizers under interpolation constraints are C^2 cubic spline curves (see [78]). The symbol $T(\cdot)$ stands for the tension spline functional (with tension parameter 6) which has been used by D. Schweikert [177] to straighten unwanted undulations in cubic C^2 splines. In that paper it is also shown that the minimizers of $T(\cdot)$ are C^2 spline curves, whose segments are parametrized according to (5.91). \square

The Thin Plate Functional with B-Spline Surfaces

Since the tension spline involves exponential terms, the solution of the combined interpolation problem is not a B-spline surface. If a B-spline representation is essential, e.g., to comply with the needs of a certain software, one uses $F(\mathbf{x})$ as a regularization functional.

Let us briefly outline the case of approximation. Suppose that a ruled surface strip is parametrized by (5.81). If both \mathbf{p} and \mathbf{q} are B-spline curves, we can interpret the pair $X = (\mathbf{p}, \mathbf{q})$ as a B-spline curve in \mathbb{R}^6 :

$$X(u) = (\mathbf{p}(u), \mathbf{q}(u)) = \sum_{i=0}^M N_i^3(u) X_i.$$

Here X_0, \dots, X_M are control points of the curve $X(u)$, with $X_i = (\mathbf{p}_i, \mathbf{q}_i)$, where \mathbf{p}_i and \mathbf{q}_i are the control points of the B-spline curves \mathbf{p} , \mathbf{q} , respectively. The cubic B-spline basis functions N_i^3 are defined by Equ. (1.95).

The regularization functional $F(\mathbf{x})$ then is a quadratic function $F^*(X_0, \dots, X_M)$ of the control points X_i . We choose parameters $u_1 < \dots < u_N$ and $\lambda > 0$, let $P_i = (\mathbf{p}(u_i), \mathbf{q}(u_i))$, and minimize the function

$$G(X_0, \dots, X_M) = \sum_{i=1}^N q(P_i - X(u_i)) + \lambda F^*(X_0, \dots, X_M), \quad (5.92)$$

where $q(\cdot)$ is given by (4.33). The parameter λ determines the influence of the smoothing functional F^* . Since the choice of the parameters usually has a heavy influence on the result, one will start with an initial guess — we might choose the chord length parametrization (5.82) or the centripetal parametrization (5.83) for the polygon P_1, \dots, P_N in \mathbb{R}^6 with respect to the metric (4.33). We subsequently improve the solution by keeping the knot vector fixed and using Hoschek's method [78] in order to iteratively change the parameters u_i until the error vectors are nearly orthogonal to the solution with respect to the metric (4.33).

Choosing Line Segments from Lines

If we have to interpolate or approximate a sequence of *lines* rather than line segments, we can choose appropriate reference points within a region of interest and make line segments from them.

Locally, after segmentation of the input data, we can choose the segments' endpoints in two parallel planes, which leads to a curve design problem in \mathbb{R}^4 with metric (4.24).

In some applications, the choice of appropriate segments of given lines or the choice of parallel reference planes can be done interactively. On the other hand, we could use variational methods to determine these unknown parameters. In view of Prop. 5.4.4, we now consider the problem of finding the 'best' central curve.

A Variational Approach to the Central Curve

We pose the following problem. Given are N lines L_i ($i = 1, \dots, N$) and parameters $u_1 < \dots < u_N$. Which curve $\mathbf{c}(u)$ (contained in a given class of curves) minimizes the linearized bending energy

$$E(\mathbf{c}) = \int_{u_1}^{u_N} \ddot{\mathbf{c}}^2 du \quad (5.93)$$

under the side condition that $\mathbf{c}(u_i) \in L_i$ for $i = 1, \dots, N$? A solution is given by the following theorem:

Theorem 5.4.5. Consider real numbers $u_1 < \dots < u_N$ and lines L_1, \dots, L_N in Euclidean three-space, such that L_i has the equation $\mathbf{x} = \mathbf{p}_i + \lambda \mathbf{g}_i$. Assume that the vectors \mathbf{g}_i are not linearly distributed in the form

$$\mu_i \mathbf{g}_i = \mu_1 \mathbf{g}_1 + (u_i - u_1) \mathbf{e}, \quad i = 2, \dots, N. \quad (5.94)$$

We let $I = [u_1, u_N]$. Then among all smooth curves $\mathbf{x} : I \rightarrow \mathbb{R}^3$ with

$$\dot{\mathbf{x}} \in AC(I), \quad \ddot{\mathbf{x}} \in L^2(I), \quad \mathbf{x}(u_i) \in L_i, \quad (i = 1, \dots, N)$$

there exists a unique curve \mathbf{c} which minimizes the functional E of Equ. (5.93). The segments $\mathbf{c}|[u_i, u_{i+1}]$ are cubic polynomials, \mathbf{c} is actually C^2 and satisfies

$$\begin{aligned} [\mathbf{c}_+^{(3)}(u_i) - \mathbf{c}_-^{(3)}(u_i)] \cdot \mathbf{g}_i &= 0, \quad i = 2, \dots, N-1 \\ \mathbf{c}_+^{(3)}(u_1) \cdot \mathbf{g}_1 &= \mathbf{c}_-^{(3)}(u_N) \cdot \mathbf{g}_N = 0, \\ \ddot{\mathbf{c}}(u_1) &= \ddot{\mathbf{c}}(u_N) = 0. \end{aligned} \quad (5.95)$$

Proof. We first assume that there is a cubic C^2 curve \mathbf{c} which satisfies (5.95) and show that for all admissible curves \mathbf{x} we have $E(\mathbf{x}) \geq E(\mathbf{c})$. We define an inner product of two curves $\mathbf{x}(u), \mathbf{y}(u)$ by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_I \ddot{\mathbf{x}}(u) \cdot \ddot{\mathbf{y}}(u) du. \quad (5.96)$$

Obviously, $\langle \cdot, \cdot \rangle$ is a symmetric positive semidefinite bilinear form which satisfies $\langle \mathbf{x}, \mathbf{x} \rangle = E(\mathbf{x})$. Bilinearity implies that

$$\langle \mathbf{x} - \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle + 2\langle \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{c}, \mathbf{c} \rangle. \quad (5.97)$$

We want to show $\langle \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle = 0$, because then we can use $\langle \mathbf{x} - \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle \geq 0$ to conclude $\langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{c}, \mathbf{c} \rangle \geq 0$. We compute $\langle \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle = \int_I \ddot{\mathbf{c}} \cdot (\ddot{\mathbf{x}} - \ddot{\mathbf{c}}) du$ by integration by parts, using the fact that $\mathbf{c}^{(4)} = 0$ because \mathbf{c} is piecewise cubic:

$$\begin{aligned}
\int_{u_i}^{u_{i+1}} \ddot{\mathbf{c}} \cdot (\ddot{\mathbf{x}} - \ddot{\mathbf{c}}) du &= \ddot{\mathbf{c}} \cdot (\dot{\mathbf{x}} - \dot{\mathbf{c}}) \Big|_{u_i}^{u_{i+1}} - \int_{u_i}^{u_{i+1}} \mathbf{c}^{(3)} \cdot (\dot{\mathbf{x}} - \dot{\mathbf{c}}) du \\
&= \ddot{\mathbf{c}} \cdot (\dot{\mathbf{x}} - \dot{\mathbf{c}}) \Big|_{u_i}^{u_{i+1}} - \mathbf{c}^{(3)} \cdot (\mathbf{x} - \mathbf{c}) \Big|_{u_i}^{u_{i+1}} + \int_{u_i}^{u_{i+1}} \mathbf{c}^{(4)} \cdot (\mathbf{x} - \mathbf{c}) du \\
&= \ddot{\mathbf{c}}(u_{i+1}) \cdot (\dot{\mathbf{x}}(u_{i+1}) - \dot{\mathbf{c}}(u_{i+1})) - \ddot{\mathbf{c}}(u_i) \cdot (\dot{\mathbf{x}}(u_i) - \dot{\mathbf{c}}(u_i)) \\
&\quad - \mathbf{c}_-^{(3)}(u_{i+1}) \cdot (\mathbf{x}(u_{i+1}) - \mathbf{c}(u_{i+1})) + \mathbf{c}_+^{(3)}(u_i) \cdot (\mathbf{x}(u_i) - \mathbf{c}(u_i)).
\end{aligned}$$

Since both $\mathbf{x}(u_i)$ and $\mathbf{c}(u_i)$ are contained in the line L_i , there is a λ_i such that $\mathbf{x}(u_i) - \mathbf{c}(u_i) = \lambda_i \mathbf{g}_i$. We use the fact that \mathbf{c} is C^2 and satisfies (5.95) to verify that $\langle \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle = 0$:

$$\begin{aligned}
\langle \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle &= \underbrace{-\ddot{\mathbf{c}}(u_1) \cdot (\dot{\mathbf{x}}(u_1) - \dot{\mathbf{c}}(u_1))}_{=0} + \sum_{i=2}^{N-1} \underbrace{(\ddot{\mathbf{c}}(u_i) - \ddot{\mathbf{c}}(u_i)) \cdot (\dot{\mathbf{x}}(u_i) - \dot{\mathbf{c}}(u_i))}_{=0} \\
&\quad + \underbrace{\ddot{\mathbf{c}}(u_N) \cdot (\dot{\mathbf{x}}(u_N) - \dot{\mathbf{c}}(u_N))}_{=0} + \lambda_1 \underbrace{\mathbf{c}_+^{(3)}(u_1) \cdot \mathbf{g}_1}_{=0} \\
&\quad + \sum_{i=2}^{N-1} \lambda_i \underbrace{\left(-\mathbf{c}_-^{(3)}(u_i) + \mathbf{c}_+^{(3)}(u_i) \right) \cdot \mathbf{g}_i}_{=0} - \lambda_N \underbrace{\mathbf{c}_-^{(3)}(u_N) \cdot \mathbf{g}_N}_{=0} = 0.
\end{aligned}$$

This shows that $E(\mathbf{c}) \leq E(\mathbf{x})$ for all admissible \mathbf{x} .

To show that the minimum is unique, if it exists, let us assume that $E(\mathbf{c}) = E(\bar{\mathbf{c}})$. Equ. (5.97) implies that $\langle \bar{\mathbf{c}} - \mathbf{c}, \bar{\mathbf{c}} - \mathbf{c} \rangle = 0$, which means that the curve $\bar{\mathbf{c}} - \mathbf{c}$ is piecewise linear. Since $\bar{\mathbf{c}} - \mathbf{c}$ is C^1 , it must be a linear function:

$$(\bar{\mathbf{c}} - \mathbf{c})(u) = (\bar{\mathbf{c}} - \mathbf{c})(u_1) + (u - u_1)\mathbf{e}.$$

Both \mathbf{c} and $\bar{\mathbf{c}}$ must satisfy the incidence conditions, which means $(\bar{\mathbf{c}} - \mathbf{c})(u_i) = \mu_i \mathbf{g}_i$ for $i = 1, \dots, N$. This implies

$$\mu_i \mathbf{g}_i = \mu_1 \mathbf{g}_1 + (u_i - u_1)\mathbf{e},$$

which has been forbidden by our assumptions. So we have finally shown that $\bar{\mathbf{c}} - \mathbf{c}$ is zero and a possible solution is unique.

To prove the *existence* of a solution \mathbf{c} , we note that finding a solution which is C^2 and satisfies Equ. (5.95) amounts to the solution of an inhomogeneous system of linear equations (we will write down this system later). If we can show that the corresponding homogeneous system of equations is uniquely solvable, then so is the inhomogeneous one.

It is easy to see (and follows also from the explicit computation given below) that by modifying the input data by letting $\mathbf{p}_1 = \dots = \mathbf{p}_N = 0$ we get a different system of equations which is just the homogeneous system corresponding to the original one. The modified input lines are incident with the origin, so the constant function $\mathbf{c}(u) = 0$ is a minimizing interpolant with $E(\mathbf{c}) = 0$.

The assumption that (5.94) does not hold is still valid for the modified problem, and therefore it is uniquely solvable, which completes the proof. \square

We see that the solution is a cubic C^2 spline curve with knot vector (u_1, \dots, u_n) and natural end conditions. In addition, Equ. (5.95) shows that for all knots u_i the difference of the left and right third derivatives is orthogonal to the corresponding data line L_i .

Explicit Computations

The explicit *computation* of the minimizing interpolant whose existence and uniqueness has been shown by Th. 5.4.5 can be done as follows. If the spline curve $\mathbf{c}(u)$ is parametrized such that u runs in $[u_1, u_N]$, we introduce a local parameter t in each cubic segment $\mathbf{c}_i = \mathbf{c}|_{[u_i, u_{i+1}]}$ by

$$t = \frac{u - u_i}{\Delta_i}, \quad \text{with } \Delta_i := u_{i+1} - u_i.$$

Then the segment \mathbf{c}_i is parametrized by $t \in [0, 1]$. Its Bézier points are denoted by $\mathbf{b}_{3i}, \dots, \mathbf{b}_{3i+3}$. We want to express these Bézier points by the still unknown values and derivatives

$$\mathbf{x}_i = \mathbf{c}(u_i), \quad \mathbf{m}_i = \frac{d\mathbf{c}}{du}(u_i)$$

of the curve $\mathbf{c}(u)$. Clearly, $\mathbf{b}_{3i} = \mathbf{x}_i$ and $\mathbf{b}_{3i+3} = \mathbf{x}_{i+1}$. By Lemma 1.4.1, we have

$$\frac{d\mathbf{c}_i}{dt}(0) = 3(\mathbf{b}_{3i+1} - \mathbf{b}_{3i}) \text{ and } \frac{d\mathbf{c}_i}{dt}(1) = 3(\mathbf{b}_{3i+3} - \mathbf{b}_{3i+2}).$$

Because $d\mathbf{c}_i/dt = \Delta_i d\mathbf{c}/du$, we have

$$\mathbf{b}_{3i+1} = \mathbf{x}_i + \frac{1}{3} \Delta_i \mathbf{m}_i, \quad \mathbf{b}_{3i+2} = \mathbf{x}_{i+1} - \frac{1}{3} \Delta_i \mathbf{m}_{i+1}. \quad (5.98)$$

By using Lemma 1.4.1 again, we compute the second derivatives

$$\ddot{\mathbf{c}}_+(u_i) = \Delta_i^{-2} \ddot{\mathbf{c}}_i(0) = \Delta_i^{-2} (-6\mathbf{x}_i + 6\mathbf{x}_{i+1} - 4\Delta_i \mathbf{m}_i - 2\Delta_i \mathbf{m}_{i+1}),$$

$$\ddot{\mathbf{c}}_-(u_i) = \Delta_{i-1}^{-2} \ddot{\mathbf{c}}_{i-1}(1) = \Delta_{i-1}^{-2} (6\mathbf{x}_{i-1} - 6\mathbf{x}_i + 2\Delta_{i-1} \mathbf{m}_{i-1} + 4\Delta_{i-1} \mathbf{m}_i).$$

Analogously, the third derivatives are computed as

$$\mathbf{c}_+^{(3)}(u_i) = \Delta_i^{-3} (12\mathbf{x}_i - 12\mathbf{x}_{i+1} + 6\Delta_i \mathbf{m}_i + 6\Delta_i \mathbf{m}_{i+1}),$$

$$\mathbf{c}_-^{(3)}(u_i) = \Delta_{i-1}^{-3} (12\mathbf{x}_{i-1} - 12\mathbf{x}_i + 6\Delta_{i-1} \mathbf{m}_{i-1} + 6\Delta_{i-1} \mathbf{m}_i).$$

After these preparations, we formulate the conditions which, according to Th. 5.4.5, the minimizing interpolant must fulfill. The incidence condition $\mathbf{c}(u_i) \in L_i$ reads

$$\mathbf{x}_i = \mathbf{p}_i + \lambda_i \mathbf{g}_i.$$

The N scalar parameters λ_i and the N first derivative vectors \mathbf{m}_i are unknowns which we are going to determine by a linear system of equations. The condition that $\mathbf{c}(u)$ is C^2 means that $\ddot{\mathbf{c}}_+(u_i) = \ddot{\mathbf{c}}_-(u_i)$ for $i = 2, \dots, N-1$. We get the equations

$$\begin{aligned} \frac{1}{3} \begin{bmatrix} \Delta_i \\ 2(\Delta_{i-1} + \Delta_i) \\ \Delta_{i-1} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{m}_{i-1} \\ \mathbf{m}_i \\ \mathbf{m}_{i+1} \end{bmatrix} + \begin{bmatrix} \Delta_i / \Delta_{i-1} \mathbf{g}_{i-1} \\ (\frac{\Delta_{i-1}}{\Delta_i} - \frac{\Delta_i}{\Delta_{i-1}}) \mathbf{g}_i \\ -\Delta_{i-1} / \Delta_i \mathbf{g}_{i+1} \end{bmatrix} \cdot \begin{bmatrix} \lambda_{i-1} \\ \lambda_i \\ \lambda_{i+1} \end{bmatrix} \\ = \frac{\Delta_i}{\Delta_{i-1}} \Delta \mathbf{p}_{i-1} + \frac{\Delta_{i-1}}{\Delta_i} \Delta \mathbf{p}_i, \quad \text{where } \Delta \mathbf{p}_i = \mathbf{p}_{i+1} - \mathbf{p}_i. \end{aligned}$$

The conditions $\ddot{\mathbf{c}}(u_1) = \ddot{\mathbf{c}}(u_N) = 0$ read as follows:

$$\begin{aligned} 2\Delta_1 \mathbf{m}_1 + \Delta_1 \mathbf{m}_2 + 3\mathbf{g}_1 \lambda_1 - 3\mathbf{g}_2 \lambda_2 &= 3\Delta \mathbf{p}_1, \\ \Delta_{N-1} \mathbf{m}_{N-1} + 2\Delta_{N-1} \mathbf{m}_N + 3\mathbf{g}_{N-1} \lambda_{N-1} - 3\mathbf{g}_N \lambda_N &= 3\Delta \mathbf{p}_{N-1}. \end{aligned}$$

The conditions on the third derivatives at the inner knots u_2, \dots, u_{N-1} are expressed as

$$\begin{aligned} \frac{1}{2} \begin{bmatrix} -\Delta_i^2 \\ \Delta_{i-1}^2 - \Delta_i^2 \\ \Delta_{i-1}^2 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{g}_i \cdot \mathbf{m}_{i-1} \\ \mathbf{g}_i \cdot \mathbf{m}_i \\ \mathbf{g}_i \cdot \mathbf{m}_{i+1} \end{bmatrix} + \begin{bmatrix} -\Delta_i^2 / \Delta_{i-1} \mathbf{g}_{i-1} \cdot \mathbf{g}_i \\ (\frac{\Delta_{i-1}^2}{\Delta_i^2} + \frac{\Delta_i^2}{\Delta_{i-1}^2}) \mathbf{g}_i^2 \\ -\Delta_{i-1}^2 / \Delta_i \mathbf{g}_{i+1} \cdot \mathbf{g}_i \end{bmatrix} \cdot \begin{bmatrix} \lambda_{i-1} \\ \lambda_i \\ \lambda_{i+1} \end{bmatrix} \\ = \frac{\Delta_{i-1}^2}{\Delta_i} \mathbf{g}_i \cdot \Delta \mathbf{p}_i - \frac{\Delta_i^2}{\Delta_{i-1}} \mathbf{g}_i \cdot \Delta \mathbf{p}_{i-1}. \end{aligned}$$

Finally, the conditions on the third derivatives at $u = u_1$ and $u = u_N$ are expressed as

$$\begin{aligned} \Delta_1 \mathbf{m}_1 \cdot \mathbf{g}_1 + \Delta_1 \mathbf{m}_2 \cdot \mathbf{g}_1 - 2\mathbf{g}_1 \cdot \mathbf{g}_2 \lambda_2 + 2\mathbf{g}_1^2 \lambda_1 &= 2\mathbf{g}_1 \cdot \Delta \mathbf{p}_1, \\ \Delta_{N-1} \mathbf{m}_N \cdot \mathbf{g}_N + \Delta_{N-1} \mathbf{m}_{N-1} \cdot \mathbf{g}_N + 2\mathbf{g}_{N-1} \cdot \mathbf{g}_N \lambda_{N-1} - 2\mathbf{g}_N^2 \lambda_N &= 2\mathbf{g}_N \cdot \Delta \mathbf{p}_{N-1}. \end{aligned}$$

These four groups of equations form together a system of $4N$ linear equations in the $4N$ unknown variables $(\lambda_1, \dots, \lambda_N, \mathbf{m}_1, \dots, \mathbf{m}_N)$ — each vector \mathbf{m}_i contains three unknowns. By Th. 5.4.5, this system of equations has a unique solution if the vectors $\mathbf{g}_1, \dots, \mathbf{g}_N$ are such that no equation of the form (5.94) holds.

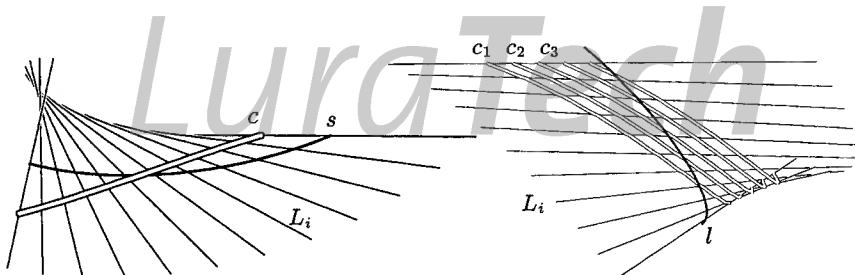


Fig. 5.38. Left: Example of curve c which minimizes the linearized bending energy. s is the striction curve. Right: One-parameter family of curves c_i which minimize the bending energy in case of a linear data arrangement.

Remark 5.4.9. In the case that the vectors $\mathbf{g}_1, \dots, \mathbf{g}_N$ fulfill a relation of the form (5.94), there is a one-parameter family of solution splines.

A simple example of this situation are lines which are contained in one of the two reguli of a hyperbolic paraboloid — all lines L_i are parallel to a plane (cf. Ex. 1.1.31 and Ex. 1.4.7) and can be written in the form $L_i = \mathbf{p}_i \vee \mathbf{q}_i$ with

$$\mathbf{p}_i = \mathbf{a}_i + u_i \mathbf{h}_i, \quad \mathbf{q}_i = \mathbf{b}_i + u_i \mathbf{k}_i.$$

The minimizing interpolants turn out to be piecewise linear (and at the same time C^2), so they must coincide with the lines of the other reguli of the same hyperbolic paraboloid.

Fig. 5.38, right shows a more interesting example. To compute a unique spline curve also in the case of a linear data arrangement, one may fix the starting point $\mathbf{c}(u_1)$ on L_1 . The uniqueness of such a solution is immediately seen from the last part of the proof of Th. 5.4.5. \diamond

Remark 5.4.10. It may be desirable to impose an additional initial and end condition on the minimizing interpolant (e.g., in order to make the system solvable also in the exceptional case, see Remark 5.4.9).

For instance, we might prescribe $\mathbf{c}(u_1)$, i.e., search for the interpolant which minimizes E and starts at a fixed point. A closer look at the proof of Th. 5.4.5 shows that the conditions mentioned by Th. 5.4.5 are the same except for the equation involving the third derivative at u_1 , which is missing. The problem is uniquely solvable for all input data.

If we prescribe the points $\mathbf{c}(u_1)$ and $\mathbf{c}(u_N)$, also the condition which involves the third derivative at $u = u_N$ goes away, and the problem remains uniquely solvable.

A *closed* minimizing interpolant with $L_1 = L_N$ is obtained if we delete all end conditions involving second and third derivatives, and add the conditions that $\dot{\mathbf{c}}(u_1) - \ddot{\mathbf{c}}(u_N) = \ddot{\mathbf{c}}(u_1) - \ddot{\mathbf{c}}(u_N) = (\mathbf{c}_-^{(3)}(u_N) - \mathbf{c}_+^{(3)}(u_1)) \cdot \mathbf{g}_i = 0$. \diamond

Approximation of a Point Cloud by Ruled Surfaces

Because ruled surfaces are generated by a one-parameter family of the simplest curves in space (i.e., lines), they arise in a variety of applications including CAD, architectural design [21], wire electric discharge machining (EDM) [166, 220] and peripheral NC milling with a cylindrical cutter [121].

In some cases the ruled surface itself is the goal, in other cases it arises in a crucial intermediate step — such as in the case of shaping a surface Φ with a cylindrical cutter. There, the axis of the cutting tool traces out an *offset* of Φ , and this offset is a ruled surface (see below).

For all these applications the following question is interesting: Given a surface or a cloud of points, how well may it be approximated by a ruled surface?

Remark 5.4.11. If a cylindrical cutter of radius r and a surface are given, the question of how well this surface can be shaped by the process of peripheral milling is the same as the question how well an offset at distance r can be approximated by

a ruled surface — this ruled surface then is the set of axis positions of the moving tool. \diamond

Chen and Pottmann [26] proposed an algorithm for the solution of this approximation problem which works as follows. In the first step, a tracing algorithm computes a finite number of candidate rulings by intersecting the given point cloud or surface with appropriate planes. In this first step we can already see whether there exists a ruled approximation surface within given tolerances at all.

These lines (or line segments) are then approximated by a ruled surface, as has been outlined above. For details, we refer the reader to [26].

Example 5.4.9. Let us look at an example. The input data are represented by a surface. This surface and the rulings found by step 1 of the algorithm are shown by Fig. C.6a.

Fig. C.6b shows the final approximant computed by a second step, and Fig. C.6c shows the intersection of the ruled approximant with the original surface. We can see how the approximant smoothes out the original surface's oscillations. \diamond

5.4.4 Offset Surfaces and their Applications

For a surface in Euclidean three-space, we consider the set of points whose distance to the surface equals d . The exact definition depends on the type of surface and on the application one has in mind. There are the following possibilities:

1. For any subset M of Euclidean three-space, and any point \mathbf{p} , the distance $d(\mathbf{p}, M)$ is defined by

$$d(\mathbf{p}, M) = \inf_{\mathbf{x} \in M} \|\mathbf{x} - \mathbf{p}\|.$$

Then M_d is the set of points such that $d(\mathbf{p}, M) = d$ (see Fig. 5.39, center).

2. If $\mathbf{x}(u, v)$ parametrizes a regular smooth surface, and $\mathbf{n}(u, v)$ is the unit normal vector field, then the two-sided offset at distance d (see Fig. 5.39, left) is given by

$$\mathbf{x}_d(u, v) = \mathbf{x}(u, v) \pm d\mathbf{n}(u, v). \quad (5.99)$$

For an oriented surface, the one-sided offset is defined analogously, but with a $+$ sign instead of \pm .

3. A surface with a piecewise smooth boundary does not have just two normal vectors $\pm\mathbf{n}(u, v)$ in its boundary points. In regular boundary points, there is a half-circle of unit normal vectors, and in a vertex of the boundary, the set of unit normal vectors is bounded by a spherical polygon. If $\mathbf{x}(u, v)$ is a singular point of the surface, then all vectors orthogonal to both $\mathbf{x}_u(u, v), \mathbf{x}_v(u, v)$ are considered normal vectors, just as in the regular case. The difference is that $\mathbf{x}_u, \mathbf{x}_v$ are parallel to each other or even zero.

The definition of offset, however, remains the same: The offset is the set of all points $\mathbf{x}(u, v) + \mathbf{n}$, where \mathbf{n} is any unit normal vector in $\mathbf{x}(u, v)$ (see Fig. 5.39, right).

We can see this offset also as the envelope of the two-parameter family of spheres with radius d and centers $\mathbf{x}(u, v)$.

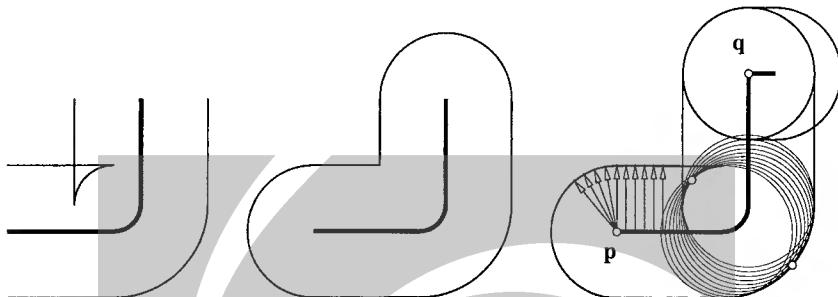


Fig. 5.39. Various definitions of ‘offset’ (see text). **p** is considered a boundary point, **q** is considered a singular point.

Both the first and the second kind of offset are contained in the third one. The difference between the second definition (Equ. (5.99)) and the third definition (envelope of spheres) is a pipe surface corresponding to the surface’s boundary, which plays a role in certain applications.

Offsets appear in various applications, which include geometric modeling of thick plates and NC machining with a ball cutter: in order to manufacture a surface \mathbf{x} with a spherical cutter of radius d , the center of the tool has to run on a one-sided offset surface \mathbf{x}_d .

Offsets of Ruled Surfaces and Peripheral Milling with a Cylindrical Cutter

We continue the discussion of the previous paragraph. If a line R is contained in the surface parametrized by $\mathbf{x}(u, v)$, we consider the spheres of radius d whose center is contained in R . The envelope of these spheres is the cylinder with axis R and radius d . If the original surface is a *ruled* surface with rulings $R(u)$, then obviously its offset is the envelope of the one-parameter family of cylinders with radius d and axis $R(u)$.

This interpretation of the offset of a ruled surface has close connection to certain applications. The first one is peripheral milling with a cylindrical cutter, where the cutter is in line contact with the final surface. The final surface therefore is part of the envelope of all cylinders with axes $R(u)$ and radius d , i.e., part of the offset of a ruled surface [121]. *Line contact* means that two surfaces touch each other in all points of a curve.

Another application is wire cut EDM (electric discharge machining), if we do not neglect the thickness of the cutting wire and model it as a cylinder of revolution [220]. Geometrically, this process is nothing but peripheral milling with a (thin) cylindrical milling tool.

We already mentioned that a moving cylinder of revolution is in line contact with its offset. We are going to describe this line of contact now. It is also called the *characteristic curve* or *characteristic* of the motion of the cylinder.

Assume that the cylinder moves smoothly such that its axis coincides with the rulings $R(u)$ of a ruled surface \mathcal{R} . It is obvious that a point \mathbf{p} of the cylinder is contained in the characteristic, if its velocity vector is tangent to the surface of the cylinder. An equivalent condition (which is more geometric because it filters out the motion parallel to $R(u)$) is that the cylinder's normal in \mathbf{p} is a surface normal of \mathcal{R} . So we can find the line of contact if we consider the surface normals in all points of $R(u)$ and intersect them with the cylinder. The result is the following:

Proposition 5.4.6. *Assume that a cylinder of revolution moves smoothly such that its axis coincides with the rulings $R(u)$ of a ruled surface \mathcal{R} . Then the characteristic curve equals*

- (i) *two lines parallel to $R(u)$ if $R(u)$ is cylindrical;*
- (ii) *two such lines plus a circle in a plane orthogonal to $R(u)$ if $R(u)$ is non-cylindrical and torsal;*
- (iii) *a rational quartic in the case that $R(u)$ is non-torsal.*

Proof. If R is a torsal generator, the surface normals in regular points are parallel. In a cuspidal point, every line orthogonal to $R(u)$ is a surface normal. This shows (i) and (ii).

If $R(u)$ is a non-torsal ruling, Lemma 5.3.1 says that the set of surface normals is a hyperbolic paraboloid A . To intersect A with the cylinder, we use the Sannia frame of Equ. (5.53) as a coordinate system.

If δ is the distribution parameter of $R(u)$, Th. 5.1.4 and Equ. (3.19) show that the surface normal of the point $(\delta \tan \phi, 0, 0)$ equals $(0, \cos \phi, -\sin \phi)$, so the characteristic curve has the parametrization

$$\mathbf{c}(t) = (\delta \tan \phi, d \cos \phi, -d \sin \phi). \quad (5.100)$$

Now (iii) follows from the change of parameters $s = \tan(\phi/2)$:

$$\mathbf{c}(s) = \left(\frac{2\delta s}{1-s^2}, \frac{d(1-s^2)}{1+s^2}, \frac{-2ds}{1+s^2} \right).$$

□

A characteristic which corresponds to a non-torsal generator is shown by Fig. 5.40, left. Note that its rationality has practical implication in NC milling of a ruled surface \mathbf{x} with a ball cutter. If rulings of \mathbf{x} are taken as cutter contact curves, the corresponding path curves of the tool center are rational quartics and thus exactly representable in Bézier form.

Example 5.4.10. If a ruled surface has only torsal rulings, Th. 5.1.7 shows that it locally is either a cylinder, a cone, or the tangent surface of a space curve.

The offsets of cylinders are cylinders again. For a cone, we consider the sphere Σ with radius d whose center coincides with the cone's vertex. The two-sided offset

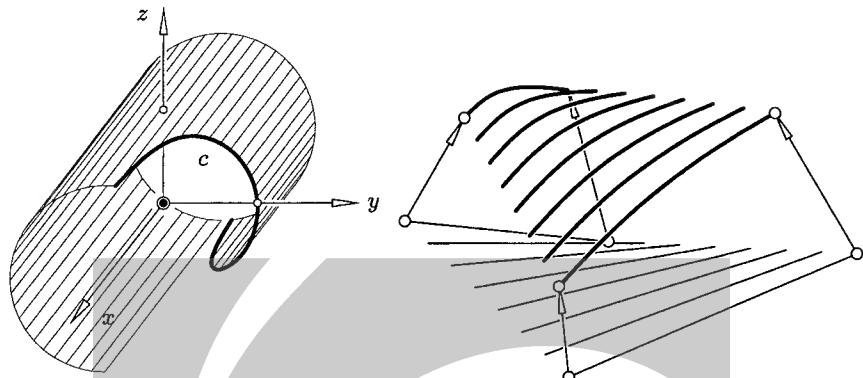


Fig. 5.40. Left: Characteristic curve c of a moving cylinder. Right: Offset of a hyperbolic paraboloid with characteristics.

surface of a cone at distance d is in general composed of a developable surface of two components which touch Σ . Part of Σ also belongs to the offset, if we use the most general definition which includes singular surface points. This part consists of great circles contained in Σ whose planes are orthogonal to the cone's generator lines.

The complete offset of a tangent surface of a space curve l consists (i) of a torsal ruled surface of two components, generated by the two straight lines mentioned in Prop. 5.4.6, and (ii) of a pipe surface of radius d whose spine curve is l , generated by the circles mentioned in Prop. 5.4.6.

The rulings of the torsal parts of the offset are parallel to the corresponding tangents of l , and the circles which generate the pipe surface part are contained in l 's normal planes. \diamond

Remark 5.4.12. Equation (5.100) shows that for a skew ruled surface of *constant* distribution parameter δ the characteristic curves are congruent, and the offset therefore is generated by a Euclidean one-parameter motion of one characteristic.

We also see that the generating motion is the motion of the Sannia frame, whose origin runs along the striction curve. \diamond

Remark 5.4.13. Equ. (5.100) further shows that even in case of non-constant δ , any two characteristic curves are affinely equivalent. The affine map is composed of a Euclidean congruence which maps one Sannia frame onto the other, and a scaling in direction of the ruling. We can therefore say that the offset of a general skew ruled surface is a kinematic surface generated by an affine one-parameter motion of a rational quartic. \diamond

Example 5.4.11. We compute a one-sheeted offset of the hyperbolic paraboloid $\mathbf{x}(u, v) = (u, v, uv)$ at distance 0.5. The unit surface normal vector is given by

$$\mathbf{n}(u, v) = \frac{1}{\sqrt{u^2 + v^2 + 1}}(-v, -u, 1).$$

The offset surface $\mathbf{x}(u, v) + 0.5\mathbf{n}(u, v)$ is shown in Fig. 5.40, right.

Fig. C.11 shows offsets at distances $-2, 2, 4, 6.4$ of the hyperboloid $(x/4)^2 + (y/3)^2 + (z/5)^2 = 1$ together with the characteristics corresponding to one family of rulings of the original hyperboloid. \diamond

Remark 5.4.14. Prop. 5.4.6 is a special case of a more general result: We consider a one-parameter family of quadratic cones or cylinders $\Gamma(t)$, not necessarily congruent, but of course projectively equivalent. This family of surfaces has an envelope Φ . A moving cylinder of revolution is a special case of this. The surface Φ touches the surfaces $\Gamma(t)$ in a curve of *contact points*.

By Prop. 1.1.39 and Ex. 1.1.28, the set $\Gamma^*(t)$ of tangent planes of $\Gamma(t)$ is a conic of dual projective space. In plane coordinates, $\Gamma^*(t)$ is a rational quadratic curve, and can be parametrized in the form

$$\mathbf{u}(s, t) = \mathbf{a}_0(t) + \mathbf{a}_1(t)s + \mathbf{a}_2(t)s^2, \quad (5.101)$$

where s runs in $\mathbb{R} \cup \{\infty\}$. Thus the union of *all* tangent planes is a surface Φ^* , which carries a one-parameter family of conics. By duality, the tangent planes of Φ^* are the contact points of Φ .

The surface Φ^* has the parametrization (5.101), with parameters s and t . To compute its tangent plane, we differentiate with respect to s and t , and compute the plane $\mathbb{R}\mathbf{u} \vee \mathbb{R}\mathbf{u}_s \vee \mathbb{R}\mathbf{u}_t$ by using the cross product (cf. Remark 1.2.12 and Ex. 2.2.8):

$$\mathbf{c}(s, t) = \mathbf{u} \times \frac{\partial \mathbf{u}}{\partial s} \times \frac{\partial \mathbf{u}}{\partial t}.$$

If $\mathbf{u}(s, t)$ is given by (5.101), then $\mathbf{c}(s, t)\mathbb{R}$ is nothing but a parametrization of Φ in point coordinates. For fixed $t = t_0$ we get the curve of contact points which corresponds to the cone $\Gamma(t)$ (i.e., the characteristic curve):

$$\mathbf{c}(s, t_0) = \mathbf{a}_0(t_0) \times \mathbf{a}_1(t_0) \times \dot{\mathbf{a}}_0(t_0) + \dots \quad (5.102)$$

Obviously it is rational and of order at most four. This shows the following result:

Theorem 5.4.7. *A one-parameter family of quadratic cones $\Gamma(t)$ has an envelope with rational characteristic curves of degree ≤ 4 .*

There is also a dual version: A surface generated by a one-parameter family of conics $c(t)$ has the property that the tangent planes in the point of a fixed conic $c(t_0)$ are the planes of a developable surface of class ≤ 4 (see Chap. 6, and, for a complete discussion, W. Degen [34]).

An application of this result is NC milling with a conical cutter. This dual approach to the computation of the envelope for a moving cylinder has been used by Xia and Ge [218] and for more general moving developable surfaces by Jüttler and Wagner [92]. \diamond

Remark 5.4.15. We ask the following question: Is it possible that the offset of a skew ruled surface Φ is again a ruled surface? By Prop. 5.4.6, the rulings on the offset cannot directly correspond to the rulings of the progenitor surface.

Th. 5.1.4 shows that the dual of a skew ruled surface is a skew ruled surface itself — the set of tangent planes in points of a generator is a pencil, i.e., a line of dual space. The offset of Φ is the envelope of a moving cylinder, so, according to Remark 5.4.14, its dual Φ^* carries a one-parameter family of conics. Surfaces Φ^* which are ruled and carry a one-parameter family of conics have been classified by H. Brauner [19]. They are the quadratic and cubic ruled surfaces, and certain rational ruled surfaces of order four. We have encountered all these surfaces in our discussion of low degree rational ruled surfaces. Their duals Φ are also of degree 2, 3 and 4, and by duality they are envelopes of one-parameter families of quadratic cones. Apparently this family of quadratic cones is never a family of congruent cylinders (which means a negative answer to our question), but no proof of this has been published as yet.

In connection with ruled surface offsets we mention that generalized offsets of ruled surfaces have been defined which are always ruled [165]. \diamond

Remark 5.4.16. We have seen the difference between offsets of torsal ruled surfaces and offsets of skew ruled surfaces. This difference also appears when we address the question of rationality of the offsets of a rational ruled surface. It has been shown by Pottmann et al. [152] that the offsets of a skew rational ruled surface always possess a rational parametrization.

The offsets of torsal rational surfaces however need not be rational. This topic is discussed in Sec. 6.3.3. \diamond

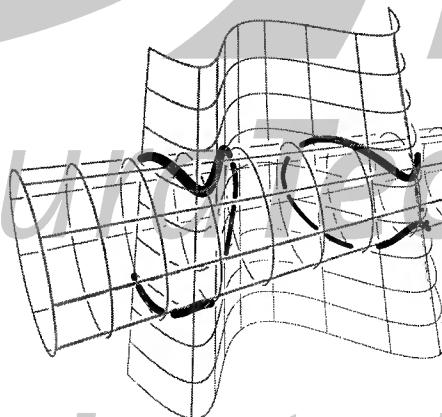


Fig. 5.41. Intersection curve of two ruled surfaces (courtesy G. Elber).

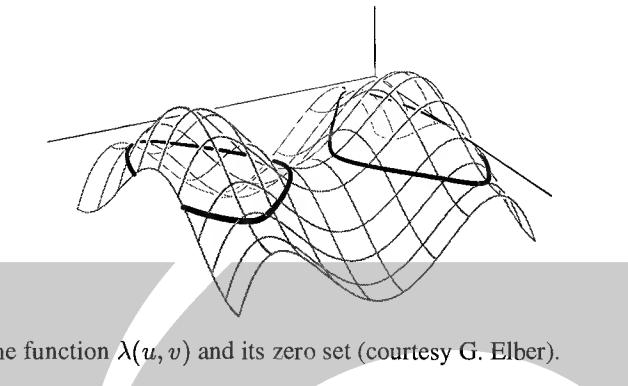


Fig. 5.42. The function $\lambda(u, v)$ and its zero set (courtesy G. Elber).

5.4.5 Intersection of Ruled Surfaces

Consider two ruled surfaces \mathcal{R} and \mathcal{S} , parametrized by

$$\mathcal{R} : \mathbf{x}(u, s) = \mathbf{a}(u) + s\mathbf{g}(u), \quad \mathcal{S} : \mathbf{y}(v, t) = \mathbf{b}(v) + t\mathbf{h}(v),$$

and rulings $R(u)$ and $S(v)$ according to this parametrization. We are interested in the intersection of the two surfaces.

Any intersection point $\mathbf{x}(u, s) = \mathbf{y}(v, t)$ is contained in the rulings $R(u)$ and $S(v)$. The intersection condition for rulings $R(u)$ and $S(v)$ reads

$$\lambda(u, v) := \det(\mathbf{g}(u), \mathbf{h}(v), \mathbf{a}(u) - \mathbf{b}(v)) = 0. \quad (5.103)$$

If we represent lines by their Plücker coordinates

$$\begin{aligned} R(u) : \quad & (\mathbf{g}(u), \bar{\mathbf{g}}(u)) = (\mathbf{g}(u), \mathbf{a}(u) \times \mathbf{g}(u)), \\ S(v) : \quad & (\mathbf{h}(v), \bar{\mathbf{h}}(v)) = (\mathbf{h}(v), \mathbf{b}(v) \times \mathbf{h}(v)), \end{aligned}$$

the intersection condition is expressed by (cf. Equ. (2.26)).

$$\mathbf{g}(u) \cdot \bar{\mathbf{h}}(v) + \mathbf{h}(v) \cdot \bar{\mathbf{g}}(u) = 0.$$

Thus the basic problem is that of finding the zero set of the bivariate function $\lambda(u, v)$. In case of rational ruled surfaces, this amounts to the computation of the zero set of a bivariate polynomial, which is an algebraic curve in the u, v -plane. In general we are unable to compute a simple parametrization for it. The computation can be done numerically in an effective way which uses the B-spline form and subdivision [42].

For a detailed discussion of singular points, redundant solutions and degenerate cases in ruled surface intersection we refer the reader to Heo, Kim and Elber [70]. Figure 5.41 shows a result of their algorithm and Fig. 5.42 illustrates the corresponding function $\lambda(u, v)$ and its zero set.

Elber, Choi and Kim [43] proposed an application of ruled surface intersection in a rendering technique called ‘ruled tracing’. This is a special organization of the traditional ray tracing technique from Computer Graphics and designed for freeform surface rendering.



LuraTech

www.luratech.com

Color Plates

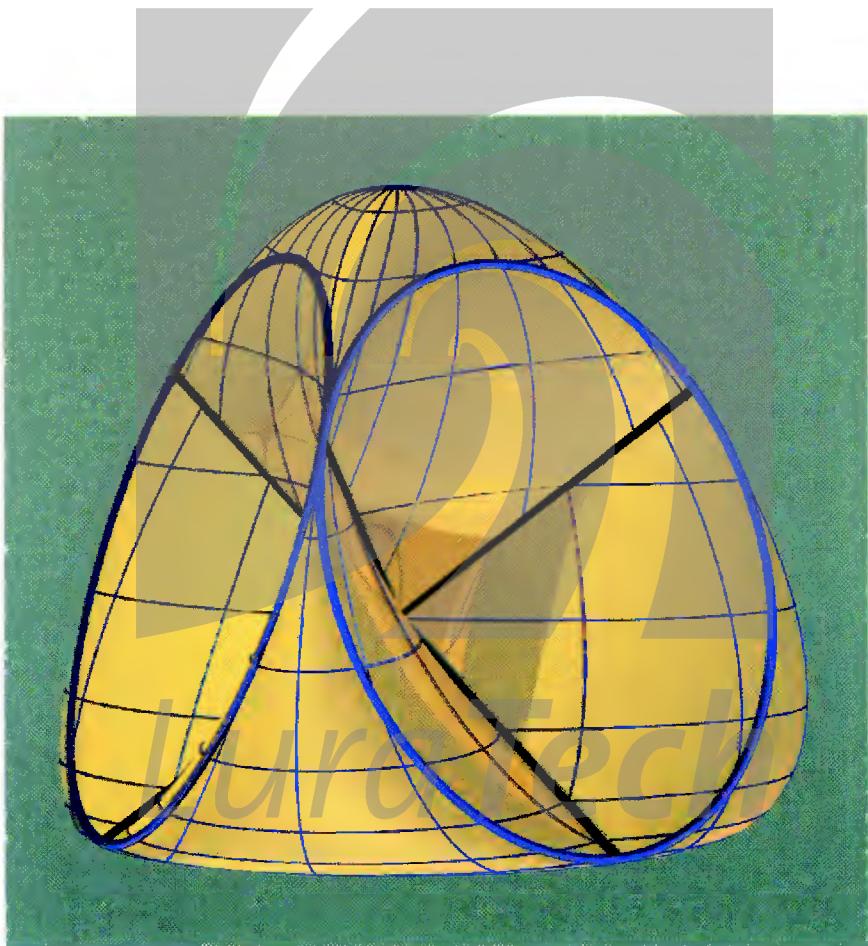


Fig. C.1. Steiner's 'Roman surface' — cf. p. 256 (courtesy H.P. Schröcker).
www.luratech.com

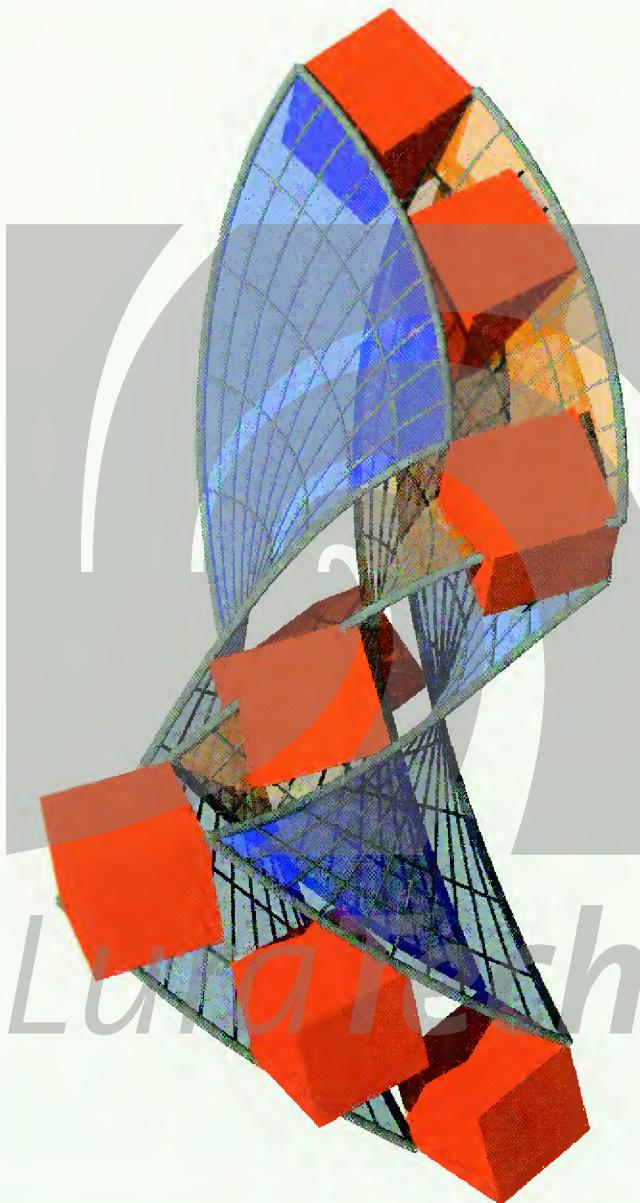
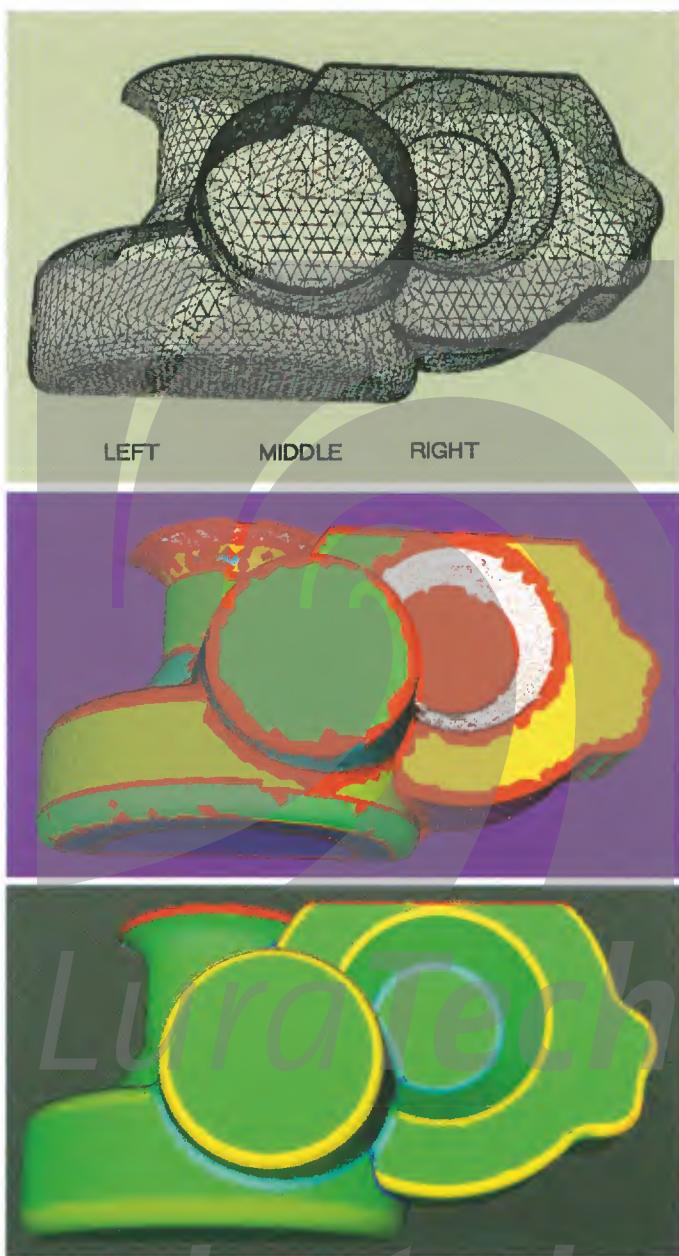


Fig. C.2. Two-parameter motion: Path surfaces and selected positions of moving object — cf. p. 187 (courtesy H.P. Schröcker).

www.luratech.com



www.luratech.com
Fig. C.3. Reverse engineering objects with conventional geometry. Top: triangulated and decimated data set Center: segmented data set Bottom: reconstructed CAD model — cf. Ex. 4.2.3 p. 209 (courtesy T. Várady).



Fig. C.4. Lie quadric of a quartic conoid — cf. p. 237 and p. 247.

www.luratech.com

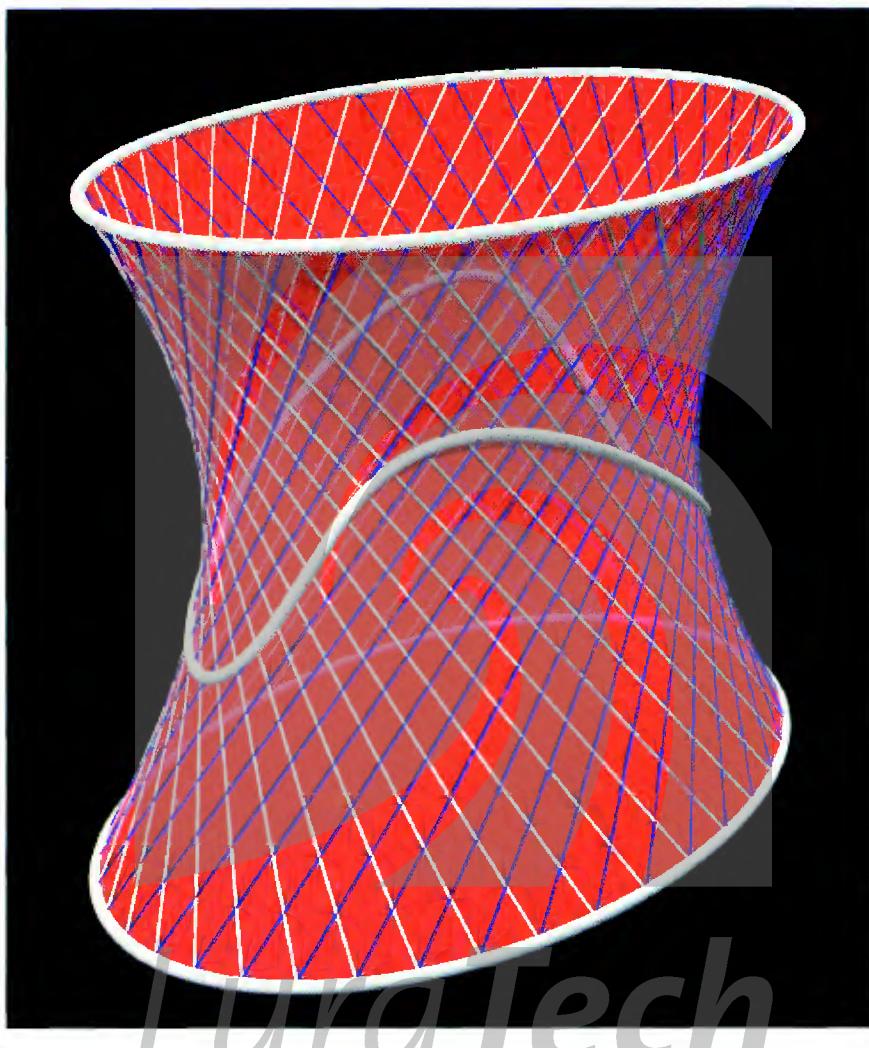


Fig. C.5. Regulus with striction curve — cf. Sec. 5.3.1.

www.luratech.com

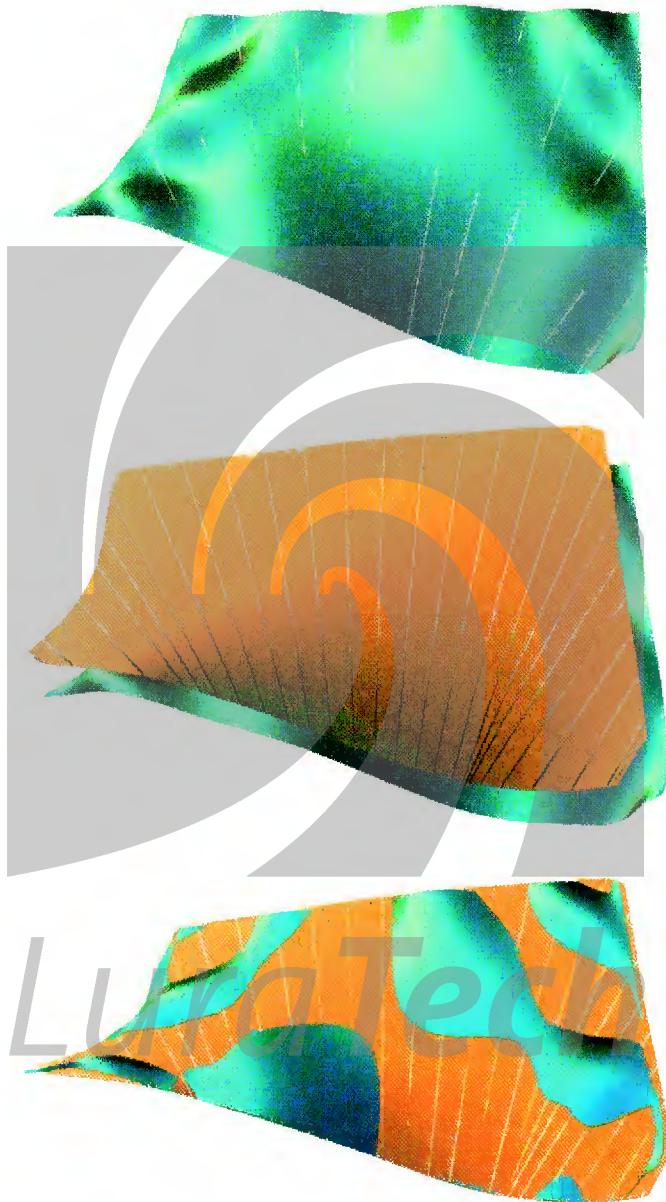


Fig. C.6. (a) Top: Surface with approximating lines; (b) Center: Ruled approximant (elevated); (c) Bottom: Intersection with approximant — cf. p. 303.



Fig. C.7. 'D-forms' — cf. p. 418 (courtesy T. Wills).

www.luratech.com

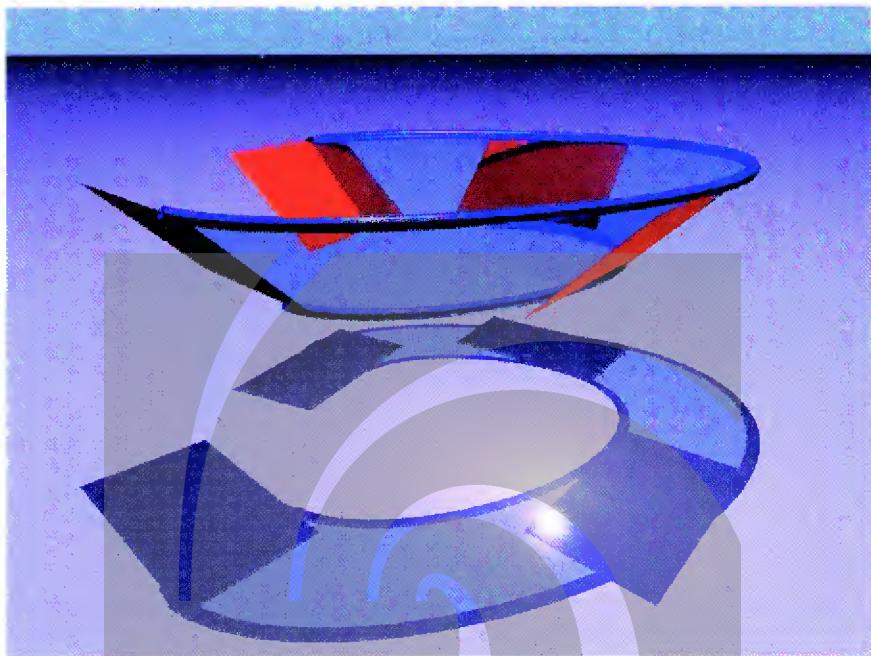


Fig. C.8. Approximation of planes by a developable NURBS surface — cf. p. 356.



Fig. C.9. Milling a star-shaped surface — cf. p. 464 (courtesy G. Glaeser).

www.luratech.com

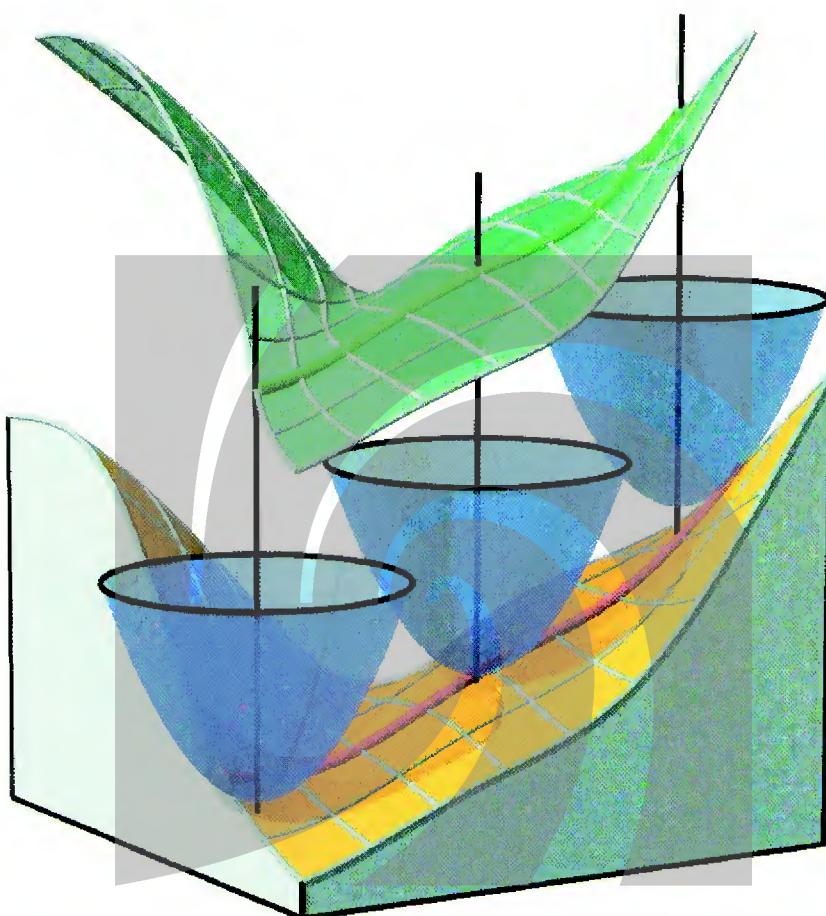


Fig. C.10. Three-axis machining of a surface and general offset surface — cf. p. 452
(courtesy G. Glaeser).

LuraTech
www.luratech.com

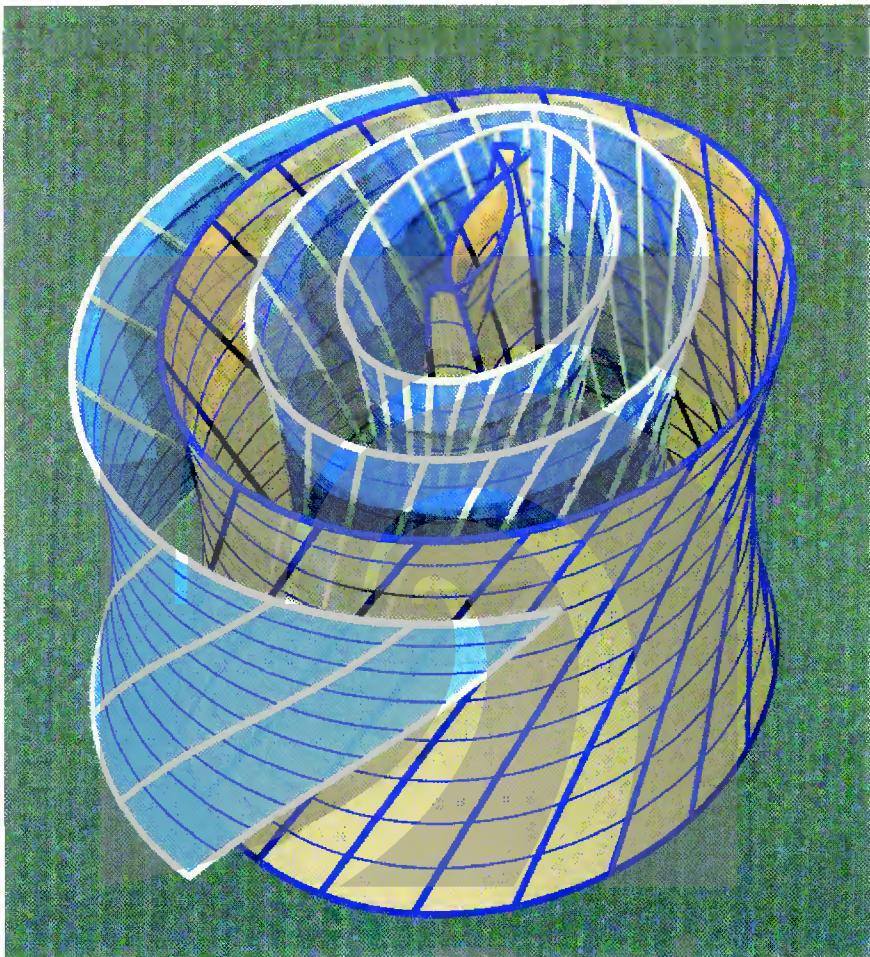


Fig. C.11. Offsets of a one-sheeted hyperboloid. One exterior and several interior offsets, one of them showing singularities — cf. pp. 307, 446 (courtesy H.P. Schröcker).

www.luratech.com

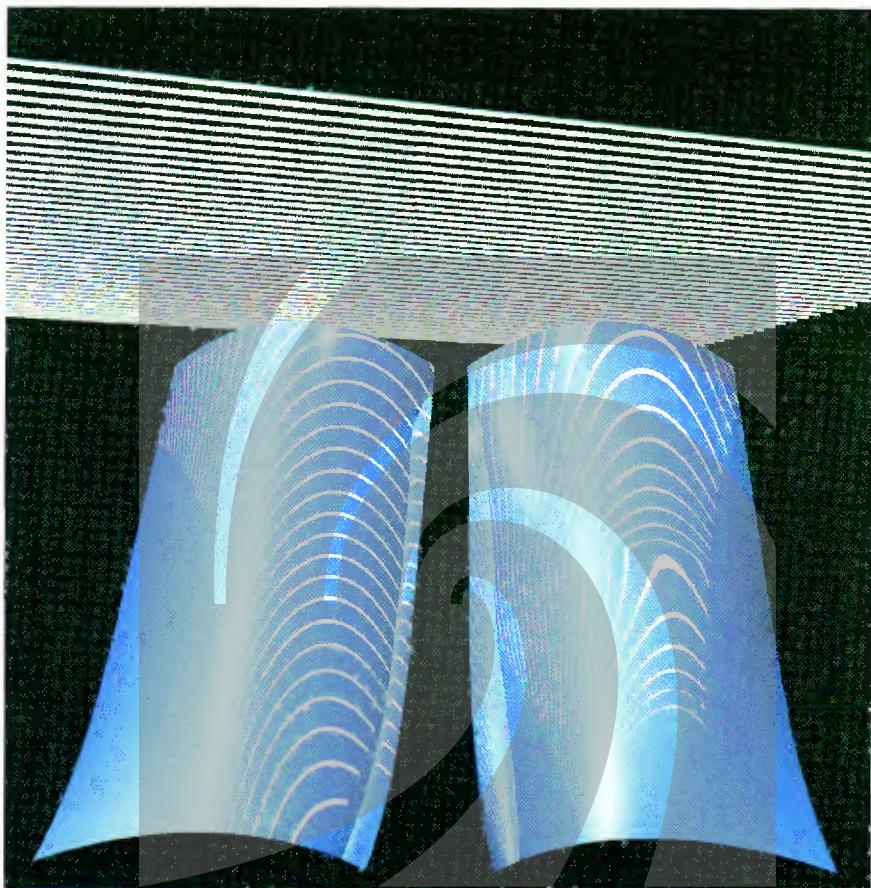


Fig. C.12. Smooth surfaces which are C^2 (left) and which are not (right). Shadows would be smooth, reflection lines are not — cf. p. 472.

Luratech

www.luratech.com

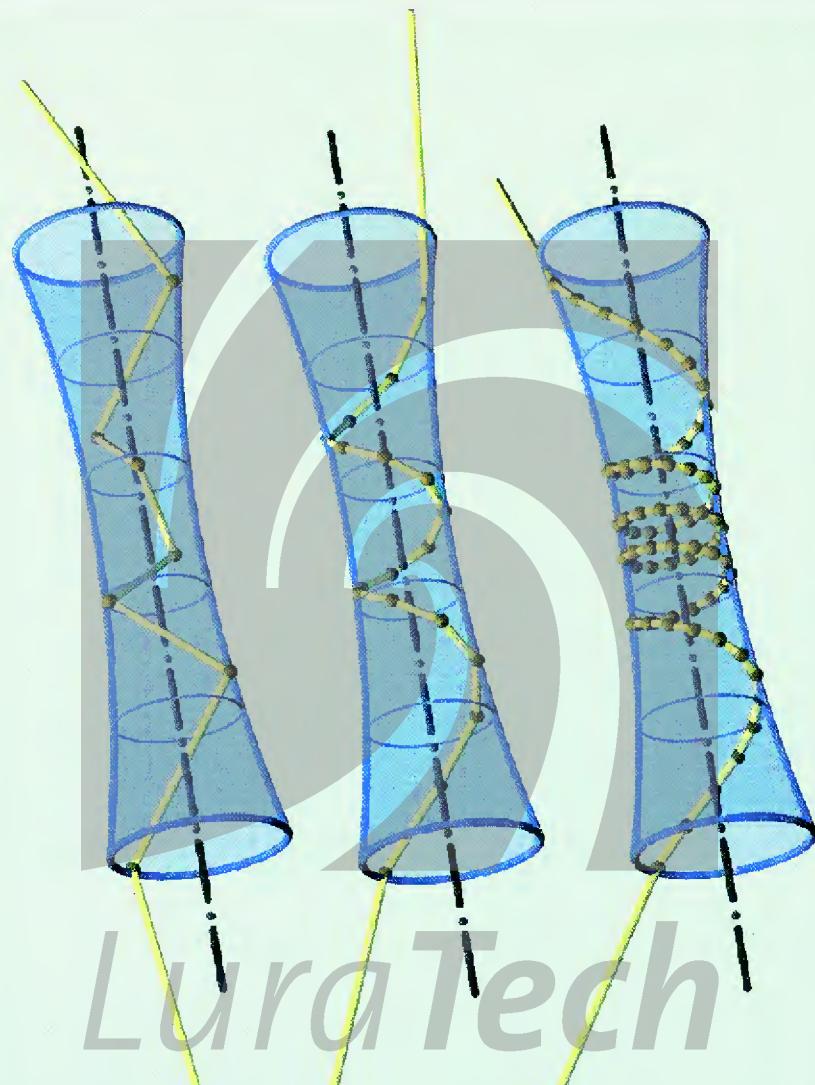


Fig. C.13. Reflection polygons which converge to geodesics — cf. pp. 491, 495
(courtesy H.P. Schröcker).

www.luratech.com

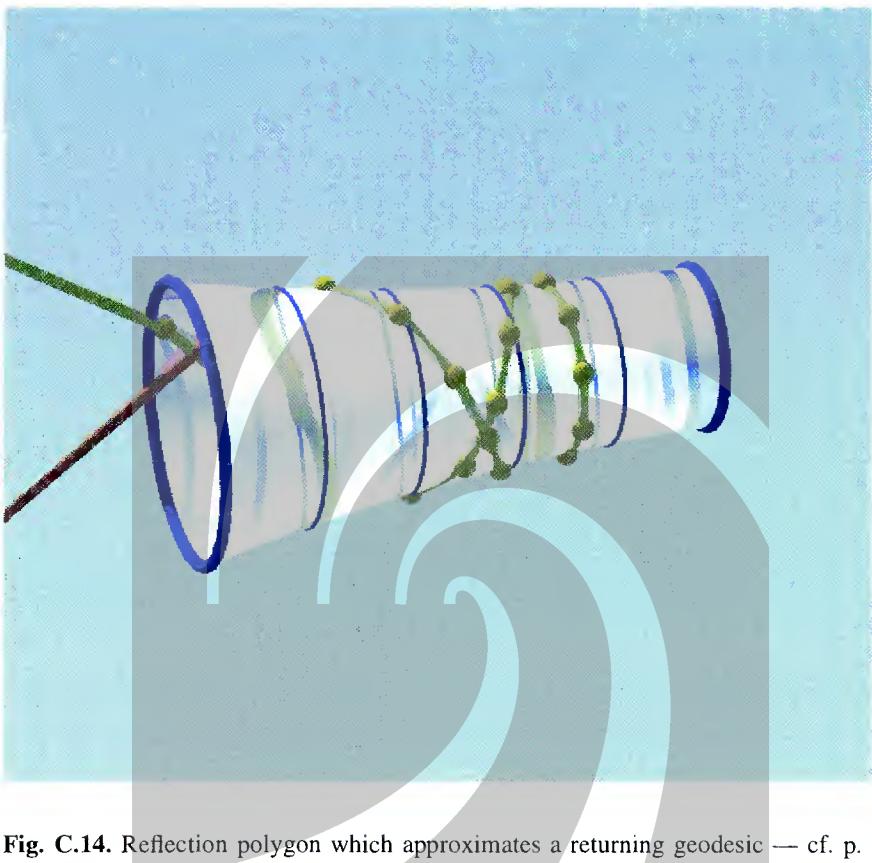


Fig. C.14. Reflection polygon which approximates a returning geodesic — cf. p. 496.

LuraTech

www.luratech.com

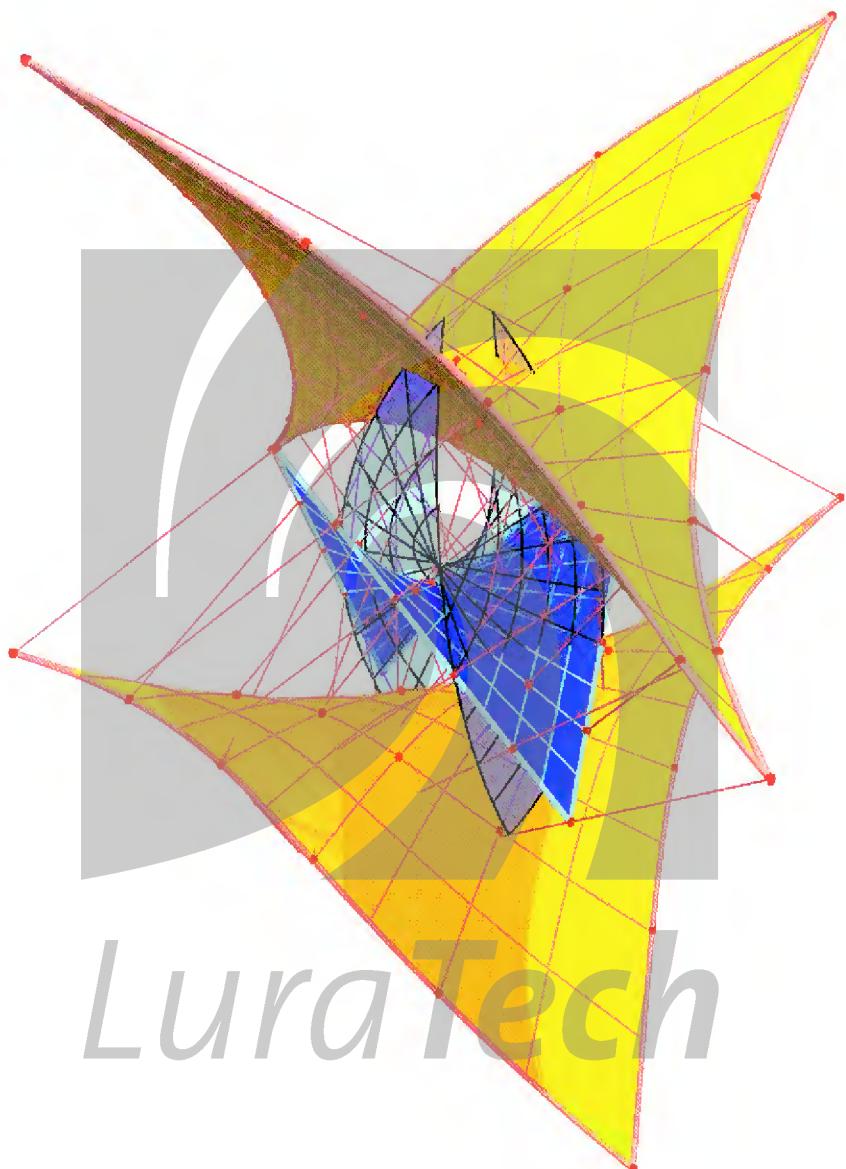


Fig. C.15. Normal congruence of a hyperbolic paraboloid (blue) with focal surfaces (orange) and central surface — cf. pp. 437, 440, 444 (courtesy H.P. Schröcker).

www.lurattech.com

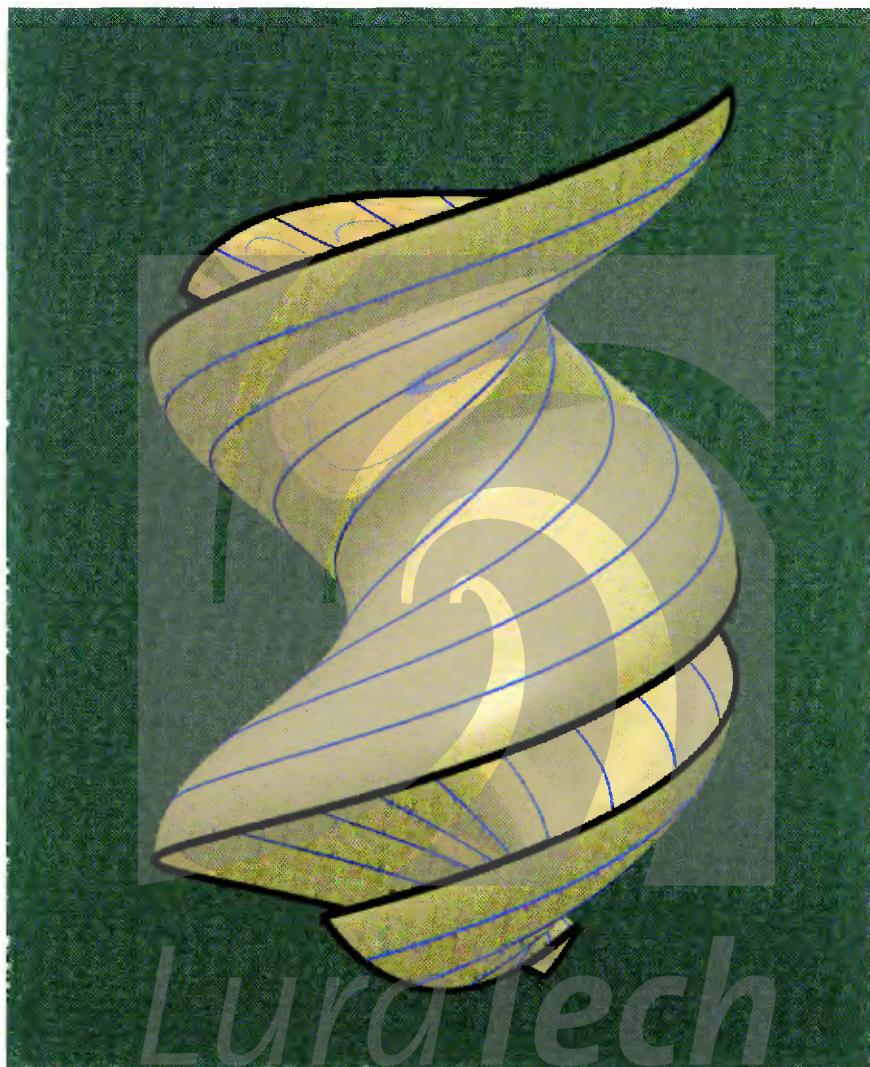


Fig. C.16. Geodesics on a helical surface — cf. p. 495

www.luratech.com

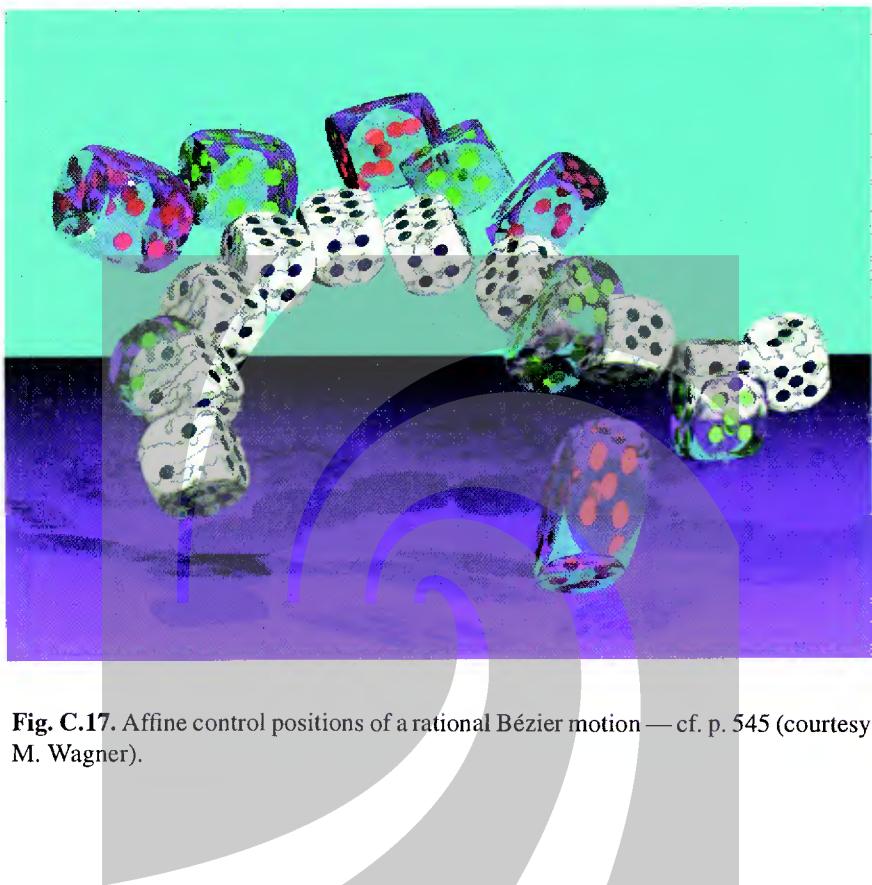


Fig. C.17. Affine control positions of a rational Bézier motion — cf. p. 545 (courtesy M. Wagner).

LuraTech

www.luratech.com

6. Developable Surfaces

Developable surfaces are surfaces in Euclidean space which ‘can be made of a piece of paper’, i.e., are isometric to part of the Euclidean plane, at least locally. If we do not assume sufficient smoothness, the class of such surfaces is too large to be useful — it includes all possible ways of arranging crumpled paper in space. For C^2 surfaces, however, developability is characterized by vanishing Gaussian curvature, and by being made of pieces of torsal ruled surfaces. We will here use ‘developable’ and ‘torsal ruled’ as synonyms, because we are most interested in the ruled surface which carries a developable surface patch.

We first have a look at the Euclidean differential geometry of developable surfaces, and then study developables as envelopes of their tangent planes. This viewpoint identifies the curves in dual projective space with the torsal ruled surfaces. We describe some fields of applications where the concept of developable surface arises naturally and knowledge of the theory leads to geometric insights. These include developables of constant slope, the cycographic mapping, medial axis computations, geometrical optics, rational curves with rational offsets, geometric tolerancing, and the use of developable surfaces in industrial processes.

6.1 Differential Geometry of Developable Surfaces

Here we first describe the relations between some equivalent characterizations of ‘developable surface’ and then study some properties of Euclidean and projective differential geometry of such surfaces. Especially we will discuss the singularities of developable surfaces.

Developable Surfaces as Ruled Surfaces

The intrinsic metric of a surface of Euclidean space is determined by the arc lengths of curves contained in it. We therefore define

Definition. A C^1 surface in Euclidean space is called *developable*, if every point has a neighbourhood which can be mapped isometrically into the Euclidean plane in the sense that arc lengths are preserved.

The following theorem shows how developability of surfaces, which is a local property, is characterized by the vanishing of a second order differential invariant:

Theorem 6.1.1. *If a surface in Euclidean space is C^2 , zero Gaussian curvature characterizes developability.*

One part of this theorem follows from the *theorema egregium* of C.F. Gauss, which says that the Gaussian curvature only depends on the intrinsic metric of the surface. Usually differential geometry textbooks show a proof which requires higher differentiability than just C^2 . The C^2 version is shown in [101]. The next theorem is a local ‘shape’ result concerning developable surfaces in Euclidean space and shows why developable surfaces belong to line geometry.

Theorem 6.1.2. *A developable C^2 surface is a torsal ruled surface in the neighbourhood of a parabolic point.*

Proof. We show a proof for C^3 surfaces: It is well known that away from umbilics and flat points the surface has a local parametrization of the form $\mathbf{x}(u, v)$, $(u, v) \in (0, 1)^2$ such that the lines $u = \text{const.}$ and $v = \text{const.}$ are principal curvature lines. If $\mathbf{n}(u, v)$ is the unit normal vector in $\mathbf{x}(u, v)$, this means that $\mathbf{n}_{,u} = -\kappa_1 \mathbf{x}_{,u}$ and $\mathbf{n}_{,v} = -\kappa_2 \mathbf{x}_{,v}$, where κ_1, κ_2 are the principal curvatures. By assumption, $\kappa_1 = 0$ and $\kappa_2 \neq 0$ (the surface is developable and has no flat point), so $\mathbf{n}_{,u} = \mathbf{0}$. This shows that the tangent plane is constant along the lines $v = \text{const.}$ Differentiating $\mathbf{n} \cdot \mathbf{x}_{,u} = 0$ shows $\mathbf{n}_{,u} \cdot \mathbf{x}_{,u} + \mathbf{n} \cdot \mathbf{x}_{,uu} = \mathbf{n} \cdot \mathbf{x}_{,uu} = 0$. Principal curvature lines are orthogonal, so $\mathbf{x}_{,u} \cdot \mathbf{x}_{,v} = 0$, consequently $\mathbf{x}_{,u} \cdot \mathbf{n}_{,v} = 0$ and $0 = \mathbf{x}_{,uu} \cdot \mathbf{n}_{,v} + \mathbf{x}_{,u} \cdot \mathbf{n}_{,vu} = \mathbf{x}_{,uu} \cdot \mathbf{n}_{,v}$. Thus we have established that $\mathbf{x}_{,uu}$ is orthogonal to \mathbf{n} and $\mathbf{n}_{,v}$, so it is parallel to $\mathbf{x}_{,u}$. This implies that the curves $v = \text{const.}$ are straight lines. \square

If the surface is only C^2 , a different proof is required (cf. the paper [68] by P. Hartman and L. Wintner). Developable surfaces of sufficient smoothness contain straight line segments also if they have flat points. A detailed discussion can be found in [186], Vol. 3, p. 347ff.

Remark 6.1.1. The definition of developability involves the metric in Euclidean space. Nevertheless, Th. 6.1.2 shows that developability of a surface is essentially equivalent to its carrying only parabolic and flat points. This property is invariant with respect to projective automorphisms. Thus developable surfaces are actually well-defined objects of projective differential geometry. \diamond



Remark 6.1.2. Surfaces which are only C^1 sometimes exhibit unusual features. By the Gauss-Bonnet theorem, the only compact surfaces embeddable into Euclidean three-space which can possibly carry a metric of zero Gaussian curvature are tori. Such flat tori are obtained, as abstract surfaces, by gluing together opposite edges of a rectangular sheet of paper. There are no C^2 flat tori in E^3 — the smallest sphere which contains such a torus touches in elliptic surface points.

The embedding theorem of J. Nash [130] states that any compact surface with an arbitrarily prescribed metric has a C^1 isometric embedding into Euclidean three-space, provided it has a smooth, not necessarily isometric, embedding. Thus there is a developable C^1 surface in \mathbb{R}^3 which is a torus (attempts to visualize this surface will fail). \diamond

Remark 6.1.3. On the one hand, a developable surface may be only a piece of a ruled surface. On the other hand, developability is obstructed by self-intersections of a surface. Indeed, it is possible to show that the only regular *geodesically complete* developable surfaces are planes or cylinders, where geodesic completeness means that geodesics can be extended arbitrarily (see [186], Vol. 3, p. 363ff.)

If we consider only the ‘ruled surface information’ contained in a developable surface patch, we of course loose information about the extent and location of this patch on the underlying ‘complete’ ruled surface. ◇

Th. 5.1.7 enumerates the possible basic types of torsal ruled surfaces: cylindrical surfaces, conical surfaces, and tangent surfaces of space curves. We are going to show that they are actually developable.

Theorem 6.1.3. *The following surfaces are developable:*

1. *cylindrical surfaces $s(u, v) = l(u) + vr$ with smooth l , and $r \neq 0$.*
2. *conical surfaces $s(u, v) = vr(u)$ with smooth r and $r(u) \neq 0$.*
3. *the tangent surface $s(u, v) = c(u) + vc(u)$ of the curve $c(u)$ (the curve of regression), if c is C^2 and nonsingular.*

Proof. The proof proceeds by computing the development of these surfaces, which is done in the following paragraphs. In all three cases we find a planar surface $\bar{s}(u, v)$ such that the coefficients of the first fundamental form are the same for s and \bar{s} . This shows that the mapping $s(u, v) \mapsto \bar{s}(u, v)$ is locally isometric (we cannot conclude that the mapping is ‘isometric’ as a whole, because it may happen that \bar{s} is not one-to-one). The reader should note that our differentiability assumptions are not the minimal ones.

Development of Cylindrical Surfaces

Without loss of generality we may assume that l is parametrized by arc length, that l is a planar curve in a plane orthogonal to r , and that $\|r\| = 1$. We consider the planar surface $\bar{s}(u, v) = (u, v)$. To show that the mapping $s(u, v) \mapsto \bar{s}(u, v)$ is isometric, we compute the coefficients $g_{jk} = s_{,j} \cdot s_{,k}$ and $\bar{g}_{jk} = \bar{s}_{,j} \cdot \bar{s}_{,k}$ of the first fundamental forms: The equations $s_{,1} = l(u)$, $s_{,2} = r$, $\bar{s}_{,1} = (1, 0)$, $\bar{s}_{,2} = (0, 1)$ imply that $g_{11} = \bar{g}_{11} = g_{22} = \bar{g}_{22} = 1$, $g_{12} = \bar{g}_{12} = 0$. Thus the intrinsic metric of the surface s is planar.

Development of Conical Surfaces

Without loss of generality we assume that $\|r(u)\| = 1$ (which implies $r \cdot \dot{r} = 0$) and that the spherical curve r is parametrized with respect to its arc length (which implies $\dot{r} \cdot \dot{r} = 1$). We define $\bar{s}(u, v) = (v \cos u, v \sin u)$ and compute $s_{,1} = vr(u)$, $s_{,2} = r(u)$, $\bar{s}_{,1} = (-v \sin u, v \cos u)$, $\bar{s}_{,2} = (\cos u, \sin u)$. Then $g_{11} = \bar{g}_{11} = v^2$, $g_{22} = \bar{g}_{22} = 1$, $g_{12} = \bar{g}_{12} = 0$, and we see that the intrinsic metric of s is flat.

Development of Tangent Surfaces

We assume that the curve \mathbf{c} is parametrized with respect to arc length, which means $\dot{\mathbf{c}} \cdot \dot{\mathbf{c}} = 1$, $\ddot{\mathbf{c}} \cdot \dot{\mathbf{c}} = 0$, and $\|\ddot{\mathbf{c}}\| = \kappa$, where κ is the curvature of \mathbf{c} . There is a planar curve $\bar{\mathbf{c}}(u)$, parametrized by its arc length, such that the curvatures $\kappa(u)$ and $\bar{\kappa}(u)$ of \mathbf{c} and $\bar{\mathbf{c}}$ are equal:

$$\bar{\mathbf{c}}(u) = \left(\int_{u_0}^u \cos \phi(x) dx, \int_{u_0}^u \sin \phi(x) dx \right) \text{ with } \phi(u) = \int_{u_0}^u \kappa(x) dx. \quad (6.1)$$

Then we let $\bar{s}(u, v) = \bar{\mathbf{c}}(u) + v \dot{\bar{\mathbf{c}}}(u)$. We compute the coefficients of the first fundamental form. Recall that $\mathbf{s}(u, v) = \mathbf{c}(u) + v \dot{\mathbf{c}}(u)$. It follows that

$$s_{,1} = \dot{\mathbf{c}}(u) + v \ddot{\mathbf{c}}(u), \quad s_{,2} = \dot{\mathbf{c}}(u), \quad (6.2)$$

$\bar{s}_{,1} = \dot{\bar{\mathbf{c}}}(u) + v \ddot{\bar{\mathbf{c}}}(u)$, $\bar{s}_{,2} = \dot{\bar{\mathbf{c}}}(u)$, $g_{11} = \bar{g}_{11} = 1 + v^2 \kappa(u)^2$, $g_{22} = \bar{g}_{22} = 1$, $g_{12} = \bar{g}_{12} = 1$. This shows that the mapping $s(u, v) \mapsto \bar{s}(u, v)$ is a local isometry, and the proof of Th. 6.1.3 is finished. \square

Computing the Development

It is desirable to compute the development of a developable ruled surface without having to compute the curve of regression first and then to integrate Equ. (6.1). A more efficient method which works for all types of developable surfaces is based on the fact that the *geodesic curvature* κ_g of a surface curve \mathbf{c} is a property of the intrinsic metric, i.e., is unchanged when developing the surface into the plane. Another intrinsic property is $\|\dot{\mathbf{c}}\|$.

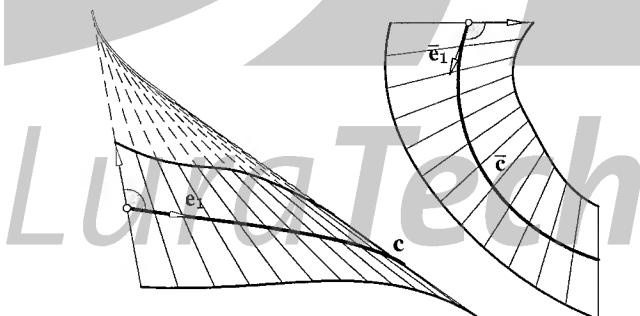


Fig. 6.1. Developable surface (left) and its development (right).

After developing the surface, we get a planar curve $\bar{\mathbf{c}}$. Comparison of the Frenet equations for the plane and the definition of geodesic curvature shows that for planar curves the latter equals the ordinary curvature of curves. Denote the unit tangent and unit normal vector of $\bar{\mathbf{c}}$ by $\bar{\mathbf{e}}_1$ and $\bar{\mathbf{e}}_2$, respectively, and temporarily let $v(t) = \|\dot{\bar{\mathbf{c}}}(t)\| = \|\ddot{\bar{\mathbf{c}}}(t)\|$. Then the Frenet equations for $\bar{\mathbf{c}}$ read

$$\dot{\bar{\mathbf{e}}}_1(t) = v(t)\kappa_g(t)\bar{\mathbf{e}}_2(t), \quad \dot{\bar{\mathbf{e}}}_2(t) = -v(t)\kappa_g(t)\bar{\mathbf{e}}_1(t).$$

This is an ordinary second order differential equation for the curve $\bar{\mathbf{c}}$. It is solved in the following way: We first let $\bar{\mathbf{e}}_1(t) = (\cos \phi(t), \sin \phi(t))$. From $\dot{\bar{\mathbf{e}}}_1(t) = \dot{\phi}\mathbf{e}_2(t)$ and $\dot{\bar{\mathbf{c}}}(t) = v(t)\bar{\mathbf{e}}_1(t)$ we conclude that

$$\phi(t) = \int v(t)\kappa_g(t)dt, \quad \bar{\mathbf{c}}(t) = \int v(t)\bar{\mathbf{e}}_1(t)dt.$$

All solutions of this differential equation are congruent curves. Having found a curve $\bar{\mathbf{c}}(t)$, we can easily find the development of the generator lines $R(u)$: The angle enclosed by the curve and the generators is preserved (see Fig. 6.1).

Difference Geometry of Developable Surfaces

We want to describe a discrete polyhedral model for developable surfaces: It is obvious how to construct such models for cylindrical or conical surfaces. A model of a tangent surface is constructed by substituting the curve of regression by a polygon, the generators of the surface by the carrier lines of the edges of this polygon, and the tangent planes by the planes spanned by two consecutive edges (cf. Fig. 6.2). Note that the polyhedral model, like the smooth surface, exhibits a sharp edge of regression.

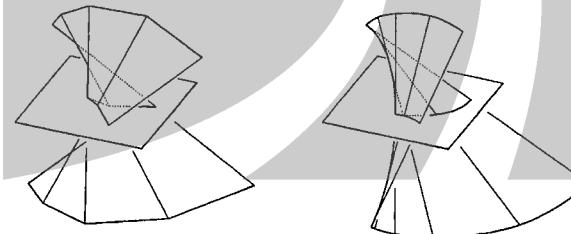


Fig. 6.2. Polyhedral model (left) of tangent surface (right) with planar sections.

Developing the polyhedral model is simply flattening the polyhedron. This procedure obviously leaves the angles between consecutive edges of the polygon of regression unchanged, which is in accordance with the proof of Th. 6.1.3, where we saw that developing does not change the curvature of the curve of regression.

This example shows already that the polyhedral model of a torsal ruled surface is helpful for experimental investigations and can easily supply motivations and ideas for differential geometric results concerning developable surfaces. If done properly, such results can also directly be proved with sequences of finer and finer polyhedral approximations (cf. [173]). Such approximations are also useful for numeric computations.

The Osculating Cone of Revolution

We already know that the tangent plane of a torsal ruled surface is constant for all points of a generator, and that the set of surface tangents consists of the field of lines of the tangent plane together with the bundle of lines incident with the cuspidal point. These two are linear objects in first order contact with a torsal ruled surface.

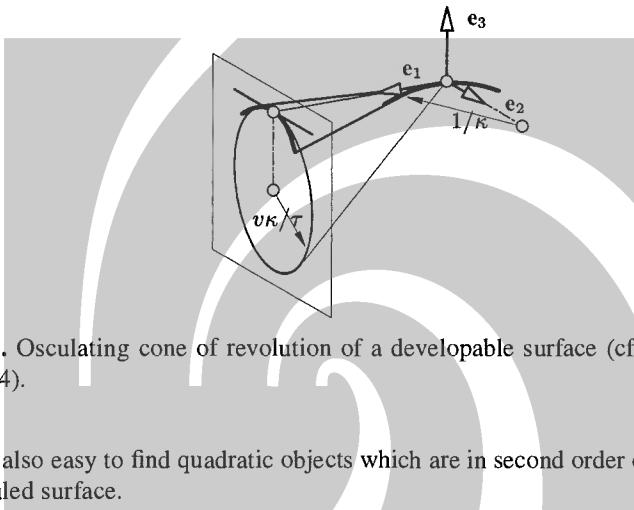


Fig. 6.3. Osculating cone of revolution of a developable surface (cf. the proof of Th. 6.1.4).

It is also easy to find quadratic objects which are in second order contact with a torsal ruled surface.

Theorem 6.1.4. *A regular generator $R(u)$ of a C^2 cylinder, cone, or tangent surface \mathcal{R} is either a flat generator — then the tangent plane is in second order contact with \mathcal{R} — or there is a cylinder or cone of revolution which is in second order contact with \mathcal{R} in all regular points of $R(u)$. If \mathcal{R} is the tangent surface of a curve, the axis of this cone is parallel to the Darboux vector $\kappa e_3 + \tau e_1$.*

The coefficients κ and τ are the curvature and torsion of the curve of regression. The quotient τ/κ is called the *conical curvature* \mathcal{R} .

Proof. If \mathcal{R} is cylindrical, an orthogonal planar section and its osculating circle show the result. If \mathcal{R} is conical, a spherical section and its spherical osculating circle do the same. If \mathcal{R} is the tangent surface of the regular C^2 curve curve \mathbf{c} , we may assume, without loss of generality, that $\|\dot{\mathbf{c}}\| = 1$. Note that the generator $R(u)$ is regular only if $\dot{\mathbf{c}}$ and $\ddot{\mathbf{c}}$ are not linearly dependent, i.e., \mathbf{c} has no inflection point at $\mathbf{c}(u)$.

We consider the Frenet frame e_1, e_2, e_3 , the curvature κ , and the torsion τ of \mathbf{c} (see Ex. 1.2.3 and Fig. 6.3). We use the parametrization $s(u, v) = \mathbf{c}(u) + v e_1(u)$ for the tangent surface and compute its principal curvatures.

The binormal vector $e_3(u)$ is the unit surface normal vector in all regular points of $R(u)$. We use the Frenet equations and the derivatives of (6.2) to compute the coefficients $h_{jk} = -s_{,j} \cdot e_{3,k}$ of the second fundamental form. The only nonzero coefficient is

$$h_{11}(u, v) = -(\mathbf{e}_1(u) + v\kappa(u)\mathbf{e}_2(u)) \cdot (-\tau(u)\mathbf{e}_2(u)) = v\kappa(u)\tau(u), \quad (6.3)$$

Because of $h_{12} = h_{21} = h_{22} = 0$, the ruling $R(u)$ is a principal curvature line, and the corresponding principal curvature is zero. A vector $\mathbf{w}(u, v)$ orthogonal to $R(u)$ in the point $s(u, v)$ is given by $\mathbf{e}_2(u)$. In order to compute its normal curvature (which equals the second principal curvature κ_2) we write \mathbf{e}_2 as linear combination of $s_{,1}$ and $s_{,2}$ and compute

$$\begin{aligned} \mathbf{e}_2 &= \mathbf{w}(u, v) = \frac{1}{v\kappa}(\mathbf{e}_1 + v\kappa\mathbf{e}_2) - \frac{1}{v\kappa}\mathbf{e}_1 = \frac{1}{v\kappa}s_{,1}(u, v) - \frac{1}{v\kappa}s_{,2}(u, v). \\ \kappa_2(u, v) &= \kappa_n(\mathbf{w}) = h(\mathbf{w}, \mathbf{w})/(\mathbf{w} \cdot \mathbf{w}) = h_{11}(\frac{1}{v\kappa})^2/1 = \frac{\tau(u)}{v\kappa(u)}. \end{aligned} \quad (6.4)$$

We see that the principal curvature radius $1/\kappa_2$ is increasing *linearly* with v from the singular point $s(u, 0)$. The unique cone of revolution with the same principal curvatures osculates \mathcal{R} in the regular points of $R(u)$. Elementary trigonometry shows that the aperture angle β of this cone is given by

$$\cot \beta = \left| \frac{\tau}{\kappa} \right|, \quad (6.5)$$

and its axis is parallel to the vector $\kappa\mathbf{e}_3 + \tau\mathbf{e}_1$. \square

Remark 6.1.4. If \mathcal{R} is a tangent surface and \mathcal{C} its osculating cone, which is in second order contact along the regular points of the generator line $R(u)$, then the Klein images $\mathcal{R}\gamma$ and $\mathcal{C}\gamma$ are not in second order contact at $R(u)\gamma$. This behaviour is in contrast to that of the Lie quadric of a non-torsal generator whose Klein image is in second order contact with the Klein image of the ruled surface (cf. Th. 5.1.9). \diamond

Circumscribed Developable Surfaces

Recall the notion of conjugate surface tangent from Sec. 1.2.2, and especially Remark 1.2.10, which defines the conjugacy relation for singular quadratic varieties Ψ : If Ψ is a parabolic cylinder, then all surface tangents are conjugate to the generator lines, and if Ψ is a plane, all surface tangents are conjugate to each other.

There is the following lemma concerning surfaces which are in line contact:

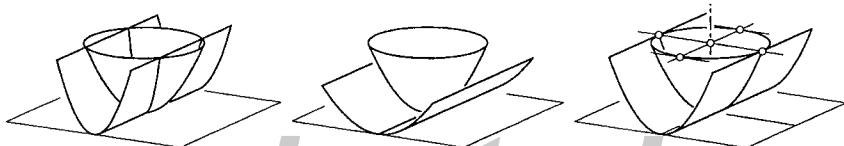


Fig. 6.4. Paraboloid and parabolic cylinder with common surface element: Transverse intersection, no intersection, line contact.

Lemma 6.1.5. *Two C^2 surfaces Φ_1, Φ_2 , which are in line contact along a curve c share a common pair of conjugate surface tangents in all points of c , and one of these tangents is tangent to c .*

Proof. We consider one such contact point, write both surfaces Φ_1, Φ_2 as graph surfaces over the common tangent plane, and approximate them by the graphs Ψ_1, Ψ_2 of their quadratic Taylor polynomials. There are three possibilities: Ψ_1 and Ψ_2 (which are paraboloids, parabolic cylinders, or planes) (i) do not intersect, (ii) intersect transversely in all points except $(0, 0)$, or (iii) are in line contact along a conic or line (see Fig. 6.4). Cases (i) and (ii) imply the same for Φ_1, Φ_2 , so we have case (iii). If both Φ_1, Φ_2 are paraboloids and l is a conic, this means that the pole of l 's carrier plane with respect to both Φ_1, Φ_2 is the same, and so l 's tangent has the same conjugate for both Φ_1, Φ_2 .

The other more degenerate cases are handled in a similar manner. It turns out that the theorem is true in all cases. Because conjugate surface tangents are defined via conjugacy with respect to an osculating quadratic variety, this implies the statement of the lemma. \square

Lemma 6.1.5 is helpful if we want to find developable surfaces circumscribed to another surface along a curve:

Theorem 6.1.6. *If the C^2 surfaces Φ_1, Φ_2 are in line contact along a curve c , and Φ_2 is a developable ruled surface, then the generator lines of Φ_2 are conjugate (in the sense of Φ_1) to the tangents of c .*

Proof. The osculating quadratic variety of Φ_2 is a parabolic cylinder or a plane. In both cases any surface tangent is conjugate to the generator lines of Φ_2 . The result follows immediately from Lemma 6.1.5. \square

A picture of the Dupin indicatrices of Φ_1 and Φ_2 is shown in Fig. 6.47.

6.2 Dual Representation

By Th. 6.1.2, developable surfaces and torsal ruled surfaces are essentially the same. The latter have the property that the tangent plane is constant for the points of a generator. Thus there is a one-parameter smooth family of tangent planes. It turns out that this family of tangent planes is a very useful tool for studying torsal ruled surfaces. One of the reasons is that the tangent planes comprise a curve in dual projective space.

6.2.1 Differential Geometry of the Dual Surface

Consider a torsal ruled surface \mathcal{R} , parametrized in the form $R(t)$. The family of its tangent planes is parametrized in the form $U(t) = \mathbb{R}u(t)$, $t \in I \subset \mathbb{R}$. We consider the surface as the envelope of its tangent planes. If we interpret $u(t)$ as coordinates of a point instead of coordinates of a plane, then $\mathbb{R}u(t)$ defines a curve in dual projective space. It will be called the *dual curve* $U(t) = \mathbb{R}u(t)$.

Envelopes of Planes

We call the representation of a developable surface as the envelope of its tangent planes $U(t) = \mathbb{R}u(t)$ its *dual representation*. Conversely, smooth families of planes define developable surfaces, if they fulfill certain non-degeneracy conditions:

Theorem 6.2.1. *Assume that $U(t) = \mathbb{R}u(t)$ ($t \in I$) is a sufficiently smooth family of planes, i.e., a curve in dual projective space. All intervals $J \subset I$ contain an interval, where one of the following holds true:*

- (i) $U(t)$ is constant, or $R(t) = U(t) \cap U^1(t)$ is a line and is constant.
- (ii) $c(t) = U(t) \cap U^1(t) \cap U^2(t)$ is a well-defined point and is constant. $R(t)$ parametrizes a cone with tangent planes $U(t)$.
- (iii) $c(t)$ is a regular curve, $R(t)$ is its tangent, and $U(t)$ is tangent to the tangent surface of c . The envelope is the tangent surface of the curve c .

Proof. The argumentation is similar to the proof of Th. 5.1.7. The computations shown in the proof are explained in more detail below. If $\{u, \dot{u}\}$ is linearly dependent in an interval, then $U(t)$ is constant. If not, $R = U \cap U^1$ is a well defined line with dual Plücker coordinates $u \wedge \dot{u}$. It may be locally constant. This is the uninteresting case (i).

Otherwise, the curve $R(t)\gamma$ is locally regular, which means that $\{u \wedge \dot{u}, u \wedge \ddot{u}\}$ is linearly independent. This implies that $\{u, \dot{u}, \ddot{u}\}$ is linearly independent and the point $c = U \cap U^1 \cap U^2$ with homogeneous coordinates $c = u \times \dot{u} \times \ddot{u}$ is well defined (for the cross product, see Remark 1.2.12 and Ex. 2.2.8).

If $c(t)$ is constant, we have case (ii). Otherwise, c locally is a regular curve. Computing its derivative $\dot{c} = u \times \dot{u} \times \ddot{u}$ shows that $R(t)$ is tangent to $c(t)$, so we have case (iii).

By the very definition of dual Plücker coordinates, the plane $\mathbb{R}u$ contains the lines with dual Plücker coordinates $u \wedge \dot{u}$ and $u \wedge \ddot{u}$. These lines span precisely the pencil mentioned in Th. 5.1.5, whose carrier plane is tangent to the ruled surface $R(t)$. This shows the statement about the tangent planes. \square

If the envelope of the planes $U(t)$ is the tangent surface of a curve c (this was case (iii)), then this curve is the *curve of regression* defined in Ex. 5.1.6.

Proposition 6.2.2. *If the envelope of planes $U(t)$ is the tangent surface of the curve $c(t)$, then $U(t_0)$ is c 's osculating plane for $t = t_0$, if U is C^4 .*

Proof. We use the notation of the proof of Th. 6.2.1. If $c = u \times \dot{u} \times \ddot{u}$, then $\dot{c} = u \times \dot{u} \times \ddot{u}$, and $\ddot{c} = u \times \ddot{u} \times \dot{u} + u \times \dot{u} \times u^{(4)}$. Obviously the plane $\mathbb{R}u$ contains the points $c\mathbb{R}$, $\dot{c}\mathbb{R}$, $\ddot{c}\mathbb{R}$. \square

Computing the Rulings

The plane $U(t)$ touches its envelope \mathcal{R} in all points of the ruling $R(t)$. It is found by intersecting $U(t)$ with its first derivative plane $U^1(t) = \mathbb{R}\dot{u}(t)$. If $u(t) = (u_0(t), u(t))$, then the Plücker coordinates $(\mathbf{r}, \bar{\mathbf{r}})$ of the ruling $R(t)$ are given by

$$(\mathbf{r}, \bar{\mathbf{r}}) = (\mathbf{u} \times \dot{\mathbf{u}}, u_0 \dot{\mathbf{u}} - \dot{u}_0 \mathbf{u}). \quad (6.6)$$

The expression $R(t) = U(t) \cap U^1(t)$ is obviously dual to the expression $T(t) = c(t) \vee c^1(t)$, which describes the construction of the tangent of a curve. The ruling $R(t)$ can therefore be seen as a dual tangent to the family $U(t)$ of planes.

The computation (6.6) fails if the vectors $\mathbf{u}(t)$ and $\dot{\mathbf{u}}(t)$ are linearly dependent, i.e., if the family $U(t)$ of planes (which is a curve in dual space) has a singularity. The geometrical meaning of this situation will be clarified later.

Computing the Curve of Regression

We want to compute the *curve $c(t)$ of regression* of the developable surface \mathcal{R} . This is done by intersecting $U(t)$ with its first and second derivative planes:

$$c(t) = U(t) \cap U^1(t) \cap U^2(t). \quad (6.7)$$

This construction is obviously dual to the definition $O = c(t) \vee c^1(t) \vee c^2(t)$ of the osculating subspace of order two of a curve (which is usually a plane). We can compute the intersection (6.7) according to Remark 1.2.12 and Ex. 2.2.8:

$$\mathbf{c}(t) = \mathbf{u}(t) \times \dot{\mathbf{u}}(t) \times \ddot{\mathbf{u}}(t), \quad (6.8)$$

In coordinates, $\mathbf{c}(t)$ equals the column vector

$$\left(\begin{vmatrix} u_1 & \dot{u}_1 & \ddot{u}_1 \\ u_2 & \dot{u}_2 & \ddot{u}_2 \\ u_3 & \dot{u}_3 & \ddot{u}_3 \end{vmatrix}, - \begin{vmatrix} u_0 & \dot{u}_0 & \ddot{u}_0 \\ u_2 & \dot{u}_2 & \ddot{u}_2 \\ u_3 & \dot{u}_3 & \ddot{u}_3 \end{vmatrix}, \begin{vmatrix} u_0 & \dot{u}_0 & \ddot{u}_0 \\ u_1 & \dot{u}_1 & \ddot{u}_1 \\ u_3 & \dot{u}_3 & \ddot{u}_3 \end{vmatrix}, - \begin{vmatrix} u_0 & \dot{u}_0 & \ddot{u}_0 \\ u_1 & \dot{u}_1 & \ddot{u}_1 \\ u_2 & \dot{u}_2 & \ddot{u}_2 \end{vmatrix} \right)^T.$$

Recall that the rulings $R(t)$ are the tangents of the curve of regression $c(t)$ (if non-singular) and that the planes $U(t)$ are the osculating planes of $c(t)$ (if it has no inflection point).

Remark 6.2.1. A torsal ruled surface \mathcal{R} is represented by a ‘curve’ in several different ways. If we consider the ruled surface representation $R(t)$, clearly $R\gamma(t)$ is a curve. We know, however, that not all curves in the Klein quadric define torsal ruled surfaces.

\mathcal{R} is locally either (i) the tangent surface of a planar curve, or (ii) the tangent surface of a twisted curve in space, or (iii) a conical surface, which includes cylinders.

Cases (i) and (ii) can be presented by the curve of regression, which is not possible for case (iii).

Cases (ii) and (iii), however, are capable of the dual representation by the family of tangent planes, which fails for case (i). Depending on the application one has in mind and on the class of developables one is interested in, one would use the ruled surface representation, the curve of regression, or the dual representation. The two latter representations of developables are duals of each other. \diamond

Example 6.2.1. Consider the family of planes $U(t) : t^n + t^m x + t^l y + z = 0$, i.e.,

$$\mathbf{u}(t) = (t^n, t^m, t^l, 1). \quad (6.9)$$

We assume that $l < m < n$, $l \geq 1$, and $m \geq 2$. By Equ. (6.6), the Plücker coordinates of the generator $R(t)$ are given by

$$(-lt^{l-1}, mt^{m-1}, (l-m)t^{l+m-1}, (m-n)t^{m+n-1}, (l-n)t^{l+n-1}, -nt^{n-1}).$$

Multiplication with the factor t^{1-l} for $t \neq 0$ does not change $R(t)$. This gives the Plücker coordinates

$$(\mathbf{r}(t), \bar{\mathbf{r}}(t)) = (-l, mt^{m-l}, (l-m)t^m, (m-n)t^{m+n-l}, (l-n)t^n, -nt^{n-l}). \quad (6.10)$$

This formula is valid also for $t = 0$, because $R(u)$ varies continuously, and $\mathbf{r}(t)$, $\bar{\mathbf{r}}(t)$ are continuous. By (6.8), the curve of regression is given by $c(t) = (\nu t^{l+m-3}, \mu t^{l+n-3}, \lambda t^{m+n-3}, -(\lambda + \mu + \nu)t^{l+m+n-3})\mathbb{R}$, with

$$\lambda = mn^2 - nm^2, \mu = nl^2 - ln^2, \nu = lm^2 - ml^2. \quad (6.11)$$

Multiplication of these homogeneous coordinates with the factor t^{3-l-m} for $t \neq 0$ and continuously extending for $t = 0$ shows that

$$c(t) = (\nu, \mu t^{n-m}, \lambda t^{n-l}, -(\lambda + \mu + \nu)t^n)\mathbb{R} \quad (6.12)$$

parametrizes the curve of regression. \diamond

Taylor Polynomials

We want to use the dual representation of a developable surface \mathcal{R} to study its local behaviour. Our basic tool will be Ex. 6.2.1, because any vector valued function $\mathbf{v}(t)$ which is k times differentiable may be approximated at $t = 0$ by its k -th order Taylor polynomial in the sense that

$$\mathbf{v}(t) = \mathbf{v}(0) + t\dot{\mathbf{v}}(0) + \frac{t^2}{2}\ddot{\mathbf{v}}(0) + \dots + \frac{t^k}{k!}\mathbf{v}^{(k)}(0) + o(t^k) \quad (t \rightarrow 0),$$

where the symbol $o(t^k)$ denotes a vector-valued smooth function with the property that $o(t^k)/t^k \rightarrow 0$ as $t \rightarrow 0$. We will use the following lemmas:

Lemma 6.2.3. *A sufficiently smooth curve c in real projective n -space has, with respect to a suitable coordinate system, the coordinate representation $c(t) = \mathbf{c}(t)\mathbb{R} = (1, t^{n_1} + o(t^{n_1}), \dots, t^{n_r} + o(t^{n_r}), o(t^q), \dots, o(t^q))\mathbb{R}$ as $t \rightarrow 0$ with $0 < n_1 < \dots < n_r$, and arbitrary (high) q .*

Proof. Consider the sequence of osculating subspaces of the curve $c(t)$, and determine numbers n_1, \dots, n_r such that the n_i -th osculating subspace $c \vee c^1 \vee \dots \vee c^{n_i}$ at $t = 0$ has dimension i , but the $(n_i - 1)$ -th osculating subspace is smaller. It is possible that this sequence terminates at $r < n$.

Choose the derivative points $c(0), c^{n_1}(0), \dots, c^{n_r}(0)$ as first $r+1$ base points of a projective coordinate system, and normalize the homogeneous coordinate representation $\mathbf{c}(t)$ of c such that $\mathbf{c}(t) = (1, \dots)$. By construction, the Taylor polynomial of the i -th coordinate of $\mathbf{c}(t)$ begins with $\lambda_i \cdot t^{n_i}$. By attaching appropriate weights to the base points (or by appropriately choosing the unit point), we can make all coefficients λ_i equal to 1. \square

Remark 6.2.2. The terms $o(t^q)$ appear if the sequence of osculating subspaces does not terminate with an n -dimensional space, but at a lower dimension. This happens if the curve is contained in a subspace (but then the lemma can be applied to this subspace), but also in other cases: The curve $t \mapsto (1, t, \exp(-1/t^2))\mathbb{R}$ is such an example: all osculating subspaces are 1-dimensional for $t = 0$. \diamond

Lemma 6.2.4. *If the dual representation $U(t) = \mathbb{R}\mathbf{u}(t)$ of a developable surface has, in a suitable coordinate system, the representation $(t^n + o(t^n), t^m + o(t^m), t^l + o(t^l), 1)$ as $t \rightarrow 0$, then its curve of regression is given by*

$$\mathbf{c}(t) = (\nu, \mu t^{n-m} + o(t^{n-m}), \lambda t^{n-l} + o(t^{n-l}), -(\lambda + \mu + \nu)t^n + o(t^n))\mathbb{R}. \quad (6.13)$$

with λ, μ, ν defined by Equ. (6.11).

Proof. This is shown by adding appropriate $o(t^j)$'s to Ex. 6.2.1. \square

Planar Sections and their Taylor Polynomials

In Ex. 6.2.1 the assumptions on l, m, n guarantee that the tangent plane $U(0)$ is the plane $(0, 0, 0, 1)\mathbb{R}$, or equivalently, the plane $z = 0$. The ruling $R(0)$ is the x -axis and the point of regression $c(0)$ is the origin of the underlying Cartesian coordinate system.

We intersect \mathcal{R} with planes in order to understand the behaviour of \mathcal{R} in the neighbourhood of a generator line. This plane shall not contain the origin, i.e., the point of regression. It is sufficient for our discussion to intersect \mathcal{R} with the ideal plane $x_0 = 0$. The intersection curve $r(t)$ is parametrized by $(0, \mathbf{r}(t))\mathbb{R}$, where $(\mathbf{r}(t), \bar{\mathbf{r}}(t))$ is the generator line $R(t)$. We use x_1, x_2, x_3 as projective homogeneous coordinates (x_1, x_2, x_3) in the ideal plane, and see that the intersection curve has the Taylor expansion

$$\mathbf{r}(t)\mathbb{R} = (-l, mt^{m-l} + o(t^{m-l}), (l-m)t^m + o(t^m))\mathbb{R} \quad (t \rightarrow 0). \quad (6.14)$$

Behaviour in the Generic Case

The generic case is $(l, m, n) = (1, 2, 3)$. The dual curve has a point of main type without a stationary osculating plane at $t = 0$, i.e., the third order osculating space of the curve $\mathbb{R}\mathbf{u}(t)$ at $t = 0$ has dimension three (in fact, this is then true for all t in a neighbourhood of 0). By Lemma 6.2.4, the curve of regression has the parametrization $(1, -3t + o(t), 3t^2 + o(t^2), -t^3 + o(t^3))$ as $t \rightarrow 0$, and therefore has a regular point at $t = 0$ without stationary osculating plane.

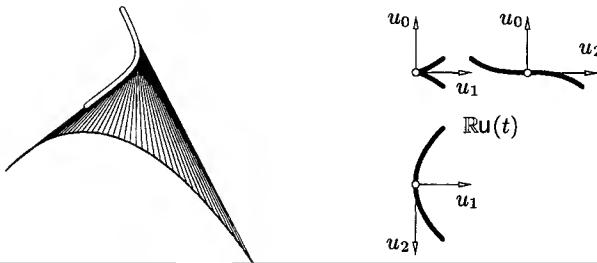


Fig. 6.5. Generic local behaviour of a developable surface — $(l, m, n) = (1, 2, 3)$.

The planar intersection $r(t)$ has exponents $(1, 2)$, which means that it has no inflection point and is regular (Fig. 6.5). We see that the developable surface \mathcal{R} has a regular ruling $R(0)$ with a non-degenerate osculating cone.

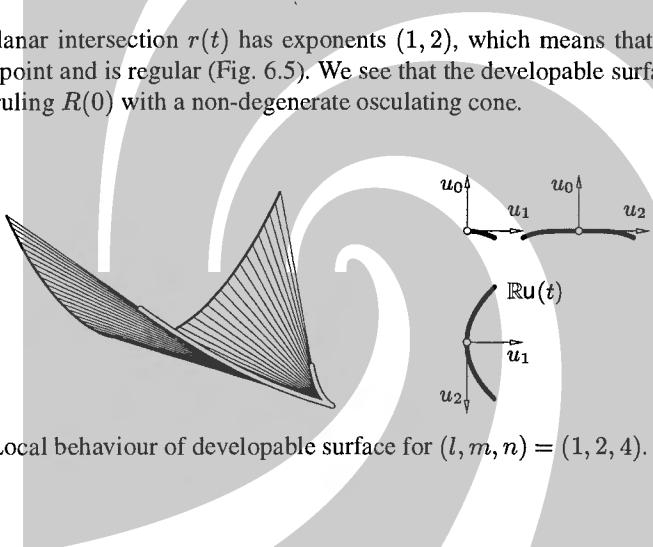


Fig. 6.6. Local behaviour of developable surface for $(l, m, n) = (1, 2, 4)$.

Inflection Generators and Their Dual

Let us now study the case $(l, m, n) = (1, 2, 4)$. This means that the dual curve $U(t)$ has a stationary osculating plane at $t = 0$. Lemma 6.2.4 shows that the Taylor expansion of the curve c of regression starts with exponents $(2, 3, 4)$, and that means that c has a cusp at $t = 0$ (Fig. 6.6). The Taylor expansion of a planar section however starts with $(1, 2)$, and so the generator line $R(0)$ is regular.

Exponents $(l, m, n) = (2, 3, 4)$ mean that the dual curve $U(t)$ has a cusp at $t = 0$. By Lemma 6.2.4, the leading exponents of c at $t = 0$ are $(1, 2, 4)$, which shows that the curve of regression has a stationary osculating plane. Note that this is dual to the previous case.

The Taylor expansion of generic planar intersections starts with $(1, 3)$, which characterizes an inflection point. There is no osculating cone, but the tangent plane osculates the developable surface along the generator $R(0)$ (see Fig. 6.7). This generator is therefore called an *inflection generator* (a detailed discussion of inflection generators has been presented by Maekawa and Chalfant [120]).

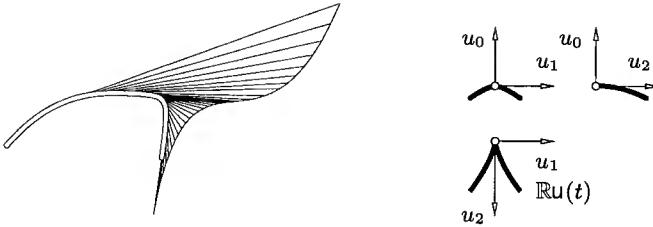


Fig. 6.7. Inflection ruling belongs to a cusp of the dual curve — $(l, m, n) = (2, 3, 4)$.

Singular Cases

An ordinary inflection point of the dual curve $U(t)$ at $t = 0$ is characterized by $(l, m, n) = (1, 3, 4)$. This gives the exponent sequence $(1, 3, 4)$ for the curve of regression, i.e., again an inflection point at $t = 0$. A generic planar intersection has the exponents $(2, 3)$ and thus possesses a cusp at $t = 0$. The corresponding ruling $R(0)$ is therefore singular (see Fig. 6.8).

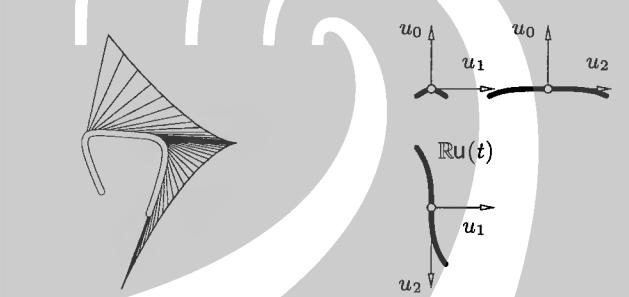


Fig. 6.8. Local behaviour of a developable surface for $(l, m, n) = (1, 3, 4)$.

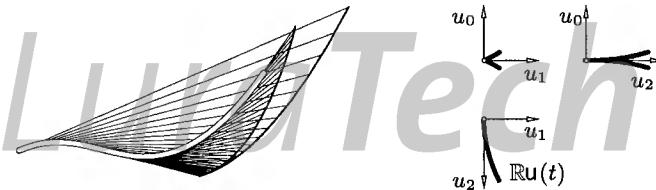


Fig. 6.9. Local behaviour of a developable surface for $(l, m, n) = (2, 4, 5)$.

Finally let us point out that most singularities of the dual curve imply singularities of planar intersections, and therefore a singular generator $R(0)$. For instance, consider the case $(l, m, n) = (2, 4, 5)$. We obtain the triple $(1, 3, 5)$ for $c(t)$, which characterizes a higher order inflection point $c(0)$ of the curve of regression. A generic planar section starts with exponents $(2, 4)$ and has therefore a *ramphoid cusp*, and $R(0)$ is a singular generator of the ‘ramphoid’ type (see Fig. 6.9).

Contact Order

The following theorem shows that contact of order k of developable surfaces is guaranteed if their duals have contact of order k .

Theorem 6.2.5. *Assume two developable ruled surfaces \mathcal{R} , $\overline{\mathcal{R}}$, which are the envelopes of their families of tangent planes $U(t)$ and $\overline{U}(t)$. Assume that the dual curves U , \overline{U} have regular points and no inflections at $t = t_0$, and that they have contact of order k at $t = t_0$. Then the surfaces \mathcal{R} , $\overline{\mathcal{R}}$ have contact of order k in all regular points of the generator $R(t_0)$.*

Proof. (Sketch) We choose a plane W and intersect it with the planes $U(t)$ and $\overline{U}(t)$. The lines $l(t) = W \cap U(t)$ and $\bar{l}(t) = W \cap \overline{U}(t)$ are tangent to the intersections $c = \mathcal{R} \cap W$ and $\bar{c} = \overline{\mathcal{R}} \cap W$. Clearly the family $l(t)$ of lines is the dual of the curve c , and likewise $\bar{l}(t)$ is the dual of \bar{c} . Computing the homogeneous line coordinates of $l(t)$ and $\bar{l}(t)$ with respect to a projective coordinate system in W from the coordinate vectors of $U(t)$ and $\overline{U}(t)$ is a linear operation, so the dual curves $l(t)$ and $\bar{l}(t)$ have contact of order k at $t = t_0$. By our assumptions on $U(t)$, both a generic intersection and its dual are regular. By Lemma 1.2.6, also the curves c and \bar{c} have this property.

We now choose planes W_u , depending smoothly on u , and repeat the construction for all u . Objects depending on u will be denoted by a subscript. We choose W_u such that the point $W_u \cap R(t_0)$ is well defined and locally parametrizes the line $R(t_0)$.

According to the proof of Lemma 1.2.6, there are parameter transforms ψ_u and $\bar{\psi}_u$, depending smoothly on u , such that the first k derivatives of the curves $c_u \circ \psi_u(t)$ and $\bar{c}_u \circ \bar{\psi}_u(t)$ agree (and depend smoothly on u). Then $s(u, t) = c_u(\psi_u(t))$ and $\bar{s}(u, t) = \bar{c}_u(\bar{\psi}_u(t))$ are parametrizations of the surfaces \mathcal{R} and $\overline{\mathcal{R}}$ which agree in their first k derivatives. \square

Remark 6.2.3. In the preceding proof, one could be tempted to write ‘by Lemma 1.2.6, the curves c and \bar{c} have contact of order k . We choose a second plane W' , and construct curves c' , \bar{c}' accordingly. Then the ruled surface parametrizations $R(u) = c(u) \vee c'(u)$ and $\overline{R}(u) = \bar{c}(u) \vee \bar{c}'(u)$ agree in their first k derivatives’. This is false, as the re-parametrizations which effect the same derivatives up to order k need not be the same for both planes W and W' . If this argumentation were true, we would also get a contradiction to Remark 6.1.4.

It should be emphasized that there need not exist ruled surface parametrizations of both \mathcal{R} and $\overline{\mathcal{R}}$ such that derivatives up to order k agree. \diamond

Example 6.2.2. We consider an important special case: A *conic* $V(t)$ of dual space, which has second order contact at $t = t_0$ with the family $U(t)$ of planes, is the dual of a quadratic cone \mathcal{S} . This cone has second order contact with the developable surface \mathcal{R} defined by $U(t)$ in all points of the ruling $R(t_0)$.

The conic $V(t)$ obviously is contained in $U(t)$ ’s osculating plane at $t = t_0$, which implies that the vertex of \mathcal{S} coincides with \mathcal{R} ’s point of regression at $t = t_0$. There are dual curves $U(t)$ which do not possess osculating conics at $t = t_0$. This

happens if and only if $U(t_0)$ is an inflection point. But if there is an osculating conic, there is an entire two-parameter family of conics, which define a two-parameter family of osculating quadratic cones. One of them is the osculating cone of revolution of Th. 6.1.4. \diamond

Remark 6.2.4. If the dual curve $U(t)$ has an ordinary cusp at $t = t_0$, the developable surface \mathcal{R} defined by $U(t)$ has an inflection generator at $t = t_0$ (cf. the discussion following Ex. 6.2.1), and therefore there is an osculating *plane* which is in second order contact with \mathcal{R} at all points of the generator $R(t_0)$. This plane, of course, is tangent to \mathcal{R} in the points of $R(t_0)$ and so equals $U(t_0)$.

The dual curve $U(t)$ has no osculating cone for $t = t_0$ (we may say that the osculating conic has shrunk to the point $U(t_0)$). \diamond

Algebraic Developable Surfaces

We consider an algebraic developable ruled surface $\mathcal{R}\gamma$ and its family $U(t)$ of tangent planes. The following result is presented without proof:

Proposition 6.2.6. *If an algebraic developable surface \mathcal{R} is non-planar and irreducible, then it is either a cone or, where it is regular, the tangent surface of an algebraic curve c . If \mathcal{R} is not planar, the set of tangent planes is an algebraic curve U of dual space, which is planar if and only if \mathcal{R} is a cone.*

We consider all algebraic curves and surfaces over the complex number field and perform complex extensions if necessary. If the degree of U as an algebraic curve in dual projective space equals k , then the generic number of intersection points of U with a plane equals k . For the original ruled surface \mathcal{R} this means that the generic number of tangent planes of \mathcal{R} which are incident with a point, equals k . We say that \mathcal{R} is a developable surface of *class* k .

A converse to Prop. 6.2.6 is also true: The set of tangents of an algebraic curve c in \mathbb{CP}^3 consists of bundles incident with the singular points of c plus an algebraic developable surface, which is called the *tangent surface* of c . We will not show the proof of this fact because it is analogous to the proof of Prop. 7.2.13. The degree of the tangent surface of a curve c is called the *rank* of c .

The classical algebraic geometry of space curves and their tangent developables has been studied extensively. Since we focus on rational curves and surfaces, we do pursue this subject further and refer the interested reader to the survey article by K. Rohn and L. Berzolari [168].

Rational Developable Surfaces

If the family $U(t)$ of tangent planes of a developable ruled surface \mathcal{R} is a *rational* curve of dual projective space, we can assume that the coordinate functions $U(t)$ of $U(t) = \mathbb{R}U(t)$ are polynomials. Equations (6.6) and (6.8) show that then also the family of rulings $R(t)$ and the curve c of regression have rational parametrizations (this includes the special cases of a cone, where c is a point and $U(t)$ is planar).

Conversely a rational developable ruled surface has a polynomial parametrization in homogeneous Plücker coordinates. Equ. (5.6) shows that the families of tangent planes and cuspidal points (points of regression) have rational parametrizations. We conclude the following result, which is important for applications in geometric design.

Theorem 6.2.7. *A developable surface is rational if and only if its dual curve is rational.*

6.2.2 Developable Bézier and B-Spline Surfaces

Everywhere where rational B-splines are used for curve and surface representations, there is a demand for efficiently computing with developable Bézier- and rational B-spline (NURBS) surfaces. Because developable surfaces are the curves of dual space, there are two possible ways to approach the problem: on the one hand we can study Bézier and NURBS surfaces and look for developable surfaces among them. On the other hand we can assume a more abstract viewpoint and simply consider Bézier and NURBS surfaces via dual space. The first possibility amounts to represent developable surfaces as tensor-product surfaces of degree $(1, n)$ (all of which are ruled surfaces) and to describe how to choose the control points such that the surface becomes torsal, i.e., developable. We will first briefly discuss this procedure (cf. [6, 23, 104]) and then focus on the dual representation (cf. [15, 80, 149, 161]).

Developable Bézier Surfaces of Bidegree $(1, n)$

Because developable surfaces are special ruled surfaces, let us start with a rational ruled surface patch expressed as a Bézier surface of bidegree $(1, n)$:

$$\mathbf{x}(u, v) = \sum_{i=0}^1 \sum_{j=0}^n B_i^1(u) B_j^n(v) \mathbf{b}_{i,j}. \quad (6.15)$$

This is a representation in homogeneous coordinates of the form (5.2) with the rational Bézier curves $\mathbf{x}(0, v)\mathbb{R}$ and $\mathbf{x}(1, v)\mathbb{R}$ as directrices. Such a surface is torsal, i.e., developable, if and only if the curve points $\mathbf{x}(0, v)\mathbb{R}$, $\mathbf{x}(1, v)\mathbb{R}$ and the derivative points $\mathbf{x}_v(0, v)\mathbb{R}$, $\mathbf{x}_v(1, v)\mathbb{R}$ are coplanar:

$$\det(\mathbf{x}(0, v), \mathbf{x}(1, v), \mathbf{x}_v(0, v), \mathbf{x}_v(1, v)) = 0. \quad (6.16)$$

Inserting representation (6.15) shows that the developability is expressed by certain nonlinear conditions imposed on the control vectors $\mathbf{b}_{i,j}$ (see Lang and Röschel [104], who also solve (6.16) explicitly in the case $n = 2$). For higher degrees the developability conditions become increasingly complicated.

Remark 6.2.5. The method of solving the nonlinear system of equations which expresses developability is more appropriate than the dual approach (described below) if one wants to find *polynomial* developable surfaces, instead of rational ones (see

[5, 23]). This is because the property of being a polynomial Bézier surface rather than a rational one is not preserved when applying a duality, and it is not so easy to distinguish the duals of the polynomial surfaces among the duals of all rational ones. \diamond

Dual Control Structure of Bézier and NURBS Developables

From Th. 6.2.7 we conclude that a piecewise rational developable surface possesses a piecewise rational dual curve U^* , which means that this dual has a piecewise polynomial parametrization in homogeneous coordinates.

We use the abbreviation NURBS (non-uniform rational B-spline) for such surfaces, because any piecewise polynomial function of degree m can be written as linear combination of normalized B-spline functions $N_i^m(t)$ of degree m over an appropriate knot vector (Th. 1.4.16). The previous paragraph means that the dual of a NURBS (i.e., piecewise rational) developable surface is a NURBS (i.e., piecewise rational) family of planes, and as such a curve of dual projective space.

Such a dual curve $U(t)$ may be written in the form

$$U(t) = \mathbb{R}u(t) = \mathbb{R} \sum_{i=0}^n N_i^m(t) u_i. \quad (6.17)$$

If it is rational (not only piecewise so), we may parametrize $U(t)$ as a dual rational Bézier curve:

$$U(t) = \mathbb{R}u(t) = \mathbb{R} \sum_{i=0}^n B_i^m(t) u_i. \quad (6.18)$$

The vectors $u_i \in \mathbb{R}^4$ are homogeneous coordinate vectors of the *control planes* U_i (Fig. 6.10). In order to encode the vectors u_i in a geometric way, we introduce dual frame points in the familiar way: The *frame planes* $F_i = \mathbb{R}f_i$ are given by

$$f_i = u_i + u_{i+1}, \quad i = 0, \dots, n-1. \quad (6.19)$$

For several algorithms it is convenient to convert (6.17) into piecewise Bézier form and process these Bézier segments further (cf. [46, 78, 142]). Therefore in the following we restrict our discussion to such segments, i.e., to *developable Bézier surfaces*.

Rulings and Curve of Regression

We consider a developable surfaces whose dual is parametrized as a rational Bézier curve $U(t)$. We assume that this dual curve $U(t)$ has control planes U_0, \dots, U_n and frame planes F_0, \dots, F_{n-1} , and we want to compute the Bézier representations of the rulings and the curve of regression.

If a rational Bézier curve has control points b_0, \dots, b_n and frame points f_0, \dots, f_{n-1} , the tangents at $t = 0$ and $t = 1$ are the lines $b_0 \vee b_1$ and $b_{n-1} \vee b_n$, respectively (Th. 1.4.10). Therefore, the rulings $R(0)$ and $R(1)$ are found by

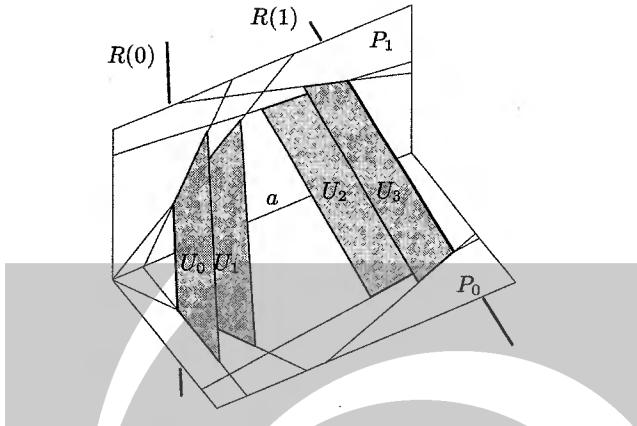


Fig. 6.10. Control planes of a developable Bézier surface.

$$R(0) = U_0 \cap U_1, \quad R(1) = U_{n-1} \cap U_n.$$

(cf. Fig. 6.10). By Th. 1.4.10, the osculating plane of a Bézier curve at $t = 1$ is spanned by the first three control points. The dual version of this statement is that the curve of regression has the end points

$$c(0) = U_0 \cap U_1 \cap U_2, \quad c(1) = U_{n-2} \cap U_{n-1} \cap U_n.$$

Lemma 6.2.8. *Assume a rational developable surface whose family $U(t)$ of tangent planes is given by (6.18). The ruling $R(t)$ has Plücker coordinates*

$$\begin{aligned} (\mathbf{r}(t), \bar{\mathbf{r}}(t)) &= \sum_{k=0}^{2m-2} B_k^{2m-2}(t)(\mathbf{b}_k, \bar{\mathbf{b}}_k) \quad \text{with} \\ (\bar{\mathbf{b}}_k, \mathbf{b}_k) &= \binom{2m-2}{k}^{-1} \sum_{i+j=k} \binom{m-1}{i} \binom{m-1}{j} \mathbf{u}_i \wedge \mathbf{u}_{j+1}. \end{aligned} \quad (6.20)$$

Proof. As $R(t) = \mathbb{R}\mathbf{u}(t) \cap \mathbb{R}\dot{\mathbf{u}}(t)$, the dual Plücker coordinates of $R(t)$ are simply $\mathbf{u}(t) \wedge \dot{\mathbf{u}}(t)$ (cf. Lemma 2.1.4), and we may replace $\mathbf{u}, \dot{\mathbf{u}}$ by any two vectors which span the tangent of the curve $\mathbf{u}(t)$. This is done by the algorithm of de Casteljau, applied to the Bézier curve (6.18): If we denote the intermediate results in the algorithm by $\mathbf{u}_i^j(t)$, then by the Cor. 1.4.7, $R(t)$ has the dual Plücker coordinates

$$\mathbf{u}_0^{m-1}(t) \wedge \mathbf{u}_1^{m-1}(t). \quad (6.21)$$

Eq. (6.20) now follows by expanding

$$\mathbf{u}_0^{m-1}(t) \wedge \mathbf{u}_1^{m-1}(t) = \left(\sum_{i=0}^{m-1} B_i^{m-1}(t) \mathbf{u}_i \right) \wedge \left(\sum_{j=0}^{m-1} B_j^{m-1}(t) \mathbf{u}_{j+1} \right).$$

□

Lemma 6.2.9. Assume a rational developable surface whose family $U(t)$ of tangent planes is given by (6.18). The point $c(t)$ of regression has coordinates

$$c(t) = \mathbf{c}(t)\mathbb{R}, \quad \text{with} \quad \mathbf{c}(t) = \sum_{l=0}^{3m-6} B_l^{3m-6}(t) \mathbf{b}_l, \quad \text{and} \quad (6.22)$$

$$\mathbf{b}_l = \binom{3m-6}{l}^{-1} \sum_{i+j+k=l} \binom{m-2}{i} \binom{m-2}{j} \binom{m-2}{k} \mathbf{u}_i \times \mathbf{u}_{j+1} \times \mathbf{u}_{k+2},$$

where the cross product of three vectors is defined by Ex. 2.2.8.

Proof. We have to compute $c(t) = \mathbf{u}(t) \times \dot{\mathbf{u}}(t) \times \ddot{\mathbf{u}}(t)$. Similarly to Cor. 1.4.7, it is easy to show that the points $\mathbf{u}_0^{m-2}, \mathbf{u}_1^{m-2}, \mathbf{u}_2^{m-2}$, which occur in the $(m-2)$ -th step of the algorithm of de Casteljau, span the osculating plane of the curve $\mathbf{u}(t)$. Thus

$$\mathbf{c}(t) = \mathbf{u}_0^{m-2}(t) \times \mathbf{u}_1^{m-2}(t) \times \mathbf{u}_2^{m-2}(t). \quad (6.23)$$

Lemma 1.4.2 shows that $\mathbf{u}_j^{m-2}(t) = \sum_{i=0}^{m-2} B_i^{m-2}(t) \mathbf{u}_{i+j}$. Expanding the cross product shows Equ. (6.22). \square

Degree and Class of Rational Developables

In general, Equ. (6.18) describes an algebraic rational developable surface of class m . Equ. (6.20) and Th. 5.2.8 show that the degree of the underlying point set is at most (and in general, equals) $2m - 2$. The degree is lower than $2m - 2$ if $(\mathbf{r}(t_0), \bar{\mathbf{r}}(t_0)) = (\mathbf{o}, \mathbf{o})$ for some t_0 , because then $\frac{1}{t-t_0}(\mathbf{r}(t_0), \bar{\mathbf{r}}(t_0))$ parametrizes the same algebraic variety. This occurs if and only if the vectors $\mathbf{u}_0^{m-1}(t_0)$ and $\mathbf{u}_1^{m-1}(t_0)$ of Equ. (6.21) are linearly dependent. The dual curve $U(t)$ then has a singularity at $t = t_0$.

Equ. (6.22) shows that the curve of regression is a rational curve of degree $\leq 3m - 6$. The degree is lower if the parametrization (6.22) has ‘base points’, i.e., $\mathbf{c}(t) = \mathbf{o}$. Equ. (6.23) shows that this occurs if and only if the vectors $\mathbf{u}_0^{m-2}, \mathbf{u}_1^{m-2}$, and \mathbf{u}_2^{m-2} are linearly dependent. The dual curve $U(t)$ then has a singularity or an inflection point, because the span of these three vectors equals $U(t) \vee U^1(t) \vee U^2(t)$. The curve of has a stationary osculating plane in this case.

Projective Invariance of the Geometric Control Structure

Lemma 1.4.12 shows that the geometric control polygon of a rational Bézier curve is invariant with respect to projective mappings. The proof does not use that this mapping is an automorphism. The result therefore is valid for *singular* projective mappings (*linear mappings*) $x\mathbb{R} \mapsto (A \cdot x)\mathbb{R}$ with A singular as well. We will see that central projections are examples of such mappings. It is interesting to note what we can get out of this result by dualizing the notion of ‘projection’:

A projection $\pi : P^3 \rightarrow P^2$ is defined by a projective plane P^2 embedded into projective three-space, and a center Z which is not contained in P^2 . Then

$$\pi(X) = (Z \vee X) \cap P^2, \quad (6.24)$$

i.e., we intersect the ray $Z \vee X$ with the plane P^2 . Dualization changes P^3 to its set of planes, changes P^2 to a point P^{2*} , and converts Z into a plane Z^* . The dual projection π^* then is the mapping of planes

$$\pi^*(X^*) = (Z^* \cap X^*) \vee P^{2*}, \quad (6.25)$$

In order to proceed at this abstract level, we recall that a *bundle of lines* is a projective plane, and a *bundle of planes* as well (they are their respective duals). An isomorphism κ of the line bundle and an ‘ordinary’ projective plane is provided by intersecting the lines of the bundle with a plane in three-space. An isomorphism κ^* of the plane bundle and an ‘ordinary’ dual projective plane is analogously provided by intersecting the planes of the bundle with a plane in three-space.

By applying appropriate isomorphisms after the projections π and π^* , we get two further mappings:

$$\kappa^{-1}\pi(X) = ((Z \vee X) \cap P^2) \vee Z = Z \vee X, \quad (6.26)$$

$$\kappa^{*-1}\pi^*(X^*) = ((Z^* \cap X^*) \vee P^{2*}) \cap Z^* = Z^* \cap X^*. \quad (6.27)$$

This shows the following:

Lemma 6.2.10. *We consider projective three-space P^3 . We choose a point Z and a plane P^2 with $Z \notin P^2$. X denotes a point and X^* a plane. Then the mappings*

$$\begin{aligned} \pi_1 : X &\mapsto Z \vee X, & \pi_2 : X &\mapsto (Z \vee X) \cap P^2, \\ \pi_3 : X^* &\mapsto P^2 \cap X^*, & \pi_4 : X^* &\mapsto (P^2 \cap X^*) \vee Z, \end{aligned}$$

are singular projective mappings.

Proof. We already saw (cf. (6.24), (6.25), (6.26), (6.27)) that π_2 and π_4 are dual versions of each other, and so are π_1 and π_3 . Moreover, (6.26) shows that π_1 is a projective mapping if and only if π_3 is. It is therefore sufficient to show that π_3 is a singular projective mapping. This follows immediately from Lemma 1.1.19. \square

Corollary 6.2.11. *The geometric control polygon of a rational Bézier curve is invariant with respect to the mappings π_1 and π_3 of Lemma 6.2.10. The geometric control structure of a dual rational Bézier curve (i.e., a developable Bézier surface) is invariant with respect to the mappings π_2 , π_4 of Lemma 6.2.10.*

Planar Sections of Rational Developables

We apply Cor. 6.2.11 in order to compute the geometric control structure of planar sections of rational developable surfaces:

Proposition 6.2.12. *Assume that $U(t)$ is the family of tangent planes of a developable Bézier surface, with control planes U_i and frame planes F_i . The intersection of this surface with a plane P is a Bézier curve $c(t)$. The control lines and frame lines of c 's dual representation are the lines $U_i \cap P$ and $F_i \cap P$.*

Proof. We consider the mapping of planes $\pi_3 : X^* \mapsto P \cap X^*$. The family of lines $\pi_3(U(t))$ is the dual of the intersection curve. By Cor. 6.2.11, the result follows. \square

Remark 6.2.6. It is possible to show this in a more direct way, by repeating the proof of Lemma 6.2.10: Choose a coordinate system such that P has equation $x_3 = 0$. The intersection of the plane $\varepsilon = \mathbb{R}(u_0, \dots, u_3)$ with $x_3 = 0$ has the equation $x_3 = u_0x_0 + u_1x_1 + u_2x_2 = 0$. If we introduce line coordinates in P this means that $P \cap \varepsilon$ has coordinates (u_0, u_1, u_2) in P .

This shows that computing the intersection means forgetting the x_3 -component, which obviously holds also for the control vectors: Equ. (6.18) is thus converted into (1.99) (see Fig. 6.10). \diamond

Cones

A cone is a special developable surface. It is the envelope of a family $U(t)$ of planes, which are incident with a fixed point v , the cone's *vertex*. It is easily seen (e.g., by multiple application of Prop. 6.2.12) that a developable rational surface is a cone with vertex v if and only if all control planes are incident with v .

We want to study the *director cone* of a developable surface, defined by a family $U(t)$ of tangent planes. Its vertex is the origin, and its rulings are parallel to those of the original surface, by definition. Lemma 5.3.2 shows that also its tangent planes are tangent to those of the original surface. The following describes how to find the geometric control figure of the director cone:

Corollary 6.2.13. *Assume that the dual geometric control figure of a developable Bézier surface consists of the control planes U_i and the frame planes F_i . Then the control and frame planes \bar{U}_i, \bar{F}_i of its director cone are parallel to U_i and F_i , respectively, and incident with the origin.*

Proof. Denote the origin with O and the plane at infinity with ω . We consider the mapping of planes $\pi_4 : X^* \mapsto (X^* \cap \omega) \vee O$. Obviously π_4 maps the family of tangent planes of the developable surface to those of the director cone, and also U_i, F_i to \bar{U}_i, \bar{F}_i , respectively. Now we apply Cor. 6.2.11. \square

Convexity of Developable Surfaces

For practical applications it is important to know where the singularities of a developable surface are, i.e., where its curve of regression is located. In most cases a surface patch which is free of singularities will be *convex* in the sense described by the next corollary (see Fig. 6.11).

Corollary 6.2.14. *Consider a developable Bézier surface patch bounded by two rulings and two curves in planes P_0, P_1 . This patch is (i) free of singularities and (ii) contained in the boundary of a convex domain, if the control planes and P_0, P_1 carry faces of a convex polyhedron Π and the frame planes are support planes of Π .*

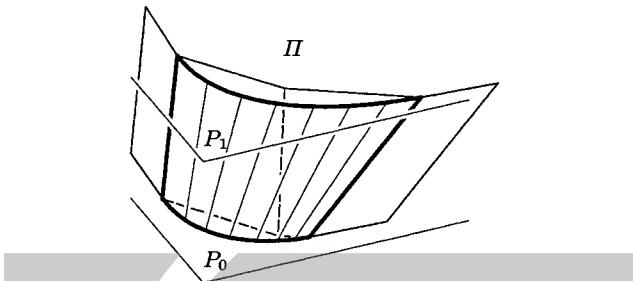


Fig. 6.11. Regular convex developable surface patch.

Proof. By Prop. 6.2.12, intersecting the dual control structure of the developable surface with P_0, P_1 gives the dual control structure of the intersection curves. Our condition means that these planar control structures satisfy the requirements of Cor. 1.4.19. \square

Remark 6.2.7. Surfaces which fulfill the condition in Cor. 6.2.14 cannot have *inflection generators* except at their boundary. Since an inflection generator belongs to a singularity of the dual representation, we can ensure an inflection ruling $R(0)$ if we choose $U_0 = U_1$ and $U_2 \supset R(0)$. If we are modeling a developable surface as a piecewise Bézier surface, inflection generators are most easily modeled by choosing the boundaries of adjacent patches there. \diamond

Conversion to Tensor Product Representation

The representation of developable surfaces by their duals is, although elegant, not supported by most software available today. Developable surfaces have to be represented as trimmed NURBS surface patches. Thus we give formulae to compute an ‘ordinary’ surface parametrization from the rational family $U(t)$ of tangent planes.

This is very easy for a surface patch whose boundary consists of the rulings $R(0)$, $R(1)$, and the intersection curves $c_0(t)$, $c_1(t)$ with two planes P_0 , P_1 (see Fig. 6.10). By Prop. 6.2.12, we get the dual control structures of $c_0(t)$ and $c_1(t)$ by intersecting the control and frame planes of $U(t)$ with P_0 and P_1 , respectively. Via the conversion formula (1.103), we compute the homogeneous coordinate vectors of the Bézier points $b_{0,0}, \dots, b_{0,2m-2}$ of $c_0(t)$ and $b_{1,0}, \dots, b_{1,2m-2}$ of $c_1(t)$.

For all t , the ruling $R(t)$ is spanned by the two points $c_0(t)$ and $c_1(t)$. This shows that the surface possesses the homogeneous parametrization

$$s(t, u) = (1 - u)c_0(t) + uc_1(t) = \sum_{i=0}^1 \sum_{j=0}^{2m-2} B_i^1(u) B_j^{2m-2}(t) b_{i,j}, \quad (6.28)$$

which is already a parametrization of the surface as a TP surface. Thus the control points of the planar intersections determine the control net of the tensor product

patch of bidegree $(1, 2m - 2)$ (Fig. 6.12). The degree of the planar intersection and thus the degree of the patch may be less than $2m - 2$; this happens for instance if the intersecting plane P contains rulings of the surface.

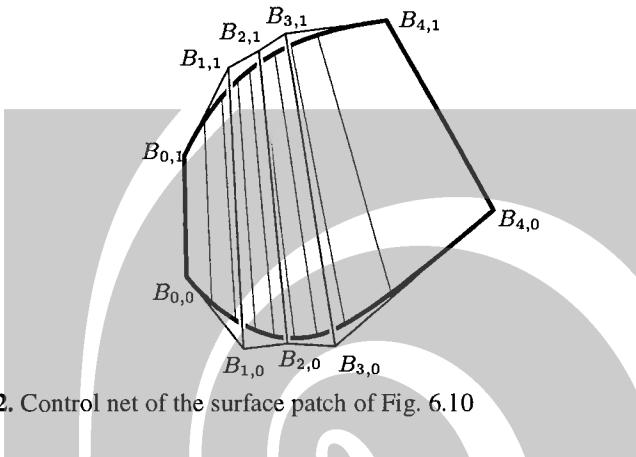


Fig. 6.12. Control net of the surface patch of Fig. 6.10

Remark 6.2.8. In case that non-planar boundary curves are necessary, trimmed patches can be used. Here one draws the preimages of the boundary curves in the parameter domain as special Bézier curves via appropriate polynomial functions $u(t)$ and maps them onto the surface. This raises the degree to $n > 2m - 2$, but we still have a Bézier tensor product representation of bidegree $(1, n)$. \diamond

Developable Surfaces with Quadratic Dual

The simplest nontrivial case is that of a surface whose family of tangent planes is a *quadratic curve*, i.e., a conic ($m = 2$). Such surfaces (dual to conics) are *quadratic cones*, which is to be understood as a term of projective geometry and includes quadratic cylinders. The three Bézier planes intersect in the vertex of the cone. The curve of regression is degenerate and consists of the vertex only. Note that our formula for the order of the regression curve shows this degeneracy, because $3m - 6 = 0$.

Patches confined by planar intersections are of degree $(1, 2)$, and clearly the degree of the surface as an algebraic variety equals $2m - 2 = 2$ (Fig. 6.13).

Developable Surfaces with Cubic Dual

The simplest case with a nondegenerate curve of regression is that of a cubic family of tangent planes ($m = 3$). The degree of these surfaces as algebraic varieties in general equals $2m - 2 = 4$ and patches confined by planar intersections are of bidegree $(1, 4)$ as tensor product surfaces. The curve of regression is in general a cubic curve, since $3m - 6 = 3$. This means that these surfaces are *tangent surfaces of twisted cubics* (see Fig. 6.14).

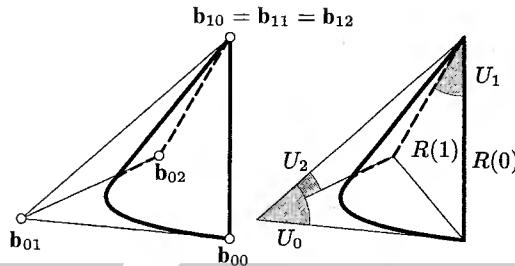


Fig. 6.13. Control points and control planes of a conical patch.

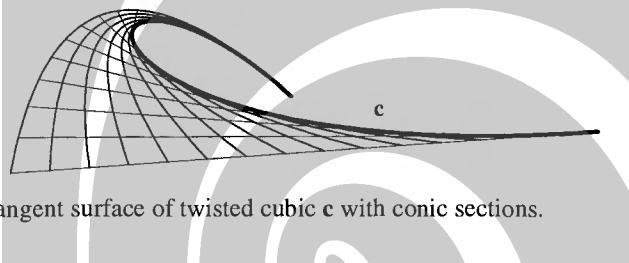


Fig. 6.14. Tangent surface of twisted cubic c with conic sections.

Remark 6.2.9. We choose a ruling $R(t_0)$ and a plane V which contains this ruling, and intersect the developable surface with V . The intersection curve is an algebraic curve (in general of degree four) which contains the line $R(t_0)$ and will therefore be reducible. The irreducible components will in general be the line $R(t_0)$ and a planar cubic curve.

If the plane V equals the tangent plane $U(t_0)$, then the intersection will contain $R(t_0)$ with multiplicity two and the remaining irreducible component of the intersection will in general be a conic.

This is also easily confirmed by direct computation. Recall our previous discussion of quartic ruled surfaces generated by a projective isomorphism of two conics. The present developable surfaces are special examples of this. \diamond

The presence of conics on developable surfaces of class $m = 3$ allows to construct Bézier patches of bidegree $(1, 2)$ on them. The boundary conics lie in tangent planes P_0, P_1 of the surface. The intersection curve $c_0(t)$ in P_0 must touch P_1 and conversely $c_1(t)$ in P_1 must touch P_0 . This means that the two conics touch the line $A = P_0 \cap P_1$ (see Fig. 6.15). The points of tangency do not coincide unless the surface degenerates to a quadratic cone. The simple geometry of the control net is described in Fig. 6.15. It is based on the generation of the surface by a projective mapping between c_0 and c_1 and on the existence of a constant tangent plane along a ruling.

Remark 6.2.10. We can construct such a patch with prescribed end rulings plus tangent planes. Therefore these surface patches are well suited for approximation of arbitrary developable surfaces: select a discrete set of rulings plus tangent planes

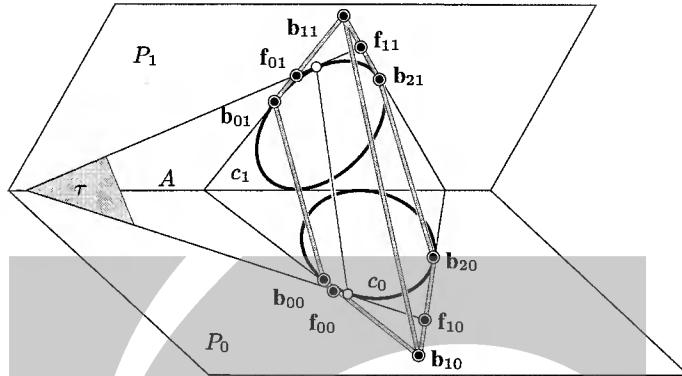


Fig. 6.15. Construction of a developable Bézier surface of bidegree (1,2).

from the given surface, and interpolate these data with developable surface segments of the kind described here (see [149], and also Sec. 6.2.3). \diamond

6.2.3 Interpolation and Approximation Algorithms with Developable Surfaces

A Class of Developable Surfaces Suited for Approximation

For approximation algorithms, it will be convenient to exclude special developable surfaces and restrict ourselves to a class of surfaces which do not contain ‘vertical’ tangent planes, where ‘vertical’ is defined with respect to some coordinate system. We consider surfaces whose family of tangent planes $U(t) = \mathbb{R}u(t)$ is of the form

$$u(t) = (u_0(t), u_1(t), u_2(t), -1). \quad (6.29)$$

This excludes tangent planes of the form $U = (u_0, u_1, u_2, 0)$. If the underlying projective coordinate system comes from a Cartesian coordinate system with axes x , y , and z , this means that planes parallel to the z -axis are excluded (‘vertical planes’).

Since both the Bernstein polynomials and the B-spline basis functions $N_i^m(t)$ form a partition of unity (Lemma 1.4.1 and Th. 1.4.14), the choice of

$$u_i = (u_{0,i}, u_{1,i}, u_{2,i}, -1)$$

ensures that both $\sum N_i^m(t)u_i$ and $\sum B_i^m(t)u_i$ are of the form (6.29). We will restrict ourselves to such control planes, and write $U = (u_0, u_1, u_2, -1)$.

The obstacle that these inhomogeneous plane coordinates do not work if the last coordinate is zero is easily overcome in practice by choosing an appropriate coordinate system.

Dual projective space without the bundle of ‘vertical’ planes $(u_0, u_1, u_2, 0)$ is an affine space, and (u_0, u_1, u_2) are affine coordinates in it. The surfaces (6.29) are ordinary polynomial B-spline curves in this affine space.

A Parametrization of Developable Surfaces Suited for Approximation

For several applications it is convenient to parametrize the surfaces in question in a special way. We consider families of planes of the form

$$U(t) = \mathbb{R}(u_0(t), u_1(t), t, -1). \quad (6.30)$$

The generators $R(t)$ of the developable surface defined by $U(t)$ are contained in the first derivative planes $U^1(t)$ which have the form

$$U^1(t) = \dot{U}(t) = \mathbb{R}(\dot{u}_0, \dot{u}_1, 1, 0). \quad (6.31)$$

Remark 6.2.11. Surfaces such that their family of tangent planes is given by (6.30) do not possess inflection generators and generators parallel to the plane $x = 0$. This is clear because the generator $R(t_0)$ is an inflection generator if and only if the dual curve $U(t)$ has a singularity. If $U(t)$ is of type (6.29) this happens if and only if $\dot{U}(t) = 0$, and it never happens for surfaces of type (6.30).

Further, direct calculation gives the intersection point of the three planes $U(t)$, $\dot{U}(t)$ and $x = 0$, which never is situated at infinity.

Because inflection generators are singularities in dual space, they need a special treatment anyway. In approximation problems we will have to cut the surface which we want to approximate into pieces which contain inflection generators at the boundary. \diamond

Let us intersect the surface (6.30) with the plane $x = c$:

Lemma 6.2.15. *We assume that $U(t)$ is a developable NURBS surface type (6.30) over an appropriate knot vector T , and that it has degree k . Then its intersection curve \mathbf{c}_c with the plane $x = c$ is a polynomial B-spline curve of degree k over a knot vector which contains the same knots as T , but with multiplicities increased by 1.*

Proof. The curve \mathbf{c}_c is the envelope of the lines $z = u_0 + cu_1 + ty$. Thus the geometric meaning of the parameter t is the *tangent slope* of the intersection curves \mathbf{c}_c . An elementary calculation gives the parametric representation of \mathbf{c}_c :

$$x = c, \quad y = -\dot{h}_c(t), \quad z = h_c(t) - t\dot{h}_c(t), \quad \text{with } h_c(t) := u_0(t) + cu_1(t). \quad (6.32)$$

We see that these are *polynomial* B-spline curves. The function \dot{h}_c is of polynomial degree less or equal k and its differentiability class at the knot values is one less than the differentiability class of the u_i . Thus Th. 1.4.15 shows the result. \square

Corollary 6.2.16. *A developable NURBS surface (6.30) can also be written as a polynomial tensor product B-spline surface of degree $(1, k)$,*

$$\mathbf{s} = (1-u)\mathbf{c}_a(t) + u\mathbf{c}_b(t), \quad (6.33)$$

where \mathbf{c}_a and \mathbf{c}_b are the intersection curves with planes $x = a$ and $x = b$ according to Lemma 6.2.15.

The computational simplicity of the surfaces whose dual $U(t)$ is of the form (6.30) is based on the following fact: Both $U(t)$ and the parametrization (6.33) are *polynomial* B-spline curves or surfaces. This facilitates the construction of linear approximation algorithms, as will be shown below.

Interpolation

Projective duality allows to treat developable surfaces like curves. This implies that any algorithm which computes interpolating curves in a projectively invariant way also solves the appropriate dual interpolation problem for developable surfaces.

An algorithm for interpolating points (plus tangents) with a NURBS curve yields an algorithm for the computation of a NURBS developable surface interpolating tangent planes (plus rulings). As an example, we have presented developable Bézier surfaces of bidegree $(1, 2)$, and Remark 6.2.10 shows how they can be applied in a Hermite like scheme for interpolating a sequence of rulings plus tangent planes.

Let us briefly discuss an algorithm for interpolating rulings plus tangent planes with a developable NURBS surface (6.17). The input planes Y_j are described by homogeneous plane coordinates y_j , ($j = 0, \dots, N$). They shall be interpolated by a family $U(t)$ of tangent planes such that for given parameter values t_j we have $U(t_j) = Y_j$, and $U(t)$ is a dual NURBS curve of degree n . This means

$$\tau_j y_j = \sum_{i=0}^n N_i^m(t_j) u_i, \quad j = 0, \dots, N. \quad (6.34)$$

In order to encode the information about rulings $R_j \subset Y_j$ with planes only, we choose additional input planes H_j with $H_j = \mathbb{R} h_j$ such that $R_j = H_j \cap V_j$. Computing the ruling $R(t)$ of the surface defined by $U(t)$ means that we have to intersect $U(t)$ with its first derivative planes $U^1(t)$. So we interpolate the given rulings if H_j , $U(t_j)$, and $U^1(t_j)$ have a line (i.e., R_j) in common:

$$\lambda_j y_j + \mu_j h_j = \sum_{i=0}^n \dot{N}_i^m(t_j) u_i, \quad j = 0, \dots, N. \quad (6.35)$$

These conditions are *linear* in the unknown plane coordinate vectors u_i of the NURBS surface and the unknown homogeneity factors λ_j , μ_j , τ_j . By the homogeneity of these factors, we may choose one of them, e.g., we let $\tau_0 = 1$. Thus, we arrive at $4(2N + 2)$ linear equations for the $4(n + 1)$ plane coordinate values and the $3(N + 1) - 1$ homogeneity factors. If

$$5N = 4n - 2, \quad (6.36)$$

the number of equations equals the number of unknowns and one can hope for a unique solution. For more details and similar interpolation problems we refer the reader to the literature [79, 80].

Distance Functions for Planes

In order to solve approximation problems in the set of planes, it is necessary to define an appropriate *distance* between two planes. Euclidean geometry does not directly provide such a distance function. All Euclidean invariants of pairs of planes are expressed in terms of the angle between planes and are inappropriate for our purposes. In view of applications, we are interested in the distances of points of the two planes which are near some region of interest, and these distances can become arbitrarily large with the angle between planes getting arbitrarily close to zero at the same time.

The set of planes of projective three-space is itself a projective space. For our purposes it will be useful to remove a bundle of planes (which will be the ‘plane at infinity’ of the dual space) in order to get an affine space of planes, and to introduce a Euclidean metric there.

As vertex of this bundle we choose the ideal point of the z -axis, so the bundle consists of all planes parallel to the z -axis. Planes not contained in this bundle have an equation of the form

$$z = u_0 + u_1x + u_2y, \quad (6.37)$$

i.e., their homogeneous plane coordinates have the form $\mathbf{u} = (u_0, u_1, u_2, -1)$. We see that (u_0, u_1, u_2) are affine coordinates in the affine space A^* of planes not parallel to the z -axis.

We will now introduce a Euclidean metric in A^* : For any positive measure μ in \mathbb{R}^2 we define the distance d_μ between planes $A : z = a_0 + a_1x + a_2y$ and $B : z = b_0 + b_1x + b_2y$ as

$$\begin{aligned} d_\mu(A, B)^2 &= \|((a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)y)\|_{L^2(\mu)}^2 \\ &= \int_D ((a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)y)^2 d\mu, \end{aligned} \quad (6.38)$$

i.e., the distance of the functions $z = a_0 + a_1x + a_2y$ and $z = b_0 + b_1x + b_2y$ in $L^2(\mu)$. Of course this makes only sense if the linear function which represents the difference between the two planes is in $L^2(\mu)$. We will always assume that the measure μ is such that all linear and quadratic functions possess finite integral.

A useful choice for μ is the Lebesgue measure $dxdy$ times the characteristic function χ_D of a *region of interest* D in the xy -plane (see Fig. 6.16). In that case we have

$$d_\mu(A, B)^2 = \int_D ((a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)y)^2 dxdy, \quad (6.39)$$

and we write $d_D(A, B)$ instead of $d_\mu(A, B)$. The subspace of $L^2(\mu)$ which contains all constant and linear functions has the basis $z(x, y) = 1$, $z(x, y) = x$ and $z(x, y) = y$. If we denote the coordinate matrix of the L^2 scalar product with respect to this basis by $G = (g_{ij})$, then

$$\begin{aligned} g_{11} &= \int 1 d\mu, & g_{12} &= \int x d\mu, & g_{13} &= \int y d\mu, \\ g_{22} &= \int x^2 d\mu, & g_{23} &= \int xy d\mu, & g_{33} &= \int y^2 d\mu, \end{aligned}$$

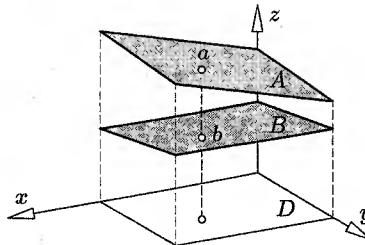


Fig. 6.16. To the definition of the deviation of two planes.

$$d\mu(A, B)^2 = (a_0 - b_0, a_1 - b_1, a_2 - b_2) \cdot G \cdot (a_0 - b_0, a_1 - b_1, a_2 - b_2)^T. \quad (6.40)$$

Another possibility is that μ equals a sum of point masses at points (x_j, y_j) (see [79]). In this case we have

$$d_\mu(A, B)^2 = \sum_j ((a_0 - b_0) + (a_1 - b_1)x_j + (a_2 - b_2)y_j)^2. \quad (6.41)$$

It is easy to show (see [161]) that the distance d_μ defines a Euclidean metric in the set of planes of type (6.37), if and only if μ is not concentrated in a straight line. In this way, approximation problems in the set of planes (with a bundle removed) are transformed into approximation problems in the set of points in Euclidean three-space, whose metric is based on some measure, defined in the region of interest. In the following, we will illustrate this by means of approximation with developable surfaces.

Approximation of a Sequence of Planes by a Developable Surface

Consider the following approximation problem. Given m planes V_1, \dots, V_m and corresponding parameter values v_i , approximate these planes by a developable surface $U(t)$, such that $U(v_i)$ is close to the given plane V_i .

The geometric meaning of ‘close’ is the following: There is a Cartesian coordinate system fixed in space such that all planes are graphs of linear functions of the xy -plane. If the distribution of the normal vectors of the given planes is uneven enough, it makes sense to determine the z -axis as best approximation of the given planar normals. For all i there is a region of interest D_i , or, more generally, a measure μ_i in the xy -plane. We want to find a developable surface with dual $U(t)$ such that

$$F_1 := \sum_{i=1}^m d_{\mu_i}(V_i, U(v_i))^2 \quad (6.42)$$

is minimal. If $U(t)$ is a NURBS surface of type (6.29), F_1 is a quadratic function in the unknown coordinates of the control planes U_i . These can then be found by solving a linear system of equations. An example of this can be seen in Fig. C.8.

Approximation of Planes, Lines and Points by Developable Surfaces

Let us now discuss how to extend the approximation method of the previous paragraph such that we can find developable surfaces whose tangent planes are close to given planes, whose generators are close to given lines, and which are close to given points.

We measure the distance $\delta(g_1, g_2)$ of two lines in analogy to Equ. (6.38), i.e., by considering both lines as graphs of functions whose domain is the real line.

We assume that the projections of g_1 and g_2 into the xy -plane have the equations $d_0 + k_0 x + y = 0$ and $d_1 + k_1 x + y = 0$, and define

$$\delta_\mu(g_1, g_2)^2 = \int (d_0 - d_1 + (k_1 - k_2)x)^2 d\mu(x), \quad (6.43)$$

where μ is some positive measure on the real line such that all linear and quadratic functions possess finite integral.

If $U(t)$ is the family of tangent planes of a developable surface, the rulings are found by $R(t) = U(t) \cap U^1(t)$. If $U(t)$ is parametrized according to (6.30), we have $U^1(t) = \mathbb{R}(\dot{u}_0, \dot{u}_1, 1, 0)$, which immediately shows that the orthogonal projection of $R(t)$ into the xy -plane has the equation $\dot{u}_0 + \dot{u}_1 x + y = 0$, and we have

$$\delta_\mu(g_0, R(t))^2 = \int (d_0 - \dot{u}_0(t) + (k_0 - \dot{u}_1(t))x)^2 \mu(x). \quad (6.44)$$

Given m tangent planes V_i plus generators g_i , we can approximate these data by a NURBS surface of the form (6.30) as follows. After an appropriate segmentation (see the discussion above) and the choice of local coordinate systems, the plane coordinates $V_i = (v_{0,i}, v_{1,i}, v_{2,i}, -1)$ with $v_{2,i} \neq v_{2,j}$ if $i \neq j$, already determine the parameters $t_j = v_{2,j}$ which have to be used in formulas like (6.42). Then, with appropriate measures μ_i from (6.42) and measures μ'_i from (6.44), we define the quadratic function

$$F_2 = \sum_{i=1}^m (d_{\mu_i}(V_i, U(v_{2,i}))^2 + \alpha_i \delta_{\mu'_i}(g_i, R(v_{2,i}))^2). \quad (6.45)$$

The parameters $\alpha_i > 0$ control the influence of the data lines. Minimization of F_2 amounts to the solution of a linear system of equations in the unknown coefficients of the B-spline functions $u_0(t)$ and $u_1(t)$.

Remark 6.2.12. The distance which we defined for lines obviously does not distinguish between lines which have the same orthogonal projection in the xy -plane. Because the generator $R(t)$ is contained in the plane $U(t)$ anyway, this does not matter. \diamond

Remark 6.2.13. If the dual of a developable surface is of type (6.30), then its curve of regression is computed by $c(t) = U(t) \cap U^1(t) \cap U^2(t) = (u_0, u_1, t, 1) \times (\dot{u}_0, \dot{u}_1, 1, 0) \times (\ddot{u}_0, \ddot{u}_1, 0, 0) \mathbb{R} = (-\ddot{u}_1, \ddot{u}_0, c_2, c_3) \mathbb{R}$, which means that the x -coordinate of the point $c(t)$ equals $-\ddot{u}_0/\ddot{u}_1$.

If we want to approximate under the side condition that the curve of regression does not enter the region $a \leq x \leq b$, this means that we must approximate under the side condition that $a \leq -\ddot{u}_0/\dot{u}_1 \leq b$.

As the set of points (x, y) with $a \leq -x/y \leq b$ is the union of convex regions, Th. 1.4.14 and the fact that the second derivative of a B-spline curve is again a B-spline curve show that minimizing F_1 or F_2 with that side condition requires the minimizing of a quadratic form on the union of two convex polyhedra in an appropriate \mathbb{R}^d . For more details, we refer to Pottmann and Wallner [161]. Fig. 6.17 shows an example. \diamond

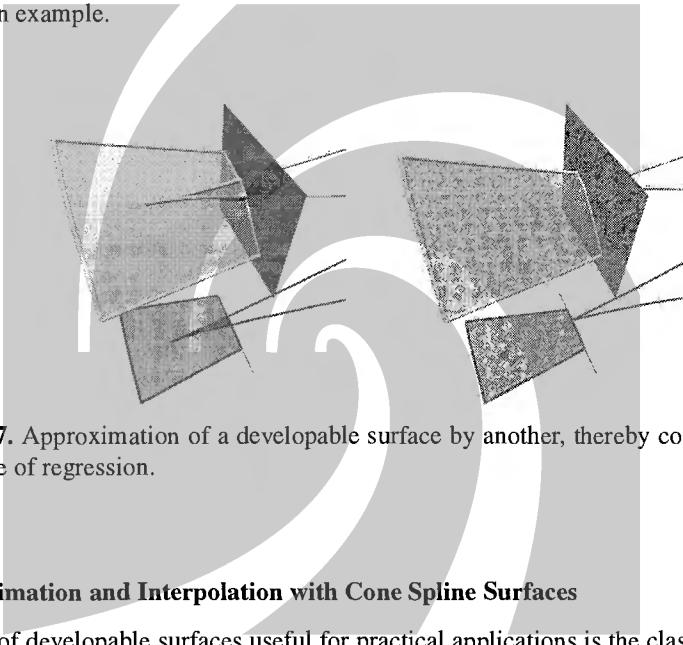


Fig. 6.17. Approximation of a developable surface by another, thereby controlling the curve of regression.

Approximation and Interpolation with Cone Spline Surfaces

A class of developable surfaces useful for practical applications is the class of surfaces composed of smoothly joined cones of revolution. These *cone spline surfaces* are of degree 2, the development can be computed in an elementary way and the offsets are of the same type. Various interpolation and approximation algorithms for these surfaces have been developed [25, 109, 110]. They work mainly with elementary geometric considerations instead of the dual representation (see Fig. 6.18).

6.3 Developable Surfaces of Constant Slope and Applications

A developable surface is said to be of *constant slope* if its rulings enclose the constant angle $\gamma \in (0, \pi/2)$ with some reference plane Π . We will see that this is equivalent to saying that all tangent planes enclose the same angle γ with the reference plane. We will also call such developable surfaces γ -*developables*. The variety of their applications and their relations to problems in geometric computing is remarkable. We will therefore discuss them in more detail and outline applications



Fig. 6.18. Developable surfaces which consist of pieces of cones of revolution and their development (courtesy S. Leopoldseder).

such as computing the medial axis, geometrical optics, designing rational curves with rational offsets and geometric tolerancing.

6.3.1 Basics

Consider a developable surface \mathcal{R} whose family of tangent planes is parametrized in the form $U(t)$. All planes enclose the angle γ with the reference plane Π . In order to support the imagination, we consider Π horizontal and choose it as the xy -plane of a Cartesian coordinate system.

The simplest example of such a surface is that of a *cone* of constant slope. It might be planar, but if it is not, it is a right circular cone with vertical axis. Its rulings likewise enclose the angle γ with Π . We call such cones γ -*cones* in the following.

Lemma 6.3.1. *The tangent planes of a smooth non-planar developable surface \mathcal{R} enclose the angle γ with the horizontal plane Π , if and only if their rulings do.*

Proof. If the surface is a cone, the lemma is true. For general surfaces we use the fact that, by Lemma 5.3.2, the director cone has rulings parallel to the rulings of \mathcal{R} , and tangent planes parallel to the tangent planes of \mathcal{R} . \square

As planar developables are an exceptional case also with respect to the dual representation, we do not consider them and give the following definition:

Definition. A torsal surface \mathcal{R} is called a *developable of constant slope γ* with respect to a base plane Π , if it is not planar; and if the rulings of \mathcal{R} enclose the angle γ with the plane Π .

Clearly the tangent surface of a curve c is of constant slope if and only if all of c 's tangents enclose the same angle γ with the horizontal plane Π . Such a curve is called a *curve of constant slope* or *generalized helix*.

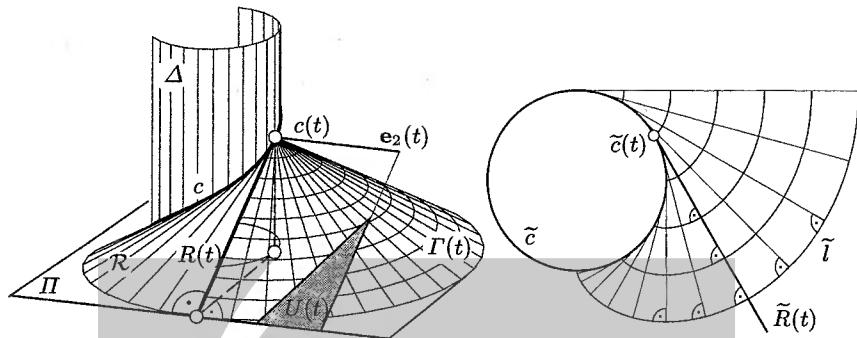


Fig. 6.19. Properties of a developable surface of constant slope. Axonometric view (left) and top view (right)

Development of the Projection Cylinder

Assume a developable surface of constant slope \mathcal{R} with tangent planes $U(t)$, rulings $R(t)$, and curve of regression $c(t)$. The vertical (z -parallel) lines through the points $c(t)$ form a cylindrical surface Δ , which is called the projection cylinder of the curve $c(t)$ (see Fig. 6.19). The projection cylinder is developable:

Lemma 6.3.2. *Developing Δ maps $c(t)$ to a straight line. $c(t)$ is a geodesic of Δ and has horizontal principal normals.*

Proof. A development of Δ into the plane maps Δ 's rulings to parallel lines and the curve $c(t)$ to a curve $\bar{c}(t)$. As the angle between $c(t)$ and the vertical lines equals $\pi/2 - \gamma$, and this angle is invariant with respect to developments, \bar{c} must be a straight line. It is therefore a geodesic, and so is $c(t)$, because being a geodesic is a property of the intrinsic geometry of surfaces. This immediately shows that the principal normals of c are the surface normals of Δ . \square

The statement about the principal normals will follow also from the proof of Lemma 6.3.5.

Level Curves

Fig. 6.19 also shows level curves of the γ -developable \mathcal{R} . Level curves are the intersection curves with planes $z = z_0$ parallel to the reference plane. A top view of several level curves is shown in Fig. 6.19, right. The next lemma summarizes some properties of level curves. Recall that the *evolute* of a planar curve is the locus of its curvature centers and at the same time the envelope of its normals.

Lemma 6.3.3. *The level curves of a γ -developable \mathcal{R} orthogonally intersect its rulings. The orthogonal projections of the level curves into Π are offsets of each other. The curve of regression projects onto the common evolute of this family of curves.*

Proof. It is elementary to see that the intersection curves of \mathcal{R} with horizontal planes $z = z_0$ are orthogonal to the rulings $R(t)$, and that the distance between two such curves l_1, l_2 at heights z_0 and z_1 , measured along corresponding rulings, equals $|z_1 - z_0| / \sin \gamma$, which is a constant (γ is the angle enclosed by \mathcal{R} and Π).

We denote the orthogonal projection into the reference plane Π by a tilde. The right angle between $R(t)$ and the curves l_i is preserved by orthogonal projection and the distance between \tilde{l}_1, \tilde{l}_2 equals $|z_1 - z_0| \cot \gamma$.

If c is the curve of regression, then the lines $\tilde{R}(t)$ are tangent to \tilde{c} , which shows the statement about the evolute. \square

Theorem 6.3.4. *A developable surface of constant slope is locally a γ -cone or the tangent surface of a curve c of constant slope.*

Proof. A cylindrical surface cannot be of constant slope unless it is planar. Planar surfaces have been excluded from the discussion. A non-planar cone of constant slope is a right circular cone. If \mathcal{R} is a tangent surface, its rulings are the tangents of c . By Th. 5.1.7, all torsal ruled surfaces are locally of these type, so the proof is complete. \square

Remark 6.3.1. The curve c of regression consists of singular surface points, which shows that the level curves have singularities where they intersect c . If this situation is projected orthogonally into Π we get the familiar result that the offsets of a curve are singular at points of the evolute. \diamond

Osculating Cones

By Th. 6.1.4, all rulings $R(t)$ of a developable surface \mathcal{R} possess a cone of revolution $\Gamma(t)$ which is in second order contact with the surface \mathcal{R} in all regular points of the ruling. The vertex of this osculating cone is the point $c(t)$ of regression.

Lemma 6.3.5. *The axes of the osculating cones $\Gamma(t)$ of a developable surface \mathcal{R} of constant slope are orthogonal to the reference plane Π and the osculating cones are γ -cones (if \mathcal{R} is sufficiently smooth).*

Proof. By Th. 6.1.4, the axis of $\Gamma(t)$ is parallel to the Darboux vector of $c(t)$, which equals $\kappa e_3 + \tau e_1$, if e_1, e_2, e_3 are the vectors of c 's Frenet frame. We let $n = (0, 0, 1)$ and differentiate $n \cdot e_1 = \text{const}$: We get $n \cdot \kappa e_2 = 0$ and $n \cdot (-\kappa e_1 + \tau e_3) = 0$. This means that $n = \lambda e_2 \times (-\kappa e_1 + \tau e_3) = \lambda(\tau e_1 + \kappa e_3)$, and we have shown that the axis of $\Gamma(t)$ is parallel to the z -axis. Because $\Gamma(t)$ touches \mathcal{R} along $R(t)$, it is itself a γ -cone (see Fig. 6.19). \square

Remark 6.3.2. The osculating cone $\Gamma(t)$ intersects all horizontal planes in circles which have second order contact with the corresponding level curves. Thus, these are the osculating circles of the level curves. Their centers are contained in the vertical axis of $\Gamma(t)$ and thus the curvature center of a level curve l is the orthogonal projection $c'(t)$ of the corresponding point of regression $c(t)$ (see Fig. 6.19, right). This is in accordance with the well known fact that the evolute of a curve is both the envelope of the curve normals and the locus of its curvature centers. \diamond

Corollary 6.3.6. *The developable surfaces of constant slope are those whose conical curvature is constant (if they are sufficiently smooth).*

Proof. (Sketch) Recall (see Th. 6.1.4 and its proof) that the conical curvature k of a tangent surface \mathcal{R} equals τ/κ , where τ and κ are curvature and torsion of the curve of regression, and that k equals the conical curvature of the surface's osculating cone of revolution. The conical curvature of a cone equals $\cot \omega$, if ω is half the aperture angle. Now Lemma 6.3.5 shows that k is constant.

The converse follows from the fact that curves are uniquely determined, up to Euclidean motions, by the curvature and torsion functions $\kappa(s)$ and $\tau(s)$, where s is the arc length. If \mathcal{R} is the tangent surface of a curve c with $\tau(s)/\kappa(s) = k = \text{const.}$, then we can construct a planar curve \tilde{c} with curvature κ , and a curve c' of constant slope which projects orthogonally onto \tilde{c} . Being of constant slope, the curvature and torsion functions of c' equal $\kappa(s)$ and $\tau(s) = k\kappa(s)$. It must therefore equal the curve c , so c is of constant slope. \square

Representations

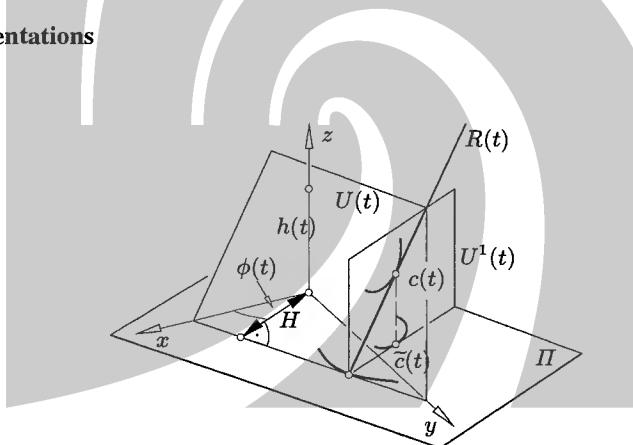


Fig. 6.20. Representation of developable surface of constant slope (see text).

We want to parametrize the set of all C^2 developables of constant slope γ with respect to some reference plane. Here C^2 refers to the dual representation. We assume that Π coincides with the xy -plane. Without loss of generality we may parametrize the family $U(t)$ of tangent planes of \mathcal{R} by

$$U(t) : x \sin \phi(t) - y \cos \phi(t) + z \cot \gamma = h(t) \cot \gamma, \quad (6.46)$$

(see Fig. 6.20). Here $\phi(t)$ is the angle enclosed by the line $U(t) \cap \Pi$ and the y -axis. The first derivative plane has the equation

$$U^1(t) : x \cos \phi + y \sin \phi = (\dot{h} \cot \gamma) / \dot{\phi}. \quad (6.47)$$

The plane $U^1(t)$ is vertical and contains the generator line $R(t) = U(t) \cap U^1(t)$ of \mathcal{R} . The points of the curve of regression $c(t)$ are contained in the second derivative plane

$$U^2(t) : -x \sin \phi + y \cos \phi = \cot \gamma \cdot (\ddot{h}\dot{\phi} - \ddot{h}\ddot{\phi})/\dot{\phi}^3. \quad (6.48)$$

The point of regression is found as $c(t) = U(t) \cap U^1(t) \cap U^2(t)$. By solving the three linear equations (6.46), (6.47), and (6.48) for x , y , and z , we find

$$\begin{aligned} x &= \cot \gamma (\cos \phi \cdot \dot{h}/\dot{\phi} - \sin \phi \cdot (\ddot{h}\dot{\phi} - \ddot{h}\ddot{\phi})/\dot{\phi}^3), \\ y &= \cot \gamma (\sin \phi \cdot \dot{h}/\dot{\phi} + \cos \phi \cdot (\ddot{h}\dot{\phi} - \ddot{h}\ddot{\phi})/\dot{\phi}^3), \\ z &= h + (\ddot{h}\dot{\phi} - \ddot{h}\ddot{\phi})/\dot{\phi}^3. \end{aligned} \quad (6.49)$$

This is an explicit representation of curves of constant slope based on two arbitrary C^2 functions $\phi(t)$, $h(t)$. All pairs (ϕ, h) define a γ -developable, if $U(t)$ is not constant or parametrizes a pencil.

Remark 6.3.3. A curve $c(t) = (x(t), y(t), z(t))$ has constant slope with respect to the xy -plane if and only if its tangent vectors enclose a constant angle with the vector $(0, 0, 1)$. This is equivalent to

$$\dot{z}^2(t) = \tan^2 \gamma (\dot{x}^2(t) + \dot{y}^2(t)). \quad (6.50)$$

If we prescribe $x(t)$ and $y(t)$, we can find $z(t)$ by integration:

$$z(t) = \tan \gamma \int \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt. \quad (6.51)$$

This means that the height function $z(t)$ of the curve c is proportional to the arc length of its top view $(x(t), y(t), 0)$. This follows also from the development of the projection cylinder through c (Lemma 6.3.2) which maps both c and its top view to straight lines. \diamond

Remark 6.3.4. Both Equ. (6.49) and (6.51) describe solutions of the differential equation (6.50). The former does this in an integral-free way. \diamond

If never $\dot{\phi} = 0$ in Equ. (6.46), we can use the angle ϕ as parameter. If $\dot{\phi} = 0$, the point of regression is at infinity or the curve $c(t)$ has an inflection point (or a higher order singularity). This can be seen by solving for the *homogeneous* coordinates of $c(t)$. Let us exclude such cases now and assume that $\phi = t$.

The family $U(\phi)$ of planes in (6.46) is generated by an *axial motion*, composed of a rotation about the angle ϕ , and whose axis is the z -axis, and a translation by the vector $(0, 0, h(\phi))$.

With $\dot{\phi} = 1$ and $\ddot{\phi} = 0$, Equ. (6.49) simplifies to

$$\begin{aligned} x &= \cot \gamma (\dot{h} \cos \phi - \ddot{h} \sin \phi), \\ y &= \cot \gamma (\dot{h} \sin \phi + \ddot{h} \cos \phi), \\ z &= h + \ddot{h}. \end{aligned} \quad (6.52)$$

The intersection curve l of the developable surface with the xy -plane Π is the envelope of the lines $L(\phi) = U(\phi) \cap \Pi$ (see Fig. 6.20):

$$L(\phi) : x \sin \phi - y \cos \phi = \cot(\gamma) h(\phi) = H(\phi). \quad (6.53)$$

The function $H(\phi)$ is the signed distance of the origin to the line L , and is called the *support function* of the curve l . The first derivative line $L^1 = U^1 \cap \Pi$ is the curve normal:

$$L^1(\phi) : x \cos \phi + y \sin \phi = \dot{H}(\phi). \quad (6.54)$$

Its distance to the origin equals $\dot{H}(\phi)$. Thus the curve $l(t) = L(t) \cap L^1(t)$ has the parametrization

$$x(\phi) = H \sin \phi + \dot{H} \cos \phi, \quad y(\phi) = -H \cos \phi + \dot{H} \sin \phi. \quad (6.55)$$

Lemma 6.3.7. *Arc length differential ds and curvature radius ρ of the curve l of Equ. (6.55) are given by*

$$\rho = H(\phi) + \ddot{H}(\phi), \quad ds = (H(\phi) + \ddot{H}(\phi)) d\phi. \quad (6.56)$$

Proof. The expression for ds follows by differentiation from Equ. (6.55). The expression for ρ is derived as follows: We consider the Frenet frame e_1, e_2 of the curve $l(t)$ (cf. Ex. 1.2.3). Then $(de_1/dt) = (d\phi/dt)e_2$. On the other hand, the Frenet equations show that $(de_1/dt) = (ds/dt)\kappa e_2$, where $s(t)$ is the arc length of the curve l . Thus $(d\phi/dt) = \kappa(ds/dt)$, i.e., $ds = \rho d\phi$. \square

Linearity of ρ and s in H is a computationally attractive feature of the support function.

Example 6.3.1. As a first example of the representation (6.46) we consider

$$h(\phi) = p\phi.$$

The corresponding γ -developable \mathcal{R} is generated by a *helical motion*. Its curve of regression (6.52) is a helix, which is also a point trajectory of the generating motion.

$$x(\phi) = p \cot \gamma \cos \phi, \quad y(\phi) = p \cot \gamma \sin \phi, \quad z(\phi) = p\phi.$$

The surface \mathcal{R} is the tangent surface of a helix (Fig. 6.21). Its intersection curves with horizontal planes are involute curves of the top view \tilde{c} of the helix c , i.e., involutes of a circle. \diamond

Example 6.3.2. A remarkable class of examples is obtained by letting

$$h(\phi) = a \cos n\phi, \quad (n \neq 0).$$

The generating motion of the corresponding family of planes is composed of a rotation about the z -axis and a harmonic oscillation with frequency n parallel to the z -axis [94]. With (6.52), we find the curve of regression:

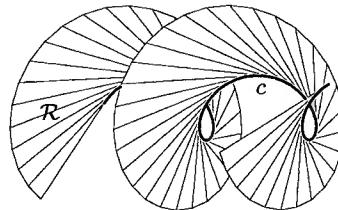


Fig. 6.21. Developable helical surface.

$$\begin{aligned} x &= an \cot \gamma (-\sin n\phi \cos \phi + n \cos n\phi \sin \phi), \\ y &= -an \cot \gamma (\sin n\phi \sin \phi + n \cos n\phi \cos \phi), \\ z &= a(1 - n^2) \cos n\phi. \end{aligned} \quad (6.57)$$

For $|n| = 1$, this parametrizes the constant curve $c(t) = (0, -a \cot \gamma, 0)$, and hence \mathcal{R} is a right circular cone. The generating motion has the property that the envelope of a plane is a cone (or a horizontal plane or a vertical cylinder). Such a motion is known as *Darboux-Mannheim motion* and is characterized by the property that all paths of points are planar (ellipses, in general) and all envelopes of planes are cones (degenerate cases included). \diamond

Example 6.3.3. (Continuation of Ex. 6.3.2) The case $|n| \neq 1$ is more interesting. Either from (6.57) or from the fact that its support function equals $H(\phi) = -a \cot \gamma \sin n\phi$ we compute the top view $\tilde{c}(t) = (x(t), y(t), 0)$ of the curve of regression. It is either an *epicycloid* for $|n| < 1$ or a *hypocycloid* for $|n| > 1$. (These curves are trajectories of the motion of a circle rolling on another circle). The intersection curves l of the surface \mathcal{R} with the xy -plane have the support function $H = a \cot \gamma \cos n\phi$, which shows that this curve also is an epicycloid or hypocycloid. Note that \tilde{c} is the evolute of l and it is also the image of l with respect to a similarity transformation (see Figs. 6.22 and 6.41). Cusps of \tilde{c} belong to points of extremal curvature of l and vice versa.

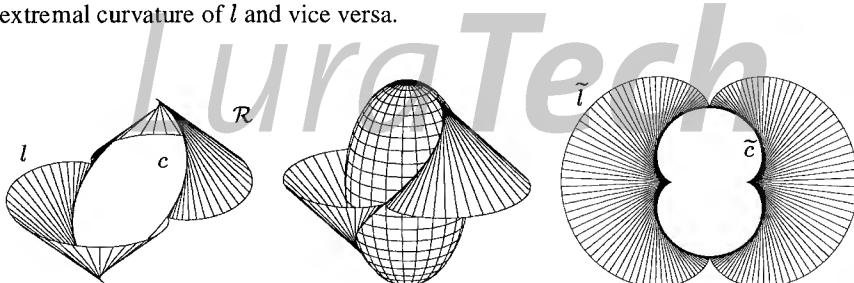


Fig. 6.22. Left: developable surface generated by $h(\phi) = \cos(\phi/2)$ (see Ex. 6.3.3). Center: ellipsoid which contains its curve of regression. Right: top view.

The curve of regression c is contained in the quadric

$$\frac{1}{a^2 \cot^2 \gamma n^2} (x^2 + y^2) + \frac{1}{a^2(1 - n^2)} z^2 = 1.$$

This quadric has rotational symmetry with respect to the z axis, and is an ellipsoid for $|n| < 1$ and a one-sheeted hyperboloid for $|n| > 1$ (see Fig. 6.22 and Fig. 6.41 for $n = 2$). Curves of constant slope contained in rotational quadrics with vertical axes appear in problems of classical geometry (see e.g., [210, 213]). \diamond

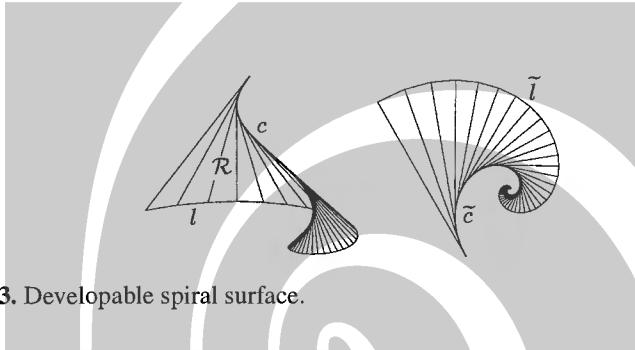


Fig. 6.23. Developable spiral surface.

Example 6.3.4. Assume a γ -developable \mathcal{R} , whose family of tangent planes is parametrized by Equ. (6.46) with

$$h(\phi) = ae^{p\phi}, \quad (p \neq 0).$$

We easily see that \mathcal{R} is the tangent surface of a *spatial logarithmic spiral* c (Fig. 6.23). Both the orthogonal projection \tilde{c} into the xy -plane and the intersection curve l with this plane are planar logarithmic spirals. The curve of regression c is contained in a cone of revolution whose vertex is $(0, 0, 0)$ and whose axis is the z -axis. The curve c encloses a constant angle with the rulings of this cone.

The surface \mathcal{R} may also be generated by a so-called group of *spiral motions*, which are actually no motions, but central similarities of Euclidean three-space (see [216] and Remark 3.1.4). Surfaces (usually non-developable) generated by such motions occur in nature as shapes of shells. \diamond

6.3.2 The Cyclographic Mapping and its Applications

At this point it is appropriate to introduce briefly the concept of *cyclographic mapping*. It is very helpful in understanding circle geometry, in particular so-called *Laguerre geometry*. For a detailed exposition of this interesting classical topic we refer to the monographs by Coolidge [31] and Müller and Krames [127]. Several problems in geometric computing possess elegant solutions based on the cyclographic mapping [100, 139, 154, 201].

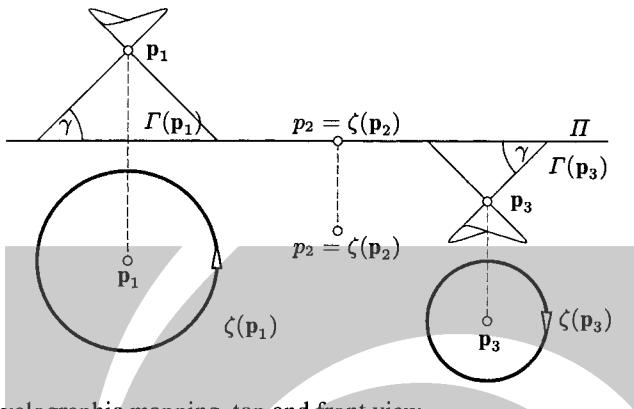


Fig. 6.24. Cyclographic mapping, top and front view.

Definition of the Cyclographic Mapping

Definition. Consider a triple $\mathbf{x} = (m_1, m_2, r)$ of real numbers, and the circle in E^2 which has the equation $(x_1 - m_1)^2 + (x_2 - m_2)^2 = r^2$. If $r > 0$, we give this circle a counterclockwise orientation, and a clockwise one for $r < 0$. It is then called a cycle and is designated by the symbol $\zeta(\mathbf{x})$. The mapping ζ is called cyclographic mapping.

Thus the domain of ζ is \mathbb{R}^3 , and the image of ζ is the set of cycles (Fig. 6.24). If $\mathbf{x} = (x_1, x_2, 0)$, then the cycle $\zeta(\mathbf{x})$ is nothing but the point (x_1, x_2) .

The cycle $\overrightarrow{\zeta} = \zeta(\mathbf{x})$ can also be obtained as follows: Consider a cone of revolution $\Gamma(\mathbf{x})$ with vertex \mathbf{x} , whose axis is parallel to the z -axis and whose generators enclose the angle $\gamma = \pi/4$ with the z -axis. Recall that we call such a cone a γ -cone. Then the cycle $\overrightarrow{\zeta}$ is the intersection of the plane Π with $\Gamma(\mathbf{x})$, given the appropriate orientation. The rulings of γ -cones, i.e., the lines with inclination angle γ , will be called γ -lines.

Remark 6.3.5. The set \mathcal{C} of γ -lines is called a *line complex of constant slope*, because it is a three-parameter family of lines. All γ -cones share a common ideal conic c_ω , which consists of the points $(0, \sin \phi, \cos \phi, 1)\mathbb{R}$. The complex \mathcal{C} obviously is the union of all line bundles whose vertex is in c_ω . Its complex cones are precisely the γ -cones (cf. p. 161).

Note that a developable surface is of constant slope γ if and only if the ideal points of its rulings are contained in the conic c_ω . \diamond

Remark 6.3.6. The cyclographic mapping ζ is a *projection* in the following sense: We choose a set of projection rays (here the complex of constant slope of Remark 6.3.5). In order to apply the projection to a point \mathbf{x} , we consider all projection rays incident with \mathbf{x} (here a γ -cone whose vertex is \mathbf{x}), and finally intersect these projection rays with a plane (here the xy -plane).

Such a projection can be defined for all sets of lines, especially for line complexes and line congruences with the property that all points of space are incident with at least one line from the set. Projection via line congruences will be discussed in Sec. 7.1.7. \diamond

Oriented Lines and Oriented Contact

A line can be given two different orientations. It is obvious how to define a *tangent* of a cycle as an oriented line which is in oriented contact with this cycle. Given an oriented line \vec{L} , we may ask for all cycles \vec{C} which are in oriented contact with \vec{L} . If we think of the Euclidean plane as the xy -plane in \mathbb{R}^3 , the ζ -preimages of these cycles are the points of a plane incident with the line \vec{L} , and whose inclination angle equals $\pi/4$ (see Fig. 6.25). Such a plane shall be called a γ -plane. This implies the following

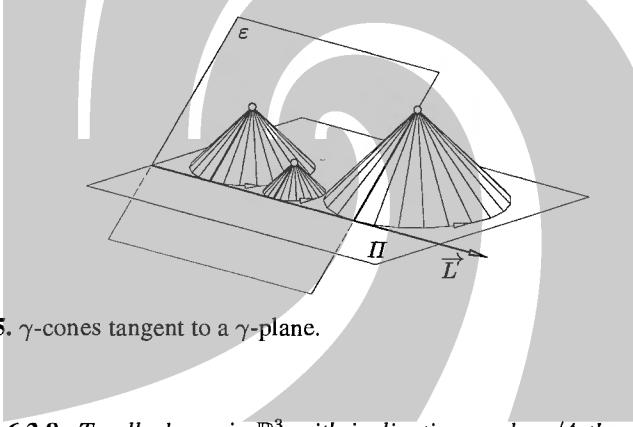


Fig. 6.25. γ -cones tangent to a γ -plane.

Lemma 6.3.8. *To all planes in \mathbb{R}^3 with inclination angle $\pi/4$ there corresponds an oriented trace line \vec{L} , and x is contained in this plane if and only if $\zeta(x)$ is in oriented contact with \vec{L} .*

Laguerre Geometry in the Cyclographic Model

The geometry of cycles and oriented lines of the Euclidean plane, together with the relation of oriented contact, is called planar *Laguerre geometry*. The interpretation of cycles and oriented lines of the Euclidean plane as points and γ -planes in \mathbb{R}^3 is called the *cyclographic model* of planar Laguerre geometry.

Consider the cycles which are in oriented contact with a given oriented line \vec{L} at a given point t : their ζ -preimages are the points of a γ -line.

Consider a cycle $\vec{C}_1 = \zeta(p_1)$. Then the points x of the γ -cone $\Gamma(p_1)$ are exactly those such that $\zeta(x)$ is in oriented contact with $\zeta(p_1) = \vec{C}_1$.

Example 6.3.5. We apply the concept of oriented contact of cycles and the cyclographic model to an ancient problem of circle geometry: We want to find circles

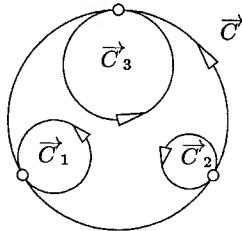


Fig. 6.26. Apollonios' problem for cycles.

which touch three given circles in the Euclidean plane. The problem is named after Apollonios of Perga (262–190 B.C.). We consider the problem for cycles instead of circles, because a solution for cycles of course gives a solution for circles, if we forget the orientation (see Fig. 6.26).

Assume three cycles $\vec{C}_i = \zeta(\mathbf{p}_i)$. Then a cycle $\vec{C} = \zeta(\mathbf{x})$ which is in oriented contact with \vec{C}_1 , \vec{C}_2 , and \vec{C}_3 , is found by intersecting the γ -cones with vertices \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3 . This leads to a quadratic equation. Since there are two different orientations for each circle, we get up to 16 solution cycles, which correspond to up to eight solution circles of the original Apollonios problem.

This spatial approach to the Apollonian problem (B. E. Cousins [32]) has initiated the use of the cyclographic mapping. \diamond

Remark 6.3.7. We should mention why three γ -cones (which are quadratic varieties), intersect in just two points — carelessly multiplying degrees would lead us to expect eight intersection points. The reason is that two γ -cones share the conic c_ω at infinity (cf. Remark 6.3.5), and so, by appropriate linear combinations of their equations, we can eliminate the quadratic terms: If the three γ -cones are given by

$$f_i(x_1, x_2, x_3) = (x_1 - p_{i,1})^2 + (x_2 - p_{i,2})^2 - (x_3 - p_{i,3})^2 = 0 \quad (i = 1, 2, 3),$$

then an equivalent system of equations is $f_1 = 0$, $f_1 - f_2 = 0$, $f_1 - f_3 = 0$, which consists of one quadratic and two linear equations. \diamond

Remark 6.3.8. For some of our applications, we will have to consider circles instead of cycles (i.e., forget about the orientations), as we have already seen in Ex. 6.3.5. In this case we can either admit only nonnegative radii and confine ourselves to the positive half-space $z \geq 0$ of the cyclographic model, or assign all possible orientations to the input circles orientations and look what happens. \diamond

Remark 6.3.9. Transformations in the set of cycles which act bijectively and preserve the relation of oriented contact are called *Laguerre transformations*. In general, they will not map cycles of zero radius (i.e., points) to cycles of zero radius.

In the cyclographic model, Laguerre transformations appear as transformations which transform γ -lines to γ -lines. It is not hard to show that such mappings are affine transformations, whose projective extension leave the conic c_ω invariant (see Remark 6.3.5).

Such affine maps are similarities in a *pseudo-Euclidean geometry* (also called *Minkowski geometry*, but beware of the different meaning of ‘Minkowski geometry’ in Sec. 6.4.4). Its metric is based on the following scalar product

$$\langle \mathbf{a}, \mathbf{b} \rangle_{\text{pe}} = a_1 b_1 + a_2 b_2 - a_3 b_3.$$

Four-dimensional pseudo-Euclidean geometry is a geometric space-time model for *special relativity* (see e.g., [12]). \diamond

Cyclographic Image of a Curve

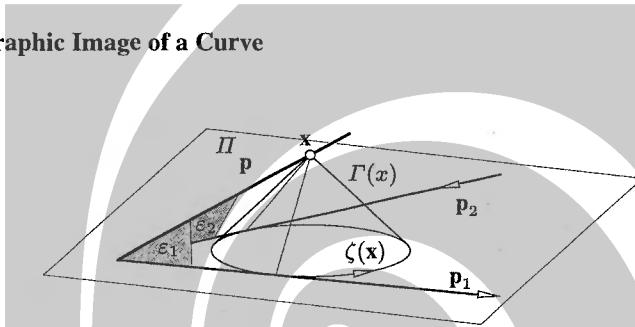


Fig. 6.27. Cyclographic image of a straight line.

Let $\mathbf{p} : I \rightarrow \mathbb{R}^3$ be a curve in \mathbb{R}^3 . If the one-parameter family $\zeta(\mathbf{p}(t))$ of cycles has an envelope, this envelope is called the *cyclographic image* of the curve $\mathbf{p}(t)$. This curve is given the orientation of the cycles whose envelope it is.

Let us start with the simplest example, that of a *straight line* in three-space, parametrized linearly by $\mathbf{p}(t)$. The angle enclosed by the line and the xy -plane will be denoted by α .

If $\alpha < \pi/4$, all cycles $\zeta(\mathbf{p}(t))$ possess two common oriented tangents \bar{p}_1, \bar{p}_2 . Such a line is called *space-like*, and its cyclographic image consists of two oriented lines (see Fig. 6.27).

If $\alpha = \pi/4$, the cycles $\zeta(\mathbf{p}(t))$ are in oriented contact in the same point. Such a family of circles does not have an envelope. This kind of lines is in this context also called *light-like*.

If $\alpha > \pi/4$ (the line is *time-like*) there are no common tangents and there is no envelope. The inclination angle of any plane incident with a time-like line is greater than $\pi/4$.

The names space-like, light-like, and time-like, come from the interpretation of the cyclographic model as space-time of special relativity, where the first two coordinates are space coordinates and the last coordinate is time. We have shown the following

Lemma 6.3.9. *The cyclographic image of a space-like line consists of two oriented lines, whereas the other types of lines have no cyclographic image curves.*

Let us now discuss the case of a γ -curve $\mathbf{p}(t)$, i.e., a curve of constant slope γ :

Lemma 6.3.10. *If $\mathbf{p}(t)$ is a γ -curve, its cyclographic image $c(\mathbf{p})$ consists of a curve in the Euclidean plane such that $\zeta(\mathbf{p}(t))$ are its osculating cycles.*

Proof. The tangent surface \mathcal{R} of $\mathbf{p}(t)$ is a γ -developable of constant slope $\gamma = \pi/4$. This tangent surface is the envelope of the family of γ -cones incident with the points $\mathbf{p}(t)$. Therefore the intersection of \mathcal{R} with the xy -plane, if given the appropriate orientation, is the envelope $c(\mathbf{p})$ of cycles $\zeta(\mathbf{p}(t))$. Lemma 6.3.5 shows that the above mentioned γ -cones are the osculating cones of \mathcal{R} , so the cycles $\zeta(\mathbf{p}(t))$ actually osculate the cyclographic image curve. \square

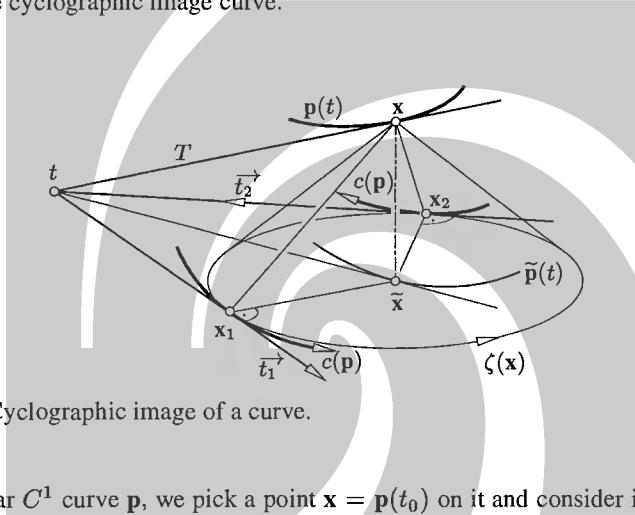


Fig. 6.28. Cyclographic image of a curve.

For a regular C^1 curve \mathbf{p} , we pick a point $\mathbf{x} = \mathbf{p}(t_0)$ on it and consider its tangent T . Since the computation of the envelope only requires first order derivatives and T is a first order approximant to $\mathbf{p}(t)$ at $t = t_0$, we can find the points where the cycle $\zeta(\mathbf{p})$ touches the cyclographic image $c(\mathbf{p})$ in the following way (see Fig. 6.28):

Compute the intersection point $\mathbf{t} = T \cap \Pi$ and find the oriented lines \vec{t}_1, \vec{t}_2 incident with \mathbf{t} and tangent to $\zeta(\mathbf{x})$. The common points of $\zeta(\mathbf{x})$ and \vec{t}_i are the desired points of the cyclographic image, and \vec{t}_i are its tangents.

The lines $\mathbf{x} \vee \mathbf{x}_1$ and $\mathbf{x} \vee \mathbf{x}_2$ are rulings $R_1(t_0), R_2(t_0)$ of γ -developables $\mathcal{R}_1, \mathcal{R}_2$ which contain the curve $\mathbf{p}(t)$.

Thus, the cyclographic mapping of general curves may be described as follows.

Lemma 6.3.11. *The cyclographic image $c(\mathbf{p})$ of a smooth curve $\mathbf{p}(t) \subset \mathbb{R}^3$ consists of the intersection curves of the γ -developables passing through $\mathbf{p}(t)$ with the xy -plane. There exist two such surfaces in an interval where the curve tangents are space-like, one in an interval where they are light-like, and none where they are time-like.*

For a thorough discussion of the cyclographic images of curves, both from the viewpoint of differential and algebraic geometry, we refer the reader to Müller and Krames [127]. We illustrate the type of geometric argumentation by means of two examples.

Cycographic Image of Conics

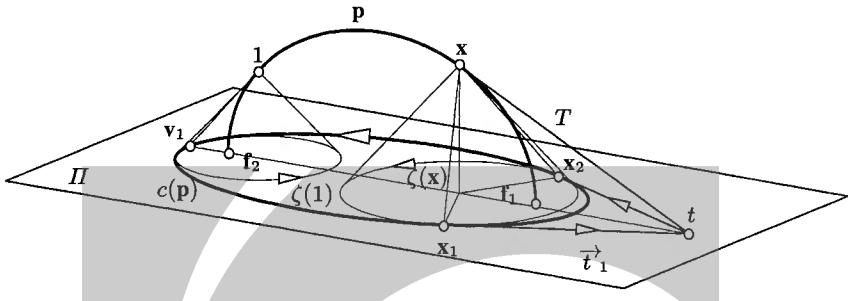


Fig. 6.29. Cycographic image of a conic.

Example 6.3.6. Consider the conic \mathbb{R}^3 which has the equation

$$x = 0, \quad ay^2 + bz^2 = 1, \quad a, b > 0 \text{ or } ab < 0.$$

It is an ellipse or hyperbola and may be parametrized in the form $\mathbf{p}(t)$, but it is easy to derive properties of its cycographic image without any computation. The curve is symmetric with respect to the reference plane Π .

The γ -developable \mathcal{R} which contains \mathbf{p} is the envelope of γ -planes which touch the given conic and, being γ -planes, touch the ideal conic c_ω (cf. Remark 6.3.5). We dualize the situation in projective three-space: The set of tangent planes of a conic is converted to the set of points of a quadratic cone. The common tangent planes of two conics are converted into the intersection curve of two quadratic cones. Thus the set of tangent planes of \mathcal{R} is dual to an algebraic curve of order four, and \mathcal{R} is an algebraic developable of class four.

The class of the intersection $c(\mathbf{p}) = \Pi \cap \mathcal{R}$ can be estimated by counting tangents of $c(\mathbf{p})$ which are incident with a test point. This test point is, generically and over the complex number field, incident with four tangent planes of \mathcal{R} . This quadruple of planes is symmetric with respect to Π , so generically two tangents of $c(\mathbf{p})$ are incident with the test point.

This shows that $c(\mathbf{p})$ is a conic, if it exists (i.e., if the intersection $\mathcal{R}_C \cap \Pi_C$ has real points). The points $\mathbf{p} \cap \Pi$ turn out to be focal points $\mathbf{f}_1, \mathbf{f}_2$ of $c(\mathbf{p})$, and the γ -tangents of \mathbf{p} intersect Π in vertices $\mathbf{v}_1, \mathbf{v}_2$ of $c(\mathbf{p})$ (see Fig. 6.29). \diamond

Example 6.3.7. We use the notation of Ex. 6.3.6, but here we consider curves which need not be symmetric with respect to the reference plane. An important special case occurs if \mathbf{p} is a so-called *pseudo-Euclidean circle*, i.e., a conic whose points at infinity are contained in c_ω . An alternative definition is that pseudo-Euclidean circles are planar intersections of γ -cones. For any pseudo-Euclidean circle \mathbf{p} , the developable surface \mathcal{R} splits into two γ -cones (one of which degenerates to a γ -plane if \mathbf{p} is a parabola). Thus the cycographic image $c(\mathbf{p})$ consists of two cycles,

or it consists of a cycle and an oriented line for a parabolic pseudo-Euclidean circle \mathbf{p} . In the symmetric case $a = -b$, the curve $c(\mathbf{p})$ consists of two points. \diamond

Offset Curves

There is an intimate relation between offsets of a planar curve c contained in the reference plane Π and the γ -developable \mathcal{R} which contains c (cf. Ex. 1.2.8 and Fig. 1.41):

Lemma 6.3.12. *The graph of the signed distance function to a curve c in the xy -plane is a developable surface \mathcal{R} of constant slope.*

Proof. This is obvious from the definition of \mathcal{R} as the γ -developable which contains c . The surface \mathcal{R} is the envelope of all γ -cones defined by cycles which touch c . \square

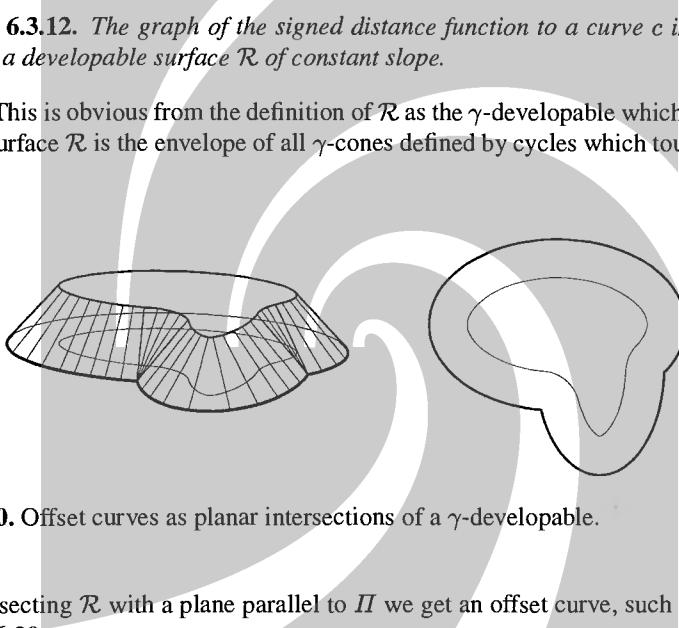


Fig. 6.30. Offset curves as planar intersections of a γ -developable.

By intersecting \mathcal{R} with a plane parallel to Π we get an offset curve, such as shown by Fig. 6.30.

Example 6.3.8. (Continuation of Ex. 6.3.6) We want to show that the offsets of ellipses and hyperbolae are non-rational algebraic curves of class four, except for circles whose offsets are again circles.

We embed the given conic as the curve $c(\mathbf{p})$ in the reference plane Π and consider the γ -developable \mathcal{R} which contains $c(\mathbf{p})$ (see Fig. 6.29). It is elementary to see that all non-circular ellipses and hyperbolae actually occur as cyclographic projections of conics \mathbf{p} .

It is possible to show that \mathcal{R} is not rational: The set of tangent planes of \mathcal{R} consists of those planes which are tangent to both conics \mathbf{p} and $c(\mathbf{p})$. By dualizing this we get a set of points which is the intersection of two quadratic cones. It is well known that such a curve is rational if and only if it has a double point and the cones have a common tangent plane. By dualizing back we see that \mathcal{R} is rational if and only if the two conics \mathbf{p} and $c(\mathbf{p})$ have a common point, which is not the case. \diamond

Example 6.3.9. (Continuation of Ex. 6.3.8) Consider a parabola \mathbf{p} with equation

$$x = 0, \quad y - z^2 = 0.$$

It is easy to see that its cyclographic image $c(\mathbf{p})$ is again a parabola (cf. the argumentation of Ex. 6.3.6). The γ -developable \mathcal{R} which contains \mathbf{p} is again of class four, but this time it is rational (the two conics \mathbf{p} and $c(\mathbf{p})$ indeed possess a common point at infinity).

The offset curves of a parabola are therefore rational curves whose class equals the class of \mathcal{R} , namely four. By Equ. (1.103) the offsets are algebraic curves of order six. \diamond

We sum up Ex. 6.3.8 and Ex. 6.3.9:

Proposition 6.3.13. *The offset curves of a conic are circles in the case of a circle, rational of order six and class four in the case of a parabola, and not rational in all other cases.*

The rationality of the offsets of a parabola has been considered as new within the CAGD community (see [55, 113]), but it follows immediately from the discussion of the cyclographic image curves of arbitrary conics (called *hypercycles*) in a paper by W. Blaschke [13]. An argument completely similar to Ex. 6.3.9 shows that the cyclographic image curves $c(\mathbf{p})$ of arbitrary parabolae \mathbf{p} are rational curves with rational offsets and have (in general) class four. Later we will return to this example.

Disk Bézier Curves

Consider a Bézier curve \mathbf{p} in E^3 whose control points \mathbf{b}_i are contained in the upper half-space H^+ which is defined by the equation $z > 0$. The cyclographic image points $\zeta(\mathbf{b}_i)$ are cycles with positive orientation. They determine disks D_i . We can see these disks as *tolerance regions* for imprecisely determined control points in the plane and ask the following question: if the control points vary in their respective tolerance regions D_i , which part of the plane is covered by the corresponding Bézier curves? We call this planar region the *tolerance region of the Bézier curve* (see Fig. 6.31). We first study the behaviour of convex combinations under the cyclographic mapping:

Definition. If A, B are subsets of \mathbb{R}^n , and $\lambda \in \mathbb{R}$, then $A + B = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in A, \mathbf{y} \in B\}$, $A - B = \{\mathbf{x} - \mathbf{y} \mid \mathbf{x} \in A, \mathbf{y} \in B\}$, and $\lambda A = \{\lambda \mathbf{x} \mid \mathbf{x} \in A\}$. The operation ‘+’ is called *Minkowski sum*.

Note that $A + A$ does not equal $2A$, and $A - A$ does not equal ‘zero’ in any sense.

Lemma 6.3.14. The cyclographic mapping commutes with convex combinations in the sense that if $\mathbf{b}' = t\mathbf{b}_1 + (1-t)\mathbf{b}_2$ ($0 \leq t \leq 1$), and \mathbf{b}_i are contained in the upper half space, then the disks D, D_1, D_2 bounded by $\zeta(\mathbf{b}')$, $\zeta(\mathbf{b}_1)$, $\zeta(\mathbf{b}_2)$ fulfill the equation $D' = tD_1 + (1-t)D_2$ as subsets of the affine plane.

Proof. We assume that $\mathbf{b}_i = (m_i, n_i, r_i)$ and use coordinates x, y, z . Then $\mathbf{b}' = (m', n', r') = (tm_1 + (1-t)m_2, \dots)$. We have to show that the points \mathbf{x}' of the disk $(x - m')^2 + (y - n')^2 \leq r'^2$ can be written in the form $\mathbf{x}' = t\mathbf{x}_1 + (1-t)\mathbf{x}_2$ with \mathbf{x}_i in the disk $(x - m_i)^2 + (y - n_i)^2 = r_i^2$. This is an elementary exercise. \square

Now it is possible to describe the tolerance region of a Bézier curve:

Proposition 6.3.15. *If the tolerance regions of the points $\tilde{\mathbf{b}}_0, \dots, \tilde{\mathbf{b}}_m$ are disks D_i , bounded by cycles $\zeta(\mathbf{b}_0), \dots, \zeta(\mathbf{b}_m)$, then the tolerance region D of the Bézier curve $\tilde{B}(t)$ defined by the control points $\tilde{\mathbf{b}}_0, \dots, \tilde{\mathbf{b}}_n$ is the cyclographic image of the Bézier curve $B(t)$ defined by control points $\mathbf{b}_0, \dots, \mathbf{b}_n$ ($0 \leq t \leq 1$).*

Proof. We consider a fixed value t_0 with $0 \leq t \leq 1$. The algorithm of de Casteljau (Lemma 1.4.2) evaluates the Bézier curve $\tilde{B}(t)$ at $t = t_0$ by an appropriate iterated convex combination of the points $\tilde{\mathbf{b}}_i$. At the same time this algorithm evaluates $B(t)$ at $t = t_0$ by the same convex combinations applied to the points \mathbf{b}_i .

Thus we can use Lemma 6.3.14 to show that the tolerance region of the curve point $\tilde{B}(t_0)$ is bounded by $\zeta(B(t_0))$. The tolerance region of the entire curve then is the union of all such disks, whose boundary is contained in the cyclographic image $c(\mathbf{b}(t))$, by definition of the cyclographic image curve. \square

An example of this can be seen in Fig. 6.31. Such *disk Bézier curves* have been studied by Lin and Rokne [111].

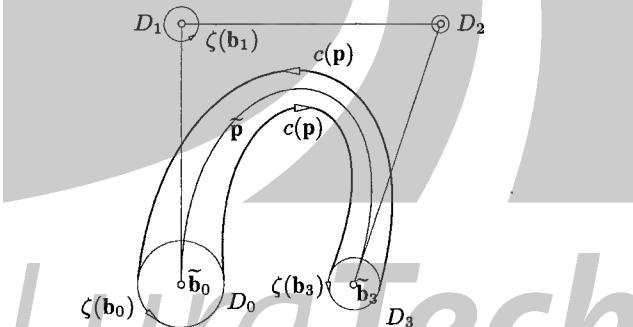


Fig. 6.31. Tolerance region of a Bézier curve with disks D_i as tolerance regions for the control points.

Remark 6.3.10. We point out that trimming of the cyclographic image curve $c(\mathbf{b})$ may be necessary in order to remove parts which are inside the tolerance region D . This occurs in case of self-intersections of $c(\mathbf{b})$ or intersections of different components of $c(\mathbf{b})$. There are 'local' types of self-intersections which typically arise in the neighbourhood of two cusps of the envelope (swallow-tail) and cannot occur if $c(\mathbf{b})$ has no singularities. Other 'global' types of self-intersections are harder to detect. Trimming may also become necessary if different components of the envelope intersect, which is caused by the parts of $\mathbf{b}(t)$ with time-like tangents. \diamond

Bisector Curves

The union of oriented planar curves c_1, c_2 can be the cyclographic image of a curve p in space: For each of the curves c_i there is a γ -developable which contains this curve, and which consists of two sheets. We choose the one which gives the right orientation to c_i . By intersecting these two surfaces we get a curve p , and obviously the cyclographic image of p is contained in the union of c_1 and c_2 .

The orthogonal projection \tilde{p} of p onto the base plane is usually called *bisector* of c_1 and c_2 (see [73, 50]). The points of \tilde{p} are centers of cycles tangent to c_1 and c_2 , and therefore possess equal distance to both curves (see Figs. 6.28 and 6.31).

The Medial Axis Transform

Assume that D is a subset of Π , whose boundary consists of a finite number of C^2 curves c_i . We assume that the boundary is oriented such that the interior of D lies to the left. For each c_i , there are two γ -developables which contain c_i . From the two possibilities, which are symmetric with respect to Π , and we choose the one which gives the right orientation to c_i . In a point x of tangent discontinuity we consider the part of the γ -cone which fills the gap between the adjacent γ -developables (see Fig. 6.32).

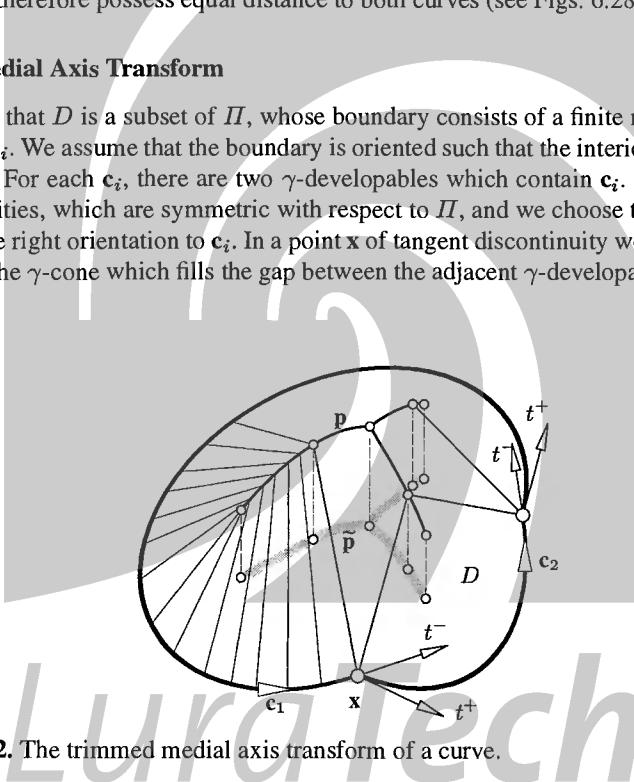


Fig. 6.32. The trimmed medial axis transform of a curve.

The set p of all intersections and self-intersections of these surfaces, together with their singular points, is called the (spatial) *medial axis transform* of D . It consists of several curves in space. The orthogonal projection onto Π maps p to the (untrimmed) *medial axis* \tilde{p} .

What usually is called the medial axis transform of a planar domain is only part of our curve \tilde{p} (cf. [27, 28]): There the medial axis is the locus of centers of disks contained in D whose boundary touches D 's boundary in two points, together with the limit points of this set.

It is not hard to show the relation between this (trimmed) medial axis and the curves \mathbf{p} and $\tilde{\mathbf{p}}$: Consider the curve \mathbf{p} and together with the γ -developables needed for its definition. The points of \mathbf{p} which are visible from below project precisely on the trimmed medial axis. A point (x, y, z) is visible from below if no point (x, y, z') with $z' < z$ is contained in one of the γ -developables mentioned before. The boundary of D can be recovered from this visible part as well:

Lemma 6.3.16. *Consider the part of the (untrimmed spatial) medial axis transform of a planar domain D which is visible from below. Then D 's boundary is the cyclo-graphic image of this part.*

This construction of the medial axis via the cyclographic model of planar circle geometry shows how to compute the medial axis with surface/surface intersection algorithms. The visible part of \mathbf{p} is easily computed by a *visibility algorithm*.

Remark 6.3.11. If one is just interested in the trimmed medial axis, only the ‘concave’ points of tangent discontinuity have to be considered, because only there the γ -cone mentioned above is contained in the upper half space. Such a point is concave if the oriented angle enclosed by the oriented tangents t^-, t^+ is negative. ◇

The word ‘medial axis transform’ suggests that this operation transforms the set of admissible planar domains to a set of curves, and that this operation is one-to-one. This is indeed the case, because Lemma 6.3.16 shows how to reconstruct D from the medial axis transform.

The medial axis transform is therefore sometimes used for shape representation. It is also very useful for the construction of trimmed interior offsets, which are needed for NC pocket machining (cf. [69]) or layered manufacturing. To construct the interior offset at distance d of the boundary of D we do the following: We consider the part of the developable surfaces used for the construction of the medial axis transform which is contained between the plane Π and the curve \mathbf{p} (Fig. 6.32 shows exactly this part). If we intersect this surface with the plane $z = d$, and project the intersection curve orthogonally into Π , we get the trimmed interior offset at distance d (see (Fig. 6.33). For more details on this approach and the mathematical theory of the medial axis transform we refer to Choi et al. [27, 28].

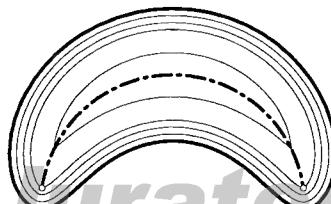


Fig. 6.33. Trimmed offsets via medial axis transform.

Remark 6.3.12. For CAD/CAM purposes it is desirable to know two classes of spline-type curves with the property that always when the medial axis transform \mathbf{p} of a planar domain D is contained in one class, the boundary of D is contained in the other, and vice versa. Even if a given planar domain is not of this type, it can be replaced by a spline approximant.

There are the following possibilities: If \mathbf{p} consists of *pseudo-Euclidean circular arcs* (see Ex. 6.3.7), then D 's boundary and all of its offsets consist of ordinary circular arcs. The two classes of spline curves are circular splines, both Euclidean planar, and pseudo-Euclidean spatial ones.

A class of spline curves with more degrees of freedom is the class of *polynomial quadratic* spline curves. If \mathbf{p} is such a curve, trimming the offsets is efficient and exact; according to Ex. 6.3.9, the boundary of D is, in general, piecewise rational of class four and degree six, and so are the offsets of D 's boundary.

More generally, the class of so-called pseudo-Euclidean Pythagorean-hodograph curves is the most general class whose cyclographic images are rational (see H. I. Choi et al. [29] and H. P. Moon [125]). \diamond

An Alternative Construction of the Medial Axis

Another spatial interpretation of the construction of the medial axis and of bisector curves has been proposed by Choi, Elber and Kim [30]. Consider a planar domain D with C^2 boundary consisting of curves \mathbf{l}_i , parametrized in the form $\tilde{\mathbf{l}}_i(t) = (x_i(t), y_i(t))$. Then consider the curves

$$\mathbf{l}_i(t) = (x_i(t), y_i(t), x_i^2(t) + y_i^2(t)),$$

which obviously are contained in the paraboloid $\Sigma : z = x^2 + y^2$, and whose orthogonal projection into the xy -plane gives the curves $\tilde{\mathbf{l}}_i$ again. Finally construct the developable surfaces Γ_i which touch Σ in the points of the curves \mathbf{l}_i . It is not hard to see that the intersections and self-intersections of these developable surfaces yield (after an appropriate trimming procedure via a visibility algorithm) the medial axis of D when projected orthogonally into the xy -plane.

An important property of this method is that it employs *rational* developable surfaces if the curves \mathbf{l}_i are rational.

Remark 6.3.13. It is (in principle) an elementary exercise to show that this algorithm actually gives the medial axis. The main ingredient of the proof is that the tangents of the paraboloid Σ which are incident with a point \mathbf{q} touch Σ in the points of an ellipse, whose orthogonal projection into the xy -plane is a circle, and the center of this circle coincides with the projection of \mathbf{q} (see Fig. 6.34).

It is interesting to note that the orthogonal projection onto the xy -plane is a stereographic projection (as defined in Sec. 4.3) when restricted to Σ . \diamond

Example 6.3.10. We want to compute the bisector curve of two circles. This bisector does not change if we replace the curves by their offsets at distance d , we may as well construct the bisector \mathbf{b} of a circle $\tilde{\mathbf{l}}_1$ and a point $\tilde{\mathbf{l}}_2$. This is shown in Fig. 6.34.

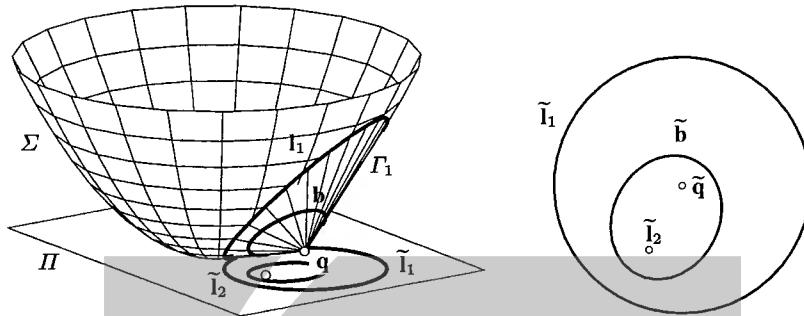


Fig. 6.34. Constructing bisectors by intersecting developable surfaces tangent to a paraboloid of revolution.

We look for curves \mathbf{l}_1 and \mathbf{l}_2 in Σ which project onto $\tilde{\mathbf{l}}_1, \tilde{\mathbf{l}}_2$. It is an elementary exercise to show that \mathbf{l}_1 is an ellipse. The developable surfaces Γ_1, Γ_2 tangent to Σ along \mathbf{l}_1 and \mathbf{l}_2 are a quadratic cone and a plane. Their intersection \mathbf{b} , which projects onto the bisector curve $\tilde{\mathbf{b}}$, is therefore a conic (cf. [49]). \diamond

Remark 6.3.14. A closer inspection shows the well known classical result that for fixed $\tilde{\mathbf{l}}_2$, there is a quadratic birational line-point transformation which transforms the set of tangents of $\tilde{\mathbf{l}}_1$ to the set of points of $\tilde{\mathbf{b}}$. This does not only explain the rationality of the bisector, but is also a useful tool for investigating the properties of $\tilde{\mathbf{b}}$ as of an algebraic curve. \diamond

Geometrical Optics

Certain problems in geometrical optics are related to the cyclographic mapping of curves: Consider a curve $\mathbf{p}(t)$ and a point $\mathbf{x} = \mathbf{p}(t_0)$ with a space-like tangent T . We denote the orthogonal projection of \mathbf{p} into the base plane by $\tilde{\mathbf{p}}$. Consider the tangent \tilde{T} to the curve $\tilde{\mathbf{p}}$ at $\tilde{\mathbf{x}} = \tilde{\mathbf{p}}(t_0)$. The construction of \mathbf{p} 's cyclographic image $c(\mathbf{p})$ (see Fig. 6.28) shows that the lines \tilde{T} and $\tilde{\mathbf{x}} \vee \mathbf{x}_1$ enclose the same angle as the lines \tilde{T} and $\tilde{\mathbf{x}} \vee \mathbf{x}_2$, and that the line $\tilde{\mathbf{x}} \vee \mathbf{x}_i$ is orthogonal to the curve \mathbf{c}_i .

If the curve normals of one branch \mathbf{c}_1 of the cyclographic image $c(\mathbf{p})$ of the curve $\mathbf{p}(t)$ are *light rays* and $\tilde{\mathbf{p}}$ is a *mirror*, then the light rays reflected in the mirror are orthogonal to the other branch \mathbf{c}_2 of the cyclographic image. This means, by definition, that \mathbf{c}_2 is an *anticaustic* of the system of reflected rays. The envelope of this system is called *caustic* — clearly the caustic is the evolute of the anticaustic. By Lemma 6.3.3, the curve of regression of γ -developable which contains \mathbf{c}_2 projects onto the caustic.

Example 6.3.11. If the curve \mathbf{c}_1 is a line, the original light rays are parallel to each other. In this case the γ -developable used in our construction which contains \mathbf{c}_1 is planar, and so \mathbf{p} is a planar curve.

If the curve \mathbf{c}_1 is a point, one sheet of the γ -developable which contains \mathbf{p} is a γ -cone whose vertex is \mathbf{c}_1 . \diamond

The cycographic interpretation solves the problem of finding a mirror which converts the pencil of light rays emanating from a given point $\tilde{\mathbf{a}}$ into another system of light rays (cf. [127]). The envelope of the given system of light rays is its caustic $\tilde{\mathbf{k}}$. We construct a curve \mathbf{k} of constant slope $\gamma = \pi/4$ whose orthogonal projection into the reference plane equals the given caustic (if $\tilde{\mathbf{k}} = (x(t), y(t))$, then $\mathbf{k} = (x(t), y(t), s(t))$ with $ds^2 = dx^2 + dy^2$, see (6.51)). Then we intersect the tangent surface \mathcal{R} of \mathbf{k} with any γ -cone whose vertex \mathbf{a} projects onto $\tilde{\mathbf{a}}$. The resulting curve \mathbf{p} has the property that its projection $\tilde{\mathbf{p}}$ is an appropriate mirror.

If the given light rays are not emanating from a point but parallel, we have to use a γ -plane instead of a γ -cone.

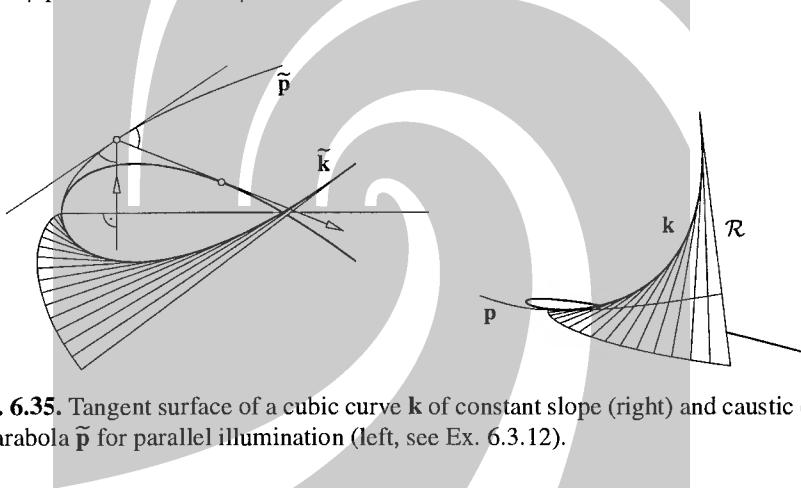


Fig. 6.35. Tangent surface of a cubic curve \mathbf{k} of constant slope (right) and caustic of a parabola $\tilde{\mathbf{p}}$ for parallel illumination (left, see Ex. 6.3.12).

Example 6.3.12. We consider the developable surface \mathcal{R} shown by Fig. 6.35. It is the tangent surface of a cubic curve \mathbf{k} of constant slope. Its intersection with one of its tangent planes (which is a γ -plane) is a parabola \mathbf{p} .

The top view of this shows the following: parallel light rays are reflected in the parabola $\tilde{\mathbf{p}}$, and the envelope of the reflected rays equals \mathbf{k} . The curve $\tilde{\mathbf{k}}$ is a so-called *Tschirnhaus cubic* (cf. also Ex. 6.3.16).

By varying the tangent plane it is possible to show that all caustics of parallel light rays not parallel to the parabola's axis are Tschirnhaus cubics. In the case of incoming light parallel to the axis we have the well known fact that the caustic is the focal point of the parabola.

It is remarkable that all caustics of a parabola with respect to parallel light rays (except the one for light rays parallel to the axis) are similar versions of each other. As a curiosity, we may give the reference [194] to Tschirnhaus' original article. \diamond

Remark 6.3.15. It is possible to study *refraction* according to Snell's law by the same method. It turns out that in order to find the curve where the light rays are refracted at, one has to intersect γ -developables with planes or cones whose inclination angle is different from γ (see Müller and Krames [127], pp. 272–275). ◇

Cyclographic Mapping of Surfaces and Isophotes

The cyclographic model represents a two-parameter family of oriented circles as a surface Φ in \mathbb{R}^3 . We are interested in the planar domain $D(\Phi)$ traced out by the circles $\zeta(x)$ for $x \in \Phi$.

Only two different kinds of points contribute to the boundary of $D(\Phi)$, if Φ is smooth (see Ex. 6.3.13 below): First, Φ 's boundary (if it has one) may have points x such that $\zeta(x)$ touches the boundary of $D(\Phi)$. All other points come from the *cyclographic silhouette* of Φ : Similar to the definition of contour and silhouette of a surface with respect to a central or parallel projection, we use the following

Definition. The cyclographic contour $c(\Phi)$ of a surface Φ is the set of all points $x \in \Phi$ such that the tangent plane at x is a γ -plane. If $\Gamma(\Phi)$ is the envelope of these tangent planes, its intersection with the reference plane is the cyclographic silhouette $s(\Phi)$ of Φ .

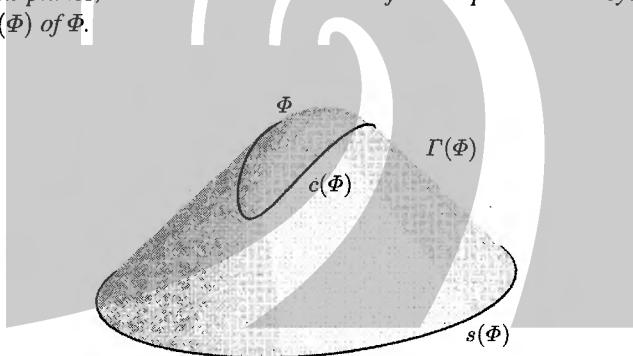


Fig. 6.36. Cyclographic contour $c(\Phi)$ and silhouette $s(\Phi)$ of a surface Φ .

In the generic case the cyclographic contour is either empty (if Φ has no tangent planes of the right slope) or it is a curve. (see Fig. 6.36). The surface $\Gamma(\Phi)$ contains the ideal conic c_ω , which consists of all points $(0, v_1, v_2, v_3) \mathbb{R}$ with $v_1^2 = v_2^2 + v_3^2$. We can say that $\Gamma(\Phi)$ is the developable surface which joins c_ω and Φ .

The set of points of a surface where the tangent planes enclose a given constant angle α with a family of parallel light rays is called an *isophote*. Obviously the cyclographic contour is an isophote for $\alpha = \pi/4$ and vertical light rays.

Example 6.3.13. If Φ is a plane, then the cycles $\zeta(x)$ for $x \in \Phi$ cover the entire reference plane, if the plane does not enclose the angle $\pi/4$ with the z -axis. If this angle equals $\pi/4$, then the cycles $\zeta(x)$ are tangent to $\Phi \cap \Pi$.

This shows that for a smooth surface only those points whose tangent plane is a γ -plane can contribute to the cyclographic contour. ◇

Example 6.3.14. The simplest surfaces which may be used for smooth approximation of smooth surfaces are *quadrics*. Therefore it is interesting to discuss the cyclographic contour and silhouette of a quadric Φ .

Duality in projective space converts the set of planes tangent to both a quadric Φ and the conic c_ω to the set of points contained in both a quadric Φ^* and a cone c_ω^* . Therefore $\Gamma(\Phi)$ is a developable surface of class at most four (and in general, non-rational of class four), which shows that the cyclographic silhouette $s(\Phi)$ is a planar curve of class at most four (in general, non-rational of class four).

We do not go into details here, but mention that studying the various cases of degree reductions is dual to studying the completely analogous cases of degree reduction of central projections of the curve $\Phi^* \cap c_\omega^*$.

Further it turns out that in the generic case $\Gamma(\Phi)$ intersects itself in four conics, one of which is c_ω . The planes of the remaining three contain the quadric's center, if it has one. The cyclographic silhouette $s(\Phi)$ therefore coincides with the cyclographic images of any of these three conics ('hypercycles', see W. Blaschke [13]).

Fig. 6.37 illustrates the case of an ellipsoid. In this example the boundary of $D(\Phi)$ has two components (in the topological sense), which correspond to the two sheets of $\Gamma(\Phi)$. It is however an irreducible algebraic curve and $\Gamma(\Phi)$ is an irreducible algebraic surface. The envelope of the 2-parameter system of cycles $\zeta(x)$ ($x \in \Phi$) contains the boundary of the domain covered by this family of cycles. \diamond

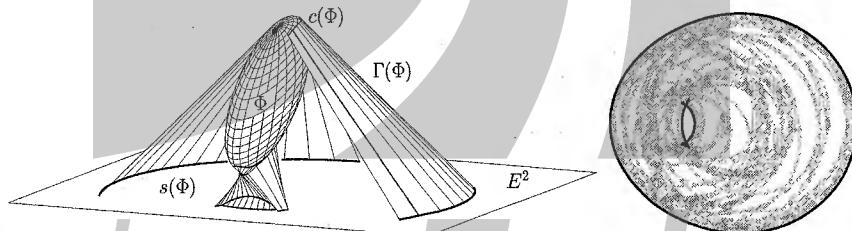


Fig. 6.37. Left: γ -developable $\Gamma(\Phi)$ tangent to an ellipsoid Φ . Right: 2-parameter family of cycles corresponding to Φ and its envelope $s(\Phi)$.

Although it is not necessary for the computation of the silhouette, it is interesting to ask for properties of the *cyclographic contour* of a quadric, and, more generally for properties of isophotes:

Proposition 6.3.17. *Isophotes (and cyclographic contours) of quadrics are algebraic curves of degree ≤ 4 .*

Proof. The polarity κ defined by the quadric Φ maps the points of Φ to its tangent planes and vice versa. Thus κ maps the planes of $\Gamma(\Phi)$ to the points of the contour $c(\Phi)$. By Ex. 6.3.14, $\Gamma(\Phi)$ is algebraic of class ≤ 4 , and so $c(\Phi)$ is algebraic of

degree ≤ 4 . The same argument holds for arbitrary isophotes (i.e., cyclographic contours with respect to an angle γ different from $\pi/4$). \square

Example 6.3.15. We use homogeneous coordinates and consider a quadric with equation $\mathbf{x}^T \cdot A \cdot \mathbf{x} = 0$. By Lemma 1.1.21, its dual Φ^* (the set of its tangent planes) has the equation $\mathbf{u}^T \cdot A^{-T} \cdot \mathbf{u} = 0$.

The conic c_ω which consists of the ideal points of lines which enclose the angle γ with the z -axis can be parametrized by $(0, \cos \phi \sin \gamma, \sin \phi \sin \gamma, \cos \gamma)$ and has the equation $\cos^2 \gamma (x_1^2 + x_2^2) - \sin^2 \gamma x_3^2 = 0$, if x_1, x_2, x_3 are coordinates in the ideal plane. The coordinate matrix of c_ω then obviously equals $\text{diag}(\cos^2 \gamma, \cos^2 \gamma, -\sin^2 \gamma)$, and by Lemma 1.1.21, its dual c_ω^* has the equation $(u_1^2 + u_2^2)/\cos^2 \gamma - u_3^2/\sin^2 \gamma = 0$, or, equivalently,

$$(u_1^2 + u_2^2) \sin^2 \gamma - u_3^2 \cos^2 \gamma = 0. \quad (6.58)$$

If $\mathbb{R}(u_0, \dots, u_3)$ is a plane of projective three-space, its intersection with the ideal plane has the equation $x_1 u_1 + x_2 u_2 + x_3 u_3 = 0$, so (6.58) is also the equation of all planes tangent to c_ω . If $B = \text{diag}(0, \sin^2 \gamma, \sin^2 \gamma, -\cos^2 \gamma)$, a plane $\mathbb{R}\mathbf{u}$ tangent to both Φ and to c_ω must fulfill $\mathbf{u}^T \cdot A^{-T} \cdot \mathbf{u} = \mathbf{u}^T \cdot B \cdot \mathbf{u} = 0$. The isophote is the κ -preimage of this set, where κ is the polarity $(\mathbf{x}\mathbb{R})\kappa = \mathbb{R}(A \cdot \mathbf{x})$ of Φ . So the isophote has the equation

$$\mathbf{x}^T \cdot A \cdot \mathbf{x} = 0, \quad \mathbf{x}^T \cdot A^T B A \cdot \mathbf{x} = 0.$$

We see that the isophote is the intersection of the original quadric with a second quadratic variety, which is a cone, because $\text{rk}(A^T B A) = \text{rk}(B) = 3$. An example is shown in Fig. 6.38. \diamond

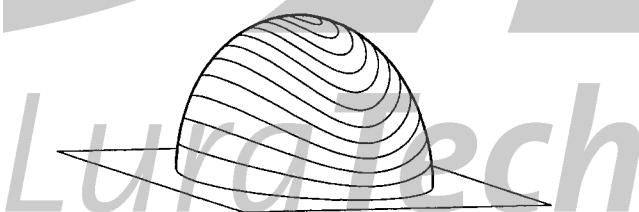


Fig. 6.38. Isophotes of a quadric (cf. Ex. 6.3.15).

6.3.3 Rational Developable Surfaces of Constant Slope and Rational Pythagorean-Hodograph Curves

In various applications including NC machining, layered manufacturing and geometric tolerancing, *offsets* play an important role. Therefore it is currently an important problem to represent offsets of curves by rational B-splines. This of course is

not always possible because many offsets of curves are not rational (cf. Ex. 6.3.8 and Prop. 6.3.13), and so in practice they get approximated (see the survey by T. Maekawa [119]). Here we are going to describe a class of curves with rational offsets.

Rational Parametrizations of the Unit Circle

Assume a planar regular curve $\mathbf{l}(t) = (x(t), y(t))$. Its unit normal vector field $\mathbf{n}(t)$ and its offset curve \mathbf{l}_d at distance d are given by

$$\mathbf{n}(t) = \frac{1}{\sqrt{\dot{x}^2(t) + \dot{y}^2(t)}} (\dot{y}(t), -\dot{x}(t)), \quad \mathbf{l}_d(t) = \mathbf{l}(t) + d\mathbf{n}(t), \quad (6.59)$$

Because of the square root, a polynomial or rational parametrization of \mathbf{l} will in general not yield a polynomial or rational parametrization of the offset.

A class of polynomial curves $\mathbf{k}(t)$ with rational offsets has been derived by Farouki and Sakkalis [53] as follows. The vector field $\mathbf{n}(t)$ of Equ. (6.59) is rational, if there exists a polynomial $\sigma(t)$ such that

$$\dot{x}^2(t) + \dot{y}^2(t) = \sigma^2(t). \quad (6.60)$$

The curve defined by the first derivative vectors $(\dot{x}(t), \dot{y}(t))$ is called the *hodograph* of \mathbf{k} . A hodograph satisfying (6.60) is called a *Pythagorean hodograph*. The curve \mathbf{k} itself, obtained by integrating the hodograph, is called a polynomial *Pythagorean-hodograph curve* or briefly *PH curve*.

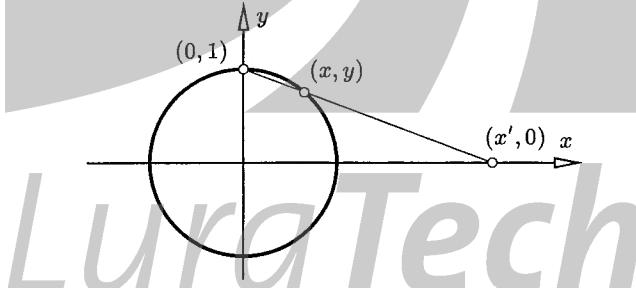


Fig. 6.39. Stereographic projection of the circle.

In order to find polynomials $x(t), y(t)$ such that (6.60) holds, we note that the curve $c(t) = (\sigma(t), \dot{x}(t), \dot{y}(t))\mathbb{R}$ parametrizes the unit circle, if and only if (6.60) holds. There is the following

Lemma 6.3.18. *For all rational parametrizations $(n_1(t), n_2(t))$ of the unit circle there exist polynomials $a(t), b(t)$ such that*

$$(n_1(t), n_2(t)) = \frac{1}{a(t)^2 + b(t)^2} (2a(t)b(t), a(t)^2 - b(t)^2). \quad (6.61)$$

If a, b are relatively prime, then so are $a^2 + b^2$, $2ab$ and $a^2 - b^2$.

Proof. We consider the stereographic projection σ with center $(0, 1)$ from the unit circle onto the x -axis (see Fig. 6.39). The mapping has the equation $(x, y) \mapsto (x/(1-y), 0)$ and its inverse has the equation $(x', 0) \mapsto (-2x', x'^2+1)/(x'^2+1)$, and so it is obvious that \mathbf{n} is rational if and only if $\sigma \circ \mathbf{n}$ is. If we choose $x'(t) = a(t)/b(t)$, we get Equ. (6.61). If the polynomials $a^2 + b^2$ and $a^2 - b^2$ have a common divisor, then so have their sum and difference, i.e., $2a^2$ and $2b^2$. This shows that also a and b have a common divisor. \square

Lemma 6.3.18 shows that a parametrization of the unit circle with homogeneous coordinates $(\sigma(t), \dot{x}(t), \dot{y}(t))$ has the form

$$\begin{aligned}\dot{x}(t) &= 2b(t)w(t), & \dot{y}(t) &= (a(t)^2 - b(t)^2)w(t), \\ \sigma(t) &= (a(t)^2 + b(t)^2)w(t).\end{aligned}\quad (6.62)$$

Thus the construction of Pythagorean-hodograph curves relies on the choice of three polynomials $a(t)$, $b(t)$, $w(t)$, and on integration of the hodograph (\dot{x}, \dot{y}) given by equation (6.62).

As a zero of $w(t)$ introduces a singularity in (6.62) and of course the choice of $w(t) \neq 1$ increases the degree of the parametrization, we often let $w = 1$. In this case the choice of polynomials a, b of degree m leads to a Pythagorean hodograph $(\dot{x}(t), \dot{y}(t))$ of degree $2m$ and a Pythagorean-hodograph curve $(x(t), y(t))$ of degree $2m + 1$. The offsets (6.59) are rational and in general of degree $\leq 4m + 1$.

Another important property of polynomial Pythagorean-hodograph curves is that their arc length function

$$s(t) = \int \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt = \int \sigma(t) dt,\quad (6.63)$$

is *polynomial*. This fact has been exploited in [56], where PH curves are applied to CNC interpolation. It turns out that the use of PH curves is advantageous for tool motion planning.

Polynomial Curves of Constant Slope

Assume $\dot{x}(t)$, $\dot{y}(t)$, and $\sigma(t)$ as in (6.62), and consider the curve

$$\mathbf{c}(t) = (2 \int abw dt, \int (a^2 - b^2)w dt, \int (a^2 + b^2)w dt).\quad (6.64)$$

This is a *polynomial curve of constant slope* whose tangents enclose the angle $\gamma = \pi/4$ with the z -axis. (cf. Equ. (6.50)). Other constant slopes can be achieved by an appropriate scaling of the z -coordinate. Obviously the orthogonal projection of \mathbf{c} into the reference plane $z = 0$ is a Pythagorean-hodograph curve \mathbf{k} and its z coordinate function coincides with the arc length function of \mathbf{k} (see (6.63)). The following result shows how PH curves and polynomial curves of constant slope are related:

Proposition 6.3.19. All polynomial curves \mathbf{c} of constant slope $\gamma = \pi/4$ with respect to the reference plane $z = 0$ can be expressed in the form (6.64), where $a(t)$, $b(t)$, $w(t)$ are arbitrary polynomials.

The orthogonal projection \mathbf{k} of \mathbf{c} into the reference plane is a Pythagorean-hodograph curve. The involutes of \mathbf{k} are rational curves with rational offsets.

Proof. We have already seen that for all polynomials a, b, w Equ. (6.64) is a polynomial curve of constant slope. Conversely, the curve $\mathbf{c}(t) = (x(t), y(t), z(t))$ is of constant slope $\pi/4$, if and only if $\dot{z}^2 = \dot{x}^2 + \dot{y}^2$, and so $(\dot{x}(t), \dot{y}(t), \dot{z}(t))\mathbb{R}$ is a rational parametrization of the unit circle. Lemma 6.3.18 shows that $\dot{x}, \dot{y}, \dot{z}$ have the form (6.62) with z instead of σ , and we have shown that (x, y, z) must be of the form (6.64). \square

The statement that $\mathbf{k} = (x(t), y(t), 0)$ is a Pythagorean-hodograph curve is now clear. Consider the developable surface \mathcal{R} of constant slope which contains the curve \mathbf{c} . By Th. 6.3.4, \mathcal{R} is \mathbf{c} 's tangent surface and therefore rational. By Lemma 6.3.3, its planar sections with planes $z = \text{const.}$ (which are rational) are the involutes of \mathbf{k} , and are offsets of each other. \square

Example 6.3.16. The simplest non-trivial choice for $a(t)$, $b(t)$, $w(t)$ is $a(t) = t$, $b(t) = 1$, $w(t) = 1$, which results in the Pythagorean triple $\dot{x}(t) = 2t$, $\dot{y}(t) = t^2 - 1$, $\sigma(t) = t^2 + 1$. The corresponding curve \mathbf{c} of constant slope is the polynomial cubic

$$\mathbf{c}(t) = (t^2, t^3/3 - t, t^3/3 + t).$$

Its projection onto the reference plane is the cubic PH curve $\mathbf{k}(t) = (t^2, t^3/3 - t)$. Fig. 6.35 shows the curve \mathbf{c} together with its tangent surface \mathcal{R} and the curve \mathbf{k} (denoted by \mathbf{k} , \mathcal{R} , and $\tilde{\mathbf{k}}$, respectively).

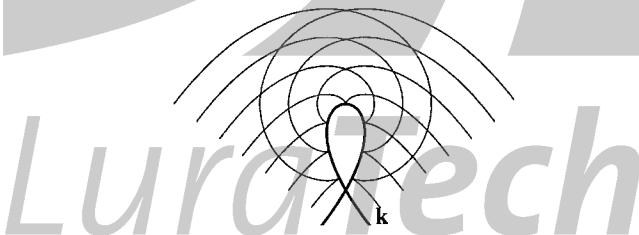


Fig. 6.40. The involutes of the Tschirnhaus cubic are rational curves of degree four and offsets of each other.

In Ex. 6.3.12 we have already discussed this curve — it is a *Tschirnhaus cubic*. It is easy to convince oneself that any other choice of linear polynomials a, b gives the same PH curve up to Euclidean similarities and linear re-parametrizations. Further it can be shown that there are no other cubic PH curves [53].

The tangent surface of \mathbf{c} is rational of order four, so the family of involutes of \mathbf{k} is a family of rational quartics, which are offset curves of each other. They are called *Tschirnhaus quartics* (see Fig. 6.40). \diamond

Rational Developables of Constant Slope

The discussion of Pythagorean-hodograph curves illustrates the close connection of this subject with the geometry of rational developable surfaces of constant slope. Therefore, we will now study these surfaces in more detail. This will lead to generalizations of polynomial Pythagorean-hodograph curves.

We restrict ourselves to the constant slope $\gamma = \pi/4$, since all other constant slopes can be obtained from this one by a scaling of z -coordinates.

Theorem 6.3.20. *A rational developable surface of constant slope $\gamma = \pi/4$ is the envelope of planes $U(t) = \mathbb{R}\mathbf{u}(t)$ with*

$$\mathbf{u} = (g, 2abf, (a^2 - b^2)f, (a^2 + b^2)f), \quad (6.65)$$

with polynomials a, b, f, g such that the greatest common divisor of a, b equals 1.

Proof. By Equ. (6.46), a developable surface of constant slope $\gamma = \pi/4$ is the envelope of planes

$$U(t) : x \sin \phi(t) - y \cos \phi(t) + z = h(t).$$

Since we are interested in rational surfaces, we must parametrize this family of planes by rational functions. The curve $(\sin \phi(t), -\cos \phi(t))$ parametrizes part of the unit circle, and so Lemma 6.3.18 shows that necessarily $(\sin \phi, \cos \phi) = (2ab, (a^2 - b^2))/(a^2 + b^2)$ with polynomials a, b , and that $U(t)$ is parametrized by

$$U(t) : \frac{2a(t)b(t)}{a(t)^2 + b(t)^2}x + \frac{a(t)^2 - b(t)^2}{a(t)^2 + b(t)^2}y + z = h(t), \quad (6.66)$$

with polynomial functions $a(t), b(t)$ and a rational function $h(t)$. By letting $h(t) \cdot (a(t)^2 + b(t)^2) = -g(t)/f(t)$ with polynomials f, g , we obtain Equ. (6.65). \square

Example 6.3.17. The simplest nontrivial case of rational surfaces of constant slope is obtained by choosing a, b, f as linear and g as a cubic polynomial. It turns out that the resulting surfaces are of class three, and they are exactly the tangent surfaces of polynomial cubics of constant slope (see Fig. 6.35). \diamond

In order to find the dual Bézier representation of a segment of such a surface, one has to express the polynomials of (6.65) in the Bernstein basis. There is a nice geometric characterization of the corresponding control and frame planes:

Theorem 6.3.21. *A rational developable Bézier surface \mathcal{R} is of constant slope if and only if the control and frame planes of its dual geometric control structure are parallel to the respective control and frame planes of a Bézier surface contained in a γ -cone.*

Proof. Consider families $U(t) = (u_0(t), \dots, u_3(t))$ and $\bar{U}(t) = (0, u_1(t), u_2(t), u_3(t))$ of planes. The planes $U(t)$ and $\bar{U}(t)$ are parallel, and $\bar{U}(t)$ contains the origin. Thus $U(t)$ is a family of tangent planes of a developable surface of constant slope γ if $\bar{U}(t)$ is tangent to a γ -cone, and vice versa.

Because the Bézier representation of an \mathbb{R}^4 -valued polynomial function works component-wise, and changing the zeroth coordinate of a homogeneous plane coordinate vector effects a translation, the result follows. (Of course, the translation is not purely arbitrary, but such that the resulting planes are still the dual control structure of a surface, i.e., F_i , U_i and U_{i+1} contain a common line). \square

Rational Curves of Constant Slope

A tangent developable of a curve is rational if and only if the curve is rational. Thus Th. 6.3.4 shows that the γ -developables are precisely the tangent surfaces of curves of constant slope, and we can determine all rational curves of constant slope by computing the curve of regression of γ -developables. A representation for the latter has been given by Th. 6.3.20.

Theorem 6.3.22. *The rational curves of constant slope $\gamma = \pi/4$ with respect to the plane $z = 0$ are curves of the form $\mathbf{c}(t) = (x(t), y(t), z(t))$, with*

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} H(b^2 - a^2) \\ 2abH \\ h(a^2 + b^2) \end{bmatrix} + \frac{\dot{H}}{2(\dot{a}b - a\dot{b})} \begin{bmatrix} ab \\ b^2 - a^2 \\ a^2 + b^2 \end{bmatrix} \quad (6.67)$$

and relatively prime polynomials a, b , a rational function $h(t)$, and

$$H = (a^2 + b^2)\dot{h} / 2(\dot{a}b - a\dot{b}).$$

Proof. We start from a rational γ -developable, whose family $U(t)$ of tangent planes is parametrized by (6.66). The first and second derivative planes $U^1(t)$ and $U^2(t)$ are then given by

$$U^1(t) : \frac{b^2 - a^2}{a^2 + b^2}x + \frac{2ab}{a^2 + b^2}y = H, \text{ with } H(t) = \frac{(a^2 + b^2)\dot{h}}{2(\dot{a}b - a\dot{b})}, \quad (6.68)$$

$$U^2(t) : \frac{-2ab}{a^2 + b^2}x + \frac{b^2 - a^2}{a^2 + b^2}y = \frac{(a^2 + b^2)\dot{H}}{2(\dot{a}b - a\dot{b})}. \quad (6.69)$$

The curve of regression is obtained by computing the intersection $\mathbf{c}(t) = U(t) \cap U^1(t) \cap U^2(t)$, which gives (6.67). \square

Example 6.3.18. We have already discussed surfaces of class three (see Ex. 6.3.16, Ex. 6.3.17) and now investigate those of class four. Equ. (6.65) is of order four if (i) both a and b are linear, f is quadratic, and g is of degree ≤ 4 , or (ii) both a and b are quadratic, f is constant, and g has degree ≤ 4 .

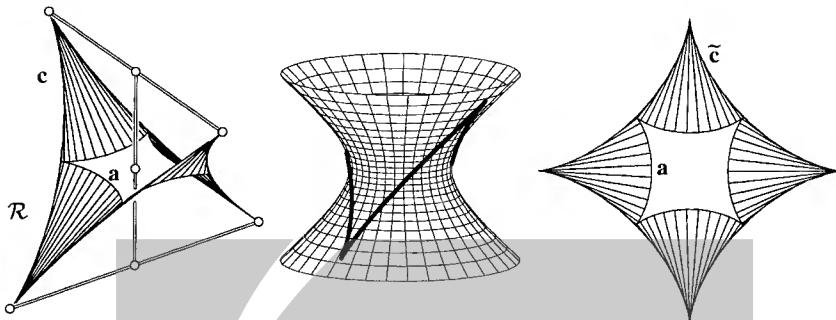


Fig. 6.41. Left: Rational γ -developable \mathcal{R} of class 4 with curve of regression \mathbf{c} and planar section \mathbf{a} . Center: \mathbf{c} is contained in a hyperboloid. Right: Top view. The curves $\tilde{\mathbf{c}}$ and \mathbf{a} are astroids.

Translating all planes $U(t) = \mathbb{R}\underline{u}(t) = \mathbb{R}(u_0(t), \dots, u_3(t))$ such that they contain the origin results in the family $\bar{U}(t) = \mathbb{R}(0, u_1(t), u_2(t), u_3(t))$ of planes, which are tangent to a γ -cone. This dual parametrization of a γ -cone (which is also referred to in the proof of Th. 6.3.21) is quartic in both cases. In case (i) it comes from a quadratic parametrization multiplied by a quadratic scalar factor, and in case (ii) it is a ‘true’ quartic parametrization.

For details we refer the reader to [145]. Here we describe just a few interesting special cases. First we choose

$$a(t) = t, \quad b(t) = 1, \quad f(t) = t^2 + 1,$$

and $g(t)$ as an arbitrary polynomial of degree four. Recall the proof of Th. 6.3.20 and the representation of the family of tangent planes $U(t)$ in the form (6.46). Our choice of a, b and f means that $h(t)$ equals $-g(t)/(1+t)^2$. Recall that $\sin \phi = 2t/(1+t^2)$, $\cos \phi = (1-t^2)/(1+t^2)$. This shows that if we use ϕ as parameter, $h(\phi)$ is of the form

$$h(\phi) = a_0 + a_1 \sin \phi + a_2 \cos \phi + a_3 \sin 2\phi + a_4 \cos 2\phi.$$

By a parallel translation we can achieve $a_0 = a_1 = a_2 = 0$, and by an appropriate choice of the x -axis we see that without loss of generality we can write

$$h(\phi) = a \cos 2\phi.$$

Such surfaces have been studied in Ex. 6.3.2 and Ex. 6.3.3. The top view of both the curve of regression and the planar section with the plane $z = 0$ are *astroids* (see Fig. 6.41). The curve of regression is of order six and is contained in a one-sheeted hyperboloid of revolution.

It can be shown that surfaces defined by $h(\phi) = a \cos(n\phi)$ are rational if and only if n is rational. The case $n = 1/2$ leads to a surface of class four and order six, shown by Fig. 6.22. \diamond

Example 6.3.19. (Continuation of Ex. 6.3.18) Another example of rational γ -developables \mathcal{R} of type (i) which we have already encountered occurs if we let $a = t$, $b = 1$, $g = -(1 + t^2)^2$, $f = 2t$. The intersection of \mathcal{R} with the plane $z = 0$ has the support function

$$h(\phi) = 1 / \sin \phi,$$

which shows that this intersection is a parabola. The surface \mathcal{R} is therefore the γ -developable which contains this parabola (cf. Ex. 6.3.9).

Remarkable examples of type (ii) belong to $a = t^2$, $b = 1 + ct^2$, $g = 1$. The corresponding surfaces \mathcal{R} are tangent surfaces of rational quartics of constant slope, whose horizontal planar sections have an inflection point, which corresponds to an ideal point of the curve of regression (see Fig. 6.42).

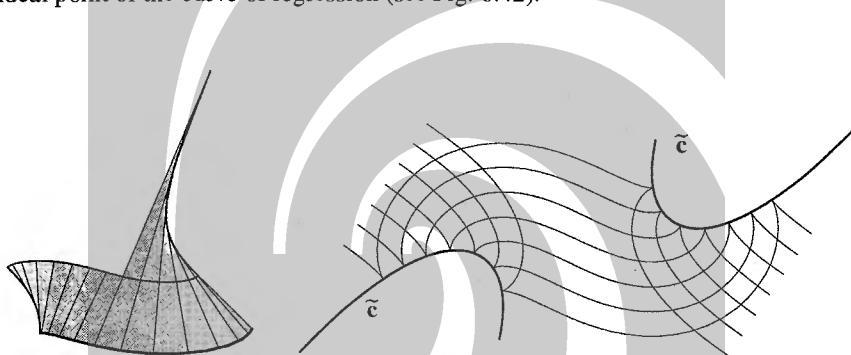


Fig. 6.42. Left: Rational γ -developable \mathcal{R} of class 4 and order 5 with a quartic curve \mathbf{c} of regression. Right: Horizontal planar sections of \mathcal{R} are rational curves and offsets of each other.

These surfaces are of degree five as projective algebraic varieties. Other surfaces of class four and degree five belong to type (i): they are the tangent surfaces of polynomial Pythagorean-hodograph quartics. Like the previous example they appear in a paper by W. Wunderlich [217] on cubic and quartic curves of constant slope. \diamond

Rational Curves with Rational Arc Length Function

We have characterized polynomial Pythagorean-hodograph curves as those with a polynomial arc length function and shown that they are orthogonal projections of polynomial curves of constant slope. The generalization to *rational* curves is described by the following theorem [144].

Theorem 6.3.23. Any rational curve $\mathbf{k}(t) = (x(t), y(t), 0)$ in the plane $z = 0$ which has a rational arc length function $s(t)$ is the orthogonal projection of a rational curve of constant slope onto this plane, and vice versa. An explicit representation is given by (6.67). The offset of \mathbf{k} at distance d is rational and is found by substituting $H + d$ for H in (6.67).

Proof. If the arc length function $s(t) = \int_{t_0}^t (\dot{x}(\tau)^2 + \dot{y}(\tau)^2)^{1/2} d\tau$ is rational, then the curve $(x(t), y(t), s(t))$ is rational and of constant slope, which shows the first statement.

We have a closer look at the proof of Th. 6.3.22. The point $\mathbf{c}(t) = (x(t), y(t), s(t))$ is found by intersecting the planes $U(t)$, $U^1(t)$, $U^2(t)$. By Equ. (6.68), the plane $U^1(t)$ is vertical and its distance to the origin equals H . We know that the ruling $R(t)$ of \mathbf{c} 's tangent surface equals $U(t) \cap U^1(t)$, which shows that the top view of $U^1(t)$ is tangent to the top view \mathbf{k} of \mathbf{c} . We can therefore parametrize an offset of this top view by substituting $H + d$ for H . \square

Rational Pythagorean-Hodograph Curves

Rational curves with a rational arc length function have the property that all their offset curves are rational. We have this in mind when we define

Definition. A curve which has a rational parametrization $\mathbf{l}(t)$ such that the corresponding unit normal vector is also rational, is called a rational Pythagorean-hodograph (PH) curve.

Note that being a PH curve is less than being a curve with rational arc length function. A PH curve is a curve with a rational arc length differential, whose integral need not be a rational function.

Remark 6.3.16. Note that the Pythagorean-hodograph nature of a rational curve cannot be decided immediately by looking at an arbitrary rational parametrization. A good example is the parabola. The usual polynomial quadratic parametrization does not have a rational normal vector field, but the following parametrization has one.

$$\mathbf{k}(t) = \frac{1}{4t^4 + 4t^2} (5t^2 - 1 + 5t^4 - t^6, -4t + 4t^5)$$

This parametrization is found if we compute a planar intersection of the rational γ -developable of Ex. 6.3.19 (see also Th. 6.3.24). \diamond

Th. 6.3.23 is not the only source of rational Pythagorean-hodograph curves. This class of curves is characterized by the following theorem:

Theorem 6.3.24. The rational Pythagorean-hodograph curves $\mathbf{l}(t)$ are the horizontal planar sections of rational γ -developables. The class of rational PH curves further coincides with the class of involutes of rational curves with rational arc length. An explicit representation of such curves is given by

$$x(t) = \frac{2ab}{a^2 + b^2} h - \frac{a^2 - b^2}{2(ab - ab)} \dot{h}, \quad y(t) = \frac{a^2 - b^2}{a^2 + b^2} h + \frac{ab}{ab - ab} \dot{h}, \quad (6.70)$$

with polynomial functions $a(t)$, $b(t)$ and a rational function $h(t)$. Substituting $h(t) + d$ for $h(t)$ in (6.70) yields the offset of \mathbf{l} at distance d .

Proof. The family $U(t)$ of tangent planes of a rational γ -developable \mathcal{R} is, according to Th. 6.3.20, given by (6.65). We intersect with the plane $z = 0$ and get the family of lines

$$L(t) : \frac{2ab}{a^2 + b^2}x + \frac{a^2 - b^2}{a^2 + b^2}y = h. \quad (6.71)$$

The intersection curve is found by computing $L(t) \cap L^1(t)$, which results in (6.70). Obviously such a curve is a PH curve. The distance of the line $L(t)$ to the origin equals $h(t)$, so by adding d to $h(t)$ we get an offset at distance d . The orthogonal projection of \mathcal{R} 's curve of regression is the evolute of this planar section, according to Lemma 6.3.3.

By Th. 6.3.23, these top projections are precisely the class of rational curves with rational arc length. It remains to show that this description exhausts all PH curves: If \mathbf{k} is a PH curve, embed it into the plane $z = 0$. Consider the offset $\mathbf{k}_1(t)$ at distance 1 (which is rational) and apply the translation $z \rightarrow z + 1$ to it, thus defining a curve $\mathbf{k}'_1(t)$. The surface whose rulings are the lines $\mathbf{k}(t) \vee \mathbf{k}'_1(t)$ is a rational γ -developable by definition, and \mathbf{k} is its horizontal planar section. \square

Dual Representation of Rational Pythagorean-Hodograph Curves

The dual of a rational PH curve has a simpler representation than the curve itself. There are the following theorems:

Theorem 6.3.25. *The family of tangents $L(t) = \mathbb{R}\mathbf{u}(t)$ of a rational Pythagorean-hodograph curve can be written in the form*

$$\mathbf{u} = (g, 2abf, (a^2 - b^2)f), \quad (6.72)$$

where a, b, f, g are arbitrary polynomials in t . The dual representation $\mathbb{R}\mathbf{u}_d(t)$ of its offset at distance d is given by

$$\mathbf{u}_d = (g - d(a^2 + b^2)f, 2abf, (a^2 - b^2)f). \quad (6.73)$$

Proof. By Th. 6.3.24, we have to intersect the rational γ -developable, whose dual is given by Equ. (6.65) with the plane $z = 0$. This results in (6.72). By increasing the distance of $L(t)$ to the origin by d we get (6.73). \square

Theorem 6.3.26. *Consider the Bézier lines B_i^* and frame lines F_i^* of a dual control structure of a circular arc. By arbitrarily translating these lines such that still B_i^* , F_i^* , and B_{i+1}^* are concurrent, we get the dual control structure of a rational PH curve, and this construction exhausts the class of PH curves.*

Proof. This is an immediate consequence of Th. 6.3.21 and Th. 6.3.24. \square

Th. 6.3.26 is illustrated by Fig. 6.43 with a circular arc of radius r . The figure also shows the Minkowski-like addition of the dual control structures of the circle $\bar{\mathbf{I}}$ and the curve \mathbf{I} to get the dual control structure of its offset \mathbf{I}_d at distance d [145]. The figure illustrates the case of dual degree (i.e., class) three. By Ex. 6.3.17, the curves \mathbf{I} and $\bar{\mathbf{I}}$ are segments of Tschirnhaus quartics.

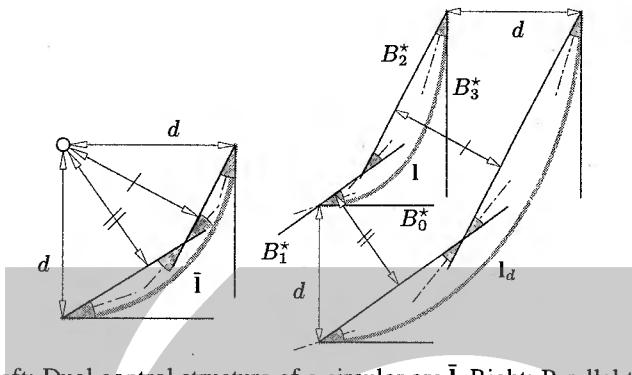


Fig. 6.43. Left: Dual control structure of a circular arc \bar{I} . Right: Parallel translation of control and frame lines gives the dual control structure of the PH curve I — here shown with its offset at distance d .

Remark 6.3.17. Rational Pythagorean-hodograph curves of class four are horizontal planar sections of rational γ -developables of class four. Special cases of degree five have been studied in Ex. 6.3.18 and Ex. 6.3.19. Such curves have been used for the design of curvature continuous (G^2) rational splines with rational offsets by H. Pottmann [145]. For the construction of tangent continuous (G^1) rational spline curves with rational offsets it is sufficient to use segments of rational Pythagorean-hodograph curves of class three, i.e., segments of Tschirnhaus quartics [174, 2]. Both classes of spline curves have the property that also their offsets belong to the class. ◇

Remark 6.3.18. At this point we should mention that *polynomial* PH curves do not have polynomial offsets, but rational ones. For example a polynomial PH cubic has rational offsets of order five, and polynomial PH curves of order five (which have been used by Farouki and Neff [51] for G^1 curve design) are embedded in an offset family of rational PH curves of degree nine.

A detailed study of the class of polynomial PH curves within the larger class of rational ones has been performed by Farouki and Pottmann [52]. The higher degree of their offsets does not reduce their attractivity for applications, since other features of polynomial PH curves, such as polynomial arc length and simple feed rate control in NC machining are not shared by rational PH curves. ◇

PH Curves in Geometrical Optics

We continue the discussion of applications in geometrical optics, which began at p. 379.

Theorem 6.3.27. *The class of non-circular rational PH curves coincides with the class of anticaustics of rational curves for parallel illumination. The corresponding caustics comprise the class of rational curves with rational arc length function.*

Proof. (Sketch) Consider a γ -developable \mathcal{R} which is not a cone and not a plane. Its curve of regression is denoted by \mathbf{k} , and orthogonal projection into the base plane will be indicated by the symbol $\tilde{\cdot}$. A planar section $\mathbf{p}(t)$ of \mathcal{R} with a γ -plane is rational. We think of the curve $\tilde{\mathbf{p}}$ as of a mirror. Then $\tilde{\mathbf{k}}$ is a caustic and any horizontal planar section \mathbf{c} gives rise to an anticaustic $\tilde{\mathbf{c}}$ (see the previous discussion in Sec. 6.3.2). By Th. 6.3.24 and Th. 6.3.23, the caustic and anticaustic have the required properties.

Conversely, we have to show that for any rational mirror curve $\tilde{\mathbf{k}}$ and corresponding curve \mathbf{k} in a γ -plane, the γ -developable which contains \mathbf{k} is rational. We leave the verification of this fact to the reader (see also [57, 139, 154]). \square

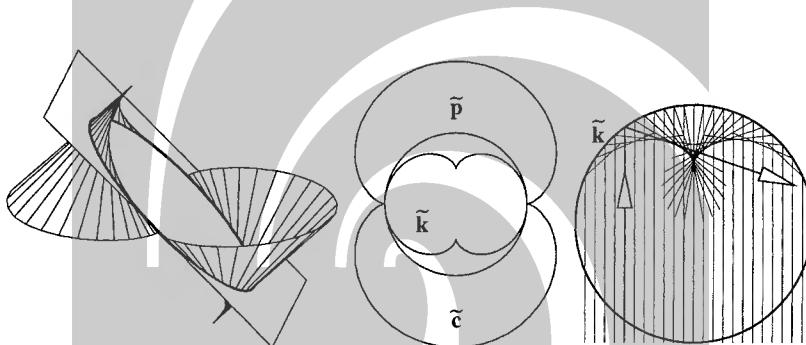


Fig. 6.44. Left: Intersection of γ -developable with γ -plane. Center and Right: Top view. The caustic $\tilde{\mathbf{k}}$ is a nephroid, and so is one anticaustic $\tilde{\mathbf{c}}$.

Example 6.3.20. We have seen that the Tschirnhaus cubic is the caustic and the Tschirnhaus quartics are anticaustics of a parabola for parallel illumination (see Ex. 6.3.12, Fig. 6.35, and Fig. 6.40).

Another simple example is shown in Fig. 6.44. The intersection of the developable surface discussed in Ex. 6.3.3 with a special γ -plane is an ellipse \mathbf{p} , whose orthogonal projection into the base plane is a circle $\tilde{\mathbf{p}}$. The curve \mathbf{e} of regression projects to a nephroid $\tilde{\mathbf{k}}$. It turns out that also one of the anticaustics is a nephroid. \diamond

Generalizations of Pythagorean-Hodograph Curves

A polynomial *Pythagorean-hodograph space curve* is a polynomial curve $\mathbf{l}(t) = (x(t), y(t), z(t))$ in Euclidean three-space whose unit tangent vector field is rational. It is necessary and sufficient that the derivative functions satisfy

$$\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t) = \sigma^2(t),$$

with a polynomial $\sigma(t)$. This generalized Pythagorean relation can be solved via the explicit representation (8.20) of rational curves on the unit sphere given by Dietz

et al. (see [36] and Th. 8.1.13). Then integration gives an explicit representation of all Pythagorean-hodograph space curves. For a geometric investigation and applications of these curves, we refer to Farouki and Sakkalis [54] and Jüttler and Mäurer [89].

Rational Pythagorean-hodograph space curves may be constructed via their duals [143]. Special cases are the rational curves of constant slope.

The construction of *rational surfaces with rational offsets* or *Pythagorean-normal (PN) surfaces* is analogous to that of planar PH curves. PN surfaces are — in general — envelopes of a two-parameter family of planes

$$n_1(u, v)x + n_2(u, v)y + n_3(u, v)z = h(u, v),$$

where $(n_1(u, v), n_2(u, v), n_3(u, v))$ is a rational parametrization of the unit sphere, and $h(u, v)$ is a rational function. Many properties are analogous to those of rational PH curves [144, 139]. We will discuss PN surfaces in Sec. 7.1.4.

Developable surfaces with rational offsets are derived as envelopes of a one-parameter family of planes with rational unit normals [144, 175]. Special polynomial surfaces with rational offsets have been investigated and used for surface design by Jüttler and Sampoli [90].

The class of Pythagorean-normal surfaces contains all regular quadrics [114, 139]. It is surprising that the offsets of regular quadrics are rational, whereas the offsets of conics (except for parabola and circle) are not. Another surprising fact is that offsets of skew rational ruled surfaces are rational [152].

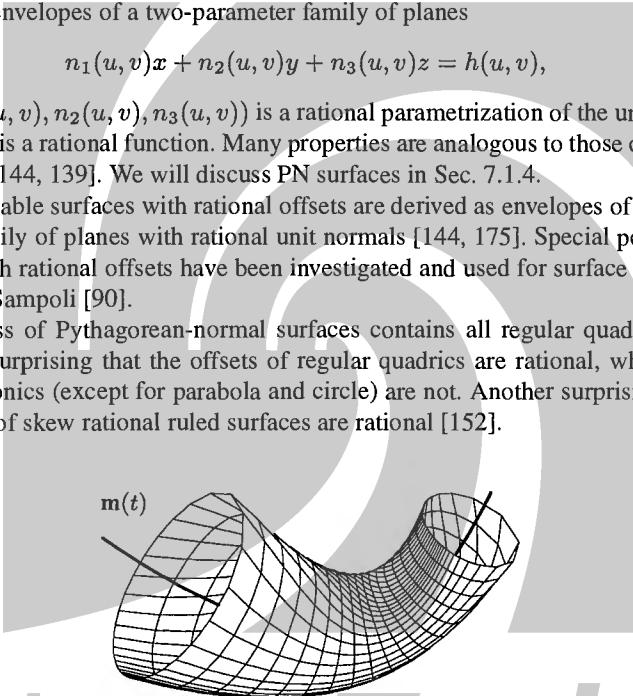


Fig. 6.45. Rational canal surface.

The envelope of a one-parameter family of spheres with centers $\mathbf{m}(t)$ and radius $r(t)$ is called *canal surface*. Canal surfaces with rational spine curve $\mathbf{m}(t)$ and rational radius function $r(t)$ have a rational parametrization [103, 135, 138, 139]. As an example, Fig. 6.45 shows the parameter lines of a rational canal surface with cubic spine curve and cubic radius function. One family of parameter lines is given by the circles contained in the surface (the spheres, which the surface is the envelope of, touch the surface there).

Envelopes of spheres of constant radius (*pipe surfaces*) are offset surfaces of their spine curve $\mathbf{m}(t)$ and if $\mathbf{m}(t)$ is rational, so is the pipe surface. This is in contrast to the behaviour of planar curves, where offsets of rational curves need not be rational.

Many of the previous results are consequences of the fact that the envelope of a rational family of quadratic cones is rational. Explicit parametrizations of these surfaces are available [135, 138].

6.4 Connecting Developables and Applications

6.4.1 Basics

The problem of finding a developable surface \mathcal{R} which contains two given curves or touches two given surfaces has many interesting applications. It is illustrated by Fig. 6.46, which shows selected generators of such a surface. We are going to describe solutions to the problem and some of the applications in greater detail.

Dualization of the Problem

A developable surface is completely determined by the family of its tangent planes. It is therefore natural to dualize the problem of finding a developable surface \mathcal{R} which touches two surfaces Φ_1, Φ_2 . We consider the set of tangent planes \mathcal{R}^* of \mathcal{R} , and Φ_1^*, Φ_2^* of Φ_1 and Φ_2 , respectively. Clearly, \mathcal{R}^* must be a subset of $\Phi_1^* \cap \Phi_2^*$.

If Φ_i is itself developable, then its dual is a *curve* of dual space. If not, its dual is a surface. It is unlikely that two space curves intersect, and a space curve and a surface intersect typically in a discrete number of points. It therefore makes no sense to ask for a developable which joins other surfaces, if they are themselves developable. We will disregard this case completely.

If neither Φ_1^* nor Φ_2^* is developable, then both duals are *surfaces* of dual space, and their intersection is likely to be a curve, i.e., the dual of a developable surface \mathcal{R} . The problem of finding \mathcal{R} is therefore equivalent to (and as complicated as) the problem of intersection of surfaces.

If we want to state conditions which ensure that \mathcal{R} is actually a regular developable surface this is equivalent to stating conditions which ensure that the intersection of two surfaces is actually a *regular curve*.

There is a large amount of literature on surface/surface intersection, whose discussion is beyond the scope of this book. However, it is worth to point out a few important geometric properties of joining developable surfaces, although in most cases we ‘just have to apply duality’.

Regularity of the Connecting Developable

By dualizing the well known situation, we get a condition for regularity of the connecting developable of two surfaces. First the original statement:

If a *curve* arises as *intersection* of two surfaces, then its *tangent* in a regular *point* is the *intersection* of the corresponding *tangent planes*. Furthermore, a common *point* of two surfaces is a regular point of the intersection, if these two surfaces

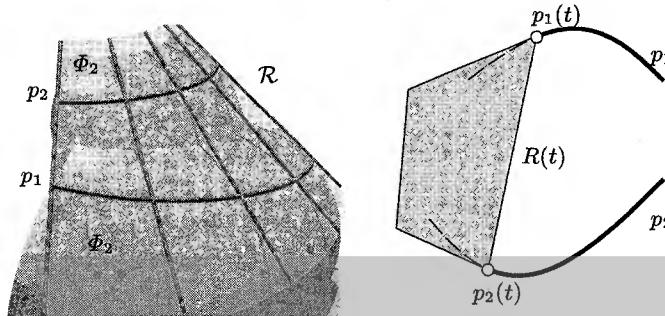


Fig. 6.46. Left: Connecting developable of two surfaces Φ_1 and Φ_2 . Right: Ruling and tangent plane of the connecting developable of two curves.

intersect transversely, i.e., they are themselves regular and their *tangent planes* do not coincide.

Now the dual one: If \mathcal{R} is the *connecting developable* of two surfaces, then the *ruling* in a regular *tangent plane* is the *span* of the corresponding *points of contact*. Furthermore, a common *tangent plane* of two surfaces is a regular plane of the connecting developable, if the duals of these two surfaces intersect transversely, i.e., they are themselves regular and their *points of contact* do not coincide.

The union of the contact points of \mathcal{R} with the surfaces Φ_i are two *contact curves* $p_i \in \Phi_i$, and obviously \mathcal{R} touches Φ_i in the points of p_i (see Fig. 6.46).

We can think of generating \mathcal{R} also in the following way: a plane is rolling on both surfaces Φ_1 and Φ_2 simultaneously. For all instants t , there are two contact points $p_1(t), p_2(t)$ in Φ_1, Φ_2 , respectively. The envelope generated by the plane during this motion is the connecting developable, and the rulings $R(t)$ join the contact points $p_1(t), p_2(t)$. If we consider a point p in the rolling plane, which at time $t = t_0$ becomes a contact point, then this point has zero velocity at $t = t_0$. This shows that the infinitesimal motion of the rolling plane is an infinitesimal rotation (there are two points with zero velocity), and the instantaneous axis at time t equals the ruling $R(t)$.

The Connecting Developable of Curves

The set of tangent planes of a space curve is the set of planes which contain one of its tangents. This set of planes is a surface in dual projective space. Indeed, the dual of a developable surface is a curve, and so the dual of a curve is a developable surface.

From the viewpoint of connecting developables, curves therefore do not play a special role. Computing the connecting developable however is easier if both Φ_1, Φ_2 are curves, and it is easiest if one of them is planar. Fig. 6.46 illustrates the situation. Of course we can identify Φ_i with the curve of contact p_i . Tangent planes of \mathcal{R} are those which touch both curves p_1, p_2 , and the ruling in a plane equals the span of the two contact points.

In Fig. 6.46, the curves p_1 and p_2 are parametrized such that the tangents in $p_1(t)$ and $p_2(t)$ are concurrent, i.e., the ruling $R(t)$ of the connecting developable equals the line $p_1(t) \vee p_2(t)$. In affine coordinates, where $p_i(t) = \mathbf{p}_i(t)$, or in projective coordinates, where $p_i(t) = \mathbf{p}_i(t)\mathbb{R}$, this means that

$$\det(\mathbf{p}_1 - \mathbf{p}_2, \dot{\mathbf{p}}_1, \dot{\mathbf{p}}_2) = 0 \iff \det(\mathbf{p}_1, \mathbf{p}_2, \dot{\mathbf{p}}_1, \dot{\mathbf{p}}_2) = 0.$$

Connecting Developables of Algebraic Curves and Surfaces

Consider two algebraic surfaces Φ_1, Φ_2 of complex projective space, and assume that they are of class m_1 and m_2 , respectively. This means that their duals Φ_1^*, Φ_2^* are algebraic surfaces of degree m_1, m_2 . We assume that $\mathcal{R}^* = \Phi_1^* \cap \Phi_2^*$ is a curve, i.e., there exists a connecting developable \mathcal{R} .

Lemma 6.4.1. *If \mathcal{R} is the connecting developable of two non-developable algebraic surfaces of class m_1, m_2 , in complex projective space, then its class is less or equal $m_1 m_2$.*

Proof. The class of \mathcal{R} , i.e., the degree of \mathcal{R}^* is the generic number of points in $\mathcal{R}^* \cap U^*$, where U^* is a test plane. As U^* intersects Φ_i^* in a planar curve of degree m_i , Th. 1.3.23 shows that $\deg(\mathcal{R}^*) \leq m_1 m_2$. \square

We have to clarify the meaning of m_i if Φ_i is a curve, not a surface. The number m_i is the degree of the dual Φ_i^* , and can be computed as the generic number of intersection points of Φ_i^* with a test line. If we look at this situation from the viewpoint of the original complex projective three-space, we see that m_i is the generic number of planes tangent to Φ_i , and incident with a test line L .

Remark 6.4.1. In the case that Φ_i is contained in a plane ε , the number m_i obviously equals the number of tangents incident with $L \cap \varepsilon$, so m_i equals the class of Φ_i as a planar algebraic curve.

In the case that Φ_i is non-planar, then (i) it is easy to see that m_i equals the generic class of cones with base curve Φ_i , and (ii) it is possible to show that m_i equals the class of Φ_i 's tangent surface (the rank of Φ_i), so m_i equals the generic number of osculating planes of Φ_i incident with a point of space. \diamond

Example 6.4.1. Consider the connecting developable \mathcal{R} of two *quadrics* Φ_1, Φ_2 . If both Φ_1, Φ_2 are regular, their duals are quadrics again, and \mathcal{R}^* is the intersection of two quadrics. This means that \mathcal{R} is, in general, of class four, but may be of class two — like the intersection curve of two regular quadrics.

There is a curve p_i of contact between \mathcal{R} and the quadric Φ_i . If we apply the polarity π_i with respect to Φ_i to the planes of \mathcal{R} , we get the set of contact points, which shows that p_i is, in general, of degree four.

Assume that Φ_i has, in homogeneous coordinates, the equation $\mathbf{x}^T \cdot A_i \cdot \mathbf{x} = 0$. Then $\pi_i : (\mathbf{x}\mathbb{R}) \mapsto \mathbb{R}(A_i \cdot \mathbf{u})$ is the polarity with respect to Φ_i . The dual Φ_i^* has, by Lemma 1.1.21, the equation $\mathbf{u}^T \cdot A_i^{-T} \cdot \mathbf{u} = 0$.

The curve $p_1 = \pi_1(\Phi_1^* \cap \Phi_2^*)$ can be computed more easily as $\pi_1(\Phi_1^*) \cap \pi_1(\Phi_2^*) = \Phi_1 \cap \pi_1(\Phi_2^*)$, where $\pi_1(\Phi_2^*)$ has the equation $\mathbf{x}^T \cdot A_1^T A_2 A_1 \cdot \mathbf{x} = 0$. An analogous formula holds for p_2 .

If Φ_i is a conic, then its dual is a quadratic cone. A special case has already been discussed in this book: If Φ_1 is the ideal curve of a cone of constant slope, then the connecting developable is of constant slope, and the curve p_1 is an isophote of Φ_2 . It can be computed as $p_2 = \Phi_2 \cap \pi_2(\Phi_1)$ (see Ex. 6.3.14). \diamond

If Φ_1 and Φ_2 are rational, their duals are rational surfaces, whose intersection however need not be rational. So we cannot expect rationality of the connecting developable of rational curves and surfaces.

Curvature of Connecting Developables

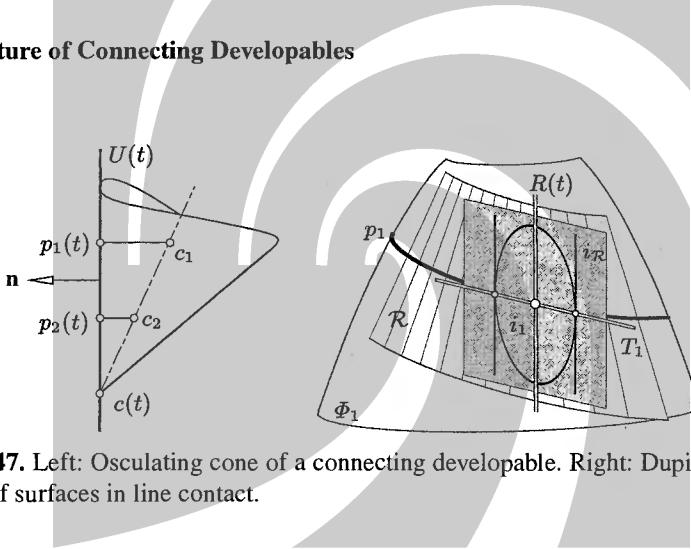


Fig. 6.47. Left: Osculating cone of a connecting developable. Right: Dupin indicatrices of surfaces in line contact.

For various purposes (e.g., for approximation) some knowledge of the curvature of a connecting developable \mathcal{R} is useful. By Th. 6.1.4, there are osculating cones of revolution, each of which is in second order contact with \mathcal{R} in the regular points of a generator line $R(t)$.

We first consider the case of a surface \mathcal{R} connecting two curves $p_1 = \Phi_1$, $p_2 = \Phi_2$. The connecting developable has parabolic surface points, and the rulings are the principal tangents of zero curvature. Assume that

- the positive real or infinite number γ_i is the radius of the osculating circle k_i of p_i at $p_i(t)$,
- the angle α_i is enclosed by the plane of k_i (p_i 's osculating plane at $p_i(t)$) and the tangent plane $U(t)$ of \mathcal{R} in the points of $R(t)$,
- the angle β_i is enclosed by the tangent T_i of p_i at $p_i(t)$ and $R(t)$.

Then the nonzero principal curvatures κ_1, κ_2 in the points $p_1(t), p_2(t)$ are expressed by

$$\frac{1}{\kappa_i} = \rho_i = \gamma_i \frac{\sin^2 \beta_i}{|\sin \alpha_i|}. \quad (6.74)$$

This follows from Euler's formula and Meusnier's theorem. If $\mathbf{n}(t)$ is a unit normal vector of $U(t)$, then we adjust the sign of ρ_i such that the vector $\rho_i \mathbf{n}$ points to the half-space bounded by $U(t)$ which contains the osculating circle k_i . Then the principal curvature center $c_i(t)$ which corresponds to the nonzero principal curvature at $p_i(t)$ is expressed in a Cartesian coordinate system as

$$\mathbf{c}_i = \mathbf{p}_i(t) + \rho_i \mathbf{n}(t). \quad (6.75)$$

This information is sufficient to determine axis and vertex of the osculating cone C of revolution. It must share the principal curvatures with \mathcal{R} , and its principal curvature centers lie in its axis. So the line $c_1(t) \vee c_2(t)$ is the axis of the osculating cone. Its vertex is found by intersecting the axis with the ruling $R(t)$ (see Fig. 6.46). The vertex coincides with $R(t)$'s point of regression.

If Φ_i is a *surface* instead of a curve, we can use Th. 6.1.6: The tangent T_i of the curve p_i of contact is conjugate to the ruling $R(t)$. We compute the normal curvature $1/\gamma_i$ for this conjugate direction, and enter the formula above with $\alpha_i = \pi/2$.

The geometric situation is visualized easily with help of the Dupin indicatrices of Φ_i and of \mathcal{R} in a contact point (see Fig. 6.47). The Dupin indicatrix $i_{\mathcal{R}}$ of \mathcal{R} is a pair of lines parallel to $R(t)$, which touches the indicatrix i_1 of Φ_1 in two diametral points. The tangent T_1 of p_1 is i_1 -conjugate to $R(t)$. This is a special case of surfaces in line contact, which will be discussed again at p. 457f.

6.4.2 Convex Hulls and Binder Surfaces

The connecting developable of a space curve with itself (a *self-connecting developable*) has interesting applications. We first consider the problem of determining the convex hull of a smooth space curve and then show an application in manufacturing.

Convex Hulls of Curves and Surfaces

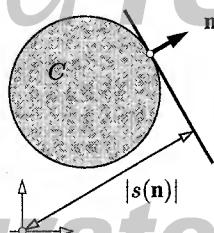


Fig. 6.48. Support plane and support function of a convex body.

The problem of finding the *convex hull* of a given set of curves and surfaces in Euclidean three-space is a very important one in computational geometry [40, 97]. It turns out that the boundary of this convex hull contains parts of the connecting developables of these curves and surfaces. We first describe some basic notions of convex geometry:

For all boundary points p of a convex body C in \mathbb{R}^3 there is a possibly not unique *support plane* $U(p)$. This plane contains p , and C is contained in one of the two closed half-spaces bounded by U . If C 's boundary is smooth, then the support plane $U(p)$ coincides with the tangent plane in p . For a non-smooth boundary point, like the vertex of a convex polyhedron, there is an infinite number of support planes.

If \mathbf{n} is a unit vector, there is a unique support plane of C which has the equation

$$\mathbf{n} \cdot \mathbf{x} = s(\mathbf{n}), \quad (6.76)$$

such that C is contained in the half-space $\mathbf{n} \cdot \mathbf{x} \leq s(\mathbf{n})$, i.e., \mathbf{n} points to the outside of C , and $|s(\mathbf{n})|$ is the distance between the origin and the support plane (see Fig. 6.48). The value $s(\mathbf{n})$, dependent on \mathbf{n} , is called the *support function* $s(\mathbf{n})$ of K . The domain of the support function is the unit sphere, it takes values in \mathbb{R} . The boundary of C is the envelope of its support planes. Thus the representation of the boundary of C by its support function is a dual representation.

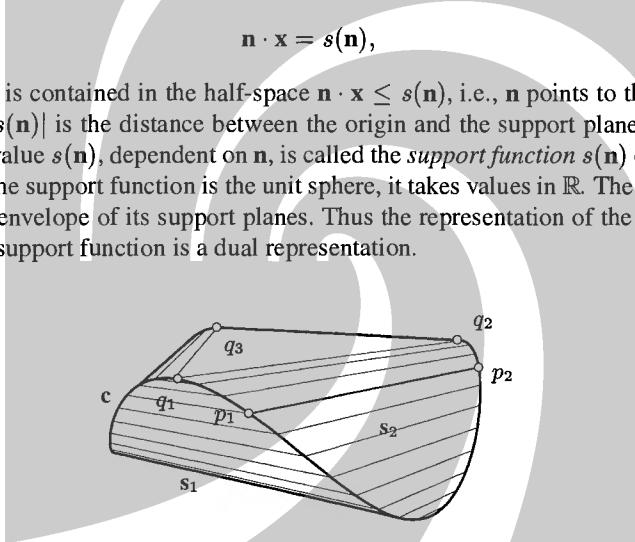


Fig. 6.49. Convex hull of a closed space curve c .

Because the intersection of convex sets is again convex, for all subsets S of \mathbb{R}^3 there is a smallest convex subset of \mathbb{R}^3 which contains S . This *convex hull* C of S equals the intersection of all convex sets which contain S . If S is closed, C equals also the intersection of all closed half-spaces which contain S . If a half-space contains C and its boundary plane intersects the boundary S , then this boundary plane is called a *support plane* of S . Thus the boundary of the convex hull is contained in the envelope of the support planes of S .

If a support plane has *two* contact points p_1, p_2 , then the line segment $\overline{p_1 p_2}$ is contained in C (see Fig. 6.49). So if there is a one-parameter family of support planes each of which has *two* contact points, the boundary of C contains a developable surface, namely the connecting developable of the two curves of contact points.

If there exists a support plane U_0 with *three or more* non-collinear contact points q_1, \dots, q_n , then the planar convex hull of q_1, \dots, q_n is part of C 's boundary (see Fig. 6.49).

Example 6.4.2. Fig. 6.49 shows the convex hull C of a closed space curve S . The boundary of C contains planar faces (such as the convex hull of q_1, q_2, q_3) arising from triple tangent planes, as well as developable surfaces.

The computation of such a self-connecting developable is dual to the computation of the self-intersection of a parametric surface. \diamond

Binder Surface Computation

In a stamping die, a ring-shaped *binder surface* or *blankholder* is needed to fix the metal sheet during the metal forming process. The binder surface together with the punch determines the position of the sheet metal before the punch closes against the die. It is desirable that this position is obtained without deformation, i.e., that the binder surface is developable. One strategy for the construction of the blankholder is by starting with its inner boundary, the *punch line*.

Since eventually the punch line p is a curve contained in the developable binder surface Φ , we can construct the latter as a self-connecting developable of p . The major difficulty here lies in the requirement that Φ has to be regular inside p and also within some distance outside (see Fig. 6.50, left). Since there is only a finite number of self-developables of p , the regularity condition will sometimes force us to weaken the interpolation condition and construct a developable surface which only approximates the punch line.

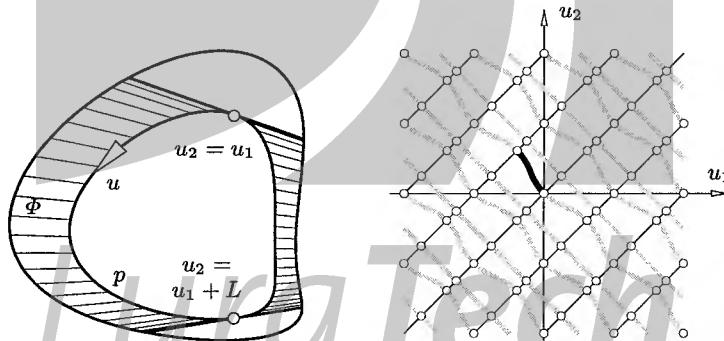


Fig. 6.50. Left: Binder surface Φ computed from a closed punch curve p . Right: Solution set of (6.77). The fat curve corresponds to the surface Φ .

In view of this application, we study the self-connecting developable of a space curve p in more detail. Such a surface is the envelope of planes $U(t)$ which touch the curve in two points $p(u_1)$ and $p(u_2)$. This means that the curve tangents are contained in a plane, and in affine coordinates with $p(t) = \mathbf{p}(t)$ this is expressed as

$$\det(\mathbf{p}(u_1) - \mathbf{p}(u_2), \dot{\mathbf{p}}(u_1), \dot{\mathbf{p}}(u_2)) = 0. \quad (6.77)$$

Finding all (u_1, u_2) such that (6.77) is fulfilled is equivalent to finding the zero set of the left hand side of this equation, which is a curve u in the u_1, u_2 -plane.

If $\mathbf{p}(t)$ is rational, then obviously u is an algebraic curve. Efficient computation is possible by using a B-spline representation of the function $\det(\mathbf{p}(u_1) - \mathbf{p}(u_2), \dot{\mathbf{p}}(u_1), \dot{\mathbf{p}}(u_2))$, and exploiting the subdivision and convex hull properties of the B-spline scheme (see Elber [42] for more details).

Clearly, the line $u_1 = u_2$ is a component of the solution of (6.77), but it does not correspond to a self-connecting developable of the curve p . If the curve p is parametrized periodically with $p(t) = p(t \pm L) = p(t \pm 2L) = \dots$, where L is the period of p , then all lines $u_1 = u_2 + kL$ with an integer k are contained in the solution set of (6.77). This set may look like the one depicted in Fig. 6.50, right.

It may happen that there is a solution branch $(u_1(t), u_2(t))$ of (6.77) where never $u_1(t) = u_2(t)$. We can imagine that $u_1(t)$ and $u_2(t)$ are running in circles without meeting each other. The resulting developable surface however is useless for practical purposes, as its curve of regression would be inside p .

We will therefore investigate solution branches of (6.77) with a limit point on the line $u_1 = u_2$, such as depicted in Fig. 6.50, left.

Suppose that $u_1 \rightarrow u_e$ and $u_2 \rightarrow u_e$. If we insert $\mathbf{p}(u) = \mathbf{p}_0 + u\dot{\mathbf{p}} + \frac{u^2}{2}\ddot{\mathbf{p}} + \frac{u^3}{6}\mathbf{p}^{(3)} + o((u - u_e)^3)$ for $u \rightarrow u_e$ into (6.77), we get

$$\det(\dot{\mathbf{p}} + o, \ddot{\mathbf{p}} + o, \mathbf{p}^{(3)} + o) = 0, \text{ with } o = o(u_1 - u_e, u_2 - u_e) \text{ for } u_1, u_2 \rightarrow u_e.$$

In the limit $u_1, u_2 \rightarrow u_e$ we have

$$\det(\dot{\mathbf{p}}(u_e), \ddot{\mathbf{p}}(u_e), \mathbf{p}^{(3)}(u_e)) = 0. \quad (6.78)$$

If $p(u_e)$ is no inflection point, then obviously the sequence of dimensions of osculating subspaces starts with $0, 1, 2, 2$. We see that $p(u_e)$ is a handle point of the curve p (cf. Th. I.2.2 and the definitions following this theorem). A Euclidean characterization of handle points is nonvanishing curvature, but vanishing torsion.

This suggests the following procedure for finding self-connecting developables without singularities: Starting from one handle point of p we follow a solution branch of (6.77), which eventually terminates in a singularity or another handle point. If we are even more lucky, the curve of regression of the self-connecting developable does not enter a pre-defined region of interest which contains p and its interior.

The Curve of Regression of a Self-Connecting Developable

For applications it is important to know the whereabouts of the curve of regression of a self-connecting developable \mathcal{R} . We investigate the situation where the curve of regression meets the punch curve p . Assume that $p(u_1)$ is the point of regression of the ruling $\mathcal{R} = p(u_1) \vee p(u_2)$. Then $p(u_1)$ must be the vertex of \mathcal{R} 's osculating cone.

If we look back at the construction of the osculating cone at p. 400, we see that this can happen only if (6.74) gives an infinite value of $1/\kappa_1$. This happens either

if the curvature radius γ_1 of p at $p(u_1)$ is infinite, or if $\alpha_1 = 0$, which means that \mathcal{R} 's tangent plane in the points of R coincides with p 's osculating plane at $p(u_1)$. On the other hand, if $\kappa_1 = 0$, $p(u_1)$ is the vertex of the osculating cone unless also $\kappa_2 = 0$, in which case the question whether $p(u_1)$ is the vertex could be decided by investigation of higher derivatives.

This shows that many closed curves will produce developable surfaces which are not useful for the use as a binder surface. If such a curve arises as the punch curve in an actual application, it is not possible to achieve the final form of the metal sheet referred to above without distortions. In this case the final position of the metal should be a surface as close as possible to a developable surface, and appropriate approximation techniques have to be employed (see Sec. 6.2.3).

The condition that never $\alpha_i = 0$ in Equ. (6.74) has a simple geometric interpretation: The tangent plane of the self-connecting developable \mathcal{R} is rolling on the punch curve p . We imagine that t is the time, and that $p(u_1(t))$, $p(u_2(t))$ are the points of tangency which determine the rulings of \mathcal{R} . If the curve is always contained in the same half-space bounded by the rolling plane, then $\alpha_i \neq 0$. A change of sides at $p(u_i(t))$ occurs precisely for $\alpha_i = 0$. In this case the rolling plane becomes the curves' osculating plane, and usually $u_i(t)$ is no longer monotonous, i.e., the one of the two points of tangency begins to run back.

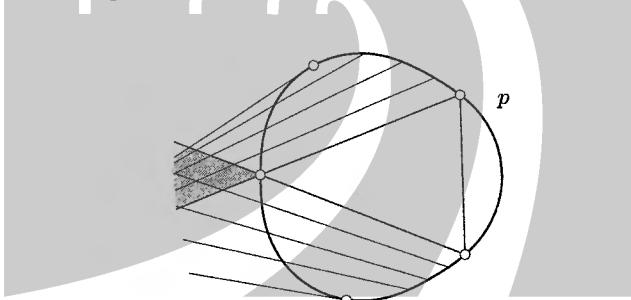


Fig. 6.51. Possible parts of a future binder surface overlap if the punch curve has a triple tangent plane.

Remark 6.4.2. The condition involving the half-space is obviously met if the curve p is contained in the boundary of its convex hull. Such a curve is called a *convex space curve*. But being a convex space curve is not sufficient for the existence of a self-connecting developable within the region of interest. An example is shown in Fig. 6.51: A triple tangent plane of p implies that p 's convex hull contains a triangle, and there is no regular extension of the self-connecting developable to the outside of p . \diamond

Remark 6.4.3. For convex space curves, V.D. Sedykh [180] has given the following generalization of the four vertex theorem: Every closed simple convex C^3 space curve p with nonvanishing curvature has at least four handle points. The proof stud-

ies the boundary of p 's convex hull. Generalizations of this result to curves with inflection points have been derived by M.C. Romero Fuster and V.D. Sedykh [169]. \diamond

6.4.3 Geometric Tolerancing

There is a fundamental question pertinent in all geometric problems arising in Computer Aided Design: How does the solution of the problem depend on the input data? A closer look at some special cases reveals a possible reason for the lack of systematic research in this area: even for simple geometric constructions the analysis of tolerances and their interdependencies is quite complex.

Requicha [167] introduced a model for geometric tolerancing based on *tolerance zones*. Tolerance zones are also a very intuitive concept useful for visualization of error propagation in geometric constructions.

Robust geometric computing using tolerance zones and *interval arithmetic* has recently received much attention (see for instance [1, 82, 83, 183]). Interval arithmetic is attractive from the computational viewpoint but is often not geometric in the sense that it does not give precise tolerance zones for the output.

Important contributions to geometric tolerancing for planar geometric constructions come from C.U. Hinze [71], who analyzed ruler and compass constructions which occur frequently. His tolerance zones for points are convex and bounded by segments of straight lines and circles.

Here we will briefly outline some relations between convex tolerance zones for affine geometric constructions and developable surfaces (see [153, 201]). It turns out that many free-form curve and surface design schemes can be effectively analyzed using ideas from convex geometry (for an overview of this large subject, see [66]).

Inner and Outer Connecting Developables of Convex Sets

Maybe the simplest and most fundamental geometric construction is a line which is spanned by two points. Let us study this construction in \mathbb{R}^3 . Assume that points p_0, p_1 , are contained in two convex tolerance zones K_0 and K_1 . We want to determine the tolerance zone of the line $L = p_0 \vee p_1$, i.e., the part of \mathbb{R}^3 which is covered by all possible lines spanned by a point of K_0 and a point of K_1 .

If K_0 and K_1 intersect (p_0 and p_1 are indistinguishable within the imposed tolerances), the tolerance zone of $p_0 \vee p_1$ is entire \mathbb{R}^3 , and from now on we disregard this case completely.

In addition, we further simplify the problem by considering only the straight line segment $\overline{p_0 p_1}$ spanned by p_0 and p_1 . The tolerance zone of line segments is easily found:

Lemma 6.4.2. *Assume that K_0, K_1 are convex subsets of \mathbb{R}^3 . Then the union of line segments $\overline{p_0 p_1}$ with $p_j \in K_j$ equals the convex hull K of the union $K_0 \cup K_1$.*

If K_0 and K_1 are compact and disjoint, the boundary of K contains the interior part of the outer connecting developable of K_0, K_1 .

The meaning of ‘interior’ and ‘outer’ is explained in the proof.

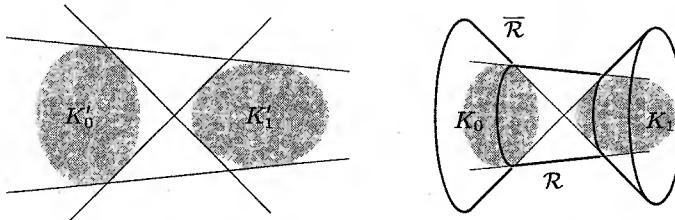


Fig. 6.52. Left: Inner and outer common tangents of disjoint compact convex sets in \mathbb{R}^2 . Right: Inner and outer connecting developable of disjoint compact convex sets in \mathbb{R}^3 . Beware of the apparent contour which makes the connecting developable look like a cone.

Proof. Clearly, the union of all line segments is contained in K . Conversely, the union of these segments is convex, because a convex combination of $\mathbf{a} = ux_0 + (1 - u)x_1$ and $\mathbf{b} = vy_0 + (1 - v)y_1$ is of the form $wz_0 + (1 - w)z_1$, with $w \in [0, 1]$ (the details are left to the reader).

By the Hahn-Banach theorem, there is a plane ε separating K_0 and K_1 . Choose a circle of unit vectors $\mathbf{n}(t)$ parallel to ε ($t \in [0, 2\pi]$). All orthogonal projections parallel to $\mathbf{n}(t)$ map K_0, K_1 to compact disjoint convex subsets K'_0, K'_1 of the plane. K is mapped to the convex hull K' of $K'_0 \cap K'_1$. There are exactly four common tangents of K'_0 and K'_1 (see Fig. 6.52), two of which contribute to the boundary of K' (the ‘outer’ ones). They are the projections of support planes $U(t)$ and $\bar{U}(t)$ of K_0, K_1 . Clearly the vector $-\mathbf{n}(t)$ results in the same support planes as $\mathbf{n}(t)$.

Thus we have shown that there is a family $U(t)$ of support planes of K_0, K_1 , each of which has a straight line segment in common with the boundary of K . The carrier lines of these segments are the rulings of the ‘outer’ connecting developable of K_0, K_1 . \square

The computation of the outer connecting developable \mathcal{R} of K_0 and K_1 is easily performed if one uses the support functions $s_0(\mathbf{n})$ and $s_1(\mathbf{n})$ of K_0 and K_1 , respectively. A common support plane with outward unit normal \mathbf{n} belongs to a zero of the difference function

$$d(\mathbf{n}) = s_0(\mathbf{n}) - s_1(\mathbf{n}). \quad (6.79)$$

Therefore the computation of the family of tangent planes \mathcal{R} amounts to the computation of the zero set of the function $d(\mathbf{n})$ defined on the unit sphere.

In order to obtain the tolerance zone for the entire straight line $p_0 \vee p_1$ with $p_j \in K_j$ we have to compute a second developable surface:

Lemma 6.4.3. *We use the notation of Lemma 6.4.2 and assume that K_0 and K_1 are compact and disjoint. The boundary of the union \bar{K} of lines $p_0 \vee p_1$ with $p_j \in K_j$ is contained in the inner and outer connecting developables of K_0 and K_1 .*

Proof. We continue the proof of Lemma 6.4.2. For all unit vectors $\mathbf{n}(t)$, the set \bar{K} projects to the polygonal domain \bar{K}' whose boundary are the four common tangents of K'_0 and K'_1 . This shows the statement of the lemma. \square

The inner connecting developable $\bar{\mathcal{R}}$ of K_0 and K_1 is the envelope of planes which are support planes $U(t)$ of both K_0 and K_1 . K_0 and K_1 are contained in the two different half-spaces bounded by $U(t)$. The normal vectors \mathbf{n} of such planes are zeros of the difference function

$$\bar{d}(\mathbf{n}) = s_0(\mathbf{n}) + s_1(-\mathbf{n}). \quad (6.80)$$

The surface $\bar{\mathcal{R}}$ has the property that the point of regression of a ruling is always located between the contact points $p_i \in K_i$ ($i = 0, 1$). This is easily seen with the osculating cones, if we assume that K_i is C^2 and strictly convex. Fig. 6.52 shows an example.

Remark 6.4.4. We did not discuss the smoothness and regularity of the connecting developable of two convex bodies K_0, K_1 . If K_i are convex and disjoint, a family of common tangent planes always exists, as has been shown above. It is not difficult to show that if K_i are C^r and disjoint, then both sheets $\mathcal{R}, \bar{\mathcal{R}}$ of the connecting developables and their curves of contact are C^{r-1} . \diamond

Convex Combination of Convex Sets

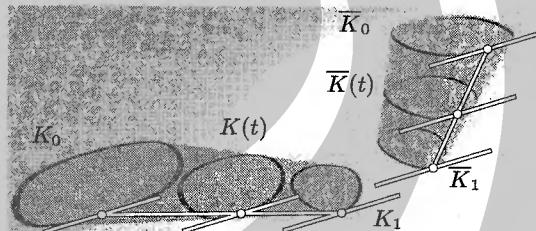


Fig. 6.53. Connecting developable of two convex curves in parallel planes and convex combination of convex domains in the plane.

Let us now consider the case of planar compact convex sets K_0 and K_1 . Further consider two planar compact convex sets \bar{K}_0, \bar{K}_1 , contained in horizontal planes $\varepsilon_0 : z = 0, \varepsilon_1 : z = 1$. We assume that there is a parallel projection which maps \bar{K}_i onto K_i (cf. Fig. 6.53).

For all support lines L of K_0 there are precisely two support lines of K_1 parallel to L , and vice versa. This shows that the connecting developable of \bar{K}_0 and \bar{K}_1 consists of an ‘inner’ sheet $\bar{\mathcal{R}}$ and an ‘outer’ sheet \mathcal{R} . Fig. 6.53 shows only the ‘outer’ connecting developable \mathcal{R} .

The part of \mathcal{R} which is between ε_0 and ε_1 is part of the boundary of the convex hull of \bar{K}_0 and \bar{K}_1 .

If we intersect this convex hull with the plane $\varepsilon : z = t$ ($0 \leq t \leq 1$), we get a convex set \bar{K} , whose boundary is the curve $\mathcal{R} \cap \varepsilon$. Obviously $\bar{K}(t) = (1-t)\bar{K}_0 + t\bar{K}_1$, and \bar{K} projects to the set

$$K(t) = (1 - t)K_0 + tK_1 \quad (6.81)$$

(for the notation, see the definition of Minkowski sum on p. 374). If K_0, K_1 are tolerance zones of the points $\mathbf{p}_0, \mathbf{p}_1$, respectively, then $K(t)$ is the tolerance zone of the point $t\mathbf{p}_0 + (1 - t)\mathbf{p}_1$, as shown by Fig. 6.53.

As the coefficients in Equ. (6.81) sum up to 1, this linear combination is an *affine combination* of the sets K_0 and K_1 . In addition, all coefficients are nonnegative, so it is actually a *convex combination*.

Lemma 6.4.4. Assume that $K(t), K_0, K_1$ are convex subsets of \mathbb{R}^n , and s_t, s_0, s_1 are their support functions. If $K = (1 - t)K_1 + tK_2$ ($0 \leq t \leq 1$), then

$$s_t(\mathbf{n}) = (1 - t)s_0(\mathbf{n}) + ts_1(\mathbf{n}). \quad (6.82)$$

Proof. We show this only for the case $n = 2$, because in that case we use developable surfaces.

We find a parallel projection and sets \bar{K}_0, \bar{K}_1 in planes $\varepsilon_0, \varepsilon_1$ which project onto K_0, K_1 , respectively, as depicted in Fig. 6.53. Consider the connecting developable \mathcal{R} of \bar{K}_0, \bar{K}_1 , whose tangent plane is the same in all points of a ruling. The rulings connect points $\mathbf{p}_0, \mathbf{p}_1$ whose support lines are parallel, and obviously also the support line incident with $\bar{\mathbf{p}}(t) = (1 - t)\bar{\mathbf{p}}_0 + t\bar{\mathbf{p}}_1$ is parallel to the other two.

This fact is not destroyed by parallel projection, and we have shown the following: If $\mathbf{p}_0, \mathbf{p}_1$ are boundary points of K_0, K_1 , respectively, with parallel support lines, then the point $\mathbf{p}(t) = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ is a boundary point of $K(t)$ and has a support line parallel to the other two. Now the oriented distances $s_0(\mathbf{n}), s_1(\mathbf{n})$, and $s(\mathbf{n})$ of the origin to these three support lines must fulfill the equation $s(\mathbf{n}) = (1 - t)s_0(\mathbf{n}) + ts_1(\mathbf{n})$. \square

The proof of Lemma 6.4.4 shows the following: In order to find the boundary of $K(t) = (1 - t)K_0 + tK_1$, search for boundary points $\mathbf{p}_0, \mathbf{p}_1$ of K_0 and K_1 which have parallel support lines. Then $\mathbf{p}(t) = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ is a boundary point of $K(t)$ (see also Fig. 6.54).

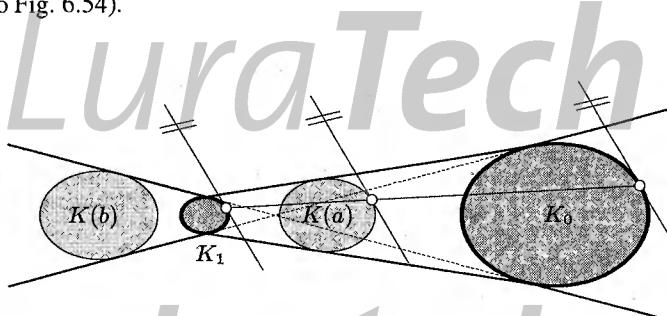


Fig. 6.54. Affine combinations $K(t) = (1 - t)K_0 + tK_1$ of convex domains K_0, K_1 for $t = a = 0.35$ and $t = b = -0.5$.

Tolerance Zones of Affine Combinations

Lemma 6.4.4 shows that the support function of a convex combination of convex bodies is just the same convex combination of their support functions. It is easy to generalize this to more general affine and linear combinations:

Lemma 6.4.5. *Assume that $K = \sum \lambda_i K_i$ with K_i convex. If s_i is the support functions of K_i and s is the support function of K , then*

$$s(\mathbf{n}) = \sum_{\lambda_i > 0} s_i(\mathbf{n}) - \sum_{\lambda_i < 0} s_i(-\mathbf{n}). \quad (6.83)$$

Proof. First we note that the support functions of the sets $\lambda_i K_i$ for $\lambda_i > 0$ and $-K_i$ are given by $\mathbf{n} \mapsto \lambda_i s_i(\mathbf{n})$ and $\mathbf{n} \mapsto -s_i(-\mathbf{n})$. This is clear from the geometric definition. The latter formula reduces the proof to the case $\lambda_i > 0$. Second, it is obviously sufficient to show the lemma for the case of two summands. We can write $\lambda_0 K_0 + \lambda_1 K_1$ as a multiple of a convex combination:

$$\lambda_0 K_0 + \lambda_1 K_1 = (\lambda_0 + \lambda_1) \left(\frac{\lambda_0}{\lambda_0 + \lambda_1} K_0 + \frac{\lambda_1}{\lambda_0 + \lambda_1} K_1 \right).$$

Now we can use the first formula and Lemma 6.4.4 to show the result. \square

We see that Minkowski sums and support functions are a basic tool for studying tolerance analysis of the linear combination of points, if the tolerance zones of the single points are convex. For computational issues of Minkowski sums in the plane, we refer the reader to Kaul and Farouki [93] and Lee et al. [107]. In the latter paper, Minkowski sums (convolutions) are studied for arbitrary planar domains. This is more subtle than Minkowski sums of convex bodies, which is one reason for our limitation to convex tolerance zones for points.

Tolerance Zones for Bézier and B-Spline Curves

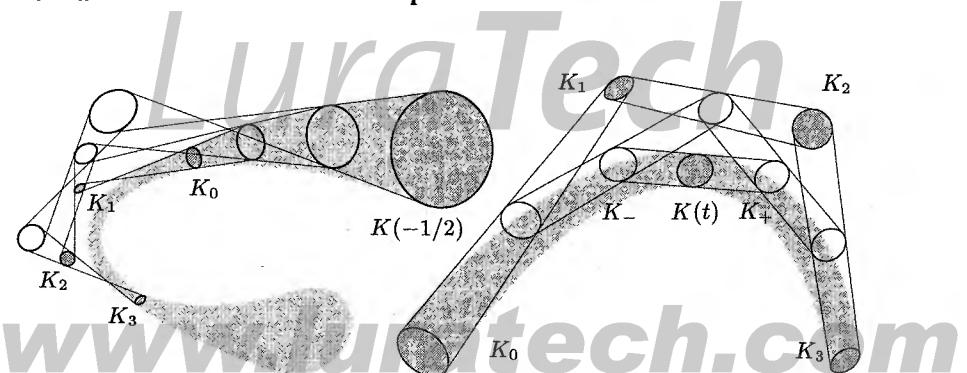


Fig. 6.55. Tolerance zones of a Bézier curve $\mathbf{b}(t)$. Left: $-1/2 \leq t \leq 3/2$. Right: $0 \leq t \leq 1$.

A large amount of curve and surface schemes in Computer Aided Geometric Design consists, in principle, of appropriate linear (over even affine or convex) combinations of control points \mathbf{b}_i . Thus tolerance analysis of many freeform curves and surfaces may be based on Lemma 6.4.5 (cf. [153, 201]).

Example 6.4.3. Consider a *Bézier curve* $\mathbf{b}(t)$ as defined by Equ. (1.89) and assume that the control point \mathbf{b}_i is contained in a convex body K_i — the tolerance zone of \mathbf{b}_i equals K_i . Then the tolerance zone of the curve point $\mathbf{b}(t)$ is given by

$$K(t) = \sum_{i=0}^n B_i^n(t) K_i.$$

Fig. 6.55, left shows the union of all $K(t)$ as t ranges in the interval $[0, 1]$. There you can also see how to compute points and tangents of the boundary of the tolerance zone of $\mathbf{b}(t)$: The points \mathbf{b}_0^{n-1} and \mathbf{b}_1^{n-1} of the algorithm of de Casteljau (cf. the definition immediately before Lemma 1.4.2) are linear combinations of the points $\mathbf{b}_0, \dots, \mathbf{b}_i$ and have appropriate tolerance zones $K_-(t), K_+(t)$. It is easy to show that the common support lines/planes of $K_-(t)$ and $K_+(t)$ are support lines/planes of both $K(t)$ and the union of all $K(t)$ as t ranges in $[0, 1]$.

Fig. 6.55, right shows the union of all $K(t)$ if t ranges in the interval $[-1/2, 3/2]$. If t is not in $[0, 1]$, the size of $K(t)$ increases rapidly with $|t|$. So the often cited optimality of the Bernstein basis is again seen to be valid only in the interval $[0, 1]$ (see [48]). \diamond

Remark 6.4.5. Note that the present general setting includes the disk Bézier curves [111] studied earlier (see Prop. 6.3.15 and Fig. 6.31) and also interval Bézier curves [178] as special cases.

A more detailed discussion of the computation of tolerance zone boundaries is found in [153, 201], but there are still many open questions in this area. In particular, it turns out that the study of tolerance zones for metric constructions (i.e., those involving angles, distances, circles, etc.) is much more subtle than the investigation of linear constructions. \diamond

6.4.4 Two-Dimensional Normed Spaces and Minkowski Offsets

A norm $\|\cdot\|_m$ in \mathbb{R}^2 is defined by the ‘unit circle’ i_m of all vectors which have norm 1. The unit circle is the boundary of the unit disk, consisting of all vectors of norm ≤ 1 . The unit disk is a centrally symmetric closed convex domain whose center is the origin,

$$i_m : \|\mathbf{x}\|_m = 1. \quad (6.84)$$

The affine geometry of this normed vector space is called the *Minkowski geometry* defined by the unit circle i_m (the *indicatrix*). The Minkowski circles are translated and scaled versions of the indicatrix:

$$\|\mathbf{x} - \mathbf{m}\|_m = r \iff \mathbf{x} - \mathbf{e} \in r \cdot i_m. \quad (6.85)$$

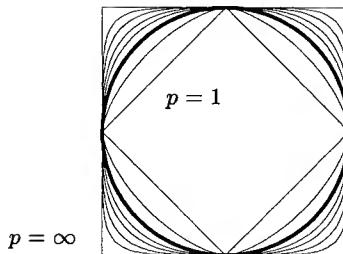


Fig. 6.56. Unit circles of some l^p norms ($p = 1, 1.5, 2, 2.5, 3, 4, 5, \infty$).

Example 6.4.4. Frequently used norms are the l^p -norms $\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p)^{\frac{1}{p}}$, where p is a real number ≥ 1 . With $p = 1$, the unit circle is a square. For $p = 2$ we get the usual Euclidean metric. In the limit $p \rightarrow \infty$ the norm $\|\cdot\|_p$ converges to $\|\mathbf{x}\|_\infty = \max(x_1, x_2)$. The unit circle is a square again (see Fig. 6.56). \diamond

The Minkowski Cyclographic Mapping

A straightforward extension of the cyclographic mapping (cf. Sec. 6.3.2) is described in what follows. We first define an *oriented Minkowski circle* or *cycle* \vec{C} in the plane as an ordinary Minkowski circle C plus an orientation, which can be thought of just as a sign of the radius or as a clockwise (negative) or counterclockwise (positive) orientation of the circle as a closed convex curve.

The point $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ is mapped to the oriented Minkowski circle $\vec{C} = \mu(\mathbf{x})$ in the plane $x_3 = 0$, whose center is $(x_1, x_2, 0)$, and whose radius equals x_3 . The circle C has radius $|x_3|$. If $x_3 = 0$, then \vec{C} has zero radius and is identified with C and \mathbf{x} .

We also generalize the notion of γ -cone. We choose a cone with vertex $\mathbf{x} = (x_1, x_2, x_3)$ ($x_3 \neq 0$) and with base curve C . The translates of this cone are called μ -cones. If $\mathbf{y} \in \mathbb{R}^3$, and $\vec{C}' = \mu(\mathbf{y})$, then C' is the intersection of the plane $x_3 = 0$ and the μ -cone with vertex \mathbf{y} (see Fig. 6.57).

After embedding affine \mathbb{R}^3 into projective three-space, all μ -cones share a common curve m_ω at infinity. In the notation of the preceding paragraph, the central projection with center \mathbf{x} maps C onto m_ω . The curve m_ω is therefore projectively equivalent to all nonzero cyclographic images and to the unit circle.

The rulings of the μ -cones are called μ -lines. A characterization of a μ -line is that it meets the ideal curve m_ω . The set of μ -lines is a line complex, and may be seen as a generalization of the line complex of constant slope (cf. Remark 6.3.5).

The properties of the cyclographic mapping ζ of Euclidean circle geometry (cf. Sec. 6.3.2) motivate to investigate similar properties of the cyclographic mapping μ of Minkowski circle geometry.

Lemma 6.4.6. *Two cycles $\mu(\mathbf{a})$ and $\mu(\mathbf{b})$ are in oriented contact, if and only if the line $\mathbf{a} \vee \mathbf{b}$ is a μ -line (Fig. 6.57).*

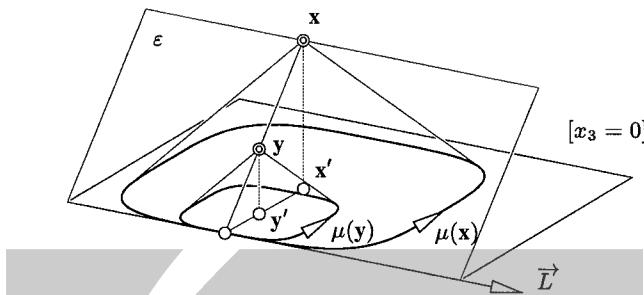


Fig. 6.57. Basic properties of the Minkowski cyclographic mapping.

The set of points \mathbf{x} such that $\mu(\mathbf{x})$ is in oriented contact with an oriented line \vec{L} is a plane tangent to the μ -cones with vertices \mathbf{x} .

The proof is a straightforward generalization of the proof of the respective properties of the cyclographic mapping ζ (see Lemma 6.3.8 and Fig. 6.57).

The planes tangent to the μ -cones are called μ -planes. Intersection with the ideal plane shows that they are precisely the planes whose ideal lines are tangent to the ideal curve m_ω .

Many properties of the classical cyclographic mapping can be extended to the present situation. We illustrate this extension by means of distance problems and Minkowski offsets.

Minkowski Offsets

Developable surfaces of constant slope or γ -developables (see Sec. 6.3) are surfaces all of whose tangent planes are γ -planes. The straightforward generalization of this notion to Minkowski geometry are developable surfaces all of whose tangent planes are μ -planes (μ -developables).

As has been mentioned in Ex. 6.3.6, there is an ideal conic c_ω such that for all γ -developables the tangent planes are tangent to c_ω and the rulings meet c_ω , i.e., the tangent planes are γ -planes and the rulings are γ -lines. It is easy to see that for any μ -developable the tangent planes are tangent to m_ω and the rulings meet m_ω , i.e., the tangent planes are μ -planes and the rulings are μ -lines.

Consider an oriented curve \vec{l} in the Minkowski plane and the set of all Minkowski cycles of a fixed radius which are in oriented contact with \vec{l} . The curve of their centers is the so-called *Minkowski offset* or *offset with respect to i_m* of \vec{l} . The proof of the following theorem is analogous to the proof of the respective properties of ordinary offset curves which have been derived with the aid of γ -developables and the Euclidean cyclographic mapping ζ .

Theorem 6.4.7. If \vec{l} is an oriented curve in the plane $x_3 = 0$, then the set of all \mathbf{x} such that $\mu(\mathbf{x})$ is in oriented contact with \vec{l} is a μ -developable \mathcal{R} which connects \vec{l} with m_ω , and which is the graph of the signed Minkowski distance of points to \vec{l} .

The level curves of \mathcal{R} project onto Minkowski offsets of \vec{l} . \mathcal{R} 's curve of regression projects onto the common Minkowski evolute of the family of Minkowski offsets, which is the locus of centers of \vec{l} 's osculating Minkowski cycles, and also the locus of the singularities of the Minkowski offsets.

A different proof of this theorem could be based on the representation of μ -developables given in the next paragraph. This representation is analogous to the representation of γ -developables given by Equ. (6.46).

Representation of μ -Developables

Consider the family of support lines $L(\phi)$ of the unit disk:

$$L(\phi) : x \sin \phi - y \cos \phi = h(\phi). \quad (6.86)$$

Recall that $h(\phi)$ is the support function of i_m which measures the distance h of i_m 's tangents to the origin in dependence of the rotation angle (cf. Equ. (6.53)). The unit disk is centrally symmetric, so we have $h(\phi) = h(\phi + \pi)$. By Lemma 6.3.7, the curvature radius of i_m is given by $h + \ddot{h}$, if h is C^2 . Here we will always require that the support function of the unit disk is twice continuously differentiable, and that the unit disk does not have cusps, which implies that $h(\phi) + \ddot{h}(\phi) > 0$. The μ -cone whose vertex is the point $(0, 0, 1)$ has the base curve curve i_m , so its tangent planes are parametrized by

$$x \sin \phi - y \cos \phi + h(\phi)z = h(\phi). \quad (6.87)$$

All of them are μ -planes, and all μ -planes are parallel to one of them. This means that an arbitrary family of μ -planes may be parametrized by

$$U(t) : x \sin \phi(t) - y \cos \phi(t) + h(\phi(t))z = k(t). \quad (6.88)$$

If $\dot{\phi}(t) \neq 0$, we locally parametrize this family of planes by

$$U(\phi) : x \sin \phi - y \cos \phi + h(\phi)z = k(\phi). \quad (6.89)$$

This family of planes defines a μ -developable \mathcal{R} , so Equations (6.88) and (6.89) are the general dual representation of a μ -developable. We see that a smooth μ -developable depends on a smooth scalar function $k(t)$ or $k(\phi)$. We compute the derivative planes using (6.89)

$$\begin{aligned} \dot{U}(\phi) : \quad & x \cos \phi + y \sin \phi + \dot{h}z = \dot{k}, \\ \ddot{U}(\phi) : \quad & -x \sin \phi + y \cos \phi + \ddot{h}z = \ddot{k}. \end{aligned}$$

\mathcal{R} 's curve of regression equals $c(\phi) = (U \cap \dot{U} \cap \ddot{U})(\phi)$. If $c(\phi) = (x(\phi), y(\phi), z(\phi))$, a simple computation gives the result

$$c(\phi) = \begin{bmatrix} k \sin \phi + \dot{k} \cos \phi \\ -k \cos \phi + \dot{k} \sin \phi \\ 0 \end{bmatrix} + \frac{k + \ddot{k}}{h + \ddot{h}} \begin{bmatrix} h \sin \phi + \dot{h} \cos \phi \\ h \cos \phi - \dot{h} \sin \phi \\ 1 \end{bmatrix}. \quad (6.90)$$

The level sets of \mathcal{R} project onto curves which are Minkowski offsets of each other. The family of tangents of such a curve \vec{l}_d is parametrized by

$$x \sin \phi - y \cos \phi = k(\phi) - dh(\phi),$$

where d indicates the z -coordinate of the original level set. Obviously $k(\phi) - dh(\phi)$ is the support function of \vec{l}_d . This shows that the Minkowski offset \vec{l}_d is a Minkowski sum of the oriented curve \vec{l}_0 corresponding to level $d = 0$ and a Minkowski cycle of radius d .

Example 6.4.5. We define functions $h(\phi) = 9 + \sin(2\phi)/2 - \sin(3\phi)/3 + \sin(4\phi)/4$ and $k(\phi) = 1$. Fig. 6.58 shows the μ -developable whose family of tangent planes is parametrized by Equ. (6.89).

Note that $h(\phi)$ is not the support function of the unit disk of a Minkowski plane in our sense, because $h(\phi) \neq h(\phi + \pi)$. The theory nevertheless is valid in the asymmetric case also. \diamond

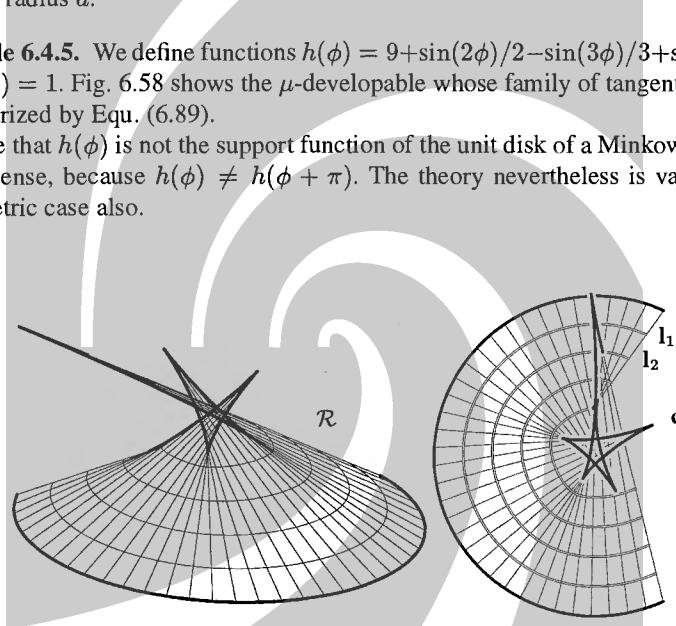


Fig. 6.58. Left: Developable μ -surface \mathcal{R} with level curves and curve of regression. Right: top view of \mathcal{R} with Minkowski offset curves and Minkowski evolute c .

Remark 6.4.6. Rational μ -surfaces can be used to study *rational* curves with *rational* Minkowski offsets [3]. Because the curve m_ω is a planar section of such a surface, m_ω and therefore i_m have to be rational. If however i_m is rational it is easy to find a rational parametrization of the tangent planes of a μ -cone, and arbitrary rational families of μ -planes.

The envelopes of such families are precisely the rational μ -developables, and their level curves are rational with rational μ -offsets. We refer to [147], where the analogous question is treated for surfaces. Minkowski offset surfaces are particularly interesting since they appear in 3-axis NC machining of surfaces. \diamond

Distance Problems in Minkowski Geometry

Th. 6.4.7 describes the graph of the Minkowski distance function to a given planar curve. This distance function is essential if we want to compute Minkowski bisector curves and Minkowski-Voronoi diagrams.

These concepts are defined analogously to the Euclidean case, but turn out to be in part more complicated. This can already be seen by means of the *bisector of two points*: The graph of the distance function to a point is the μ -cone whose vertex is this point, and the intersection of two such μ -cones projects onto the bisector curve. This intersection however need not be planar at all and so we cannot expect that the bisector is a straight line.

On the other hand, the bisector of two oriented lines is a straight line, because the graph of the distance function to an oriented line is a μ -plane.

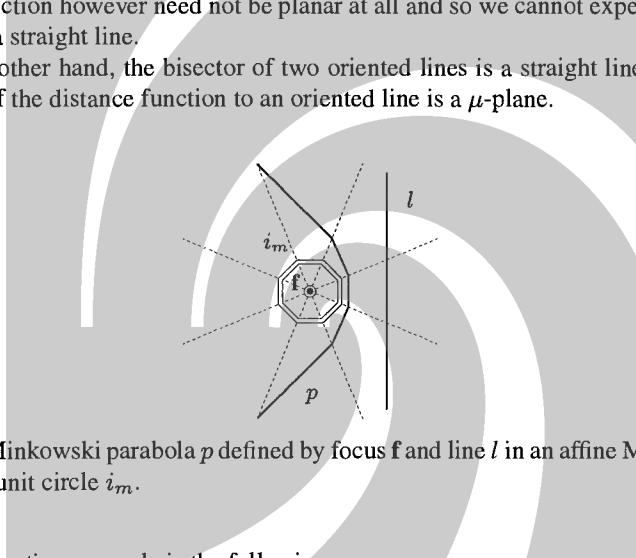


Fig. 6.59. Minkowski parabola p defined by focus f and line l in an affine Minkowski plane with unit circle i_m .

An interesting example is the following:

Lemma 6.4.8. *A Minkowski parabola p , defined as bisector of a line l and a non-incident point f , is a curve which is projectively equivalent to a Minkowski circle and is tangent to the line at infinity (if the unit circle is smooth).*

Proof. The graph of the distance function to the straight line is a μ -plane ε , and the graph of the distance function to f is a μ -cone $M(f)$ whose vertex is f . The top view of this intersection is the Minkowski parabola p (see Fig. 6.59). Thus i_m and p are projectively equivalent, as both are planar sections of the same cone.

Because ε is parallel to a tangent plane of $M(f)$, the intersection $\varepsilon \cap M(f)$ is tangent to the ideal plane, and therefore p is tangent to the ‘top view’ of the ideal plane, i.e., the ideal line of the horizontal plane. \square

This behaviour is completely analogous to that of the Euclidean parabola, which is a curve projectively equivalent to the Euclidean unit circle and touches the ideal line.

Remark 6.4.7. If the unit circle is not smooth, there is still a projective automorphism of the plane which carries p to i_m . The image of the ideal line is a support line of i_m (this is the case in Fig. 6.59).

For a study of bisectors and Voronoi diagrams in Minkowski geometries (and generalizations of those) we refer to a book of R. Klein [97]. The reader of the present will easily find out which parts could benefit from the use of the Minkowski cyclographic mapping μ . \diamond

6.5 Developable Surfaces with Creases

Let us briefly recall (see Sec. 5.3) some basic facts from differential geometry. Assume that $s(u, v)$ parametrizes a developable surface, $c(t) = s(u(t), v(t))$ is a curve in this surface, and $\bar{s}(u, v)$ is the development of the surface. We would like to compute the curvature of the development $\bar{c}(t) = \bar{s}(u(t), v(t))$.

Recall that the geodesic curvature $\kappa_g(t)$ of the curve (cf. Equ. (5.35)) is a property of the intrinsic surface metric and does not change if we apply isometric mappings. The curvature κ , the geodesic curvature κ_g and the angle α enclosed by c 's osculating plane and s 's tangent plane are related by

$$\kappa_g = \kappa \cos \alpha. \quad (6.91)$$

By taking inverses, we get the following relation between the curvature radii $\rho = 1/\kappa$ and $\bar{\rho} = 1/\kappa_g$ of c and \bar{c} :

$$\rho = \bar{\rho} \cos \alpha. \quad (6.92)$$

Remark 6.5.1. Equ. (6.92) says that the axis of $c(t)$'s osculating circle $c_{osc}(t)$ intersects the tangent plane U of the surface s in some point p which has the property that the distance $pc(t) = \bar{\rho}$ (see Fig. 6.60). It is easy to see that $\bar{\rho}$ equals the curvature radius of c 's orthogonal projection onto U . \diamond

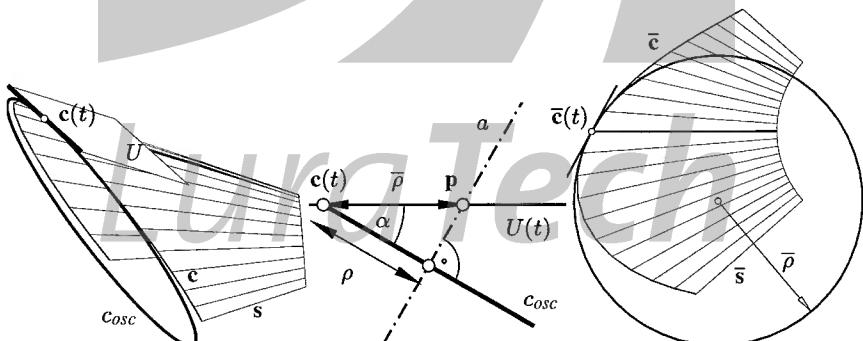


Fig. 6.60. Left: Developable surface s , curve c , and osculating circle c_{osc} . Center: View in direction of the curve tangent. Right: development into the tangent plane U .

Let us point to a special case. The curve \bar{c} has zero curvature (has an inflection point), if either c itself has zero curvature (then c is tangent to a ruling, or the surface has a flat point), or the angle α in Equ. (6.91) equals $\pi/2$.

Two Developable Surfaces which Contain the Same Curve

We consider two developable surfaces, parametrized by s_1 and s_2 , which both contain the curve $c = s_1(u_1(t), v_1(t)) = s_2(u_2(t), v_2(t))$. The tangent planes of the two surfaces in the point $c(t)$ will in general differ from each other (Fig. 6.61). If the surfaces actually *touch* each other in all points $c(t)$, then they would be essentially equal, because a developable surface is uniquely determined by the tangent planes in the points of a director curve.

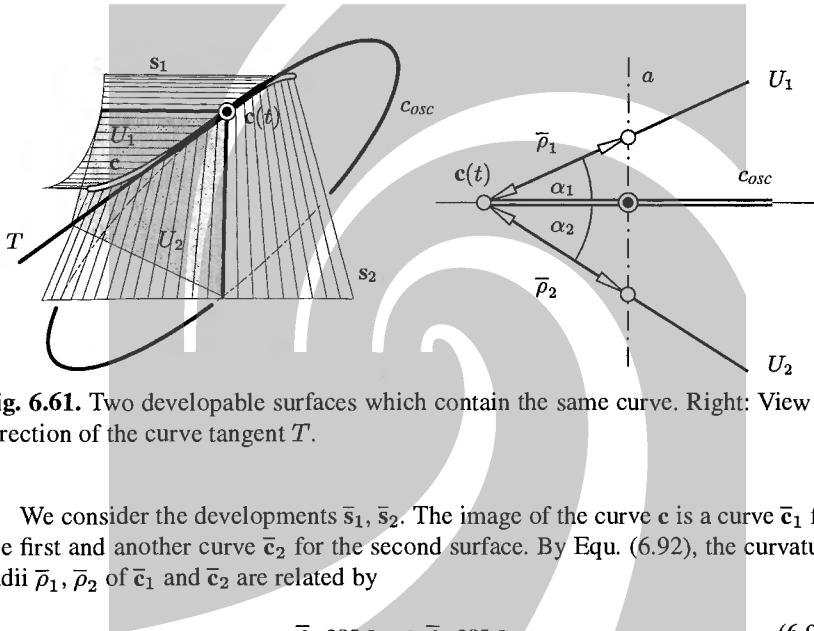


Fig. 6.61. Two developable surfaces which contain the same curve. Right: View in direction of the curve tangent T .

We consider the developments \bar{s}_1, \bar{s}_2 . The image of the curve c is a curve \bar{c}_1 for the first and another curve \bar{c}_2 for the second surface. By Eq. (6.92), the curvature radii $\bar{\rho}_1, \bar{\rho}_2$ of \bar{c}_1 and \bar{c}_2 are related by

$$\bar{\rho}_1 \cos \alpha_1 = \bar{\rho}_2 \cos \alpha_2, \quad (6.93)$$

where α_i is the angle enclosed by c 's osculating plane and the tangent plane U_i of the surface s_i ($i = 1, 2$).

This simple fact allows us to solve the following problem. Given are a developable surface s_1 , whose boundary is a curve c , and the development \bar{s}_2 of a second developable surface, whose boundary is a curve \bar{c}_2 . Find a surface s_2 such that it contains c and the development of c is the given curve \bar{c}_2 (the surface s_1 has no influence on the solution of this problem, only the curve c).

We could say that we are given a sheet of paper with a well-defined boundary curve, and we have to attach this sheet of paper to a given space curve.

Equations (6.92) and (6.93) show how to solve this problem: We know c 's radius of curvature $\rho = \bar{\rho}_1 \cos \alpha_1$ and from \bar{c}_2 's radius of curvature $\bar{\rho}_2$ we can compute the angle α_2 which is enclosed by c 's osculating plane and the unknown tangent plane U_2 of the surface s_2 . If we compute $U_2(t)$ for all points $c(t)$, we get the surface s_2 as envelope of its tangent planes.

D-Forms

An interesting but presumably difficult problem which fits into the current discussion has been posed by Tony Wills, London. Wills takes two planar sheets of unstretchable material (such as paper), which are bounded by closed convex curves \bar{c}_1 and \bar{c}_2 of equal total arc length. Then he attaches these two sheets to each other, starting at arbitrary points p_1 in the first and p_2 in the second curve. Finally the two sheets are glued together along their common boundaries. Mathematically speaking, we have to find two developable surfaces s_1, s_2 , bounded by the same closed curve c , whose developments \bar{c}_1, \bar{c}_2 are given. The body bounded by these two surfaces is called a *D-form*. After some experiments we found that, surprisingly, both s_1 and s_2 were free of creases, but we do not know whether this will be so in all cases.

Figure 6.49 shows an example of a D-form. This follows from the fact that the developments of both s_1, s_2 map c to a convex curve, or from the fact that neither development of c has inflection points — c 's osculating plane is nowhere orthogonal to the surface s_i ($i = 1, 2$).

It is not clear under what conditions a D-form is the convex hull of a space curve. Even more difficult seems the task of actually computing the D-form from its development. The reason for this difficulty is that the problem is of a global nature. One can nicely observe that when building a model of a D-form from paper. The sheets are changing their shape all the time until the procedure is complete and the final shape of the D-form appears.

A design by Tony Wills using a truncated D-form is shown in Fig. C.7.

Developable Surfaces with Creases

The existence of crumpled paper suggests that there are developable surfaces which are not smooth. We will study here developable surfaces with *creases*, which in our context will mean developable surfaces which are smooth everywhere except in a finite number of smooth curve segments.

To create such shapes, we first look at one such segment. The following is a fundamental property of developable surfaces with creases:

Proposition 6.5.1. *In all points of a crease curve c in a developable surface, the osculating plane of c in the point $c(t)$ is a bisector plane of the two tangent planes of the surface in the point $c(t)$.*

Proof. We consider a neighbourhood of the point $c(t)$. Developing the two smooth developable surfaces s_1, s_2 at either side of the curve c must result in two *congruent* developments \bar{c}_1, \bar{c}_2 of the curve c . In the notation of Equ. (6.93), this is equivalent to the equality of the curvature radii $\bar{\rho}_1(t)$ and $\bar{\rho}_2(t)$ for all t . By (6.93), the angles α_1 and α_2 enclosed by the tangent planes U_1, U_2 of s_1, s_2 and c 's osculating plane must be equal. This is illustrated by Fig. 6.61. □

Prop. 6.5.1 says that if s_1 and c are given, we find the tangent planes $U_2(t)$ of s_2 by reflecting the tangent planes $U_1(t) = \mathbb{R}u_1(t)$ in the osculating planes $P(t) = \mathbb{R}p(t)$ of c . If we use homogeneous plane coordinates, we have

$$u_2 = 2 \frac{\mathbf{p} \cdot \mathbf{u}_1}{\mathbf{p}^2} \mathbf{p} - \mathbf{u}_1, \quad \text{where } \mathbf{p} = (p_0, \mathbf{p}), \mathbf{u} = (u_{10}, \mathbf{u}_1). \quad (6.94)$$

Proposition 6.5.2. *If s_1 is a developable surface and $\mathbf{c}(t)$ is a curve in it, then Equ. (6.94) computes the dual representation of the unique developable surface s_2 which extends s_1 with crease \mathbf{c} . The composite surface has a tangent discontinuity in those points of \mathbf{c} where the osculating plane is not orthogonal to the tangent plane of s_1 .*

Proof. Equ. (6.94) ensures that when developing s_1 and s_2 into the plane, the curvatures of the developments \bar{s}_1 and \bar{s}_2 of the curve \mathbf{c} are equal in respective points. The development is isometric, so the arc length differential of \bar{s}_1 equals that of \bar{s}_2 . This is sufficient for congruence of the curves \bar{s}_1 and \bar{s}_2 , so s_2 indeed extends s_1 .

As developable surfaces are uniquely determined by their tangent planes, s_2 is unique. The tangent plane $U_2(t)$ of s_2 equals the tangent plane $U_1(t)$ of s_1 if and only if the reflection described by Equ. (6.94) takes U_1 to U_2 , i.e., if the osculating plane is orthogonal to the tangent plane. \square

The explicit dual representation (6.94) of the surface s_2 shows that s_2 is rational if (i) s_1 is rational and (ii) $\mathbf{c} = s_1(u(t), v(t))$ is rational in s_1 in the sense that $u(t)$ and $v(t)$ are rational functions. We can estimate the degree of s_2 by

$$\deg(u_2) = 2 \deg(\mathbf{p}) + \deg(\mathbf{u}_1).$$

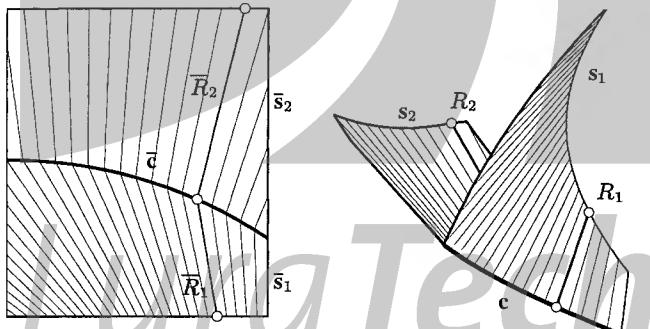


Fig. 6.62. Developable surface with a crease (right) and its development (left).

Example 6.5.1. Fig. 6.62 shows an example of a developable surface with a crease. We start from the development (left) of the surfaces s_1, s_2 , which have the common boundary \bar{c} . We choose a curve \mathbf{c} in space and the developable s_1 . Then s_2 is constructed as envelope of its tangent planes $U_2(t)$, which are found by reflecting the tangent planes $U_1(t)$ of s_1 in the osculating planes of \mathbf{c} . We compute the rulings $R_1(t) = U_1(t) \cap U_2^1(t)$ and $R_2(t) = U_2(t) \cap U_1^1(t)$. Finally we find the developments $\bar{R}_1(t)$ and $\bar{R}_2(t)$ of the rulings. \diamond

More details on the computation of models for crumpled paper may be found in papers by Kergosien et al. [95] and Ben Amar and Pomeau [11]. Inflection points $\bar{c}(t)$ in the planar development occur if the planes $P(t)$ and $U_1(t)$ are orthogonal. In this case the reflection of U_1 in P gives U_1 again and we see that in such points there is no tangent discontinuity. In particular, the crease can never be a geodesic line, because then we would have orthogonality between the osculating plane and the surface in all points.

Geometry of the Shoulder of a Packaging Machine

The shoulder of a packaging machine is a surface that should guide the material (paper or plastic sheet) from a horizontal roll into a vertical circular cylinder (see Fig. 6.63).

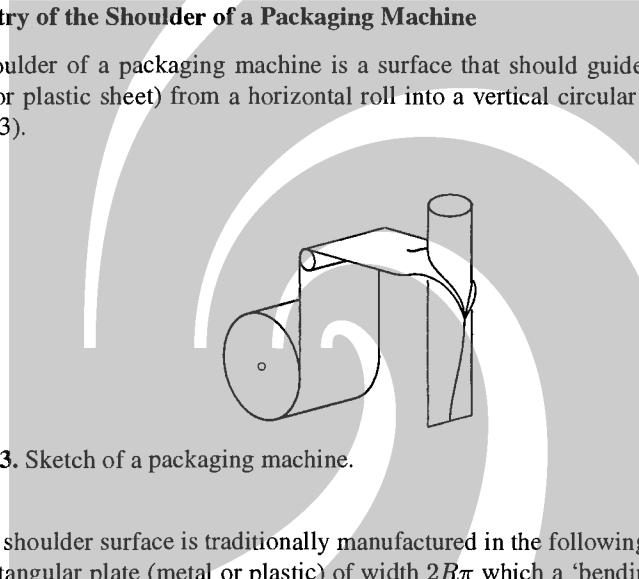


Fig. 6.63. Sketch of a packaging machine.

The shoulder surface is traditionally manufactured in the following way: Take a thin rectangular plate (metal or plastic) of width $2R\pi$ which a ‘bending curve’ \bar{c} is carved into. Then the part below \bar{c} is wrapped around a right circular cylinder s_1 of radius R . Thereby \bar{c} is mapped to a curve c in s_1 , and the upper part takes on the shape of the shoulder surface s_2 (see Fig. 6.64).

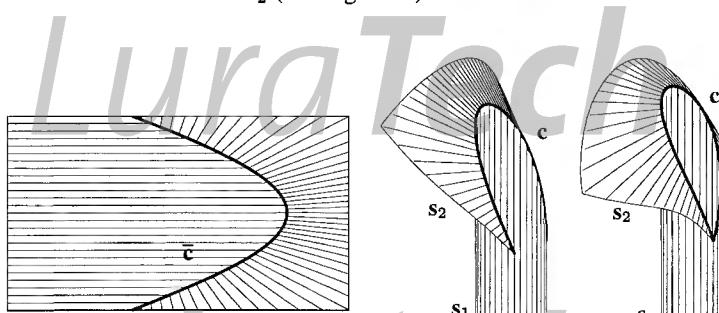


Fig. 6.64. Bending a planar sheet into a surface consisting of a cylinder and the shoulder surface (courtesy S. Leopoldseder).

The construction of shoulders subject to various geometric constraints is an interesting topic, which has been discussed by Boersma and Molenaar [16]. Even simpler than the approach taken there, we can analyze it as follows. We immediately see that the object consisting of the cylinder and the shoulder is a developable surface with crease curve \mathbf{c} . For simplicity, we assume $R = 1$, and we parametrize the planar curve $\bar{\mathbf{c}}$ and the crease \mathbf{c} in the form

$$\bar{\mathbf{c}}(t) = (t, h(t)), \quad \mathbf{c}(t) = (\cos t, \sin t, h(t)). \quad (6.95)$$

A vector orthogonal to \mathbf{c} 's osculating plane is given by $\dot{\mathbf{c}} \times \ddot{\mathbf{c}}$. Thus the osculating plane has the homogeneous plane coordinates

$$\mathbf{p}(t) = (-h - \ddot{h}, \dot{h} \cos t + \dot{h} \sin t, \ddot{h} \sin t - \dot{h} \cos t, 1).$$

The dual representation of the cylinder reads $\mathbf{u}_1(t) = (-1, \cos t, \sin t, 0)$, and Equ. (6.94) computes the dual representation $\mathbf{u}_2(t)$ of the shoulder surface by

$$\begin{aligned} \mathbf{u}_2(t) = & (1 + \dot{h}^2 - 2h\ddot{h} - \ddot{h}^2, 2\dot{h}\ddot{h} \sin t - (1 + \dot{h}^2 - \ddot{h}^2) \cos t, \\ & -2\dot{h}\ddot{h} \cos t - (1 + \dot{h}^2 - \ddot{h}^2) \sin t, 2\ddot{h}). \end{aligned} \quad (6.96)$$

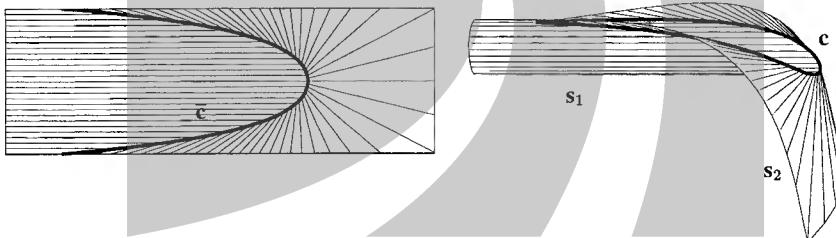


Fig. 6.65. Shoulder surface orthogonal to the cylinder — see Ex. 6.5.2 (courtesy S. Leopoldseder).

Example 6.5.2. As a first interesting example, we compute a shoulder surface s , which encloses a constant angle 2γ with the cylinder s_1 in all points of the crease \mathbf{c} . Therefore, by Prop. 6.5.1, the osculating planes of \mathbf{c} must enclose the constant angle $\gamma < \pi/2$ with both the cylinder and the shoulder.

A curve \mathbf{c} with this property is called a *pseudo-geodesic* — if the angle equals $\pi/2$, the curve is a geodesic. Pseudo-geodesics have been introduced by W. Wunderlich [212]. In particular, he investigated pseudo-geodesics of the right circular cylinder, which are exactly the bending curves we are looking for. According to Wunderlich, a pseudo-geodesic of the cylinder s_1 is characterized by

$$h(t) = \frac{1}{a} \cosh at, \quad a = \cot \gamma. \quad (6.97)$$

Equ. (6.96) now gives the dual representation of the shoulder surface:

$$\begin{aligned} \mathbf{u}_2(t) = & -(1 + a^2) \cosh at, 2a \sinh at \sin t + (a^2 - 1) \cosh at \cos t, \\ & -2a \sinh at \cos t + (a^2 - 1) \cosh at \sin t, 2a). \end{aligned} \quad (6.98)$$

A case which is particularly simple is $\gamma = \pi/4$, i.e., $a = 1$. Then we have

$$\mathbf{u}_2(t) = (-\cosh t, \sinh t \sin t, -\sinh t \cos t, 1). \quad (6.99)$$

This case is illustrated in Fig. 6.65. The curve of regression of the surface s_2 has the homogeneous parametrization $\mathbf{r}(t) = \mathbf{u}_2(t) \times \dot{\mathbf{u}}_2(t) \times \ddot{\mathbf{u}}_2(t)$:

$$\begin{aligned} \mathbf{r}(t) = & (2 \cosh t, \cosh t \cos t + \sinh t \sin t, \\ & \cosh t \sin t - \sinh t \cos t, 1 + \cosh^2 t). \end{aligned} \quad \diamond$$

The curve of regression restricts the size of a regular developable with a crease. In reality, if we force a thin plate to assume the shape of s_2 , it is highly unlikely that this is possible beyond the line of regression — the material will develop further creases.

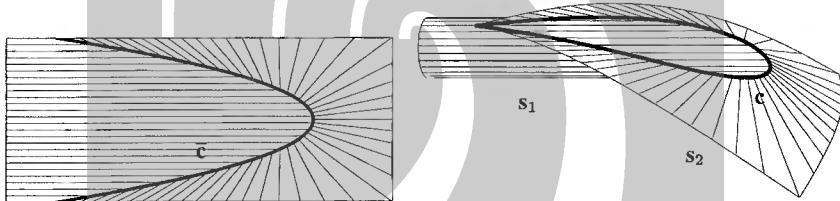


Fig. 6.66. Conical shoulder surface (courtesy S. Leopoldseder).

Example 6.5.3. As a second example, which has been treated in detail by Boersma and Molenaar [16], we investigate *conical shoulders*. The dual approach yields the result very quickly: The shoulder is a cone with vertex $\mathbf{v} = (v_1, v_2, v_3)$, if \mathbf{v} lies in all tangent planes of the shoulder. This incidence condition, which reads

$$(1, v_1, v_2, v_3) \cdot \mathbf{u}_2(t) = 0,$$

is actually a differential equation for the function $h(t)$. By an appropriate choice of the coordinate system, we can achieve $v_2 = v_3 = 0$, i.e., the vertex of the cone is contained in the x -axis. The differential equation mentioned above reads

$$1 + \dot{h}^2 - 2h\ddot{h} - \ddot{h}^2 + v_1[2\dot{h}\ddot{h} \sin t - (1 + \dot{h}^2 - \ddot{h}^2) \cos t] = 0. \quad (6.100)$$

We are interested in the special case of a conical shoulder whose vertex is contained in the cylinder's axis, i.e., $v_1 = 0$. The solution of the corresponding differential equation $1 + \dot{h}^2 - 2h\ddot{h} - \ddot{h}^2 = 0$ is the quadratic polynomial $h(t) = a_0 + a_1 t + a_2 t^2$ whose coefficients satisfy $1 + a_1^2 - 4a_0 a_2 - 4a_2^2 = 0$. Thus the crease's development \bar{c} is a *parabola*. Such a conical shoulder is illustrated in Fig. 6.66. \diamond

7. Line Congruences and Line Complexes

This chapter is devoted to two- and three-dimensional manifolds of lines — so-called line congruences and line complexes. We have already encountered them in their simplest forms: Chap. 3 deals with linear complexes and linear congruences of lines, and in Sec. 6.3 we considered the three-dimensional manifold of lines of constant slope.

Here we study line congruences and complexes from the viewpoint of projective and Euclidean differential geometry, together with applications in curve and surface theory in three-space. We show how the set of normals of a surface in Euclidean space — which is a line congruence — arises naturally in collision problems in five-axis milling. Further, we study algebraic and rational congruences and complexes, and their relations to geometrical optics.

7.1 Line Congruences

7.1.1 Projective Differential Geometry of Congruences

Two-parameter families \mathcal{K} of lines in projective three-space P^3 are called *line congruences*. We use the Klein image to give the following definition:

Definition. A set \mathcal{K} of lines in P^3 is called a C^r line congruence, if its Klein image $\mathcal{K}\gamma$ is a C^r 2-surface contained in the Klein quadric M_2^4 .

The Klein image of a *linear* line congruence is a quadric or a quadratic cone contained in a three-dimensional projective subspace of P^5 (cf. Th. 3.2.4, Th. 3.2.7, Th. 3.2.9), and such sets can be parametrized as 2-surfaces. Thus the linear line congruences of Sec. 3.2 are indeed special instances of line congruences.

As to notation, we will denote the *lines* of the congruence by K (or $K(u^1, u^2)$) to emphasize the dependence on two parameters), and the Plücker coordinates of K by $(\mathbf{k}, \bar{\mathbf{k}})$. We will abbreviate (u^1, u^2) with the symbol u .

Being a 2-surface means that $\mathcal{K}\gamma$ is locally parametrized as a C^r 2-surface according to our definition of surfaces in Sec. 1.2.2:

$$(u^1, u^2) \in D \mapsto K\gamma(u^1, u^2) \in M_2^4. \quad (7.1)$$

A line $K(u)$ of the congruence is called *regular*, if the point $K\gamma(u)$ is a regular point of the surface $K\gamma$, which means that the three points

$$(K\gamma)(u), \quad (K\gamma)_{,1}(u), \quad (K\gamma)_{,2}(u)$$

are projectively independent. The symbol $(K\gamma)_{,i}$ denotes the point $(\mathbf{k}_{,i}, \bar{\mathbf{k}}_{,i})\mathbb{R}$, and the subscript with a comma indicates differentiation with respect to the i -th parameter. In case of regularity, the tangent space $K\gamma \vee (K\gamma)_{,1} \vee (K\gamma)_{,2}$ is two-dimensional. Since $\mathcal{K}\gamma(u) = (\mathbf{k}(u), \bar{\mathbf{k}}(u))\mathbb{R}$ is entirely contained in the Klein quadric, we may differentiate the Plücker identity $\mathbf{k} \cdot \bar{\mathbf{k}} = 0$ to get

$$0 = \mathbf{k}_{,j} \cdot \bar{\mathbf{k}} + \mathbf{k} \cdot \bar{\mathbf{k}}_{,j} = \mathbf{k}_{,jk} \cdot \bar{\mathbf{k}} + \mathbf{k}_{,j} \cdot \bar{\mathbf{k}}_{,k} + \mathbf{k}_{,k} \cdot \bar{\mathbf{k}}_{,j} + \mathbf{k} \cdot \bar{\mathbf{k}}_{,jk} = \dots \quad (7.2)$$

Ruled Congruence Surfaces

We are interested in ruled surfaces \mathcal{R} contained in a congruence \mathcal{K} . Such a surface is called a *congruence surface* and its Klein image $\mathcal{R}\gamma$ is a curve contained in the surface $\mathcal{K}\gamma$. If \mathcal{K} is parametrized as in (7.1), we consider a curve $u(t)$ in the parameter domain D . Then $K(u(t))$ parametrizes a congruence surface. It turns out that the ‘direction’ of this surface within the congruence is important for its first order properties.

Definition. *If $K(u)$ parametrizes a line congruence \mathcal{K} as in Equ. (7.1), and $u(t)$ is a curve in K 's parameter domain with tangent vector $\dot{u} = (\dot{u}^1, \dot{u}^2)$ at $t = t_0$, then the ratio $\dot{u}^1 : \dot{u}^2$ is called direction in \mathcal{K} at the line $K(u(t_0))$.*

Lemma 7.1.1. *If \mathcal{K} , as parametrized by (7.1), is a line congruence and $R(t) = K(u(t))$ is a congruence surface in \mathcal{K} , then the tangent $L = R\gamma(t_0) \vee (R\gamma)^1(t_0)$ depends only on the direction $\dot{u}^1(t_0) : \dot{u}^2(t_0)$.*

Proof. Differentiation shows that the derivative point $(R\gamma)^1$ equals $(\dot{u}^1(\mathbf{k}_1, \bar{\mathbf{k}}_1) + \dot{u}^2(\mathbf{k}_2, \bar{\mathbf{k}}_2))\mathbb{R}$, so $R\gamma \vee (R\gamma)^1$ depends only on the ratio $\dot{u}^1 : \dot{u}^2$. \square

Lemma 7.1.2. *If the congruence \mathcal{K} is parametrized by (7.1), then the congruence surface $R(t) = K(u(t))$ is regular if its lines are regular lines of \mathcal{K} and the curve $u(t)$ in the parameter domain is regular. Two such congruence surfaces $R_i(t)$ ($i = 1, 2$) which share a common direction at $t = t_0$ are in first order contact in the points of the common ruling $R_1(t_0) = R_2(t_0)$.*

Proof. Consider the derivative point $(R\gamma)^1$, which has the homogeneous coordinates $\dot{u}^1(\mathbf{k}_1, \bar{\mathbf{k}}_1) + \dot{u}^2(\mathbf{k}_2, \bar{\mathbf{k}}_2)$. Regularity of the curve $u(t)$ (i.e., $(\dot{u}^1, \dot{u}^2) \neq (0, 0)$) together with regularity of \mathcal{K} obviously implies that $(R\gamma)^1(t_0) \neq R\gamma(t_0)$, so $R(t)$ is regular as a ruled surface.

If R_1 and R_2 are as above, Lemma 7.1.1 implies that the tangents $R_1\gamma \vee (R_1\gamma)^1$ and $(R_2\gamma) \vee (R_2\gamma)^1$ coincide. Thus $R_1\gamma$ and $R_2\gamma$ are in first order contact, and Th. 5.1.3 shows the statement of the lemma. \square

First Order Differential Properties

The simplest projective invariants of surfaces computable by differentiating once are tangents. If we consider surfaces Φ contained in another surface Φ' , then Φ 's tangents are of course not contained in Φ' in general, but they are always in first order contact with Φ' . Contact of higher order is possible, and the set of Φ 's tangents which are in higher order contact with Φ are another first order differential invariant of the surface Φ as a submanifold of Φ' .

Here we study surfaces $K\gamma$ contained in the Klein quadric. If a line is in second order contact with a quadric, it is entirely contained in this quadric. Thus we define:

Definition. Assume a line congruence \mathcal{K} as parametrized by (7.1). A direction $\dot{u}^1 : \dot{u}^2$ is called *torsal*, if the tangent L of the surface $K\gamma$ defined by this direction according to Lemma 7.1.1 is contained in the Klein quadric, and *non-torsal* otherwise.

All congruence surfaces $R(t) = K(u(t))$ which share this direction at t_0 , also share the tangent $L = (R\gamma)(t_0) \cup (R\gamma)^1(t_0)$. If the direction is torsal, the ruling $R(t_0)$ is torsal, and vice versa by the definition of a torsal ruling (cf. Sec. 5.1.1).

In order to determine how many of $K\gamma$'s tangents are contained in the Klein quadric, we study the shapes of the intersection of a plane T^2 tangent to the Klein quadric with this quadric.

Obviously the intersection $T^2 \cap M_2^4$ is a quadratic variety in T^2 , and because T^2 is tangent to M_2^4 , this variety is contained in a tangential section of the Klein quadric (see Ex. 1.1.29). Thus $T^2 \cap M_2^4$ may be a single point, one line, two lines, or may even coincide with entire T^2 .

Definition. A plane T^2 tangent to the Klein quadric in a point $R\gamma$ is called *elliptic*, *parabolic*, *hyperbolic*, or *torsal*, if 0, 1, 2, or all lines of T^2 incident with $R\gamma$ are contained in the Klein quadric.

A regular line $K(u_0)$ of a line congruence \mathcal{K} is called *elliptic*, *parabolic*, *hyperbolic*, or *torsal*, if the tangent plane of $K\gamma$ at $K\gamma(u_0)$ has this property, i.e., if 0, 1, 2, or all directions are torsal.

Assume that \mathcal{K} is parametrized by (7.1). A surface tangent L of $K\gamma$ corresponds to a direction $\dot{u}^1 : \dot{u}^2$ at $u = u_0$ in the parameter domain. It is spanned by $K\gamma(u_0)$ and the point

$$(\dot{u}^1(\mathbf{k}_{,1}, \bar{\mathbf{k}}_{,1}) + \dot{u}^2(\mathbf{k}_{,2}, \bar{\mathbf{k}}_{,2})) \mathbb{R}.$$

We have $L \subset M_2^4$ if and only if this point is contained in M_2^4 . This condition reads

$$\begin{aligned} q_{u_0}(\dot{u}) &= (\dot{u}^1 \mathbf{k}_{,1} + \dot{u}^2 \mathbf{k}_{,2}) \cdot (\dot{u}^1 \bar{\mathbf{k}}_{,1} + \dot{u}^2 \bar{\mathbf{k}}_{,2}) = \sum_{j,k=1,2} \bar{a}_{jk} \dot{u}^j \dot{u}^k = 0 \\ \text{with } a_{jk} &= \mathbf{k}_{,j}(u_0) \cdot \bar{\mathbf{k}}_{,k}(u_0), \quad \bar{a}_{jk} = (a_{jk} + a_{kj})/2. \end{aligned} \quad (7.3)$$

The function q_{u_0} defined by Equ. (7.3) is a quadratic form with coordinate matrix (\bar{a}_{jk}) . Depending on the zero set of q_{u_0} , we have the following four cases:

- A. $\det(\bar{a}_{jk}) < 0$: The quadratic form q_{u_0} is indefinite. $T^2 \cap M_2^4$ consists of two lines and there are two torsal directions at u_0 . $K(u_0)$ is a hyperbolic line of \mathcal{K} .
- B. $\det(\bar{a}_{jk}) > 0$: q_{u_0} is positive or negative definite and $T^2 \cap M_2^4$ consists only of one point. There are no torsal directions. The line $K(u_0)$ is elliptic.
- C. $\text{rk}(\bar{a}_{jk}) = 1$: $T^2 \cap M_2^4$ consists of one line and there is one torsal direction. The line $K(u_0)$ is parabolic.
- D. $\bar{a}_{jk} = 0$ for all j, k : T^2 is contained in M_2^4 and all directions are torsal. The line $K(u_0)$ is torsal.

Example 7.1.1. The different kinds of linear congruences we studied in Sec. 3.2.1 furnish examples of line congruences where all lines are of the same type: All lines R in a hyperbolic linear congruence \mathcal{N} are hyperbolic, because there are two line pencils in \mathcal{N} which contain R (and R is of course a torsal generator of these pencils).

We have $\mathcal{N}\gamma = G^3 \cap M_2^4$ with a projective subspace G^3 of dimension three, and $\mathcal{N}\gamma$ is a ruled quadric in G^3 . Any tangent plane T^2 is contained in G^3 , so $T^2 \cap M_2^4 = T^2 \cap (G^3 \cap M_2^4) = T^2 \cap \mathcal{N}\gamma$. Because the intersection of a ruled quadric with its tangent plane consists of two lines, this shows again that all lines of \mathcal{N} are hyperbolic.

An example of a line congruence which contains only elliptic lines is an elliptic linear congruence \mathcal{N} , because $\mathcal{N}\gamma$ is an oval quadric, and its tangent planes have only one point in common with $\mathcal{N}\gamma$.

The lines of a parabolic linear congruence are parabolic, except its axis, which is a singular line. This is seen in the same way. $\mathcal{N}\gamma$ is a quadratic cone, and the intersection of a tangent plane with this cone consists of one line.

A bundle of lines and a field of lines are line congruences which contain only torsal lines, and it can be shown that these are the only connected C^2 congruences with this property. \diamond

Definition. Assume a torsal direction $\dot{u}^1 : \dot{u}^2$ in a congruence \mathcal{K} , which defines a line L contained in the Klein quadric (Lemma 7.1.1). There is a line pencil \mathcal{R} with $\mathcal{R}\gamma = L$. Its vertex is called *focal point* and its plane *focal plane* of this torsal direction.

Lemma 7.1.3. Consider a congruence surface R of \mathcal{K} which has a torsal generator $R(t_0)$, corresponding to a torsal direction $\dot{u}^1 : \dot{u}^2$. The cuspidal point and the torsal plane of $R(t_0)$ are the focal point and focal plane of this torsal direction.

Proof. Consider the line pencil mentioned in the definition immediately before this lemma. All ruled surfaces in \mathcal{K} which share the given torsal direction are in first order contact with this line pencil (by Th. 5.1.3). This means that they have the same common torsal plane and cuspidal point (cf. the proof of Th. 5.1.5). \square

Torsal congruence surfaces are called *congruence developables*. If the line congruence \mathcal{K} is parametrized by $K(u)$, we can find congruence developables $R(t) = K(u(t))$ by integrating the field of torsal directions.

Parabolic Congruences

Here we report briefly on some ‘classification’ results concerning line congruences with only parabolic lines, which are called *parabolic congruences*.

There are essentially two types of parabolic congruences: The first type is constructed from a surface Φ in P^3 , which has only hyperbolic surface points. Locally there are two well-defined families of asymptotic tangents. One family comprises a parabolic congruence \mathcal{K} . Its focal points are the points of Φ , and the focal planes are Φ 's tangent planes. The tangent surfaces of the asymptotic curves are the congruence developables of \mathcal{K} . This construction yields a non-degenerate line congruence only if the asymptotic curves themselves are no straight lines.

The second type consists of lines in P^3 which touch a given surface Φ in the points of a curve $c \subset \Phi$. Here the set of focal points is c and Φ 's tangent planes in the points of c are the focal planes. The regular congruence developables are the pencils of surface tangents in the points of c . Singular congruence developables are c 's tangent surface and the developable surface tangent to Φ along c .

An example of the latter type is a *linear* parabolic congruence, which arises if Φ is a ruled surface and c is a non-torsal ruling (see Fig. 3.12, left).

It is possible to show that all regular parabolic congruences \mathcal{K} of sufficient smoothness, which are parametrized in the form (7.1) have the following property: All open subsets U of D contain an open subset U' such that $K(U')$ is of one of the two types described above.

Hyperbolic Congruences

Hyperbolic congruences, which have only hyperbolic lines, fall basically into the following three types:

The first kind of congruence consists of all lines which are tangent to two given surfaces Φ_1 and Φ_2 . These surfaces are the sets of focal points and at the same time the envelopes of the focal planes. They are therefore called *focal surfaces*. There are two families of torsal congruence surfaces: their curve of regression is contained in one focal surface, and their rulings are tangent to the second (see Fig. 7.1).

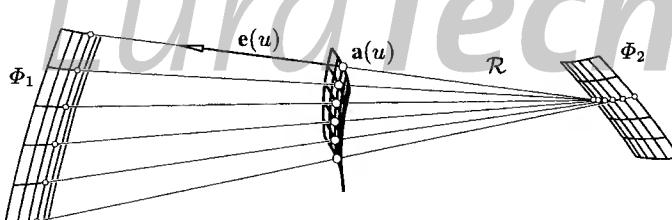


Fig. 7.1. Hyperbolic line congruence with directrix a (cf. Equ. (7.4)), focal surfaces Φ_1, Φ_2 , and congruence developable \mathcal{R} .

The other two types arise if one or both focal surfaces are replaced by a curve. An example of the latter is the hyperbolic linear congruence which has two focal lines. More examples will be given in Sec. 7.1.4, where we consider the normal congruence of a surface.

Elliptic Congruences

Elliptic congruences are those which have only elliptic lines. They contain no torsal ruled surfaces. However, if \mathcal{K} is real analytic, it may be extended to complex projective space. This extension $\mathcal{K}_{\mathbb{C}}$ is clearly hyperbolic — if the zero set of a quadratic form q in \mathbb{R}^2 is only the zero vector, then the zero set of its complex extension $q_{\mathbb{C}}$ consists of two conjugate complex focal lines.

The set of focal points of $\mathcal{K}_{\mathbb{C}}$ is invariant under complex conjugation, which means that we have two basic types of surfaces: (i) \mathcal{K} is the set of real lines L whose complex extensions $L_{\mathbb{C}}$ are tangent to two conjugate complex surfaces $\Phi, \bar{\Phi}$, and (ii) the same with two curves c, \bar{c} instead of $\Phi, \bar{\Phi}$.

An example is an elliptic linear congruence, which consists of all real lines which meet a pair of conjugate complex lines.

7.1.2 Rational Congruences and Trivariate Bézier Representations

Congruences in Affine Space and Computation of Focal Points

A two-parameter family $K(u^1, u^2)$ of lines of affine 3-space may be described by the following parametric representation:

$$\begin{aligned}\mathbf{a}, \mathbf{e} : D \subset \mathbb{R}^2 &\rightarrow \mathbb{R}^3, \quad \mathbf{x} : D \times \mathbb{R} \rightarrow \mathbb{R}^3, \\ \mathbf{x}(u^1, u^2, v) &= \mathbf{a}(u^1, u^2) + v\mathbf{e}(u^1, u^2).\end{aligned}\tag{7.4}$$

The line $K(u^1, u^2)$ is incident with the point $\mathbf{a}(u^1, u^2)$ and is parallel to $\mathbf{e}(u^1, u^2)$. The set of lines $K(u^1, u^2)$ is a line congruence \mathcal{K} , and \mathcal{K} is the set of v -parameter lines of $\mathbf{x}(u^1, u^2, v)$ (cf. Fig. 7.1).

This parametrization even describes a two-parameter family of *oriented* lines. The surface $\mathbf{a}(u^1, u^2)$ is called *director surface* or *directrix*. Clearly for all functions $\lambda : D \rightarrow \mathbb{R}$ the surface

$$\mathbf{a}'(u^1, u^2) = \mathbf{a}(u^1, u^2) + \lambda(u^1, u^2)\mathbf{e}(u^1, u^2)\tag{7.5}$$

is again a director surface and \mathbf{a}', \mathbf{e} describe the same line congruence.

The following lemma shows how to compute the focal points of a line in \mathcal{K} . Differentiation with respect to the three parameters u^1, u^2 , and v is denoted by the symbols $\mathbf{x}_{,1}, \mathbf{x}_{,2}$, and $\mathbf{x}_{,v}$, respectively.

Lemma 7.1.4. *Assume that a line congruence \mathcal{K} is determined by a parametrization of the form (7.4). The point $\mathbf{x}(u^1, u^2, v)$ is a focal point of the line $K(u^1, u^2)$, if*

$$\det(\mathbf{x}_{,1}, \mathbf{x}_{,2}, \mathbf{x}_{,v}) = 0.\tag{7.6}$$

Proof. By Lemma 7.1.3, for a focal point P contained in the line $K(u_0^1, u_0^2)$ there exists a regular curve $(u^1(t), u^2(t))$ in the parameter domain D , which assumes the value (u_0^1, u_0^2) for $t = 0$, such that the ruled surface $\mathcal{R}(t) = K(u^1(t), u^2(t))$ has a torsal generator at $t = 0$, and P is its cuspidal point. This ruled surface, parametrized by $\mathbf{x}(u^1(t), u^2(t), v)$ has a singularity in P .

This happens if and only if the three partial derivative vectors of \mathbf{x} are linearly dependent, i.e., (7.6) is fulfilled. \square

Remark 7.1.1. We consider (7.4) as a one-parameter family of surfaces $\mathbf{x}(u^1, u^2, v)$ with v fixed. If condition (7.6) is fulfilled for some value (u_0^1, u_0^2, v_0) , then the surface $v = v_0$ is tangent to the congruence line $u^1 = u_0^1, u^2 = u_0^2$ in the point $\mathbf{x}(u_0^1, u_0^2, v_0)$. This is clear because the tangent plane to the surface $v = v_0$ is spanned by $\mathbf{x}_{,1}$ and $\mathbf{x}_{,2}$.

On the other hand, it is well known that the envelope of a one-parameter family of surfaces $\mathbf{x}(u^1, u^2, v)$, v being the family parameter, is found by solving (7.6). This means that the focal surface (the surface of all focal points) is the envelope of the above mentioned one-parameter family of surfaces, and the lines of the congruence are tangent to the focal surface. \diamond

If we insert (7.4) in (7.6), we get the following equation, whose solutions v correspond to focal points:

$$\begin{aligned} &v^2 \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}) + v [\det(\mathbf{e}_1, \mathbf{a}_2, \mathbf{e}) + \\ &+ \det(\mathbf{a}_1, \mathbf{e}_2, \mathbf{e})] + \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}) = 0. \end{aligned} \quad (7.7)$$

For a line $K(u^1, u^2)$ of \mathcal{K} this is, in general, a quadratic equation, but it may be linear or even a contradictory equation of degree zero. Depending on the number of solutions, there are the following possibilities:

If all coefficients of (7.7) vanish, we have, of course, a torsal line of \mathcal{K} . If (7.7) is quadratic without solutions, $K(u^1, u^2)$ has no focal points. Thus $K(u^1, u^2)$ is an elliptic congruence line.

For a torsal congruence line, all directions are torsal. The focal point depends on the direction in a continuous manner, so it is not possible that there are exactly two focal points. This shows that in the case of two solutions of (7.7), the line $K(u^1, u^2)$ is a hyperbolic congruence line.

If there is exactly one solution, $K(u^1, u^2)$ has exactly one focal point, and is therefore not elliptic. It is also easy to show that the two focal points of a hyperbolic congruence line cannot coincide, so the case of one solution corresponds to a parabolic or torsal line $K(u^1, u^2)$.

Remark 7.1.2. The remaining two cases are (i) the coefficient of v^2 vanishes, and the equation becomes linear, and (ii) both the coefficients of v^2 and v vanish. In both cases there is a focal point at infinity. In case (i) the line is hyperbolic, in case (ii) it is parabolic or torsal. \diamond

We see that the computation of the focal surfaces amounts to a contouring problem in (u^1, u^2, v) -parameter space. We have to compute the zero contour for the trivariate function on the left hand side of equation (7.7). To do this numerically, we may apply standard contouring algorithms like ‘marching cubes’ [112]. However, the special structure of the present function suggests to use the following alternative: For a sufficiently dense grid of points in the (u^1, u^2) -parameter domain, we compute the corresponding focal points $\mathbf{x}(u^1, u^2, v_1)$ and $\mathbf{x}(u^1, u^2, v_2)$ (if they exist) and then compute B-spline approximants to the functions $v_1(u^1, u^2)$ and $v_2(u^1, u^2)$ (if they exist). We may have to perform a segmentation of the parameter domain into parts where \mathcal{K} is hyperbolic, elliptic, or parabolic. The parabolic region usually is the common boundary of the elliptic and hyperbolic parts. These ‘parabolic curves’ in the parameter domain correspond to congruence surfaces which separate the hyperbolic and elliptic parts of \mathcal{K} .

Line Congruences via Trivariate Bézier Representations

We call a line congruence *rational*, if it is represented by (7.4) with rational functions \mathbf{e} and \mathbf{a} . For the sake of brevity and simplicity, we confine ourselves to *polynomial congruences*, which means that \mathbf{e} and \mathbf{a} are polynomial functions. The extension to homogeneous representations and general rational congruences is straightforward.

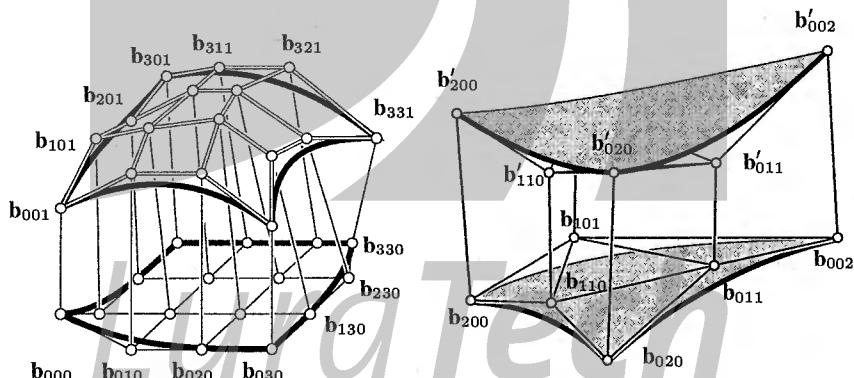


Fig. 7.2. Left: Line congruence defined by a trivariate tensor product Bézier representation of degree $(3, 3, 1)$. Right: Line congruence defined by a pentahedral Bézier solid of degree two.

A trivariate tensor product Bézier function represents a polynomial line congruence:

$$\mathbf{x}(u^1, u^2, v) = \sum_{k=0,1} \sum_{i=0}^m \sum_{j=0}^n B_k^1(v) B_i^m(u^1) B_j^n(u^2) \mathbf{b}_{ijk}. \quad (7.8)$$

\mathbf{x} is linear in the parameter v . Restriction of (7.8) to $v = 0$ gives the surface \mathbf{a} , evaluation at $v = 1$ gives $\mathbf{a} + \mathbf{e}$, which are both (m, n) -tensor product surfaces. Representations of this type, especially the natural extension to B-splines, provide a sufficiently rich class of congruences which is useful for practical computations and approximation problems for line congruences.

If $\mathbf{x}(u^1, u^2, v)$ is a polynomial function, then the left hand side of the focal point equation (7.7) is polynomial. It can even be made polynomial, without loss of generality, if $\mathbf{x}(u^1, u^2, v)$ is rational (by multiplication with a common denominator). This shows that the focal surfaces are the image of certain algebraic surfaces in (u^1, u^2, v) -parameter space. The knowledge of this fact is useful in applications. For instance we can employ special algorithms like the one of G. Elber proposed in [42], which uses B-spline subdivision and convex hull properties.

Remark 7.1.3. Of course we can evaluate (7.8) for arbitrary real values of u^1 , u^2 , and v .

On the other hand the Bézier representation is often used if the interesting part of the domain requires evaluation of the Bernstein polynomials only in the interval $[0, 1]$. The image of the cube $[0, 1]^3 \subset \mathbb{R}^3$ under the mapping (7.8) is called a Bézier solid.

If \mathbf{x} is regular (i.e., $\det(\mathbf{x}_{11}, \mathbf{x}_{12}, \mathbf{x}_{1v}) \neq 0$) and one-to-one, the boundary of the Bézier solid consists of 6 Bézier surfaces, the images of the faces of the parameter cube. In our case four of them are ruled surfaces.

If the restriction of \mathbf{x} to the cube $[0, 1]^3$ is not regular, part of the focal surface of the corresponding line congruence may be part of the boundary of its image. For an efficient computation of this nontrivial part of the boundary, which is an envelope of parameter surfaces defined by condition (7.6), we refer the reader to Joy and Duchaineau [87]. \diamond

A different interpretation of the tensor product representation (7.8) of a polynomial congruence is the following: We have two Bézier tensor product surfaces (see Fig. 7.2, left) and connect corresponding points. Analogously, we may prescribe two surfaces in triangular Bézier form and connect points with the same parameter values. This gives the so-called *pentahedral Bézier representation* [78]: Assume that $\mathbf{s}(\mathbf{u})$ and $\mathbf{s}'(\mathbf{u})$ are two triangular Bézier surfaces defined by control points $\mathbf{b}_{i_0 i_1 i_2}$ and $\mathbf{b}'_{i_0 i_1 i_2}$, according to Equ. (1.111). \mathbf{u} stands for a point in the two-dimensional parameter domain (which enters the generalized Bernstein polynomials not directly, but via its barycentric coordinates, see Equ. (1.110)). Then the pentahedral Bézier representation is given by

$$\mathbf{x}(\mathbf{u}, v) = v\mathbf{s}(\mathbf{u}) + (1 - v)\mathbf{s}'(\mathbf{u})$$

(cf. Fig. 7.2, right). We will illustrate this in a simple case.

Example 7.1.2. We consider two *linear* triangular Bézier surfaces \mathbf{a} and \mathbf{b} , which are just linear parametrizations of planes. We do not use the notation of (1.111), but denote the control points of \mathbf{a} and \mathbf{b} by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$.

We choose a base triangle in the (u^1, u^2) parameter domain and use barycentric coordinates (r, s, t) with respect to this triangle (cf. Equ. (1.109)). Then \mathbf{a} and \mathbf{b} are parametrized as follows:

$$\mathbf{a}(r, s, t) = r\mathbf{a}_1 + s\mathbf{a}_2 + t\mathbf{a}_3, \quad \mathbf{b}(r, s, t) = r\mathbf{b}_1 + s\mathbf{b}_2 + t\mathbf{b}_3, \quad (r + s + t = 1).$$

The mapping $\mathbf{a}(r, s, t) \mapsto \mathbf{b}(r, s, t)$ is an affine mapping. We do not restrict evaluation to the interior of the two triangles, i.e., we admit negative r, s, t as well. The congruence \mathcal{K} is described by the parametrization

$$\mathbf{x}(r, s, t, v) = (1 - v)\mathbf{a}(r, s, t) + v\mathbf{b}(r, s, t) \quad (r + s + t = 1). \quad (7.9)$$

For a fixed value of v , the points $\mathbf{x}(r, s, t, v)$ are contained in a plane $U(v)$, which

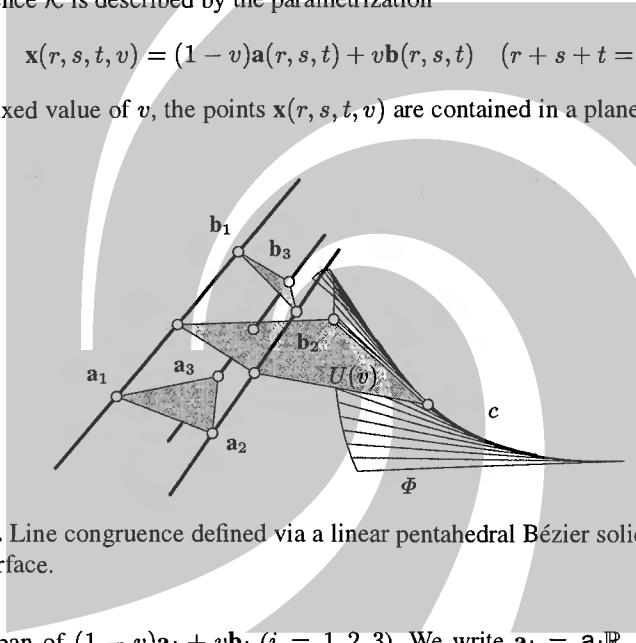


Fig. 7.3. Line congruence defined via a linear pentahedral Bézier solid with part of focal surface.

is the span of $(1 - v)\mathbf{a}_i + v\mathbf{b}_i$ ($i = 1, 2, 3$). We write $\mathbf{a}_i = \mathbf{a}_i \mathbb{R}$, $\mathbf{b}_i = \mathbf{b}_i \mathbb{R}$ in homogeneous coordinates and compute the homogeneous plane coordinates $\mathbb{R}u$ of the plane $U(v)$:

$$u = ((1 - v)\mathbf{a}_1 + v\mathbf{b}_1) \times ((1 - v)\mathbf{a}_2 + v\mathbf{b}_2) \times ((1 - v)\mathbf{a}_3 + v\mathbf{b}_3). \quad (7.10)$$

Thus $u(v)$ is a polynomial function in v , of degree at most three.

The envelope of the planes $U(v)$ coincides with the focal surface of \mathcal{K} . This can be shown by using the results on rational developable surfaces derived in Chap. 6. The focal surface is in general of degree four as an algebraic surface. After a change of coordinates $U(v)$ is parametrized by (6.9). By (6.12), the curve of regression of the focal surface is cubic. After the change of parameters $v \leftrightarrow 1/v$ in (7.10) we see that the ideal plane is contained in the family $U(v)$, so the curve of regression is a *polynomial cubic*.

It is interesting that we get only one focal surface instead of two. All lines which carry two focal points touch the focal surface *twice*. What are locally two sheets connects to an irreducible algebraic surface. \diamond

Example 7.1.3. (Continuation of Ex. 7.1.2) There are some special cases of pentahedral Bézier representations:

- If for some $v = v_0$ the three points in Equ. (7.10) are contained in a line F , then all lines of the congruence meet the line F . Thus F is a focal *line* — the focal surface is degenerate.
For $v = v_0$, the cross product (7.10) equals zero, so the polynomial (7.10) is divisible by $v - v_0$. If the remaining quadratic factor does not have a further zero over the complex number field, it represents a parabolic cylinder Φ .
The complete focal surface of \mathcal{K} consists of Φ and the line F , which happens to be tangent to Φ .
- Collinearity of the points of (7.10) may occur for *two* real values $v = v_0$ and $v = v_1$. In this case there are two focal lines, and \mathcal{K} is a hyperbolic linear congruence (see Sec. 3.2.1).
- If the complex extension of (7.10) has two conjugate complex zeros, \mathcal{K} is an elliptic linear congruence.
- If (7.10) is divisible by $(v - v_0)^2$, \mathcal{K} is a parabolic linear congruence
- If the three points in (7.10) are identical, \mathcal{K} is a bundle of lines. \diamond

Ex. 7.1.2 and Ex. 7.1.3 can be carried out with rational Bézier representations instead of polynomial ones. We get line congruences whose lines connect points P and $P\kappa$, where κ is a projective mapping instead of an affine one. The following result is a straightforward generalization of the affine case:

Theorem 7.1.5. *If $\kappa : U \rightarrow U'$ is a projective mapping of a plane U onto a plane U' in P^3 , then the set of lines $\mathcal{K} = \{X \vee X\kappa \mid X \in U\}$ has, in general, a focal surface Φ which is a cubic developable.*

After a complex extension, the lines of κ are tangents of Φ , generically in two points.

Degenerate cases of Φ are: (i) a quadratic cone plus one of its tangents, (ii) two lines, (iii) one line, (iv) one point. In cases (ii)–(iv) \mathcal{K} is a linear congruence. In cases (ii) and (iii) κ leaves $U \cap U'$ invariant, \mathcal{K} is hyperbolic, elliptic, or parabolic. In case (iv) the points of $U \cap U'$ are fixed points of κ and \mathcal{K} is a bundle of lines.

Remark 7.1.4. If we restrict evaluation of (7.9) to $r, s, t \geq 0$ and $v \in [0, 1]$ ($r + s + t = 1$), we obtain a so-called linear pentahedral Bézier solid (see Fig. 7.3). Such spatial finite elements have been used by Barnhill et al. [10] and Bajaj and Xu [7] to describe a surface and some neighbourhood of it ('fat surface'). The applications concern the computation of interpolating functions defined on surfaces and reconstruction and deformation of fat surfaces. In these applications, regularity in the following sense is essential: There shall be no self-intersections of the solid (i.e., x is one-to-one) and the boundary shall consist exactly of the two planar triangles $a_1a_2a_3$, $b_1b_2b_3$ and the three quadrilateral faces $a_1a_2b_2b_1$, $a_2a_3b_3b_2$, $a_1a_3b_3b_1$, which are in general pieces of hyperbolic paraboloids (in the polynomial case) or general ruled quadrics (in the rational case). Regularity is tested by looking at the location of the focal points. \diamond

Remark 7.1.5. The following generalization of Th. 1.1.47 follows from Ex. 7.1.2 and Ex. 7.1.3: If G, H, K are lines of projective three-space, P is a point of G , and $\kappa : G \rightarrow H, \lambda : G \rightarrow K$ are projective isomorphisms, then the planes $P \vee P\kappa \vee P\lambda$ are (except for degenerate cases similar to Th. 7.1.5) tangent to a developable surface Φ of class three.

If κ and λ are *affine* mappings, this is the result of Ex. 7.1.2. The general case is a straightforward generalization. The three lines G, H, K are not distinguished among all double tangents of Φ . \diamond

The Envelope of a Two-Parameter Family of Congruent Cylinders of Revolution

In Sec. 5.4.4, we considered a one-parameter family of congruent cylinders of revolution. The axes of these cylinders form a ruled surface \mathcal{R} , and the envelope of the cylinders themselves is an offset surface of \mathcal{R} . For a *two-parameter* family of congruent cylinders of revolution, the axes are not a ruled surface, but a line congruence \mathcal{K} . The surface which is the envelope of all lines of a congruence is its focal surface (cf. Remark 7.1.1). The following result is mentioned without proof, which would require a more precise definition of ‘envelope’ and ‘offset’:

Proposition 7.1.6. *Consider a smooth two-parameter family of cylinders of revolution, such that all radii are equal to $d > 0$. The envelope of this family of cylinders is an offset at distance d of the focal surface of \mathcal{K} , where \mathcal{K} is the congruence of all axes of cylinders.*

Motivated by applications in NC milling, Bézier representations of a two-parameter family of congruent cylinders and the corresponding envelopes have recently been studied by Xia and Ge [219] with help of the dual representation of the cylinders.

7.1.3 Euclidean Differential Geometry of Line Congruences

We are going to describe some properties of line congruences which are invariants of Euclidean geometry. We first describe how to represent line congruences in Euclidean space and what a ‘geometric property’ of a line congruence is. This presentation follows [75].

A two-parameter family $K(u^1, u^2)$ of lines of Euclidean space is defined by a mapping of the form (7.4), where we now additionally assume $\mathbf{e}^2 = 1$. The mapping $K(u^1, u^2) \mapsto \mathbf{e}(u^1, u^2)$ is called *spherical mapping* of the congruence, and $\mathbf{e}(D)$ is the *spherical image* of \mathcal{K} .

The First and Second Fundamental Forms

We assume a line congruence \mathcal{K} parametrized by a director surface \mathbf{a} and vector field \mathbf{e} as in (7.4). Properties of this parametrization which are actually properties of \mathcal{K} must not change if we change the director surface according to (7.5) and if we replace \mathbf{e} by $-\mathbf{e}$.

In the following we use the Einstein convention on summation: if the same letter, say i , occurs both as subscript and as superscript in a product, this means summation for $i = 1, \dots, n$, where n is the dimension of the domain of the functions involved. Here $n = 2$. Following E. Kummer (1860, [102]), we give the following definition.

Definition. For a parametrization of the form (7.4), its first and second fundamental forms are defined as follows: If $\dot{u} = (u^1, \dot{u}^2)$, $\dot{v} = (\dot{v}^1, v^2)$ are two tangent vectors, then

$$(I) : g(\dot{u}, \dot{v}) = g_{ik} \dot{u}^i \dot{v}^k \quad \text{with} \quad g_{ik} = \mathbf{e}_{,i} \cdot \mathbf{e}_{,k}, \quad (7.11)$$

$$(II) : G(\dot{u}, \dot{v}) = G_{ik} \dot{u}^i \dot{v}^k \quad \text{with} \quad G_{ik} = -\mathbf{e}_{,i} \cdot \mathbf{a}_{,k}. \quad (7.12)$$

Remark 7.1.6. The first fundamental form g does not change if we change the director surface, because this surface does not enter the definition, and it does not change if we replace \mathbf{e} by $-\mathbf{e}$. Thus it is a geometric property of the congruence.

This is in contrast to the behaviour of the second fundamental form, which actually is no property of the congruence, but a property of the pair \mathbf{a}, \mathbf{e} . Both fundamental forms are invariant with respect to Euclidean congruence transformations. ◇

Note that g is the first fundamental form of the spherical image if we consider it as an ordinary two-surface in Euclidean space. It is therefore a symmetric positive semi-definite bilinear form. If the spherical image is regular (i.e., $\text{rk}(\mathbf{e}_{,1}, \mathbf{e}_{,2}) = 2$), then g is even positive definite. A ruling $K(u^1, u^2)$ of the congruence where \mathbf{e} is singular, will be called *cylindrical*. G is usually not symmetric.

Euclidean Invariants of Congruence Surfaces

We consider a line congruence \mathcal{K} parametrized by (7.4), and study first order Euclidean differential invariants of ruled surfaces contained in \mathcal{K} . They will be expressed by the first and second fundamental form of the parametrization and will eventually lead to geometric invariants of the congruence \mathcal{K} .

Assume a congruence surface $\mathcal{R}(t) = K(u(t))$. Its point set then is parameterized by

$$\mathbf{x}(t, v) = \mathbf{a}(u(t)) + v\mathbf{e}(u(t)).$$

We use Equ. (5.43) to compute its striction points $\mathbf{a}(u(t)) + v_s(t)\mathbf{e}(t)$

$$v_s = -\frac{\dot{\mathbf{e}} \cdot \dot{\mathbf{a}}}{\dot{\mathbf{e}}^2} = -\frac{\mathbf{e}_{,i} \dot{u}^i \cdot \mathbf{a}_{,k} \dot{u}^k}{(\mathbf{e}_{,i} \dot{u}^i)^2} = \frac{G_{ik} \dot{u}^i \dot{u}^k}{g_{ik} \dot{u}^i \dot{u}^k}. \quad (7.13)$$

We see that the striction point of a congruence surface depends only on the direction $\dot{u}^1 : \dot{u}^2$ (which must be so because of Lemma 7.1.2).

The central normal vector of the congruence surface is, according to (5.53), given by

$$\mathbf{e}_2 = \frac{\dot{\mathbf{e}}}{\|\dot{\mathbf{e}}\|} = \frac{\mathbf{e}_{,i} \dot{u}^i}{\sqrt{g_{ik} \dot{u}^i \dot{u}^k}}. \quad (7.14)$$

To compute the distribution parameter we use (5.45) and find

$$\delta = \frac{\det(\mathbf{a}_i \dot{u}^i, \mathbf{e}, \mathbf{e}_k \dot{u}^k)}{g_{ik} \dot{u}^i \dot{u}^k}. \quad (7.15)$$

If we let

$$\gamma = g_{11}g_{22} - g_{12}^2, \quad (7.16)$$

then $\sqrt{\gamma}\mathbf{e} = \mathbf{e}_{,1} \times \mathbf{e}_{,2}$, and $\mathbf{e} \times \mathbf{e}_{,1} = (-g_{12}\mathbf{e}_{,1} + g_{11}\mathbf{e}_{,2})/\sqrt{\gamma}$, $\mathbf{e} \times \mathbf{e}_{,2} = (-g_{22}\mathbf{e}_{,1} + g_{12}\mathbf{e}_{,2})/\sqrt{\gamma}$. We expand (7.15) using the identity $\det(\mathbf{a}_{,i}, \mathbf{e}, \mathbf{e}_{,k}) = \mathbf{a}_{,i} \cdot (\mathbf{e} \times \mathbf{e}_{,k})$ and see that

$$\delta = -\frac{1}{\sqrt{\gamma} g_{ik} \dot{u}^i \dot{u}^k} \begin{vmatrix} g_{1k} \dot{u}^k & g_{2k} \dot{u}^k \\ G_{1k} \dot{u}^k & G_{2k} \dot{u}^k \end{vmatrix}. \quad (7.17)$$

Equations (7.13), (7.15), and (7.14) depend only on the ratio $\dot{u}^1 : \dot{u}^2$. This means that ruled congruence surfaces which share the same direction also share their striction point, Sannia frame, and distribution parameter and are therefore in first order contact. This shows again Lemma 7.1.2.

The Central Surface of a Congruence

Assume a line congruence \mathcal{K} parametrized by (7.4). We choose the congruence line $u_0 = (u_0^1, u_0^2)$ and consider the set of possible striction points, which depend on a direction at u_0 . The *extremal positions* of the striction point on the line $K(u_0)$ are, by (7.13), given by the minimum and maximum of the quotient $G(\dot{u}, \dot{u})/g(\dot{u}, \dot{u})$.

This problem is formally equivalent to determining the minimum and maximum normal curvature in the point of a surface in Euclidean space, with the difference that the matrix G_{ik} need not be symmetric. This obstacle is easily removed by letting

$$\bar{G}_{ik} = \frac{1}{2}(G_{ik} + G_{ki}). \quad (7.18)$$

Obviously $\bar{G}_{ik} \dot{u}^i \dot{u}^k = G_{ik} \dot{u}^i \dot{u}^k$. Let us abbreviate the matrices (g_{ik}) and (\bar{G}_{ik}) by g and \bar{G} . Then

$$\begin{aligned} \text{(I)} : \quad g(\dot{u}, \dot{u}) &= \dot{u}^T \cdot g \cdot \dot{u}, \\ \text{(II)} : \quad G(\dot{u}, \dot{u}) &= \dot{u}^T \cdot \bar{G} \cdot \dot{u}. \end{aligned}$$

We transform the problem of finding the extremal values of G/g to the problem of finding the extremal values of $\dot{u}^T \cdot G \cdot \dot{u}$ subject to the constraint $\dot{u}^T \cdot g \cdot \dot{u} = 1$. With a Lagrangian multiplier λ we have to solve

$$(\bar{G} - \lambda g) \cdot \dot{u} = 0, \quad \dot{u}^T \cdot g \cdot \dot{u} = 1. \quad (7.19)$$

Since the left hand equation is homogeneous, there is a nontrivial solution if and only if $\det(\bar{G} - \lambda g) = 0$. If the generator $K(u_0)$ of the congruence is non-cylindrical and therefore g is regular, this is equivalent to

$$\det(\bar{G} \cdot g^{-1} - \lambda I) = 0. \quad (7.20)$$

This characteristic equation of the symmetric matrix $M = \bar{G} \cdot g^{-1}$ has two (possibly equal) real solutions v_1, v_2 for λ . The case that $v_1 = v_2$ of a so-called *isotropic congruence line* shall be excluded here and will be addressed later.

If $\lambda = v_i$ is an eigenvalue, and \dot{u}_i the associated eigenvector which satisfies $(\bar{G} - v_i g) \cdot \dot{u}_i = 0$, then the corresponding v -value of the striction point equals

$$v = \frac{G(\dot{u}, \dot{u})}{g(\dot{u}, \dot{u})} = \frac{\dot{u}_i^T \cdot \bar{G} \cdot \dot{u}_i}{\dot{u}_i^T \cdot g \cdot \dot{u}_i} = \frac{\dot{u}_i^T \cdot v_i g \cdot \dot{u}_i}{\dot{u}_i^T \cdot g \cdot \dot{u}_i} = v_i. \quad (7.21)$$

We have shown that the extremal striction points $\mathbf{a}(u_0) + v_i \mathbf{e}(u_0)$ are given by the eigenvalues v_i of the matrix $\bar{G} \cdot g^{-1}$. The midpoint

$$\mathbf{c} = \mathbf{a} + \frac{1}{2}(v_1 + v_2)\mathbf{e}$$

of the two extremal striction points is called *central point* of the congruence line. The sum of eigenvalues equals twice the trace of $\bar{G} \cdot g^{-1}$, so

$$v_c = \frac{1}{2} \operatorname{tr}(\bar{G} \cdot g^{-1}) = \frac{1}{2} \bar{G}_{ik} g^{ik}, \text{ with } (g^{ik}) = g^{-1} = \frac{1}{\gamma} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \quad (7.22)$$

The surface of the central points,

$$\mathbf{c}(u^1, u^2) = \mathbf{a}(u^1, u^2) + v_c(u^1, u^2)\mathbf{e}(u^1, u^2),$$

is called *central surface* of the line congruence. Fig. C.15 shows an example.

Principal Directions

We continue the discussion of extremal striction points. The directions \dot{u}_i which correspond to the extremal values of v_s are called *principal*. They are found as eigenvectors of the matrix $\bar{G} \cdot g^{-1}$ (cf. (7.21)). A ruled surface $R(u(t))$ in the congruence $K(t)$ which has the property that \dot{u} indicates a principal direction, is called a *principal surface*. Such surfaces play the role of principal curvature lines. There are the following results, which describe this analogy.

Proposition 7.1.7. *Assume that L is a nonisotropic and non-cylindrical line in the C^2 congruence K parametrized by (7.4), and that $R(t) = K(u(t))$ is a ruled congruence surface of K . The striction point of $R(t)$ assumes its extremal position in L if $\dot{u}^1 : \dot{u}^2$ is a principal direction.*

The central normals of two surfaces $R_i(t)$ corresponding to the two principal directions are orthogonal to each other.

Proof. We have already discussed the problem of extremal position $v_s = G(\dot{u}, \dot{u}) / g(\dot{u}, \dot{u})$ of the striction point (cf. Equ. (7.19)).

It is well known that the solution vectors \dot{u}_1, \dot{u}_2 of the generalized eigenvalue problem $(\bar{G} - \lambda g)\dot{u} = 0$ with symmetric matrices \bar{G}, g are g -orthogonal in the sense that $\dot{u}_1^i \dot{u}_2^k g_{ik} = 0$. By (7.14), the two corresponding central normals $\mathbf{e}_2^{(1)}, \mathbf{e}_2^{(2)}$ are orthogonal: $(\dot{u}_1^i \mathbf{e}_{,i}) \cdot (\dot{u}_2^k \mathbf{e}_{,k}) = \dot{u}_1^i \dot{u}_2^k g_{ik} = 0$. \square

The central planes corresponding to the two orthogonal central normals mentioned in Prop. 7.1.7 are called *principal planes*.

Theorem 7.1.8. *If L is a nonisotropic and non-cylindrical line in the C^3 congruence \mathcal{K} , then in a neighbourhood of L there exists a parametrization of \mathcal{K} in the form (7.4) such that the surfaces with equation $u^1 = \text{const}$ and also the surfaces with equation $u^2 = \text{const}$ are principal surfaces.*

Proof. (Sketch) The condition that $(\bar{G} - \lambda g) \cdot \dot{u} = 0$ is equivalent to the linear dependence of the vectors $\bar{G} \cdot \dot{u}$ and $g \cdot \dot{u}$, which means that

$$\det(\bar{G} \cdot \dot{u}, g \cdot \dot{u}) = 0. \quad (7.23)$$

This is a quadratic homogeneous differential equation. In the neighbourhood of a non-cylindrical congruence line it splits into the product of two linear homogeneous differential equations and there are *two* families of regular integral curves, which lead to two families of principal surfaces. It is further possible to show that there is a local change of parameters such that the integral curves of (7.23) become the parameter lines. This is completely analogous to a parametrization of a surface in Euclidean space whose parameter lines are the principal curvature lines. \square

If we have a *principal parametrization* as described by Th. 7.1.8, the principal directions are given by $\dot{u}_1 = (1, 0)$ and $\dot{u}_2 = (0, 1)$. Such a parametrization is therefore characterized by

$$g_{12} = 0, \quad \bar{G}_{12} = 0 \quad (\iff G_{12} + G_{21} = 0). \quad (7.24)$$

If we compute the central normals which correspond to the directions $(1, 0)$ and $(0, 1)$ by Equ. (7.14), we get

$$\mathbf{e}_2^{(1)} = \mathbf{e}_{,1}/\sqrt{g_{11}}, \quad \mathbf{e}_2^{(2)} = \mathbf{e}_{,2}/\sqrt{g_{22}}. \quad (7.25)$$

Example 7.1.4. We compute the principal surfaces of an elliptic linear congruence with rotational symmetry: We start with the parametrization of the form $\mathbf{x}(u^1, u^2, v) = \mathbf{a}(u^1, u^2) + v\mathbf{e}(u^1, u^2)$ and let

$$\mathbf{a}(u^1, u^2) = \begin{bmatrix} u^1 \cos u^2 \\ u^1 \sin u^2 \\ 0 \end{bmatrix}, \quad \mathbf{e}(u^1, u^2) = \frac{1}{\sqrt{1 + (u^1)^2}} \begin{bmatrix} u^1 \sin u^2 \\ -u^1 \cos u^2 \\ 1 \end{bmatrix}.$$

Comparison with Equ. (3.15) shows that the union of lines $\mathbf{a}(u^1, u^2) + v\mathbf{e}(u^1, u^2)$ indeed coincides with an elliptic linear congruence without its line at infinity. Equations (7.11), (7.12), and (7.18) show how to compute the matrices g and \bar{G} , which are symmetric. Their nonzero coefficients are

$$g_{11} = \frac{1}{(1 + (u^1)^2)^2}, \quad g_{22} = \frac{(u^1)^2}{1 + (u^1)^2}, \quad \bar{G}_{12} = -\frac{1}{2} \frac{(u^1)^3}{(1 + (u^1)^2)^{3/2}}.$$

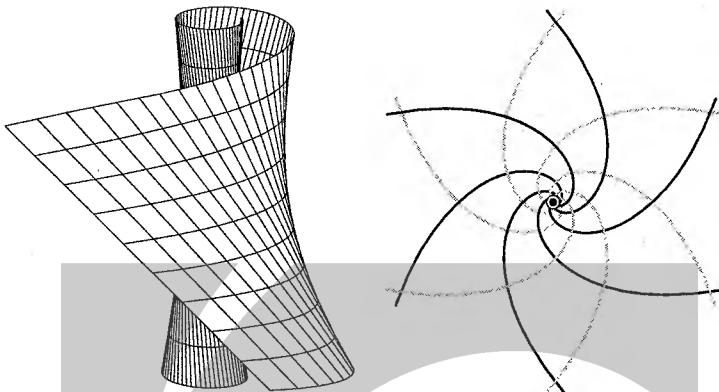


Fig. 7.4. Left: Principal surface of an elliptic linear congruence. Right: Planar sections of principal surfaces.

The differential equation (7.23) of principal directions is integrable:

$$\dot{u}^2 = \pm \frac{\dot{u}^1}{u^1 \sqrt{1 + (\dot{u}^1)^2}} \implies u^2 = \mp \operatorname{artanh} \frac{1}{\sqrt{1 + (\dot{u}^1)^2}} + C,$$

We let $C = 0$, $u = u^2$, choose the plus sign, and insert into $\mathbf{x}(u^1, u^2, v)$ to get the following parametrization of one particular principal surface:

$$\mathbf{x}(u, v) = \frac{1}{\sinh u} \begin{bmatrix} \cos u \\ \sin u \\ 0 \end{bmatrix} + v \begin{bmatrix} \sin u / \sinh u \\ -\cos u / \sinh u \\ \tanh u \end{bmatrix}.$$

This surface is shown in Fig. 7.4, left. Fig. 7.4, right, shows the section $x_3 = 0$ of some principal surfaces of both families. These curves are spirals which have the equation $r(\phi) = 1 / \sinh(\pm\phi + C)$ in polar coordinates. \diamond

Torsal Directions

Like in the previous discussion of the principal directions in a line congruence \mathcal{K} , we assume a parametrization in the form (7.4).

The ruled surfaces which correspond to *torsal* directions $(\dot{u}^1 : \dot{u}^2)$ have vanishing distribution parameter. We use Equ. (7.17), to show that the column vector $\dot{u} = (\dot{u}^1 : \dot{u}^2)$ must satisfy a relation of the form

$$\det(G \cdot \dot{u}, g \cdot \dot{u}) = 0 \implies G \cdot \dot{u} = \lambda g \cdot \dot{u}. \quad (7.26)$$

We have seen earlier (see Remark 5.3.5) that Equ. (7.13), which computes the v -value of the striction point for a non-torsal generator, computes the cuspidal point of a torsal one. The difference is that the result may be undefined in the case of a

cuspidal point at infinity. So we can use (7.13) to compute the finite *focal points* $\mathbf{a} + v\mathbf{e}$ of torsal directions. We get

$$v = \frac{\dot{u}^T \cdot G \cdot u}{\dot{u}^T \cdot g \cdot \dot{u}} = \lambda \implies G \cdot \dot{u} = vg \cdot \dot{u}.$$

This is equivalent to

$$\det(G \cdot g^{-1} - vI) = 0. \quad (7.27)$$

Obviously there are (up to a scalar factor) two nonzero solutions of (7.27) for a hyperbolic congruence line, one for a parabolic one, none for an elliptic one, and in case of a torsal congruence line, (7.27) is fulfilled for all v .

If we have two finite focal points, we can speak of their midpoint. Its v -value is given by

$$\frac{1}{2} \operatorname{tr}(G \cdot g^{-1}) = \frac{1}{2} \operatorname{tr}(\bar{G} \cdot g^{-1}). \quad (7.28)$$

Equations (7.28) and (7.22) show that

Lemma 7.1.9. *For a hyperbolic congruence line, the midpoint of the focal points coincides with the central point.*

Example 7.1.5. With Lemma 7.1.9 we can easily find the central surface of a linear hyperbolic congruence with focal lines F_1 and F_2 : The set of midpoints $(\mathbf{a}_1 + \mathbf{a}_2)/2$ with $\mathbf{a}_i \in F_i$ is a plane parallel to both focal lines. Another example is shown by Fig. C.15. \diamond

If one focal point is at infinity, (7.13) shows that there exists a direction \dot{u} with $g \cdot \dot{u} = 0$, which means that g is not regular and the congruence line under consideration is cylindrical.

Examples — Spherical Tractrices

Example 7.1.6. We consider a line congruence \mathcal{K} whose focal surfaces are a *quadric* Φ and a *conic*, parametrized in the form $c(t)$. \mathcal{K} consists of all lines which meet c and which touch Φ .

The congruence developables in \mathcal{K} are (i) quadratic cones Δ_t with vertex $c(t)$ and tangent to Φ , and (ii) tangent surfaces Ψ_t of curves $k \subset \Phi$ whose rulings meet c . By Prop. 1.1.27, Δ_t is tangent to Φ in the points of a conic l_t . This is illustrated in Fig. 7.5, left with a sphere Φ and a circle c .

We apply a projective transformation κ such that $c\kappa = c_\omega$, where c_ω is the ideal conic of Remark 6.3.5, consisting of the ideal points of lines of constant slope. Hence the curves $k\kappa$ are of constant slope (cf. Ex. 6.3.3 and Ex. 6.3.18). Such curves have been studied in detail by W. Wunderlich [209]. Because of this property, curves k in Φ whose tangents meet c are called *projective curves of constant slope* (with respect to c).

Consider points $\mathbf{q} \in \Phi$, $\mathbf{p} = c(t)$, such that $L = \mathbf{q} \vee \mathbf{p}$ is in \mathcal{K} . The ruling L of Δ_t is, by Lemma 6.1.5, conjugate to l_t 's tangent in the point \mathbf{q} . As L is tangent to k

in \mathbf{q} , this shows that the curves k and the curves l_t have conjugate tangents wherever they meet. \diamond

Remark 7.1.7. The conjugacy mentioned in Ex. 7.1.6 is valid for all hyperbolic congruences. For a focal surface, consider (i) the curves whose tangent surfaces are congruence developables, and (ii) the curves where the other congruence developables (whose lines of regression are contained in the other focal surface) touch this focal surface. If a curve of type (i) meets a curve of type (ii), they have conjugate tangents in the common point. \diamond

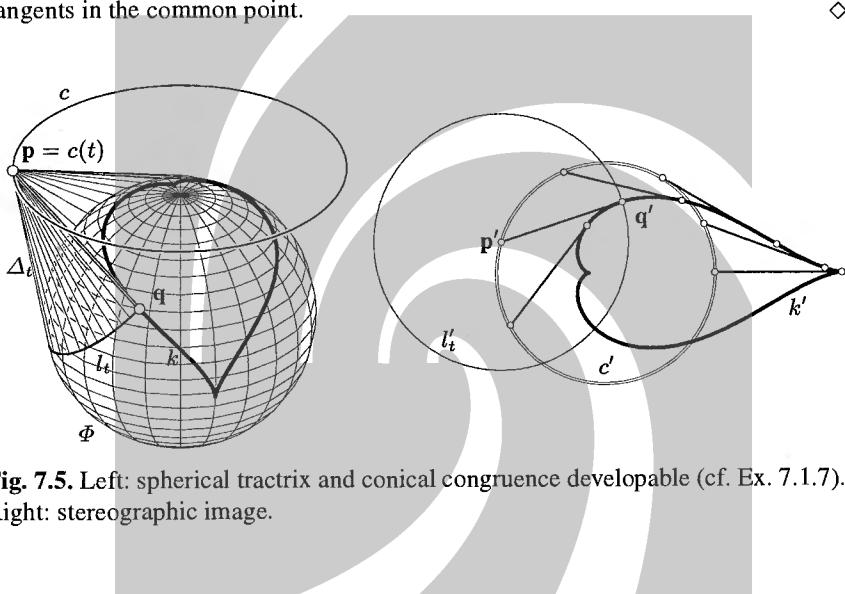


Fig. 7.5. Left: spherical tractrix and conical congruence developable (cf. Ex. 7.1.7). Right: stereographic image.

Example 7.1.7. (Continuation of Ex. 7.1.6) We assume that Φ is the unit sphere of a Cartesian coordinate system and c is a circle with center $(0, 0, z_0)$ and radius r , whose axis is the z -axis. Both the system of curves l_t and the system of curves k have rotational symmetry about the z -axis.

The distance $\overline{\mathbf{p}\mathbf{q}}$ (see Fig. 7.5, left) obviously is constant if \mathbf{q} moves in l_t and \mathbf{p} is fixed. Because it remains constant if we rotate about the z -axis, $\overline{\mathbf{p}\mathbf{q}}$ is constant for all \mathbf{p}, \mathbf{q} . This shows that the segment of k 's tangent between k and c has constant length if we move along k . The curve k is therefore a *spherical tractrix*.

Conjugacy of a sphere's tangents is orthogonality, so the circles l_t and the curves k are orthogonal. If we project c radially onto Φ we get a circle \bar{c} , containing the spherical centers of the circles l_t . Thus the curves k are also tractrices in the sense of spherical geometry.

We compute the curves k by applying the projective mapping

$$\kappa : x' = x/(\varepsilon z - 1), y' = y/(\varepsilon z - 1), z' = z/(\varepsilon z - 1) \quad (\varepsilon = 1/z_0).$$

It is verified by computation that $c\kappa = c_\omega$ and that $\Phi\kappa$ has the equation $x'^2 + y'^2 + z'^2 = (\varepsilon z' - 1)^2$. $\Phi\kappa$ is an ellipsoid, or paraboloid, or two-sheeted hyperboloid, if $|z_0| > 1$, $|z_0| = 1$, or $|z_0| < 1$, respectively.

For one special case of z_0 , the curves $k\kappa$ coincide with the curves of Ex. 6.3.3. The corresponding curves k are shown in Fig. 7.5, left. The reader may try to verify that horizontal planar sections of k 's tangent surface are offsets of a nephroid. For more information on spherical tractrices and other remarkable examples of projective curves of constant slope we refer to W. Wunderlich [214].

Finally we mention that the central surface of \mathcal{K} is a sphere concentric with Φ . \diamond

Remark 7.1.8. We use the notation of Ex. 7.1.6 and Ex. 7.1.7. Central projection (denoted by a prime) from $(0, 0, 1)$ onto $z = 0$ is the stereographic projection when restricted to the unit sphere. It preserves rotational symmetry and orthogonality of l_t and k (see Fig. 7.5, right). The rulings of Δ_t are mapped to lines which intersect in p' and are orthogonal to l'_t , so p' is the center of l'_t . This shows that the tangential distance $p'q'$ is constant, and the curves k' are planar tractrices of the circle c' [210]. The curve k' shown in Fig. 7.5, right is rational of degree six. \diamond

Isotropic Congruences

A line congruence is called *isotropic*, if there exists $\lambda \in \mathbb{R}$ such that $g = \lambda \bar{G}$. In this case principal directions are undefined. By (7.13), all ruled surfaces which contain the same congruence line have the same striction point there. This striction point obviously coincides with the central point.

We choose the central surface as directrix. Then (7.13) must result in $v = 0$ for all \dot{u} , and \bar{G} must be the zero matrix:

$$G_{11} = G_{22} = 0, \quad G_{12} + G_{21} = 0. \quad (7.29)$$

To compute torsal directions \dot{u} , we expand (7.26) using (7.29) and get the equation

$$\dot{u}^T \cdot g \cdot \dot{u} = 0 \quad (G_{12} \neq 0). \quad (7.30)$$

We see that torsal directions do not exist if the congruence line under consideration is not cylindrical. An isotropic (non-cylindrical) congruence is therefore elliptic.

Normal Congruences

If the matrix G is symmetric, we have $G = \bar{G}$. Equations (7.23) and (7.26) show that every principal direction is a torsal direction and vice versa. By Equ. (7.20) and (7.27), limit points and focal points are the same. These congruences are called *normal congruences*. The name will be justified by Th. 7.1.10:

Theorem 7.1.10. A C^2 line congruence \mathcal{K} is the set of surface normals of a surface if and only if limit points and focal points are the same. If \mathcal{K} is parametrized by (7.4) this happens if and only if G is symmetric.

Proof. We assume that \mathcal{K} is parametrized by (7.4). By Equ. (7.20) and (7.27), limit and focal points are the same if and only if $G = \overline{G}$, which means that G is symmetric.

Assume now that G is symmetric. We want to find a surface $s = \mathbf{a} + v(u^1, u^2)\mathbf{e}$ orthogonal to the lines of \mathcal{K} , which means

$$0 = s_i \cdot \mathbf{e} = \mathbf{a}_{,i} \cdot \mathbf{e} + v_{,i} = \mathbf{a}_{,i} \cdot \mathbf{e} + v_{,i}, \quad i = 1, 2. \quad (7.31)$$

This system of partial differential equations is solvable if and only if the integrability condition $(\mathbf{a}_{,1} \cdot \mathbf{e})_{,2} = -v_{,12} = -v_{,21} = (\mathbf{a}_{,2} \cdot \mathbf{e})_{,1}$ is satisfied. Equ. (7.12) shows that the integrability condition is satisfied if

$$G_{12} = G_{21}. \quad (7.32)$$

If there is such a surface s , it is not unique, because adding a constant to the function v yields another orthogonal surface, all such surfaces being offsets of each other.

Conversely, if there is a surface orthogonal to the congruence lines, we can use (7.31) to show (7.32). \square

Remark 7.1.9. If the parametrization (7.4) is C^k ($k \geq 2$), then the functions $\mathbf{a}_{,i} \cdot \mathbf{e}$ in Equ. (7.31) are C^{k-1} , and so are $v_{,1}$ and $v_{,2}$. This means that v is C^k . An orthogonal trajectory surface s has therefore the same differentiability class as \mathbf{a} and \mathbf{e} , but in addition it has surface normals which are C^k . However it does not have to be regular. If it is regular, it is actually (after a suitable C^k change of parameters) a C^{k+1} surface.

Conversely, the family of surface normals of a regular C^{k+1} surface obviously is a C^k normal congruence, but so is the family of surface normals of a (not necessarily regular) C^k surface with C^k normals. \diamond

We use an orthogonal trajectory surface of the congruence lines as directrix. To emphasize that \mathbf{e} is the unit normal of \mathbf{a} , we write \mathbf{n} instead of \mathbf{e} . By differentiating the identity $\mathbf{a}_{,i} \cdot \mathbf{n} = 0$ we see that

$$\mathbf{a}_{,i} \cdot \mathbf{n}_{,k} + \mathbf{a}_{,ik} \cdot \mathbf{n} = 0, \quad (7.33)$$

which shows that

$$G_{ik} = -\mathbf{a}_{,i} \cdot \mathbf{n}_{,k} = \mathbf{a}_{,ik} \cdot \mathbf{n} = h_{ik}. \quad (7.34)$$

This means that the second fundamental form of the pair \mathbf{a}, \mathbf{n} coincides with the second fundamental form of the surface \mathbf{a} used in elementary differential geometry of surfaces in Euclidean E^3 (cf. Sec. 5.3).

For $\dot{u} = (\dot{u}^1, \dot{u}^2)$ we consider the corresponding derivative vectors $\dot{\mathbf{a}}$ and $\dot{\mathbf{n}}$. If \dot{u} determines a *torsal* direction, then Equ. (5.41) shows that $\{\dot{\mathbf{a}}, \dot{\mathbf{n}}, \mathbf{n}\}$ is linearly dependent.

Since both $\dot{\mathbf{a}}$ and $\dot{\mathbf{n}}$ are orthogonal to \mathbf{n} , this means that $\dot{\mathbf{a}} + \lambda \dot{\mathbf{n}} = \mathbf{o}$. This is a characterization of principal curvature directions of surfaces in Euclidean space (O. Rodrigues, 1815), and λ equals the corresponding principal curvature radius R_i :

$$\dot{\mathbf{a}} + R_i \mathbf{n} = \mathbf{0}, \quad i = 1, 2. \quad (7.35)$$

We use (5.43) to compute the focal points $\mathbf{a} + v_i \mathbf{n}$ of these torsal directions:

$$v_i = -\frac{\dot{\mathbf{a}} \cdot \mathbf{n}}{\mathbf{n}^2} = R_i. \quad (7.36)$$

Theorem 7.1.11. *If \mathcal{K} is the normal congruence defined by the surface \mathbf{a} , then the principal surfaces of \mathcal{K} are developable and consist of the surface normals along principal curvature lines of \mathbf{a} . The focal points of \mathcal{K} are the principal curvature centers of \mathbf{a} . The curve of regression of a developable surface \mathcal{R} in \mathcal{K} is contained in one focal surface, and \mathcal{R} touches the other one along a curve.*

Proof. The discussion above shows that \mathcal{K} 's torsal directions, \mathcal{K} 's principal directions, and \mathbf{a} 's principal directions coincide. Equ. (7.36) shows the statement about the focal points.

Assume a principal parametrization of \mathbf{a} (cf. Th. 7.1.8). Then one family of principal surfaces has the form $\mathbf{p}(u_1, v) = \mathbf{a}(u_1, \alpha) + v \mathbf{n}(u_1, \alpha)$ with some constant α . A focal surface is parametrized by $\mathbf{f}_i(u_1, u_2) = \mathbf{a}(u_1, u_2) + v_i(u_1, u_2) \mathbf{n}(u_1, u_2)$ with v_i from (7.36). It is an elementary exercise to verify that the surface \mathbf{p} is the tangent surface of the curve $\mathbf{c}(u_1) = \mathbf{p}(u_1, v_1(u_1, \alpha)) = \mathbf{f}_1(u_1, \alpha)$, and that the surfaces \mathbf{p} and \mathbf{f}_2 are tangent in the points $\mathbf{p}(u_1, v_2(u_1, \alpha)) = \mathbf{f}_2(u_1, \alpha)$. \square

Example 7.1.8. Figure C.15 shows a piece of a hyperbolic paraboloid and the corresponding parts of the central and focal surfaces of its normal congruence. \diamond

Example 7.1.9. Focal surfaces of normal congruences can be degenerate. If \mathcal{K} , defined by the surface \mathbf{a} , has a focal curve, a point \mathbf{x} of this focal curve corresponds to a set $\mathbf{a}(V)$ of surface points such that $\mathbf{a}(u) + R_i(u) \mathbf{n}(u) = \mathbf{x}$ with $i = 1$ or $i = 2$ for $u \in V$. Clearly the surface is orthogonal to the line $\mathbf{a}(u) - \mathbf{x}$ in such points and therefore on connected components of V the principal curvature R_i is constant. This means that the surface \mathbf{a} is the envelope of all possible spheres with center \mathbf{x} in the focal curve and corresponding radius R_i .

The envelope of a one-parameter family of spheres is called a *canal surface* (cf. Sec. 6.3.3 and Fig. 6.45). Its focal curve is the curve of centers of the spheres, which is also called spine curve. We see that if \mathbf{a} has a focal curve, it consists of canal surfaces. \diamond

Example 7.1.10. It is obvious how to generalize the notion of ‘focal curve’ to a curve at infinity. Arguments similar to those in Ex. 7.1.9 show that a surface \mathbf{a} which has a focal curve at infinity consists of surfaces which are envelopes of one-parameter families of planes, i.e., torsal ruled surfaces. \diamond

Example 7.1.11. The mapping $\mathbf{x} \mapsto \mathbf{m} + (\mathbf{x} - \mathbf{m})/\|\mathbf{x} - \mathbf{m}\|^2$ from Euclidean space E^n into E^n is called *inversion* with center \mathbf{m} .

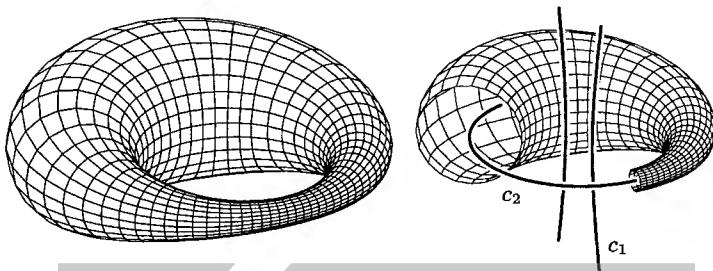


Fig. 7.6. Dupin cyclide with principal curvature lines (circles) and focal conics c_1 , c_2 .

The center of the inversion does not have an image. We can, however, extend the affine space by adding *one* point at infinity which corresponds to the center. This so-called conformal closure or conformal extension is different from the projective extension. It is an easy exercise to verify that a sphere or plane is mapped to a sphere or plane by an inversion. Within the conformal extension of Euclidean space we add the point at infinity to each plane and say that planes are spheres which contain the point at infinity. This is consistent with the properties of inversion.

The class of spheres and planes is distinguished among the connected C^2 smooth surfaces by the property that the focal surfaces of the normal congruence degenerates into a point. \diamond

Example 7.1.12. (Continuation of Ex. 7.1.11) We consider the class of surfaces which is generated by inversions from cylinders of revolution, cones of revolution, and tori. They are called *Dupin cyclides* (see Fig. 7.6).

It turns out that this class of surfaces, which has several remarkable properties, is precisely the class of surfaces which have *two focal curves*: All connected (in the sense of the parameter domain) C^2 surfaces with this property are pieces of Dupin cyclides or tori, cones, or cylinders of revolution.

A complete Dupin cyclide is a double canal surface (cf. Sec. 6.3.3), i.e., the envelope of *two* one-parameter families of spheres. The spine curve of each family of spheres is a conic. These two conics c_1 , c_2 are the focal curves of Φ .

The conics c_1 , c_2 are ‘focal conics’ also in the following sense: Their carrier planes are orthogonal, and they have a common axis A . The focal points of c_1 , contained in A , are vertices of c_2 , and vice versa.

The spheres which generate Φ as an envelope surface touch Φ in a family of circles. Thus Φ carries two families of circles. They are the principal curvature lines of Φ , because their normal surfaces are cones. The tangent developables along these circles are, in general, also cones of revolution (or cylinders of revolution or planes).

Because of their nice properties, Dupin cyclides are used in geometric modeling for the construction of smooth transition surfaces (blending surfaces) between cylinders, cones and planes (see e.g. [162]). \diamond

Remark 7.1.10. A scalar function $f(u^1, u^2)$ defined on a surface $\mathbf{a}(u^1, u^2)$ is sometimes visualized by the following graph surface

$$\mathbf{g}(u^1, u^2) = \mathbf{a}(u^1, u^2) + f(u^1, u^2)\mathbf{n}(u^1, u^2), \quad (7.37)$$

where \mathbf{n} is the unit normal vector field of \mathbf{a} (see e.g. [67]). However, this generalization of the graph of a function defined in a planar domain has some drawbacks. As we know from the discussion above, the graph surface will either touch the focal surface (this is the generic case) or have a singularity there. This may not be the intended behaviour and is likely to mislead the viewer.

A simple example is the graph surface of a constant function, which is nothing but the offset surface. It has singularities whenever the function value equals a principal curvature radius, i.e., whenever it has a point in common with the focal surface (see Fig. C.11).

Another type of graph surfaces for functions on surfaces is based on a graph in \mathbb{R}^4 and is visualized via a perspective image in \mathbb{R}^3 [151]; this approach does not have the problems of (7.37). \diamond

7.1.4 Normal Congruences and Geometrical Optics

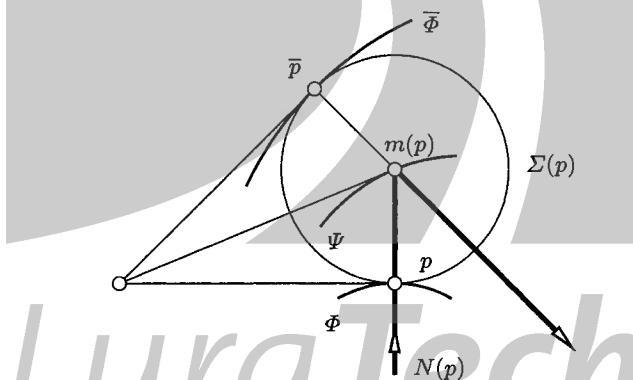


Fig. 7.7. The theorem of Malus and Dupin (see text).

Geometrical optics has a close relation to normal congruences and sphere geometry. In this context, the following theorem of Malus and Dupin is of importance.

Theorem 7.1.12. (Malus and Dupin) If the lines of a normal congruence \mathcal{K} are reflected in a smooth surface Ψ , then the set $\bar{\mathcal{K}}$ of reflected lines is again a normal congruence.

If \mathcal{K} is the normal congruence of the surface Φ , then surfaces $\bar{\Phi}$ orthogonal to the lines of $\bar{\mathcal{K}}$ (i.e., surfaces which $\bar{\mathcal{K}}$ is the normal congruence of) are called *anticaustics* of Φ .

Proof. There are several methods to prove this theorem. We will give a sphere-geometric proof, since it is a rather straightforward extension of our discussion of planar circle geometry (see Sec. 6.3.2). We assume that \mathcal{K} is the normal congruence of a surface Φ , $p \in \Phi$, and $N(p) \in \mathcal{K}$ is the surface normal in p . Since the theorem is of a local nature, we assume without loss of generality that all lines of \mathcal{K} intersect the mirror Ψ in exactly one point.

The sphere $\Sigma(p)$ whose center is $m(p) = N(p) \cap \Psi$ and which contains p actually touches Φ in p . The two-parameter family $\Sigma(p)$ of spheres has an envelope of which Φ is a part. For all p , the sphere $\Sigma(p)$ touches the envelope in p and in a further point \bar{p} which is symmetric to p with respect to the mirror's tangent plane in $m(p)$. This is easy to verify by direct computation (see also the computation given in the following and Fig. 7.7).

Therefore the second component of the envelope is a surface $\bar{\Phi}$ orthogonal to the reflected rays. \square

The assumption that incoming rays are perpendicular to a surface Φ is motivated by the important special cases of illumination by a point source (Φ is a sphere) and illumination by parallel light rays (Φ is a plane).

Remark 7.1.11. The theorem of Malus-Dupin is also valid for refraction according to Snell's law. \diamond

Remark 7.1.12. Note the relation of Th. 7.1.12 to *bisector surfaces*. The mirror surface Ψ is a bisector of the two surfaces $\Phi, \bar{\Phi}$ (cf. also the discussion of bisectors and medial axes in Sec. 6.3.2):

The computation of bisectors in space is much more tedious than in the plane. A direct generalization of Lemma 6.3.16 shows that we can compute the bisector surface by introducing a cyclographic mapping whose domain is \mathbb{R}^4 . We eventually have to intersect two hypersurfaces of \mathbb{R}^4 (the graphs of the distance functions of $\Phi, \bar{\Phi}$), which are counterparts of γ -developables, and whose geometry is closely related to normal congruences.

A numerical procedure for the computation of a bisector surface has been proposed by Elber and Kim [45]. A class of simple surfaces (e.g., spheres, cylinders, cones), whose bisectors are explicitly computable has been studied by M. Peternell [136]. \diamond

The *focal surfaces* of the congruence $\bar{\mathcal{K}}$ are the spatial counterpart of caustics of ray systems in the plane (see the discussion p. 379f). They play a role in reflector design. An application of the Malus-Dupin theorem to *photometric ray tracing* has been given by Lukács and Andor [117].

Parallel Light Rays

We use the notation of Th. 7.1.12 and imagine that the lines of \mathcal{K} are light rays, which are reflected in the surface Ψ . Here we discuss the special case of *parallel*

light rays, i.e., \mathcal{K} is a bundle of parallel lines, and Φ is a plane orthogonal to the lines of \mathcal{K} .

We use a Cartesian coordinate system such that Φ has the equation $x_3 = 0$. The mirror surface shall be parametrized by $\mathbf{m}(u^1, u^2) = (m_1(u^1, u^2), \dots, m_3(u^1, u^2))$. Then the family of spheres mentioned in the proof of Th. 7.1.12 is parameterized by

$$\Sigma(u^1, u^2) : (x_1 - m_1)^2 + (x_2 - m_2)^2 + x_3^2 - 2x_3m_3 = 0, \quad (7.38)$$

where m_1, \dots, m_3 are functions of u^1, u^2 . In order to compute the envelope, we differentiate with respect to u^i , ($i = 1, 2$):

$$(x_1 - m_1)m_{1,i} + (x_2 - m_2)m_{2,i} - x_3m_{3,i} = 0, \quad i = 1, 2.$$

Clearly these equations are satisfied for the contact points $\mathbf{p}(u^1, u^2) = (m_1(u^1, u^2), m_2(u^1, u^2), 0)$, and they express the fact that the line spanned by the two contact points is orthogonal to the mirror's tangent plane.

If $\mathbf{n}(u^1, u^2)$ is a normal vector field of the mirror (e.g., $\mathbf{n} = \mathbf{m}_{,1} \times \mathbf{m}_{,2}$), we can write the second contact point $\bar{\mathbf{p}}$ in the form $\bar{\mathbf{p}} = (m_1, m_2, 0) + \lambda \mathbf{n}$. By intersecting this line with the sphere $\Sigma(u^1, u^2)$, we get the following parametrization of an anticaustic surface $\bar{\Phi}$:

$$\bar{\Phi} : \bar{\mathbf{p}}(u^1, u^2) = (m_1, m_2, 0) + \frac{2n_3m_3}{\|\mathbf{n}\|^2} \mathbf{n}. \quad (7.39)$$

All anticaustic surfaces are offsets of each other. A parametrization is easily found from (7.39), because the surface normals of $\bar{\Phi}$ are the lines of $\bar{\mathcal{K}}$. Obviously $\bar{\mathbf{n}} = (\bar{\mathbf{p}} - \mathbf{m})/m_3$ is a unit normal vector of $\bar{\Phi}$ in $\bar{\mathbf{p}}$, and the family of anticaustics is parameterized by

$$\bar{\Phi}_c : (u^1, u^2) \mapsto \bar{\mathbf{p}} + \frac{c}{m_3}(\bar{\mathbf{p}} - \mathbf{m}). \quad (7.40)$$

The representations are particularly simple if we use a mirror which is the graph of a bivariate function f . If we write (u, v) instead of (u^1, u^2) , the mirror is parametrized by

$$\mathbf{m}(u, v) = (u, v, f(u, v)),$$

and an anticaustic is given by

$$\bar{\Phi} : (u, v) \mapsto \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} + \frac{2f(u, v)}{f_{,u}(u, v)^2 + f_{,v}(u, v)^2 + 1} \begin{bmatrix} -f_{,u}(u, v) \\ -f_{,v}(u, v) \\ 1 \end{bmatrix}. \quad (7.41)$$

Example 7.1.13. We consider the paraboloid Ψ which is the graph of the function

$$f(u, v) = (au^2 + bv^2)/2, \quad ab \neq 0.$$

The anticaustic surface $\bar{\Phi}$ possesses the simple rational parametrization

$$\bar{\varPhi} : (u, v) \mapsto \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} + \frac{au^2 + bv^2}{a^2u^2 + b^2v^2 + 1} \begin{bmatrix} -au \\ -bv \\ 1 \end{bmatrix}.$$

In order to further investigate this surface, we intersect its normals (the lines $\mathbf{m} \vee \bar{\mathbf{p}}$) with the symmetry planes $x_1 = 0$ and $x_2 = 0$. A simple computation reveals the surprising fact that all intersection points are contained in *curves*, namely the parabolae

$$\begin{aligned}\mathbf{f}_1(v) &= (1/2a)(0, 2v(a-b), (ab-b^2)v^2+1), \\ \mathbf{f}_2(u) &= (1/2b)(2u(b-a), 0, (ab-a^2)u^2+1).\end{aligned}$$

This shows that the congruence $\bar{\mathcal{K}}$ of reflected lines has two focal *curves*, and that all anticaustic surfaces are Dupin cyclides (cf. Ex. 7.1.11). The cyclides which occur here are special cases: they are cubic and are called *parabolic Dupin cyclides* (see Fig. 7.8, left). The parabolae \mathbf{f}_1 and \mathbf{f}_2 are the so-called *focal parabolae* of the mirror paraboloid Ψ . Tangent cones of Ψ with vertex in \mathbf{f}_1 or \mathbf{f}_2 are cones of revolution.

Clearly, in the special case of a paraboloid of revolution the reflected rays are incident with the focal point, the congruence $\bar{\mathcal{K}}$ is a bundle, and the anticaustic surfaces are spheres. \diamond

Proposition 7.1.13. *The anticaustic surfaces of a mirror paraboloid (which has no rotational symmetry) with respect to incoming light rays parallel to the paraboloid's axis are parabolic Dupin cyclides.*

Proof. see Ex. 7.1.13. \square

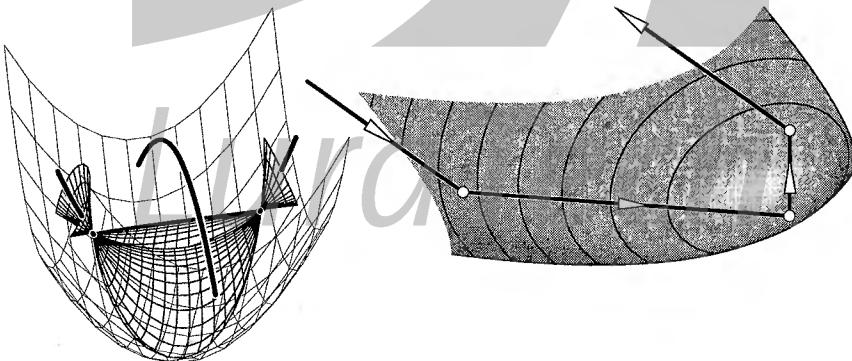


Fig. 7.8. Reflection in a paraboloid. Left: focal parabolae and anticaustics (parabolic Dupin cyclides). Right: Threefold reflection in a paraboloid. Light rays parallel to the axis are reflected into rays parallel to the axis.

Example 7.1.14. (Continuation of Ex. 7.1.13) The paraboloid Ψ is a translational surface, generated by translating a parameter line $u = u_0$ along a parameter line $v = v_0$, and vice versa. We want to show that the mapping $\mathbf{m}(u, v) \mapsto \bar{\mathbf{p}}(u, v)$ from Ψ onto the anticaustic maps this family of parabolae onto the *principal curvature lines* of $\bar{\Phi}$, and that the latter are circles (or lines in special cases). The same is then true for the parabolae $v = v_0$.

In Ex. 7.1.13 it is clear that all rays which meet the same parameter line $v = v_0$ of the paraboloid Ψ are reflected into the same point $\mathbf{f}_1(v_0)$ of the focal parabola \mathbf{f}_1 . There is a parabolic cylinder Δ tangent to Ψ in all points of the line $v = v_0$. Clearly the angle enclosed by the reflected rays and the cylinder's rulings is constant (it is before the reflection), and so the reflected rays form a cone of revolution, which meets the anticaustic in a curve c which is orthogonal to the reflected rays. The curve c therefore is a circle. Because the surface normals of the points of c form a developable surface (the cone mentioned above), c is a principal curvature line of $\bar{\Phi}$. \diamond

Remark 7.1.13. If Ψ is an elliptic paraboloid we can show the focal property of Prop. 7.1.13 as follows: Ψ is a translational surface whose parameter lines are parabolae in orthogonal planes. For the parameter line $v = v_0$ there exists a paraboloid of revolution $\Gamma(p)$ which touches Ψ there, and an incoming light ray which meets this parameter line is reflected in the same way for both $\Gamma(p)$ and Ψ — the reflected ray is incident with the focal point of $\Gamma(p)$. The set of these focal points is the focal parabola \mathbf{f}_1 . \diamond

Remark 7.1.14. The mapping $\mathbf{m}(u, v) \mapsto \bar{\mathbf{p}}(u, v)$ from the mirror to the anticaustic surface has been used in a modeling scheme which employs Dupin cyclides proposed by Peternell and Pottmann [137]: Triangular patches contained in paraboloids with parallel axes, and whose union is a smooth surface are mapped to triangular patches contained in parabolic Dupin cyclides, whose union is a smooth surface again. \diamond

Remark 7.1.15. The process of reflection in a surface may be iterated. We start with a line congruence \mathcal{K} , reflect its lines in Ψ , get a congruence $\bar{\mathcal{K}}$, reflect in Ψ , get a congruence $\bar{\bar{\mathcal{K}}}$, and so on.

For an elliptic paraboloid Ψ W. Wunderlich [211] has shown that $\bar{\bar{\mathcal{K}}} = \bar{\mathcal{K}}$. This implies the following remarkable result:

Proposition 7.1.14. *Light rays parallel to the axis of an elliptic paraboloid become parallel to the axis again after being reflected three times (see Fig. 7.8, right).*

Lines in the symmetry planes become parallel to the axis after a double reflection (because of the focal property of a parabola). If Ψ is a paraboloid of revolution, all lines are contained in symmetry planes. \diamond

Rational Surfaces with Rational Offsets

Parabolic Dupin cyclides are rational surfaces with rational offsets. This is just a special case of the following result, which immediately follows from (7.40):

Theorem 7.1.15. *The anticaustics with respect to reflection of parallel light rays in a rational mirror surface are rational surfaces with rational offsets.*

Peternell and Pottmann [139] have shown that all rational surfaces with rational offsets (*PN surfaces*, i.e., rational surfaces with a rational unit normal vector field) can be generated in this way.

Cyclides are PN surfaces with planar principal curvature lines (namely circles). Surfaces with these two properties have been investigated in [139, 160]. There is the following result:

Theorem 7.1.16. *Consider a translational surface Ψ generated by planar rational curves in orthogonal planes, and light rays parallel to both planes. Then the anticaustic surfaces for reflection in Ψ are rational PN surfaces with planar rational principal curvature lines.*

Proof. We assume that the mirror Ψ is parametrized by

$$\mathbf{m}(u, v) = (a(u), b(v), c(u) + d(v)), \quad (7.42)$$

with rational functions a, b, c, d . The corresponding parametrization of an anticaustic $\bar{\Phi}$ shall be $\bar{\mathbf{p}}(u, v)$. We show that the mapping $\mathbf{m}(u, v) \mapsto \bar{\mathbf{p}}(u, v)$ maps the u parameter lines to planar principal curvature lines. The proof for the v parameter lines is analogous.

We first show that the ruled surface \mathcal{R} of reflected rays along a u parameter line is a γ -developable. This is a first order differential property of \mathcal{R} . In all points of the u -line there is a paraboloid which osculates Ψ , for which it holds true (see Ex. 7.1.14), so it is true for \mathcal{R} also.

The curve c in $\bar{\Phi}$ which corresponds to the u -line is orthogonal (by definition of $\bar{\Phi}$) to the rulings of \mathcal{R} , and therefore planar. The surface \mathcal{R} consists of the surface normals of $\bar{\Phi}$ in the points of c , which shows that c is a principal curvature line (as \mathcal{R} is developable), and completes the proof. \square

There are PN surfaces with planar principal curvature lines which cannot be constructed as anticaustics of translational surfaces [160]. They are a subject of sphere geometry rather than of line geometry and geometrical optics. Although they are more flexible than Dupin cyclides, these surfaces are still not fully suitable to realize a concept in geometric design proposed by R. Martin and others (see [162]): The goal is to design surfaces consisting of patches whose boundaries are principal curvature lines — the intention of this being control over the curvature behaviour.

Example 7.1.15. We let $a(u) = u$, $b(v) = v$, $c(u) = (u - u^3)/2$, and $d(v) = 2(v - v^3/4)/3$ in Equ. (7.42). The resulting surface together with an anticaustic is shown by Fig. 7.9. \diamond

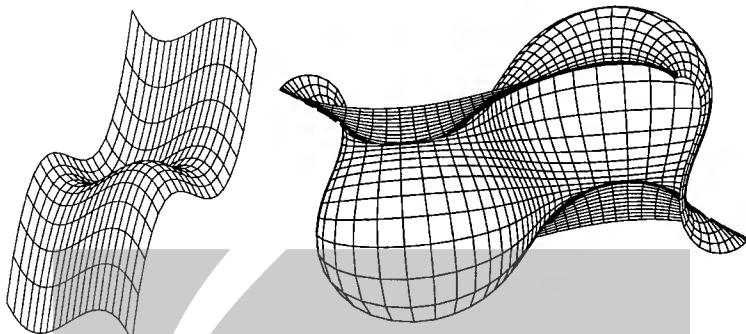


Fig. 7.9. Left: Translational surface. Right: Anticaustic surface according to Th. 7.1.16.

7.1.5 Singularities of Motions Constrained by Contacting Surfaces and Applications in Sculptured Surface Machining

Collision avoidance is a fundamental issue in NC machining of sculptured surfaces [121]. There is a variety of algorithmic contributions on this topic, but very little has been known about the underlying general geometric relations. Recently some global geometric results have been obtained [150, 65, 203], and we would like to describe them briefly.

The *cutter* shall be the surface of revolution generated by the spinning motion of the cutting tool around its axis. Although in practice both the cutter and the work-piece may be moving, it is more elegant to consider only the motion of the cutter relative to the surface and consider the workpiece as fixed. A rotation about the cutter's axis does not play any role and can be neglected completely.

If after these abstractions the cutter performs a purely translational motion, i.e., the cutter's axis remains parallel to a fixed direction, we speak of *3-axis machining* (see Fig. C.10). In *5-axis machining*, also the direction of the axis may change.

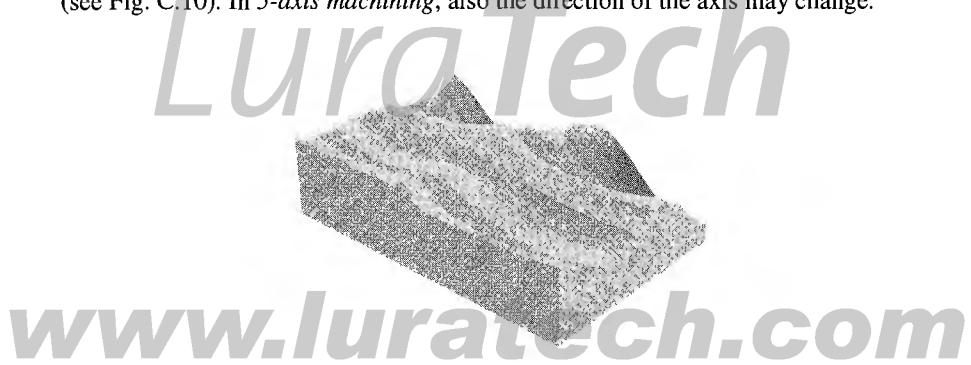


Fig. 7.10. Surface generated during NC machining.

From the geometric viewpoint the cutter moves such that it is always tangent to the final sculptured surface. It is evidently impossible to find a (rectifiable) cutter path which covers the entire surface, so we have to find a sequence of one-parameter motions of the cutter which cover the workpiece in a sufficiently dense way. The surface generated by such a milling process will, in principle, always look like the surface shown in Fig. 7.10.

The deviations from the ideal surface should be as small as possible, and thus appropriate tool selection and cutter motion planning are important. A larger cutter may give a better surface approximation, but could cause interference in some places and therefore be unable to shape the entire surface. Therefore a good understanding of the essential geometric properties of the problem is important.

Local and Global Interference

We say that a surface Φ is *locally millable* by a cutter Σ in a point $p \in \Phi$, if the cutter, while being tangent to Φ in p , does not interfere with Φ in some neighbourhood of p . Interference means intersection of the solids bounded by the surfaces Φ and Σ in a point different from the contact point p .

Local millability can be decided in almost all cases using only curvature information from Φ and Σ in the point p (see [121] and Remark 7.1.18).

Global millability means that for all positions of Σ which it assumes during the manufacturing process there is no interference.

For a convex cutter and 3-axis machining Pottmann et al. [150] have found results of the following type: If Φ is a surface which fulfills some global shape condition which is easily checked for (such as being accessible in the cutting direction), and Φ is *locally* millable by a cutter Σ , then Φ is also *globally* millable by the same cutter.

If such results are applicable, then the test for global interference can be performed by testing for local interference only, which is done by comparing curvatures. As an application, an algorithm for geometrically optimal selection of cutting tools can be developed [65].

We now follow [159] in investigating *motions constrained by a contacting surface pair* and characterize their singularities.

Normal Congruences and Motions

We have shown in Sec. 3.4 that a one-parameter motion has a velocity vector field of the form $\mathbf{v}(\mathbf{x}) = \bar{\mathbf{c}} + \mathbf{c} \times \mathbf{x}$, which is uniquely determined by the pair $(\mathbf{c}, \bar{\mathbf{c}})$. This vector and all its scalar multiples $(\lambda\mathbf{c}, \lambda\bar{\mathbf{c}})$ determine the same path normal complex \mathcal{C} whose Klein image $\mathcal{C}\gamma^*$ equals the point $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} \in P^5$ (Th. 3.4.2).

In what follows we will briefly call $(\mathbf{c}, \bar{\mathbf{c}})$ an *infinitesimal motion* and $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ its *Klein image* (in Sec. 3.4.1, the velocity vector field $\mathbf{v}(x)$ determined by $(\mathbf{c}, \bar{\mathbf{c}})$ was called an infinitesimal motion). For an infinitesimal rotation or translation, the image point is contained in the Klein quadric and represents a line — the axis of the rotation or the ideal line orthogonal to the translation.

Since we will study motions constrained by surface contact, we need more information on surfaces in a line geometric context. This link between surface theory and line geometry is furnished by the normal congruence \mathcal{N} of a C^2 surface Φ .

The Klein image $\mathcal{N}\gamma$ of \mathcal{N} is a two-dimensional surface in the Klein quadric M_2^4 . The simplest example is the normal congruence of a sphere or plane, which is just a bundle of lines. $\mathcal{N}\gamma$ then is a plane contained in M_2^4 , and the tangent space T^2 of $\mathcal{N}\gamma$ in all of its points coincides with $\mathcal{N}\gamma$ itself. In particular, $T^2 \subset M_2^4$.

This is not the typical case, however. By Th. 7.1.11, a line of \mathcal{N} is a hyperbolic congruence line (see the definition p. 425), if the two principal curvatures in the corresponding surface point are not equal. Thus T^2 intersects the Klein quadric in two lines. These lines are the Klein images of the two *principal curvature pencils*, whose vertex is a principal curvature center, and whose plane is spanned by the corresponding principal tangent and the surface normal.

Infinitesimal Gliding Motions

Consider a *flag* $F = (\mathbf{p}, L, \pi)$ in Euclidean three-space, which is a point \mathbf{p} , a line L , and a plane π with $\mathbf{p} \subset L \subset \pi$. Such a flag is also called a *surface element*. We investigate the gliding motion of such a flag along some surface Φ . This is the three-parameter set of positions of the flag such that $\mathbf{p} \in \Phi$ and π is tangent to Φ . We fix an instant of this motion, i.e., a position of F , and study the infinitesimal motion at this instant (cf. the discussion of k -parameter motions at p. 187).

This means that we consider all possible velocity vector fields which belong to one-parameter gliding motions of the flag. For simplicity, we write (\mathbf{p}, L, π) for the position of F we are studying. The surface normal of Φ shall be denoted by N .

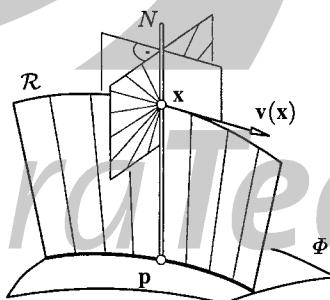


Fig. 7.11. Instantaneous behaviour of the gliding motion of a surface element.

The surface normal N in \mathbf{p} generates a ruled surface \mathcal{R} as \mathbf{p} moves. \mathcal{R} is contained in Φ 's normal congruence (see Fig. 7.11), and any ruled surface \mathcal{R} contained in \mathcal{N} leads to a motion of \mathbf{p} . If $\mathbf{x} \in N$, then its velocity vector $v(\mathbf{x})$ is obviously tangent to \mathcal{R} . As N is a path normal for the point \mathbf{p} , it is a path normal also for \mathbf{x} (see Remark 3.4.3). The vector $v(\mathbf{x})$ is therefore orthogonal to N , and the set of path

normals at \mathbf{x} , which is the pencil orthogonal to $\mathbf{v}(\mathbf{x})$, is orthogonal to \mathcal{R} and contains N (see Fig. 7.11).

This simple observation helps to show the following theorem. Note that an *umbilic* is a point of a surface where the two principal curvatures are equal.

Theorem 7.1.17. *Consider an instant of the gliding motion of a surface element along a C^2 surface Φ , and consider the surface Φ^* generated from Φ by rotation through $\pi/2$ about the contact normal N . The normal congruence of Φ^* is denoted by \mathcal{N}^* .*

Then the Klein image of infinitesimal one-parameter gliding motions at this instant coincides with the tangent plane of the surface \mathcal{N}^γ at $N\gamma$. If the contact point is not an umbilic or flat point of Φ , the axes of infinitesimal rotations contained in this set are the lines of the principal curvature pencils of Φ^* .*

Proof. We continue the discussion immediately before the theorem. A star indicates the rotation mentioned above. We consider an infinitesimal one-parameter gliding motion $(\mathbf{c}, \bar{\mathbf{c}})$. The path normals of a point \mathbf{x} are orthogonal to $\mathbf{v}(\mathbf{x})$. We have seen that the rotation with axis N and angle $\pi/2$ maps the surface tangents of \mathcal{R} in \mathbf{x} to the set of path normals in \mathbf{x} (see Fig. 7.11).

Conversely, $(\mathbf{c}, \bar{\mathbf{c}})$ is an infinitesimal gliding motion, if the velocity vectors $\mathbf{v}(\mathbf{x})$ are tangent to some ruled surface \mathcal{R} and orthogonal to N for all points $\mathbf{x} \in N$. The other points of space are not important.

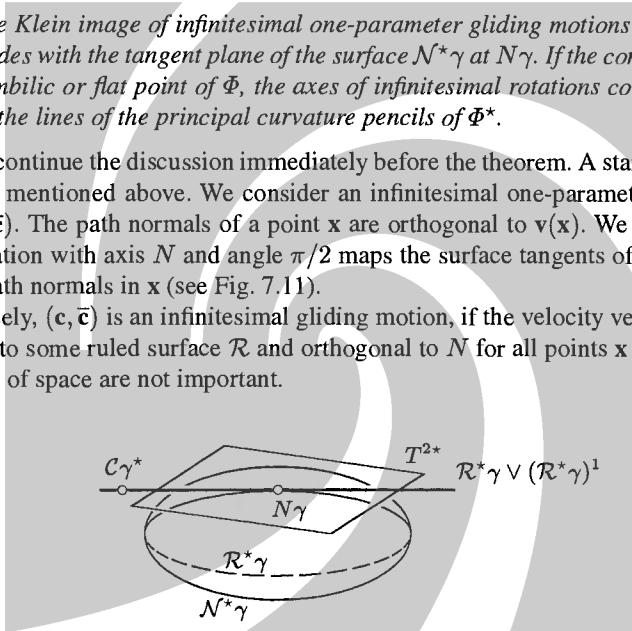


Fig. 7.12. Klein image of path normal complexes (see the proof of Th. 7.1.17).

By Th. 3.4.2, the path normal complex \mathcal{C} of the infinitesimal motion $(\mathbf{c}, \bar{\mathbf{c}})$ is given by $\mathcal{C}\gamma^* = (\mathbf{c}, \bar{\mathbf{c}})\mathcal{R}$. Th. 5.1.4 says that the set of surface tangents of \mathcal{R}^* in points of N is a parabolic line congruence \mathcal{M} , which is the carrier of the pencil $\mathcal{R}^*\gamma \vee (\mathcal{R}^*\gamma)^1$ of linear complexes, i.e., $\mathcal{C} \supset \mathcal{M} \Leftrightarrow \mathcal{C}\gamma^* \in \mathcal{R}^*\gamma \vee (\mathcal{R}^*\gamma)^1$ (see Fig. 7.12).

A linear complex \mathcal{C} or the corresponding infinitesimal motion $(\mathbf{c}, \bar{\mathbf{c}})$ is therefore admissible if $\mathcal{C}\gamma^*$ is contained in the tangent of $\mathcal{R}^*\gamma$ for any ruled surface $\mathcal{R}^* \subset \mathcal{N}^*$, which means $\mathcal{C}\gamma^* \in T^{2*}$. This shows the first part of the theorem.

The statement about the infinitesimal rotations follows from the fact that $(\mathbf{c}, \bar{\mathbf{c}})$ is an infinitesimal rotation, if and only if $(\mathbf{c}, \bar{\mathbf{c}})\mathcal{R} \in M_2^4$, from Th. 7.1.11, and the definitions of ‘torsal direction’ and ‘focal surface’. \square

Remark 7.1.16. It is possible to derive this result without referring to the Klein image. By using elementary differential geometry (Rodrigues’ equations) we can

show that special infinitesimal gliding motions of a flag are (i) the rotation about the contact normal N and (ii) the rotations about the principal axes of curvature (or translations parallel to the principal tangents in case of vanishing principal curvatures). These three infinitesimal rotations are linearly independent and therefore they span a three-dimensional linear space — their Klein image points span a two-dimensional projective plane. This space has the right dimension already and thus actually is the space of infinitesimal gliding motions.

The Klein model however gives additional insights which will be useful later. It also shows how to compute the space of infinitesimal motions without computing the principal curvatures: If $(\mathbf{n}^*(u, v), \bar{\mathbf{n}}^*(u, v))$ is a parametrization of Φ^* 's normal congruence in Plücker coordinates, the space of infinitesimal gliding motions at parameter values (u_0, v_0) is spanned by

$$(\mathbf{n}^*, \bar{\mathbf{n}}^*)(u_0, v_0), (\mathbf{n}_{,u}^*, \bar{\mathbf{n}}_{,u}^*)(u_0, v_0), (\mathbf{n}_{,v}^*, \bar{\mathbf{n}}_{,v}^*)(u_0, v_0).$$

◇

Singular Instants of the Gliding Motion

We consider the motion of the surface Φ_2 gliding on another surface Φ_1 . We assume that both surfaces are C^2 . We introduce a flag $F = (\mathbf{p}, L, \pi)$ tangent to both surfaces in the contact point as a third moving system (This actually restricts the possible motions of Φ_2 because we assume that the contact point moves smoothly). Of course, only the point \mathbf{p} and the plane π are uniquely determined by the position of Φ_2 .

The motion Φ_2/Φ_1 can be decomposed into the motions Φ_2/F and F/Φ_1 . The instantaneous behaviour of the motion Φ_2/F is the same as that of the motion F/Φ_1 , and thus we know the behaviour of both parts in this decomposition.

There are the normal congruences \mathcal{N}_i and their rotated versions \mathcal{N}_i^* according to Th. 7.1.17 ($i = 1, 2$), and the tangent planes T_i^* of \mathcal{N}_i^* in the point $N\gamma$, where N is the contact normal. If $(\mathbf{c}, \bar{\mathbf{c}})$ is an admissible infinitesimal motion of Φ_2/F , and $(\mathbf{c}', \bar{\mathbf{c}}')$ is an admissible infinitesimal motion of F/Φ_1 , then Cor. 3.4.7 shows that the sum $(\mathbf{c} + \mathbf{c}', \bar{\mathbf{c}} + \bar{\mathbf{c}}')$ is an admissible infinitesimal motion for Φ_2/Φ_1 . This means that the Klein images of these three infinitesimal motions are collinear. Since all scalar multiples of admissible infinitesimal motions are admissible as well, the entire line

$$(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} \vee (\mathbf{c}', \bar{\mathbf{c}}')\mathbb{R}$$

consists of admissible infinitesimal motions of Φ_2/Φ_1 . We have shown the following:

Lemma 7.1.18. *The Klein image space of infinitesimal motions of Φ_2/Φ_1 equals the space spanned by T_1^* and T_2^* (which is, in general, four-dimensional).*

It is a consequence of Lemma 7.1.18 that there is in general a five-dimensional linear space of infinitesimal gliding motions of Φ_2 on Φ_1 — the gliding motion Φ_2/Φ_1 has five infinitesimal degrees of freedom. If we identify infinitesimal motions which are

multiples of each other, we get a four-dimensional projective space of infinitesimal motions.

A *singular* position of Φ_2/Φ_1 occurs, by definition, if the dimension of the space $T_1^* \vee T_2^*$ is less than four.

Remark 7.1.17. If $(\mathbf{n}_1(u, v), \bar{\mathbf{n}}_1(u, v))$ and $(\mathbf{n}_2(u, v), \bar{\mathbf{n}}_2(u, v))$ are parametrizations of the normal congruences of Φ_1 and Φ_2 in Plücker coordinates, respectively, then the infinitesimal degree of freedom of the gliding motion Φ_2/Φ_1 is computed as the dimension of the linear space spanned by

$$(\mathbf{n}_1, \bar{\mathbf{n}}_1) = (\mathbf{n}_2, \bar{\mathbf{n}}_2), (\mathbf{n}_{1,u}, \bar{\mathbf{n}}_{1,u}), (\mathbf{n}_{1,v}, \bar{\mathbf{n}}_{1,v}), (\mathbf{n}_{2,u}, \bar{\mathbf{n}}_{2,u}), (\mathbf{n}_{2,v}, \bar{\mathbf{n}}_{2,v}).$$

It is not necessary to perform the rotation about $\pi/2$ mentioned in Th. 7.1.17 in order to determine the dimension. In view of unavoidable inaccuracies in computation, this computation should actually be done within a geometrically formulated regression problem, using ideas on approximation in line space given in Chap. 4. \diamond

Second Order Line Contact of Surfaces

Here we are studying different concepts of second order contact of surfaces and how they are related to contact of their normal congruences.

Consider a C^2 surface parametrized by $\mathbf{f}(u, v)$, and denote its unit normal vector field by $\mathbf{n}(u, v)$. Differentiation with respect to t will be denoted by a dot. Recall that a surface curve $\mathbf{f}(u(t), v(t))$ with tangent vector $\mathbf{t} = \dot{\mathbf{f}}$ has the normal curvature

$$\kappa_n(\mathbf{t}) = \dot{\mathbf{n}} \cdot \dot{\mathbf{f}} / \|\dot{\mathbf{f}}\|^2.$$

If the surface curve is parametrized by arc length, the normal curvature is the angular velocity of the normal vector. The polar diagrams i_+ of the inverse square root of the normal curvature and i_- of the inverse square root of the negative normal curvature are conics or pairs of parallel lines, or void, and are called the *signed Dupin indicatrix* of the surface. Two surfaces are in second order contact in a point if and only if their Dupin indicatrices coincide.

We say that two line congruences $\mathcal{N}_1, \mathcal{N}_2$ are in *k-th order contact*, if their Klein images $\mathcal{N}_1\gamma, \mathcal{N}_2\gamma$ are in *k-th order contact* as surfaces in projective space. For the next lemma, we need some preparations:

Consider a surface Φ which contains the point $(0, 0, 0)$ and has the normal vector $(0, 0, 1)$ there. If $\mathbf{c}(t)$ is a surface curve and $\mathbf{n}(t)$ is the unit normal vector in $\mathbf{c}(t)$, we intersect the surface normal with the plane $x_3 = 0$ and get the point $\mathbf{d}(t) = (c_1 - n_1 c_3/n_3, c_2 - n_2 c_3/n_3, 0)$. If $\mathbf{c}(0) = (0, 0, 0)$, then obviously $\mathbf{d}(0) = \dot{\mathbf{c}}(0)$, because $\dot{c}_3(0) = n_1(0) = n_2(0) = 0$.

If we are given a line congruence \mathcal{N} , parametrized in Plücker coordinates by $(\mathbf{n}(u, v), \bar{\mathbf{n}}(u, v))$, such that $\mathbf{n}(0, 0) = (0, 0, 1)$, we consider the intersection point $\mathbf{d}(u, v)$ of the line $N(u, v)$ with the plane $x_3 = 0$. Obviously we can parametrize \mathcal{N} by the first two coordinates of \mathbf{d} , so without loss of generality we may assume that $N(u, v) = (\mathbf{n}(u, v), (u, v, 0) \times \mathbf{n}(u, v))$.

Lemma 7.1.19. Consider two C^2 surfaces Φ, Φ' which are in first order contact. Their normal congruences $\mathcal{N}, \mathcal{N}'$ are in first order contact if and only if Φ, Φ' are in second order contact.

Proof. The ‘if’ part is clear because the surface normals are computed by differentiating once. The ‘only if’ part is shown as follows: Consider a surface curve $\mathbf{c}(t)$ in Φ such that $\mathbf{c}(0)$ is the contact point. The unit normal vector in $\mathbf{c}(t)$ shall be denoted by $\mathbf{n}(t)$. N is the surface normal in the contact point.

The text immediately before the lemma shows that by re-parametrizing the normal congruence we can compute $\dot{\mathbf{n}}(0)$ (and therefore $\kappa_n(\dot{\mathbf{c}}(0))$) from $\dot{\mathbf{c}}(0)$, using only first order derivatives of \mathcal{N} , i.e., the tangent plane of $\mathcal{N}\gamma$.

Now the tangent planes of $\mathcal{N}\gamma, \mathcal{N}'\gamma$ coincide in the point $N\gamma$, the surfaces Φ, Φ' have equal normal curvatures for all surface tangents, and so they are in second order contact. \square

Lemma 7.1.20. Assume that surfaces Φ_1, Φ_2 are in second order contact, and that surface curves $\mathbf{c}_i(t)$ in Φ_i touch each other in the contact point. The ruled surfaces \mathcal{S}_i , whose generators are the surface normals $N_i(t)$ in $\mathbf{c}_i(t)$, are in first order contact in all points of the contact normal.

Proof. By Th. 5.1.4 and Th. 5.1.5, the tangent behaviour of \mathcal{S}_1 and \mathcal{S}_2 is completely determined by the tangents of the curves $N_1\gamma(t)$ and $N_2\gamma(t)$. The discussion immediately before Lemma 7.1.19 shows that these tangents are determined by $\dot{\mathbf{c}}_i(0)$ and the tangent plane of $\mathcal{N}_i\gamma$, where \mathcal{N}_i is the normal congruence of Φ_i . These tangent planes are equal by Lemma 7.1.19, so we are done. \square

We say that two surfaces Φ_1, Φ_2 are in *line contact*, if they have a common surface curve \mathbf{c} , and touch each other in all points of \mathbf{c} . Obviously the normal surfaces $\mathcal{S}_1, \mathcal{S}_2$ along this curve coincide. We say that two surfaces are in *second order line contact*, if they are in first order contact, and there are surface curves $\mathbf{c}_1, \mathbf{c}_2$ in Φ_1, Φ_2 , respectively, such that their normal surfaces are in first order contact.

As the proof of Lemma 7.1.20 shows, the tangent behaviour of the normal surface in the points of the contact normal depends only on $\dot{\mathbf{c}}_1$ or $\dot{\mathbf{c}}_2$.

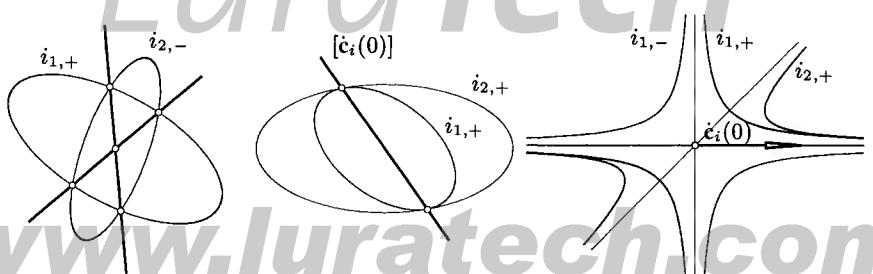


Fig. 7.13. Dupin indicatrices of a pair of surfaces which are in first order contact. The second and third pair of surfaces is in second order line contact.

There are several conditions which equivalently express second order line contact of surfaces.

1. If both Φ_1 and Φ_2 are parametrized as graph surfaces $(x, y, z_i(x, y))$, such that the contact point is $(0, 0, 0)$, and the contact normal is $(0, 0, 1)$, we consider the Hessian matrices

$$H_i = \begin{bmatrix} z_{i,xx}(0, 0) & z_{i,xy}(0, 0) \\ z_{i,xy}(0, 0) & z_{i,yy}(0, 0) \end{bmatrix}. \quad (7.43)$$

It is easily seen that second order line contact is equivalent to

$$\det(H_1 - H_2) = 0. \quad (7.44)$$

2. Another more geometric characterization is in terms of Dupin indicatrices of the two surfaces in question (which, incidentally, have equations $(x, y) \cdot H_i \cdot (x, y)^T = 1$):

The surfaces Φ_1, Φ_2 are in second order line contact if and only if the indicatrices touch each other in two opposite points, or if they hyperosculate at infinity. Hyperosculating conics at infinity means that the two conics are hyperbolae which have a common asymptote, and there is a shear transformation parallel to this asymptote which transforms one hyperbola into the other (see Fig. 7.13, right).

3. We use the notation of the definition of second order line contact, and consider the unit normal vectors $\mathbf{n}_i(t)$ in the points $\mathbf{c}_i(t)$. Second order line contact means that the equality $\dot{\mathbf{c}}_1(0) = \dot{\mathbf{c}}_2(0)$ implies the equality $\dot{\mathbf{n}}_1(0) = \dot{\mathbf{n}}_2(0)$. This implies that the normal curvatures $\kappa_n(\dot{\mathbf{c}}_i(0))$ are equal. So are the surface tangents *conjugate* to $\dot{\mathbf{c}}_i(0)$, because they are computed by $\mathbf{n} \times \dot{\mathbf{n}}$.

If the normal curvatures are nonzero and equal, and if the tangents conjugate to $\dot{\mathbf{c}}_i(0)$ coincide, then also $\dot{\mathbf{n}}_1(0) = \dot{\mathbf{n}}_2(0)$, so we have second order line contact. If the normal curvatures are equal and zero, equality of the conjugate tangent is not sufficient. It is necessary (and sufficient) that normal curvature and *geodesic torsion*

$$\tau_g(\dot{\mathbf{c}}_i) = \frac{1}{\dot{\mathbf{c}}_i^2} \det(\dot{\mathbf{c}}_i, \mathbf{n}, \dot{\mathbf{n}}_i) \quad (7.45)$$

are the same for both surfaces.

Remark 7.1.18. Recall the criterion for local millability mentioned in the beginning of this section: A surface Φ is locally millable by a convex cutter Σ in point \mathbf{p} , if the indicatrix $i_{\Sigma,+}$ of Σ is contained in the interior of the indicatrix $i_{\Phi,+}$ of Φ (see Fig. 7.14).

In an elliptic surface point all normal curvatures have the same sign, so the indicatrix i_+ is an ellipse and i_- is void, or vice versa. In Fig. 7.14, the cutter has an elliptic surface point, and we assume that $i_{\Sigma,+}$ is the non-void indicatrix. \diamond

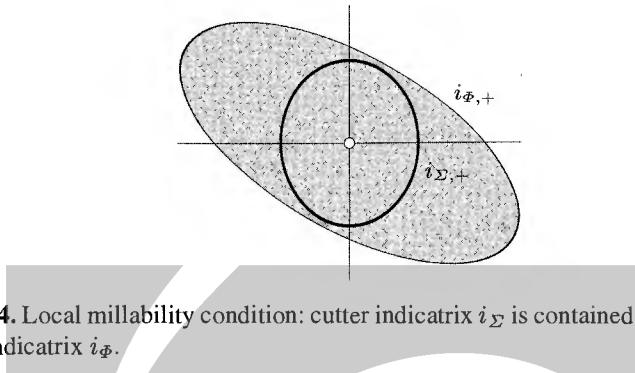


Fig. 7.14. Local millability condition: cutter indicatrix i_Σ is contained in the interior of the indicatrix i_Φ .

Characterization of Singular Positions

It turns out that there is another equivalent characterization of second order line contact of surfaces:

Theorem 7.1.21. *The gliding motion of a surface Φ_2 on a surface Φ_1 has five infinitesimal degrees of freedom (i.d.o.f's). An instant of this motion is singular (the number of i.d.o.f's is less or equal four), if and only if Φ_1 and Φ_2 are in second order line contact.*

Φ_1, Φ_2 are in second order contact if and only if the number of i.d.o.f's equals three.

Proof. We resume the discussion which led to Lemma 7.1.18, and use the symbols \mathcal{N}_i for the normal congruence of Φ_i , N for the contact normal, T_i for the tangent plane of $\mathcal{N}_i\gamma$ in the point $N\gamma$.

The dimension of $T_1 \vee T_2$ equals two (which means three i.d.o.f's), if and only if $T_1 = T_2$, i.e., \mathcal{N}_1 and \mathcal{N}_2 are in first order contact. By Lemma 7.1.19, this happens if and only if Φ_1, Φ_2 are in second order contact.

The dimension of $T_1 \vee T_2$ is less or equal three (which means at most four i.d.o.f's), if these two planes intersect in a line G^1 . A curve $S_i\gamma$ in $\mathcal{N}\gamma$ which contains $N\gamma$ and touches G^1 corresponds to a ruled surface S_i consisting of surface normals of Φ_i . Because $S_1\gamma, S_2\gamma$ touch each other in the point $N\gamma$, Th. 5.1.4 and Th. 5.1.5 show that S_1 and S_2 are in first order contact in all points of the contact normal N . By definition, Φ_1 and Φ_2 are in second order line contact. \square

Remark 7.1.19. We can locally parametrize Φ_1 and Φ_2 as graph surfaces $(x, y, z_i(x, y))$ such that the contact point is $(0, 0, 0)$ and the contact normal is parallel to $(0, 0, 1)$. We compute the Hessians H_1, H_2 according to (7.43). Then the infinitesimal degree of freedom of the gliding motions equals $3 + \text{rk}(H_1 - H_2)$. \diamond

Corollary 7.1.22. *Assume that Φ_2 performs a gliding motion on Φ_1 . If Φ_1, Φ_2 are not in second order line contact, the linear space of admissible infinitesimal gliding motions $(\mathbf{c}, \bar{\mathbf{c}})$ has the equation*

$$\mathbf{n} \cdot \bar{\mathbf{c}} + \bar{\mathbf{n}} \cdot \mathbf{c} = 0, \quad (7.46)$$

where $(\mathbf{n}, \bar{\mathbf{n}})$ are the Plücker coordinates of the contact normal.

Proof. The contact normal is a path normal for any admissible infinitesimal gliding motion $(\mathbf{c}, \bar{\mathbf{c}})$, so it must fulfill (7.46). By Th. 7.1.21, the linear space of all admissible $(\mathbf{c}, \bar{\mathbf{c}})$ has dimension five, therefore coincides with the solution space of (7.46). \square

By considering pipe surfaces of radius r and letting $r \rightarrow 0$, we are able to derive information on the infinitesimal degrees of freedom of the gliding motion of a curve on a surface, or of a curve on a curve. This is also relevant for applications in NC machining, because both the cutter and the workpiece are likely to have edges (a flat end mill, for instance). We state only the results which are obtained from Th. 7.1.21 by a limit argument.

Theorem 7.1.23. *The gliding motion of a curve c on a surface Φ has five i.d.o.f's if c does not have second order contact with Φ . Otherwise, the number of i.d.o.f's equals four.*

Obviously for all points of Φ a pipe surface with spine curve c is not in second order contact with Φ if the pipe's radius is small enough. This fact is responsible for the fact that the three i.d.o.f's do not occur in Th. 7.1.23.

Theorem 7.1.24. *An instant of the gliding motion of a curve c_1 on a curve c_2 — which means motion of c_2 such that it always meets c_1 — has five i.d.o.f's if and only if c_1 and c_2 do not touch each other.*

The motion has three i.d.o.f's if and only if c_1 and c_2 are in second order contact, and four i.d.o.f's if and only if c_1 and c_2 are in first order contact without being in second order contact.

An immediate application of these results is the investigation of infinitesimal degrees of freedom of kinematic chains which contain gliding joints. The simplest case is that of a rigid body (the moving system), which is connected to the 'fixed' system by a number of joints. The space of infinitesimal motions of this body is found by intersecting the spaces of possible infinitesimal motions of the single joints.

A smooth one-parameter motion, which has one infinitesimal degree of freedom in all its instants, is obtained, in general, if the moving body is connected to the fixed system by five gliding joints. This is because the intersection of five five-dimensional linear subspaces of a six-dimensional linear space is of dimension one, if they are in general position. In this general case it is very easy to find the intersection: An infinitesimal motion $(\mathbf{c}, \bar{\mathbf{c}})$ must have all contact normals N_1, \dots, N_2 , as path normals, so the path normal complex \mathcal{C} is found by $\mathcal{C}\gamma = N_1\gamma \vee \dots \vee N_2\gamma$. If $(\mathbf{n}_i, \bar{\mathbf{n}}_i)$ are the Plücker coordinates of the contact normals, we have to solve the equations $\mathbf{n}_i \cdot \bar{\mathbf{c}} + \bar{\mathbf{n}}_i \cdot \mathbf{c} = 0$ (cf. Cor. 7.1.22).

If one of the joints is in a singular position, however, the dimension of the space of admissible infinitesimal motions drops, and we must expect zero infinitesimal degrees of freedom, which is called a *rest position*.

Remark 7.1.20. For a study of singular positions of motions defined by *rolling* of surfaces as well as applications in geometric design, we refer the reader to [159]. It should be emphasized, however, that rolling of one surface on another is a motion which has more degrees of freedom than it has infinitesimally, but this need not be true for motions constrained by multiple rolling conditions. \diamond

Motion by Translations and General Offsets

Consider the motion by *translations* of a surface Σ such that it touches another surface Φ . Such motions occur in three-axis milling, where the cutter Σ is a surface of revolution and Φ is the surface which is eventually shaped by the cutter.

A reference point in Σ traces out a general offset surface (see Fig. C.10) of Φ with respect to Σ . Th. 7.1.21 specializes to a result in [147]:

Corollary 7.1.25. *The general offset surface defined by a translatory gliding motion of a surface Σ on a surface Φ is regular if Φ and Σ are not in second order line contact.*

Proof. We have to show that the space of infinitesimal translations admissible for the gliding pair Σ/Φ is two-dimensional. The linear space T of infinitesimal translations has dimension three, and is never contained in the linear space U of admissible infinitesimal motions, because the translation parallel to the contact normal is certainly not admissible. By Th. 7.1.21, $\dim(U) = 5$, which implies $\dim(U \cap T) = 2$. \square

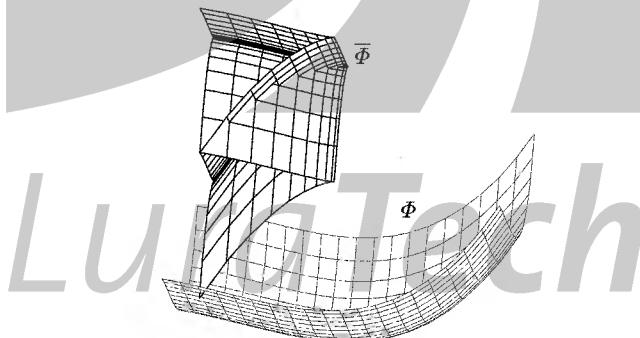


Fig. 7.15. The surface $\Phi : z = (x^2 + y^2)/2 + (x^3 - y^3)/3$ and its general offset $\bar{\Phi}$ with respect to the paraboloid $\Sigma : z = (x^2 + y^2)/2$.

Example 7.1.16. Assume that Σ is a paraboloid of revolution which has the equation $2z = x^2 + y^2$, and that Φ is a graph surface parametrized by $(x, y, f(x, y))$. The tangent plane of Φ in the point $(x, y, f(x, y))$ is parallel to the tangent plane of Σ in the point

$$\mathbf{s}(x, y) = (f_{,x}, f_{,y}, \frac{1}{2}(f_{,x}^2 + f_{,y}^2)).$$

We choose the point $\mathbf{r} = (0, 0, h)$ as reference point. Its path surface under the gliding motion Σ/Φ has the parametrization

$$\bar{\Phi} : \bar{\mathbf{f}}(x, y) = \mathbf{f} - \mathbf{s} + \mathbf{r} = (x - f_{,x}, y - f_{,y}, f - (f_{,x}^2 + f_{,y}^2)/2 + h). \quad (7.47)$$

Figure 7.15 illustrates the surface $\Phi : z = (x^2 + y^2)/2 + (x^3 - y^3)/3$ with $h = 1$. The general offset is parametrized by

$$\bar{\Phi} : \bar{\mathbf{f}}(x, y) = (-x^2, y^2, \frac{2}{3}(y^3 - x^3) - \frac{1}{2}(x^4 + y^4) + 1).$$

The difference of the Hessians of the two surfaces Φ and Σ equals

$$H_1 - H_2 = \begin{bmatrix} 2x & 0 \\ 0 & -2y \end{bmatrix}.$$

This shows that we lose one degree of freedom in points with $x = 0$ or $y = 0$ and two degrees of freedom if $x = y = 0$. The two surfaces Φ and Σ are in second order contact at $x = y = 0$. Thus the parameter lines $x = 0$ and $y = 0$ are edges of regression of $\bar{\Phi}$. They are quartic curves which have a cusp at $x = y = 0$, indicating the loss of two degrees of freedom.

Both Φ and Σ are translational surfaces generated by translating a curve in the plane $x = 0$ along a curve in the plane $y = 0$ and vice versa. This implies that the general offset is likewise a translational surface. The quartics of regression can be used as generator curves. \diamond

Applications to the Geometry of Sculptured Surface Machining

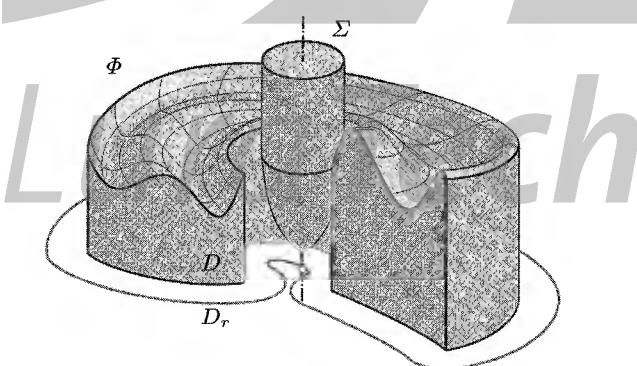


Fig. 7.16. Situation which allows to conclude global millability from the local one.

Imagine a cutter Σ which moves by translations such that it touches the surface Φ . Φ is globally millable, if the interiors of the solids bounded by Σ and Φ do not

intersect. Imagine Σ approaching Φ and beginning its motion. If Σ , while being tangent to Φ in one point, intersects the interior of Φ somewhere else, it must have been tangent to Φ in a second point before. If we recall the definition of the general offset surface $\bar{\Phi}$, this means that for two different contact points, the reference point which traces out $\bar{\Phi}$ is in the same place, so $\bar{\Phi}$ has a self-intersection. So all we have to do is to find conditions that $\bar{\Phi}$ does not self-intersect.

This is basically the idea used in [150] to prove millability results. One particular situation where local millability implies global millability is illustrated in Fig. 7.16: (i) the surface Φ is accessible in milling direction, which is assumed vertical; and (ii) the orthogonal projection D of Φ onto a horizontal plane has an offset curve D_r at distance r which is free of self-intersections, where r is the maximum radius of the cutter.

To illustrate how to use our results in the discussion of global millability for 5-axis sculptured surface machining, consider the following special milling strategy. Assume a cutter, represented by a strictly convex surface of revolution Σ . Further assume a closed surface Φ which is strictly star-shaped with respect to a point \mathbf{o} in the sense that all rays emanating from \mathbf{o} intersect Φ transversely in exactly one point (see Fig. C.9).

An actual milling motion in practice would consist of a series of sufficiently close one-parameter motions, which we would like to embed in a two-parameter gliding motion of Σ on Φ . A simple example is that Σ 's axis is always incident with \mathbf{o} .

This gliding motion is a three-parameter motion: two parameters for the variation of the contact point and one parameter for rotation about Σ 's axis A , which is actually negligible in geometric considerations. It is not difficult to show that for this example local millability implies global millability:

Theorem 7.1.26. *Assume that Φ , Σ are piecewise C^2 surfaces, that Σ is a strictly convex surface of revolution, and Φ is strictly star-shaped with respect to a point \mathbf{o} in the sense defined above.*

Then the gliding motion of Σ on Φ such that Σ 's axis is incident with \mathbf{o} is globally collision-free, if the local millability criterion is fulfilled in all instants.

Proof. (Sketch) Local millability means that the Dupin indicatrix of Σ is contained in the interior of the Dupin indicatrix of Φ in all possible contact points, which shows that Σ and Φ are not in second order line contact.

By Cor. 7.1.22, the linear space U of infinitesimal motions $(\mathbf{c}, \bar{\mathbf{c}})$ admissible for the pair Σ/Φ has the equation $\mathbf{n} \cdot \bar{\mathbf{c}} + \bar{\mathbf{n}} \cdot \mathbf{c} = 0$, if $(\mathbf{n}, \bar{\mathbf{n}})$ are Plücker coordinates of the contact normal.

If \mathbf{p} is a point in Σ 's axis A , then A has Plücker coordinates $(\mathbf{p}, \mathbf{o} \times \mathbf{p}) = (\mathbf{p}, \mathbf{o})$. Any infinitesimal rotation with axis A is in U , so $\mathbf{p} \cdot \bar{\mathbf{n}} = 0$.

Consider the four-parameter motion of Σ such that A is incident with \mathbf{o} . Denote the linear space of its infinitesimal motions by V . An element $(\mathbf{c}, \bar{\mathbf{c}}) \in V$ assigns to the origin the velocity vector $\bar{\mathbf{c}} + \mathbf{c} \times \mathbf{o} = \bar{\mathbf{c}} = \lambda \mathbf{p}$. Φ is strictly star-shaped, so $\mathbf{n} \cdot \mathbf{p} \neq 0$, and we compute $U \cap V = \{(\mathbf{c}, -(\bar{\mathbf{n}} \cdot \mathbf{c} / \mathbf{n} \cdot \mathbf{p}) \mathbf{p})\}$, which is three-dimensional.

We consider the surface $\bar{\Phi}$ traced out by \mathbf{p} . In order to show that it is actually a surface and it is regular, we have to show that the linear space of velocities of \mathbf{p} (which eventually span the tangent plane of $\bar{\Phi}$) is exactly two-dimensional. We consider the mapping of $U \cap V$ to \mathbb{R}^3 which maps $(\mathbf{c}, \bar{\mathbf{c}})$ to $\mathbf{v}(\mathbf{p}) = \bar{\mathbf{c}} + \mathbf{c} \times \mathbf{p}$. Its kernel is easily verified to be the linear span of (\mathbf{p}, \mathbf{o}) , so its rank equals two, what we wanted to show.

Analogously we can show that $\bar{\Phi}$ is not tangent to A . Now consider the unit sphere centered in \mathbf{o} , and define a mapping ψ as follows: For $\mathbf{x} \in S^2$ we find the unique point $\lambda \mathbf{x} \in \Phi$, find the position of Σ such that it touches Φ there, find the position \mathbf{p} of the reference point, and define $\psi(\mathbf{x}) = \mathbf{p}/\|\mathbf{p}\|$.

The mapping ψ is smooth and regular (because all components are smooth and regular), and by shrinking Σ to a point it can be deformed smoothly into the identity mapping. It is well known that such a mapping is one-to-one and onto. It follows that there is no position of Σ which touches Φ twice, and so Σ cannot intersect the interior of the solid bounded by Φ during the gliding motion.

This shows the statement of the theorem if both Σ and Φ are C^2 . For piecewise C^2 surfaces, see [150, 200]. \square

Extensions of this result have been shown by Wallner and Pottmann [203]. As in the proof of Th. 7.1.26, the actual cutter motion, which consists of a sequence of one-parameter motions, is embedded into a two-parameter motion (or a three-parameter motion with negligible rotations in it). Then the line congruence \mathcal{K} of cutter axis positions plays an essential role: The cutter should be contained in a neighbourhood of Φ where the lines of \mathcal{K} define a *fibration*, i.e., all points are incident with precisely one line of \mathcal{K} . Then local millability implies global millability if some technical conditions are fulfilled. In Th. 7.1.26, \mathcal{K} was a bundle of lines, which satisfies the fibration condition everywhere except in \mathbf{o} .

7.1.6 Numerical Geometry of Line Congruences

Much of Sec. 5.4, which concerns the numerical geometry of ruled surfaces, can be extended to line congruences. Often this extension is straightforward. Thus we will just briefly show the generalization by means of a few illustrative examples.

One motivation for the construction of interpolating and approximating line congruences comes from motion planning in five-axis NC machining (cf. the previous section). Having found a number of locally and globally collision free cutter positions, one might want to embed their axes into a congruence. This congruence usually uniquely determines a two-parameter gliding motion of a given cutter on the workpiece, which is to be decomposed into a sequence of one-parameter cutter paths.

For this application it is sufficient to work with line segments contained in an area of interest. The mapping (4.31) transforms the interpolation or approximation problem on line segments into a respective problem on points in \mathbb{R}^6 , whose solution is based on standard methods of CAGD, which work in higher dimensions as well.

Methods which use a Euclidean scalar product and distance are appropriately modified (cf. Sec. 5.4), e.g., by using the scalar product (4.33). We can also view the interpolation/approximation problem in \mathbb{R}^6 as two coupled interpolation/approximation problems in \mathbb{R}^3 , one for each set of endpoints of the given line segments.

Variational Design of Two-Parameter Families of Line Segments

Variational design of a two-parameter family of line segments is done analogously to variational design of one-parameter families. The latter are strips of ruled surfaces, and the former can be imagined as layers in line congruences. We parametrize such a layer in the form

$$\mathbf{x}(u, v, w) = \mathbf{c}(u, v) + w\mathbf{g}(u, v), \quad (u, v) \in D \subset \mathbb{R}^2, w \in [-1, 1]. \quad (7.48)$$

The two-parameter family of line segments which corresponds to this parametrization is $\mathbf{x}(u, v, -1)\mathbf{x}(u, v, 1)$. We call \mathbf{c} the central surface of \mathbf{x} . Consider the functional,

$$\tilde{F}(\mathbf{x}) = \int_D \int_{-1}^1 (\mathbf{x}_{,uu}^2 + \mathbf{x}_{,vv}^2 + \mathbf{x}_{,ww}^2 + 2\mathbf{x}_{,uv}^2 + 2\mathbf{x}_{,uw}^2 + 2\mathbf{x}_{,vw}^2) d(u, v) dw, \quad (7.49)$$

which decomposes into the sum

$$\begin{aligned} \tilde{F}(\mathbf{x}) &= 2 \int_D (\mathbf{c}_{,uu}^2 + 2\mathbf{c}_{,uv}^2 + \mathbf{c}_{,vv}^2) du dv + \\ &+ \frac{2}{3} \int_D [\mathbf{g}_{,uu}^2 + 2\mathbf{g}_{,uv}^2 + \mathbf{g}_{,vv}^2 + 6(\mathbf{g}_{,u}^2 + \mathbf{g}_{,v}^2)] du dv = \\ &= F(\mathbf{c}) + \tilde{T}(\mathbf{g}). \end{aligned} \quad (7.50)$$

The first term is the value of the thin plate spline functional (5.89) on the central surface \mathbf{c} . It may be seen as a linearized bending energy and its minimizers under interpolation constraints are Duchon's thin plate splines [78]. The second term involves only \mathbf{g} and its minimizers are thin plate splines with tension, which have been studied by R. Franke [58]. Both Duchon's and Franke's splines are expressed with radial basis functions. For their computation we refer to the literature [78].

The following result sums up the previous paragraphs and is a counterpart of Prop. 5.4.4, which dealt with ruled surface strips.

Proposition 7.1.27. *If $\mathbf{x}(u, v, w) = \mathbf{c}(u, v) + w\mathbf{g}(u, v)$ minimizes the functional \tilde{F} of Equ. (7.49), and interpolates given line segments, then \mathbf{c} is a thin plate spline surface and \mathbf{g} is a thin plate spline surface with tension.*

Although an explicit computation of the solution is possible, it is apparently faster to use a B-spline representation for \mathbf{c} and \mathbf{g} , and convert the interpolation problem into a least square approximation problem with (7.49) as a regularization term. This is a straightforward extension of the method described in Sec. 5.4.3.

If we are dealing with lines rather than line segments, a good central surface may be found similar to the ruled surface case via an appropriate variational problem.

Interpolatory Subdivision of Discrete Congruences

A discrete model of a parametrized line congruence $N(u, v)$ is a double-indexed finite collection R_{ij} of lines.

We assume that an area of interest is bounded by two director surfaces Φ and Ψ , and we let $\mathbf{p}_{ij} = R_{ij} \cap \Phi$, $\mathbf{q}_{ij} = R_{ij} \cap \Psi$. We define new points $\bar{\mathbf{p}}_{ij}$ by

$$\begin{aligned}\bar{\mathbf{p}}_{ij} = & \frac{1}{2}(\mathbf{p}_{i,j} + \mathbf{p}_{i+1,j+1}) + 2w(\mathbf{p}_{i+1,j} + \mathbf{p}_{i,j+1}) - \\ & - w(\mathbf{p}_{i,j-1} + \mathbf{p}_{i-1,j} + \mathbf{p}_{i+1,j+2} + \mathbf{p}_{i+2,j+1}).\end{aligned}\quad (7.51)$$

The number w serves as tension parameter. This procedure is called the *butterfly scheme* [39] and is a generalization of the interpolatory subdivision algorithm of Dyn et al. for curves (cf. Ex. 5.4.4). Fig. 7.17 shows how it works. The same is done with the points \mathbf{q}_{ij} to obtain $\bar{\mathbf{q}}_{ij}$. The result of subdivision are the lines $\bar{R}_{ij} = \bar{\mathbf{p}}_{ij} \vee \bar{\mathbf{q}}_{ij}$.

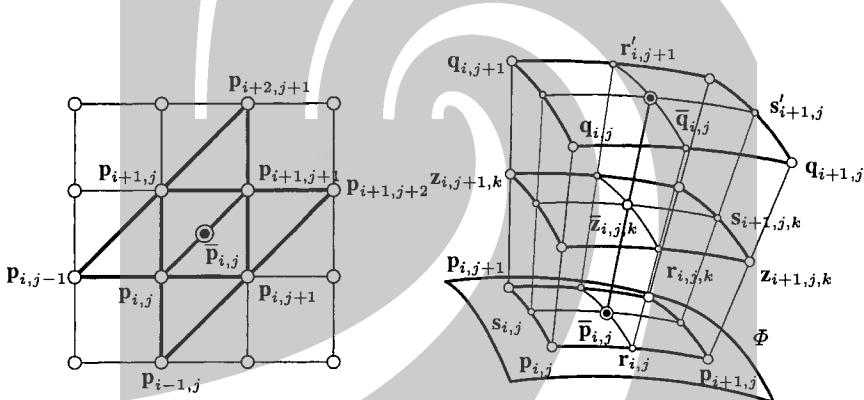


Fig. 7.17. Left: The butterfly scheme. Right: Subdividing a discrete line congruence, notations.

Variational Subdivision of Discrete Line Congruences

We are going to show how to combine variational methods with interpolatory subdivision (cf. [132]).

We assume that a discrete line congruence is given by lines L_{ij} with $i = 1, \dots, N_1$ and $j = 1, \dots, N_2$, and each line is spanned by two points \mathbf{p}_{ij} and \mathbf{q}_{ij} .

By applying the variational subdivision scheme described in Sec. 5.4.1 to the discrete ruled surfaces L_{ij} , j fixed, we insert, for all j , lines $R_{ij} = \mathbf{r}_{ij} \vee \mathbf{r}'_{ij}$ ($i = 1, \dots, N_1 - 1$). Similarly, by applying the same procedure to the discrete ruled surfaces L_{ij} , i fixed, we insert, for all i , new lines $S_{ij} = \mathbf{s}_{ij} \vee \mathbf{s}'_{ij}$ ($j = 1, \dots, N_2 - 1$) (see Fig. 7.17).

In a second step we insert lines $\bar{L}_{ij} = \bar{\mathbf{p}}_{ij} \vee \bar{\mathbf{q}}_{ij}$ ($i = 1, \dots, N_1 - 1, j = 1, \dots, N_2 - 1$) as follows: We divide the line segments $\overline{\mathbf{p}_{ij}\mathbf{q}_{ij}}, \overline{\mathbf{r}_{ij}\mathbf{r}'_{ij}}, \overline{\mathbf{s}_{ij}\mathbf{s}'_{ij}}$ into K equal parts by defining the points

$$\mathbf{z}_{ijk} = \lambda_k \mathbf{p}_{ij} + \mu_k \mathbf{q}_{ij}, \quad \mathbf{a}_{ijk} = \lambda_k \mathbf{r}_{ij} + \mu_k \mathbf{r}'_{ij}, \quad \mathbf{b}_{ijk} = \lambda_k \mathbf{s}_{ij} + \mu_k \mathbf{s}'_{ij},$$

with $\mu_k = \frac{k}{K}$, $\lambda_k = 1 - \mu_k$ for $k = 0, \dots, K$. Assume for the moment that we already know $\bar{\mathbf{p}}_{ij}$ and $\bar{\mathbf{q}}_{ij}$, and let

$$\bar{\mathbf{z}}_{ijk} = (1 - \frac{k}{K})\bar{\mathbf{p}}_{ij} + \frac{k}{K}\bar{\mathbf{q}}_{ij}. \quad (7.52)$$

We define the *umbrella vectors*

$$\begin{aligned} \mathbf{u}_{ijk} &= \bar{\mathbf{z}}_{ijk} - (1-w)(\mathbf{z}_{ijk} + \mathbf{z}_{i+1,j,k} + \mathbf{z}_{i+1,j+1,k} + \mathbf{z}_{i,j+1,k})/4 \\ &\quad - w(\mathbf{r}_{ijk} + \mathbf{r}_{i,j+1,k} + \mathbf{s}_{ijk} + \mathbf{s}_{i+1,j,k})/4, \end{aligned} \quad (7.53)$$

for $i = 1, \dots, N_1 - 1, j = 1, \dots, N_2 - 1, k = 0, \dots, K$. The parameter w determines the influence of the lines R_{ij} and S_{ij} . We define a sort of discrete energy functional by

$$F = \sum_{i,j,k} \mathbf{u}_{ijk}^2, \quad (7.54)$$

which is quadratic in the $(N_1 - 1)(N_2 - 1)$ unknown points $\bar{\mathbf{p}}_{ij}$ and $\bar{\mathbf{q}}_{ij}$. Minimization of F amounts to solving a system of linear equations.

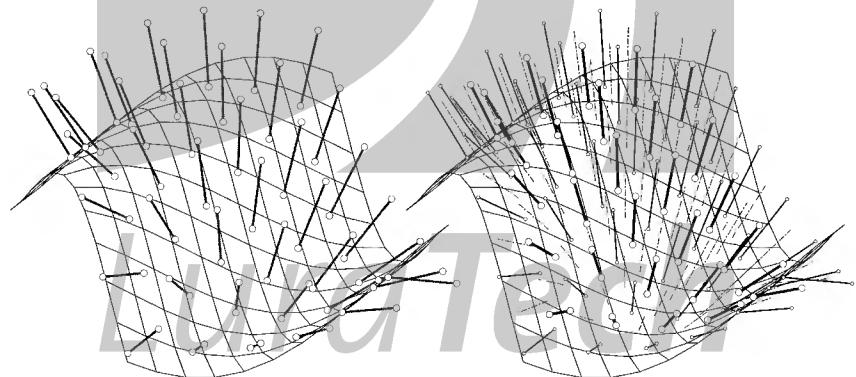


Fig. 7.18. Left: Discrete line congruence with director surface. Right: Result of interpolatory subdivision.

Example 7.1.17. We determine a line congruence $K(u_1, u_2)$ according to Equ. (7.4) by letting $\mathbf{a}(u_1, u_2) = (u_1, u_2, 0.5 \sin(2u_1 + 1.5u_2))$ and $\mathbf{x}(u_1, u_2) = (u_1, u_2, \sin u_1 + 1.3 \cos u_2)$, and we choose the 7×7 lines $L_{i,j}$ evenly distributed in the parameter domain $[-1.1, 1.1]^2$. The result of variational subdivision can be seen in Fig. 7.18. \diamond

7.1.7 Projection via Line Congruences

Motivated by the applications which are described below, we investigate the following mapping: Assume that a line congruence \mathcal{K} has the property that there is a director surface Φ which intersects all lines $L \in \mathcal{K}$ exactly once, and there is an open subset $D \subseteq P^3$ such that for all $\mathbf{x} \in D$ there is exactly one line $L_x \in \mathcal{K}$ which is incident with \mathbf{x} . Then we consider the mapping

$$\pi_{\mathcal{K}} : D \rightarrow \Phi, \mathbf{x} \mapsto L_x \cap \Phi,$$

which is called the projection via \mathcal{K} from D to Φ (see Fig. 7.19).

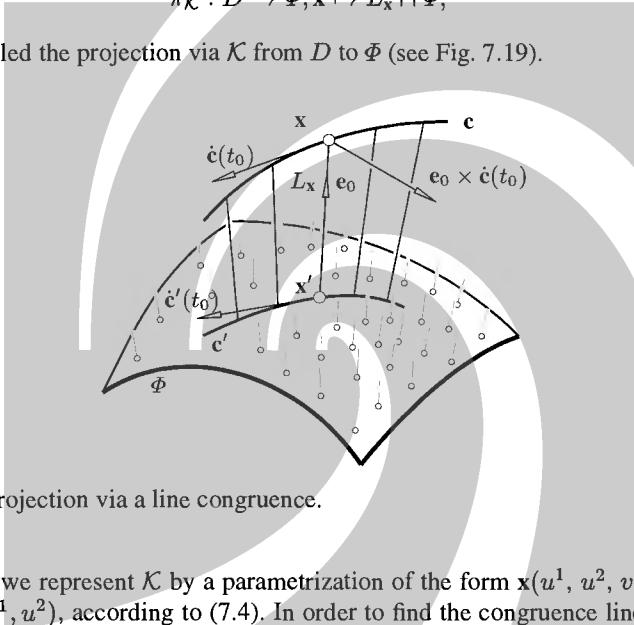


Fig. 7.19. Projection via a line congruence.

Locally we represent \mathcal{K} by a parametrization of the form $\mathbf{x}(u^1, u^2, v) = \mathbf{a}(u^1, u^2) + v\mathbf{e}(u^1, u^2)$, according to (7.4). In order to find the congruence line incident with a point \mathbf{x} , we have to solve

$$(\mathbf{x} - \mathbf{a}(u^1, u^2)) \times \mathbf{e}(u^1, u^2) = 0. \quad (7.55)$$

These are three scalar equations which are not independent. If no line of \mathcal{K} is parallel to the plane $x_3 = 0$, it is sufficient to consider the two equations which express the vanishing of the first two coordinates. We may also use the equivalent incidence conditions with Plücker coordinates. The image point \mathbf{x}' is found by evaluating $\mathbf{x}' = \mathbf{a}(u_s^1, u_s^2)$, where (u_s^1, u_s^2) satisfies (7.55). In practice this problem will have to be solved numerically.

If we want to compute numerically the projection of a *curve* $\mathbf{c}(t)$ onto the surface Φ via the congruence \mathcal{K} , we may compute a sufficiently dense sequence of image points and use interpolation afterwards. For that it is useful to know the tangent vectors of the image curve \mathbf{c}' (see Fig. 7.19).

In order to compute the tangent vector of \mathbf{c}' in the point $\mathbf{x}' = \mathbf{a}(u_s^1, u_s^2)$, assume that $\mathbf{c}'(t) = \mathbf{a}(u^1(t), u^2(t))$ and $\mathbf{c}(t) = \mathbf{c}'(t) + v(t)\mathbf{e}(u^1(t), u^2(t))$. Differentiation gives the equation

$$\dot{u}^1(\mathbf{a}_{,1} + v\mathbf{e}_{,1}) + \dot{u}^2(\mathbf{a}_{,2} + v\mathbf{e}_{,2}) + v\mathbf{e} = \dot{\mathbf{c}}.$$

To compute $\dot{u}^1(t_0)$, $\dot{u}^2(t_0)$, we multiply with $\mathbf{e}_0 = \mathbf{e}(u^1(t_0), u^2(t_0))$ and get the linear system of equations

$$\dot{u}^1(\mathbf{a}_{,1} + v\mathbf{e}_{,1}) \times \mathbf{e}_0 + \dot{u}^2(\mathbf{a}_{,2} + v\mathbf{e}_{,2}) \times \mathbf{e}_0 = \dot{\mathbf{c}}(t_0) \times \mathbf{e}_0, \quad (t = t_0)$$

which has rank two and allows to compute $\dot{u}^1(t_0)$ and $\dot{u}^2(t_0)$. With

$$\dot{\mathbf{c}}' = \dot{u}^1 \mathbf{a}_{,1} + \dot{u}^2 \mathbf{a}_{,2},$$

we have computed $\dot{\mathbf{c}}'(t_0)$.

Remark 7.1.21. A simple case which is important both theoretically and in applications is projection via a line congruence. It occurs naturally in the study of linear line mappings (see Chap. 8). \diamond

Remark 7.1.22. The computation of the image curve \mathbf{c}' of a curve \mathbf{c} under projection via a line congruence \mathcal{K} has the following interpretation: We intersect the complex \mathcal{C} of those lines which meet \mathbf{c} with the projection congruence \mathcal{K} . The intersection is the ruled surface \mathcal{R} of congruence lines incident with \mathbf{c} . Finally $\mathbf{c}' = \mathcal{R} \cap \Phi$.

If the complex \mathcal{C} has a simple equation of the form $\mathcal{C} : F(l_{01}, \dots, l_{12}) = 0$ in Plücker coordinates, the computation is easy. If $(l_{01}(u^1, u^2), \dots, l_{12}(u^1, u^2))$ is a parametrization of \mathcal{K} in Plücker coordinates, then

$$F(l_{01}(u^1, u^2), \dots, l_{12}(u^1, u^2)) = 0$$

is the preimage of \mathcal{R} in the parameter domain of \mathcal{K} . Computing this preimage amounts to computing the zero contour of a bivariate function. \diamond

Rolling Ball Blends

The following examples discuss several applications of the projection via a line congruence.

Example 7.1.18. Consider a sphere Φ of radius r which simultaneously rolls on two surfaces Ψ_1, Ψ_2 (see Fig. 7.20). The center of the sphere has distance r to both Ψ_1 and Ψ_2 , and the trajectory \mathbf{c} of the center is contained in the intersection of the offset surfaces of Ψ_1 and Ψ_2 at distance r (see Fig. 7.20). Of course, the intersection can consist of several branches — here we consider only one smooth branch.

The envelope surface generated by the rolling sphere is a *pipe surface* with spine curve \mathbf{c} , which touches each surface Ψ_i in the points of a curve \mathbf{c}'_i . This curve is the projection of the spine curve \mathbf{c} onto Ψ_i via the normal congruence \mathcal{N}_i of Ψ_i .

In geometric modeling a part of the pipe surface bounded by $\mathbf{c}'_1, \mathbf{c}'_2$ is used as smooth transition surface (*rolling ball blend*) between Ψ_1 and Ψ_2 near their intersection. \diamond

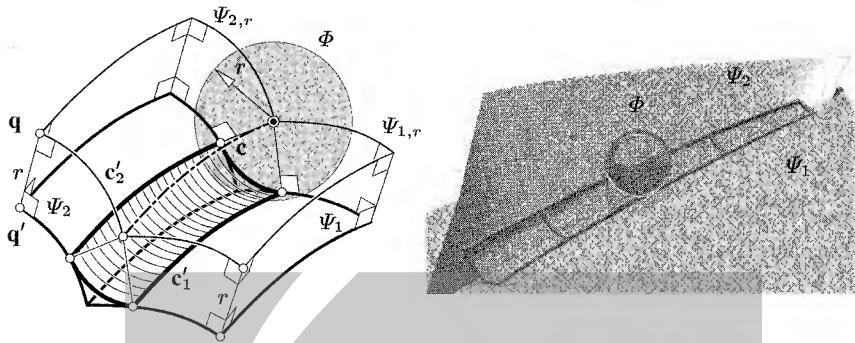


Fig. 7.20. Left: Construction of a rolling ball blend. Right: Rendered image of a rolling ball blend (courtesy N. Pomaroli).

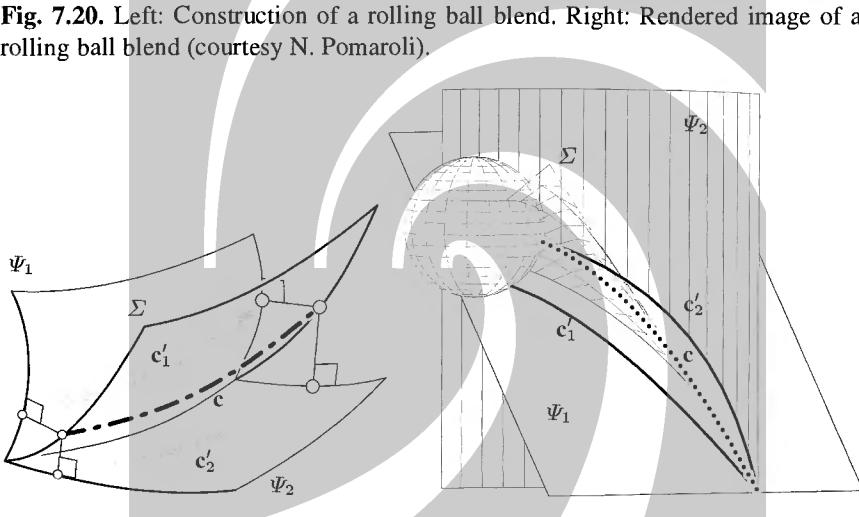


Fig. 7.21. Left: Construction of a variable radius rolling ball blend. Right: Variable radius rolling ball blend which joins a cylinder and a plane.

Example 7.1.19. A generalization of the rolling ball blend is the so-called *variable radius rolling ball blend* (see [116] and the references therein). There, the moving sphere also changes its size, while maintaining tangency to Ψ_1 and Ψ_2 . The center of the sphere is at equal distance to both surfaces Ψ_1 and Ψ_2 , and is therefore contained in the bisector surface Σ of Ψ_1 , Ψ_2 .

For the construction of the blending surface we can prescribe an appropriate curve c in Σ . This defines a family of spheres, which are centered in c and touch both Ψ_1 and Ψ_2 . The envelope of this sphere family is a *canal surface*. It is tangent to Ψ_i in the points of a curve c'_i , which is obtained by projection of c into Ψ_i via the normal congruence \mathcal{N}_i of Ψ_i ($i = 1, 2$). The curve c'_i is the transition curve between the blending surface and the given surfaces Ψ_i .

The computation of the bisector surface (see Remark 7.1.12) is in general possible only in a numerical way, and the same holds for the projection via the normal

congruence. In special cases, however, we can get simple variable radius rolling ball blends, for example blending surfaces which are part of Dupin cyclides [162]. \diamond

Reflection Lines and Isophotes

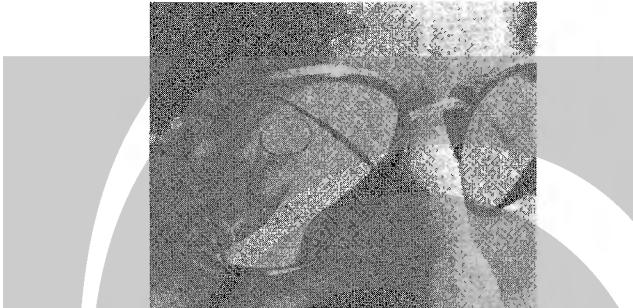


Fig. 7.22. Left: Reflection in a sphere (courtesy G. Glaeser).

Example 7.1.20. Consider a surface Φ and the line congruence \mathcal{K} obtained by reflecting a bundle with vertex v in Φ . Let L be a straight line and L' its projection via \mathcal{K} into Φ . This curve is called a *reflection line* for obvious reasons (see Fig. 7.22, [63]).

These reflection lines are very sensitive to changes in the derivatives of the surface Φ , and thus they are used to inspect the quality of a designed surface in automotive industry [78]. If the surface is C^r for $r \geq 1$, then the reflection lines are C^{r-1} . Fig. C.12 shows how the reflection of a sequence of parallel lines can be used to detect curvature discontinuities of a surface.

The computation of a reflection line is a contouring problem as described in Remark 7.1.22, since the lines of \mathcal{K} meeting L fulfill a linear equation in Plücker coordinates. \diamond

Example 7.1.21. Sometimes also *isophotes*, which we have discussed earlier, are used for quality inspection of surfaces [78]. In fact, isophotes are also images under projection via a suitable line congruence: Fix a line L , an angle α and a surface Φ . Then the isophote of Φ for parallel illumination parallel to L and angle α is the set of points of Φ such that the angle enclosed by the surface normal and the line L equals α .

The set of lines G with $\sphericalangle(G, L) = \alpha$ is a line complex \mathcal{C} (a complex of constant slope, see Remark 6.3.5). The ideal points of these lines are contained in a conic c_ω at infinity. Thus the isophote equals the projection of c_ω onto Φ via Φ 's normal congruence. \diamond

Accessibility Areas

Our final example concerns a recent contribution by G. Elber [44] on verification in 5-axis freeform milling environments. We assume a cutter, which consists of a cutting head and a cylindrical cutting shaft (Fig. 7.23, left) of radius r . The surface which is to be shaped by the cutter is denoted by Ψ , and additionally we have an obstacle surface (check surface) Φ . We want to know which parts of the surface Ψ are accessible for the cutter without collision between shaft and check surface, if the cutter axes are contained in some prescribed line congruence \mathcal{K} .

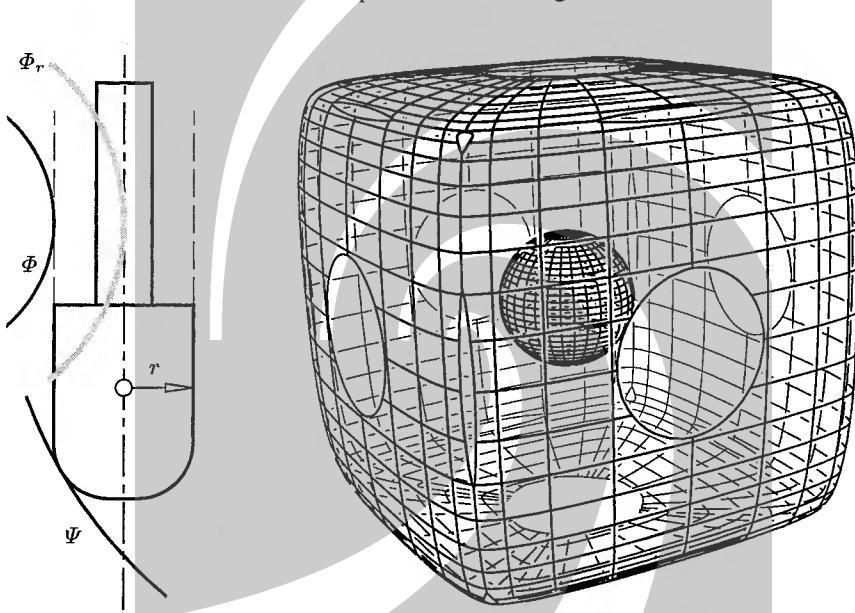


Fig. 7.23. Left: Cutter with cylindrical shaft and principle of check surface. Right: Spherical check surface Φ and rounded box surface Ψ to be machined with cutter axes orthogonal to Ψ . The inaccessible regions are trimmed away (courtesy G. Elber)

In order to eliminate the influence of the shape of the cutter head, we define the *accessible region* of Ψ as the union of points of Ψ which can be incident with the cutter axis. We want to determine the accessible region, especially its boundary curve.

In order to convert this problem into a simpler one, we shrink the cutter shaft to its axis, and simultaneously enlarge Φ to its offset surface Φ_r at distance r — if the actual cutter touches Φ , its axis touches Φ_r .

Therefore the boundary of the accessible region for NC machining with axis congruence \mathcal{K} is obtained as follows: compute all axis positions which are tangent

to Φ_r and intersect the obtained ruled surface with Ψ . The intersection curve is the *silhouette* of Φ_r for projection via \mathcal{K} .

Computation leads to a contouring problem (cf. Remark 7.1.22), if the line complex of Ψ_r 's surface tangents is described by one equation $F(l_{01}, \dots, l_{12})$ in Plücker coordinates. If not, we have to solve the system of equations which expresses tangency of a line in \mathcal{K} with Ψ_r [44]. A result of this algorithm is shown in Fig. 7.23, right.

7.2 Line Complexes

Three-parameter families of lines in projective three-space P^3 are called *line complexes*. We will study them both from the smooth and from the algebraic viewpoint.

7.2.1 Differential Geometry of Line Complexes

We first consider *smooth* line complexes and use the notion of ‘ k -surface’ (see Equ. (1.84)) to define

Definition. A set \mathcal{C} of lines is called a C^r line complex, if its Klein image $\mathcal{C}\gamma$ is a C^r 3-surface in the Klein quadric.

Remark 7.2.1. Sometimes a set \mathcal{C} of lines has the property that $\mathcal{C}\gamma$ locally is a regular 3-surface, but there is no global regular parametrization of $\mathcal{C}\gamma$ in the form (1.84). An example of such a set of lines is a linear line complex. Its Klein image $\mathcal{C}\gamma$ is a quadric, which has locally a parametrization as a regular 3-surface.

When we call such a surface a C^r line complex, we actually mean a subset \mathcal{C}' such that $\mathcal{C}'\gamma$ is a 3-surface. We will only consider *local* properties of line complexes, so this imprecise way of speaking will do no harm. \diamond

A parametrization of a smooth line complex \mathcal{C} has the form

$$(u^1, u^2, u^3) \in D \mapsto C\gamma(u^1, u^2, u^3) \in \mathcal{C}\gamma \subset M_2^4 \subset P^5. \quad (7.56)$$

Here $C(u^1, u^2, u^3)$ is a line of \mathcal{C} and $C\gamma$ is its Klein image. We use the abbreviation $u = (u^1, u^2, u^3)$ and assume that

$$C\gamma(u^1, u^2, u^3) = (\mathbf{c}(u^1, u^2, u^3), \bar{\mathbf{c}}(u^1, u^2, u^3))\mathbb{R}, \quad (7.57)$$

with two vector-valued trivariate C^r functions \mathbf{c} and $\bar{\mathbf{c}}$. If $(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R}$ is the Klein image of a line, it must fulfill the relations

$$\begin{aligned} 0 &= \mathbf{c} \cdot \bar{\mathbf{c}} = \mathbf{c}_{,j} \cdot \bar{\mathbf{c}} + \mathbf{c} \cdot \bar{\mathbf{c}}_{,j} = \mathbf{c}_{,jk} \cdot \bar{\mathbf{c}} + \mathbf{c}_{,j} \cdot \bar{\mathbf{c}}_{,k} + \mathbf{c}_{,k} \cdot \bar{\mathbf{c}}_{,j} + \mathbf{c} \cdot \bar{\mathbf{c}}_{,jk} \\ &= \mathbf{c}_{,jkl} \bar{\mathbf{c}} + \mathbf{c}_{,jk} \bar{\mathbf{c}}_{,l} + \mathbf{c}_{,jl} \bar{\mathbf{c}}_{,k} + \mathbf{c}_{,j} \bar{\mathbf{c}}_{,kl} + \mathbf{c}_{,kl} \bar{\mathbf{c}}_{,j} + \mathbf{c}_{,k} \bar{\mathbf{c}}_{,jl} + \mathbf{c}_{,l} \bar{\mathbf{c}}_{,jk} + \mathbf{c} \bar{\mathbf{c}}_{,jkl} \\ &= \dots, \end{aligned}$$

the first of which is the Plücker identity (2.3) and the others are found by differentiation.

We say that a linear complex \mathcal{C} is *regular* at a line $L \in \mathcal{C}$, if $\mathcal{C}\gamma$ is a regular 3-surface in the point $L\gamma$. This means that the derivative points span a projective three-space, the *tangent space* of $\mathcal{C}\gamma$ in the point $L\gamma$. If we assume that \mathcal{C} is parametrized by (7.57), then its tangent space T^3 equals

$$(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} \vee (\mathbf{c}, 1, \bar{\mathbf{c}}, 1)\mathbb{R} \vee (\mathbf{c}, 2, \bar{\mathbf{c}}, 2)\mathbb{R} \vee (\mathbf{c}, 3, \bar{\mathbf{c}}, 3)\mathbb{R}.$$

T^3 is a projective subspace of P^5 , but it cannot be used to define a ‘linear line complex tangent to \mathcal{C} ’, because it is only three-dimensional.

However there is a linear congruence \mathcal{T} such that $\mathcal{T}\gamma = T^3 \cap M_2^4$. We recall from Sec. 3.2 that we can determine the type of \mathcal{T} by studying the polar subspace $T^3\mu_2^4$. The inclusion $L\gamma \in T^3$ implies $T^3\mu_2^4 \subset L\gamma\mu_2^4$. The tangent hyperplane of the Klein quadric in $L\gamma$ is $L\gamma\mu_2^4$. Thus we have $T^3 \subset L\gamma\mu_2^4$, which implies $L\gamma \in T^3\mu_2^4$. As T^3 is tangent to the Klein quadric in $L\gamma$, its intersection with this quadric has a singular point. There are two possibilities:

1. $T^3\mu_2^4 \not\subset T^3$: The space T^3 intersects the Klein quadric in a quadratic cone, and the congruence \mathcal{T} is parabolic with axis L .
2. $T^3\mu_2^4 \subset T^3$: The space T^3 intersects the Plücker quadric in two generator planes, and \mathcal{T} consists of a bundle and a field of lines.

We call the linear congruence \mathcal{T} *tangent* to the complex \mathcal{C} at the line L .

Intersection of Line Complexes

It is well known that the intersection of a regular m -surface Φ and a regular l -surface Ψ in an n -dimensional affine or projective space is a regular $(m + l - n)$ -surface, if for all points of $\Phi \cap \Psi$ the span of the two tangent spaces of Φ and Ψ has dimension n . We say that Φ and Ψ intersect *transversely*. Then the tangent space of $\Phi \cap \Psi$ is the intersection of the respective tangent spaces of Φ and Ψ in the point of intersection.

We will use this result to obtain information on the intersection of line complexes, but first we mention another important application: Since the Klein quadric has dimension four, and a line complex has dimension three, we expect that *one* homogeneous equation of the form $F(l_{01}, \dots, l_{12}) = F(\mathbf{c}, \bar{\mathbf{c}}) = 0$ in Plücker coordinates determines a line complex:

$$\mathcal{C}\gamma = \{(\mathbf{c}, \bar{\mathbf{c}})\mathbb{R} \mid \mathbf{c} \cdot \bar{\mathbf{c}} = F(\mathbf{c}, \bar{\mathbf{c}}) = 0\}. \quad (7.58)$$

Proposition 7.2.1. *Equ. (7.58) defines a regular line complex if F is homogeneous (i.e., $F(\lambda\mathbf{c}, \lambda\bar{\mathbf{c}}) = \lambda^m F(\mathbf{c}, \bar{\mathbf{c}})$ for some m), and if the vectors*

$$(\bar{\mathbf{c}}, \mathbf{c}), \text{ grad}_F(\mathbf{c}, \bar{\mathbf{c}}) \quad (7.59)$$

are linearly independent.

Proof. The equation $F(\mathbf{c}, \bar{\mathbf{c}}) = 0$ determines a regular five-surface $\tilde{\Phi}$ in \mathbb{R}^6 , if $\text{grad}_F \neq \mathbf{0}$. If $(\mathbf{x}, \bar{\mathbf{x}}) \in \tilde{\Phi}$, then also $(\lambda \mathbf{x}, \lambda \bar{\mathbf{x}}) \in \tilde{\Phi}$, because F is homogeneous. Therefore $F = 0$ also defines a regular four-surface Φ in P^5 . Its (projective) tangent space U is four-dimensional and has the equation $(\mathbf{x}, \bar{\mathbf{x}}) \cdot \text{grad}_F = 0$. The tangent space of the Klein quadric in $(\mathbf{c}, \bar{\mathbf{c}})$ has the equation $\mathbf{x} \cdot \bar{\mathbf{c}} + \bar{\mathbf{x}} \cdot \mathbf{c} = (\mathbf{x}, \bar{\mathbf{x}}) \cdot (\bar{\mathbf{c}}, \mathbf{c}) = 0$. This shows that the intersection of $\Phi \cap M_2^4$ is transverse if and only if the vectors $(\bar{\mathbf{c}}, \mathbf{c})$ and $\text{grad}_F(\mathbf{c}, \bar{\mathbf{c}})$ are linearly independent. \square

It is straightforward to generalize this result to intersection of several line complexes.

Proposition 7.2.2. *Assume that the line complex \mathcal{C}_i ($i = 1, 2, \dots$) is regular and is parametrized according to (7.57) by $(\mathbf{c}_i(u), \bar{\mathbf{c}}_i(u))\mathbb{R}$. Its tangent space T_i is given by*

$$T_i = (\mathbf{c}_i, \bar{\mathbf{c}}_i)\mathbb{R} \vee (\mathbf{c}_{i,1}, \bar{\mathbf{c}}_{i,1})\mathbb{R} \vee \dots \vee (\mathbf{c}_{i,3}, \bar{\mathbf{c}}_{i,3})\mathbb{R},$$

and is contained in the tangent space T_0 of the Klein quadric, whose equation is $\mathbf{x} \cdot \bar{\mathbf{c}} + \bar{\mathbf{x}} \cdot \mathbf{x} = 0$. $\mathcal{C}_1 \cap \mathcal{C}_2$ is a regular line congruence, if $T_1 \vee T_2 = T_0$. The intersection $\mathcal{C}_1 \cap \mathcal{C}_2 \cap \mathcal{C}_3$ is a regular ruled surface, if $T_1 \vee T_2 = T_0$ and $(T_1 \cap T_2) \vee T_3 = T_0$.

If the complexes \mathcal{C}_i are given by the equations $F_i(\mathbf{c}, \bar{\mathbf{c}}) = 0$, the same result holds if the vectors

$$(\bar{\mathbf{c}}, \mathbf{c}), \text{ grad}_{F_1}(\mathbf{c}, \bar{\mathbf{c}}), \dots, \text{ grad}_{F_i}(\mathbf{c}, \bar{\mathbf{c}}),$$

are linearly independent for $i = 2$ or $i = 3$, respectively.

Proof. The proof is analogous to the proof of Prop. 7.2.1. We have to verify repeatedly that the linear span of two tangent spaces is five-dimensional, and write down the intersection of these two tangent spaces. The details are left to the reader. \square

Complex Cones

In order to get an impression how a three-parameter manifold of lines looks like, we study some objects associated with a line complex \mathcal{C} . We have already defined some of them in Chap. 3, p. 161, but we will repeat the definitions here.

A ruled surface \mathcal{R} which is contained in a line complex \mathcal{C} is called a *complex surface* of \mathcal{C} . The Klein image $\mathcal{R}\gamma$ is a curve contained in the three-surface $\mathcal{C}\gamma$.

A special complex surface is the following: The set \mathcal{C}_p of lines in \mathcal{C} which is incident with a point p is called the *complex cone* with vertex p . If the bundle of lines with vertex p is denoted by \mathcal{K}_p , then the Klein image of the complex cone equals the intersection $\mathcal{C}_p\gamma = \mathcal{K}_p\gamma \cap \mathcal{C}\gamma$. This is a planar section of $\mathcal{C}\gamma$, because $\mathcal{K}_p\gamma$ is a plane in the Klein quadric.

Example 7.2.1. A linear line complex \mathcal{C} which is regular in the sense of Chap. 3 has a complex cone \mathcal{C}_p which is a line pencil. If \mathcal{C} is singular, then it consists of all lines which meet an axis G . If $p \notin G$, the cone \mathcal{C}_p is a pencil again, and if $p \in G$, the set \mathcal{C}_p is an entire bundle of lines. \diamond

This example shows that the set \mathcal{C}_p need not be a regular cone, not even if \mathcal{C} is regular as a three-surface. There is the following result, which describes the first order behaviour of a line complex in terms of its complex cones.

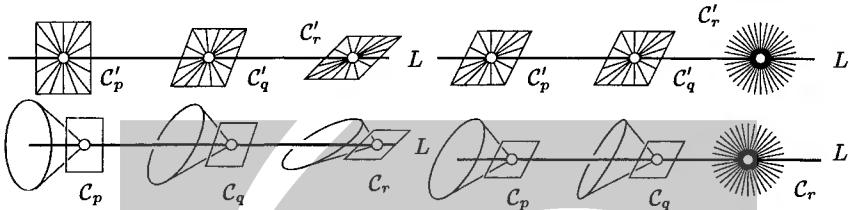


Fig. 7.24. Complex cones for points $p, q, r \in L$. Top: Complex cones for regular (left) and for singular (right) linear line complex C' . Bottom: Complex cones for line complex C — regular case (left) and singular case (right).

Lemma 7.2.3. Consider a line complex \mathcal{C} such that \mathcal{C}_γ is regular at L_γ , and such that the linear congruence \mathcal{T} tangent to \mathcal{C} in L is parabolic. Then the complex cone \mathcal{C}_p is regular in a neighbourhood of L for all points $p \in L$, and the pencils of \mathcal{T} are tangent to the cones \mathcal{C}_p .

If \mathcal{T} consists of a bundle with vertex p_0 and the field of lines of a plane π , then the complex cones \mathcal{C}_p are regular in a neighbourhood of L for all $p \in L, p \neq p_0$, and π is their common tangent plane.

Proof. As all properties involved depend on the first derivatives, we may replace \mathcal{C} by a linear line complex \mathcal{C}' whose tangent space intersects the Klein quadric in the same set of lines as \mathcal{C} does. It follows from Ex. 7.2.1 and the detailed description of linear line complexes in Sec. 3.1.3 that the lemma is valid for linear complexes. \square

Planar Complex Curves

The set \mathcal{C}_π of lines of a line complex \mathcal{C} which are contained in a plane π corresponds, via the Klein mapping, to the intersection of \mathcal{C}_γ with a plane \mathcal{K}_γ , where \mathcal{K} is the field of lines of the plane π . In this way \mathcal{C}_π is an object *dual* to the complex cone \mathcal{C}_p .

By planar duality, \mathcal{C}_π is the set of tangents of a curve, which is called the *planar complex curve* of the complex \mathcal{C} in the plane π . Of course this is valid only if \mathcal{C}_π is actually a regular smooth one-parameter family of lines, so its envelope can be defined.

The first order behaviour of the planar complex curves in the neighbourhood of a line L is described analogously to Lemma 7.2.3. Here ‘neighbourhood of L ’ means the following: The line L is tangent to the planar complex curve, and the neighbourhood is defined in terms of tangents.

Lemma 7.2.4. Consider a line complex \mathcal{C} such that \mathcal{C}_γ is a regular three-surface in the point L_γ , and the linear congruence \mathcal{T} tangent to \mathcal{C} in L is parabolic.

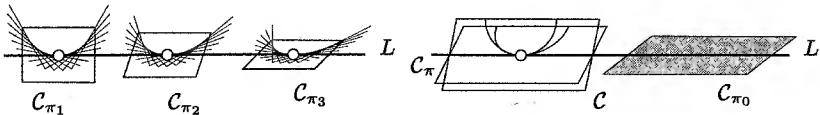


Fig. 7.25. Complex curves for a line complex \mathcal{C} . regular case (left) and singular case (right).

Then \mathcal{C}_π is regular in a neighbourhood of L as a dual curve for all planes $\pi \supset L$. If $p(\pi)$ is L 's contact point with \mathcal{C}_π , the mapping $p \mapsto \pi(p)$ is a projective mapping.

If \mathcal{T} consists of a bundle with vertex p_0 and the field of lines of the plane π_0 , then \mathcal{C}_π is regular in a neighbourhood of L for all $\pi \supset L, \pi \neq \pi_0$. L 's contact point is p_0 .

Proof. The proof involves first order derivatives, so we can replace \mathcal{C} by any linear line complex \mathcal{C}' which has the same tangent space. For linear complexes, the lemma is true. \square

Complex Curves and Complex Developables

A ruled surface \mathcal{R} contained in a line complex may be torsal. If it is, its curve of regression is called a *complex curve*. Conversely, a curve c which has the property that all of its tangents are contained in \mathcal{C} defines a tangent developable \mathcal{R} in \mathcal{C} .

Special cases are the complex cone \mathcal{C}_p , whose curve of regression is degenerate (it is the constant curve p), and the planar ruled surface \mathcal{C}_π , whose curve of regression is the planar complex curve in π .

A complex curve c is, by its definition, tangent to the complex cone \mathcal{C}_p if $p \in c$. We can show even more:

Proposition 7.2.5. *If c is a complex curve of the line complex \mathcal{C} , then the osculating plane of c in a point $p \in c$ is tangent to the complex cone \mathcal{C}_p .*

Proof. This is a statement which involves second order derivatives. However it is easy to transform it into a statement which involves only first order derivatives: Consider c 's tangent developable \mathcal{R} . Then \mathcal{R}_γ is a curve in \mathcal{C}_γ . By definition of torsal generator, \mathcal{R}_γ 's tangent is contained in the Klein quadric and we denote it by \mathcal{R}'_γ . Then \mathcal{R}' is a pencil which is in first order contact with \mathcal{R} . The plane of this pencil is \mathcal{R} 's tangent plane in the points of L , and it equals c 's osculating plane.

Obviously $\mathcal{R}' \subset \mathcal{T}$, where \mathcal{T} is the linear congruence tangent to \mathcal{C} , and we have already seen that the pencils in \mathcal{T} are tangent to the complex cones. \square

We are going to describe some examples of complex curves which lead to nice corollaries.

Example 7.2.2. We consider a *linear* line complex \mathcal{C} and the null polarity ν associated with \mathcal{C} . By Cor. 3.1.7, there is a helical motion such that $p\nu$ is the path normal plane of the point p , for all $p \in E^3$.

Consider a complex curve c . Prop. 7.2.5 says that the osculating plane of c in $p \in c$ is tangent to the complex cone \mathcal{C}_p , i.e., to the pencil incident with p and contained in the plane $\pi = p\nu$. This is a complicated way to say that $p\nu$ is c 's osculating plane in the point p . \diamond

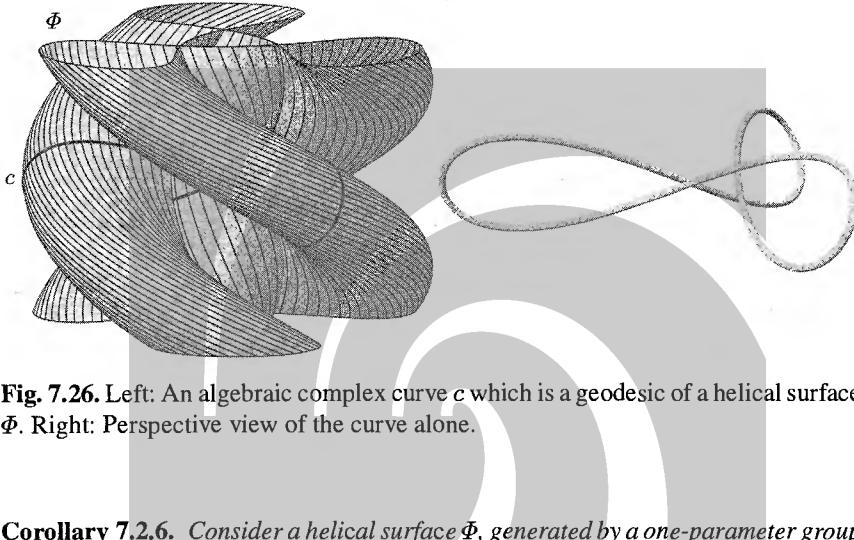


Fig. 7.26. Left: An algebraic complex curve c which is a geodesic of a helical surface Φ . Right: Perspective view of the curve alone.

Corollary 7.2.6. Consider a helical surface Φ , generated by a one-parameter group of helical motions $\alpha(t)$. If c is an orthogonal trajectory of the helices in Φ , then c is a geodesic (see Fig. 7.26).

Proof. Consider the linear complex \mathcal{C} of path normals of $\alpha(t)$. Obviously c is a complex curve of \mathcal{C} . We have seen (cf. Ex. 7.2.2) that c 's osculating planes are orthogonal to Φ . This is a characterization of geodesics. \square

Example 7.2.3. The polynomial cubic $c : u \mapsto c(u) = \mathbf{c}(u)\mathbb{R} = (1, u, u^2, u^3)$ is a complex curve for the linear complex \mathcal{C} with $\mathcal{C}\gamma^* = (\mathbf{a}, \bar{\mathbf{a}})\mathbb{R} = (0, 0, 3, 0, 0, -1)\mathbb{R}$. To see this, we compute the tangent $T(u)$ of c in $c(u)$: $T(u)\gamma = (\dot{\mathbf{c}}(u), \mathbf{c}(u)) \times \ddot{\mathbf{c}}(u)\mathbb{R} = (1, 2u, 3u^2, u^4, -2u^3, u^2) = (\mathbf{t}(u), \bar{\mathbf{t}}(u))\mathbb{R}$. Obviously $\mathbf{a} \cdot \bar{\mathbf{t}} + \bar{\mathbf{a}} \cdot \mathbf{t} = 0$.

We have seen in Ex. 7.2.2 that c 's osculating plane in the point $c(u)$ equals $c(u)\nu$, where ν is the null polarity associated with \mathcal{C} . By Lemma 3.1.2 and Equ. (3.3), $(p_0, \dots, p_3)\mathbb{R}\nu$ equals the plane $\mathbb{R}(-p_3, 3p_2, -3p_1, p_0)$. \diamond

Corollary 7.2.7. The inflection points of a rational planar cubic are collinear.

Proof. We first show that a central projection (denoted by a prime) maps the cubic c of Ex. 7.2.3 to a planar cubic c' with collinear inflection points, if the projection center Z is not contained in the cubic (see Fig. 7.27).

Prop. 1.2.3 says that I' is an inflection point of c' if Z is contained in the osculating plane of c in I . By Ex. 7.2.3, this osculating plane equals $I\nu$. So I' is an

inflection point if and only if $I \in Z\nu$. As $Z \in Z\nu$, the plane $Z\nu$ projects to a line $Z\nu'$, which contains all inflection points. The statement we wanted to prove now follows immediately from Lemma 1.4.11. \square

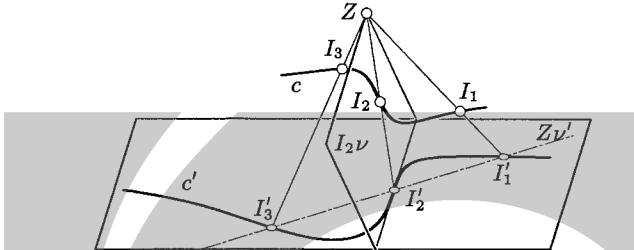


Fig. 7.27. Rational cubic c and rational planar cubic c' with three collinear inflection points I'_1, I'_2, I'_3 .

Euclidean Differential Geometry of Complexes

A smooth line complex C has Euclidean differential invariants which are functions of its lines, like differential invariants of surfaces, which are functions of its points. We will not discuss these invariants in general.

One invariant however can be defined without much effort: The linear congruence tangent to C in a line L is an invariant of projective differential geometry (see p. 475), and its distribution parameter is an invariant of Euclidean differential geometry. We call it the distribution parameter of C in L .

7.2.2 Algebraic Complexes and Congruences

In Sec. 5.2 we defined the notion of ‘algebraic ruled surface’. Here we are going to define, more generally, algebraic congruences and complexes. When studying algebraic complexes, we almost always assume that the ground field is \mathbb{C} .

Definition. Consider real or complex projective three-space, its set of lines \mathcal{L} , and the corresponding Klein mapping γ . Then $C \subset \mathcal{L}$ is algebraic, if $C\gamma$ is an algebraic variety. It is called irreducible, if $C\gamma$ is. An irreducible algebraic set of lines is called ruled surface, line congruence, or line complex, if its dimension equals one, two, or three, respectively.

We expect that one homogeneous polynomial equation in Plücker coordinates defines an algebraic complex, if this equation is no power of the Plücker identity. Even the converse is true, which is a consequence of the following results, the second of which is presented without proof:

Lemma 7.2.8. An $(n - 1)$ -dimensional algebraic variety $M \subseteq \mathbb{C}P^n$ ($n \geq 1$) is the zero set of one homogeneous polynomial, i.e., is an algebraic hypersurface (this has already been stated in Prop. 1.3.20).

Proof. (Sketch) Assume that M is irreducible and is defined by equations $F_1 = \dots = F_k = 0$. We may assume that none of the F_i is redundant, and that all F_i are irreducible. Let $M' = V(F_1)$. Then $M \subseteq M'$ and M' is irreducible. The chain of irreducible varieties $M \subseteq M' \subsetneq \mathbb{C}P^n$ and $\dim M = n - 1$ show that $M = M'$, so $M = V(F_1)$.

To show the lemma for reducible M , we show it for each component separately, and multiply the respective polynomials. \square

Lemma 7.2.9. (F. Klein) If Φ is a quadric in $\mathbb{C}P^n$ ($n \geq 4$) defined by the quadratic equation $F(\mathbf{x}) = 0$, and $\Psi \subset \Phi$ is an $(n - 2)$ -dimensional algebraic variety, then there is a homogeneous polynomial G such that Ψ has the equation $F(\mathbf{x}) = G(\mathbf{x}) = 0$.

Remark 7.2.2. Lemma 7.2.9 is not true for $n = 3$, because a cubic c contained in a ruled quadric Φ is *not* the intersection of Φ with another algebraic variety Φ' :

Almost all test planes intersect the cubic in three points, the quadric Φ in a conic, and the surface Φ' in a planar algebraic curve of degree m . By Bézout's theorem (Th. 1.3.23), a test plane intersects $\Phi \cap \Psi$ in $2m$ points, counting multiplicities. As 3 is an odd number, $c \neq \Phi \cap \Psi$. \diamond

Proposition 7.2.10. For an algebraic line complex \mathcal{C} in complex projective three-space there is a homogeneous polynomial F of six indeterminates such that precisely the lines of \mathcal{C} have Plücker coordinates $(\mathbf{x}, \bar{\mathbf{x}})$ which fulfill $F(\mathbf{x}, \bar{\mathbf{x}}) = 0$.

Conversely, any such polynomial which is not a power of the Plücker relation Ω_q (cf. Equ. (2.3)) defines an algebraic line complex.

Proof. The first part follows immediately from Lemma 7.2.9, because $\mathcal{C}\gamma$ is a three-dimensional algebraic variety contained in the Klein quadric.

The converse is easy: If we denote the polynomial $\mathbf{x} \cdot \bar{\mathbf{x}}$ with Ω_q , then $\Omega_q(\mathbf{x}, \bar{\mathbf{x}}) = 0$ is the equation of the Klein quadric. Clearly Ω_q is irreducible, so Prop. 1.3.10 implies that a homogeneous polynomial F which vanishes in all points of M_2^4 is a power of Ω_q . Thus there is a point $L\gamma \in M_2^4$ which is not in the zero set $M = V(F, \Omega_q)$. Since M_2^4 is irreducible, $\dim(M) \leq 3$, and because M is defined by two equations, $\dim(M) \geq 3$. Thus $\dim(M) = 3$, and the proof is complete. \square

This result considerably simplifies the discussion of algebraic line complexes.

If \mathcal{C} is an algebraic line complex, then $\mathcal{C}\gamma$ is the zero set $F = \Omega_q = 0$, and therefore equals the zero set of the ideal

$$I_{\mathcal{C}\gamma} = \langle F, \Omega_q \rangle_{\mathbb{C}[\mathbf{x}, \bar{\mathbf{x}}]} = \{gF + h\Omega_q \mid g, h \in \mathbb{C}[\mathbf{x}, \bar{\mathbf{x}}]\}.$$

Without loss of generality we may assume that F has minimal degree among all homogeneous polynomials which generate $I_{\mathcal{C}\gamma}$ together with the Plücker relation. Its degree then is called the *degree* $\deg(\mathcal{C})$ of the algebraic complex \mathcal{C} .

Remark 7.2.3. If \mathcal{C} is an algebraic complex of degree n , then the set $\mathcal{C}\gamma$ has degree $2n$ as a projective algebraic variety. Therefore Prop. 7.2.10 shows that all three-dimensional subvarieties of the Klein quadric are of even degree. \diamond

Complex Cones of Algebraic Complexes

We look for a geometric interpretation of the degree of an algebraic complex \mathcal{C} . Clearly a test two-space in \mathbb{CP}^5 intersects $\mathcal{C}\gamma$ in $\deg(\mathcal{C}\gamma)$ points, and a test three-space in \mathbb{CP}^5 intersects $\mathcal{C}\gamma$ in an algebraic curve of degree $\deg(\mathcal{C}\gamma)$. We have already mentioned (Remark 7.2.3) that this degree is twice the degree of \mathcal{C} .

If \mathcal{K}_p is the bundle of lines incident with p , then the complex cone \mathcal{C}_p is defined as $\mathcal{C}_p\gamma = \mathcal{C}\gamma \cap \mathcal{K}_p\gamma$. The set $\mathcal{K}_p\gamma$ is a three-space, but we cannot use only these special three-spaces as test spaces to compute the degree of $\mathcal{C}\gamma$ — they are far from constituting ‘almost all’ three-dimensional subspaces of \mathbb{CP}^5 . Indeed, the following lemma shows that the planes $\mathcal{K}_p\gamma$ are special planes with respect to $\mathcal{C}\gamma$. Recall that the degree of a cone is its degree as a planar curve in dual projective space.

Lemma 7.2.11. *If \mathcal{C} is an algebraic complex of degree h , then almost all complex cones \mathcal{C}_p are of degree h .*

Proof. We assume that \mathcal{C} has the equation $F(\mathbf{c}, \bar{\mathbf{c}}) = \mathbf{c} \cdot \bar{\mathbf{c}} = 0$ in Plücker coordinates. We apply a coordinate transform κ such that $p\kappa = (1, 0, 0, 0)\mathbb{C}$. The polynomial F is transformed to another polynomial \tilde{F} of the same degree.

A line is incident with $p\kappa$ if its Plücker coordinates are (\mathbf{c}, \mathbf{o}) , and $\mathcal{C}_{p\kappa}$ has the equation $\tilde{F}'(\mathbf{c}) = \tilde{F}(\mathbf{c}, \mathbf{o}) = 0$. For almost all coordinate transforms, the polynomial \tilde{F} is of degree h in the first three variables, so $\deg(\tilde{F}') = \deg(\tilde{F})$. This shows that for almost all p the cone \mathcal{C}_p is of degree h . \square

The following result is dual to Lemma 7.2.11:

Lemma 7.2.12. *If \mathcal{C} is an algebraic complex of degree h , then almost all planar complex curves in planes π (whose duals are denoted by \mathcal{C}_π), are of class h .*

Example 7.2.4. A linear complex \mathcal{C} is defined by a linear equation in Plücker coordinates. It is therefore of degree one. The set $\mathcal{C}\gamma$ is a quadric, which is of degree two.

If \mathcal{C} is a regular linear complex, the complex cones are pencils, i.e., algebraic cones of degree one. If \mathcal{C} is a singular linear complex, still almost all complex cones are pencils. The same is true for the planar complex curves. \diamond

Line Complexes Defined by Algebraic Curves and Surfaces

We will show that the set of surface tangents of an algebraic surface $\Phi \subset \mathbb{CP}^3$ is an algebraic line complex, and so is the set of lines which meet an algebraic curve.

Proposition 7.2.13. *If $\Phi \subset \mathbb{CP}^3$ is an algebraic surface, then the set \mathcal{C} of its surface tangents is an algebraic line complex, the degree of \mathcal{C} equals the degree of Φ 's generic tangent cones and the class of Φ 's generic planar sections.*

Proof. We define a relation in $\mathbb{C}P^3 \times \mathbb{C}P^5$ by $p \sim q$ if and only if $p \in \Phi$, $q \in M_2^4$, and the line L whose Klein image is q is contained in Φ 's tangent plane at p . Obviously the projection onto the second factor maps ' \sim ' to $\mathcal{C}\gamma$.

If Φ is the zero set of a polynomial F , and $p = p\mathbb{C}$ is in Φ , then the tangent plane has the plane coordinates $\mathbb{C} \text{grad}_F(p)$. As all defining relations of ' \sim ' are polynomials, which are homogeneous in both groups of indeterminates separately (cf. Equ. (2.24)), ' \sim ' is an algebraic relation. Prop. 1.3.16 shows that $\mathcal{C}\gamma$ is a projective algebraic variety.

The statement about the dimension will not be proved here. The rest follows directly from Lemma 7.2.11 and Lemma 7.2.12. \square

Proposition 7.2.14. *If $c \in \mathbb{C}P^3$ is an algebraic curve, then the set \mathcal{C} of lines which meet c is an algebraic line complex, and $\deg(c) = \deg(\mathcal{C})$.*

This line complex is called the *secant complex* of c .

Proof. The proof is completely analogous to that of Prop. 7.2.13. We define $p \sim q$ if and only if $p \in c$, $q \in M_2^4$, and p is contained in the line whose Klein image is q . Equ. (2.23) shows that ' \sim ' is an algebraic relation and Prop. 1.3.16 implies that $\mathcal{C}\gamma$ is a projective algebraic variety. The dimension of $\mathcal{C}\gamma$ equals three, because \mathcal{C} contains bundles.

The statement about the degree of \mathcal{C} follows from Lemma 7.2.11, because a test plane incident with p intersects \mathcal{C}_p in as many generators as it has intersection points with c . \square

Example 7.2.5. As an example, we compute the complex of lines which meet the circle $c : x^2 + y^2 = 1, z = 0$ in Euclidean three-space E^3 . We embed E^3 into complex projective space $\mathbb{C}P^3$ via $x = x_1/x_0, y = x_2/x_0, z = x_3/x_0$.

Equ. (2.23) shows that a line $(\mathbf{c}, \bar{\mathbf{c}})$ which meets c in the point $(x_0, x_1, x_2, 0)\mathbb{C}$ must fulfill the relations $x_1 c_{23} + x_2 c_{31} = 0$ and $x_0 \bar{\mathbf{c}} = (x_1, x_2, 0) \times \mathbf{c}$.

Elimination of x_0, \dots, x_3 according to the procedure described in Ex. 1.3.14 gives the following results: (i) The Plücker relation and $c_{03}^2 - c_{23}^2 - c_{31}^2 = 0$ if $x_0 \neq 0$; (ii) The Plücker relation and $c_{03} = 0, c_{23}^2 + c_{31}^2 = 0$ for $x_0 = 0, x_1 \neq 0$, and (iii) nothing for $x_0 = x_1 = x_2 = 0, x_3 \neq 0$. Thus $\mathcal{C}\gamma$ has, besides the Plücker relation, the equation $c_{03}^2 - c_{23}^2 - c_{31}^2 = 0$. \diamond

Example 7.2.6. We use the result of Ex. 7.2.5 to compute the shortest distance of two circles c_1, c_2 in space. We do not claim that this is the most elegant way to do this — it should merely illustrate how to work with algebraic line complexes. If this shortest distance is assumed for points $p_1 \in c_1, p_2 \in c_2$, then the line $p_1 \vee p_2$ is orthogonal to both c_1 and c_2 . This means that it intersects the axes a_1, a_2 of c_1, c_2 .

We first derive the equation of the complex \mathcal{C} of lines which intersect the circle with radius r , center \mathbf{p} , and axis is parallel to the unit vector \mathbf{n} . If $\mathbf{o} = \mathbf{p}$, $\mathbf{n} = (0, 0, 1)$, and $r = 1$, we have the result of Ex. 7.2.5. If we change \mathbf{n} , \mathcal{C} has the equation $(\mathbf{c} \cdot \mathbf{n})^2 + (\bar{\mathbf{c}} \cdot \mathbf{n})^2 - \bar{\mathbf{c}}^2 = 0$. This is true because (i) this equation is obviously invariant with respect to the group of rotations, and (ii) it reduces to the

result of Ex. 7.2.5 if $\mathbf{n} = (0, 0, 1)$. Scaling by the factor r induces the transformation $(\mathbf{c}, \bar{\mathbf{c}}) \mapsto (\mathbf{c}, r\bar{\mathbf{c}})$ in line space. The translation $\mathbf{o} \mapsto \mathbf{p}$ induces the transformation $(\mathbf{c}, \bar{\mathbf{c}}) \mapsto (\mathbf{c}, \bar{\mathbf{c}} + \mathbf{p} \times \mathbf{c})$. Thus the general equation is

$$\mathcal{C} : r^2(\mathbf{c} \cdot \mathbf{n})^2 + ((\bar{\mathbf{c}} - \mathbf{p} \times \mathbf{c}) \cdot \mathbf{n})^2 - (\bar{\mathbf{c}} - \mathbf{p} \times \mathbf{c})^2 = 0. \quad (7.60)$$

We return to the original problem. We look for a line L with Plücker coordinates $(\mathbf{c}, \bar{\mathbf{c}})$ which intersects a_1, a_2, c_1 , and c_2 , and use a coordinate system such that c_1 is the circle of Ex. 7.2.5. We assume that c_2 is given by r, \mathbf{p} and \mathbf{n} , as indicated above. We get the equations $c_{12} = 0$ (L meets a_1) and $c_{03}^2 - c_{23}^2 - c_{31}^2 = 0$ (L meets c_1). Together with the Plücker relation, $\mathbf{c} \cdot (\mathbf{p} \times \mathbf{n}) + \bar{\mathbf{c}} \cdot \mathbf{n} = 0$ (L meets a_2) and (7.60) (L meets c_2) this is a system of five homogeneous equations for six homogeneous unknowns. The Plücker coordinates of $p_1 \vee p_2$ are among the solutions. \diamond

Example 7.2.7. If c is the cubic in $\mathbb{C}P^3$ whose affine part is parametrized by $c(u) = (1, u, u^2, u^3)$, then the set \mathcal{C} of lines which meet c is an algebraic line complex. We eliminate x_0, \dots, x_3 from the defining equations (cf. Ex. 1.3.14) and see that $\mathcal{C}\gamma$ is determined by the equations $c_{01}c_{23} + c_{02}c_{31} + c_{03}c_{12} = 0$ (the Plücker relation), and $-3c_{02}c_{31}c_{12} + c_{02}^2c_{23} - c_{01}c_{31}^2 + c_{12}^3 - 2c_{03}c_{12}^2 + c_{31}c_{02}c_{03} + c_{03}^2c_{12}$. Obviously \mathcal{C} is a cubic line complex, which confirms Prop. 7.2.14. \diamond

Algebraic Congruences

An algebraic congruence \mathcal{K} is a set of lines whose Klein image is a two-dimensional projective algebraic subvariety of the Plücker quadric.

There are several ways how algebraic congruences occur naturally. One is described by Prop. 7.2.16. Before, we state a lemma concerning polynomials with common zeros.

Lemma 7.2.15. *The set of equivalence classes of scalar multiples of nonzero complex homogeneous bivariate polynomials of degree k is identified with $\mathbb{C}P^k$ via*

$$F(u_0, u_1) = \sum_{i=0}^k a_i u_0^i u_1^{k-i} \rightarrow (a_0, \dots, a_k) \mathbb{C}.$$

Then the property of having a common zero (or having m common zeros, counting multiplicities) is an algebraic relation between points of $\mathbb{C}P^k$.

Proof. (Sketch) Denote this relation by ‘~’. The symbol $F\mathbb{C}$ means the set of scalar multiples of the polynomial F , i.e., a point in the projective space $\mathbb{C}P^k$. If $F\mathbb{C} \sim G\mathbb{C}$, then both F and G have a common linear factor $H(u_0, u_1) = b_0 u_0 + b_1 u_1$. The relation ‘~’, considered as a surface, is therefore parametrized by the mapping $\psi(H\mathbb{C}, G_1\mathbb{C}, G_2\mathbb{C}) = (HG_1\mathbb{C}, HG_2\mathbb{C})$, which takes a triple of projective points (representing polynomials) to a pair of projective points (representing polynomials).

It is now not difficult to show (using Prop. 1.3.16, analogous to the proof of Prop. 5.2.10) that ~ is an algebraic relation. \square

Proposition 7.2.16. *If c is an algebraic curve in $\mathbb{C}P^3$, then the set \mathcal{K} of c 's tangents plus the lines which meet c twice is an algebraic congruence.*

This congruence is called the *bisecant congruence* of c .

Proof. We confine ourselves to show that \mathcal{K} is algebraic. We use Plücker coordinates (l_{01}, \dots, l_{12}) for lines. In view of Lemma 5.2.1, it is sufficient to show that the set of lines in \mathcal{K} with $l_{01} \neq 0$ is an affine algebraic variety, then the same for $l_{02} \neq 0$, and so on.

We consider a line L with $l_{01} \neq 0$, and its points $s_0\mathbb{C}$ and $s_1\mathbb{C}$ with coordinates $s_0 = (0, l_{01}, l_{02}, l_{03})$ and $s_1 = (-l_{01}, 0, l_{12}, -l_{31})$. Because $l_{01} \neq 0$, these two points are well-defined and not equal.

We assume that c is the zero set of the homogeneous polynomials F_1, \dots, F_k . The line L intersects c twice or is a tangent, if the k homogeneous bivariate polynomials $\tilde{F}_i(u_0, u_1) = F_i(u_0 s_0 + u_1 s_1)$ ($i = 1, \dots, k$) in the indeterminates u_0, u_1 have two zeros $u_0 : u_1$ in common (which are possibly equal).

By Lemma 7.2.15, for each pair i, j , this is a polynomial condition on the coefficients of \tilde{F}_i, \tilde{F}_j , which in turn are polynomials in the Plücker coordinates l_{01}, \dots, l_{12} . Thus the set of lines in \mathcal{K} with $l_{01} \neq 0$ is an affine algebraic variety. The argument for $l_{02} \neq 0, \dots$ is completely analogous. \square

Example 7.2.8. We consider the bisecant congruence \mathcal{K} of the cubic c , whose affine part is parametrized by $c(u) = (u, u^2, u^3)$. The procedure described in the proof of Prop. 7.2.16 simplifies here because c is a rational curve, and we can parametrize the bisecant congruence by $(u, v) \mapsto c(u) \vee c(v)$. Implicitization gives the following equations which are a Gröbner basis of an ideal whose zero set is $\mathcal{K}\gamma$ (with respect to the ordering $c_{01} < \dots < c_{12}$): $(c_{12} + c_{03})c_{01} = c_{02}^2, c_{01}c_{23} = c_{12}^2, c_{31}c_{01} + c_{12}c_{02} = 0, c_{23}c_{02} + c_{12}c_{31} = 0, (c_{12} + c_{03})c_{12} + c_{02}c_{31} = 0, c_{23}(c_{03} + c_{12}) = c_{31}^2$. This Gröbner basis does not contain the Plücker relation explicitly (which would have been pure coincidence), but certainly implies it. \diamond

Bundle and Field Degree of an Algebraic Congruence

If \mathcal{K} is an algebraic congruence, the set $\mathcal{K}\gamma$ has a degree as an algebraic variety. From the viewpoint of line geometry, there are the following more natural definitions of ‘degree’:

We say that an algebraic congruence \mathcal{K} in complex projective three-space has the *bundle degree* k , if for almost all points $p \in \mathbb{C}P^3$ exactly k lines of \mathcal{K} are incident with p . The *field degree* of \mathcal{K} is k , if almost all planes in $\mathbb{C}P^3$ are incident with exactly k lines of \mathcal{K} . It can be shown that all algebraic congruences indeed have a finite bundle degree and a finite field degree which need not be equal.

Proposition 7.2.17. *Assume that \mathcal{C}, \mathcal{D} are algebraic line complexes of degree m and n , respectively. If $\mathcal{K} = \mathcal{C} \cap \mathcal{D}$, then \mathcal{K} is an algebraic congruence whose bundle and field degree equals the product mn .*

Proof. (Sketch) Without loss of generality $p = (1, 0, 0, 0)\mathbb{R}$. We consider the bundle with vertex p as a projective plane and use the plane $x_3 = 0$ for coordinatization (that is, the line $p \vee (0, x_1, x_2, x_3)\mathbb{R}$ has homogeneous coordinates $(x_0 : x_1 : x_3)$). Algebraic cones with vertex p have an equation of the form $F(x_1, x_2, x_3) = 0$, which means that they are algebraic curves in this projective plane. Almost all complex cones \mathcal{C}_p (\mathcal{D}_p) are of degree m (n , resp.) as planar algebraic curves. Th. 1.3.23 shows that these curves have mn points, i.e., lines incident with p , in common, if we count multiplicities.

The same argument for the planar complex curves shows the statement about the field degree. \square

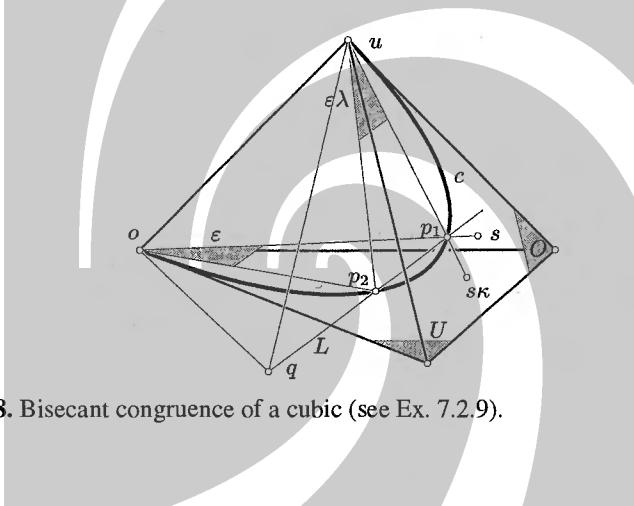


Fig. 7.28. Bisecant congruence of a cubic (see Ex. 7.2.9).

Example 7.2.9. We consider again the bisecant congruence \mathcal{K} of the twisted cubic, whose affine part is parametrized by $c(t) = c(t)\mathbb{R} = (1, t, t^2, t^3)\mathbb{R}$. Since the determinant $\det(c(t_1), c(t_2), c(t_3), c(t_4))$ equals the product of all differences $(t_i - t_j)$, four different points of c are projectively independent.

A plane ϵ which contains no tangent of c intersects c in three points, which are projectively independent. Thus ϵ contains three lines of \mathcal{K} . The bundle degree of \mathcal{K} therefore equals three.

On the other hand consider the projective mapping $\kappa : (0, t_1, t_2, t_3) \mapsto (t_1, t_2, t_3, 0)$ of the plane $O : x_0 = 0$ onto the plane $U : x_3 = 0$. Define a projective mapping λ from the bundle with vertex $o = (1 : 0 : 0 : 0)$ onto the bundle with vertex $u = (0 : 0 : 0 : 1)$ by $G \mapsto (G \cap O)\kappa \vee u$. It is elementary to verify that $p \in c$ if and only if $(o \vee p)\lambda = u \vee p$ (see Fig. 7.28).

A line L which meets two points p_1, p_2 of c (a bisecant) is thus contained in $\epsilon = o \vee p_1 \vee p_2$ and also in its λ -image $\epsilon\lambda = u \vee p_1 \vee p_2$, so $L = \epsilon \cap \epsilon\lambda$. This shows that ϵ (and therefore L) is uniquely determined by any point $q \in L$: We must have $q \vee o \in \epsilon$ and $q \vee u \in \epsilon\lambda$, so $\epsilon = (q \vee o) \vee (q \vee u)\lambda^{-1}$.

\mathcal{K} is algebraic, so for all points $p \in \mathbb{CP}^3$ there is at least one line $L \in \mathcal{K}$ incident with p . L may be a tangent, but if p is not contained in c 's tangent surface, L is a

bisecant and is unique. This shows that the bundle degree of \mathcal{K} equals one. As one is an odd number, all real points are incident with a real line of \mathcal{K} , even if this line is spanned by two conjugate complex points of c . \diamond

Remark 7.2.4. Ex. 7.2.9 and Prop. 7.2.17 show that not all algebraic congruences are the intersection of *two* algebraic complexes. Especially the bisecant congruence of a cubic is not. \diamond

Ruled Surfaces as Intersection of Algebraic Complexes

We have already studied algebraic ruled surfaces (cf. Sec. 5.2). If \mathcal{R} is such an algebraic ruled surface, then $\mathcal{R}\gamma$ is an algebraic curve in M_2^4 , defined by homogeneous polynomials, each of which defines an algebraic complex or vanish for all lines, such as the Plücker relation. An algebraic ruled surface is therefore the intersection of a finite number of algebraic complexes.

If an irreducible algebraic ruled surface \mathcal{R} is the intersection of three algebraic complexes which have degrees m_1, m_2, m_3 , then it is in general of degree $2m_1m_2m_3$.

Example 7.2.10. The surface of Ex. 5.1.4 and Ex. 5.2.5 is the intersection of three complexes. These are the quadratic complex \mathcal{C}_1 of lines which intersect the directrix a , the singular linear complex \mathcal{C}_2 of lines meeting the directrix b , which is the y -axis, and the singular linear complex \mathcal{C}_3 of lines parallel to the plane $y = 0$. This again shows the result that the surface is algebraic of degree $2 \cdot 2 \cdot 1 \cdot 1 = 4$. \diamond

7.2.3 Special Quadratic Complexes

A *quadratic* complex \mathcal{C} is an algebraic complex of degree two. By definition, its Klein image is defined by the Plücker relation and one further quadratic equation $F(\mathbf{l}, \bar{\mathbf{l}}) = 0$, which is not a power of the Plücker relation.

We consider a *pencil* of quadratic varieties in P^5 , whose equations have the form

$$F_{\lambda_0:\lambda_1} : \lambda_0 F(\mathbf{x}, \bar{\mathbf{x}}) + \lambda_1 \mathbf{x} \cdot \bar{\mathbf{x}} = 0, \quad (\lambda_0, \lambda_1) \neq (0, 0). \quad (7.61)$$

If $(\lambda_0 : \lambda_1) \neq (0 : 1)$, the quadratic polynomial $F_{\lambda_0:\lambda_1}$ together with the Plücker relation defines the same quadratic complex as F and the Plücker relation. This shows that any pencil $F_{\lambda_0:\lambda_1}$ of the form (7.61) corresponds to a quadratic line complex and vice versa. This correspondence is one-to-one only up to multiplication of F with a scalar factor.

The *classification* of quadratic complexes means enumeration of all possible different types of quadratic complexes, where the meaning of ‘different’ is ‘projectively nonequivalent’. A systematic treatment of the theory is beyond the scope of this book. In principle, we have to classify pencils of quadrics which contain the Klein quadric. A discussion of forty-nine cases can be found in [222]. We will restrict ourselves to illustrative and interesting examples instead, especially such ones important for applications of the theory.

The Tangent Complex of a Helical Motion

Remarkable examples of quadratic line complexes occur in the investigation of helical motions. We have already shown that the path normals of a uniform helical motion are the non-ideal lines of a linear line complex (see Th. 3.1.6). The set of path tangents is again a line complex, but this one is quadratic:

Proposition 7.2.18. *The path tangents of a uniform helical motion are the non-ideal lines of a quadratic line complex.*

Proof. We assume a Cartesian coordinate system such that the uniform helical motion is parametrized by (3.6). If $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$, Equ. (3.7) shows that its velocity vector $\mathbf{v}(\mathbf{x}) = (-y, x, p)$. The path tangent in \mathbf{x} therefore has Plücker coordinates

$$(l_{01} : \dots : l_{12}) = (\mathbf{v}, \mathbf{x} \times \mathbf{v}) = (-y, x, p, yp - xz, -yz - xp, x^2 + y^2), \quad (7.62)$$

which satisfy the relation

$$\mathcal{C} : p(l_{01}^2 + l_{02}^2) - l_{03}l_{12} = 0. \quad (7.63)$$

It is left to the reader to show the converse. Thus, the set of path tangents is the quadratic line complex defined by the Plücker relation and (7.63). \square

It is not difficult to see that the complex cones of the path tangent complex of a uniform helical motion are, in general, quadratic cones with circular sections in planes orthogonal to the axis of the helical motions. The generic planar complex curves are parabolae.

If we let $p \rightarrow 0$ in Prop. 7.2.18, the uniform helical motion becomes a uniform rotation, and Equ. (7.62) becomes

$$(l_{01} : \dots : l_{12}) = (-y, x, 0, -xz, -yz, x^2 + y^2).$$

This is a parametrization of the linear complex

$$\mathcal{C} : l_{03} = 0,$$

which consists of all lines orthogonal to the axis of the rotation. The third kind of one-parameter group of Euclidean motions, the uniform translation, has path tangents which are parallel to the translations. The set of path tangents therefore is a bundle of lines with vertex at infinity.

By Th. 3.4.2, the velocity vector field of a general one-parameter motion in Euclidean space is that of a uniform helical motion, a uniform rotation, or a uniform translation, or is zero. The set of path tangents is a quadratic complex, or a linear complex, or a bundle of lines, or is empty, respectively.

Remark 7.2.5. The path tangent complex of a helical motion occurs as a limit case of the following construction of quadratic complexes: Consider a Euclidean congruence transformation κ in E^3 . Then it is not difficult to show that the lines $p \vee p\kappa$

with $p \in E^3$ are contained in a quadratic complex. This complex plays a role in ‘two positions theory’ for spatial kinematics and in certain problems concerning the synthesis of mechanisms [18]. If $\kappa(t)$ is a one-parameter motion, the path tangents are the limits $\lim_{t \rightarrow t_0} p \vee p\kappa(t)$.

We also obtain a quadratic complex if κ is not a motion of Euclidean space, but more generally a projective automorphism of P^3 . Such complexes are called *collineation complexes*. \diamond

Remark 7.2.6. In Ex. 3.2.1 we have considered the set of path tangents of a uniform helical motion for the points of a plane orthogonal to the axis A of this motion. We have shown that this set is an elliptic linear congruence with rotational symmetry about A (see Fig. 3.7).

Therefore, the path tangent complex of a uniform helical motion is the union of a one-parameter family of linear line congruences, all of which exhibit rotational symmetry about A , and which are permuted by translations parallel to A . This may help to visualize the tangent complex of a helical motion. \diamond

The Tangent Complex of a Quadric

By Prop. 7.2.13, the set of tangents of an algebraic surface Φ in $\mathbb{C}P^3$ is an algebraic complex, and the degree of this complex equals the degree of Φ 's tangent cones. Therefore the set of surface tangents of a quadric Φ in complex projective three-space is a quadratic line complex.

There are several ways to compute its equation. One is to use a coordinate system such that Φ has a simple equation (e.g., an auto-polar simplex, found as described in Ex. 1.1.22). An alternative interpretation of the same procedure is to apply a projective automorphism κ such that $\Phi\kappa$ has a simple equation, e.g., $\Phi\kappa$ is a sphere. Then we compute the tangent complex $\tilde{\mathcal{C}}$ of $\Phi\kappa$ (see below), and use Equ. (2.20) to compute the κ -preimage \mathcal{C} of $\tilde{\mathcal{C}}$.

Another way is the following: We use homogeneous coordinates and assume that Φ has the equation (1.33), i.e., $x^T \cdot C \cdot x = 0$. The line $L = p\mathbb{R} \vee q\mathbb{R}$ is parametrized by $x\mathbb{R} \in L \iff x = \lambda p + \mu q$. To compute $L \cap \Phi$ we have to solve

$$0 = (\lambda p + \mu q)^T \cdot C \cdot (\lambda p + \mu q) = \lambda^2(p^T \cdot C \cdot p) + 2\lambda\mu(p^T \cdot C \cdot q) + \mu^2(q^T \cdot C \cdot q).$$

The line L is tangent to Φ if this quadratic equation has a double solution, which happens if and only if

$$(p^T \cdot C \cdot q)^2 - (p^T \cdot C \cdot p)(q^T \cdot C \cdot q) = 0. \quad (7.64)$$

According to Equ. (2.4), a line with Plücker coordinates $(l_{01} : \dots : l_{12})$ contains the points $s_0 = (0 : l_{01} : l_{02} : l_{03})$ and $s_1 = (-l_{01} : 0 : l_{12} : -l_{31})$. These are well-defined and not equal if $l_{01} \neq 0$. Substituting $p = s_0$ and $q = s_1$ in (7.64) results in a quadratic equation for the Plücker coordinates (l_{01}, \dots, l_{12}) , which characterizes the lines of the tangent complex among the lines with $l_{01} \neq 0$. Because we already know that there is one quadratic equation which defines the tangent complex, this equation actually characterizes the lines tangent to Φ .

An interesting special case is the tangent complex \mathcal{C} of a sphere Φ . For simplicity we consider only the unit sphere. A real line is tangent to Φ if and only if its distance to \mathbf{o} equals one.

It follows from the discussion at p. 155 that in *normalized* Plücker coordinates, the equation of \mathcal{C} is

$$\mathcal{C} : \|\mathbf{c}\|^2 = 1, \|\bar{\mathbf{c}}\| = 1.$$

This equation is easily made homogeneous and is then valid for arbitrary Plücker coordinates:

$$\mathcal{C} : \|\mathbf{c}\|^2 - \|\bar{\mathbf{c}}\|^2 = 0.$$

Geodesics on Cones and the Tangent Complex of a Sphere

The complex curves of a sphere's tangent complex have a remarkable property:

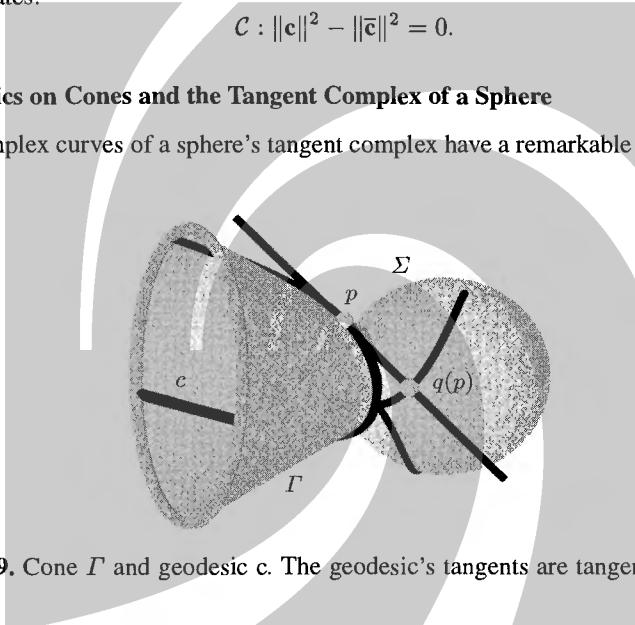


Fig. 7.29. Cone Γ and geodesic c . The geodesic's tangents are tangent to a sphere Σ .

Proposition 7.2.19. *If c is a geodesic (but not a ruling) on a cone with vertex \mathbf{o} , then c 's tangents are also tangent to a sphere Σ centered in \mathbf{o} . Conversely, if a curve c has this property, then c is a geodesic on the cone Γ with vertex \mathbf{o} and directrix c (see Fig. 7.29).*

Proof. We first show the second part. Denote the tangent complex of Σ by \mathcal{C} and c 's tangent developable by \mathcal{R} . Then $\mathcal{R} \subset \mathcal{C}$ and c 's osculating plane $\varepsilon(p)$ in the point $p \in c$ is \mathcal{R} 's tangent plane. This shows that $\varepsilon(p)$ is tangent to Σ in a point $q(p)$ (cf. Prop. 7.2.5).

In order to show that c is a geodesic curve on Γ , we have to show that $\varepsilon(p)$ is orthogonal to Γ 's tangent plane $\tau(p)$ in p . This is obvious, since $\tau(p)$ contains the line $\mathbf{o} \vee q(p)$, which is orthogonal to $\varepsilon(p)$.

The first part is shown as follows: Both geodesics and complex curves constrained to a given surface are uniquely determined by an initial point p plus the

tangent T there (this follows from an appropriate theorem on existence and uniqueness of the solution of a certain differential equation and will not be shown here). As any complex curve is a geodesic, this shows that also any geodesic is a complex curve for the tangent complex \mathcal{C} of the sphere Σ which is tangent to T . \square

The following result is a discrete version of Prop. 7.2.19:

Proposition 7.2.20. *Consider a cone surface Γ with vertex \mathbf{o} and a light ray, which is repeatedly reflected in Γ . Then all legs of the resulting reflection polygon P are tangent to a sphere centered in \mathbf{o} . Moreover, the reflection polygon is a geodesic on the pyramid surface Ψ with vertex \mathbf{o} and directrix P .*

Proof. Since all tangent planes of Γ contain \mathbf{o} , the law of reflection implies that the distance of a light ray to \mathbf{o} is the same before and after reflection. This shows the first statement. The second follows from the fact that developing Ψ into a plane transforms P into a straight line, which is also an immediate consequence of the law of reflection (see Fig. 7.30). \square

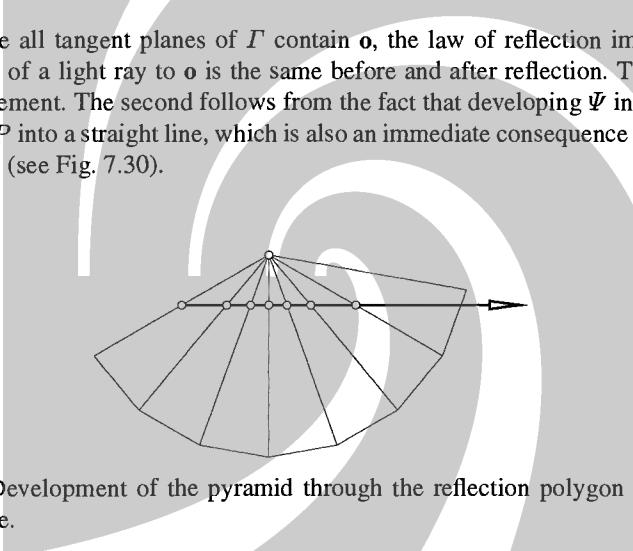


Fig. 7.30. Development of the pyramid through the reflection polygon of a right circular cone.

Remark 7.2.7. Prop. 7.2.19 is a limit case of Prop. 7.2.20 in the following sense. If the angle enclosed by the incoming light ray and the cone surface Γ tends to zero, the reflection polygon converges to a smooth curve c , which is Γ 's geodesic. The legs of the polygon converge to c 's tangents.

This is true not only for cones and is an immediate consequence of the fact that incoming light ray, reflected light ray, and surface normal are co-planar — in the limit this means that c 's osculating plane contains the surface normal. Another argument is that geodesics are locally shortest curves, and light takes the shortest possible path between points. A sequence of reflection polygons which converges to a geodesic is shown by Fig. C.13. \diamond

Example 7.2.11. Consider a right circular cone Γ with vertex \mathbf{o} , and a polygon P obtained by repeated reflection in Γ . It is an easy exercise to see that the pyramid whose vertex is \mathbf{o} and whose directrix is P is regular in the sense that the angle enclosed by adjacent faces is constant (see Fig. 7.30).

This, together with the fact that it is a geodesic curve on the pyramid, considerably simplifies the computation of such a polygon. As a limit case we obtain a geodesic curve on Γ . \diamond

The Complex of Constant Slope

In Sec. 6.3 we have encountered the set \mathcal{C}' of lines which enclose the constant angle γ with some reference plane. If we choose a Cartesian coordinate system such that the reference plane is $z = 0$, a proper line with Plücker coordinates $(l_{01} : \dots : l_{12})$ is in \mathcal{C}' if and only if

$$l_{03}^2 - \tan^2 \gamma(l_{01}^2 + l_{02}^2) = 0. \quad (7.65)$$

We see that \mathcal{C}' is the set of proper lines of the quadratic line complex \mathcal{C} defined by (7.65).

The complex cones with proper vertex are cones of revolution and translates of each other. The planar complex curves are degenerate: depending on the inclination of the plane ε , the set \mathcal{C}_ε consists of two pencils of parallel lines, one pencil of parallel lines, or is empty. \mathcal{C} consists of all lines which meet a certain conic of the ideal plane (cf. Remark 6.3.5), and so \mathcal{C} is projectively equivalent to the complex of Ex. 7.2.5.

The complex curves, i.e., the curves of constant slope, have been studied in detail in Sec. 6.3.3. Their properties are similar to those described by Prop. 7.2.19: A curve of constant slope is a geodesic on a cylinder whose rulings are orthogonal to the reference plane. The reflection polygon defined by such a cylinder consists of line segments of constant slope.

Cyclic Complexes

Reflection in cones and cylinders is tied to special quadratic complexes in a rather elementary way (cf. Prop. 7.2.20, Remark 7.2.7). W. Wunderlich [215] showed analogous properties of surfaces of revolution and helical surfaces. These are a little less obvious and shall be briefly discussed here.

Assume a uniform helical motion, which in a suitable Cartesian coordinate system is parametrized by Eqn. (3.6). The velocity vector $\mathbf{v}(\mathbf{p})$ of a point $\mathbf{p} = (x, y, z)$ has coordinates $(-y, x, p)$. We want to attach a cone of revolution $\Gamma(\mathbf{p})$ to all points of Euclidean space, such that the union of their generators is a line complex. This is described by the following

Theorem 7.2.21. (W. Wunderlich) *Assume a uniform helical motion parametrized by (3.6). For all points $\mathbf{p} = (x, y, z)$ of Euclidean space we consider the cone of revolution $\Gamma(\mathbf{p})$ whose vertex is \mathbf{p} , whose axis is \mathbf{p} 's path tangent, and whose aperture angle $2\gamma(\mathbf{p})$ satisfies*

$$\sqrt{x^2 + y^2 + p^2} \cos \gamma = c \neq 0. \quad (7.66)$$

The union of rulings of the cones $\Gamma(\mathbf{p})$ consists of the proper lines of the quadratic complex defined by

$$\mathcal{C} : c^2(l_{01}^2 + l_{02}^2 + l_{03}^2) = (pl_{03} + l_{12})^2, \quad (7.67)$$

This quadratic complex is called *cyclic complex*.

Proof. Consider a ruling L of $\Gamma(\mathbf{p})$, which has the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$. Clearly, $\langle \mathbf{l}, \mathbf{v}(\mathbf{p}) \rangle = \gamma$, so we have $\mathbf{l} \cdot \mathbf{v} = \|\mathbf{l}\| \|\mathbf{v}\| \cos \gamma$. This equation expands to

$$(-l_{01}y + l_{02}x + l_{03}p)^2 = (l_{01}^2 + l_{02}^2 + l_{03}^2)(x^2 + y^2 + p^2) \cos^2 \gamma. \quad (7.68)$$

From $\mathbf{p} \in L$ we get $-l_{01}y + l_{02}x = l_{12}$, and the result follows. \square

If $c = 0$ in Equ. (7.67), the angle γ of Th. 7.2.21 equals $\gamma = \pi/2$ and the cones become pencils of path normals. Their union is the linear complex

$$\mathcal{C}_0 : pl_{03} + l_{12} = 0. \quad (7.69)$$

Remark 7.2.8. Note that cyclic complexes already appeared in Sec. 4.1, where we fitted linear complexes to lines, and defined the *moment* of a line with respect to a linear complex \mathcal{C}_0 . If we compare Equ. (7.66) with Equ. (4.1), we see that the set of lines L which satisfy $m(L, \mathcal{C}_0) = \text{const.}$ is a cyclic complex. \diamond

It is easy to determine the complex cones and planar complex curves of a cyclic complex:

Proposition 7.2.22. *The complex cones of a cyclic complex (defined over the field of real numbers) are cones of revolution (possibly degenerate or void) and the planar complex curves are circles.*

Proof. The first part follows immediately from Th. 7.2.21. Whether or not there is a non-void complex cone depends on whether

$$\cos \gamma = \frac{c}{\sqrt{x^2 + y^2 + p^2}}.$$

has a solution γ or not. If \mathbf{p} is contained in the cylinder

$$\Delta : x^2 + y^2 = c^2 - p^2, \quad (7.70)$$

the cone \mathcal{C}_p degenerates to a path tangent. Of course, Δ is the empty set if $c^2 - p^2 < 0$. If Δ exists, points inside Δ have no complex cone, because then $x^2 + y^2 + p^2 < c^2$, which would imply $\cos \gamma > 1$.

In order to determine the planar complex curves, we consider a plane α which is not parallel to the z -axis. We are going to show that the lines in α have the same distance to some point of α .

If the lines L, L' are contained in the plane α and not parallel, then both are contained in the complex cone \mathcal{C}_p of their intersection point $p = L \cap L'$. As \mathcal{C}_p 's axis is p 's path tangent, one of the bisectors of L, L' within α , call it K , is contained in p 's path normal plane. Thus $K \in \mathcal{C}_0$ and $\alpha\nu \in K$, where ν is the null polarity associated with \mathcal{C}_0 . This shows that the distances $d(\alpha\nu, L)$ and $d(\alpha\nu, L')$ are equal, and that \mathcal{C}_α is a circle with center $\alpha\nu$. \square

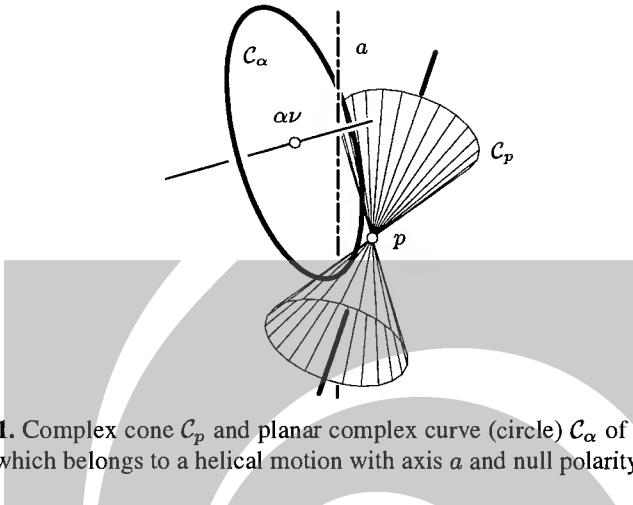


Fig. 7.31. Complex cone C_p and planar complex curve (circle) C_α of a cyclic complex \mathcal{C} , which belongs to a helical motion with axis a and null polarity ν .

Remark 7.2.9. If \mathcal{C} is a complex quadratic line complex, the set of points p of P^3 such that C_p is no quadratic cone is called the *singularity surface* of \mathcal{C} .

If \mathcal{C} has a real equation, we also consider the set of real lines in \mathcal{C} , and complex cones C_p of real lines in \mathcal{C} incident with a real point p of the singularity surface. As \mathcal{C} is quadratic, C_p may consist of a bundle, of two pencils, of one pencil, or of one line (the intersection of two conjugate complex pencils).

The singularity surface of a quadratic complex is in general an algebraic surface of order four and class four, known as the *Kummer surface*. If \mathcal{C} is a cyclic complex, the singularity surface is given by the surface Δ whose equation is (7.70) plus the ideal plane. If $c^2 > p^2$, (7.70) defines a right circular cylinder with real points, if $c^2 < p^2$, a right circular cylinder without real points, and if $c^2 = p^2$ a pair of conjugate complex planes whose intersection is the z -axis.

It can also be shown that both the complex cones and the complex circles of a cyclic complex \mathcal{C} touch Δ in two points. \diamond

Reflection in Rotational and Helical Surfaces

W. Wunderlich [215] has shown the following connection between cyclic complexes and geometrical optics.

Theorem 7.2.23. Consider a surface Φ generated by a uniform rotational or helical motion $\alpha(t)$, and the reflection polygon P generated by repeated reflection of an incoming light ray L in Φ .

If L is orthogonal to the path curve of a point under $\alpha(t)$, then the edges of P are contained in the path normal complex defined by $\alpha(t)$. If not, all edges are contained in a cyclic complex.

Proof. There is a unique cyclic or linear complex \mathcal{C} which contains L and is attached to the uniform motion $\alpha(t)$ as described by Th. 7.2.21.

Assume that L intersects Φ in the point p and is reflected into L' . Then Φ 's surface normal N in p is a bisector of L, L' . The path normal plane of p (which contains N) is a symmetry plane of \mathcal{C}_p , which shows that $L' \in \mathcal{C}_p$ and thus $L' \in \mathcal{C}$. Iteration of this argument shows that all edges of the reflection polygon are contained in \mathcal{C} . \square

As an example, Figure C.13 shows a reflection polygon for a surface of revolution.

Theorem 7.2.24. *If Φ is a surface generated by a uniform helical or rotational motion $\alpha(t)$, then a geodesic on Φ is a complex curve of a cyclic or linear complex \mathcal{C} attached to $\alpha(t)$ according to Th. 7.2.21.*

Conversely, a complex curve c of \mathcal{C} is a geodesic of the surface generated by c under the action of $\alpha(t)$.

Proof. We use Remark 7.2.7 to convert Th. 7.2.23, which is a statement on reflection polygons, into a statement about geodesics. The details are left to the reader. \square

The explicit formula of Th. 7.2.21 shows the following result of R. Sauer, which generalizes an old result of A. Clairaut, and transforms the second order differential equation of geodesics into a family of first order differential equations:

Theorem 7.2.25. *(of A. Clairaut and R. Sauer) Assume that Φ is a surface generated by a uniform helical or rotational motion of parameter p , and $c(t) = (c_1(t), c_2(t), c_3(t))$ is a geodesic on Φ . There is the following relation between the radius $r = \sqrt{c_1^2 + c_2^2}$ and the angle α enclosed by $\dot{c}(t)$ and the path tangent of $c(t)$:*

$$\sqrt{r^2 + p^2} \cos \gamma = c = \text{const.} \quad (7.71)$$

The case $p = 0$ is the classical Clairaut theorem.

The geodesics described by Cor. 7.2.6 are special cases of Th. 7.2.24 — the corresponding complex is the linear path normal complex ($c = 0$ in Equ. (7.66) and Equ. (7.71)). If the uniform motion which generates Φ has zero parameter, then Φ is a surface of revolution. The geodesics which belong to the case $c = 0$ then are the meridian curves of Φ .

Example 7.2.12. A simple example of Th. 7.2.25 is the planar complex circle of a cyclic complex: It is a geodesic of the helical pipe surface which it generates. A more complicated example is the helical surface

$$\mathbf{x}(u, v) = \begin{bmatrix} a \cos(u+v) + b \cos(u-v) \\ a \sin(u+v) - b \sin(u-v) \\ q_2 u + p v + q_2 \sin(2u) \end{bmatrix}, \quad \text{with}$$

$$q_1 = \frac{p^2(a^2 + b^2) - (a^2 - b^2)^2}{2p(a^2 - b^2)}, \quad q_2 = \frac{abp}{2(b^2 - a^2)}.$$

The value p is the pitch of the helical surface. The lines $v = \text{const}$ are geodesics, and the top view of such geodesics are ellipses with principal axes $a + b$ and $a - b$. Fig. C.16 shows the surface for $a = 2, b = 1, p = 1$. \diamond

Geometrical Fiber Optics

Repeated reflection in a tube-like surface, such as a surface of revolution, occurs, at least approximately, in optical fibers capable of more than one mode. We will not attempt to discuss the transmission of light through optical fibers here, nor will we make any claim as to the range of validity of the approximation by reflection of an ideal geometrical ray in the boundary of the fiber. This approximation must break down for small diameters.

Let us discuss a simple model, namely a surface of revolution. First consider a conical surface. Fig. 7.30 shows clearly that a light ray which does not aim directly at the vertex will inevitably eventually be reflected backwards, regardless of the dimensions of the cone.

Now consider a surface of revolution, with axis A , and bounded by circles c_0, c_1 of radius r_0, r_1 , respectively, with $r_1 < r_0$. We assume that between c_0 and c_1 the radius of the surface is monotonically decreasing. Suppose that an incoming light ray L_0 is reflected in a point $\mathbf{p}_0 \in c_0$ and encloses the angle γ_0 with the circle c_0 .

By Th. 7.2.23, there is a cyclic complex \mathcal{C} with axis A , which contains L_0 and all edges of the reflection polygon generated by L_0 . If \mathbf{p} is contained in an edge L of this polygon, let $r = \overline{\mathbf{p}A}$ and γ equal the angle enclosed by L and the circle whose axis is A and which contains \mathbf{p} . Th. 7.2.21 shows that $r \cos \gamma$ is a constant.

Suppose the reflection polygon leaves the surface at the narrower end, in a point \mathbf{p}_e in the interior or possibly in c_1 , and further assume that r and γ have the values r_e, γ_e there. Obviously $r_e \leq r_1$. Then

$$\cos \gamma_e = \frac{r_0}{r_e} \cos \gamma_0 \geq \frac{r_0}{r_1} \cos \gamma_0$$

This is a contradiction if $r_0 \cos \gamma_0 > r_1$, so the reflection polygon does *not* leave the surface at the narrower end if

$$\cos \gamma_0 > \frac{r_1}{r_0} \quad (7.72)$$

Even if $\cos \gamma_0 \leq r_1/r_0$ the ray need not come through, which can be seen in the case of a light ray which meets the axis (then $\cos \gamma = 0$). The case of the cone suggests that nevertheless the likelihood that a ray does come through is related to the ratio r_1/r_0 .

As a limit case we get results on the behaviour of geodesics of surfaces of revolution with monotonically decreasing radius. It is remarkable that the knowledge of the angle γ_0 at c_0 already determines the radius r_r of a circle c_r , which the geodesic will touch and then return: Obviously $\gamma_r = 0$, which implies

$$r_r = r_0 \cos \gamma_0.$$

The specific shape of the surface between c_0 and c_r has no influence on the radius r_r . An example of a returning reflection polygon can be seen in Fig. C.14.

8. Linear Line Mappings — Computational Kinematics

In Sec. 2.1 we have identified the lines of projective three-space P^3 with the points of the Klein quadric M_2^4 , which is contained in five-dimensional projective space P^5 . This identification has been performed by the introduction of Plücker coordinates, and makes it possible to apply concepts of projective geometry to lines. An example of this is the discussion of line complexes, linear manifolds of line complexes, and related sets of lines in Chap. 3.

This chapter is devoted to the study of mappings which are *linear* in Plücker coordinates, but are singular in the sense that some nonzero Plücker coordinates are mapped to zero. We know that all these mappings are generalizations of the familiar concept of *central projection*. The ordinary central projection, which produces a two-dimensional image of three-dimensional objects meets the old desire to make planar pictures of spatial objects, such that linear arrangements of points are preserved. The same is true for linear mappings of higher dimensions: What makes them useful is that the image is contained in a projective space of lower dimension than the original.

It turns out that linear mappings of this kind are not only tools for visualization, but some of them have interesting links to planar and spherical motions, to the quaternion unit sphere, to rational curves in quadratic surfaces, and to problems of surveying.

8.1 Linear Line Mappings and Visualization of the Klein Model

Recall the properties of singular projective mappings from Sec. 1.1.2. Following H. Brauner [20], we define

Definition. A mapping μ from the set \mathcal{L} of lines into a projective space P^d is called *linear* if it is the composition of the Klein mapping γ and a linear mapping λ from P^5 onto P^d .

An immediate consequence of this definition is that in Plücker coordinates the mapping μ is described by

$$(\mathbb{L}\mathbb{R})\mu = (A \cdot L)\mathbb{R}, \quad A \in \mathbb{R}^{(d+1) \times 6}. \quad (8.1)$$

We will always assume that $d \leq 4$ and that the rank of the matrix A actually equals its number of rows (so that the image of μ is not contained in a proper subspace of

P^d). Then the linear mapping λ has a kernel subspace Z of projective dimension $4-d$. Assume we have factored λ as the product of a central projection and a projective isomorphism. All possible linear mappings which share the same projection center Z are indistinguishable from the viewpoint of projective geometry.

Example 8.1.1. In Sec. 4.3 we have used a mapping σ of lines to points of \mathbb{R}^4 . If the line L is spanned by the points $(x_1, x_2, 0)$ and $(x_3, x_4, 1)$, σ shall map it to the point $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. As the Plücker coordinates of such a line L are given by $L\gamma = (x_3 - x_1, x_4 - x_2, 1, x_2, -x_1, x_1 x_4 - x_2 x_3)\mathbb{R}$, the linear mapping $\mu : (l_{01}, \dots, l_{12})\mathbb{R} \mapsto (l_{03}, -l_{31}, l_{23}, l_{01}-l_{31}, l_{02}+l_{23})\mathbb{R}$ obviously has the property that $L\gamma\mu = L\sigma$, after embedding \mathbb{R}^4 into projective space P^4 .

According to Lemma 1.1.19, the mapping σ is the composition of a central projection onto a four-dimensional subspace of P^5 and a projective isomorphism. The central projection part of σ has the center $Z = (0, 0, 0, 0, 0, 1)\mathbb{R}$. In this special case the center itself is the Klein image of a line, namely the ideal line of the xy -plane. Thus σ is essentially a *stereographic projection*. \diamond

We are now going to describe interesting linear line mappings into projective spaces of dimension two and three. In every case we will show a definition of the mapping which employs only projective three-space and does not use the Klein quadric.

8.1.1 Linear Line Mappings into P^2

Linear line mappings whose image is a projective plane have been studied in detail. For a classification and further references to the literature the reader is referred to H. Brauner [20].

We assume that the linear line mapping in question has the form $\gamma\lambda$, where γ is the Klein mapping. $\lambda : P^5 \rightarrow P^2$ is a linear mapping such that $\lambda = \pi\alpha$, with a central projection π and a projective isomorphism α . Then the center Z of π must be two-dimensional. Depending on the intersection of Z with the Klein quadric, there are different projective types of linear line mappings. We first show two basic examples.

Example 8.1.2. If Z' is a point and U' is a plane in projective three-space such that $Z' \notin U'$, the *projection* μ of lines with center Z' and image plane U' is defined by

$$L \mapsto L\mu = (Z' \vee L) \cap U', \text{ for } Z' \notin L. \quad (8.2)$$

If $Z' \in L$, we leave $L\mu$ undefined. This is a mapping of lines to lines. The set of lines in U' is a projective plane, so μ is indeed a mapping from \mathcal{L} into some projective plane P^2 . Yet another ‘projection’ $\bar{\mu}$ of lines is the mapping

$$L \mapsto L\bar{\mu} = Z' \vee L, \text{ for } Z' \notin L. \quad (8.3)$$

Here lines are mapped to planes, but the bundle of planes incident with Z' is a projective plane \bar{P}^2 itself (see p. 5). A projective isomorphism of P^2 and \bar{P}^2 is

provided by the mapping π_3 of Lemma 6.2.10, which maps a plane of the bundle \bar{P}^2 to its intersection with U' .

To show that μ or $\bar{\mu}$ are linear, we introduce a projective coordinate system such that $Z' = (1, \mathbf{o})\mathbb{R}$. A plane incident with Z' has plane coordinates $\mathbb{R}(0, \mathbf{u})$, and the vector \mathbf{u} serves as homogeneous coordinate vector of a point of \bar{P}^2 (i.e., of a plane). If L has Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$, (2.16) shows that $Z' \vee L = \mathbb{R}(0, -\bar{\mathbf{l}})$. Thus $\bar{\mu}$ is a linear line mapping, and so is μ .

The exceptional set of both μ and $\bar{\mu}$ is the set of lines incident with Z' . The Klein image of this bundle is a plane contained in the Klein quadric. This plane is the center of the linear mappings which correspond to μ and $\bar{\mu}$. \diamond

Example 8.1.3. Dual to the mapping $\bar{\mu}$ of Ex. 8.1.2 is the mapping

$$L \mapsto L\bar{\mu}^* = (L \cap U') \text{ for } L \not\subset U',$$

The mapping is left undefined for the lines of the plane U' . The Klein image Z of this exceptional set of lines is a plane contained in the Klein quadric and equals the kernel of $\bar{\mu}^*$. \diamond

The Right and Left Image of a Line

Let us proceed with a substantial example. We consider Euclidean space E^3 and its projective extension P^3 . It will be convenient to use a Cartesian coordinate system for points, together with the corresponding Plücker coordinates for lines. The mapping μ^+ from the space of lines to the bundle P^2 of lines incident with the origin is defined by

$$L\mu^+ = O \vee P, \quad \text{with} \quad P = \mathbf{l}^+ = \mathbf{l} + \bar{\mathbf{l}}, \quad (8.4)$$

if L has the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$. Obviously μ^+ is a linear line mapping, because the vector \mathbf{l}^+ serves as homogeneous coordinate vector of the line $L\mu^+$. In order to better understand this mapping, we investigate its kernel and its fibers.

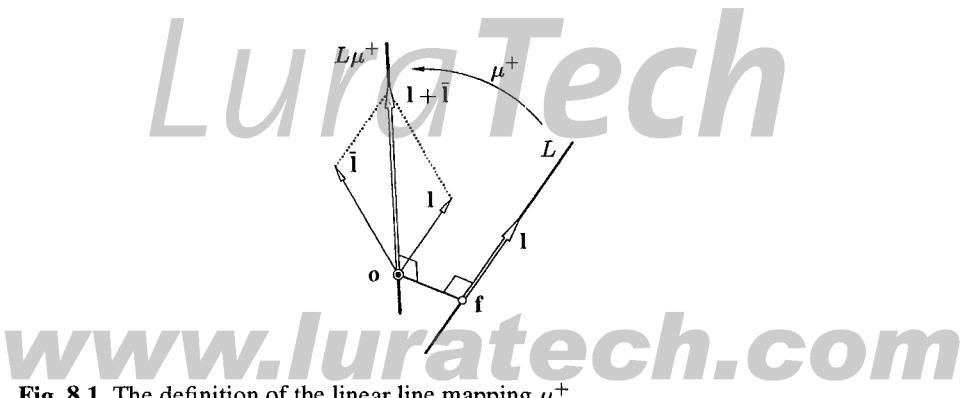


Fig. 8.1. The definition of the linear line mapping μ^+ .

Lemma 8.1.1. *The mapping μ^+ is defined for all lines of P^3 . The extension of μ^+ to the lines of complex three-space $\mathbb{C}P^3$ is undefined for the lines of a regulus \mathcal{R}^+ , which is contained in the quadric*

$$\Lambda : \mathbf{x} \cdot \mathbf{x} = -x_0^2. \quad (8.5)$$

Proof. The symbol $(0, \mathbf{l}^+)\mathbb{R}$ of (8.4) designates no point if \mathbf{l}^+ is zero, i.e., if

$$\mathbf{l} + \bar{\mathbf{l}} = \mathbf{o}, \text{ or } l_{01} + l_{23} = l_{02} + l_{31} = l_{03} + l_{12} = 0. \quad (8.6)$$

Each of these three linearly independent equations describes, together with the Plücker relation $\mathbf{l} \cdot \bar{\mathbf{l}} = 0$, a linear complex. The intersection of these complexes is empty, because $\mathbf{l} = -\bar{\mathbf{l}}$ implies that $\mathbf{l} \cdot \bar{\mathbf{l}} = -\|\mathbf{l}\|^2$, which is nonzero unless $\mathbf{l} = -\bar{\mathbf{l}} = \mathbf{o}$.

The complex version of μ^+ is described by the same equation (8.4) as the real version, but with \mathbb{C} instead of \mathbb{R} . This time the solution of Equ. (8.6) is not void, but equals a regulus \mathcal{R}^+ . By Equ. (2.23), a point $(x_0, \mathbf{x})\mathbb{C}$ incident with a line $(\mathbf{l}, \bar{\mathbf{l}})\mathbb{C}$ of \mathcal{R}^+ must fulfill the relations $\mathbf{x} \cdot \mathbf{l} = 0$ and $x_0\mathbf{l} + \mathbf{x} \times \mathbf{l} = \mathbf{o}$. This implies that $\mathbf{o} = x_0(\mathbf{x} \times \mathbf{l}) + \mathbf{x} \times (\mathbf{x} \times \mathbf{l}) = -x_0^2\mathbf{l} + (\mathbf{x} \cdot \mathbf{l})\mathbf{x} - (\mathbf{x} \cdot \mathbf{x})\mathbf{l} = -(x_0^2 + \mathbf{x} \cdot \mathbf{x})\mathbf{l}$, which shows that \mathcal{R}^+ is contained in the quadric (8.5). \square

The quadric Λ of (8.5) has a real polarity, but no real points. It has an interpretation as a sphere whose radius is a purely imaginary number. Let us now look at the *fibers* of μ^+ (i.e., sets of the form $(\mu^+)^{-1}(K)$, where K is a line incident with the origin). First we see that the line $L\mu^+$ of (8.4) has the Plücker coordinates $(\mathbf{l} + \bar{\mathbf{l}}, \mathbf{o})$, which shows that

$$L\mu^+ \mu^+ = L\mu^+. \quad (8.7)$$

Lemma 8.1.2. *For all lines G of the bundle incident with the origin, the fiber $(\mu^+)^{-1}(G)$ is an elliptic linear congruence which contains G . The focal lines of the fiber are contained in \mathcal{R}^+ 's complementary regulus, \mathcal{R}^- .*

Proof. Suppose that $\mu^+ = \gamma\lambda$ where λ is a linear mapping with center Z . If L is an arbitrary line, then precisely the lines K with $K\gamma \in L\lambda \vee Z$ have the property that $L\mu^+ = K\mu^+ = G$. Thus a fiber \mathcal{F}^+ is the intersection of the Klein quadric with the three-dimensional subspace $F^3 = L\lambda \vee Z$, i.e., a linear congruence. Equ. (8.7) shows that $G \in \mathcal{F}^+$.

If a line of P^5 is contained in $\mathcal{F}^+\gamma$, then this line must intersect Z , as Z is a plane contained in the three-space F^3 . By Lemma 8.1.1, such a line cannot be contained in the Klein quadric. Thus \mathcal{F}^+ contains no pencils, and is elliptic.

If we embed both P^3 and P^5 into $\mathbb{C}P^3$ and $\mathbb{C}P^5$, a fiber \mathcal{F}^+ of a real point becomes a hyperbolic linear congruence, and its focal lines A, B are conjugate complex (cf. Th. 3.2.7). By Lemma 3.2.1 and Th. 3.2.4, $\{A\gamma, B\gamma\} = M_2^4 \cap F^3\mu_2^4$, where μ_2^4 is the polarity of the Klein quadric. As F^3 contains the center Z , the line $F^3\mu_2^4$ is contained in the plane $Z\mu_2^4$.

By Th. 3.3.3 and Lemma 3.2.1, the regulus \mathcal{R}^- defined by $\mathcal{R}^-\gamma = Z\mu_2^4 \cap M_2^4$ is complementary to the regulus \mathcal{R}^+ , and is therefore likewise contained in the quadric Λ . This completes the proof. For an alternative proof, see Remark 8.1.5. \square

We compute an equation of the fiber $(\mu^+)^{-1}(G)$: If $L\mu^+ = G$ with $G = (1, \mathbf{o})\mathbb{R} \vee (0, \mathbf{g})\mathbb{R}$, then the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ of L must be such that $\mathbf{l} + \bar{\mathbf{l}}$ is a multiple of \mathbf{g} , i.e., $\mathbf{g} \times (\mathbf{l} + \bar{\mathbf{l}}) = \mathbf{o}$. This is a set of three linear equations, two of which are linearly independent. If for instance $g_2 \neq 0$, then the two equations

$$g_2(l_{01} + l_{23}) - g_1(l_{02} + l_{31}) = 0, \quad g_3(l_{02} + l_{31}) - g_2(l_{03} + l_{12}) = 0, \quad (8.8)$$

describe the fiber $(\mu^+)^{-1}(G)$.

The existence of reguli \mathcal{R}^+ and \mathcal{R}^- suggests to define a second linear line mapping where \mathcal{R}^+ and \mathcal{R}^- change their roles. It is easily verified that this is accomplished by defining

$$L\mu^- = O \vee P, \quad \text{with} \quad P = \mathbf{l}^- = \mathbf{l} - \bar{\mathbf{l}}, \quad (8.9)$$

if $(\mathbf{l}, \bar{\mathbf{l}})$ are the Plücker coordinates of L . We call $L\mu^-$ and $L\mu^+$ *right* and *left image* of the line L .

Remark 8.1.1. The name ‘left image’ for $L\mu^+$ and ‘right image’ for $L\mu^-$ is perhaps contrary to the familiar picture that the left hand part of the real number line consists of the negative numbers, and the right hand part of the positive ones. ◇

The following relation between lines plays a role in elliptic geometry. Here we are going to show its connections to the right and left image of a line and to the spherical kinematic mapping.

Definition. Two lines L_1, L_2 are called *left Clifford parallel*, if $L_1\mu^+ = L_2\mu^+$, and *right Clifford parallel*, if $L_1\mu^- = L_2\mu^-$.

Lemma 8.1.3. If L is a line in P^3 , then the set of lines left Clifford parallel to L is an elliptic congruence. For all points $P \in P^3$ there is a unique line L' incident with P and left Clifford parallel to L . The line $L\mu^+$ is the line left Clifford parallel to L and incident with the origin.

Proof. This lemma is an immediate consequence in turn of the definition of Clifford parallelity and Lemma 8.1.2, of Prop. 3.2.6, and of Equ. (8.7). ◇

Remark 8.1.2. The quadric Λ of (8.5) defines an elliptic geometry in projective three-space according to Ex. 1.1.39. We call its two reguli \mathcal{R}^+ and \mathcal{R}^- the left and the right regulus.

By Lemma 8.1.2, two lines G, H are left Clifford parallel, if both G and H meet the same lines K, L of \mathcal{R}^- , and the same holds for right Clifford parallelity and \mathcal{R}^+ . The name ‘parallel’ is motivated by the fact that all points of G have the same elliptic distance to H , if G, H are right or left Clifford parallel (cf. Remark 8.2.3). ◇

Modified Right and Left Images — the Eckhart-Rehbock Mapping

Remark 8.1.3. It may help to understand the role of \mathcal{R}^+ and \mathcal{R}^- better if we consider a linear line mapping where the analogues of these reguli are real. We begin with a ruled quadric $\tilde{\Lambda}$ and pick a point $O \notin \tilde{\Lambda}$. The two reguli of $\tilde{\Lambda}$ are denoted by $\tilde{\mathcal{R}}^+$ and $\tilde{\mathcal{R}}^-$ (see Fig. 8.2). The mapping ν^+ is defined as follows: If two rulings F_1, F_2 of $\tilde{\mathcal{R}}^-$ intersect the line L , then $L\nu^+$ is the unique line incident with O which intersects the same rulings F_1, F_2 .

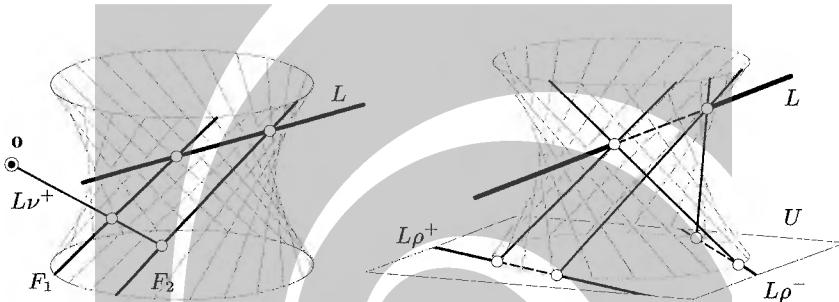


Fig. 8.2. Definition of the mappings ν^+ , ρ^+ and ρ^- by ruled quadrics.

Obviously all lines of the linear congruence with focal lines F_1, F_2 have the same image. To extend this mapping to lines which do not intersect $\tilde{\Lambda}$ in two points, we use the complex extension (for lines which intersect in two conjugate complex points), and limits (for lines which are tangent to $\tilde{\Lambda}$ and lines of $\tilde{\mathcal{R}}^-$) — we do not write down the details here. Only the lines of $\tilde{\mathcal{R}}^+$ are not given an image.

We get a linear line mapping ν^+ quite similar to μ^+ , but with the difference that the set of exceptional lines is a regulus $\tilde{\mathcal{R}}^+$, and the complementary regulus $\tilde{\mathcal{R}}^-$ equals the set of focal lines of fibers. ◇

Remark 8.1.4. We define a mapping ρ^+ similar to the mapping ν^+ of Remark 8.1.3 by choosing a plane U and defining $L\rho^+ = (F_1 \cap U) \vee (F_2 \cap U)$, if F_1, F_2 are two rulings of $\tilde{\mathcal{R}}^-$ which intersect L . The mapping ρ^+ is called the Eckhart-Rehbock mapping. The mapping ρ^- is defined analogously, with $\tilde{\mathcal{R}}^+$ instead of $\tilde{\mathcal{R}}^-$ (see Fig. 8.2). ◇

Kinematic Interpretation of Right and Left Image Line

There is an interesting connection between the mappings μ^+ and μ^- and uniform helical motions (see Fig. 8.3).

Lemma 8.1.4. *If L is the axis of a uniform helical motion of pitch 1 or -1 , then $L\mu^+$ and $L\mu^-$ are the respective path tangents of the origin.*

Proof. Without loss of generality we assume that the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ of L are such that $\|\mathbf{l}\| = 1$. It is a consequence of Th. 3.1.9 that the two helical motions with pitch ± 1 belong to the linear complexes \mathcal{C}^+ , \mathcal{C}^- with generalized Plücker coordinates $(\mathbf{c}^+, \bar{\mathbf{c}}^+) = (\mathbf{l}, \bar{\mathbf{l}} + \mathbf{l})$ and $(\mathbf{c}^-, \bar{\mathbf{c}}^-) = (\bar{\mathbf{l}}, \bar{\mathbf{l}} - \mathbf{l})$, respectively. By Prop. 3.4.1, the velocity vector of the origin equals $\bar{\mathbf{l}} + \mathbf{l}$ or $\bar{\mathbf{l}} - \mathbf{l}$, respectively. \square

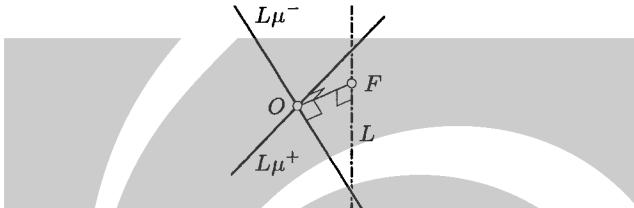


Fig. 8.3. Right and left image $L\mu^-$ and $L\mu^+$ as path tangents of the origin.

Remark 8.1.5. One of the interesting properties of Lemma 8.1.4 is that we can interchange the role of path tangent and helical axis: Not only is $L\mu^+$ a path tangent for a uniform helical motion with axis L , but L is a path tangent for a uniform helical motion with axis $L\mu^+$ (the value of the pitch does not change). The point where L is tangent at is the footpoint F of the common perpendicular (see Fig. 8.3). This is shown by the elementary fact that $O \vee F$ is the common perpendicular of L and $L\mu^+$, and there is a Euclidean motion which exchanges L for $L\mu^+$, and O for F .

We can use this fact to show that the fibers of μ^+ are elliptic linear congruences: A line K with $K\mu^+ = L\mu^+$ is a path tangent in some point F' , with $F'O$ orthogonal to $L\mu^+$, i.e., F' is in the plane ε incident with O and orthogonal to $L\mu^+$. The converse is also true — any such path tangent K has the property $K\mu^+ = L\mu^+$. We have seen in Ex. 3.2.1 that the set of path tangents of points in ε is an elliptic linear congruence. \diamond

Th. 8.1.5 below is the reason why the mappings μ^+ and μ^- are important in spherical kinematics (we are going to study this relationship in more detail and from a different point of view in Sec. 8.2). We first extend the mappings μ^+ and μ^- to *oriented* lines. We use normalized Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ with $\mathbf{l}^2 + \bar{\mathbf{l}}^2 = 1$ for an oriented line \vec{L} . Then the line L is oriented by the vector \mathbf{l} parallel to L if L is a proper line. Note that $\|\mathbf{l}^+\| = 1$, because $(\mathbf{l} + \bar{\mathbf{l}}) \cdot (\mathbf{l} + \bar{\mathbf{l}}) = \mathbf{l}^2 + \bar{\mathbf{l}}^2 = 1$, and analogously $\|\mathbf{l}^-\| = 1$. We define

Definition. If \vec{L} is an oriented line with normalized Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$, then the left and right image points of \vec{L} are defined as

$$\vec{L}\mu^+ = \mathbf{l}^+ = \mathbf{l} + \bar{\mathbf{l}}, \quad \vec{L}\mu^- = \mathbf{l}^- = \mathbf{l} - \bar{\mathbf{l}}. \quad (8.10)$$

The mapping μ^+ maps the space $\vec{\mathcal{L}}$ of oriented lines onto the unit sphere.

Theorem 8.1.5. *The oriented lines \vec{G} and \vec{H} intersect if and only if*

$$\sphericalangle(\vec{G}\mu^+, \vec{H}\mu^+) = \sphericalangle(\vec{G}\mu^-, \vec{H}\mu^-). \quad (8.11)$$

Proof. We assume that the Plücker coordinates $(\mathbf{g}, \bar{\mathbf{g}})$ of \vec{G} are normalized such that $\mathbf{g}^2 + \bar{\mathbf{g}}^2 = 1$. Then both $\mathbf{g}^+ = \mathbf{g} + \bar{\mathbf{g}}$ and $\mathbf{g}^- = \mathbf{g} - \bar{\mathbf{g}}$ are unit vectors. The same holds for \vec{H} . Equ. (8.11) shows that $\mathbf{g}^- \cdot \mathbf{h}^- = \mathbf{g}^+ \cdot \mathbf{h}^+$, which is equivalent to $\bar{\mathbf{g}} \cdot \mathbf{h} + \mathbf{g} \cdot \bar{\mathbf{h}} = 0$. This expresses the fact that \vec{G}, \vec{H} intersect (cf. Equ. (2.26)). \square

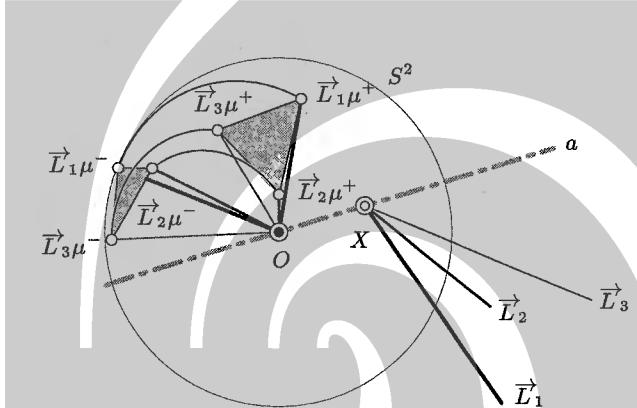


Fig. 8.4. The spherical kinematic mapping and the linear line mappings μ^+, μ^- .

The right and left images of oriented lines are intimately related to the *spherical kinematic mapping*, which is defined as follows and will be studied in Sec. 8.2. Consider a bundle of oriented lines \vec{L}_i concurrent in a point X of projective three-space P^3 . By Th. 8.1.5, the angle enclosed by $\vec{L}_i\mu^+$ and $\vec{L}_j\mu^+$ equals the angle enclosed by $\vec{L}_i\mu^-$ and $\vec{L}_j\mu^-$. This shows that there is a Euclidean congruence transformation κ which transforms the left image points into the right image points (see Fig. 8.4).

It turns out that this spherical motion κ is orientation-preserving, i.e., is a rotation. It is then clear that the axis of this rotation must coincide with the line OX , because $OX\mu^+ = OX\mu^-$. Further it is possible to show that in this way we get a one-to-one correspondence between the points of projective three-space and spherical motions, which is then called the spherical kinematic mapping.

If we consider the right and left images of all lines incident with a plane instead of all lines incident with a point, we get a congruence transformation again, but this one reverses orientation.

Angles of Planes in Four-Space and the Elliptic Metric in P^3

Here we describe an application of the mappings μ^+ and μ^- . We consider the bundle model of P^3 in four-dimensional Euclidean space E^4 , where the point $(x_0, x_1,$

$x_2, x_3)\mathbb{R}$ of P^3 is represented as the line which contains the origin O and which is parallel to the vector (x_0, x_1, x_2, x_3) . Obviously the *angle* between two lines which correspond to points $\mathbf{x}\mathbb{R}, \mathbf{y}\mathbb{R}$ equals the elliptic distance of the points $\mathbf{x}\mathbb{R}, \mathbf{y}\mathbb{R}$, as defined by Equ. (1.79).

Lines of P^3 correspond to *planes* of the bundle. The relative position of two planes in four-space is not as easily described by an angle as it is in three-space — it is necessary to define *two* angles between such planes.

We first have a closer look at the angle enclosed by two planes U, U' in three-space, which have the point O in common. A line L incident with O and contained in U , and a line L' incident with O and contained in U' enclose an angle $\sphericalangle(L, L')$, and for all L there is an L' such that $\sphericalangle(L, L')$ is minimal. We denote this minimal angle by $\sphericalangle(L, U')$. If L happens to be the line $U \cap U'$, of course $\sphericalangle(L, U') = 0$. On the other hand, if L is orthogonal to $U \cap U'$, $\sphericalangle(L, U')$ assumes its maximum, which is denoted by $\sphericalangle(U, U')$. So $\sphericalangle(L, U')$ has *two* stationary values, namely 0 and $\sphericalangle(U, U')$.

We repeat this procedure for planes U, U' in Euclidean four-space, which have a point O in common. Usually $U \cap U'$ is no line but we have $O = U \cap U'$, so the minimum and maximum of $\sphericalangle(L, U')$ are both nonzero. We call them $\sphericalangle_1(U, U')$ and $\sphericalangle_2(U, U')$.

If we use lines of three-space instead of planes of four-space, and the elliptic distance of points instead of the angle between concurrent lines, we get precisely an analogous result, namely that there are *two* distances $d_1(G, G')$ and $d_2(G, G')$ between lines (see Fig. 8.5). It turns out (cf. Remark 8.2.3) that $d_1(G, G') = d_2(G, G')$ means that G, G' are Clifford parallel. The planes in \mathbb{R}^4 which correspond to these lines are called *equi-angular*.

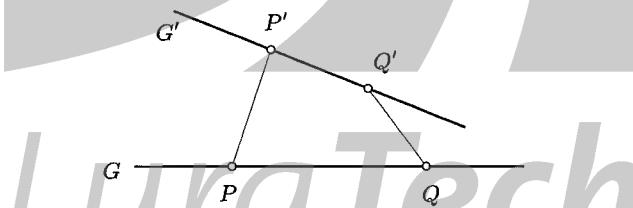


Fig. 8.5. Distances between lines in elliptic space: $d(P, P') = d_1(G, G') = \sphericalangle_1(U, U')$, $d(Q, Q') = d_2(G, G') = \sphericalangle_2(U, U')$.

Lemma 8.1.6. Assume that G, G' are two lines of projective three-space, which is endowed with the elliptic distance (1.79), and that U, U' are the planes of Euclidean four-space which represent G, G' in the bundle model. Then

$$\sphericalangle_1(U, U') = \frac{1}{2}(\beta_1 + \beta_2), \quad \sphericalangle_2(U, U') = \frac{1}{2}|\beta_1 - \beta_2|,$$

$$\beta_1 = \sphericalangle(G\mu^+, G'\mu^+), \quad \beta_2 = \sphericalangle(G\mu^-, G'\mu^-),$$

The proof is left to the reader as an exercise.

Remark 8.1.6. Lemma 8.1.6 can be used to construct packings of planes in E^4 (N.J.A. Sloane [185]). The question is how to arrange (in four-space) N planes incident with the fixed point O such that they are as far apart as possible. As a distance measure between two planes U, U' in E^4 , Sloane uses

$$d(U, U')^2 = \sin^2 \sphericalangle_1(U, U') + \sin^2 \sphericalangle_2(U, U'), \quad (8.12)$$

A set $\{U_1, \dots, U_N\}$ of concurrent planes in E^4 corresponds to a set $\{G_1, \dots, G_N\}$ of lines in P^3 , which is represented by a ‘binocular code’ of the pairs $(G_i\mu^-, G_i\mu^+)$ of right and left image lines. We have a list of N (not necessarily distinct) right lines in the ‘right code’ and accordingly N lines of the ‘left code’.

Sloane et al. have derived various optimal packings. An example is the optimal packing for $N = 6$, where the right and left code lines intersect the unit sphere S^2 in the 12 vertices of an icosahedron. The correspondences between right and left image are illustrated in Fig. 8.6. It is easily seen from the arrangement of right and left image that all pairs of planes $U_i \neq U_j$ have the same distance $d(U_i, U_j)$, and a simple computation shows that $d^2 = 6/5$. \diamond

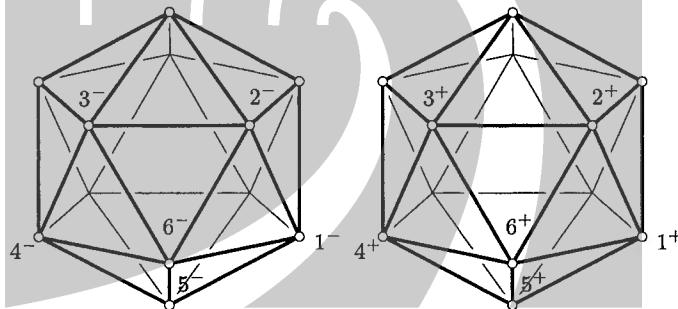


Fig. 8.6. Right and left image points $1^-, \dots, 6^-$, $1^+, \dots, 6^+$ of an optimal packing of six concurrent planes in E^4 .

Remark 8.1.7. One particular application of the concepts of concurrent planes in four-space (cf. Remark 8.1.6) — or more generally, packings of concurrent subspaces in E^n — is the following. Having data points in E^n , whose coordinates represent certain quantities taken from a sample of objects, one wants to analyze possible relations between these quantities. For this purpose, projections to lower-dimensional spaces (such as planar images) are of interest. In order to optimally traverse the set of two-dimensional images, one may try to arrange several image planes as uniformly as possible and take orthogonal projections onto them. In this way, we obtain a number of views which are as different from each other as possible. An interactive visualization tool would then need simple ways to get refined view directions. At least for the case of four-space, this can be accomplished by subdivision techniques on the unit sphere, which lead to ‘nearly’ optimal arrangements. \diamond

The Blaschke-Grünwald Mapping

The linear line mappings μ^- , μ^+ discussed above lead in a natural way to a mapping of points in P^3 to spherical motions (see Sec. 8.2). There is another pair of linear line mappings, denoted by β^- and β^+ , which are analogously related to motions of the Euclidean plane.

We fix a Cartesian coordinate system in E^3 and consider horizontal all planes $z = \text{const}$.

Definition. Suppose that L is a non-horizontal line in E^3 , which contains the point $(s_1, s_2, 0)$ and is parallel to the vector $(l_1, l_2, 1)$. Then (see Fig. 8.7)

$$L\beta^+ = (s_1 - l_2, s_2 + l_1, 0), \quad L\beta^- = (s_1 + l_2, s_2 - l_1, 0). \quad (8.13)$$

Thus β^+ and β^- are mappings from the space of lines onto the horizontal plane $z = 0$.

Remark 8.1.8. If L is not horizontal, there is a unique uniform helical motion of pitch 1 and with a vertical axis, such that L is a path tangent for the point $(s_1, s_2, 0) \in L$. Obviously the axis of this helical motion is incident with $L\beta^+$. Analogously we get $L\beta^-$ if we look for a helical motion of pitch -1 . \diamond

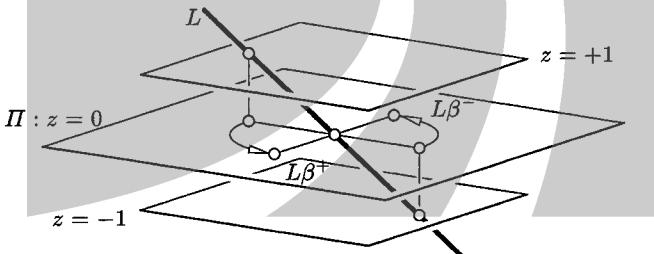


Fig. 8.7. Construction of right and left image points $L\beta^-$ and $L\beta^+$.

In order to verify that both β^+ and β^- are linear line mappings, we assume that L contains the point $s = (s_1, s_2, 0)$ and is parallel to the vector $(l_1, l_2, 1)$. We compute the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}})$ of L and express $L\beta^+$ and $L\beta^-$ in terms of Plücker coordinates:

$$\begin{aligned} \mathbf{l} &= (l_{01}, l_{02}, l_{03}) = (l_1, l_2, 1), \quad \bar{\mathbf{l}} = (l_{23}, l_{31}, l_{12}) = (s_2, -s_1, s_1 l_2 - s_2 l_1), \\ L\beta^+ &= (l_{03}, -l_{02} - l_{31}, l_{01} + l_{23}, 0)\mathbb{R}, \\ L\beta^- &= (l_{03}, l_{02} - l_{31}, -l_{01} + l_{23}, 0)\mathbb{R}. \end{aligned} \quad (8.14)$$

This shows that β^+ and β^- are linear line mappings. As to their kernels and fibers, we do the following:

Remark 8.1.8 shows that the set of lines L with $L\beta^+ = G$ equals the set of path tangents of the same uniform helical motion with axis G . The same is true for β^- instead of β^+ . So Ex. 3.2.1 shows that the fibers of β^- and β^+ are elliptic linear congruences. The focal lines of these congruences are also described there, and an elementary computation shows that the focal lines of the fibers of β^+ are exceptional lines for β^- and vice versa. The projection centers for β^+ and β^- are two-dimensional. As the description of focal lines in Ex. 3.2.1 tells us, these planes intersect the Klein quadric in two conjugate complex lines, which correspond to two conjugate complex pencils of exceptional lines.

The following result, which is a counterpart of Th. 8.1.5, is the reason why β^+ and β^- are important in planar kinematics.

Theorem 8.1.7. *Two non-horizontal lines G, H of Euclidean space E^3 intersect in P^3 if and only if $d(G\beta^-, H\beta^-) = d(G\beta^+, H\beta^+)$.*

Proof. We use Plücker coordinates $(\mathbf{g}, \bar{\mathbf{g}})$ and $(\mathbf{h}, \bar{\mathbf{h}})$ for G and H with $g_{03} = h_{03} = 1$. The condition on the distances reads $(g_{02} - g_{31} - h_{02} + h_{31})^2 + (-g_{01} + g_{23} + h_{01} - h_{23})^2 = (-g_{02} - g_{31} + h_{02} + h_{31})^2 + (g_{01} + g_{23} - h_{01} - h_{23})^2$. The Plücker relation implies that this equation is equivalent to $\mathbf{g} \cdot \mathbf{h} + \bar{\mathbf{g}} \cdot \bar{\mathbf{h}} = 0$. \square

We can use Th. 8.1.7 to find a *planar kinematic mapping* (or *Blaschke-Grünwald mapping*) in just the same way as we used Th. 8.1.5 to find the spherical kinematic mapping. A point of P^3 , which is not contained in the horizontal line at infinity, is the carrier of a bundle of lines L , which have the property that $d(L\beta^+, L\beta^-)$ is a constant. Therefore there is a planar Euclidean congruence transformation which transforms $L\beta^-$ into $L\beta^+$ for all L incident with X . It turns out that we get a one-to-one correspondence between the rotations and translations of the Euclidean plane and the points of P^3 (except those contained in the horizontal line at infinity). The discussion of this mapping will be continued in Sec. 8.2.3.

Ivory's Theorem

If $\mu = \gamma\lambda$ is a linear line mapping onto a projective plane, the center Z of λ is two-dimensional. As the Klein image $\mathcal{R}\gamma$ of a regulus \mathcal{R} is a planar conic, $\mathcal{R}\mu$ is a line (if the plane of $\mathcal{R}\gamma$ meets Z) or a conic (if the intersection of Z and the plane of $\mathcal{R}\gamma$ is void). We are especially interested in the cases $\mu = \beta^+$ and $\mu = \beta^-$, and consider the reguli $\mathcal{R}_e, \mathcal{R}_f$ contained in the ruled quadric $\Phi : x^2/a^2 + y^2/b^2 - z^2 = 1$, which possesses the two parametrizations

$$(x, y, z) = (a(\cos u - v \sin u), \sin u + v \cos u, bv), \\ (x, y, z) = (a(\cos u + v \sin u), \sin u - v \cos u, bv).$$

The lines of \mathcal{R}_e and \mathcal{R}_f are the lines $u = \text{const.}$ of these parametrizations, so in the notation of Equ. (8.13), a line $E(u)$ of \mathcal{R}_e (or a line $F(u)$ of \mathcal{R}_f) is characterized by $s_1 = a \cos u, s_2 = \sin u, l_1 = \mp(a/b) \sin u, l_2 = \pm(1/b) \cos u$, which shows that

$$\begin{aligned} E(u)\beta^- &= F(u)\beta^+ = ((a + 1/b) \cos u, (1 + a/b) \sin u, 0), \\ E(u)\beta^+ &= F(u)\beta^- = ((a - 1/b) \cos u, (1 - a/b) \sin u, 0). \end{aligned} \quad (8.15)$$

We see that the points $E(u)\beta^-$ and $F(u)\beta^+$ are contained in the same ellipse with semiaxes $A = a + 1/b$ and $B = 1 + a/b$, and that the points $E(u)\beta^+$, $F(u)\beta^-$ are contained in a second ellipse with semiaxes $A' = a - 1/b$, $B' = 1 - a/b$ (see Fig. 8.8). These ellipses have the same axes of symmetry and even the same focal points, because $A^2 - B^2 = A'^2 - B'^2$. Furthermore, an elementary computation shows that the points $E(u)\beta^+$ and $E(u)\beta^-$ are contained in a hyperbola which again has the same focal points. The same is true for $F(u)$ instead of $E(u)$.

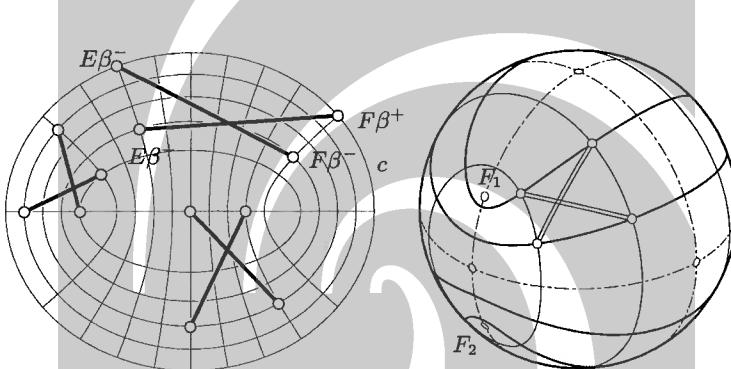


Fig. 8.8. Left: Images of the reguli on a ruled quadric and Ivory's diagonal property in a system of confocal conics. Right: A confocal family of spherical conics and the spherical version of Ivory's theorem (courtesy H. Stachel).

A system of confocal conics (ellipses and hyperbolae) is shown in Fig. 8.8, left. It is well known that all intersections of conics in this system are orthogonal. By Th. 8.1.7, the length of the two diagonals in the ‘curvilinear rectangle’ $E(u_1)\beta^-$, $E(u_1)\beta^+$, $F(u_2)\beta^-$, $F(u_2)\beta^+$ are equal.

Theorem 8.1.8. (J. Ivory) *The two straight diagonals of a curvilinear rectangle in a system of confocal conics have equal length.*

Proof. (Sketch) In case of ellipses and hyperbolae we have done the greater part of the proof already. It remains to show that all possible curvilinear rectangles occur in the way described in the previous paragraph. This is done by adjusting parameters a, b, u_1, u_2 . We omit the details.

The case of confocal parabolae (see Fig. 8.9) is similar and is based on the two reguli of a hyperbolic paraboloid. \square

Our example deals only with a rather special case of a ruled quadric. There is, however, the following result:

Theorem 8.1.9. *Assume that \mathcal{R}_e and \mathcal{R}_f are complementary reguli. Then the curve pairs $(\mathcal{R}_e\beta^-, \mathcal{R}_e\beta^+)$ and $(\mathcal{R}_f\beta^+, \mathcal{R}_f\beta^-)$ are congruent pairs of confocal conics.*

There is an elegant proof of Th. 8.1.9 which uses the kinematic mapping of Blaschke and Grünwald.

The Trilateration Problem in Surveying

The following beautiful application of the mapping pair (β^-, β^+) to a problem in surveying has been given by H. Stachel [188]. He considers the *planar trilateration problem*: Determine the relative position of six pairwise distinct points $e_1, e_2, e_3, f_1, f_2, f_3$ in the Euclidean plane E^2 , if the nine distances $\overline{e_i f_j}$ ($i, j \in \{1, 2, 3\}$) are known. It turns out that this amounts to finding the zeros of some polynomial of degree eight. We are interested in *unstable* positions of the framework which is defined by vertices e_1, \dots, f_3 and rigid rods of length $\overline{e_i f_j}$ between them (cf. p. 218). Such an instance of the trilateration problem is also called *unstable*.

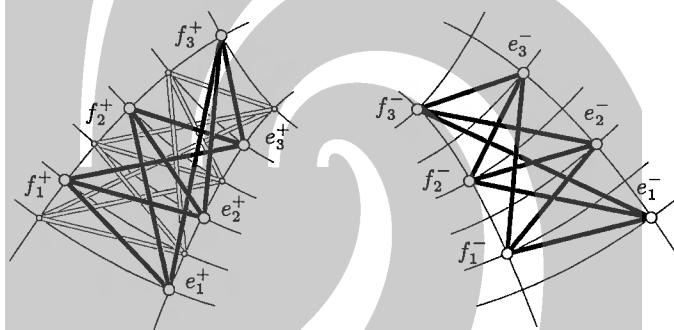


Fig. 8.9. Two solutions of the trilateration problem. α (see text) is the reflection in the vertical axis of symmetry.

Theorem 8.1.10. *An instance of the trilateration problem is unstable, if and only if the points $e_1, e_2, e_3, f_1, f_2, f_3$ of its solution are contained in a conic or the union of two lines.*

Proof. (Sketch) The proof according to Stachel [188] starts with two different solutions e_1^-, \dots, f_3^- and e_1^+, \dots, f_3^+ of some instance of the trilateration problem, i.e.,

$$\overline{e_i^- f_j^-} = \overline{e_i^+ f_j^+}, \quad \text{but} \quad \overline{e_i^- e_j^-} \neq \overline{e_i^+ e_j^+}, \quad \overline{f_i^- f_j^-} \neq \overline{f_i^+ f_j^+}, \quad i, j = 1, 2, 3. \quad (8.16)$$

Thus the two solutions cannot be transformed into each other by a Euclidean congruence transformation. There are lines E_1, \dots, E_3 such that $e_i^+ = E_i \beta^+$ and $e_i^- = E_i^-$, and analogously lines F_1, \dots, F_3 defined by the points f_i^+, f_i^- . Equ. (8.16) shows that E_i intersects F_j , for all i and j , but the lines E_i are pairwise skew, and so are the lines F_i . Thus there are two complementary reguli \mathcal{R}_e and \mathcal{R}_f with $E_i \in \mathcal{R}_e$ and $F_i \in \mathcal{R}_f$. It is possible to show that there is a Euclidean congruence

transformation α such that $\mathcal{R}_e\beta^+ = \mathcal{R}_f\beta^- \alpha$ and $\mathcal{R}_e\beta^- = \mathcal{R}_f\beta^+ \alpha$ are confocal conics (cf. Th. 8.1.9). The points $e_i, f_i \alpha$ look like in Fig. 8.9.

If we perturb an unstable instance of the trilateration problem we get instances which have two solutions lying closely together. Thus for an unstable solution we must have $e_i^+ = e_i^-$ and $f_i^+ = f_i^-$, both contained in the same conic. The case that the right hand set of inequalities in (8.16) does not hold is similar. We get the result that the points are distributed in the union of two lines. \square

Note that the proof above shows more than just the unstable solutions, it describes in general the relation between two different solutions of the trilateration problem.

Remark 8.1.9. The *spherical* trilateration problem is treated in much the same way by replacing the mappings β^+, β^- by the mappings μ^+, μ^- .

Unstable solutions of the trilateration problem are characterized by being contained in a *spherical conic*, which is an algebraic curve of degree four, symmetric with respect to the sphere's center, and consisting of two spherical ovals. Each of the components of a spherical conic may be defined as the locus of points whose spherical distances to two fixes points F_1 and F_2 sum up to a constant (see Fig. 8.8, right).

Note that there is no need to define a ‘spherical hyperbola’: If $d(p, q)$ denotes the geodesic distance, then $d(p, F_1) - d(p, F_2) = \text{const}$ is the same as $d(p, F_1) + d(p, \bar{F}_2) = \text{const}$, where \bar{F}_2 and F_2 are antipodal points.

All spherical conics are the intersection of the sphere with a quadratic cone whose vertex is the center of the sphere. \diamond

8.1.2 Linear Line Mappings into P^3

Linear line mappings to P^3 have been studied by Ch. Lübbert [115], who gave a classification and some results based on the Klein model. Only recently it has been investigated how to generate these mappings geometrically in P^3 [133].

If μ is a linear line mapping onto P^3 and $\mu = \gamma\lambda$ with a linear mapping λ , then λ 's center Z is one-dimensional. The classification of projective types of μ depends on the intersection of Z with the Klein quadric. Here we consider only the case where the intersection is void — after performing the complex extension the intersection $Z_{\mathbb{C}} \cap M_2^4(\mathbb{C})$ consists of two conjugate complex points. For the remaining cases we refer to [133].

It turns out that such linear line mappings are closely related to the first nontrivial example of a Hopf mapping, which plays a role e.g. in the theory of fiber bundles. Dietz et al. [36] used this mapping to find a representation of rational curves and surfaces contained in a sphere. Because they (and we) define this mapping in a more line-geometric way, it is called a *generalized stereographic projection*.

This section is organized in the following way: first we study *projections via linear line congruences*, especially elliptic ones. Adding a stereographic projection will yield the generalized stereographic projection, and finally we discuss the linear line mapping associated with it. We also briefly show applications in computer aided geometric design.

Projection via a Linear Congruence

Linear line congruences can be used to map points of projective space P^3 to points of a projective plane P^2 . This is a special case of a projection onto a surface via a line congruence (cf. Sec. 7.1.1).

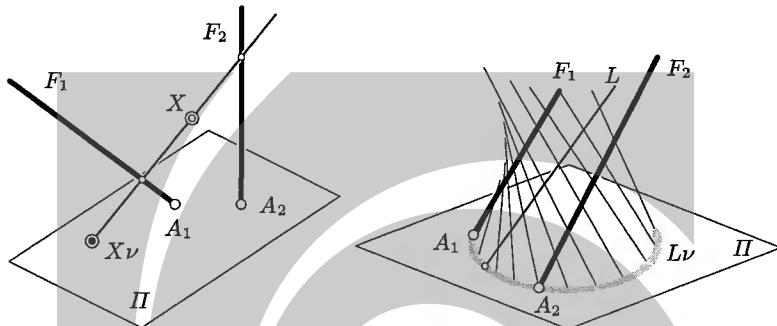


Fig. 8.10. Projection via a hyperbolic linear line congruence

We first discuss the case of projection ν via a hyperbolic linear line congruence \mathcal{N} with focal lines F_1, F_2 onto a plane Π . The left hand part of Fig. 8.10 shows how to find the image $X\nu$ of a point X : The unique line which contains X and meets both F_1 and F_2 intersects the plane Π in the point $X\nu$. If $X \in F_1$ or $X \in F_2$, $X\nu$ is undefined.

Lemma 8.1.11. *If a line L is contained in \mathcal{N} , $L\nu$ is a point. If $L \notin \mathcal{N}$, $L\nu$ is a straight line or a conic which contains $F_1 \cap \Pi$ and $F_2 \cap \Pi$.*

Proof. The statement about the case $L \in \mathcal{N}$ is obvious. If $L \notin \mathcal{N}$, then the set of lines in \mathcal{N} incident with L is precisely the set of lines which meet L, F_1, F_2 , and is therefore a regulus \mathcal{R} . If all lines of \mathcal{R} intersect Π in a unique point, $L\nu = \{(G \cap \Pi) \mid G \in \mathcal{R}\}$, and so $L\nu$ is a conic (see Fig. 8.10, right). If some line $G_0 \in \mathcal{R}$ happens to be contained in Π , $L\nu = \{(G \cap \Pi) \mid G \in \mathcal{R}, G \neq G_0\}$, which is a straight line (cf. the discussion of tangential intersections of quadrics in Sec. 1.1.5).

If $R \in \mathcal{R}$, all points of R have the same ν -image. If we choose R incident with $A_i = F_i \cap \Pi$ ($i = 1, 2$), we see that $(R \cap L)\nu = A_i\nu = A_i$. This shows that $F_i \cap \Pi \in L\nu$. \square

We see that ν is no linear mapping, but a quadratic one. If \mathcal{N} is elliptic, we use the complex extension to make it hyperbolic and we are able to show the same properties. In the case that \mathcal{N} is elliptic the mapping ν is defined for all real points, because the focal lines are non-real. This is a property which distinguishes it from the projection via hyperbolic congruences.

For the following it will be useful to have an analytic description of the projection via the linear congruence \mathcal{N} with pitch 1 of Ex. 3.2.1, which has rotational symmetry. We use a Cartesian coordinate system such that Π is the xy -plane.

The line of \mathcal{N} incident with the point $(x_0, y_0, 0)$ has the parametric representation $(x_0, y_0, 0) + t(-y_0, x_0, 1)$, and therefore the congruence projection maps a point (x, y, z) according to

$$\nu : (x, y, z) \mapsto \frac{1}{1+z^2}(x+yz, y-xz, 0). \quad (8.17)$$

Using Equ. (8.17), it is straightforward to verify that a line L with Plücker coordinates (l_{01}, \dots, l_{12}) is mapped onto the curve

$$L\nu : l_{03}(x^2 + y^2) + (l_{31} - l_{02})x + (l_{01} - l_{23})y + l_{12} = 0. \quad (8.18)$$

This is a circle if $l_{03} \neq 0$, i.e., if L is not parallel to Π , and a straight line otherwise. This follows from the fact that the restriction of ν to a horizontal plane is a similarity transformation. The coefficients in Equ. (8.18) are linear combinations of L 's Plücker coordinates. This suggests to construct a linear line mapping into some ‘projective space of circles’, where the coefficients of the circles’ equations serve as homogeneous coordinates.

This result also suggests that the projective geometry of lines is not the appropriate tool to study the mapping ν . We should use the geometry of circles and lines (which are circles incident with a unique point at infinity), i.e., planar Möbius geometry.

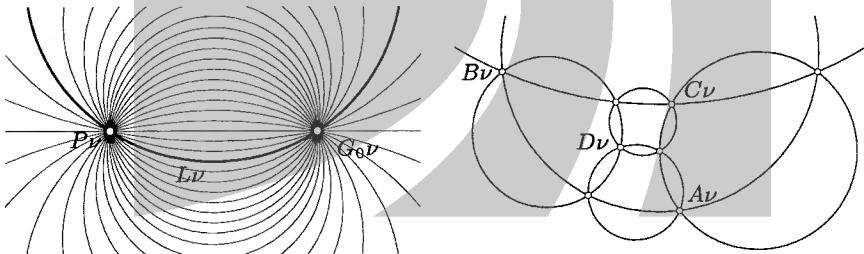


Fig. 8.11. Left: The projection via an elliptic linear congruence with rotational symmetry maps a pencil of lines to a pencil of circles. Right: Miquel’s circle configuration is the ν -image of a tetrahedron $ABCD$.

Example 8.1.4. Equ. (8.18) shows that ν maps lines in P^3 (except the horizontal line at infinity) to circles in Π , where ‘circle’ means either an ordinary circle or a line, incident with the plane’s ‘point at infinity’. A *pencil of lines* with vertex P and carrier plane U is mapped onto a *pencil of circles*. There is a line G_0 of \mathcal{N} which is contained in U . If L is a line of the pencil, $L\nu$ is incident with $P\nu$ and $(L \cap G_0)\nu = G_0\nu$ (see Fig. 8.11). \diamond

Example 8.1.5. We continue the discussion of Ex. 8.1.4. The ν -image of a tetrahedron is the famous *Miquel configuration* of circles (see Fig. 8.11). It is a configuration of six circles (the images of the tetrahedron’s edges), which intersect in eight

points (the images of the tetrahedron's vertices plus the images of congruence lines in the tetrahedron's faces).

Since each edge is incident with two vertices and two faces, each circle of the Miquel configuration contains four of the eight points. Each vertex or face of the tetrahedron is incident with three edges, and thus each of the eight points is contained in three circles. \diamond

The Hopf Mapping

We consider the stereographic projection σ from a unit sphere of Euclidean three-space onto its equator plane (cf. Ex. 1.1.37, Lemma 4.3.1, Ex. 4.3.1). If we use homogeneous coordinates for a point of the unit sphere, its inverse reads

$$\sigma^{-1} : (x, y, 0) \mapsto (x^2 + y^2 + 1 : 2x : 2y : x^2 + y^2 - 1). \quad (8.19)$$

Recall (or see the paragraph preceding Equ. (8.21)) that circles and straight lines are mapped to circles in the sphere by σ^{-1} . It is convenient to extend the Euclidean plane by one point at infinity, denoted by the symbol ' ∞ ', which is incident with all lines. If we let $\sigma^{-1}(\infty) = (1 : 0 : 0 : 1) = N$, σ becomes one-to-one and onto, and maps circles to circles, if we now regard lines as special cases of circles.

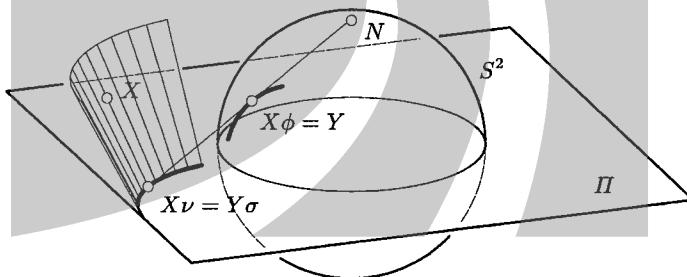


Fig. 8.12. Connection between the Hopf mapping ϕ , the stereographic projection σ , and the congruence projection ν .

Consider the congruence projection ν defined by Equ. (8.17) and the mapping $\phi : X \mapsto X\nu\sigma^{-1}$. It is called the *Hopf mapping*. If $X\nu$ is no ideal point (i.e., X is not contained in the horizontal line at infinity), it is easy to check that

$$\phi : \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} x_0^2 + x_1^2 + x_2^2 + x_3^2 \\ 2(x_0x_1 + x_2x_3) \\ 2(x_0x_2 - x_1x_3) \\ -x_0^2 + x_1^2 + x_2^2 - x_3^2 \end{bmatrix}. \quad (8.20)$$

We use Equ. (8.20) as definition of the Hopf mapping. Some of its properties are easily found by the factorization into congruence projection and inverse stereographic

projection. Clearly the fibers of ϕ are the lines of the elliptic congruence \mathcal{N} discussed in the previous paragraph. The image of a line L with $L \notin \mathcal{N}$ is a circle, because $L\nu$ is a circle or a line.

Remark 8.1.10. We consider the two-dimensional complex vector space \mathbb{C}^2 and identify it with \mathbb{R}^4 . With $z_0, z_1 \in \mathbb{C}$, the expression $(z_0, z_1)\mathbb{R}$ means a point of real projective space P^3 .

The expression $(z_0, z_1)\mathbb{C}$ on the other hand means a point of the complex projective line, and obviously there is a natural ‘Hopf’ mapping

$$(z_0, z_1)\mathbb{R} \rightarrow (z_0, z_1)\mathbb{C}.$$

If $(z_0 : z_1)$ does not equal the point $(0 : 1)$, we can identify it with the point $z = z_1/z_0$ of the affine complex line $\mathbb{C}^1 = \mathbb{R}^2$, and project it stereographically onto the sphere. The ideal point $(0 : 1)$ projects, by definition, to the north pole. If we let $z_0 = u_0 + iv_0$, $z_1 = u_1 + iv_1$, the result of this mapping is $(u_0^2 + v_0^2 + u_1^2 + v_1^2, 2(u_1u_0 - v_1v_0), 2(v_1u_0 + u_1v_0), -u_0^2 - v_0^2 + u_1^2 + v_1^2)$. Apart from a permutation of coordinates, this is the Hopf mapping (8.20). \diamond

A Linear Line Mapping Related to the Hopf Mapping

We consider a circle or straight line c in the plane $\Pi : z = 0$ which has the equation

$$A(x^2 + y^2) + Bx + Cy + D = 0.$$

In order to compute $c\sigma^{-1}$ we use the cone with base curve c and vertex $(0, 0, 1)$, which has the equation $A(x^2 + y^2) + Bx(1 - z) + Cy(1 - z) + D(1 - z)^2 = 0$, and intersect it with the unit sphere $x^2 + y^2 = 1 - z^2$ by substituting $1 - z^2$ for $x^2 + y^2$. This yields the equation $(1 - z)(A(1 + z) + Bx + Cy + D(1 - z)) = 0$ and shows that apart from a non-real component in the plane $z = 1$, the circle $c\sigma^{-1}$ is contained in the plane

$$A(1 + z) + Bx + Cy + D(1 - z) = 0. \quad (8.21)$$

Thus σ^{-1} maps circles and lines to circles. We consider the set $L\nu$ defined by (8.18), where ν is the congruence projection of Eqn. (8.17). Obviously, if L has the Plücker coordinates $(l_{01} : \dots : l_{12})$, then the curve $L\phi = L\nu\sigma^{-1}$ is the intersection of the unit sphere S^2 with the plane $(u_0 : u_1 : u_2 : u_3) = (A + D : B : C : A - D) = (l_{03} + l_{12} : l_{31} - l_{02} : l_{01} - l_{23} : l_{03} - l_{12})$. This shows that ϕ induces a linear mapping from the space of lines to the space of planes in P^3 . In order to get a mapping from \mathcal{L} to P^3 itself, we may add any duality. It seems natural to use the polarity π defined by the quadric S^2 itself:

$$\pi : \mathbb{R}(u_0, u_1, u_2, u_3) \mapsto (-u_0, u_1, u_2, u_3)\mathbb{R}. \quad (8.22)$$

So we finally have a mapping μ from \mathcal{L} into P^3 , which is defined by the formula

$$L\mu = (l_{03} + l_{12}, l_{02} - l_{31}, l_{23} - l_{01}, l_{12} - l_{03})\mathbb{R}. \quad (8.23)$$

The following theorem sums up properties of the mappings π , ϕ , and μ :

Theorem 8.1.12. Consider the Hopf mapping ϕ , the polarity π with respect to the unit sphere S^2 , and the linear line mapping μ , which are defined by (8.20), (8.22), and (8.23), respectively. Then $\mu = \phi\pi$. If L is a line and U_L is the plane with $U_L \cap S^2 = L\phi$, then $L\mu = U_L\pi$. The image of μ consists of all points of P^3 which are not inside S^2 . A linear line mapping into P^3 defined for all lines is projectively equivalent to μ .

Proof. The equation $\mu = \phi\pi$ follows directly from our derivation of the mapping μ . The statement about μ 's image follows from the fact that $L\phi$ can be any circle in the unit sphere.

A linear line mapping $\mu = \gamma\lambda$ into P^3 which is defined for all lines has the property that λ 's center Z is one-dimensional and does not intersect M_2^4 , so $Z\mu_2^4 = \mathcal{N}\gamma$, with an elliptic linear congruence \mathcal{N} . As all elliptic linear congruences are projectively equivalent, \mathcal{N} determines Z by $Z = [\mathcal{N}\gamma]\mu_2^4$, and since linear mappings with the same center have projectively equivalent images, the theorem follows. \square

The class of linear line mappings mentioned in Th. 8.1.12 is called of *elliptic type*.

Remark 8.1.11. Recall from Sec. 1.1.5 the definition of ‘silhouette’ of a surface Φ with respect to a central projection with center Z . If Φ is a quadric, the silhouette is the image of $Z\pi \cap \Phi$, where π is the polarity defined by Φ (cf. Prop. 1.1.42).

In the notation of the proof of Th. 8.1.12, the silhouette of M_2^4 with respect to the linear mapping μ equals the image of $Z\mu_2^4 \cap M_2^4 = \mathcal{N}\gamma$. It is not difficult to identify \mathcal{N} with the congruence \mathcal{N} used in the definition of ν and ϕ . So we can say that the μ -silhouette of the Klein quadric is the image of $\mathcal{N}\mu$, i.e., the unit sphere S^2 . This is in accordance with the fact that S^2 actually equals the boundary of the set $M_2^4\mu$, as is to be expected of silhouettes. \diamond

Rational Curves and Surfaces Contained in Quadrics

If $c(t)$ is a rational curve of degree d , and ϕ is the Hopf mapping, then $c(t)\phi$ is a rational curve of degree $2d$, as ϕ 's defining equation (8.20) is quadratic. The fibers $\phi^{-1}(X)$ of points are the lines of the linear congruence \mathcal{N} used in the definition of ϕ . If $\tilde{c}(t)$ is another curve, such that for all t , $c(t) \vee \tilde{c}(t) \in \mathcal{N}$, then obviously $c(t)\phi = \tilde{c}(t)\phi$. Thus there are many curves whose ϕ -image is the same. At least one of them has a very special property:

Theorem 8.1.13. For all rational curves $d(t)$ of degree $2n$ contained in the unit sphere, there exists a rational curve $c(t)$ of degree n such that $c(t)\phi = d(t)$.

For a proof, see [36]. Curves of odd degree cannot be contained in a sphere, because such curves must intersect all planes of space.

Lemma 8.1.14. An elliptic linear congruence \mathcal{N} admits a one-parameter group G of projective automorphisms, which leaves every line of \mathcal{N} invariant, and which is isomorphic to the group of rotations in \mathbb{R}^2 . If $L \in \mathcal{N}$ and $X, Y \in L$, then there is exactly one $\kappa \in G$ with $X\kappa = Y$.

Proof. All elliptic linear congruences are projectively equivalent, so we can restrict ourselves to the linear congruence \mathcal{N} with equations $l_{01} + l_{23} = l_{02} + l_{31} = 0$ (this is case $p = 1$ in Ex. 3.2.1, and equals the congruence used in the definition of ν and ϕ). We consider the set G of projective automorphisms κ_u , defined by

$$\kappa_u(x\mathbb{R}) = (A_u \cdot x)\mathbb{R}, \quad A_u = \begin{bmatrix} \cos u & \cos u & -\sin u & -\sin u \\ \sin u & \sin u & \cos u & \cos u \end{bmatrix}. \quad (8.24)$$

It is easily verified that $\kappa(u)\kappa(v) = \kappa(u+v)$ (so G is a group), and that $\kappa(u) = \kappa(v)$ if and only if $u - v$ is an integer multiple of π (so G is isomorphic to SO_2). An elementary computation shows that for all points p , $p \vee p\kappa_u \in \mathcal{N}$. Further, If $b = A_{\pi/2} \cdot a\mathbb{R}$, then

$$(\cos va + \sin vb)\mathbb{R}\kappa_u = (\cos(u+v)a + \sin(u+v)b)\mathbb{R}.$$

This shows that all lines of \mathcal{N} are invariant lines; and that for all $L \in \mathcal{N}$, $X, Y \in L$ we can find $\kappa(u)$ such that $X\kappa(u) = Y$. \square

Remark 8.1.12. The fact that the group SO_2 acts transitively on the fibers of the elliptic linear congruence \mathcal{N} makes \mathcal{N} a principal fiber bundle $\phi: P^3 \rightarrow S^2$ and is therefore interesting in differential geometry and topology. \diamond



Corollary 8.1.15. *If \mathcal{N} is an elliptic linear congruence and \mathcal{R} is a congruence surface of degree $2n$, then the point set of \mathcal{R} is the union of a one-parameter family of rational curves of degree n .*

Proof. Consider the linear line mapping μ defined by (8.23) and the Hopf mapping ϕ . The curve $\mathcal{R}\gamma$ is a rational curve of degree $2n$ in the oval quadric $\mathcal{N}\gamma$ (cf. Th. 5.2.8). If $L \in \mathcal{N}$, the curve $L\phi$ is a point, and Th. 8.1.12 shows that $\mathcal{R}\phi = \mathcal{R}\mu$, i.e., $\mathcal{R} = \phi^{-1}(c)$ with $c = \mathcal{R}\mu$. The curve c is a linear image of $\mathcal{R}\gamma$, so it is rational of degree $2n$ and by Th. 8.1.13 there is a curve d of degree n with $d\phi = c$. Obviously d meets every line of \mathcal{R} .

To find a one-parameter family of curves which covers \mathcal{R} , we apply the projective mappings κ_u of Lemma 8.1.14 to d . \square

The construction of rational curves and surface patches on the sphere with aid of the Hopf mapping has various applications to curve and surface design problems. Here we show how to define an invariant control structure for spherical Bézier and NURBS curves [146]:

If $X_1, X'_1, X_2, X'_2, \dots$ is a spherical polygon, then we consider the lines $L_i = \phi^{-1}(X_i)$ and $L'_i = \phi^{-1}(X'_i)$ and choose a point $P_1 \in L_1$. As all lines L_i, L'_j are skew, there is a unique polygon $P_1, P'_1, P_2, P'_2, \dots$ such that $P_i \in L_i, P'_j \in L'_j$, and that $\{P_i, P'_i, P_{i+1}\}$ is collinear. This polygon has an interpretation of a geometric control polygon of a NURBS curve $c(t)$ in P^3 , and we define $c(t)\phi$ to be the spherical NURBS curve defined by the original polygon X_1, \dots (see Fig. 8.13).

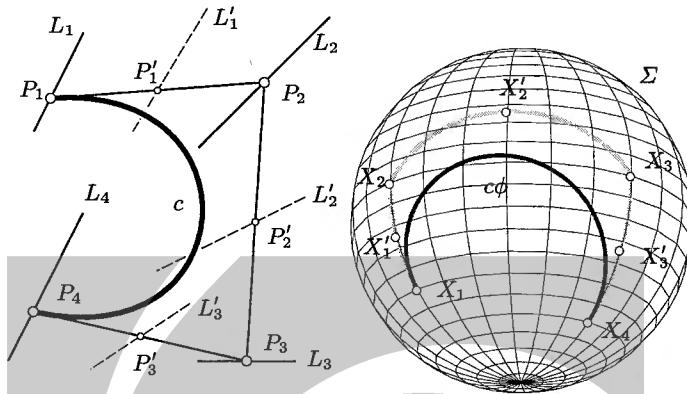


Fig. 8.13. Right: example of a spherical curve $c\phi$ with spherical control structure X_1, X'_1, \dots . Left: The ϕ -preimage of the spherical control structure defines a curve c .

Lemma 8.1.16. *If the curve $c(t)$ defined by the spherical polygon X_1, X'_1, X_2, \dots and the point $P_1 \in \phi^{-1}(X_1)$ is a NURBS curve of degree d over the knot vector t_0, t_1, \dots with multiplicities μ_0, μ_1, \dots , then $c(t)\phi$ is a NURBS curve of degree $2d$ over a knot vector with the same knots, but with multiplicities $\mu_k + d$. The curve $c(t)\phi$ does not depend on the choice of P_1 .*

If the points X_i, X'_j undergo a spherical motion (indeed, any projective automorphism of the sphere), the curve $c(t)\phi$ transforms accordingly.

Proof. If $c(t)$ is of degree d , then $c(t)\phi$ has degree $2d$. It has the same continuity and differentiability properties as $c(t)$, because ϕ is a rational mapping, so Th. 1.4.15 shows the statement about the multiplicities.

If we choose a different point \bar{P}_1 instead of P_1 , thereby defining a curve $\bar{c}(t)$, there is a unique projective automorphism κ_u , defined by Equ. (8.24), such that $P_1\kappa_u = \bar{P}_1$. As the construction of the polygon P_1, P'_1, \dots is performed by taking spans and intersections of subspaces and is therefore projectively invariant, we have $\bar{c}(t) = c(t)\kappa_u$ and $\bar{c}(t)\phi = c(t)\kappa_u\phi = c(t)\phi$.

As to the last statement, we recall (see the proof of Cor. 8.1.15) that $c\phi = \mathcal{R}\lambda$ with a linear mapping λ , and \mathcal{R} being the ruled congruence surface with director curve c . A projective automorphism κ of the unit sphere induces, via λ and γ^{-1} , a projective automorphism of the congruence \mathcal{N} . The construction of c is projectively invariant, which implies that the curve defined by the spherical polygon $X_1\kappa, X'_1\kappa, \dots$ equals $c(t)\phi\kappa$. \square

The nice thing about this spherical control structure is that – unlike the ordinary NURBS control points of spherical curves — there are no interdependencies of control points, and a designer of spherical NURBS curves may choose the control polygon arbitrarily.

An example of a spherical NURBS curve (actually a rational Bézier curve) together with its control structure is shown in Fig. 8.13. An analogous procedure is possible for surfaces.

Clearly this control structure is not confined to the sphere, but can be applied to any oval quadric by applying a suitable projective isomorphism. Analogous ‘Hopf mappings’ for other quadratic varieties (ruled and singular) are possible [35]. An application to the construction of rational *blending surfaces* between quadrics has

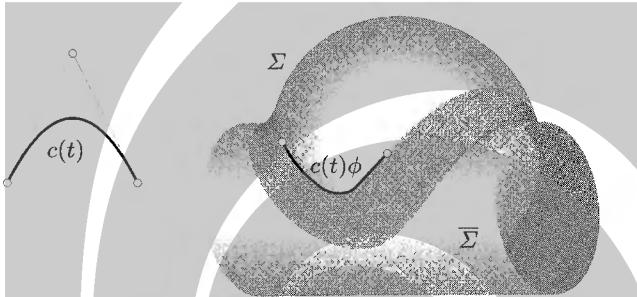


Fig. 8.14. Rational blending surface which joins two quadrics.

been proposed in [202]. The spherical (or cylindrical, or whatever) control polygon of the trim lines can be prescribed arbitrarily and a smooth NURBS tensor product transition surface of lowest possible degree between quadrics is constructed. For G^1 continuity we need degree (4, 3). An example is shown in Fig. 8.14.

8.1.3 Visualization of the Klein Image

The linear line mappings discussed in Sec. 8.1.1 and Sec. 8.1.2 produce perspective images of the Klein quadric and can be used for visualization of the Klein image.

By means of a few examples we illustrate how to use the mapping μ , which has been defined by Equ. (8.23), for visualization. Recall that we can regard the unit sphere as the silhouette of the Klein quadric with respect to μ (see Remark 8.1.11).

Example 8.1.6. In Ex. 5.1.9 we considered the ruled surface \mathcal{R} , parametrized in Plücker coordinates by

$$(\mathbf{r}, \bar{\mathbf{r}})(u) = (1 + u - u^2, -1 + u + u^2, 1, 1 - u^2, u, -1 + u^2 - u^4),$$

and computed its Lie quadric \mathcal{S} at $u_0 = -1/2$. To visualize the second order contact of \mathcal{R} and its Lie quadric \mathcal{S} , we use the linear line mapping μ . The surface $\mathcal{R}\mu$ is, in homogeneous coordinates, parametrized by

$$(-u^2(u-1)(u+1) : (u-1)(u+1) : -u : -2 + u^2 - u^4).$$

We see that $\mathcal{R}\mu$ intersects the plane $x_0 = 0$, but not the plane $x_3 = 0$, so we use the affine coordinates $x_0/x_3, x_1/x_3, x_2/x_3$. Fig. 8.15 shows two-dimensional images

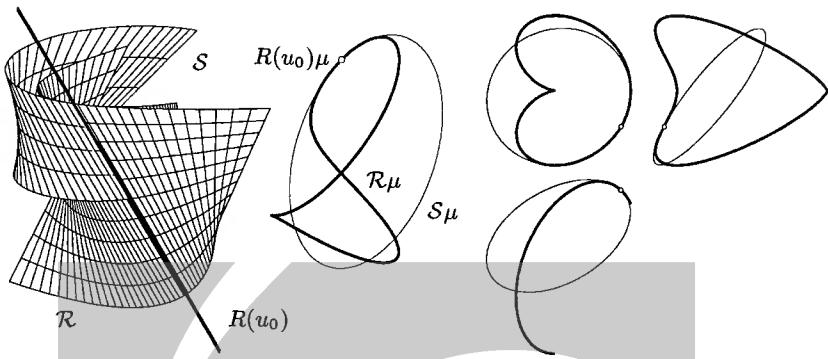


Fig. 8.15. Left: Quartic ruled surface \mathcal{R} and osculating regulus \mathcal{S} for the generator $R(u_0)$. Center: Linear images $\mathcal{R}\mu$, $\mathcal{S}\mu$ and $R(u_0)\mu$. Right: top view, front view, side view of the linear image.

(axonometric view, top view, front view, side view) of the three-dimensional object $\mathcal{R}\mu$. The image of the Lie quadric is a conic which is in second order contact with $\mathcal{R}\mu$ in the point $R(u_0)\mu$. \diamond

Example 8.1.7. As a very simple example of a line congruence, we consider a hyperbolic linear congruence \mathcal{N} with focal lines F_1, F_2 (Fig. 3.11). We know (cf. Th. 3.2.4) that the Klein image $\mathcal{N}\gamma$ is a ruled quadric contained in a three-dimensional projective subspace $G^3 \subset P^5$.

Consider a linear line mapping μ . Depending on the intersection of μ 's center Z with U , the image $\mathcal{N}\mu$ is a ruled quadric (if $\dim(Z \cap U) = 0$) or a subset of a plane (if $\dim(Z \cap U) \geq 1$). The rulings of $\mathcal{N}\mu$ are the images of pencils contained in \mathcal{N} (the sets \mathcal{P}_1 and \mathcal{P}_2 in Fig. 3.11). The image of a regulus \mathcal{R} is a conic.

As $\mathcal{N}\gamma$ is contained in a three-space G^3 , visualization of \mathcal{N} would require only a projective isomorphism of G^3 onto P^3 and a subsequent central projection onto a plane (see Ex. 3.3.1). \diamond

Example 8.1.8. A less trivial example is the normal congruence \mathcal{N} of a cone of revolution, which is parametrized by $\mathbf{c}(u, v) = (cv, v \cos u, v \sin u)$ ($u, v \in \mathbb{R}$, $c \neq 0$). The surface normals $N(u, v)$ and the image points $N(u, v)\mu$ are parametrized by

$$\begin{aligned} N\gamma &= (-1, c \cos u, c \sin u, 0, -v \sin u - vc^2 \sin u, c^2 v \cos u + v \cos u) \mathbb{R}, \\ N\mu &= (v(1 + c^2) \cos u + c \sin u, v(1 + c^2) \sin u + c \cos u, \\ &\quad 1, v(1 + c^2) \cos u - c \sin u) \mathbb{R}. \end{aligned}$$

With the substitution $\cos u = (1 - t^2)/(1 + t^2)$, $\sin u = 2t/(1 + t^2)$, $N\gamma$ is easily recognized as a rational ruled surface of degree four. The parameter lines $v = \text{const.}$ and $u = \text{const.}$ are the images of quadratic cones $R(u)$ and pencils $P(v)$ contained in \mathcal{N} (see Fig. 8.16). \diamond

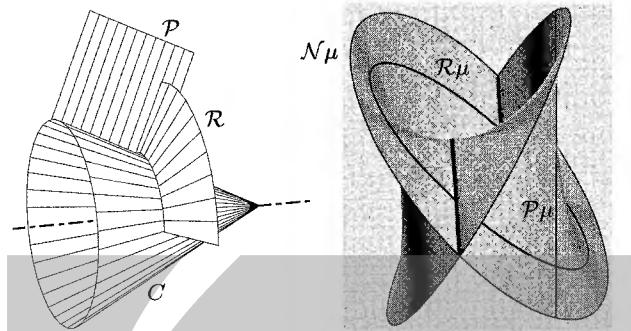


Fig. 8.16. Left: cone C together with a pencil \mathcal{P} and a regulus \mathcal{R} contained in its normal congruence \mathcal{N} . Right: μ -image.

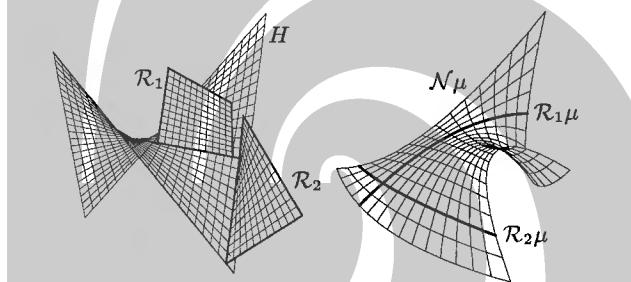


Fig. 8.17. Left: Two hyperbolic paraboloids \mathcal{R}_1 , \mathcal{R}_2 contained in the normal congruence \mathcal{N} of a third hyperbolic paraboloid H . Right: Image with respect to the linear line mapping μ .

Example 8.1.9. Another example is the normal congruence of the hyperbolic paraboloid H parametrized by $(u, v, -uv)$ ($u, v \in \mathbb{R}$). The surface normals $N(u, v)$ and their image points $N(u, v)\mu$ are given by

$$\begin{aligned} N\gamma &= (v, u, 1, v + u^2v, -uv^2 - u, u^2 - v^2)\mathbb{R}, \\ N\mu &= (1 + u^2 - v^2, 2u + uv^2, u^2v, u^2 - v^2 - 1)\mathbb{R}. \end{aligned}$$

If we use affine coordinates $x_1/x_0, x_2/x_0, x_3/x_0$, we get the surface

$$\mathbf{x}(u, v) = \frac{1}{1 + u^2 - v^2} (2u + uv^2, u^2v, u^2 - v^2 - 1).$$

The parameter lines $u = \text{const.}$ and $v = \text{const.}$ of the surface $N(u, v)\mu$ are conics — they are μ -images of the hyperbolic paraboloids which consist of the surface normals in the points of a ruling of H (see Fig. 8.17).

Since the union of surface normals in the points of a non-torsal generator is a hyperbolic paraboloid for all ruled surfaces, linear images of normal congruences always carry at least one one-parameter family of conics. \diamond

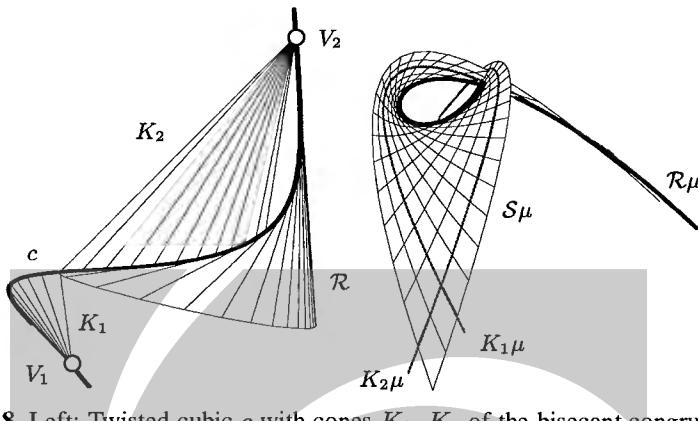


Fig. 8.18. Left: Twisted cubic c with cones K_1 , K_2 of the bisecant congruence \mathcal{S} , and part of c 's tangent developable \mathcal{R} . Right: Linear image.

Example 8.1.10. This example visualizes the bisecant congruence \mathcal{S} of the twisted cubic $c(u) = (1, u, u^2, u^3)\mathbb{R}$ (cf. Prop. 7.2.16 and Ex. 7.2.8). The line $L(u, v) = c(u) \vee c(v)$ ($u \neq v$) and its linear image are given by

$$\begin{aligned} L\gamma &= (1, v + u, v^2 + vu + u^2, u^2v^2, -uv^2 - vu^2, vu)\mathbb{R}, \quad (u \neq v) \\ L\mu &= ((v + u)^2, (v + u)(vu + 1), (vu - 1)(vu + 1), -v^2 - u^2)\mathbb{R}. \end{aligned}$$

The parametrizations above make sense also for the case $u = v$. It is easily checked that $L(u, u)$ is the line tangent to c in $c(u)$. These lines are also contained in the bisecant congruence by definition. The parameter lines of the image surface $\mathcal{S}\mu$ are conics, which are the images of quadratic cones with vertices $c(u)$ or $c(v)$ and director curve c (see the surfaces K_1 and K_2 in Fig. 8.18).

The curve's tangent developable \mathcal{R} is also contained in \mathcal{S} . It is of degree four as an algebraic ruled surface, so $\mathcal{R}\mu$ is a curve of degree four. It is not difficult to verify that the points of $\mathcal{R}\mu$ are singular points of $\mathcal{S}\mu$, and that the congruence \mathcal{S} contains a two-parameter family of reguli, which implies that $\mathcal{S}\mu$ contains a two-parameter family of conics. \diamond

8.2 Kinematic Mappings

Here we continue the discussion of the spherical kinematic mapping, which has been mentioned in Sec. 8.1.1. We show how to use quaternions to describe spherical motions and study again, from the quaternion point of view, the right and left image of lines, Clifford parallelity, and the Hopf mapping. Finally we consider briefly kinematic mappings for the planar and spatial Euclidean motion groups.

8.2.1 Quaternions

We consider the real vector space \mathbb{R}^4 , where addition of vectors and multiplication of real numbers with vectors are defined in the usual way. Coordinates of points will be numbered from 0 to 3, and we write $(a^0, a^1, a^2, a^3) \in \mathbb{R}^4$. We are going to define multiplication of vectors such that \mathbb{R}^4 becomes a non-commutative number field. Then the elements of \mathbb{R}^4 will be called *quaternions*, and \mathbb{R}^4 itself will be denoted by the symbol \mathbb{H} . Note that we write the components of a with upper indices. This is because we have to consider lists of quaternions a_0, a_1, \dots later. We could of course use a vector symbol for the quaternion a , but we don't in order to emphasize that the set \mathbb{H} of quaternions is a number field.

First we identify a real number a^0 with the vector $(a^0, 0, 0, 0)$. So the real number 1 corresponds to the vector $e_0 = (1, 0, 0, 0)$ of the canonical basis. For the other three basis vectors, we use the symbols i , j , and k . An element a of \mathbb{R}^4 is a linear combination of the four basis vectors $1, i, j$, and k :

$$a = (a^0, a^1, a^2, a^3) = a^0 + ia^1 + ja^2 + ka^3.$$

The *product* of quaternions a, b is defined by

$$ab = \begin{bmatrix} a^0 & -a^1 & -a^2 & -a^3 \\ a^1 & a^0 & -a^3 & a^2 \\ a^2 & a^3 & a^0 & -a^1 \\ a^3 & -a^2 & a^1 & a^0 \end{bmatrix} \cdot \begin{bmatrix} b^0 \\ b^1 \\ b^2 \\ b^3 \end{bmatrix} = \begin{bmatrix} a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3 \\ a^1b^0 + a^0b^1 - a^3b^2 + a^2b^3 \\ a^2b^0 + a^3b^1 + a^0b^2 - a^1b^3 \\ a^3b^0 - a^2b^1 + a^1b^2 + a^0b^3 \end{bmatrix}, \quad (8.25)$$

where the product of the matrix with the vector expands to a column vector, which is identified with a quaternion in the obvious way. The following multiplication rules for the basis vectors follow directly from this definition:

$$\begin{aligned} i^2 &= j^2 = k^2 = -1 \\ ij &= -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{aligned} \quad (8.26)$$

If λ is a real number or a quaternion which corresponds to the real number λ , then $\lambda a = a\lambda$.

Obviously quaternion multiplication is not commutative. We will now show that \mathbb{H} with addition and multiplication has all the properties of a number field except commutativity of the multiplication. It is therefore called the *skew field* of quaternions.

Proposition 8.2.1. *Addition and multiplication of quaternions have the following properties: If we assign to a quaternion $a = (a^0, a^1, a^2, a^3)$ the complex 2×2 -matrix*

$$M_a = \begin{bmatrix} a^0 + i_{\mathbb{C}}a^1 & a^2 + i_{\mathbb{C}}a^3 \\ -a^2 + i_{\mathbb{C}}a^3 & a^0 - i_{\mathbb{C}}a^1 \end{bmatrix} \text{ with } i_{\mathbb{C}}^2 = -1, \quad (8.27)$$

then $M_{a+b} = M_a + M_b$ and $M_{ab} = M_a M_b$. Quaternion multiplication is associative, i.e., $(ab)c = a(bc)$, and it is distributive with respect to addition, i.e., $(a+b)c = ac + bc$ and $a(b+c) = ab + ac$.

Proof. The statement about the matrices is verified by direct computation. The second statement follows from the first, because matrix addition and multiplication have the required properties. \square

We further define the quaternion \bar{a} conjugate to a , its norm $N(a)$, and its multiplicative inverse a^{-1} by

$$\begin{aligned}\bar{a} &= (a^0, -a^1, -a^2, -a^3), \quad N(a) = (a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2, \\ a^{-1} &= \frac{1}{N(a)}\bar{a}.\end{aligned}\tag{8.28}$$

Lemma 8.2.2. *There are the following relations:*

$$N(a) = a\bar{a} = \bar{a}a, \quad \bar{ab} = \bar{b}\bar{a}, \tag{8.29}$$

$$a^{-1}a = aa^{-1} = 1. \tag{8.30}$$

Proof. (8.29) is verified directly. (8.30) follows by computing $a^{-1}a = N(a)^{-1}\bar{a}a = N(a)^{-1}N(a) = 1$ and $aa^{-1} = aN(a)^{-1}\bar{a} = N(a)^{-1}a\bar{a} = 1$. \square

The following property of the norm is very important:

Proposition 8.2.3. *The norm is multiplicative, i.e.,*

$$N(a)N(b) = N(ab), \quad N(a^{-1}) = N(a)^{-1}. \tag{8.31}$$

The scalar product $\langle a, b \rangle$ of quaternions as of vectors of \mathbb{R}^4 is given by

$$\langle a, b \rangle = a^0b^0 + a^1b^1 + a^2b^2 + a^3b^3 = \frac{1}{2}(a\bar{b} + b\bar{a}) = \frac{1}{2}(\bar{a}b + \bar{b}a). \tag{8.32}$$

Proof. We compute $N(ab) = ab\bar{a}\bar{b} = a\bar{b}\bar{a}b = aN(b)\bar{a} = N(b)a\bar{a} = N(b)N(a)$. The equation $a^{-1}a = 1$ implies $N(a^{-1})N(a) = N(1) = 1$. The statement about the scalar product is shown as follows: Both expressions in Equ. (8.32) are scalar quaternions, because they equal their respective conjugates, they are bilinear symmetric functions of the vectors a and b , and they equal $N(a)$ for $a = b$. This uniquely characterizes the scalar product corresponding to a norm. \square

Inner Automorphisms

If a is a nonzero quaternion, the mapping

$$\nu_a : x \mapsto a^{-1}xa, \tag{8.33}$$

is called the *inner automorphism* defined by a . We embed \mathbb{R}^3 into \mathbb{R}^4 by letting

$$\mathbf{x} = (x_1, x_2, x_3) \mapsto (0, x_1, x_2, x_3) = ix_1 + jx_2 + kx_3. \tag{8.34}$$

A quaternion which corresponds to a vector of \mathbb{R}^3 in this way is called *vectorial*. A quaternion of the form $(a^0, 0, 0, 0)$, which corresponds to a scalar a^0 , is called a *scalar quaternion*.

Lemma 8.2.4. *The mappings $x \mapsto bx$, $x \mapsto xa$ are orthogonal linear mappings for $N(b) = N(a) = 1$. The inner automorphism ν_a is an orthogonal linear mapping for all nonzero a , and it transforms the subspace of vectorial quaternions onto itself. For all nonzero a, b , we have $\nu_{ab} = \nu_b \circ \nu_a$.*

Proof. The linearity of $x \mapsto bx$ follows directly from Equ. (8.25), which shows the coordinate matrix of this mapping. The argument for $x \mapsto xa$ is similar. Composition of these two mappings for $b = a^{-1}$ gives ν_a , so ν_a is linear.

To show orthogonality of a linear mapping, it is sufficient to show that it is norm-preserving: If $N(a) = N(b) = 1$, then $N(ax) = N(a)N(x) = N(x)$ and $N(xb) = N(x)N(b) = N(x)$. If a is nonzero, the multiplicativity of the quaternion norm implies that $N(a^{-1}xa) = N(a^{-1}) N(x) N(a) = N(a)^{-1} N(x) N(a) = N(x)$.

If $x = x^0$ is a scalar quaternion, then $a^{-1}xa = xa^{-1}a = x$, so ν_a fixes the subspace of scalar quaternions. As ν_a is orthogonal, it leaves invariant also the orthogonal complement of this subspace, which is the subspace of vectorial quaternions. This is also symbolically depicted in Fig. 8.19.

To verify $\nu_{ab} = \nu_b \circ \nu_a$, we compute $\nu_{ab}(x) = (ab)^{-1}x(ab) = b^{-1}a^{-1}xab$. \square

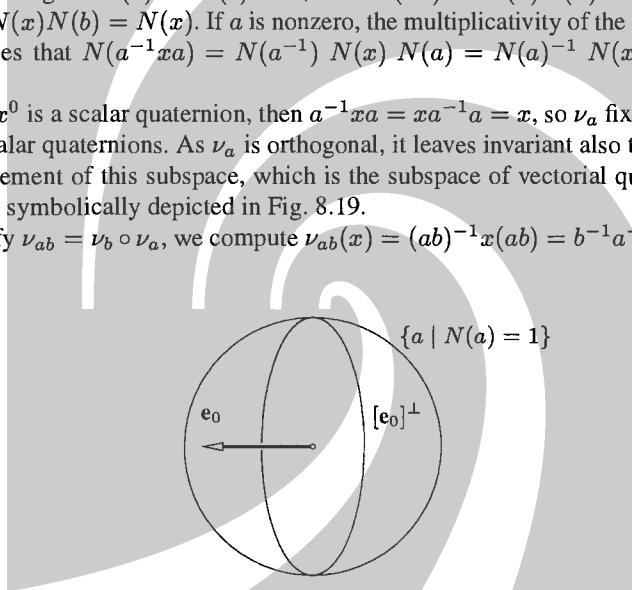


Fig. 8.19. The subspaces of scalar and vectorial quaternions, and the quaternion unit sphere.

Spherical Motions

It turns out that the inner automorphisms ν_a are indeed spherical motions when restricted to the subspace of vectorial quaternions. This will lead to a complete description of the set of spherical displacements, or motions, by quaternions.

Theorem 8.2.5. *The restriction of ν_a to the subspace \mathbb{R}^3 of vectorial quaternions is a spherical motion. The coordinate matrix of this restriction equals*

$$\frac{1}{m_{00}} \begin{bmatrix} m_{11} & 2(a^1a^2 + a^0a^3) & 2(a^1a^3 - a^0a^2) \\ 2(a^1a^2 - a^0a^3) & m_{22} & 2(a^2a^3 + a^0a^1) \\ 2(a^1a^3 + a^0a^2) & 2(a^2a^3 - a^0a^1) & m_{33} \end{bmatrix}, \quad (8.35)$$

$$\begin{aligned}m_{00} &= (a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2, \quad m_{11} = (a^0)^2 + (a^1)^2 - (a^2)^2 - (a^3)^2, \\m_{22} &= (a^0)^2 - (a^1)^2 + (a^2)^2 - (a^3)^2, \quad m_{33} = (a^0)^2 - (a^1)^2 - (a^2)^2 + (a^3)^2.\end{aligned}$$

We have $\nu_a = \nu_b$ if and only if a is a scalar multiple of b .

Proof. By Lemma 8.2.4, ν_a 's restriction to \mathbb{R}^3 is an orthogonal linear mapping. The columns of ν_a 's coordinate matrix are the images

$$\frac{1}{N(a)}\bar{a}ia, \quad \frac{1}{N(a)}\bar{a}ja, \quad \frac{1}{N(a)}\bar{a}ka$$

of the basis vectors i, j, k , so (8.35) follows by direct computation. The determinant of this matrix equals 1, so ν_a acts as a spherical motion.

Clearly ν_a is the identity mapping if a is a scalar quaternion. To show the converse, we observe that the equation $\nu_a(x) = x$ is equivalent to $ax = xa$. Only scalar quaternions a fulfill $ai = ia$, $aj = ja$, $ak = ka$, so the last statement of the theorem is true for the special case $a = 1$.

The general case is reduced to this special case as follows: $\nu_a = \nu_b$ is equivalent to $\text{id} = \nu_b^{-1} \circ \nu_a = \nu_{ab^{-1}}$, and this is equivalent to ab^{-1} being a scalar quaternion, i.e., a and b being scalar multiples of each other. \square

Th. 8.2.5 says that we can assign a spherical motion to a one-dimensional linear subspace $a\mathbb{R}$ of \mathbb{R}^4 , or, equivalently, that we can assign a spherical motion to a point $a\mathbb{R}$ of projective three-space. We will show that this correspondence is onto, and that it actually is the spherical kinematic mapping mentioned in Sec. 8.1.1.

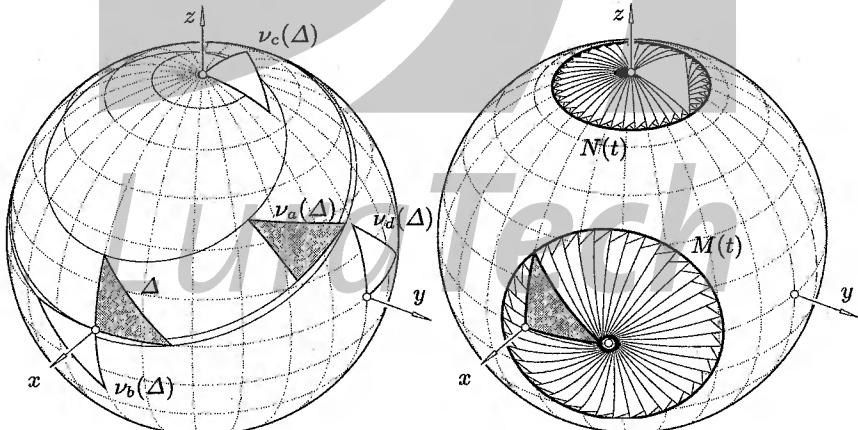


Fig. 8.20. Left: Spherical motions ν_a, ν_b, \dots which correspond to quaternions $a = 3 - i - 2k, b = -i, c = 1 + j, d = 1 - k$. Right: The rotation $M(t)$ about the axis $(2, 1, 0)$ and the coset $N(t)$ of motions which take $(1, 0, 0)$ to $(0, 0, 1)$.

Example 8.2.1. Fig. 8.20 shows some motions ν_a, \dots, ν_d for special quaternions a, \dots, d . We can use Th. 8.2.5 to produce orthogonal matrices whose entries are small rational numbers: The matrix

$$\frac{1}{7} \begin{bmatrix} 3 & -6 & 2 \\ 6 & 2 & -3 \\ 2 & 3 & 6 \end{bmatrix},$$

for instance, is the coordinate matrix of ν_a for $a = 3 - i - 2k$. \diamond

8.2.2 The Spherical Kinematic Mapping

By Th. 3.4.3, a nontrivial one-parameter subgroup $\gamma(t)$ of the special orthogonal group SO_3 of spherical motions is a uniform rotation about an axis. The following theorem describes how to find the quaternions which correspond to this uniform spherical motion.

Theorem 8.2.6. *If $\mathbf{x} = (x_1, x_2, x_3)$ is a unit vector in \mathbb{R}^3 , then consider the great circle of unit quaternions $a(t) = \cos t + (ix_1 + jx_2 + kx_3) \sin t$. Then $\nu_{a(t)}$ is a rotation through an angle of $2t$ about an axis spanned by \mathbf{x} .*

Proof. The coordinate matrix $M(t)$ of $\nu_{a(t)}$ is given by (8.35). It is directly verified that $M(t) \cdot \mathbf{x} = \mathbf{x}$, so $M(t)$ actually is a rotation about the axis \mathbf{x} .

Further, it is elementary to verify that $a(t)a(s) = a(t+s)$, which implies $M(t) \cdot M(s) = M(t+s)$. Thus $M(t)$ is a one-parameter group of rotations. According to Th. 3.4.3, the parameter t is automatically proportional to the rotation angle ϕ , so $\phi = kt$ with some real number k .

To show that actually $k = \pm 2$ it is sufficient to consider one special case: By Th. 8.2.5, $M(t)$ is the identity matrix if and only if $a(t) = \pm a(0)$, i.e., if and only if t is an integer multiple of π . Thus full rotations (the angle is an integer multiple of 2π) occur every time the parameter t reaches an integer multiple of π . This shows that $k = \pm 2$. \square

Corollary 8.2.7. *If a is a nonzero quaternion, then the axis of the spherical motion ν_a is indicated by the ‘vectorial’ part (a^1, a^2, a^3) of a . If ϕ is the angle of ν_a , then $\cos(\phi/2) = |a^0|/\sqrt{N(a)}$.*

Proof. We divide a by the square root of its norm and use Th. 8.2.6. \square

Example 8.2.2. The spherical motion ν_a of Ex. 8.2.1 corresponds to the quaternion $a = 3 - i - 2k$. By Cor. 8.2.7, the axis of this motion is spanned by the vector $(1, 0, 2)$, and the angle ϕ satisfies $\cos(\phi/2) = a^0/\sqrt{N(a)} = 3/\sqrt{14}$.

Fig. 8.20, right, shows the one-parameter group of motions $M(t) = \nu_a(t)$ where $a(t) = \cos t + \sin t(-2i - j)/\sqrt{5}$. This is a uniform rotation about the axis spanned by the unit vector $(2, 1, 0)/\sqrt{5}$. \diamond

Remark 8.2.1. The components of one of the two unit quaternions $\pm a$ which correspond to the spherical motion ν_a are called the *Euler parameters* of this motion. \diamond

The following sums up Th. 8.2.5 and Th. 8.2.6:

Corollary 8.2.8. *The correspondence $\pm a \longleftrightarrow \nu_a$ of unit quaternions a and spherical motions ν_a is two-to-one and onto. The correspondence $a\mathbb{R} \longleftrightarrow \nu_a$ of points $a\mathbb{R}$ of P^3 and spherical motions $\nu_a \in SO_3$ is one-to-one and onto.*

A one-parameter subgroup of spherical rotations with axis \mathbf{x} is the set of motions which fix the vector \mathbf{x} . A slight generalization is the following:

Definition. *The set of spherical motions which map a unit vector \mathbf{x} to another unit vector \mathbf{y} is called the coset determined by \mathbf{x} and \mathbf{y} .*

An example of such a coset is shown in Fig. 8.20, right: The coset of motions which take $(1, 0, 0)$ to $(0, 0, 1)$ is parametrized in the form $N(t)$ — the figure shows the images of a spherical triangle one of whose vertices equals $(1, 0, 0)$.

The Left and Right Image Point of an Oriented Line

We defined the left and right image lines $L\mu^+$ and $L\mu^-$ of a line L in projective three-space by Equ. (8.4) and (8.9). We also defined the left and right image points $\vec{L}\mu^+$ and $\vec{L}\mu^-$ of an oriented line \vec{L} , which are contained in the unit sphere.

A line L , as a two-dimensional subspace of \mathbb{R}^4 , can always be spanned by orthogonal unit quaternions a, b .

Lemma 8.2.9. *If L is spanned by orthogonal unit quaternions a, b , then the Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}}) = a \wedge b$ are normalized, i.e., $\mathbf{l}^2 + \bar{\mathbf{l}}^2 = 1$. The expressions $\bar{ab} = -\bar{ba}$ and $b\bar{a} = -\bar{a}b$ are vectorial unit quaternions, and as such they equal the unit vectors $\mathbf{l}^- = \mathbf{l} - \bar{\mathbf{l}}$ and $\mathbf{l}^+ = \mathbf{l} + \bar{\mathbf{l}}$, respectively.*

Proof. Equ. (8.32) and $\langle a, b \rangle = 0$ show that $\bar{ab} = -\bar{ba}$ and $b\bar{a} = -\bar{a}b$. All four vectors are unit vectors, as a and b are.

We compute $(\mathbf{l}, \bar{\mathbf{l}}) = a \wedge b = (a^0 b^1 - a^1 b^0, \dots, a^1 b^2 - a^2 b^1)$, and $\mathbf{l}^+ = (a^0 b^1 - a^1 b^0 + a^2 b^3 - a^3 b^2, a^0 b^2 - a^2 b^0 + a^3 b^1 - a^1 b^3, a^0 b^3 - a^3 b^0 + a^1 b^2 - a^2 b^1)$. Computing \bar{ab} (cf. also Equ. (8.25)) gives a quaternion whose last three components are equal to $\mathbf{l} - \bar{\mathbf{l}}$, and whose first component equals $\langle a, b \rangle$, i.e., zero. The computation for $\mathbf{l} + \bar{\mathbf{l}}$ and $b\bar{a}$ is analogous.

As $\mathbf{l} + \bar{\mathbf{l}}$ has been established as a unit vector, we can compute $\mathbf{l}^2 + \bar{\mathbf{l}}^2 = \mathbf{l}^2 + 2\mathbf{l} \cdot \bar{\mathbf{l}} + \bar{\mathbf{l}}^2 = (\mathbf{l} + \bar{\mathbf{l}})^2 = 1$. So $(\mathbf{l}, \bar{\mathbf{l}})$ are normalized Plücker coordinates. \square

If a, b is a pair of orthogonal unit quaternions which span a line L , then there are two smooth families of pairs of orthogonal unit quaternions which span this line. These two families are parametrized by

$$(a \cos t + b \sin t, -a \sin t + b \cos t), \\ (a \cos t - b \sin t, -a \sin t - b \cos t).$$

Obviously the pairs a, b and $a, -b$ give rise to two different normalized Plücker coordinates $(\mathbf{l}, \bar{\mathbf{l}}) = a \wedge b$ and $(-\mathbf{l}, -\bar{\mathbf{l}}) = a \wedge (-b)$. Pairs of the same family, however,

must result in the same normalized Plücker coordinate vector, because this vector varies continuously with t , and it cannot jump between its two possible values. Thus a pair of orthogonal unit quaternions a, b actually determines an oriented line. The following is an immediate consequence of Lemma 8.2.9

Corollary 8.2.10. *If \vec{L} is the oriented line determined by the orthogonal unit quaternions a, b , its left and right images $\vec{L}\mu^+$ and $\vec{L}\mu^-$ are equal to the vectorial unit quaternions $\bar{ba} = -ab$ and $\bar{ab} = -\bar{ba}$, respectively.*

The Spherical Kinematic Mapping

We have already mentioned the spherical kinematic mapping. Because we have not yet given an exact definition, we do it here, and show afterwards that it has the properties described in Sec. 8.1.1.

Definition. *We identify \mathbb{R}^4 and \mathbb{H} and use quaternions for homogeneous coordinate vectors of points in P^3 . The mapping $a\mathbb{R} \mapsto \nu_a$ of P^3 into the spherical motion group SO_3 is called the spherical kinematic mapping.*

By Cor. 8.2.8, the spherical kinematic mapping is one-to-one and onto. It has been considered first by C. Stephanos [189] and has been studied in detail by W. Blaschke and H. R. Müller [129].

Theorem 8.2.11. *If an oriented line \vec{L} of P^3 is incident with a point $a\mathbb{R}$, then the spherical motion ν_a takes \vec{L} 's left image point to \vec{L} 's right image point: $\nu_a(\vec{L}\mu^+) = \vec{L}\mu^-$. The same is true for the right and left image lines of non-oriented lines.*

Proof. Without loss of generality we may assume that a is a unit quaternion. \vec{L} is determined by the point $a\mathbb{R}$ and a second point $b\mathbb{R}$ with a unit quaternion b orthogonal to a . With Cor. 8.2.10, it is easy to compute $\nu_a(\vec{L}\mu^+) = \bar{a}(b\bar{a})a = (\bar{ab})(\bar{aa}) = \bar{ab} = \vec{L}\mu^-$.

The statement about the right and left image of non-oriented lines follows immediately from the respective statement for oriented lines. \square

Corollary 8.2.12. *Assume that x, y are points of the unit sphere. The coset of spherical motions which take x to y corresponds, via the spherical kinematic mapping, to the line L with Plücker coordinates $(x + y, x - y)$.*

Proof. These Plücker coordinates actually define an oriented line \vec{L} with $\vec{L}\mu^+ = x$, $\vec{L}\mu^- = y$. Now we can use Th. 8.2.11. \square

Example 8.2.3. Fig. 8.21 shows a coset which takes the point $x = \vec{L}\mu^+$ to the point $y = \vec{L}\mu^-$. The oriented line \vec{L} has the Plücker coordinates $(x + y, x - y)$. \diamond

Th. 8.2.11 shows that the spherical kinematic mapping indeed has the properties promised earlier. To sum up, we have shown the following:

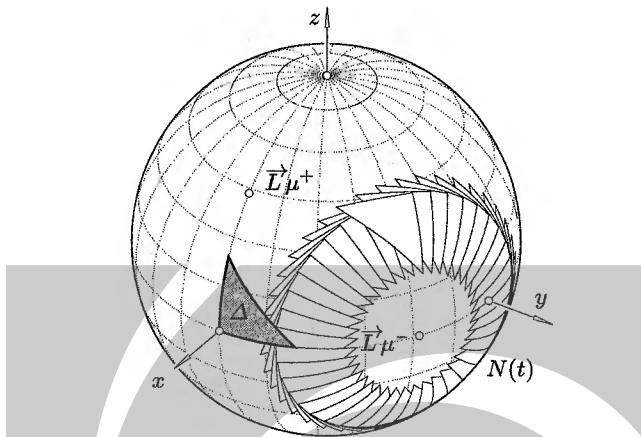


Fig. 8.21. Coset corresponding to an oriented line. All motions $N(t)$ take $\vec{L}\mu^+$ to $\vec{L}\mu^-$.

1. Points of P^3 correspond to spherical motions. The coordinate matrix of this spherical motion is given by Th. 8.2.5.
2. If axis and angle of a spherical motion are given, we can find a quaternion representing this motion by Th. 8.2.6. If a quaternion is given, we can find axis and angle of the corresponding motion by Cor. 8.2.7.
3. A line L of projective space can be endowed with an orientation. The spherical motions which correspond to L 's points map $\vec{L}\mu^+$ to $\vec{L}\mu^-$.
4. Conversely, the points of P^3 which correspond to the spherical motions transforming x to y , comprise a line L , whose Plücker coordinates can be found by Cor. 8.2.12.

The Elliptic Metric in P^3

A spherical motion has a well-defined axis and an *angle*, which is a real number between 0 and π . The angle can be interpreted as the distance of a given spherical motion to the identity transformation, which alone has zero angle. The angle $\sphericalangle(\nu_a, \nu_b)$ enclosed by two spherical motions ν_a and ν_b is defined as the angle of the motion $\nu_b \circ \nu_a^{-1} = \nu_{a^{-1}b}$ (see Fig. 8.22). As both a motion and its inverse have the same angle, $\sphericalangle(\nu_a, \nu_b) = \sphericalangle(\nu_b, \nu_a)$.

The angle enclosed by two one-dimensional subspaces spanned by quaternions $a\mathbb{R}, b\mathbb{R}$ is defined as the angle $\sphericalangle(a, b)$ enclosed by the vectors a and b , if $\sphericalangle(a, b) \leq \pi/2$, and $\sphericalangle(-a, b) = \pi - \sphericalangle(a, b)$ otherwise. Thus the angle of subspaces takes values in the interval $[0, \pi/2]$.

In Ex. 1.1.39 we defined *elliptic three-space* as a projective three-space endowed with an elliptic metric. Especially the metric

$$d(a\mathbb{R}, b\mathbb{R}) = \sphericalangle(a\mathbb{R}, b\mathbb{R}) \quad (8.36)$$

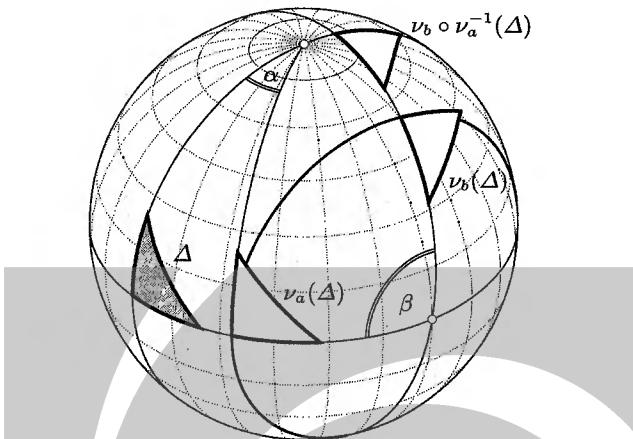


Fig. 8.22. Spherical motions ν_a , ν_b , and the angles $\alpha = \sphericalangle(\nu_a)$, $\beta = \sphericalangle(\nu_a, \nu_b)$.

is elliptic. If we speak of ‘the’ elliptic metric, we will always mean the metric defined by Equ. (8.36), and if we speak of ‘the’ elliptic three-space, it will always be the projective three-space whose points are one-dimensional subspaces $a\mathbb{R}$ of \mathbb{R}^4 , with a nonzero quaternion a .

Theorem 8.2.13. *If ν_a , ν_b are the spherical motions defined by quaternions a and b , then $2\sphericalangle(a\mathbb{R}, b\mathbb{R}) = \sphericalangle(\nu_a, \nu_b)$. The angle of motions ν_a , ν_b , is twice the elliptic distance of the points $a\mathbb{R}$, $b\mathbb{R}$ of elliptic space.*

Proof. It is sufficient to consider unit quaternions. The special case $a = 1$ follows directly from Th. 8.2.6, which describes the spherical kinematic image of a great circle.

By definition, $\sphericalangle(\nu_a, \nu_b) = \sphericalangle(\nu_{ca}, \nu_{cb})$. By Lemma 8.2.4, $\sphericalangle(a, b) = \sphericalangle(ca, cb)$. As $\nu_{ca} = \nu_a \circ \nu_c$ and $\nu_{cb} = \nu_b \circ \nu_c$, we can use $c = a^{-1}$ to convert the general case into the special one. \square

Clifford Parallelity

We define two lines being left (right, resp.) Clifford parallel, if their left (right, resp.) image lines coincide.

By Cor. 8.2.10, the lines of projective three-space are in one-to-one correspondence with the cosets of the spherical motion group.

Proposition 8.2.14. *Consider the coset of spherical motions which take x to y and the line $L_{x,y}$ which corresponds to this coset.*

Then the set of lines $L_{x,y}$, x fixed, is an equivalence class of right Clifford parallels; and the set of lines $L_{x,y}$, y fixed, is an equivalence class of left Clifford parallels.

Proof. This follows from the definition of Clifford parallelity and Th. 8.2.11. \square

Remark 8.2.2. Assume that a, b, c are unit quaternions, and that a, b are orthogonal. Then the pairs a, b and ac, bc define oriented lines \vec{L}, \vec{L}' . Since $\bar{a}\bar{c}bc = \bar{c}(\bar{a}b)c$ and $b\bar{c}\bar{a} = b\bar{c}\bar{a} = b\bar{a}$, by Cor. 8.2.10 we have $\vec{L}'\mu^- = \nu_c(\vec{L}\mu^-)$ and $\vec{L}'\mu^+ = \vec{L}\mu^+$. This shows that L, L' are left Clifford parallel. By varying c , we can achieve that $\nu_c(\vec{L}\mu^-)$ equals any point of the unit sphere, so all left Clifford parallels to \vec{L} can be constructed in this way.

The same is true for the lines $(ca)\mathbb{R} \vee (cb)\mathbb{R}$ and right Clifford parallelity. Left multiplication does not change right image points, and acts on left images via ν_c^{-1} . \diamond

Remark 8.2.3. If $L = a\mathbb{R} \vee b\mathbb{R}$, with orthogonal unit quaternions a, b and K is left Clifford parallel to L , then K is spanned by $ac\mathbb{R}$ and $bc\mathbb{R}$, for some unit quaternion c . As L is alternatively spanned by $(\cos \phi a + \sin \phi b)\mathbb{R}$ and $(-\sin \phi a + \cos \phi b)\mathbb{R}$, for all ϕ , and K likewise by $(\cos \phi ac + \sin \phi bc)\mathbb{R}$ and $(-\sin \phi ac + \cos \phi bc)\mathbb{R}$, the minimum distance of any point of L to K is the same for all points. The same holds true for right Clifford parallel lines. Such lines are called equi-distant (cf. Remark 8.1.2).

The converse is also true: If all points of L have the same elliptic distance to K , then L and K are left or right Clifford parallel. \diamond

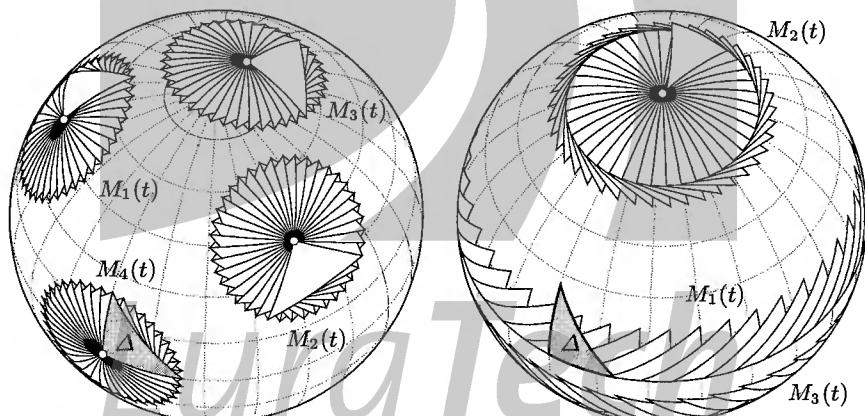


Fig. 8.23. Left: Cosets $M_i(t)$ which correspond to left Clifford parallels. Right: Cosets $M_i(t)$ which belong to right Clifford parallels.

Fig. 8.23, left, illustrates cosets $M_1(t), \dots, M_4(t)$, which correspond to left Clifford parallel lines L_1, \dots, L_4 in P^3 . $M_i(t)$ consists of those spherical motions which take a point x to points y_i — the point x is the same for all cosets.

Fig. 8.23, right illustrates cosets $M_1(t), \dots, M_3(t)$ which correspond to right Clifford parallel lines L_1, \dots, L_3 in P^3 . One coset $M_i(t)$ consists of those spherical motions which take points x_i to a point y — the point y is the same for all cosets.

Remark 8.2.4. The pair of a ‘left’ and a ‘right’ unit sphere serves as a model for the set of oriented lines, if we describe an oriented line \vec{L} by its left and right images.

It turns out that an elliptic congruence transformation, which permutes the set of oriented lines, acts on the left and right image sphere as a spherical motion, and that any pair of spherical motions defines an elliptic congruence transformation. This is interesting also from a group-theoretic viewpoint, because it implies that the elliptic motion group equals the product $\mathrm{SO}_3 \times \mathrm{SO}_3$. Such a decomposition is only possible in dimension three, all other elliptic motion groups are simple. \diamond

Remark 8.2.5. The Hopf mapping ϕ , defined by Equ. 8.20 is related to the spherical kinematic mapping as follows: We consider the point $(x_0, \dots, x_3)\mathbb{R}$ of P^3 and the quaternion $a = (x_0, x_1, x_2, x_3)$. Then $N(a)$ times the last column of ν_a equals $(2(x_1x_3 - x_0x_2), 2(x_2x_3 + x_0x_1), x_0^2 - x_1^2 - x_2^2 + x_3^2)$. Apart from the sign and a permutation of coordinates, this is the Hopf mapping (8.20).

This shows that apart from a coordinate transformation, the Hopf image of a curve $a(t)\mathbb{R}$ in P^3 equals the trajectory of the north pole under the spherical motion $\nu_{a(t)}$. \diamond

Remark 8.2.6. The fibers of the Hopf mapping (8.20) are the lines of the elliptic linear congruence \mathcal{N} defined by the equations $l_{01} + l_{23} = 0$, $l_{02} + l_{31} = 0$. This is clear from the very definition of the Hopf mapping.

On the other hand, Remark 8.2.5 shows that the fibers of the Hopf mapping must be the set of lines whose left image equals the north pole. Indeed, the equation $L\mu^+ = \mathbf{e}_3$ is equivalent to the defining equations of \mathcal{N} . \diamond

Infinitesimal Motions

We consider smooth spherical motions, which are matrix-valued smooth functions $M(t)$ such that $M(t) \in \mathrm{SO}_3$ for all t . The trajectory of a point \mathbf{x}^0 then is $\mathbf{x}(t) = M(t) \cdot \mathbf{x}^0$. We have studied the velocity vector field of a smooth motion in Sec. 3.4.1. By Prop. 3.4.1, it has the form

$$\dot{\mathbf{x}}(t) = \dot{M}(t)\mathbf{x}^0 = \mathbf{c}(t) \times \mathbf{x}(t), \quad (8.37)$$

where $\mathbf{c}(t)$ is the Darboux vector. Equ. (3.25) shows how the Darboux vector and the derivative $\dot{M}(t)$ of $M(t)$ are related to each other. As the velocity distribution is that of a certain uniform rotation about an axis spanned by the Darboux vector, we call the line $R(t)$ spanned by $\mathbf{c}(t)$ the *instantaneous axis* of the spherical motion. The smooth family of axes is called the *fixed axis cone* or *fixed polhode*. The lines $L(t)$ spanned by $M(t)^{-1} \cdot \mathbf{c}(t)$ comprise the *moving axis cone* or *moving polhode*. We can think of $L(t)$ as a cone in the ‘moving’ system, and $R(t)$ as a cone in the ‘fixed’ system. By construction, the motion $M(t)$ takes $L(t)$ to $R(t)$.

The following property of axis cones shows that the motion $M(t)$ is essentially defined by them:

Proposition 8.2.15. *During the motion $M(t)$, the moving axis cone $L(t)$ is rolling on the fixed axis cone $R(t)$ (for t in an interval such that neither $L(t)$ nor $R(t)$ have singular rulings).*

Proof. By construction, $R(t)$ is spanned by $\mathbf{c}(t)$, $L(t)$ is spanned by $\mathbf{d}(t) = M(t)^{-1} \cdot \mathbf{c}(t)$. We have to show that the curves $\mathbf{c}(t)$ and $M(t_0) \cdot \mathbf{d}(t)$ have the same tangent vectors at $t = t_0$, i.e., $M(t_0) \cdot \dot{\mathbf{d}}(t) = \dot{\mathbf{c}}(t)$. By differentiating the defining equation $\mathbf{c} = M \cdot \mathbf{d}$, we get $\dot{\mathbf{c}} = \dot{M} \cdot \mathbf{d} + M \cdot \dot{\mathbf{d}}$. The vector $\dot{M} \cdot \mathbf{d}$ equals, by (8.37), the vector $\mathbf{c} \times (M \cdot \mathbf{d}) = \mathbf{c} \times \mathbf{c} = \mathbf{0}$, so the result follows. \square

The simplest spherical motions are one-parameter groups and cosets, and they also have the simplest axis cones:

Lemma 8.2.16. *If the spherical motion $M(t)$ parametrizes a coset which corresponds, via the spherical kinematic mapping, to a line \vec{L} , then the Darboux vector is a multiple of $\vec{L}\mu^-$, for either orientation of L . The axis cones are degenerate: $L(t) = L\mu^+$, $R(t) = L\mu^-$.*

Proof. The coset consists of those spherical motions which take $\vec{L}\mu^+$ to $\vec{L}\mu^-$ (see also Fig. 8.21). \square

We use this simple lemma to show how the axis cones and the spherical kinematic mapping are related:

Proposition 8.2.17. *Assume that $a(t)\mathbb{R}$ is a smooth curve in P^3 , and that $T(t)$ is its tangent at $t = t_0$. Then the corresponding spherical motion $\nu_{a(t)}$ has the polhodes $R(t) = T(t)\mu^-$ and $L(t) = T(t)\mu^+$.*

Proof. The curve $a'(t)\mathbb{R} = (a(t_0) + (t - t_0)\dot{a}(t_0))\mathbb{R}$ parametrizes the tangent $T(t_0)$, and is in first order contact with the curve $a(t)\mathbb{R}$ at $t = t_0$. Thus the spherical motions $\nu_{a(t)}$ and $\nu_{a'(t)}$ have the same velocity vector fields for $t = t_0$, and also the same instantaneous axes. By Lemma 8.2.16, $L(t) = T(t)\mu^+$ and $R(t) = T(t)\mu^-$. \square

Thus we have shown how the cones $R(t)$ and $L(t)$ arise as linear line images of a tangent surface.

Example 8.2.4. We consider a spherical motion $\nu_{a(t)}$ which is defined by the rolling of a plane on a cone. A particular example of such a motion is the spherical motion of the Sannia frame, which has been discussed in Sec. 5.3.2, and especially Remark 5.3.8.

The tangent surface $T(t)$ of the curve $a(t)\mathbb{R}$ has, according to Prop. 8.2.17, the property that $T(t)\mu^+$ is contained in a plane. This condition is essentially a linear equation for the Plücker coordinates of $T(t)$, so the curve $a(t)\mathbb{R}$ is a complex curve and $T(t)$ is a complex developable for a certain linear line complex. \diamond

8.2.3 Other Kinematic Mappings

We have studied the spherical kinematic mapping in great detail. Of course there are other groups of motions as well, such as the group of planar or spatial Euclidean motions. They independently deserve interest. We will only briefly describe some results, because they are similar to the spherical case. The group of planar Euclidean motions appears as a limit case of the group of spherical motions. The group of spatial Euclidean motions has a kinematic mapping which reflects the composition of motions from a translational and a spherical part.

The Kinematic Mapping of Blaschke and Grünwald

The group of motions of the Euclidean plane is denoted by OA_2 . If we choose a Cartesian coordinate system, then a motion $\alpha \in \text{OA}_2$ has the form

$$\mathbf{x} \mapsto M \cdot \mathbf{x} + \mathbf{b}, \quad \text{with} \quad M = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}. \quad (8.38)$$

In homogeneous coordinates, Equ. (8.38) reads

$$\mathbf{x}\mathbb{R} \mapsto (A \cdot \mathbf{x})\mathbb{R}, \quad A = \begin{bmatrix} 1 & \mathbf{o}^T \\ \mathbf{b} & M \end{bmatrix}. \quad (8.39)$$

The *kinematic mapping of Blaschke and Grünwald* is a correspondence between points of real projective three-space P^3 and planar Euclidean motions. It is defined as follows:

$$d\mathbb{R} \in P^3 \quad \mapsto \quad \begin{bmatrix} d_0^2 + d_3^2 & 0 & 0 \\ 2(d_0d_1 - d_2d_3) & d_3^2 - d_0^2 & 2d_0d_3 \\ 2(d_1d_3 + d_0d_2) & -2d_0d_3 & d_3^2 - d_0^2 \end{bmatrix} \in \text{OA}_2 \quad (8.40)$$

Note that the image of a point with coordinates $(0, d_1, d_2, 0)$ is not a Euclidean motion. We therefore call the line $x_0 = x_3 = 0$ the *absolute line* and consider the kinematic mapping defined in projective space without the absolute line.

There is the following more geometric interpretation of the kinematic mapping of Blaschke and Grünwald:

It is an elementary exercise to verify that a rotation with angle ϕ and center x_m, y_m corresponds to the point $(1, x_m, y_m, -\cot(\phi/2))\mathbb{R}$, and that the translation $\mathbf{x} \mapsto \mathbf{x} + \mathbf{b}$ corresponds to the point $(0, b_2, -b_1, 2)\mathbb{R}$. This is illustrated in Fig. 8.24, which shows an affine part of projective three-space.

Remark 8.2.7. For all real numbers α, β we can define ‘generalized quaternions’ by the following multiplication rules:

$$i^2 = -\alpha, \quad j^2 = -\beta, \quad k^2 = -\alpha\beta \quad (8.41)$$

$$ij = -ji = k, \quad jk = -kj = \beta i, \quad ki = -ik = \alpha j.$$

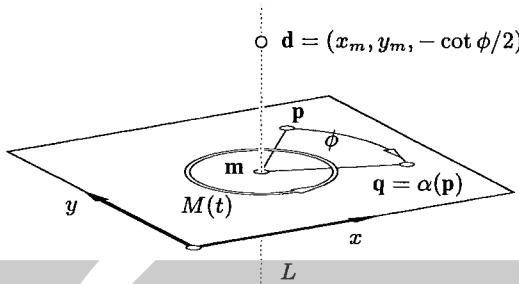


Fig. 8.24. A planar rotation α with center $m = (x_m, y_m)$ transforms p to q and has the kinematic image point d .

For ‘ordinary’ quaternions we let $\alpha = \beta = 1$. If we let $\alpha = 1$ and $\beta = 0$ and repeat the considerations of Sec. 8.2, we see that we get the kinematic mapping of Blaschke and Grünwald. The corresponding geometry is not the elliptic one, but so-called quasi-elliptic geometry. The interested reader is referred to [14].

If we let $\alpha = 1, \beta = -1$, we get a kinematic mapping for the group of motions of the hyperbolic plane. \diamond

Recall that a one-parameter group of Euclidean motions is either the set of rotations about a center, or the set of translations parallel to a fixed vector — we could say the set of motions which fixes a *direction*. Motions which map a point p to another point q , or motions which map a direction to another direction comprise, by definition, a *coset*.

The following theorem sums up properties of the kinematic mapping of Blaschke and Grünwald. We show no detailed proof, because we have dwelt on the spherical kinematic mapping long enough. The planar Euclidean motion group, although simpler than SO_3 in some aspects, is more complicated in others.

Recall the linear line mappings β^+ and β^- and their properties, especially Th. 8.1.7. We call a line horizontal if it meets the absolute line.

Theorem 8.2.18. *The kinematic mapping of Blaschke and Grünwald as defined by Equ. (8.40) has the following properties:*

1. *It is a one-to-one correspondence between the planar Euclidean motion group and projective three-space without the absolute line. The identity transformation corresponds to the point $(0, 0, 0, 1)\mathbb{R}$.*
2. *One-parameter subgroups correspond to ‘vertical’ lines incident with $(0, 0, 0, 1)\mathbb{R}$.*
3. *If a non-horizontal line L is incident with a point $p \in P^3$, then the motion which corresponds to p maps $L\beta^-$ to $L\beta^+$.*
4. *Cosets correspond to lines of P^3 . The motions which correspond to the points of a non-horizontal line L take $L\beta^-$ to $L\beta^+$.*
5. *The family of cosets which map x to y , x fixed, corresponds to the lines of an elliptic linear congruence, and so does the family of cosets for y fixed.*

Proof. (Sketch)

1. Clearly the mapping is homogeneous. It is an elementary exercise to show that it is well defined in both directions, and that the matrix of Equ. (8.40) actually corresponds to a planar Euclidean motion.
2. This follows from the explicit description of one-parameter subgroups.
3. This follows from the geometric description of the mappings β^+ and β^- (cf. Fig. 8.7, Th. 8.1.7) and the geometric description of the kinematic mapping (cf. Fig. 8.24).
4. This follows directly from 3.
5. Both the fibers of $L\beta^+$ and of $L\beta^-$ are elliptic linear congruences (see the discussion following Remark 8.1.8), so 5. follows from 4. \square

For more details, the interested reader is referred to [14].

Motions of Euclidean Three-Space

A motion of Euclidean space, which is the same as a Euclidean congruence transformation, maps points according to

$$\mathbf{x} \mapsto M \cdot \mathbf{c} + \mathbf{b}, \quad (M^T \cdot M = E, \det(M) = 1). \quad (8.42)$$

The translation $\mathbf{x} \mapsto \mathbf{x} + \mathbf{b}$ is called the translational part of this motion, and the spherical motion $\mathbf{x} \mapsto M \cdot \mathbf{x}$ is called its spherical part. Note that the translational part changes if we translate the coordinate system, whereas the spherical part does not. We use a homogeneous Cartesian coordinate system and let $\mathbf{x} = (1, \mathbf{x})$. Equ. (8.42) then becomes

$$\mathbf{x}\mathbb{R} \mapsto (A \cdot \mathbf{x})\mathbb{R}, \quad A = \begin{bmatrix} 1 & \mathbf{o}^T \\ \mathbf{b} & M \end{bmatrix}. \quad (8.43)$$

(cf. Equ. (1.80)). We already know how to represent the spherical part by quaternions — the translational part does not need such an elaborate theory, because there are no restrictions on the components of the vector \mathbf{b} . Thus a vector $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$ and a nonzero quaternion $a = (a^0, \dots, a^3) \in \mathbb{H}$ represent a Euclidean motion:

$$\mathbf{x} \mapsto \nu_a(\mathbf{x}) + \mathbf{b}. \quad (8.44)$$

With respect to a homogeneous Cartesian coordinate system, the coordinate matrix of this transformation is given by

$$\begin{bmatrix} N(a) & \mathbf{o}^T \\ N(a)\mathbf{b} & \widetilde{M} \end{bmatrix}. \quad (8.45)$$

where \widetilde{M} is the matrix of Equ. (8.35) without the denominator $m_{00} = N(a)$. We may regard the correspondence between the pair (a, \mathbf{b}) and the Euclidean motion defined by Equ. (8.45) as a kinematic mapping for the group of Euclidean spatial motions.

Remark 8.2.8. A kinematic mapping for spatial Euclidean motions has been defined and extensively studied by E. Study [191]. It is based on the Study map of oriented lines to \mathbb{D}^2 , where \mathbb{D} is the ring of dual numbers, and which has been discussed in Sec. 2.3. Especially, see Th. 2.3.2 and Remark 2.3.1. \diamond

8.3 Motion Design

Our aim is to describe some methods useful for modeling and control of *smooth motions*, which are paths in a certain motion group. We are going to describe how to design spherical, planar Euclidean, and spatial Euclidean motions. For these three motion groups we have found a ‘kinematic mapping’ which is a correspondence between a projective space (or a subset of projective space) and the motion group. Then, a smooth path in the motion group corresponds, by definition, to a smooth curve in projective space. Thus the design and control of smooth motions is equivalent to the design and control of smooth curves in projective space. A recent survey on the use of kinematic mappings for motion design is O. Röschel’s paper [170].

Rational Motions

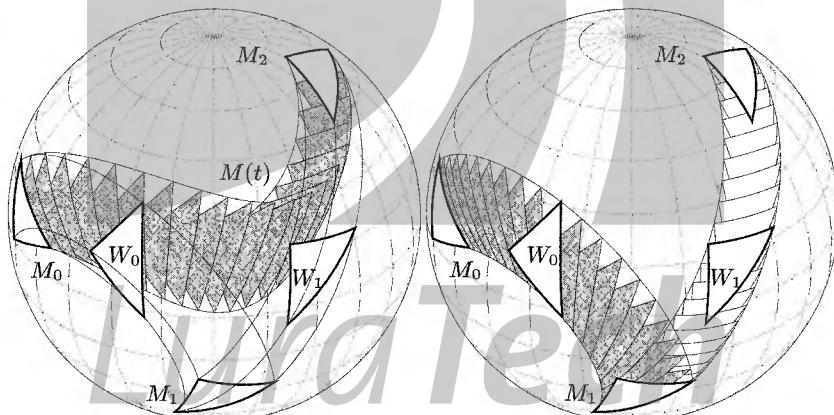


Fig. 8.25. Control structure for a ‘quadratic’ spherical Bézier motion (actually of degree four). The quaternion control vectors are $5 - i + j - 2k$, $5 + i - 4j$, $5 + 5i + 2j + 3k$.

If the group of motions is a group of matrices (such as the spherical, planar, spatial Euclidean groups) it makes sense to define a rational motion $M(t)$ as a motion whose coordinate matrix contains rational functions. The curve in projective space which corresponds, via an appropriate kinematic mapping, to the motion $M(t)$ may

also be rational. It will turn out that these different notions of rationality are actually the same.

We consider rational motions for several reasons. One reason is that low degree rational motions furnish important examples of motions with special properties (e.g., concerning the trajectories of points). Another reason is that the whole Bézier and B-spline machinery is available for modeling and controlling rational curves, which gives hope that we can apply it in some way to rational motions as well.

We are most interested in spherical motions, because we have treated the spherical kinematic mapping in detail, and because they are the main object of study in spatial kinematics. Fortunately, rationality is compatible with the spherical kinematic mapping, so this will enable us to combine Sec. 1.4 with Sec. 8.2 and find geometric control structures for rational and piecewise rational spherical motions.

Rational Spherical Motions

A *rational spherical motion* is a matrix-valued rational function $M(t)$, such that $M(t)$ is an orthogonal matrix of determinant 1 for all t . The existence of the spherical kinematic mapping shows that there is a curve $a(t)\mathbb{R}$, with a quaternion-valued function $a(t)$, such that $\nu_{a(t)} = M(t)$.

The correspondence between motions and quaternions is described by Th. 8.2.5 and Th. 8.2.6. Equ. (8.35) implies that a rational function $a(t)$ defines a rational spherical motion $M(t)$. The converse is not so obvious — Cor. 8.2.7 shows how to find $a(t)$ if we know axis and angle of $M(t)$, but this procedure does not preserve rationality.

There is, however, the following theorem (B. Jüttler, [88]), which shows that for all rational motions $M(t)$ there exists indeed a quaternion-valued rational function $a(t)$ such that $\nu_a(t) = M(t)$. In addition, we can always choose $a(t)$ such that its degree equals half the degree of $M(t)$:

Theorem 8.3.1. Assume that $M(t) = \frac{1}{v(t)} \widetilde{M}(t)$ is a rational spherical motion with polynomial functions $v(t)$, $\widetilde{M}(t)$ of degree m , and the greatest common divisor $\gcd(\widetilde{M}(t), v(t)) = 1$. Then m is even and there exists a quaternion-valued polynomial function $a(t)$ of degree $m/2$ such that $M(t) = \nu_{a(t)}$.

The proof is based on two lemmas.

Lemma 8.3.2. There is a rational curve $a'(t)\mathbb{R}$ such that $\nu_{a'(t)} = M(t)$.

Proof. We denote the l -th column of $M(t)$ by the symbol $\mathbf{m}_l(t)$. As $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$ and $\mathbf{m}_3(t) = \mathbf{m}_1(t) \times \mathbf{m}_2(t)$, it is sufficient to find $a'(t)$ such that $\nu_{a'(t)}(\mathbf{e}_i) = \mathbf{m}_i(t)$ for $i = 1, 2$. The oriented lines $\vec{L}_1(t)$ and $\vec{L}_2(t)$ with Plücker coordinates $(\mathbf{e}_i + \mathbf{m}_i(t), \mathbf{e}_i - \mathbf{m}_i(t))$ ($i = 1, 2$) have the property that $\vec{L}_i(t)\mu^+ = \mathbf{e}_i$ and $\vec{L}_i(t)\mu^- = \mathbf{m}_i(t)$. By Th. 8.2.11, $a'(t)\mathbb{R} = L_1(t) \cap L_2(t)$. We can use Equ. (2.19) to compute this intersection, so $a'(t)\mathbb{R}$ is a rational curve. \square

Lemma 8.3.3. If $a(t) = (a^0(t), \dots, a^3(t))$ is a quaternion-valued polynomial of degree n and $\gcd(a^0(t), \dots, a^3(t)) = 1$, then $\deg \nu_{a(t)} = 2n$.

Proof. $a(t)$ has, by definition, degree n if at least one of $a^i(t)$ has degree n . Obviously then $\deg N(a(t)) = 2n$, so the matrix representation (8.35) has degree $2n$. Suppose there is a common factor of the nine numerators and the denominator $N(t)$. As the polynomials $(a^0(t))^2, \dots, (a^3(t))^2$ are linear combinations of the diagonal entries and the denominator, it would follow that there is a common factor of $(a^0(t))^2, \dots, (a^3(t))^2$, and therefore a common factor of $a^0(t), \dots, a^3(t)$. \square

Proof. (of Th. 8.3.1) Lemma 8.3.2 shows the existence of a rational curve $a'(t)\mathbb{R}$ with $\nu_{a'(t)} = M(t)$. By multiplying with a common denominator and dividing by common factors of $a'^0(t), \dots, a'^3(t)$ we get a polynomial function $a(t)$ with $a(t)\mathbb{R} = a'(t)\mathbb{R}$, and which fulfills the requirements of Lemma 8.3.3. \square

Example 8.3.1. We are going to give a few examples of low degree rational motions. By Th. 8.3.1, there are no *linear* spherical rational motions (indeed, no spherical rational motion has odd degree). This is also clear from the fact that a trajectory of odd order intersects all planes and is therefore not confined to a sphere.

Thus the motions $M(t)$ of lowest degree are quadratic. Th. 8.3.1 shows that there is a *linear* quaternion-valued function $a(t)$ with $\nu_{a(t)} = M(t)$. Then $a(t)$ parametrizes an *oriented line* \vec{L} in P^3 , and Th. 8.2.11 shows that $M(t)$ parametrizes the coset of spherical motions which take $\vec{L}\mu^+$ to $\vec{L}\mu^-$. Examples of such quadratic paths in the motion group are shown in Fig. 8.25, right.

All other rational spherical motions are of degree at least four. The simplest of them, those of degree four, correspond to quadratic curves $a(t)$, i.e., *conics* $a(t)\mathbb{R}$ in projective three-space. An example of such a quartic path in SO_3 is shown in Fig. 8.25, left. \diamond

Example 8.3.2. To show some computations, we parametrize the rotations about the axis $(1, 2, 0)$ by a quadratic rational function: We choose $a(t) = (1-t) + t(-i - 2j)$, which parametrizes the oriented line \vec{L} with $\vec{L}\mu^+ = \vec{L}\mu^- = (-1, -2, 0)$. We get

$$\nu_a(t) = \frac{1}{1 - 2t + 6t^2} \begin{bmatrix} 1 - 2t - 2t^2 & 4t^2 & -4(t-1)t \\ 4t^2 & 1 - 2t + 4t^2 & 2(t-1)t \\ 4(t-1)t & -2(t-1)t & 1 - 2t - 4t^2 \end{bmatrix}. \quad \diamond$$

Rational Motions in Euclidean Plane and Space

A planar or spatial motion is said to be rational, if its matrix representation (8.39) or (8.43) is rational. By multiplying with a common denominator we may, without loss of generality, assume that a rational motion has the form

$$\mathbf{x}\mathbb{R} \mapsto (A(t) \cdot \mathbf{x})\mathbb{R} \text{ with } A(t) = \begin{bmatrix} v(t) & \mathbf{o}^T \\ \tilde{\mathbf{b}}(t) & \tilde{M}(t) \end{bmatrix} \quad (8.46)$$

with a polynomial function $v(t)$, a vector-valued polynomial function $\tilde{\mathbf{b}}(t)$, and a matrix-valued polynomial function $\tilde{M}(t)$. The trajectory of a point $\mathbf{x} \in \mathbb{R}^2$ or $\mathbf{x} \in \mathbb{R}^3$ is parametrized by

$$\mathbf{x} \mapsto \frac{1}{v(t)} \widetilde{M}(t) \cdot \mathbf{x} + \frac{1}{v(t)} \tilde{\mathbf{b}}(t). \quad (8.47)$$

In the planar case, the matrix $M(t) = \widetilde{M}(t)/v(t)$ must be of the form described in Equ. (8.38), and in the spatial case it must fulfill $M^T \cdot M = E$, $\det(M) = 1$.

Proposition 8.3.4. *Rational planar motions correspond, via the kinematic mapping of Blaschke and Grünwald, to rational curves in projective three-space, and vice versa.*

Proof. (Sketch) A rational curve in three-space defines a rational motion. This is clear from Equ. (8.40). To show the converse, we consider the inverse kinematic mapping, which takes a rotation $\mathbf{x} \mapsto A \cdot \mathbf{x} + \mathbf{b}$ to the point $(1, x_m, y_m, -\cot(\phi/2))\mathbb{R}$, where (x_m, y_m) is the center of rotation. We essentially have to show that x_m , y_m and $\cot(\phi/2)$ depend on the coefficients of A and \mathbf{b} in a rational way. This follows from the equation $\cot(\phi/2) = (\cos \phi + 1)/\sin \phi$ and from the fact that the center \mathbf{m} of the motion equals $(E - A)^{-1}\mathbf{b}$. \square

By Th. 8.3.1, for all rational spatial motions of degree k there are a quaternion-valued polynomial $a(t)$ of degree $l \leq k/2$, a real-valued polynomial $w(t)$ of degree $k - 2l$, and a vector-valued polynomial $\tilde{\mathbf{b}}(t)$ of degree k , such that

$$A(t) = \begin{bmatrix} w(t)N(a(t)) & \mathbf{o}^T \\ \tilde{\mathbf{b}}(t) & w(t)\widetilde{M}(t) \end{bmatrix}, \quad \nu_{a(t)} = \frac{1}{N(a(t))} \widetilde{M}(t), \quad (8.48)$$

where $\widetilde{M}(t)$ is the matrix of Equ. (8.35) without the denominator $N(a)$.

Example 8.3.3. The trajectories of points which undergo rational motions are rational, which is clearly seen from Equ. (8.47). Particularly interesting is the case of quadratic spatial motions with a nontrivial spherical part — these are examples of non-planar motions with quadratic and therefore *planar* trajectories.

We can find all such motions from the representation (8.48): We have to choose the polynomials $\tilde{\mathbf{b}}(t)$, $a(t)$, $w(t)$ quadratic, linear, and constant, respectively. Such motions are called *Darboux motions*. For an example of a Darboux motion, see Ex. 6.3.2. \diamond

Remark 8.3.1. If a line parallel to the unit vector $\mathbf{v} = (v_1, v_2)$ undergoes a planar rational motion $A(t)$, it is parallel to $M(t) \cdot \mathbf{v}$, where $M(t)$ is the rotational part of the motion defined by Equ. (8.38). A unit vector orthogonal to this line is the vector $M(t) \cdot (-v_2, v_1)$. Thus the envelope of the lines $L(t)$ is a rational curve with rational normal vector field, i.e., a PH curve (see p. 384).

Conversely, when traversing a PH curve $c(t)$, the motion defined by the unit tangent and unit normal vectors of $c(t)$ is rational (cf. [198]). \diamond

Remark 8.3.2. As an instructive example how we can exploit line geometry in kinematics we consider the problem of curves $c(t)$ in Euclidean space, such that the motion of the Frenet frame is rational. Recall (cf. Ex. 1.2.3) that the Frenet

frame consists of the unit tangent vector \mathbf{e}_1 , the principal normal vector \mathbf{e}_2 , and the binormal vector $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$. This frame is attached to the point $\mathbf{c}(t)$. The spherical part of its motion is quite analogous to that of the Sannia frame, especially the moving polhode is planar (cf. Remark 5.3.8).

We have seen in Ex. 8.2.4 that spherical motions $\nu_{a(t)}$ with planar moving polhode have the property that $a(t)\mathbb{R}$ is a complex curve of a certain linear complex \mathcal{C} . If we disregard the case that $a(t)\mathbb{R}$ parametrizes a line (the motion of the Frenet frame would be planar), the next possible degree is three — conics are planar and the planar complex curves of \mathcal{C} are not conics.

Thus the first interesting example of such a motion comes from a cubic complex curve $a(t)\mathbb{R}$, and is therefore of degree six (cf. [199]). \diamond

Geometric Control Structures for Motions

We first consider the spherical motions, as this is the most important and interesting case. By Th. 8.3.1, there is a one-to-one correspondence between rational curves in P^3 of degree d and rational spherical motions of degree $2d$. So we will be able to use schemes for modeling curves to spherical motions as well.

Recall that a rational curve in P^3 has the form

$$a(t)\mathbb{R} = (\sum B_i^n(t)a_i)\mathbb{R}, \quad (8.49)$$

with control coefficients $a_i \in \mathbb{R}^4$. We identify \mathbb{R}^4 with the quaternion number field. The sequence a_0, a_1, \dots can be reconstructed (almost always, cf. p. 113), up to a common scalar factor, from the points

$$a_0\mathbb{R}, w_0\mathbb{R} = (a_0 + a_1)\mathbb{R}, a_1\mathbb{R}, w_1\mathbb{R} = (a_1 + a_2)\mathbb{R}, \dots, a_n\mathbb{R},$$

which constitute the *geometric control polygon* of the rational curve $a(t)$. The points $a_i\mathbb{R}, w_i\mathbb{R}, a_{i+1}\mathbb{R}$ are collinear.

By Th. 8.2.5, a spherical motion ν_a defines the quaternion a up to a scalar factor. So we can define the *spherical Bézier motion* $M(t) = \nu_{a(t)}$ by its *control motions*

$$M_0 = \nu_{a_0}, W_0 = \nu_{w_0}, M_1 = \nu_{a_1}, W_1 = \nu_{w_1}, \dots, M_n = \nu_{a_n}, \quad (8.50)$$

which are arbitrary, but subject to the condition that M_i, W_i, M_{i+1} belong to the same coset (see e. g., Fig. 8.26, right). Instead of a control motion we could also speak of a control *position* — the motion means the transformation, and the position means the result of the transformation. Mathematically, there is of course no difference, but ‘position’ appeals to the intuition of the reader, and it is preferable in a context where ‘motion’ could also mean a continuous path in a motion group.

Recall that a coset is the set of motions which take a point x to another point y , and a coset corresponds to a straight line L in projective three-space. The points x and y are the left and right images $\vec{L}\mu^+$ and $\vec{L}\mu^-$, if L is oriented in one of the two possible ways.

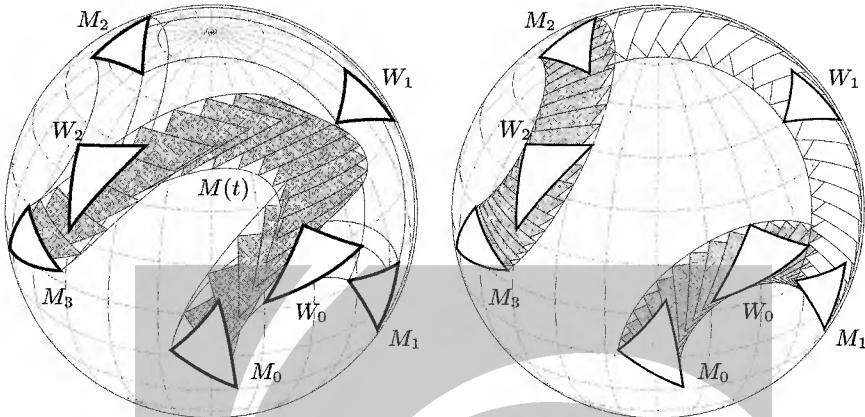


Fig. 8.26. Control structure for a ‘cubic’ spherical Bézier motion (which is actually of degree six). The quaternion Bézier vectors are $5 - 2.5i + 3j + k$, $5 + 5i + j - 3k$, $5 + 2.5i + 2.5j$, $5 + 2.5i - 3j + 4k$.

Example 8.3.4. Fig. 8.25 and Fig. 8.26 show spherical Bézier motions and their geometrical control structures, consisting of positions M_0, W_0, \dots, M_2 and M_0, W_0, \dots, M_3 , respectively. The positions/motions M_i, W_i, M_{i+1} belong to the same coset. These cosets are shown in the right hand part of the figures.

The sequence of control positions yields a sequence of control points in projective three-space. It defines a rational Bézier curve, which in turn defines a rational Bézier motion $M(t)$, shown in the left hand part of the figures. ◇

Remark 8.3.3. All intermediate points which occur in the algorithm of de Casteljau when evaluating quaternion Bézier curves are quaternions, and so they correspond to spherical motions as well, if they happen to be nonzero. Assume that a continuous spherical motion is induced by the curve $a(t)\mathbb{R}$, where $a(t)$ is an ordinary quaternion-valued Bézier curve with control vectors a_0, \dots, a_n as described by Equ. (8.49). Then we choose a real number t and compute points a_{ij} ($i = 0, \dots, n$, $j = 0, \dots, n - i$) by the recursion

$$a_{0j} = a_j, \quad a_{ij} = (1-t)a_{i-1,j-1} + ta_{i-1,j}.$$

By Lemma 1.4.2, $a_{n0} = a(t)$.

Fig. 8.27 shows the spherical Bézier motion $M(t)$ corresponding to such a curve, and the motions $M_{ij} = \nu_{a_{ij}}$ which correspond to the intermediate steps in the algorithm of de Casteljau. The motions W_{ij} corresponding to the weight points $w_{ij} = a_{i,j} + a_{i,j+1}$ are also shown. Note that w_{ij} is contained in the span of $w_{i-1,j}$ and $w_{i-1,j+1}$, so the motions $W_{ij}, W_{i-1,j}$ and $W_{i-1,j+1}$ belong to the same coset. ◇

Remark 8.3.4. There are other ways to control spherical motions. The spherical kinematic mapping does not only provide a one-to-one correspondence between

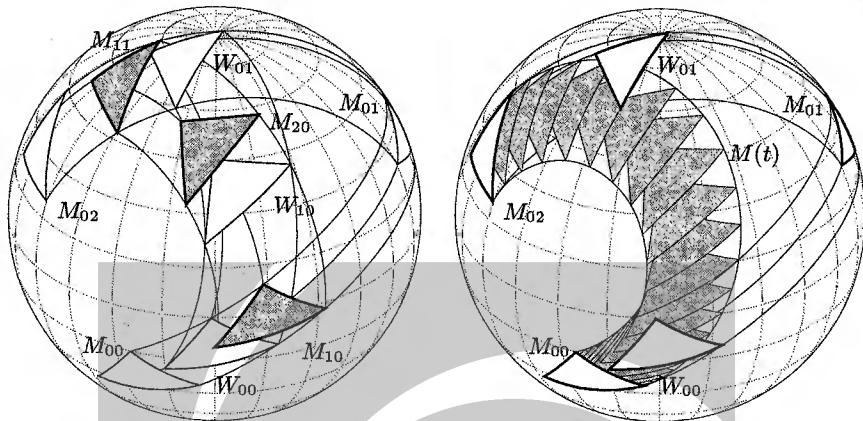


Fig. 8.27. Algorithm of de Casteljau for spherical motions. The quaternion control points are $(5 - 2i + 4j - 2k)\mathbb{R}$, $(5 - i - 1.5j + 3j)\mathbb{R}$, $(5 + 2.5i - 1.5j - 2.5k)\mathbb{R}$ with weights 2, 0.7, and 1, respectively. The algorithm is shown for $t = 0.65$.

SO_3 and projective space, but also a two-to-one correspondence between SO_3 and the quaternion unit sphere. We can use a spherical variant of the algorithm of de Casteljau to define freeform curves in three-space, which do not have many of the properties of polynomial or rational Bézier curves (cf. [184]). \diamond

Spatial Euclidean Motions

As all polynomial curves are Bézier curves, all rational motions of Euclidean three-space are actually *polynomial Bézier motions* of the form

$$A(t) = \sum B_i^n(t) A_i = \sum B_i^n(t) \begin{bmatrix} v_i & \mathbf{o}^T \\ \mathbf{b}_i & M_i \end{bmatrix}. \quad (8.51)$$

The block matrices A_i , which occur as coefficients in Equ. (8.51) are the coordinate matrices of certain *affine* transformations, possibly even of Euclidean motions. The converse of this procedure is not so obvious: If we choose affine transformations A_i and insert them in Equ. (8.51), then the resulting smooth family of affine transformations will in general not be a Euclidean motion.

Example 8.3.5. The same is true not only for rational motions, but also for piecewise rational (NURBS) motions: In this case we have

$$A(t) = \sum N_i^n(t) A_i = \sum N_i^n(t) \begin{bmatrix} v_i & \mathbf{o}^T \\ \mathbf{b}_i & M_i \end{bmatrix}, \quad (8.52)$$

where $N_i^n(t)$ are B-spline basis functions defined over a certain knot vector (cf. Equ. (1.95)).

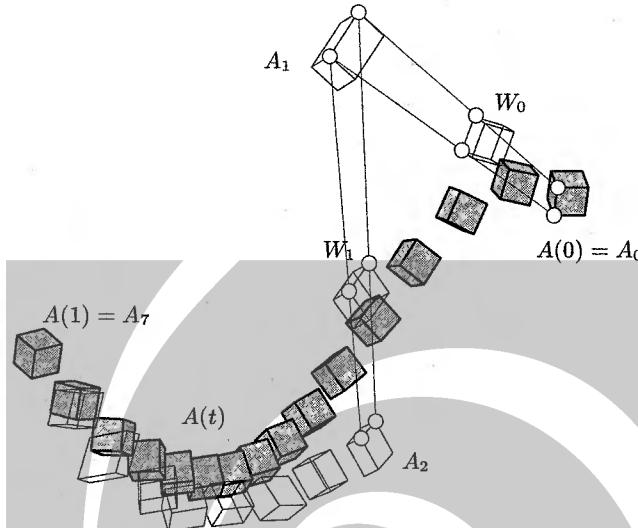


Fig. 8.28. NURBS coefficient matrices $A_0, W_0 = A_0 + A_1, A_1, \dots, A_7$ of a spatial Euclidean NURBS motion $A(t)$ (courtesy B. Jüttler).

An example of a spatial Euclidean NURBS motion $A(t)$ is illustrated in Fig. 8.28, which shows also the affine transformations $A_i, W_i = A_i + A_{i+1}$. It is clearly visible that the transformations A_1, \dots, A_6 are no Euclidean motions. Another example of affine control positions is shown in Fig. C.17. \diamond

Note that the sequence of affine transformations A_i cannot be called a control structure for spatial motions, because we cannot choose them arbitrarily in order to get a Euclidean motion.

In order to actually determine spatial Euclidean motions we use the representation (8.48). By letting $w(t) = 1$ we get the following method for the design of spatial motions, which is not the most general one, but is sufficient for most purposes. We will describe the procedure for rational, i.e., Bézier motions. The modifications for piecewise rational, i.e., NURBS motions, are obvious.

1. First we design the spherical part of the spatial motion — we get a polynomial quaternion curve $a(t)$ such that $\nu_{a(t)} = M(t)$ is a spherical motion. We let $\tilde{M}(t) = N(a(t))M(t)$, so that $\tilde{M}(t)$ is a matrix-valued polynomial of even degree $2n$.
2. We write the polynomial $N(a(t))$ and the polynomial matrix function $\tilde{M}(t)$ in Bézier form:

$$N(a(t)) = \sum B_i^{2n}(t) w_i, \quad \tilde{M}(t) = \sum B_i^{2n}(t) M_i.$$

This can be done by computing the polar forms and applying Th. 1.4.3 (for B-splines, Th. 1.4.15).

3. We let $w(t) = 1$ in Equ. (8.48) and consider the spatial motion

$$A(t) = \sum B_i^n(t) A_i = \sum B_i^n(t) \begin{bmatrix} w_i & \mathbf{o}^T \\ \mathbf{b}_i & M_i \end{bmatrix}, \quad (8.53)$$

with unknown coefficients \mathbf{b}_i .

4. If two affine transformations A_i , A'_i differ only in the vector \mathbf{b}_i , then we get A'_i from A_i by adding a translation. Thus we can control the vectors \mathbf{b}_i (interactively, if we want) by adding appropriate translations to the affine transformations A_i .

Approximation and Interpolation Problems

The problem of interpolating a discrete sequence of positions of a rigid body by a smooth motion can be transformed into an interpolation problem for curves if we apply a kinematic mapping. Interpolation with rational spline motions in the sphere and in Euclidean space has been discussed in [198, 91, 74]. As to the approximation problem, the Bézier and especially the B-spline motions described above are one particular solution. An approximation problem is to find a smooth motion which interpolates given positions within certain tolerances.

A slightly different problem is the following: Assume that $\mathbf{c}(t)$ is a curve in Euclidean space and that $\mathbf{c}_1(t)$ is its unit tangent vector field. Find vector fields $\mathbf{c}_2(t)$ and $\mathbf{c}_3(t)$ such that $\mathbf{c}_1(t), \mathbf{c}_2(t), \mathbf{c}_3(t)$ is an orthonormal basis, and that the angular velocity of the motion defined by this basis is minimal (i.e., this motion is geodesic parallel transport in \mathbf{c} 's normal bundle). For the case that $\mathbf{c}(t)$ is a helix, trajectories of points are shown in Fig. 4.1, left. Rational approximation of these motions has been studied in [92].

LuraTech

www.luratech.com

References

1. Abrams, S. L. et al.: Efficient and reliable methods for rounded-interval arithmetic. *Computer-Aided Design* **30** (1998), 657–665.
2. Ait Haddou, R., Biard, L.: G^2 approximation of an offset curve by Tschirnhausen quartics. In: *Mathematical Methods for Curves and Surfaces* (M. Dæhlen, T. Lyche, L. L. Schumaker, eds). Vanderbilt Univ. Press, Nashville, 1995, pp. 1–10.
3. Ait Haddou, R., Biard, L., Slawinski, M.A.: Minkowski isoperimetric-hodograph curves. *Comp. Aided Geometric Design* **17** (2000), 835–861.
4. Apéry, F.: La surface de Boy. *Adv. Math.* **61** (1986), 185–266.
5. Aumann, G.: Approximate development of skew ruled surfaces. *Computers & Graphics* **13** (1989), 361–366.
6. Aumann, G.: Interpolation with developable Bézier patches. *Comp. Aided Geometric Design* **8** (1991), 409–420.
7. Bajaj, C.L., Xu, G.: Smooth adaptive reconstruction and deformation of free-form surfaces. TICAM Report 99-08, Univ. of Texas, Austin, 1999.
8. Ball, R. S.: Theory of Screws. Cambridge University Press, 1900.
9. Barnhill, R. E.: Representation and approximation of surfaces. In: *Mathematical Software III* (J. R. Rice, ed). Academic Press, New York 1977, pp. 68–119.
10. Barnhill, R. E., Opitz, K., Pottmann, H.: Fat surfaces: a trivariate approach to triangle based interpolation on surfaces. *Comp. Aided Geometric Design* **9** (1992), 365–378.
11. Ben Amar, M., Pomeau, Y.: Crumpled paper. *Proc. Royal Soc. London A* **453** (1997), 729–755.
12. Benz, W.: *Geometrische Transformationen*. BI-Wiss. Verlag, Mannheim 1992.
13. Blaschke, W.: Untersuchungen über die Geometrie der Speere in der Euklidischen Ebene. *Monatshefte für Mathematik und Physik* **21** (1910), 3–60.
14. Blaschke, W., Müller, H. R.: *Ebene Kinematik*. Oldenbourg, München 1956.
15. Bodduluri, R. M. C., Ravani, B.: Design of developable surfaces using duality between plane and point geometries. *Computer-Aided Design* **25** (1993), 621–632.
16. Boersma, J., Molenaar, J.: Geometry of the shoulder of a packaging machine. *SIAM Review* **37** (1995), 406–422.
17. Bol, G.: *Projektive Differentialgeometrie I, II, III*. Vandenhoeck & Ruprecht, Göttingen 1950, 1954, 1967.
18. Bottema, O., Roth, B.: *Theoretical Kinematics*. Dover Publ., New York 1990.
19. Brauner, H.: Die windschiefen Kegelschnittflächen. *Math. Annalen* **183** (1969), 33–44.
20. Brauner, H.: Eine geometrische Kennzeichnung linearer Abbildungen. *Monatsh. Math.* **77** (1973), 10–20.
21. Brauner, H., Kickinger, W.: *Baugeometrie 1*. Bauverlag, 1977.
22. Chaikin, G.: An algorithm for high-speed curve generation. *Comp. Graphics Image Processing* **3** (1974), 346–349.
23. Chalfant, J., Maekawa, T.: Design for manufacturing using B-spline developable surfaces. *J. Ship Research* **42** (1998), 207–215.

24. Chazelle, B., Edelsbrunner, H., Guibas, L. J., Sharir, M., Stolfi, J.: Lines in space: combinatorics and algorithms. *Algorithmica* **15**, 428–447 (1996).
25. Chen, H.-Y., Lee, I.-K., Leopoldseder, S., Pottmann, H., Randrup, T., Wallner, J.: On surface approximation using developable surfaces. *Graphical Models and Image Processing* **61** (1999), 110–124.
26. Chen, H.-Y., Pottmann, H.: Approximation by ruled surfaces. *J. Comput. Applied Math.* **102** (1999), 143–156.
27. Choi, H.-J., Choi, S.-W., Moon, H.-P.: Mathematical theory of medial axis transform. *Pacific J. Math.* **181** (1997), 57–88.
28. Choi, H.-J., Choi, S.-W., Moon, H.-P.: New algorithm for medial axis transform of plane domain. *Graphical Models and Image Processing* **59** (1997), 463–483.
29. Choi, H.-J., Han, C.-Y., Moon, H.-P., Roh, K.-H., Wee, N.-S.: Medial axis transform and offset curves by Minkowski Pythagorean hodograph curves. *Computer-Aided Design* **31** (1999), 59–72.
30. Choi, J.-J., Kim, M.-S., Elber, G.: Computing planar bisector curves based on developable SSI. Preprint, POSTECH, Korea 1997.
31. Coolidge, J. L.: *A Treatise on the Circle and the Sphere*. Clarendon Press, Oxford 1916.
32. Cousinery, B. E.: *Géométrie Perspective*. Paris 1828.
33. de Boor, C.: *A Practical Guide to Splines*. Springer Verlag, New York 1978.
34. Degen, W.: Zur projektiven Differentialgeometrie der Flächen, die von einer einparametrischen Schar von Kegelschnitten erzeugt werden I, II. *Math. Annalen* **155** (1964), 1–34 und **170** (1967), 1–36.
35. Dietz, R.: *Rationale Bézier-Kurven und Bézier-Flächenstücke auf Quadriken*. Diss., TH Darmstadt, 1995.
36. Dietz, R., Hoschek, J., Jüttler, B.: An Algebraic Approach to Curves and Surfaces on the Sphere and on Other Quadrics. *Comp. Aided Geometric Design* **10** (1993), 211–229.
37. Driver, R. D., Downing, J. N., Leskowitz, G. M.: Evanescent-wave spectroscopy down infrared transmitting optical fibers. In: *Infrared Fiber Optics III*. SPIE Vol. 1591, 1991, pp. 168–179.
38. Dyn, N., Levin, D., Gregory, J. A.: A 4-point interpolatory subdivision scheme for curve design. *Comp. Aided Geometric Design* **4** (1987), 257–268.
39. Dyn, N.: A butterfly subdivision scheme for surface interpolation with tension control. *ACM Trans. on Graphics* **9** (1990), 160–169.
40. Edelsbrunner, H.: *Algorithms in Combinatorial Geometry*. Springer, Berlin/Heidelberg/New York, 1987.
41. Edge, W. L.: *The Theory of Ruled Surfaces*. Cambridge University Press, Cambridge, UK 1931.
42. Elber, G.: Rational constraint solver using multivariate spline functions. Preprint, Dept. of Computer Science, Technion, Haifa 2000.
43. Elber, G., Choi, J.-J., Kim, M.-S.: Ruled tracing. *The Visual Computer* **13** (1997), 78–94.
44. Elber, G., Cohen, E.: A unified approach to verification in 5-axis freeform milling environments. *Computer-Aided Design* **31** (1999), 795–804.
45. Elber, G., Kim, M.-S.: A computational model for nonrational bisector surfaces: curve-surface and surface-surface bisectors. In: *Proceedings Geometric Modeling and Processing 2000*, IEEE, Los Alamitos, CA, pp. 364–372.
46. Farin, G.: *Curves and Surfaces for CAGD*. Academic Press, 1993.
47. Farouki, R. T.: Pythagorean-hodograph curves in practical use. In: *Geometry Processing for Design and Manufacturing* (R. E. Barnhill, ed). SIAM, Philadelphia, 1992, pp. 3–33.
48. Farouki, R. T., Goodmann, T. N. T.: On the optimal stability of the Bernstein basis. *Math. Comp.* **216** (1996), 1553–1566.
49. Farouki, R. T., Johnstone, J.: The bisector of a point and a plane parametric curve. *Comp. Aided Geometric Design* **11** (1994), 117–151.

50. Farouki, R. T., Johnstone, J.: Computing point/curve and curve/curve bisectors. In: Mathematics of Surfaces V (R. B. Fisher, ed). Oxford Univ. Press, 1994, pp. 327–354.
51. Farouki, R. T., Neff, C. A.: Hermite interpolation by Pythagorean-hodograph quintics. *Mathematics of Computation* **64** (1995), 1589–1609.
52. Farouki, R. T., Pottmann, H.: Polynomial and rational Pythagorean-hodograph curves reconciled. In: The Mathematics of Surfaces VI (G. Mullineux, ed). Oxford Univ. Press, 1996, pp. 355–378.
53. Farouki, R. T., Sakkalis, T.: Pythagorean hodographs. *IBM J. Res. Develop.* **34** (1990), 736–752.
54. Farouki, R. T., Sakkalis, T.: Pythagorean-hodograph space curves. *Advances in Comp. Math.* **2** (1994), 41–66.
55. Farouki, R. T., Sederberg, T.: Analysis of the offset to a parabola. *Comp. Aided Geometric Design* **12** (1995), 639–645.
56. Farouki, R. T., Shah, S.: Real-time CNC interpolators for Pythagorean-hodograph curves. *Comp. Aided Geometric Design* **13** (1996), 583–600.
57. Fiorot, J. C., Gensane, T.: Characterization of the set of rational parametric curves with rational offsets. In: Curves and Surfaces in Geometric Design (P. J. Laurent, A. Le Méhauté, L. L. Schumaker, eds). A K Peters, Wellesley, MA, 1994, 153–160.
58. Franke, R.: Thin plate splines with tension. In: Surfaces in CAGD '84 (R. E. Barnhill, W. Boehm, eds.). North Holland, 1985, 87–95.
59. Ge, Q. J., Ravani, B.: On representation and interpolation of line segments for computer aided geometric design. *ASME Design Automation Conf.*, Vol. 96/1 (1994), pp. 191–198.
60. Ge, Q. J., Ravani, B.: Geometric Design of rational Bézier line congruences and ruled surfaces using line geometry. *Computing* **13** (1998), 101–120.
61. Giering, O.: Vorlesungen über höhere Geometrie. Vieweg, Braunschweig 1982.
62. Giering, O., Hoschek, J. (eds): Geometrie und ihre Anwendungen. Hanser, München und Wien 1994.
63. Glaeser, G.: Reflections on spheres and cylinders of revolution. *J. Geometry and Graphics* **3** (1999), 121–139.
64. Glaeser, G., Stachel, H.: Open Geometry. Springer-Verlag, New York 1999.
65. Glaeser, G., Wallner, J., Pottmann, H.: Collision-free 3-axis milling and selection of cutting tools. *Computer-Aided Design* **31** (1999), 225–232.
66. Gruber, P., Wills, J. (eds): *Handbook of Convex Geometry*, Vol. 1,2. North Holland, Amsterdam 1993.
67. Hagen, H. et al.: Curve and surface interrogation. In: Focus on Scientific Visualization (H. Hagen, H. Müller, G. Nielson, eds). Springer, Berlin Heidelberg 1993, pp. 243–258.
68. Hartman, P., Wintner, L.: On the fundamental equations of differential geometry. *Amer. J. Math.* **72** (1950), 757–774.
69. Held, M.: On the computational geometry of pocket machining. *Lecture Notes in Computer Science*, Vol. 500, Springer Verlag, Berlin 1991.
70. Heo, H.-S., Kim, M.-S., Elber, G.: The intersection of two ruled surfaces. *Computer-Aided Design* **31** (1999), 33–50.
71. Hinze, C. U.: A contribution to optimal tolerancing in 2-dimensional Computer Aided Design. Dissertation, Kepler University Linz, 1994.
72. Hlavaty, V.: Differential Line Geometry. P. Nordhoff Ltd., Groningen 1953.
73. Hoffmann, C.: Computer vision, descriptive geometry and classical mechanics. In: Computer Graphics and Mathematics (B. Falcidieno et al., eds). Springer Verlag, Berlin 1992, pp. 229–243.
74. Horsch, T., Jüttler, B.: Cartesian spline interpolation for industrial robots. *Computer-Aided Design* **30** (1998), 217–224.
75. Hoschek, J.: Liniengeometrie. Bibliograph. Institut, Zürich 1971.

76. Hoschek, J.: Dual Bézier curves and surfaces. In: Surfaces in Computer Aided Geometric Design (R. E. Barnhill, W. Boehm, eds). North Holland, 1983, 147–156.
77. Hoschek, J.: Detecting regions with undesirable curvature. Comp. Aided Geometric Design 1 (1984), 183–192.
78. Hoschek, J., Lasser, D.: Fundamentals of Computer Aided Geometric Design. A. K. Peters, Wellesley, MA 1993.
79. Hoschek, J., Pottmann, H.: Interpolation and approximation with developable B-spline surfaces. In: Mathematical Methods for Curves and Surfaces (M. Daehlen, T. Lyche, L.L. Schumaker, eds). Vanderbilt University Press, Nashville (TN) 1995, pp. 255–264.
80. Hoschek, J., Schneider, M.: Interpolation and approximation with developable surfaces. In: Curves and Surfaces with Applications in CAGD (A. Le Méhauté, C. Rabut, L. L. Schumaker, eds). Vanderbilt University Press, Nashville (TN) 1997, pp. 185–202.
81. Hoschek, J., Schwanecke, U.: Interpolation and approximation with ruled surfaces. In: The Mathematics of Surfaces VIII (R. Cripps, ed). Information Geometers, 1998, pp. 213–231.
82. Hu, C.-Y., Patrikalakis, N. M., Ye, X.: Robust interval solid modeling: Part I, Representations. Computer-Aided Design 28 (1997), 807–817.
83. Hu, C.-Y., Patrikalakis, N.M., Ye, X.: Robust interval solid modeling: Part II, Boundary evaluation. Computer-Aided Design 28 (1997), 819–830.
84. Hunt, K. H.: Kinematic geometry of mechanisms. Clarendon Press, Oxford 1978.
85. Hunt, K. H.: Special configurations of robot-arms via screw theory. Robotica 4 (1986), 171–179.
86. Hustý, M., Karger, A., Sachs, H., Steinhilper, W.: Kinematik und Robotik. Springer Verlag, Berlin u.a. 1997.
87. Joy, K., Duchaineau, M. A.: Boundary determination for trivariate solids. Proceedings Pacific Graphics '99. Seoul, Korea 1999, 82–91.
88. Jüttler, B.: Über zwangsläufige rationale Bewegungsvorgänge. Sitzungsber. Österr. Akad. Wiss. 202 (1993), 117–132.
89. Jüttler, B., Mäurer, C.: Cubic Pythagorean hodograph spline curves and applications to sweep surface modeling. Computer-Aided Design 31 (1999), 73–83.
90. Jüttler, B., Sampoli, M. L.: Hermite interpolation by piecewise polynomial surfaces with rational offsets. Comp. Aided Geometric Design 17 (2000), 361–385.
91. Jüttler, B., Wagner, M.: Computer Aided Design with spatial rational B-spline motions. ASME J. Mech. Design 118 (1996), 193–201.
92. Jüttler, B., Wagner, M.: Rational motion-based surface generation. Computer-Aided Design 31 (1999), 203–213.
93. Kaul, A., Farouki, R. T.: Computing Minkowski sums of plane curves. Int. J. Computational Geometry & Applications 5 (1995), 413–432.
94. Kautny, W.: Zur Geometrie des harmonischen Umschwungs. Monatsh. Math. 60 (1956), 66–82.
95. Kergosien, Y. L., Gotoda, H., Kunii, T. L.: Bending and creasing virtual paper. IEEE Computer Graphics & Applications 14 (1994), 40–48.
96. Klein, F.: Die allgemeine lineare Transformation der Linienkoordinaten. In: Gesammelte math. Abhandlungen I, Springer Verlag, Berlin 1921.
97. Klein, R.: Algorithmische Geometrie. Addison-Wesley-Longman, Bonn 1997.
98. Kobbelt, L.: Iterative Erzeugung glatter Interpolanten. Dissertation, Karlsruhe 1994.
99. Kós, G., Martin, R. R., Várdy, T.: Methods to recover constant radius rolling ball blends in reverse engineering. Comp. Aided Geometric Design 17 (2000), 127–160.
100. Krasauskas, R., Mäurer, C.: Studying cyclides with Laguerre geometry. Comp. Aided Geometric Design 17 (2000), 101–126.
101. van Kampen, E. R.: The theorems of Gauss-Bonnet and Stokes. Amer. J. Math. 60 (1938), 129–138.

102. Kummer, E. E.: Allgemeine Theorie der geradlinigen Strahlensysteme. *J. Reine Angew. Math.* **57** (1860), 189–230.
103. Landsmann, G., Schicho, J., Winkler, F., Hillgarter, E.: Symbolic parametrization of pipe and canal surfaces. *Proceedings ISSAC-2000*, ACM Press, 2000, 194–200.
104. Lang, J., Röschel, O.: Developable $(1, n)$ Bézier surfaces. *Comp. Aided Geometric Design* **9** (1992), 291–298.
105. Lee, E. T. Y.: Choosing nodes in parametric curve interpolation. *Computer-Aided Design* **21** (1989), 363–370.
106. Lee, I.-K.: Curve reconstruction from unorganized points. *Comp. Aided Geometric Design* **17** (2000), 161–177.
107. Lee, I.-K., Kim, M.-S., Elber, G.: Polynomial/rational approximation of Minkowski sum boundary curves. *Graphical Models and Image Processing* **60** (1998), 136–165.
108. Lee, I.-K., Wallner, J., Pottmann, H.: Scattered data approximation with kinematic surfaces. *Proceedings of SAMPTA'99 Conference*, Løn, Norway 1999, pp. 72–77.
109. Leopoldseder, S.: Cone spline surfaces and spatial arc splines. Dissertation, Vienna Univ. of Technology, 1998.
110. Leopoldseder, S., Pottmann, H.: Approximation of developable surfaces with cone spline surfaces. *Computer-Aided Design* **30** (1998), 571–582.
111. Lin, Q., Rokne, J.: Disk Bézier curves. *Comp. Aided Geometric Design* **15** (1998), 721–737.
112. Lorensen, W. E., Cline, H. E.: Marching cubes: a high resolution 3D surface construction algorithm. *Computer Graphics (SIGGRAPH)* **21** (1987), 163–169.
113. Lü, W.: Offset-rational parametric plane curves. *Comp. Aided Geometric Design* **12** (1995), 601–616.
114. Lü, W.: Rational parameterization of quadrics and their offsets. *Computing* **57** (1996), 135–147.
115. Lübbert, Ch.: Über eine Klasse von linearen Geradenabbildungen. Preprint No. 341, TH Darmstadt, 1977.
116. Lukács, G.: Differential geometry of G^1 variable radius rolling ball blends. *Comp. Aided Geometric Design* **15** (1998), 585–613.
117. Lukács, G., Andor, L.: Photometric ray tracing. In: *The Mathematics of Surfaces VIII* (R. Cripps, ed). Information Geometers, 1998, 325–337.
118. Lukács, G., Marshall, A. D., Martin, R. R.: Faithful least-squares fitting of spheres, cylinders, cones and tori for reliable segmentation. In: *Computer Vision - ECCV '98. Lecture Notes in Computer Science*, Springer, 1998.
119. Maekawa, T.: An overview of offset curves and surfaces. *Computer-Aided Design* **31** (1999), 165–173.
120. Maekawa, T., Chalfant, J.S.: Computation of inflection lines and geodesics on developable surfaces. *Math. Engineering in Industry* **7** (1998), 251–267.
121. Marciniak, K.: *Geometric Modelling for Numerically Controlled Machining*. Oxford University Press, New York 1991.
122. Martin, R. R., Stroud, I. A., Marshall, A. D.: Data reduction for reverse engineering. In: *The Mathematics of Surfaces VII* (T.N.T. Goodman, ed). Information Geometers, 1997, 85–100.
123. Merlet, J.-P.: Singular Configurations of Parallel Manipulators and Grassmann Geometry. *Int. J. Robotics Research* **8** (1992), 45–56.
124. Minding, F. A.: Über die Biegung gewisser Flächen. *J. Reine Angew. Math.* **18** (1838), 297–302.
125. Moon, H.-P.: Minkowski Pythagorean hodographs. *Comp. Aided Geometric Design* **16** (1999), 739–753.
126. Mørken, K.: Some identities for products and degree raising of splines. *Constructive Approximation* **7** (1991), 195–208.

127. Müller, E., Krames, J.: Vorlesungen über Darstellende Geometrie II: Die Zyklographie. Deuticke, Leipzig, Wien 1929.
128. Müller, E., Krames, J.: Vorlesungen über Darstellende Geometrie III: Konstruktive Behandlung der Regelflächen. Deuticke, Leipzig, Wien 1931.
129. Müller, H. R.: Sphärische Kinematik. VEB Deutscher Verlag der Wissenschaften, Berlin 1962.
130. Nash, J.: On C^1 isometric imbeddings. Annals of Math., II. Ser. **60** (1954), 383–396.
131. Nielson, G. M.: A survey of applications of an affine invariant metric. In: Mathematical Methods in Computer Aided Geometric Design (L. L. Schumaker, T. Lyche, eds). Academic Press, 1989, pp. 445–468.
132. Odehnal, B., Pottmann, H.: Computing with discrete models of ruled surfaces and line congruences. To appear in the proceedings of the 2nd workshop on computational kinematics, Seoul, May 20–22, 2001.
133. Odehnal, B., Pottmann, H.: Three-dimensional linear images of line space. Preprint, Vienna Univ. of Technology, 2001.
134. Ohwovorile, M. S., Roth, B.: An Extension of Screw Theory. J. Mech. Des. **103** (1981) 725–735.
135. Peternell, M.: Rational parameterizations for envelopes of quadric families. Dissertation, Vienna Univ. of Technology, 1997.
136. Peternell, M.: Geometric properties of bisector surfaces. Graphical Models and Image Processing **62** (2000), 202–236.
137. Peternell, M., Pottmann, H.: Designing rational surfaces with rational offsets. In: Advanced Topics in Multivariate Approximation (F. Fontanella, K. Jetter and P. J. Laurent, eds). World Scientific, 1996, pp. 275–286.
138. Peternell, M., Pottmann, H.: Computing rational parametrizations of canal surfaces. J. Symbolic Computation **23** (1997), 255–266.
139. Peternell, M., Pottmann, H.: A Laguerre geometric approach to rational offsets. Comp. Aided Geometric Design **15** (1998), 223–249.
140. Peternell, M., Pottmann, H.: Interpolating functions on lines in 3-space. In: Curve and Surface Fitting: Saint Malo 1999 (P. J. Laurent, P. Sablonnière, L. L. Schumaker, eds). Vanderbilt University Press, Nashville (TN) 2000, pp. 351–358.
141. Peternell, M., Pottmann, H., Ravani, B.: On the computational geometry of ruled surfaces. Computer-Aided Design **31** (1999), 17–32.
142. Piegl, L., Tiller, W.: The NURBS book. Springer Verlag, New York 1995.
143. Pottmann, H.: Applications of the dual Bézier representation of rational curves and surfaces. In: Curves and Surfaces in Geometric Design (P. J. Laurent, A. LeMehaute and L. L. Schumaker, eds). A K Peters, Wellesley, MA, 1994, pp. 377–384.
144. Pottmann, H.: Rational curves and surfaces with rational offsets. Comp. Aided Geometric Design **12** (1995), 175–192.
145. Pottmann, H.: Curve design with rational Pythagorean-hodograph curves. Advances in Comp. Math. **3** (1995), 147–170.
146. Pottmann, H.: Studying NURBS curves and surfaces with classical geometry. In: Mathematical Methods for Curves and Surfaces (M. Daehlen, T. Lyche and L. L. Schumaker, eds). Vanderbilt University Press, Nashville (TN) 1995, pp. 413–438.
147. Pottmann, H.: General offset surfaces. Neural, Parallel and Scientific Computations **5** (1997), 55–80.
148. Pottmann, H., DeRose, T.D.: Classification using normal curves. In: Curves and Surfaces in Computer Vision and Graphics II. SPIE Proceedings 1610, Boston 1991, pp. 217–228.
149. Pottmann, H., Farin, G.: Developable rational Bezier and B-spline surfaces. Comp. Aided Geometric Design **12** (1995), 513–531.
150. Pottmann, H., Glaeser, G., Wallner, J., Ravani, B.: Geometric criteria for gouge-free three-axis milling of sculptured surfaces. ASME J. Mech. Design **121** (1999), 241–248.

151. Pottmann, H., Hagen, H., Divivier, A.: Visualizing functions on a surface. *J. Visualization and Computer Animation* **2** (1991), 52–58.
152. Pottmann, H., Lü, W., Ravani, B.: Rational ruled surfaces and their offsets. *Graphical Models and Image Processing* **58** (1996), 544–552.
153. Pottmann, H., Odehnal, B., Peternell, M., Wallner, J., Ait Haddou, R.: On optimal tolerancing in computer-aided design. In: *Geometric Modeling and Processing 2000*. IEEE, Los Alamitos, CA, pp. 347–363.
154. Pottmann, H., Peternell, M.: Applications of Laguerre Geometry in CAGD. *Comp. Aided Geometric Design* **15** (1998), 165–186.
155. Pottmann, H., Peternell, M., Ravani, B.: Approximation in line space: applications in robot kinematics and surface reconstruction. In: *Advances in Robot Kinematics: Analysis and Control* (J. Lenarčič, M. Hustý, eds). Kluwer, 1998, pp. 403–412.
156. Pottmann, H., Peternell, M., Ravani, B.: Contributions to computational line geometry. *Proceedings Int. Conf. Differential/Topological Techniques in Geometric Modeling and Processing*, Pohang, Korea, 1998, pp. 43–81.
157. Pottmann, H., Peternell, M., Ravani, B.: An introduction to line geometry with applications. *Computer-Aided Design* **31** (1999), 3–16.
158. Pottmann, H., Randrup, T.: Rotational and helical surface reconstruction for reverse engineering. *Computing* **60** (1998), 307–322.
159. Pottmann, H., Ravani, B.: Singularities of motions constrained by contacting surfaces. *Mechanism and Machine Theory* **35** (2000), 963–984.
160. Pottmann, H., Wagner, M.: Principal surfaces. In: *The Mathematics of Surfaces VII* (T. Goodman, R. R. Martin, eds). Information Geometers, 1997, pp. 337–362.
161. Pottmann, H., Wallner, J.: Approximation algorithms for developable surfaces. *Comp. Aided Geometric Design* **16** (1999), 539–556.
162. Pratt, M. J.: Cyclides in Computer Aided Geometric Design I,II. *Comp. Aided Geometric Design* **7** (1990), 221–242; and *Comp. Aided Geometric Design* **12** (1995), 131–152.
163. Pratt, V.: Direct least-squares fitting of algebraic surfaces. *Computer Graphics (SIGGRAPH)* **21** (1987), 27–31.
164. Randrup, T.: Approximation by cylinder surfaces. *Computer-Aided Design* **30** (1998), 807–812.
165. Ravani, B., Ku, T. S.: Bertrand offsets of ruled and developable surfaces. *Computer-Aided Design* **23** (1991), 145–151.
166. Ravani, B., Wang, J. W.: Computer aided design of line constructs. *J. Mech. Des.* **113** (1991) 363–371.
167. Requicha, A. A. G.: Towards a theory of geometric tolerancing. *Int. J. of Robotics Research* **2** (1983), 45–60.
168. Rohn, K., Berzolari, L.: Algebraische Raumkurven und abwickelbare Flächen. In: *Enzyklopädie d. Math. Wiss.* III C 9. Teubner, Leipzig, 1926.
169. Romero Fuster, M. C., Sedykh, V. D.: On the number of singularities, zero curvature points and vertices of a simple convex space curve. *J. Geometry* **52** (1995), 168–172.
170. Röschel, O.: Rational motion design — a survey. *Computer-Aided Design* **30** (1998), 169–178.
171. Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. Wiley, New York 1987.
172. Rozenfel'd, B. A.: A history of non-Euclidean geometry. Springer, New York 1988.
173. Sauer, R.: Differenzengeometrie. Springer, Berlin/Heidelberg 1970.
174. Schickentanz, R.: Interpolating G^1 splines with rational offsets. *Comp. Aided Geometric Design* **14** (1997), 881–885.
175. Schickentanz, R.: Interpolation und Approximation durch B-Spline-Flächen mit rationalen Offsets. Diss., TU Darmstadt, 1999.
176. Schumaker, L. L.: Spline Functions: Basic Theory. Wiley, New York 1981.

177. Schweikert, D. G.: An interpolation curve using a spline in tension. *J. Math. and Phys.* **45** (1966), 312–317.
178. Sederberg, T., Farouki, R. T.: Approximation by interval Bézier curves. *IEEE Computer Graphics & Appl.* **12** (1992), 87–95.
179. Sederberg, T. W., Saito, T.: Rational-ruled surfaces: Implicitization and section curves. *Graphical Models and Image Processing* **57** (1995), 334–342.
180. Sedykh, V. D.: Four vertices of a convex space curve. *Bull. London Math. Soc.* **26** (1994), 177–180.
181. Selig, J. M.: Geometrical Methods in Robotics. Springer, New York 1996.
182. Shafarevich, I. R.: Basic Algebraic Geometry I, II. Springer Verlag, Berlin 1977, 1994.
183. Shen, G., Patrikalakis, N.: Numerical and Geometric Properties of Interval B-Splines. *Int. J. Shape Modeling* **4** (1998), 31–62.
184. Shoemake, K.: Animating rotations with quaternion curves. *Computer Graphics (SIGGRAPH)* **19** (1985), 245–254.
185. Sloane, N. J. A.: Packing planes in four dimensions and other mysteries. In: *Algebraic Combinatorics and Related Topics* (E. Bannai, M. Harada, M. Ozeki, eds). Yamagata University, Dept. of Mathematics, 1998.
186. Spivak, M.: A Comprehensive Introduction to Differential Geometry. Publish or Perish, Boston 1975.
187. Sprott, K., Ravani, B.: Ruled surfaces, Lie groups and mesh generation. *Proc. ASME Design Eng. Techn. Conf.* 1997, No. DETC97/DAC-3966.
188. Stachel, H.: Eine Anwendung der kinematischen Abbildung. *Anzeiger Österr. Akad. Wiss., Math.-Naturwiss. Kl.* **5** (1981), 108–111.
189. Stephanos, C.: Sur la théorie des quaternions. *Math. Ann.* **22** (1883), 589–592.
190. Stolfi, J.: Oriented Projective Geometry. Academic Press, San Diego 1991.
191. Study, E.: Von den Bewegungen und Umlegungen (I. u. II. Abhandlung). *Math. Ann.* **34** (1891), 441–566.
192. Study, E.: Geometrie der Dynamen. Leipzig 1903.
193. Taubin, G.: Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Trans. on Pattern Anal. and Mach. Int.* **13** (1991), 1115–1138.
194. Tschirnhaus, W. von: Methodus curvas determinandi, quae formantur a radiis reflectis, quorum incidentes ut paralleli considerantur. *Acta Eruditorum*, Feb. 1690, 68–75.
195. Várády, T., Benkő, P., Kós, G.: Reverse engineering regular objects: simple segmentation and surface fitting procedures, *Int. J. Shape Modeling* **4** (1998), 127–141.
196. Várády, T., Martin, R. R., Cox, J.: Reverse engineering of geometric models — an introduction. *Computer-Aided Design* **29** (1997) 255–268.
197. Waerden, B.L. van der: Einführung in die algebraische Geometrie. Springer Verlag, Berlin u.a. 1973.
198. Wagner, M., Pottmann, H.: Geometric motion design. In: *Modelling and Planning for Sensor Bases Intelligent Robot Systems* (H. Bunke, T. Kanade, H. Noltmeier, eds). World Scientific, 1995, pp. 104–119.
199. Wagner, M., Ravani, B.: Curves with rational Frenet-Serret motion. *Comp. Aided Geometric Design* **15** (1997), 79–101.
200. Wallner, J.: Self-intersections and smoothness of general offset surfaces. *J. Geometry* **70** (2001), 176–190.
201. Wallner, J., Krasauskas, R., Pottmann, H.: Error propagation in geometric constructions. *Computer-Aided Design* **32** (2000), 631–641.
202. Wallner, J., Pottmann, H.: Rational blending surfaces between quadrics. *Comp. Aided Geometric Design* **14** (1997), 407–419.
203. Wallner, J., Pottmann, H.: On the geometry of sculptured surface machining. In: *Curve and Surface Design: Saint Malo 1999* (P. J. Laurent, P. Sablonnière, L. L. Schumaker, eds). Vanderbilt University Press, Nashville (TN) 2000, pp. 417–432.

204. Wang, W., Joe, B.: Interpolation on quadric surfaces with rational quadratic spline curves. *Comp. Aided Geometric Design* **14** (1997), 207–230.
205. Weiss, E.A.: *Einführung in die Liniengeometrie und Kinematik*. Teubner, Leipzig/Berlin 1935.
206. Weiß, G.: Zur Euklidischen Liniengeometrie I. *Sitzungsber. Österr. Akad. Wiss.* **187** (1978), 417–436.
207. Weiß, G.: Zur Euklidischen Liniengeometrie II. *Sitzungsber. Österr. Akad. Wiss.* **188** (1980), 343–359;
208. Weiß, G.: Zur Euklidischen Liniengeometrie III. *Sitzungsber. Österr. Akad. Wiss.* **189** (1980), 19–39.
209. Wunderlich, W.: Über die Böschungslinien auf Flächen 2. Ordnung. *Sitzungsber. Österr. Akad. Wiss.* **155** (1947), 309–331.
210. Wunderlich, W.: Über die Schleppkurven des Kreises, *Sitzungsber. Österr. Akad. Wiss.* **156** (1948), 155–173.
211. Wunderlich, W.: Spiegelung am elliptischen Paraboloid. *Monatsh. Math.* **52** (1948), 13–37.
212. Wunderlich, W.: Pseudogeodätische Linien auf Zylinderflächen. *Sitzungsber. Österr. Akad. Wiss.* **158** (1950), 61–73.
213. Wunderlich, W.: Pseudogeodätische Linien auf Kegelflächen. *Sitzungsber. Österr. Akad. Wiss.* **158** (1950), 75–105.
214. Wunderlich, W.: Beispiele für das Auftreten projektiver Böschungslinien auf Quadriken. *Mat. Tidsskrift* (1951), 9–26.
215. Wunderlich, W.: Zyklische Strahlkomplexe und geodätische Linien auf euklidischen und nichteuklidischen Dreh- und Schraubflächen. *Math. Zeitschrift* **88** (1964), 407–418.
216. Wunderlich, W.: *Darstellende Geometrie II*. Bibliograph. Institut, Mannheim 1967.
217. Wunderlich, W.: Algebraische Böschungslinien dritter und vierter Ordnung. *Sitzungsber. Österr. Akad. Wiss.* **181** (1973), 353–376.
218. Xia, J., Ge, Q. J.: On the exact representation of the boundary of the swept volume of a cylinder undergoing rational Bézier and B-spline motions. *Proceedings ASME Design Eng. Techn. Conf.*, Las Vegas 1999.
219. Xia, J., Ge, Q. J.: On the exact computation of the swept surface of a cylindrical surface undergoing two-parameter rational Bézier motions. *Proceedings ASME Design Automation Conference*, 2000.
220. Yang, M., Lee, E.: NC verification for wire-EDM using an R-map, *Computer-Aided Design* **28** (1996), 733–740.
221. Zindler, K.: *Liniengeometrie mit Anwendungen*. 2 volumes, De Gruyter, Berlin 1902, 1906.
222. Zindler, K.: *Algebraische Liniengeometrie*. In: *Enzyklopädie d. Math. Wiss.* III, 2.2. Teubner, Leipzig, 1926, pp. 973–1228.

List of Symbols

| | | | |
|---------------------------------------|--|--|--------------------------------|
| \bigcap , 9 | A^n , 56 | g_{jk} , 262 | $N(a)$, 524 |
| \bigvee , 9 | \bar{a} , 524 | GL_n , 12 | $N_i^n(t)$, 116 |
| \cap , 4 | a^0, a^1, a^2, a^3 , 523 | $G_{n,k}$, 148 | |
| \trianglelefteq , 87 | $\mathbf{a}_1 \times \dots \times \mathbf{a}_n$, 82 | $\mathrm{grad}_g(\mathbf{x})$, 104 | OA_n , 60, 535 |
| \square , 86 | $A \cdot \mathbf{x}$, 11 | $G(\dot{u}, \dot{v}), g(\dot{u}, \dot{v})$, 435 | O_n , 60 |
| \star , 150 | | | $o(t^k)$, 337 |
| \subset , 9 | | | |
| \times , 82 | | | |
| \vee , 4 | | | |
| \wedge , 133 | | | |
| $\triangleleft_i(U, U')$, 505 | | | |
| $ \alpha $, 86 | | | |
| β^+, β^- , 507 | | | |
| $\Gamma(\mathbf{x})$, 367 | | | |
| γ , 141 | | | |
| γ^* , 169 | | | |
| $\Delta_i, \Delta \mathbf{p}_i$, 300 | | | |
| $\delta(u)$, 265 | | | |
| ε , 154 | | | |
| $\zeta(\mathbf{x})$, 367 | | | |
| κ , 272, 332 | | | |
| κ_1, κ_2 , 262 | | | |
| κ_g, κ_n , 262 | | | |
| $\Lambda^2 \mathbb{R}^4$, 134 | | | |
| $\Lambda^k V$, 145 | | | |
| $\mu(\mathbf{x})$, 411 | | | |
| μ^+, μ^- , 499 | | | |
| μ_2^4 , 169 | | | |
| ν^+ , 502 | | | |
| ν_a , 524 | | | |
| Π_b, Π_f , 143 | | | |
| ρ^+, ρ^- , 502 | | | |
| Σ, Σ^0 , 185 | | | |
| σ , 273 | | | |
| $\sigma : \Phi \rightarrow Q$, 211 | | | |
| τ , 272, 332 | $\mathbf{e}_{\wedge i_1 \dots i_k}$, 145 | M_2^4 , 141 | $\ \mathbf{x}\ $, 60 |
| \varPhi_V , 33 | $\mathbf{e}_i \wedge \mathbf{e}_j$, 134 | $M_2^4(\mathbb{C})$, 143 | \mathbf{x}^0 , 185 |
| Ω , 31, 140 | \mathbf{f}_p , 192 | M^k , 187 | \mathbf{x}^α , 86 |
| Ω_q , 31, 135 | $\widehat{\mathbf{g}}$, 155 | $m(L, C)$, $m(\vec{L}, \vec{C})$, 195 | $\mathbf{x}_{\bullet i}$, 129 |
| ω , 1 | GA_n , 57 | | $\ \mathbf{x}\ _m$, 410 |
| A^T, A^{-T} , 12 | \gcd , 539 | \mathcal{N} , 173 | \mathbf{xR} , 2 |

Index

- A**
accessibility (by cutter), 473
algebra, exterior, 145
algorithm
– of Buchberger, 94
– of Chaikin, 287
– of de Casteljau, 107, 543
– of Dyn, Levin, Gregory, 288
– subdivision-, 110, 287, 289, 466, 467
angle, 60, 505
– dual, 157
– of spherical motion, 530
anticaustic, 379, 393, 446ff
approximation, 195ff,
 291 ff, 316, 318, 352ff
arc length, 76, 261, 390
astroid, 389
asymptote, 35
automorphism
– inner, 524
– of Klein quadric, 144, 170
– of quadric, 32, 42
– projective, *see* mapping, projective
axis
– instantaneous, 186, 397, 533
– medial, 376
– of helical motion, 164
– of linear complex, 161, 167
– of linear congruence, 177
– of projective mapping, 16, 22
– of quadric, 36, 44
- B**
barycenter, 64
basis
– Bernstein, 121, 247, 410
– Gröbner, 88
bending, 277
Bézier, Pierre, 105
Bézout, Etienne, 102
biarc, 292
binormal, 77, 332
bisector, 376, 415
blankholder, 402
- Blaschke, Wilhelm, 507, 535
Bolyai, Janos, 67
Boy, W., 6
Buchberger, Bruno, 94
bundle
– of linear complexes, 181, 199
– of lines, 5, 142
– of planes, 10
butterfly scheme, 467
- C**
camera obscura, 18
carrier, 171
Casteljau, Paul de Faget de, 107, 543
caustic, 379, 393, 447
Cayley, Arthur, 1, 66, 260
center
– of collineation, 16
– of linear mapping, 27, 28, 498
– of perspectivity, 22
– of quadric, 36
Chaikin, George, 287
characteristic, 305
circle, 513
– Minkowski, 411
– oriented, 368
– pseudo-Euclidean, 372
Clairaut, Alexis-Claude, 495
class, 105, 125, 241
– of developable, 342, 346, 398
Clifford, William K., 501, 531
collinear, 4
collineation, *see* mapping, projective
combination, affine/convex, 106, 408
complex, 474ff
– algebraic, 161, 480
– collineation-, 489
– cyclic, 196, 492
– linear, 159ff
– complex, 174
– fitting, 195ff, 215
– linear manifold of, 171
– oriented, 195
- regular, 161
– singular, 161, 166
– of constant slope, 367, 492
– quadratic, 487
– secant, 483
– smooth, 161
– tangent-, 482
– of helical motion, 488
– of quadric, 489
- concurrent, 4
cone, 232, 279
– γ -, 359, 367
– μ -, 411
– asymptotic, 35
– axis-, 533
– complex-, 161, 476, 482, 494
– director-, 263, 359
– of constant slope, 359
– osculating, 281, 282, 332, 361, 400
– quadratic, 33, 46, 350
– rational, 348
– tangent-, 34
confocal, 509
congruence, 423ff
– algebraic, 484
– Bézier, 430
– bisecant-, 485, 522
– discrete, 467
– elliptic, 175, 428
– hyperbolic, 173, 256, 427
– isotropic, 442
– linear, 172 ff, 256, 260, 438, 475
– fitting, 216
– normal-, 324, 442, 446ff, 453, 520
– parabolic, 177, 260, 427
– rational, 428ff
– tangent, 184, 475
congruence transformation,
 see transformation
conic, 15, 31, 53
– absolute, 64
– confocal, 509
– ideal, 35
– normal form, 40
– of dual space, 341

- osculating, 78
- spherical, 511
- conoid, 268
- Plücker-, 180, 258
- right, 268
- Zindler-, 259
- contact
 - line-, 304, 458
 - of curves, 75
 - of developables, 332, 341
 - of ruled surfaces, 227, 234
 - of surfaces, 79, 262, 333, 452
 - oriented, 368, 411
 - second order line-, 458
- contact order, *see* order of contact
- contour, 50
- cycographic, 382
- coordinates
 - affine, 56
 - axis-, 137, 162
 - barycentric, 127, 247, 431
 - Cartesian, 2, 14
 - change of, 139
 - Grassmann, 148
 - dual, 151
 - homogeneous, 2 ff
 - hyperplane-, 8
 - line-, 3
 - of linear complex, 160
 - Plücker, 133 ff, 155
 - projective, 13
- correlation, 29
- coset, 528, 536
- couple of forces, 191
- crease, 416
- cross ratio, 20
- cubic, 90, 101, 386
 - bisecant congruence of, 485, 522
 - inflection points of, 479
 - secant complex of, 484
 - tangent surface, 350
 - Tschirnhaus, 380, 386
- curvature, 77, 332
 - conical, 280, 332, 362
 - Gaussian, 262, 278, 327
 - geodesic, 262, 280, 330, 416
 - normal, 262, 457
 - of ruled surface, 272
 - principal, 262, 443, 444
- curve, 68 ff
 - γ -, 370
 - algebraic, 93, 342
 - asymptotic, 81, 234, 427
 - B-spline, 116 ff, 288, 353
 - Bézier, 105 ff, 410
 - disk, 374
 - dual, 121, 344
 - bisector, 376, 415
 - central, 298
 - characteristic, 305
 - complex-, 477, 534
 - planar, 161, 477, 494
 - convex, 111, 123, 404
 - cuspidal, *see* curve of regression
 - director, *see* directrix
 - dual, 121, 334, 344, 396
 - focal, 444
 - level, 360
 - normal, 113
 - NURBS, 119, 344, 353
 - of constant slope, 359, 370, 385, 390, 440, 492
 - of regression, 232, 335, 336, 346
 - offset, 84, 373
 - PH, 384, 391 ff, 394
 - rational, 105 ff, 112, 388, 390, 516
 - in sphere, 516
 - of constant slope, 388
 - with rational arc length, 390
 - spine, 208, 395
 - spline, 116 ff, 288
 - striction, 267, 270, 315
 - surface-, 262
 - cusp, 73, 104, 339
 - cutter, 452
 - cycle, 368, 411
 - cyclide, Dupin-, 445, 449
 - cyclography, *see* mapping, cycographic
 - cylinder, 203, 279
 - projection-, 360
 - cylindroid, 180, 258
 - D**
 - D-form, 317, 418
 - Darboux, Jean Gaston, 186, 280, 332, 365, 533, 541
 - degree
 - bundle-, 485
 - field-, 485
 - of algebraic complex, 481
 - of algebraic hypersurface, 93
 - of algebraic variety, 100
 - of developable, 346
 - of freedom, 188, 218, 456
 - of parallelity, 58
 - of polynomial, 89
 - of rational curve, 125
 - of rational ruled surface, 243
 - of spline, 116
 - total, 86
 - developable, 327
 - γ , 358, 451
 - μ -, 412, 413
 - algebraic, 342, 346
 - asymptotic, 267
 - Bézier, 343
 - central, 267
 - complex-, 478, 534
 - congruence-, 426
 - connecting, 396
 - convex, 348
 - cubic, 350
 - NURBS, 344, 353
 - of constant slope, 358 ff, 383 ff, 451
 - polynomial, 343
 - quadratic, 350
 - rational, 342, 343
 - self-connecting, 400
 - Taylor expansion, 337
 - with crease, 416
 - development, 329, 360, 416
 - approximate, 286
 - dimension, 6, 99, 481
 - direction
 - in congruence, 424
 - principal, 262, 437, 443
 - torsal, 439
 - directrix, 225, 245, 253, 428
 - distance
 - elliptic, 67, 530
 - Euclidean, 60
 - of circles, 483
 - of line segment and linear complex, 201
 - of line segments, 221
 - of lines, 156, 212
 - in elliptic space, 505

- of planes, 355
- division, polynomial, 93
- duality, 4, 104, 150
- duality (mapping), 29
- Dupin, Charles, 400, 445, 446, 449, 457
- Dyn, Nira, 288

- E**
- Eckhart, Ludwig, 502
- eigenvector, 15, 36, 262, 437
- Einstein, Albert, 261, 434
- elation, 16
- element, surface-, 454
- elimination, 95, 98
- ellipse, 41
 - confocal, 509
- ellipsoid, 41, 59, 366, 382
- embedding
 - of projective plane, 6
 - Segre, 97
- envelope
 - of circles, 370 ff
 - of cylinders, 434
 - of lines, 50, 121, 353, 360, 541
 - of normals, 360
 - of planes, 85, 267, 282, 334 ff, 427
 - of quadratic cones, 307
 - of spheres, 204, 208, 304
 - of surfaces, 429
- epicycloid, 365
- Erlangen program, 11
- Euler, Leonhard, 527
- evolute, 360
 - Minkowski-, 412
- extension, of space
 - complex, 54, 174
 - conformal, 445
 - projective, 1, 56

- F**
- fiber (preimage), 500, 533
- fiber, optical, 496
- field
 - of lines, 143
 - of points, 10
 - skew-, 523
- flag, 72, 454
- flexion, 219
- force, 191
- form
 - bilinear, 31
 - fundamental, 262, 435
 - normal, 38, 40, 163, 258
 - polar, 107, 116, 129
 - quadratic, 31
 - formula
 - of Laguerre, 64
 - of Lamarle, 278
 - frame
 - affine, 56
 - Darboux, 280
 - Frenet, 76, 332, 541
 - projective, 13
 - Sannia, 271, 534
 - framework, 219
 - Frenet, Jean Frédéric, 76, 332, 541
 - function
 - analytic, 154
 - B-spline basis, 116
 - support-, 364, 401

 - G**
 - Gauss, Carl Friedrich, 67, 278
 - generator, *see* ruling
 - geodesic, 263, 325, 479, 490 ff
 - pseudo-, 421
 - geometry
 - absolute, 67
 - affine, 56
 - Cayley-Klein, 66
 - difference, 283, 331
 - elliptic, 67, 501, 530
 - equi-affine, 58
 - equiformal, 64
 - Euclidean, 59
 - hyperbolic, 66
 - Laguerre, 368
 - Möbius, 66, 215, 513
 - Minkowski, 370, 410
 - pseudo-Euclidean, 369
 - gradient, 104, 216, 475
 - Grassmann, Hermann, 133
 - Gröbner, Wolfgang, 88
 - group
 - equi-affine, 58
 - fundamental, 11
 - one-parameter-, 164, 187, 202, 527
 - orthogonal, 60
 - projective linear, 11

- special orthogonal, 60, 527

Grünwald, Josef, 507, 535

H

- harmonic, 24
- helix, 164
 - generalized, 359
- Hermite, Charles, 292
- Hilbert, David, 88, 100
- hodograph, Pythagorean, 384
- homology, 16
 - harmonic, 26
- homothety, 12, 214
- Hopf, Heinz, 514
- horizon, 37
- hull, convex, 106, 400, 418
- hyperbola, 41, 44, 459
 - confocal, 509
- hyperboloid, 41, 177, 278, 307, 366, 389
- hypercycle, 374
- hyperosculating, 459
- hyperplane
 - axis-, 16
 - ideal, 6
 - polar, 31
 - tangent, 34, 47
 - vanishing, 15
- hypersurface, 82, 481
 - algebraic, 93
 - dual, 82 ff
- hypocycloid, 365

I

- ideal, 87 ff, 481
- identity, Plücker, 135
- image, *see also* mapping
 - left, right, 501, 528
 - spherical, 272, 434
- implicitization, 96
- incidence, 3, 8, 139
- independence, projective, 7
- index of quadric, 39, 169
- indicatrix, 410
 - Dupin-, 400, 457
- instant, 186
- interference, 453
- interpolation, 291 ff, 352, 354 ff
 - of scattered data, 63, 217
- intersection

- of algebraic varieties, 99, 481, 487
- of complexes, 475
- of ruled surfaces, 309
- of smooth surfaces, 396
- of subspaces, 8, 138, 153
- planar, of quadric, 48
- interval arithmetic, 405
- invariants, 11, 64
- discretized, 283
- Euclidean, 60, 270 ff, 435, 480
- projective, 69, 425
- inversion, 444
- involute, 273, 364, 386, 391
- involution, 24
 - of conjugate lines, 43
 - of conjugate points, 43
 - of conjugate tangents, 81
 - right angle-, 25
 - skew, 27
- isometry, *see* mapping, isometric
- isomorphism, projective, *see* mapping, projective
- isophote, 381, 472
- Ivory, James, 508

- J**
- joint, 219, 461

- K**
- kernel, 28, 498
- Klein, Felix, 66, 141, 481
- knot, 116
- Kummer, Ernst Eduard, 494

- L**
- Laguerre, Edmund Nicolas, 64, 368
- Lamarle, Ernest, 278
- lemma, of F. Klein, 481
- Lie, Sophus, 236, 314
- line, 1, *see also* ruling
 - γ -, 367
 - μ -, 411
 - absolute, 535
 - Bézier, *see* line, control-
 - complex, 515
 - conjugate, 43
 - control-, 121, 122, 347, 392
 - focal, 173, 177
 - frame, *see* line, control-
 - geodesic, *see* geodesic
 - ideal, 1
 - imaginary, 55
 - isotropic, 62
 - light-like, 370
 - minimal, 62
 - null, 160
 - oriented, 155, 368
 - principal curvature-, 262, 444, 450
 - punch, 402
 - skew, 9
 - space-like, 370
 - time-like, 370
- line complex, *see* complex
- line congruence, *see* congruence
- line element, 5
- Lobachevskij, Nicolai I., 67
- local control, 118
- Lüroth, Jakob, 101, 125, 243

- M**
- Malus, Etienne Louis, 446
- manifold, linear, of complexes, 171
- mapping, *see also* transformation
 - annulator-, 9
 - cyclographic, 366 ff, 447
 - Minkowski, 411
 - Eckhart-Rehbock, 502
 - Hopf, 514, 533
 - isometric, 262, 277, 285, 327, 329, *see also* development
 - kinematic, 158, 522 ff
 - of Blaschke and Grünwald, 507, 535
 - spherical, 529
 - kinematic, spherical, 504
 - Klein, 141, 169
 - linear
 - of lines, 497 ff
 - of projective space, 27, 346, 497
 - local, 211
 - polynomial, 97
 - projective, 11 ff, 253
 - of subspaces, 17
 - perspective, 16
 - singular, 27, 346
 - rational, 97
 - spherical, 434
 - Study, 156
 - matrix
 - orthogonal, 40, 525
 - skew-symmetric, 30
 - symmetric, 31
 - measure, 355
 - mechanism, 219, 285
 - metric
 - elliptic, 67, 530
 - Euclidean, 60, 215, 355
 - intrinsic, 262, 279, 327
 - milling, 304, 319, 452 ff, 463 ff, 473
 - Minding, Ferdinand A., 277
 - Minkowski, Hermann, 374, 407, 410
 - Miquel configuration, 513
 - Möbius, August Ferdinand, 66, 215, 513
 - model
 - bundle-, 5, 504
 - cyclographic, 368
 - Klein, 133 ff
 - of line space, 133, 156, 533
 - of projective plane, 5, 255
 - Study, 156
 - moment, 140, 155, 191
 - of line and complex, 195, 490, 493
 - monomial, 86, 129
 - motion
 - axial, 363
 - Bézier-, 542
 - control-, 542
 - Darboux-, 365, 541
 - dual spherical, 157
 - gliding, 454
 - helical, 164, 186, 187, 364, 488
 - infinitesimal, 186, 188, 219, 397, 453
 - NURBS-, 544
 - k -parameter, 185, 312, 464
 - rational, 538 ff
 - regular, 188
 - spherical, 525 ff, 539
 - spiral, 165, 366
 - surface-surface contact, 452 ff

- uniform, 164, 187, 202, 527
 - motor, 191
 - multiplication
 - quaternion, 523
 - multiplication, exterior, 145
 - multiplicity, intersection-, 99
- N**
- Nash, John, 328
 - norm, 60, 410, 524
 - normal
 - central, 178, 264
 - curve-, 361
 - path-, 164, 185 ff
 - normalization, 69
 - number, dual, 154
 - NURBS, 131, 344, 353, 544
- O**
- offset, 84, 320, 383
 - general, 462
 - Minkowski, 412
 - rational, 451
 - optics, 379 ff, 393, 446 ff
 - order
 - of algebraic variety, *see* degree
 - of contact, 75, 79, 103
 - of spline, 116
 - ordering, monomial, 86
 - orientability, 11
 - orthogonality, 60
- P**
- packing, of planes, 506
 - pair, reciprocal, 162, 193
 - parabola, 41, 50
 - confocal, 509
 - Minkowski-, 415
 - semicubic, 74, 235
 - paraboloid
 - elliptic, 41, 48, 450
 - hyperbolic, 42, 50, 302, 305, 306, 433, 521
 - parallelity, 58
 - Clifford-, 501, 531
 - geodesic, 262
 - parameter
 - distribution-, 178, 265, 306, 480
 - Euler, 527
 - parametrization
 - arc length-, 76
 - Lüroth, 102, 125, 243
 - principal, 262, 328, 438, 444
 - pencil
 - of complexes, 172 ff
 - of directrices, 245
 - of linear complexes, 199
 - of lines, 5, 141
 - of planes, 10
 - of quadratic complexes, 487
 - principal curvature-, 454
 - pentahedral representation, 431
 - perspectivity, 18, 21, 22
 - pitch, 164, 167, 178, 180, 197, 198, 266
 - plane
 - γ -, 368
 - asymptotic, 178, 264
 - central, 178, 264
 - control, 344
 - director-, 49
 - equi-angular, 505
 - focal, 426
 - frame, 344
 - hyperbolic, 66
 - null, 160
 - osculating, *see* subspace, osculating
 - osculation, 237
 - principal, 438
 - projective, 1 ff, 255
 - support, 401
 - torsal, 230, 426
 - plate, thin, 296, 466
 - platform, 220
 - Plücker, Julius, 133, 180, 258
 - Poincaré, Louis, 192
 - point, 1
 - absolute, 62
 - at infinity, 1
 - base, 127
 - central, 26, 178, 264, 437
 - complex, 54
 - conjugate, 29, 43
 - control, 105, 113, 119, 128
 - dual, 344, 392
 - cuspidal, 230, 426
 - elliptic, 80, 262
 - fixed, 15
 - of projectivity, 23
 - flat, 73, 80, 262
 - focal, 426, 428, 440
 - frame, 113
 - fundamental, 13
 - handle, 73, 403
 - hyperbolic, 80, 262
 - ideal, 1
 - inflection, 73, 479
 - Laguerre, 25
 - null, 160
 - of main type, 73
 - orthogonal, 67
 - osculation, 237
 - parabolic, 80, 262
 - real, 54
 - regular, 70, 73, 104, 475
 - singular, 70, 104, 226
 - striction, 264
 - unit, 13
 - polar form, *see* form, polar
 - polarity, 28 ff
 - absolute, 61
 - elliptic, 31
 - hyperbolic, 31
 - induced, 32
 - induced in subspace, 42
 - null, 30, 159 ff
 - control-, 248
 - pole, 31
 - polhode, 273, 533, 541
 - polygon
 - control, 105, 119
 - dual, 121
 - geometric, 113, 542
 - reflection, 322, 323, 491
 - polynomial, 86
 - Bernstein, 105, 127, 247
 - division, 93
 - Hilbert, 100
 - ring, 86
 - position, control-, 326, 542
 - product
 - exterior, 133
 - of projective spaces, 97
 - scalar, *see* norm, distance
 - vector-, 82, 153, 336
 - projection
 - central, 1, 18, 27, 50
 - cyclographic, 367
 - of algebraic varieties, 95
 - of lines, 498
 - parallel, 51

- stereographic, 65, 131, 211 ff, 378, 442, 498
- generalized, 511
- of complex, 214
- via congruence, 469 ff, 512
- projectivity, 20 ff, 22
 - elliptic/hyperbolic/parabolic, 23
 - contact-, 229
 - fundamental, 178
- property, variation
 - diminishing, 111, 115, 120, 122

Q

- quadric, 28 ff, 489
 - absolute, 61
 - base of quadratic cone, 33
 - dual, 45
 - in Euclidean space, 35
 - Klein, 141 ff
 - Lie, 236, 333, 519
 - Lie-, 314
 - osculating, 80
 - oval, 41
 - ruled, 41
 - sections, 48
 - singular, 45
- quadrilateral, 18
- quaternion, 523 ff
 - generalized, 535

R

- range of points, 5
- rank, of curve, 342, 398
- ratio, 21, 58
- reflection
 - affine, 26
 - in cone, 490
 - in curve, 379
 - in helical surfaces, 494
 - in surface, 321, 446 ff, 491
 - projective, 26, 32, 170
- refraction, 380, 447
- regulus, 42, 182, 315
 - complementary, 183
 - fitting, 217
 - right, left, 501
- Rehbock, Fritz, 502
- relation, algebraic, 98, 239
- relativity, 370

- remainder, 93
 - representation, dual, 124, 335, 401
 - rest position, 461
 - resultant, 102
 - reverse engineering, 202, 313
 - robot, 218
 - rolling, 273, 397, 461, 533
 - rolling ball, 209, 470
 - rotation
 - infinitesimal, 186, 397
 - uniform, 164, 187
 - ruled surface, *see* surface, ruled
 - ruling
 - cylindrical, 266, 435
 - elliptic, 425
 - hyperbolic, 425
 - inflection-, 280, 339
 - isotropic, 437
 - non-torsal, 228, 236
 - parabolic, 425
 - regular, 423
 - singular, 227, 235, 340
 - torsal, 228, 230, 237, 254, 266, 425
-
- S**
 - Sannia, Gustavo, 271, 275, 534
 - Sauer, Robert, 495
 - scale
 - metric, 21
 - projective, 15, 20
 - scattered data, 63, 202, 217
 - screw, 192
 - segment, 201
 - projective, 114
 - Segre, Corrado, 97
 - self-dual, 5
 - set
 - fundamental, 13
 - singular, 171
 - silhouette, 50, 241, 474, 516
 - cyclographic, 381
 - similarity, 16
 - simplex, 128
 - auto-polar, 37
 - singularity, 70, 226, 227, 235
 - degree of, 220
 - of developable, 339 ff
 - of framework, 219
 - of motion, 452 ff, 456, 460
 - slope, 353
 - snapping, 220
 - Snell, Willebrord v. R., 380, 447
 - space, *see also* subspace
 - elliptic, 67, 530
 - fixed, moving, 185
 - projective, 6 ff
 - complex, 7, 54
 - dual, 9
 - product of, 97
 - tangent, 79, 104, 475
 - vertex-, 33
 - span, 8, 138, 153
 - spear, 155
 - sphere, 59, 490, 503, 514
 - Study, 154 ff
 - spiral, 165, 366
 - spline
 - B-, 116 ff, 288
 - circular, 378
 - motion-, 544
 - ruled surface-, 292
 - thin plate, 296, 466
 - w. rat. offsets, 393
 - stability, 219 ff, 510
 - stationary, 186
 - Steiner, Jacob, 256, 311
 - Stewart platform, 220
 - striction, 273
 - striction line, *see* curve, striction-
 - structure, mechanic, 219
 - Study, Eduard, 156
 - subdivision, 110, 287, 289, 466, 467
 - subspace, *see also* space
 - exceptional, 28, 498
 - osculating, 72 ff, 112, 117, 519
 - parallel, 58
 - projective, 7
 - complementary, 9
 - tangent to quadric, 45
 - sum, Minkowski-, 374, 407
 - surface
 - algebraic, 93
 - Bézier, 126 ff
 - rational, 131
 - ruled, 251

- tensor product, 251, 343, 349
 - triangular, 128
 - binder, 402
 - bisector-, 447
 - blending-, 209, 470
 - Boy's, 6
 - canal, 395, 444, 471
 - Cayley-, 260
 - central, 436
 - complex-, 248, 476
 - cone spline, 358
 - congruence-, 248, 424, 435
 - cubic, 252 ff, 253
 - developable, *see* developable
 - director-, *see* directrix
 - dual, 82 ff, 396
 - focal, 427, 447
 - helical, 203, 479, 494
 - hyper-, *see* hypersurface
 - invariant, 203, 204
 - kinematic, 202 ff
 - Kummer, 494
 - NURBS, 131, 344, 353
 - of revolution, 494
 - offset, 303, 304, 383, 451
 - piecewise quadratic, 292
 - pipe-, 208, 395, 470
 - PN, 395, 451
 - principal, 437, 444
 - rational, 105 ff, 242 ff, 269
 - Roman, 256, 311
 - ruled, 127, 223 ff
 - Bézier, 247 ff
 - algebraic, 238 ff
 - conoidal, 268, 286
 - cubic, 252 ff, 253
 - of constant conical curvature, 282
 - of constant slope, 269, 282
 - quadratic, 41
 - rational, 242 ff, 269
 - regular, 225
 - skew, 232, 240
 - spline, 292
 - torsal, 232, 240, 270
 - with const. κ, τ, σ , 275
 - shoulder, 420
 - singularity-, 494
 - Steiner's, 256, 311
 - tangent-, 85, 232, 270
 - algebraic, 342
 - development, 330
 - tensor product, 126
 - symmetry, projective, 32
 - system, fixed/moving, 185
- T**
- tangent, 70, 79, 104
 - asymptotic, 81, 427
 - conjugate, 81, 262, 333, 459
 - of quadric, 34
 - tangent subspace, 45
 - term, leading, 86
 - theorem
 - dimension-, 8
 - inertia, 39
 - of Bézout, 102
 - of Brianchon, 5
 - of Clairaut and Sauer, 495
 - of Ivory, 508
 - of Lüroth, 101
 - of Malus and Dupin, 446
 - of Nash, 328
 - of Pappos, 5
 - of Poinsot, 192
 - of Sannia, 275
 - of Steiner, 52
 - of Sylvester, 39
 - of Thales, 53
 - theorema egregium, 328
 - tolerance, 374, 405
 - topology, 11, 72, 143, 517
 - torsion, 77, 272, 332
 - geodesic, 459
 - trace, 24
 - transformation, *see also* mapping
 - affine, 57
 - elliptic, 67, 533
 - Euclidean congruence, 60
 - Laguerre, 369
 - Möbius, 66
 - projective, *see* mapping, projective
 - shear, 459
 - translation, 16
 - infinitesimal, 186
 - uniform, 187
 - transversality, 103, 104, 475
 - trilateration, 510
- U**
- umbilic, 262, 455
 - umbrella vector, 468
- V**
- variation, *see* property, variation diminishing
 - variety
 - algebraic, 89 ff
 - dual, 104
 - Grassmann, 148
 - irreducible, 91
 - quadratic, 45
 - Segre, 97
 - vector
 - binormal, 77, 332
 - Darboux, 186, 272, 332, 533
 - dual, 155
 - knot, 116
 - moment, *see* moment
 - principal normal, 77
 - simple, 135, 147
 - umbrella, 468
 - velocity, 186
 - vertex, 232
 - of cone, 348
 - of quadratic cone, 33
 - of quadric, 36
 - volume, 58
- W**
- weight, 114, 201
 - work, virtual, 193
- Z**
- Zindler, Konrad, 259