de Gruyter Studies in Mathematics 29

# de Gruyter Studies in Mathematics

1 Riemannian Geometry, 2nd rev. ed., *Wilhelm P. A. Klingenberg*
2 Semimartingales, *Michel Métivier*
3 Holomorphic Functions of Several Variables, *Ludger Kaup and Burchard Kaup*
4 Spaces of Measures, *Corneliu Constantinescu*
5 Knots, *Gerhard Burde and Heiner Zieschang*
6 Ergodic Theorems, *Ulrich Krengel*
7 Mathematical Theory of Statistics, *Helmut Strasser*
8 Transformation Groups, *Tammo tom Dieck*
9 Gibbs Measures and Phase Transitions, *Hans-Otto Georgii*
10 Analyticity in Infinite Dimensional Spaces, *Michel Hervé*
11 Elementary Geometry in Hyperbolic Space, *Werner Fenchel*
12 Transcendental Numbers, *Andrei B. Shidlovskii*
13 Ordinary Differential Equations, *Herbert Amann*
14 Dirichlet Forms and Analysis on Wiener Space, *Nicolas Bouleau and Francis Hirsch*
15 Nevanlinna Theory and Complex Differential Equations, *Ilpo Laine*
16 Rational Iteration, *Norbert Steinmetz*
17 Korovkin-type Approximation Theory and its Applications, *Francesco Altomare and Michele Campiti*
18 Quantum Invariants of Knots and 3-Manifolds, *Vladimir G. Turaev*
19 Dirichlet Forms and Symmetric Markov Processes, *Masatoshi Fukushima, Yoichi Oshima and Masayoshi Takeda*
20 Harmonic Analysis of Probability Measures on Hypergroups, *Walter R. Bloom and Herbert Heyer*
21 Potential Theory on Infinite-Dimensional Abelian Groups, *Alexander Bendikov*
22 Methods of Noncommutative Analysis, *Vladimir E. Nazaikinskii, Victor E. Shatalov and Boris Yu. Sternin*
23 Probability Theory, *Heinz Bauer*
24 Variational Methods for Potential Operator Equations, *Jan Chabrowski*
25 The Structure of Compact Groups, *Karl H. Hofmann and Sidney A. Morris*
26 Measure and Integration Theory, *Heinz Bauer*
27 Stochastic Finance, *Hans Föllmer and Alexander Schied*
28 Painlevé Differential Equations in the Complex Plane, *Valerii I. Gromak, Ilpo Laine and Shun Shimomura*

Werner Fenchel · Jakob Nielsen

# Discontinuous Groups of Isometries in the Hyperbolic Plane

Edited by
Asmus L. Schmidt

Walter de Gruyter
Berlin · New York 2003

*Editor*
Asmus L. Schmidt
University of Copenhagen
Institute for Mathematical Sciences
Universitetsparken 5
2100 Copenhagen
Denmark



*Mathematics Subject Classification 2000:* Primary: 51-02; secondary: 30F35, 30F60, 51M10, 57M60, 11Fxx

*Keywords:* Discontinuous group, discrete group, hyperbolic geometry, non-euclidean geometry, Riemann surface, low dimensional topology

*Photo credit:* p. xvii, p. xviii – courtesy of The Royal Danish Academy of Sciences and Letters (© The Royal Danish Academy of Sciences and Letters)

♾ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

*Library of Congress – Cataloging-in-Publication Data*

Fenchel, W. (Werner), 1905–
Discontinuous groups of isometries in the hyperbolic plane / Werner Fenchel, Jakob Nielsen ; edited by Asmus L. Schmidt.
p. cm. – (De Gruyter studies in mathematics ; 29)
Includes bibliographical references and index.
ISBN 3 11 017526 6 (cloth : alk. paper)
1. Discontinuous groups. 2. Isometries (Mathematics). I. Nielsen, Jakob, 1890–1959. II. Schmidt, Asmus L. III. Title. IV. Series.
QA612.14 .F46 2002
514'.2–dc21 2002068997

ISBN 3-11-017526-6

*Bibliographic information published by Die Deutsche Bibliothek*

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.


Printed in Germany.
Cover design: Rudolf Hübler, Berlin.
Typeset using the authors' TEX files: I. Zimmermann, Freiburg.
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen.

# Preface

This book by Jakob Nielsen (1890–1959) and Werner Fenchel (1905–1988) has had a long and complicated history. In 1938–39, Nielsen gave a series of lectures on discontinuous groups of motions in the non-euclidean plane, and this led him – during World War II – to write the first two chapters of the book (in German). When Fenchel, who had to escape from Denmark to Sweden because of the German occupation, returned in 1945, Nielsen initiated a collaboration with him on what became known as the Fenchel–Nielsen manuscript. At that time they were both at the Technical University in Copenhagen. The first draft of the Fenchel–Nielsen manuscript (now in English) was finished in 1948 and it was planned to be published in the Princeton Mathematical Series. However, due to the rapid development of the subject, they felt that substantial changes had to be made before publication.

When Nielsen moved to Copenhagen University in 1951 (where he stayed until 1955), he was much involved with the international organization UNESCO, and the further writing of the manuscript was left to Fenchel. The archives of Fenchel now deposited and catalogued at the Department of Mathematics at Copenhagen University contain two original manuscripts: a partial manuscript (manuscript 0) in German containing Chapters I–II (§§1–15), and a complete manuscript (manuscript 1) in English containing Chapters I–V (§§1–27). The archives also contain part of a correspondence (first in German but later in Danish) between Nielsen and Fenchel, where Nielsen makes detailed comments to Fenchel's writings of Chapters III–V. Fenchel, who succeeded N. E. Nørlund at Copenhagen University in 1956 (and stayed there until 1974), was very much involved with a thorough revision of the curriculum in algebra and geometry, and concentrated his research in the theory of convexity, heading the International Colloquium on Convexity in Copenhagen 1965. For almost 20 years he also put much effort into his job as editor of the newly started journal Mathematica Scandinavica. Much to his dissatisfaction, this activity left him little time to finish the Fenchel–Nielsen project the way he wanted to.

After his retirement from the university, Fenchel – assisted by Christian Siebeneicher from Bielefeld and Mrs. Obershelp who typed the manuscript – found time to finish the book *Elementary Geometry in Hyperbolic Space*, which was published by Walter de Gruyter in 1989 shortly after his death. Simultaneously, and with the same collaborators, he supervised a typewritten version of the manuscript (manuscript 2) on discontinuous groups, removing many of the obscure points that were in the original manuscript. Fenchel told me that he contemplated removing parts of the introductory Chapter I in the manuscript, since this would be covered by the book mentioned above; but to make the Fenchel–Nielsen book self-contained he ultimately chose not to do so. He did decide to leave out §27, entitled *The fundamental group*.

As editor, I started in 1990, with the consent of the legal heirs of Fenchel and Nielsen, to produce a TEX-version from the newly typewritten version (manuscript 2). I am grateful to Dita Andersen and Lise Fuldby-Olsen in my department for having done a wonderful job of typing this manuscript in AMS-TEX. I have also had much help from my colleague Jørn Børling Olsson (himself a student of Käte Fenchel at Aarhus University) with the proof reading of the TEX-manuscript (manuscript 3) against manuscript 2 as well as with a general discussion of the adaptation to the style of TEX. In most respects we decided to follow Fenchel's intentions. However, turning the typewritten edition of the manuscript into TEX helped us to ensure that the notation, and the spelling of certain key-words, would be uniform throughout the book. Also, we have indicated the beginning and end of a proof in the usual style of TEX.

With this TEX-manuscript I approached Walter de Gruyter in Berlin in 1992, and to my great relief and satisfaction they agreed to publish the manuscript in their series Studies in Mathematics. I am most grateful for this positive and quick reaction. One particular problem with the publication turned out to be the reproduction of the many figures which are an integral part of the presentation. Christian Siebeneicher had at first agreed to deliver these in final electronic form, but by 1997 it became clear that he would not be able to find the time to do so. However, the publisher offered a solution whereby I should deliver precise drawings of the figures (Fenchel did not leave such for Chapters IV and V), and then they would organize the production of the figures in electronic form. I am very grateful to Marcin Adamski, Warsaw, Poland, for his fine collaboration concerning the actual production of the figures.

My colleague Bent Fuglede, who has personally known both authors, has kindly written a short biography of the two of them and their mathematical achievements, and which also places the Fenchel–Nielsen manuscript in its proper perspective. In this connection I would like to thank The Royal Danish Academy of Sciences and Letters for allowing us to include in this book reproductions of photographs of the two authors which are in the possession of the Academy.

Since the manuscript uses a number of special symbols, a list of notation with short explanations and reference to the actual definition in the book has been included. Also, a comprehensive index has been added. In both cases, all references are to sections, not pages.

We considered adding a complete list of references, but decided against it due to the overwhelming number of research papers in this area. Instead, a much shorter list of monographs and other comprehensive accounts relevant to the subject has been collected.

My final and most sincere thanks go to Dr. Manfred Karbe from Walter de Gruyter for his dedication and perseverance in bringing this publication into existence.

Copenhagen, October 2002 *Asmus L. Schmidt*

# Contents

# Life and work of the Authors

Jakob Nielsen[1] was born on October 15, 1890 in the village Mjels in Northern Schleswig (then under Germany), where his father owned a small farm. After attending the village school Jakob was taken to Rendsburg in 1900, where he went to the Realgymnasium. In 1908 he entered the University of Kiel and attended lectures in physics, chemistry, geology, biology, and literature. Only after some terms did mathematics take a prominent place. When Max Dehn was attached to the university at the end of 1911 he introduced Jakob Nielsen to topology and group theory at the level of current research. Their contact developed into a lifelong friendship.

In the summer of 1913 Nielsen graduated from the university with the doctor's thesis "Kurvennetze auf Flächen", which already points towards his later achievements. But shortly afterwards he was called up for service in the German navy, attached to the coast defence artillery. The war had broken out, and he was sent first to Belgium and then in April 1915 to Constantinople as one of the German officers functioning as advisers to the Turkish government on the defence of the Bosporus and the Dardanelles. He found time to write two short papers, published in 1917 and 1918, in continuation of his thesis and dealing with finitely generated free groups.

Back in Germany after the war had ended, Nielsen spent the summer term of 1919 in Göttingen, where he met Erich Hecke and later accompanied him to Hamburg as his assistent and "Privatdozent"; they too became close friends. From that period we have two papers of Nielsen both dealing with the fixed point problem for surface mappings.

Already in 1920 Jakob Nielsen was appointed professor at the Institute of Technology in Breslau, where he resumed contact with Dehn. In lectures here Nielsen formulated clearly the central problem he had set himself to solve: to determine and investigate the group of homotopy classes of homeomorphisms of a given surface. One link of this investigation, namely the proof that every automorphism of the fundamental group of a closed surface is induced by a homeomorphism, had been communicated to him by Dehn, who never published it. It is characteristic of Nielsen that whenever he needed this theorem, or merely some idea resembling its proof, he would stress his debt to Dehn.

The stay in Breslau became a brief one, for later in 1920 North Schleswig was reunited with Denmark after a referendum, and Jakob Nielsen opted for Denmark. He moved to Copenhagen the year after together with his wife Carola (née von Pieverling), and here he became a lecturer at the Royal Veterinary and Agricultural College. Quickly he became a treasured member of the Danish mathematical community. He met frequently with Harald Bohr and Tommy Bonnesen, and they followed each other's work with keen interest.

[1] What is written above about Jakob Nielsen and his work is largely an extract of Werner Fenchel's comprehensive memorial address at a meeting in the Danish Mathematical Society on 7 December 1959, printed in Acta Mathematica **103** (1960), vii–xix.

In a purely group theoretic paper by Nielsen, from 1921, a major result is that every subgroup of a finitely generated free group is likewise free. His proof is based on an ingenious method of reduction of systems of generators. In 1927 the theorem was extended by Otto Schreier to arbitrary free groups, and under the name of the Nielsen–Schreier theorem it contributes now one of the bases of the theory of infinite groups. Two other papers, from 1924, continue earlier investigations of the group of automorphisms of a given group.

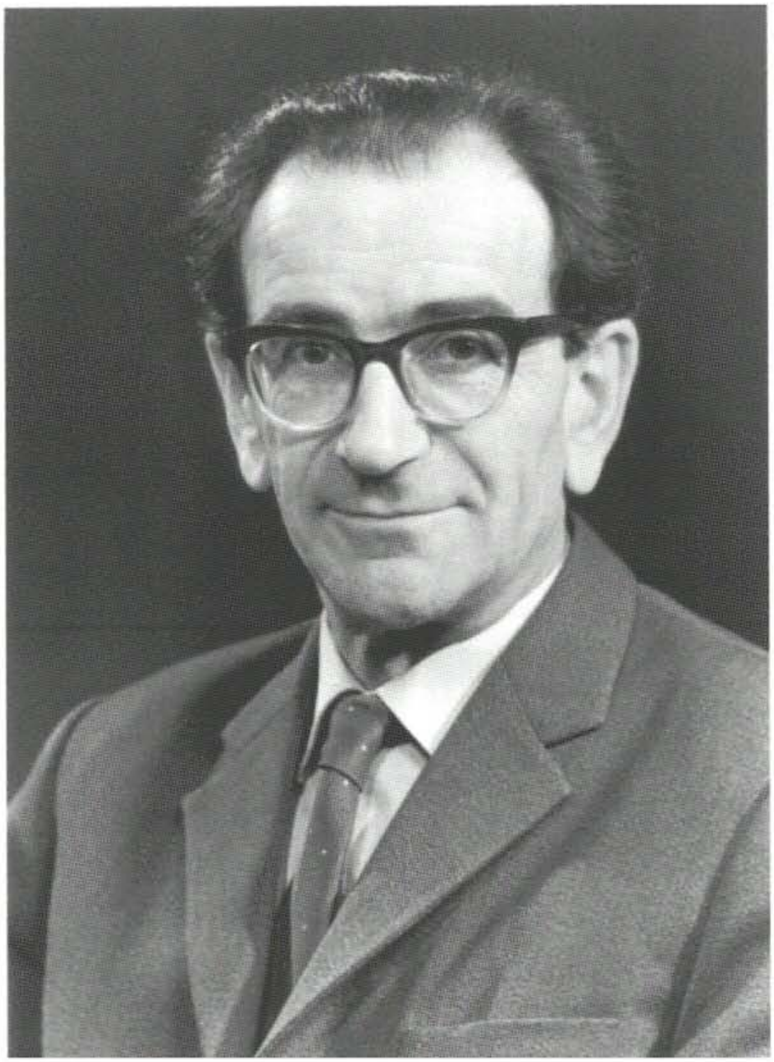
Along with these and other investigations Jakob Nielsen took up the study of discontinuous groups of isometries of the non-euclidean plane and devoted several papers (1923, 1925, 1927) to this subject. His interest in it arose from the fact that the fundamental group of a surface of genus greater than 1 admits representations by such discontinuous groups.

These apparently somewhat desultory investigations turned out to be stones that went to the erection of an impressive building. Hints of this are to be found in some lectures given by Nielsen in Hamburg in 1924 and in Copenhagen in 1925, at the 6th Scandinavian Congress of Mathematicians. But in its final form it appeared in three long memoirs (300 pages in all) from the years 1927, 1929, and 1932 in Acta Mathematica under the common title "Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen". Here we find again the notions and methods he had previously used or developed: the universal covering surface interpreted as the non-euclidean plane, the latter represented by the conformal model in the interior of the unit circle; the fundamental group as a discontinuous group of isometries of the non-euclidean plane; the mappings of the latter onto itself which lie over a given surface mapping, and the automorphisms induced by them. As an essential new tool comes here the following theorem: Every mapping of the non-euclidean plane onto itself which lies over some surface mapping can be extended continuously to the points of the unit circle, representing the points at infinity of the non-euclidean plane, and the mapping of the circumference which arises in this way depends only on the homotopy class of the surface mapping. A two-dimensional problem is hereby reduced to a one-dimensional one. – With these memoirs Jakob Nielsen had broken new ground, and they gave him great international reputation.

In 1925 Nielsen became professor of theoretical mechanics at the Technical University in Copenhagen. Here he worked out his textbook on that subject in two volumes, published in 1933–34, in which he exploited more recent mathematical tools. A third volume about aerodynamics was added later. Nielsen's lectures demanded much of his students; he had an unusual power of expressing himself with great lucidity, but also with great terseness.

It is not possible here to mention the many papers, about 20, among them several comprehensive ones, which Jakob Nielsen published in the years after 1935, most of them carrying on his investigations on surface mappings. By means of the powerful tools developed in the previous papers, he succeeded in solving a series of related problems. In 1937 he gave a complete classification of the periodic mappings of a surface onto itself, and in 1942 a fourth great memoir, "Abbildungsklassen endlicher

Jakob Nielsen



Werner Fenchel

Ordnung", was published in Acta Mathematica. It deals with a problem to which he had been led in the third of the above mentioned Acta papers, and which he had solved there in some special cases: Does every homotopy class of surface mappings which is of finite order, in the sense that a certain power of it is the class of the identity mapping, contain a periodic mapping, that is, a mapping the same power of which is the identity? The proof that this is the case is extremely difficult and makes up all the 90 pages long paper. One cannot but admire the intellectual vigour with which this investigation is carried out. Finally I shall mention one more large paper: "Surface transformation classes of algebraically finite type" from 1944, in which more general classes of surface mappings are thoroughly investigated.

On several occasions Jakob Nielsen lectured at the Mathematical Institute of Copenhagen University to a small circle of young mathematicians on subjects that occupied him in connection with his research. Of special interest is a series of lectures on discontinuous groups of isometries of the non-euclidean plane, given in the year 1938–39; here he took up the theory for a certain class of these groups for its own sake. He realized the need for investigating the theory of discontinuous groups of motions in the non-euclidean plane in its full generality and from the bottom, in view of its many important fields of applications. Gradually it became clear, however, that this task, which Jakob Nielsen took up together with Werner Fenchel, was considerably more extensive than anticipated.

Although his heart was at this project, Jakob Nielsen could only devote to it a moderate part of his great working power, for since the end of the 1939–1945 war he was deeply engaged in international cooperation, especially the work of UNESCO, where he was a highly esteemed member of the Executive Board from 1952 to 1958.

In 1951 Jakob Nielsen was nominated Harald Bohr's successor at the University of Copenhagen. Here he lectured with delight and zeal to young mathematicians on subjects close to his heart. But the growing demands made upon him by his UNESCO work, with frequent journeys abroad, which interrupted his lectures, caused him in 1955 to resign his professorship; and after finishing his UNESCO work he could devote himself wholeheartedly to the work on the monograph with Fenchel. Jakob Nielsen succeeded in surmounting a difficulty which had long prevented a satisfactory conclusion. But already in January 1959 he was stricken with the disease which carried him off on the 3rd of August.

Werner Fenchel was born on the 3rd of May 1905 in Berlin, son of a representative. Already in highschool his deep interest in physics led him into mathematical studies far beyond the school curriculum. Aged eighteen he entered the University of Berlin, where he attended lectures by Einstein among others. With the growing demands of mathematical knowledge needed to understand the theory of relativity, Fenchel eventually concentrated foremost on mathematics. Towards the end of his study he succeeded in proving that the total curvature of a closed curve in space is at least $2\pi$. He presented his result in the mathematics colloquium, and afterwards Erhard Schmidt decided right away that this would be suitable for a doctoral thesis.

Soon after graduating from the university in 1928 Fenchel was lucky to become assistent of Edmund Landau in Göttingen. At this leading centre of mathematics, counting Hilbert among its professors, Werner Fenchel met Harald Bohr, who was guest lecturing, and also briefly Jakob Nielsen for the first time. A Rockefeller stipend allowed Fenchel to spend a semester in Rome, studying differential geometry with Levi-Cività, and also to visit Bohr in Copenhagen in the spring of 1931. Here he also met Bonnesen, with whom he wrote in the following years the Ergebnisse tract "Theorie der konvexen Körper", published in 1934. Reprinted in 1976, it has become a classic in convexity theory.

In 1933 Werner Fenchel, like so many others, had to leave Germany. Invited by Harald Bohr he went to Copenhagen with his wife Käte (née Sperling), a group theorist. Here he continued assisting Otto Neugebauer in editing the Zentralblatt für Mathematik. He also translated and adapted Jakob Nielsen's textbook on theoretical mechanics to German. Inspired by Bohr's theory of almost periodic functions Fenchel wrote with him a paper on stable almost period motions (1936); and in a paper with Jessen (1935) he showed that every almost periodic motion on certain types of surfaces can be deformed continuously and almost periodically into a periodic motion. A paper by Fenchel from 1937 deals with motions in a euclidean space which are almost periodic modulo isometries. Retrospectively, these investigations of almost periodic motions may be seen as forerunners to the theory of dynamical systems.

The cooperation with Bonnesen led Fenchel to new contributions to the theory of convex bodies as developed by Brunn and Minkowski. He succeeded in solving a long standing problem about extension of Minkowski's inequalities for mixed volumes (1936). The Brunn-Minkowski theory had been developed in two extreme cases, the convex body being either smoothly bounded or a polytope. In a memoir from 1938 Fenchel and Jessen succeeded, independently of A. D. Alexandrov, in extending the theory to general convex bodies.

The German occupation of Denmark during the 1939–45 war forced in 1943 Werner and Käte to leave their new home country. Helped by Marcel Riesz they found refuge in Lund, together with their little son Tom. After the end of the war they returned to Denmark, where Fenchel in 1947 had his first tenure position, at the Technical University in Copenhagen, and here he succeeded in 1951 Jakob Nielsen as professor of theoretical mechanics.

Werner Fenchel visited the United States with his family in 1950–51, staying at U.S.C. in Los Angeles with his close friend Herbert Busemann, next at Stanford with Polýa and Szegö, and finally in Princeton at the Institute for Advanced Study and Princeton University. In a short paper from 1949 Fenchel had sketched ideas which were to lead to a far-reaching development in convexity theory. He associated with each convex function on a euclidean space a conjugate function, likewise convex, and established the basic properties of this concept of duality. This theory entered naturally in a series of lectures he gave at Princeton University, and mimeographed notes were written. These certainly ought to have been properly published, but copies soon began to circulate widely and had a great impact on research in this field.

Back in Denmark, Werner Fenchel seems to have put the duality theory aside, his publications from the 1950's dealing with other aspects of convexity and with geometrical and topological topics. In the light of the development in the theory of topological vector spaces it was, however, clear to Fenchel that it was desirable to extend his theory of conjugate convex functions to these very general spaces, and thereby widen its applicability. Thus encouraged, one of his students, Arne Brøndsted, carried out that project in a comprehensive paper published in 1964.

In a pioneering monograph "Convex Analysis" from 1970, R. T. Rockafellar applied the duality theory to create a theory of convex optimization based on the ideas of Kuhn and Tucker. This aspect of mathematical optimization has become an integral part of theoretical economics. Earlier, the author had spent a year in Copenhagen with Fenchel. In the preface to his book Rockafellar emphasizes the great impact Fenchel's lecture notes from Princeton had on his own perception of convexity theory, and he writes: "It is highly fitting, therefore, that this book be dedicated to Fenchel as honorary co-author".

In 1956 Fenchel had succeeded N.E. Nørlund as professor at the University of Copenhagen. He was an inspiring lecturer, with a delightful ability of vizualizing his subject. The newly started Journal Mathematica Scandinavica had Fenchel as a very dedicated editor during nearly twenty years. Likewise for many years he was secretary of the Danish Mathematical Society, and from 1958 to 1962 its chairman. In 1965 he organized a big international colloquium on convexity theory in Copenhagen.

After the war Werner Fenchel had joined Jakob Nielsen in pursuing the study of discontinuous groups of isometries of the hyperbolic plane. This led to a joint paper in 1948, and in the same year Fenchel published two more articles on that topic. As described in the above outline of Nielsen's work, their project of developing the theory from its basis with the aim of giving a comprehensive presentation of it turned out to be much bigger than foreseen. Provisional sketches of their work had circulated in a few copies among researchers in the field and excited keen interest.

After Jakob Nielsen's death in 1959 Werner Fenchel continued the project alone – no less so after his retirement from the university in 1974. In periods he was assisted by younger colleagues: Asmus Schmidt, Nils Andersen, Troels Jørgensen (then in Copenhagen), and Christian Siebeneicher in Bielefeld. And late in his life Fenchel succeeded in completing the body of the manuscript.

While working on the Nielsen project, Werner Fenchel had realized the need for a comprehensive exposition of the underlying hyperbolic geometry, also in higher dimensions and based on the conformal model. And shortly before his death on 24 January 1988 he had completed the manuscript to the monograph "Elementary Geometry in Hyperbolic Space", which was published the year after in the de Gruyter Studies.

# Chapter I
# Möbius transformations and non-euclidean geometry

## §1 Pencils of circles – inversive geometry

**1.1 Notations.** The following considerations are based on the plane of all complex numbers, this plane being closed as usual by a point at infinity, in other words on Riemann's sphere of complex numbers. In general, the points of the plane as well as the corresponding complex numbers are denoted by small Latin letters, real numbers by Greek letters. The straight lines of the plane are considered as circles passing through the point at infinity; even single points will occasionally be included among the circles and are then spoken of as *zero-circles*. The circles of the plane in this general sense – as well as other subsets of the plane – will be denoted by calligraphic capitals. In this paragraph some definitions and theorems concerning *pencils* of circles are enumerated for subsequent use.

**1.2 Three kinds of pencils.** An *elliptic pencil*[1] consists of all circles passing through two different points $u$ and $v$, the common points of the pencil. Each point of the plane

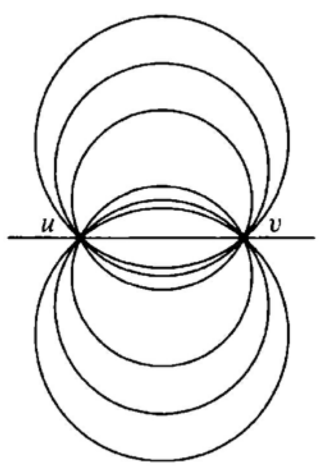



Figure 1.1

other than $u$ and $v$ lies on exactly one circle of the pencil. If one of the common points,

[1] Editor's note: the names elliptic and hyperbolic pencil have been switched as compared with the first edition of the Fenchel–Nielsen manuscript. It is now in accordance with common usage, cf. [15], [31], [55]. Earlier the expression coaxal circles were in use, cf. [25], [45]

in the sequel usually $u$, is termed the negative and the other the positive, the pencil is said to be *directed*. It is often appropriate to think of a directed elliptic pencil as made up of circular arcs joining $u$ and $v$ and directed from $u$ towards $v$.

A *parabolic pencil* consists of all circles touching each other in some definite point $u$, the common point or zero-circle of the pencil. A direction of the circles in $u$ is called the direction of the pencil. Each point of the plane other than $u$ lies on exactly one circle of the pencil.

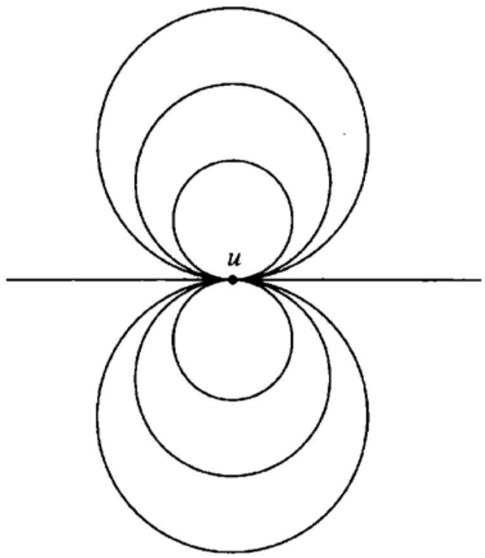



Figure 1.2

A *hyperbolic pencil* consists of all circles which are orthogonal to all circles of an elliptic pencil. The common points $u$ and $v$ of the elliptic pencil are included in the hyperbolic pencil as zero-circles. If none of these is at infinity, the hyperbolic pencil is made up of all apollonian circles for the points $u$ and $v$, i.e. each circle of the pencil is the locus of all points whose distances from $u$ and $v$ are in a fixed ratio. If one of the zero-circles is at infinity, the pencil consists of all circles with the other zero-circle as their common centre. The two zero-circles are separated by every other circle of the pencil. Each point of the plane lies on exactly one circle of the pencil. If one of the zero-circles is termed the negative and the other the positive, the pencil is said to be directed. In that case the circles of the pencil are directed in accordance with the usual orientation of the complex plane when seen from the positive zero-circle.

**1.3 Determination of pencils.** The hyperbolic and elliptic pencil based on two different points $u$ and $v$ are called *conjugate*. The conjugate of a parabolic pencil is a parabolic pencil with the same common point and with a direction at right angles to the direction of the first pencil. Two different circles determine exactly one pencil of which they are members. This pencil is elliptic, parabolic, or hyperbolic according as

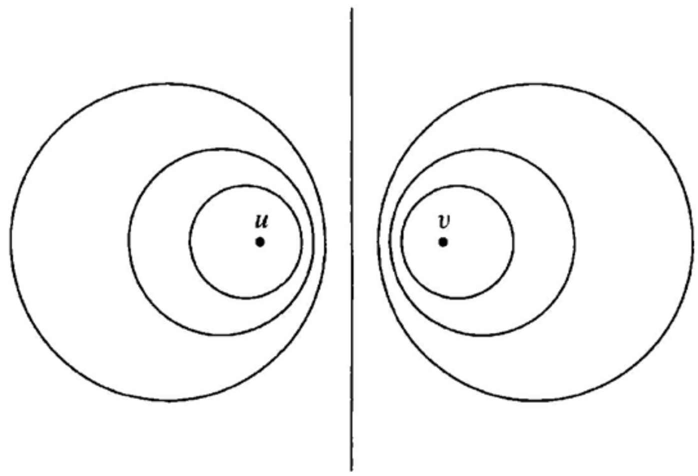



Figure 1.3

the two given circles intersect, touch, or have no point in common. (In this connection a zero-circle touches another circle if the point lies on the circle.)

Two different circles determine exactly one pencil of circles to which they are orthogonal. This pencil is hyperbolic, parabolic, or elliptic, according as the two given circles intersect, touch, or have no point in common. This pencil is conjugate to the pencil of which the two given circles are members. In this connection a zero-circle is orthogonal to another circle, if the point lies on the circle.

If three circles do not belong to one and the same pencil, the necessary and sufficient condition for the existence of exactly one circle orthogonal to all three is the following: If at least two of the circles intersect, the pair of intersection points of the first circle with the second are not separated on the first circle by the pair of its intersection points (if any) with the third. (In this connection a zero-circle has to be considered as orthogonal to itself.)

**1.4 Inverse points.** Two points are called *inverse with respect to a circle $\mathcal{K}$*, which is not a zero-circle, if they are the zero-circles of a hyperbolic pencil to which $\mathcal{K}$ belongs; in other words if they are the common points of an elliptic pencil orthogonal to $\mathcal{K}$. To every point $x$ not on $\mathcal{K}$ there exists exactly one inverse, the second zero-circle of the hyperbolic pencil determined by the zero-circle $x$ and the circle $\mathcal{K}$. Two points inverse with respect to $\mathcal{K}$ are separated by $\mathcal{K}$. The inverse of a point on $\mathcal{K}$ is, by definition, the point itself. The mapping which assigns to a point of the plane its inverse with respect to $\mathcal{K}$ is called the *inversion with respect to $\mathcal{K}$*.

# §2 Cross-ratio

**2.1 Definition and identities.** The *cross-ratio* of two pairs of points $x_1, y_1$ and $x_2, y_2$ (thus of four points $x_1, y_1, x_2, y_2$ given in this order) is denoted by $(x_1 y_1 x_2 y_2)$ and defined as the complex number

$$(x_1 y_1 x_2 y_2) = \frac{x_2 - x_1}{x_2 - y_1} : \frac{y_2 - x_1}{y_2 - y_1} = \frac{(x_2 - x_1)(y_2 - y_1)}{(x_2 - y_1)(y_2 - x_1)}.$$

This definition has a meaning if no three among the four points coincide. The cross-ratio assumes the special values 0, $\infty$ and 1 in the following cases respectively, and in these cases only: If the two first or the two second points of the pairs coincide; if the first point of one pair coincides with the second of the other; if the two points of one pair coincide. Given any three different points $x_1, y_1, x_2$ there exists exactly one point $y_2$ such that the cross-ratio assumes a prescribed value. The following relations hold:

$$(x_2 y_2 x_1 y_1) = (y_1 x_1 y_2 x_2) = (x_1 y_1 x_2 y_2) \tag{1}$$

$$(y_1 x_1 x_2 y_2) = (x_1 y_1 y_2 x_2) = \frac{1}{(x_1 y_1 x_2 y_2)} \tag{2}$$

$$(x_1 x_2 y_1 y_2) = (y_2 y_1 x_2 x_1) = 1 - (x_1 y_1 x_2 y_2) \tag{3}$$

$$(x_1 y_1 x_2 z)(x_1 y_1 z\, y_2) = (x_1 y_1 x_2 y_2). \tag{4}$$

**2.2 Amplitude and modulus.** First, let the four points $x_1, y_1, x_2, y_2$ be different and none of them at infinity. Let $x_1$ and $y_1$ be joined by two circular arcs passing through $x_2$ and $y_2$ respectively, and let the half-tangents $y_1 s$ and $y_1 t$ of these circular arcs at $y_1$ by drawn (see Fig. 2.1). Counting the sign of angles in accordance with the

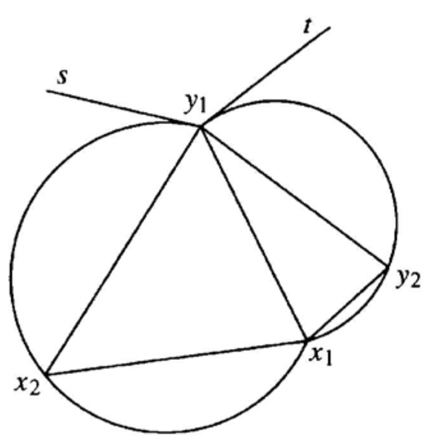



Figure 2.1

orientation of the complex plane, Fig. 2.1 illustrates the following relation:

$$
\begin{aligned}
sy_1t &= sy_1x_2 + x_2y_1y_2 + y_2y_1t \\
&= x_2y_1y_2 + y_1x_1x_2 + y_2x_1y_1 = x_2y_1y_2 + y_2x_1x_2 \\
&= \operatorname{amp}\frac{y_2-y_1}{x_2-y_1} + \operatorname{amp}\frac{x_2-x_1}{y_2-x_1} = \operatorname{amp}(x_1y_1x_2y_2).
\end{aligned}
$$

Hence the amplitude of the cross-ratio of two pairs of points equals the angle between the two circular arcs joining the points of one pair and passing each through one point of the other pair. In particular, the condition for the cross-ratio being real is that all four points are on one circle, the cross-ratio being negative or positive according as the pairs $x_1$, $y_1$ and $x_2$, $y_2$ separate or do not separate each other on that circle.

Moreover, $|x_2-x_1|/|x_2-y_1|$ and $|y_2-x_1|/|y_2-y_1|$ equals the ratio of distances of the points $x_2$ and $y_2$ respectively from the points $x_1$ and $y_1$. The first ratio remains unaltered if $x_2$ is displaced on the apollonian circle for $x_1$ and $y_1$ passing through $x_2$, thus on a circle of the hyperbolic pencil determined by $x_1$ and $y_1$ as zero-circles; and equally for $y_2$. In particular, $x_2$ and $y_2$ may be replaced by the intersection points $x_2'$ and $y_2'$ of these two circles with any circular arc joining $x_1$ and $y_1$. Hence

$$|(x_1y_1x_2y_2)| = (x_1y_1x_2'y_2')$$

this cross-ratio being positive. The condition for

$$|(x_1y_1x_2y_2)| = 1$$

is that $x_2$ and $y_2$ are on the same apollonian circle for $x_1$ and $y_1$.

The necessary and sufficient conditions for these special cases may be so formulated: The cross-ratio for two pairs of points is real (in particular: positive) if the points of one pair are situated on one circle (in particular: circular arc) of the elliptic pencil which is determined by the other pair as common points; in this case the two pairs are called *concyclical*. The cross-ratio for two pairs of points has modulus 1, if the points of one pair are situated on one circle of the hyperbolic pencil which is determined by the other pair as zero-circles.

It is easily seen that this holds even if the point at infinity or coincidences of points are admitted, with the restriction that coincident points of one pair cannot, of course, play the rôle of common points or zero-circles of the above pencils.

**2.3 Harmonic pairs.** Two pairs of points are called *harmonic*, if

$$(x_1y_1x_2y_2) = -1. \tag{5}$$

If $x_1$ and $y_1$ are chosen as common points of an elliptic and as zero-circles of a hyperbolic pencil the necessary and sufficient condition for the validity of (5) is that $x_2$ and $y_2$ are the intersection points of a circle of one pencil with a circle of the other.

Now, conjugate pencils are orthogonal. Hence, if $\mathcal{K}$ is the circle passing through two harmonic, and thus concyclical pairs $x_1, y_1$ and $x_2, y_2$ and $\mathcal{K}_1$ and $\mathcal{K}_2$ are circles orthogonal to $\mathcal{K}$ and passing through $x_1$ and $y_1$ and through $x_2$ and $y_2$ respectively, then $\mathcal{K}_1$ and $\mathcal{K}_2$ are mutually orthogonal. $x_1$ and $y_1$ are inverse with respect to $\mathcal{K}_2$, and so are $x_2$ and $y_2$ with respect to $\mathcal{K}_1$. Conversely, if three circles are mutually orthogonal, each of them cuts the two others in harmonic pairs.

If the points $x$ and $x'$ are inverse with respect to the circle $\mathcal{K}$, every circle through $x$ and $x'$ will cut $\mathcal{K}$ in a pair of points which is harmonic with the pair $x, x'$.

If none of the four points is at infinity, equation (5) may by written

$$(x_2 - x_1)(y_2 - y_1) + (x_2 - y_1)(y_2 - x_1) = 0$$

or

$$2(x_1 y_1 + x_2 y_2) - (x_1 + y_1)(x_2 + y_2) = 0.$$

This equation obviously holds in the case when three of the four points coincide, in which case no cross-ratio is defined. In the sequel it is appropriate to include this case in the term *harmonic pairs*.

## §3 Möbius transformations, direct and reversed

**3.1 Invariance of the cross-ratio.** The set of linear fractional transformations

$$x \mapsto x' = \frac{ax + b}{cx + d}, \quad ad - bc \neq 0 \tag{1}$$

with complex coefficients constitute a group of bijective mappings of the closed complex plane onto itself. Multiplication of the matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ of coefficients of the transformation (1) by a factor $\neq 0$ does not affect the transformation, and so by a suitable choice of such a factor the determinant $ad - bc$ can be given the value 1.

It is easily seen that all angles are preserved under the mapping by the *Möbius transformation* (1): If $x = x(\tau)$ is a parametric representation of some differentiable curve, amp $\frac{dx}{d\tau}$ is equal to the angle from the real axis to the tangent of the curve directed in the sense of the increase of $\tau$. For the image of the curve the corresponding angle is

$$\operatorname{amp} \frac{dx'}{dt} = \operatorname{amp} \frac{d}{dt} \frac{ax + b}{cx + d} = \operatorname{amp} \frac{dx}{dt} + \operatorname{amp} \frac{ad - bc}{(cx + d)^2}. \tag{2}$$

Thus amp $\frac{dx}{dt}$ is increased by an amount which depends on the point considered but which is independent of the direction of the curve at that point. Hence the angle subtended at the intersection point of two curves remains unaltered by the mapping both in magnitude and in sign. This remains valid for the point at infinity if angles are measured at the point $x = 0$ after performing the transformation $x \mapsto x' = \frac{1}{x}$.

Let $x_1, y_1, x_2, y_2$ be any four points, no three of which coincide, and $x_1', y_1'$, $x_2', y_2'$ their images under the Möbius transformation (1). If none of the four points is $-\frac{d}{c}$ or $\infty$, then

$$\frac{x_2' - x_1'}{x_2' - y_1'} = \frac{\frac{ax_2+b}{cx_2+d} - \frac{ax_1+b}{cx_1+d}}{\frac{ax_2+b}{cx_2+d} - \frac{ay_1+b}{cy_1+d}} = \frac{x_2 - x_1}{x_2 - y_1} \cdot \frac{cy_1 + d}{cx_1 + d},$$

and likewise, since the second factor of the right-hand member does not depend on $x_2$,

$$\frac{y_2' - x_1'}{y_2' - y_1'} = \frac{y_2 - x_1}{y_2 - y_1} \cdot \frac{cy_1 + d}{cx_1 + d}.$$

Hence

$$(x_1' y_1' x_2' y_2') = (x_1 y_1 x_2 y_2),$$

showing the invariance of the cross-ratio under the transformation (1). Continuity then shows this even holds in the special cases excluded above.

Since the reality of the cross-ratio characterizes the concyclical disposition of four points, any Möbius transformation maps circles onto circles. Combining this property with the property of isogonality, it follows that the circles of a pencil are mapped onto the circles of a pencil of the same kind. In particular, two points inverse with respect to some circle are mapped onto two points which are inverse with respect to the image of that circle.

**3.2 Determination by three points.** Let $x_1, x_2, x_3$ and $x_1', x_2', x_3'$ be any two triples each made up of three different points. Then there is exactly one Möbius transformation (1) carrying $x_1$ into $x_1'$, $x_2$ into $x_2'$ and $x_3$ into $x_3'$: Denoting by $x'$ the image-point of an arbitrary point $x$, the invariance of the cross-ratio yields the equation

$$(x_1' x_2' x_3' x') = (x_1 x_2 x_3 x)$$

from which $x'$ is calculated as a linear fractional function of $x$ with the required property; the determinant of this transformation is

$$(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)(x_1' - x_2')(x_1' - x_3')(x_2' - x_3')$$

and thus does not vanish in virtue of the conditions stated. Consequently, a Möbius transformation leaving three points fixed is the identical transformation.

**3.3 Reversed transformations.** Transformations like

$$x \mapsto x' = \frac{a\overline{x} + b}{c\overline{x} + d}, \quad ad - bc \neq 0, \tag{3}$$

$\overline{x}$ denoting the conjugate of $x$, produce bijective mappings of the closed complex plane onto itself reversing orientation. They are the composition of an inversion with respect to the real axis and a Möbius transformation and may be called *reversed Möbius transformations*. Angles are left unaltered in magnitude but are reversed in sign. Cross-ratios are replaced by their conjugate values. Circles are mapped onto circles and pencils of circles onto pencils of the same kind. Since the product of two reversed Möbius transformations is a direct Möbius transformation, the set of all Möbius transformations, direct and reversed, constitute a group.

For any two prescribed triples $x_1, x_2, x_3$ and $x'_1, x'_2, x'_3$ each consisting of three different points there exists exactly one reversed Möbius transformation carrying $x_1$ into $x'_1$, $x_2$ into $x'_2$ and $x_3$ into $x'_3$, this transformation being calculated from the equation

$$(x'_1 x'_2 x'_3 x') = (\bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}). \tag{4}$$

**3.4 Inversions.** If the transformation (4) leaves fixed the points $x_1, x_2, x_3$, i.e. if $x_1 = x'_1$, $x_2 = x'_2$, $x_3 = x'_3$, each point of the circle $\mathcal{C}$ passing through these three points remains fixed; for if $x$ is a point on this circle the cross-ratios in (4) are real and hence

$$(x_1 x_2 x_3 x') = (x_1 x_2 x_3 x).$$

This equation implies $x' = x$. In consequence of the properties of reversed transformations described above, every circle orthogonal to $\mathcal{C}$ is mapped onto itself. The pair of common points $u$ and $v$ of the pencil must then be mapped onto itself. Now, $u$ and $v$ cannot be left fixed individually, since in that case every point of every circle of the pencil would be invariant, and the transformation cannot be identical since it is reversed. Hence $u$ and $v$ are interchanged. The transformation thus carries every point of the plane into its inverse with respect to $\mathcal{C}$ and is called *inversion* with respect to $\mathcal{C}$.

A reversed Möbius transformation leaving three points fixed is the inversion with respect to the circle passing through these three points.

## §4 Invariant points and classification of Möbius transformations

**4.1 The multiplier.** The invariant points of a Möbius transformation (3.1) with matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ are determined by the equation

$$x = \frac{ax + b}{cx + d}$$

or

$$cx^2 + (d - a)x - b = 0.$$

In case $c = 0$ the point $x = \infty$ has to be included among its roots. In case $c = d - a = b = 0$, the transformation is the identity; this case needs no consideration.

The equation then has one or two roots according as

$$D = (d-a)^2 + 4bc = (\mathrm{tr}\, A)^2 - 4 \det A$$

is equal to zero or different from zero. Let $u$ and $v$ denote the invariant points of the mapping (different or equal), $x$ any other point and $x'$ its image. In virtue of (3.3) and the fact that $u = u'$, $v = v'$, a short calculation yields

$$k = (uvxx') = \frac{d + a - \sqrt{D}}{d + a + \sqrt{D}} \tag{1}$$

(or the reciprocal value dependent on the choice of $u$ and $v$ after a definite value for $\sqrt{D}$ has been fixed). Thus $k$ is an invariant of the transformation. It is called the *multiplier* of the transformation (3.1). One has:

$$k + 2 + k^{-1} = \frac{(\mathrm{tr}\, A)^2}{\det A}.$$

Moreover, from (3.2) one can calculate the increase of the amplitude in an invariant point:

$$\operatorname{amp} \frac{ad - bc}{(cu + d)^2} = -\operatorname{amp} k, \quad \operatorname{amp} \frac{ad - bc}{(cv + d)^2} = +\operatorname{amp} k. \tag{2}$$

**4.2 Two invariant points.** At first, let $D$ be different from zero, thus $u$ and $v$ different. In virtue of the invariance of the cross-ratio one has

$$(uvx_0'x') = (uvx_0x).$$

On multiplying by $(uvxx_0')$ one gets from (4) in §2

$$(uvxx') = (uvx_0x_0'),$$

showing once more the invariance of $k$. This constant is neither 0 nor 1 nor $\infty$, since all four points are different, $x$ not being invariant. Conversely, under the same conditions the equation

$$(uvxx') = k \tag{3}$$

determines a Möbius transformation with two different invariant points $u$ and $v$. Obviously, the multiplier of the product of two such transformations with the same invariant points is the product of the corresponding multipliers.

The image of any circular arc joining $u$ and $v$ is a circular arc joining $u$ and $v$; hence the circular arcs of the elliptic pencil with $u$ and $v$ as common points are interchanged. In consequence of the isogonality this also holds for the circles of the conjugate hyperbolic pencil with $u$ and $v$ as zero-circles. As stated in §2 and confirmed by calculation in Section 1, the amplitude of $k$ measures the angle through which the circular arcs of the elliptic pencil are rotated about $u$ or $v$; likewise the modulus of $k$ characterizes the displacement of the circles of the hyperbolic pencil.

The necessary and sufficient condition for $k$ being positive is that $x$ and $x'$ are on the same circular arc of the elliptic pencil; each of these circular arcs is then mapped onto itself. These transformations are called *hyperbolic*. The elliptic pencil with $u$ and $v$ as common points is called the *fundamental pencil* of the hyperbolic transformation.

The necessary and sufficient condition for the modulus of $k$ being 1 is that $x$ and $x'$ are on the same circle of the hyperbolic pencil; each of these circles is then mapped onto itself. These transformations are called *elliptic*. The hyperbolic pencil with $u$ and $v$ as zero-circles is called the *fundamental pencil* of the elliptic transformation. Among these elliptic transformations is included the particular case $k = -1$, in which $u$, $v$ and $x$, $x'$ are harmonic; in the two conjugate pencils determined by $u$ and $v$ the two circles passing through $x$ intersect again in $x'$. Thus this transformation is involutory since it interchanges the intersection points. It will be called the *involution with respect to the pair of points* $u$, $v$. A transformation (3.1) with two invariant points $u$ and $v$, which interchanges two different points $x$ and $x'$ is the involution with respect to $u$, $v$. For the equation

$$k = (uvxx') = (uvx'x) = \frac{1}{k}$$

yields $k = -1$, since $k \neq 1$.

If neither $k > 0$ nor $|k| = 1$, no circular arc of the elliptic pencil and no circle of the hyperbolic pencil is mapped onto itself and the transformation is called *loxodromic*.

**4.3 One invariant point.** Secondly, let $D$ be zero, thus $u$ and $v$ coincide. These transformations with only one invariant point, $u$, are called *parabolic*. Since in this case $D = (\operatorname{tr} A)^2 - 4 \det A = 0$ one gets $\operatorname{tr} A \neq 0$ and (1) yields $k = 1$, $\operatorname{amp} k = 0$. Thus the directions in $u$ are left unaltered. Hence any circle through $u$ is mapped onto a circle touching the former in $u$. Any parabolic pencil with $u$ as common point is mapped onto itself.

Let $x$ be any point other than $u$ and $x'$ its image, and draw the circle through $u$, $x$ and $x'$. Its image must touch it in $u$ and pass through $x'$ and therefore coincides with the circle itself. Hence every point other than $u$ lies on a circle which passes through $u$ and coincides with its image. Two such circles have only $u$ in common, since a second common point obviously would be invariant. These circles therefore form a parabolic pencil with $u$ as common point. Thus there exists exactly one parabolic pencil with $u$ as common point, whose circles are mapped onto themselves individually. It is called the *fundamental pencil* of the parabolic transformation. The direction of this pencil in $u$ is called the *fundamental direction* of the parabolic transformation. Conversely, a Möbius transformation which reproduces the circles of a parabolic pencil individually, is parabolic, or the identity. For the common point $u$ of the pencil is invariant, and if there is another invariant point $v$ the circles of the elliptic pencil with $u$ and $v$ as common points are reproduced individually, since the directions in $u$ either remain fixed or are reversed. Every other point of the plane is the intersection of a circle of the parabolic pencil and a circle of the elliptic pencil and thus remains fixed. The transformation, therefore, is the identity.

A parabolic transformation is uniquely determined by the invariant point $u$, another point $x$ and its image $x'$. For the image $y'$ of any other point $y$ is situated both on the circle through $x'$ which touches the circle through $u, x$ and $y$ in $u$ and on the circle through $y$ which touches the circle through $u, x$ and $x'$ in $u$; the latter belongs to the fundamental pencil. – If, in particular, $y$ is on the circle through $u, x$ and $x'$, one may first construct the image $z'$ of an arbitrary point $z$ outside that circle, and then let $z$ and $z'$ play the role of $x$ and $x'$.

No parabolic transformation can interchange two points. For if $u$ is the invariant point, $x$ any other point, and $x'$ its image, the circle through $u, x$ and $x'$ is mapped onto itself in such a way that all its points are displaced in a definite direction without passing through $u$. Hence the image of $x'$ is separated from $x$ by $x'$ and $u$ and, therefore, cannot coincide with $x$. In reviewing the different types investigated it comes out that the involution with respect to a pair of points is the only type of transformation which interchanges two points.

**4.4 Transformations with an invariant circle.** Which are the Möbius transformations (other than the identity) which map a prescribed circle $\mathcal{K}$ onto itself and each of the two regions determined by $\mathcal{K}$ in the plane onto itself?

First, let it be assumed that $\mathcal{K}$ contains no invariant point of the transformation. Let $u$ be an invariant point and denote by $v$ its inverse with respect to $\mathcal{K}$. Since $\mathcal{K}$ is mapped onto itself, a pair of inverse points with respect to $\mathcal{K}$ are mapped onto a pair of inverse points with respect to $\mathcal{K}$. Since $u$ is left fixed, $v$ must be so too. So there is one invariant point in each of the two regions. $\mathcal{K}$ belongs to the hyperbolic pencil with $u$ and $v$ as zero-circles and, since $\mathcal{K}$ is mapped onto itself, the transformation is elliptic.

Secondly, let $\mathcal{K}$ contain two invariant points. $\mathcal{K}$ belongs to the elliptic pencil determined by these points as common points. Since $\mathcal{K}$ is mapped onto itself, so is every other circle of this pencil. Moreover, since the regions are reproduced individually, the same holds for every circular arc of the pencil. The transformation, therefore, is hyperbolic, $k > 0$.

Thirdly, let $\mathcal{K}$ contain one invariant point. If there were another outside $\mathcal{K}$, its inverse with respect to $\mathcal{K}$ would be invariant too and there would be more than two in all, which is impossible. The transformation, therefore, is parabolic, and $\mathcal{K}$ together with the invariant point as zero-circle determines the parabolic pencil whose circles are mapped onto themselves individually.

These three cases form a complete list of direct Möbius transformations of the required nature. In each case the circle $\mathcal{K}$ belongs to the fundamental pencil of the transformation. As far as reversed transformations are concerned, these can be characterized in the following way: Since each of the regions is mapped onto itself with orientation reversed, $\mathcal{K}$ must be so too. Hence there are exactly two invariant points $u$ and $v$ on $\mathcal{K}$. Let $\mathcal{C}$ denote the circle through $u$ and $v$ at right angles to $\mathcal{K}$. In virtue of the isogonality and of the invariance of $u$ and $v$, $\mathcal{C}$ is mapped onto itself and, in particular, in consequence of the conservation of the regions, each of the two arcs into which it falls by $u$ and $v$ is mapped onto itself. Let $x$ be a point on such an arc and $x'$ its image. If one combines the hyperbolic transformation which has $u$ and $v$ as invariant points and carries $x$ into $x'$ with the inversion with respect to $\mathfrak{C}$, one gets a reversed Möbius transformation which maps the three points $u, v, x$ in the same way as the transformation considered and therefore is that transformation itself. In the particular case $x = x'$ the hyperbolic transformation is the identity and the transformation considered is the inversion with respect to $\mathfrak{C}$.

**4.5 Permutable transformations.** In the sequel Möbius transformations, both direct and reversed, will by denoted by small *gothic* characters except for the identity for which the symbol 1 is generally used. For products of such symbols it is understood that the transformation indicated by the last symbol is the first performed, then the preceding one and so on. The inverse of a Möbius transformation $\mathfrak{f}$ is a Möbius transformation of the same kind; it is denoted by $\mathfrak{f}^{-1}$. Let $\mathfrak{f}$ and $\mathfrak{g}$ be two such transformations and $u$ an invariant point for $\mathfrak{f}$. Then the transformation $\mathfrak{g}\mathfrak{f}\mathfrak{g}^{-1}$ evidently has $\mathfrak{g}u$ as invariant point, i.e. the image of $u$ by $\mathfrak{g}$. From this is inferred:

*If two transformations commute then each maps the set of invariant points of the other onto itself.* □

There are several possibilities for two transformations to commute:

Let first both be direct Möbius transformations. If one is parabolic, the other must be so too and with the same invariant point; for a parabolic transformation can neither leave the invariant points of a non-parabolic fixed individually nor interchange them. If both are non-parabolic, each must leave the invariant points of the other fixed individually or interchange them. In the first case they have their invariant points in common; in the second case both must be involutions with respect to pairs of points, and these pairs must be harmonic.

These above necessary conditions prove also to be sufficient: For two parabolic transformations, $\mathfrak{f}$ and $\mathfrak{g}$, with the same invariant point $u$ but with different fundamental directions this is inferred from the above mentioned invariance of the parabolic pencils with $u$ as common point under the transformations $\mathfrak{f}$ and $\mathfrak{g}$ (cf. Fig. 4.1): The image $\mathfrak{C}_1$ of the circle through $u$, $x$ and $\mathfrak{f}x$ by $\mathfrak{g}$ contains $\mathfrak{g}x$ and $\mathfrak{g}\mathfrak{f}x$; since $\mathfrak{C}_1$ belongs to the fundamental pencil for $\mathfrak{f}$, and thus is mapped onto itself by $\mathfrak{f}$, it also contains $\mathfrak{f}\mathfrak{g}x$.

The circle $\mathfrak{C}_2$ through $u$ and $\mathfrak{f}x$ which belongs to the fundamental pencil of $\mathfrak{g}$, and thus is mapped onto itself by $\mathfrak{g}$, contains $\mathfrak{g}\mathfrak{f}x$. As it is the image of the circle through $u$, $x$ and $\mathfrak{g}x$ by $\mathfrak{f}$, it also contains $\mathfrak{f}\mathfrak{g}x$. Hence $\mathfrak{g}\mathfrak{f}x$ and $\mathfrak{f}\mathfrak{g}x$ coincide with the intersection of $\mathfrak{C}_1$ and $\mathfrak{C}_2$. In addition it is seen that the product $\mathfrak{f}\mathfrak{g} = \mathfrak{g}\mathfrak{f}$ is again parabolic, since for every point $x \neq u$ the points $x$ and $\mathfrak{f}\mathfrak{g}x$ lie on two different circles which touch at $u$ and thus cannot coincide. If $\mathfrak{f}$ and $\mathfrak{g}$ have the same fundamental direction and thus the same fundamental pencil, it is evident that $\mathfrak{f}\mathfrak{g}$ and $\mathfrak{g}\mathfrak{f}$ are parabolic, since they reproduce individually the circles of that pencil. Now, let $\mathfrak{h}$ be a parabolic transformation with the same invariant point $u$ but with another fundamental direction; $\mathfrak{h}$ is thus permutable with $\mathfrak{f}$ and $\mathfrak{g}$. Then $\mathfrak{f}\mathfrak{h}^{-1}$ and $\mathfrak{h}\mathfrak{g}$ are parabolic with $u$ as invariant point, and their fundamental directions are different, since for $x \neq u$ (cf. Fig. 4.2) the points $x$, $\mathfrak{h}\mathfrak{g}x$ and $\mathfrak{f}\mathfrak{h}^{-1}(\mathfrak{h}\mathfrak{g}x) = \mathfrak{f}\mathfrak{g}x$ are not situated on a circle through $u$. They are, therefore,

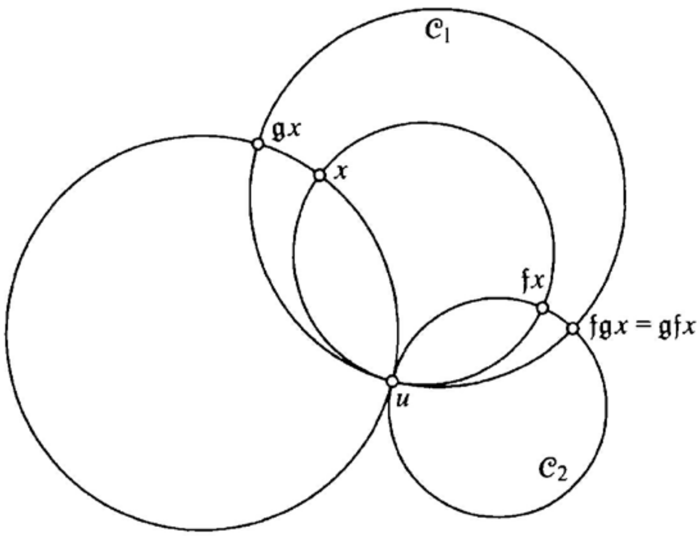



Figure 4.1

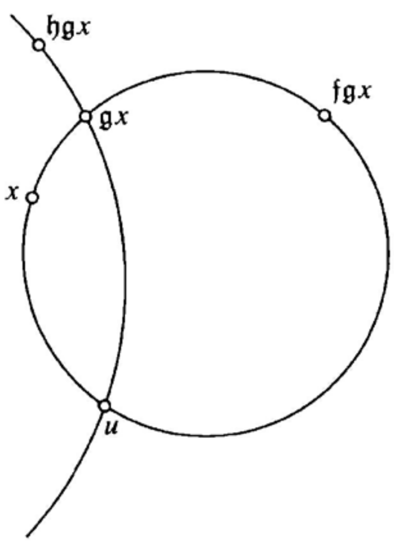



Figure 4.2

permutable in consequence of the case already dealt with and it follows that

$$\mathfrak{f}\mathfrak{g} = \mathfrak{f}\mathfrak{h}^{-1} \cdot \mathfrak{h}\mathfrak{g} = \mathfrak{h}\mathfrak{g}\mathfrak{f}\mathfrak{h}^{-1} = \mathfrak{g}\mathfrak{f}.$$

That two direct non-parabolic transformations $\mathfrak{f}$ and $\mathfrak{g}$ with the same invariant points are permutable, is inferred from the fact that in consequence of (3) the multipliers of $\mathfrak{fg}$ and $\mathfrak{gf}$ are both equal to the product of the multipliers of $\mathfrak{f}$ and $\mathfrak{g}$ and that a direct transformation is uniquely determined by its multiplier and its invariant points. – That two involutions $\mathfrak{f}$ and $\mathfrak{g}$ with respect to two pairs of harmonic points are permutable can be seen as follows: Since $\mathfrak{fg}$ interchanges the invariant points of $\mathfrak{g}$, $\mathfrak{fg}$ is involutory, hence $\mathfrak{fgfg} = 1$. In virtue of $\mathfrak{f} = \mathfrak{f}^{-1}$, $\mathfrak{g} = \mathfrak{g}^{-1}$ this can be written $\mathfrak{fgf}^{-1}\mathfrak{g}^{-1} = 1$, thus $\mathfrak{fg} = \mathfrak{gf}$.

As far as reversed transformations are concerned, only inversions with respect to circles are taken into account. If the inversion with respect to a circle $\mathfrak{C}$ is to be permutable with a direct transformation, the latter must map $\mathfrak{C}$ onto itself, since $\mathfrak{C}$ consists of invariant points of the inversion. This necessary condition is also sufficient for the permutableness, since the direct transformation then carries any two points which are inverse with respect to $\mathfrak{C}$ into two points which are also inverse with respect to $\mathfrak{C}$. – The inversions with respect to two different circles $\mathfrak{C}$ and $\mathfrak{C}'$ can only be permutable if the circle $\mathfrak{C}'$ which consists of the invariant points of the second inversion, is reproduced by the inversion with respect to $\mathfrak{C}$, i.e. if $\mathfrak{C}$ and $\mathfrak{C}'$ are orthogonal. On the other hand, this is sufficient; for the inversions with respect to $\mathfrak{C}$ and $\mathfrak{C}'$ are involutory, and their product is the involution with respect to the pair of intersection points of $\mathfrak{C}$ and $\mathfrak{C}'$, hence also involutory. – In all, the following result is obtained, the indicated conditions being necessary and sufficient:

*Two direct Möbius transformations are permutable if they have their invariant points in common, or if they are involutions with respect to harmonic pairs of points. A direct transformation is permutable with the inversion with respect to a circle if it maps that circle onto itself. The inversions with respect to two circles are permutable if the circles are mutually orthogonal.* □

## §5 Complex distance of two pairs of points

**5.1 Definition.** Let any two pairs of points $x$, $y$ and $x'$, $y'$ be given in this order, the points of the single pairs likewise being given in the indicated order. It is first assumed that the two pairs have no point in common, whereas coincidence of the points of the single pairs is not excluded. It will first be shown that there exists exactly one pair of points $u$, $v$ which is harmonic with both of the given pairs; in the case of coincidences harmonicity is taken in the generalized sense indicated at the end of §2. If $x \neq y$, there exists exactly one direct Möbius transformation carrying $x$ into $y$, $y$ into $x$ and $x'$ into $y'$. Since that transformation interchanges two points, it is the involution with respect to a certain pair of points $u$, $v$ (§4.3); therefore it also carries $y'$ into $x'$. In case $x' = y'$, this point will at the same time be one of the points $u$ and $v$. – If $x = y$ but $x' \neq y'$, one can start with $x'$, $y'$ in an analogous way. – If both $x = y$ and $x' = y'$, the required solution is found by putting $x = y = u$ and $x' = y' = v$, or conversely.

In all cases $u$ and $v$ are different. Since the two pairs $x, y$ and $x', y'$ are assumed without common point, in such expressions as e.g. $(uvxx')$ no three points coincide, and the cross-ratio therefore has a meaning.

First, consider the normal case of no coincidence, thus $x, y, x', y'$ being four different points. Then

$$(uvxx') = (uvyx)(uvxx')(uvx'y') = (uvyy'),$$

since the two factors added in the intermediate term assume the value $-1$ in consequence of harmonicity. This expression is the multiplier (4.3) of the Möbius transformation with $u$, $v$ as invariant points which carries $x$ into $x'$; that it also carries $y$ into $y'$ is also evident from the fact that it must carry two points which are harmonic with $u$, $v$ into two points which again are harmonic with $u$, $v$.

The logarithm, taken with reversed sign, of this multiplier is called *the complex distance of the pair of points $x$, $y$ and $x'$, $y'$*, given in this order:

$$a = \delta + \varphi i = -\log(uvxx') = -\log(uvyy'). \tag{1}$$

After a fixed choice of the notation $u$ and $v$ has been made, $a$ is uniquely determined except for multiples of $2\pi i$. If the pairs or the points $u$ and $v$ are interchanged, the sign of $a$ is reversed. From the relations

$$\begin{aligned}(uvxy') &= (uvxy)(uvyy') = -(uvyy')\\ (uvyx') &= (uvyx)(uvxx') = -(uvxx')\end{aligned} \tag{2}$$

it is seen that $a$ is increased by $\pi i$ if the points of one pair are interchanged.

The values of $\delta$ and $\varphi$ may be deduced from the equation

$$\log(uvxx') = \log|uvxx'| + i\ \mathrm{amp}(uvxx')$$

using the results of §2.2:

Draw through $x$ and $x'$ the circles $\mathcal{K}$ and $\mathcal{K}'$ of the hyperbolic pencil with $u$ and $v$ as zero-circles and the circular arcs $\mathcal{B}$ and $\mathcal{B}'$ of the conjugate elliptic pencil, and draw an arbitrary arc $\mathcal{B}_0$ of the latter cutting $\mathcal{K}$ and $\mathcal{K}'$ at points $x_0$ and $x_0'$ (Fig. 5.1). Then

$$\delta = -\log(uvx_0x_0'),$$

and $\varphi$ is the angle from $\mathcal{B}$ to $\mathcal{B}'$ when measured in $u$ (not in $v$ as in §2.2 because of the reversed sign).

If coincidence takes place in at least one pair, $x = y$ say, this point is at the same time $u$ or $v$, and there is no transformation with $u$ and $v$ as invariant points carrying the first pair into the second. But the cross-ratios in (1) exist and evidently take the values 0 or $\infty$. Accordingly one has to put $a = 0$ or $a = \infty$.

The definition of complex distance has to be extended to the case, hitherto excluded, where the two pairs have at least one point in common. Let for instance $x = x'$. If

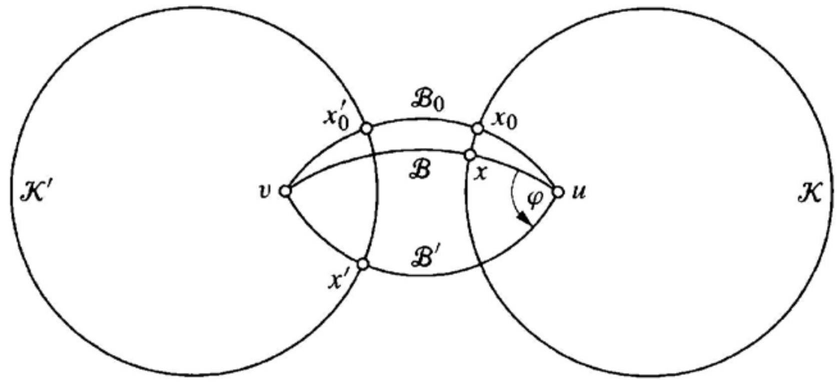



Figure 5.1

both $u$ and $v$ are chosen in this point the condition of common harmonicity (in the generalized sense of §2) is maintained; this is true whether $y$ and $y'$ coincide or not. In this case the symbol $(uvxx')$ is devoid of meaning, but another cross-ratio in (1), $(uvyy') = (uuyy')$, exists provided $y$ and $y'$ do not coincide with $x = x' = u = v$ and it takes the value 1. Thus $a = 0$: If the two first points or the two second points of the pairs coincide, the distance is zero. – If the first point of one pair coincides with the second of the other, $x = y'$ or $y = x'$, and if again both $u$ and $v$ are chosen in this common point in order to fulfill the condition of common harmonicity, both cross-ratios in (1) are devoid of meaning. This case may be treated by the remark that it reduces to the former case by the interchange of the points of one pair. According to a previous statement the distance then has to be $a = \pi i$. These are the only cases in which the distances 0 and $\pi i$ occur. In both cases they may be justified by considerations of continuity.

**5.2 Relations between distances.** Let five pairs of points $x_1, y_1 : x_2, y_2 : x_3, y_3 : x_4, y_4 : x_5, y_5 :$ be given. The order of the two points in the single pairs is as indicated. As to the succession of the pairs, only their *cyclical* order matters. This order is indicated by the subscripts these being, in the sequel, only taken into account modulo 5. Let furthermore any two neighbouring pairs be harmonic; this is expressed by the equation

$$(x_\nu y_\nu x_{\nu+1} y_{\nu+1}) = -1 \tag{3}$$

or

$$2(x_\nu y_\nu + x_{\nu+1} y_{\nu+1}) - (x_\nu + y_\nu)(x_{\nu+1} + y_{\nu+1}) = 0, \tag{4}$$

$\nu$ ranging over all values modulo 5. Now, the complex distance of the two pairs next to the pair $x_\nu, y_\nu$ is, according to Section 1

$$a_\nu = -\log(x_\nu y_\nu x_{\nu-1} x_{\nu+1}) = -\log(x_\nu y_\nu y_{\nu-1} y_{\nu+1}).$$

The relations governing these five distances $a_\nu$ will now be established. Since cross-ratios and, therefore, the relations looked for are invariant under Möbius transformations, it can be assumed that $x_3 = 0,\ \ y_3 = \infty$.

From (3), taken for $\nu = 2$ and $\nu = 3$, it is inferred that

$$x_2 + y_2 = 0, \quad x_4 + y_4 = 0.$$

This together with (4), taken for $\nu = 1$ and $\nu = 4$ yields

$$x_1 y_1 = x_2^2, \quad x_5 y_5 = x_4^2. \tag{5}$$

Now, from the definition of $a_\nu$,

$$e^{-a_3} = (x_3 y_3 x_2 x_4) = (0\,\infty\, x_2 x_4) = \frac{x_2}{x_4},$$

hence from (4), taken for $\nu = 5$, and (5)

$$\cosh a_3 = \frac{e^{a_3} + e^{-a_3}}{2} = \frac{x_2^2 + x_4^2}{2x_2x_4} = \frac{x_1y_1 + x_5y_5}{2x_2x_4} = \frac{x_1 + y_1}{2x_2} \cdot \frac{x_5 + y_5}{2x_4}.$$

On the other hand

$$e^{-a_2} = (x_2 y_2 x_1 x_3) = (x_2\ \ {-x_2}\, x_1\, 0) = \frac{x_2 - x_1}{x_2 + x_1},$$

hence from (5)

$$\coth a_2 = \frac{e^{a_2} + e^{-a_2}}{e^{a_2} - e^{-a_2}} = \frac{x_1^2 + x_2^2}{2x_1x_2} = \frac{x_1 + y_1}{2x_2}.$$

Likewise

$$e^{-a_4} = (x_4 y_4 x_3 x_5) = (x_4\ \ {-x_4}\, 0\, x_5) = \frac{x_4 + x_5}{x_4 - x_5},$$

hence from (5)

$$\coth a_4 = \frac{e^{a_4} + e^{-a_4}}{e^{a_4} - e^{-a_4}} = -\frac{x_4^2 + x_5^2}{2x_4x_5} = -\frac{x_5 + y_5}{2x_4}.$$

From these formulae one gets

$$\cosh a_3 = -\coth a_2 \coth a_4.$$

Since all distances are defined by cross-ratios, the result is independent of the above special choice of $x_3$, $y_3$ hence one gets generally by permutation of subscripts

$$\cosh a_\nu = -\coth a_{\nu-1} \coth a_{\nu+1} \quad (\nu \bmod 5). \tag{6}$$