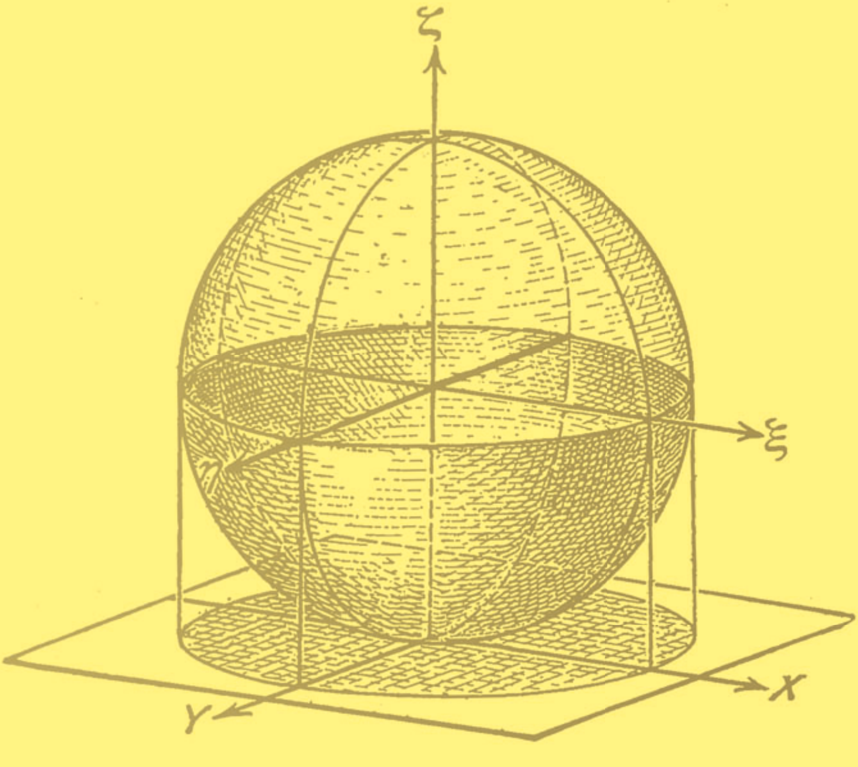
Raymond O. Wells, Jr.

Differential and Complex Geometry: Origins, Abstractions and Embeddings



Springer

Differential and Complex Geometry: Origins, Abstractions and Embeddings

Raymond O. Wells, Jr.

Differential and Complex Geometry: Origins, Abstractions and Embeddings

Springer

Raymond O. Wells, Jr.
University of Colorado Boulder
Boulder, CO
USA

and

Jacobs University Bremen
Bremen
Germany

ISBN 978-3-319-58183-5 ISBN 978-3-319-58184-2 (eBook)
DOI 10.1007/978-3-319-58184-2

Library of Congress Control Number: 2017943176

Mathematics Subject Classification (2010): 01-02, 30-02, 32-02, 14-02, 51-02, 53-02, 55-02

Cover illustration reproduced from 'Vorlesungen über Nicht-Euklidische Geometrie' by Felix Klein (Volume 26 of the series Die Grundlehren der Mathematischen Wissenschaften), 1967, p. 300. With permission of Springer Nature.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my friend, Howard Resnikoff

Preface

About ten years ago, I had the idea of writing up a survey of the major embedding theorems of the twentieth century. This book represents the culmination of this idea, and I'm quite happy to be able to finally publish it after all this time. The embedding theorems represent important ideas in the modern fields of differential topology, differential geometry, complex manifold theory, and the general theory of functions of several complex variables, as well as the overall concept of manifolds in general. I thought it would be useful to review the origins of these various concepts as a way of hoping to give a deeper understanding of the theorems themselves.

Consequently, I spent a fair amount of time these past years looking at a number of contributions by mathematicians during the seventeenth through the nineteenth centuries, where almost all of these concepts first appeared and then developed. In my book, I have tried to give the reader some sense of the language and understanding of these earlier mathematicians as they gave voice to the many issues at hand. For instance, the developments of projective geometry and intrinsic differential geometry both evolved at the same time in the first half of the nineteenth century, but in reading the literature of the time, it seems as if they were hardly aware of each other. Only in the last half of the nineteenth century did these seemingly disparate sets of ideas come to be part of a mathematical whole.

I would not have been able to peruse these papers and books from these earlier times had it not been for the Internet and the fact that the great libraries of the world put time and effort into digitizing their collections. I am very thankful that these ideas can be so readily shared today.

I have had the support of three academic institutions over the past decades, where it has been my privilege to hold various academic appointments, and I want to thank them all for their continued support over the years: Rice University in Houston; Jacobs University in Bremen, Germany; and the University of Colorado in Boulder, Colorado, where I now live.

Springer is the publisher of two of my earlier books, and I am very happy that they are bringing this new work of mine to the public. I want to thank, in particular, Rémi Lodh, who encouraged me and helped bring this book to fruition.

The comments of his reviewers were very helpful to me. Anne-Kathrin Birchley-Brun, also in the London Springer office, has been very helpful in the process of managing the digital files and ushering them into the production process.

I want to thank Ina Mette, formerly of Springer and now an editor for the American Mathematical Society, for her encouragement for this project over many years now.

I have dedicated this book to my very close friend, Howard Resnikoff. He has been an inspiration for me for over fifty years, and we have shared many things together. His reading of various drafts of this book and his encouraging words have been very important to me.

Finally, I want to thank my wife, Rena, for her continuous support in so many ways. In particular, she read a final draft and her comments and editorial pen were so very useful, as always.

Boulder, CO, USA
March 2017

Raymond O. Wells, Jr.

Contents

Introduction

In 1913, Hermann Weyl wrote a very influential book that one can perceive of as a bridge between the geometry of earlier centuries and the geometry that evolved in the twentieth century. More specifically, using the newly discovered theories of set theory and point set topology, he created a theory of differentiable and complex manifolds, in particular, in the case of Riemann surfaces. In 1936, Hassler Whitney ushered in a new era of geometry when he formulated and proved the first embedding theorem for differentiable manifolds. Namely, he showed that any differentiable manifold can be embedded as a closed submanifold of a higher-dimensional Euclidean space.

This was followed up over the next two decades by various mathematicians who provided similar characterizations: real-analytic submanifolds of Euclidean space (Grauert 1958), differentiable submanifolds with a Riemannian metric induced from the ambient Euclidean space (Nash 1956), complex submanifolds of complex Euclidean space (Remmert, Narasimhan, Bishop, 1956–1961), and complex submanifolds of complex projective space (Kodaira 1954). All of these were embedding theorems of one sort or another. Their formulations and proofs depended on a variety of mathematical ideas, many of which had also evolved in the twentieth century.

This book outlines a survey of roughly three centuries of mathematical work concerned with differential and complex geometry, culminating in the twentieth-century embedding theorems. The book is divided into four parts, which are described below in more detail. In Parts I–III, we provide an overview of many of the geometric ideas that play an important role in the twentieth-century embedding theorems and which arose in various guises in the previous three centuries, and Part IV describes the embedding theorems in some detail. Our major source for this survey of mathematical ideas has been to look in some detail at the original papers and monographs of the principal authors whose works are the cornerstones of these developments. We have tried to look at the writings of these authors in the context of the mathematical knowledge known at the time.

Part I looks at the way the geometry of curves and surfaces in two- and three-dimensional Euclidean space began to interact with the simultaneously evolving theories of analysis that grew out of the late seventeenth century with the discoveries of differential and integral calculus. Here, the notion of tangent vectors and tangent spaces, first-order approximations to curves and surfaces, and curvature, measuring how far curves and surfaces deviated from straight lines and planes, all evolved in a systematic fashion. This became the essence of what we now call extrinsic differential geometry, where the notion of the distance between points is inherited from the ambient Euclidean space.

Part II describes two parallel theories that evolved in the nineteenth century. The first was the discovery of intrinsic differential geometry that has become the foundation of contemporary differential geometry (Gauss, Riemann). The second was the creation of projective geometry, a generalization of classical Euclidean geometry that transcended the usual two- and three-dimensional space by asking new types of geometric questions and introducing points at infinity (Monge, Poncelet, and many others). This evolved into our contemporary notion of projective space and became the basis for much of algebraic geometry in the twentieth century.

Part III is an outline of the origins of what became known in the twentieth century as complex geometry. It has its roots in the generalizations of trigonometric functions and their properties: Euler's addition theorems for elliptic integrals (Legendre, Abel), elliptic functions, and their generalizations, Abelian functions (Abel, Jacobi, Weierstrass, Riemann). Moreover, the development of function theory over many decades, starting with the pioneering work of Cauchy in the 1820s (Riemann, Weierstrass), and in the innovative work of Riemann in his creation of the theory of Riemann surfaces in the mid-nineteenth century, led to the developments of algebraic topology and the theory of manifolds in general at the end of the nineteenth century (Riemann, Betti, Poincaré, Weyl). The work of Klein and Lie on transformation groups in the latter half of the century was a very important contribution for modern geometry as well.

Part IV of this book outlines in some detail the major twentieth-century embedding theorems. They are all philosophically related: A manifold of some sort can be embedded as a submanifold in some higher-dimensional Euclidean space or projective space, and the embedding characterizes all such submanifolds. Technically, they involve a broad range of mathematical tools and, for the most part, solved problems that had been formulated earlier and involved quite technical and often very difficult proofs. More specifically, they involve differentiable and real-analytic manifolds, arising from the work in Part II, and they involve complex manifolds, arising from the work in Part III.

Each Part of this book has its own more detailed introduction to the material in that set of chapters. Here, we have given only a very brief overview of the whole book.

This survey over several centuries tries to show how these various strands of mathematical thought have culminated in the powerful embedding theorems of the mid-twentieth century. Of course, there were many other areas of development of geometric ideas in the same time period which are not included in our survey, but we feel we have chosen a coherent family of ideas that have contributed greatly to our mathematical culture in the twentieth century.

Part I
Geometry in the Age of Enlightenment

Introduction

The *Age of Enlightenment* is a term that refers to a time of dramatic changes in Western society in the arts, in science, in political thinking, and, in particular, in philosophical discourse. It is generally recognized as being the period from the mid-seventeenth century to the latter part of the eighteenth century. It was a successor to the Renaissance and Reformation periods and was followed by what is termed the Romanticism of the nineteenth century. In his book *A History of Western Philosophy* [205], Bertrand Russell (1872–1970) gives a very lucid description of this time period in intellectual history, especially in Book III, Chapter VI–Chapter XVII. He singles out René Descartes as being the founder of the era of new philosophy in 1637 and describes other philosophers who made significant contributions to mathematics, such as Newton and Leibniz. This time of intellectual fervor also included literature (e.g., Voltaire), music, and the world of visual arts. One of the most significant developments was perhaps in the political world: Here, the absolutism of the church and of the monarchies were questioned by the political philosophers of this era, ushering in the Glorious Revolution in England (1689), the American Revolution (1776), and the bloody French Revolution (1789). All of these were culminations of this Age of Enlightenment, which permanently changed the shape of Western civilization from the absolutism of the Middle Ages. A most important development of this time was the rise of science as it began to play an increasingly important role in the world at large, along with the technological advances which accompanied it. Russell describes this advance in science very succinctly in his book.

Mathematics experienced, as a part of this intellectual development, exciting growth with numerous new sets of ideas. In this Part I of our book, we outline some of the important developments and new ideas in geometry which were a part of this era. There were, of course, many important mathematical developments during this period, such as in analysis, number theory, algebra, applied mathematics. In Parts II and III, we have described a number of the very innovative geometric ideas that arose in the fertile nineteenth century and which extended beyond the usual study of geometry in two- and three-dimensional space that was inspired by the mathematicians of the Greek era. However, in the Age of Enlightenment there were crucial new discoveries concerning curves and surfaces in the plane and in ordinary three space which became crucial building blocks for nineteenth-century geometry.

Between the time of the Greeks (which lasted about one millennium from *c.* 600 BCE until *c.* 400 CE) and the rise of mathematical thought in Western Europe, in the seventeenth and eighteenth centuries, there were numerous mathematical developments in the Arabic and Indian cultures, primarily in arithmetic and algebra. One of the most significant accomplishments of the Arabic world after the fall of the Roman empire and preceding the time of the Renaissance was the actual preservation of a substantive amount of the accomplishments of the Greek mathematicians. Pappus, whose work plays an important role in this context, came toward the very end of Greek mathematical culture in the fourth-century CE. For more than 1000 years, mainly from the time of Pappus to the extraordinarily original work of Descartes in the seventeenth century, geometry seemed to be at a standstill.

There are several major areas of contemporary geometry which have their roots in the Age of Enlightenment. *Projective geometry* was primarily developed in the nineteenth century. Its roots stemmed from works of Pascal from 1639 and from Desargues in 1642, and their contributions were rediscovered by the nineteenth-century geometers. *Algebraic topology* was hinted at in a letter of Leibniz to Huygens in 1679, which was cited by Euler in his famous paper on the Königsberg bridges. Leibniz and then Euler used the phrase "analysis situs" to describe a relationship between geometric and algebraic quantities. The final paper worth mentioning from this period, and which is very important for algebraic topology, was Euler's singular paper [62] from 1752, which described for the first time what has become known as the *Euler characteristic* for a surface.

All of these ideas from the seventeenth and eighteenth century are described in their context in Parts II and III.

We will consider in this Part I two important areas of geometry developed in this time period, namely *algebraic geometry* and *differential geometry*, both of which had substantive growth in the eighteenth century.

Section 1.2 describes some of the work of Pappus which played an important role in these geometric developments. We then turn to Descartes, whose work includes a solution to a problem posed by Pappus, and included his celebrated coordinate geometry, which transformed mathematics in so many ways. Both Descartes and Fermat were able to successfully classify the algebraic curves of degree two, as described in this section. Fermat's work on number theory is much more well known, but he also contributed significantly to the study of curves in the plane.

Following up on the work of Descartes and Fermat, Newton gave a detailed classification of real-algebraic curves of degree three in the plane. Some 50 years later, Euler gave a similar classification. Here, analysis played an important role in both the work of Newton and the work of Euler, in particular in their use of infinite series in their descriptions of asymptotic behavior. Section 1.3 gives a brief overview of this initial work of Newton.

In Chap. 2, we note how the study of transcendental functions led to many geometric objects which were not necessarily defined algebraically (this was pointed out quite explicitly in Euler's *Introductio* in 1748) [61]. Earlier, Newton first formulated the notion of curvature of a curve in the plane in terms of calculus, which followed up on ideas of Apollonius who looked at curvature of conic sections and the more general work of Huygens from the seventeenth century. These ideas are all discussed in Sect. 2.2. The curvature of curves in space was initiated by Clairaut in the early eighteenth century and brought into its final form in the mid-nineteenth century by Frenet and Serret (Sect. 2.3). The final topic in this chapter considers the work of Euler from 1767, who studied the curvature of a surface by analyzing the curvature of the curves arising from intersections of planes normal to the surface with the surface itself.

The ideas discussed above all play a major role in modern mathematics. For instance, the classification of algebraic curves in the plane of degree two and three is an important predecessor of what has become a major theme of contemporary geometry: to classify geometric objects. Moduli of Riemann surfaces and classification of topological and differentiable manifolds of various dimensions, including the famous Poincaré conjecture, Thurston's classification conjecture for three-manifolds, Kodaira and Spencer's classification of compact two-dimensional complex manifolds, and many other examples, are all instances of the classification of geometric objects.

The work on differential geometry for curves and surfaces in three space that we describe here was an important prelude to the work of Gauss and Riemann on curvature of differentiable manifolds in the nineteenth century, which we describe in Part II. This has developed into the very rich field of differential geometry of the twentieth century. We note that, for instance, Grigori Perelman's recent solution of the three-dimensional Poincaré conjecture used as a tool the full power of differential geometry to solve this topological problem.

Chapter 1
Algebraic Geometry

1.1 Introduction

One of the last major figures in Greek mathematics was Pappus of Alexandria (*ca.* 290 CE–350 CE), who published a major work entitled the *Collection*. Book I and the introduction to Book VII of this work are missing, but that which has been preserved gives a good survey of many mathematical discoveries of his predecessors, the originals of which have been lost, and, in addition, he contributed significantly to solutions of a number of geometric and arithmetic problems (see e.g., Boyer [25], pp. 205–213, for a summary of the important contributions in the *Collection*). We want to single out one particular contribution that has played such an important role in the history of geometry. This is now referred to as the *problem of Pappus* and was described in the beginning of Book VII of the *Collection*. This problem was treated by Euclid, Apollonius and others who preceded Pappus, and, as will be seen later, Descartes.

Apollonius (*c.* 262 BCE–*c.* 190 BCE), one of the great geometers of the Greek era of mathematics, wrote in the preface to his *Conics* [98] the following (using the translation in Boyer [25], p. 167):

> The third book contains many remarkable theorems useful for the synthesis of solid loci and determinations of limits; the most and prettiest of these theorems are new and, when I had discovered them, I observed that Euclid had not worked out the synthesis of the locus with respect to three and four lines, but only a chance portion of it and that not successfully: for it was not possible that the synthesis could have been completed without my additional discoveries.

Here Apollonius was referring to his discoveries concerning conic sections, which transcended substantially the work of the earlier Greek mathematicians. The "synthesis of the locus with respect to three and four lines" is a special case of what has come to be called the problem of Pappus, which is discussed more explicitly in the following paragraphs.

We first formulate the original version, which was solved by Apollonius, and then later its generalizations. The *Problem of Pappus*: given three or four lines in

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_1

the Euclidean plane, find the locus of points such that the square of the distance to one line (in the three-line case) is proportional to the product of the distances to the remaining two lines. In the case of four lines, one asks for the locus of points with the property that the product of the distances to two of the lines is proportional to the product of the distances to the remaining two lines.[1] In all cases the distance to a given line is measured at a given angle to the given line (thus the given data is the set of lines, the set of angles and the proportionality factor). Apollonius shows that the resulting curves are indeed given by conic sections, which, of course, is the primary topic of his book. He implies in the quote above that Euclid did not have the detailed results needed concerning conic sections in order to solve this problem.

Using the language and ideas of analytic geometry, one can easily verify Apollonius's result (see Boyer [25], pp. 167–168). If, in the case of three lines given by the equations

$$A_1x + B_1y + C_1 = 0,$$
$$A_2x + B_2y + C_2 = 0,$$
$$A_3x + B_3y + C_3 = 0,$$

and if the angles used for measuring distance are given by θ_1, θ_2, and θ_3, then the locus is given as the set of points (x, y) satisfying:

$$\frac{(A_1x + B_1y + C_1)^2}{(A_1^2 + B_1^2)\sin^2\theta_1} = K\frac{(A_2x + B_2y + C_2)}{\sqrt{A_2^2 + B_2^2}\sin\theta_2} \cdot \frac{(A_3x + B_3y + C_3)}{\sqrt{A_3^2 + B_3^2}\sin\theta_3}.$$

Since the locus is the solution of a quadratic equation in the plane, it follows that it is a conic section, which is what Apollonius had discovered using his methodology.

The general problem of Pappus is to be given an arbitrary number of lines and angles and to ask the same question. Here is a quote from Pappus concerning the more general problem. First he notes (following Boyer [25], p. 209) that for six lines, the locus can be considered as a solid which is in fixed ratio to another solid (here "solid" refers to products of three lengths, i.e., homogeneous polynomial terms of degree three, using modern language). However, higher-degree terms were a mystery to him, as (quoting Pappus)

> there is not anything contained by more than three dimensions

and, he continued,

> men a little before our time have allowed themselves to interpret such things, signifying nothing at all comprehensible, speaking of the product of the content of such and such lines by the square of this or the content of those. These things might however be stated and shown generally by means of compounded proportions.

---

[1] This set of problems is similar in spirit to the characterization of a circle being the locus of all points (in the plane) whose distance to a given point is constant, or an ellipse being the set of all points such that the sum of the distances to two distinct points is constant.

Pappus did not study the higher-degree case (higher than six), but he did make the important observation that the loci were curves in the plane. As Boyer observes, Pappus was a geometer and Diophantus, a contemporary, was an algebraist (who did consider higher powers and who had the notation to tackle the higher-degree problems), but it required a mathematician who was familiar with both algebra and geometry to make the next step, and that turned out to be Descartes, some 1300 years later.

1.2 Descartes and Fermat

René Descartes (1596–1650) published a slim volume in 1637 entitled *La Géométrie* [55], which initially was an appendix to a longer work in philosophy but was also published independently. Descartes's work turned out to be revolutionary, and when the next generation of mathematicians began to write general texts concerning what today is called analytic geometry the impact of his work spread throughout the mathematical world of Europe and became fully developed in the eighteenth century. Until Descartes, and actually long after as well, Euclid's *Elements* were definitive on almost all things concerning geometry. Descartes's new view of geometry was very important, but Euclid's ideas were still very valid. Only with non-Euclidean geometry in the nineteenth century was Euclid challenged in a fundamental way. A very brief but succinct survey of Descartes's *Géométrie* is given by Serfati [213].

Descartes is most well known to mathematicians for having discovered *analytic geometry* or probably more appropriately named *coordinate geometry*, which has been taught in twentieth-century high schools and on up to the present day around the world. The label "analytic geometry" as applied to Descartes is slightly a misnomer. What Descartes showed was how common problems of geometry as described by Greek geometers could be described by using algebraic equations and conversely. The most important historical example of this is Descartes's theorem that solutions of algebraic equations of degree two in two variables corresponds precisely to the conic sections studied byApollonius and others. What was much more important, aside from the new coordinate system point of view, was that he *defined* geometric objects to be solutions of algebraic equations. In particular, he considered algebraic equations in two variables of arbitrary degree, which became the nucleus of *algebraic geometry*, and which, in its modern form as developed in the nineteenth and twentieth centuries, is definitely *not* taught in high schools around the world.

However, the way analytic geometry is taught today involves not only algebraic functions, but also the standard transcendental functions, $\sin x$, e^x, etc., as well, and students also learn about the curves that these functions can represent in the plane. This is something that Descartes absolutely rejected. He had learned from Greek authors that there were three types of curves studied by mathematicians: *plane curves*, i.e., curves that could be described with a straight edge and compass in the plane; *solid curves*, that is curves that could be described in three-space by intersections of simple surfaces with a plane, the simplest being the full family of conic sections, i.e., intersections of a cone with a plane; and the third category was

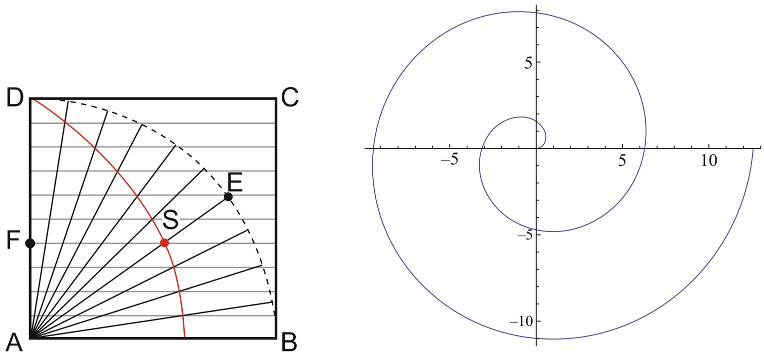


Fig. 1.1 The quadratrix (trisectrix) curve of Hippias of Elis and the spiral of Archimedes. In polar coordinates, $r = \frac{2a\theta}{\pi \sin\theta}$, $r = a\theta$, respectively

linear curves, i.e., everything else. This last category included the quadratrix (often called the trisectrix), the spiral of Archimedes, and other transcendental curves (to use modern language). These are illustrated in Fig. 1.1.

Of course, this terminology (which one finds in Euclid, Apollonius and Pappus) would be totally confusing today. Descartes pleads for his reader to have only two categories of curves: *geometric* and *mechanical*. His definition of geometric was simply anything described by algebra, i.e., *algebraic* curves (and surfaces, which he did not really address in his book), and *mechanical* being all others. Hence he extended the Greeks planar and solid curves to include curves of arbitrary degree. On the other hand, he excluded from the study of geometry the mechanical curves which include the quadratrix and spiral. The conchoid is an example of an algebraic curve of higher degree

$$(x - a)^2(x^2 - y^2) = b^2x^2,$$

and this curve was used for cube duplication and angle trisection problems. The squaring of the circle required trigonometric functions (for instance, using the quadratrix). Today one uses the terminology of Leibniz: *algebraic curves* and *transcendental curves*.

The main reason Descartes made this distinction was that he thought that one could always calculate solutions to algebraic equations but not solutions of transcendental equations. He was familiar with the formulas for solutions of the third- and fourth-degree equations of Cardano, and it seems that he assumed that such formulas were true in general. Descartes spent almost all of Book III (the third and final chapter of *La Géométrie*) discussing the explicit solution of equations, and in particular he developed a theory of roots of polynomial equations in one variable. But in any event, if he had a polynomial equation of degree d of the form $P(x, y) = 0$, he argued that if you fixed a particular value of the variable y then you obtained a polynomial in one variable which could always be solved (he believed the fundamental theorem of algebra was indeed true, but he also believed roots could be found simply by extracting roots, which Abel and Galois later showed to be false, in general). He then argued that one could not compute the values of the function whose graph is the quadratrix curve for arbitrary values of the variable in question, but only for certain special values. Fundamentally, this is equivalent to the fact that one can compute by elementary geometry special values of the trigonometric functions, but not all of them. This was beyond the scope of Descartes at the time, and for this reason he rejected the study of transcendental functions as an object of study in geometry. In the eighteenth century this would change radically. Also Descartes excluded solutions of equations which involved complex numbers with nonzero imaginary part, and said simply there were no solutions in those cases. It would be two more centuries before mathematicians became comfortable with complex numbers (we discuss this in some detail in Chap. 6).

Descartes had strong opinions on what was true (or worthy of study) and what was not. One more example, which is of historic importance, concerns arc length. In the development of trigonometric functions, arc length is a critical ingredient (the relation between the length of an arc to the length of the chord subtended by it is how trigonometry was originally introduced, and the sine and cosine functions are simply modern variations of this). Descartes was convinced that for no algebraic curve (e.g., the circle) could one ever find precisely the length of the arc in terms of the length of the chord (this clearly relates to the difficulty of computing π). As he put it on p. 32 of *La Géométrie*,

> ...car encore qu'on n'y puisse recevoir aucunes lignes qui semblent à des cordes, c'est-à-dire qui deviennent tantôt droites et tantôt courbes, à cause que la proportion qui est entre les droites et les courbes n'étant pas connue, et même, je crois, ne le pouvant être par les hommes, on ne pourroit rien conclure de là qui fût exact et assuré.[2]

What Descartes did do, and it seems to have been a major part of the inspiration for writing his book, was to give a new and extensive solution to the Problem of Pappus (both Books I and II of Descartes's book are devoted to this topic, among other things). He showed that for an arbitrary number of straight lines and associated angles in the plane the locus of points such that the products of the distances to half the set of lines is proportional to the products of the distances to the other half of the set of lines, all distances being measured at the given angles, is an algebraic curve in the plane.[3] He computes that, for the classical case of three or four lines (the original problem solved byApollonius), the curve is an algebraic curve of degree 2, for the case of 5–9 lines the curve is an algebraic curve of degree 4, and for 10–13 the curve is of degree 6, etc. He refers to the curves of degree 2, 4, and 6, etc., as curves of *genre* 1, 2, and 3, etc. Descartes considered various special cases where odd degree polynomials could appear, but he lumped them in his genre classification with the even degree cases.

[2] "... because one should not be able to consider lines (or curves) that are like strings, in that they are sometimes straight and sometimes curved, since the ratios between straight and curved lines are not known, and I believe cannot be discovered by human minds, and therefore no conclusion based upon such ratios can be accepted as rigorous and exact".

[3] For the case of an odd number of lines, one takes the distance to one of the lines twice in this proportionality.

But, in addition to showing that the solution to the problem was algebraic curves, for the classical case where he obtained algebraic curves of degree two, he showed that all of these curves were conic sections. In his proof he showed how each of the polynomials of degree two arising in this context could be put in a normal form by a suitable change of coordinates, and then he was able to use Apollonius's characterization of the conic sections in terms of suitable coordinates to determine that the solutions were indeed conic sections. The major distinction between Apollonius and Descartes in this context was that Apollonius started with a given conic section and produced coordinates which helped describe it, while Descartes started with the coordinate system and the equation and was able to put it in canonical form and identify it in a suitable manner. This work of Descartes is equivalent to Apollonius's result that any section of a skew cone is one of the three classical conic sections (see Heath's translation of Apollonius's book on conic sections [98]).

Descartes recognized that different equations could describe the same geometric curve, and he pointed out the need to find the "simplest" algebraic function that could represent a given curve. This, of course, was the key question for the classification of algebraic curves (and later higher-dimensional manifolds in a variety of categories, to use a small pun), which has been a consistent and important theme in the following centuries. Descartes classified the algebraic curves of degree two, and Newton followed up with his major work on the classification of algebraic curves of degree three [166] which will be discussed in more detail in the next section.

Pierre de Fermat (1601–1665) played an important and less recognized role in this development of geometric ideas. First, in his only published paper in his lifetime,[4] he showed that one could explicitly compute the arc length of a specific algebraic curve, which, as noted above, Descartes claimed was impossible. More precisely, he showed that for the algebraic curve $y^2 = x^3$, the semicubical parabola, the arc length could be explicitly computed. Namely, if one takes the positive branch of this curve $y = x^{\frac{3}{2}}$ on the interval $[0, b]$, then one can verify that the arc length integral is

$$\int_0^b \sqrt{1 + [y'(x)]^2}dx = \frac{8}{27}[(1 + 9b/4)^{\frac{3}{2}} - 1],$$

as any calculus student today can do (and is often asked to do!). However, Fermat, who played a major role in the development of differential and integral calculus, was not aware of the fundamental theorem of calculus or of this arc-length formula, which makes the calculation somewhat more difficult!

Moreover, Fermat also proved, independently of Descartes, that algebraic curves of degree two are conic sections (see his biography, which discusses this, among many other things [150]; see also [194]). A major difference between these two historic figures on this particular point was that Fermat expressed his work in the classical language of Euclidean geometry $\overline{OA}$, $\overline{OB}$, etc., representing the lengths

[4] See [150], p. 267 for this reference *De linearum curvarum cum lineis rectis comparatione dissertatio geometrica* (Geometrical dissertation on the comparison of curved lines with straight lines), which appears as an appendix in a book by Antoine de Lalouvère from 1660.

of the line segments O to A or O to B in the Euclidean plane. Descartes used this notation as well, but he adroitly introduced the variables x, y, and z to denote unknown quantities (lengths of segments in the problems he was considering), and symbols a, b, c, etc., from the first part of the alphabet to represent known quantities in a given computation, a practice that has been followed ever since. In this way he reduced geometric problems to algebraic problems. He was very concerned that his new way of looking at things should be well connected with classical Greek geometry. As an example in Book I of *La Géométrie* he goes to great pains to show that solutions of a quadratic equation such as

$$z^2 = az + b^2$$

can be constructed by straight edge and compass. Figure 1.2 illustrates his construction.

We want to mention one final historical note that is not that well known to the mathematical public (or public in general). One of the most important innovations in Descartes's *La Géométrie* was the invention of *exponential notation*. He used x^3, x^4, etc., freely throughout the book. In the 1886 edition we have been quoting from, the editors point out two modernizations that they introduced to make the book more readable for the nineteenth-century reader. The first was the use of the "$=$" sign instead of the symbol ∞ that Descartes used for equality, and the second, surprisingly, was the use of x^2 instead of Descartes's preferred notation xx. This certainly seems strange to a modern reader, as he used the higher-power exponential notation with no hesitation. Before this innovation of Descartes mathematicians used *different symbols* for different powers of the unknown variable x, which would make a formula like the law of exponents

$$x^{m+n} = (x^m)(x^n)$$

somewhat difficult to formulate (see, e.g., [194] and the references therein).

1.3 Newton and Euler

The development of differential and integral calculus and other ideas in analysis (e.g., the theory of infinite series) in the late seventeenth century by Isaac Newton (1642–1746) and Gottfried Wilhelm Leibniz (1646–1716) and their successors have been some of the most important developments in all of mathematics. These ideas have been well documented (see, e.g. Boyer [25], Kline [125], and other such references). We won't try to give any historical background on this important topic, as we want to concentrate on the interaction between analysis and the developments in geometry as it evolved in this Age of Enlightenment. Newton published in 1704 for the first time two pivotal works on geometry and calculus which were quite independent of each other and which were chapters in a larger book on optics [167].

LIVRE PREMIER. 5

C'est pourquoi je me contenterai ici de vous avertir que, pourvu qu'en démêlant ces équations, on ne manque point à se servir de toutes les divisions qui seront possibles, on aura infailliblement les plus simples termes auxquels la question puisse être réduite.

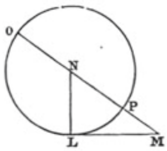
Et que si elle peut être résolue par la géométrie ordinaire, c'est-à-dire en ne se servant que de lignes droites et circulaires tracées sur une superficie plate, lorsque la dernière équation aura été entièrement démêlée, il n'y restera tout au plus qu'un carré inconnu, égal à ce qui se produit de l'addition ou soustraction de sa racine multipliée par quelque quantité connue, et de quelque autre quantité aussi connue. Quels sont les problèmes plans.

Et lors cette racine, ou ligne inconnue, se trouve aisément; car si j'ai par exemple Comment ils se résolvent.

$$z^2 = az + b^2,$$

je fais le triangle rectangle NLM (*fig. 3*), dont le côté LM est égal à b,

Fig. 3.



racine carrée de la quantité connue b^2, et l'autre LN est $\frac{1}{2}a$, la moitié de l'autre quantité connue qui étoit multipliée par z, que je suppose être la ligne inconnue; puis prolongeant MN, la base de ce triangle, jusques à O, en sorte que NO soit égale à NL, la toute OM est z, la ligne cherchée; et elle s'exprime en cette sorte :

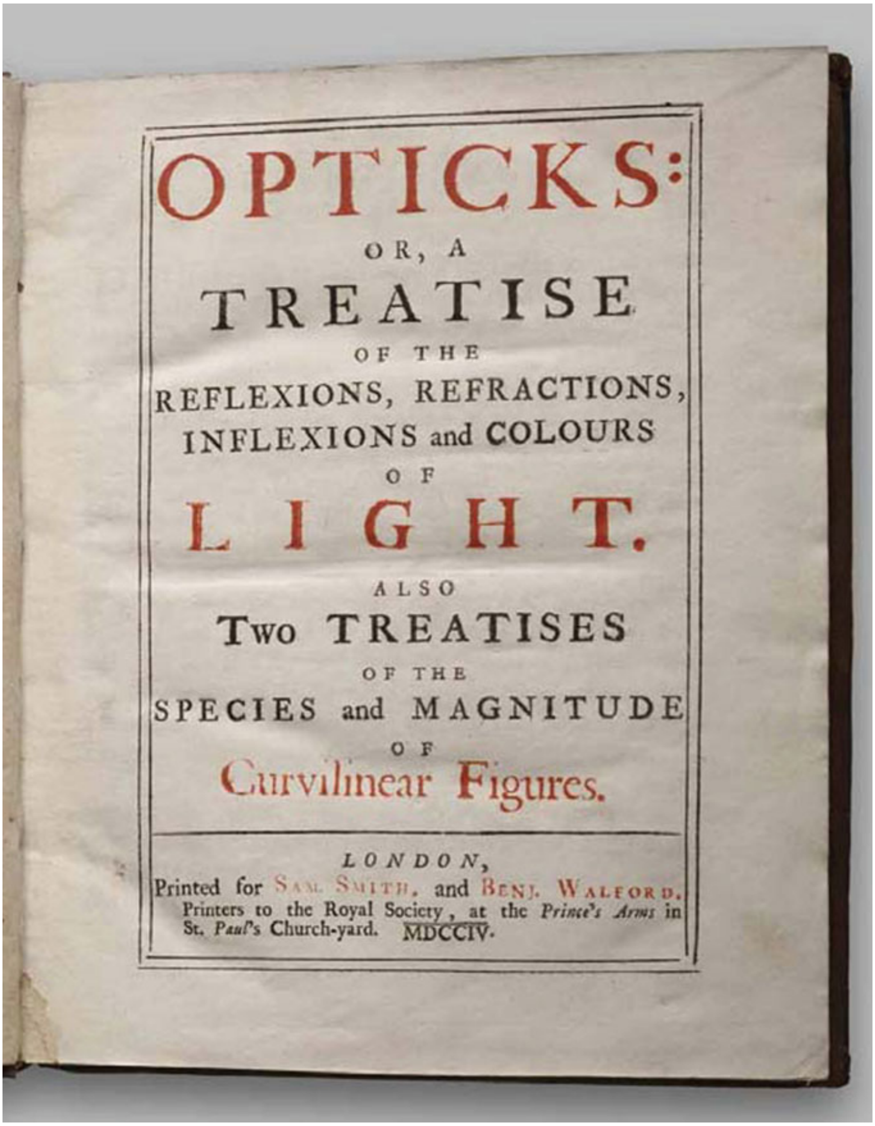
$$z = \frac{1}{2}a + \sqrt{\frac{1}{4}a^2 + b^2}.$$

Que si j'ai $y^2 = -ay + b^2$, et que y soit la quantité qu'il faut trouver, je fais le même triangle rectangle NLM, et de sa base MN j'ôte NP égale

Fig. 1.2 Page 5 of Descartes's *La Géométrie*

Figure 1.3 shows the title page of this very important work; here the two mathematical treatises are labeled "Also Two Treatises of the Species and Magnitude of Curvilinear Figures."

The first of these mathematical treatises deals with the classification of algebraic curves of degree three [166], and the second concerns itself with calculus and measuring the area under a curve [168]. Both of these mathematical chapters are written in Latin, whereas the main part of the book is written in English, and concerns

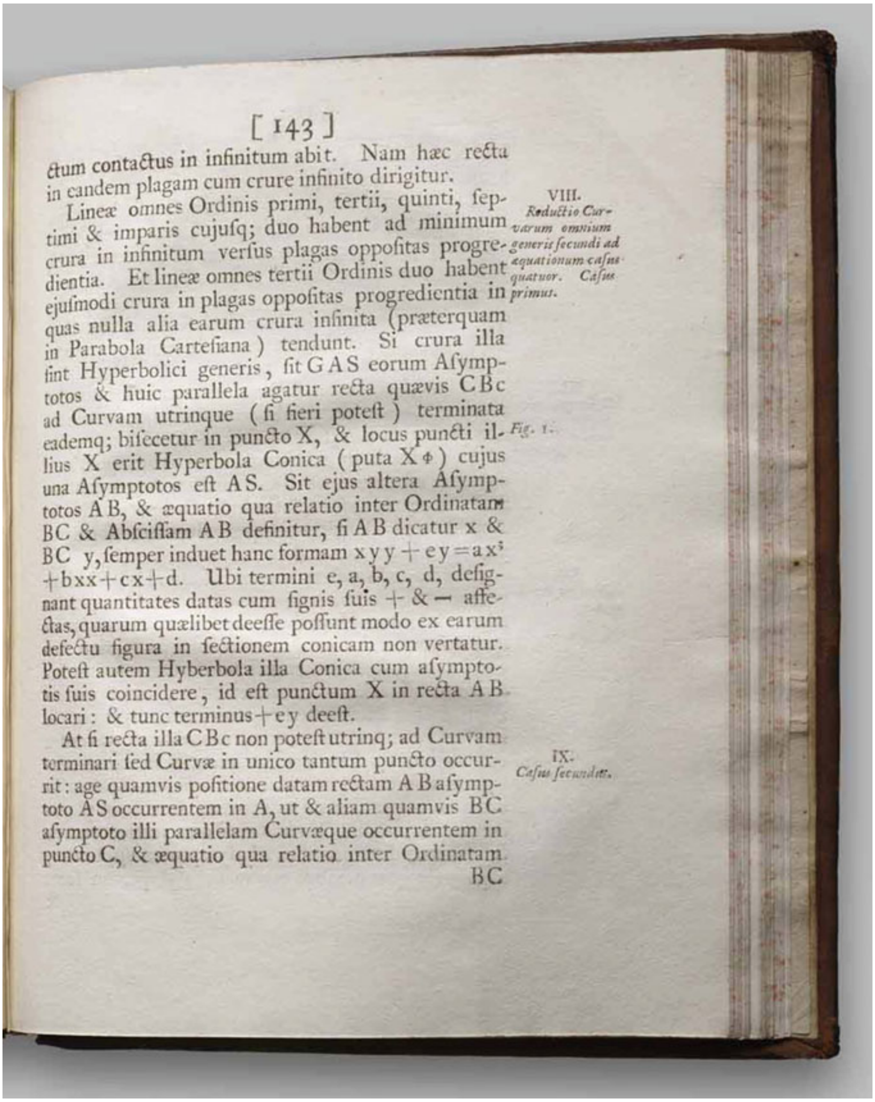


OPTICKS:
OR, A
TREATISE
OF THE
REFLEXIONS, REFRACTIONS,
INFLEXIONS and COLOURS
OF
LIGHT.
ALSO
Two TREATISES
OF THE
SPECIES and MAGNITUDE
OF
Curvilinear Figures.

LONDON,
Printed for SAM. SMITH, and BENJ. WALFORD,
Printers to the Royal Society, at the Prince's Arms in
St. Paul's Church-yard. MDCCIV.

Fig. 1.3 Title Page of Newton's *Opticks* from 1704

itself with physics, and optics in particular. Note that Newton's *Principia Mathematica* was published earlier in 1687 and the definitive third edition was published in 1726 (see the annotated copy of this third edition of Newton's most important work [170], edited by Khoyré and Cohen, which was published in 1972). This book uses the ideas of calculus in a fundamental way, even if the language used to express them is somewhat cumbersome, describing everything in terms of limiting processes of



[143]

ctum contactus in infinitum abit. Nam hæc recta in eandem plagam cum crure infinito dirigitur.

Lineæ omnes Ordinis primi, tertii, quinti, septimi & imparis cujusq; duo habent ad minimum crura in infinitum versus plagas oppositas progredientia. Et lineæ omnes tertii Ordinis duo habent ejusmodi crura in plagas oppositas progredientia in quas nulla alia earum crura infinita (præterquam in Parabola Cartesiana) tendunt. Si crura illa sint Hyperbolici generis, sit GAS eorum Asymptotos & huic parallela agatur recta quævis CBc ad Curvam utrinque (si fieri potest) terminata eademq; bisecetur in puncto X, & locus puncti illius X erit Hyperbola Conica (puta XΦ) cujus una Asymptotos est AS. Sit ejus altera Asymptotos AB, & æquatio qua relatio inter Ordinatam BC & Abscissam AB definitur, si AB dicatur x & BC y, semper induet hanc formam $xyy + ey = ax^3 + bxx + cx + d$. Ubi termini e, a, b, c, d, designant quantitates datas cum signis suis + & − affectas, quarum quælibet deesse possunt modo ex earum defectu figura in sectionem conicam non vertatur. Potest autem Hyberbola illa Conica cum asymptotis suis coincidere, id est punctum X in recta AB locari: & tunc terminus + ey deest.

VIII. *Reductio Curvarum omnium generis secundi ad æquationum casus quatuor. Casus primus.*

Fig. 1.

At si recta illa CBc non potest utrinq; ad Curvam terminari sed Curvæ in unico tantum puncto occurrit: age quamvis positione datam rectam AB asymptoto AS occurrentem in A, ut & aliam quamvis BC asymptoto illi parallelam Curvæque occurrentem in puncto C, & æquatio qua relatio inter Ordinatam

IX. *Casus secundus.*

BC

Fig. 1.4 Page 143 of Newton's *Linearum* from 1704

Euclidean geometric objects. His work on algebraic curves is, however, not a part of *Principia Mathematica*.

Newton gave a quite precise classification of algebraic curves of degree three. This is a direct generalization of the case of curves of degree two, the conic sections. The basic tool used was to analyze the highest-order homogeneous terms of degree three and their possible factorizations over the real numbers. This led to various types

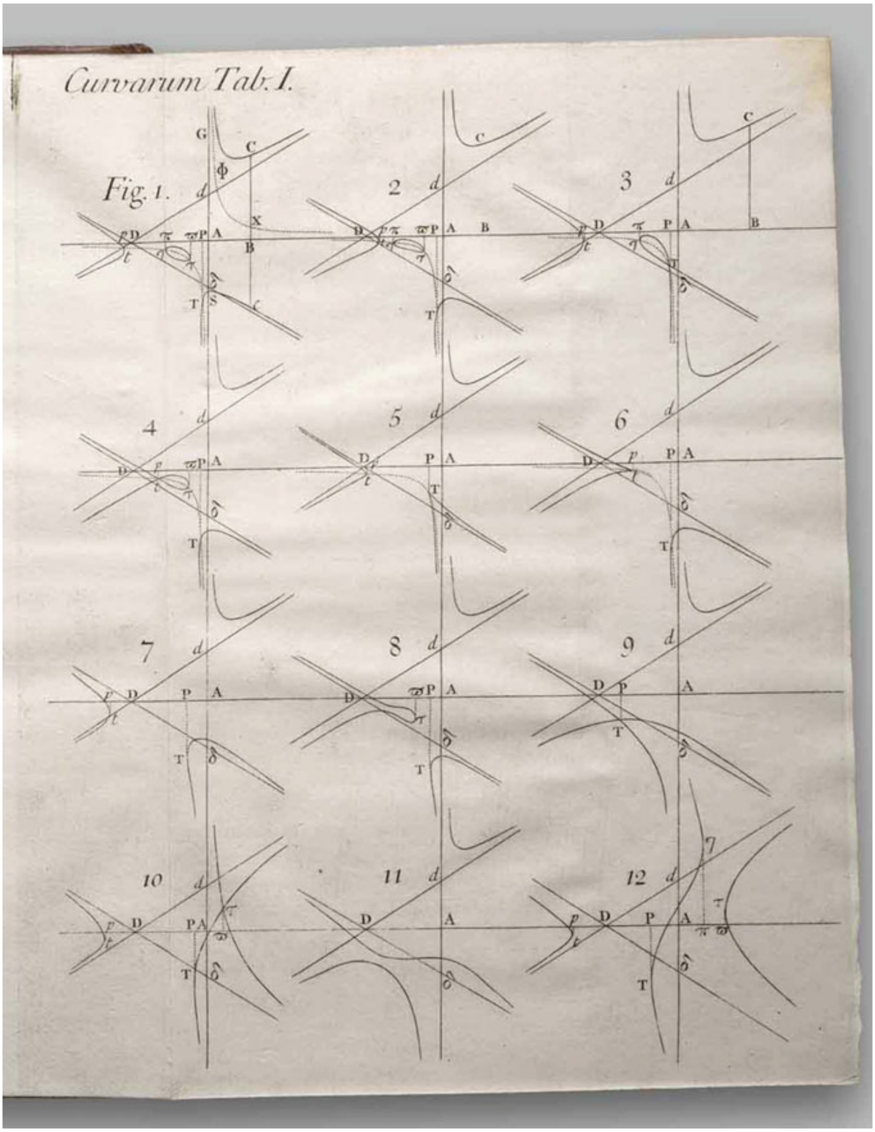


Fig. 1.5 First page of Newton's figures from *Linearum* including Fig. 1 referenced on p. 143 of the same book

of branches that are unbounded (and to various types of asymptotes) that can arise, and they become an important part of the classification. For instance, for curves of degree two, one can see that an ellipse has no infinite branches, a parabola has two infinite branches, and a hyperbola has four infinite branches including two straight lines which are asymptotes. This behavior at infinity completely distinguishes these three classes of curves.

Newton's analysis yielded the classification of 78 different types of curves. In fact, he only described 72, having missed six types. He classified them algebraically, and then provided beautiful drawings of the typical curve of the specified classification type. In Figs. 1.4 and 1.5 we see some examples of the classification and the corresponding drawings. His analysis of how he arrived at the classification is very terse indeed. In fact, there is very little description of his analysis. The monograph is not much more than a simple listing of his findings. Proofs of his results were published later by mathematicians who analyzed and generalized his results. He also indicates that one can carry out such an analysis for curves of higher degree by the same method of analysis.

Many mathematicians over the centuries have, of course, analyzed Newton's results in great detail and one can find the classification in a variety of monographs (see, for example, Walker [227] for a modern treatment of real algebraic curves). However, one of the most beautiful and thorough analyses of Newton's work from the classical literature is that of Leonhard Euler (1707–1783) in [62]. This very influential book, *Introductio in Analysin Infinitorum*, was first published in Latin in 1748 and then reappeared later in many editions and in various languages.

What we have described above very briefly is Newton's classification of algebraic curves of degree three from the beginning of the eighteenth century. Such developments carried on into modern times, in particular with similar classifications of Riemann surfaces (considered as algebraic curves of complex dimension one of various degrees), and complex manifolds of higher complex dimension. In general the study of complex manifolds or complex algebraic varieties of various dimensions has turned out to be simpler than the study of real algebraic manifolds and varieties over the real numbers due to the closure of the field of complex numbers. But it all started in the real algebraic setting, as that's what the mathematical community was familiar with in the early eighteenth century. They knew about complex numbers, but they were not yet familiar with complex geometry.

Chapter 2
Differential Geometry

2.1 Introduction

A second major development in geometry in the eighteenth century was the study of curves and surfaces in $\mathbf{R}^2$ and $\mathbf{R}^3$ defined by not necessarily algebraic functions. These included two not quite independent developments that took place more or less simultaneously. The first was the development of the now standard elementary transcendental functions: the trigonometric, exponential, and logarithmic functions. In Euler's textbook from 1748 [62] these functions and their algebraic and analytic properties (e.g.,

$$\frac{d}{dx}\sin x = \cos x, \quad \sin(x+y) = \sin x \cos y + \cos x \sin y.$$

etc.) were fully developed and correspond to what one learns in contemporary precalculus and calculus courses in high school today. The second development involved the solution of differential equations (primarily ordinary differential equations) which provided a large variety of functions for analysis and geometrical representation. This led to a large class of special functions that went by the names of the mathematicians who created and developed them: Hermite, Legendre, Bessel, Euler's Gamma function and many others. These functions were tabulated for computational use and their various algebraic and analytical properties were developed, similar to those properties illustrated above for trigonometric functions. Over the course of time these mathematical tools became very important for the applications of mathematics to the worlds of chemistry, physics, biology and other areas of scientific understanding. These methods preceded by one or two centuries contemporary techniques for scientific analysis made available through the use of computers and simulation tools involving modern numerical analysis, which were to diminish the once important role of special functions.

In the latter half of the eighteenth century the differential geometry of curves and surfaces began to develop and flourish. First we consider the development of what

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_2

became known as planar curves and space curves (i.e., smooth curves in $\mathbf{R}^2$ and $\mathbf{R}^3$). *Differential geometry* was named as a concept by Bianchi in 1894 (as noted by Kline [125] on p. 554). This naming of the discipline came long after the most significant developments in the field. It came to mean precisely manifolds equipped with a Riemannian (or more general) metric, or more generally a connection, and where the concepts of curvature played a central role. Indeed, the interaction of differential analysis (i.e. calculus, differential equations, all aspects of analysis involving infinite processes) with geometry is much older and broader than the more precise notion of differential geometry as it is employed today. For instance, the notion of differential topology, which developed in the mid-twentieth century, certainly involves manifolds and analysis, but doesn't formally use the notion of a differential-geometric metric as in differential geometry *per se*. Archimedes knew how to compute areas by the method of exhaustion, and Fermat understood both differentiation of functions (finding maxima and minima and tangents) and how to compute the area under some curves, but he did not know the fundamental theorem of calculus (see [194] for a discussion of these issues). All of these are indeed an interaction of analysis with geometry, and are parts of the foundation of what became differential geometry two centuries later.

2.2 Huygens and Newton

The first important task in differential geometry was to be able to efficiently compute the tangent line to a given curve at a given point and, as any beginning student of calculus knows, this is one of the first applications of the notion of the derivative. A deeper question that we explore in greater detail in this section is: what is curvature? More precisely, what is the curvature of a curve in a plane or in three-dimensional space? What is the curvature of a surface in three-dimensional space? Finally, what is curvature of an abstract two-dimensional or higher-dimensional manifold? This last question is a key part of the geometric developments in the nineteenth century and will be discussed in Part II.

Consider first the simple case of a curve in the plane defined by the graph of a function as in Fig. 2.1. Then one learns in calculus that the curvature of the curve at $P = (x, y)$ is given by

$$K_P = \pm \frac{f''(x)}{[1 + (f'(x)^2]^{\frac{3}{2}}}, \tag{2.1}$$

where the sign is chosen to be positive if the normal vector to the curve at P intersects the approximating circle and is negative otherwise. In the illustration in Fig. 2.1, the normal vector to the curve at P using the usual orientation would be pointing upwards in the figure, away from the approximating circle, whose radius is $1/|K_P|$, and hence in this case the curvature would be negative.

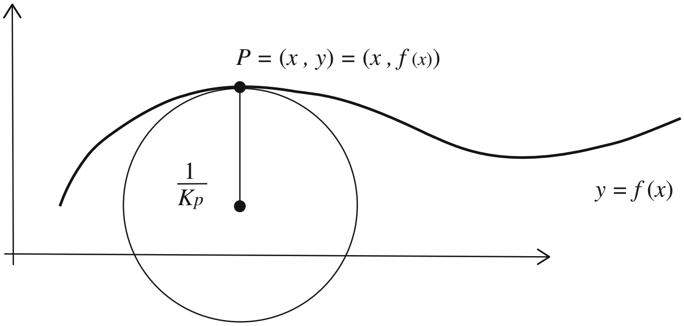


Fig. 2.1 Radius of curvature of a curve at a point

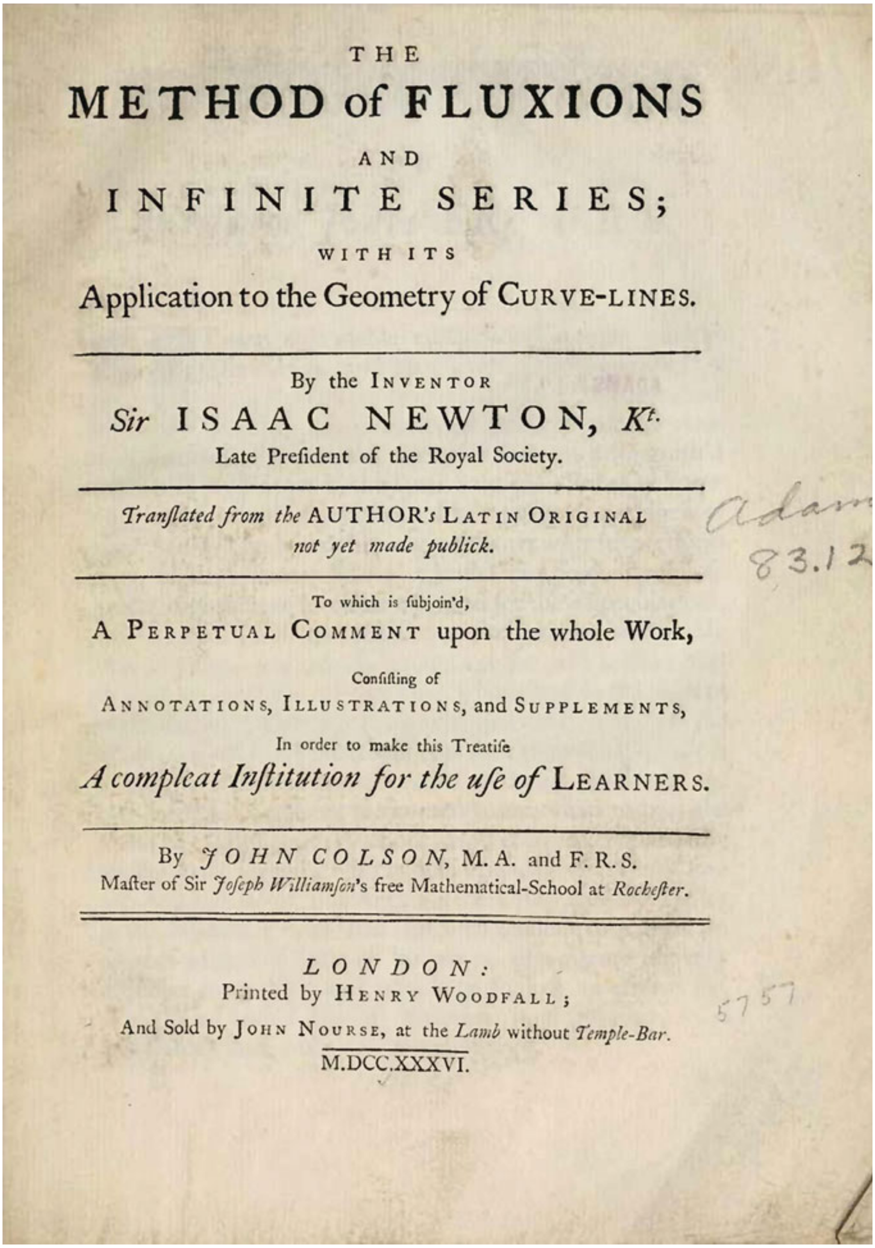
This formula is given for the first time in Newton's monograph of 1736 [169], which was published as an English translation of his original Latin manuscript from 1671, which was never published, but was privately circulated among some of Newton's colleagues. This monograph, published in 1736 after Newton's earlier death, was part of the basis for the controversy between adherents of Newton and Leibniz on who had first invented (or discovered) calculus. Figure 2.2 shows the cover page of this singular monograph, and Fig. 2.3 shows the table of contents, where the curvature of a curve stands out so very distinctly as an object of study. The formula (2.1) appears in the text of Newton's monograph.

The first published account of the curvature of a general curve was due to Christiaan Huygens (1629–1695) in 1673 [114]. In both Newton and Huygens the fundamental definition of the center of curvature (center of the osculating circle at a given point) is the intersection of normal lines to the curve near the given point on the curve (see the figures in Huygens p. 84 [114] and Newton on p. 60 [169], reproduced here in Figs. 2.4 and 2.5).

Huygens didn't have calculus *per se* at his disposal, but he made estimates in terms of normals at an approximating point (like the estimates of slopes of an approximating secant to a tangent line in differential calculus), and using these estimates he was able to compute the curvature for a variety of examples (cycloid, conic sections, etc.).

An interesting historical point is how Huygens came to study this phenomenon. Some 16 years before the appearance of his monograph [114] he had built one of the most important clocks in history: a pendulum whose motion is isochronous. That is, the swing of the pendulum has a constant period of repetition. Huygens showed that a simple pendulum, whose pendant moves in a circular arc, has a period that depends on the size of the oscillations, whereas if the pendant moves in the arc of a cycloid, then the period is fixed independent of the size of the oscillation.

The method Huygens used for making the pendant move in a cycloidal path (which he patented in 1657) was to have the path be the *involute* of a curved plate (which was also a cycloid), i.e., the curve traced out by a fixed string moving from a center attached to a given curve, where initially the fixed string lies along the given curve and moves away from it, with the free straight line portion of the string being



THE

METHOD of FLUXIONS

AND

INFINITE SERIES;

WITH ITS

Application to the Geometry of CURVE-LINES.

By the INVENTOR

Sir ISAAC NEWTON, *K^t.*

Late Prefident of the Royal Society.

Tranflated from the AUTHOR'*s* LATIN ORIGINAL

not yet made publick.

To which is fubjoin'd,

A PERPETUAL COMMENT upon the whole Work,

Confifting of

ANNOTATIONS, ILLUSTRATIONS, and SUPPLEMENTS,

In order to make this Treatife

A compleat Inftitution for the ufe of LEARNERS.

By *JOHN COLSON*, M. A. and F. R. S.

Mafter of Sir *Jofeph Williamfon*'s free Mathematical-School at *Rochefter*.

LONDON:

Printed by HENRY WOODFALL;

And Sold by JOHN NOURSE, at the *Lamb* without *Temple-Bar*.

M.DCC.XXXVI.

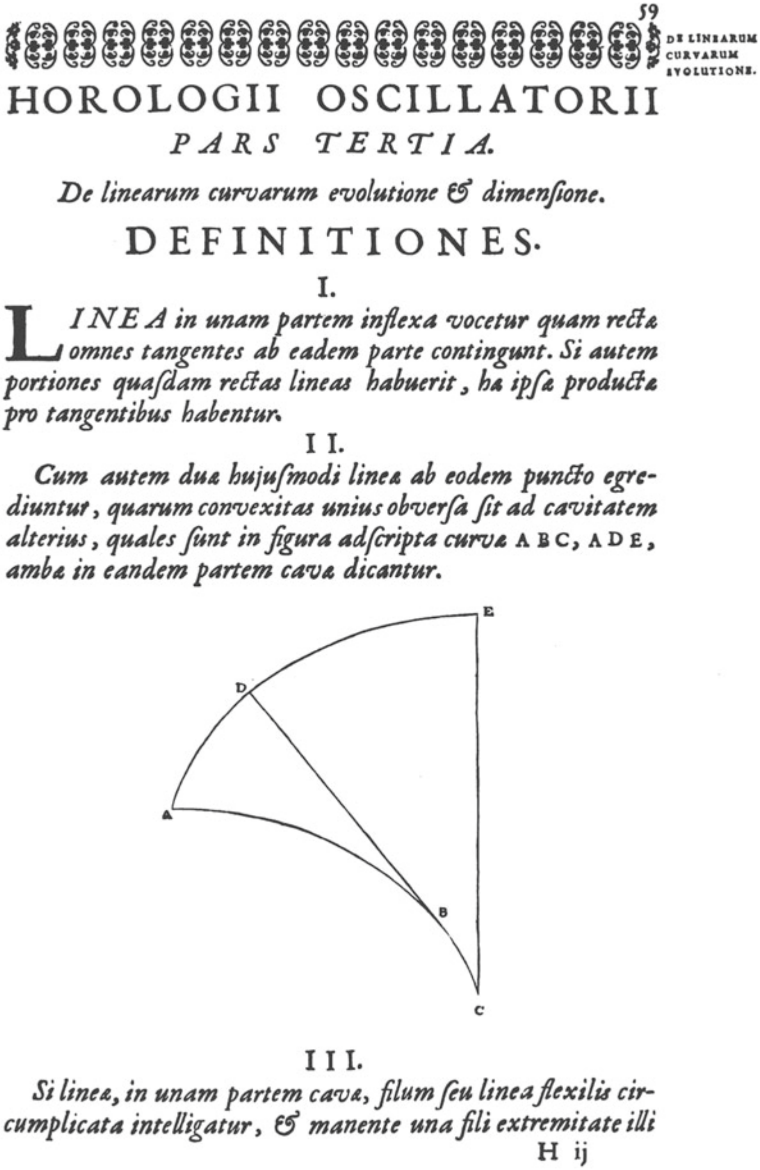
Fig. 2.2 Title page of Newton's 1736 Monograph on Fluxions

(xxiv)

THE

CONTENTS.

Fig. 2.3 Table of contents of Newton's 1736 monograph on fluxions


59

DE LINEARUM CURVARUM EVOLUTIONE.

HOROLOGII OSCILLATORII

PARS TERTIA.

De linearum curvarum evolutione & dimenſione.

DEFINITIONES.

I.

L*INEA in unam partem inflexa vocetur quam rectæ omnes tangentes ab eadem parte contingunt. Si autem portiones quaſdam rectas lineas habuerit, hæ ipſæ productæ pro tangentibus habentur.*

II.

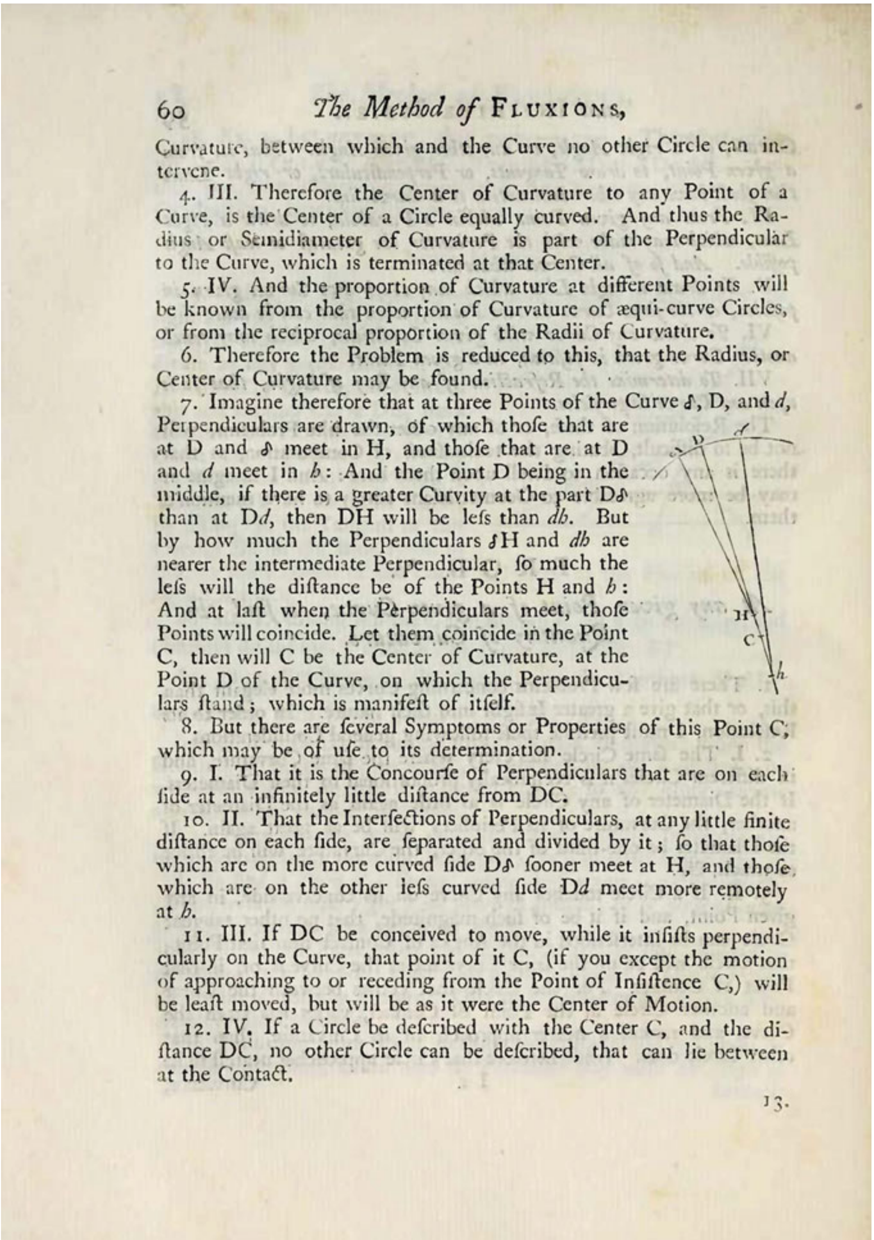
Cum autem duæ hujuſmodi lineæ ab eodem puncto egrediuntur, quarum convexitas unius obverſa ſit ad cavitatem alterius, quales ſunt in figura adſcripta curvæ ABC, ADE, *ambæ in eandem partem cavæ dicantur.*

III.

Si lineæ, in unam partem cavæ, filum ſeu linea flexilis circumplicata intelligatur, & manente una fili extremitate illi

H ij

Fig. 2.4 Huygens's center of curvature from *Horologium Oscillatorium*



60 *The Method of* FLUXIONS,

Curvature, between which and the Curve no other Circle can intervene.

4. III. Therefore the Center of Curvature to any Point of a Curve, is the Center of a Circle equally curved. And thus the Radius or Semidiameter of Curvature is part of the Perpendicular to the Curve, which is terminated at that Center.

5. IV. And the proportion of Curvature at different Points will be known from the proportion of Curvature of æqui-curve Circles, or from the reciprocal proportion of the Radii of Curvature.

6. Therefore the Problem is reduced to this, that the Radius, or Center of Curvature may be found.

7. Imagine therefore that at three Points of the Curve δ, D, and *d*, Perpendiculars are drawn, of which thoſe that are at D and δ meet in H, and thoſe that are at D and *d* meet in *h*: And the Point D being in the middle, if there is a greater Curvity at the part Dδ than at D*d*, then DH will be leſs than *dh*. But by how much the Perpendiculars δH and *dh* are nearer the intermediate Perpendicular, ſo much the leſs will the diſtance be of the Points H and *h*: And at laſt when the Perpendiculars meet, thoſe Points will coincide. Let them coincide in the Point C, then will C be the Center of Curvature, at the Point D of the Curve, on which the Perpendiculars ſtand; which is manifeſt of itſelf.

8. But there are ſeveral Symptoms or Properties of this Point C, which may be of uſe to its determination.

9. I. That it is the Concourſe of Perpendiculars that are on each ſide at an infinitely little diſtance from DC.

10. II. That the Interſections of Perpendiculars, at any little finite diſtance on each ſide, are ſeparated and divided by it; ſo that thoſe which are on the more curved ſide Dδ ſooner meet at H, and thoſe which are on the other leſs curved ſide D*d* meet more remotely at *h*.

11. III. If DC be conceived to move, while it inſiſts perpendicularly on the Curve, that point of it C, (if you except the motion of approaching to or receding from the Point of Inſiſtence C,) will be leaſt moved, but will be as it were the Center of Motion.

12. IV. If a Circle be deſcribed with the Center C, and the diſtance DC, no other Circle can be deſcribed, that can lie between at the Contact.

13.

Fig. 2.5 Newton's center of curvature from *Method of Fluxions*

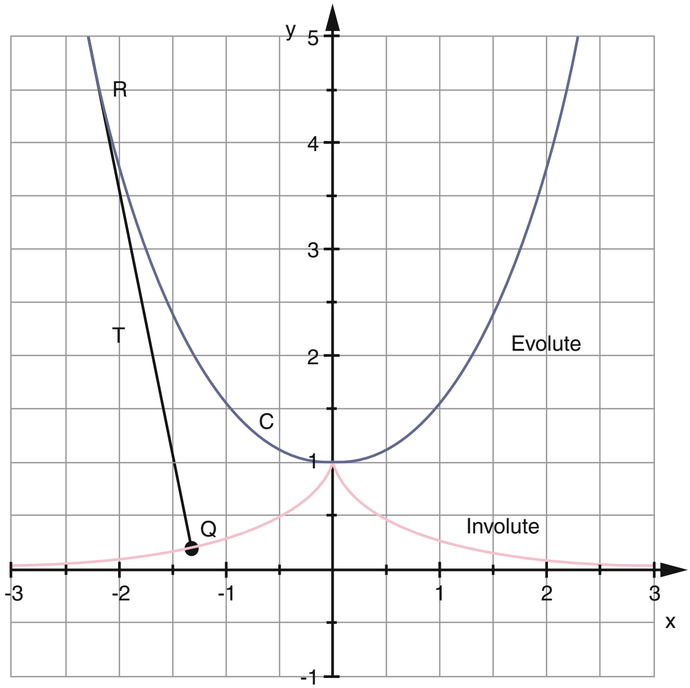


Fig. 2.6 An involute being generated by a string attached to the curve C (called the evolute)

continuously tangent to the given curve (see the illustration in Fig. 2.6). The curve C in Fig. 2.6 is called the *evolute* (which generates the involute traced out by the point Q by the motion of the string). The problem Huygens posed and solved was: given the involute, find the evolute, i.e., find the generating curve. Now the straight line T is normal to the involute at the point Q (as Huygens showed), and, at the point of contact at point R, T is tangent to C. Thus T is normal to the involute at Q, and R can be seen to be the intersections of the normals close to Q (as both Huygens and Newton showed). Hence R is the center of curvature of the involute at the point Q, and the evolute C is the locus of centers of curvature of the involute at points near Q.

In the second illustration of an involute in Fig. 2.7, one sees two "parallel" involutes, the curves C' and C'' being generated from the curve C, and one can see that the involutes are orthogonal to the generating string at the intersection points (as was proved by Huygens). Looking at the illustration from p. 4 (Fig. 2.8) of Huygens's book [114] one sees in Fig. II of the diagrams in Fig. 2.8 the cycloid-shaped curve from which the pendant of the pendulum sweeps out the involute, which is the cycloidal motion of the pendant. Huygens calculated the evolutes for a number of examples, independent of the specific example he used in the design of his clock.

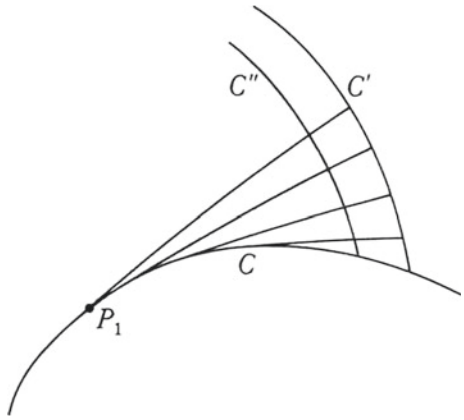


Fig. 2.7 Involutes are orthogonal to the generating string

Some 2000 years earlier, in Book V of his famous work *Conics*, Apollonius was able to compute the curvature of the classical conic sections. Apollonius was in fact trying to solve a different set of problems, and curvature was not explicitly discussed. In Heath's translation [98], he shows what Apollonius did in modern notation. More particularly, on p. 171 one finds that for the parabola of the form

$$\frac{1}{2a}y^2 = x,$$

the evolute (locus of centers of curvature) of this parabola has the form:

$$27ay^2 = 4(x - 2a)^3,$$

which is a semicubical parabola. He finds similar formulas for the ellipse and hyperbola.

Here Apollonius was studying the behavior of normals to conic sections. He showed that each conic section has a unique normal passing through each point. He *defined* a normal as being a straight line which was either a local maximum or a local minimum-length straight line from some point not on the curve. He then showed that such a line was indeed perpendicular to the tangent line at the given point. This leads, by an interesting argument, to the conclusion that Apollonius has calculated the points of the evolute, as Heath points out very explicitly.

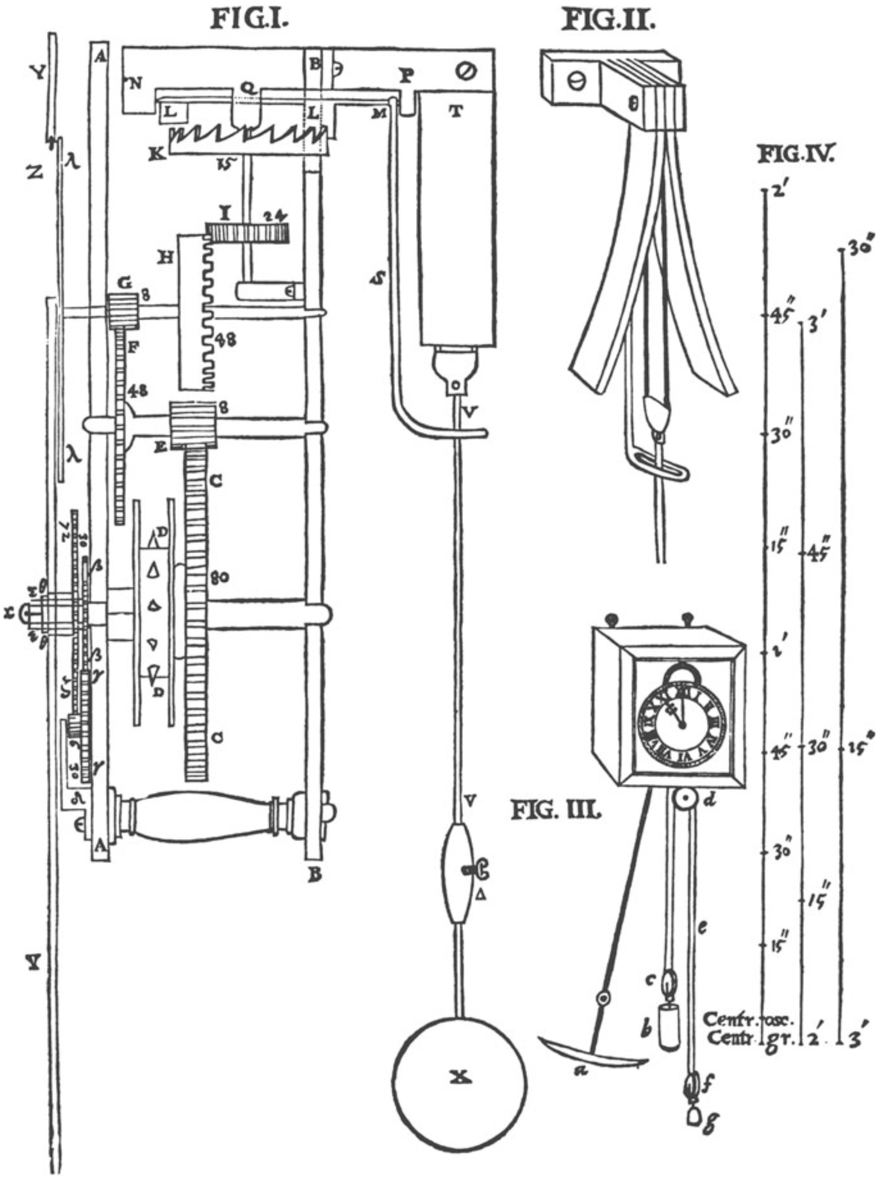


Fig. 2.8 Page 4 of Huygens's book *Horologium Oscillatorium* [114]

2.3 Curves in Space: *Courbes à double courbure*

Since the time of Newton, curvature of a curve in the plane became a standard object of mathematical investigation. The first step in investigating the differential geometry of curves in $\mathbf{R}^3$ was taken by Alexis Claude Clairaut (1713–1765) in his book *Recherches sur les courbes à double courbure* [48], written when he was only 16 years old and published two years later, following up on work he had started when he was 12 years old. We know this from the "Approbation" at the beginning of the book, written by two of the reviewers of the book; and the page where this appears, following the Preface, is the only place Clairaut's name appears in the book, not on the title page! See Fig. 2.9. Clairaut called curves in $\mathbf{R}^3$ "courbes à double courbure",[1] and he says in his book that he was inspired by Descartes, who suggested space curves could be studied in terms of their projections on two orthogonal planes. Clairaut studied the tangent line to a curve, its arc length and the infinite variety of normal lines in the plane perpendicular to the tangent line.

The next steps in the study of space curves were taken by Euler, who primarily looked at space curves which were defined as the intersections of surfaces in $\mathbf{R}^3$ (see Volume 2 of Euler's *Introductio* of 1748 [62]). Michel Ange Lancret (1774–1807) singled out in 1806 the three principal directions of a space curve at any point (tangent, normal, and binormal), and formulated the additional notion of torsion of a curve [132].

The final steps in the study of space curves were taken by Augustin-Louis Cauchy (1789–1857) in 1826 in his *Leçons sur les Applications du Calcul Infinitésimal à la Géométrie* [38], and by Serret [216] and Frenet [75] in their back-to-back papers in 1851 and 1852. Cauchy gave us the formulation of space curves we use today (without the vector notation), and Serret and Frenet gave the final form to the structure equations (which today bear their name, the Frenet–Serret equations), which brought together the formal characterization of space curves in terms of the three principal directions of a curve and its curvature and torsion.

2.4 Curvature of a Surface: Euler in 1767

The concept of the curvature of a curve in $\mathbf{R}^3$ was well understood at the end of the eighteenth century, and the later work of Cauchy, Serret and Frenet completed this set of investigations begun by the young Clairaut a century earlier. The problem arose: how can one define the curvature of a surface defined either locally or globally in $\mathbf{R}^3$? An important contribution is made by Euler in his paper entitled "Recherches sur la courbure des surfaces"[2] [65] from 1767 (note this article is written in French,

[1] "curves with double curvature". The expression "courbes à double courbure" was used to describe space curves for a long time by many mathematicians after the initial impetus of Clairaut.

[2] "Research on the curvature of surfaces".

APPROBATION.

J'AI lû par ordre de M. le Garde des Sceaux un Manuſcrit intitulé *Recherches ſur les Courbes à double courbure*, compoſé par M. CLAIRAUT. Ce Traité que les plus habiles Geometres de notre temps & des ſiecles paſſés ſe ſeroient fait honneur d'avoir compoſé, & qui eſt certainement l'Ouvrage d'un jeune homme de ſeize ans, qui dès l'âge de douze avoit déja donné des marques publiques de ſon habileté dans les Mathematiques, ne merite pas ſeulement d'être imprimé, mais d'être admiré comme un prodige d'imagination, de conception & de capacité. A Paris ce 3 Juin 1730.

J. DE MOLIERES.

EXTRAIT des Regiſtres de l'Academie Royale des Sciences, du 20 Août 1729.

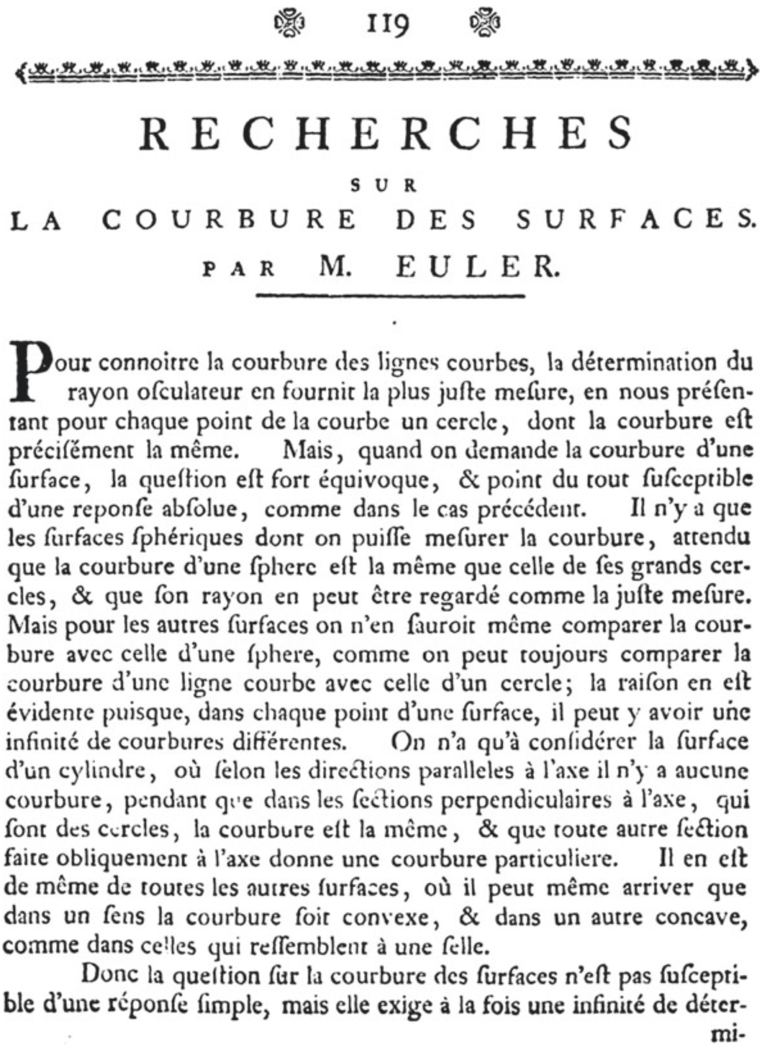
MEſſieurs de Mairan & Nicole qui avoient été nommés pour examiner un Ouvrage de M. CLAIRAUT le Fils, intitulé, *Recherches ſur les Courbes à double courbure*, en ayant fait leur rapport, la Compagnie a jugé que cet Ouvrage contenoit beaucoup de choſes curieuſes & nouvelles ſur ces ſortes de courbes, & montroit non-ſeulement de l'invention dans l'Auteur, qui n'eſt âgé que de ſeize ans, mais encore beaucoup de connoiſſance du Calcul differentiel, & de l'Integral. Fait à Paris ce 23 Août 1729.

FONTENELLE, Sec. perp. de l'Ac. Roy. des Sc.

Fig. 2.9 Excerpt from the beginning of Clairaut's book *Recherches sur les courbes à double courbure* [48]

not like his earlier works, most of which were written in Latin). Figure 2.10 shows the first page of the article and we quote the translation here:

> In order to know the curvature of a curve, the determination of the radius of the osculating circle furnishes us the best measure, where for each point of the curve we find a circle whose curvature is precisely the same. However, when one looks for the curvature of a surface, the question is very equivocal and not at all susceptible to an absolute response, as in the case above. There are only spherical surfaces where one would be able to measure the curvature, assuming the curvature of the sphere is the curvature of its great circles, and whose radius could be considered the appropriate measure. But for other surfaces one doesn't know even how to compare a surface with a sphere, as when one can always compare the curvature of a curve with that of a circle. The reason is evident, since at each point of a surface there are an infinite number of different curvatures. One has to only consider a cylinder, where along the directions parallel to the axis, there is no curvature, whereas in the directions perpendicular to the axis, which are circles, the curvatures are all the same, and all other oblique sections to the axis give a particular curvature. It's the same for all other surfaces, where it can happen that in one direction the curvature is convex, and in another it is concave, as in those resembling a saddle.



119

RECHERCHES

SUR

LA COURBURE DES SURFACES.

PAR M. EULER.

Pour connoitre la courbure des lignes courbes, la détermination du rayon oſculateur en fournit la plus juſte meſure, en nous préſentant pour chaque point de la courbe un cercle, dont la courbure eſt préciſément la même. Mais, quand on demande la courbure d'une ſurface, la queſtion eſt fort équivoque, & point du tout ſuſceptible d'une reponſe abſolue, comme dans le cas précédent. Il n'y a que les ſurfaces ſphériques dont on puiſſe meſurer la courbure, attendu que la courbure d'une ſphere eſt la même que celle de ſes grands cercles, & que ſon rayon en peut être regardé comme la juſte meſure. Mais pour les autres ſurfaces on n'en ſauroit même comparer la courbure avec celle d'une ſphere, comme on peut toujours comparer la courbure d'une ligne courbe avec celle d'un cercle; la raiſon en eſt évidente puisque, dans chaque point d'une ſurface, il peut y avoir une infinité de courbures différentes. On n'a qu'à conſidérer la ſurface d'un cylindre, où ſelon les directions paralleles à l'axe il n'y a aucune courbure, pendant que dans les ſections perpendiculaires à l'axe, qui ſont des cercles, la courbure eſt la même, & que toute autre ſection faite obliquement à l'axe donne une courbure particuliere. Il en eſt de même de toutes les autres ſurfaces, où il peut même arriver que dans un ſens la courbure ſoit convexe, & dans un autre concave, comme dans celles qui reſſemblent à une ſelle.

Donc la queſtion ſur la courbure des ſurfaces n'eſt pas ſuſceptible d'une réponſe ſimple, mais elle exige à la fois une infinité de déter-

mi-

Fig. 2.10 The opening page of Euler's work on curvature [65]

In this paper Euler describes quite clearly the problem of formulating a concept of curvature of a surface in $\mathbf{R}^3$. In particular, in the quote above one sees that Euler recognized the difficulties in defining curvature for a surface at any given point. He does not resolve this issue in this paper, but he makes extensive calculations and several major contributions to the subject. He considers a surface S in $\mathbf{R}^3$ defined as a graph

$$z = f(x, y)$$

near a given point $P = (x_0, y_0, z_0)$. At the point P he considers planes in $\mathbf{R}^3$ passing through the point P which intersect the surface in a curve in that given plane. For each such plane and corresponding curve he computes explicitly the curvature of the curve at the point P in terms of the given data.

He then restricts his attention to planes which are normal to the surface at P (planes containing the normal vector to the surface at P). There is a one-dimensional family of such planes E_θ, parametrized by an angle θ. He computes explicitly the curvature of the intersections of E_θ with S as a function of θ, and observes that there is a maximum and minimum κ_1 and κ_2 of these curvatures at P, corresponding to two planes E_1 and E_2. These curvatures are called the *principal curvatures* of the surface at the point P. In the generic case, Euler shows that the two planes E_1 and E_2 are orthogonal to each other. Moreover, he shows that the curvature κ_θ for the plane E_θ can be computed in terms of the principal curvatures, namely

$$\kappa_\theta = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta.$$

This is as far as he goes, but it is a great step forward in understanding the curvature of a surface. He does *not* use this data to define what we now call the *curvature* of the surface S at the point P. This step was taken by Gauss in a visionary and extremely important paper some 60 years later [81].

Part II
Differential and Projective Geometry in the Nineteenth Century

Introduction

In the nineteenth century, there were a number of parallel developments of innovative ideas all of which played a major role in twentieth-century mathematics. In this part, we want to discuss two major themes that are important for our study of embedding theorems in the mid-twentieth century.

The first theme is that of *projective geometry*, which began in 1799 with the work of Monge in Paris and culminated with the work of many mathematicians in the second half of the nineteenth century. This led to our contemporary notion of projective space, an important adjunct to the classical Euclidean space that came to us from our Greek mathematical heritage. We survey in Chap. 3 the work of a number of mathematicians involved in developing projective geometry. There were two schools of thought: synthetic projective geometry (the French school) and analytic projective geometry (the German school), which came together towards the end of the century to be the projective geometry we understand today.

Our second theme is the creation of intrinsic differential geometry, the concept of higher-dimensional manifolds, and the introduction of a metric and a curvature tensor on a manifold. This is primarily the work of Carl Friedrich Gauss (1777–1855) in 1828 and of Bernhard Riemann (1826–1866) in 1854. We outline their two fundamental papers in Chaps. 4 and 5. This is the foundational work for contemporary *differential geometry*.

In Part III, we will examine a third major theme of the nineteenth century, namely the origin of *complex geometry*.

Chapter 3
Projective Geometry

The first major development that we want to discuss, having its beginnings at the start of the nineteenth century, is the creation of *projective geometry*; which led in the latter half of the nineteenth century to the important concept of a *projective space* $\mathbf{P}_n$ over the real or complex numbers (or more general fields). This became at that time a centerstage for many developments in algebraic geometry and complex manifolds which carried over vigorously into the twentieth century. In this chapter we will outline the very interesting story of how and why projective geometry developed.

Projective geometry started as a school of mathematics in France around 1800. Throughout the eighteenth century, as we mentioned above, coordinate geometry and its interaction with differential and integral calculus—what became known later as differential geometry—dominated mathematical research in geometry. In particular, the classical ideas of what became known as *synthetic geometry* in the spirit of Euclid's *Elements* began to fall to the wayside in mathematical research, even though Newton and others had often resorted to synthetic geometric arguments early in the eighteenth century as a complement to (and often as a check on) the analytic methods using coordinates.

Some of the proponents of projective geometry wanted to create a type of geometry that could provide an important alternative to coordinate geometry for a variety of interesting geometric problems. This led to new considerations of geometric equivalence for geometric figures, and led to an extension of Euclidean space to include points at infinity, as well as other important innovations. A parallel development in projective geometry was the use of coordinate systems of various types to provide alternative proofs of some of the fundamental results in the subject. This led in particular to the concept of homogeneous coordinates.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_3

3.1 Monge and Descriptive Geometry

It is generally recognized that projective geometry is the singular creation of Ernst Gaspard Monge[1] (1746–1818) and his pupils, principally Carnot, Poncelet, and Chasles, to mention just three of them. Monge published a book [159], *Geométrie Descriptive*, which inspired major developments by his pupils and, in turn, resulted in several substantive books; Poncelet's book *Propriétés Projectives* [190] in 1822 being perhaps the most influential. The book by Michel Chasles (1793–1880) [43] is a two-part work, the first of which is a brilliant historical treatise on the whole history of geometry up to that point in time, 1831. The second part is his own treatment of projective geometry, following up on the work of Monge, Carnot, Poncelet and others.

Lazare Nicolas Marguerite Carnot (1753–1823) wrote in 1803 a significant book; *Géométrie de Position*[2] [31], which played an important role in projective geometry. This book, along with Monge's original book on descriptive geometry from 1799 [159], was a major influence on Poncelet. We will say more about this later, but first we want to give an overview of how and why projective geometry came to be and what these first authors believed they had achieved.

It is quite fascinating to read the 1827 edition of Monge's book [159], which was edited by his pupil Bernabé Brisson and published after Monge's death in 1818. This was the fifth edition of the book which first appeared in 1799 and was the result of his lectures at the *École Normale*, which were very inspirational, according to several testimonials published in this edition. Monge was very concerned about secondary education and the creation of a new generation of educated citizens who could help in the development of the Industrial Revolution in France.

He believed that descriptive geometry, which he developed in this book, would be a tool for representing three-dimensional objects in terms of their projections onto one or more planes in three-dimensional space, and that this should be a major part of the educational development of students. The applications he uses as examples came from architecture, painting, the representation of military fortifications, and elsewhere. Monge was also interested in developing an alternative approach to classical synthetic geometry in the three-dimensional setting in order to develop a geometric method which would be useful in engineering. Today we use his ideas in the form of blueprints for industrial design with such two-dimensional drawings representing horizontal and vertical projections of the object being designed or manufactured.

The basic thesis of plane geometry as formulated by the Greeks and brought down to us in the book of Euclid is a solution of geometric problems in the plane by the use of the straight edge and compass. Monge's fundamental thesis in his book is to reduce problems of geometry in three-dimensional space to plane geometry problems (in the Euclidean sense) on the *projections* of the problems to two (or more) independent

---

[1]Monge was a major scientific advisor to Napoleon on his Egyptian expedition and helped create the Ecole Polytechnique, the first engineering school in France, as well as the metric system.

[2]"Geometry of Position".

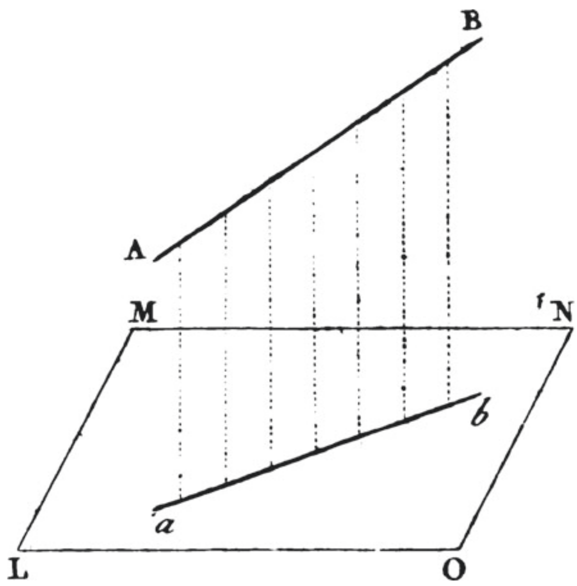


Fig. 3.1 Figure 1 in Monge's *Géométrie Descriptive* [159]

planes. Considering the straight edge and compass as engineering tools, a designer or engineer could work on three-dimensional problems in the two-dimensional medium.

Let us illustrate this with a simple example from his book. In Fig. 3.1 we see the projection of the line segment AB in $\mathbf{R}^3$ onto the line segment ab in the plane $LMNO$. Let us visualize these line segments as representing straight lines extended infinitely in both directions. Let's change notation slightly from that used by Monge in this figure. Let L be the given line in $\mathbf{R}^3$ and let E_1 and E_2 be two non-parallel planes in $\mathbf{R}^3$. Let L_1 and L_2 be the perpendicular projections of the line L onto the planes E_1 and E_2, respectively. We see that L_1 and L_2 represent L in this manner. Can we reverse the process? Indeed, given L_1 and L_2, two lines in E_1 and E_2, let H_1 and H_2 be the planes in $\mathbf{R}^3$ which are perpendicular to E_1 and E_2 and which pass through the lines L_1 and L_2. Then the intersection $H_1 \cap H_2$ is the desired line in $\mathbf{R}^3$ whose projections are L_1 and L_2. This simple example is the main idea in the projection of a point in $\mathbf{R}^3$ onto the three orthogonal axes (x-axis, y-axis, and z-axis), giving the Cartesian coordinates (x, y, z) of a given point.

Monge formulates and solves a number of problems using these types of ideas. Here is a quite simple example from his book, paralleling a similar question in plane geometry. Given a line L in $\mathbf{R}^3$ and a point P not on the line, construct a line through P parallel to L. One simply chooses two reference planes E_1 and E_2, as above, and projects orthogonally L and P onto the planes, obtaining lines L_1 and L_2 and points P_1 and P_2 in the reference planes. Then, in this plane geometry setting, find parallel lines $l_1 \subset E_1$ and $l_2 \subset E_2$ to the lines L_1 and L_2 passing through the points P_1 and P_2. The lines l_1 and l_2 determine a line $l \subset \mathbf{R}^3$, which solves the problem.

Monge considers in great depth for most of his book much more complicated problems concerning various kinds of surfaces in $\mathbf{R}^3$ and shows how to construct tangents, normals, principal curvatures and solutions to many other such problems. He uses consistently the basic idea of translating a three-dimensional problem into several two-dimensional problems. This is the essence of descriptive geometry, but it becomes the basis for the later developments in projective geometry, as we will now see.

3.2 Poncelet's "Propriétés Projectives"

The most influential figure in the development of projective geometry in the first half of the nineteenth century (aside from the initial great influence of Monge, as described above) was undoubtedly Jean-Victor Poncelet (1788–1867). He was an engineer and mathematician who served as Commandant of the Ecole Polytechnique in his later years. His major contributions to projective geometry were his definitive books *Traité des Propriétés Projectives des Figures*, Volumes 1 and 2 [190, 191], which were first published in 1822 and 1824 and reappeared as second editions in 1865 and 1866.

In addition, he published in 1865 a fascinating historical book *Applications d'Analyse et de Géométrie* [189], the major portion of which is the reproduction of notebooks Poncelet wrote in a Russian prisoner-of-war camp in Saratov 1813–1814. He was interned there for about two years after a major battle (at Krasnoi) which Napoleon lost in November of 1812 towards the end of his disastrous Russian campaign to Moscow, and where Poncelet and others had been left on the field as dead. Poncelet had been serving as an *Officier de génie* (engineering officer). He had no books or notes with him and wrote out and developed further ideas he had learned from the lectures and writings of Monge and Carnot, in particular the works cited above [31, 159]. He attributes the long time between his first book (1822) and their revisions and his historical book which appeared some 40–50 years later to the extensive time commitment his administrative career demanded. He also published numerous mathematical papers early in his career, most of which play an important role in his books on geometry. Moreover, he was the author of several engineering monographs as well.

Poncelet promoted in his writings a specific doctrine for the development of geometry which became, with time, known as *projective geometry*, a term we didn't see him using in our reading of his works. He always used the expression "*propriétés projectives*" (projective properties) of figures, by which he meant those geometric properties of geometric figures that could be derived via projection methods by the new methodology that he was developing. He believed that the use of coordinate systems for the study of geometric problems was overvalued and led to difficulties of geometric understanding when negative and imaginary (complex) numbers appeared as solutions to equations. A real positive number could represent the length of a segment, or an area or volume of the figure, but what did negative and complex numbers represent geometrically? Only in the latter half of the nineteenth century did satisfactory answers to these questions arise.

As an example of this questioning of the use of algebra and geometry it is interesting to point to the opening 30 pages or so of the book by Carnot [31], which is solely dedicated to showing that negative numbers do not exist. Towards the end of his polemical assertions about the non-existence of negative numbers, Carnot introduces the notions of "direct" and "inverse" of numbers to justify the proliferation of plus and minus signs (the usual formulas of algebra and trigonometry) in his book. He also gives tantalizing hints of what became "analysis situs" in the late nineteenth century, which is now called simply *topology*. This early work of Carnot included specifically the notion of points and lines at infinity as well as the notion of duality and dual problems, which we will say more about later.

An important theorem of Apollonius that we discussed earlier in Sect. 1.2 asserts that any section of a skew cone could be considered as a section of a right circular cone, i.e. one of the classical conic sections (this is mathematically equivalent to the work of Newton and Fermat representing second-degree curves in the plane as one of the three classical conic sections). Now let's use the same picture but from a different perspective. Let E_1 and E_2 be two (non-parallel) planes in $\mathbf{R}^3$, let Γ_1 be a circle in E_1, let P be a point not on either plane, and let C be the cone formed by the pencil of lines emanating from P, the *point of perspective*, and passing through the points of Γ_1. Consider the intersection Γ_2 of the cone C with the second plane E_2. Then, according to Apollonius, this curve is again a conic section.

From the point of view of projective geometry; one says that the curves Γ_1 and Γ_2 are *projectively equivalent*, and that they represent the "same curve" projected onto different planes. One can imagine a figure being projected onto one plane from one perspective point and also being in the same plane from a different perspective point, and then from the second perspective point being projected onto a third version of the figure on a different plane. All of these changes of perspective and projections correspond to the natural mappings of projective space onto itself in modern terms, but this was the initial way these things were looked at by Poncelet and his school.

They also took the time to represent these changes of perspective and projection onto different planes as fractional-linear mappings of one plane onto another in terms of coordinate systems. For instance; in a supplementary article in Poncelet's book [189], written by one of his collaborators (M. Moutard, presumably a student), one finds the diagram on p. 512 (see Fig. 3.2) which shows a typical projection onto two different planes with coordinate systems (x, y) on the one plane and (x', y') on the second plane. After three pages of calculations; the author finds the coordinate transformations on p. 515 as given in Fig. 3.3.

This use of coordinate systems was used in this context as a way of making sure the computations done via projections (without coordinates) agreed with the results being obtained by Möbius, Plücker and others (with coordinates). In fact, the title of the article by Moutard (see p. 509 of [189]) is: "Rapprochement divers entre le

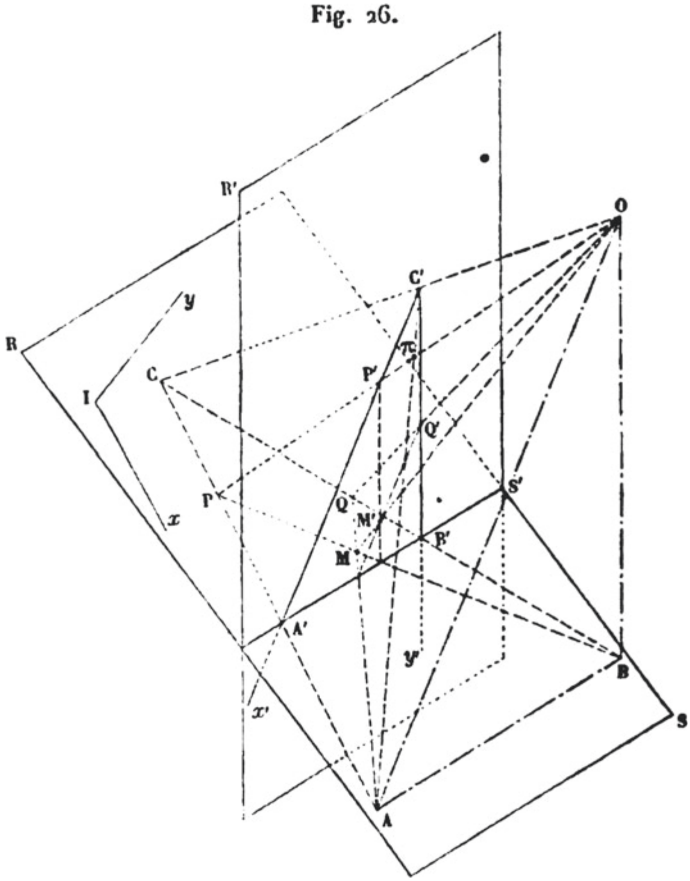


Fig. 3.2 Figure 26 in Poncelet's *Applications d'Analyse et Géométrie*, p. 512 [189]

$$(2) \qquad x = \frac{ax' + by' + c}{a''x' + b''y' + c''}, \qquad y = \frac{a'x' + b'y' + c'}{a''x' + b''y' + c''},$$

Fig. 3.3 Equation 2 on p. 615 in Poncelet's Fig. 1 in *Applications d'Analyse et Géométrie* [189]

principales méthodes de la géométrie pure et celles de l'analyse algébrique".[3] Here "*rapprochement*" means reconciliation, and note the use of "*géométrie pure*" for what later became known as synthetic projective geometry.

The equivalence of geometric objects as being different projections of the same object is one of the essential points of Poncelet's geometry. In the introductory remarks in his books, he formulates the basic principle that guides his investigations: given a particular geometric configuration and an associated problem in this context

[3] "Various reconciliations between the principal methods of pure geometry and those of analytic algebra".

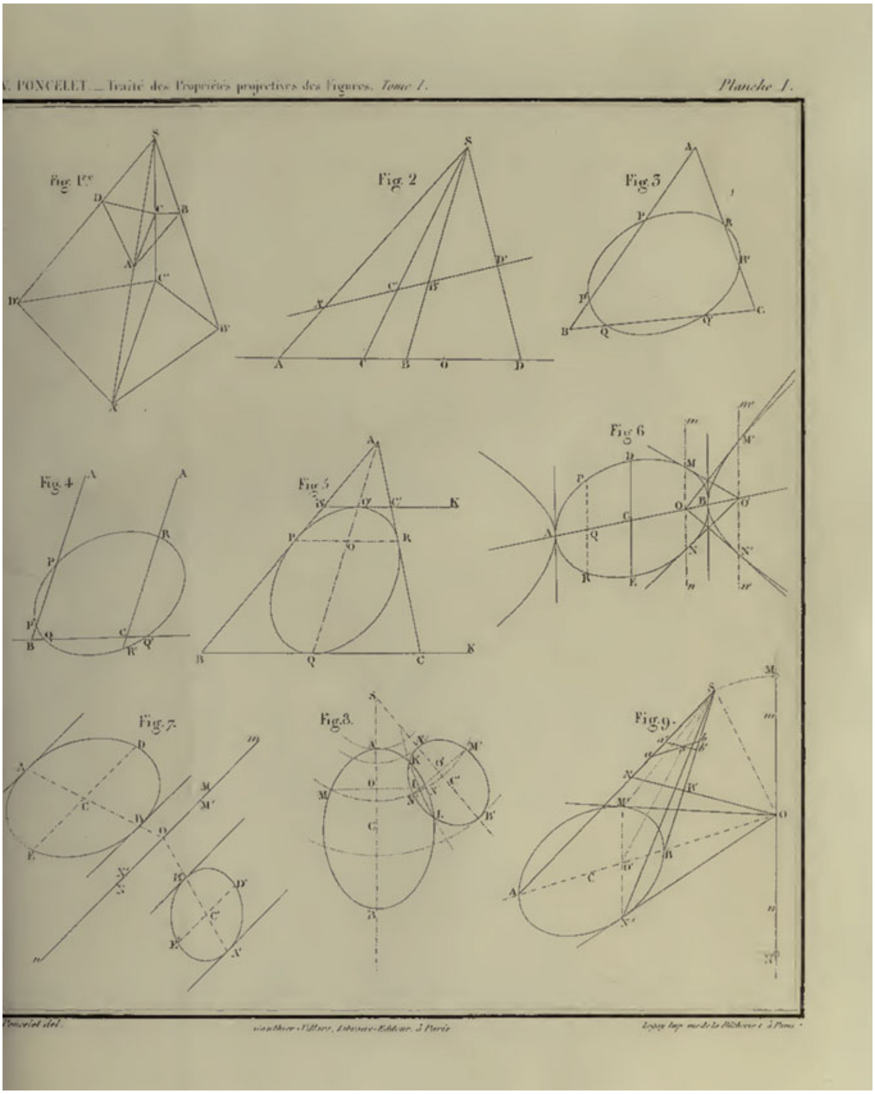


Fig. 3.4 Figures on Plate 1 in Poncelet's *Traité des Propriétés Projectives des Figures, Tome 1* [190]

(often in a planar context), find the simplest projective representation of this figure, by means of which one can resolve the problem, and thus it becomes resolved for all other configurations, including the one in the initial context. We want to illustrate this principle by looking at two simple examples of projectively equivalent figures from Poncelet's book [190] (Figs. 1 and 2 on plate 1, reproduced here in Fig. 3.4).

In the first figure (Fig. 1); we see that the quadrilateral (assumed in a plane) $ABCD$ is projectively equivalent to the quadrilateral $A'B'C'D'$ (again in a plane).

In Fig. 2 the segmented line $ABCD$ is projectively equivalent to the segmented line $A'B'C'D'$. Moreover, Poncelet (and many others in the time period) proved that the *cross-ratios* of these two sets of collinear points satisfy

$$\frac{AB}{CB}\frac{CD}{CE} = \frac{A'B'}{C'B'}\frac{C'D'}{A'D'}. \tag{3.1}$$

That is, this numerical quantity doesn't change for these variable projective views of this segmented interval.

It turns out that the cross-ratio is one of the most important numerical invariants of projective geometry. It was first described as lengths of intervals as in (3.1) (and was used by the Greeks in their mathematics work; see the book by Milne [153] for a history of its use in geometry). Initially it was simply the ratio of lengths of line segments and later was extended to lengths with a sign (when a line was assigned a direction and the order of the points determined the sign, as was introduced by Carnot [31]). This cross-ratio in this context was called in French *rapport anharmonique* and in German *Doppelverhältnis*. The modern English terminology *cross-ratio* was introduced by Clifford in his study of mechanics [50].

Most contemporary students of mathematics first come across the term cross-ratio in a first course on complex analysis, where for four points z_1, z_2, z_3, z_4 on the extended complex plane $\overline{\mathbf{C}}$ $(= \mathbf{C} \cup \infty)$, the cross-ratio is defined by[4]

$$\frac{z_1 - z_2}{z_1 - z_4}\frac{z_3 - z_4}{z_3 - z_2}.$$

This cross-ratio of four points in the extended complex plane is invariant under fractional-linear mapping transformations (Möbius transformations) of the form

$$w = \frac{az + b}{cz + d},$$

as one learns in a first course in complex analysis (see, e.g., Ahlfors [5]). These fractional-linear mapping transformations are simply reformulations of the projective-linear transformations of the projective space $\mathbf{P}_1(\mathbf{C}) \cong \overline{\mathbf{C}}$.

Poncelet used the word *homologous* to describe figures that are projectively equivalent in the sense that we are using here.[5] A second fundamental principle of Poncelet was that of the *loi de continuité* (law of continuity). On p. xiii of the introduction in his basic monograph [190] he formulates this principle as an "axiom" used by

[4]Note that there are 24 permutations of these four symbols, and each is called a cross-ratio and satisfies the properties outlined here. For each such permutation, there are three others with the same values, and hence there are 6 cross-ratios with distinct values. For more details see, for instance, the very informative book by Milne [153].

[5]Note the similarity to the modern use of the word "homologous" in topology, which was introduced by Poincaré [185]. In both cases these mathematicians wanted to describe two objects as being similar in their topological or graphical aspects, but not necessarily in their metric aspects (*à la* "congruence" or later "isometry").

illustrious mathematicians of the past. In fact, he formulates it as a question, as we see in this excerpt from this page:

> Considérons une figure quelconque, dans une position générale et en quelque sorte indéterminée, parmi toutes celles qu'elle peut prendre sans violer les lois, les conditions, la liaison qui subsistent entre les diverses parties du système; supposons que, d'après ces données, on ait trouvé une ou plusieurs relations ou propriétés, soit métriques, soit descriptives, appartenant à la figure, en s'appuyant sur le raisonnement explicite ordinaire, c'est-à dire par cette marche que, dans certains cas, on regarde comme seule rigoureuse. N'est-il pas évident que si, en conservant ces mêmes données, on vient à faire varier la figure primitive par degrés insensibles, ou qu'on imprime à certaines parties de cette figure un mouvement continu d'ailleurs quelconque, n'est-il pas évident que les propriétés et les relations, trouvées pour le premier système, demeureront applicables aux états successifs de ce système, pourvu toutefois qu'on ait égard aux modifications particulières qui auront pu y survenir, comme lorsque certaines grandeurs se seront évanouies, auront changé de sens ou de signe, etc., modifications qu'il sera toujours aisé de reconnaître *à priori*, et par des règles sûres?[6]

This is indeed the definition of the law of continuity Poncelet used in his book, and from our point of view, we would say that this definition is very *vague* (to be kind to M. Poncelet), but nineteenth-century authors used this principle in very many particular cases to derive new results which were later established by more rigorous means. There were indeed critics of this at the time, e.g. Cauchy was such a critic, and he, during the same period of time, was primarily responsible for formulating many of our current ideas concerning continuity for functions and mappings in general.

We want to give an important historical example which illustrates both projective equivalence and the law of continuity, and this is the well-known Desargues's theorem. In Fig. 3.5 we see an illustration of the theorem in the plane. Namely, if the two triangles are in perspective with respect to a perspective point (in this case the "center of perspectivity" in Fig. 3.5), then the intersections of the homologous line segments lie on a straight line (called the axis of perspectivity in the figure). This is Desargues's theorem, and the converse is also true. Let us use Fig. 22 of Poncelet's book ([190], reproduced in Fig. 3.6) to illustrate the proof. In this figure we see that the two triangles ABC and $A'B'C'$ are in perspective with respect to the point S and we assume that they are located in two non-parallel planes E and E' in $\mathbf{R}^3$. Since both of the lines represented by AB and $A'B'$ lie in the plane represented by the triangle ASB, then they must necessarily intersect (and for simplicity we assume the intersection is not at infinity). This intersection is denoted by MK (and the point

[6]"Consider an arbitrary figure in a general position and in some sense indeterminant, among all those that one can take without violating the laws, the conditions, the relations that takes place among the various parts of the system; suppose that, according to this given data, one can find one or more relations or properties, either metric or descriptive, pertaining to the figure, using ordinary explicit reasoning, that is to say, by those steps that, in certain cases, one regards as completely rigorous. Isn't it evident that if, in preserving the same given data, one can make the figure vary by imperceptible degrees, or where one imposes to certain parts of that figure a movement that is continuous and moreover arbitrary, isn't it evident that the properties and the relations, found for the first system, remain applicable to the successive states of the system, provided that one takes regard of particular modifications that could arise when certain quantities vanish, having a change of directions or sign, etc., modifications that will always be easy to recognize *à priori*, and according to well-determined rules?".

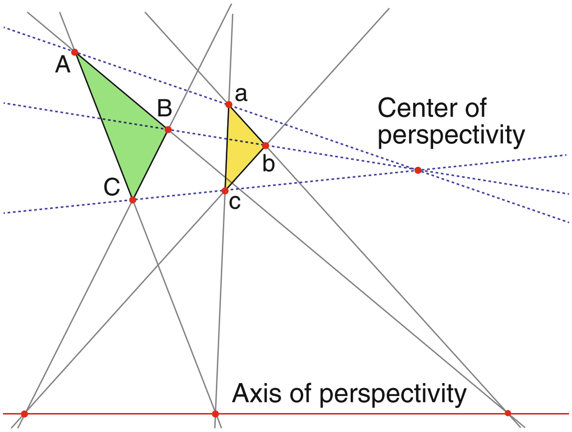


Fig. 3.5 Illustration of Desargues's theorem

M is not pictured in Fig. 22 in Fig. 3.6, since the scan of this page cut off some of the left-hand side of the page, but the intersection is clear on this page). This intersection point M lies on the intersection of the two planes $E \cap E'$, since $AB \subset E$ and $A'B' \subset E'$. The same is true for the intersections L of the lines represented by BC and $B'C'$ and for the intersection K of the lines represented by AC and $A'C'$. Thus K, L and M all lie on the intersection $E \cap E'$, which is a straight line. This is simply Desargues's theorem in this three-dimensional setting. Now suppose we have two triangles in perspective in a plane, and we envision them as being continuous limits of triangles in perspective in three-dimensional space, not in the same plane, as above, then the limit of the axes of perspectivity for the three-dimensional case will yield the desired axis of perspectivity in the two-dimensional case.

This fundamental result of projective geometry is due to Girard Desargues (1591–1661) in the seventeenth century. Unfortunately, most of his work was lost. However, fragments, reworkings and references to his work by others were discovered in the beginning of the nineteenth century, in particular after his theorem had been rediscovered by the projective geometry school. One work by Desargues is available today [54], which is a draft of three small articles, the first dealing with geometry.

The basic ideas of Desargues are contained in a series of books published from 1643 to 1648 by Abraham Bosse (1603–1676), who was a student of Desargues learning about architecture and engraving. We cite the last of this series here [24], as it is the most mathematical and contains at the end of the book a *proposition géométrique* which is Desargues's theorem as discussed above. Poncelet gives great credit to Desargues and Blaise Pascal (1623–1662), whose work we discuss below, for having the initial ideas in projective geometry (see, in particular, p. xxv in the Introduction in [190]).

The final principle of projective geometry we want to mention is that of *duality*. This was first introduced formally in the work of Carnot in 1803 [31]. In the simplest case in the plane, by using a nondegenerate quadratic form, one can associate in a

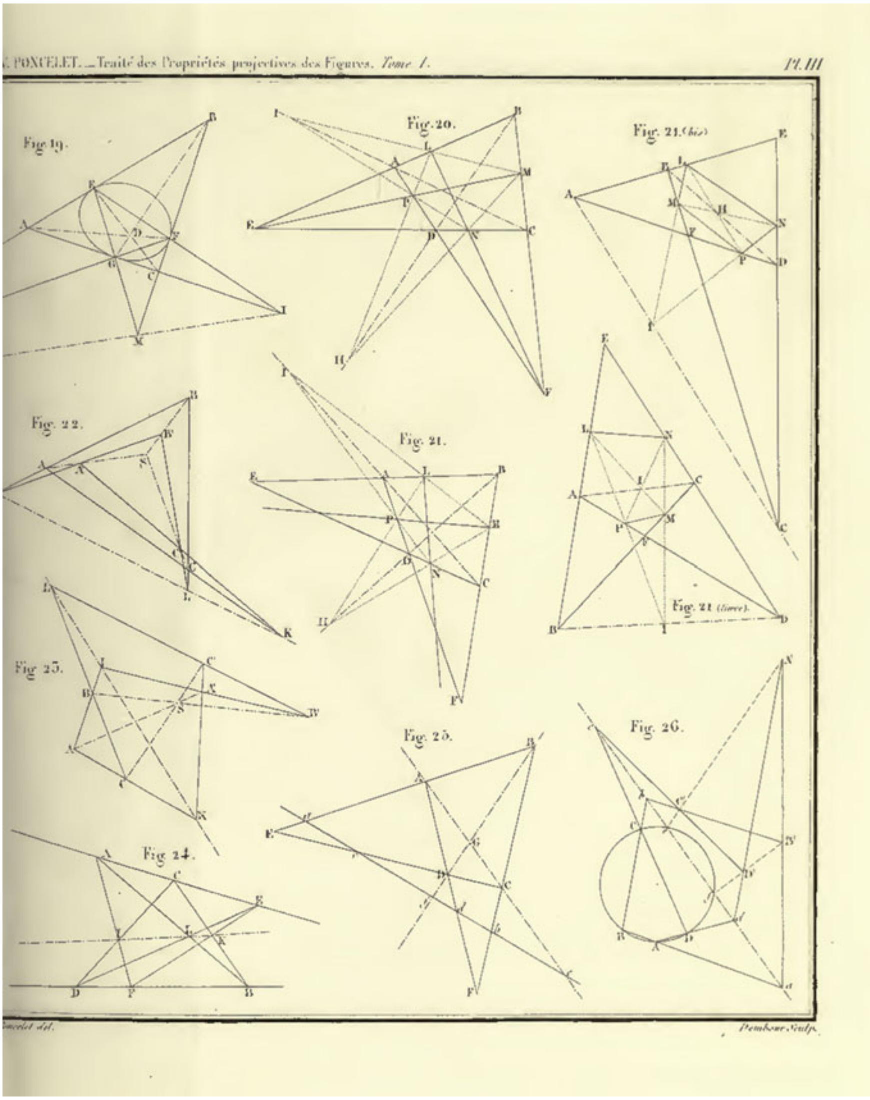


Fig. 3.6 Figures on Plate III in Poncelet's *Traité des Propriétés Projectives des Figures, Tome 1.* [190]

one-to-one manner points to lines and conversely. Namely, two points determine a line, and two lines intersect in a point (adding points at infinity here). Propositions utilizing these concepts have dual formulations.

A very classical example of this is illustrated by Pascal's theorem and its dual formulation due to Brianchon. Pascal's theorem asserts that if we consider a hexagram

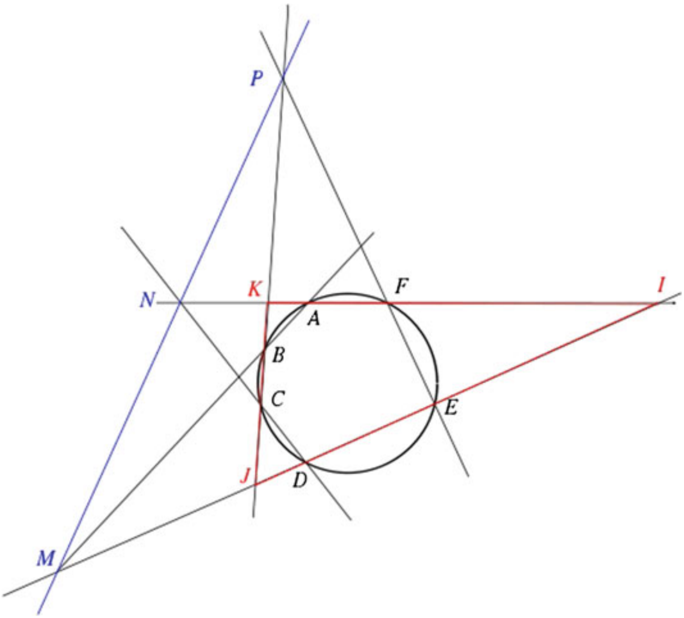


Fig. 3.7 Illustration of Pascal's theorem

inscribed in a conic section on a plane, then the intersections of the opposite sides of the hexagon lie on a straight line (see Fig. 3.7 for an illustration of the theorem).

The first published reference to Pascal's theorem is in his collected works, Vol. 5 [178], published in 1819. This volume contains a letter of Leibniz written in 1676 (pp. 429–431 in [178]) discussing Pascal's papers which he had access to, and including the statement and a proof of Pascal's theorem. Pascal had written some years earlier a set of essays on conic sections that Leibniz had access to and which he wrote about in this same letter.[7]

The dual of Pascal's theorem is the theorem of Brianchon (Charles-Julien Brianchon (1783–1864)) [27], which asserts that a circumscribing hexagon to a conic section has diagonals that intersect at a point (see Fig. 3.8 for an illustration of this theorem). Brianchon's monograph [28] published a decade later discusses this and related theorems and gives a succinct history of the subject, including references to the work of Pascal (the reference here is simply the title of the paper discussed by Leibniz) and Desargues (there was no mention of Pascal in his paper [27]). Note that the tangent lines in Fig. 3.8 correspond to the points on the conic section in Pascal's setting, and the intersections of the diagonals, a *point*, corresponds to the *line* on which the intersections of opposite sides in Pascal's setting lie.

[7]The editor (unnamed) of the volume [178] said that he searched for these papers referred to by Leibniz (written in the mid-seventeenth century) and was not able to find them for this volume of the collected works in 1819.

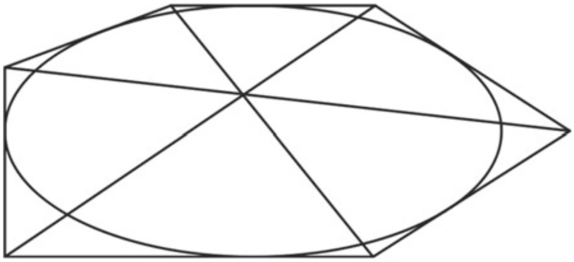


Fig. 3.8 Illustration of Brianchon's theorem

3.3 Analytic Projective Geometry

The final point we want to discuss with respect to projective geometry is the use of coordinate systems. In the first decades of the nineteenth century two schools of projective geometry developed in parallel. The first school, championed by Poncelet and what we might call the *French school*, developed what became known as *synthetic projective geometry*, which was a natural extension of classical Euclidean geometry in its methodology, notation, etc. (with the added emphasis on the new ideas in projective geometry, as outlined in the previous section). The second school was interested in using coordinate systems and analytic notation to prove the essential projective-geometric theorems. This was called *analytic projective geometry* (in the spirit of analytic geometry being the alternative to Euclidean geometry in schools). These two schools complemented each other, competed in a scholarly manner in the academic publications of the time, and learned from each other. The primary authors in the analytic schools happened to be German, and this could be called the *German school*. We will discuss several prominent German contributions in the following paragraphs, and we have mentioned the primary participants in the French school earlier.

August Ferdinand Möbius (1790–1868) is the first of the German school we want to mention. In his book published in 1827, *Der barycentrische Calcul*[8] [158], Möbius was very interested in showing that coordinate systems could be useful in proving the primary theorems of projective geometry (as well as some new results of his own). He saw the difficulty of using only the standard (x, y, z) coordinate system, which could often be too cumbersome or not as elegant as possible (one of the complaints of the French school). Möbius developed a new coordinate-type system which in its details is very much like the representation of vectors in a vector space using linear combinations of vectors. In $\mathbf{R}^3$ he took four points not lying on a plane and took "linear combinations" of these points (using the notion of center of mass from mechanics). Using the formalism he developed, he was able to successfully develop the fundamental results in projective geometry. Today his ideas are still used in barycentric subdivision in the triangulation of topological spaces, and it is a quite

[8] *The Barycentric Calculus.*

original, very readable, and interesting book to peruse. In this book he classifies geometric structures according to:

- *congruence* (he uses the word "*Gleichheit*"), that is, equivalent under Euclidean motions (translations and rotations),
- *affine equivalence*, that is, equivalence under translations and linear transformations of $\mathbf{R}^3$, and
- *collinearity*, equivalence under mappings preserving lines, projective equivalence as described earlier in the previous section.

The second author we turn to, whose work turned out to be very fundamental, is Julius Plücker (1801–1868). He wrote two books, *Analytisch-geometrische Entwicklungen* Vol. 1 [180] in 1828 and Vol. 2 [182] in 1831. In the first volume Plücker developed an abridged analytic notation to make the analytic proofs more transparent and more concise (abridgment of the classical coordinate system notation in $\mathbf{R}^3$). In Vol. 2 from 1831, he excitedly told his readers that he had developed yet another and new notation which simplified the understanding of projective geometry, and this was his introduction of *homogeneous coordinates*, which he had announced in a research paper in Crelle's Journal in 1830 [181].

Later, after a successful career in the spectroscopy of gases, he returned to the study of geometry and developed the notion of *line geometry*, the study of the lines in $\mathbf{R}^3$ as basic objects of study, and which could be parametrized as a quadric surface in $\mathbf{P}_5(\mathbf{R})$. This work was first developed in a paper published in 1865 [183] and developed into a two-part monograph as a new way of looking at geometrical space. This monograph [184] was published after Plücker's death and was edited by A. Clebsch with a great deal of assistance by the young Felix Klein, who was a student at the University of Bonn studying with Plücker at the time. The fundamental idea was to make lines and their parametrizations the principal object of study (in contrast to having points in space being the primary objects). One obtained points as intersections of lines, just as in classical geometry one obtained a line passing through two points. This led naturally (among many other things) to the definition of two-dimensional real projective space as the space of all lines passing through a given point in $\mathbf{R}^3$.

The most original contribution to the analytic side of projective geometry (indeed to geometry itself!) in this time period was that of Hermann Grassmann (1809–1877) [84], who, in a singular work, formulated what became known today as *exterior algebra*. His work also laid the foundation for the theory of vector spaces, one of the building blocks of exterior algebra. His work led to the development of Grassmannian manifolds, which regarded all of the planes of a fixed dimension in a Euclidean space as a geometric space, which was a generalization of the line geometry of Plücker. His lengthy philosophical introduction to his work is both a challenge and at the same time a pleasure to read. His work was not that well recognized in his lifetime (as so often happened in the history of mathematics), but turned out in the hands of Elie Cartan (1869–1951), for instance, to be a powerful tool for studying geometric problems in the twentieth century.

Finally, we note that Felix Klein in his writings put projective geometry on the firm footing we see today. In particular, his *Erlangen Program*[9] [121] from 1872 and his lectures on non-Euclidean geometry from the 1890s, published posthumously in 1928 [123], are marvels of exposition and give us, for instance, the use of the term *projective space* as a concept.

[9]This was written as his research program when he took up a new professorship in Erlangen in 1872.

Chapter 4
Gauss and Intrinsic Differential Geometry

4.1 Gaussian Curvature

In 1828 Gauss published his landmark paper concerning the differential geometry of surfaces entitled *Disquisitiones circa superficies curvas*[1] [81]. He had published the year before a very readable announcement and summary (written in German) [80] of his major results in the much longer paper [81], which was written in Latin. We will quote from this announcement paper somewhat later, letting Gauss tell us in his own words what he thinks the significance of his discoveries is. For the moment, we will simply say that this paper laid the foundation for doing intrinsic differential geometry on a surface and was an important first step in the creation of what has become known as Riemannian geometry on a manifold.

We now want to outline Gauss's study of curvature of a surface as presented in his original paper. We consider a locally defined surface S in $\mathbf{R}^3$ which, as Gauss points out, can be represented in one of three ways (he uses all three methods in his extensive computations). We will follow his notation so that the interested reader could refer to the original paper for more details. First we consider S as the zero set of a smooth function

$$w(x, y, z) = 0, \;\; w_x^2 + w_y^2 + w_z^2 \neq 0.$$

Secondly, we consider S as the graph of a function

$$z = f(x, y),$$

by the implicit function theorem.[2] Finally, we consider S to be the image of a parametric representation

[1]"Investigations of curved surfaces".

[2]Gauss used here simply $z = z(x, y)$, and we have modified his notation in this one instance for clarity.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_4

$$x = x(p, q),$$
$$y = y(p, q),$$
$$z = z(p, q).$$

In his computations, Gauss freely goes back and forth between the three representations, using what he needs and has developed previously.

At each point P of the surface S, there is a unit normal N_P associated with a given orientation of the surface, and we define the mapping

$$g : S \to S^2 \subset \mathbf{R}^3,$$

where S^2 is the standard unit sphere in $\mathbf{R}^3$,

$$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\},$$

and where

$$g(P) := N_P.$$

This mapping, first used in this paper, is now referred to as the *Gauss map*. If U is an open set in S, then the area on the two-sphere of $g(U)$ is called by Gauss the *total curvature* of U. For instance, if U is an open set in a plane, then all the unit normals to U would be the same, $g(U)$ would be a single point, its measure would be zero, and thus the total curvature of the planar set would be zero, as it should be.

Gauss then defines the *curvature* of S at a point P to be the *derivative of the mapping* $g : S \to S^2$ *at the point* P. He expresses this as the ratio of infinitesimal areas of the image to the infinitesimal area of the domain, and this is, as Gauss shows, the same as the Jacobian determinant of the mapping in the language of classical calculus. Today we refer to the curvature at a point of a surface as defined by Gauss to be the *Gaussian curvature*. Note that this definition is *a priori* extrinsic in nature, i.e., it depends on the surface being embedded in $\mathbf{R}^3$ so that the notion of a normal vector to the surface at a point makes sense.Gauss proceeds to compute the curvature of a given surface in each of the three representations above. In each case he expresses the normal vector in terms of the given data and explicitly computes the required Jacobian determinant. We will look at the first and third cases in somewhat more detail.

We start with the simplest case of a graph

$$z = f(x, y),$$

and let

$$df = t dx + u dy,$$

where $t = f_x$, $u = f_y$. Similarly, we have

$$dt = T dx + U dy,$$
$$du = U dx + V dy,$$

where

$$T := f_{xx}, U := f_{xy}, V := f_{yy}.$$

Thus the unit normal vectors to S have the form

$$(X, Y, Z) = (-tZ, -uZ, Z),$$

where

$$Z^2 = \frac{1}{1 + t^2 + u^2}.$$

By definition the curvature is the Jacobian determinant

$$k = \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} - \frac{\partial Y}{\partial x}\frac{\partial X}{\partial y},$$

which Gauss computes to be

$$k = \frac{TV - U^2}{(1 + t^2 + u^2)},$$

or in terms of f,

$$k = \frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2}, \tag{4.1}$$

and we see a strong similarity to the formula for the curvature of a curve in the case of a graph of a function as given by Newton (2.1).

Gauss considers the case where the tangent plane at (x_0, y_0) is perpendicular to the z-axis (i.e., $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$), obtaining

$$k = f_{xx} f_{yy} - f_{xy}^2,$$

and by making a rotation in the (x, y) plane to get rid of the term $f_{xy}(x_0, y_0)$, he obtains

$$k = f_{xx} f_{yy},$$

which he identifies as being the product of the two principal curvatures of Euler (as discussed in Sect. 2.4).

Next Gauss proceeds to compute the curvature of S in terms of an implicit representation of the form $w(x, y, z) = 0$, and he obtains a complicated expression which we won't reproduce here. It has the form

$$(w_x^2 + w_y^2 + w_z^2)k = h(w_x, w_y, w_z, w_{xx}, w_{xy}, w_{yy}, w_{xz}, w_{yz}, w_{zz}),$$

where h is an explicit homogeneous polynomial of degree 4 in 9 variables. Here, of course, $w_x^2 + w_y^2 + w_z^2 \neq 0$ on S, so this expresses k as a rational function of the derivatives of first and second order of w, similar to (4.1) above. This explicit formula is given on p. 232 of [81].

Then Gauss moves on to the representation of curvature in terms of a parametric representation of the surface. This we present in more detail, as we did with the first case. He starts with the representation of the local surface as

$$\begin{aligned} x &= x(p, q), \\ y &= y(p, q), \\ z &= z(p, q), \end{aligned}$$

and he gives notation for the first and second derivatives of these functions. Namely, let

$$\begin{aligned} dx &= adp + a'dq, \\ dy &= bdp + b'dq, \\ dz &= cdp + c'dq, \end{aligned}$$

where, of course, $a = x_p(p, q)$, $b = y_p(p, q)$, etc., using the subscript notation for partial derivatives. In the same fashion we define the second derivatives

$$\begin{array}{l} \alpha := x_{pp},\ \alpha' := x_{pq},\ \alpha'' := x_{qq}, \\ \beta := y_{pp},\ \beta' := y_{pq},\ \beta'' := y_{qq}, \\ \gamma := z_{pp},\ \gamma' := z_{pq},\ \gamma'' := z_{qq}. \end{array}$$

Now the vectors (a, b, c), (a', b', c') in $\mathbf{R}^3$ represent tangent vectors to S at $(x(p, q), y(p, q), q(p, q)) \in S$, and thus the vector

$$(A, B, C) := (bc' - cb', ca' - ac', ab' - ba') \tag{4.2}$$

is a normal vector to S at the same point (cross product of the two tangent vectors). Hence one obtains

$$Adx + Bdy + Cdz = 0$$

on S, and therefore we can write

$$dz = -\frac{A}{C}dx - \frac{B}{C}dy,$$

assuming that $C \neq 0$, i.e., we have the graphical representation $z = f(x, y)$, as before. Thus

$$t := z_x = -\frac{A}{C}, \tag{4.3}$$

$$u := z_y = -\frac{B}{C}. \tag{4.4}$$

Now since

$$dx = adp + a'dq,$$
$$dy = bdb + b'dq,$$

we can solve these linear equations for dp and dq, obtaining (recalling that C is defined in (4.2))

$$Cdp = b'dx - a'dy, \tag{4.5}$$

$$Cdq = -bdx + ady. \tag{4.6}$$

Now we differentiate the two Eqs. (4.3) and (4.4), and using (4.5) and (4.6) one obtains

$$C^3dt = \left(A\frac{\partial C}{\partial p} - C\frac{\partial A}{\partial p}\right)(b'dx - a'dy) + \left(C\frac{\partial A}{\partial q} - A\frac{\partial C}{\partial q}\right)(bdx - ady),$$

and one can derive a similar expression for C^3du.

Using the expression for the curvature in the graphical case (4.1), and by setting

$$D = A\alpha + B\beta + C\gamma,$$
$$D' = A\alpha' + B\beta' + C\gamma',$$
$$D'' = A\alpha'' + B\beta'' + C\gamma'',$$

Gauss obtains the formula for the curvature in this case

$$k = \frac{DD'' - (D')^2}{(A^2 + B^2 + C^2)^2}, \tag{4.7}$$

where again this is a rational function of the first and second derivatives of the parametric functions $x(p, q)$, $y(p, q)$, $z(p, q)$. Note that the numerator is homogeneous of degree 4 in this case, and again this formula depends on the explicit representation of S in $\mathbf{R}^3$. One might think that he was finished at this point, having obtained three different representations for curvature corresponding to the three representations of the locally defined surface. But he goes on to make one more very ingenious change of variables that leads to his celebrated discovery.

4.2 Gauss's *Theorema Egregrium*

Continuing with the computations of the previous section, Gauss defines new functions of the previously defined functions of the first and second derivatives of the defining functions (in terms of the various representations of the surface). We simply list them here as Gauss did in his paper. He defines

$$\begin{aligned}
E &= a^2 + b^2 + c^2, \\
F &= aa' + bb' + cc', \\
G &= (a')^2 + (b')^2 + (c')^2, \\
\Delta &= A^2 + B^2 + C^2 = EG - F^2,
\end{aligned}$$

and additionally,

$$\begin{aligned}
m &= a\alpha + b\beta + c\gamma, \\
m' &= a\alpha' + b\beta' + c\gamma', \\
m'' &= a\alpha'' + b\beta'' + c\gamma'', \\
n &= a'\alpha + b'\beta + c'\gamma, \\
n' &= a'\alpha' + b'\beta' + c'\gamma', \\
n'' &= a'\alpha'' + b'\beta'' + c'\gamma''.
\end{aligned}$$

He is able to show that (using the subscript notation for differentiation again, i.e., $E_p = \frac{\partial E}{\partial p}$, etc.)

$$\begin{aligned}
E_p &= 2m, & E_q &= 2m', \\
F_p &= m' + n, & F_q &= m' + n', \\
G_p &= 2n', & G_q &= 2n'',
\end{aligned}$$

from which one obtains, by solving these linear equations,

$$\begin{aligned}
m &= \tfrac{1}{2}E_p, & n &= F_o - \tfrac{1}{2}E_q, \\
m' &= \tfrac{1}{2}E_q, & n' &= \tfrac{1}{2}G_p, \\
m'' &= F_q - \tfrac{1}{2}G_p, & n'' &= \tfrac{1}{2}G_q.
\end{aligned}$$

Gauss also showed that the numerator which appears in (4.7) can be expressed in terms of these new variables as

$$\begin{aligned} DD'' - (D')^2 = \Delta[&\alpha\alpha'' - \beta\beta'' + \gamma\gamma'' - (\alpha')^2 - (\beta')^2 - (\gamma')^2] \\ &+E([(n')^2 - nn''] + F[nm'' - 2m'n' + mn''] \\ &+G[(m')^2 - mm'']. \end{aligned} \tag{4.8}$$

Finally, one confirms that

$$\begin{aligned} \frac{\partial n}{\partial q} - \frac{\partial n'}{\partial p} &= \frac{\partial m''}{\partial p} - \frac{m'}{\partial q} \\ &= \alpha\alpha'' - \beta\beta'' + \gamma\gamma'' - (\alpha')^2 - (\beta')^2 - (\gamma')^2. \end{aligned} \tag{4.9}$$

Substituting (4.9) and (4.8) into the curvature formula (4.7) Gauss obtains the expression for the curvature he was looking for:

$$\begin{aligned} 4(EG - FF)^2 k = {} & E(E_q G_q - 2F_p G_q + G_p^2) \\ &+F(E_p G_q - E_q G_p - 2E_q F_q + 4F_p F_q - 2F_p G_p) \\ &+G(E_p G_p - 2E_p F_q + E_q^2) \\ &-2(EG - FF)(E_{qq} - 2F_{pq} + G_{pp}). \end{aligned} \tag{4.10}$$

If we consider the metric on S induced by the metric on $\mathbf{R}^3$, we have, in terms of the parametric representation,

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= (adp + a'dq)^2 + (bdp + b'dq)^2 + (cdp + c'd1)^2 \\ &= Edp^2 + 2Fdpdq + Gdq^2. \end{aligned} \tag{4.11}$$

We can now observe that in (4.10) Gauss has managed to express the curvature of the surface in terms of E, F, G and their first and second derivatives with respect to the parameters of the surface; that is, the curvature is a function of the *line element* (4.11) and its first and second derivatives.[3]

Gauss called this his *Theorema Egregrium* (see pp. 236 and 237 from [81] given in Figs. 4.1 and 4.2, where the formula (4.10) is on p. 236 and the statement of the theorem is on p. 237). What the text on p. 237 of Gauss's paper says, in so many words, is that if two surfaces can be represented by the same parametrization such that the induced metrics are the same, then the curvature is preserved. In more modern language, an isometric mapping of one surface to another will preserve the Gaussian curvature.

Gauss understood full well the significance of his work and the fact that this was the beginning of the study of a new type of geometry (which later generations have

[3] We note that the three representations of curvature in Sect. 4.1 depend on first and second derivatives of the defining functions for the surface, whereas the Gaussian curvature in (4.10) depends on three derivatives of the defining functions. This comes about since the change of variables in (4.9) involves first derivatives.

Combinatio huius aequationis cum aequatione (10) producit

$$DD''-D'D' = (\alpha\alpha''+\beta\beta''+\gamma\gamma''-\alpha'\alpha'-\beta'\beta'-\gamma'\gamma')\Delta$$
$$+E(n'n'-nn'')+F(nm''-2m'n'+mn'')+G(m'm'-mm'')$$

Iam patet esse

$$\frac{dE}{dp}=2m,\quad \frac{dE}{dq}=2m',\quad \frac{dF}{dp}=m'+n,\quad \frac{dF}{dq}=m''+n',\quad \frac{dG}{dp}=2n',\quad \frac{dG}{dq}=2n''$$

sive

$$m=\tfrac{1}{2}\frac{dE}{dp},\qquad m'=\tfrac{1}{2}\frac{dE}{dq},\qquad m''=\frac{dF}{dq}-\tfrac{1}{2}\frac{dG}{dp}$$
$$n=\frac{dF}{dp}-\tfrac{1}{2}\frac{dE}{dq},\qquad n'=\tfrac{1}{2}\frac{dG}{dp},\qquad n''=\tfrac{1}{2}\frac{dG}{dq}$$

Porro facile confirmatur, haberi

$$\alpha\alpha''+\beta\beta''+\gamma\gamma''-\alpha'\alpha'-\beta'\beta'-\gamma'\gamma' = \frac{dn}{dq}-\frac{dn'}{dp}=\frac{dm''}{dp}-\frac{dm'}{dq}$$
$$=-\tfrac{1}{2}\cdot\frac{ddE}{dq^2}+\frac{ddF}{dp.dq}-\tfrac{1}{2}\cdot\frac{ddG}{dp^2}$$

Quodsi iam has expressiones diversas in formula pro mensura curvaturae in fine art. praec. eruta substituimus, pervenimus ad formulam sequentem, e solis quantitatibus E, F, G atque earum quotientibus differentialibus primi et secundi ordinis concinnatam:

$$4(EG-FF)^2k = E\left(\frac{dE}{dq}\cdot\frac{dG}{dq}-2\frac{dF}{dp}\cdot\frac{dG}{dq}+\left(\frac{dG}{dp}\right)^2\right)$$
$$+F\left(\frac{dE}{dp}\cdot\frac{dG}{dq}-\frac{dE}{dq}\cdot\frac{dG}{dp}-2\frac{dE}{dq}\cdot\frac{dF}{dq}+4\frac{dF}{dp}\cdot\frac{dF}{dq}-2\frac{dF}{dp}\cdot\frac{dG}{dp}\right)$$
$$+G\left(\frac{dE}{dp}\cdot\frac{dG}{dp}-2\cdot\frac{dE}{dp}\cdot\frac{dF}{dq}+\left(\frac{dE}{dq}\right)^2\right)$$
$$-2(EG-FF)\left(\frac{ddE}{dq^2}-2\frac{ddF}{dp.dq}+\frac{ddG}{dp^2}\right)$$

12.

Quum indefinite habeatur

$$dx^2+dy^2+dz^2 = Edp^2+2Fdp.dq+Gdq^2$$

patet, $\sqrt{(Edp^2+2Fdp.dq+Gdq^2)}$ esse expressionem generalem elementi linearis in superficie curva. Docet itaque analysis in art. praec. explicata, ad inveniendam mensuram curvaturae haud opus esse formulis finitis, quae coordina-

Fig. 4.1 The formula for Gaussian curvature in its intrinsic form

CIRCA SUPERFICIES CURVAS. 237

tas x, y, z tamquam functiones indeterminatarum p, q exhibeant, sed sufficere expressionem generalem pro magnitudine cuiusvis elementi linearis. Progrediamur ad aliquot applicationes huius gravissimi theorematis.

Supponamus, superficiem nostram curvam explicari posse in aliam superficiem, curvam seu planam, ita ut cuivis puncto prioris superficiei per coordinatas x, y, z determinato respondeat punctum determinatum superficiei posterioris, cuius coordinatae sint x', y', z'. Manifesto itaque x', y', z' quoque considerari possunt tamquam functiones indeterminatarum p, q, unde pro elemento $\sqrt{(\mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2)}$ prodibit expressio talis

$$\sqrt{(E'\mathrm{d}p^2 + 2F'\mathrm{d}p\,.\,\mathrm{d}q + G'\mathrm{d}q^2)}$$

denotantibus etiam E', F', G' functiones ipsarum p, q. At per ipsam notionem *explicationis* superficiei in superficiem patet, elementa in utraque superficie correspondentia necessario aequalia esse, adeoque identice fieri

$$E = E', \quad F = F', \quad G = G'$$

Formula itaque art. praec. sponte perducit ad egregium

THEOREMA. *Si superficies curva in quamcunque aliam superficiem explicatur, mensura curvaturae in singulis punctis invariata manet.*

Manifesto quoque *quaevis pars finita superficiei curvae post explicationem in aliam superficiem eandem curvaturam integram retinebit.*

Casum specialem, ad quem geometrae hactenus investigationes suas restrinxerunt, sistunt superficies in planum explicabiles. Theoria nostra sponte docet, talium superficierum mensuram curvaturae in quovis puncto fieri $= 0$, quocirca, si earum indoles secundum modum tertium exprimitur, ubique erit

$$\frac{\mathrm{dd}z}{\mathrm{d}x^2} \cdot \frac{\mathrm{dd}z}{\mathrm{d}y^2} - \left(\frac{\mathrm{dd}z}{\mathrm{d}x\,.\,\mathrm{d}y}\right)^2 = 0$$

quod criterium, dudum quidem notum, plerumque nostro saltem iudicio haud eo rigore qui desiderari posset demonstratur.

13.

Quae in art. praec. exposuimus, cohaerent cum modo peculiari superficies considerandi, summopere digno, qui a geometris diligenter excolatur. Scilicet quatenus superficies consideratur non tamquam limes solidi, sed tamquam soli-

Fig. 4.2 The *Theorema Egregrium* of Gauss called *intrinsic differential geometry*). We quote here from his announcement of his results published some months earlier (in his native German) from pp. 344–345 of [80]:

> Diese Sätze führen dahin, die Theorie der krummen Flächen aus einem neuen Gesichtspunkte zu betrachten, wo sich der Untersuchung ein weites noch ganz unangebautes Feld öffnet. Wenn man die Flächen nicht als Grenzen von Körpern, sondern als Körper, deren eine Dimension verschwindet, und zugleich als biegsam, aber nicht als dehnbar betrachtet, so begreift man, dass zweierlei wesentlich verschiedene Relationen zu unterscheiden sind, theils nemlich solche, die eine bestimmte Form der Fläche im Raum voraussetzen, theils solche, welche von den verschiedenen Formen, die die Fläche annehmen kann, unabhängig sind. Die letztern sind es, wovon hier die Rede ist: nach dem, was vorhin bemerkt ist, gehört dazu das Krümmungsmaass; man sieht aber leicht, dass eben dahin die Betrachtung der auf der Fläche construirten Figuren, ihrer Winkel, ihres Flächeninhalts und ihrer Totalkrümmung, die Verbindung der Punkte durch kürzeste Linien u. dgl. gehört. [4]

The results described in this short announcement (and the details in the much longer Latin papers on the subject) formed the basis of most of what became modern differential geometry. An important point that we should make here is that Gauss did significant experimental work on measuring the curvature of the earth in the area around Göttingen, where he spent his whole scientific career. This involved measurements over hundreds of miles, and involved communicating between signal towers from one point to another. He developed his theory of differential geometry as he was conducting the experiments, and at the end of all of his papers on this subject one finds descriptions of the experimental results (including, in particular, the short paper in German [80]).

[4] "These theorems lead us to consider the theory of curved surfaces from a new point of view, whereby the investigations open to a quite new undeveloped field. If one doesn't consider the surfaces as boundaries of domains, but as domains with one vanishing dimension, and at the same time as bendable but not as stretchable, then one understands that one needs to differentiate between two different types of relations, namely, those which assume the surface has a particular form in space, and those that are independent of the different forms a surface might take. It is this latter type that we are talking about here. From what was remarked earlier, the curvature belongs to this type of concept, and moreover figures constructed on the surface, their angles, their surface area, their total curvature as well as the connecting of points by curves of shortest length, and similar concepts, all belong to this class."

Chapter 5
Riemann's Higher-Dimensional Geometry

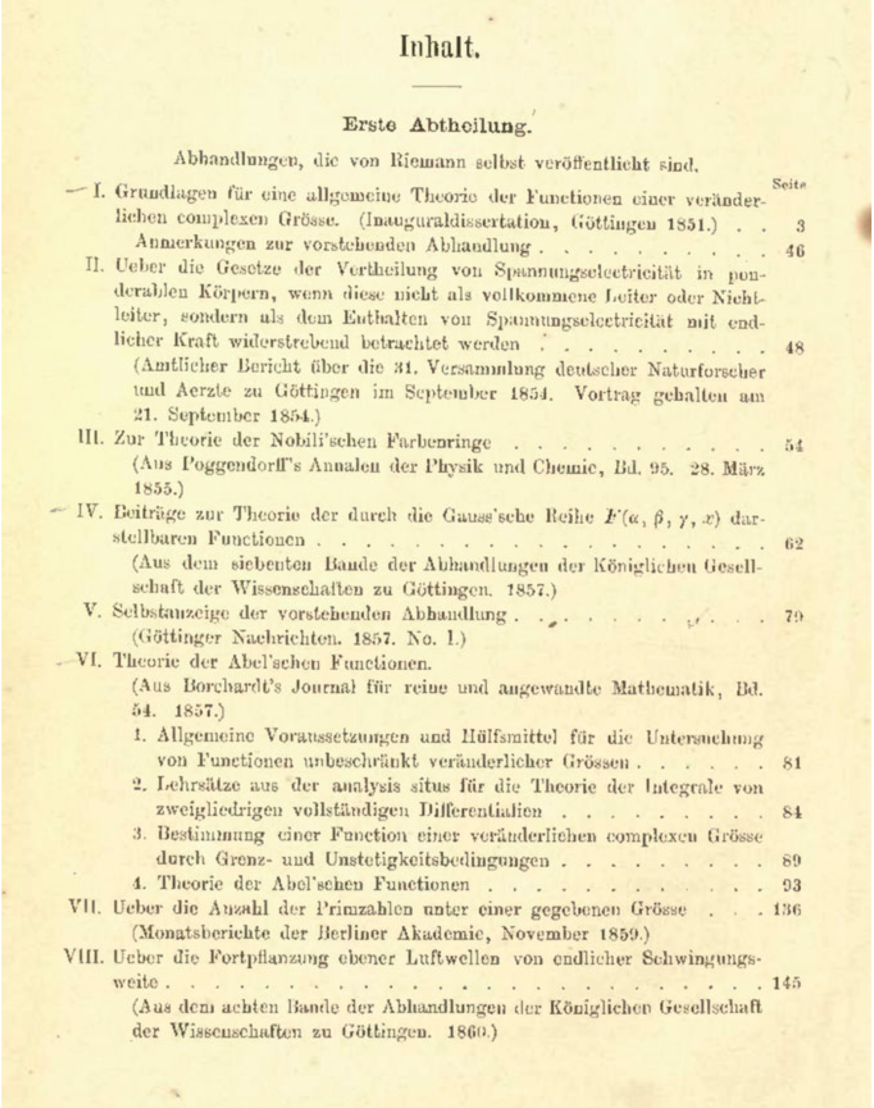
5.1 The Legacy of Riemann

In mathematics we sometimes see striking examples of brilliant contributions or completely new ideas that change the ways mathematics develops in a significant fashion. A prime example of this is the work of Descartes [55], which completely changed how mathematicians looked at geometric problems. But it is rare that a single mathematician makes as many singular advances in his lifetime as did Bernhard Riemann in the middle of the nineteenth century. In this section we will discuss in some detail his fundamental creation of the theory of higher-dimensional manifolds and the additional creation of what is now called Riemannian geometry. In Part III we will review his contributions to complex analysis and complex geometry.

However, it is worth noting that he only published nine papers in his short lifetime (he lived to be only 40 years old); and several other important works, including those that concern us in this section, were published posthumously from the writings he left behind. His collected works (including in particular these posthumously published papers) were edited and published in 1876 and are still in print today [200].

In Figs. 5.1 and 5.2 we have reproduced the table of contents of Riemann's collected works [200]. Looking through the titles one is struck by the wide diversity as well as the originality. Let us give a few examples here. In Paper I (his dissertation) he formulated and proved the Riemann mapping theorem and dramatically moved the theory of functions of one complex variables in new directions. In Paper VI, in order to study Abelian functions, he formulated what became known as Riemann surfaces and this led to the general theory of complex manifolds in the twentieth century. In Paper VII he introduced the Riemann zeta function as a tool for studying the Prime Number Theorem and formulated the Riemann hypothesis, which is surely the outstanding mathematical problem in the world today. In Paper XII he formulated the first rigorous definition of a definite integral (the Riemann integral) and applied it to trigonometric series, setting the stage for Lebesgue and others in the early twentieth century to develop many consequences of the powerful theory of Fourier analysis. In Papers XIII and XXII he formulated the theory of higher-dimensional manifolds,

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_5



Inhalt.

Erste Abtheilung.

Abhandlungen, die von Riemann selbst veröffentlicht sind.

Fig. 5.1 Table of contents, p. VI, Riemann's collected works [200]

including the important concepts of Riemannian metric, normal coordinates and the Riemann curvature tensor, which we will visit very soon in the sections below. Paper XVI contains correspondence with Enrico Betti leading to the first higher-dimensional topological invariants beyond those Riemann had earlier developed for two-dimensional manifolds.

Inhalt. VII

Fig. 5.2 Table of contents, p. II, Riemann's collected works [200]

This will suffice. The reader can glance at the other titles to see their further diversity. His contributions to the theory of partial differential equations and various problems in mathematical physics were also quite significant.

5.2 Higher-Dimensional Manifolds and a Quadratic Line Element

Riemann's paper "Ueber die Hypothesen, welche der Geometrie zu Grunde liegen"[1] [201] (Paper XIII above) is a posthumously published version of a public lecture Riemann gave as his *Habilitationsvortrag* in 1854. This was part of the process for obtaining his *Habilitation*, a German advanced degree beyond the doctorate necessary to qualify for a professorship in Germany at the time (such requirements are still in place at most German universities today as well as in other European countries, e.g., France and Russia; it is similar to the research requirements in the US to be qualified for tenure). This paper, being a public lecture, has very few formulas, is at times quite philosophical and is amazing in its depth of vision and clarity. On the other hand, it is quite a difficult paper to understand in detail, as we shall see.

Before this paper was written, manifolds were all one- or two-dimensional curves and surfaces in $\mathbf{R}^3$, including their extension to points at infinity, as discussed in Chap. 3. In fact, some mathematicians who had to study systems parametrized by more than three variables declined to call the parametrization space a manifold or give such a parametrization a geometric significance. In addition, these one- and two-dimensional manifolds always had a differential geometric structure which was induced by the ambient Euclidean space (this was true for Gauss, as well).

In Riemann's paper [201] he discusses the distinction between discrete and continuous manifolds, where one can make comparisons of quantities by either counting or by measurement, and gives a hint, on p. 256, of concepts from set theory, which was only developed later in a single-handed effort by Cantor. Riemann begins his discussion of manifolds by moving a one-dimensional manifold, which he intuitively describes, in a transverse direction (moving in some type of undescribed ambient "space") to obtain a surface, and inductively, generating an n-dimensional manifold by moving an $n-1$-manifold transversally in the same manner. Conversely, he discusses having a nonconstant function on an n-dimensional manifold, and the set of points where the function is constant is (generically) a lower-dimensional manifold; and by varying the constant, one obtains a one-dimensional family of $n-1$-manifolds (similar to his construction above).[2]

Riemann formulates local coordinate systems $(x^1, x^2, \dots, x^n)$ on a manifold of n dimensions near some given point, taken here to be the origin. He formulates

[1] "On the hypotheses, which are the basis for geometry".

[2] He alludes to some manifolds that cannot be described by a finite number of parameters; for instance, the manifold of all functions on a given domain, or all deformations of a spatial figure. Infinite-dimensional manifolds, such as these, were studied in great detail a century later.

a curve in the manifold as being simply n functions $(x^1(t), x^2(t), \ldots, x^n(t))$ of a single variable t. The concepts of set theory and topological space were developed only later in the nineteenth century, and so the global nature of manifolds is not really touched on by Riemann (except in his later work on Riemann surfaces and his correspondence with Betti, mentioned above). It seems clear on reading his paper that he thought of n-dimensional manifolds as being extended beyond Euclidean space in some manner, but the language for this was not yet available.

At the beginning of this paper Riemann acknowledges the difficulty he faces in formulating his new results. Here is a quote from the second page of his paper (p. 255) in [200]:

> Indem ich nun von diesen Aufgaben zunächst die erste, die Entwicklung des Begriffs mehrfach ausgedehnter Grössen, zu lösen versuche, glaube ich um so mehr auf eine nachsichtige Beurtheilung Anspruch machen zu dürfen; da ich in dergleichen Arbeiten philosophischer Natur, wo die Schwierigkeiten mehr in den Begriffen, als in der Construction liegen, wenig geübt bin und ich ausser einigen ganz kurzen Andeutungen, welche Herr Geheimer Hofrath Gauss in der zweiten Abhandlung über die biquadratischen Reste in den Göttingenschen gelehrten Anzeigen und in seiner Jubiläumsschrift darüber gegeben hat, und einigen philosophischen Untersuchungen Herbart's, durchaus keine Vorarbeiten benutzen konnte.[3]

The paper of Gauss that he cites here [78] refers to Gauss's dealing with the philosophical issue of understanding the complex number plane after some thirty years of experience with its development. We will mention this paper more explicitly in Sect. 6.3. Hebart was a philosopher whose metaphysical investigations influenced Riemann's thinking. Riemann was very aware of the speculative nature of his theory, and he used this philosophical point of view, as the technical language he needed (set theory and topological spaces) was not yet available. This was very similar to Gauss's struggle with the complex plane, as we shall see later.

As mentioned earlier, measurement of the length of curves goes back to the Archimedean study of the length of a circle. The basic idea there and up to the work of Gauss was to approximate a given curve by straight line segments and take a limit. The *length* of each straight line segment was determined by the Euclidean ambient space, and the formula, using calculus for the limiting process, became, in the plane for instance,

$$\int_\Gamma ds = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}dt,$$

[3] "In that my first task is to try to develop the concept of a multiply spread-out quantity [he uses the word 'Mannigfaltikeit' (manifold) later], I believe even more in being allowed an indulgent evaluation, as in such works of a philosophical nature, where the difficulties are more in the concepts than in the construction, wherein I have little experience, and except for the paper by Mr. Privy Councilor Gauss in his second commentary on biquadratic residues in the Göttingen gelehrten Anzeigen [1831] and in his Jubiläumsschrift and some investigations by Hebart, I have no precedents I could use."

where $ds^2 = dx^2 + dy^2$ is the line element of arc length in $\mathbf{R}^2$. As we saw in Chap. 4, Gauss formulated in [81] on a two-dimensional manifold with coordinates (p, q) the line element

$$ds^2 = Edp^2 + 2Fdpdq + Gdq^2, \tag{5.1}$$

where E, F, and G are induced from the ambient space. He didn't consider any examples of such a line element (5.1) that weren't induced from an ambient Euclidean space, but his remarks (see Gauss's quote in Sect. 4.2) clearly indicate that this could be a ripe area for study, and this could well include allowing coefficients of the line element (5.1) to be more general than induced from an ambient space.

Since Riemann formulated an abstract n-dimensional manifold (with a local coordinate system) with no ambient space, and since he wanted to be able to measure the length of a curve on his manifold, he formulated, or rather postulated, an independent measuring system which mimics Gauss's formula (5.1). Namely, he prescribes for a given local coordinate system a metric (line element) of the form

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}(x)dx^i dx^j, \tag{5.2}$$

where $g_{ij}(x)$ is, for each x, a symmetric positive-definite matrix, and he postulates, by the usual change of variables formulas,

$$ds^2 = \sum_{i,j=1}^{n} \tilde{g}_{ij}(\tilde{x})d\tilde{x}^i d\tilde{x}^j, \tag{5.3}$$

where $\tilde{g}_{ij}(\tilde{x})$ is the transformed positive-definite matrix in the new coordinate system $(\tilde{x}_1, \ldots, \tilde{x}^n)$. This has the form

$$g_{kl}(x) = \sum_{ij} \tilde{g}_{ij}(\tilde{x}(x))\frac{\partial \tilde{x}_i}{\partial x_k}\frac{\partial \tilde{x}_j}{\partial x_l}. \tag{5.4}$$

Using the line element (5.2), the length of a curve is defined by

$$l(\Gamma) := \int_a^b \sqrt{\sum_{i,j=1}^{n} g_{ij}(x(t))\frac{dx^i}{dt}(t)\frac{dx^j}{dt}(t)dt}.$$

The line element (5.2) is what is called a *Riemannian metric* today, and the two-form ds^2 is considered as a positive-definite bilinear form giving an inner product on the tangent space $T_p(M)$ for p a point on the manifold M. This has become the basis for almost all of modern differential geometry (with the extension to Lorentzian type spaces where $g_{ij}(x)$ is not positive-definite *à la* Minkoswki space). Riemann merely says on p. 260 of his paper (no notation here at all),

> ich beschränke mich daher auf die Mannigfaltigkeiten, wo das Linienelement durch die Quadratwurzel aus einem Differentialausdruck zweiten Grades ausgedrückt wird.[4]

Earlier he had remarked that a line element should be homogeneous of degree 1 and one could also consider the fourth root of a differential expression of fourth degree, for instance. Hence his restriction in the quote above.

5.3 Geodesic Normal Coordinates and a Definition of Curvature

The next step in Riemann's paper is his formulation of curvature. This occurs on a single page (p. 261 of [201], which we reproduce here in Fig. 5.3). It is extremely dense and not at all easy to understand. In the published collected works of Riemann one finds an addendum to Riemann's paper which analyzes this one page in seven pages of computations written by Julius Wilhelm Richard Dedekind (1831–1916). This is an unpublished manuscript that appeared only in these collected works of Riemann, pp. 384–391. In Volume 2 of Spivak's three-volume comprehensive introduction to and history of differential geometry [218], we find a detailed analysis of Riemann's paper (as well as Gauss's papers that we discussed earlier and later important works of the nineteenth century in differential geometry, including translations into English of the most important papers).

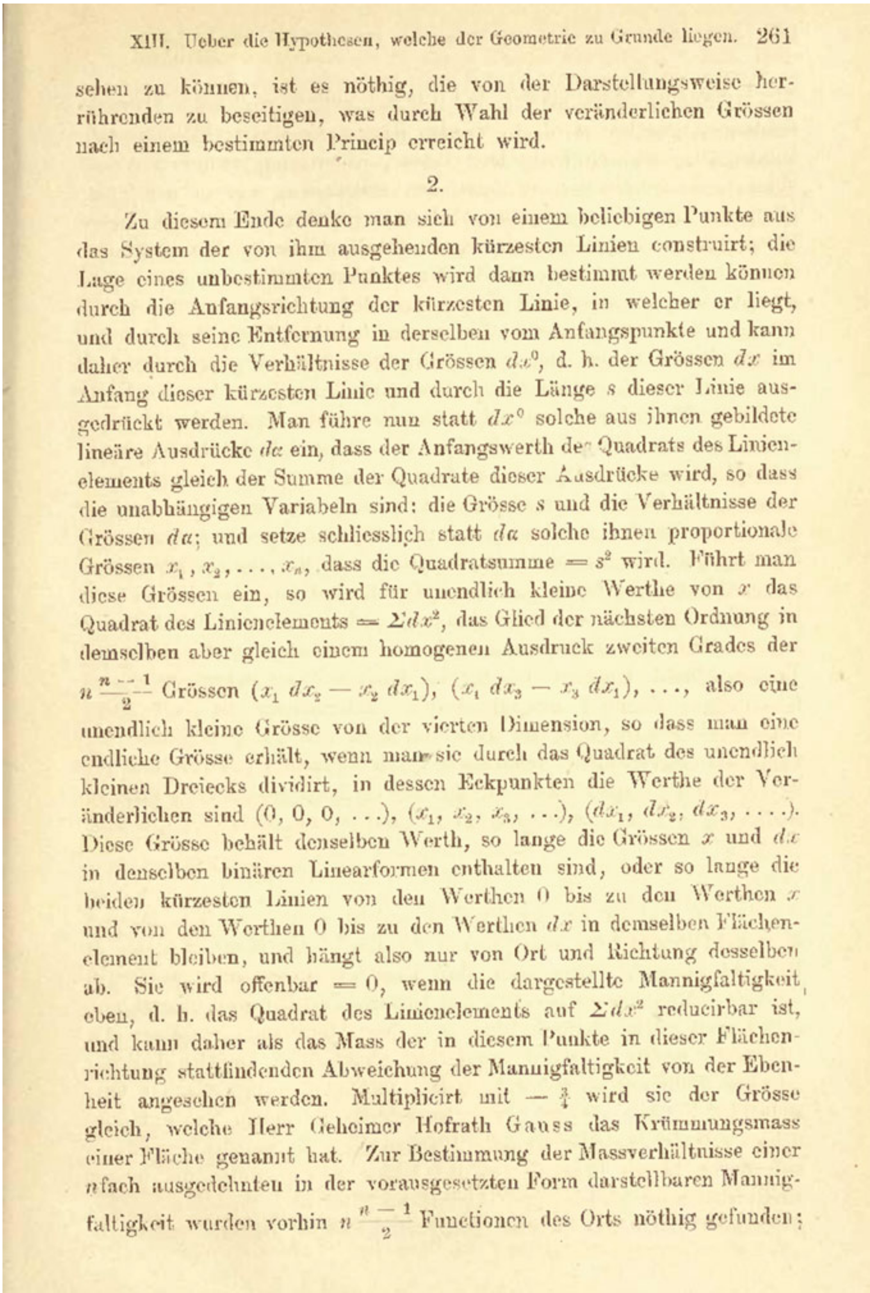
We want to summarize what Riemann says on p. 261 (again, see Fig. 5.3). He starts by introducing near a given point p on his manifold M *geodesic normal coordinates*, that is, coordinates which are geodesics emanating from the given point and whose tangent vectors at p are an orthonormal basis for T_p (this orthogonality and the geodesics use, of course, the given Riemannian metric). In this coordinate system $(x^1, \dots, x^n)$, the metric ds^2 has a Taylor expansion through second-order terms of the form

$$ds^2 = \sum_{i=1}^{n} dx^i dx^i + \frac{1}{2} \sum_{ijkl}^{n} \frac{\partial^2 g_{ij}}{\partial x^k \partial x^l}(0) x^k x^l dx^i dx^j. \tag{5.5}$$

The first-order terms in this expansion involve terms of the form $\frac{\partial g_{ij}}{\partial x^k}(0)$, all of which vanish, which follows from the geodesic coordinates condition. Letting now

$$c_{ijkl} := \frac{\partial^2 g_{ij}}{\partial x^k \partial x^l}(0),$$

[4] "I restrict myself therefore to manifolds where the line element is expressed by the square root of a differential expression of second degree."



XIII. Ueber die Hypothesen, welche der Geometrie zu Grunde liegen. 261

sehen zu können, ist es nöthig, die von der Darstellungsweise herrührenden zu beseitigen, was durch Wahl der veränderlichen Grössen nach einem bestimmten Princip erreicht wird.

2.

Zu diesem Ende denke man sich von einem beliebigen Punkte aus das System der von ihm ausgehenden kürzesten Linien construirt; die Lage eines unbestimmten Punktes wird dann bestimmt werden können durch die Anfangsrichtung der kürzesten Linie, in welcher er liegt, und durch seine Entfernung in derselben vom Anfangspunkte und kann daher durch die Verhältnisse der Grössen dx^0, d. h. der Grössen dx im Anfang dieser kürzesten Linie und durch die Länge s dieser Linie ausgedrückt werden. Man führe nun statt dx^0 solche aus ihnen gebildete lineäre Ausdrücke $d\alpha$ ein, dass der Anfangswerth des Quadrats des Linienelements gleich der Summe der Quadrate dieser Ausdrücke wird, so dass die unabhängigen Variabeln sind: die Grösse s und die Verhältnisse der Grössen $d\alpha$; und setze schliesslich statt $d\alpha$ solche ihnen proportionale Grössen $x_1, x_2, \ldots, x_n$, dass die Quadratsumme $= s^2$ wird. Führt man diese Grössen ein, so wird für unendlich kleine Werthe von x das Quadrat des Linienelements $= \Sigma dx^2$, das Glied der nächsten Ordnung in demselben aber gleich einem homogenen Ausdruck zweiten Grades der $n\frac{n-1}{2}$ Grössen $(x_1\,dx_2 - x_2\,dx_1)$, $(x_1\,dx_3 - x_3\,dx_1)$, $\ldots$, also eine unendlich kleine Grösse von der vierten Dimension, so dass man eine endliche Grösse erhält, wenn man sie durch das Quadrat des unendlich kleinen Dreiecks dividirt, in dessen Eckpunkten die Werthe der Veränderlichen sind $(0, 0, 0, \ldots)$, $(x_1, x_2, x_3, \ldots)$, $(dx_1, dx_2, dx_3, \ldots)$. Diese Grösse behält denselben Werth, so lange die Grössen x und dx in denselben binären Linearformen enthalten sind, oder so lange die beiden kürzesten Linien von den Werthen 0 bis zu den Werthen x und von den Werthen 0 bis zu den Werthen dx in demselben Flächenelement bleiben, und hängt also nur von Ort und Richtung desselben ab. Sie wird offenbar $= 0$, wenn die dargestellte Mannigfaltigkeit eben, d. h. das Quadrat des Linienelements auf Σdx^2 reducirbar ist, und kann daher als das Mass der in diesem Punkte in dieser Flächenrichtung stattfindenden Abweichung der Mannigfaltigkeit von der Ebenheit angesehen werden. Multiplicirt mit $-\frac{3}{4}$ wird sie der Grösse gleich, welche Herr Geheimer Hofrath Gauss das Krümmungsmass einer Fläche genannt hat. Zur Bestimmung der Massverhältnisse einer $n$fach ausgedehnten in der vorausgesetzten Form darstellbaren Mannigfaltigkeit wurden vorhin $n\frac{n-1}{2}$ Functionen des Orts nöthig gefunden;

Fig. 5.3 Page 261 of Riemann's foundational paper on differential geometry [201] we have the natural symmetry conditions

$$c_{ijkl} = c_{jikl} = c_{jilk},$$

due to the symmetry of the indices in g_{ij} and in the commutation of the second-order partial derivatives. Moreover, and this is *not* easy to verify, the coefficients *also* satisfy

$$\begin{aligned} c_{ijkl} &= c_{klij}, \\ c_{lijk} + c_{ljki} + c_{lkij} &= 0. \end{aligned} \tag{5.6}$$

This is proved in six pages of computation in Spivak's Vol. 2 (pp. 172–178 of [218]), and we quote from the top of p. 174: "We now proceed to the hardest part of the computation, a hairy computation indeed." These symmetry conditions use the fact that the coordinates are specifically linked to the metric (our geodesic coordinates). For instance, on p. 175 Spivak points out that

$$x^i = \sum_{j=1}^{n} g_{ij} x^j,$$

illustrating vividly the relation between the coordinates and the metric.

Let now

$$Q(x, dx) := \sum_{ijkl} c_{ijkl} x^k x^l dx^i dx^j \tag{5.7}$$

be the biquadratic form defined by the second-order terms in (5.5), then Riemann asserts on p. 61 of [201] that $Q(x, dx)$ can be expressed in terms of the $n(\frac{n-1}{2})$ expressions $\{(x^1 dx^2 - x^2 dx^1), (x^1 dx^3 - x^3 dx^1), \ldots\}$ (see Fig. 5.3), that is,

$$Q(x, dx) = \sum_{ijkl}^{n} C_{ijkl}(x^i dx^j - x^j dx^i)(x^k dx^l - d^l dx^k). \tag{5.8}$$

Spivak proves that the conditions (5.6) are necessary and sufficient for $Q(x, dx)$ to be expressed in the form (5.8), and he shows, moreover, that

$$C_{ijkl} = \frac{1}{3} c_{ijkl}.$$

Riemann simply asserts that this is the case, which is, of course, indeed true!

The expression $Q(x, dx)$ defined by (5.8) is Riemann's definition of *curvature* for the manifold at the point 0 defined by the metric (5.2) using geodesic normal coordinates. This has become known as *Riemannian curvature* ever since.

Let's look at the special case where the manifold has two dimensions. In this case we see that there is only one coefficient of the nonzero term $(x^1 dx^2 - x^2 dx^1)^2$ which has the form

$$Q(x, dx) = \frac{1}{3}[c_{2211} + c_{1122} - c_{2112} - c_{1221}](x^1 dx^2 - x^2 dx^1)^2.$$

Now using Gauss's notation for the Riemannian metric (5.1), that is, $g_{11} = E$, $g_{12} = g_{21} = F$, and $g_{22} = G$, we see that

$$\begin{aligned} c_{2211} &= \frac{1}{2} G_{xx}, \\ c_{1122} &= \frac{1}{2} E_{yy}, \\ c_{2112} &= \frac{1}{2} F_{xy}. \\ c_{1221} &= \frac{1}{2} F_{xy}, \end{aligned}$$

and thus we have

$$Q(x, dx) = \frac{1}{6}[G_{xx} + E_{yy} - 2F_{xy}].$$

Looking at Gauss's formula for Gaussian curvature at the point 0 (4.10), we see that, since the first derivatives of the metric vanish at the origin, the curvature at $x = 0$ is

$$k = -\frac{1}{2}(G_{xx} G_{yy} - 2F_{xy}), \tag{5.9}$$

and hence

$$Q(x, dx) = -\frac{k}{3}(x^1 dx^2 - x^2 dx^1)^2. \tag{5.10}$$

Thus the coefficient of the single term $(x^1 dx^2 - x^2 dx^1)^2$ in the biquadratic form $Q(x, dx)$ is, up to a constant, the Gaussian curvature. As Riemann asserts it (and we paraphrase here): 'divide the expression $Q(x, dx)$ by the square of the area of the (infinitesimal) triangle formed by the three points $(0, x, dx)$, and the result of the division is $-\frac{4}{3}k$'. The factor 4 appears since the square of the area of the infinitesimal parallelogram[5] is $(x^1 dx^2 - x^2 dx^1)^2$, and thus the square of the area of the infinitesimal triangle is $\frac{1}{4}(x^1 dx^2 - x^2 dx^1)^2$. This yields the relation between Riemann's coefficient in (5.10) and Gaussian curvature (one can see this coefficient of $-\frac{3}{4}$ near the bottom of p. 261 in Fig. 5.3). Namely, except for a constant factor, Riemann's curvature expressed in normal coordinates on a two-dimensional manifold coincides with Gaussian curvature.

Riemann then considers the biquadratic form $Q(x, dx)$ in an n-dimensional manifold M and its restriction to any two-dimensional submanifold N passing through the point p, obtaining a curvature (constant multiple of the Gaussian

[5] Riemann visualizes the parallelogram formed by the points $(0, x, dx, x + dx)$ in $\mathbf{R}^2$, and the area of such a rectangle is simply given by the cross product $\|x \times dx\| = \|x^1 dx^2 - x^2 dx^2\|$, and the area of the triangle formed $(0, x, dx)$ is $\frac{1}{2}\|x^1 dx^2 - x^2 dx^1\|$.

curvature as we saw above) for the submanifold at that point. This is the sectional curvature of Riemann, introduced on this same p. 261.

In the remainder of the paper he discusses questions of flat manifolds, manifolds of positive or negative constant curvature, and numerous other questions.

The coefficients $\{c_{ijkl}\}$ in (5.7) or C_{ijkl} in (5.8) are effectively the components of the Riemannian curvature tensor for this special type of coordinate system (geodesic normal coordinates). How does one define such a curvature tensor for n-dimensional manifolds with a Riemannian metric in a general coordinate system (in the spirit of Gauss's curvature formula (4.10))? Clearly this will involve the first derivatives of the Riemannian metric as well. In a paper written in Latin for a particular mathematical prize in Paris (Paper No. XXII in Fig. 5.2), Riemann provides the first glimpse of the general Riemann curvature tensor, and this is again translated and elaborated on by Spivak [218]. The purpose of this paper was to answer a question in the Paris competition dealing with the flow of heat in a homogeneous solid body.

Riemann's ideas in these two posthumously published papers were developed and expanded considerably in the following decades in the work of Christoffel, Levi-Cevita, Ricci, Beltrami and many others. This is all discussed very elegantly in Spivak's treatise [218], and we won't elaborate on this any further at this point. The main point of our discussion has been that Riemann created on these few pages the basic idea of an n-dimensional manifold not considered as a subset of Euclidean space *and* of the independent concept of a Riemannian metric and the Riemann curvature tensor. What is missing at this point in time is the notion of a topological space on the basis of which one could formulate the contemporary concepts of a differentiable manifold or a Riemannian manifold.

Part III
Origins of Complex Geometry

Introduction

One of the most beautiful and profound developments in the nineteenth century is *complex geometry*. We mean by this a constellation of discoveries that led to the modern theory of complex manifolds (and more generally complex spaces: complex manifolds with specified types of singularities) and modern algebraic geometry, both of which have played an important role in the twentieth century.

The primary aspects of the theory of complex manifolds are the geometric structure itself, its topological structure, coordinate systems, etc., and holomorphic functions and mappings and their properties. Algebraic geometry over the complex number field uses polynomial and rational functions of complex variables as the primary tools, but the underlying topological structures are similar to those that appear in complex manifold theory, and the nature of singularities in both the analytic and algebraic settings is also structurally very similar.

Algebraic geometry uses the geometric intuition which arises from looking at varieties over the complex and real case to deduce important results in arithmetic algebraic geometry where the complex number field is replaced by the field of rational numbers or various finite number fields. This has led to important results in the latter half of the twentieth century, most notably Wiles's proof of Fermat's Last Theorem.

Complex geometry includes such diverse topics as Hermitian differential geometry, which plays an important role in Chern classes of holomorphic vector bundles, for instance, Hermitian symmetric domains or more generally homogeneous spaces with complex structure, or real differentiable manifolds with some complex structure in the tangent bundle such as almost complex manifolds and CR (Cauchy–Riemann) manifolds, and many other examples. Of course, a domain[1] in the complex plane **C** was an initial example of a complex manifold, much studied in the nineteenth century, and that will be an important part of the story.

---

[1]We will use the generic word *domain* to mean a connected open set.

As we saw in Part I of this book, during the seventeenth and eighteenth centuries, mathematics experienced major developments in geometry and analysis, specifically the geometry of curves and surfaces in $\mathbf{R}^3$, following the pioneering work of Descartes and Fermat, and the flourishing of analysis after the creation of differential and integral calculus by Leibniz, Newton and others. In all of this work, geometry was restricted to real geometric objects in the Euclidean plane and three space. Complex numbers, on the other hand, were developed and referred to as *imaginary numbers*, as they were called for several centuries, and they arose as solutions of specific polynomial algebraic equations. In the eighteenth century, they became part of the standard tools of analysis, especially in the development of fundamental elementary functions, which is epitomized in Euler's famous formulation of the complex exponential function

$$e^z := e^x(\cos y + i \sin y),$$

for a complex variable $z = x + iy$. However, the study of the geometry of curves and surfaces in $\mathbf{R}^3$ did not include complex numbers in any substantive manner, whereas in the twentieth century, complex geometry has become one of the main themes of twentieth-century mathematical research.

The purpose of this part of the book is to highlight key ideas developed in the nineteenth century which became the basis for twentieth-century complex geometry. We shall do this by looking in some detail at some of the innovators and their initial publications on a selection of research topics that, in the end, contributed in various ways to what we now call complex geometry.

In Chap. 6, we discuss the work of the Norwegian surveyor Wessel, the French mathematician Argand and the German astronomer and mathematician Gauss, all of whom contributed to our understanding of the complex plane as the usual Euclidean plane with complex coordinates $z = x + iy$, including its polar coordinate representation and expressing the distance between points in terms of complex coordinates. Over the course of the century, this understanding became universally adopted, but at the beginning of the century, it was quite unknown.

In Chap. 7, we look in some detail at Abel's fundamental paper concerning what is now known as Abel's theorem, which generalizes the addition theorem for elliptic integrals due to Euler. This paper became a primary motivation for major work by Riemann, Weierstrass and many others in the second half of the nineteenth century, as we discuss later in this part.

In the next chapter (Chap. 8), we discuss two fundamental papers by Abel and Jacobi which created the theory of elliptic functions, the nineteenth-century generalization of trigonometric functions. These new functions were doubly periodic (in two independent directions in the complex plane). Elliptic functions utilized the geometry of the complex plane in a fundamental manner, for instance in the role of the period parallelogram, whose translates cover the complex plane. This theory was developed further in the work of Cauchy, Liouville and Weierstrass, among many others, and we trace this development in some detail, as it became quite standard material in texts at the end of the century.

A key development in the nineteenth century was the creation of a theory of complex-valued functions that were intrinsically defined on domains in the complex plane, and this is the theory of holomorphic and meromorphic functions. The major steps in this theory were taken by Cauchy, in his theory of the Cauchy integral theorem and its consequences, by Riemann, in his use of partial differential equations, in particular, the study of harmonic functions, and by Weierstrass with his powerful use of power series (pun intended!). The unification of all three points of view towards the end of the nineteenth century had created what was then called *function theory* and has in the meantime over the course of the twentieth century become known as *complex analysis*. It is now a standard part of the contemporary mathematical curriculum. In Chap. 9, we shall look at some of the initial papers by these innovators and see how the point of view for this important topic evolved over time.

Finally, in Chap. 10, we come to a pivotal development in complex geometry, namely Riemann's creation of Riemann surfaces. Riemann's paper of 1857, which we discuss in some detail in this section, takes some of the main ideas from Abel's paper on Abel's theorem concerning multivalued functions of one real variable and creates a theory of single-valued holomorphic functions on an abstractly defined surface with complex coordinates. These surfaces are looked at from the point of view of analysis, from algebraic geometry as the solution of algebraic equations of two complex variables and from the point of view of topology, including the important notion of connectivity of a surface, which led to later developments in algebraic topology.

The concluding chapter of this part outlines some topics which are today important for complex geometry and which were also developed during the latter part of the nineteenth century. These include the theory of transformation groups of Lie and Klein, the development of set theory by Cantor and the subsequent developments of topological spaces by Hausdorff and Kuratowski, and the fundamental work on the foundations of algebraic topology by Poincaré. We conclude this chapter by discussing briefly the creation of abstract topological, differentiable and complex manifolds in the definitive book by Hermann Weyl in 1913, who used all of the topics discussed above, and which became the cornerstone of what became complex geometry in the twentieth century.

Chapter 6
The Complex Plane

6.1 Introduction

The well-known quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

as a solution to the quadratic equation

$$ax^2 + bx + c = 0,$$

is attributed to the Babylonians during their very creative period of mathematical discovery (circa 1800 BCE to 300 BCE) (see [194, 223] for discussions of the splendid mathematical accomplishments of the Babylonians, mostly preceding and greatly influencing the Greek mathematicians and astronomers). Of course they used different notation, but their understanding was clear. This formula led to the problem of understanding what one means by the square root in the cases where $b^2 - 4ac$ happens to be negative. This problem has been a part of mathematical culture ever since. By the eighteenth century numbers involving $\sqrt{-1}$ were used by numerous mathematicians in the solutions of a variety of problems, and Euler introduced the well-known notation of i to represent[1] $\sqrt{-1}$ and gave us his famous formula involving our basic mathematical constants

$$e^{\pi i} = -1.$$

[1]However, we note that in the work of a number of several nineteenth-century mathematicians, the notation $\sqrt{-1}$ was used for emphasis, for instance in the well-known dissertation of Riemann from 1851 [199], which we will discuss below.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_6

These numbers became known as *imaginary numbers*, indicating clearly that they were figments of the imagination in some sense, but weren't real mathematical objects. The mathematicians of the eighteenth century, many of whom were very interested in questions of geometry, including Euler, missed the opportunity to come up with a geometric interpretation of these imaginary numbers. This opportunity was not missed at the beginning of the nineteenth century.

There were three independent contributions to the creation of the *complex plane* at the beginning of the nineteenth century, namely by Caspar Wessel (1745–1818) in 1797 [240], Jean-Robert Argand (1768–1822) in 1806 [9], and Carl Friedrich Gauss in 1831 [78]. We can cite this creation of the geometric complex plane as having been the birth of complex geometry, and it took some time for this new perspective to become an ordinary part of mathematical discourse.

Wessel and Argand both wrote definitive papers on the geometric representation of complex numbers in the Cartesian plane $\mathbf{R}^2$, and neither paper was recognized at the time of publication for the great breakthrough they both represented. In the extreme case, Wessel's paper was not recognized until a century later when it was translated into French (from the original Danish). Today this paper is available in a beautiful book [240] (translated into English), along with a personal and mathematical biography of Wessel. [2]

6.2 Caspar Wessel's Cartography

Wessel was a geographical and trigonometrical surveyor who surveyed large parts of Denmark and one section of Germany (Duchy of Oldenburg, northwest of Bremen, at the time under control of the Danish crown). In fact, Gauss, in his survey of the land southeast of Oldenburg (Bremen to Göttingen), used some of Wessel's survey data to lend accuracy to his own measurements. Wessel came upon his idea of representing complex numbers[3] in a geometric manner as a tool for simplifying trigonometrical calculations, which were so prevalent in his surveying work. He described a complex number as a length and a direction from a given point and a given axis passing through that point, just as we do today. More importantly, he described how to add and multiply numbers using this language. In Fig. 6.1 we see Wessel using a polar coordinate system involving complex numbers as coordinates, and in Fig. 6.2 we

[2] In addition, the book contains an excellent detailed article by Kirsti Andersen entitled *Wessel's Work on Complex Numbers and its Place in History*, which concerns the history of the use of the plane to represent complex numbers from Wessel to Hamilton, including the contributions of numerous other mathematicians including Argand and Gauss.

[3] The term *complex numbers* was introduced by Gauss in 1831 [78], although the term *imaginary numbers* was used till the latter half of the nineteenth century by many mathematicians, including, in particular, Cauchy.

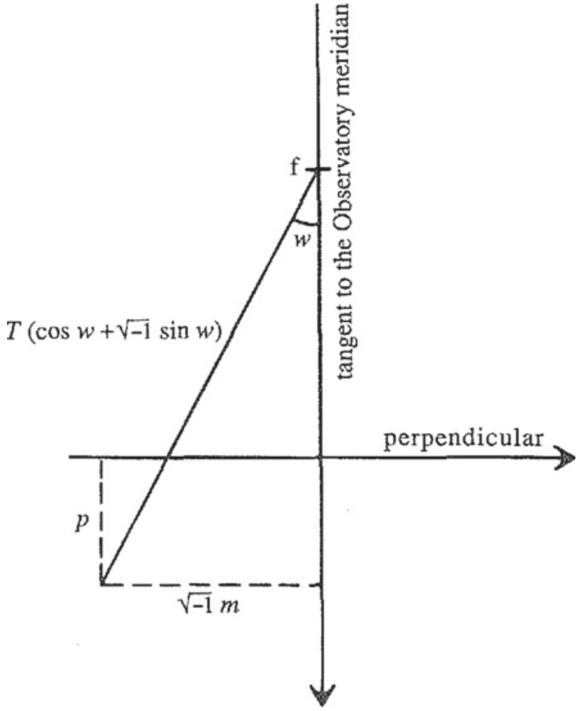


Fig. 6.1 A figure taken from a manuscript of Wessel called *Trigonometric Calculations* from 1779 (on p. 46 of [240]) illustrating the use of complex coordinates. *This illustration is reprinted with the permission of The Royal Danish Academy of Sciences and Letters*

show a page of the English translation from his paper of 1797, where he describes addition and multiplication of complex numbers. Note that at the bottom of the page in Fig. 6.2 he identifies his geometric quantity ε, a unit vector perpendicular to the real axis, as $\sqrt{-1}$. He uses this representation to give a complete description of the n roots of unity of degree n in the form:

$$\{1, \cos(2\pi/n) + \varepsilon \sin(2\pi/n), \cos(4\pi/n) + \varepsilon \sin(4\pi/n), \cdots\},$$

where ε is his notation for $\sqrt{-1}$. Finally, his main task in the remainder of this paper is to tackle problems of spherical geometry in three dimensions. We note that his product of two directed line segments (see Fig. 6.2) from a common point *lies on a plane spanned by the two segments*, indicating that he has been conceptualizing his ideas in three dimensions from the beginning. We conclude this discussion of Wessel by including in Fig. 6.3 a beautiful map from his earlier work, showing his skill as a cartographer.

106 C. Wessel MfM 46:1

thogonal planes, the segment has the same effect on the distance of the point from each of the three planes; consequently, one of several added lines contributes the same to the position of the last point of the sum, whether it is the first, the last, or has any other number among the addends; thus the order in the addition of straight lines is immaterial, and the sum always remains the same, because the initial point is assumed to be given, and the last point always attains the same position.

Hence, in this case one may also denote the sum by inserting the sign + between the lines to be added. For instance, when in a quadrilateral the first side is drawn from a to b, the second from b to c, the third from c to d, and the fourth from a to d: then one can write $ad = ab + bc + cd$.

§ 3.

If the sum of several lengths, widths, and heights = 0, then the sum of the lengths, that of the widths, and that of the heights, each sum separately = 0.

§ 4.

The product of two straight lines should in every respect be formed from the one factor, in the same way as the other factor is formed from the positive or absolute unit line that is set = 1, that is:

First, the factors must have such directions, that they can both be included in the same plane as the positive unit.

Next, concerning the length of the product, it must be to the one factor as the other is to the unit; and

Finally, if the positive unit, the factors, and the product are given a common initial point, then the product, with respect to the direction must lie in the plane of the unit and the factors, and the product must deviate as many degrees from the one factor, and to the same side, as the other factor deviates from the unit, so that the directional angle of the product or its deviation from the positive unit is the sum of the directional angles of the factors.

§ 5.

Let +1 denote the positive, rectilinear unit, and $+\varepsilon$ a certain different unit, perpendicular to the positive unit, and with the same initial point; then the directional angle of +1 is 0, of −1 it is 180°, of $+\varepsilon$ it is 90°, and of $-\varepsilon$ it is −90° or 270°; and according to the rule that the directional angle of the product is the sum of those of the two factors, one gets

$$(+1)\cdot(+1) = +1,\ (+1)\cdot(-1) = -1,\ (-1)\cdot(-1) = +1,\ (+1)\cdot(+\varepsilon) = +\varepsilon,\ (+1)\cdot(-\varepsilon) = -\varepsilon,$$

$$(-1)\cdot(+\varepsilon) = -\varepsilon,\ (-1)\cdot(-\varepsilon) = +\varepsilon,\ (+\varepsilon)\cdot(+\varepsilon) = -1,\ (+\varepsilon)\cdot(-\varepsilon) = +1,\ (-\varepsilon)\cdot(-\varepsilon) = -1.$$

From this it follows that ε becomes $= \sqrt{-1}$, and the deviation of the product is determined so that not a single one of the usual rules of operation is violated.

Fig. 6.2 Wessel's notion of sum and product of complex numbers from his paper of 1797 (English translation [240], p. 106). *This illustration is reprinted with the permission of The Royal Danish Academy of Sciences and Letters*

Fig. 6.3 Wessel's map of Denmark from 1768 ([240], plate following p. 21). *This illustration is reprinted with the permission of The Royal Danish Academy of Sciences and Letters*

6.3 Argand and Gauss

We turn now to Argand, who published a small pamphlet [9] in a limited print edition in 1806 entitled *Essai sur un manière de représenter les quantités imaginaires dans les constructions géométrique*.[4] This was later reprinted in an influential mathematical journal edited by Joseph Diaz Gergonne (Annales de Mathématiques pures et appliquées) in 1813, which included papers by Jacques-Frédéric Français (1775–1833), François Joseph Servois (1768–1847), responses by Argand, and some commentary by Gergonne concerning the new ideas in Argand's work.[5] Argand also gave in this paper the first definitive proof of the fundamental theorem of algebra using his geometric representation. Indeed, he formulated the theorem in the form that for any polynomial equation of degree n with complex numbers as coefficients,

[4] *Essay on a manner to represent imaginary quantities in a geometric construction*.

[5] See the article by Andersen [7] for a detailed analysis of this interesting mathematical discussion in Gergonne's journal.

$$p(z) = a_0 + a_1 z + \ldots + a_n z^n = 0, \; a_j \in \mathbf{C},$$

there exists at least one complex number $z_0 \in \mathbf{C}$ such that $p(z_0) = 0$. This proof is not constructive and is a proof by contradiction, like many other proofs given later by others, and it utilizes substantively the notion of the *modulus* of a complex number,

$$|x + iy| := \sqrt{x^2 + y^2},$$

which was first introduced by Argand in his paper and is the length of the directed line segment used by Wessel.

Gauss had thought about the issue of the geometric representation of complex numbers for some decades at the beginning of the nineteenth century, but didn't publish anything on the subject until his "Theoria residuorum biquadraticorum, Commentatio secunda"[6] [78] in 1831, in which he specifically defined a complex number z of the form $x + iy$ to correspond to the point (x, y) in the Euclidean two-plane $\mathbf{R}^2$, and the usual arithmetic (addition and multiplication) of complex numbers

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$
$$(a + ib)(c + id) = (ac - bd) + i(ad + bc)$$

corresponded to new specific points in the plane. In this paper, he did not consider a polar coordinate representation of complex numbers, so that multiplication corresponded to multiplying moduli and adding angles of complex numbers as did Wessel and Argand, although he surely was aware of this by this time. He was more concerned with emphasizing that this relation of arithmetic and geometry was a valid way of doing mathematics, and that had such numbers not been called "imaginary" centuries earlier, they would have been accepted much earlier. His main purpose in this short note is to indicate that a number of his number-theoretic results from his well-known treatise on number theory from 1801, *Disquistiones Arithmeticae* [77], could be extended to the setting of complex numbers; and specifically he discusses complex numbers of the form $a + ib$, where a and b are integers (Gaussian integers), but emphasizing that such numbers were points in a two-dimensional plane. The only earlier reference by Gauss to a geometric representation of complex numbers was in a detailed letter from Gauss to Bessel in 1811.

Here is what Gauss had to say in that letter [82]:

> Was soll man sich nun bei $\int \varphi x \cdot dx$ für $x = a + bi$ denken? Offenbar, wenn man von klaren Begriffen ausgehen will, muss man annehmen, dass x durch unendlich kleine Incremente (jedes von der Form $\alpha + \beta i$) von demjenigen Werthe, für welche das Integral 0 sein soll, bis zu $x = a + bi$ übergeht und dann alle $\varphi x \cdot dx$ summiert. So ist der Sinn vollkommen festgesetzt. Nun aber kann der Übergang auf unendlich viele Arten geschehen: so wie man sich das ganze Reich aller reellen Grössen durch eine unendliche gerade Linie denken kann,

[6] "Second Commentary on Quadratic Residues"; note: the title is in Latin, referring to his earlier work on quadratic residues, and the paper itself is written in German.

> so kann man das ganze Reich aller Grössen, reeller und imaginärer Grössen sich durch eine unendliche Ebene sinnlich machen, worin jeder Punkt, durch Abcisse $= a$, Ordinate $= b$ bestimmt, die Grösse $a + bi$ gleichsam repräsentiert. Der stetige Übergang von einem Werthe von x zu einem andern $a + bi$ geschieht demnach durch eine Linie und ist mithin auf unendlich viele Arten möglich. Ich behaupte nun, dass das Integral $\int \varphi x \cdot dx$ nach zweien verschiednen Übergängen immer einerlei Werth erhalte, wenn innerhalb des zwischen beiden die Übergänge repräsentirenden Linien eingeschlossenen Flächenraumes nirgends $\varphi x = \infty$ wird. Dies ist ein sehr schöner Lehrsatz, dessen eben nicht schweren Beweis ich bei einer schicklichen Gelegenheit geben werde. [7]

In this quote we also see Gauss's quite specific understanding of what became known as the *Cauchy integral theorem*, which we will discuss later in Chap. 9.

[7] "What should one understand by $\int \varphi x \cdot dx$ for $x = a + bi$? Obviously, if we want to start from clear concepts, we have to assume that x passes from the value for which the integral has to be 0 to $x = a + bi$ through infinitely small increments (each of the form $x = a + bi$), and then to sum all the $\varphi x \cdot dx$. Thereby the meaning is completely determined. However, the passage can take place in infinitely many ways: Just like the realm of all real magnitudes can be conceived as an infinite straight line, so can the realm of all magnitudes, real and imaginary, be made meaningful by an infinite plane, in which every point, determined by abscissa $= a$ and ordinate $= b$, represents the quantity $a + bi$. The continuous passage from one value of x to another $a + bi$ then happens along a curve and is therefore possible in infinitely many ways. I claim now that after two different passages the integral $\int \varphi x \cdot dx$ acquires the same value when φx never becomes equal to ∞ in the region enclosed by the two curves representing the two passages. This is a very beautiful theorem whose not exactly difficult proof I shall give at a suitable occasion." (This reference is from Andersen [7].)

Chapter 7
Elliptic and Abelian Integrals

7.1 Introduction

In the eighteenth century, trigonometric functions (often called circular functions), and the related logarithmic and exponential functions, became fundamental tools of analysis. The trigonometric functions first appeared in the work of Hipparchus of Nicaea (c. 190 BCE – c. 120 BCE) in the context of spherical trigonometry for use in astronomy, and later plane trigonometry was developed and used for practical engineering and building problems. In Euler's well-known text on analysis from 1748 [62] we see these functions used in the form we are familiar with today. These functions and others like them were called *transcendental functions* in that they were a more general class of functions than the *rational functions*, which were ratios of polynomial functions or *algebraic functions*, which are solutions of algebraic equations, such as $y = \sqrt{x}$. It is important to note that almost all of the important transcendental functions of the eighteenth century, including many of the newer transcendental functions of the nineteenth century (e.g., Bessel functions, Riemann zeta function, etc.) were accompanied by *numerical tables* of their values, so that they could be used in applied computational settings. Only with the advent of computers in the mid-twentieth century did the use of such tables become obsolete.

Calculus became an important tool involving calculating with symbols which could often reduce a complicated problem to a simpler one before tables of values or approximation tools (such as power series) had to be used. As was known from the beginning of the use of calculus, it was most often much simpler to differentiate a given function than to find its integral, i.e., a formula for its antiderivative. Definite integrals of specific functions which didn't seem to have an antiderivative were studied extensively in the first half of the nineteenth century by the means of integration in the complex plane using Cauchy residue theory, as we will see in Chap. 9. But toward the end of the eighteenth century and the first half of the nineteenth century a great deal of effort went into understanding specific classes of indefinite integrals. In fact, the notation most often used, $\int f(x)dx$, meant $\int f(x)dx$ was a function whose

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_7

derivative was $f(x)$, and often a constant of integration was implied or explicitly mentioned.

In the eighteenth century it was well known that the trigonometric functions and logarithm and exponential functions can be defined as integrals of specific rational or algebraic functions or inverses of such functions. For instance,

$$\log(x) = \int \frac{dx}{x}, \ \arcsin(x) = \int \frac{dx}{\sqrt{1-x^2}}, \ \arctan(x) = \int \frac{dx}{1+x^2}, \tag{7.1}$$

i.e., the derivatives of these transcendental functions are these specific rational and algebraic functions. A function such as $\sqrt{1-x^2}$ was often referred to in the literature of the time as an *irrational* function, i.e., an algebraic function (involving possible roots of a rational function) which was not rational.[1]

The question of understanding integrals of various classes of functions became an important topic in the eighteenth and first half of the nineteenth century, and this led to very important work in complex geometry, as we shall see.

First, since the creation of calculus and the fundamental theorem of calculus, it was well known how to integrate a polynomial, i.e.,

$$\int x^n dx = \frac{1}{n+1} x^{n+1}.$$

Moving up one step in complication, let

$$r(x) = \frac{p(x)}{q(x)}$$

be a rational function of one real variable x, where p and q are polynomials. Then, by basic algebra, namely, using the fact that any polynomial with real coefficients could be factored into linear and irreducible quadratic terms,[2] and the method of partial fractions, one was able to write:

$$\int r(x)dx = \int p(x)dx + \int \sum \frac{a_j dx}{x-b_j} + \int \sum \frac{(e_k x + d_k)dx}{x^2 + e_k x + f_k}, \tag{7.2}$$

where $p(x)$ is a polynomial. Hence each integral of the form $\int r(x)dx$ can be reduced to a rational function and integrals of the form:

$$\int \frac{dx}{x} = \log(x), \int \frac{dx}{1+x^2} = \arctan(x),$$

[1] The notion of irrational function as used at the time didn't seem to refer to transcendental functions, which, of course, are also not rational functions.

[2] This was well known and used regularly throughout the eighteenth century, but the *proofs* of the fundamental theorem of algebra didn't appear until the nineteenth century.

two transcendental functions. This general principle was formulated by Johann Bernoulli (1667–1748), who published a short paper on this topic in 1703 [15] in which he outlined the process described above as a general algorithm for integrals of rational functions.[3]

7.2 Euler's Addition Theorem

If we now consider a rational function $r(x, y)$ of two real variables (again a ratio of two polynomials $p(x, y), q(x, y)$ of the two variables x and y), and let x and y be related by the quadratic equation

$$y^2 = a + bx + cx^2,$$

and hence,

$$y(x) = \pm\sqrt{a + bx + cx^2},$$

then the question arose in the eighteenth century: can one reduce an integral of the form

$$\int r(x, y(x))dx \tag{7.3}$$

to a sum of rational and elementary transcendental functions (i.e., trigonometric and logarithmic functions)?

Special cases of this were known for some time, as in (7.1) for $\int \frac{dx}{\sqrt{1-x^2}}$, for instance, where $r(x, y) = 1/y$ and $y^2 = 1 - x^2$. These kinds of problems arose in a variety of problems in elasticity, astronomy, and other sciences, and provided an important motivation for finding general solutions (see Kline [125], Chap. 19, for an overview of this intertwined scientific and mathematical development in the eighteenth century).

In 1768 Euler published an important book on integral calculus [68],[4] which solved this particular problem and also set the stage for the work of Abel and Jacobi some 50 years later. Euler proved, by making a judicious change of variables of the form $x = x(t)$, where $x(t)$ was an explicit rational function of t, that the integral (7.3) became

$$\int r(x, y(x))dx = \int g(t)dt, \tag{7.4}$$

[3]In fact, in this paper Bernoulli assumed simple complex roots of a polynomial reducing (7.2) to simply logarithmic terms.

[4]This was the first of three volumes; Vol. 2 was published in 1769 and Vol. 3 was published in 1770.

where $g(t)$ was a rational function, and hence the problem was reduced to the older one. Thus, such an integral reduced to a sum of a rational function and elementary functions, as before.

This change of variables, due to Euler, later became known as the rational parametrization of an algebraic curve of degree two, which we want to illustrate here due to its simplicity. Suppose we have an algebraic curve in $\mathbf{R}^2$ of degree two of the form:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

First, we use a translation in the plane to make the constant term vanish, and we then have (in the new coordinates)

$$ax^2 + bxy + cy^2 + dx + ey = 0. \tag{7.5}$$

Thus, the origin (0, 0) is a point on the curve, and we can consider the one-parameter family of straight lines of the form

$$y = tx, \text{ for } t \in \mathbf{R},$$

which will intersect the curve at both the origin and one other point on the curve for a fixed t. Substituting $y = tx$ into (7.5), we obtain

$$ax^2 + btx^2 + ct^2x^2 + dx + etx = 0. \tag{7.6}$$

Solving for x in terms of t, we find the parametrization of the curve in terms of t to be:

$$x(t) = \frac{-(d + et)}{a + bt + ct^2}, \tag{7.7}$$

$$y(t) = t\left(\frac{-(d + et)}{a + bt + ct^2}\right). \tag{7.8}$$

From (7.7), we see that $dx(t) = s(t)dt$, where $s(t)$ is a rational function of t. It follows then that

$$\int r(x, y)dx = \int r(x(t), y(t))s(t)dt, \tag{7.9}$$

when x and y are related by (7.5). This verifies that such an integral is computable in terms of rational functions and elementary functions, Euler's result from 1768.

An algebraic curve which has a parametrization in terms of rational functions of the form (7.7) and (7.8) is called a *rational curve*, and there are many examples of polynomials $f(x, y)$ of degree higher than two which also define rational curves.[5]

[5]For instance, there is the *folium of Descartes* given by $x^3 + y^3 - 3axy = 0$, which is parametrized by the rational functions $x = \frac{3at}{1+t^3}$, $y = \frac{3at^2}{1+t^3}$.

In the same book from 1768 [69] Euler discussed the more difficult problem of the form

$$\int \frac{dx}{\sqrt{A + Bx + Cx^2 + Dx^3 + Ex^4}}, \tag{7.10}$$

or, more generally,

$$\int r(x, y)dx, \tag{7.11}$$

where

$$y^2 = A + Bx + Cx^2 + Dx^3 + Ex^4,$$

and $r(x, y)$ is a rational function.

Functions of the type (7.11) have been known since the eighteenth century as *elliptic integrals*, as they originally arose in the context of computing via integration the lengths of arcs of an ellipse, just as the classical trigonometric functions arose in conjunction with measuring the lengths of circular arcs. Note that elliptic integrals are functions of the variable x, indeed, they are transcendental functions, just as the elementary functions are, even though they are referred to as integrals.

In his original paper in [66] and in the text [69], Euler discovered algebraic relations between elliptic integrals having the same form. For instance, the differential equation

$$\frac{dx}{\sqrt{A + Bx + Cx^2 + Dx^3 + Ex^4}} = \frac{dy}{\sqrt{A + By + Cy^2 + Dy^3 + Ey^4}} \tag{7.12}$$

has a solution as a complete algebraic integral (an algebraic one-parameter family of algebraic curves in $\mathbf{R}^2$ which satisfies the differential equation).

Let's give an outline of Euler's solution to this problem. First Euler makes a change of variables of the form

$$x = \frac{\alpha t + \beta}{\gamma t + \delta},$$

to get rid of the linear and cubic terms, reducing the problem to

$$\frac{dx}{\sqrt{A + Cx^2 + Ey^4}} = \frac{dy}{\sqrt{A + Cy^2 + Ey^4}}. \tag{7.13}$$

Then, by several more quite nontrivial (and nonlinear) changes of variables and integrating, he is able to produce the integral of this equation as a very specific polynomial function of degree 4 with coefficients which depend on A, C, and E and an arbitrary constant f. His solution has the form

$$A(x^2 + y^2) = f^2(A + Ex^2y^2) + 2xy\sqrt{A(A + Cf^2 + Ef^4)}, \tag{7.14}$$

where f is a constant of integration. See §15 of [66], and he has a number of variations of this solution in this paper; we shall see a special case of this below. This type of relationship between these variables (solution of the differential equation (7.13)) became known as an *Euler addition theorem for elliptic integrals*.

Let's illustrate this concept in the simpler case of

$$\frac{dx}{\sqrt{A+Cx^2}} = \frac{dy}{\sqrt{A+Cy^2}},$$

which Euler had discussed earlier in his text [70]. He obtained a solution of the form

$$y = x\sqrt{\frac{A+Cb^2}{A}} + b\sqrt{\frac{A+Cx^2}{A}}, \tag{7.15}$$

having solved for y in terms of the other variables (here b is the constant of integration) from his solution. Let's assume the special case of $A = 1, C = -1$, and then we have the function

$$y = x\sqrt{1-b^2} + b\sqrt{1-x^2} \tag{7.16}$$

is the solution of

$$\frac{dx}{\sqrt{1-x^2}} = \frac{dy}{\sqrt{1-y^2}}. \tag{7.17}$$

If we integrate both sides we find that

$$\int_0^y \frac{dt}{\sqrt{1-t^2}} = \int_0^x \frac{dt}{\sqrt{1-t^2}} + \text{constant}. \tag{7.18}$$

But from (7.16) we see that for $x = 0$, we must have $y = b$, and thus the constant in (7.18) has the form

$$\text{constant} = \int_0^b \frac{dt}{\sqrt{1-t^2}},$$

and hence

$$\int_0^y \frac{dt}{\sqrt{1-t^2}} = \int_0^x \frac{dx}{\sqrt{1-t^2}} + \int_0^b \frac{dt}{\sqrt{1-t^2}},$$

where x, y, and b are related by (7.16). By relabeling the variables, as did Euler, we find the familiar formula

$$\int_0^x \frac{dt}{\sqrt{1-t^2}} + \int_0^y \frac{dt}{\sqrt{1-t^2}} = \int_0^b \frac{dt}{\sqrt{1-t^2}},$$

where

$$b = x\sqrt{1-y^2} + y\sqrt{1-x^2}. \tag{7.19}$$

This is the classical addition formula for the inverse sine function:

$$\arcsin(x) + \arcsin(y) = \arcsin(b),$$

where (x, y, b) satisfies (7.19). This is, in turn, the same as the classical addition formula for the sine function

$$\sin(x + y) = \cos(x)\sin(y) + \sin(x)\cos(y). \tag{7.20}$$

Thus Euler's solution of the Eq. (7.17) yields the classical addition formula (7.20) for circular functions, which was known to the ancient trigonometers. The corresponding half-angle formulas allowed the Greek astronomers to compute the trigonometric tables which were so critical for their astronomical research. Euler's generalization of (7.19) for the elliptic integrals preceded any knowledge of elliptic functions (inverse functions to the elliptic integrals), which were discovered considerably later. As we will see in Chap. 8, where we discuss the discovery of elliptic functions, Euler's addition theorems for elliptic integrals did then provide addition formulas for elliptic functions similar to (7.20).

7.3 Abel's Addition Theorem

Niels Henrik Abel (1802–1829) in his very short lifetime[6] wrote a number of quite important papers, several of which came to play a significant role in the development of complex geometry. We will discuss two of these papers in some detail. The first paper from 1826 [2] concerns what is now known as Abel's theorem in algebraic geometry, and we will explore this paper in this section. Abel's second paper [3] is his foundational paper on elliptic functions, which we will discuss in the next chapter (Chap. 8).

Both of these papers were influenced by the work of Euler that was described in the previous section, as well as the follow-up to Euler's work by Adrien-Marie Legendre (1752–1833) in his several-decades-long study of elliptic integrals and their applications.

Legendre's principal contributions were contained in three monographs he published in the decade before Abel's work. These three volumes were entitled *Exercices de Calcul Intégral*. Volume 1 [138] in 1811 was his major theoretical work on elliptic integrals, which showed how all elliptic integrals of a general kind could be reduced, via algebra and calculus, to three specific types of integrals, which Legendre referred to as integrals of the first, second and third kind, which we shall see shortly. Volume 2 [140] from 1817 contained a major survey of approximation methods, methods of creating tables and numerous applications to geometry and applied mathematics, in particular to mechanics. Volume 3 [139] (which was actually published in 1816

[6] He was not yet 27 years old when he died.

before Volume 2) contains detailed tables for elliptic functions of the first and second kind and their logarithms, as well as a discussion of the issues of reducing computations of some integrals of the third kind to those of the first and second kind (there were too many free parameters in these transcendental functions of the third kind to allow the creation of reasonable tables). After the groundbreaking work of Abel and Jacobi in 1826 and 1827 Legendre continued his surveys of the development of what has now become the theory of elliptic functions, which we describe in the next chapter.

We will now look at Abel's 1826 paper, "Mémoire sur une propriété générale d'une classe trés étendue de fonctions transcendantes"[7] [2]. This paper was presented to the French Academy of Science in 1826 and was finally published posthumously in 1841. It gives a vast generalization of Euler's addition formula for elliptic integrals, which was discussed in the previous section, and is now called *Abel's theorem* in algebraic geometry.[8]

Let us preface our formulations of Abel's theorem[9] with a specific version of Euler's addition formula for elliptic integrals. Namely, in 1761 [64] Euler studied the differential equation

$$\frac{dx}{\sqrt{1-x^4}} = \frac{dy}{\sqrt{1-y^4}}, \tag{7.21}$$

a special case of (7.12) discussed briefly above, and he finds the complete algebraic integral to be

$$x^2 + y^2 + c^2x^2y^2 = c^2 + 2xy\sqrt{1-c^2}, \tag{7.22}$$

where c is the constant of integration.

Now consider the specific elliptic integral

$$E(x) := \int_0^x \frac{dx}{\sqrt{1-x^4}}, \tag{7.23}$$

where we choose the lower limit of integration to be $x = 0$. Then one finds by integrating each side of (7.21) that

$$E(x) = E(y) + C,$$

[7] "Memoir concerning a general property of a very extended class of transcendental functions".

[8] There are a number of theorems known as Abel's Theorem in different parts of mathematics, e.g., on the convergence of power series, on the unsolvability of quintic polynomial equations, etc.

[9] There is more than one algebraic-geometric theorem referred to historically over the past century as *Abel's theorem*. The very informative paper by Stephen Kleiman entitled "What is Abel's Theorem anyway?" [119] discusses four variants of what have been called Abel's theorem. This paper is an article in a beautiful book [10] representing the proceedings of a conference held in honor of the mathematical legacy of Abel in 2002, 200 years after his birth in 1802.

where C is a constant. From the complete integral of (7.21) given by (7.22) we see that when $x = 0$, then $y = c$ (we take the positive square root in this case, for convenience), and hence

$$0 = E(0) = E(c) + C,$$

and hence $C = -E(c)$, yielding

$$E(x) = E(y) - E(c),$$

or

$$E(x) + E(c) = E(y).$$

Changing the names of the variables $x_3 = y$, $x_1 = x$, $x_2 = c$, we obtain the addition theorem for this particular elliptic integral of the form

$$E(x_1) + E(x_2) = E(x_3), \tag{7.24}$$

where

$$x_1{}^2 + x_3{}^2 + x_1{}^2 x_2{}^2 x_3{}^2 = x_2{}^2 + 2x_1 x_3\sqrt{1 - x_2{}^2},$$

which gives after squaring

$$4x_1{}^2 x_3{}^2(1 - x_2{}^2) = (x_1{}^2 + x_3{}^2 + x_1{}^2 x_2{}^2 x_3{}^2 - x_2{}^2)^2,$$

a polynomial relation of degree 12 among the arguments of the three transcendental functions $E(x_1)$, $E(x_2)$, $E(x_3)$ (see [119], p. 20 for various references to this formula).

Note that for the arcsine addition formula (7.2), which we can write as

$$\int_0^{x_1} \frac{dx}{\sqrt{1 - x^2}} + \int_0^{x_1} \frac{dx}{\sqrt{1 - x^2}} = \int_0^{x_3} \frac{dx}{\sqrt{1 - x^2}},$$

we have the same sort of algebraic relation which takes the (familiar) form

$$x_3 = x_1\sqrt{1 - x_2{}^2} + x_2\sqrt{1 - x_1{}^2},$$

which we discussed earlier, and which when squared twice yields a polynomial relation among the three arguments of these three transcendental functions of degree six.

Let now $r(x, y)$ be a rational function, let $f(x, y)$ be a polynomial, and let $y(x)$ be the implicit (multivalued) function defined by $f(x, y) = 0$. Then the general *Abelian integral* is defined to be

$$A(x) := \int_{x_0}^{x} r(x, y(x))dx, \tag{7.25}$$

for some lower limit of integration x_0. What Abel originally meant by this was an antiderivative (as did Euler), i.e. $A(x)$ is a function whose derivative is $r(x, y(x))$, and we are expressing this as a definite integral from an initial point x_0 to an upper limit x, using the same symbol x as the variable of integration.

Let us now formulate a first version of Abel's theorem.

Theorem 7.1 *Let $A(x)$ be an Abelian integral as defined in (7.25). If $g(x, y)$ is an auxiliary polynomial, and if the curve $g(x, y) = 0$ intersects the curve $f(x, y) = 0$ in the points $(x_1, y_1), \cdots, (x_N, y_N)$, then there are rational functions $u, v_1, \cdots, v_r$ of the variables $x_1, \cdots, x_N$ and the coefficients of the polynomial $g(x, y)$ such that*

$$A(x_1) + A(x_2) + \ldots + A(x_N) = u + k_1 \log v_1 + \ldots + k_r \log v_r, \tag{7.26}$$

where $k_1, \cdots, k_r$ are constants.

This says that the left-hand side of (7.26), a sum of N transcendental functions, is an *elementary function*, i.e., in this case a sum of a rational function and logarithmic terms. Thus (7.26) says that such a sum of Abelian integrals is an elementary function. Note that this is a generalization of the much simpler case where the integral

$$\int_{x_0}^{x} r(x, y(x))dx$$

is the sum of elementary functions, when $r(x, y) = 1/y$ and $y^2 = Ax^2 + Bx + C$, i.e., in the trigonometric case (Euler's theorem, see (7.9)).[10] This version of Abel's theorem (7.26) is sometimes referred to as the *elementary addition theorem*, i.e., a specific sum of Abelian integrals is an elementary function (see Kleiman [119]).

The more general version of Abel's theorem, often known as the Abel addition theorem (see again [119]) has the following form:

Theorem 7.2 *Let $A(x)$ be given by (7.25), defined in terms of the rational function $r(x, y)$ and the polynomial function $f(x, y)$, then there is an integer $p \geq 0$, depending only on f, such that, for any set of points $\{x_1, , \cdots, x_\alpha\}$, there are points $\{y_1, \cdots, y_p\}$ so that*

$$A(x_1)+A(x_2)+\ldots+A(x_\alpha) = e(x_1, \cdots, x_\alpha)+A(y_1)+A(y_2)+\ldots+A(y_p), \tag{7.27}$$

where e is an elementary function of $(x_1, \cdots, x_\alpha)$, and $y_1, \cdots, y_p$ are algebraic functions of $(x_1, \cdots, x_\alpha)$.

Note that in (7.26) we have only elementary functions on the right-hand side, and in the special elliptic integral case $r(x, y) = 1/y$, $f(x, y) = y^2 - x^4 - 1$, (7.24), there is only one elliptic integral on the right-hand side (no elementary functions). In this case we had $\alpha = 2$, but we could have iterated (7.24) and had any number of terms on the left-hand side and still had one term on the right-hand side. Thus, in

[10] Note that there is no auxiliary polynomial $g(x, y)$ in this simple case.

this case, p, for $f = x^2 - x^4 - 1$, seems to be equal to 1, and that indeed turns out to be the case. We will discuss the significance of the integer p in Abel's theorem (7.27) somewhat later in this section.

There is one important issue in understanding or interpreting Abel's two theorems here. The first is the multivalued nature of $y(x)$ as implicitly defined by the equation $f(x, y) = 0$, and the second is: what does the integral

$$\int_{x_0}^{x} r(x, y(x))dx \tag{7.28}$$

mean? Here we are now thinking of the integral in (7.28) as a definite integral from some fixed point x_0 to some variable end point x. Namely, the implicitly defined function $y(x)$ defined by the equation $f(x, y) = 0$ is *a priori* a *multivalued function*, and hence the integral (7.28) is also a multivalued function.

Abel dealt with these issues in a straightforward manner, and, as we mentioned earlier, he thought in terms of antiderivatives and differentiation, and his proofs involve differentiation; the fundamental theorem of calculus; the implicit function theorem; and, quite importantly, the general fact, apparently quite well known at the time, that a rational symmetric function of the roots of a polynomial is a rational function of the coefficients of the polynomial. This is a result due to Vandermonde [224], as pointed out by Kleiman [119]. It was used repeatedly by Abel to reexpress various (symmetric) functions of the multivalued functions as single-valued functions.

Abel's work in this early part of the nineteenth century led to vigorous work in the latter half of that same century to understand better this issue of multivalued functions appearing in his work; the most decisive next steps were taken by Bernhard Riemann (1826–1866) [202] in 1857, as we shall see later in Chap. 10. One aspect of the integration issue that was recognized by Abel, and which was definitively pursued by Riemann, was the fact that the integral $\int_{x_0}^{x} r(x, y(x))dx$ could have different values depending on the path one took from the initial point x_0 to the final point x. On the real line there seems to be only one path, but one could specify which signs to use in any formula for $y(x)$ involving various combinations of radicals, for instance.

The possible ambiguities in this integral became known as *periods* of the integral, as differences of two such integrals were specific multiples of fixed entities. At the time of Riemann and later, the variables (x, y) were interpreted as complex numbers, and the integral (7.28) was considered as a complex path integral from x_0 to x along some complex path γ. Whether the integral along two different paths was the same or not became a major subject of study in complex analysis (Cauchy's integral theorem and residue theory) and in what became algebraic topology (whether the two paths bounded a simply-connected domain or not). Both topics became major research areas in the second half of the nineteenth century.

Finally, we want to discuss the significance of the integer p in Abel's theorem (7.27), which is the number of Abelian integrals on the right-hand side of (7.27). First, let us quote from p. 172 of Abel's paper [2], where he denoted the Abelian integrals $A(x_j)$ as $\psi_j x_j$, and p was the difference of the two integers: μ, the total number of Abelian integrals appearing in the theorem (on both sides of the equation in (7.27)), and the integer α, the number of Abelian integrals appearing on the left-hand side of the theorem. In Abel's words:

> Dans cette formule les nombre des fonctions $\psi_{\alpha+1}x_{\alpha+1}$,$\psi_{\alpha+1}x_{\alpha+2}$,...,$\psi_{\mu}x_{\mu}$ est très-remarquable. Plus il est petit, plus la formule est simple. Nous allons, dans ce qui suit, chercher la moindre valeur dont ce nombre, qui est eprimé par $\mu - \alpha$, est susceptible.[11]

Strangely enough, Abel never expressed this number, which we have called p, by a single symbol, in spite of the significance he attributed to this integer, which only depended on the polynomial $f(x, y)$. Note that Abel asserts that p is *small*. What he means by this is that the left-hand side of (7.27) can have an arbitrarily large number of Abelian integral terms relative to the right-hand side, which has a fixed number, p, of Abelian integrals. Abel proceeds to derive formulas which allow him to compute this number in various special cases, and we mention three such cases here.

The first is the most complex. Namely, consider a polynomial $f(x, y)$ of degree 13, i.e.,

$$f(x, y) = p_0 + p_1 y + p_2 y^2 + \ldots + p_{12} y^{12} + y^{13},$$

where the degrees of the polynomials (in the variable x) $p_0, p_1, \cdots, p_{12}$ are

$$2, 3, 2, 3, 4, 5, 3, 4, 2, 3, 4, 1, 1.$$

In this case, after four pages of computation (pp. 181–185 of [2]), Abel obtains $p = 38$.

This number p turns out to be the celebrated *genus* of the algebraic curve defined by $f(x, y) = 0$, and is a topological invariant of the Riemann surface (and topological manifold) defined by the algebraic curve. Riemann formulated the genus in the more modern sense a half-century later. Note that the definition of genus as defined by Abel was an invariant of the analytical data he had at his disposal, and later became a topological invariant in the hands of Riemann.

In the case when

$$f(x, y) = y^2 - \varphi(x),$$

the *hyperelliptic* case, which was studied extensively by Abel in [4], one finds that if $d = \deg \varphi$, where we assume that φ has distinct roots, then

$$p = \begin{cases} (d-1)/2, & \text{if } d \text{ is odd,} \\ (d-2)/2, & \text{if } d \text{ is even.} \end{cases}$$

So, if we have an *elliptic curve* in this hyperelliptic case, i.e., $d = 3$ or 4, then $p = 1$, which means topologically (as we learn later from Riemann [202]) that the elliptic curve is a two-dimensional torus. In this case any sum of Abelian integrals

[11] In this formula the number of functions $\psi_{\alpha+1}x_{\alpha+1}$, $\psi_{\alpha+1}x_{\alpha+2}$, ..., $\psi_{\mu}x_{\mu}$ is very remarkable. Moreover, it is small and the formula is simple. We shall, in that which follows, search for the smallest value for which this number, which is expressed by $\mu - \alpha$, can be attained.

(these would be now elliptic integrals) is the sum of one such elliptic integral plus an elementary function (as in the special case of (7.24) above).

Our final and simplest example is the case $y^2 = Ax^2 + Bx + C$, which gives $p = 0$. This means that the underlying Riemann surface is the Riemann sphere, which is, topologically, a simple two-sphere. We mention again in this very special hyperelliptic case that since $p = 0$, the right-hand side of Abel's theorem (7.27) contains no Abelian integrals, only elementary functions, as we know from the earlier work of Euler discussed earlier (7.9) on the rational parametrization of an algebraic curve of degree two.

One final note is that an Abelian integral is called of the *first kind* if the integral is finite for all x. This terminology was introduced by Legendre in the case of elliptic integrals in [138]. For instance, the following Abelian integrals in the hyperelliptic case (where $f(x, y) = y^2 - \varphi(x)$ and $\varphi(x)$ has distinct roots) are of the first kind, where p is again the genus of the hyperelliptic curve $f(x, y) = 0$,

$$\int_{x_o}^{x} \frac{dx}{\sqrt{\varphi(x)}}, \int_{x_o}^{x} \frac{xdx}{\sqrt{\varphi(x)}}, \ldots, \int_{x_o}^{x} \frac{x^{p-1}dx}{\sqrt{\varphi(x)}}. \tag{7.29}$$

In this case these p Abelian integrals of the first kind in (7.29) are linearly independent and they span the space of all such Abelian integrals of the first kind (see Markushevich [151]). We will see these Abelian integrals in greater detail later in Chap. 10. Note that the genus p appears here explicitly, and the dimension of this vector space of all Abelian integrals of the first kind can be used as a second and equivalent definition of genus in this case.

Chapter 8
Elliptic Functions

8.1 Introduction

Abel's second major work [3], which we are exploring in this book, published in Volumes 2 and 3 of Crelle's journal in 1827 and 1828, was a definitive and foundational paper on elliptic functions. The title of this paper, *Recherches sur les fonctions elliptiques*, is misleading, and at the same time, so very appropriate. What he meant in the title by "elliptic functions," as he explains in his paper, were the transcendental functions studied by Euler and Legendre, etc., which were defined by and known as elliptic integrals, and which he generalized in the concept of what we now call Abelian integrals, as described in our previous chapter. In this paper Abel introduced, for the first time, the *inverses* of the elliptic integral functions, and these became the now familiar doubly-periodic meromorphic functions on the complex plane known as elliptic functions, which we will discuss in the forthcoming sections. So the title is absolutely correct in modern times, even if Abel didn't know it at the time!

This was followed one year later by the equally definitive and independent work of Carl Gustav Jacob Jacobi (1804–1851) [116] on precisely the same subject (Jacobi had published a shorter introduction to his work at the end of 1827 [115]). These two long papers by Abel and Jacobi laid the foundation for the rich development of the theory of doubly-periodic functions in the complex plane that was pursued by numerous mathematicians throughout the nineteenth century in a wide variety of forms (complex analysis, algebraic geometry, number theory, etc.).

However, before we look at Abel's and Jacobi's work, let's briefly review what functions of a complex variable meant to mathematicians at the beginning of the nineteenth century. As we saw in Chap. 6, the geometric representation of complex numbers in the complex plane had not yet been developed. Complex numbers were simply algebraic combinations of real numbers with the imaginary unit $i = \sqrt{-1}$ of the form $a + ib$ manipulated according to the well-known rules of addition and multiplication of such numbers. In reading through the work of Euler from the mid-eighteenth century [62], which we have cited in Chap. 6, one sees that imaginary numbers arose from solving algebraic equations and were manipulated by the usual

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_8

rules of algebra. A rational function f of a complex variable $x + iy$ computed $f(x + iy)$ by algebra, i.e.,

$$(x + iy)^2 = x^2 - y^2 + i(2xy),$$

and a series of the form

$$\sum_{n=0}^{\infty} a_n (x + iy)^n$$

would be expressed in terms of its real and imaginary parts by term-by-term algebra.

For transcendental functions we find a pregnant remark of Euler on p. 96 of [67] (the 1796 French edition of his analysis book from 1748) which says (in English translation), where here x is a real number:

> Since $\sin^2 x + \cos^2 x = 1$, in decomposing into factors, one would have
>
> $$(\cos x + \sqrt{-1} \sin x)(\cos x - \sqrt{-1} \sin x) = 1. \tag{8.1}$$
>
> These factors, although imaginary, are of great usage in the combination and multiplication of arcs [radian angles].

A few pages later in the same book Euler observes that (now letting $i = \sqrt{-1}$, for convenience), by letting

$$e^{ix} = \cos x + i \sin x,$$

the first factor in (8.1), then

$$\cos x = \frac{e^{ix} + e^{-ix}}{2},$$
$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Using the addition formula for exponentials he then obtains (by definition)

$$e^{x+iy} := e^x e^{iy} = e^x(\cos x + i \sin y),$$

with similar expressions for the transcendental functions of a complex variable $\sin(x+iy)$, $\cos(x+iy)$, etc. These are then examples of transcendental functions of a complex variable represented as algebraic combinations (involving the imaginary unit i) of real-valued functions of real variables.

This was the stage that was set for Abel and Jacobi as they set out to create their theories of elliptic functions (which would also be formulated initially as algebraic combinations of real-valued functions, just as Euler did with the trigonometric functions).

Let us now formulate what an elliptic function is in the standard language of complex analysis. Namely, let ω_1 and ω_2 be two fixed complex numbers such that

Im $(\omega_1/\omega_2) \neq 0$, then an *elliptic function* $f(z)$ with the two periods ω_1 and ω_2 is a meromorphic function on the complex plane $\mathbf{C}$ such that

$$f(z + m\omega_1 + n\omega_2) = f(z), \text{ for all } n, m \in \mathbf{Z},$$

where $\mathbf{Z}$ denotes the ring of integers. We say that such a function is *doubly-periodic with the two periods* ω_1 and ω_2. This is completely analogous to the simply-periodic functions from trigonometry, where, for instance,

$$\sin(x + 2\pi n) = \sin(x), \text{ for all } n \in \mathbf{Z},$$

with the period 2π.

This formulation of an abstract family of functions with double-periodicity is due to the work of Joseph Liouville (1809–1882). In 1844 he announced [147] the theorem that a doubly-periodic function which is holomorphic must be a constant; and in the same issue of the journal, Cauchy gave a proof of Liouville's theorem using his Cauchy integral theorem. Liouville published considerably later, in 1880 [148], an in-depth paper that he had written in 1847, which included many properties of doubly-periodic functions including the theorem just mentioned.

In Sect. 9.5 we give a brief outline of Weierstrass's theory of doubly-periodic functions that he presented in the second half of the nineteenth century, primarily in his Berlin lectures in 1863, which were published in 1915 [235], but which included results from earlier in his career as well.

Abel and Jacobi gave the first *examples* of such doubly-periodic functions (that is, there *are* doubly-periodic functions as defined above), and they proved many of their important properties, as well as giving a variety of ways to represent such functions (power series, infinite products, etc.). The theory of trigonometric functions was a model for both of them. In Sect. 8.2 we look at Abel's work on elliptic functions, and then in Sect. 8.3 we will see how Jacobi covered most of the same material with one important innovation, namely the theory of theta functions.

8.2 Abel's *Recherches sur les fonctions elliptiques*

Let's start with Abel's paper [4], and we will follow the notation and normalizations used in his paper, although the formalism and notation of Jacobi became the standard in the literature in the following decades. Due to Abel's early death, he was not able to participate in the later developments. The basic idea of both mathematicians was to study the *inverse* of the elliptic integral functions that had been studied extensively by their predecessors. In this manner the addition theorems for elliptic integrals, à la Euler, became addition formulas for the elliptic functions, which generalized the addition formulas for trigonometric functions.

Let us note that if one starts with the transcendental function

$$\arcsin(x) := \int_0^x \frac{dx}{\sqrt{1-x^2}},$$

then one can define its inverse $\sin(x)$ and obtain the full theory of trigonometric functions. This is, in effect, what Abel and Jacobi do in the elliptic integral context.

Abel begins in [4] by recalling the work of Euler and Legendre that we discussed in the preceding paragraphs. He notes that every elliptic integral of the form

$$\int \frac{R(x)dx}{\sqrt{\alpha + \beta x + \gamma x^2 + \delta x^3 + \varepsilon x^4}},$$

where $R(x)$ is a rational function, can be reduced to

$$\int \frac{P(y)dy}{\sqrt{a + by^2 + cy^4}},$$

where $P(y)$ is a rational function of y^2. This can, in turn, be reduced to the form

$$\int \frac{A + By^2}{C + Dy^2} \frac{dy}{\sqrt{a + by^2 + cy^4}},$$

and by yet one more change of variables, this can be reduced to the trigonometric form

$$\int \frac{A + B\sin^2\theta}{C + D\sin^2\theta} \frac{d\theta}{\sqrt{1 - c^2\sin^2\theta}},$$

where c is real and $|c| < 1$. Finally, Abel notes that (all of this is from Legendre's book [138]), every elliptic integral, by this type of reduction, can be reduced to the three cases:

$$\int \frac{d\theta}{\sqrt{1-c^2\sin^2\theta}}, \int d\theta\sqrt{1-c^2\sin^2\theta}, \int \frac{d\theta}{(1+n^2\sin^2\theta)\sqrt{1-c^2\sin^2\theta}},$$

which Legendre calls elliptic integrals of the first, second and third kind. Abel decides to concentrate on the elliptic integrals of the first kind; and on p. 164 of [4], after the brief introduction outlined above, he says;

> Je me propose, dans ce mémoire, de considérer le fonction inverse, c'est à dire la fonction $\varphi\alpha$, detérminée par les équations[1]
>
> $$\alpha = \int \frac{d\theta}{\sqrt{1-c^2\sin^2\theta}},$$
> $$\sin\theta = \varphi\alpha = x.$$

[1] "I propose, in this memoir, to consider the inverse function, that is to say the function $\varphi\alpha$ determined by the equations".

Abel then considers specifically the elliptic integral of the first kind in the form

$$\alpha(x) = \int_0^x \frac{dtx}{\sqrt{1-t^2}\sqrt{1-c^2t^2}}, \tag{8.2}$$

in terms of the variable x, where again $c^2 > 0$.[2] Now Abel makes two changes in notation to suit his purposes. He replaces c^2 by $-e^2$ and replaces the term $\sqrt{1-x^2}$ by $\sqrt{1-c^2x^2}$ for symmetry, and finally considers the specific elliptic integral of the first kind in the form

$$\alpha(x) = \int_0^x \frac{dt}{\sqrt{1-c^2t^2}\sqrt{1+e^2t^2}}. \tag{8.3}$$

We let $x(\alpha)$ be the inverse of $\alpha(x)$ given by (8.3), which is well defined near $x = 0$, and Abel defines $\varphi(\alpha)$ to be this inverse $x(\alpha)$ on a suitable interval containing $x = 0$. The derivative of $\alpha(x)$ is simply given by

$$\alpha'(x) = \frac{1}{\sqrt{1-c^2x^2}\sqrt{1+e^2x^2}};$$

and, by the inverse function theorem, the derivative of $\varphi(\alpha)$ is given by

$$\varphi'(\alpha) = \sqrt{1-c^2\varphi(\alpha)^2}\sqrt{1+e^2\varphi(\alpha)^2}. \tag{8.4}$$

Then Abel introduces two additional functions of α defined by

$$f(\alpha) := \sqrt{1-c^2\varphi(\alpha)^2},\ F(\alpha) := \sqrt{1+e^2\varphi(\alpha)^2}, \tag{8.5}$$

which appear in (8.4), yielding $\varphi'(\alpha) = f(\alpha)F(\alpha)$. These *three* functions of a real variable[3] α are the generalizations of the *two* trigonometric functions $\sin(\alpha)$ and $\cos(\alpha)$, and, as Abel says on p. 265 of his paper:

> Plusieurs propriétés de ces fonctions se dédusierent immédiatement des propriétés connues de la fonction elliptique [elliptic integral] de la première espèce, mais d'autres sont plus cachées. Par exemple on démontre que les équations $\varphi\alpha = 0$, $f\alpha = 0$, $Fa = 0$ ont un nombre infini de racines, qu'on peut trouver toutes. Une des les plus remarquables est qu'on peut exprimer rationellement $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$ (m un nombre entier) en $\varphi\alpha$, $f\alpha$, Fa. Aussi rien n'est plus facile que de trouver $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$, lorsqu'on connaît $\varphi\alpha$, $f\alpha$, $F\alpha$; mais le problème inverse, savoir de déterminer $\varphi\alpha$, $f\alpha$, $F\alpha$ en $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$, est plus difficile, parcequ'il dépend d'une equation d'un degré élevé (savoir du degré m^2).

[2] Abel doesn't distinguish between the upper limit of the integral and the variable of integration, but we do to clarify the discussion.

[3] The inverse function $\varphi(\alpha)$ and its related functions $f(\alpha)$ and $F(\alpha)$ are well defined locally near $\alpha = 0$ by the inverse function theorem. The extension to the full real line is discussed later in this section.

> La résolution de cette équation est l'objet principal de ce mémoire. D'abord on fera voir, comment on peut trouver toutes les racines, au moyen des fonctions φ, f, F. On traitera ensuite de la résolutions algébrique de l'équation en question, et on parviendra à ce résultat remarquable, que $\varphi\frac{\alpha}{m}$, $f\frac{\alpha}{m}$, $F\frac{\alpha}{m}$ peuvent être exprimées en $\varphi\alpha$, $f\alpha$, $F\alpha$, par une formule qui, par rapport à α, ne contient d'autre irrationalité que des radicaux. Cela donne une classe très générale d'équations qui sont résoluble algébriquement.[4]

We note that this last comment of Abel's about solvability of high-degree equations by means of extracting roots relates to one of his first papers [1] in which he shows for the first time the unsolvability in terms of radicals of generic algebraic equations of degree 5 or higher, a problem that had been outstanding for a long time. The definitive work on whether a given polynomial equation was solvable in terms of radicals was due to Évariste Galois (1811–1832) in his work which established the now well-known Galois theory. This was published in 1846 [71], 14 years after his very early death at the age of 20.

Let us now turn to Abel's construction of his version of elliptic functions and their first fundamental properties. He first defines each of these functions for all real values of α in a specific interval around the origin, and then proceeds to define them as functions of a complex variable $\alpha + i\beta$ on the entire complex plane in a sequence of steps. First he sets

$$\frac{\omega}{2} := \int_0^{\frac{1}{c}} \frac{dt}{\sqrt{1-c^2t^2}\sqrt{1+e^2t^2}},$$

where it is simple to verify that the limiting integral at the singular point $x = \frac{1}{c}$ is well-defined. Thus one sees that $\varphi(\alpha) > 0$ on $(0, \omega/2)$, and $\varphi(0) = 0$ and $\varphi(\omega/2) = 1/c$. Also, from the definition of $\varphi(\alpha)$, one sees that $\varphi(-\alpha) = -\varphi(\alpha)$, and thus we have $\varphi(\alpha)$ well defined on $[-\omega/2, \omega/2]$, and similarly for $f(\alpha)$ and $F(\alpha)$. Now Abel wants to define these functions for imaginary numbers of the form $i\beta$.

For this he formally substitutes iy for x in (8.3) and integrates the integrand of the elliptic integral in (8.3) on the imaginary axis from 0 to iy, obtaining

$$i\int_0^{y} \frac{dt}{\sqrt{1+c^2t^2}\sqrt{1-e^2t^2}},$$

[4] "Several properties of these functions are deducible immediately from the known properties of the elliptic function [elliptic integral] of the first kind, but others are more hidden. For example, one can show that the equations $\varphi\alpha = 0$, $f\alpha = 0$, $Fa = 0$ have an infinite number of roots, where one can find all of them. One of the most remarkable properties is that one can express rationally $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$ (m an integer) in $\varphi\alpha$, $f\alpha$, Fa. Also, nothing is more simple than to find $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$, when one knows $\varphi\alpha$, $f\alpha$, $F\alpha$; but the inverse problem, to know how to determine $\varphi\alpha$, $f\alpha$, $F\alpha$ in $\varphi(m\alpha)$, $f(m\alpha)$, $F(m\alpha)$, is more difficult, since it depends on an equation of high degree (more specifically of degree m^2).

The solution of this equation is the principal object of this memoir. At first one can see how one can find all the roots, by means of the functions φ, f, F. Then one treats the algebraic solution of the equation in question, and one comes to this remarkable result, that $\varphi\frac{\alpha}{m}$, $f\frac{\alpha}{m}$, $F\frac{\alpha}{m}$ can be expressed in $\varphi\alpha$, $f\alpha$, $F\alpha$, by a formula, which, with respect to α, doesn't contain any irrationality except radicals. This gives a very general class of equations which are solvable algebraically."

where we see that the roles of e and c have been interchanged. Let

$$\beta(y) := \int_0^y \frac{dt}{\sqrt{1+c^2t^2}\sqrt{1-e^2t^2}},$$

which is again a monotone increasing function on the interval $[-\frac{\tilde{\omega}}{2}, \frac{\tilde{\omega}}{2}]$, where

$$\frac{\tilde{\omega}}{2} := \int_0^{\frac{1}{e}} \frac{dx}{\sqrt{1+c^2x^2}\sqrt{1-e^2x^2}},$$

and we let the inverse of $\beta(y)$ on this interval be denoted by $y(\beta)$.

We have already defined $\varphi(\alpha)$ to be $x(\alpha)$ on $[-\omega/2, \omega/2]$, and now we define similarly $\varphi(i\beta) := iy(\beta)$ on the interval $[-i\frac{\tilde{\omega}}{2}, i\frac{\tilde{\omega}}{2}]$ on the imaginary axis. We then define on this same interval

$$f(i\beta) := F(\beta), \text{ and } F(i\beta) = f(\beta),$$

using the interchange of c and e in the definition of $\alpha(x)$ and $\beta(y)$. We note that $\varphi(\pm\frac{\omega}{2}) = \pm\frac{1}{c}$ and $\varphi(\pm i\frac{\tilde{\omega}}{2}) = \pm i\frac{1}{e}$.

Thus, at this point $\varphi(\alpha)$ and $\varphi(i\beta)$ are defined for $\omega/2 \le \alpha \le \omega/2$, and $-\tilde{\omega}/2 \le \beta \le \tilde{\omega}/2$. The problem remains to define $\varphi(\alpha)$ and $\varphi(i\beta)$ for all α and β, and to then define $\varphi(\alpha + i\beta)$ for all complex numbers $\alpha + i\beta$.

For both of these tasks Abel needs a tool, and that is a specific generalization of the usual addition formulas for sines and cosines. Abel formulates these new addition formulas for the three functions $\varphi(\alpha)$, $f(\alpha)$, $F(\alpha)$ as follows:

$$\varphi(\alpha+\beta) = \frac{\varphi(\alpha)f(\beta) + \varphi(\beta)f(\beta)f(\alpha)F(\alpha)}{1+e^2c^2\varphi^2(\alpha)\varphi^2(\beta)}, \tag{8.6}$$

$$f(\alpha+\beta) = \frac{f(\alpha)f(\beta) - c^2\varphi(\alpha)\varphi(\beta)F(\alpha)F(\beta)}{1+e^2c^2\varphi^2(\alpha)\varphi^2(\beta)}, \tag{8.7}$$

$$F(\alpha+\beta) = \frac{F(\alpha)F(\beta) + e^2\varphi(\alpha)\varphi(\beta)f(\alpha)f(\beta)}{1+e^2c^2\varphi^2(\alpha)\varphi^2(\beta)}. \tag{8.8}$$

We recall briefly the classical formulas for trigonometric functions (as one finds in Euler's *Introductio* [62] from 1748, for instance):

$$\sin(\alpha+\beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta), \tag{8.9}$$

$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta), \tag{8.10}$$

$$\tan(\alpha+\beta) = \frac{\tan(\alpha)+\tan(\beta)}{1-\tan(\alpha)\tan(\beta)}, \tag{8.11}$$

which have the same type of rational expressions as in (8.6), (8.7), and (8.8).

Abel points out that these addition formulas follow from Legendre's theory of elliptic integrals [138], which follows up on the Euler addition theorem for elliptic integrals that we discussed earlier. He also gives a simple and elegant proof which we can sketch here (the same proof will work for the trigonometric formulas listed above as well). First, using the fact that

$$f^2(\alpha) = 1 - c^2\varphi^2(\alpha),$$
$$F^2(\alpha) = 1 + e^2\varphi^2(\alpha),$$

then, by differentiating, we obtain

$$f(\alpha)f'(\alpha) = -c^2\varphi(a)\varphi'(\alpha), \tag{8.12}$$
$$F(\alpha)F'(\alpha) = 1 + e^2\varphi(\alpha)\varphi'(\alpha), \tag{8.13}$$

and from (8.3) we have

$$\varphi'(\alpha) = \sqrt{1 - c^2\varphi^2(\alpha)}\sqrt{1 + e^2\varphi^2(\alpha)} = f(\alpha)F(\alpha). \tag{8.14}$$

Substituting (8.14) in (8.12) and (8.13), we find that

$$\varphi'(\alpha) = f(\alpha)F(\alpha),$$
$$f'(\alpha) = -c^2\varphi(\alpha)F(\alpha),$$
$$F'(\alpha) = c^2\varphi(\alpha)f(\alpha),$$

the elliptic function analogue to $(\sin(\alpha))' = \cos(\alpha)$, etc.

Now for the proof of, for instance (8.6), we denote the right-hand side of (8.6) by $r(\alpha, \beta)$, and compute both $\frac{\partial r}{\partial \alpha}$ and $\frac{\partial r}{\partial \beta}$ using the differentiation formulas above. It turns out that α and β appear symmetrically in these expressions and that one verifies by inspection that

$$\frac{\partial r}{\partial \alpha} = \frac{\partial r}{\partial \beta}. \tag{8.15}$$

As was known at the time, a solution of the partial differential equation (8.15) is a function of the sum $\alpha + \beta$, and hence there is a function ψ of one variable such that

$$r(\alpha, \beta) = \psi(\alpha + \beta).$$

One can find ψ by looking at particular values, and for instance, for $\beta = 0$, we have $\varphi(0) = 0$, $f(0) = 1$, $F(0) = 1$, and hence

$$r(\alpha, 0) = \varphi(a) = \psi(\alpha),$$

so

$$r(\alpha, \beta) = \psi(\alpha + \beta) = \varphi(\alpha + \beta),$$

and (8.6) is proved. The addition formulas (8.7) and (8.8) can be proved in the same manner.[5]

Abel uses these addition formulas to define in a natural manner the evaluation of the elliptic functions on the real line for $|\alpha| > \omega/2|$ and on the imaginary axis for $|i\beta| > \tilde{\omega}/2$. Then he also invokes the addition formula to define, for instance,

$$\varphi(\alpha + i\beta) = \frac{\varphi(\alpha)f(i\beta) + \varphi(i\beta)f(i\beta)f(\alpha)F(\alpha)}{1 + e^2c^2\varphi^2(\alpha)\varphi^2(i\beta)},$$
$$= \frac{\varphi(\alpha)F(\beta) - i\varphi(\beta)F(\beta)f(\alpha)F(\alpha)}{1 - e^2c^2\varphi^2(\alpha)\varphi^2(\beta)},$$

and similarly for the other two elliptic functions f and F.

After having used the addition formulas in this manner, Abel says on p. 279 of [4],

> Des formules (8.6), (8.7), (8.8) on peut déduire une foule d'autres.[6]

In Fig. 8.1 we see a sample of the plethora of formulas that he derives from the basic addition theorems. Here he has used the abbreviation

$$R = 1 + e^2c^2\varphi^2(\alpha)\varphi^2(\beta).$$

After two more pages of calculations we find on p. 272 of his paper (reproduced in Fig. 8.2) the first formulation of the doubly-periodic nature of his elliptic functions. This is equation no. 20 on this page in Fig. 8.2. At the top of the same page we see in the second equation that these elliptic functions all have a pole at the point $(\frac{\omega}{2}, i\frac{\tilde{\omega}}{2})$ (and at the suitable translates of this point as well). This is the first instance in the literature of a doubly-periodic function of a single complex variable.

What is significant for us here is that one cannot formulate this notion of double-periodicity without the use of complex variables, and in the decades that followed, these functions and others related to them, became important objects in the study of meromorphic functions in the complex plane. Later in his paper Abel found many different kinds of representations of these functions. An important historical point is that these functions played a role in applied mathematics as well.

In the remainder of his paper [4] Abel goes on to establish a variety of identities and properties for the elliptic functions he created in this paper, along with applications to the transformations of elliptic integrals and to the special case of the elliptic integral

$$\int \frac{d}{\sqrt{1 - x^4}}$$

[5] Of course this proof depends on *knowing* what the right-hand side of such an addition formula looks like, and this knowledge stems from the work of Euler and Legendre.

[6] "From the formulas (8.6), (8.7), (8.8) one can deduce many others".

$$(12)\quad\left\{\begin{aligned}
\varphi(\alpha+\beta)+\varphi(\alpha-\beta)&=\frac{2\,\varphi\alpha\,.\,f\beta\,.\,F\beta}{R},\\
\varphi(\alpha+\beta)-\varphi(\alpha-\beta)&=\frac{2\,\varphi\beta\,.\,f\alpha\,.\,F\alpha}{R},\\
f(\alpha+\beta)+f(\alpha-\beta)&=\frac{2\,f\alpha\,.\,f\beta}{R},\\
f(\alpha+\beta)-f(\alpha-\beta)&=\frac{-2\,c^2\,.\,\varphi\alpha\,.\,\varphi\beta\,.\,F\alpha\,.\,F\beta}{R},\\
F(\alpha+\beta)+F(\alpha-\beta)&=\frac{2\,F\alpha\,.\,F\beta}{R},\\
F(\alpha+\beta)-F(\alpha-\beta)&=\frac{2\,e^2\,.\,\varphi\alpha\,.\,\varphi\beta\,.\,f\alpha\,.\,f\beta}{R}.
\end{aligned}\right.$$

En formant le produit de $\varphi(\alpha+\beta)$ et $\varphi(\alpha-\beta)$, on trouvera

$$\begin{aligned}
\varphi(\alpha+\beta)\,.\,\varphi(\alpha-\beta)&=\frac{\varphi\alpha\,.\,f\beta\,.\,F\beta+\varphi\beta\,.\,f\alpha\,.\,F\alpha}{R}\;\frac{\varphi\alpha\,.\,f\beta\,.\,F\beta-\varphi\beta\,.\,f\alpha\,.\,F\alpha}{R}\\
&=\frac{\varphi^2\alpha\,.\,f^2\beta\,.\,F^2\beta-\varphi^2\beta\,.\,f^2\alpha\,.\,F^2\alpha}{R^2},
\end{aligned}$$

ou, en substituant les valeurs de $f^2\beta$, $F^2\beta$, $f^2\alpha$, $F^2\alpha$ en $\varphi\beta$ et $\varphi\alpha$,

$$\begin{aligned}
\varphi(\alpha+\beta)\,.\,\varphi(\alpha-\beta)&=\frac{\varphi^2\alpha-\varphi^2\beta-e^2c^2\varphi^2\alpha\,.\,\varphi^4\beta+e^2c^2\varphi^2\beta\,.\,\varphi^4\alpha}{R^2}\\
&=\frac{(\varphi^2\alpha-\varphi^2\beta)(1+e^2c^2\varphi^2\alpha\,.\,\varphi^2\beta)}{R^2};
\end{aligned}$$

or $R=1+e^2c^2\varphi^2\alpha\,.\,\varphi^2\beta$, donc

$$(13)\qquad \varphi(\alpha+\beta)\,.\,\varphi(\alpha-\beta)=\frac{\varphi^2\alpha-\varphi^2\beta}{R}.$$

On trouvera de même

$$(14)\quad\left\{\begin{aligned}
f(\alpha+\beta)\,.\,f(\alpha-\beta)&=\frac{f^2\alpha-c^2\varphi^2\beta\,.\,F^2\alpha}{R}=\frac{f^2\beta-c^2\varphi^2\alpha\,.\,F^2\beta}{R}\\
&=\frac{1-c^2\varphi^2\alpha-c^2\varphi^2\beta-c^2e^2\varphi^2\alpha\,.\,\varphi^2\beta}{R}=\frac{f^2\alpha\,.\,f^2\beta-c^2(c^2+e^2)\,\varphi^2\alpha\,.\,\varphi^2\beta}{R},\\
F(\alpha+\beta)\,.\,F(\alpha-\beta)&=\frac{F^2\alpha+e^2\varphi^2\beta\,.\,f^2\alpha}{R}=\frac{F^2\beta+e^2\varphi^2\alpha\,.\,f^2\beta}{R}\\
&=\frac{1+e^2\varphi^2\alpha+e^2\varphi^2\beta-e^2c^2\varphi^2\alpha\,.\,\varphi^2\beta}{R}=\frac{F^2\alpha\,.\,F^2\beta-e^2(c^2+e^2)\,\varphi^2\alpha\,.\,\varphi^2\beta}{R}.
\end{aligned}\right.$$

Fig. 8.1 Page 270 of Abel's paper on elliptic functions [4]

that Euler had studied, which describes the arc length of a leminiscate ([4], pp. 361–362). In addition, he obtains a variety of representations of theelliptic functions in terms of infinite series and infinite products.

$$(18)\quad \left\{\begin{aligned} &\varphi\left(\alpha \pm \frac{\omega}{2}\right)\varphi\left(\alpha + \frac{\varpi}{2}i\right) = \pm\frac{i}{ce};\quad F\left(\alpha \pm \frac{\omega}{2}\right)F\alpha = \frac{\sqrt{e^2+c^2}}{c};\\ &\qquad\qquad f\left(\alpha \pm \frac{\varpi}{2}i\right)f\alpha = \frac{\sqrt{e^2+c^2}}{e}.\end{aligned}\right.$$

En faisant $\alpha = \frac{\omega}{2}$ et $\frac{\varpi}{2}i$, on en déduit

$$\varphi\left(\frac{\omega}{2} + \frac{\varpi}{2}i\right) = \tfrac{1}{0},\quad f\left(\frac{\omega}{2} + \frac{\varpi}{2}i\right) = \tfrac{1}{0},\quad F\left(\frac{\omega}{2} + \frac{\varpi}{2}i\right) = \tfrac{1}{0}.$$

En mettant ensuite dans les trois premières équations (17) $\alpha + \frac{\omega}{2}$ au lieu de α, et dans les trois dernières $\alpha + \frac{\varpi}{2}i$ au lieu de α, on obtiendra les suivantes

$$(19)\quad \left\{\begin{aligned} &\varphi(\alpha+\omega) = -\varphi\alpha;\ f(\alpha+\omega) = -f\alpha;\ F(\alpha+\omega) = F\alpha;\\ &\varphi(\alpha+\varpi i) = -\varphi\alpha;\ f(\alpha+\varpi i) = f\alpha;\ F(\alpha+\varpi i) = -F\alpha;\end{aligned}\right.$$

et en mettant $\alpha+\omega$ et $\alpha+\varpi i$ au lieu de α:

$$(20)\quad \left\{\begin{aligned} &\varphi(2\omega+\alpha) = \varphi\alpha;\ \varphi(2\varpi i+\alpha) = \varphi\alpha;\ \varphi(\omega+\varpi i+\alpha) = \varphi\alpha;\\ &f(2\omega+\alpha) = f\alpha;\quad f(\varpi i+\alpha) = f\alpha;\\ &F(\omega+\alpha) = F\alpha;\ F(2\varpi i+\alpha) = F\alpha.\end{aligned}\right.$$

Ces équations font voir que les fonctions $\varphi\alpha$, $f\alpha$, $F\alpha$ sont des fonctions *périodiques*. On en déduira sans peine les suivantes, où m et n sont deux nombres entiers positifs ou négatifs:

$$(21)\quad \left\{\begin{aligned} &\varphi[(m+n)\omega+(m-n)\varpi i+\alpha] = \varphi\alpha;\\ &\varphi[(m+n)\omega+(m-n+1)\varpi i+\alpha] = -\varphi\alpha;\\ &f(2m\omega+n\varpi i+\alpha) = f\alpha;\ f[(2m+1)\omega+n\varpi i+\alpha] = -f\alpha.\\ &F(m\omega+2n\varpi i+\alpha) = F\alpha;\ F[m\omega+(2n+1)\varpi i+\alpha] = -F\alpha.\end{aligned}\right.$$

Ces formules peuvent aussi s'écrire comme il suit:

$$(22)\quad \left\{\begin{aligned} &\varphi(m\omega+n\varpi i\pm\alpha) = \pm(-1)^{m+n}\varphi\alpha,\\ &f(m\omega+n\varpi i\pm\alpha) = (-1)^m f\alpha,\\ &F(m\omega+n\varpi i\pm\alpha) = (-1)^n F\alpha.\end{aligned}\right.$$

On peut remarquer comme cas particuliers:

Fig. 8.2 Page 272 of Abel's paper on elliptic functions [4]

8.3 Jacobi's *Fundamenta Nova*

As we mentioned earlier, Jacobi had announced his discovery of elliptic functions in a short paper in December of 1827 [115] and followed up with his foundational 190 page paper [116]. Interestingly, both of these papers were published in the *Astronomische Nachrichten*, edited by Heinrich Christian Schumacher (1780–1850), an important astronomer at that time. Applications to astronomy of this new theory seemed to have been an important motivation for Jacobi.

We will look at some of the innovations in Jacobi's paper [116]. First, he proceeds in a similar manner to what Abel did at roughly the same time. Namely, he considers the inverse of the elliptic integral

$$u(x) = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{8.16}$$

to be

$$\begin{aligned} \varphi &= \operatorname{am} u, \\ x &= \sin \operatorname{am} u. \end{aligned}$$

Jacobi defines

$$K := \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}},$$

using the substitution $x = \sin\varphi$. He then defines a number of other functions related to $\sin \operatorname{am}(u)$, which have now become standard in the theory of elliptic functions.

We discuss the most important ones here briefly. Namely, we have the two additional functions $\cos \operatorname{am} u$ and

$$\Delta \operatorname{am} u := \frac{d}{du}(\operatorname{am} u) = \sqrt{1-k^2 \sin \operatorname{am}^2 u}.$$

This was the notation of Jacobi, and towards the end of the nineteenth century it has become standard to write

$$\begin{aligned} \operatorname{sn} u &:= \sin \operatorname{am} u, \\ \operatorname{cn} u &:= \cos \operatorname{am} u, \\ \operatorname{dn} u &:= \Delta \operatorname{am} u, \end{aligned}$$

for these three functions, which are the analogues of the three elliptic functions of Abel, φ, f, and F. These satisfy the properties

$$\begin{aligned} \operatorname{sn}^2 u + \operatorname{cn}^2 u &= 1, \\ \operatorname{sn}^2 u + k^2 \operatorname{dn}^2 u &= 1, \end{aligned}$$

the analogues of $\sin^2 x + \cos^2 x = 1$ in this context, and they satisfy addition formulas, which are formulated explicitly by Jacobi for these three functions and other related functions (see [247] or [113] for proofs of these addition formulas). Jacobi does not prove these formulas, but depends on the earlier work of Legendre on elliptic integrals [138] for proofs in this elliptic-functions context.

Jacobi then defines $\operatorname{sn} iv$, $\operatorname{cn} iv$, and $\operatorname{dn} iv$, in the same manner as Abel, and using the addition formulas extends his elliptic functions to be functions of a complex variable, e.g., $\operatorname{sn}(u + iv)$. These functions are doubly-periodic, which follows easily from the addition formulas. For instance, letting K' be defined by

$$K^2 + (K')^2 = 1,$$

one finds that

$$\operatorname{sn}(u + iv + 4K) = \operatorname{sn}(u + iv), \text{ and } \operatorname{sn}(u + iv + i2K') = \operatorname{sn}(u + iv),$$

which shows that $\operatorname{sn}(u + iv)$ has two independent periods, $4K$ along the real axis, and $i2K'$ along the imaginary axis. One can find a complete set of these period relations for all of the Jacobi elliptic functions in [247].

8.4 Jacobi's Theta Functions

In his *Fundamenta Nova* paper Jacobi obtains an extensive set of properties for the elliptic functions, many of which are similar to those derived by Abel (representation in terms of power series, infinite products, solutions of certain differential equations, etc.). Then on p. 198 of [116] he defines for the first time a new concept, which gives an important method of representing the Jacobi elliptic functions, and which becomes intrinsically very important in mathematics, independent of the theory of elliptic functions. This is Jacobi's discovery of *theta functions*, as they have been called ever since. A theta function is a rapidly convergent Fourier series with quasiperiodic properties, and the quotient of two such functions can represent an elliptic function. In Fig. 8.3 we see his definition of $\Theta(u)$ in the middle of p. 198 in his *Fundamenta Nova* paper [116].

Let us give an example of two such theta functions whose quotient is an elliptic function. Our notation differs from that used byJacobi, but it is the same thing mathematically. Let

$$\theta(z;\tau) := \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau + 2\pi i z}$$

be a theta function which is defined for $z \in \mathbf{C}$, $\tau \in \mathbf{C}$ with $\operatorname{Im}(\tau) > 0$. We consider $\theta(z;\tau)$ as a function of z, with τ as a parameter. Since $\operatorname{Im} \tau > 0$, it follows that

198 EVOLUTIO FUNCTIONUM ELLIPTICARUM.

a $x = 0$ usque ad $x = x$, prodit:

$$\frac{2K}{\pi}\int_0^x Z\left(\frac{2Kx}{\pi}\right)dx = -2\left\{\frac{q\cos 2x}{1-q^2} + \frac{q^2\cos 4x}{2(1-q^4)} + \frac{q^3\cos 6x}{3(1-q^6)} + \cdots\right\} + \text{const.}$$
$$= \log[(1-2q\cos 2x+q^2)(1-2q^3\cos 2x+q^6)(1-2q^5\cos 2x+q^{10})\cdots] + \text{const.},$$

ubi *constans*, ita determinata, ut integrale pro $x = 0$ evanescat, fit:

$$= 2\left\{\frac{q}{1-q^2} + \frac{q^2}{2(1-q^4)} + \frac{q^3}{3(1-q^6)} + \cdots\right\} = -\log[(1-q)(1-q^3)(1-q^5)\cdots]^2,$$

ideoque:

$$(1.)\quad \frac{2K}{\pi}\int_0^x Z\left(\frac{2Kx}{\pi}\right)dx = \log\left\{\frac{(1-2q\cos 2x+q^2)(1-2q^3\cos 2x+q^6)(1-2q^5\cos 2x+q^{10})\cdots}{[(1-q)(1-q^3)(1-q^5)\cdots]^2}\right\}.$$

Designabimus in posterum per characterem $\Theta(u)$ expressionem:

$$\Theta(u) = \Theta(0)e^{\int_0^u Z(u)du},$$

designante $\Theta(0)$ constantem, quam adhuc indeterminatam relinquimus, dum commodam eius determinationem infra obtinebimus; erit ex (1.):

$$(2.)\quad \frac{\Theta\left(\frac{2Kx}{\pi}\right)}{\Theta(0)} = \frac{(1-2q\cos 2x+q^2)(1-2q^3\cos 2x+q^6)(1-2q^5\cos 2x+q^{10})\cdots}{[(1-q)(1-q^3)(1-q^5)\ldots]^2},$$

unde formula (4.) §. 51. in hanc abit:

$$\Pi\left(\frac{2Kx}{\pi}, \frac{2KA}{\pi}\right) = \frac{2Kx}{\pi} Z\left(\frac{2KA}{\pi}\right) + \frac{1}{2}\log\frac{\Theta\left(\frac{2K}{\pi}(x-A)\right)}{\Theta\left(\frac{2K}{\pi}(x+A)\right)},$$

sive, rursus posito $\frac{2Kx}{\pi} = u$, $\frac{2KA}{\pi} = a$:

$$(3.)\qquad \Pi(u, a) = uZ(a) + \frac{1}{2}\log\frac{\Theta(u-a)}{\Theta(u+a)} = u\frac{\Theta'(a)}{\Theta(a)} + \frac{1}{2}\log\frac{\Theta(u-a)}{\Theta(u+a)},$$

siquidem ponitur: $\frac{d\Theta(u)}{du} = \Theta'(u)$. Quae est commoda expressio integralis elliptici Π per transcendentem novam Θ.

Facile constat, esse $\Theta(-u) = \Theta(u)$, unde, commutatis inter se a et u e (3.) prodit:

Fig. 8.3 Jacobi's introduction oftheta functions, p. 198 of his *Fundamenta Nova* paper [115]

$$|e^{\pi i n^2 \tau}| \leq e^{-\pi n^2 \mathrm{Im}\,\tau},$$

which shows that, for fixed τ, the coefficients of the Fourier series converge to zero very rapidly, and hence $\theta(z;\tau)$ is a holomorphic function of z for fixed τ. Moreover, it is clear from the definition that

$$\theta(z+m;\tau) = \theta(z;\tau),$$

so $\theta(z;\tau)$ is periodic with period 1 with respect to its first argument.

Now we compute the behavior of θ with respect to the τ direction in the complex plane, which is not along the real axis. Namely, we want to calculate $\theta(z+k\tau;\tau)$. We find

$$\begin{aligned}\theta(z+k\tau;\tau) &= \sum_{n=-\infty}^{\infty} e^{i\pi n^2\tau+2\pi i n(z+k\tau)},\\ &= e^{-i\pi\tau k^2} \sum_{n=-\infty}^{\infty} e^{i\pi\tau(n^2+2nk+k^2)+2\pi i n z},\end{aligned}$$

and letting $l = n+k$, we have

$$\begin{aligned}\theta(z+k\tau;\tau) &= e^{-i\pi\tau k^2-2\pi i k z} \sum_{l=-\infty}^{\infty} e^{i\pi\tau l^2+2\pi i l z},\\ &= e^{-i\pi\tau k^2-2\pi i k z}\theta(z;\tau).\end{aligned} \tag{8.17}$$

This is the quasiperiodicity alluded to above. Except for the factor $e^{-i\pi\tau k^2-2\pi k z}$, $\theta(z;\tau)$ seems to be periodic in the direction τ. How can we exploit this?

Let's consider a second such function

$$\theta_1(z;\tau) := \theta(z+\frac{1}{2}, m\tau).$$

This is also holomorphic and periodic with period 1 in z. What is the periodicity in the direction τ? Again we compute and find

$$\begin{aligned}\theta_1(z+k\tau;\tau) &= \sum_{n=-\infty}^{\infty} e^{\pi i n^2+2\pi i n(z+\frac{1}{2}+k\tau}),\\ &= e^{-\pi i k^2} \sum_{n=-\infty}^{n\infty} e^{\pi i\tau(n^2+2nk+k^2)+2\pi i n(z+\frac{1}{2})},\end{aligned}$$

which gives, letting $n = l-k$, as before,

$$\theta_1(z + k\tau; \tau) = e^{-\pi i \tau k^2 - 2\pi i k z} \dot{e}^{-\pi i k} \theta_1(z; \tau),$$
$$= (-1)^k e^{-\pi i \tau k^2 - 2\pi i k z} \theta_1(z; \tau).$$

Thus the multiplicative factor here is the same as in (8.17), except for the factor of $(-1)^k$. Therefore, if we form the quotient

$$e(z; \tau) := \frac{\theta(z, \tau)}{\theta_1(z; \tau)},$$

we see that

$$e(z + m + k2\tau; \tau) = e(z; \tau).$$

Thus we see that $e(z; \tau)$ is a doubly-periodic function with periods $(1, 2\tau)$, where $\operatorname{Im} \tau > 0$. By modifying suitably the choice of such theta functions, one can construct all of the Jacobi elliptic functions (again, see either [247], or [113] or any other standard reference on elliptic functions).

Chapter 9
Complex Analysis

9.1 Introduction

So far in this Part of the book we have seen the development of the complex plane, Abel's theorems concerning the generalization of elliptic integrals and the creation of the theory of elliptic functions of a complex variable. We now turn to a set of ideas which also started in the early decades of the nineteenth century and which would develop into a subject of great importance. This was the creation of what was often called function theory at the time, but which we call today complex analysis The fundamental concept is the study of special classes of complex-valued functions of a complex variable which are now known as *holomorphic* and *meromorphic functions*. We will see how these concepts arose out of the work of various mathematicians over a long period of time.

The fundamental innovators in the creation of function theory were Cauchy, Riemann and Karl Weierstrass (1815–1897), and we will discuss their respective contributions in some detail below. Today a course in complex analysis is considered an essential part of undergraduate education, and over the course of the twentieth century (and indeed towards the end of the nineteenth century) a number of texts evolved to explain this important subject, for instance, Hurwitz and Courant from 1922 [113], the classic text by Ahlfors [5], and there are many other fine more recent texts on the subject.

9.2 Cauchy in 1814

We start with the fundamental contributions of Cauchy, who contributed to the development of complex analysis throughout most of his very productive career. His collected works consist of two series, each with about 12 volumes and approximately 500 pages per volume; this includes his published papers as well as a number of monographs and textbooks. He worked on numerous fields of mathematics, includ-

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_9

ing differential geometry, number theory, mathematical physics, and a variety of other areas. His first paper in complex analysis [37] was presented to the Academie des Sciences in 1814 and finally published in 1827. Several footnotes added to the published version indicate some conceptual progress he had made in going from the real to the complex setting.

In this paper [37] Cauchy considers a function f of a *real* variable z, and shows that if z is considered to be a function of two other real variables x and y,[1] then

$$\frac{\partial}{\partial x}\left(f(z)\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial y}\left(f(z)\frac{\partial y}{\partial x}\right), \tag{9.1}$$

which is easy to verify. Namely,

$$\frac{\partial}{\partial x}\left(f(z)\frac{\partial z}{\partial y}\right) = f'(z)\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + f(z)\frac{\partial^2 z}{\partial x \partial y}, \tag{9.2}$$

$$\frac{\partial}{\partial y}\left(f(z)\frac{\partial y}{\partial x}\right) = f'(z)\frac{\partial z}{\partial y}\frac{\partial z}{\partial x} + f(z)\frac{\partial^2 z}{\partial y \partial x}, \tag{9.3}$$

and since

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x},$$

we see that (9.1) is satisfied.

Now we let z be a *particular* function of the two real variables x and y using complex numbers, namely let

$$z = x + iy,$$

where $i = \sqrt{-1}$, and let $f(z)$ take on complex values, so that

$$f = u + iv,$$

for real-valued functions u and v. Then (9.1) becomes, noting that $\frac{\partial z}{\partial x} = 1$, and $\frac{\partial z}{\partial y} = i$,

$$i\frac{\partial f}{\partial x}(z) = 1\frac{\partial f}{\partial y}(z), \tag{9.4}$$

that is,

$$i\frac{\partial}{\partial x}(u + iv) = 1\frac{\partial}{\partial y}(u + iv),$$

which becomes, upon setting real and imaginary parts equal to each other,

[1] We've used the now standard notation z, x, and y for these variables, where $z = x + iy$; Cauchy used y, z and x with $y = z + \sqrt{-1}x$ in this paper. In his later papers he used the now standard notation, and in his earlier papers he used $\sqrt{-1}$ instead of the symbol i for the imaginary unit, which he also used later.

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \tag{9.5}$$
$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \tag{9.6}$$

and this is the first appearance in Cauchy's work of the well-known *Cauchy–Riemann equations*.[2] Cauchy remarks at this point in his paper (p. 338):

> Ces deux équations renferment toute la théorie du passage du réel à l'imaginaire, et il ne nous reste plus qu'à indiquer la manière de s'en servir.[3]

Thus Cauchy indicates that he understood the significance of these equations, and his work over the next 30 years certainly bears this out. The implicit assumption that Cauchy makes here is that the derivative $f'(z)$ in (9.2) and (9.3) *exists*, as a generalization of $f'(z)$, when z was a real variable. This means that the limit

$$f'(z) = \lim_{\varepsilon \to 0} \frac{f(z+\varepsilon) - f(z)}{\varepsilon}$$

exists and is well defined, for small complex-valued ε. Functions which have this property became known in time as *holomorphic functions*, as we will see later in the book. We shall return to this point later when we look at Riemann's work.

The next step Cauchy takes is to integrate both sides of (9.4) over a rectangle in $\mathbf{R}^2$, which we take to be the rectangle R defined as the product of the two intervals $[0, X]$ on the x-axis and $[0, Y]$ on the y-axis, as pictured in Fig. 9.1. Thus we have

$$i\int_R \frac{\partial f}{\partial x} dxdy = \int_R \frac{\partial f}{\partial y} dxdy,$$

and, assuming that these partial derivatives are continuous on the rectangle R, we can evaluate these area integrals in terms of iterated integrals, obtaining,

$$i\int_0^Y \left(\int_0^X \frac{\partial f}{\partial x} dx\right) dy = \int_0^X \left(\int_0^Y \frac{\partial f}{\partial y} dy\right) dx,$$

which gives, using the fundamental theorem of calculus,

$$i\int_0^Y [f(X, y) - f(0, y)]dy = \int_0^X [f(x, Y) - f(x, 0]dy. \tag{9.7}$$

---

[2] These equations had appeared earlier in the work of d'Alembert in the context of fluid dynamics and in the work of Euler and Laplace for the evaluation of certain definite integrals. See Kline [125], pp. 626–628 for a discussion of this point. This is the beginning of his very interesting chapter on the history of function theory.

[3] "These two equations contain all the theory of passing from the real to the imaginary, and it only remains for us to indicate how this can be used".

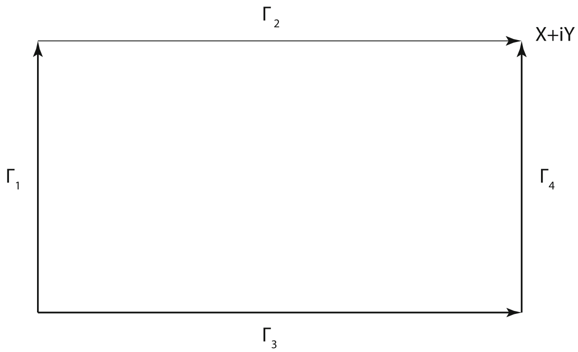


Fig. 9.1 A rectangle R in the complex plane

If we denote by $\Gamma_1 + \Gamma_2$ and $\Gamma_3 + \Gamma_4$ the two paths along the edges of the rectangle from 0 to $X + iY$ as indicated in Fig. 9.1, then we see that (9.7) becomes

$$\int_0^X f(x, 0)dx + \int_0^Y f(X, y)d(iy) = \int_0^Y f(0, y)d(iy) + \int_0^X f(x, Y)dx,$$

which is the same as

$$\int_{\Gamma_3+\Gamma_4} f(z)dz = \int_{\Gamma_1+\Gamma_2} f(z)dz. \tag{9.8}$$

This equation says that the path integrals of $f(z)$ in the complex plane along the two paths $\Gamma_1 + \Gamma_2$ and $\Gamma_3 + \Gamma_4$ have the same value. If we let Γ be the closed path $\Gamma_1 + \Gamma_2 - \Gamma_3 - \Gamma_4$ around the boundary of the rectangle R, then we have

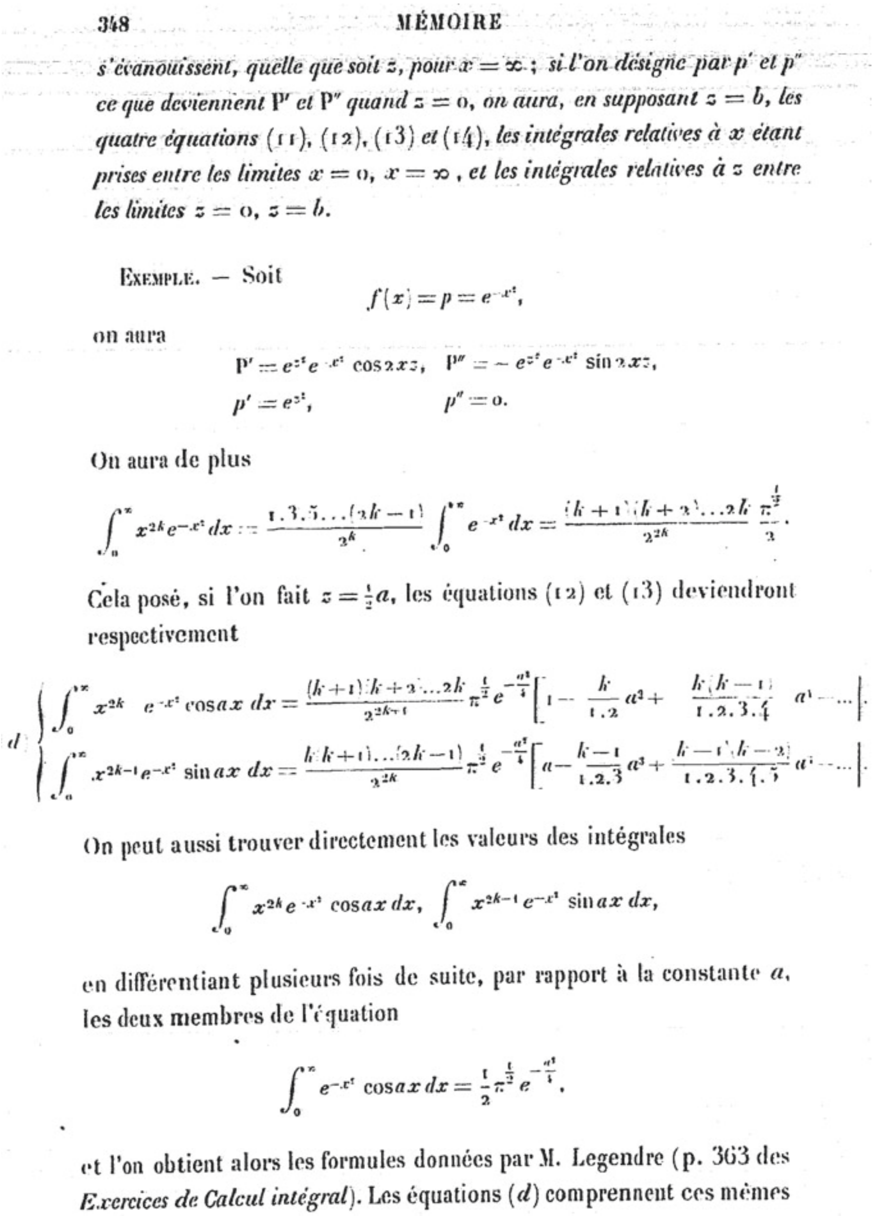
$$\int_\Gamma f(z)dz. \tag{9.9}$$

This is the famous *Cauchy integral theorem* in this important special case.

Cauchy expressed this theorem in terms of the real-valued functions u and v, and only later, when this 1814 paper was printed in 1827, he added footnotes indicating how the work could be simplified using the complex variable notation, as we have done here. He used these results to compute various examples of definite integrals, usually of the form $\int_{-\infty}^{\infty} f(z)dz$, where, for instance, the vertical integrals vanished asymptotically, and one was left with something like

$$\int_{-\infty}^{\infty} f(x + ib)dx = \int_{-\infty}^{\infty} f(x)dx,$$

the integration being shifted from the x-axis to a translate of the x-axis in the complex plane, which could often be simpler to compute. He also concerned himself with a variety of singular integrals, and proper values of integrals. In Fig. 9.2 we see a sampling of the evaluation of such integrals.


348 MÉMOIRE

s'évanouissent, quelle que soit z, pour $x=\infty$; si l'on désigne par p' et p'' ce que deviennent P′ *et* P″ *quand $z=0$, on aura, en supposant $z=b$, les quatre équations* (11), (12), (13) *et* (14), *les intégrales relatives à x étant prises entre les limites $x=0$, $x=\infty$, et les intégrales relatives à z entre les limites $z=0$, $z=b$.*

Exemple. — Soit

$$f(x)=p=e^{-x^2},$$

on aura

$$\mathrm{P}'=e^{z^2}e^{-x^2}\cos 2xz,\quad \mathrm{P}''=-e^{z^2}e^{-x^2}\sin 2xz,$$
$$p'=e^{z^2},\qquad p''=0.$$

On aura de plus

$$\int_0^\infty x^{2k}e^{-x^2}\,dx=\frac{1.3.5\ldots(2k-1)}{2^k}\int_0^\infty e^{-x^2}\,dx=\frac{(k+1)(k+2)\ldots 2k}{2^{2k}}\,\frac{\pi^{\frac12}}{2}.$$

Cela posé, si l'on fait $z=\frac12 a$, les équations (12) et (13) deviendront respectivement

$$(d)\left\{\begin{aligned}\int_0^\infty x^{2k}\,e^{-x^2}\cos ax\,dx&=\frac{(k+1)(k+2)\ldots 2k}{2^{2k+1}}\pi^{\frac12}e^{-\frac{a^2}{4}}\left[1-\frac{k}{1.2}a^2+\frac{k(k-1)}{1.2.3.4}a^4-\ldots\right],\\ \int_0^\infty x^{2k-1}e^{-x^2}\sin ax\,dx&=\frac{k(k+1)\ldots(2k-1)}{2^{2k}}\pi^{\frac12}e^{-\frac{a^2}{4}}\left[a-\frac{k-1}{1.2.3}a^3+\frac{(k-1)(k-2)}{1.2.3.4.5}a^5-\ldots\right].\end{aligned}\right.$$

On peut aussi trouver directement les valeurs des intégrales

$$\int_0^\infty x^{2k}e^{-x^2}\cos ax\,dx,\quad \int_0^\infty x^{2k-1}e^{-x^2}\sin ax\,dx,$$

en différentiant plusieurs fois de suite, par rapport à la constante a, les deux membres de l'équation

$$\int_0^\infty e^{-x^2}\cos ax\,dx=\frac12\pi^{\frac12}e^{-\frac{a^2}{4}},$$

et l'on obtient alors les formules données par M. Legendre (p. 363 des *Exercices de Calcul intégral*). Les équations (d) comprennent ces mêmes

Fig. 9.2 Page 348 of Cauchy's paper of 1814 [37] showing the evaluation of definite integrals using the first version of Cauchy's integral theorem

9.3 Cauchy's 1825 *Mémoire*

Cauchy wrote a number of papers and books on this topic over the decades following his seminal 1814 paper, and his writings became the basis for a significant amount of what became known as function theory. We will briefly discuss some of these in this section. Perhaps his most important paper is a fundamental paper written in 1825 entitled *Mémoire sur les intégrales, définies entre les limites imaginaires*[4] [36], which, due to its significance, was reprinted in 1874. Curiously, it doesn't appear in his collected works. We recall that at this time mathematicians still used the term "imaginary number" for what we now call complex numbers.

Cauchy considered in [36] complex-valued functions of a complex variable to have a well-defined derivative at each point where it had a finite value: and at any point where the function became infinite, he considered the function to be locally the reciprocal of a function with a zero of finite order. Today we call such a function a *meromorphic function*. For such a function $f(z)$, he defined the path integral

$$\int_{z_0}^{Z} f(z)dz$$

to be

$$\int_{t_0}^{T} f(\varphi(t))\varphi'(t)dt,$$

where $\varphi(t)$ is a smooth curve in the complex plane parametrized by a parameter t varying between t_0 and T, and where $z_0 = \varphi(t_0)$, and $Z = \varphi(T)$, provided the function is finite at all points of the path.[5]

Then Cauchy shows by a calculus of variations technique that, if one perturbs the path suitably, then the first variation of the perturbation vanishes, indicating to Cauchy that the path integral for the perturbed paths are the same as for the original unperturbed path, an infinitesimal version, so to speak, of the Cauchy integral theorem in this case. Then Cauchy considers the case where the path encounters a point where $f(z)$ becomes infinite, and he introduces the notion of the *proper value* for such a singular integral, being a limit of a specific perturbation of the integral.

For a function which is infinite at a point z_0 of finite order, which means that $f(z)$ can be represented near z_0 in the form

$$f(z) = \frac{a_{-m}}{(z-z_0)^m} + \cdots + \frac{a_{-1}}{(z-z_0)} + g(z), \tag{9.10}$$

where $g(z)$ is finite at z_0, Cauchy defines the *residue of* $f(z)$ at z_0 to be the coefficient a_{-1} in (9.10). Let us denote this residue by $\mathrm{Res}\, f(z_0)$ (a notation similar to that which Cauchy uses in his later papers).

[4] "Memoir concerning integrals defined between imaginary limits".

[5] Recall Gauss's letter to Bessel from 1811 quoted at the end of Sect. 6.3.

Using various perturbations of path integrals, he formulates the *Cauchy residue theorem* for a large rectangle R containing a finite number of singular points of $f(z)$, $z_0, z_1, \ldots, z_N$, to be

$$\int_{\partial R} f(z)dz = 2\pi i \sum_{k=0}^{N} \mathrm{Res}\, f(z_k).$$

He uses this for numerous examples of calculations of specific definite integrals, similar to what he had done in his 1814 paper.

This was a very important breakthrough for complex analysis, and he further developed this theory with numerous examples in his four-volume work *Exercices de Mathématiques* [38], published between 1826 and 1829. In a slightly later work, *Exercices d'Analyse et de Physique Mathématique* [42], he developed the *Cauchy integral formula* as

$$f(z_0) = \frac{1}{2\pi i} \int_{\partial\Delta} \frac{f(z)dz}{z - z_0},$$

where Δ is a small circular disc centered at z_0, and where $f(z)$ has finite values on the closure of the disc. In addition, he created similar integral formulas for all of the coefficients of the Laurent series of a function $f(z)$ with an isolated singularity at z_0. Such a series was first formulated for an isolated singularity by Pierre Alphonse Laurent (1813–1854) in 1843 [134], and was developed in full in a paper published posthumously in 1863 [135]. The series has the form

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n,$$

and the coefficients can be computed by

$$a_n = \frac{1}{2\pi i} \int_{\partial\Delta} \frac{f(z)dz}{(z - z_0)^{n+1}},$$

which became an alternative to the usual formula for the Taylor series coefficients $\frac{f^{(n)}(z_0)}{n!}$ in terms of derivatives, which don't make any sense when $f(z)$ is singular at z_0.

In 1846 [40] Cauchy formulated but did not prove what has become known as *Green's theorem*,[6] namely, for two continuously differentiable real-valued functions $P(x, y)$ and $Q(x, y)$ defined on the closure of a bounded domain D in $\mathbf{R}^2$,

$$\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy = \int_{\partial D} Pdx + Qdy. \tag{9.11}$$

[6]George Green (1793–1841) in [89] formulated and proved a version of Green's theorem in three dimensions, now known as the "divergence theorem" or often Gauss's theorem, all of which are special cases of the general Stokes's theorem in n-dimensions, see e.g., [217].

He showed that this formula will imply a proof of the Cauchy integral theorem, and, as it is so simple and instructive, we indicate its proof here. Let $f(z)$ be a function which satisfies the Cauchy–Riemann equations on the closure of a domain D, then we see that

$$\begin{aligned}\int_{\partial D} f(z)dz &= \int_{\partial D}(u+iv)(dx+idy)\\ &= \int_{\partial D}(udx - vdy) + i\int_{\partial D}(vdx+udy)\\ &= \int_D\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)dxdy + i\int_D\left(\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right)dxdy\\ &= 0 + i0,\end{aligned}$$

by Green's theorem (9.11) above. This shows the direct relationship between a function satisfying the Cauchy–Riemann equations and the Cauchy integral theorem. Riemann proved Green's theorem, and hence the Cauchy integral theorem, in his dissertation [199].

Finally, we mention that Cauchy wrote several papers dealing with the generalizations of his ideas to the case of multivalued functions, including applications to the study of elliptic integrals and functions (see, e.g., [39]) and in 1851 (the same year Riemann's dissertation appeared, as we discuss in the next section), Cauchy extended [41] his theory of multivalued function to be single-valued functions spread over a complex plane, a concept that Riemann introduced at the same time. Cauchy introduced the notion of branch points and branch cuts, which again Riemann would develop more fully later. Cauchy's paper is one of several in which he dealt with path integrals of multivalued functions.

9.4 Riemann's Dissertation from 1851

Riemann's doctoral dissertation from 1851 [199] was a foundational work in complex analysis, and we will survey a number of its most important results. Riemann starts his paper with the introduction of surfaces spread over domains in the complex plane (branched coverings), and elementary notions of connectedness for open sets of such surfaces. This was followed up in his 1857 paper on Abelian functions [202], and this became the theory of Riemann surfaces and eventually developed into the theory of complex manifolds in general, as we will discuss in some detail later in the book.

He starts, as did Cauchy, by considering the class of complex-valued functions of a complex variable on an open set that have a well-defined derivative at each point. He computes the derivative of a function $f(z) = u(z) + iv(z)$ as

$$\frac{d(u+iv)}{dx+iy} = \frac{\left(\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}\right)dx + i\left(\frac{\partial v}{\partial y}-i\frac{\partial u}{\partial y}\right)dy}{dx+idy}, \tag{9.12}$$

and he argues that this is a well-defined complex number if and only if

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \tag{9.13}$$

This gives

$$f'(z) = \frac{df}{dz} = \frac{d(u+iv)}{dx+iy} = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + i\frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

as the value of the derivative at the point z in terms of the real-valued partial derivatives, when this derivative exists. This is Riemann's version of the Cauchy–Riemann equations. Riemann says simply in his paper that a function f which satisfies the Cauchy–Riemann equations (9.13) or which has a well-defined derivative at each point is a "function of a complex variable." Today, we say that a function that satisfies these equations is a *holomorphic function*.

A holomorphic function $w = f(z)$ can be considered as a *mapping* from a region in the z-plane to a subset of the w-plane. We will often refer to a holomorphic function viewed geometrically like this as a *holomorphic mapping*. Riemann shows early in his dissertation that a holomorphic function f is conformal and orientation-preserving at each point of the domain where $f'(z) \neq 0$. *Conformal* means that if two smooth curves meet at a point z_0, then the angle between their tangents at that point is preserved under the mapping of the plane to the plane by the functions $f(z)$ near the point z_0. We say in this context that such an f is a *conformal mapping*.[7] The mapping preserves orientation if the direction from one tangent line to another is also preserved. Gauss had found necessary and sufficient conditions for a local mapping of $\mathbf{R}^2$ to be conformal in 1825 [79], which turned out, in the language of holomorphic mappings, *à la* Riemann in the paper we are discussing here, to mean a holomorphic (or anti-holomorphic) mapping.[8]

That a holomorphic mapping f at a point where $f'(z_0)$ is not zero preserves orientation is, as Riemann points out, easy to prove, since the Jacobian determinant of the mapping at that point is given by

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix},$$

which, by the Cauchy–Riemann equations (9.13), has the value

[7] Conformality was called by Gauss [79] and by Riemann [199] "in kleinsten Theilen ähnlich," or in English we might say "infinitesimally similar". We will use the now common term *conformal*.

[8] A conformal mapping in this same context could also be an *anti-holomorphic mapping*, i.e., a mapping $f(z)$, such that $\frac{\partial f}{\partial z}$ is zero, instead of the holomorphic mapping where $\frac{\partial f}{\partial \bar{z}}$ vanishes, using the contemporary notation that $\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)$ and $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)$.

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=\left(\frac{\partial v}{\partial y}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2>0,$$

unless $f'(z_0)=0$. Riemann proves later in this paper that a holomorphic mapping from an open set in $\mathbf{C}$ to $\mathbf{C}$ has an image which is an open set. Any mapping that maps open sets to open sets is called an *open mapping*, and this theorem of Riemann is often referred to in function theory as the *open mapping theorem*. In this context, one speaks of a holomorphic function defined on a domain where $f'(z)\neq 0$ as providing a *conformal mapping* from one domain to another.

Riemann stated and gave a proof of Green's theorem (9.11) for a domain in $\mathbf{R}^2$ and showed, as Cauchy had done, that the Cauchy integral theorem is valid and that an integral of a holomorphic function in a simply-connected domain is path independent. He also used a version of Green's theorem to develop some fundamental properties of harmonic functions, which play an important role in his later development in this same paper of what we now call the Riemann mapping theorem, which we will discuss shortly. *Harmonic functions* are solutions of the equation

$$\Delta u=0,$$

where

$$\Delta:=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$$

is the *Laplacian differential operator* in two dimensions. This terminology, which is now standard, was not used by Riemann and was introduced by William Thomson (Lord Kelvin, 1824–1907) in the mid-nineteenth century (see Kline [125], p. 685).

The main tool Riemann uses is a variation of what are now known as *Green's formulas*, which he proved as a consequence of Green's theorem. Namely, let X and Y be two continuously differentiable functions on a domain $T\subset\mathbf{R}^2$, which satisfies on T the equation

$$\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}=0.$$

It then follows from Green's theorem (9.11) that the boundary integral

$$\int_{\partial T} Y dx - X dy = 0.$$

Riemann introduces normal and tangential coordinates along a neighborhood of the boundary curve ∂T in the following manner. He lets s be the arc length from a fixed point on the boundary to a variable point P on the boundary and lets p be the distance from that point P along an inner directed normal to a point $z=x+iy$ on the interior of T. Then, by letting ξ be the angle the normal at P makes with the x-axis, and η the angle the normal makes with the y-axis, the version of Green's theorem that Riemann uses becomes

$$\int_T \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dT = -\int_{\partial T} (X\cos\xi + Y\cos\eta) ds.$$

Riemann then computes the change of variables formulas

$$\frac{\partial x}{\partial p} = \cos\xi, \ \frac{\partial y}{\partial p} = \cos\eta,$$
$$\frac{\partial x}{\partial s} = \cos\eta, \ \frac{\partial y}{\partial s} = -\cos\xi,$$

using a positive orientation of the coordinates (s, p) with respect to the standard orientation of the (x, y) plane. Green's theorem then becomes

$$\int_T \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dt = -\int_{\partial T} \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds \tag{9.14}$$
$$= \int_{\partial T} \left(X\frac{\partial y}{\partial s} - Y\frac{\partial x}{\partial s}\right) ds. \tag{9.15}$$

If $\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0$ in T, then these boundary integrals are zero. If $\frac{\partial X}{\partial x}$ or $\frac{\partial Y}{\partial y}$ have singular points at some finite set of points in T, say $z_1, \ldots, z_N$, and

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0 \text{ on } T - \{z_1, \ldots, z_N\},$$

then

$$\int_{\partial T} \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds = -\sum_{j=1}^{N} \int_{\partial\Delta_j} \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds, \tag{9.16}$$

where Δ_j are small nonintersecting discs centered at z_j, such that the closure of each disc is contained in the open set T. This is quite parallel to Cauchy's residue theorem in this context, where Cauchy gave meaning to the localized integrals in terms of residues of the holomorphic functions at such singular points.

Riemann then considers two functions u and $\tilde{u}$ which are C^2 on T and which have a continuous extensions along with their first derivatives to ∂T.[9] Now suppose that both u and $\tilde{u}$ are harmonic on T, then setting

$$X = u\frac{\partial \tilde{u}}{\partial x} - \tilde{u}\frac{\partial u}{\partial x},$$
$$Y = u\frac{\partial \tilde{u}}{\partial y} - \tilde{u}\frac{\partial u}{\partial y},$$

[9] Here we use the notation C^k to denote functions which have continuous derivatives of order k, and C^∞ will mean continuously differentiable of all orders.

it follows that

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = u\Delta\tilde{u} - \tilde{u}\Delta u,$$

and using Green's theorem (9.15) above gives

$$\int_{\partial T} \left(u\frac{\partial \tilde{u}}{\partial p} - \tilde{u}\frac{\partial u}{\partial p} \right) ds = 0. \tag{9.17}$$

Riemann considers two particular cases for $\tilde{u}$, which lead to interesting and useful results. The first case is to simply take $\tilde{u} \equiv 1$, and it follows that

$$\int_{\partial T} \frac{\partial u}{\partial p} ds = 0. \tag{9.18}$$

In the second case, for any particular point $z_0 \in T$, Riemann chooses polar coordinates $z - z_0 = r^{i\varphi}$ and sets

$$\tilde{u}(z) := \log r = \log |z - z_0|.$$

Then using the extension of Green's theorem to the case of singularities (9.16), it follows that

$$\int_{\partial T} \left(u\frac{\partial \log r}{\partial p} - \log r\frac{\partial u}{\partial p} \right) ds = \int_{\Delta_\varepsilon} \left(u\frac{\partial \log r}{\partial p} - \log r\frac{\partial u}{\partial p} \right) ds,$$

where Δ_ε is a small disc of radius ε, centered at z_0. Note that on the boundary of Δ_ε,

$$\frac{\partial \log r}{\partial p} = -\frac{\partial \log r}{\partial r} = -\frac{1}{r},$$

and thus

$$\begin{aligned} \int_{\partial T} \left(\log r\frac{\partial u}{\partial p} - u\frac{\partial \log r}{\partial p} \right) ds &= \int_0^{2\pi} u(\varepsilon e^{i\varphi}) df + \log r \int_{\Delta_\varepsilon} \frac{\partial u}{\partial p} ds \\ &= \int_0^{2\pi} u(\varepsilon e^{i\varphi}) d\varphi, \end{aligned} \tag{9.19}$$

since the second term on the right-hand side vanishes by (9.18) (where we let T be Δ_ε).

Now letting $\varepsilon \to 0$ in (9.19), Riemann obtains the following formula

$$u(z_0) = \frac{1}{2\pi} \int_{\partial T} \left(\log |z - z_0| \frac{\partial u}{\partial p} - u\frac{\partial \log |z - z_0|}{\partial p} \right) ds, \tag{9.20}$$

which represents the value of the harmonic function $u(z)$ at an interior point z_0 of T in terms of the boundary integral on ∂T.[10] This result and the use of the potential function $\log r$ is similar to the work of Green from 1828 [89], in which the three-dimensional potential function $1/r$ is used in $\mathbf{R}^3$.

Suppose we restrict the harmonic function $u(z)$ above to a disc Δ_ε of radius ε centered at z_0 whose closure is contained in the domain T. Then the formula (9.20) becomes

$$\begin{aligned} u(z_0) &= \frac{1}{2\pi}\log r \int_{\partial\Delta_\varepsilon} \frac{\partial u}{\partial p} ds + \frac{1}{2\pi}\int_0^{2\pi} u(\varepsilon e^{i\varphi})d\varphi \\ &= \frac{1}{2\pi}\int_0^{2\pi} u(\varepsilon e^{i\varphi})d\varphi, \end{aligned} \tag{9.21}$$

since the first term on the right-hand side vanishes, by (9.18). Thus $u(z_0)$ is the *mean value* of its values on the boundary of the disc Δ_ε. Equation (9.21) is known as the *mean-value theorem* for harmonic functions (and this is true in all dimensions). It has numerous consequences, as Riemann shows in his paper, and we list some of them here. We refer the reader to, e.g., Ahlfors [5], for proofs of these results, as well as to this paper of Riemann.

Let u be a harmonic function in a domain T.

- *Removable singularity theorem*: If u is potentially singular or undefined at a point z_0, and if ρ is the distance from a neighboring point z to z_0, and if

$$\rho\frac{\partial u}{\partial x} \text{ and } \rho\frac{\partial u}{\partial y} \to 0, \text{ as } \rho \to 0,$$

  then u can be continued as a continuous function to z_0 and u is harmonic in a neighborhood of z_0.
- *Smoothness*: The harmonic function u is C^∞ in all of T (this follows from (9.20) by differentiation under the integral sign).
- *Maximum principle*: The harmonic function u cannot have a local maximum or minimum at any interior point of T unless u is a constant function near that point.
- *Identity theorem*: The harmonic function u in T is determined by the values of u and $\frac{\partial u}{\partial p}$ on any arc segment in T, and moreover, if on a segment of an arc in T $u \equiv 0$ and $\frac{\partial u}{\partial p} \equiv 0$, then $u \equiv 0$ in T.

Riemann remarks that many of these properties of harmonic functions carry over to holomorphic functions in a natural manner. For instance,

- *Riemann removable singularity theorem*: If f is holomorphic on a punctured disc centered at z_0, $\Delta - \{z_0\}$, and if $(z - z_0)f(z) \to 0$, as $z \to z_0$, then f extends as a holomorphic function to Δ.

[10] In more contemporary literature, the normal derivative of data along ∂T is usually denoted by $\frac{\partial}{\partial n}$.

- *Smoothness*: A holomorphic function in a domain is infinitely differentiable.
- *Maximum principle*: The modulus of a holomorphic function f in a domain can take on a local maximum in the interior of the domain only if f is constant in the domain.[11]

In Sect. 15 of his paper ([199], p. 28), Riemann formulates and proves the *open mapping theorem* for holomorphic functions, mentioned earlier. Namely, let T be a domain in $\mathbf{C}$, and U be any open subset of T, and let $f : U \to \mathbf{C}$ be a nonconstant holomorphic function defined on U, then $f(U)$, the image of f under the mapping f, is an open set in $\mathbf{C}$.[12] This is a strong property of holomorphic functions, and it is also proved in any standard complex analysis text (again, e.g., Ahlfors [5]). We note that, in contrast, this is not true for real-valued smooth (or real-analytic) functions. As a simple example, the mapping $f(x) : \mathbf{R} \to \mathbf{R}$ given by $f(x) = \sin x$ has as the image of the open set $\mathbf{R}$ the closed set $[-1, 1]$.

In the last section of this very innovative paper Riemann comes to what is undoubtedly its deepest result, the *Riemann mapping theorem*. We note first that in the beginning of the paper, Riemann formulated the concept of what is now known as a Riemann surface spread over a region of the complex plane, a branched covering of an open set in $\mathbf{C}$, and most of his results in this paper are formulated in this more general context. We have chosen to formulate his results for domains in the complex plane, as that is simpler.

In Sects. 9.5 and 9.6 of Riemann's dissertation he formulates and proves a number of elementary results concerning the Riemann surfaces he introduces. These include the important notions of connectedness of domains in the plane, for instance, simply-connectedness, which we have already had occasion to use, and more general connectedness of order n; we will come back to these concepts in Chap. 10. Riemann defines a domain in the plane to be simply-connected if it has the property that any curve in the plane joining any two boundary points will split the domain into two parts that are not connected to each other (he always means path-wise connected).

Now we can give, in his own words, Riemann's formulation of the Riemann mapping theorem:

> Zwei gegebene einfach zusammenhängende ebene Flächen können stets so auf einander bezogen werden, dass jedem Punkte der einen Ein mit ihm stetig fortrückender Punkt der andern entspricht und ihre entsprechenden kleinsten Theile ähnlich sind; und zwar kann zu Einem innern Punkte und zu Einem Begrenzungspunkte der entsprechende beliebig gegeben werden; dadurch aber ist für alle Punkte die Beziehung bestimmt.[13]

---

[11] We recall that a domain is a connected open set.

[12] At the time of Riemann, the notion of open set was not yet a mathematical concept. He formulated his theorem in terms of neighborhoods of points. We are giving the modern formulation of this important result.

[13] "Two given simply-connected domain plane surfaces can always be related to one another, so that each point of one corresponds in a continuous manner to each point of the other and such that the corresponding smallest parts are infinitesimally similar [conformal]; and indeed such that a given inner point and a given boundary point correspond to a specified interior and boundary point; with this last condition, the relationship is determined for all points".

Riemann immediately reduces this formulation to the simpler statement that any simply-connected domain can be conformally mapped onto the unit disc, with one interior point mapping to the center of the disc, and a boundary point mapping to a specified boundary point of the unit disc (e.g., $z = 1$).

The proof that Riemann gives is based on what he terms the *Dirichlet principle*, a name he gave to this principle in his follow-up paper on Abelian functions (which used the same principle for additional results) in 1857 [202]. We will discuss this principle here, and then indicate how it became a problem for Riemann and his proof of the Riemann mapping theorem, and how it was finally resolved some 50 years later by David Hilbert (1862–1943). We will return to the Riemann mapping theorem in Sect. 11.2.

We first formulate an important special case of the Dirichlet principle. Let T be a bounded domain in $\mathbf{R}^2$ with a smooth boundary, and let f be a continuous function on ∂T. Consider the family $\mathcal{F}$ of real-valued functions that are continuously differentiable in T and continuous on $\overline{T}$, such that $u|_{\partial T} = f$. This is an infinite-dimensional family of functions. Consider the Dirichlet integral

$$D(u) := \int_T \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dxdy, \text{ for } u \in \mathcal{F}. \tag{9.22}$$

From the definition it is clear that

$$0 \leq D(u) < \infty.$$

It follows that

$$m := \inf_{u \in \mathcal{F}} D(u)$$

is well defined and $m \geq 0$. The Dirichlet principle asserts that there exists a unique $u \in \mathcal{F}$ such that

$$D(u) = m,$$

and moreover, that u is harmonic (hence C^∞ on T), and that u is continuous on T and that $u|_{\partial T} = f$. This last statement says that u is a solution to the *Dirichlet problem*: find a harmonic function with given boundary values on ∂T.

Today, there are many different proofs of the solution to the Dirichlet problem, but in the mid-nineteenth century, these did not exist, and mathematicians and physicists used this principle to solve many difficult problems. Specifically, Green and Thomson formulated and used this principle: Green in 1835 [90] in the context of gravitational attraction, and Thomson in 1848 [221] as a general mathematical principle (called Thomson's principle in England for some time, see Kline [125], p. 685). In lectures in Göttingen in 1856 concerning inverse square forces, which very likely Riemann attended, Peter Gustav Lejeune Dirichlet (1805–1859) used this minimization principle for the existence and uniqueness of specific harmonic functions, i.e., to solve the Dirichlet problem. As noted above, Riemann used this principle in both his 1851 dissertation [199] and his seminal paper on Abelian functions [202], wherein he denoted this principle as the *Dirichlet principle*, and it has been called that ever since (in spite of the earlier work of Green and Thomson).

Then in 1870 Weierstrass gave a lecture at the Berlin Academy of Sciences (which was published in his collected works in 1895 [231]) entitled "Über das sogennante Dirichlet'sche Princip".[14] As Weierstrass notes in this paper, he had a handwritten copy of lecture notes from Dirichlet's lectures, which he had received from Dedekind. Weierstrass quotes several pages from these notes, and then points out that Dirichlet's proof of the Dirichlet principle (namely the proof of the existence of a minimizing function) was incomplete and not rigorous, and then produces an example of a similar type of calculus of variations problem in one dimension in which he showed that the minimum value of a specific energy integral was *not* assumed by any function in the class of functions being considered.

In the same year as Weierstrass's lecture, Hermann Amandus Schwarz (1843–1921), a student of Weierstrass, published a paper [212] which included a rigorous proof of the Riemann mapping theorem in the special case of simply-connected domains in the complex plane (as stated above). This proof did not apply to the more general case of a simply-connected domain in a Riemann surface, which Riemann had formulated but also did not prove. This was proved later in the uniformization theorem of Koebe, a student of Schwarz, and we will discuss this in Sect. 11.2. Schwarz followed the outline of Riemann's proof, replacing the Dirichlet-principle argument with a convergent iterative argument involving a sequence of harmonic functions defined on specified open subsets of the given domain. Finally, in 1904 Hilbert [102] gave a proof of this disputed Dirichlet principle in the special context of one of Riemann's existence theorems in his paper on Abelian functions [202], thus justifying Riemann's original argument.

9.5 The Lectures of Weierstrass

Now we turn more specifically to the third of our major contributors to function theory, whom we have mentioned several times already, namely, Karl Weierstrass. His work stretched over a number of decades in the latter half of the nineteenth century and set standards of rigor and methodology that became a major force in how function theory (and more generally limiting processes and analysis in general) was perceived and used in the twentieth century. The first papers of Weierstrass in the 1840s concerned themselves with specific problems in the theory of elliptic functions following up on the pioneering work of Abel and Jacobi. During these early years Weierstrass wrote several fundamental papers which were published only later.

The main influence of Weierstrass in function theory came via his lecture courses in Berlin in the 1860s, which were published at the time, and which are all included

[14] "On the so-called Dirichlet principle".

in his collected works. The first three volumes of his collected works [229, 232, 234] contain primarily his original papers over his professional lifetime (including those papers mentioned earlier that weren't published when they were written), and the following three volumes contain reprints of his lecture notes from his lectures on Abelian functions [233], his lectures on elliptic functions [235], and his lectures on applications of elliptic functions [236].

Let us mention here some of his principal results which have become part of the standard repertoire in function theory. Weierstrass defined a holomorphic function[15] to be a locally defined function of a complex variable defined near a point $z_0 \in \mathbf{C}$ of the form

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

which converges in some disc of radius R centered at z_0. If two such functions f_1 and f_2 are defined in discs Δ_1 and Δ_2 centered at two points z_1 and z_2, and if these two discs intersect, and if

$$f_1|_{\Delta_1 \cap \Delta_2} = f_2|_{\Delta_1 \cap \Delta_2},$$

then f_2 is said to be the *analytic continuation* of f_1, and vice versa. Moreover,

$$f := \begin{cases} f_1, z \in \Delta_1, \\ f_2, z \in \Delta_2, \end{cases}$$

is a holomorphic function in $\Delta_1 \cup \Delta_2$. More generally, a holomorphic function in a domain D is a function which admits such a power-series expansion near each point of D.[16] This definition of holomorphic as formulated by Weierstrass is equivalent to the definition used by Cauchy and Riemann as solutions of the Cauchy–Riemann equations.

Weierstrass brought much needed rigor to mathematical analysis, not only in function theory. For instance, he showed that a sequence of continuous functions on a domain $D \subset \mathbf{R}^n$,

$$f_1(x), f_2(x), \ldots, f_k(x), \ldots$$

which converges uniformly on compact subsets of the domain D has a limit that is continuous on D. In the holomorphic setting he showed that if

$$f_1(z), f_2(z), \ldots, f_k(z), \ldots$$

[15]Weierstrass used the term analytic function instead of holomorphic function, which was used regularly in the twentieth century as well. Later these became known as complex-analytic functions to contrast with the similarly defined real-analytic functions defined as a locally convergent real power series of real variables. Today holomorphic refers to the complex-analytic case, and one still uses the term real-analytic for the case of analytic functions of a real variable or variables.

[16]This definition extends naturally to Riemann surfaces spread over domains in $\mathbf{C}$ as well.

is a sequence of holomorphic functions in a domain $D \subset \mathbf{C}$ which converges uniformly on compact subsets of D, then the limiting function is holomorphic in D.

Although our emphasis in this section has been on holomorphic functions of one variable, Weierstrass (and others) considered holomorphic functions of several variables as well. For instance, the definition of a holomorphic function $f(z_1, \ldots, z_n)$ of several complex variables $(z_1, \ldots, z_n)$ can either be that the function has a convergent power-series expansion of the form

$$f(z_1, \ldots, z_n) = \sum_{i_1, \ldots, i_n} a_{i_1 \ldots i_n} z_1^{i_1} \ldots z_n^{i_n},$$

near each point of a domain, where $(z_1, \ldots, z_n)$ are coordinates in $\mathbf{C}^n$, or, alternatively, one can require that the Cauchy–Riemann equations in $\mathbf{C}^n$ are satisfied, i.e.,

$$\frac{\partial f}{\partial \overline{z}_j}(z_1, \ldots, z_n) = 0,$$

where

$$\frac{\partial}{\partial \overline{z}_j} := \frac{1}{2}\left(\frac{\partial}{\partial x_j} + i\frac{\partial}{\partial y_j}\right), \; j = 1, \ldots, n.$$

Riemann, Weierstrass and others were very interested in *Abelian functions*, which were functions of several complex variables in $\mathbf{C}^n$ which generalized elliptic functions of one complex variable and which we will encounter in the next chapter.

A given holomorphic function with a power series expansion at a given point has a *radius of convergence* for the series, and there are various criteria and descriptions of how to compute this due to Cauchy and others. In particular, if the radius of convergence of a function f at a given point z_0 is a finite number $R < \infty$, then there is a least one boundary point z_1 on the disc of radius R centered at z_0 where f is singular and doesn't admit any analytic continuation to a larger open set containing that point. For instance, the function

$$f(z) = \frac{1}{z-1}$$

has an expansion in the unit disc (the geometric series), and this does not converge at the point $z = 1$, and the function cannot be analytically continued beyond (or through) that point. However, this function does have an analytic continuation through all other points on the boundary of the unit disc, as is very easy to see, since $f(z)$ is holomorphic on $\mathbf{C} - \{1\}$. Weierstrass was the first to describe a function holomorphic on the unit disc which is singular at every boundary point of the unit disc.

Here's a simple example of such a function given by a lacunary series,

$$f(z) = \sum_{n=0}^{\infty} z^{2^n},$$

and it easy to see that the series is divergent on the dense set of roots of unity of all orders on the boundary of the unit disc. Namely, for $z = 1$,

$$f(1) = 1 + 1 + 1 + \cdots + 1 \cdots ,$$

which diverges, and for the two roots $\varepsilon_1, \varepsilon_2$ of $z^2 = 1$,

$$\begin{aligned} f(\varepsilon_j) &= \varepsilon_j^1 + \varepsilon_j^2 + \varepsilon_j^4 + \cdots + \varepsilon_j^{2^n} + \cdots \\ &= \varepsilon_j + 1 + 1 + \cdots + 1 + \cdots , \end{aligned}$$

and for the four roots $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ of $z^{2^2} = 1$ we see that

$$\begin{aligned} f(\varepsilon_j) &= \varepsilon_j^1 + \varepsilon_j^2 + \varepsilon_j^4 + \varepsilon_j^8 + 1 + \cdots + \varepsilon_j^{2^n} + \ldots \\ &= \varepsilon_j + e_j^2 + 1 + 1 + \cdots + 1 + \cdots . \end{aligned}$$

Using an induction argument, we see the series is divergent on this dense set, and hence at each of the boundary points (since no point where a function is holomorphic can be a limit point of singular points).

By the Riemann mapping theorem, it follows that any simply-connected domain has a function singular at every boundary point, and by results of Weierstrass and Gösta Mittag-Leffler (1846–1927), which we discuss in the next section, this is true for all domains in **C**. However, a striking result for holomorphic functions of several complex variables, due to Friedrich Moritz Hartogs (1874–1943) [96] at the beginning of the twentieth century, shows that this is *not* true for functions of two or more variables, and this led to a major new direction of research in complex analysis for functions of several complex variables. In Sect. 15.2 in Part IV of the book, we give an overview of this subject, which became an important area of research in the twentieth century.

As we mentioned earlier, Weierstrass started his mathematical career by studying elliptic functions as were formulated by Jacobi, and we met these functions earlier as $\operatorname{sn} z$, $\operatorname{cn} z$, and $\operatorname{dn} z$. In his later work on elliptic functions Weierstrass (see his lectures on elliptic functions [235]) introduced a new way to describe elliptic functions, which has now become one of the two standard approaches to these functions (the other being that of Jacobi).

Let us now briefly summarize how Weierstrass approached elliptic functions, which is quite different from the original constructions of Abel and Jacobi.

If ω_1, ω_2 are two complex numbers with $Im\omega_1/\omega_2 \neq 0$, then Weierstrass defines the *Weierstrass $\wp$-function* as

$$\wp(z) := \frac{1}{z^2} + \sum_{m^2+n^2>0} \frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2}. \tag{9.23}$$

The series converges due to the extra term

$$\frac{1}{(m\omega_1 + n\omega_2)^2},$$

which is added to insure convergence. It is easy to see that this function has the following properties (assuming the convergence, which requires some work):

- $\wp(z + m\omega_1 + n\omega_2) = \wp(z)$, i.e., $\wp(z)$ is doubly-periodic in $\mathbf{C}$ with periods ω_1 and ω_2.
- $\wp(z)$ is a meromorphic function on $\mathbf{C}$ with a single double pole in each period parallelogram.[17]
- The derivative $\wp'(z)$ of the Weierstrass $\wp$-function is also a doubly-periodic function with periods ω_1 and ω_2.

The two functions $\wp(z)$ and $\wp'(z)$ play an analogous role in elliptic-function theory to the original doubly-periodic functions $\mathrm{sn}\,(z)$ and $\mathrm{cn}\,(z)$ of Jacobi. A classical reference for this infinite-series approach to elliptic functions is the text by Karl Boehm [23]. See Hurwitz and Courant [113] or Whittaker and Watson [247] for a complete discussion of classical elliptic-function theory, including Weierstrass's contributions.

The final result of Weierstrass we want to mention in this section is often referred to as the *Weierstrass factorization theorem* or the *Weierstrass product theorem*, which was published in 1876 [228]. This result is an important generalization of the fundamental theorem of algebra, which asserts that any polynomial can be expressed in terms of factors

$$p(z) = c(z - a_1)^{m_1} \cdots (z - a_k)^{m_l}, \tag{9.24}$$

where $a_1, \ldots, a_k$ are the roots of the polynomial with multiplicities $m_1, \ldots, m_k$, and c is a constant. Let $f(z)$ be a holomorphic function in the complex plane $\mathbf{C}$ (such an f is called an *entire function*) with zeros at a possibly infinite set of points $a_1, \ldots, a_k, \ldots$, then the Weierstrass factorization theorem asserts that f can be represented as an infinite product similar to the finite product in (9.24),

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right) e^{\left(\frac{z}{a_n}\right) + \frac{1}{2}\left(\frac{z}{a_n}\right) + \cdots + \frac{1}{m_n}\left(\frac{z}{a_n}\right)^{m_n}}, \tag{9.25}$$

where m, m_n are integers, and g is an entire function (see Ahlfors [5] or any standard complex-analysis text for a discussion and proof of this theorem). The result is often formulated in the following manner. Let $a_1, \ldots, a_k, \ldots$ be any sequence of points in the plane such that $\lim a_k = \infty$, then there exists an entire function with zeros at precisely these points (namely use the formula (9.25)). Weierstrass introduced the exponential factors in the infinite product to insure its convergence.

[17] The period parallelograms defined by the periods ω_1, ω_2 are the translates in the complex plane by integers of the form $m + in$ of the fundamental period parallelogram with the four vertices $0, \omega_1, \omega_2$, and $\omega_1 + \omega_2$ (see, e.g., [113]).

9.6 The Mittag-Leffler Theorem

A consequence of the Weierstrass factorization theorem, as observed by Weierstrass, is that any meromorphic function in **C** can be expressed as the quotient of two entire functions. These results of Weierstrass were generalized by Mittag-Leffler in 1884 [157] from the case of the complex plane to an arbitrary domain in the following sense. Let D be a domain in the complex plane and let $a_1, \ldots, a_k, \ldots$ be an infinite sequence of points in D with no accumulation points in D, i.e., each point a_k is isolated in D, then there exists a function $f(z)$ holomorphic in D which has zeros precisely at the points a_k. If we consider a domain D with a boundary ∂D, and let a_k be a set of points in D again with no accumulation points in D and which is dense on the boundary of D, then the Weierstrass function with these zeros has no analytic continuation beyond any boundary point (a result we alluded to above in the context of the Riemann mapping theorem and lacunary series for simply-connected domains).

We now turn to the final topic of this section on holomorphic functions, the Mittag-Leffler theorem. As we just saw, the Weierstrass factorization theorem showed that for given prescribed zeros, one can find a holomorphic function with those zeros. By taking reciprocals, one could find a meromorphic function with poles of a certain order at those same points by prescribing the multiplicity of the zeros or the order of the poles. A variation on this question was raised and solved in two of Mittag-Leffler's earlier papers in 1877 [155, 156], which we formulate here.

A meromorphic function $f(z)$ near a pole at a point z_0 of order m has a Laurent expansion at z_0 of the form

$$\frac{a_{-m}}{(z-z_0)^m} + \cdots + \frac{a_{-1}}{(z-z_0)} + \sum_{n=0}^{\infty} a_n (z-z_0)^n,$$

where the infinite series converges in the neighborhood of z_0, and the finite number of terms of powers of $\frac{1}{(z-z_0)}$ all converge to ∞ at z_0 and represent what is called the *principal part* of the meromorphic function at z_0. We saw this earlier, when we identified a_{-1} as the residue of $f(z)$ at z_0. More generally, if the meromorphic function $f(z)$ has poles at z_k in a domain $D \subset \mathbf{D}$, then there are functions $p_k(z)$ and $g_k(z)$ defined near each point z_k where $p_k(z)$ has the form

$$p_k(z) = \frac{a^k_{-m_k}}{(z-z_k)^{m_k}} + \cdots + \frac{a^k_{-1}}{(z-z_k)}, \tag{9.26}$$

where $p_k(z)$ is the principal part of the meromorphic function $f(z)$ at z_k, and $f(z) - p_k(z) := g_k(z)$ is holomorphic near z_k.

The question Mittag-Leffler raised was the following: given a discrete set of points z_k in a domain D and, for each point z_k, given a polynomial of the form (9.26), does there exist a meromorphic function $f(z)$ in D and locally defined holomorphic functions $g_k(z)$ defined near each point z_k such that near z_k one has

$$f(z) - p_k(z) = g_k(z)?$$

Mittag-Leffler showed that this was true, and the result is known as the *Mittag-Leffler theorem.*

Let us sketch Mittag-Leffler's proof of this in the simple case where z_k is a discrete sequence in the complex plane and the principal parts which are given are simple poles of the form

$$p_k(z) = \frac{a_k}{(z - z_k)}.$$

Assume the points z_k are ordered such that

$$|z_1| \leq |z_2| \leq \cdots \leq |z_k| \cdots ,$$

and let Δ_k be concentric discs centered at the origin whose radii increase in such a fashion that

$$z_j \notin \overline{\Delta_k}, \text{ for } j \geq k,$$

where we use the standard notation $\overline{K}$ to mean the closure of a set K in $\mathbf{R}^n$. Thus each $p_k(z)$ is holomorphic on a neighborhood of Δ_k. This implies that $p_k(z)$ can be expanded in a power series centered at the origin which converges on a neighborhood of $\overline{\Delta}_k$, and hence $p_k(z)$ can be approximated on $\overline{\Delta_k}$ by a polynomial $h_k(z)$ such that

$$|p_k(z) - h_k(z)| < \frac{1}{2^k}, \text{ for } z \in \overline{\Delta_k}.$$

It follows that the series

$$\begin{aligned} f(z) &= \sum_{k=1}^{\infty}(p_k(z) - h_k(z)) \\ &= \sum_{k=1}^{\mathbf{N}}(p_k(z) - h_k(z)) + \sum_{k=N+1}^{\infty}(p_k(z) - h_k(z)) \end{aligned}$$

converges uniformly on Δ_n, $N = 1, 2, \ldots$, and $f(z)$ is a well-defined meromorphic function on $\mathbf{C}$ with principal parts $p_k(z)$ near each pole z_k.

Mittag-Leffler first proved this result in his two papers (written in Swedish) [155, 156] for the case of $D = \mathbf{C}$, and in his longer *Acta Mathematica* paper[18] [157] for arbitrary domains. Carl Runge (1856–1927) gave a new and simpler proof of Mittag-Leffler's theorem in 1885 [205], which involved a new approximation theorem, now called Runge's theorem or the Runge approximation theorem, which showed how one can approximate holomorphic functions on a multiply-connected domain D uniformly on compact subsets of the domain by rational functions with poles in the

[18]*Acta Mathematica* was founded by Mittag-Leffler in 1882, and initially, for a number of years, all the papers were in French, the principal international language of its time.

bounded components of the complement of D. If D is simply-connected, then the function can be approximated by a polynomial in the same manner. These results of Weierstrass, Mittag-Leffler, and Runge all have generalizations to holomorphic functions of more than one complex variable, and they play an important role in the further development of this field of research as it developed in the twentieth century. In particular, the question of when a theorem of Mittag-Leffler type might be true for domains in $\mathbf{C}^n$, for $n > 1$, became an important research topic.

The work of Cauchy, Riemann, and Weierstrass generated immense interest in the mathematicians of the second half of the nineteenth century. Initially there were more or less three schools of thought following these three innovators (the integral theorems of Cauchy, the differential equations of Riemann, and the power series of Weierstrass), but at the end of the nineteenth century the concept of *function theory*, the theory of holomorphic functions of a complex variable, reached a significant stage of maturity and used all of the tools available to study new levels of problems which arose. At the turn of the nineteenth to the twentieth century various mathematicians began the study of holomorphic functions of several complex variables, and new phenomena (Hartogs's theorem) changed the direction of research in this higher-dimensional setting.

Chapter 10
Riemann Surfaces

10.1 Riemann's Multilayered Surfaces

One of the most important milestones in our study of the origins of complex geometry is the creation of the theory of Riemann surfaces. This singular creation by Riemann in his dissertation of 1851 [201] and his papers on Abelian functions in 1857 [195–198][1] developed over the next century into the very rich subject of *complex manifolds* of arbitrary dimension (Riemann surfaces being the case of a one-dimensional complex manifold), with strong overlaps with algebraic geometry, as we will indicate later.

Riemann's motivation for his creation of Riemann surfaces arose from the study of multivalued functions, and in particular in the multivalued functions that arose in Abel's work on generalizations of elliptic integrals that we discussed in Sect. 7.3. Multivalued functions had been a topic that had occupied mathematicians a great deal during the several centuries preceding Riemann's work, and the ambiguities that arose were a major concern. A fundamental example that arose in Abel's work was the study of a $y(x)$ which occurred as a solution of the algebraic equation

$$y^n + a_{n-1}y^{n-1} + \ldots + a_0(x) = 0,$$

where the $a_k(x)$ are polynomials in x. A different and familiar set of examples is given by the inverses of elementary transcendental functions such as e^z and $\sin z$, which we denote by $\log z$ and $\arcsin z$, and which were intensely studied in the eighteenth century. These particular multivalued functions have an infinite number of different values at a given point.

A key ingredient in Riemann's creation of Riemann surfaces was the notion of analytic continuation of a holomorphic function, which we discussed briefly earlier.

[1]In Riemann's collected works [200] these four papers are published together under the heading of a single paper entitled *Theorie der Abel'schen Functionen*. The first three papers summarize and clarify concepts developed in his dissertation from 1851 as tools for his detailed study of Abelian integrals and Abelian functions in the fourth paper. We will often refer simply to his Abelian functions paper of 1857.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_10

As Riemann observes on p. 102 of his Abelian functions paper [196], the function $\log(z-a)$, a well-known multivalued function, when continued analytically on a simple closed path around the point a, increases or decreases its value by $2\pi i$, depending on the direction of the path. If we let $z - a = re^{i\theta}$ be polar coordinates at the point a, then

$$\log(z - a) = \log re^{i\theta} = \log r + i\theta,$$

and as θ varies from 0 to 2π, $\log(z - a)$ varies from $\log r$ to $\log r + 2\pi i$. This well-known phenomenon played a major role in Riemann's work.

Riemann considered the possible different analytic continuations of a given holomorphic function to be *branches* (Zweige) of the function, and he defines a *branch point* (Verzweigungsstelle) as a point around which one branch moves into another (in our example a is a branch point for the multivalued function $\log(z - a)$). He then describes (on pp. 103–104) of [196] the surfaces he wants to consider:

> Für manche Untersuchungen, namentlich für die Untersuchung algebraischer und Abel'scher Functionen ist es vortheilhaft, die Verzweigungsart einer mehrwerthigen Function in folgender Weise geometrisch darzustellen. Man denke sich in der (x, y)-Ebene eine andere mit ihr zusammenfallende Fläche (oder auf der Ebene einen unendlich dünnen Körper) ausgebreitet, welche sich so weit und nur so weit erstreckt, als die Function gegeben ist. Bei Fortsetzung dieser Function wird also diese Fläche ebenfalls weiter ausgedehnt werden. In einem Theile der Ebene, für welchen zwei oder mehrere Fortsetzungen der Function vorhanden sind, wird die Fläche doppelt oder mehrfach sein; sie wird dort aus zwei oder mehreren Blättern bestehen, deren jedes einen Zweig der Function vertritt. Um einen Verzweigungspunkt der Function herum wird sich ein Blatt der Fläche in ein anderes fortsetzen, so dass in der Umgebung eines solchen Punktes die Fläche als eine Schraubenfläche mit einer in diesem Punkte auf der (x, y)-Ebene senkrechten Axe und unendlich kleiner Höhe des Schraubenganges betrachtet werden kann. Wenn die Function nach mehren Umläufen des um den Verzweigungswerth ihren vorigen Werth wieder erhält (wie z.B. $(z - a)^{\frac{m}{n}}$, wenn m, n relative Primzahlen sind, nach n Umläufen von z um a), muss man dann freilich annehmen, dass sich das oberste Blatt der Fläche durch dieübrigen hindurch in das unterste fortsetzt.
>
> Die mehrwerthige Function hat für jeden Punkt einer solchen ihre Verzweigungsart darstellenden Fläche nur einen bestimmten Werth und kann daher als eine völlig bestimmte Function des Orts in dieser Fläche angesehen werden.[2]

[2]"For many investigations, namely for the investigation of algebraic and Abelian functions it is advantageous to geometrically represent the branching nature of a multivalued function in the following manner. One imagines a surface (or an infinitesimally thin body) coinciding with and spread over the (x, y)-plane, which is extended as far as, and only as far as, the function is given. As the function is analytically continued, the surface will be further extended. In a region of the plane, for which two or more continuations are present, the surface will be covering twice or more times the region; it will consist of two or more sheets, each of which will represent a branch of the function. Around a branch point the function will continue from one sheet of the surface so that in the neighborhood of such a point the surface can be considered as a helicoid with a vertical axis through this point, and infinitesimally small heights of the screw thread from one revolution to another. If the function comes back to its same value after several such revolutions (as happens, for instance with $(z - a)^{\frac{m}{n}}$, if m, n are relatively prime numbers, after n cycles of z around a), then one has to assume that the upper sheet moves through the other sheets to the bottom sheet.

The multivalued function has, for every point of such a surface representing its branching, only one definite value, and can thereby be a completely well-determined function of position on this surface.".

This description of a Riemann surface and Riemann's further use of it in these four papers became the standard way (a branched covering) to describe Riemann surfaces for the next half century until Hermann Weyl (1885–1955) introduced the first abstract version of a Riemann surface as a topological manifold with a complex structure in 1913 [241]. Fundamentally, the Riemann surface as a branched covering of the extended complex plane gave local coordinates at each point of the surface except at the branch points, and at a branch point a of the type described in the quote above from Riemann one can use $\zeta = (z - a)^{\frac{1}{n}}$ as a local coordinate chart at this point. Riemann also added the point at infinity, ∞, to each sheet of the Riemann surface, thus giving rise to a closed or compact Riemann surface, with the local coordinate system $\zeta = 1/z$ at the point at infinity. The system of local coordinate charts for points of a Riemann surface was formalized by Hermann Weyl in his book mentioned above, but the nineteenth century mathematicians worked quite well with the structure Riemann set up and which we have summarized here.

10.2 The Analysis Situs of Riemann

Riemann's second paper in this series of four [198] has the title "Lehrsätze aus der analysis situs für die Theorie der Integrale von zweigliedrigen vollständigen Differentialen".[3] *Analysis situs* was a somewhat common name in the nineteenth century for what became *topology* (or what became *algebraic topology*, more specifically) in the twentieth century (we will discuss the origin of the word "topology" shortly). The term *analysis situs* originated in work of Leibniz which was contained in correspondence between Leibniz and Huygens, with the first and most fundamental letter being from Leibniz to Huygens on 8 September 1679 [141] in which he compared the geometry of magnitude with the geometry of position (*situm*), and felt that he could contribute to this new way of thinking by expressing positions of geometric objects and their relationships with symbols, just as algebra used symbols to represent relationships between numbers. Leibniz felt this was very important, but Huygens remained skeptical of this optimistic young mathematician's ideas in this direction, while recognizing the significance of his work on infinitesimal analysis. The recent book by Vincenzo Risi, *Geometry and Monadology* [203], has a very interesting analysis of Leibniz's work on analysis situs.

The work of Leibniz forms part of the inspiration for the book *Géométrie de Position* by Carnot [31], which had a major influence on projective geometry and where an important impulse was to investigate geometric phenomena that were not dependent on measurement of distances. The first definitive work on topology after the initial impetus by Leibniz came from Euler in 1735 in his famous solution of the *Seven Bridges of Königsberg* problem [61]. Note the title of this paper, *Solutio problematis ad geometriam situs*, contains the phrase *geometriam situs*, almost exactly the term used by Leibniz, which Euler cited and which Carnot used in the title of his

[3] "Theorems from analysis situs for the theory of integrals of two-fold complete differentials".

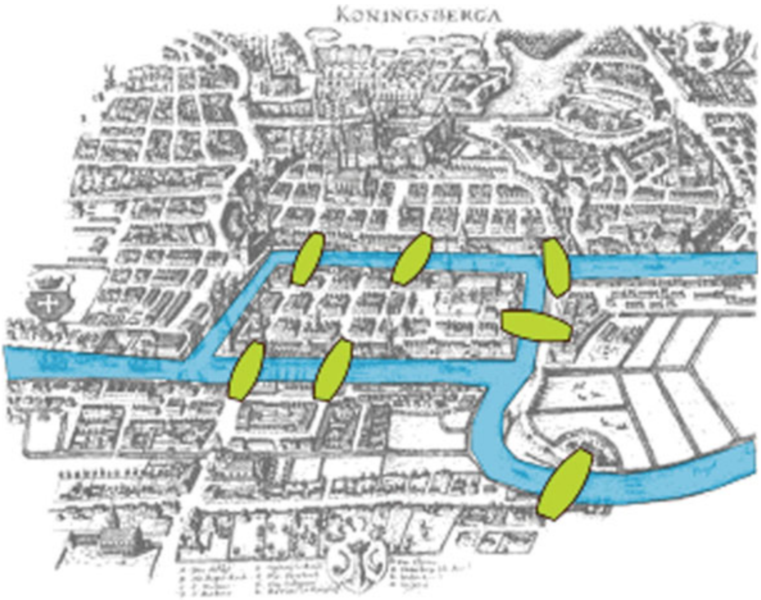


Fig. 10.1 The Königsberg Bridges problem

book. Euler was able to abstract the problem (Can one find a path crossing all seven bridges precisely once? See Fig. 10.1) to be a problem in what became graph theory and showed there was no such path.

Somewhat later Euler gave the first example of what became the *Euler characteristic* of a polyhedral surface [63]. Euler used a triangulation of a two-sphere in three-space, a covering of the sphere by triangles (faces), edges, and vertices. Figure 10.2 shows a number of illustrations of such triangulations from his paper. In all cases, the Euler characteristic χ, defined to be

$$\chi = F - E + V,$$

was equal to 2, where F is the number of faces, E is the number of edges, and V is the number of vertices in the triangulation. This Euler characteristic, and all of its generalizations in algebraic topology and other forms of geometry, has played an important role in mathematics.

In the nineteenth century Johann Benedict Listing (1808–1882) published a short book entitled *Vorstudien zur Topologie* [149], which followed up on the work of Euler and developed a theory of knots, a special topic in algebraic topology today concerning curves embedded in (usually) some Euclidean space, and in Listing's case, in $\mathbf{R}^3$. Then a few years later came Riemann's work that we are discussing here on the algebraic topology of surfaces.

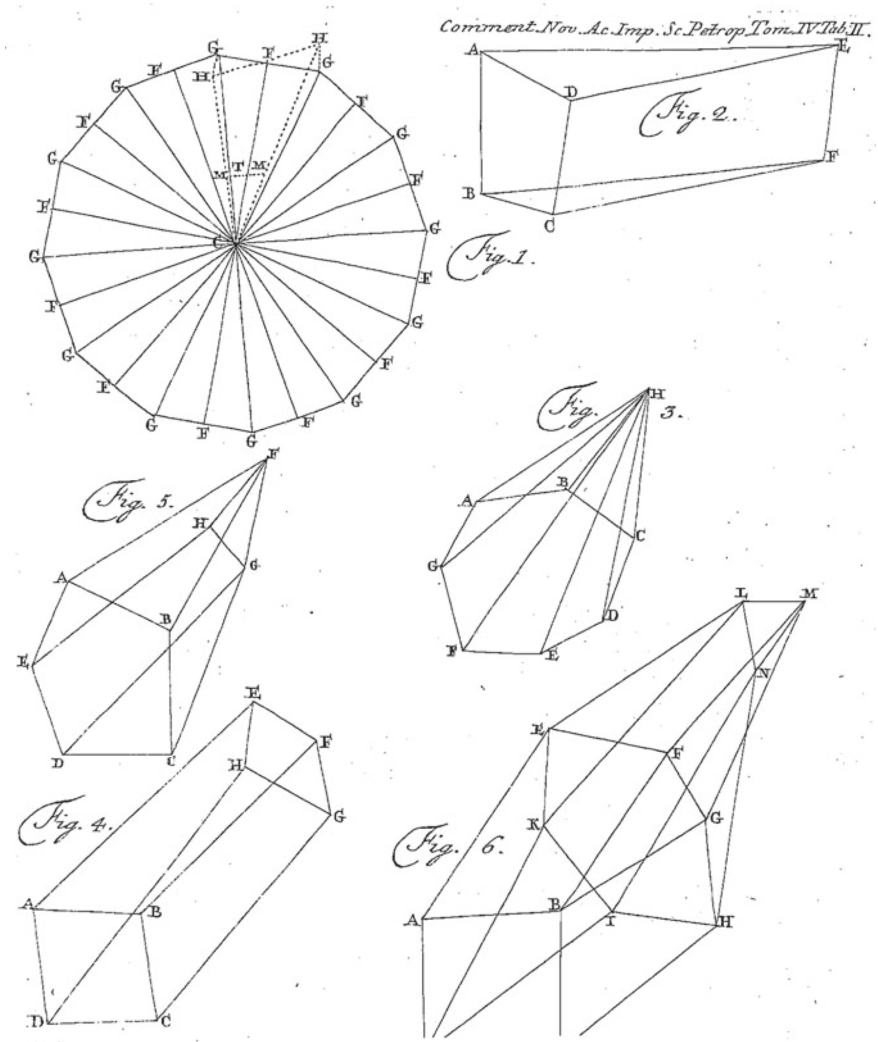


Fig. 10.2 Page 33 of Euler's paper on triangulations of a two-sphere [63]

Riemann defined both in his dissertation [199] and in the second paper in the Abelian functions series [198] the notion of *connectivity* of a surface. First he defined a *simply-connected surface* S to be any surface such that the complement of any closed curve (or a curve from one boundary point to another for a surface with a boundary) in the surface consisted of two components. He then defined an $n+1$-connected surface to be a surface S whereby n suitably chosen curves deleted from the surface would give a simply-connected subdomain $S' \subset S$. He showed that this concept was well defined and used it extensively in the remainder of his Abelian functions papers.

In his dissertation he had used the notion of a simply-connected domain to formulate his Riemann mapping theorem, which we discussed earlier. In his Abelian functions paper, the non-simply-connected compact surfaces play the most important role, as that is where the theory of Abelian functions can be developed. In Fig. 10.3 we see diagrams from Riemann's paper [198], which illustrate the notion of connectivity. Note that the illustration at the bottom of Fig. 10.3 shows the possibility of two sheets of a surface overlapping as described in his definition of a Riemann surface as a branched covering.

Let's consider a simple example illustrating the topology of a compact Riemann surface arising from an algebraic function. Let the Riemann surface S be defined by the polynomial

$$w^2 = (z - a_1)(z - a_2)(z - a_3)(z - a_4), \tag{10.1}$$

where the points a_1, a_2, a_3, a_4 are distinct, and let λ and μ be two nonoverlapping cuts joining a_1 to a_2 and a_3 to a_4, as illustrated in the top part of Fig. 10.4. Here $\lambda^{\pm}$, $\mu^{\pm}$ indicate the two sides of the cuts λ and m. By opening the cuts as indicated in the bottom portion of Fig. 10.4, we see that we can glue the two copies of the Riemann sphere E and E' along cylinders joining them to create the torus on the lower right-hand side of the figure. Thus we conclude that the topology of the Riemann surface defined by (10.1) is that of a torus.

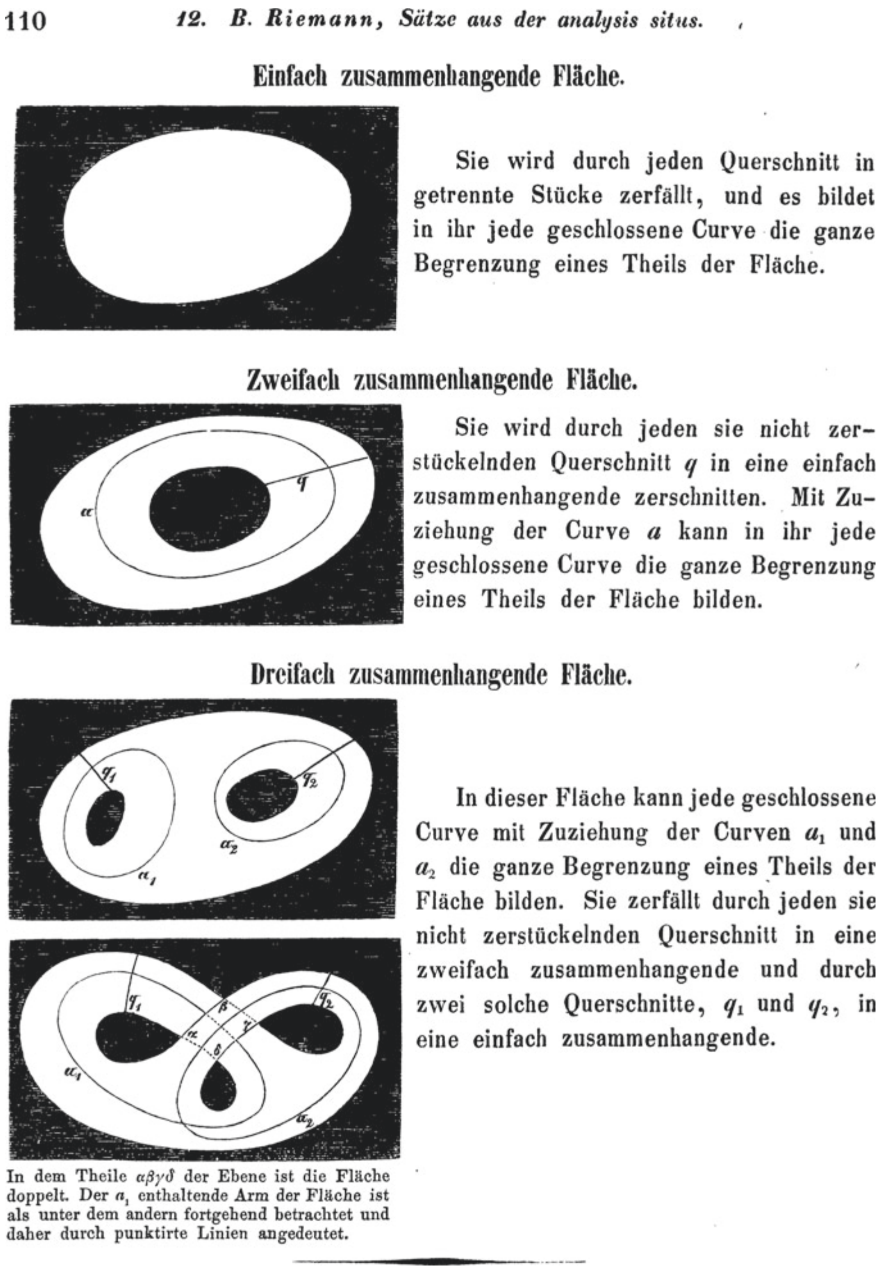
In the third of these preliminary papers [197] Riemann revisits the Dirichlet principle and shows that there exist meromorphic functions with prescribed poles or logarithmic singularities. Strictly speaking, a meromorphic function, by definition, does not have logarithmic singularities, but Riemann is interested also in the *integrals* of meromorphic functions from a specified point to an indefinite point. That is, he is interested in Abelian integrals on a Riemann surface, and such integrals can have a logarithmic singularity, e.g.,

$$\int_{z_0}^{z} \frac{dz}{z} = \log(z - z_0).$$

This is an important existence theorem for Riemann surfaces, which is considered in some detail by later authors, in particular in the book by Hermann Weyl [241]. We will discuss Riemann's existence theorem for meromorphic functions on a Riemann surface in Sect. 10.4.

10.3 Abelian Integrals and Abelian Functions

In the fourth paper in this series, *Theorie der Abel'schen Functionen* [195], Riemann develops his theory of *Abelian functions*, a vast generalization of elliptic functions which are defined as several-variable inverses of Abelian integrals on Riemann surfaces of genus >1. Riemann's work followed up on some announced results of Weierstrass whose proofs hadn't yet appeared at the time Riemann wrote his paper.



110 12. B. Riemann, Sätze aus der analysis situs.

Einfach zusammenhangende Fläche.

Sie wird durch jeden Querschnitt in getrennte Stücke zerfällt, und es bildet in ihr jede geschlossene Curve die ganze Begrenzung eines Theils der Fläche.

Zweifach zusammenhangende Fläche.

Sie wird durch jeden sie nicht zerstückelnden Querschnitt q in eine einfach zusammenhangende zerschnitten. Mit Zuziehung der Curve a kann in ihr jede geschlossene Curve die ganze Begrenzung eines Theils der Fläche bilden.

Dreifach zusammenhangende Fläche.

In dieser Fläche kann jede geschlossene Curve mit Zuziehung der Curven a_1 und a_2 die ganze Begrenzung eines Theils der Fläche bilden. Sie zerfällt durch jeden sie nicht zerstückelnden Querschnitt in eine zweifach zusammenhangende und durch zwei solche Querschnitte, q_1 und q_2, in eine einfach zusammenhangende.

In dem Theile $\alpha\beta\gamma\delta$ der Ebene ist die Fläche doppelt. Der a_1 enthaltende Arm der Fläche ist als unter dem andern fortgehend betrachtet und daher durch punktirte Linien angedeutet.

Fig. 10.3 Riemann's connectivity in his analysis situs paper [198]

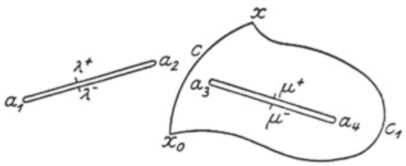


Abb. 61.

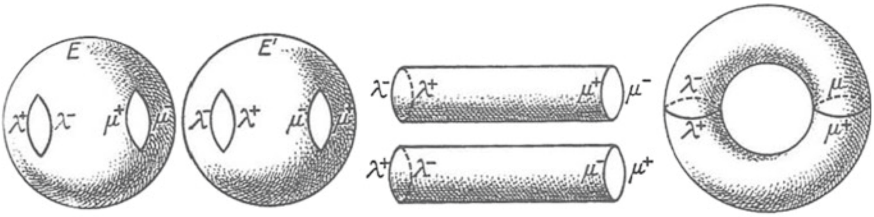


Abb. 62. Abb. 63. Abb. 64.

Fig. 10.4 Example of the Riemann surface of $w^2 = (z - a_1)(z - a_2)(z - a_3)(z - a_4)$ from pp. 236–237 of Hurwitz and Courant [113]. *This illustration is reprinted with the permission of Springer*

The basic theory which evolved became known as the theory of Abelian functions (this name was adopted by Riemann in the paper we are discussing), and was further developed by numerous mathematicians in the following century. We will discuss this briefly later in this section.

We want to discuss some aspects of this paper which directly relate to the theory of complex manifolds in the twentieth century. Let now S be a Riemann surface defined as a branched covering defined by the polynomial function

$$F(z, w) = w^n + a_{n-1}(z)w^{n-1} + \ldots + a_0(z) = 0, \tag{10.2}$$

where $F(z, w)$ is an irreducible polynomial of degree n in w and of degree m in z. Then the function $w(z)$ defined by (10.2) is a single-valued function on S, and S is a compact Riemann surface, where we have added a point at infinity to each sheet of the Riemann surface, as did Riemann. Let us suppose that S is $(2p + 1)$-connected, for $p \geq 0$, and we say that S has *genus* p. This is the topological definition of genus. That the connectivity is odd for compact Riemann surfaces was shown by Riemann in his paper. In Fig. 10.5 we see illustrations of surfaces of genus 0, 1, 2, and 3.

Now consider the Abelian integral

$$A(z) = \int_{z_0}^{z} r(z, w(z))dz,$$

along any path in S joining z_0 to z, where $r(z, w)$ is a rational function of z and w, as we discussed in Sect. 7.3. Note that in Sect. 7.3 we dealt with real integrals (with singularities in the integrand), and here we are dealing with path integrals in

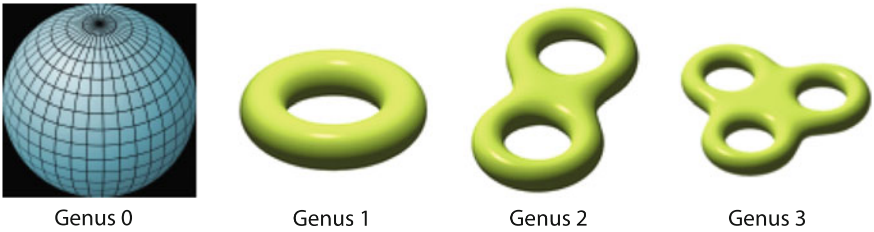


Fig. 10.5 Surfaces of genus 0, 1, 2, and 3

the complex plane or on a Riemann surface spread over the complex plane (also with possible singularities in the integrand, depending on the path).

As before, we define $A(z)$ to be an *Abelian integral of the first kind* if $A(z)$ is finite for all points $z \in S$. We will see examples of this momentarily. We note also that $A(z)$ is in principle a *multivalued* function if the genus $p > 0$, since different paths along portions of a cycle which is not homologous to zero can have different values.

If $p = 0$, then there are no Abelian integrals of the first kind. To see this, suppose that $A(z)$ is an Abelian integral on S, where $p = 0$. Then S is simply-connected and hence $A(z)$ is single-valued and holomorphic at each point of S. But S is compact and therefore at some point of S, by continuity, $|A(z)|$ assumes a maximum value, which is necessarily an interior point, and hence, by the maximum principle that Riemann proved in his dissertation (Sect. 9.4), $A(z)$ must be constant. However, a constant cannot be an Abelian integral, since its derivative would be zero, which cannot be the integrand of an Abelian integral as defined above.

Abelian integrals of the second kind are characterized by $A(z)$ having poles at some points of S, and Abelian integrals of the third kind are characterized by $A(z)$ having logarithmic singularities at some points of S. We will not be discussing these further here, but we will see that Abelian integrals of the first kind are intimately connected to the topology of the surface S.

Riemann shows in his paper [195] that on the Riemann surface S of genus p as defined above, there are precisely p linearly independent Abelian integrals of the first kind $A_1(z), A_2(z), \cdots, A_p(z)$, and that if $A(z)$ is any Abelian integral of the first kind on S, then there are constants $\alpha_1, \cdots, \alpha_p$ such that

$$A(z) = \alpha_1 A_1(z) + \ldots \alpha_p A_p(z) + \text{const.}$$

Thus the number of linearly independent Abelian integrals of the first kind[4] is the genus p. Let's give a couple of examples to illustrate this.

---

[4]In a later text on Abelian functions by Clebsch and Gordan written some 10 years after Riemann's work [49], they compute the number p for a Riemann surface defined by a polynomial of the type $F(z, w)$ in terms of degrees of the polynomial and numbers of double points and cusps, and refer to this simply as the number of Abelian integrals of the first kind.

We consider $F(z, w)$ to be of the form, for some $k > 0$, $|k| \neq 1$,

$$F(z, w) = w^2 - (z^2 - 1)(z^2 - k^2). \tag{10.3}$$

We let $R(z, w) = 1/w$, and we define

$$A_1(z) = \int_0^z \frac{dz}{w(z)},$$

(this is the elliptic integral considered by Abel (8.2) in our discussion of Abel's work on elliptic functions). If $z \neq \pm 1, \pm k$, then $A_1(z)$ is finite.

To see this, let γ be a path from 0 to ∞ not passing through these same points, then

$$\lim_{z\to\infty} A(z) = v, \text{ and } |v| < \infty,$$

which is easy to verify. Namely, any path γ going from 0 to ∞ and missing the points $\pm 1, \pm(1/k)$ can be modified to be a path from 0 to a point $R_0 > \max\{1, 1/k\}$ and from R_0 to ∞ via a path on the positive real axis. Then the resulting path integral part on the real axis would be

$$\int_{R_0}^{\infty} \frac{dx}{\sqrt{(x^2 - 1)(x^2 - k^2)}}.$$

But, for $x \in [R_0, \infty)$, there is a constant K such that

$$\frac{1}{\sqrt{(x^2 - 1)(x^2 - k^2)}} \leq K \frac{1}{x^2},$$

and thus $\lim_{z\to\infty} A_1(z)$ exists and is finite.

Now we have to examine the behavior of $A_1(z)$ at the singular points of the integrand. Consider the point $z = 1$. We want to show that $\lim_{z\to 1} A_1(z)$ exists and is finite. To do this we make a change of variable at this point of the form

$$\zeta = (z - 1)^{\frac{1}{2}},$$

then

$$d\zeta = \frac{1}{2}(z - 1)^{-\frac{1}{2}} dz,$$

which gives

$$\begin{aligned} dz &= 2\zeta d\zeta, \\ z &= 1 + \zeta^2, \end{aligned}$$

and hence

$$A_1(z) = A_1(1+\zeta^2) = \int_i^{1+\zeta^2} \frac{2\zeta d\zeta}{\sqrt{\zeta^2(\zeta^2+2)(\zeta^2+1+k)(\zeta^2+1-k)}}$$
$$= \int_i^{1+\zeta^2} \frac{2d\zeta}{\sqrt{(\zeta^2+2)(\zeta^2+1+k)(\zeta^2+1-k)}},$$

which has a nonsingular integrand near $\zeta = 0$, and hence $A_1(z)$ is holomorphic in a neighborhood of $z = 1$. A similar argument holds for the other singular points $\{-1, k, -k\}$, and hence $A_1(z)$ is finite at all points of S, as Legendre and others knew, long before Riemann, but in the real variable context.

Now the Riemann surface for the polynomial (10.3) has genus $p = 1$, as we discussed earlier, and thus, according to Riemann, the only Abelian integrals of the first kind in this case are constant multiples of $A_1(z)$ plus a possible constant.

A second example is the case of a hyperelliptic curve (hyperelliptic Riemann surface) defined by an equation of the form

$$F(z, w) = w^2 - (z-a_1)(z-a_2)\cdots(z-a_{2p}),$$

where $a_k \neq 0$ are distinct complex numbers, then

$$A_k(z) := \int_0^z \frac{z^k dz}{w(z)}, k = 0, \cdots, p-1, \tag{10.4}$$

are p distinct Abelian integrals of the first kind, as we mentioned in our discussion of Abel's work in Sect. 7.3. The proof that these are Abelian integrals of the first kind is similar to that given above in the elliptic case, and it is easy to verify that the genus for this Riemann surface is also p. Namely, there are $2p$ branch points for this two-sheeted Riemann surface. Hence the Abelian integrals in (10.4) form a basis for the vector space of Abelian integrals of the first kind.

Let's look at the integrands of these Abelian integrals of the first kind. Notice that originally in the work of Euler, Legendre, Abel and others, the integrals were integrals on the real axis with singularities (with the inherent multivaluedness). For instance, consider the integral

$$A_1(x) = \int_0^x \frac{dx}{\sqrt{(x^2-1)(x^2-k^2)}},$$

in the elliptic case. The integrand is

$$\frac{dx}{\sqrt{(x^2-1)(x^2-k^2)}},$$

and this becomes, going to the complex plane,

$$\frac{dz}{\sqrt{(z^2 - a)(z^2 - k^2)}},$$

which has the form

$$f(z)dz, \tag{10.5}$$

where $f(z)$ is a multivalued meromorphic function on $\mathbf{C}$. We shall see below how (10.5) can be interpreted as a single-valued and, in fact, holomorphic one-form on the Riemann surface defined by $w^2 = (z^2 - 1)(z^2 - k^2)$.

Let S be a Riemann surface spread over the extended complex plane of the type described by Riemann with local coordinates: $\zeta = z - a$ at nonbranching points a, $\zeta = (z - a)^{\frac{1}{k}}$ at a branching point a of order k, and $\zeta = 1/z$ at ∞, assuming, for simplicity that the point at infinity is not a branch point. A *meromorphic one-form* on S is a one-form defined with respect to any local coordinate chart ζ as above to be of the form

$$\omega = f(\zeta)d\zeta,$$

where $f(\zeta)$ is meromorphic in ζ. And, if we transform from one coordinate system z to another by a change of coordinates $\tilde{\zeta}(\zeta)$, then there is a meromorphic function $\tilde{f}(\tilde{\zeta})$ such that

$$\omega = \tilde{f}(\tilde{z})d\tilde{\zeta} = f(\zeta)d\zeta,$$

where $d\tilde{\zeta} = \tilde{\zeta}'(\zeta)d\zeta$, and $\tilde{\zeta}'(\zeta)$ is a holomorphic function of ζ.

We will say that a meromorphic one-form ω on S is a *holomorphic one-form* on S if, for each local coordinate ζ, $\omega = f(\zeta)d\zeta$, where the coefficient function $f(\zeta)$ is holomorphic.

If we look at the integrands of the examples of Abelian integrals that we discussed above, then it is easy to check that

$$\omega_1 = \frac{dz}{\sqrt{(z^2 - 1)(z^2 - k^2)}},$$
$$\omega_k = \frac{z^k dz}{\sqrt{(z - a_1)(z - a_2)\cdots(z - a_{2p})}}, k = 0, \ldots, p - 1$$

are indeed holomorphic one-forms on the respective Riemann surfaces of genus 1 and genus p. First of all, these are indeed meromorphic one-forms as they are defined, and it remains to show that they are holomorphic near any of the singular points (where these one-forms have potential poles). The calculations are essentially the same as those we used to show that these integrals had well-defined values at the singular points.

For instance, near the singular point $z = a_1$ for the one form ω_k in the second case, we have

$$\zeta = (z - a_1)^{\frac{1}{2}},$$
$$d\zeta = \frac{1}{2}(z - a_1)^{-\frac{1}{2}} dz,$$

which gives

$$z = \zeta^2 + a_1,$$
$$dz = 2\zeta d\zeta,$$

and hence

$$\omega_k = \frac{2(\zeta^2 + a_1)^k d\zeta}{\sqrt{(\zeta^2 + a_1 - a_2)(\zeta^2 + a_1 - a_3) \cdots (\zeta^2 + a_1 - a_{2p})}},$$

which is holomorphic near $z = 0$. It still has to be checked that ω_k is holomorphic at ∞, and this is a similar calculation.

Using this terminology we see that the number of linearly independent Abelian integrals of the first kind on a Riemann surface S is the same as the number of linearly independent holomorphic one-forms on S. This point of view became standard in modern treatments of Riemann surfaces using differential forms to represent the connectivity (the genus) using de Rham's theorem, and using the holomorphic one-forms to be a way of representing this topology when the complex structure is assumed. This is part of the essence of Hodge theory on general Kähler manifolds, as we see in Sect. 14.4.

In the remainder of his main Abelian function paper [195], Riemann formulated and solved several geometric problems which have come to have major significance in the subsequent development of Riemann surfaces, algebraic geometry and complex manifolds. In addition, he formulated and gave his version of a solution to the Jacobi inversion problem. This was concerned with the generalizations of the inverses of elliptic integrals to inverses of Abelian integrals, which are now called Abelian functions. Finally, he showed that on any compact Riemann surface there exist nonconstant meromorphic functions. This important result became the basis for the Riemann–Roch theorem that we outline in the following section. We will look at the geometric problems first, return to the Jacobi inversion problem, and conclude with our next section concerned with the Riemann–Roch theorem.

The first geometric problem he formulated and resolved was to show that the Riemann surfaces of the kind he had described as a branched covering of $\overline{\mathbf{C}}$ could be realized as the Riemann surface defined by a polynomial function $F(z, w)$. In modern terms the question he raised could be reformulated to ask if a compact Riemann surface could be realized as a projective algebraic submanifold of complex projective space. The answer to this question is that it is indeed possible, and this is a special case and simple consequence of the Kodaira embedding theorem for compact complex manifolds (see Chap. 14).

The second problem he formulated was to consider the birational equivalence of solutions of equations of the form $F(z, w) = 0$ for different choices of the polynomial $F(z, w)$. Let's formulate this question somewhat more precisely. Let

$$C = \{(z, w) \in \mathbf{C}^2 : F(z, w) = 0\},$$

where $F(z, w)$ is a polynomial in the variables z and w. We call such a C an *algebraic curve*, and we include in C points at infinity (which is easy to do using homogeneous coordinates, and became standard in algebraic geometry shortly after the time of Riemann).[5]

Two algebraic curves

$$\begin{aligned} C &= \{F(z, w) = 0\}, \\ C_1 &= \{F_1(z_1, w_1) = 0\} \end{aligned}$$

are *birationally equivalent* if there exist rational mappings

$$z_1(z, w), z_2(z, w),$$

with inverse rational mappings

$$z(z_1, w_1), w(z_1, w_1),$$

such that

$$F_1(z_1, w_1) = F(z(z_1, w_1), w(z_1, w_1))$$

and

$$F(z, w) = F_1(z_1(z, w), w_1(z, w)).$$

Riemann formulated the problem of classifying equivalence classes of birationally equivalent algebraic curves, and, by dimensional analysis of the parameters, he concluded that for a given algebraic curve of genus p (defined by an equation $F(z, w) = 0$):

- For $p = 0$, all curves are birationally equivalent.
- For $p = 1$, there is a one-complex-parameter family of birationally inequivalent algebraic curves.
- For $p > 1$, there is a $3p-3$-complex-parameter family of birationally inequivalent algebraic curves.

Riemann termed these parameters the *moduli* of the algebraic curves, and a major problem in mathematical research over the next century became to understand

[5] For instance, the text by Clebsch and Gordan on Abelian functions [49], published in 1866, nine years after Riemann's fundamental papers of 1857, formulated this theory in terms of homogeneous coordinates.

the nature and representations of these moduli. For Riemann the classification of algebraic curves was equivalent to the classification of Riemann surfaces associated to these algebraic curves.

In algebraic geometry the classification of algebraic varieties of one or more dimensions has been a very important research topic ever since the time of Riemann. In the theory of complex manifolds (most complex manifolds do not correspond to solutions of algebraic equations), the deformations of complex structures on manifolds of one or more dimensions has been an equally rich field of research. The theory of Teichmüller spaces plays an important role in the contemporary theory of moduli of Riemann surfaces (an abstraction of Riemann's moduli of algebraic curves; see Bers [18] for an overview of moduli theory in the mid-twentieth century).

Finally, we discuss briefly the main topic Riemann was addressing in this series of papers, the Jacobi inversion problem. Namely, let

$$F(z, w) = 0$$

define a Riemann surface S of genus p, and let $A_1(z), \cdots, A_p(z)$ be p linearly independent Abelian integrals of the first kind on S. Each of these are multivalued functions and the value of each of these functions at a given point z depends on the path of integration from some fixed initial point to the point on the Riemann surface whose coordinate in the extended complex plane is z, which is implicit in the definition of each A_k. Define the functions

$$\begin{aligned} v_1(z_1, \cdots, z_p) &= A_1(z_1) + \ldots + A_1(z_p), \\ v_2(z_1, \cdots, z_p) &= A_2(z_1) + \ldots + A_2(z_p), \\ &\vdots \\ v_p(z_1, \cdots, z_p) &= A_p(z_1) + \ldots + A_p(z_p). \end{aligned} \tag{10.6}$$

The *Jacobi inversion problem* is to find an inverse to this mapping and determine its properties. An inverse mapping here would be p functions

$$\begin{gathered} z_1(v_1, \cdots, v_p), \\ z_2(v_1, \cdots, v_p), \\ \vdots \\ z_p(v_1, \cdots, v_p), \end{gathered} \tag{10.7}$$

which provide an inverse to the mapping described in (10.6).

For the case of $p = 1$, we have only one Abelian integral to deal with, and this is an elliptic integral, and its inverse is an elliptic function as discovered by Abel and Jacobi, and which we discussed at length in Chap. 8.

Thus the inverse functions

$$\{z_1(v_1, \cdots, v_p), z_2(v_1, \cdots, v_p), \cdots, z_p(v_1, \cdots, v_l p)\}$$

in (10.7) would each be a function of p complex variables which would then be generalizations of elliptic functions, presumably with some periodicity properties of the same sort as elliptic functions have. We shall see that the inverse functions do exist and are meromorphic functions in $\mathbf{C}^p$ with $2p$ independent periods. Riemann called such functions *Abelian functions* in honor of Abel, who had studied in the various versions of Abel's theorem sums of Abelian integrals of the sort that appear in (10.6) (as we discussed in Sect. 7.3).

One measure of the multivalued nature of the Abelian integrals used in (10.6) is to use what are called the *periods of an Abelian integral*, and these will turn out to be the periods of the corresponding Abelian functions described briefly above. Let $A_k(z)$ be k linearly independent (over the real numbers) Abelian integrals of the first kind on a Riemann surface S of genus p, $k = 1, \cdots, p$, and let $\gamma_1, \cdots, \gamma_{2p}$ be closed curves (*cycles*) on S which represent cuts which reduce the Riemann surface S to a simply-connected open subset $S' \subset S$. Then let

$$\omega_k^j := \int_{\gamma_j} \alpha_k(z), \tag{10.8}$$

where $\alpha_k(z)$ is the *integrand* of $A_k(z)$ considered as a holomorphic one-form on S, as was illustrated in our examples earlier. These are the *periods* of these Abelian integrals $A_k(z)$ for this choice of cycles $\gamma_1, \cdots, \gamma_{2p}$. It follows that if γ and $\tilde{\gamma}$ are any two paths joining an initial point z_0 to a variable point z on S, and if $A_k(z)$ represents the value of the Abelian integral along the path γ, and $\tilde{A}_k(z)$ represents the value of the same Abelian integral along the path $\tilde{\gamma}$, then

$$\tilde{A}_k(z) = A_k(z) + m_k^1\omega_k^1 + \ldots + m_k^{2p}\omega_k^{2p},$$

where m_k^j are integers. Thus, theperiods $\omega_k^1, \cdots, \omega_k^{2p}$ represent precisely the multivalued nature of the Abelian integral $A_k(z)$.

We now define a *meromorphic function* $f(z) = f(z_1, \cdots, z_p)$ on an open domain $D \subset \mathbf{C}^p$, $p \geq 1$, to be a holomorphic function on $D - S$, where S is a closed lower-dimensional subset of D such that near any point $z^0 = (z_1^0, \ldots, z_p^0)$ of D, f can be represented as the quotient of two holomorphic functions. The singular points S correspond to the points where the local holomorphic function in the denominator is zero. Thus the set S is generically a $p-1$-dimensional locally defined complex submanifold of $\mathbf{C}^p$. For instance, if we set

$$f(z_1, z_2) = \frac{z_1 - z_1^0}{z_2 - z_2^0},$$

then $f(z)$ is meromorphic on $\mathbf{C}^2$, it has zeros on the line $z_1 = z_1^0$, it has poles along the line $z_2 = z_2^0$, and it has no well defined value at the singular point (z_1^0, z_2^0).

Consider a meromorphic function $f(z)$ on $\mathbf{C}^p$. Let $\omega \in \mathbf{C}^p$ be a fixed complex p-tuple, $(\omega_1, \cdots, \omega_p) \neq 0$. Then we say the function $f(z)$ is *periodic with respect* *to the period* ω if

$$f(z+\omega) = f(z), \text{ for all } z \in \mathbf{C}^p.$$

Let us define the vectors

$$\omega^j = (\omega_1^j, \cdots, \omega_p^j),\ j = 1, \cdots, 2p,$$

where the ω_k^j are defined as the periods of the Abelian integrals as in (10.8). Riemann, Weierstrass and others showed that there exist meromorphic functions $f(z)$ on $\mathbf{C}^p$, such that

$$f(z + m_1\omega^1 + \cdots + m_{2p}\omega^{2p}) = f(z), z \in \mathbf{C}^p,$$

where the m_j are arbitrary integers. More precisely, there exist functions with $2p$ independent periods (namely, the periods defined above are linearly independent over the real numbers), and they also showed there are no functions with more than $2p$ periods.

Riemann used theta functions in [195] to demonstrate the existence of Abelian functions, in a manner similar to our discussion of the use of theta functions to represent elliptic functions in Sect. 8.4. Weierstrass formulated (by differentiation) (10.6) as differential equations and solved these using power-series methods. Both solutions gave great impetus to further research in the rich theory of Abelian functions during the latter half of the nineteenth century. We recommend highly the interesting book by Markushevich [151] which gives a detailed and very well written history of the early development of both elliptic and Abelian functions.

10.4 The Riemann–Roch Theorem

The final topic in Riemann's Abelian function paper [202] that we want to discuss is the question of the existence of meromorphic functions on a compact Riemann surface. If we look at Riemann's construction of a Riemann surface S as a multisheeted branched covering of the extended complex plane $\overline{\mathbf{C}}$, then the covering mapping $\pi : S \to \overline{\mathbf{C}}$ is indeed a meromorphic function on S. Riemann asked if there were any other meromorphic functions[6] on S.

We note first that there cannot be any nonconstant holomorphic functions on a compact Riemann surface. This follows from the maximum principle, just as in our proof in the previous section that there are no Abelian functions of the first kind on a Riemann surface of genus 0.

[6]The Riemann surfaces Riemann considered when he wrote his paper were all solutions of an algebraic equation $F(z, w) = 0$, solving for w in terms of z, the variable in the complex plane. In this context, $w(z)$ was a multivalued function of z which yielded the Riemann surface S with branch points. Riemann was looking for "gleich-verzweigte algebraische Funktionen" on the complex plane, i.e., functions of z with the same branching pattern as $w(z)$, and thus were single-valued functions on S.

For the remainder of this section, we assume that S is a compact connected Riemann surface of genus p. The Abelian integrals $A_1, \ldots, A_p$ that we discussed in the previous section are all multivalued holomorphic functions on S. Riemann's construction of nonconstant meromorphic functions on S involve linear combinations of specific multivalued meromorphic functions on S and these Abelian integrals. We recall that Abel often used the fact that a symmetric rational function of the roots of a polynomial (multivalued function) is a symmetric rational function of the coefficients of the polynomial (single-valued function). Riemann's construction is similar in spirit: constructing single-valued functions from combinations of multivalued functions.

Riemann uses the Dirichlet principle to construct specific multivalued meromorphic functions that we will utilize. Let $p_1, \cdots, p_m$ denote m distinct points on S. At one such point p_l, he uses the Dirichlet principle to construct a function u_λ which is harmonic on $S - \{p_l\}$ and which has the local behavior at p_l (in a local coordinate system centered at p_l),

$$u_l(z) = \mathrm{Re}\left(\frac{1}{z}\right) + \varphi_l(z),$$

where $\varphi_l(z)$ is harmonic near $z = 0$.

Let now $z = x + iy$ be any local coordinate system on S, and define the one-form

$$*u_l := \frac{\partial u_l}{\partial x}dy - \frac{\partial u_l}{\partial y}dx.$$

It is easy to check that this one-form is independent of the coordinate system. It has a singularity at the point p_l. Let now x_0 be a fixed point on S, disjoint from the points $p_1 \cdots, p_m$, and let γ be a smooth path from x_0 to any point $x \in S - \{p_l\}$, then define the *harmonic conjugate* of u_l on $S - \{p_l\}$ to be the integral of $*u$ along the path γ from the point x_0 to the point x. The function v_l is then a multivalued harmonic function[7] defined on $S - \{p_l\}$, and such that

$$f_l := u_l + iv_l$$

is a well-defined multivalued meromorphic function on S with a simple pole with residue 1 at the point p_l.

We can assume that the $2p$ cycles γ_j that describe the topology[8] of S do not intersect the m points $p_1, \cdots, p_m$, and we can define the periods of f_l to be the integrals

$$c_l^j := \int_{\gamma_j} df_l, \tag{10.9}$$

[7]Riemann uses cuts (curves) C_j on S so that $S - \cup_j C_j$ is simply-connected, and defines the values of v_l on a cut to be a limit from one side or another of a given cut, making the function v_l harmonic and multivalued on $S - \{p_l\}$.

[8]This is the $(2p+1)$-connectedness in Riemann's language, or a basis of the homology group in more modern terms.

and the values of f_l at any given point differ from one another by integral multiples of these periods.

Now Riemann considers the sum

$$f := a_1 f_1 + \ldots + a_m f_m + b_1 A_1 + \ldots + b_p A_p, \tag{10.10}$$

which is a new multivalued meromorphic function on S. Let us denote the periods of f by π_f^j, defined as in (10.9). These are linear combinations of the periods of f_l and the periods of A_k, namely,

$$\pi_f^j = a_1 c_1^j + \ldots b_m c_m^j + b_1 \omega_1^j + \ldots + b_p \omega_p^j. \tag{10.11}$$

Thus if we let $(\pi_f^1, \ldots, \pi_f^{2p})$ be a vector in $\mathbf{C}^{2p}$, then the linear equation (10.11) defines a linear mapping

$$L : \mathbf{C}^{m+p} \to \mathbf{C}^{2p}$$

given by

$$(a_1, \cdots, a_m, b_1, \cdots, b_p) \mapsto (\pi_f^1, \cdots, \pi_f^{2p}).$$

Riemann observes that, for $m + p > 2p$, then necessarily $\ker L \neq 0$, and hence there are coefficients a_l, b_k in (10.10) such that all of the periods of f are zero. It follows that f is a nonconstant single-valued meromorphic function on S, which was what Riemann had set out to prove. Namely, on any Riemann surface of genus p, there exist nonconstant meromorphic functions. For genus $p = 0$, these are, of course, the rational functions on the extended complex plane $\overline{\mathbf{C}}$, which were well known. Also, for genus $p = 1$, the elliptic functions of Abel, Jacobi and Weierstrass, which we discussed in Chap. 8, were all nonconstant meromorphic functions on these Riemann surfaces (now called, almost universally, elliptic curves).

Looking at the coefficients a_l in (10.10), which yield such a nonconstant meromorphic function f, one sees that at least one, but not necessarily all, of these coefficients must be nonzero (linear combinations of the Abelian integrals alone could not yield a nonconstant single-valued function, as we saw above by the maximum principle). Thus f is a meromorphic function with simple poles at some of the points p_l. Riemann asks the question: how many meromorphic functions on S are there with at most simple poles at the points p_l? This question and the variety of answers to this question became the basis for the Riemann–Roch theorem in its various formulations over the past century and a half.

In the late nineteenth century, language from algebraic number theory was adapted to the geometry of Riemann surfaces to describe the formulation of this type of problem (this was first described in the monograph of Hensel and Landsberg [100], later used by Weyl in 1913 [241], and has been standard ever since). We will now use this language to describe the results of Riemann's solution to this problem.

Let p_l be a discrete set of points on S. We let the formal sum

$$D = \sum_l n_l p_l, \ \ n_l \in \mathbf{Z},$$

be a *divisor* [9] on S. We can add divisors by concatenating the points and adding the corresponding coefficients. Thus the divisors form an Abelian group. We define the *degree of the divisor* D to be the sum of the coefficients in D, namely,

$$\deg(D) := \sum_l n_l.$$

Let now $\mathcal{M}(S)$ denote the meromorphic functions on S. Although the fact that this set of functions has the structure of a field will not be that important for us here, it is an important concept in the algebraic study of Riemann surfaces defined by algebraic functions, for instance. We note that the degree function in this context is additive, i.e., $\deg(D_1 + D_2) = \deg(D_1) + \deg(D_2)$. We say that a divisor is *nonnegative* or *positive* if all of its coefficients are nonnegative or positive.

If $f \in \mathcal{M}(S)$ is a meromorphic function on S, then we let $z_1, \cdots, z_m$ be the zeros of f with multiplicities $\mu_1, \cdots, \mu_n$, and we let $p_1, \cdots, p_n$ be the poles of f with multiplicities $\nu_1, \cdots, \nu_n$. Then we let (f) denote the divisor associated to f defined by

$$(f) := \sum_k \mu_k z_k + \sum_l \nu_l p_l.$$

We note that the number of zeros of f, counting multiplicities, must be the same as the number of poles of f, also counting multiplicities. This is a consequence of the Cauchy residue theorem. Thus, it follows that $\deg((f)) = 0$.

Now, for any divisor D on S, we can let $L(D)$ be the vector space defined by

$$L(D) := \{f \in \mathcal{M}(S) : (f) + D \geq 0\}.$$

It is straightforward to verify that this is indeed a vector space. Riemann's fundamental contribution to what we now know as the Riemann–Roch theorem was a lower-bound estimate on the dimension of this vector space in terms of the degree of the divisor D and the genus of the Riemann surface, which we formulate in the following theorem.

Theorem 10.3 (Riemann [202]) *Let S be a compact Riemann surface of genus p, then, for any divisor D on S,*

$$\dim L(D) \geq \deg(D) - p + 1.$$

We note that if $\deg(D) < p$, then this theorem has no content, of course.

---

[9] We will see divisors again in Chaps. 14 and 15.

Let's look at the situation we discussed earlier, and let D be the divisor defined by our system of m points $p_1, \cdots, p_m$ on S,

$$D = \sum_{l=1}^{m} p_l,$$

where $m > p$. Then $\deg(D) = m$, and Riemann's inequality in Theorem 10.3 gives us

$$\dim L(D) \geq m - p + 1 \geq 2,$$

which again shows that there are nonconstant meromorphic functions with at least simple poles at the points p_l.

A few years after Riemann published his Abelian functions paper with the inequality in Theorem 10.3, Gustav Roch (1839–1866) published a significant improvement on this result. We will formulate his result below, but first we need to introduce one additional concept which is critical for Roch's work.

Suppose that α is a meromorphic one-form on S. For instance, since we know there are nonconstant meromorphic functions f on S, simply taking the exterior derivative df will be an example of such a meromorphic one-form. The one-form α will have poles and zeros with multiplicities and will define a divisor K, which we will call a *canonical divisor* on S. Suppose that α_1 and a_2 are any two meromorphic one-forms on S, then one can form the quotient α_1/α_2 (do this in local coordinates), and this quotient yields a nonconstant meromorphic function f. Let K_1 and K_2 be the divisors (canonical) of α_1 and α_2, then we have $\alpha_1 = f\alpha_2$, and it follows that

$$K_1 = K_2 + (f),$$

which implies that

$$\deg(K_1) = \deg(K_2) + \deg((f)),$$

but as we noted earlier, $\deg((f)) = 0$, and hence $\deg(K_1) = \deg(K_2)$.

Thus the canonical divisors on S all have the same degree and are considered equivalent in the sense that any two of them differ from each other by the divisor of a meromorphic function (this is referred to as *linear equivalence* in the literature). One refers to "the canonical divisor K" on S as being any divisor associated to a meromorphic one-form in this fashion.

An important ingredient in understanding the proof of the Riemann–Roch theorem is the following lemma, which provides a link between the canonical divisor and the genus of the surface S.

Lemma 10.4 *Let K be the canonical divisor on the Riemann surface S with genus p, then*

$$\deg(K) = 2 - 2p.$$

There are many different proofs of this, going back to Roch, but we won't include this in our discussion here. See, for instance, Weyl's book from 1913 [241]. We give a proof of this in our discussion of the modern version of the Riemann–Roch theorem in Sect. 14.7.

We can now formulate the Riemann–Roch theorem, which replaces the equality in Theorem 10.3 with an equality, making the result much more precise.

Theorem 10.4 (Riemann–Roch theorem) *Let S be a compact Riemann surface of genus p, and let D be a divisor on S, then,*

$$\dim L(D) - \dim L(K - D) = \deg(D) - p + 1, \tag{10.12}$$

where K is the canonical divisor on S.

We see here that Roch supplied the missing information in Riemann's inequality, and the canonical divisor plays a critical role.

We note that the left-hand side of (10.12) involves dimensions of vector spaces of functions and the right-hand side involves topological invariants of the divisor and the Riemann surface. In Sect. 14.7, we will see how this theorem evolved in the twentieth century to higher-dimensional manifolds and settings and became an important part of complex geometry relating analysis and topology on complex manifolds. Weyl's 1913 book on Riemann surfaces [241] has a complete proof of the results above as well as many modern references (e.g. Farkas and Kra [72], and the lecture notes by Bers [16] and Gunning [92]).

Chapter 11
Complex Geometry at the End of the Nineteenth Century

11.1 Klein and Lie

In the preceding sections of this Part III of the book, we have seen how the complex plane evolved into the concept of a Riemann surface and how the special class of holomorphic functions began to play an important role in analysis. There were several other significant ideas which arose in the nineteenth century which play an important role in complex geometry.

The first of these was the development of projective geometry and more specifically the notion of projective space, a generalization of classical Euclidean space which evolved over numerous decades of the nineteenth century. This was described in Chap. 3 in Part II. For an outstanding historic reference to this development, we recommend Felix Klein's beautiful lectures on non-Euclidean geometry from the end of the nineteenth century, which were published in 1928 [123]. In modern complex geometry, complex n-dimensional projective space $\mathbf{P}_n(\mathbf{C})$ plays a very important role.

At the end of the nineteenth century all the ingredients had been developed which allowed Hermann Weyl (1885–1955) to develop in 1913 the first theory of manifolds in the important special case of abstract Riemann surfaces. We will discuss this in more detail towards the end of this chapter. We now discuss some fundamental developments that preceded Weyl's work.

The first development, the theory of transformation groups, has become more well-known under its modern appellation of *Lie groups*. The first study of Lie groups arose as transformation groups of specific geometric spaces in various papers of Felix Klein (1849–1925) and Sophus Lie (1842–1899). For instance, in 1871 they wrote a joint paper [124], which referred to two earlier papers each of them had written that dealt with transformation groups on quite specific geometric spaces [120, 144]. In 1872 Klein wrote his famous *Erlangen Program* paper [121], in which he outlined, among other things, the role he foresaw for transformation groups, or more generally, continuous groups and their subgroups (in particular discrete subgroups) to play in geometry. This turned out to be a very significant paper, and the study of the action

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_11

of Lie groups on manifolds became an important topic in the twentieth century. In 1880 Lie published the first of his fundamental papers on what became the theory of Lie groups entitled "Theorie der Transformationsgruppen I"[1] [145]. In this paper he classified the Lie groups acting on Euclidean spaces of one and two dimensions and developed the tools of Lie algebras as a means of determining the classification. He gives a summary in this paper of all earlier references known to him at the time concerning this generic topic. One can leaf through the collected works of both Lie and Klein [122, 146], which are all available on-line today, to get a good overview of the development of transformation groups and their role in studying geometry in a wide variety of contexts.

11.2 The Uniformization Theorem for Riemann Surfaces

A very important relation between transformation groups and complex geometry came at the end of the nineteenth century in what became known as the *uniformization theorem* of Riemann surfaces, representing all compact Riemann surfaces as quotients of the three distinct simply-connected Riemann surfaces (the Riemann sphere, the complex plane, and the unit disc) by discrete groups of the biholomorphic automorphisms of these spaces. Let us give a brief summary of this important work, which became a role model for many similar questions in higher dimensions.

The history of non-Euclidean geometry has been well documented (see the classical treatment by Klein [123], for instance). In the nineteenth century there were various discoveries of geometries that were not Euclidean with abstract axiomatic systems which had variations on the parallel axiom and, more particularly, specific *models* of a given geometry. Here we will only mention that the developments of complex geometry in the nineteenth century led to three very specific models of the three types of geometries that have evolved.

Namely, first the complex plane $\mathbf{C}$ with its Euclidean metric

$$ds^2 = dzd\overline{z} = dx^2 + dy^2$$

is a model for the classical Euclidean plane *plane geometry*, with its geodesics being the classical straight lines, and the Euclidean translations and rotations being given by $z \mapsto z + a$, and rotations $z \mapsto ze^{i\theta}$.

The second case of *elliptic geometry* is represented in terms of the two-sphere S^2, which can be described in complex terms as the Riemann sphere, that is, the one-point compactification of the complex plane $\overline{\mathbf{C}} = \mathbf{C} \cup \infty$, which is biholomorphic to one-dimensional complex projective space $\mathbf{P}_1(\mathbf{C})$. If we let the complex plane be the standard coordinate chart (the complement of the point at infinity) with the metric

[1] "Theory of transformation groups".

$$ds^2 = \frac{dz d\overline{z}}{(1 + |z|^2)^2} = \frac{dx^2 + dy^2}{(1 + (x^2 + y^2))^2}$$

then the geodesics are the great circles, and the transformation group of (holomorphic) isometries is the orientation-preserving rotations of the sphere $SO(3) \cong PSU(2)$, in terms of real and complex coordinates, respectively. As a model for non-Euclidean geometry, it is necessary to consider the two-sphere with antipodal points identified (which gives two-dimensional real projective space, so that through any two points there is only one geodesic joining them, one of the axioms of all the geometries). This projective space structure doesn't preserve the complex structure.

Finally, we have the very important case of *hyperbolic geometry*, which can be modeled in terms of the Poincaré disk Δ, which is the unit disk in the complex plane

$$\Delta := \{z \in \mathbf{C} : |z| < 1\},$$

equipped with the Poincaré metric

$$ds^2 = \frac{dz d\overline{z}}{(1 - |z|^2)^2} = \frac{dx^2 + dy^2}{(1 - (x^2 + y^2))^2}.$$

Here the geodesics are the arcs of circles in the unit disc which have endpoints on the unit circle and which are orthogonal to the unit circle at those points. The transformation group of holomorphic isometries is $\mathrm{SU}(1, 1)$, which can be represented as the set of Möbius transformations of the form

$$z \mapsto e^{i\theta} \left(\frac{z - a}{-\overline{a}z + 1} \right), \theta \in \mathbf{R}, |a| < 1.$$

As is well known, these three two-dimensional models of Euclidean and non-Euclidean geometry $\mathbf{C}$, $\overline{\mathbf{C}}$, and Δ, also give the complete classification of the connected and simply-connected complex manifolds of dimension one, which was finally proved satisfactorily at the beginning of the twentieth century by Henri Poincaré (1854–1912) [186] and Paul Koebe (1882–1945) [129]. This theorem is referred to in the literature as the *uniformization theorem.*[2] Moreover, any compact Riemann surface of genus 1 is equivalent to the quotient of the complex plane by a lattice (a complex torus of dimension one), and any compact Riemann surface of genus $g > 1$ is equivalent to Δ/Γ, where Γ is a properly discontinuous subgroup of $\mathrm{SU}(1, 1)$, the automorphisms of the unit disc (see, e.g., the extensive survey paper by Lipman Bers [18]).

[2]See the very informative historical paper by Jeremy Gray [88] on the history of both the Riemann mapping theorem and the uniformization theorem.

11.3 Point Set and Algebraic Topology

The final developments of the nineteenth century critical for complex geometry primarily concerned the development of topology, both point set topology and algebraic topology, and finally the abstraction of the notion of a manifold.

First came the development of *set theory* by Georg Cantor (1845–1918) in the 1870s, which led to the development of abstract topological spaces in the early twentieth century. Cantor's work turned out to be revolutionary for all of mathematics as well as leading to the famous continuum hypothesis and fundamental questions in the foundations of mathematics which we won't discuss here (see, for instance, the collected works of Cantor [30] with its interesting introduction by Ernst Zermelo (1871–1953), as well as modern surveys of this important topic).

This led to the development of *abstract topological spaces*. Maurice Fréchet (1878–1973) was the first to formulate an abstract topological space [74] (he used the notion of spaces of type (L), which had axiomatically sequences of elements which either converged or didn't); and a few years later more general notions of a topological space were formulated by Felix Hausdorff (1868–1942) [97] using axioms of neighborhoods as the fundamental notion, including, in addition, his Hausdorff separation axiom. Then in 1922 Casimir Kuratowski (1896–1980) [131] provided the most general theory of topological spaces (using axioms concerning closed sets as a basis for the theory). Today we use axioms for open sets as the basis for the theory of topological spaces, which is equivalent to Kuratowski's theory.

Second was the development of the *algebraic topology of manifolds*. This started with the work of Riemann on Riemann surfaces (as discussed in Chap. 10), where he developed the notion of connectivity for Riemann surfaces. This was extended by Enrico Betti (1823–1892) in 1871 [19], who generalized Riemann's connectivity for two-dimensional manifolds to what are now called Betti numbers of higher-dimensional manifolds. Finally, Poincaré, in a fundamental series of papers at the end of the nineteenth century, formulated the fundamental principles of what has become known as *algebraic topology* (see the collection of papers on topology in Volume VI of Poincaré's collected works [187] and in particular the translation of Poincaré's topology papers into English by John Stillwell [188] with its lucid introduction to the whole topic). In these papers Poincaré considered manifolds which were smooth submanifolds of Euclidean space of any dimension with or without boundary, and if with boundary, the boundaries were piece-wise smooth, all defined in terms of defining functions in the ambient space.

11.4 Weyl's Book, *Die Idee der Riemannschen Fläche*, in 1913

The final geometric development of this time period concerns the creation of the notion of an *abstract manifold* in a mathematically satisfactory way. In Riemann's original paper concerning higher dimensional manifolds [201], he discussed this concept in a philosophical but not technical manner. For Riemann at the time it was simply something (not defined) which had local coordinate charts with suitable transition functions, and he worked from there.

The fundamental creation of an abstract manifold (with topological, differentiable or complex structures) was taken by Hermann Weyl in the first edition of his famous book on Riemann surfaces (*Die Idee der Riemannschen Fläche* [241]). He formulates for the first time in a rigorous manner the notion of an abstract topological manifold in the two-dimensional setting. He notes that he could work in more dimensions, but he was concerned with a new way of looking at Riemann surfaces. He describes a (two-dimensional) manifold as any set M with a set of neighborhoods satisfying suitable axioms and such that for each point $p \in M$ there is a neighborhood U of p which is homeomorphic to a disc in $\mathbf{R}^2$.[3]

The key here is that he starts with an abstract set, and this would have been possible only after Cantor created set theory in a manner that could be used in all parts of mathematics. He makes a further assumption to those made above, namely that the manifold is triangulated.[4] He goes on to define a Riemann surface to be a triangulated topological manifold which has local coordinate systems which map to discs in the complex plane whose overlap transformations are holomorphic. He makes a point that algebraic topology will play an important role in the theory of manifolds (hence the use of triangulations), noting the earlier work of Riemann and Poincaré. The notion of a triangulation of a surface went back at least to Euler's first description of an Euler characteristic of a surface of a two-sphere in $\mathbf{R}^3$, as we described in Sect. 10.2. Weyl worked in the context of a real two-dimensional surface, and his triangulation consisted of a covering of the surface with a disjoint union of (homeomorphic copies of) open triangles (faces), open line segments (edges), and point (vertices), where the edges and vertices formed the boundaries of the triangles in a natural manner. In Fig. 11.1 we see how Weyl introduced triangulation in his 1913 book.

His second edition [242] uses the now more common version of being a topological space (set with neighborhoods satisfying axioms) which is Hausdorff and has a countable basis for the topology.[5]

Part IV of this book concerns itself with differentiable, real-analytic, and complex manifolds, all of which are twentieth-century generalizations of Weyl's definition of a Riemann surface in 1913.

We close by quoting from Weyl's 1913 book about his new way of looking at what has become known as the theory of manifolds.

---

[3] Weyl had the key to an abstract topological space here, but he did not pursue it further.

[4] We discuss this triangulation hypothesis in more detail in the Introduction to Part IV of this book, which immediately follows this section.

[5] A topological space is *Hausdorff* if for any two distinct points x and y there are two open sets $U \ni x$ and $V \ni y$ such that $U \cap V = \emptyset$ and is *second countable* or has a *countable basis* if there is a countable set of open sets which generates (by countable unions) all open sets.

22 Begriff und Topologie der Riemannschen Flächen.

Punkte 1:0:0, 0:1:0, 0:0:1 werden als die drei **Ecken** des Dreiecks zu bezeichnen sein; die drei **Kanten** desselben werden durch $\xi_1 = 0$, bezw. $\xi_2 = 0$, bezw. $\xi_3 = 0$ geliefert.

Die Forderung der Möglichkeit einer Triangulation werden wir nun [im engsten Anschluß an Brouwers fundamentale Arbeiten[1]) und in einiger Übereinstimmung mit der in der Enzyklopädie entwickelten Dehn-Heegaardschen Theorie[2])] so zu fassen haben. Es seien auf einer Mannigfaltigkeit $\mathfrak{F}$ endlich- oder unendlichviele Dreiecke Δ (die „Elementardreiecke" der Triangulation) definiert, sodaß jeder Punkt der Mannigfaltigkeit mindestens einem der Dreiecke Δ angehört und außerdem folgende Bedingungen erfüllt sind:

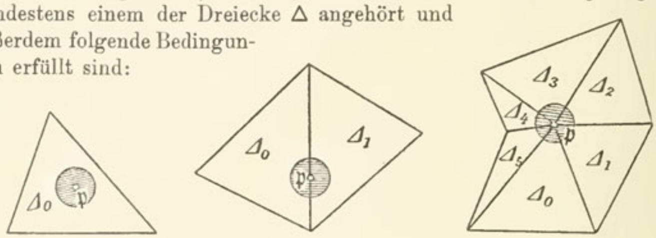


Fig. 1. Innerer Punkt, Kantenpunkt, Eckpunkt einer Triangulation.

1) Ist $\mathfrak{p}$ ein dem Δ-Dreieck Δ_0 angehöriger, nicht auf den Kanten von Δ_0 gelegener Punkt, so gehört weder $\mathfrak{p}$ noch irgend ein Punkt einer gewissen Umgebung von $\mathfrak{p}$ zu einem der von Δ_0 verschiedenen Dreiecke Δ.

2) Ist $\mathfrak{p}$ auf einer Kante k von Δ_0 gelegen, jedoch kein Eckpunkt von Δ_0, so gibt es ein weiteres Δ-Dreieck Δ_1, dem $\mathfrak{p}$ außerdem noch angehört; Δ_0, Δ_1 haben alle Punkte auf k, aber keine weiteren gemein; weder $\mathfrak{p}$ noch irgend ein Punkt einer gewissen Umgebung von $\mathfrak{p}$ gehört zu einem von Δ_0 und Δ_1 verschiedenen Δ-Dreieck.

3) Ist $\mathfrak{p}$ Eckpunkt eines Dreiecks Δ, so gibt es endlich viele Δ-Dreiecke: $\Delta_0, \Delta_1, \cdot\cdot, \Delta_h$, denen der Punkt $\mathfrak{p}$ angehört; in allen diesen Dreiecken ist $\mathfrak{p}$ ein Eckpunkt, und sie hängen in der Weise in einem einzigen Zykel zusammen, daß Δ_0 mit Δ_1 genau eine Kante, Δ_1 mit Δ_2 eine Kante, $\cdot\cdot\cdot$, schließlich Δ_h wieder mit Δ_0 eine Kante gemein hat[3]); weder $\mathfrak{p}$ noch irgend ein Punkt einer gewissen Umgebung von

1) S. namentlich L. E. J. Brouwer, Über Abbildung von Mannigfaltigkeiten, Math. Ann. Bd. 71 (1912), S. 97 ff. Ferner J. Hadamard „Sur quelques applications de l'indice de Kronecker" im 2. Bande von J. Tannery, Introduction à la théorie des fonctions d'une variable, 2. Aufl. (Paris 1910). Die Auffassung einer beliebigen geschlossenen Raumfläche als eines Polyeders findet sich klar entwickelt bei Möbius, namentlich in den nachgelassenen Papieren „Zur Theorie der Polyeder und der Elementarverwandtschaft" (1861), Werke Bd. II, S. 517 ff.

2) Enzyklopädie III AB 3, S. 153 ff.

3) Diese Konfiguration bezeichnen wir als einen **Dreiecks-Stern**.

Fig. 11.1 Page 22 of Weyl's 1913 book on Riemann surfaces [241]

> Eine solche strenge Darstellung, die namentlich auch bei Begründung der fundamentalen, in die Funktionentheorie hineinspielenden Begriffe und Sätze der Analysis situs sich nicht auf anschauliche Plausibilität beruft, sondern mengentheoretisch exakte Beweise gibt, liegt bis jetzt nicht vor. Die wissenschaftliche Arbeit, die hier zu erledigen blieb, mag vielleicht als Leistung nicht sonderlich hoch bewertet werden. Immerhin glaube ich behaupten zu können, daß ich mit Ernst und Gewissenhaftigkeit nach den einfachsten und sachgemäßesten Methoden gesucht habe, die zu dem vorgegebenen Ziele führen; und an manchen Stellen habe ich dabei andere Wege einschlagen müssen als diejenigen, die in der Literatur seit dem Erscheinen von C. Neumanns klassischem Buche "Über Riemanns Theorie der Abelschen Integrale" (1865) traditionell geworden sind.[6]

[6]"Such a rigorous presentation, which, namely by establishing the fundamental concepts and theorems in function theory and using theorems of the analysis situs which don't just depend on intuitive plausibility, but have set-theoretic exact proofs, does not exist. The scientific work that remains to be done in this regard may perhaps not be particularly highly valued. But, nevertheless, I believe I can maintain that I have tried in a serious and conscientious manner to find the simplest and most appropriate methods that lead to the asserted goal; and at many points I have had to proceed in a different manner than that which has become traditional in the literature since the appearance of C. Neumann's classical book about *Riemann's theory of Abelian Integrals*."

Part IV
Twentieth-Century Embedding Theorems

Introduction

In the first three Parts of this book, we have seen how the differential-geometric study of curves and surfaces in Euclidean two- and three-dimensional space in the seventeenth and eighteenth centuries evolved into several threads of abstraction in the nineteenth century, leading to the modern theory of manifolds: topological, differentiable, and complex manifolds, to mention the prominent families. In addition, the ideas of Riemannian geometry from the work of Gauss and Riemann became part of mainstream mathematics as well.

A major problem arose in the early twentieth century: Could these abstractions of various kinds be realized as submanifolds of Euclidean space of some dimension, and thus, in effect, philosophically return to the extrinsic geometry of the eighteenth century that was looked at in some detail in Part I of this book? To be sure, this would not revert to the two- and three-dimensional problems studied at that time.

This became the question of embedding a manifold with specific properties and a specific dimension into a Euclidean space of some higher dimension, preserving these same properties. For some cases, such as compact complex manifolds, it was necessary to consider embedding into a suitable projective space, a natural compactification of Euclidean space.

The first such embedding questions arose in the late nineteenth century in the context of locally defined Riemannian manifolds, following up on the work of Riemann from a few years earlier. After the first formulation of the notion of a globally defined abstract manifold by Weyl in 1913, the embedding question extended to the full range of geometries that evolved during the first decades of the twentieth century.

Part IV of this book provides a detailed survey of five different embedding theorems that were formulated and proved in the mid-twentieth century. These five theorems are all philosophically very similar, as we shall see: a given manifold of a certain class could be embedded as a submanifold of Euclidean or projective space, and the embedding provided a characterization of all such submanifolds. However, these theorems are technically very diverse in the nature of the mathematical tools utilized for their proofs. In fact, the mathematical ideas that evolved during this period of time and which were used in proving these theorems consisted of certainly many of the most important developments in geometry of the first half of the twentieth century. In our survey of these theorems, we will introduce a number of these concepts and show how they link together to prove these theorems. In the remainder of this introduction, we will introduce these theorems and we will describe some of the basic language and concepts for embedding theorems in general.

In 1936, Hassler Whitney (1907–1989) initiated the new era of what are now called *embedding theorems*. As we discussed above, the overall question that he raised was the following: Given a particular class of manifolds with a given geometric structure, could any such manifold be embedded into a Euclidean space as a submanifold which inherits the given geometric structure from the Euclidean space?

In summary form, we formulate these fundamental theorems here.

- In 1936, Hassler Whitney proved that any differentiable (C^∞) manifold can be embedded as a closed differentiable submanifold of $\mathbf{R}^N$ [245].
- In 1954, Kunihiko Kodaira proved that any compact complex manifold with a Hodge metric can be embedded as a complex submanifold of complex projective space $\mathbf{P}_N(\mathbf{C})$ [127].
- In 1956, John Nash showed that any smooth Riemannian manifold can be isometrically embedded onto a smooth submanifold of an open subset of $\mathbf{R}^N$ equipped with the induced Riemannian metric from $\mathbf{R}^N$ [164].
- In 1956, Reinhold Remmert announced the result that any Stein manifold can be holomorphically embedded as a closed complex submanifold of $\mathbf{C}^N$, and this was subsequently proved independently by Raghavan Narasimhan in 1960 and Errett Bishop in 1961 [20, 162, 192].
- In 1958, Hans Grauert proved that any real-analytic manifold can be real-analytically embedded as a closed real-analytic submanifold of $\mathbf{R}^N$ using methods of several complex variables, and Charles B. Morrey proved in the same year the special case of compact real-analytic manifolds using methods of partial differential equations [85, 159].

In the theorems listed above, N is a sufficiently large positive integer, depending on the given manifold. Each of these theorems gave a complete characterization in the abstract setting of the submanifolds described.

In the following chapters, we will give some background and a discussion of the key ideas that went into the proofs of these theorems. Each of the results mentioned above generated numerous refinements, new proofs, and special variations of these theorems over the last half-century, and we do not attempt to follow these developments up to the present time. We indicate some advances in proofs in certain cases, but we do not attempt to be conclusive on subsequent developments after these fundamental theorems were proven.

We start with the work of Whitney and Nash in Chaps. 12 and 13, which we group together as they use fundamental concepts in real analysis, including, in particular, developments in Lebesgue measure theory and functional analysis (Banach spaces, etc.) which were also developed in the early part of the twentieth century. These have become quite standard tools for most students of mathematics these days, and we assume the reader is familiar with these concepts.

In the final two chapters of the book, we outline the proofs of the complex geometry-oriented embedding theorems. There were significant developments in twentieth century complex analysis and geometry which may be less familiar to the reader, and we have included sections which outline the principal developments in the theory of several complex variables and the theory of compact complex manifolds which are essential tools in the proofs of these embedding theorems, all of which involve complex analysis and complex geometry in one way or another.

As we discussed in Sect. 3.3, Grassmann and Plücker first studied the family of manifolds now known as projective space or Grassmannian manifolds. None of these examples, due to Plücker, Grassmann and others (except projective space itself) are defined as submanifolds of either Euclidean or projective space. Plücker showed in his book [183] that the set of planes in $\mathbf{C}^4$ (which is the same as the set of lines in the projective space $\mathbf{P}_3(\mathbf{C})$) is equivalent to a quadric[1] in $\mathbf{P}_5(\mathbf{C})$. Here, we mean homeomorphic (and in this case biholomorphic, achieved by an algebraic equivalence). This is the first instance of what will later be known as the *Kodaira embedding theorem*, which we will discuss later in Chap. 14, and this is one of the first examples of the type of embedding theorem we are discussing in this Part IV of the book.

We now want to discuss two other developments concerning the general structure of manifolds which followed up on the work of Hermann Weyl.

Weyl included in the 1913 edition of his book [241] an additional hypothesis that a Riemann surface should be triangulated, whereas in his 1955 edition of the same book [242], he made the hypothesis that a Riemann surface should have a countable basis for its topology, and he proved later in this same edition that one did not need this countable basis hypothesis.[2] What had happened in the meantime was that T. Radó in 1925 [191] asked the question as to why Weyl had made the hypothesis of triangulation, as that did not seem in the spirit of the philosophy of Weyl's topological space formulation of a Riemann surface. Radó noted that for a connected two-dimensional topological manifold, having a triangulation was equivalent to having a countable basis for its topology. He pointed out that various contemporary authors sometimes assumed that manifolds had a countable basis and sometimes they did not.

He posed the question: Can a Riemann surface, defined as Weyl did, but without the triangulation hypothesis, be triangulated? He answered this question in his paper affirmatively, a landmark result in the theory of Riemann surfaces. In the introduction to this paper, he noted that the "concrete" Riemann surfaces described by Riemann, and used by mathematicians for decades before Weyl's book came out, were all triangulated. Namely, they were all finite or countable sheets over the complex plane.

---

[1] A hypersurface defined by a quadratic equation in homogeneous coordinates. This equivalence is true for the real numbers and other fields as well.

[2] Recall that a topological space X is said to have a *countable basis (base) for its topology* if there is a countable family C of open sets in X such that every open set in X is a union of open sets from C.

He also included in his paper an example of a two-manifold which does *not* have a countable basis for its topology, and he attributed this example to a colleague of his, Heinz Prüfer. He further noted that the insight that he obtained from this example was critical for the proof of his theorem. He alluded also to the "long line" paper of Alexandroff, which had appeared a year earlier [6]. This was an example of a one-dimensional topological manifold without a countable basis (the long line is essentially an uncountable number of copies of the interval (0,1] glued together appropriately).

Following upon this thread of ideas, Calabi and Rosenlicht showed in 1953 [29] that there exist two-dimensional complex manifolds without a countable basis. Thus, one-dimensional complex manifolds (Riemann surfaces) differ from higher-dimensional complex manifolds in this important topological sense, just as one-dimensional complex function theory differs from the theory of functions of several complex variables due to Hartogs's theorem [95]. This theorem asserts, among other things, that there are no isolated singularities for holomorphic functions of two or more complex variables. Hartogs's theorem and its consequences are discussed in more detail in Sect. 15.2.

The three basic target spaces for our embedding theorems are $\mathbf{R}^n$, $\mathbf{C}^n$, and $\mathbf{P}^n$, and these all have a countable topology. If a manifold M is to be embedded into any of these spaces as a closed submanifold homeomorphic to M (as we will be considering in this book), then M must necessarily have had a countable basis for its topology. In this book, we will assume from here on that all manifolds we consider have a countable basis for its topology. As we have seen above, not all manifolds have this property.

The notion of paracompactness of a topological space was introduced in 1944 by Dieudonné [55]. A topological space S is *paracompact* if any covering of S by open sets has a locally finite subcover, i.e., any point $x \in S$ is an element of only a finite number of the open sets in the subcover. Dieudonné showed that a connected topological manifold is paracompact if and only if it has a countable basis. This will allow us to use partitions of unity for the manifolds we consider, which is a very useful concept, as we will see later.

A final step in the creation of a suitable formulation of differentiable or complex manifolds during this period is the important paper of Veblen and Whitehead in 1931 [225], which was followed up by their monograph [226] in 1932. This became the framework for the theory of manifolds for the twentieth century, and we briefly present here the formal language of manifolds as introduced by them, all of which is a natural generalization of the work of Weyl from 1913.

We define a *topological manifold M* of dimension n to be a topological space with a covering of open sets U_α such that there are homeomorphisms

$$h_\alpha : U_\alpha \to B_\alpha \subset \mathbf{R}^N,$$

where B_a is an open ball in $\mathbf{R}^n$ with coordinates $(x_1, \ldots, x_n)$, which is referred to as a *local coordinate system* or simply *local coordinates* for the manifold M at any point of U_α. The open sets U_α and the mappings h_α are referred to as *coordinate charts* for M. We define the *transition functions*

$$h_{\alpha\beta} := h_\alpha \circ h_\beta^{-1} : h_\beta(U\alpha \cap U_\beta) \to h_\alpha(U_\alpha \cap U_\beta),$$

defined for all nontrivial intersections $U_\alpha \cap U_b \neq \emptyset$.

The coordinate charts can often be chosen to have various kinds of smoothness properties of the transition functions, and that is how we define the types of manifolds we are interested in. *A priori* the transition functions are homeomorphisms, and this is then the definition of a topological manifold. If we require the transition functions to be C^k diffeomorphisms (i.e., the transition functions and their inverses have k continuous derivatives), then we have a C^k manifold. If the transition functions and their inverses are C^∞ functions (an infinite number of continuous derivatives), then we have a C^∞ manifold, and we will simply say in this book *differentiable manifold* for a C^∞ manifold. We define a *real-analytic manifold* in the same way, where we require the transition functions and their inverses to be real-analytic functions.

If the transition functions are required to have positive Jacobian matrices, then the manifold is said to be *orientable*. If n is even, say $n = 2m$, and if $\mathbf{R}^{2m}$ is given a complex structure isomorphic to $\mathbf{C}^m$, and if the transition functions are biholomorphic,[3] then the manifold is said to be a *complex manifold*, and there are numerous other examples. This last example is an explicit generalization of Weyl's definition of a Riemann surface. The main purpose of Veblen and Whitehead's paper and monograph is to illustrate how to formulate and study Riemannian geometry on a differentiable manifold, and in doing so, they provided the background for the study of a large class of manifolds of different types, an important elaboration on the original work of Weyl from 1913.

One final note here is that this family of manifolds as described in the previous paragraph and put forward by Veblen and Whitehead is completely consistent with Riemann's definition of a manifold in 1854 [200] and his definition of Riemann surfaces in 1857 [201]. In his 1854 paper on manifolds and differentiable geometry, he explicitly used transition functions, but did not have a *set* where the coordinate charts U_α could be defined or located. In the case of Riemann surfaces in 1857, he did have the full notion of coordinate charts since the global manifold (the Riemann surface) in that case was described as sheets covering parts of the complex plane, and the U_α were subsets of the sheets.

We have seen many examples of manifolds, but the simplest class of manifolds are simply submanifolds of a domain in $\mathbf{R}^n$ or $\mathbf{C}^n$. To make things more precise, we say that a subset S of a domain $D \subset \mathbf{R}^n$ is a *differentiable submanifold* of D of dimension k if S is a closed subset of D such that, for each point $x \in D$, there is a neighborhood U of x in D and $n - k$ differentiable functions $f_1, \ldots, f_{n-k}$ defined in U such that

[3] A *biholomorphic mapping* is a holomorphic mapping with an inverse which is also holomorphic.

$$S|_U = \{x \in U: f_j(x) = 0, j = 1, \ldots, n-k\},$$

and where the Jacobian matrix

$$\frac{\partial(f_1,\ldots,f_{n-k})}{\partial(x_1,\ldots,x_n)}$$

has rank $n-k$ at each point of $S|_U$. We define real-analytic submanifolds in the same manner, and the definition of a *complex submanifold* of $\mathbf{C}^n$ has the same formulation, where one uses holomorphic functions in the place of differentiable or real-analytic functions.

At the close of this introduction, we come now to the Whitney embedding[4] theorem of 1936 [245], the first of the major embedding theorems we want to consider in this book. In Fig. 12.1 we see how Whitney formulated the first general embedding theorem. The first paragraph set the stage with two different definitions of manifolds, and then, in the first sentence of the second paragraph, we read:

> The first fundamental theorem is that the first definition is no more general than the second; any differentiable manifold may be imbedded in Euclidean space.

In this paper, Whitney discusses and proves embedding theorems of this type for $C^1, C^2, \ldots, C^\infty$ manifolds. We will concentrate on the C^∞ case, which we are calling differentiable manifolds. We will discuss more of what he did in the next chapter, but first, we want to formulate the more general embedding theorem problems that we consider in this book. We will be interested in the embedding of differentiable manifolds, complex manifolds, and real-analytic manifolds into real or complex Euclidean space or complex projective space, depending on the context.

Let M be an n-dimensional differentiable manifold, and let $f : M \to \mathbf{R}^N$ be a continuous mapping. Then, f is a *proper mapping* if the inverse image of any compact set $K \subset \mathbf{R}^N$ is compact in M. If M itself is compact, then any such f will, of course, be proper. But in general, not all such continuous mappings are proper.[5] An important lemma that is used in many of the embedding theorems, and which is not that difficult to prove, is that if f is proper, then the image $f(M)$ is closed in $\mathbf{R}^N$.

We will be considering mappings from a given n-dimensional manifold M into Euclidean space $\mathbf{R}^N$, and we want to establish some terminology that we will use throughout the remainder of the book (the same terminology would apply for a mapping from one manifold to another, but we choose to concentrate on this family of mappings, as it is the setting for most of the embedding theorems). A mapping

$$f = (f_1, \ldots, f_N): M \to \mathbf{R}^N$$

---

[4]The spelling "embedding" has now come to be standard instead of the spelling "imbedding" used by Whitney and several of his contemporaries; see [217].

[5]This is easy to see by simple examples of the form $f : \mathbf{R}^1 \to \mathbf{R}^1$, e.g., take f to be a branch of $\tan^{-1}$.

is *differentiable* if the components f_j of f are C^∞ functions on M (this means that they are infinitely differentiable in terms of any local coordinate system on M). If we consider any other type of smoothness (e.g., C^k) in a given context, then this will be made explicit at the time. A real-analytic mapping is defined in the same manner.

Let M be an n-dimensional differentiable manifold, then a differentiable mapping $f : M \to \mathbf{R}^N$ is said to be a *regular mapping* at $x \in M$ if the derivative df_x of the mapping f at x has maximal rank (in local coordinates at x the Jacobian matrix $\frac{\partial(f_1,\ldots,f_N)}{\partial(x_1,\ldots,x_n)}$ has maximal rank). We say that f is a *regular mapping on M* if f is regular at each point of M. If $f : M \to \mathbf{R}^N$ is a regular mapping on M and if $N \geq n$, then we say that f is a *differentiable immersion*, or often, simply an *immersion*, the smoothness being clear from the context. The notion of regularity and immersion is defined for a real-analytic mapping in the same manner.

We will also consider *holomorphic mapping*s of the form

$$f = (f_1, \ldots, f_N) : X \to \mathbf{C}^N,$$

where X is an n-dimensional complex manifold and the components f_j of the mapping f are holomorphic functions on X (i.e., holomorphic with respect to the local coordinate systems on X). Such a holomorphic mapping is said to be a *regular mapping* at $x \in X$ if, in a local coordinate system at x, say $(z_1, \ldots, z_n)$, the Jacobian matrix $\frac{\partial(f_1,\ldots,f_N)}{\partial(z_1,\ldots,z_n)}$ has maximal rank, and the mapping f is *regular on X* if it has maximal rank at each point of X. As before, a holomorphic mapping $f : X \to \mathbf{C}^N$ is a *holomorphic immersion* if f is regular on X, and $\dim_{\mathbf{C}} X \leq N$.

We note that we are using the same word "regular" to denote this maximal rank condition in all three contexts (differentiable, real-analytic, or holomorphic), and its use will be clear in context.

The implicit function theorem is valid in all three categories we are considering, and thus, if $f : M \to \mathbf{R}^N$ is a differentiable or real-analytic immersion, then the image of a neighborhood of each point $x \in M$ is a differentiable or real-analytic submanifold of a neighborhood of $f(x) \in \mathbf{R}^N$. Similarly, by the implicit function theorem for holomorphic functions, if $f : X \to \mathbf{C}^N$ is a holomorphic immersion, then f, restricted to a suitable neighborhood U of a given point $x \in M$, maps onto a holomorphic submanifold defined in a neighborhood of $f(x)$ in $\mathbf{C}^N$.

If f is a differentiable immersion and is a homeomorphism onto its image in $\mathbf{R}^N$, then it is an *embedding* into $\mathbf{R}^N$. The same definition of embedding can be made for real-analytic mappings into $\mathbf{R}^N$ or holomorphic mappings into $\mathbf{C}^N$. Finally, if f is a proper mapping and an embedding (in any of the three categories), then its image is a closed submanifold of $\mathbf{R}^N$ or $\mathbf{C}^N$. This final result, again not very difficult to prove, is used in almost all of the embedding theorems we discuss.

In Chap. 14, we will discuss the somewhat more subtle nature of mappings of complex manifolds into complex projective spaces, since mappings defined by vector-valued functions, which we have used up till now, are not adequate, as will then become apparent.

Chapter 12
Differentiable Manifolds

12.1 Introduction

In this chapter we want to describe Whitney's proof of the differentiable embedding theorem. First we outline two different tools which play an important role in this theorem and which will be useful later as well.

The first tool is that of *Lebesgue measure theory*, which plays such an important role in many branches of mathematics today. This was the creation of Henri Lebesgue (1875–1941) in 1902 [136] (see also his important monograph from 1904 [137]). The study of Lebesgue measure and integration theory is now a standard subject in undergraduate curricula. All we need from this theory is the concept of a set of measure[1] zero in $\mathbf{R}^n$, which Whitney and others used very effectively in their embedding theorems. Let's define a ball of radius r at a point $x \in \mathbf{R}^n$ to be

$$B(x, r) := \{y \in \mathbf{R}^n : |x - y| < r\}.$$

Consider any set $S \subset \mathbf{R}^n$, then we say that S has *measure zero* if, for any $\varepsilon > 0$, there is a countable cover of S by balls $B_i = B(x_i, r_i)$ such that

$$\sum_i \mathrm{Vol}(B_i) < \varepsilon,$$

where $\mathrm{Vol}(B)$ is the volume of a ball $B \subset \mathbf{R}^n$. We note that if B is a ball of radius r, then $\mathrm{Vol}(B) = K_n r^n$, where K_n is a constant[2] depending only on the dimension n. An important property of sets of measure zero is that if $S \subset \mathbf{R}^n$ has measure zero, then the *complement of* S ($\mathbf{R}^n - S$) is dense in $\mathbf{R}^n$, which is easy to verify.

We now want to mention a key measure-theoretic result which is very useful in the various embedding theorems, and which is quite easy to prove. Whitney describes

[1] Here we refer only to Lebesgue measure in $\mathbf{R}^n$.

[2] In fact, $K_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_12

this explicitly in Sect. 17 of his paper ([245], pp. 660–661). Namely, let

$$f : \mathbf{R}^n \to \mathbf{R}^n$$

be a C^1 mapping, and let S be a set of measure zero in $\mathbf{R}^n$, then the image set $f(S)$ under the mapping f has measure zero in $\mathbf{R}^n$.

To prove this, simply show first that for any compact set $K \subset \mathbf{R}^n$, the measure of $f(K \cap S)$ is 0, and since S can be covered by a countable number of such compact sets K, and the measure of a countable union of sets of measure zero is also measure zero, then $f(S)$ would necessarily have measure zero. Now fix such a compact subset K of $\mathbf{R}^n$, then there exists a Lipschitz estimate of the form

$$|f(x) - f(y)| \leq M|x - y|, \text{ for all } x, y \in K$$

for some constant M which depends on estimates of the first derivatives of the mapping f on the compact set K. Then if B_i are balls of radius r_i in $\mathbf{R}^n$ such that $\{B_i\}$ cover $S \cap K$ and

$$\sum_i \mathrm{Vol}(B_i) < \varepsilon,$$

then, for any two points x, y in B_i, we see that

$$|f(x) - f(y)| \leq M|x - y| < Mr_i.$$

It follows that

$$f(S \cap K) \subset \bigcup_i f(B_i) \subset \bigcup_i B(f(x_i), Mr_i),$$

and

$$\begin{aligned} \text{meas } f(S \cap K) &< \sum_i K_n M^n r_i^n, \\ &< M^n \sum_i \mathrm{Vol}(B_i), \\ &< M^n \varepsilon, \end{aligned}$$

and hence $f(S \cap K)$ has measure zero since ε was arbitrarily small.

A simple corollary of this which we will use later is to note that if $S \subset \mathbf{R}^m$ and f is a C^1 mapping,

$$f : \mathbf{R}^m \to \mathbf{R}^n, \text{ where } m < n,$$

then meas $(S) = 0$ in $\mathbf{R}^n$. Consider the inclusion $i : \mathbf{R}^m \to \mathbf{R}^n$ given by

$$i(x) = (x, 0) \in \mathbf{R}^m \times \mathbf{R}^{n-m}.$$

Since $i(\mathbf{R}^m)$ has measure zero in $\mathbf{R}^n$ (it is a lower-dimensional subspace!), then defining

$$F : \mathbf{R}^n \to \mathbf{R}^n \text{ by } F(x, y) = (f(x), y),$$

we see that $f(S) = F(i(S))$, which must have measure zero since $i(S)$ has measure zero.

A second tool that arose in the first half of the twentieth century, and which has been used in a wide variety of contexts, is the concept of a *cut-off function*. We will give here a simple example of how this is useful in Whitney's proof of the embedding theorem, but the concept has many broad applications (smoothing operators in partial differential equations, the theory of distributions, and in the general theory of differential topology, as a few examples).

Let us define the function

$$\chi(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ e^{-\frac{1}{x^2}} & \text{if } x > 0. \end{cases}$$

The function χ has the properties that it is C^∞ on $\mathbf{R}$ and at the point $x = 0$, χ is *not* real-analytic. Whitney has a similar construction in his 1936 paper [245]. This type of example shows that the class of C^∞ functions is different from the class of real-analytic functions, and was likely known in the nineteenth century.[3] However, the study of C^∞ functions and mappings was primarily developed in the twentieth century (e.g., by studying manifolds of varying degrees of smoothness as initiated in the book by Veblen and Whitehead [226]).

Consider now three concentric balls in $\mathbf{R}^n$, centered at the origin:

$$B(0, 1) \subset B(0, 2) \subset B(0, 3).$$

We want to find a C^∞ function $\varphi(x)$, symmetric about the origin such that:

$$\begin{aligned} &\varphi(x) \equiv 1 \text{ for } |x| \leq 1, \\ &\varphi(x) \in (0, 1) \text{ for } 1 < |x| < 2, \\ &\varphi(x) \equiv 0 \text{ for } |x| > 2. \end{aligned} \tag{12.1}$$

Such a function is called a *cut-off function* in this particular geometric setting. We simply define

$$\varphi(x) := \frac{\chi(2 + |x|)\chi(2 - |x|)}{\chi(2 + |x|)\chi(2 - |x|) + \chi(|x| - 1)}, \tag{12.2}$$

and one can verify easily that $\varphi(x) \in \mathbf{C}^\infty(\mathbf{R}^n)$ and satisfies the desired properties in (12.1). There are, of course, many such examples of cut-off functions.

[3] We have not been able to ascertain who first described such an example, but we recall that Weierstrass formulated in 1872 an example of a continuous function which is nowhere differentiable [230], so he or others at that time might have known such an example.

ANNALS OF MATHEMATICS
Vol. 37, No. 3, July, 1936

DIFFERENTIABLE MANIFOLDS[1]

BY HASSLER WHITNEY

(Received February 10, 1936)

INTRODUCTION

The main purpose of this paper is to provide tools of a purely analytic character for a general study of the topology of differentiable manifolds, and maps of them into other manifolds. A differentiable manifold is generally defined in one of two ways; as a point set with neighborhoods homeomorphic with Euclidean space E_n, coördinates in overlapping neighborhoods being related by a differentiable transformation,[2] or as a subset of E_n, defined near each point by expressing some of the coördinates in terms of the others by differentiable functions.[3]

The first fundamental theorem is that the first definition is no more general than the second; any differentiable manifold may be imbedded in Euclidean space. In fact, it may be made into an analytic manifold in some E_n. As a corollary, it may be given an analytic Riemannian metric. The second fundamental theorem (when combined with the first) deals with the smoothing out of a manifold. Let f be a map of any character (continuous or differentiable, without an inverse) of a differentiable manifold M of dimension m into another, N, of dimension n. (Either manifold might be an open subset of Euclidean space.) Then if $n \geqq 2m$, we may alter f as little as we please, forming a regular map F. (A map is *regular* if, near each point, it is differentiable and has a differentiable inverse.) Moreover, if $n \geqq 2m + 1$, F may be made (1-1). We show in Theorem 6 that if $n \geqq 2m + 2$, then any two regular maps f_0, f_1 of M into E_n are equivalent, in the following sense. $f_0(M)$ may be deformed into $f_1(M)$ by maps $f_t(0 \leqq t \leqq 1)$ so that the path crossed by the manifold is the regular map of an $(m + 1)$-dimensional manifold. Moreover, if $n \geqq 2m + 3$, and $f_0(M)$ and $f_1(M)$ are non-singular, so is the $(m + 1)$-manifold.

A fundamental unsolved problem is the following: *Can any analytic manifold be mapped in an analytic manner into Euclidean space?*[4]

---

[1] Presented to the Am. Math Soc. Sept. 1935. An outline of the paper will be found in Proc. Nat. Ac. of Sci., vol. 21 (1935), pp. 462–463.

[2] Differentiable manifolds have been studied for instance by O. Veblen and J. H. C. Whitehead, *The foundations of differential geometry*, Cambridge Tracts, 1932. An example of a differentiable (in fact, analytic) manifold is the manifold of k-planes through a point in n-space. See §24.

[3] Manifolds in E_n which are defined by the vanishing of a set of differentiable functions are of a special character; see H. Whitney, *The imbedding of manifolds* ⋯ , in the October 1936 issue of these Annals.

[4] This seems quite probable. It is proved for some special analytic manifolds in §§23–24.

645

Fig. 12.1 First page from Whitney's 1936 embedding paper for differentiable manifolds [245]. *Reprinted with the permission of the Annals of Mathematics*

We now want to formally state Whitney's embedding theorem, and then we will outline the key steps in the proof.

Theorem 12.6 (Whitney 1936 [245]) *Let M be an n-dimensional differentiable manifold, then there exists a proper differentiable embedding*

$$f : M \to \mathbf{R}^{2n+1}.$$

In Fig. 12.1 we see the first page of Whitney's embedding theorem paper. The proof of this theorem splits into several natural parts, which we discuss in the following sections.

12.2 The Local Immersion Approximation

The first step in the proof of the embedding theorem is to show that any differentiable mapping $f : M \to \mathbf{R}^N$ can be approximated near a given point of M by a mapping which is an immersion near that same point, provided that $N \geq 2n$. As we have seen earlier, any immersion of a manifold into Euclidean space is locally a differentiable embedding into the same Euclidean space, so our first step here will provide a local differentiable embedding into $\mathbf{R}^{2n}$, and we will see why the hypothesis that $N \geq 2n$ becomes an important part of the proof at this local level.

By restricting to a local coordinate system, we can formulate the following problem in Euclidean space. We consider a neighborhood U of $0 \in \mathbf{R}^n$ and any differentiable mapping

$$f : U \subset \mathbf{R}^n \to \mathbf{R}^N, N \geq 2n.$$

For instance, f could be simply a constant mapping. For convenience, we represent points in $\mathbf{R}^n$ and $\mathbf{R}^N$ as column vectors and we consider a perturbation of the given mapping f by a linear mapping represented by an $N \times n$ matrix A, i.e., let

$$F(x) := f(x) + Ax,$$

where A is to be determined.

Our goal is to choose A sufficiently small in size (e.g., $\|A\| < \varepsilon$), for any small ε, and such that

$$DF(x) = Df(x) + A$$

is regular (has rank n in this case) at each point of U. Here $DF(x)$ and $Df(x)$ are the $N \times n$ Jacobian matrices of the two mappings at the point $x \in U$. The way we do this is to choose A so that it avoids the cases where $DF(x)$ might have less than maximal rank.

Let $M(N, n)$ be the set of all $N \times n$ matrices. It is clear that as differentiable manifolds (or simply as vector spaces)

$$M(N, n) \cong \mathbf{R}^{Nn}.$$

Let

$$M_k(N, m) := \{A \in M(N, n) : \text{ rank } A = k\}.$$

We claim that $M_k(N, n)$ is a submanifold of $M(N, n)$ of dimension $k(n - k + N)$. This is easy to verify by observing that a local coordinate system for a point in $M_k(N, n) \subset M(N, n)$ can be described by matrices of the form

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix},$$

where A has rank k. We note that these dimensions increase from 0, for $k = 0$, to $(n - 1)(N + 1)$, for $k = n - 1$. Therefore,

$$\dim M_k(N, n) \leq (n - 1)(1 + N) \text{ for } k < n. \tag{12.3}$$

Now consider the mapping

$$\Phi_k : U \times M_k(N, n) \to M(N, n)$$

defined by

$$\Phi_k(x, B) := B - Df(x).$$

The dimension of the image space here is $\dim M(N, n) = Nn$. How large is

$$\dim(U \times M_k(N, n))?$$

By our estimate (12.3) above, we see that, for each $k = 0, \ldots, (n - 1)$,

$$\begin{aligned} \dim U \times M_k(N, n) &\leq n + (n - 1)(N + 1), \\ &\leq nN + (2n - N) - 1, \\ &< nN, \end{aligned}$$

since $N \geq 2n$.

Thus the images

$$\Phi_k(U \times M_k(N, n)) \subset M(N, n), k = 0, \ldots, n - 1,$$

all have measure zero, and their finite union must have measure zero as well. Hence we can choose a matrix A arbitrarily small and such that

$$A \notin \operatorname{Im} \Phi_k, \text{ for } k = 0, \ldots, n - 1.$$

For this choice of A it follows that

$$DF(x) = Df(x) + A$$

is regular for all $x \in U$. Thus F is an immersion on U, as desired.

12.3 Whitney's Embedding Theorem

Let M be a differentiable manifold of dimension n and consider an arbitrary differentiable mapping

$$f : M \to \mathbf{R}^N, \;\; N \geq 2n.$$

The next step in the proof of the embedding theorem is to show that f can be approximated on M by a mapping $\tilde{f}$ which is a *differentiable immersion* of M into $\mathbf{R}^N$. As we saw in the previous section it was useful to have $N \geq 2n$ in order to find an approximation of a local mapping by a local immersion, and we make the same assumption here.

To start this process, we need a convenient method of approximating mappings of a manifold M into Euclidean space $\mathbf{R}^N$. There are many such methods of approximating, but in this case we choose a simple pointwise approximation. Namely, let $\eta(x)$ be a continuous function defined on M such that $\eta(x) > 0$ for all $x \in M$. We say that a function $\tilde{f}$ is an *η approximation to f* if

$$|f(x) - \tilde{f}(x)| < \eta(x), \;\; x \in M,$$

where $|y|$ is the usual Euclidean norm of a vector in $\mathbf{R}^N$.

We now define a specific countable locally finite covering $\{U_i\}$ of the manifold M, where

$$h_i : U_i \to \mathbf{R}^n$$

is a coordinate mapping (diffeomorphism) onto the ball $B(0, 3)$. We define

$$V_i = h_i^{-1}(B(0, 2)),$$
$$W_i = h_i^{-1}(B(0, 1)),$$

and we require further that $\{W_i\}$ is also a countable locally finite covering[4] of M.

---

[4]It is easy to see that such a covering exists. Simply take any countable covering $\{\tilde{U}_j\}$ of M with coordinate chart mappings

$$\tilde{h}_j : \tilde{U}_j \to \mathbf{R}^n,$$

and then consider the countable collection of balls $\{B_{\mu j}\}$ in the open set

$$\tilde{h}_j(U_j) \subset \mathbf{R}^n$$

Let φ be a cut-off function for

$$B(0,1) \subset B(0,2) \subset B(0,3),$$

as we constructed earlier (see (12.1)), and let

$$\varphi_i := \varphi \circ h_i, \tag{12.4}$$

which are differentiable functions on M with compact support in U_i.

We now use this covering $\{U_i\}$ and the associated cut-off functions φ_i to define a sequence of local approximations which will converge to the desired global immersion. We will simply illustrate one simple step in the approximation procedure. Suppose that f is given as an immersion on an open set $U \subset M$. Let the sets $\{U_i\}$ be indexed by integers $i \in \mathbf{Z}$, and suppose that for $i < 0$,

$$W_i \subset U,$$

and for $i \geq 0$, this is not the case. We restrict our attention to $U_0 \cup U$, which is an extension of U to a somewhat larger open set, and define an approximation to f by

$$\tilde{f} := f + \varphi_0 h_0^{-1}(Ax), \tag{12.5}$$

where A is an $N \times n$ matrix to be chosen, and again the vector x is considered as a column vector in $\mathbf{R}^n$.

We want to choose A small enough so that

$$|\tilde{f}(x) - f(x)| < \eta(x),$$

for a given measure of approximation $\eta(x)$. Expressing this perturbation in terms of the local coordinates $x \in \mathbf{R}^n$, we have

$$\tilde{f}(x) = f(x) + \varphi(x)Ax,$$

and we require the Jacobian matrix

$$D\tilde{f} = Df + Ax \cdot D\varphi(x) + \varphi(x)A$$

to have maximal rank on $B(0,1)$ (here $D\varphi(x)$ is a row vector). Note that on $B(0,1)$, $\varphi(x) \equiv 1$, and $D\varphi(x) \equiv 0$, so that using the arguments as in Sect. 12.2 we can choose an A small enough to ensure that $D\tilde{f}$ has maximal rank on a neighborhood

(Footnote 4 continued)
of rational radii and rational center points. The collection of open sets $\{\tilde{h}_j(B_{\mu j})\}$ will provide a covering of M from which one can construct the desired locally finite covering.

of $\overline{B}(0,1)$. Hence on M, $\tilde{f}$ is regular on a neighborhood of $\overline{W_0}$. This argument can be extended[5] inductively to all of M.

There is a similar argument with a slightly different local perturbation that provides a proof that an immersion can be modified to become a differentiable embedding, provided that $N \geq 2n+1$. Suppose that we have an immersion

$$f : M \to \mathbf{R}^N,$$

where $N \geq 2n+1$, then we want to indicate how to show that it can be approximated by an embedding.

Suppose that f is one-to-one on an open set $U \subset M$, and let us use the same type of covering

$$W_j \subset V_j \subset U_j$$

as was used in the previous paragraphs, where we assume that f is an embedding on U and when restricted to each U_i (any immersion is an embedding on the neighborhood of each point, as we have seen earlier). Let us consider the first approximation in U_0 to be of the form

$$\tilde{f}(x) := f(x) + \varphi_0(x)b,$$

where b is a vector in $\mathbf{R}^n$ to be chosen and φ_0 is the cut-off function for $W_0 \subset V_0 \subset U_0$. Let

$$N := \{(x,y) \in M \times M : \varphi_0(x) \neq \varphi_0(y)\}.$$

Consider the mapping

$$\Phi : N \to \mathbf{R}^N$$

defined by

$$\Phi(x,y) = \frac{f(x) - f(y)}{\varphi_0(x) - \varphi_0(y)}.$$

Since $\dim N = 2n$, and $N > 2n$, if follows that

$$\text{meas}\, \Phi(N) = 0 \text{ in } \mathbf{R}^N.$$

By choosing a vector b sufficiently small and not in the image of Φ, it follows easily that the mapping $\tilde{f}$ is an embedding on $U \cup V_0$. This can then be extended step-by-step inductively to the remaining elements of the covering U_i, $i > 0$.

The final step of the proof of the embedding theorem is to show that there is a *proper embedding* of a manifold of dimension n in the Euclidean space $\mathbf{R}^N$. We recall that the sequence of approximations above started with a given arbitrary continuous mapping

[5] The lecture notes on differential topology by Milnor [154] have a very readable and simplified account of Whitney's original arguments in [245].

$$f : M \to \mathbf{R}^N,$$

where $N \geq 2n+1$, and we obtained an approximation which was an embedding. We noted at the time that the initial mapping could well have been a constant mapping.

Let us consider a somewhat different initial mapping as a starting point, namely a proper mapping of a specific kind. Let $\{U_i\}$ be a locally finite covering of the type used above, where now we assume the index set consists of the positive integers, and let $\varphi_i(x)$ be the cut-off functions as in (12.4). Let

$$f(x) := \sum_n n\varphi_n(x)$$

be a differentiable function defined on M and let f denote the mapping

$$f : M \to \mathbf{R} \subset \mathbf{R}^N$$

given by the natural inclusion of the real line $\mathbf{R}$ in the higher-dimensional Euclidean space $\mathbf{R}^N$.

It is easy to verify that f is a proper mapping. Namely, any compact set K in $\mathbf{R}^N$ is bounded, and if $|y| \leq L$ for y in K where L a positive integer, then evidently $|f(x)| \leq L$, for any x in $f^{-1}(K)$. It follows that

$$f^{-1}(K) \subset \bigcup_{i=0}^{L} W_i \subset \bigcup_{i=0}^{L} \overline{W_i},$$

since $f(x) > L$ on all other W_i, and hence $f^{-1}(K)$ is compact, as desired.

Now taking this f as our initial mapping and finding an approximation $\tilde{f}$ of f which is an embedding as in the previous paragraphs such that

$$|f(x) - \tilde{f}(x)| < 1 \text{ on } M,$$

it will follow that $\tilde{f}$ is also a proper mapping. This completes our outline of Whitney's original proof of his embedding theorem from 1936.

12.4 Concluding Remarks

There were several unanswered questions which arose out of the fundamental results of Whitney's 1936 embedding paper. First of all, we remark that there were many aspects of this wide-ranging paper that we have not discussed here. For instance, he was able to show that a differentiable submanifold of $\mathbf{R}^N$ can be deformed in $\mathbf{R}^N$ in a differentiable manner to a real-analytic submanifold of $\mathbf{R}^N$. Hence, his paper gives a proof that an abstract real-analytic manifold M of dimension n has a C^∞ embedding of M onto a real-analytic submanifold of $\mathbf{R}^{2n+1}$. However, he was not able to find a real-analytic embedding. We quote from the first page of his paper ([245], p. 645):

> A fundamental problem is the following: *Can any analytic manifold be mapped in an analytic manner into Euclidean space?*

This problem remained unresolved until 1958, when Hans Grauert (1930–2011) provided a complete solution, as we will see in Sect. 15.7.

A second problem arose naturally from Whitney's paper. Could the embedding dimension $2n+1$ for an n-dimensional manifold be diminished, and if so, how much? In 1944 Whitney [246] was able to show that if an n-dimensional manifold M were immersed in $\mathbf{R}^{2n}$ in a manner that the two intersecting submanifolds at a double point were transversal (which is the generic case, and which he had proved in his 1936 paper), then by an ingenious sequence of deformations of the immersed submanifold, called the "Whitney trick," he was able to deform the immersed submanifold to an embedded submanifold, having removed the double points. Hence the *strong Whitney embedding theorem* asserts: any n-dimensional differentiable manifold can be embedded in $\mathbf{R}^{2n}$.

It had been known for some time at the time of Whitney's work that the Klein bottle and real two-dimensional projective space could not be embedded in $\mathbf{R}^3$, but could be embedded in $\mathbf{R}^4$, consistent with Whitney's general work (this related to their non-orientable nature). A later result of Franklin Peterson in 1957 [179] generalized these examples. Namely, he showed that real projective space of dimension 2^n cannot be embedded in $\mathbf{R}^{2^{n+1}-1}$. Hence it is not true that any n-manifold embeds in $\mathbf{R}^{2n-1}$, and the embedding dimension $2n$ is the best possible in general.

There are, however, many examples of embedding and non-embedding theorems of a more specialized nature. These questions are mostly related to problems in algebraic topology (more specifically, characteristic class questions of the associated vector bundles arising in the embedding problems). The survey paper by Sanderson in 1964 [207] illustrates a number of such embedding results. There is a listing on the Internet of a large number of such embedding and immersion theorems (including the paper by Sanderson, just quoted, for instance) by Don Davis of Lehigh University [51].

Chapter 13
Riemannian Manifolds

13.1 Introduction

In the previous chapter, we have described how a differentiable manifold can be embedded in a Euclidean space. In this chapter, we want to discuss the isometric embedding theorem of John Nash (1928–2015) [165]. First we recall that a *Riemannian manifold* is a differentiable manifold M with a Riemannian metric g which is a smoothly varying inner product on the tangent space $T_x(M)$ at each point $x \in M$, which we denote by $g(X, Y)_x$, where X and Y are tangent vectors at x. Locally such a metric is represented by a symmetric positive-definite matrix of differentiable functions

$$g_{ij}(x) = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right)_x,$$

where $\frac{\partial}{\partial x_i}$ and $\frac{\partial}{\partial x_j}$ are tangent vectors at x in terms of local coordinates $(x_1, \ldots, x_n)$.

Let now $\mathbf{R}^N$ be equipped with its usual Euclidean structure, which makes it into a Riemannian manifold as described above. Classically this is described by

$$ds^2 = \sum_{j=1}^{N} dy_j^2,$$

where $(y_1, \ldots, y_N)$ are coordinates in $\mathbf{R}^N$. In this case the matrix g_{ij} is the constant identity matrix, and we designate this Riemannian structure on $\mathbf{R}^N$ by the symbol ρ.

If $f : M \to N$ is a differentiable mapping from a differentiable manifold M to a differentiable manifold N, and N is equipped with a Riemannian metric g, then there is a natural pullback Riemannian metric on M defined by

$$f^*(g)(X, Y) = g(D_f X, D_f Y)_{f(x)},$$

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_13

where $D_f X$ and $D_f Y$ represent the derivative of f mapping tangent vectors at $x \in M$ to tangent vectors at $f(x) \in N$.

If

$$f : M \to \mathbf{R}^N$$

is a differentiable mapping, then the pullback Riemannian metric on M induced by the mapping $f = (f_1, \dots, f_N)$ and the Euclidean metric ρ on $\mathbf{R}^N$ is given by

$$g_f := f^*(\rho) = \sum_\alpha (df_\alpha)^2,$$

or in local coordinates $(x_1, \dots, x_n)$ on M

$$(g_f)_{ij} = \sum_\alpha \frac{\partial f_\alpha}{\partial x_i}\frac{\partial f_\alpha}{\partial x_j}.$$

The *isometric embedding problem* is: given a Riemannian manifold (M, g), find a differentiable embedding

$$f :\to \mathbf{R}^N,$$

so that

$$f^*(\rho) = g.$$

In other words, can an embedding f be found so that the given metric g on M is the same as the metric induced by the ambient Euclidean metric on the submanifold $f(M) \subset \mathbf{R}^N$?

This question in its global form could not have been formulated until we had the notion of an abstract manifold as discussed earlier in this book. However, local Riemannian geometry has been a subject of intense study since the advent of Riemann's innovative paper of 1954 [201]. In 1871 Ludwig Schaefli (1814–1895) formulated and announced a theorem that a local Riemannian manifold M of dimension n could be locally isometrically embedded in $\mathbf{R}^{\frac{n(n+1)}{2}}$ [208]. This was the first instance of an isometric embedding theorem, which then became an object of focussed investigation in the twentieth century.

In the local situation, where we have an n-dimensional manifold with coordinates $(x_1, \dots, x_n)$ to be embedded in $\mathbf{R}^N$, if we look at the equation

$$g_{ij}(x) = \sum_{\alpha=1}^{N} \frac{\partial f_\alpha(x)}{\partial x_i}\frac{\partial f_\alpha(x)}{\partial x_j},$$

where g_{ij} is given and the embedding functions $(f_1, \dots, f_N)$ are unknown, then we see that there are $\frac{n(n+1)}{2}$ differential equations (recalling that g_{ij} is a symmetric $n \times n$ matrix) with N unknowns.

So, in principle, the optimal number of unknown functions defining an embedding would be $N = \frac{n(n+1)}{2}$, and this is what Schaefli formulated in his paper. This result was proved in the real-analytic case for $n = 2$ by Maurice Janet (1888–1983) in 1926 [117] and immediately generalized in 1927 to the general case of a local real-analytic isometric embedding theorem by Elie Cartan [32]. Cartan utilized the theory of Pfaffian systems as a tool which yielded this result quite simply. The analytic basis for solving these types of geometric families of partial differential equations was the Cauchy–Kovalevskaya theorem, which required real-analyticity of the differential equations involved.

Hermann Weyl attempted to solve a global type of isometric embedding problem when he tried to show that a Riemannian two-sphere with positive curvature could be isometrically embedded into $\mathbf{R}^3$. He was not successful, but he outlined a method of proof that he thought could some day yield a proof if carried out completely. Louis Nirenberg (1925–) was able to complete Weyl's proof in 1953 [171]. There were also results developed during this time period which showed that isometric embeddings for specific dimensional situations were not possible if certain curvature constraints were imposed. For instance, Hilbert showed in 1901 that a two-dimensional surface of constant negative curvature could not be isometrically embedded in $\mathbf{R}^3$ as a smooth hypersurface [101]. For other such references, we refer to Nash's paper [165].

Now we turn to Nash's paper of 1956, where he proved a quite general isometric embedding theorem [165], which we formulate here. In Fig. 13.1 we see the opening page of Nash's celebrated paper.

Theorem 13.1 (Nash 1956 [165]) *Let (M, g) be an n-dimensional Riemannian manifold, then there is an integer $N > n$ and an embedding*

$$f : M \to \mathbf{R}^N$$

such that

$$g = f^*(\rho),$$

where ρ is the Euclidean metric on $\mathbf{R}^N$.

In the next section we will give a summary of the key ideas in the proof of Nash's theorem. The principal theorem was first proved by Nash in his paper for the case of a *compact* Riemannian manifold, and then at the end of his paper, he shows by a relatively simple device how he was able to extend this result to the noncompact case. We will mention this briefly at the end of the next section. However, we will only consider the isometric embedding theorem for the case of compact Riemannian manifolds in the remainder of this chapter.

ANNALS OF MATHEMATICS
Vol. 63, No. 1, January, 1956
Printed in U.S.A.

THE IMBEDDING PROBLEM FOR RIEMANNIAN MANIFOLDS

BY JOHN NASH

(Received October 29, 1954)
(Revised August 20, 1955)

Introduction and remarks

History. The abstract concept of a Riemannian manifold is the result of an evolution in mathematical attitudes [1, 2]. In an earlier period mathematicians thought more concretely of surfaces in 3-space, of algebraic varieties, and of the Lobatchevsky manifolds. As the more abstract view of manifolds came into favor a question naturally arose: To what extent are the abstract Riemannian manifolds a more general family than the sub-manifolds of euclidean spaces?

This question has been considered in various specializations and with assorted side conditions. In 1873 Schlaefli [3] discussed the local form of this imbedding problem. He conjectured that a neighborhood in an n-manifold would generally require an imbedding space of $(n/2)\ (n + 1)$ dimensions. In 1901 Hilbert [4] obtained a negative result, showing that the Lobatchevsky plane is not realizable as a smooth surface in E^3. Some contemporary negative theorems are due to Tompkins [5] and to Chern and Kuiper [6]. For example, a flat n-torus is not realizable in less than $2n$ dimensions.

Janet [7] solved the local problem for two-manifolds with analytic metric in 1926, and Cartan [8] immediately extended the result to n-manifolds, treating it as an application of his theory of Pfaffian forms. The dimensionality requirement was $(n/2)\ (n + 1)$, as conjectured by Schlaefli. This number is a plausible one, being the number of components of the metric tensor. The proof depended on power series development, so it was limited to local results and it required that the metric be analytic.

There are some theorems on the existence of isometric imbeddings in infinite dimensional spaces. This is a much simpler problem.

A recent discovery [9, 10] is that C^1 isometric imbeddings of Riemannian manifolds can be obtained in rather low dimensional spaces. At first glance some of these C^1 results seem inconsistent with the negative theorems, such as Hilbert's. Apparently C^1 imbeddings are very different from the smoother ones.

Until recently the only general results on imbeddings in the large were proved for the problem of Weyl. This problem is to realize in E^3 all two-manifolds with everywhere positive Gaussian curvature. Alexandrov [13] and Pogorelov [14] have been successful with a geometrical approach based on polyhedral approximations. H. Lewy [12] and L. Nirenberg [15] have treated the problem from the viewpoint of partial differential equations. These results can probably be sharpened with respect to differentiability, but dimension-wise they are clearly optimal.

Rigidity theory concerns the metric preserving perturbations of an imbedding.

20

Fig. 13.1 First page from Nash's 1956 isometric embedding paper [165]. *Reprinted with the permission of the Annals of Mathematics*

13.2 Summary of the Proof of Nash's Embedding Theorem

There are three key steps in the proof of Nash's theorem, and the second step is by far the most difficult. We will outline these steps here and show how this proves the embedding theorem. Then we will discuss each of them in more detail in the concluding three sections of this chapter.

We assume now that an n-dimensional compact Riemannian manifold M with a Riemannian metric g are given, and then consider the ideas leading to the proof of the embedding theorem.

Step 1: Nondegenerate Embeddings. Nash's first step is to formulate and show the existence of a particular kind of nondegenerate embedding of M into a Euclidean space. Let

$$f : M \to \mathbf{R}^N$$

be a given embedding that we know exists by Whitney's embedding theorem (we could take N to be $2n$, for instance). As an embedding, we know that the Jacobian matrix (in local coordinates $(x_1, \ldots, x_n)$)

$$\frac{\partial(f_1, \ldots, f_N)}{\partial(x_1, \ldots, x_n)} \tag{13.1}$$

has maximal rank at each point of M. Here we consider this Jacobian matrix as having N columns indexed by the f_i and n rows indexed by the first-order partial derivatives $\frac{\partial}{\partial x_j}$.

We now want to generalize the maximum rank condition involving first derivatives to take into account nondegeneracy involving second derivatives as well. Now we assume we have an embedding f of M into a Euclidean space of dimension N, where $N \geq n + \frac{n(n+1)}{2}$, and consider the following generalization of the Jacobian matrix (13.1)

$$\begin{pmatrix} \frac{\partial f_\alpha}{\partial x_i} \\ \frac{\partial^2 f_\alpha}{\partial x_i \partial x_j} \end{pmatrix}, \tag{13.2}$$

for $i \leq j$. Here the columns are indexed by the embedding functions $(f_1, \ldots, f_N)$ as in (13.1), the first n rows are given by the first-order partial derivatives (again as in (13.1)), and the next $\frac{n(n+1)}{2}$ rows are given by the second-order derivatives in some specific order. We say that f is *nondegenerate* if the generalized Jacobian matrix in (13.2) has rank $n + \frac{n(n+1)}{2}$ at each point of M. It is not difficult to check that this condition is valid under changes of coordinate systems on M.

Nash showed that there exist nondegenerate embeddings of M into a Euclidean space $\mathbf{R}^N$, where $N = (n/2)(n + 5)$, and we will describe his proof below. This is the simplest of his three steps, as we will see.

Step 2: A Perturbation Theorem. Now we want to discuss a topological structure on the vector space of all C^s functions on M, for $s \geq 1$. Let us choose a finite covering of M by coordinate charts U_i of the form

$$\chi_i : U_i \to B(0, 3)$$

and such that if $V_i = \chi_i^{-1}(B(0, 2))$, then $\{\overline{V}_i\}$ also covers M (as we did in Sect. 12.2). We define, for any $f \in \mathbf{C}^s(M)$,

$$\|f\|_s = \sup_i (\sup_{|\alpha| \leq s} \sup_{x \in \overline{V}_i} |D^\alpha f(x)|),$$

where

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_n}}(x),$$

and where α is a multiindex of the form $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. The norm $\| \cdot \|$ makes $C^s(M)$ into a *Banach space*.[1] This norm depends on the choice of coordinate charts on M, but any other choice of coordinate charts gives an equivalent topology on the vector space $C^s(M)$.

We can create the same sort of norm on the vector space of vector-valued functions, and we let $\mathrm{Map}^s(M, \mathbf{R}^N)$ be the Banach space of C^s mappings from M to $\mathbf{R}^N$. We can denote the C^s embeddings of M into $\mathbf{R}^N$ by $\mathrm{Em}^s(M, \mathbf{R}^N)$ and similarly the nondegenerate embeddings of M into $\mathbf{R}^N$ can be denoted by $\mathrm{ND}^s(M, \mathbf{R}^N)$. Thus we have inclusions:

$$\mathrm{ND}^s(M, \mathbf{R}^N) \subset \mathrm{Em}^s(M, \mathbf{R}^N) \subset \mathrm{Map}^s(M, \mathbf{R}^N),$$

and it is easy to see that (assuming that $s \geq 2$) each of these inclusions is an open subset of the next Banach space in this series, since the regularity and one-to-one nature of these mappings is preserved under C^2 perturbations.

Let us now consider the vector space of covariant symmetric two-tensors on M, and let us denote this vector space by $\mathrm{Sym}(M)$ for convenience.[2] Our given Riemannian metric g is a tensor of this type. We can put a Banach space structure

---

[1]We refer to any standard text on real analysis that describes the basic concepts of elementary functional analysis that we need in this book, i.e., Banach spaces and Hilbert spaces are natural generalizations to infinite dimensions of Euclidean space that we utilize. See, for instance, the classic text by Royden [204], which has the basic language of Banach spaces and Hilbert spaces that we need in this book.

[2]Tensor fields in geometry are generalizations of vector fields and differential forms and go back to Riemann's original description of a Riemannian metric, as described in Sect. 5.2. In particular, a covariant symmetric two-tensor is simply, in terms of a local coordinate system $x_1, \ldots, x_n$), a differentiable matrix $t_{ij}(x)$ which is symmetric and behaves under coordinate transformations just as a Riemannian metric does, as described in (5.4). See Helgason's excellent monograph on differential geometry [99] for a description of tensor fields in terms of multilinear algebra and vector fields on a differentiable manifold, along with their varied local descriptions, as we have used here.

with a specific degree of smoothness on these tensors, just as we did for the smooth functions above. We note that a given tensor on a coordinate chart is a vector-valued function. We can put a C^s norm on each V_i in the same manner as we did for scalar functions in the previous paragraph and then maximize over all coordinate charts obtaining a C^s norm on a given tensor t, which we denote again by $\|t\|_s$. Let $\mathrm{Sym}^s(M)$ be the space of C^s tensors on M. This again depends on the coordinate charts, but the topology does not. The set of Riemannian metrics on M which are locally C^s vector-valued functions on M is then an open set in $\mathrm{Sym}^s(M)$, which we denote by $\mathrm{Met}^s(M)$. Thus we have the inclusion

$$\mathrm{Met}^s(M) \subset \mathrm{Sym}^s(M).$$

We see that if $f \in \mathrm{Map}^s(M, \mathbf{R}^N)$, then

$$g_f = f^*(\rho) \in \mathrm{Met}^{s-1}(M),$$

since $f^*(\rho)$ involves first derivatives of the mapping f. So, in particular, there is a mapping

$$F : \mathrm{ND}^s(M, \mathbf{R}^N) \to \mathrm{Met}^{s-1}(M). \tag{13.3}$$

Nash's second step in his proof is to show that the mapping (13.3) is an *open mapping* (provided $s \geq 3$). In particular, if $g_0 = f_0^*(\rho)$ is a particular Riemannian metric in $\mathrm{Met}^{s-1}(M)$ induced by a nondegenerate embedding f_0, then *every* Riemannian metric in a neighborhood of g_0 is induced by a nondegenerate embedding as well. That is, all metrics in such a neighborhood come from isometric embeddings into $\mathbf{R}^N$. This perturbation result of Nash is the deepest and most difficult part of his paper, and we will discuss it in greater detail below.

Step 3: An Approximation Theorem. The third step in the proof is to show that for a given metric g on M, one can find a C^3-approximation to g by a metric $\tilde{g}$ which is induced by some mapping $y : M \to \mathbf{R}^L$, i.e., $\tilde{g} = g_y$, where $L = 2n^2 + 3n$ (note that a mapping can induce a metric even if the mapping is not an embedding!).

Outline of Proof: Putting these three steps together, as Nash does, involves a very nice trick which we outline briefly here, and then we will discuss each of these steps in more detail below. By the hypothesis in Nash's theorem, we are given a metric g on the manifold M. By Step 1, we can find a nondegenerate embedding f which induces a second metric g_f on M. But the two metrics on M, g and g_f, have no *a priori* relation to each other, and, for instance, the distance between them $\|g - g_f\|$ in some C^k-norm might be quite large. We know from Step 2 that metrics close to g_f can be induced by an embedding, so we need to find a way of reducing the problem to this situation.

Nash solves this problem by considering a suitable additional mapping y into a perhaps different dimensional Euclidean space which can effectively move the problem into a suitable neighborhood of g_f, as we will see. Suppose we have two such differentiable mappings

$$f : M \to \mathbf{R}^N,$$
$$y : M \to \mathbf{R}^L,$$

where f is an embedding.[3] Then we can define a new mapping, which is then also an embedding:

$$e : M \to \mathbf{R}^{N+L}$$

by

$$e = (f_1, \dots, f_N, y_1, \dots, y_L).$$

It then follows that there is an addition of induced metrics on M, namely,

$$\begin{aligned} e^*(\rho) &= f^*(\rho) + y^*(\rho) \\ &= g_f + g_y. \end{aligned}$$

Nash chooses f to be a nondegenerate embedding as above, with the additional property that $g - g_f$ is positive-definite (which is possible by changing the scale in $\mathbf{R}^N$ for the mapping f). Then he chooses the mapping y so that

$$g_f + g_y \text{ approximates } g.$$

More specifically, he chooses the mapping y so that

$$g_y \text{ approximates } g - g_f,$$

which is possible by Step 3, since $g - g_f$ is a metric on M. We choose this approximation in the C^3-norm.

Putting this together, we now have

$$\|(g - g_y) - g_f\|_3 < \varepsilon,$$

and hence $g - g_y$ is realizable as an embedding

$$g - g_h = g_{\tilde{f}},$$

for a nondegenerate embedding

$$\tilde{f} :\to \mathbf{R}^N,$$

by Step 2. Consequently,

$$g = g_{\tilde{f}} + g_y = g_{\tilde{e}},$$

[3]Nash uses the notation (z, y) for this pair of mappings, where z is an embedding. We have used the notation f in general in this book to denote an embedding.

where
$$\tilde{e} = (\tilde{f}, y) : M \to \mathbf{R}^{N+L},$$

which shows that the given metric g on M is given as an induced metric from the embedding $\tilde{e}$, and this concludes Nash's proof of his theorem (assuming Steps 1, 2, and 3 above).

In Step 1, Nash was able to find $N = (n/2)(n+5)$, and in Step 3, he obtained $L = 2n^2 + 3n$, obtaining for a compact Riemannian manifold of dimension n an isometric embedding of the form

$$f : M \to \mathbf{R}^{N+L} = \mathbf{R}^{(n/2)(3n+11)}.$$

For noncompact manifolds his embedding dimension is larger, and it is of the form

$$f : M \to \mathbf{R}^{(n/2)(n+1)(3n+11)}.$$

13.3 Nondegenerate Embeddings

As we have seen in Sect. 12.3, Whitney's embedding theorem tells us that if M is an n-dimensional differentiable manifold, then there is an embedding

$$f : M \to \mathbf{R}^T,$$

for some $T > n$ (in fact, T can be taken to be $2n$). We want to show that there exist nondegenerate embeddings in the sense described in (13.2). Consider the following embedding which uses the mapping functions from f. Namely, let

$$e = (e_1, \ldots, e_N) : M \to \mathbf{R}^N$$

be given by

$$e = (f_1, \ldots, f_T, f_1^2, \ldots, f_T^2, f_1 f_2, \ldots, f_1 f_T,$$
$$f_2 f_3, \ldots, f_2 f_T, \ldots, f_{T-1} f_T),$$

where we see that $N = T + \frac{T(T+1)}{2}$. We claim that e is a nondegenerate embedding.

To see this, we let $p \in M$, and let $(x_1, \ldots, x_n)$ be local coordinates near $p = (0, \ldots, 0)$. Suppose, by reordering the mapping functions $(f_1, \ldots, f_T)$, the first n functions have nonzero Jacobian determinant

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}$$

near p. By the implicit function theorem, we can express the mappings f and e locally near p as a graph over the coordinate chart with coordinates $x = (x_1, \dots, x_n)$. Namely, f becomes

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, f_{n+1}(x_1, \dots, x_n), \dots)$$

or

$$x \mapsto (x, f_{n+1}(x), \dots, f_T),$$

and the mapping e becomes, in the same manner,

$$x \mapsto (x, f_{n+1}(x), \dots, f_T(x), x_1^2, \dots, x_n^2, f_{n+1}^2(x), \dots).$$

By reordering again, and starting with the explicit linear and quadratic terms, we have that e has the form

$$x \mapsto (x_1, \dots, x_n, x_1^2, \dots, x_n^2, x_1x_2, \dots, x_1x_n, x_2x_3, \dots, x_{n-1}x_n, f_{n+1}(x), \dots).$$

The second-order matrix of derivative of this mapping will have the form

$$\begin{pmatrix} \frac{\partial f^\alpha}{\partial x_i} \\ \frac{\partial^2 f^\alpha}{\partial x_i \partial x_j} \end{pmatrix},$$

as described above in (13.2). There are $n + \frac{n(n+1)}{2}$ rows in this generalized Jacobian matrix, and using the explicit form of the mapping above, when we compute this matrix at $x = 0$, we obtain

$$\begin{pmatrix} I & * & * * \\ 0 & 2I & * * \\ 0 & 0 & I * \end{pmatrix},$$

where I denotes an identity matrix of the appropriate size. This shows that the mapping e is nondegenerate near the point $p \in M$, and since this argument is valid for any such point on M, we see that the mapping is nondegenerate, as desired.

This argument shows that there exist nondegenerate embeddings of the form

$$f : M \to \mathbf{R}^{2n+n(2n+1)} = \mathbf{R}^{2n^2+3n}.$$

Nash has a perturbation of this argument[4] in his paper which lowers this embedding dimension substantively, and he obtains a nondegenerate embedding of the form

$$f : M \to \mathbf{R}^{2n+\frac{n(n+1)}{2}} = \mathbf{R}^{(n/2)(n+5)}.$$

[4] In Sect. 13.5 there is a perturbation argument which lowers the mapping dimension in that context, which is very similar to the argument which is omitted here.

13.4 Nash's Implicit Function Theorem

In our summary of the proof of Nash's isometric embedding theorem in Sect. 13.2, we discussed what we called Step 2 of this proof, namely that the mapping

$$F : \mathrm{ND}^s(M, \mathbf{R}^N) \to \mathrm{Met}^{s-1}(M),$$

as formulated in (13.3), is an open mapping. This specific result is now referred to as the *Nash implicit function theorem*, and its proof has been adapted to many other contexts, as we will briefly mention later. We want to understand this in the context of more classical implicit function theorems, and then we will proceed to a discussion of the proof itself.

The mapping F in (13.3) is actually defined on the Banach spaces which include the specific open sets $\mathrm{ND}^s(M, \mathbf{R}^N)$ and $\mathrm{Met}^{s-1}(N)$. Namely,

$$F : \mathrm{Map}^s(M, \mathbf{R}^N) \to \mathrm{Sym}^{s-1}(M) \tag{13.4}$$

is given by the same local coordinates formula

$$(F(f_1, \ldots, f_N))_{i,j} = \sum_{\alpha=1}^{N} \frac{\partial f_\alpha}{\partial x_i}\frac{\partial f_\alpha}{\partial x_j},\ \ i, j = 1, \ldots, n, \tag{13.5}$$

which is a nonlinear mapping from one Banach space to another. Note that it is nonlinear simply because $F(\lambda f) = \lambda^2 F(f)$.

In the finite-dimensional situation, if we have a smooth mapping

$$F : U \subset \mathbf{R}^n \to \mathbf{R}^m,\ \ n \geq m,$$

where U is an open subset of $\mathbf{R}^n$, then the classical implicit function theorem asserts that if DF, the derivative of the mapping F, is surjective at a point $x \in U$, then a neighborhood of x is mapped by F onto a neighborhood of $F(p)$.

It follows that in the infinite-dimensional case we are led to look at the derivative of the mapping F. There are several notions of differentiation in the infinite-dimensional case, but a common and very important one is the Fréchet derivative, which is patterned for Banach spaces after the finite-dimensional case. Namely, let V and W be Banach spaces, and if $F : V \to W$, then the *Fréchet derivative* $F'(u)$ is defined to be a bounded linear operator[5] $A : V \to W$ such that

$$\lim_{h \to 0} \frac{\|F(u + h) - F(u) - Ah\|}{\|h\|} = 0.$$

[5]A *bounded linear operator* is a linear mapping $L : V \to W$, where V and W are Banach spaces, and such that there is a constant K such that $\|L(v)\| \leq K\|v\|$, for all $v \in V$.

We will consider a special case of mappings of Banach spaces which serves our purpose here. Namely, F is said to be a *quadratic form* on V if

$$B(u, v) := F(u + v) - F(u) - F(v)$$

is bilinear and

$$F(\lambda u) = \lambda^2 F(u).$$

It follows easily that such an F has a continuous Fréchet derivative at each point of V and that

$$F'(u)h = B(u, h).$$

Going back to our mapping F in the Nash setting (13.4), we easily see that F is a quadratic form on $\mathrm{Map}^s(M, \mathbf{R}^N)$, and the Fréchet derivative is given in local coordinates by

$$(F'(f)h)_{i,j} = \sum_\alpha \frac{\partial f_\alpha}{\partial x_i}\frac{\partial h_\alpha}{\partial x_j} + \sum_\alpha \frac{\partial h_\alpha}{\partial x_i}\frac{\partial f_\alpha}{\partial x_j}, 1 \le i, j \le n. \tag{13.6}$$

Let us now give an example of an infinite-dimensional implicit function theorem which is a model for the more difficult Nash implicit function theorem (see Theorem 2.1 in [210]).

Theorem 13.2 *Let B be the unit ball in a Banach space V and let*

$$F : B \to V$$

be a mapping with first and second Fréchet derivatives which are bounded by a constant $M > 2$, and suppose that there is a mapping

$$L : B \to \mathcal{B}(V),$$

where $\mathcal{B}(V)$ denotes the bounded operators on V, such that

$$F'(u)L(u)h = h,\ u \in B, h \in B, \tag{13.7}$$

$$|L(u)h| \le M|h|,\ u \in B,\ h \in V. \tag{13.8}$$

Then, if

$$|F(0) \le M^{-5},$$

it follows that $F(B)$ contains a neighborhood of the origin.

Equation (13.7) in Theorem 13.2 asserts that the derivative $F'(u)$ of F has a right inverse at each point of B. This is the same as hypothesizing that the derivative mapping should be surjective, as was hypothesized in the finite-dimensional setting.

It is easy to see, by using translations in both the domain and image Banach spaces, that this theorem implies that if $F(u_0) = v_0$, for a given $u_0 \in B$, then the image of B in V contains a neighborhood of $v_0 \in V$, just as in the finite-dimensional case, i.e., that F is an open mapping.

The proof of Theorem 13.2, as given in Schwartz's lecture notes [210], uses an infinite-dimensional version of the classical Newton method for approximating the roots of a function in one variable. Let us recall the classical situation to clarify this. Suppose we are given a suitable function $f : \mathbf{R} \to \mathbf{R}$, and we want to find a root, i.e., a point x on the real line where $f(x) = 0$. Assume the derivative of $f \neq 0$ and choose an initial approximation x_0, and let

$$\begin{aligned} x_1 &:= x_0 - \tfrac{f(x_0)}{f'(x_0)}, \\ &\vdots \\ x_{n+1} &:= x_n - \tfrac{f(x_n)}{f'(x_n)}. \end{aligned}$$

Under appropriate hypotheses this set of successive approximations converges to a point x on the real line where $f(x) = 0$. We note that each of the reciprocals of the derivative used in these approximations, $1/f'(x_n)$, is the *inverse* to the one-dimensional linear mapping $f'(x_n) : \mathbf{R} \to \mathbf{R}$.

We can mimic this procedure for a proof of Theorem 13.2 by using the right inverses $L(u_n)$ to the linear mappings $F'(u_n) : B \to B$. Namely, we obtain a set of successive approximations of the form

$$u_{n+1} := u_n - L(u_n)F(u_n), \tag{13.9}$$

where we set $u_0 = 0$. This does indeed provide a proof of the theorem, as shown in [210], pp. 42–44.

Let us return to the Nash mapping (13.4), whose derivative is given in the formula (13.6). We write this symbolically as

$$F'(f)h = k, \tag{13.10}$$

and we need to try to solve this linear equation for $h \in \mathrm{Map}(M, \mathbf{R}^N)$, for a given $k \in \mathrm{Sym}(M)$ (ignoring the orders of differentiability for the moment). This is a *linearization* (at a specific point) of the original problem: given a metric g, find an embedding f such that $F(f) = g$.

To try to use the ideas in the Banach space implicit theorem above (Theorem 13.2), we need to try to find a right inverse to the linear mapping (13.10). In order to do this, Nash introduces additional linear equations to this under-determined set of linear equations to simplify the problem. Namely, consider the additional constraint on the vector h, in our given local coordinate system:

$$\sum_\alpha \frac{\partial f_\alpha}{\partial x_i} h_\alpha = 0, \ i = 1, \ldots, n. \tag{13.11}$$

Nash phrases this condition as: "requires the perturbation $\{h_\alpha\}$ to be normal to the imbedding" ([165], p. 31). This condition can be expressed in a coordinate-free manner, but we won't bother with that here. Thus we are trying to solve the system of equations:

$$\sum_\alpha \frac{\partial f_\alpha}{\partial x_i}\frac{\partial h_\alpha}{\partial x_j} + \frac{\partial h_\alpha}{\partial x_i}\frac{\partial f_\alpha}{\partial x_j} = k_{ij},\, 1 \le i, j, n, \tag{13.12}$$

$$\sum_\alpha \frac{\partial f_\alpha}{\partial x_i} h_\alpha = 0,\;\; i = 1, \dots, n. \tag{13.13}$$

Again, here we are trying to solve for h_α for a given k_{ij}. Note this set of equations involves derivatives of the unknown functions h_α.

By differentiating (13.11) we obtain

$$\sum_\alpha \frac{\partial^2 f_\alpha}{\partial x_i \partial x_j} h_\alpha + \sum_\alpha \frac{\partial f_\alpha}{\partial x_i}\frac{\partial h_\alpha}{\partial x_j} = 0,\;\; 1 \le i, j \le n. \tag{13.14}$$

We can now substitute (13.14) into (13.12) and obtain a set of equations of the form

$$-2\sum_\alpha \frac{\partial^2 f_\alpha}{\partial x_i \partial x_j} h_\alpha = k_{ij,},\;\; 1 \le i, j \le n, \tag{13.15}$$

$$\sum_\alpha \frac{\partial f_\alpha}{\partial x_i} h_\alpha = 0, \tag{13.16}$$

which is a *linear algebraic system* of equations relating h_α and k_{ij} for a fixed mapping f. Most importantly, there are no derivatives of h_α in these equations. A solution h of this second system (13.15) and (13.16) yields a solution of (13.12) and (13.13), and conversely.

Let us write the equations (13.15) and (13.16) as a matrix equation of the form

$$AH = K, \tag{13.17}$$

where $K = \begin{pmatrix} k_{ij} \\ 0 \end{pmatrix}$ is a vector in $\mathbf{R}^{\frac{n(n+1)}{2}+n}$ and $H = (h_\alpha)$. We now assume that our given mapping f is a nondegenerate embedding, $f \in \mathrm{ND}^s(M, \mathbf{R}^N)$, and we see that this implies that the matrix A in (13.17) has maximal rank (of rank $n + \frac{n(n+1)}{2}$). Since the matrix A has maximal rank, we know that there is at least one solution to the Eq. (13.17). By the linearity of the equation, the set of all solutions is therefore a nonempty convex set. There is a unique solution to (13.17) which satisfies

$$|H|^2 = \sum_\alpha h_\alpha^2 \text{ is a minimum,} \tag{13.18}$$

as is easy to show.

We now want to find an explicit formula for this unique solution in order to see its dependence on the embedding function f. Since A is maximal rank, it follows that the associated Grammian matrix AA^t is nonsingular. Let D be the unique solution to

$$(AA^t)D = K,$$

and let

$$H = A^t D.$$

Then we see that H satisfies (13.17). Namely, $D = (AA^t)^{-1}K$ and hence

$$AH = A(A^t D) = AA^t(AA^t)^{-1}K = K.$$

It is also true that $H = A^t D$ has minimal norm. To see this, let $(\cdot, \cdot)$ be the Euclidean inner product on $\mathbf{R}^N$, and suppose that $\tilde{H}$ is any other solution. We see that

$$\begin{aligned}
(\tilde{H}, \tilde{H}) - (H, H) &= (\tilde{H}, \tilde{H}) - (H, H) - 2(H, \tilde{H} - H) \\
&= (\tilde{H} - H, \tilde{H} - H) - 2(A^t D, \tilde{H} - H) \\
&= (\tilde{H} - H, \tilde{H} - H) - 2(D, A\tilde{H} - AH) \\
&= (\tilde{H} - H, \tilde{H} - H) - 2(D, 0) \\
&\geq 0,
\end{aligned}$$

and hence H is indeed the minimum norm solution, as desired.

This solution H determines a solution h of $F'(f)h = k$ in this coordinate system. If we had two overlapping coordinate systems U_1 and U_2 and solutions h_1 and h_2 in each such coordinate system, then on the overlapping open set $(U_1 \cap U_2)$, these two solutions would have to coincide, by the minimal norm condition, and hence there is a globally defined and well-defined solution h to the Eq. (13.10). Note that the local solution H involves second derivatives of f, since the coefficients of the matrix A involve second derivatives of f, and the solution H involves linear combinations of the coefficients of A as well as the reciprocal of its determinant (which is nonzero). In conclusion, if $k \in \mathrm{Sym}^s(M)$, and if $f \in \mathrm{ND}^{s+2}(M, \mathbf{R}^N)$, then $h \in \mathrm{Map}^s(M, \mathbf{R}^N)$. We define the right inverse $L(f)$ to $F'(f)$ to be this solution h, and we have

$$F'(f)L(f)k = k, \text{ for } k \in \mathrm{Sym}^s(M).$$

It would be nice if we could use this right inverse $L(f)$ for $F'(f)$ as we did above in the proof of the Banach space implicit function theorem using Newton's method of successive approximations as in (13.9). However, there is a big problem in doing this. Namely, $L(f)$ involves a loss of two derivatives, i.e.,

$$L(f) : \mathrm{Sym}^s(M) \to \mathrm{Map}^{s-2}(M), \tag{13.19}$$

so a sequence of approximations such as in (13.9) would successively become less and less smooth, and it would not be possible to obtain a limit of such a sequence which was smooth. This was the problem that Nash attacked in a subtle and powerful manner. He introduced *smoothing operators* acting on the various function spaces that are involved.

Let's briefly review the idea of a smoothing operator. A simple first-order differential operator reduces the smoothness of a function, for instance

$$\frac{d}{dx} : C^k(\mathbf{R}) \to C^{k-1}(\mathbf{R})$$

decreases the differentiability as in the right inverse $L(f)$ in (13.19). Let now φ be a C^∞ compactly-supported function in $\mathbf{R}^n$ (such as in (12.2)), and let

$$S(f)(x) := \int_{\mathbf{R}} f(y)\varphi(x-y)dy$$

be the convolution of φ with a function f. If f is, for instance, simply continuous on $\mathbf{R}^n$, then $S(f)$ is C^∞ on $\mathbf{R}^n$, i.e., S is a smoothing operator, and we have

$$S : C^0(\mathbf{R}^n) \to C^\infty(\mathbf{R}^n).$$

In 1938, Sergei Sobolev (1908–1989) created the theory of *Sobolev spaces* and, in that context, introduced a large variety of smoothing operators. A Sobolev space $W^s(\mathbf{R}^n)$ can be defined intuitively as the space of generalized functions (distributions) with distributional derivatives of order s which are in $L^2(\mathbf{R}^n)$, the Lebesgue square-integrable functions in $\mathbf{R}^n$. The Sobolev spaces are a generalization of the spaces $C^s(\mathbf{R}^n)$ for s being a nonnegative integer, to s being an arbitrary real number. Sobolev's fundamental theorem asserts that the spaces of smooth functions are embedded in his general family of Sobolev spaces. Namely,

$$W^s(\mathbf{R}^n) \subset C^k(\mathbf{R}^n), \text{ for } s > [n/2] + k + 1. \tag{13.20}$$

Sobolev spaces are most simply defined by using the Fourier transform. There are many references for the theory of Sobolev spaces (e.g., Hörmander's classic monograph on partial differential equations [111]; in Chap. IV of Wells [239], the theory of Sobolev spaces, operators of various orders, and the theory of differential and pseudodifferential operators that is suitable for studying elliptic differential operators on manifolds is developed). We will see the use of this theory for compact differentiable manifolds in the next chapter of this book. In our context here, an *operator of order r* is defined to be a continuous linear mapping

$$L : W^s(\mathbf{R}^n) \to W^{s-r}(\mathbf{R}^n), \; r \in \mathbf{R}.$$

Note that r can be any real number here. Differential operators of order k, for k a positive integer, are operators of order k in this context (so operators of order r are generalizations of differential operators, and the notion of order of an operator generalizes the concept of order of a differential operator).

This language allows one to easily define a smoothing operator in this context. Quite simply, a *smoothing operator* is an operator[6] of order k, where $k < 0$. Namely, the smoothing operator increases the smoothness of a given function, and the order indicates the amount of smoothing. This is what Nash introduced in order to create a sequence of successive approximations which converged to a solution of his problem. In particular, Nash introduced suitable smoothing operators T_n so that he could use successive approximations of the following form:

$$f_{n+1} = f_n - T_n L(f_n) f_n. \tag{13.21}$$

He was then able to show that such a sequence converged to a solution of

$$f^*(\rho) = g,$$

as desired, where g is a metric on N sufficiently close (in the C^3-norm) to the metric $f_0^*(\rho)$, for the initial nondegenerate embedding we started with.

The details of this convergence process, as carried out by Nash in his paper, are quite detailed and complex (pp. 37–42 of [165]), involving numerous delicate estimates. The successive approximation sequence (13.21) is a simplification of the set of estimates that Nash actually formulated and used in his proof. We illustrate this by reproducing p. 43 of his paper in Fig. 13.2.

After Nash's paper appeared, J.T. Schwartz (1930–2009) [209] gave a generalization and somewhat simpler proof of Nash's implicit function theorem. This was followed a year later by Jürgen Moser (1928–1999) [161], who gave a substantive generalization of this theorem. Today this implicit function theorem is referred to as the Nash–Moser implicit function theorem. Both Schwartz and Moser showed the utility of this type of implicit function theorem for solving a broad variety of nonlinear partial differential equations. Nash's original work on this type of implicit function theorem, which we have outlined here, involved solving the very specific and quite difficult partial differential equation

$$F(f) = g,$$

or, in local coordinates,

$$\sum_\alpha \frac{\partial f^\alpha}{\partial x_i}\frac{\partial f^\alpha}{\partial x_j} = g_{ij},$$

for $|g|$ being sufficiently small.

[6]The word *operator* is used in functional analysis to indicate a continuous linear mapping or transformation from one topological vector space to another, most often a Banach or Hilbert space.

where $r = 0, 1, 2, 3, 4$ and either $s = r - 5$ (which makes $r - s - 5 = 0$) or s satisfies $-2 \leqq s \leqq r - 2$. These two alternatives correspond to the two expressions $[\theta_2 \mid {}^{-3,1}_{0,4}]$ and $[\theta_3 \mid {}^{-3,1}_{0,4}]$ which are added in T_2. These integrals give varied terms and we can handle the situation most clearly by simply listing all cases. (This is done in Figure 1.)

	$r = 0$	$r = 1$	$r = 2$	$r = 3$	$r = 4$
$s = r - 5$	$\frac{1}{4}\theta_0^{-4}$	$\frac{1}{3}\theta_0^{-3}$	$\frac{1}{2}\theta_0^{-2}$	θ_0^{-1}	$\log(\theta_3/\theta_0)$
$s = -2$	$\theta_0^{-1}\theta_3^{-3}$	$\theta_0^{-1}\theta_3^{-2}$	$\theta_0^{-1}\theta_3^{-1}$	θ_0^{-1}	$\theta_0^{-1}\theta_3$
-1		$\log(\theta_3/\theta_0)\theta_3^{-3}$	$\log(\theta_3/\theta_0)\theta_3^{-2}$	$\log(\theta_3/\theta_0)\theta_3^{-1}$	$\log(\theta_3/\theta_0)$
0			θ_3^{-2}	θ_3^{-1}	1
1				$\frac{1}{2}\theta_3^{-1}$	$\frac{1}{2}$
2					$\frac{1}{3}$
majorizer	θ_0^{-4}	θ_0^{-3}	θ_0^{-2}	θ_0^{-1}	$\theta_0^{-1}\theta_3$

FIG. 1

By using the majorizing terms listed at the bottom of the chart we can say

$$T_2 \precsim P_7(\xi)(\xi + \gamma)(\mu^* + \delta + \lambda^*)\theta_0^{-1}[\theta_3 \mid {}^{0,1}_{3,4}].$$

Because the pattern of powers of θ_0 does not fit into our notation scheme we have simply used the highest power (θ_0^{-1}) which occurs.

Now if we add the T_1 and T_2 estimates we get an estimate for $z(\theta_3) - z_0$:

$$\text{(B43)} \qquad \begin{aligned} z(\theta_3) - z_0 &\precsim P_8(\xi)(1 + \xi + \gamma)(\mu^* + \delta + \lambda^*)[\theta_3 \mid {}^{0,1}_{3,4}], \quad \text{or} \\ z - z_0 &\precsim \beta^*[{}^{0,1}_{3,4}]. \end{aligned}$$

Since $z = z_0 + (z - z_0)$ we can say

$$\text{(B44)} \qquad \begin{aligned} z &\precsim (\alpha + \beta^*)[{}^{0,1}_{3,4}] \\ &\precsim \xi^*[{}^{0,1}_{3,4}]. \end{aligned}$$

To conclude the rederivation of bounds we must consider the requirement (B22). We have stipulated $\theta_0 \geqq \theta_a$ so that

$$S_\theta z_0 - z_0 \leqq \varepsilon/2[{}^0_2].$$

Applying (A15) to (B43) we see that

$$\zeta - S_\theta z_0 = S_\theta(z - z_0) \leqq C_{15}\beta^*[{}^{-3,1}_{0,4}].$$

Adding inequalities,

$$\text{(B45)} \qquad \begin{aligned} \zeta - z_0 = (\zeta - S_\theta z_0) + (S_\theta z_0 - z_0) &\leqq (C_{15}\beta^* + \varepsilon/2)[{}^0_2] \\ &\leqq \varepsilon^*[{}^0_2]; \end{aligned}$$

we hope, of course, that $\varepsilon^* \leqq \varepsilon$.

Fig. 13.2 Page 43 from Nash's 1956 paper [165]. *Reprinted with the permission of the Annals of Mathematics*

We give here a more precise example of a formal statement of this type of implicit function theorem in a special case (this is from Schwartz's lecture notes [210], which illustrates the complexity of the estimates involved). We continue to assume that M is a compact differentiable manifold of dimension n, and we let C^m denote the Banach space of C^m real-valued functions on M.

Theorem 13.3 *Let F be a mapping from the unit ball B in C^m into $C^{m-\alpha}$. Suppose that:*

- *F has two continuous Fréchet derivatives, both bounded by M.*
- *There exists a mapping L with domain B and range in the space $\mathcal{B}(C^m, C^{m-\alpha})$ of bounded operators from C^m to $C^{m-\alpha}$, such that:*

$$|L(f)h|_{m-\alpha} \leq M|h|, \;\; f \in B, h \in C^m, \tag{13.22}$$

$$F'(f)L(f)h = h, \;\; f \in B, h \in C^{m+\alpha}, \tag{13.23}$$

$$|L(f)F(f)|_{m+9\alpha} \leq M(1 + |f|_{m+10\alpha}), \;\; f \in C^{m+10\alpha}. \tag{13.24}$$

Then, if

$$|F(0)|_{m+9\alpha} \leq 2^{-40} M^{-202},$$

F(B) contains a neighborhood of the origin.

The proof of this theorem (see [210]) uses a family of suitably-defined smoothing operators T_n indexed by an integer n and the successive approximations (13.21) and the initial estimate $f_0 = 0$.

13.5 Approximation of a Metric by an Induced Metric

Step 3 of Nash's proof of his embedding theorem consists of showing that a given Riemannian metric g on a compact differentiable manifold M of dimension n can be approximated in the C^3-norm by a Riemannian metric $y^*(\rho)$ induced by a mapping $y : M \to \mathbf{R}^L$, where ρ is the Euclidean metric on $\mathbf{R}^L$. We will outline this approximation, following Nash's exposition in his paper.

First, Nash uses an aspect of Whitney's embedding theorem that we mentioned earlier in Sect. 12.4. Namely, Whitney not only showed that a C^k-manifold could be embedded by a C^k-mapping

$$f : M \to \mathbf{R}^{2n},$$

as a C^k-submanifold of $\mathbf{R}^{2n}$, but that the *image* $f(M)$ could be realized as a *real-analytic submanifold* of $\mathbf{R}^{2n}$. Even if M had a real-analytic structure, it was unknown until 1958 if there was a real-analytic mapping onto its image $f(M)$, as we will see in Sect. 15.7. However, we will use (as Nash did) the real-analytic structure of its image

$f(M)$, i.e., represented by the zeros of real-analytic functions with nonvanishing Jacobian determinants at each point of $f(M) \subset \mathbf{R}^{2n}$.

We now suppose that this is the case, and we let A denote the image $f(M)$ as a real-analytic submanifold of $\mathbf{R}^a$, where for convenience, we let $a = 2n$. Our given metric g then becomes a C^k-metric on the submanifold A, and we need to find a mapping y of A into $\mathbf{R}^L$, for some L, which induces a metric $\tilde{g}$ on A that is a C^3 approximation of g on A.

The metric g on A is a symmetric two-tensor, represented in terms of local coordinates $(x_1, \ldots, x_n)$ on A as a symmetric matrix $g_{ij}(x)$. Let $(u_1, \ldots, u_a)$ be Euclidean coordinates in the ambient space $\mathbf{R}^a$. We want to use the ambient coordinates to create a basis at each point of A for the symmetric tensors at that point, and this will then yield a specific representation of the metric g_{ij} in terms of this basis and the ambient coordinates. We will continue to work in the coordinate chart on A with the coordinates $(x_1, \ldots, x_n)$, but the assertions we will be making will hold for all points of the submanifold A.

Let us illustrate the kind of basis for symmetric two-tensors we will be using by giving a simple example in $\mathbf{R}^2$. Suppose that (u_1, u_2) are coordinates in $\mathbf{R}^2$, and consider the functions

$$f^{ij} = u_i + u_j, i \leq j.$$

We can enumerate these three functions with a single index r in the following manner:

$$\begin{aligned} \psi^1 &= f^{11} = u_1 + u_1, \\ \psi^2 &= f^{12} = u_1 + u_2, \\ \psi^3 &= f^{22} = u_2 + u_2. \end{aligned}$$

Each of the functions $\psi^r(u)$ induces a symmetric two-tensor of the form

$$M_{ij}^r = \frac{\partial \psi^\alpha}{\partial u_i} \frac{\partial \psi^\alpha}{\partial u_j}, \ \alpha = 1, 2, 3, \ 1 \leq i, j \leq 2.$$

Writing this out explicitly in this case, we see that

$$\begin{aligned} M_{ij}^1 &= \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \\ M_{ij}^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \\ M_{ij}^3 &= \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

It is clear that these matrices are a basis at each point of $\mathbf{R}^2$ for the three-dimensional vector space of all symmetric 2×2 matrices. This will be a model for what we now proceed to carry out in the general case.

For $A \subset \mathbf{R}^a$, we let $\tilde{\psi}^r(u)$ be a denumeration of the $\frac{a(a+1)}{2}$ functions on $\mathbf{R}^a$,

$$f^{kl} := u_k + u_l,\ 1 \le k \le l \le a. \tag{13.25}$$

We can restrict the functions $\tilde{\psi}^r$ to the submanifold A and by differentiation create the symmetric matrices

$$\tilde{M}^r_{ij} = \frac{\partial \tilde{\psi}^r}{\partial x_i} \frac{\partial \tilde{\psi}^r}{\partial x_j}. \tag{13.26}$$

Let

$$s = \frac{n(n+1)}{2}.$$

This integer denotes the dimension of the vector space of symmetric two-tensors at each point of A. Note that in (13.26) we are differentiating the ambiently defined functions $\tilde{\psi}^r$ with respect to the local coordinates on A (in the intrinsic geometry of A), and these symmetric matrices span the s-dimensional vector space of symmetric two-tensors at each point of this coordinate chart on A (one can choose n of the coordinate functions u_k to be coordinates for the submanifold A near each point of A).

At this point we could use these functions $\{\tilde{\psi}^r\}$ and the associated tensors $\{\tilde{M}^r_{ij}\}$ to conclude Nash's proof of Step 3 of his embedding theorem proof, but, as he remarks, this would increase the embedding dimension for the final theorem unnecessarily. Therefore he introduces an analogously defined, but smaller set of functions ψ^r, which will work just as well.

Namely, Nash proposes to use a more general set of linear combinations of the coordinate functions[7] of the form

$$\psi^r = \sum_\rho C^r_\rho u_p, r = 1, \dots, n+s,$$

where $\beta = 1, \dots, a$, and $s = \frac{n(n+1)}{2}$, as above. The coefficients $C^r_\rho \in \mathbf{R}$ are yet to be chosen. We choose these coefficients so that the $n+s$ functions ψ^r have the property that

$$M^r_{ij} := \frac{\partial \psi^r}{\partial x_i} \frac{\partial \psi^r}{\partial x_j}$$

forms a spanning set of vectors for $\mathrm{Sym}(A)$ at each point of A. By a careful dimensional analysis, Nash shows that there are a sufficient number of coefficients $C^r_\rho \in \mathbf{R}$, so that the tensors M^r_{ij} forms a basis at each point of A. We omit this argument here. We note that the number of ψ^α is $n + s = n + \frac{n(n+1)}{2}$, which is less than the number of $\tilde{\psi}^\alpha$, which is $\frac{a(a+1)}{2} = \frac{2n(2n+1)}{2}$.

We can use the M^r_{ij} to represent our given metric g_{ij} at each point of A, namely,

[7] This is the same sort of perturbation Nash used in Step 1 to lower the embedding dimension for a nondegenerate embedding.

$$g_{ij}(x) = \sum_r \alpha_r(x) M^r_{ij},$$

where the $\alpha_r(x)$ are to be chosen as smooth C^k-coefficients of this pointwise representation. There is *a priori* not a unique set of such coefficients, but if we proceed as we did in Sect. 13.4 (see Eq. (13.18)), we can choose the coefficients $\{\alpha_r\}$ so that

$$\sum_r |\alpha_r|^2 \text{ is minimum.}$$

Minimizing the norm in this way gives a unique set of coefficients well defined on all of A.

Since A is a real-analytic compact submanifold of $\mathbf{R}^a$, it follows by an approximation theorem of Whitney [244] that each of the C^k-functions $\alpha_r(x)$ can be approximated by real-analytic functions $a_r(x)$ in the C^3 norm. Thus we obtain a C^3 approximation to the metric g_{ij} of the form

$$\tilde{g}_{ij} = \sum_r a_r M^r_{ij} = \sum_r a_r \frac{\partial \psi^r}{\partial x_i} \frac{\partial \psi^r}{\partial x_j} \approx g_{ij}. \tag{13.27}$$

Now we need to find a mapping of A to a Euclidean space that will in turn yield a good approximation to $\tilde{g}_{ij}$.

Nash introduces a nice trick at this point. Consider the $2(n+s)$ functions of the form

$$y_r = \frac{(a_r)^{\frac{1}{2}}}{\lambda} \sin(\lambda \psi^r),$$

$$\tilde{y}^r = \frac{(a_r)^{\frac{1}{2}}}{\lambda} \cos(\lambda \psi^r),$$

where λ is a positive parameter and a_r and ψ^r are defined as above. The mapping

$$y = (y_r, \tilde{y}_r) : A \to \mathbf{R}^{2(n+s)}$$

induces a metric on A of the form

$$g_y = \sum_r \frac{\partial y_r}{\partial x_i} \frac{\partial y_r}{\partial x_j} + \sum_r \frac{\partial \tilde{y}^r}{\partial x_i} \frac{\partial \tilde{y}^r}{\partial x_j}.$$

Expanding this expression, there is a nice cancellation: the λ^{-1} terms involve $\sin(\lambda\psi^r)\cos(\lambda\psi^r)$ and $-\sin(\lambda\psi^r)\cos(\lambda\psi^r)$ and these terms cancel; and the remaining terms combine using $\sin^2 + \cos^2 = 1$. This yields

$$g_y = \sum_r a_r \frac{\partial \psi^r}{\partial x_i} \frac{\partial \psi^r}{\partial x_j} + \lambda^{-2} h_{ij},$$

where $\{h_{ij}\}$ are real-analytic functions on A. Then we have, from (13.27),

$$(g_y))ij = \tilde{g}_{ij} + \lambda^{-2} h_{ij}.$$

By choosing λ sufficiently large, we find that

$$g_y \approx \tilde{g},$$

in the C^3-norm.

Thus we have the sequence of C^3 approximations

$$g \approx \tilde{g} \approx g_y,$$

and we obtain the final result that our given metric g on M can be approximated on M in the C^3-norm by a metric on M induced from a mapping. Note that for this final step, we are composing the two mappings:

$$M \xrightarrow{f} A \xrightarrow{y} \mathbf{R}^{2(n+s)},$$

where $A = f(M) \subset \mathbf{R}^{2n}$ and f is our original Whitney embedding that started this argument. This completes our outline of Step 3 of Nash's proof and hence finishes our summary of Nash's proof of the isometric embedding theorem for a compact Riemannian manifold M.

We conclude by observing that for a compact n-dimensional Riemannian manifold, we have obtained in this Step 3 a mapping of the form

$$y : M \to \mathbf{R}^{2(n+s)} = \mathbf{R}^{n^2+3n}.$$

13.6 Closing Remarks

In this section, we have concentrated on Nash's embedding theorem paper of 1956 [165]. Two years earlier Nash published a proof of a C^1 isometric embedding theorem [164] which was simpler and did not involve his deep implicit function theorem. Note that the paper we have considered yielded C^k embeddings for all $k \geq 3$, including $k = \infty$.

In 1962 Serge Lang (1927–2005) presented a simplification in a Seminaire Bourbaki lecture [133] of the proof of the embedding theorem (Lang concentrated on the implicit function theorem), and he used an idea that had been suggested by Adriano Garsia that one could consider the special case of an isometric embedding of a torus (all of this is for the compact case). Namely, if one is given a Riemannian manifold M, then one can, by Whitney, embed it as a submanifold of a Euclidean space $\mathbf{R}^{2n}$ (which we continue to call M), and one can consider this Euclidean space to be an affine piece of a real $2n$-torus T. The differentiable metric g on M can be extended as a smooth metric $\tilde{g}$ to all of T by the usual partition of unity[8] arguments. If one then had an isometric embedding f of the torus T with the extended metric $\tilde{g}$ into some Euclidean space $\mathbf{R}^N$, then the restriction of the mapping f to M yields the desired isometric embedding of M. Any such torus has one coordinate chart covering the manifold (using the periodicity), and this simplifies some of the arguments involving smoothing operators, etc., which Lang exploits.

In 1964–65 Schwartz gave lectures at NYU on nonlinear functional analysis [210], in which he surveyed earlier infinite-dimensional implicit function theorems and gave a complete proof of Nash's implicit function theorem as well as the embedding theorem. He also exploited the idea of embedding a torus.

In the meantime, there have been, of course, many other developments concerning such isometric embedding theorems, which we won't try to discuss here. See, for instance, the recent survey by Andrews [8] and the book by Han and Hong [95].

[8] See, e.g., Milnor [154], Hirzebruch [104], Wells [239], or other references concerning manifolds, for a description of partition of unity and its use in geometry.

Chapter 14
Compact Complex Manifolds

14.1 Introduction

In the previous two chapters, we have looked at the embedding of real smooth manifolds into Euclidean space: Whitney in the general case and Nash for the case of isometric embeddings. In this and the following chapter, we are going to be dealing with complex manifolds and holomorphic embeddings. The case of real-analytic embeddings will be considered as the restriction of holomorphic mappings to real-analytic submanifolds of suitable types, as we will see in Sect. 15.7.

We will consider in this and the next chapter two fundamentally different kinds of holomorphic embedding theorems. We note that if X is a compact complex manifold, then there cannot be an embedding into a complex Euclidean space, as any holomorphic function on a compact complex manifold is necessarily a constant (a simple consequence of the maximum principle for holomorphic functions). And yet, there are many important examples of compact complex manifolds, most notably compact Riemann surfaces. The simplest examples of compact complex manifolds are simply closed holomorphic submanifolds of complex projective space $\mathbf{P}_N$.

So the question arose: which compact complex manifolds of dimension n could be embedded into a projective space of some dimension, as this is the natural extension of complex Euclidean space to a compact complex manifold. In 1954 Kunihiko Kodaira (1915–1997) gave a necessary and sufficient condition on a compact complex manifold that it be embeddable holomorphically into $\mathbf{P}_N$, which we will describe in some detail in this chapter. This had been conjectured by William Vallance Douglas Hodge (1903–1975) at the International Congress of Mathematicians in 1950 [110].

Kodaira published his embedding theorem in 1954, and the very same year Fritz Hirzebruch announced his generalization of the Riemann–Roch theorem for projective algebraic manifolds of arbitrary dimension. Both Kodaira and Hirzebruch used extensively the new tools of sheaf cohomology theory that developed after the second world war and which we survey in Sect. 14.3 below.

In Sect. 14.7 we show how the classical Riemann–Roch theorem from the nineteenth century that we reviewed in Sect. 10.4 can be reformulated in terms of sheaf

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_14

cohomology. In particular, we outline Serre's elegant proof of this theorem from 1955. We then give Hirzebruch's theorem, which is formulated in terms of sheaf cohomology and which used the theory of Chern classes of holomorphic vector bundles to represent the algebraic topological part of the theorem. We conclude this section with a description of the Atiyah–Singer index theorem from 1963 which shows as a special case that Hirzebruch's theorem is valid for arbitrary compact complex manifolds.

The remaining question as to which noncompact complex manifolds can be holomorphically embedded in $\mathbf{C}^N$ was solved a few years later by Reinhold Remmert (1930–2016), where again a complete characterization of such manifolds was given. We will discuss this in the concluding chapter of the book, Chap. 15.

We will need to introduce some mathematical concepts and results in order to formulate Kodaira's and Hirzebruch's theorems. Our main interest will be compact complex manifolds, which are a special case of oriented compact differentiable manifolds. Let M be a connected compact orientable differentiable manifold of dimension m. As we discussed in Sect. 11.3, Poincaré introduced at the end of the nineteenth century the fundamental concepts of algebraic topology that were to play such an important role in the theory of manifolds in the twentieth century. This includes the theory of homology, cohomology and homotopy, in particular. The fundamental cohomology groups of M can be described in terms of triangulations, or more commonly, singular homology (and numerous other methods of describing these topological invariants). For our purposes, we want to consider cohomology groups with integral coefficients[1] as well as with complex coefficients, which are denoted by

$$H^q(M, \mathbf{Z}) \overset{i}{\hookrightarrow} H^q(M, \mathbf{C}), \;\; q = 0, \ldots, m, \tag{14.1}$$

where the injective mapping i is the natural inclusion mapping, and all higher-degree cohomology groups vanish. Since M is orientable, and using Poincaré duality, we know that

$$H^0(M, \mathbf{Z}) \cong H^m(M, \mathbf{Z}) \cong \mathbf{Z}. \tag{14.2}$$

In 1931 Georges de Rham (1903–1990) formulated and proved his well-known and important de Rham theorem for differentiable manifolds [52], which we will discuss shortly (he discussed this and other topics at length in his later monograph [53]). In 1933 Erich Kähler (1906–2000) introduced a Hermitian metric on a complex manifold with a "remarkable" property ("Über eine bemerkenswerte Hermitsche Metrik",[2] [118]), which we now call a Kähler metric on a complex manifold. Using

---

[1] See, for instance, the classic textbook by Eilenberg and Steenrod [59] or any standard text on algebraic topology for the concepts and theory of cohomology groups with constant coefficients; we will discuss and use extensively sheaf cohomology, a substantive generalization of classic cohomology, in these last two chapters of this book.

[2] "Concerning a remarkable Hermitian metric".

this special type of metric, Kähler was able to derive an interesting set of properties for the algebraic topology of projective algebraic manifolds.[3]

One year later, Hodge published the first in a series of three papers [106–108], which, using the work of de Rham, created a theory of harmonic differential forms that could represent cohomology in a unique manner using the additional structure of a metric on a given differentiable manifold. This work was used by Hodge in his important monograph applying his theory of harmonic forms to projective algebraic manifolds [109], and this was then extended to the more abstract setting by later authors to Kähler manifolds, in particular in the work of Kodaira, which we are discussing in this chapter.

We now summarize some of the results we need concerning de Rham's theorem and Kähler geometry to formulate the statement of Kodaira's embedding theorem.[4]

Let M be a compact orientable differentiable manifold of dimension m as before, and let $\mathcal{E}^k(M)$ denote the differential forms of degree k on M. Locally such a differential form $\varphi(x)$ is represented in a coordinate chart U with coordinates $(x_1, \ldots, x_n)$ as

$$\varphi(x) = \sum a_{i_1 \ldots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{x_k}, \tag{14.3}$$

where

$$1 \leq i_1 < \cdots < i_k \leq m$$

are increasing multiindices and $a_{i_1 \ldots i_k}(x)$ are complex-valued differentiable (C^∞) functions on U. Let d represent the exterior differentiation operator which maps k-forms to $k + 1$-forms. We have the complex

$$0 \to \mathcal{E}^0(M) \xrightarrow{d} \mathcal{E}^1(M) \xrightarrow{d} \mathcal{E}^2(M) \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{E}^m(M) \xrightarrow{d} 0, \tag{14.4}$$

where we recall that the exterior derivative is defined by

$$d\varphi = \sum d(a_{i_1 \ldots i_k}(x)) \wedge dx_{i_1} \ldots dx_{i_k}, \tag{14.5}$$

and

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i,$$

[3] In 1949 Wei-Liang Chow (1911–1995) [47] showed that any holomorphic submanifold of complex projective space is, in fact, algebraic, i.e., defined by algebraic functions on projective space. We will simply say projective algebraic manifold for any such holomorphic submanifold of complex projective space.

[4] Basic references for Hodge theory on Kähler manifolds for this chapter concerning Kodaira's embedding theorem include the books by Weil [238], Griffiths and Harris [91], and Wells [239]. These last two references include proofs of Kodaira's embedding theorem.

for a function f, and the multiplication in (14.5) is carried out using exterior algebra products.

We define the *closed k-forms* on M to be

$$Z^k(M) := \{f \in \mathcal{E}^k(M) : d\varphi = 0\},$$

and the *exact k-forms* on M to be

$$B^k(M) := \{\varphi \in \mathcal{E}^k(M) : \varphi = d\psi \text{ for some } \psi \in \mathcal{E}^{k-1}(M)\}.$$

It is well known that the composed differential operator $d^2 = 0$ in the sequence (14.4), which is also easy to verify (which is why we called the sequence (14.4) a complex).

We now define the *de Rham groups* on M to be

$$H^k_{\mathrm{dR}}(M) := Z^k(M)/B^k(M). \tag{14.6}$$

The fundamental de Rham theorem in his 1931 paper [52] asserts that

$$H^k(M, \mathbf{C}) \cong H^k_{\mathrm{dR}}(M). \tag{14.7}$$

Note that this differential forms representation of cohomology is sufficient for computing the Betti numbers of the manifold M, but it does not detect the more subtle torsion elements of cohomology over the integers.

We now move back to the complex manifold setting, and we consider a local coordinate system $(z_1, \dots, z_n)$ on our given complex manifold X. We have the standard conventions for real and complex coordinates and differentiation:

$$\begin{aligned}
z_\mu &= x_\mu + i y_\mu, \\
\overline{z}_\mu &= x_\mu - i y_\mu, \\
dz_\mu &= dx_\mu + i dx_\mu, \\
d\overline{z}_\mu &= dx_\mu - i dx_\mu, \\
\frac{\partial}{\partial z_\mu} &= \frac{1}{2}\left(\frac{\partial}{\partial x_\mu} - i\frac{\partial}{\partial y_\mu}\right), \\
\frac{\partial}{\partial \overline{z}_\mu} &= \frac{1}{2}\left(\frac{\partial}{\partial x_\mu} + i\frac{\partial}{\partial y_\mu}\right),
\end{aligned}$$

and the standard exterior derivative operators:

$$d = \partial + \overline{\partial},$$

where

$$\partial = \sum_{\mu} \frac{\partial}{\partial z_\mu} dz_\mu,$$

$$\overline{\partial} = \sum_{\mu} \frac{\partial}{\partial \overline{z}_\mu} d\overline{z}_\mu.$$

If a real differential form on X is expressed as a linear combination (with smooth coefficients) of wedge products of the real differentials dx_μ and dy_μ as in (14.3), then this can be reexpressed in terms of the differentials dz_μ and $d\overline{z}_\mu$. Thus it follow that any differential form φ of degree k on X can be expressed as a sum of differential forms of type (p, q), where $p + q = k$, as in the following expression:

$$\varphi = \sum a_{\mu_1 \ldots \mu_p \nu_1 \ldots \nu_q} dz_{\mu_1} \cdots \wedge dz_{\mu_p} \wedge d\overline{z}_{\nu_1} \wedge \cdots \wedge d\overline{z}_{\nu_q}, \tag{14.8}$$

with the increasing double multiindex $\mu_1 \ldots \mu_p \nu_1 \ldots \nu_q$. We will have occasion to use this concept again further below, and we introduce here the simplified multiindex notation,

$$\varphi = \sum a_{M,N} dz_M \wedge d\overline{z}_N, \tag{14.9}$$

where $M = (\mu_1, \ldots, \mu_p)$ and $N = (\nu_1, \ldots, \nu_q)$ represent the increasing multi-indices in a natural manner. We will let $|M| = p$, $|N| = q$, as is customary.

We define the vector space of all forms of type (p, q) on X to be $\mathcal{E}^{p,q}(X)$, and we have the direct sum decomposition

$$\mathcal{E}^k(X) = \sum_{p+q=k} \mathcal{E}^{p,q}(X). \tag{14.10}$$

This direct sum will play an important role in the theory of Kähler manifolds, which we will be discussing shortly. Note that it only depends on the complex structure of the complex manifold X. The exterior derivative operator $\overline{\partial}$ maps (p, q)-forms to $(p, q + 1)$ forms, i.e.,

$$\overline{\partial} : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q+1}(X), \tag{14.11}$$

as is easy to verify.

We let $\Omega^p(X)$ denote the *holomorphic* p*-forms* on X, $0 \leq p \leq n$. Locally $\varphi \in \Omega^p(X)$ has the form

$$\varphi(x) = \sum_{|I|=p} a_I(x) dz^I,$$

where the coefficients $a_I(x)$ are holomorphic functions. In particular,

$$\Omega^p(X) = \ker \overline{\partial} : \mathcal{E}^{p,0}(X) \to \mathcal{E}^{p,1}(X).$$

We now suppose that our complex manifold X is equipped with a Hermitian metric. This is simply a smoothly varying Hermitian inner product $h(X, Y)_x$ on the tangent space $T_x(X)$ at each point $x \in X$ (which is a complex vector space of dimension n). On $\mathbf{C}^n$ the standard Hermitian metric is given by (using matrix multiplication where $z \in \mathbf{C}^n$ is a row vector)

$$(z, w) = z\overline{w}^t,$$

and a more general and varying Hermitian metric on $\mathbf{C}^n$ is given by

$$(z, w)_x = zH(x)\overline{w}^t,$$

where $H(x)$ is a smoothly varying Hermitian matrix on $\mathbf{C}^n$. Such local Hermitian metrics can be patched together on any complex manifold with a partition of unity, just as in the case of Riemannian manifolds, and any complex manifold can therefore be equipped with a Hermitian metric.

If X has a Hermitian metric h, which we express locally by

$$ds^2 = \sum_{\mu,\nu} h_{\mu\nu}(x) dz_\mu d\overline{z}_\nu,$$

where $h_{\mu\nu}(x)$ is a smoothly varying Hermitian matrix, then we define the *fundamental two-form* Ω *associated to* h to be

$$\Omega = \frac{i}{2} \sum_{\mu\nu} h_{\mu\nu}(x) dz_\mu \wedge d\overline{z}_\nu. \tag{14.12}$$

This is a globally defined two-form on X uniquely associated to the given Hermitian metric h on X. Note that it is a two-form of type $(1, 1)$. One can formulate both the Hermitian metric and its associated two-form in an invariant manner, and we leave such a description to the reference books we mentioned earlier.

An Hermitian metric h on a complex manifold X is said to be a *Kähler metric* if the fundamental two-form Ω is a closed form, that is

$$d\Omega = 0. \tag{14.13}$$

We say that a complex manifold is a *Kähler manifold* if it is equipped with a Kähler metric. The standard Hermitian metric on $\mathbf{C}^n$ has the associated fundamental two-form

$$\Omega = \frac{i}{2}\sum_{\mu} dz_{\mu} \wedge d\bar{z}_{\mu},$$

and this satisfies $d\Omega = 0$, since the coefficients of Ω are constants. This is therefore, trivially, a Kähler manifold.

A more complicated example is given by the standard *Fubini–Study metric* on $\mathbf{P}_n$ (letting $w_{\mu} = \frac{\xi_{\mu}}{\xi_0}$ be affine coordinates in the coordinate chart on $\mathbf{P}_n$ defined by $\xi_0 \neq 0$ in the homogeneous coordinates $(\xi_0, \ldots, \xi_n)$):

$$\Omega(w) = \frac{i}{2}\frac{(1+|w|^2)\sum_{\mu=1}^{n} dw_{\mu} \wedge dw_{\mu} - \sum_{\mu,\nu=1}^{n} \overline{w}_{\mu} w_{\nu} dw_{\mu} \wedge d\overline{w}_{\nu}}{(1+|w|^2)^2}. \qquad (14.14)$$

This can also be shown to be a closed two-form of type (1, 1) on $\mathbf{P}_n$ (see p. 190 of [239]). Thus complex Euclidean space and complex projective space provide two fundamental examples of Kähler manifolds. It is easy to verify that any complex submanifold of a Kähler manifold is a Kähler manifold, providing many more examples (in particular, any projective algebraic manifold is a Kähler manifold).

We now define an important special case of a type of Kähler manifold called a *Hodge manifold*. The formulation of the notion of a Hodge manifold is due to André Weil (1906–1998) [237] and it plays a critical role in Kodaira's theorem, as we will see shortly. Namely, if X is a Kähler manifold with fundamental form Ω, then Ω defines by the de Rham theorem an element of the cohomology group $\omega \in H^2(X, \mathbf{C})$, and we say that it is a *Hodge metric* on X if ω is in the image of the inclusion mapping

$$H^2(X, \mathbf{Z}) \overset{i}{\hookrightarrow} H^2(X, \mathbf{C}).$$

We say then that a Hodge manifold is a complex manifold equipped with a Hodge metric h. We can now formulate Kodaira's embedding theorem.

Theorem 14.1 (Kodaira Embedding Theorem [128]) *Let X be a compact complex manifold. Then there is an embedding of X into $P_N(\mathbf{C})$ for some integer N if and only if X is a Hodge manifold.*

In Fig. 14.1 we present the first page of Kodaira's embedding theorem paper.

If X is a projective algebraic manifold, then restricting the fundamental Fubini–Study two-form Ω on projective space (14.14) to X makes X into a Hodge manifold.

To show that a Hodge manifold admits an embedding is much more difficult, and we will outline the proof of that in the next several sections. In Sect. 14.6 we will outline the basic embedding proof, which uses sections of specific line bundles to realize the embedding. This will depend on the existence of sufficient numbers of sections of the line bundles to ensure an embedding, just as in the Whitney theorem, we needed a sufficient number of smooth functions for the embedding. The existence of a sufficient number of sections of such bundles depends on a specific theorem, also due to Kodaira, which is outlined in Sect. 14.5 and is referred to as the Kodaira vanishing theorem and depends on the use of curvature estimates and Hodge's theory

ANNALS OF MATHEMATICS
Vol. 60, No. 1, July, 1954
Printed in U.S.A.

ON KÄHLER VARIETIES OF RESTRICTED TYPE* (AN INTRINSIC CHARACTERIZATION OF ALGEBRAIC VARIETIES)

BY K. KODAIRA

(Received March 18, 1954)

A compact complex analytic variety V is called a *Kähler variety of restricted type*[1] or a *Hodge variety*[2] if V carries a Kähler metric $ds^2 = 2\sum g_{\alpha\bar\beta}(dz_\alpha d\bar z_\beta)$ such that the associated exterior form $\omega = i\sum g_{\alpha\bar\beta}dz_\alpha d\bar z_\beta$ belongs to the cohomology class of an *integral* 2-cocycle on V. In what follows such a metric will be called a *Hodge metric*. The main purpose of the present paper is to prove that *every Hodge variety is* (*bi-regularly equivalent to*) *a non-singular algebraic variety imbedded in a projective space.*[3] Since every non-singular algebraic variety in a projective space carries a Hodge metric, our main result may also be stated in the following form:

A compact complex analytic variety V is (*bi-regularly equivalent to*) *a non-singular algebraic variety imbedded in a projective space if and only if V can carry a Hodge metric.*

This paper[4] is divided into four sections. In Section 1 we give a summary of some known results concerning complex line bundles over Kähler varieties. Section 2 is concerned with quadratic transformations. The proof of our main theorem (Theorem 4) is given in the following Section 3. The final Section 4 is devoted to several applications of our main theorem. In particular, we prove that *a compact complex analytic variety is a non-singular algebraic variety imbedded in a projective space if it carries a Hermitian metric whose Ricci curvature is everywhere negative* [*or positive*] *definite* (Theorem 5). We also derive the following result: *Let $\mathfrak{B}$ be a bounded domain in the space of n complex variables and let Δ be a discontinuous group of analytic automorphisms without fixed points of $\mathfrak{B}$ such that $\mathfrak{B}/\Delta$ is compact. Then the factor space $\mathfrak{B}/\Delta$ is a non-singular algebraic variety imbedded in a projective space* (Theorem 6).

The author wishes to express his sincere thanks to A. Borel for valuable suggestions given during the period of the preparation of this paper.

---

* This work was supported by a research project at Princeton University sponsored by the Office of Ordnance Research, U. S. Army.

[1] W. V. D. Hodge, *A special type of Kähler manifolds*, Proc. London Math. Soc., 1 (1951), pp. 104–117.

[2] A. Weil, *On Picard varieties*, Amer. J. Math., 74 (1952), pp. 865–894.

[3] This gives, in some sense, an answer to a problem proposed by Hodge. See W. V. D. Hodge, *The topological invariants of algebraic varieties*, Proc. International Congress of Math., 1 (1950), pp. 182–192, §8.

[4] The results of the present paper have been announced in Proc. Nat. Acad. Sci., U. S. A., Vol. 40 (1954), pp. 313-316.

28

Fig. 14.1 First page from Kodaira's projective algebraic manifold characterization paper from 1954 [128]. Reprinted with the permission of the Annals of Mathematics of harmonic differential forms. Sheaf theory and sheaf cohomology play a critical role in this proof along with Hodge theory on Kähler manifolds, and these will be outlined in Sects. 14.3 and 14.4, as needed for the development of the proof.

Remark At the beginning of this Introduction we noted that embedding theorems for complex manifolds consisted of two distinct cases: compact complex manifolds (Kodaira's embedding theorem) and noncompact complex manifolds (the Stein manifold embedding theorem, which will be discussed in the next chapter). There are many examples of compact complex manifolds which do not satisfy the Hodge manifold criterion in Kodaira's theorem, and there are many examples of noncompact complex manifolds that do not satisfy the criterion to be a Stein manifold.

However, in the important case of a complex manifold X of complex dimension 1, a Riemann surface, one of the two criteria is always satisfied. Namely, for some suitable $N \geq 3$:

1. If X is compact, then X is a Hodge manifold, and hence is embeddable as a closed submanifold of $\mathbf{P}_N$.
2. If X is noncompact, then X is Stein, and hence is embeddable as a closed submanifold of $\mathbf{C}^N$.

More simply put, a Riemann surface X is embeddable as a closed submanifold of either $\mathbf{P}_N$ or $\mathbf{C}^N$, depending on whether X is compact or not.

For a compact Riemann surface X, this is quite easy to see, as was pointed out in Kodaira's original paper [128]. Namely, if h is any Hermitian metric on X (which, as we have seen before, always exists), then let Ω be its associated fundamental $(1, 1)$-form. Since $\dim_{\mathbf{R}} X = 2$, it follows that $d\Omega = 0$, and hence h is a Kähler metric. Moreover, if

$$c = \int_X \Omega,$$

then $c \neq 0$, since X is orientable. Letting $\tilde{\Omega} = (1/c)\Omega$, we see that $\tilde{\Omega}$ is a fundamental form associated to a Hodge metric, and hence X is a Hodge manifold.

If X is noncompact, then a fundamental existence theorem for Riemann surfaces is that there exists a nonconstant meromorphic function f on X (this is a consequence of the Riemann–Roch theorem, see Sect. 10.4). This was proved in Weyl's 1913 book [241] (see also the book by Farkas and Kra [72], which has a modernized proof of this using the Dirichlet principle ideas, which Weyl also utilized). The function f exhibits X as a branched covering of an open subset of $\mathbf{P}_1$, which is Riemann's original description of a Riemann surface. In this context, Behnke and Stein showed in 1947 [14] that there is a generalization of Runge's approximation theorem for domains in the complex plane (as was discussed in Sect. 9.6). It follows from this that X is a Stein manifold, which is formally defined in the following chapter, Chap. 15.

14.2 Holomorphic Line Bundles

The primary tool Kodaira uses for the construction of an embedding from a compact complex manifold into projective space is the theory of holomorphic line bundles. This theory is developed in several of the references we have cited, and we will introduce some of the terminology and fundamental ideas briefly here.

Let X be a complex manifold. A *holomorphic line bundle* E over X is a complex manifold E with a mapping $\pi : E \to X$, such that π is regular and surjective, and each fibre of the mapping $E_x := \pi^{-1}(x)$, for $x \in X$, has the structure of a one-dimensional complex vector space (these are the "lines" of the line bundle). Moreover, it is required that there is an open covering U_α of X such that there are biholomorphic mappings h_α satisfying

$$\begin{aligned}&h_\alpha : \pi^{-1}(U_\alpha) \cong U_\alpha \times \mathbf{C},\\ &h_{\alpha|E_x} : E_x \to \{x\} \times \mathbf{C} \text{ is complex-linear.}\end{aligned}$$

A *section* of a holomorphic line bundle (differentiable or holomorphic section) on an open set $U \subset X$ is a mapping (differentiable or holomorphic)

$$s : U \to E,$$

such that $\pi \circ s = \text{id}$ is the identity mapping on U. A holomorphic line bundle $E \to X$ is said to be *trivial* if $E \cong X \times \mathbf{C}$, and the sections of a trivial bundle are simply complex-valued functions on open sets on X. The mappings h_α are called *local trivializations* of the line bundle E (and play the role of coordinate charts on a manifold), and sections of the line bundle can be considered locally as ordinary functions by means of the trivializations (again, just as functions on a manifold are functions of coordinates in a coordinate chart on a manifold). Let $\mathcal{O}(U, E)$ ($\mathcal{E}(U, E)$) denote the holomorphic (differentiable, i.e., C^∞) sections of a holomorphic line bundle $E \to X$ on the open set $U \subset X$.

We define *transition functions* for a holomorphic line bundle $E \to X$ by

$$h_{\alpha\beta} := h_\alpha \circ h_\beta^{-1} : U_\alpha \cap U_\beta \times \mathbf{C} \to U\alpha \cap U_\beta \times \mathbf{C},$$

which are nonvanishing holomorphic functions on $U_\alpha \cap U_\beta$, assuming this is a non-empty intersection. These transition functions satisfy

$$\begin{aligned}&h_{\alpha\beta} \circ h_{\beta\gamma} \circ h_{\gamma\alpha} = \text{id}, \ \text{ on } U_\alpha \cap U_\beta \cap U_\gamma,\\ &h_{\alpha\alpha} = \text{id on } U_\alpha.\end{aligned} \tag{14.15}$$

It is easy to show that if one has a set of transition functions on X satisfying (14.15), then one can construct an associated holomorphic line bundle E on X.

The notion of a holomorphic line bundle is a special case of the more general notion of a vector bundle on a manifold (differentiable or holomorphic), which has fibres modeled on a vector space (real or complex) of rank r. The transition functions will be nonsingular matrix-valued functions of the appropriate smoothness (differentiable or holomorphic, for instance). More general classes of fibre bundles include where the fibre is a Lie group or a homogeneous space (e.g. a sphere). The book by Norman Steenrod (1910–1971) [219], *The Topology of Fibre Bundles*, from 1951 provided the first systematic description of fibre bundles after its vigorous study over the previous decades. In particular, the theory of characteristic classes of fibre bundles is developed, where a particular characteristic class of a vector bundle $E \to M$, for instance, is an element of an appropriate cohomology group of its base manifold M. Characteristic classes help measure how far a given bundle deviates from being a trivial bundle.

The primary example of a vector bundle is the tangent bundle to a manifold, where the fibres are simply the tangent space at each point. The tangent bundle to a Riemann surface X, denoted by $T(X)$, is a holomorphic line bundle as we defined above, and it is trivial (in the case of a compact Riemann surface) if and only if X has genus 1, i.e., X is a complex torus of dimension one.

Vector bundles are modeled on vector spaces, and as such, most standard vector space operations carry over to vector bundles. For instance, if E and F are vector bundles on a manifold M, then $E \oplus F$ and $E \otimes F$ are well-defined vector bundles. For the case of a line bundle E on a complex manifold X, we will be interested in tensor products of higher order, and we will write, for instance, $E \otimes E$ as E^2, and E^n will denote higher tensor powers. Also we let E^* denote the dual line bundle (each fibre E_x^* is defined as the dual vector space of E_x). Finally, we note that if $T(M)$ is the tangent bundle of a general differentiable manifold M, then the differentiable sections of the exterior algebra bundle $\wedge^k T^*(M)$ on an open set U are the differential forms of degree k on U.

We want to give a few more concrete examples of holomorphic line bundles that are important for our understanding of Kodaira's theorem. First, we describe how to associate a holomorphic line bundle with any holomorphic submanifold Y of codimension one of a complex manifold X (a hypersurface in X). A hypersurface $Y \subset X$ can be described by an open covering U_α of X and holomorphic functions $f \in \mathcal{O}(U_\alpha)$ such that

$$Y \cap U_\alpha = \{x \in U_\alpha : f_\alpha(x) = 0\}, \tag{14.16}$$

and

$$df_\alpha \neq 0 \text{ on } U_\alpha. \tag{14.17}$$

Defining transition functions for a holomorphic line bundle E_Y by

$$g_{\alpha\beta} = \frac{f_\alpha}{f_\beta} \text{ on } U_\alpha \cap U_\beta,$$

we see easily that the conditions (14.15) are satisfied. We note that the nonsingular condition (14.17) which is required in order that Y be a submanifold is not necessary to define the line bundle E_Y. More generally, (14.16) without the nonsingular condition defines a holomorphic subvariety of codimension one, and is an example of a *divisor* which plays an important role in algebraic and analytic geometry. Certain equivalence classes of divisors are in one-to-one correspondence with equivalence classes of holomorphic line bundles (using the natural notion of two vector bundles being equivalent; see [238] or [91], for instance).

Now let Y be a hyperplane in $\mathbf{P}_n$ which is defined by a homogeneous linear equation in homogeneous coordinates (such a hyperplane is referred to as a *hyperplane section*). For instance, if $(\xi_0, \ldots . \xi_n)$ are homogeneous coordinates on $\mathbf{P}_n$, let

$$Y = \{(\xi_0, \ldots, \xi_n\} : \xi_0 = 0\}$$

be an example of such a hyperplane section. Using the standard covering of $\mathbf{P}_n$ by the open sets

$$U_\alpha := \{(\xi_0, \ldots, \xi_n) : \xi_\alpha \neq 0\},$$

then the defining function for Y in each coordinate chart U_α are given by (letting "$\hat{*}$" denote an omitted entry)

$$f_\alpha\left(\frac{\xi_0}{\xi_\alpha}, \ldots, \frac{\hat{\xi}_\alpha}{\xi_\alpha}, \ldots, \frac{\xi_n}{\xi_\alpha}\right) = \frac{\xi_0}{\xi_\alpha},\ \alpha \neq 0,$$

$$g_{\alpha\beta} = \frac{\left(\frac{\xi_0}{\xi_\alpha}\right)}{\left(\frac{\xi_0}{\xi_\beta}\right)} = \frac{\xi_\beta}{\xi_\alpha} \text{ on } U_\alpha \cap U_\beta.$$

Any other hyperplane section would have yielded an equivalent line bundle. We refer simply to the *hyperplane section bundle* H on $\mathbf{P}_n$ for all such equivalent line bundles.

We note that for any holomorphic line bundle E on a complex manifold X with transition functions $\{g_{\alpha\beta}\}$, the transition functions for E^k are given by

$$\{g_{\alpha\beta}^k\},$$

and the transition functions for E^* are given by

$$\{g_{\alpha\beta}^{-1}\}.$$

It follows that the transition functions for H^k and H^* on $\mathbf{P}_n$ are given by

$$g^n_{\alpha\beta} = \left(\frac{\xi_\beta}{\xi_\alpha}\right)^n,$$
$$g^{-1}_{\alpha\beta} = \frac{\xi_\alpha}{\xi_\beta}.$$

Finally, an important holomorphic line bundle on any complex manifold X of dimension n is the *canonical bundle* defined by

$$K_X := \wedge^n T^*(X).$$

The holomorphic sections of K_X are simply the holomorphic n-forms on X, locally given by

$$\varphi = f(z)dz_1 \wedge \cdots \wedge dz_n,$$

where $f(z)$ is a holomorphic function in a local coordinate system $(z_1, \ldots, z_n)$ on X.

The canonical bundle on $\mathbf{P}_n$ can be shown to be

$$K_{\mathbf{P}_n} = (H^*)^{n+1},$$

where H is the hyperplane section bundle on $\mathbf{P}_n$, as described above (see, e.g., [239] pp. 224–225, for a proof of this).

As we mentioned at the beginning of this section, we want to use holomorphic sections of holomorphic line bundles to obtain an embedding. As we saw, the only globally defined holomorphic functions on a compact complex manifold X are constants, and these can be considered as holomorphic sections on X of the trivial line bundle $X \times \mathbf{C}$. An important question is whether a given line bundle has any nontrivial global sections that could be used for an embedding. Let's look at our examples above in this context. It is not difficult to verify that, for $n > 0$,

$$\mathcal{O}(\mathbf{P}_n, H^n) \cong \{\text{polynomials of degree } n+1 \text{ on } \mathbf{C}^{n+1}\},$$

and

$$\mathcal{O}(P_n, (H^*)^n) \cong 0,$$

where here 0 denotes the zero-dimensional complex vector space. Thus, some line bundles have sections and some do not. We will formulate a criterion below for holomorphic line bundles, due to Kodaira, that ensures that there are sufficiently many sections of such bundles that will eventually provide an embedding.

For the moment, consider a compact complex manifold with a holomorphic line bundle E, where $\mathcal{O}(X, E)$ has $N+1$ sections $s_0, \ldots, s_N$, such that for each $x \in X$,

$$s_j(x) \neq 0, \ \text{for some } j, \ 0 \leq j \leq N. \tag{14.18}$$

Then these sections provide a holomorphic mapping

$$\Phi : X \to \mathbf{P}_N. \tag{14.19}$$

Namely, let

$$h_\alpha : E_{|U_\alpha} \to U_\alpha \times \mathbf{C}$$

be trivializations for E on open sets $U_\alpha \subset X$. For $x \in U_\alpha$, define $\Phi_\alpha(x)$ to be the point in $\mathbf{P}_N$ which is the one-dimensional subspace of $\mathbf{C}^{n+1}$ spanned by complex multiples of the nonzero vector

$$(h_\alpha(s_0)(x), \ldots, h_\alpha(s_N)(x)) \in \mathbf{C}^{N+1} - \{0\}.$$

Note that Φ_α is well defined by (14.18). If $x \in U_\alpha \cap U_\beta$, then, letting $g_{\alpha\beta}$ denote the transition functions for E, we have

$$\begin{aligned}(h_\alpha(s_0)(x), \ldots, h_a(s_N)(x)) &= (g_{\alpha\beta}(x)h_\beta(s_0)(x), \ldots, g_{a\beta}(x)h_\beta(s(_N)(x)) \\ &= g_{\alpha\beta}(x)(h_\beta(s_0)(x), \ldots, h_\beta(s_N)(x)),\end{aligned}$$

and hence $\Phi_\alpha(x) = \Phi_\beta(x)$ on $U_a \cap U_\beta$. We can then define

$$\Phi(x) := \Phi_a(x), \ \text{for } x \in U_\alpha,$$

and this is then a well-defined holomorphic mapping from X to $\mathbf{P}_N$, as desired.

The basic problem for Kodaira was to show that for a given Hodge manifold X, there is a holomorphic line bundle $E \to X$ such that the mapping F in (14.19) is well-defined and that, moreover, it is one-to-one and regular at each point of X. We now need to introduce the tools of sheaf theory and Hodge theory in the next two sections, which lead up to Kodaira's vanishing theorem in Sect. 14.5. This will then be used to find an embedding of the form (14.19) in Sect. 14.6.

14.3 Sheaf Theory

Three major developments in algebraic topology were single-handedly developed in a prisoner-of-war camp during the second world war. Namely, a well-known French mathematician, Jean Leray (1906–1998), who had worked on nonlinear partial differential equations, including fixed-point theorems in the infinite-dimensional context (Schauder–Leray fixed-point theorem), was an officer in the French army and taken prisoner and placed in a prisoner-of-war camp in Austria during 1940–1945. He didn't want his captors to know that he knew something of applied mathematics, and so he spent his time studying and developing a new theory of algebraic topology.

After the war, it was difficult for others to understand what he had developed, as the language of the new mathematics was difficult to absorb at the time. What he had developed was the theory of sheaves, sheaf cohomology and spectral sequences, all of which became standard tools in algebraic topology, several complex variables and, most decisively, algebraic geometry. The story of Leray's prisoner-of-war incarceration and the evolution of his ideas into the tools they became is beautifully described by Haynes Miller [152]. Leray's experiences are very similar to those of Poncelet in a Russian prisoner-of-war camp, 1813–1814, during which time he wrote up what became a draft of his major monograph on projective geometry, as we described in Sect. 3.2.

The seminars in Paris after the second world war run by Henri Cartan (1904–2008) played a very important role in synthesizing these ideas (Seminaire Cartan, 1948–1964), all of which were published and are now available online (www.numdam.org). We will summarize the basic concepts of sheaf theory and sheaf cohomology that we need for Kodaira's work in this chapter as well as for the next chapter concerning noncompact complex manifolds, Chap. 15. We refer to standard references again for details and references (e.g., Hirzebruch [104], Griffiths and Harris [91], and Wells [239]).

Sheaf theory and sheaf cohomology theory are well defined on topological spaces in general, but we restrict ourselves to complex manifolds, as that is our main focus in these last two chapters of the book, and it the nature of most of our important examples as well. Let X denote a complex manifold. A sheaf $\mathcal{F}$ on X is an assignment of an Abelian group $\mathcal{F}(U)$ to each open set U on X satisfying certain properties. Let us give several examples to illustrate this. Namely, consider the assignments, for U an open set in X,

$$\begin{aligned} U &\mapsto C^\infty(U), \\ U &\mapsto \mathcal{E}^k(U), \\ U &\mapsto \mathcal{O}(U). \end{aligned}$$

These assignments define the sheaves C^∞, $\mathcal{E}^k$, and $\mathcal{O}$ on X. These Abelian groups (on each open set U) are spaces of functions (or differential forms), and they all happen to be vector spaces, but for our immediate purposes we think of them as Abelian groups under the operation of addition. Sheaves have restriction mappings which are group homomorphisms of the form

$$r_{UV} : \mathcal{F}(U) \to \mathcal{F}(V), \text{ for } V \subset U,$$

which for our examples are simply pointwise restrictions from the larger to the smaller open set. There are two simple axioms for sheaf theory which all of our examples (and the ones to come) clearly satisfy, and we leave this theoretical aspect to the references.

The elements of the Abelian group $\mathcal{F}(U)$ are called *sections of the sheaf* on (or over) the set U, and in our examples sections are simply functions or differential forms on the set U. One more set of examples is the so-called *constant* sheaves. These are simply of the form, where again, U is an arbitrary open set on X

$$\begin{gathered} U \mapsto \mathbf{Z}, \\ U \mapsto \mathbf{C}, \\ \dots \end{gathered}$$

A section of the sheaf $\mathbf{Z}$ on an open set U is simply the assignment of an integer (a constant function) on each connected component of U.

A sheaf $\mathcal{F}$ on X has information at the global level, e.g., sections on all of X, $\mathcal{F}(X)$, and at the local level, e.g., $\mathcal{F}(U)$, for U a small neighborhood of a point. A prime motivation for Leray in the creation of sheaf theory was to create a theory that would help understand how to piece together local data to obtain global mathematical objects of a similar type, and, in particular, to formalize obstructions to such a process. We will see an example of what we mean by this in the paragraphs below.

An important ingredient in this process is to formalize the local information of a sheaf at a given point. Let $x \in X$, and consider all of the sections $\mathcal{F}(U)$ defined near x, namely all $s \in \mathcal{F}(U)$ for $U \ni x$. If $s_1 \in \mathcal{F}(U_1)$ and $s_2 \in \mathcal{F}(U_2)$, where $x \in U_1 \cap U_2$, then we say that s_1 and s_2 are *equivalent at* x if there is an open set $W \subset U_1 \cap U_2$ with $x \in W$, and such that

$$r_{U_1 W}(s_1) = r_{U_2 W}(s_2),$$

i.e., the restrictions of s_1 and s_2 to some smaller neighborhood of x coincide. Using this equivalence relation we can form the direct limit of all $\mathcal{F}(U)$ for $U \ni x$ modulo this equivalence relation, and we denote this by $\mathcal{F}_x$. This is called the *stalk of* $\mathcal{F}$ at the point x, and it is an Abelian group, where the addition is obtained by adding representatives of equivalence classes in suitable neighborhoods of x.

For instance, if $\mathcal{O}$ is the sheaf of holomorphic functions on X (i.e., $\mathcal{O}(U)$, for all open sets U in X), then $\mathcal{O}_x$ at $x \in X$ can be identified with all convergent power series at x in any coordinate chart containing x. The elements of the stalk $\mathcal{F}_x$ are referred to as the *germs of the sheaf* $\mathcal{F}$ at x.

Now we turn to sheaf cohomology. A *sequence* of Abelian groups is a discrete set of Abelian groups with linking homomorphisms of the form

$$\cdots \xrightarrow{\alpha} A \xrightarrow{\beta} B \xrightarrow{\gamma} C \xrightarrow{\delta} \cdots . \tag{14.20}$$

The sequence is said to be *exact* at B if $\ker \gamma = \operatorname{im} \beta$.

A *sheaf homomorphism*

$$\mathcal{F} \xrightarrow{h} \mathcal{G}$$

is a set of homomorphisms of the form

$$\mathcal{F}(U) \xrightarrow{h_U} \mathcal{G}(U),$$

for each open set $U \subset X$. A *sequence of sheaves* is a discrete set of sheaves linked by homomorphisms, just as in (14.20),

$$\cdots \xrightarrow{\alpha} \mathcal{A} \xrightarrow{\beta} \mathcal{B} \xrightarrow{\gamma} \mathcal{C} \xrightarrow{\delta} \cdots . \tag{14.21}$$

Finally, a sequence of sheaves, such as in (14.21), is *exact* at $\mathcal{B}$ if the induced sequence of stalks (which are all Abelian groups)

$$\cdots \xrightarrow{\alpha} \mathcal{A}_x \xrightarrow{\beta} \mathcal{B}_x \xrightarrow{\gamma} \mathcal{C}_x \xrightarrow{\delta} \cdots ,$$

is exact at $\mathcal{B}_x$ for each $x \in X$. Note that we are not requiring exactness at $\mathcal{B}(U)$ for all open sets $U \subset X$, but only at the localizations of the sheaves near each point of X (which is represented by the stalks).

Let's illustrate this last notion with an important example. Consider the sequence of sheaves of differential forms on X linked by the homomorphisms of exterior differentiation

$$0 \to \mathbf{C} \xrightarrow{i} \mathcal{E}^0 \xrightarrow{d} \cdots \to \mathcal{E}^{k-1} \xrightarrow{d} \mathcal{E}^k \xrightarrow{d} \mathcal{E}^{k+1} \to \cdots , \tag{14.22}$$

i.e.,

$$\cdots \to \mathcal{E}^{k-1}(U) \xrightarrow{d} \mathcal{E}^k(U) \xrightarrow{d} \mathcal{E}^{k+1}(U) \to \cdots ,$$

for each open set $U \subset X$, and i is the natural inclusion of constants in smooth functions. By using the Poincaré lemma[5] at each point $x \in X$, one easily sees that the sequence (14.22) is exact at $\mathcal{E}^k$, for each $k \geq 0$. However, if X is compact and has nontrivial cohomology for some k, $0 < k < \dim_{\mathbf{R}} X$ (for instance, a compact Riemann surface of genus $g > 1$, where $H^1(X, \mathbf{C}) \neq 0$), then the sequence

$$\cdots \to \mathcal{E}^{k-1}(X) \xrightarrow{d} \mathcal{E}^k(X) \xrightarrow{d} \mathcal{E}^{k+1}(X) \to \cdots$$

is not exact at $\mathcal{E}^k(X)$ by de Rham's theorem (14.7), i.e., there is a d-closed form φ on X which is not an exact form.

[5] See any of the references for this important lemma from advanced calculus, which generalizes the fundamental theorem of calculus.

This is a prime example of the difference between local exactness (which has been formalized in the exactness of a sheaf sequence) and global exactness. Sheaf cohomology theory provides a measure of the difference being exact near a point (like the Poincaré lemma) versus being globally exact (as exemplified by de Rham's theorem) for a larger class of sheaves than the differential forms in this example.

A *short exact sequence of sheaves* is a sequence of sheaves of the form

$$0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \to 0,$$

which is exact at $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. That is to say, the homomorphism α is injective, the homomorphism β is surjective, and $\ker \beta = \operatorname{im} \alpha$. Let us give an important example of such a short exact sequence that will be useful later. On a complex manifold X, let $\mathcal{O}^*$ denote the sheaf of nonvanishing holomorphic functions with the Abelian group structure given by multiplication. Then consider the sequence of sheaves

$$0 \to \mathbf{Z} \xrightarrow{i} \mathcal{O} \xrightarrow{\exp} \mathcal{O}^* \to 0, \tag{14.23}$$

where

$$\exp(f) := e^{2\pi i f}, \ \ f \in \mathcal{O}(U),$$

for U an open subset of X. It is easy to verify that this is a short exact sequence of sheaves. In particular, if $g(z)$ is holomorphic and nonzero in the neighborhood of a point $x_0 \in X$, then choosing a branch of the logarithm function near $g(x_0)$ will provide a local inverse to the exponential mapping in the sequence. On the other hand, the sequence

$$\mathcal{O}(\mathbf{C} - \{0\}) \xrightarrow{\exp} \mathcal{O}^*(\mathbf{C} - \{0\}) \to 0$$

is *not* exact at $\mathcal{O}^*(\mathbf{C} - \{0\})$, since there is no branch of the logarithm function on the punctured plane.

Now we proceed to formulate the notion of *sheaf cohomology groups* for a sheaf $\mathcal{F}$ on a topological space X. Namely, to each sheaf $\mathcal{F}$, there is an assignment of a sequence of Abelian groups of the form

$$(X, \mathcal{F}) \mapsto H^q(X, \mathcal{F}), \ \ q \in \mathbf{Z}, q \geq 0. \tag{14.24}$$

These Abelian groups $H^q(X, \mathcal{F})$ satisfy specific axioms which we will not discuss in detail here. However, we will describe their most important property, namely, the sheaf cohomology groups have the property that, for each short exact sequence of sheaves

$$0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0, \tag{14.25}$$

there is a long exact sequence of cohomology groups of the form

$$\begin{aligned} 0 \to H^0(X, \mathcal{A}) \to H^0(X, \mathcal{B}) \to H^0(X, \mathcal{C}) \to H^1(X, \mathcal{A}) \to \\ H^1(X, \mathcal{B}) \to H^1(X, \mathcal{C}) \to H^2(X, \mathcal{A}) \to H^2(S, \mathcal{B}) \to \cdots , \end{aligned} \tag{14.26}$$

where

$$\begin{aligned} H^0(X, \mathcal{A}) &= \mathcal{A}(X), \\ H^0(X, \mathcal{B}) &= \mathcal{B}(X), \\ H^0(X, \mathcal{C}) &= \mathcal{C}(X). \end{aligned} \tag{14.27}$$

See the references we have cited at the beginning of this section for more details.

One shows that there is a unique cohomology theory satisfying the axioms, and there are several existence proofs for the cohomology groups, including Čech theory (using open coverings), resolution of sheaves by specific types of sheaves, and in terms of category theory, for instance. Classical cohomology theory from algebraic topology ($H^q(X, \mathbf{Z})$, $H^q(X, \mathbf{R})$, etc.), which also has an axiomatic formulation as in the text by Eilenberg and Steenrod [59], coincides with the more general sheaf cohomology theory we are using here, using constant sheaves ($\mathbf{Z}$, $\mathbf{R}$, etc.) on X. In working with cohomology groups, just as in classical algebraic topology, one uses homological algebra to compute unknown cohomology groups in terms of known cases, just as in calculus one computes derivatives using properties of differentiation (e.g., product rule, chain rule, etc.).

Let us look at the example of a short exact sequence we introduced above (14.23) and use the long exact sequence (14.26), which yields, for the case of a domain D in the complex plane, using (14.27),

$$\mathcal{O}(D) \overset{\exp}{\to} \mathcal{O}^*(D) \to H^1(D, \mathbf{Z}),$$

which is exact at the center. If $H^1(D, \mathbf{Z}) = 0$, that is, in this case, if D is simply-connected, then the exponential mapping is surjective. This is the classical statement that one can choose a branch of a logarithm on a simply-connected domain in the complex plane.

We now give an important analogue of de Rham's theorem on a complex manifold that we will use later. Let X be an n-dimensional complex manifold and consider the sequence of sheaves

$$0 \to \Omega^p \overset{i}{\to} \mathcal{E}^{p,0} \overset{\overline{\partial}}{\to} \mathcal{E}^{p,1} \overset{\overline{\partial}}{\to} \cdots \overset{\overline{\partial}}{\to} \mathcal{E}^{p,n} \to 0,$$

using (14.11). Pierre Dolbeault (1924–2015) showed that this is an exact sequence of sheaves and he was able to formulate and prove an analogue to de Rham's theorem context which is now known as *Dolbeault's Theorem*: [58]:

$$H^q(X, \Omega^p) \cong \frac{\ker \overline{\partial} : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q+1}(X)}{\operatorname{im} \overline{\partial} : \mathcal{E}^{p,q-1}(X) \to \mathcal{E}^{p,q}(X)}. \tag{14.28}$$

The sheaf cohomology in Dolbeault's theorem, $H^q(X, \Omega^p)$, is formulated on the complex manifold X in terms of holomorphic differential forms, and the theorem represents this cohomology in terms of solutions on X of partial differential equations involving smooth differential forms. We will see how this can be further refined in the next section, using Hodge theory. For convenience later on, we define the right-hand side in Dolbeault's theorem as

$$H^{p,q}(X) := \frac{\overline{\partial}\text{-closed}(p,q)\text{-forms}}{\overline{\partial}\text{-exact}(p,q)\text{-forms}}. \tag{14.29}$$

These are called the *Dolbeault cohomology groups*. Thus Dolbeault's theorem becomes, with this notation,

$$H^q(X, \Omega^p) \cong H^{p,q}(X),$$

and we note that the order of the two symbols p and q is reversed here.

We give one final example of an application of sheaf cohomology that will be of use later. Namely, consider again the long exact sequence of cohomology groups associated to the exponential short exact sequence (14.23)

$$\cdots \to H^1(X, \mathcal{O}) \to H^1(X, \mathcal{O}^*) \xrightarrow{\delta} H^2(X, \mathbf{Z}) \to \cdots. \tag{14.30}$$

By using Čech cohomology[6] one can identify an element $E \in H^1(X, \mathcal{O}^*)$ with an equivalence class of holomorphic line bundles on X. Namely, a representative cocycle in this cohomology group can be defined by nonvanishing holomorphic functions $h_{\alpha\beta} \in \mathcal{O}^*(U_\alpha \cap U_\beta)$ satisfying the cocycle conditions (14.15), which is then a representation of a line bundle E. The image of E under the mapping δ in (14.30)

$$c_1(E) := \delta(E) \in H^2(X, \mathbf{Z}) \tag{14.31}$$

is defined to be the *Chern class* $c_1(E)$ of the holomorphic line bundle[7] E. It is a primary topological obstruction for the line bundle E being trivial or not. This is well-defined in the category of differentiable as well as topological manifolds using the same mechanism with sheaves of differentiable or continuous functions. We will discuss later in this section a differential-geometric representation for the Chern class of a line bundle, and it will play an important role in Kodaira's theorem.

We will need one further sheaf-theoretic tool that will turn out to be quite important for us later, and that is the notion of a *quotient sheaf*. Suppose that $\mathcal{G} \subset \mathcal{F}$ is a subsheaf defined in some natural manner (e.g., on a complex manifold, the sheaf of

---

[6]We will give an overview of Čech cohomology in Sect. 15.6; see Hirzebruch [104] for a detailed discussion of sheaf cohomology theory from this open covering point of view.

[7]It is denoted as the first Chern class $c_1(E)$, since for higher-rank vector bundles there are higher-order Chern classes.

holomorphic functions $\mathcal{O}$ is a subsheaf of the sheaf $\mathcal{E}$ of differentiable functions). We can form the quotient Abelian groups

$$\mathcal{F}(U)/\mathcal{G}(U) \tag{14.32}$$

for each open subset $U \subset X$. However, this assignment of Abelian groups to open sets U may not satisfy the sheaf axioms (which we have not explicitly stated).

One can, on the other hand, form the quotients of the stalks of $\mathcal{F}$ and $\mathcal{G}$ at a given point x. These quotients have the form

$$\mathcal{F}_x/\mathcal{G}_x, \tag{14.33}$$

at each point $x \in X$. It is possible to take the union

$$\bigcup_{x \in X} \mathcal{F}_x/\mathcal{G}_x,$$

to form a sheaf called the *quotient sheaf* $\mathcal{F}/\mathcal{G}$. This is one of the more delicate and important constructions in sheaf theory, which the references can explain more fully. The sections of the quotient sheaf over an open set U, $(\mathcal{F}/\mathcal{G})(U)$, will contain $\mathcal{F}(U)/\mathcal{G}(U)$, but, in principle, will be a larger group. However, the stalks of $\mathcal{F}/\mathcal{G}$ at $x \in X$ will be as in (14.33). There is a canonical exact sequence for quotient sheaves of the form

$$0 \to \mathcal{G} \to \mathcal{F} \to \mathcal{F}/\mathcal{G} \to 0. \tag{14.34}$$

We will have occasion to use this type of exact sequence in Sect. 14.6 when we construct the Kodaira embedding, as well as in Chap. 15.

Now we suppose that X is a complex manifold, and we observe that $\mathcal{O}$ is a sheaf of rings on X. Namely, each Abelian group $\mathcal{O}(U)$ also has the structure of a ring in a natural manner, as products and sums of holomorphic functions are well-defined and satisfy the properties of a ring. In a similar manner, one can define a sheaf of modules over the sheaf of rings $\mathcal{O}$, and these are called *analytic sheaves*. A *locally free* analytic sheaf $\mathcal{F}$ is defined by the property that in a neighborhood W of a given point the sheaf has the form

$$\mathcal{F}_{|W} \cong \mathcal{O}^r,$$

where $\mathcal{O}^r$ is the direct sum of r copies of $\mathcal{O}$, and r is called the rank of the (locally free) sheaf $\mathcal{F}$. If we consider a holomorphic vector bundle E on X, then the sheaf of holomorphic sections of E is a locally free sheaf of rank r, where $r = \dim_{\mathbb{C}} E$. Namely, locally the sections of E are simply vector-valued holomorphic functions, and that is the nature of the sections of a locally free analytic sheaf, by definition. All locally free analytic sheaves are simply sheaves of a holomorphic vector bundle and vice-versa. This is simply a change in language, which is often very useful.

In addition, there are important analytic sheaves which are not sections of holomorphic vector bundles. The most important category of such analytic sheaves are the *coherent analytic sheaves*,[8] generalizations of locally free sheaves, introduced by Henri Cartan in his seminars in the 1950s, and we will see examples of such sheaves as we proceed. If we have two analytic sheaves $\mathcal{F}$ and $\mathcal{G}$ on X, then, as modules over the sheaf of rings $\mathcal{O}$, we can define the tensor product

$$\mathcal{F} \otimes_{\mathcal{O}} \mathcal{G},$$

which is defined by the tensor product

$$\mathcal{F}(U) \otimes_{\mathcal{O}(U)} \mathcal{G}(U),$$

for each open set U. The tensor product of two analytic sheaves is a well-defined analytic sheaf on X, and we will have need of this in Sect. 14.6. If we have two locally free sheaves $\mathcal{F}$ and $\mathcal{G}$ on X corresponding to two holomorphic vector bundles F and G on X, then

$$\mathcal{F} \otimes_{\mathcal{O}} \mathcal{G} \cong \mathcal{O}(F \otimes G),$$

which will also turn out to be useful for us.

Let us close with one important example. Let p be a point in a complex manifold X, and define the subsheaf m_p of $\mathcal{O}$ on X by

$$m_p(U) := \begin{cases} \{f \in \mathcal{O}(U) : f(p) = 0\}, \text{ if } p \in U, \\ \mathcal{O}(U), \text{ if } p \notin U. \end{cases} \tag{14.35}$$

The sheaf m_p is an example of an *ideal sheaf* on X (more generally, an ideal sheaf is a subsheaf of $\mathcal{O}$ which vanishes on a given analytic subvariety of X). It is called an ideal sheaf, since the stalk of m_p at p is an ideal in the ring $\mathcal{O}_p$. This ideal sheaf m_p is an example of a coherent analytic sheaf.

The quotient sheaf of two coherent analytic sheaves is also a coherent analytic sheaf.

[8] A *coherent analytic sheaf* $\mathcal{F}$ on X is a sheaf which is locally finitely generated, i.e., there is, in a neighborhood of any point, a surjective homomorphism

$$\mathcal{O}^p \to \mathcal{F},$$

and the kernel of such a mapping is also locally finitely generated. We will not have the need to use this concept explicitly, but we will see its use in special cases. See the references, for instance, Gunning and Rossi [93], Hörmander [112] and Krantz [130], for more detailed information about coherent analytic sheaves.

Let us now look at the quotient sheaf $\mathcal{O}/m_p$ in this example. It is easy to see that

$$(\mathcal{O}/m_p)_x \cong \begin{cases} \mathbf{C}, & x = p, \\ 0, & x \neq p. \end{cases}$$

Moreover, there is an exact sequence of the form

$$0 \to m_p \to \mathcal{O} \to \mathcal{O}/m_p \to 0.$$

This type of exact sequence will turn out to be very important in Kodaira's embedding proof, as we will see in Sect. 14.6 as well as in Chap. 15.

14.4 Hodge Theory

As we mentioned earlier, Hodge created a theory of harmonic differential forms on manifolds that proved to be very useful in studying the algebraic topology of projective algebraic manifolds. We now want to summarize some of the key aspects of this theory as applied to the investigation of Kähler manifolds, which play a critical role in Kodaira's theorem.

We review first the elements of the theory for compact differentiable manifolds. Let M be a compact orientable differentiable manifold of m dimensions, and let g be a Riemannian metric on M. The metric g is an inner product on the tangent space $T_x(M)$ at each point $x \in M$, and it naturally induces a smoothly varying inner product on the dual spaces $T_x^*(M)$.

Let U be an open set in M. A *frame*[9] for the cotangent bundle $T^*(M)$ on U is a set of one-forms $(e_1, \ldots, e_m)$ such that the differential forms $\{e_\mu\}$ are linearly independent at each point of U. If U is a coordinate chart on M with coordinates $(x_1, \ldots, x_m)$, any one-form φ is a linear combination of the form

$$\varphi = \sum a_\mu(x) dx_\mu,$$

where $a_\mu(x)$ are differentiable functions on U. So, in this case, $(dx_1, \ldots, dx_n)$ is a frame for $T^*(M)$ in the coordinate chart U. We can choose suitable linear combinations of the differentials dx_μ to be a frame as well, i.e., let

$$e_\mu = \sum_\nu e_{\mu\nu}(x) dx_\nu, \tag{14.36}$$

[9] A *frame* for a vector bundle E on M is a set of sections defined in a neighborhood U of a given point which are linearly independent and which span the vector spaces E_x at each point $x \in U$. A vector bundle always has a frame defined near each point by the local trivializations of the bundle, and it has a frame defined on all of M only if it is a trivial vector bundle.

where we require the matrix $e_{\mu\nu}(x)$ to be nonsingular for $x \in U$, then $(e_1, \ldots, e_m)$ is also a frame.

By the Gram–Schmidt orthogonalization process, we can choose a frame e_μ of the form (14.36) so that

$$(e_\mu(x), e_\nu(x))_x = \delta_{\mu\nu}, 1 \leq \mu, \nu, \leq m,$$

i.e., the frame e_μ forms an orthonormal basis at each point $x \in U$. A frame with this orthonormality property is called an *orthonormal frame* or a *moving frame*.[10]

Let $(e_1, \ldots, e_m)$ be an orthonormal frame for $T^*(M)$ on an open set $U \subset M$. Then the k-forms

$$e_{\mu_1} \wedge \cdots \wedge e_{\mu_k}, 1 \leq \mu_1 < \cdots < \mu_k \leq m, \tag{14.37}$$

form a basis for the k-forms on U. We define the $*$-operator on this basis as follows:

$$*\,(e_{\mu_1} \wedge \cdots \wedge e_{\mu_k}) = \pm e_{\nu_1} \wedge \cdots \wedge e_{\nu_{m-k}}, \tag{14.38}$$

where

$$\{e_{\mu_1}, \ldots, e_{\mu_k}, e_{\nu_1}, \ldots, e_{\nu_{m-k}}\}$$

is a permutation of

$$\{e_1, \ldots, e_m\}, \tag{14.39}$$

and the sign in (14.38) is chosen to be positive for an even permutation and negative if the permutation is odd. Note that the ordering in (14.39) corresponds to a choice of an orientation on the manifold M.

Using this definition of the $*$-operator on the basis (14.37) we can extend the definition of $*$-operator to any k-form on U. Hodge shows that if there are orthonormal frames on overlapping open sets U_1 and U_2, then the definition of $*$-operator must agree on the overlap $U_1 \cap U_2$, and hence the $*$-operator is defined for k-forms on M. Thus we have a linear mapping of vector spaces

$$* : \mathcal{E}^k(M) \to \mathcal{E}^{m-k}(M),$$

which is called the *Hodge $*$-operator*, which depends on the Riemannian metric on M. As we want to extend this theory to complex manifolds, we choose to take differential forms to have complex-valued coefficients. If

[10] Elie Cartan pioneered the use of moving frames ("repère mobile") in the early part of the twentieth century, and this tool allowed one to simplify many calculations and formulations of concepts in differential geometry and its various generalizations.

$$\eta(x) = \sum_{|I|=k} \eta_I(x) dx^I$$

is a complex-valued differential form on M, then the complex conjugate of η is defined by

$$\overline{\eta}(x) := \sum_{|I|=k} \overline{\eta_I(x)} dx^I.$$

We now define the *Hodge inner product* on $\mathcal{E}^k(M)$ as

$$(\varphi, \psi) := \int_M \varphi \wedge *\overline{\psi}. \tag{14.40}$$

This L^2-inner product on differential forms is a generalization of the classical L^2-Hermitian inner product on complex-valued functions on $\mathbf{R}^m$,

$$(f, g) = \int_{\mathbf{R}^m} f(x)\overline{g(x)} dx. \tag{14.41}$$

Let $\mathcal{D}(\mathbf{R}^m)$ denote the compactly-supported differentiable functions on $\mathbf{R}^m$, and consider a linear differential operator

$$L : \mathcal{D}(\mathbf{R}^m) \to \mathcal{D}(\mathbf{R}^m),$$

then we define the *adjoint differential operator*[11] to be a linear operator L^* which satisfies

$$(Lf, g) = (f, L^* g), \;\; f, g \in \mathcal{D}(\mathbf{R}^m),$$

for the L^2-inner product defined by (14.41). For instance, in the simplest case, using integration by parts,

$$\int_{\mathbf{R}} \left(\frac{df}{dx}\right) g dx = \int_{\mathbf{R}} f \left(-\frac{dg}{dx}\right) dx, \;\; f, g \in \mathcal{D}(\mathbf{R}), \tag{14.42}$$

that is, $-\frac{d}{dx}$ is the adjoint of $\frac{d}{dx}$ in this very simple case.

We now use this model to define the adjoint d^* of the exterior differentiation operator d using the Hodge inner product. Formally, we define d^* by the equality

$$(d\varphi, \psi) = (\varphi, d^*\psi), \tag{14.43}$$

[11] This is the *formal adjoint* in the theory of linear differential equations; the word adjoint in that theory involves boundary conditions as well.

where on the left-hand side we use the Hodge inner product on $\mathcal{E}^{k+1}(M)$ and on the right-hand side the Hodge inner product on $\mathcal{E}^k(M)$.

In particular, the formula (14.43) indicates that d^* is an operator of the form

$$d^* : \mathcal{E}^{k+1}(M) \to \mathcal{E}^k(M).$$

Indeed, this is the case, and Hodge shows, by proving suitable straightforward properties of the $*$-operator, that we can define d^* by

$$d^*\varphi = (-1)^{m+mk+1} * d * \varphi, \text{ for } \varphi \in \mathcal{E}^k(M), \tag{14.44}$$

which does satisfy (14.43). Note that the $-$ sign in (14.44) is reminiscent of the $-$ sign in (14.42), which is in fact a consequence of the use of Stokes's theorem in this more general case.

Using the adjoint operator d^* we can now define the *Laplacian operator* Δ acting on k-forms. Namely, we set

$$\Delta = dd^* + d^*d : \mathcal{E}^k(M) \to \mathcal{E}^k(M).$$

For $k = 0$, d^* is undefined, and we simply take $\Delta = d^*d$ in this case. In the case of differential forms on $\mathbf{R}^m$ with the simple Euclidean metric, the Laplacian operator Δ is the usual Laplacian operator

$$\Delta = \sum_{\mu=1}^{m} \frac{\partial^2}{\partial x_\mu^2},$$

acting on the coefficients of the differential forms. The primary differential equation on M that one needs to solve is:

$$\Delta\varphi = \psi. \tag{14.45}$$

In the long history of the theory of differential equations, the fundamental problem has always been: given a differential equation, e.g.,

$$\frac{df}{dx} = g, \tag{14.46}$$

find a solution f which satisfies the given equation. In this case, and in most cases, we use an integration process to find the solutions. For the simple differential equation (14.46), we write

$$f(x) = \int_{x_0}^{x} g(t)dt,$$

and, by the fundamental theorem of calculus, $f(x)$ is the desired solution. In classical potential theory, integral operators of the form

$$\int_{\mathbf{R}^3} \frac{g(y)}{|x-y|} dy \tag{14.47}$$

were used to solve problems of the form

$$\Delta u = g,$$

where Δ is the three-dimensional Laplacian (and $1/r$ is the kernel of the integral operator, in this case the potential). This was important for the solution of many problems of physics in the eighteenth and nineteenth centuries (fluid mechanics, electromagnetism, etc.). Many such integral operators of that time were called Green's operators.

Hodge was able to generalize the theory of Green's operators to this differential-geometric setting on a compact differentiable manifold. Without going into the details of the integral operators involved, we summarize the fundamental results of Hodge here in the following theorem. First we define the vector space of *harmonic k-forms* to be

$$\mathcal{H}^k(M) := \{\varphi \in \mathcal{E}^k(M) : \Delta\varphi = 0\}. \tag{14.48}$$

It is not difficult to show (in this case where M is a compact manifold) that a k form φ is harmonic if and only if

$$d\varphi = 0 \text{ and } d^*\varphi = 0. \tag{14.49}$$

Theorem 14.2 (Hodge decomposition theorem [109]) *Let M be a compact Riemannian manifold, then there is a continuous linear mapping*

$$G : \mathcal{E}^k(M) \to \mathcal{E}^k(M),$$

and an orthogonal decomposition of the form

$$\begin{aligned} \mathcal{E}^k(M) &= \mathcal{H}^k(M) \oplus dd^*G\mathcal{E}^k(M) \oplus d^*dG\mathcal{E}^k(M) \\ &= \mathcal{H}^k(M) \oplus Gdd^*\mathcal{E}^k(M) \oplus Gd^*d\mathcal{E}^k(M), \end{aligned} \tag{14.50}$$

and moreover,

$$\dim \mathcal{H}^k(M) < \infty.$$

The operator G in this theorem is referred to as a *Green's operator*. We define the orthogonal projection (the *harmonic projection*)

$$H : \mathcal{E}^k(M) \to \mathcal{H}^k(M)$$

to be the orthogonal projection onto the harmonic differential forms (first term in the sum on the right-hand side of (14.50)). An important fact that we use below is that the Green's operator and the Laplacian commute and respect the degrees of the forms (i.e., $\Delta G = G\Delta$, and, in particular, $dd^*G = dd^*G$ and $d^*dG = Gd^*d$). We note here that the left-hand side of (14.50) depends only on the differentiable structure of M, and the right-hand side depends explicitly on the Riemannian metric on M, as well. The orthogonal direct sum in (14.50) is referred to as the *Hodge decomposition* on the manifold M.

What the Hodge decomposition (14.50) shows is that the inhomogeneous Laplacian operator on the infinite-dimensional vector space $\mathcal{E}^k(M)$ has an inverse, modulo the finite-dimensional space of harmonic forms. In fact the Green's operator is such a pseudoinverse and can be defined by pseudodifferential operators, generalizations of differential operators which include integral operators of the form (14.47) (see e.g., Wells [239]). It can also be defined in terms of the spectral theory of elliptic differential equations. In both cases, the proof of the Hodge decomposition involves including the spaces of smooth differential forms in corresponding Sobolev spaces of differential forms, which are then Hilbert spaces which allow us to prove that limiting processes converge in the appropriate sense. Fundamentally, in the theory of elliptic linear differential equations, and (14.45) is an example of such an equation, one shows first that there exists a weak distribution solution in some appropriate space of distributions (usually a Sobolev space of some sort), and then one shows that such a weak solution is, in fact, smooth.

An immediate corollary of the Hodge decomposition is Hodge's representation of de Rham cohomology on M. Namely, there is an isomorphism

$$H^k_{\mathrm{dR}}(M) \cong \mathcal{H}^k(M). \tag{14.51}$$

To see this, let φ be a closed k-form which represents a cohomology class $[\varphi] \in H^k_{\mathrm{dR}}(M)$, then, by the Hodge decomposition

$$\varphi = H\varphi + dd^*G\varphi,$$

since $d\varphi = 0$. The mapping

$$\varphi \mapsto H\varphi$$

induces a mapping

$$Z^k(M) \to \mathcal{H}^k(M).$$

This mapping is surjective, since any harmonic k-form is closed. If $H\varphi = 0$, then

$$\begin{aligned}\varphi &= H\varphi + dd^*G\varphi + Gd^*d\varphi \\ &= d(d^*g\varphi),\end{aligned}$$

and hence $\varphi \in B^k(M)$. It follows that the induced mapping

$$H^k_{\mathrm{dR}}(M) = Z^k(M)/B^k(M) \to \mathcal{H}^k(M)$$

is an isomorphism.

Let us now extend the theory of harmonic forms to a compact complex manifold X with a Hermitian metric h. The Hermitian metric induces naturally a Riemannian metric on the underlying differentiable manifold, which we still denote as X. Thus there is a well-defined Hodge $*$-operator, Hodge inner product, and Laplacian operator on X, and these yield harmonic differential forms on X. We need to investigate how these mathematical objects interact with the complex structure on X. We begin with the following fundamental result (again we refer to the references). The $*$-operator is a **C**-linear isomorphism

$$* : \mathcal{E}^{p,q}(X) \xrightarrow{\cong} \mathcal{E}^{n-q,n-p}(X). \tag{14.52}$$

The proof of this involves a careful linear algebra analysis of the exterior algebra

$$\wedge^k T_x^*(X) = \sum_{p+q=k} \wedge^{p,q} T_x^*(X),$$

at each point $x \in X$ and its interaction with the Hermitian inner product on T_x.

We now use the Hodge inner product on the differential forms on X, and we find that the direct sum decomposition

$$\mathcal{E}^k(X) = \sum_{p+q=k} \mathcal{E}^{p,q}(X) \tag{14.53}$$

is an *orthogonal* direct sum decomposition. This is an easy consequence of (14.52). Namely, if

$$\varphi \in \mathcal{E}^{p,q}(X),\ \ \psi \in \mathcal{E}^{r,s}(X),\ \ p+q = r+s,$$

then

$$\varphi \wedge *\overline{\psi}$$

is of type $(n-r+p, n-s+q)$. Now $\overline{\psi}$ is of type (s, r), and by (14.52) $*\overline{\psi}$ is of type $(n-r+p, n-s+q)$, as desired. But $\varphi \wedge *\overline{\psi}$ must be of type (n, n) in order that the inner product

$$\int \varphi \wedge *\overline{\psi}$$

be nonzero, and this is only the case if $p = r$ and $q = s$. Otherwise, the inner product is zero, and this shows the orthogonality of (14.53), as desired.

We now want to make a small modification of the $*$-operator for notational convenience in this complex manifold setting. Let

$$\overline{*}(\varphi) := *(\overline{\varphi}), \text{ for } \varphi \in \mathcal{E}^k(X).$$

We recall the differential operator

$$\overline{\partial} : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q+1}(X),$$

and one can compute its adjoint with respect to the Hodge inner product and obtain

$$\overline{\partial}^* = -\overline{*}\partial\overline{*} : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q-1}(X).$$

We define the associated Laplacian operator to be

$$\overline{\Delta} = \overline{\partial}\overline{\partial}^* + \overline{\partial}^*\overline{\partial} : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q}(X).$$

This is again a linear elliptic differential operator and we can define *harmonic* (p, q)-forms in this context by

$$\overline{\Delta}(\varphi) = 0.$$

Note that these harmonic forms may not necessarily be harmonic in the sense that $\Delta\varphi = 0$. We will come back to this important point shortly.

We define the vector space of (p, q)-harmonic forms to be

$$\mathcal{H}^{p,q}(X),$$

and this is also finite-dimensional, as in Theorem 14.2. We recall Dolbeault's theorem (14.28), and Hodge's harmonic theory yields the harmonic representation of the sheaf cohomology (this is due to Kodaira [127]),

$$H^q(X, \Omega^P) \cong H^{p,q}(X) \cong \mathcal{H}^{p,q}(X).$$

This is completely analogous to Hodge's representation of the de Rham groups and the standard cohomology of algebraic topology that we saw earlier in this section.

We now have the two sets of cohomology groups on X, each represented by harmonic forms:

$$H^k(X, \mathbf{C}) \cong \mathcal{H}^k(X),$$
$$H^{p,q}(X) \cong \mathcal{H}^{p,q}(X),$$

all of which are finite-dimensional, and we set

$$b^k := \dim_{\mathbf{C}} H^k(X, \mathbf{C}),$$
$$h^{p,q} = \dim_{\mathbf{C}} H^{p,q}(X),$$

where the b^k are the *Betti numbers* of the underlying compact differentiable manifold X. The numbers $h^{p,q}$ are called the *Hodge numbers* of the compact complex manifold X. However, we note that if φ is a d-closed form on X of degree k, then there is a decomposition

$$\varphi = \varphi^{k,0} + \varphi^{k-1,1} + \cdots + \varphi^{0,k}$$

of φ into forms of type (p, q), and none of these forms need be $\overline{\partial}$-closed; and similarly, if φ is a $\overline{\partial}$-closed form on X, it need not be d-closed.

There is, however, a spectral sequence due to Fröhlicher [76] which yields a relation between these dimensions. It has the form

$$H^{p,q}(X) \Rightarrow H^k(X, \mathbf{C}),$$

which yields a representation of the Euler characteristic $\chi(X)$ in terms of the Hodge numbers

$$\chi(X) := \sum_k (-1)^k b^k = \sum_k \sum_{p+q=k} (-1)^{p+q} h^{p,q}. \tag{14.54}$$

Now we make the assumption that X is a Kähler manifold with a fundamental form Ω. With this assumption the relation between the de Rham and Dolbeault groups simplifies considerably. By a subtle family of computations concerning the fundamental form Ω and the $*$-operator, one can derive the important relation between the two Laplacians on X, namely,

$$\Delta = 2\overline{\Delta}. \tag{14.55}$$

There are a number of such identities, but this suffices for us to be able to say that the two different notions of harmonic form that we introduced above coincide when we have a Kähler metric (see the references Weil [238], Griffiths and Harris [91], or Wells [239]).

An important consequence of this is that the orthogonal decomposition

$$\mathcal{E}^k(X) = \sum_{p+q=k} \mathcal{E}^{p,q}(X)$$

induces an orthogonal decomposition of harmonic forms,

$$\mathcal{H}^k(X) = \sum_{p+q=k} \mathcal{H}^{p,q}(X), \tag{14.56}$$

which is called the *Hodge structure* or the *Hodge decomposition*[12] of the Kähler manifold X. We note that the $*$-operator $\overline{*}$ induces a conjugate-linear isomorphism

$$\overline{*} : \mathcal{H}^{p,q} \overset{\cong}{\to} \mathcal{H}^{q,p}, \tag{14.57}$$

and hence the Hodge numbers satisfy

$$h^{p,q} = h^{q,p},$$

which implies from the Hodge decomposition that the odd Betti numbers of X are, in fact, even numbers, i.e.,

$$\begin{aligned} b^1 &= h^{1,0} + h^{0,1} = 2h^{1,0}, \\ b^3 &= h^{3,0} + h^{2,1} + h^{1,2} + h^{0,3} = 2(h^{3,0} + h^{2,1}), \\ &\cdots \end{aligned}$$

which is an important topological restriction on the Kähler manifold X. The Euler characteristic formula of Fröhlicher (14.54) is a simple consequence of the Hodge decomposition (14.56) in the case where X is a Kähler manifold.

The references give examples of compact complex manifolds which have, for instance, $b^1 = 1$. An example is a Hopf manifold, homeomorphic to $S^1 \times S^3$, which therefore cannot be a Kähler manifold, and hence, by Kodaira's embedding theorem, could not be a projective algebraic manifold.

14.5 Kodaira's Vanishing Theorem

As we saw in the previous section, the assumption that a compact complex manifold has a Kähler metric puts a restriction on the Betti numbers of the manifold. In a similar vein, Salomon Bochner (1899–1982) showed in 1948 [22] that a compact Riemannian manifold which has a Ricci curvature which is positive in a certain precise sense must have its first Betti number $b^1 = 0$. Under additional such hypotheses, Bochner showed that higher Betti numbers vanished as well. Bochner used Hodge's representation of the de Rham cohomology groups in terms of harmonic forms as a main tool in his proof. Kodaira was able to generalize Bochner's ideas to show

[12] Note that the term "Hodge decomposition" is often used for (14.50) for a differentiable manifold as well as (14.56) on a Kähler manifold. It is usually clear from the context which is meant.

that certain generalized Hodge numbers vanish, and this has become known as the Kodaira vanishing theorem, which we will outline here.

First, we need to generalize the Dolbeault groups to include the cohomology of (p, q)-forms with coefficients in a holomorphic line bundle. Let $E \to X$ be a holomorphic line bundle on X and let $\wedge^{p,q} T^*(X)$ be the holomorphic vector bundle whose differentiable sections are the (p, q)-forms we have been using. Then define the tensor product bundle

$$\wedge^{p,q} T^*(X) \otimes E \to X.$$

We will call the differentiable sections of this vector bundle differential forms of type $(p, q))$ with coefficients in E, which we denote by $\mathcal{E}^{p,q}(X, E)$, for sections defined on all of X. If $g_{\alpha\beta}$ is a set of transition functions for E with respect to a covering by open sets U_α, then a differential form of this type is simply a differential form φ_a of type (p, q) defined in each U_α which satisfies

$$\varphi_\alpha = g_{\alpha\beta}\varphi_\beta \text{ in } U_\alpha \cap U_\beta,$$

for each nonempty intersection $U_\alpha \cap U_\beta$. Similarly, we can define a holomorphic p-form on X with coefficients in E to be defined in the same manner, where the transition functions act on holomorphic p-forms instead of differentiable forms. A holomorphic section of E is simply the special case of such a holomorphic p-form, for $p = 0$.

We denote by $\mathcal{O}(E)$ the sheaf of holomorphic sections of the holomorphic line bundle E on X, and we denote by $\Omega^p(E)$ the sheaf of holomorphic p-forms on X with coefficients in E. Thus we can consider the cohomology groups $H^q(X, \mathcal{O}(E))$ and $H^q(X, \Omega^p(E))$, which will occur in Kodaira's vanishing theorem below.

To apply the theory of harmonic forms in this setting, we need a Hermitian metric l on the line bundle E. We suppose that we have a set of transition functions $g_{\alpha\beta}$ for E defined on a covering of X by open sets U_α. We define a Hermitian metric l on E to be a set of positive functions $l_\alpha \in \mathcal{E}(U_\alpha)$ which satisfy

$$\frac{l_\alpha}{l_\beta} = |g_{\alpha\beta}|^2 \text{ on } U_\alpha \cap U_\beta. \tag{14.58}$$

Such metrics always exist (using, e.g., a partition of unity, just as in the case of a metric on the tangent bundle). The metric l is a Hermitian inner product on the fibres E_x, and it induces a conjugate-linear isomorphism

$$\tau : E \to E^*,$$

defined on each fibre of the line bundle E.

If we assume, as before, that X is a Kähler manifold with Kähler metric h, and that E has a Hermitian metric l, then we can define the Hodge inner product in this setting. Namely, if

$$\varphi_x \otimes s_x \in \wedge^{p,q} T_x(X) \otimes E_x,$$

then we define the generalized Hodge $*$-operator

$$\overline{*}_E(\varphi_x \otimes s_x) := \overline{*}\varphi_x \otimes \tau(s_x).$$

If we have two forms φ and ψ in $\mathcal{E}^{p,q}(X, E)$, then we can define the Hodge inner product by

$$(\varphi, \psi) = \int \varphi \wedge \overline{*}_E(\psi). \tag{14.59}$$

The wedge product under the integral sign is defined at a point $x \in X$ in the following manner. If $v \otimes a, w \otimes b \in \wedge^{p,q} T_x \otimes E_x$, then

$$(v \otimes a) \wedge (\overline{*}w \otimes \tau(b)) := v \wedge \overline{*}w < a, \tau(b) >,$$

where $< \ , \ >$ denotes the pairing between E_x and its dual space E_x^*. Hence $\varphi \wedge \overline{*}_E(\psi)$ is a scalar differential form, and the integrand in (14.59) makes sense.

It is easy to verify that the mapping

$$\overline{\partial} : \mathcal{E}^{p,q}(X, E) \to \mathcal{E}^{p,q+1}(X, E)$$

is well-defined, since $\overline{\partial}$ annihilates the transition functions of E. We define the analogue of the Dolbeault groups in this case to be

$$H^{p,q}(X, E) = \frac{\ker \overline{\partial} : \mathcal{E}^{p,q}(X, E) \to \mathcal{E}^{p,q+1}(X, E)}{\operatorname{im} \overline{\partial} : \mathcal{E}^{p,q-1}(X, E) \to \mathcal{E}^{p,q}(X, E)}.$$

The adjoint of $\overline{\partial}$ is given by

$$\overline{\partial}_E^* = -\overline{*}_E \partial \overline{*}_E : \mathcal{E}^{p,q+1}(D, E) \to \mathcal{E}^{p,q}(X, E),$$

just as in the scalar case. Now we can define the Laplacian operator

$$\overline{\Delta}_E = \overline{\partial}\overline{\partial}_E^* + \overline{\partial}_E^* \overline{\partial} : \mathcal{E}^{p,q}(X, E) \to \mathcal{E}^{p,q}(X, E),$$

and we define the harmonic forms as before:

$$\mathcal{H}^{p,q}(X, E) := \{\varphi \in \mathcal{E}^{p,q}(X, E) : \overline{\Delta}_E(\varphi) = 0\}.$$

Using the harmonic theory as we have earlier, we have the representation of the sheaf cohomology group and its corresponding Dolbeault group with coefficients in E in terms of harmonic forms

$$H^q(X, \Omega^p(E)) \cong H^{p,q}(X, E) \cong \mathcal{H}^{p,q}(X, E),$$

which is again a result of Kodaira [126].

We can now use this harmonic representation of the Dolbeault groups to obtain the vanishing theorem that we need for the embedding theorem. For this we will need to use the *Chern class* of the holomorphic line bundle E which we introduced earlier (see (14.31)), and which we will now reformulate with a differential-geometric definition. In 1946 Shiing-Shen Chern (1911–2004) [44] showed how curvature of Hermitian vector bundles (including the tangent bundle of a given Hermitian complex manifold, for instance) could be used to create characteristic classes for such a bundle, which were topological obstructions to the triviality of the bundle. These characteristic classes are now called *Chern classes*. This was very analogous to Hodge using differential geometry and harmonic forms to represent cohomology groups of a manifold, and it proved to be an important advance in our geometric understanding of vector bundles. We will see the use of Chern classes for holomorphic vector bundles in Sect. 14.7.

Let us now define Chern classes for holomorphic line bundles via differential geometry. Let E be a holomorphic line bundle on a compact complex manifold, as before. We define a Hermitian metric l on E (by using transition functions $g_{\alpha\beta}$ for E) by letting l_α be a positive differentiable function defined on each U_α which satisfies (14.58). We define on each U_α the $(1, 1)$-form

$$c_\alpha = \frac{i}{2\pi}\overline{\partial}\partial \log l_\alpha.$$

We need to show that the c_α agree on overlapping open sets $U_\alpha \cap U_\beta$, and this is easy to do. On $U_\alpha \cap U_\beta$ we have

$$\begin{aligned}\overline{\partial}\partial \log l_\alpha &= \overline{\partial}\partial \log(g_{\alpha\beta} l_\beta), \\ &= \overline{\partial}\partial \log g_{\alpha\beta} + \overline{\partial}\partial \log l_\beta,\end{aligned}$$

but

$$\overline{\partial}(\partial \log g_{\alpha\beta}) = 0 \text{ on } U_\alpha \cap U_\beta,$$

since the transition functions $g_{\alpha\beta}$ are holomorphic. Thus, we can define

$$c(E, l) := \left\{\frac{i}{2\pi}\overline{\partial}\partial \log l_\alpha(x), x \in U_\alpha\right\}, \tag{14.60}$$

and this is a well-defined (1, 1)-form on X, which we will call the *Chern form* on X with respect to the Hermitian metric l. We want to show that the Chern form is a closed differential form on X.

For this we introduce a useful notational device due to Weil [238]. We define

$$C : \mathcal{E}^{p,q}(X) \to \mathcal{E}^{p,q}(X),$$

by setting

$$C(\varphi^{p,q}) := i^{p-q}\varphi^{p,q}.$$

We then define

$$d^C := C^{-1}dC,$$

and it is then an easy exercise to verify that

$$\begin{aligned} d &= \partial + \overline{\partial}, \\ id^C &= \partial - \overline{\partial}, \\ 2i\overline{\partial}\partial &= dd^C. \end{aligned}$$

So, in each U_α, the Chern form becomes

$$-\frac{1}{\pi}dd^C \log l_\alpha,$$

and hence is a closed two-form on X.

Suppose that l and $\tilde{l}$ are two different metrics on E. We want to know that they define the same cohomology class on the de Rham group $H^2_{\mathrm{dR}}(X)$. Consider the difference of the two Chern forms in U_α

$$-\frac{1}{\pi}(dd^C \log l_\alpha - dd^C \log \tilde{l}_\alpha) = -\frac{1}{\pi}\left(dd^C \log \frac{l_\alpha}{\tilde{l}_\alpha}\right).$$

But

$$\begin{aligned} l_\alpha &= g_{\alpha\beta}l_\beta, \\ \tilde{l}_a &= g_{\alpha\beta}\tilde{l}_\beta, \end{aligned}$$

and thus

$$\frac{l_a}{\tilde{l}_\alpha} = \frac{l_\beta}{\tilde{l}_b},$$

and it follows that

$$-\frac{1}{\pi}\left(d^C \log \frac{l_\alpha}{\tilde{l}_\alpha}\right) = -\frac{1}{\pi}\left(d^C \log \frac{l_\beta}{\tilde{l}_\beta}\right),$$

on $U_\alpha \cap U_\beta$, and thus defines a global one-form ψ on X with

$$c(E, l) - c(E, \tilde{l}) = d\psi,$$

and hence $c(E, h)$ gives a well-defined cohomology class $[c(E, l)]$ in $H^2_{\mathrm{dR}}(X)$.

One can show[13] that the cohomology class defined by the Chern form $[c(E, l)]$ coincides with the cohomology class $c_1(E)$, which we called the Chern class of E defined by (14.31). Here $c_1(E)$ is considered as an element in $H^2(X, \mathbf{C})$ (assuming the identification of this cohomology group with the de Rham group $H^2_{\mathrm{dR}}(X)$) under the inclusion

$$c_1(E) \in H^2(X, \mathbf{Z}) \hookrightarrow H^2(X, \mathbf{C}).$$

We now define the concept of a positive line bundle on X. Let ψ be a differential form of type (1,1) on X, then ψ is said to be a *positive differential form* if locally (on some coordinate chart U)

$$\psi = i \sum_{\mu\nu} \psi_{\mu\nu} dz_\mu \wedge d\bar{z}_\nu,$$

and the coefficient matrix $\psi_{\mu\nu}(x)$ is a positive-definite matrix at each $x \in U$. A holomorphic line bundle $E \to X$ is said to be a *positive line bundle* if there is a real positive d-closed $(1, 1)$-form $\psi \in c_1(E)$, i.e., ψ is a representative of the Chern class of E. A line bundle E is said to be *negative* if its dual bundle E^* is positive. Now we can formulate Kodaira's vanishing theorem.

Theorem 14.3 (Kodaira vanishing theorem [126]) *Let X be an n-dimensional compact complex manifold, and let E be a holomorphic line bundle on X, then:*

(a) If $E \otimes K^*$ is a positive line bundle, then

$$H^q(X, \mathcal{O}(E)) = 0, \ \ q > 0.$$

(b) If E is a negative line bundle, then

$$H^q(X, \Omega^p(E)) = 0, \ \ p + q < n.$$

[13] See, for instance, Wells [239], Chap. III, Sect. 4. for a proof of this. It uses Čech cohomology to represent the sheaf cohomology group $H^2(X, \mathbf{Z})$.

We want to outline Kodaira's proof of his vanishing theorem. First we will need to introduce some differential-geometric concepts that will prove useful for this. In the late nineteenth century development of differential geometry the concepts of a connection, covariant differentiation, and the curvature of a connection were developed, generalizing the original ideas of Riemann. In the early twentieth century, Elie Cartan formulated the ideas of Riemannian geometry and numerous variations thereof in terms of differential forms. Chern, in his theory of characteristic classes, developed Cartan's ideas further for Hermitian vector bundles on a complex manifold. We will use Chern's methodology in our context for holomorphic line bundles.

Let E be a holomorphic line bundle on a compact complex manifold X, and let D be a *connection* [14] on E. That is, D is a linear mapping

$$D : \mathcal{E}^0(X, E) \to \mathcal{E}^1(X, E)$$

satisfying

$$D(f\varphi) = df \cdot \varphi + f D\varphi,$$

for $f \in \mathcal{E}(X)$, $\varphi \in \mathcal{E}^0(X, E)$. Using the complex structure of X, we have

$$\mathcal{E}^1(X, E) = \mathcal{E}^{1,0}(X, E) \oplus \mathcal{E}^{0.1}(X, E),$$

and thus see that the connection D is the sum of two linear mappings

$$D = D' + D'',$$

where

$$D'(\varphi) \in \mathcal{E}^{1,0}(X, E),$$
$$D''(\varphi) \in \mathcal{E}^{0,1}(X, E),$$

for $\varphi \in \mathcal{E}^0(X, E)$.

If l is now a Hermitian metric on E, then there is a unique connection D on E satisfying

$$d(\varphi, \psi) = (D\varphi, \psi) + (\varphi, D\psi),$$

for $\varphi, \psi \in \mathcal{E}^k(X, E)$, and

$$D''\varphi = 0, \text{ if } \varphi \in \Omega^p(X, E).$$

[14] We will use the specific notation and terminology from Wells [239]; see also Chern [45], Griffiths and Harris [91].

Here we have used the natural generalization of the Hermitian metric on E to differential form-valued sections of E.

It follows that locally (for a suitable covering U_α of X and transition functions $g_{\alpha\beta}$ for E) we can represent $D = D' + D''$ in the form

$$\begin{aligned} D'\varphi_a &= \partial\varphi_a + \theta_\alpha \wedge \varphi_a, \\ D''\varphi &= \overline{\partial}\varphi_\alpha, \end{aligned}$$

where θ_α is a $(1,0)$-form given by

$$\theta_\alpha = l_\alpha^{-1}\partial l_\alpha,$$

a local connection one-form.

The *curvature* of E, defined by

$$\Theta = D^2,$$

is a (1, 1)-form on X defined locally by

$$\Theta_\alpha = \overline{\partial}(l_\alpha^{-1}\partial l\alpha).$$

Thus we see that the Chern form that we defined earlier is defined by the curvature of the connection D, namely

$$c(E,l) = \frac{i}{2\pi}\Theta,$$

and the properties of curvature yield corresponding properties of the characteristic class (Chern class) of the line bundle, a principal idea of Chern.

Now we need to compute the adjoints of these operators in terms of the Hermitian metric l on E and a Hermitian metric h on X, using the Hodge inner product. We find that (locally)

$$\begin{aligned} (D')^*_E &= = - * \overline{\partial} * = \partial^*, \\ (D'')^*_E &= - * \partial * + w * \theta_\alpha *, \end{aligned}$$

where the mapping w is a $\mathbf{C}$-linear mapping defined by

$$w = ** = (-1)^k\varphi, \text{ for } \varphi \in \mathcal{E}^k(X, E).$$

Let now Ω be the fundamental form on X associated to the Hermitian metric h on X, and define the $\mathbf{C}$-linear mapping

$$L : \mathcal{E}^k(X) \to \mathcal{E}^{k+2}(X)$$

by

$$L(\varphi) = \Omega \wedge \varphi, \ \ \varphi \in \mathcal{E}^k(X),$$

and let L^* denote its adjoint. This operator has many important roles in the geometry of Hermitian and Kähler manifolds, and we will need to use it briefly, as we see below. We first need to describe a few elementary properties[15] of L. First, it is easy to see that

$$L^* : \mathcal{E}^k(X) \to \mathcal{E}^{k-2}(X), \ \text{ and } L^*(\varphi) = w * L*, \varphi \in \mathcal{E}^k(X).$$

Secondly, if φ is a (p, q)-form on X, then

$$(L^*L - LL^*)\varphi = (n - p - q)\varphi. \tag{14.61}$$

Now we assume that X is a compact Kähler manifold with a fundamental form Ω. Then there are a number of identities relating the operators $d = \partial + \overline{\partial}$ and L and L^* on X (see, for instance, Wells [239], Sect. V.4)). Some of these were used in proving $2\overline{\Delta} = \Delta$, which we made use of earlier. We single out one that we will need in our proof below:

$$\overline{\partial} L^* - L^* \overline{\partial} = i\partial^*. \tag{14.62}$$

We have two fundamental differential-geometric inequalities that are the key to Kodaira's proof. Let $\varphi \in \mathcal{H}^{p,q}(X, E)$, then

$$\left(\frac{i}{2}\right)(\Theta \wedge L^*\varphi, \varphi) \leq 0, \tag{14.63}$$

$$\left(\frac{i}{2}\right)(L^*\Theta \wedge \varphi, \varphi) \geq 0. \tag{14.64}$$

These inequalities involving harmonic forms are a version in this complex-analytic setting of Bochner's use of similar inequalities in his paper on curvature and Betti numbers from 1948 [22]. It was generalized to holomorphic vector bundles by Shigeo Nakano in 1955 [162] after Kodaira's work came out in 1954.

We can now outline the proof of these inequalities. First we note that

$$D^2 = (d + \theta)^2 = \Theta,$$

[15] See e.g., Weil [238], Chap. 1 or Wells [239], Sect. V.1 for more details about Hermitian linear algebra.

where the second term is the square of the local representation of the connection D. If follows that, if η is an E-valued differential form of type (p, q), then

$$\Theta \wedge \eta = D^2\eta = (D'\overline{\partial} + \overline{\partial}D')\eta,$$

where we note that $(D')^2\eta = 0$, by comparing types. Namely, $\Theta \wedge \eta$ must be of type $(p+q, q+1)$.

Thus we have, for our harmonic form φ,

$$i(\partial^*\varphi, \partial^*\varphi) = ((\overline{\partial}L^* - L^*\overline{\partial})\varphi, \partial^*\varphi),$$

by (14.62). Since φ is harmonic, we have $\overline{\partial}\varphi = 0$ and $\overline{\partial}_E^* = 0$, and it follows that

$$\begin{aligned} i(\partial^*\varphi, \partial^*\varphi) &= (\overline{\partial}L^*\varphi, \partial^*\varphi) \\ &= (L^*\varphi, \overline{\partial}_E^*\partial^*\varphi) \\ &= (L^*\varphi, \overline{\partial}_E^*\partial^*\varphi + \partial^*\overline{\partial}_E^*\varphi), \end{aligned}$$

since $\overline{\partial}_E^*\varphi = 0$. Taking adjoints twice in this expression, we obtain

$$\begin{aligned} i(\partial^*\varphi, \partial^*\varphi) &= ((D'\partial + \partial D')L^*\varphi, \varphi) \\ &= (\Theta \wedge L^*\varphi, \varphi), \end{aligned}$$

and this yields the inequality (14.63). The second inequality (14.64) is proved in a similar fashion.

To continue now with the proof of Kodaira's vanishing theorem, we let E be a negative line bundle on X and let

$$\Omega = -\left(\frac{i}{2}\right)\Theta$$

be a fundamental form for a Kähler metric on X. We subtract (14.64) from (14.63), obtaining

$$\left(\frac{i}{2}\right)((\Theta \wedge L^* - L^*\Theta)\varphi, f) \leq 0,$$

for a harmonic (p, q) form $\varphi \in \mathcal{H}^{p,q}(X, E)$. Now we use

$$L\varphi = -\left(\frac{i}{2}\right)\Theta \wedge \varphi,$$

obtaining

$$((L^*L - LL^*)\varphi, \varphi) \leq 0.$$

Now, using the commutator relation for L and L^* (14.61), we obtain

$$(n - p - q)(\varphi, \varphi) \leq 0,$$

which gives

$$\varphi = 0, \text{ for } p + q < n,$$

and hence

$$\mathcal{H}^{p,q}(X, E) = 0, \text{ for } p + q < n,$$

and this completes our outline of a proof of part (b) of Theorem 14.3.

To prove part (a) of the same theorem, we note that there is a conjugate-linear isomorphism

$$\mathcal{H}^{p,q}(X, E) \cong \mathcal{H}^{n-p,n-q}(X, E^*), \tag{14.65}$$

due to Kodaira [126], which is a special case of Serre duality [215] for compact complex manifolds. Kodaira proves this in a straightforward manner using the theory of harmonic forms that we have developed in this chapter. Serre's theorem is valid for all complex manifolds (using cohomology with compact supports) and uses the natural duality between distributions and functions with compact support. We now let $E \otimes K^*$ be positive, as in the hypothesis of part (a) of Theorem 14.3, and hence $E^* \otimes K$ is a negative line bundle. Therefore by part (b) we have

$$\mathcal{H}^{p,q}(X, E^* \otimes K) = 0, \quad p + q < n,$$

which means by (14.65) that

$$\mathcal{H}^{n-p,n-q}(X, E \otimes K^*) = 0, \quad p + q < n.$$

Let now $p = 0$, and we have

$$\mathcal{H}^{n,n-q}(X, E \otimes K^*) = 0, \quad q < n. \tag{14.66}$$

Now we recall the Dolbeault representation that

$$\mathcal{H}^{n,n-q}(X, F) \cong H^q(X, \Omega^n(F)) = H^q(X, K \otimes F),$$

for a holomorphic line bundle, since the canonical bundle K is the line bundle whose holomorphic sections are holomorphic n-forms. Using this, we see that (14.66) becomes

$$\begin{aligned}0 &= \mathcal{H}^{n,n-q}(X, E \otimes K^*)\\ &= \mathcal{H}^{q,n-q}(X, K \otimes E \otimes K^*)\\ &= \mathcal{H}^{0,n-q}(X, E)\end{aligned}$$

for $q < n$, which means, by relabeling the index q,

$$\mathcal{H}^{0,q}(X, E) = 0, \ q > 0,$$

which gives part (a) of the theorem.

14.6 The Kodaira Embedding

We formulated the Kodaira embedding theorem in Theorem 14.1, and we now want to outline the proof of this theorem, where we will use the various tools from sheaf theory and the theory of harmonic forms that we have discussed in the previous sections.

Suppose now that X is a Hodge manifold with a closed fundamental form Ω which represents an integral cohomology class in $H^2(X, \mathbf{Z})$. Consider the exact sequence

$$\cdots \to H^1(X, \mathcal{O}^*) \xrightarrow{\delta} H^2(X, \mathbf{Z}) \xrightarrow{\kappa} H^2(X, \mathcal{O}) \to \cdots \tag{14.67}$$

coming from

$$0 \to \mathbf{Z} \to \mathcal{O} \to \mathcal{O}^* \to 0.$$

Let $[\Omega]$ denote the cohomology class of Ω in $H^2(X, \mathbf{Z})$. Any two-form on X has a decomposition of the form

$$\varphi = \varphi^{2,0} + \varphi^{1,1} + \varphi^{0,2}.$$

Thus the $(1, 1)$-form Ω has $\Omega^{2,0} = \Omega^{0,2} = 0$.

Since

$$H^{0,2}(X) \cong H^2(X, \mathcal{O}),$$

it then follows that $\kappa([\Omega]) = 0$ in (14.67), and thus there is a holomorphic line bundle $E \to X$ such that

$$c_1(E) = \delta(E) = [\Omega].$$

Now E is a positive line bundle, since it has a positive Chern form of type $(1, 1)$ representing its Chern class. We will see below that a suitable tensor power of E, of the form $F = E^{\mu}$, for some $\mu > 0$, will be a line bundle that will give an embedding of X into $\mathbf{P}_N$ for a suitable N.

We now need to define several specific subsheaves of the sheaf $\mathcal{O}$ of holomorphic functions on X. We recall our example (14.35) of an ideal sheaf m_p of holomorphic functions that vanish at a point $p \in X$. We now define $m_{p,q}$ to be the subsheaf of $\mathcal{O}$ of holomorphic functions that vanish at two points p and q. If the points p and q coincide, then this is the sheaf of functions that vanish at p to second order, which we denote by m_p^2. It is easy to verify that the quotient sheaves have the form

$$(\mathcal{O}/m_{p,q})_x \cong \begin{cases} \mathbf{C}, & \text{if } x = p \text{ or } q, \\ 0, & \text{if } x \neq p \text{ or } q, \end{cases}$$
$$(\mathcal{O}/m_p^2)_x \cong \begin{cases} \mathbf{C} \oplus \mathbf{C}^n, & \text{if } x = p, \\ 0, & \text{if } x \neq p. \end{cases}$$

We have the corresponding short exact sequences,

$$\begin{aligned} &0 \to m_{p,q} \to \mathcal{O} \to \mathcal{O}/m_{p,q} \to 0, \\ &0 \to m_p^2 \to \mathcal{O} \to \mathcal{O}/m_p^2 \to 0, \end{aligned}$$

and we can tensor these with the sheaf $\mathcal{O}(F)$, obtaining

$$\begin{aligned} &0 \to m_{p,q} \otimes \mathcal{O}(F) \to \mathcal{O}(F) \to \mathcal{O}/m_{p,q} \otimes \mathcal{O}(F) \to 0, \\ &0 \to m_p^2 \otimes \mathcal{O}(F) \to \mathcal{O}(F) \to \mathcal{O}/m_p^2 \otimes \mathcal{O}(F) \to 0. \end{aligned}$$

These short exact sequences induce the long exact sequences

$$\begin{aligned} &\cdots \to H^0(X, \mathcal{O}(F)) \xrightarrow{r} H^0(X, \mathcal{O}/m_{p,q} \otimes \mathcal{O}(F)) \to H^1(X, m_{p,q} \otimes \mathcal{O}(F)) \to \cdots, \\ &\cdots \to H^0(X, \mathcal{O}(F)) \xrightarrow{s} H^0(X, \mathcal{O}/m_p^2 \otimes \mathcal{O}(F)) \to H^1(X, m_p^2 \otimes \mathcal{O}(F)) \to \cdots. \end{aligned} \tag{14.68}$$

We will show that if the two mappings r and s in (14.68) are *surjective*, then a basis for the finite-dimensional vector space of holomorphic sections

$$H^0(X, \mathcal{O}(F))$$

yields an embedding of X into $\mathbf{P}_N$, where

$$\dim H^0(X, \mathcal{O}(F)) = N + 1.$$

This will then give the required Kodaira embedding.

If we knew that

$$\begin{aligned} &H^1(X, m_{p,q} \otimes \mathcal{O}(F)) = 0, \\ &H^1(X, m_p^2 \otimes \mathcal{O}(F)) = 0, \end{aligned} \tag{14.69}$$

then it would follow that r and s were surjective, and the Kodaira embedding theorem would follow. The vanishing of the two cohomology groups (14.69) was proved in 1962 by Grauert [87], which, at the time, provided a new proof of the Kodaira embedding theorem. In fact, Grauert extended Kodaira's results to compact complex spaces (allowing singularities) with an alternative definition of positive line bundles in the more general case. This work followed up on some of the methodology of Grauert's proof of the real-analytic embedding theorem that we will see in Sect. 15.7.

Kodaira used the theory of harmonic forms to show that the mappings r and s in (14.68) are surjective, but he had to do this somewhat indirectly, as the harmonic theory using partial differential equations was only valid for locally free holomorphic sheaves, and the sheaves in (14.69) are not locally free. We will see how Kodaira did this later in this section. First, we will show that if r and s are surjective, then we obtain an embedding.

Let us look at the mapping r in (14.68), and we assume that it is surjective. It follows that given any two points in the stalks $(\mathcal{O}(F)/m_{p,q})_p$ and $(\mathcal{O}(F)/m_{p.q})_q$, there is a section s which takes on these values at these two points. It follows that we can find two sections

$$s_1, s_2 \in H^0(X, \mathcal{O}(F))$$

such that

$$\begin{aligned} s_1(p) = 0, s_1(q) \neq 0, \\ s_2(p) \neq 0, s_2(q) = 0. \end{aligned} \tag{14.70}$$

If we now let $\{s_0, s_1, \ldots, s_N\}$ be a basis for $H^0(X, \mathcal{O}(F))$, then it follows from (14.70) that, for each $x \in X$, there is at least one section s_j with the property that $s_j(x) \neq 0$. Hence the Kodaira mapping defined in (14.19)

$$\Phi : X \to \mathbf{P}_N,$$

is a well-defined holomorphic mapping, and it follows easily from (14.70) that Φ is a one-to-one mapping.

Now we suppose that the mapping s in (14.68) is surjective. Let's see what this looks like locally at the point p. Let $s \in H^0(X, \mathcal{O}(F))$, then in a coordinate chart U near p, where we assume that $p = (0, \ldots, 0)$, and where we assume that the section s is represented by a holomorphic function f, then the first two terms of the Taylor series at p have the form

$$f(0) + \sum_{\mu=1}^{n} \frac{\partial f}{\partial z_\mu}(0) z^\mu.$$

This is then a representative of the section s in the stalk of $\mathcal{O}(F)/m_p^2$ at p. To say that the mapping s in (14.68) is surjective means that locally at p, the mapping

$$f \mapsto f(0) + df(0)$$

is surjective, and this implies that the induced mapping locally at p from the Kodaira mapping $\Phi : X \to \mathbf{P}_N$ is regular. Thus the mapping Φ is an embedding, as required.

Now we still need to understand *why* these mappings r and s in (14.68) are surjective. The problem we face is that the two sheaves $m_{p,q}$ and m_p^2 are not locally free. Thus they don't correspond to holomorphic vector bundles, and the theory of harmonic forms wouldn't apply. Kodaira introduces a new trick at this point, which allows him to use the vanishing theorem that he proved using harmonic forms. The basic idea is to blow up the manifold X at the points p and q in a suitable manner which converts the problem to one involving locally free sheaves. To "blow up a manifold at a point" is a specific classical procedure from algebraic geometry which replaces the point p by a copy of $\mathbf{P}_{n-1}$ in such a way that X remains a smooth manifold. This has been used for a long time by algebraic geometers to resolve singularities. More formally, it is called the quadratic transform of the manifold X at the point p, which we now briefly describe.

Let Y be an n-dimensional complex manifold, and suppose $p \in Y$. Let U be a coordinate chart near p, where we let $p = (0, \ldots, 0)$ in these coordinates. Consider

$$W = \{(x, t) \in U \times \mathbf{P}_{n-1} : t_\alpha z_\beta - t_\beta z_\alpha = 0, \alpha, \beta = 1, \ldots, n\}, \tag{14.71}$$

where $(t_1, \ldots, t_n)$ are homogeneous coordinates for $\mathbf{P}_{n-1}$. Then W is a holomorphic submanifold of $U \times \mathbf{P}_{n-1}$, and there are natural projections

$$\pi : W \to U,$$

given by

$$\pi(z, t) = z,$$

and

$$\sigma : W \to \mathbf{P}_{n-1},$$

given by

$$\sigma(z, t) = t.$$

It follows from the above construction that

$$S_p := \pi^{-1}(0) = \{0\} \times \mathbf{P}_{n-1},$$

and that if we restrict π to $W - S_p$, we see that

$$\pi_{|W-S_p} : W - S_p \to U - \{0\}$$

is a biholomorphic mapping.

We now define the *quadratic transform* of Y at p to be:

$$Q_p Y := \begin{cases} W, & x \in U, \\ Y - U, & x \in Y - U, \end{cases}$$

and we extend the projection π to all of $Q_p Y$ in a natural manner:

$$\pi_p : Q_p Y \to Y,$$

being defined by π in W and by the identity mapping on $Y - W$.

Now we return to our compact complex manifold and the mappings r and s in (14.68). We consider the two given points p and q, where $p \neq q$, and we can perform the double quadratic transform with its projection

$$\pi_{p,q} : Q_p Q_q X \to X,$$

which is well-defined since each quadratic transform is local near p and q, respectively. We set $\tilde{X} := Q_p Q_q X$ and let $\tilde{\mathcal{O}}$ be the sheaf of holomorphic functions on $\tilde{X}$. We define the two hypersurfaces in $\tilde{X}$ by

$$\begin{aligned} S_p &= \pi_{p,q}^{-1}(p), \\ S_q &= \pi_{p,q}^{-1}(q). \end{aligned} \tag{14.72}$$

Then we let

$$S := S_p \cup S_q$$

be the holomorphic submanifold of codimension one in $\tilde{X}$ defined by the blowups at the points p and q (14.72).

We want to reduce the surjectivity of r and s in (14.68) to a different sheaf-theoretic vanishing theorem where we will be able to utilize the Kodaira vanishing theorem. We will first illustrate how this is done for the mapping r, and then indicate how this procedure can be modified to obtain a similar result for the mapping s.

We can now define the ideal sheaf $\mathcal{I}$ of holomorphic functions on $\tilde{X}$ which vanish on the hypersurface S.

We recall that part of our hypothesis is that we have a positive line bundle $E \to X$, and that $F = E^{\mu}$ is defined as a power of E for a not yet specified power μ. We let

$$\tilde{F} := \pi^* F$$

be the pullback line bundle of F on $\tilde{X}$, and we then have the short exact sequences

$$\begin{array}{ccccccccc} 0 \to & \mathcal{I} \otimes \tilde{\mathcal{O}}(\tilde{F}) & \to & \tilde{\mathcal{O}}(\tilde{F}) & \to & \tilde{\mathcal{O}}(\tilde{F})/\mathcal{I} & \to 0, \\ & \uparrow \pi_1^* & & \uparrow \pi^* & & \uparrow \pi_2^* & \\ 0 \to & m_{p,q} \otimes \mathcal{O}(F) & \to & \mathcal{O}(F) & \to & \mathcal{O}(F)/m_{p,q} & \to 0, \end{array}$$

where π_1 and π_2 are the naturally induced mappings from $\pi_{p,q}$. This then yields the pair of long exact sequences

$$\begin{array}{ccccccc} \cdots \to H^0(\tilde{X}, \tilde{\mathcal{O}}(\tilde{F})) & \xrightarrow{\tilde{r}} & H^0(\tilde{X}, \tilde{\mathcal{O}}(\tilde{F}) \otimes \tilde{\mathcal{O}}/\mathcal{I}) & \to & H^1(\tilde{X}, \tilde{\mathcal{O}}(\tilde{F}) \otimes \mathcal{I}) & \to \cdots, \\ \alpha \downarrow \uparrow \pi^* & & \uparrow \pi_2^* & & \uparrow \pi_1^* & \\ \cdots \to H^0(X, \mathcal{O}(F)) & \xrightarrow{r} & H^0(X, \mathcal{O}(F) \otimes \mathcal{O}/m_{p,q}) & \to & H^1(X, \mathcal{O}(F) \otimes m_{p,q}) & \to \cdots. \end{array} \tag{14.73}$$

We can now show that the first vertical mapping π^* in this diagram is an isomorphism. We have denoted the inverse by α, and we now show that it is indeed a well-defined inverse. We note first that for $n = 1$, $\tilde{X} = X$, so there is nothing to show, since π is then biholomorphic. If $n > 1$, then we see that

$$\pi : \tilde{X} - S \to X - \{p, q\}$$

is biholomorphic, and we can define the inverse

$$\alpha := ((\pi^{-1})^* : H^0(\tilde{X} - S, \tilde{\mathcal{O}}(\tilde{F})) \to H^0(X - \{p, q\}, \Omega(F))$$

on these open sets which exclude the points $\{p, q\}$ and their blowups. We need to show that this mapping extends across the exceptional sets. Suppose that $s \in \mathrm{H}^0(\tilde{X} - S, \tilde{\mathcal{O}}(\tilde{F}))$, and consider $\alpha(s)$. Then locally, near either p or q, $\alpha(s)$ can be represented as a holomorphic function on a punctured neighborhood of such a point, and by Hartogs's theorem[16] the function analytically continues across the point. It follows that π^* in (14.73) is indeed an isomorphism with inverse α.

Moreover, it is easy to show that the mapping π_2^* in (14.73) is an injective mapping, and we will use this fact below.

Now, we can check that if

$$H^1(\tilde{X}, \mathcal{I} \otimes \tilde{\mathcal{O}}(\tilde{F})) = 0, \tag{14.74}$$

then the mapping r in (14.73) (which is the same as the original mapping r in (14.68)) is surjective. Namely, if $s \in H^0(X, \mathcal{O}(F)/m_{p,q})$, then, by the vanishing theorem

[16]Hartogs proved that there are no isolated singularities of holomorphic functions of more than one complex variable [96]. See Sect. 15.2 for a discussion of the important role this theorem played in the theory of functions of several complex variables in the first half of the twentieth century.

(14.74), $\tilde{r}$ in (14.73) is surjective, and there is then an element $\tilde{s} \in H^0(\tilde{X}, \tilde{\mathcal{O}}(\tilde{F}))$ such that

$$\tilde{r}(\tilde{s}) = \pi^*(s),$$

and thus we obtain

$$r(\alpha(\tilde{s}) = s,$$

and thus r is surjective (using the fact that π_2^* is injective).

One can prove in the same manner that the mapping s in (14.68) is surjective by using the ideal sheaf $\mathcal{I}^2$ of holomorphic functions on $\tilde{X}$ that vanish to second order on S. For this, one would need to have a similar vanishing theorem of the form

$$H^1(\tilde{X}, \tilde{\mathcal{O}}(\tilde{F}) \otimes \mathcal{L}^2) = 0. \tag{14.75}$$

As we saw in Sect. 14.2, to any hypersurface in a complex manifold, we can associate a holomorphic line bundle. Let L be the line bundle on $\tilde{X}$ associated to the hypersurface $S = S_p \cup S_q$, the blowups of the points p and q. The ideal sheaf of holomorphic functions on $\tilde{X}$ which vanish on the hypersurface S, which we have denoted by $\mathcal{I}$ and used extensively in the last few paragraphs, corresponds to the sheaf of sections of the line bundle L^*. More precisely,

$$\mathcal{I} \cong \tilde{\mathcal{O}}(L^*),$$

as is easy to verify.

Consider now the line bundle

$$\pi^*(E^\mu) \otimes L^* \otimes K_{\tilde{X}}^*, \tag{14.76}$$

for the case of two distinct points p and q, as above. If we can show that this line bundle is positive for some sufficiently large μ, then it would follow from the Kodaira vanishing theorem (Theorem 14.3) that

$$\mathrm{H}^1(\tilde{X}, \tilde{\mathcal{O}}(\pi^* E^\mu) \otimes \mathcal{I}) = 0,$$

which is what we needed in (14.74) above.

In order to show that the tensor product (14.76) is positive, we will need to have metrics on each of the factors in the product and compute and compare the curvatures. First, we have, by hypothesis, a metric l on E such that the curvature Θ_E is positive on X. It follows that l induces a metric on $\pi^* E^\mu$, and it is easy to see that

$$\Theta_{\pi^* E^\mu} = \mu \Theta_{\pi^* E}.$$

This curvature is positive on $\tilde{X} - S$, and, since S is lower-dimensional, it is positive semidefinite on $\tilde{X}$.

Now we let W_p and W_q be coordinate neighborhoods of S_p and S_q where we have the representation of the quadratic transforms of X at the points p and q as in (14.71), and we let $W = W_p \cup W_q$. It is not difficult to show, using the coordinates (z, t) in W_p, that

$$L^*_{|W_p} \cong \sigma^*(H),$$

where H is the hyperplane section bundle on $\mathbf{P}_{n-1}$. We define a metric on H by

$$k_\alpha = \log \frac{|t_\alpha|^2}{(|t_1|^2 + \cdots + t_n|^2)} \text{ in } V_\alpha = \{t_\alpha \neq 0\} \subset \mathbf{P}_{n-1}.$$

The curvature

$$\Theta_H = \overline{\partial}\partial \log k_a$$

is the Fubini–Study (1, 1)-form given by (14.14) (ignoring the factor of $i/2$), and is positive. This induces a metric k_p and curvature on $\sigma^*(H)$ which is positive semidefinite and is positive-definite in the t direction in the open set W_p. We can perform the same construction in W_q.

Now let ρ be a cutoff function which is $\equiv 1$ near S_p and S_q and is $\equiv 0$ on a neighborhood of the compact set $\tilde{X} - W$. The line bundle L^* is trivial on $\tilde{X} - S$, and we can put a constant metric k_0 on $L^*_{\tilde{X}-S}$. We can then define a metric on L^* by

$$k = \rho k_p + \rho k_q + (1 - \rho) k_0.$$

The curvature of L^* with respect to this metric is positive-semidefinite near S_p and S_q and is positive-definite in the t-directions, also near S_p and S_q.

Now we need a good representation of $K_{\tilde{X}}$. Again, using the local coordinate charts W_p and W_q, one can verify that

$$K_{\tilde{X}} \cong (L^*)^{n-1} \otimes \pi^* K_X,$$

where the factor $(L^*)^{n-1}$ is the contribution to $K_{\tilde{X}}$ coming from the quadratic transforms of X at p and q. The metric h on X induces a metric on the canonical bundle $K = \wedge^n T^*(X)$, and hence on the pullback bundle $\pi^* K_X$.

Our tensor product (14.76) now has the form

$$\pi^* E^\mu \otimes (L^*)^n \otimes \pi^* K_X. \tag{14.77}$$

Let's rewrite this as

$$\pi^* E^{\mu_1} \otimes (L^*)^n \otimes \pi^* E^{\mu_2} \otimes \pi^* K_X, \tag{14.78}$$

where $\mu = \mu_1 + \mu_2$ and μ_1 and μ_2 are both positive integers.

We note the important fact that if $G \otimes H$ is a tensor product of holomorphic line bundles with Hermitian metrics, then the curvature of the tensor product is the sum of the curvatures of the factors, i.e.,

$$\Theta_{G\otimes H} = \Theta_G + \Theta_H,$$

as is easy to see from the Chern form representation of curvature (14.60).

Now consider the sum of curvatures of the first two factors in (14.78),

$$\mu_1 \Theta_{\pi^* E} + n\Theta_{L^*}. \tag{14.79}$$

The first term is positive-semidefinite on $\tilde{X}$; and in the coordinate neighborhood W_p near the hypersurface S_p the first term is positive-definite in the z-direction and the second term is positive-definite in the t-direction. This is also true near the hypersurface S_q, and hence their sum is positive-definite in a neighborhood of $S = S_p \cup S_q$. The first term is positive-definite on the closure of the set

$$\{x \in \tilde{X} : \rho(x) < 1\},$$

and hence we can choose a μ_1^0 such that the sum in (14.79) is positive-definite for any $\mu_1 \geq \mu_1^0$.

Let us now consider the last two factors in (14.78). We have

$$\mu_2 \Theta_{\pi^* E} + \Theta_{\pi^* K_X} = \pi^*(\mu_2 \Theta_E + \Theta_{K_X}).$$

Since Θ_E is a positive-definite $(1,1)$-form on the compact manifold X, there is a positive integer m_2^0 such that

$$\mu_2 \Theta_E + \Theta_{K_X}$$

is positive on X for all $\mu_2 \geq \mu_2^0$. The pullback form

$$\pi^*(\mu_2 \Theta_E + \Theta_{K_X})$$

will then be a positive semidefinite $(1,1)$-form on $\tilde{X}$; thus we obtain, for $\mu_1 \geq \mu_1^0$, $\mu_2 \geq \mu_2^0$, that the tensor product (14.78) is a positive line bundle.

The above argumentation was for a fixed p and q on X. By continuity, the estimates would also be true for points near p and q. By covering X by a finite covering of suitable open sets, we would be able to conclude that there is an integer μ^0 such that the tensor product (14.78) is positive for $\mu \geq \mu^0$, for all pairs of distinct points p and q on X.

A variation of this argument will yield the same result for the tensor product of line bundles

$$\pi^* E^\mu \otimes (L^*)^2 \otimes K_{\tilde{X}},$$

where here L^* corresponds to the ideal sheaf $\mathcal{I}$ for the hypersurface S_p for a single point p. This then concludes our outline of the proof of the Kodaira embedding theorem.

14.7 Riemann–Roch Theorems in Higher Dimensions

At the end of Chap. 10 we discussed the celebrated Riemann–Roch theorem for Riemann surfaces. We want to revisit this theorem and show how it can be expressed in terms of sheaf theory, sheaf cohomology and contemporary algebraic topology. These are the same tools we have used in this chapter on the Kodaira embedding theorem. This reformulation has led to simpler proofs as well as to quite important generalizations to higher-dimensional manifolds that we want to outline in this section.

Let now X be a compact one-dimensional complex manifold of genus g[17] (a compact Riemann surface), and let D be a divisor on X as defined in Sect. 10.4, namely,

$$D = \sum_l n_l p_l,$$

where p_l is a discrete set of points in X, and n_l are integers. We can associate to this divisor a holomorphic line bundle. We can describe this divisor as the zeros and poles with multiplicities of meromorphic functions defined in the neighborhood of each point p_l (e.g., $f_\lambda \in \mathcal{M}(U_l)$, $l = 1, \ldots, L$, where U_l is a neighborhood of p_l and $f_0 \equiv 1$ in the open set $U_0 = X - \cup p_l$). In the intersections $U_\alpha \cap U_\beta$, we let $g_{\alpha\beta} = f_\alpha / f_\beta$, which defines a holomorphic line bundle E_D on X. Linearly equivalent divisors ($D - D' = (f)$, for some globally defined meromorphic function f) correspond to holomorphically equivalent holomorphic line bundles ($E_D \cong E'_D$).

We then have the following vector space isomorphism:

$$L(D) = \{f \in \mathcal{M}(X) : (f) + D \geq 0\} \cong H^0(X, \mathcal{O}(E_D)), \tag{14.80}$$

[17] We now use the common designation of g for genus for a Riemann surface, as opposed to the notation p used by Riemann and his successors at the end of the nineteenth century.

which is quite easy to verify. This links the language of the nineteenth century with the sheaf-theoretic language of the mid-twentieth century. Here we are following the paper of Jean-Pierre Serre (1926–) [215], which makes this transition in language and which gives a very elegant proof of the Riemann–Roch theorem that we will outline below. For simplicity, we will let E denote the line bundle associated with the given divisor D in the paragraphs below.

Now let K be the canonical divisor on X (divisor defined by a meromorphic one-form, as in Sect. 10.4), and let K also denote the canonical bundle ($T^*(X)$ in this case of a Riemann surface) on X. We use the same notation for the canonical bundle, which we have used extensively earlier in this chapter, as for the holomorphic line bundle associated to the divisor K. The meromorphic sections of the canonical bundle K are precisely the meromorphic one-forms used in the definition of the canonical divisor K.

The divisors in the Riemann–Roch theorem form an Abelian group which are described additively as $D + D'$, and the corresponding holomorphic line bundles form an Abelian group described in this chapter multiplicatively as $E \otimes E'$. And thus we would have $K - D$, as in the Riemann–Roch theorem, corresponding to $K \otimes E^*$, recalling that the multiplicative inverse of a holomorphic line bundle is its dual bundle. Thus we have the isomorphism

$$L(K - D) \cong H^0(X, \mathcal{O}(K \otimes E^*)).$$

In Serre's paper that we are following here [215], he proves his well-known Serre duality theorem for complex manifolds. A special case of this theorem was proved by Kodaira for compact complex manifolds, which we used earlier in this chapter; see Eq. (14.65). In this special case we have the conjugate-linear isomorphism

$$H^1(X, \mathcal{O}(E)) \cong H^0(X, \Omega^1(E^*)),$$

which is also isomorphic to $H^0(X, \mathcal{O}(K \otimes E^*))$, since the sheaves $\Omega^1(E^*)$ and $\mathcal{O}(K \otimes E^*)$ are isomorphic. Thus we obtain

$$\dim H^1(X, \mathcal{O}(E)) = \dim H^0(X, \mathcal{O}(K \otimes E^*)). \tag{14.81}$$

We recall now the statement of the Riemann–Roch theorem from Sect. 10.4 (Theorem 10.4)

$$\dim L(D) - \dim L(K - D) = \deg(D) + 1 - g, \tag{14.82}$$

where g is the genus of X. Using the isomorphism above,[18] we obtain an equivalent form of the Riemann–Roch theorem

[18] R.C. Gunning's lecture notes on Riemann surfaces [92] have a very readable and much more detailed account of Serre's work concerning the classical and sheaf-theoretic versions of the Riemann–Roch theorem.

$$\dim H^0(X, \mathcal{O}(E)) - \dim H^1(X, \mathcal{O}(E)) = \deg(D) + 1 - g, \tag{14.83}$$

where here E is the line bundle corresponding to the divisor D.

We can now outline Serre's elegant proof of the Riemann–Roch theorem (in the form of (14.83)) using sheaf cohomology. The basic idea is to show that the expression

$$R(D) := \dim H^0(X, \mathcal{O}(E)) - \dim H^1(X, \mathcal{O}(E)) - \deg(D)$$

is a constant. To see this, suppose that D is any divisor, and let $D' = D + p$ for any point $p \in X$, and we simply have to show that $R(D) = R(D')$. Letting E and E' be the line bundles associated to D and D', we have the short exact sequence of sheaves:

$$0 \to \mathcal{O}(E) \to \mathcal{O}(E') \to \mathcal{O}(E')/\mathcal{O}(E) \to 0.$$

This yields the long exact sequence

$$\begin{aligned} 0 \to H^0(X, \mathcal{O}(E)) \to H^0(X, \mathcal{O}(E')) \to H^0(X, \mathcal{O}(E')/\mathcal{O}(E)) \to \\ H^1(X, \mathcal{O}(E)) \to H^1(X, \mathcal{O}(E')) \to H^1(X, \mathcal{O}(E')/\mathcal{O}(E)) \to \cdots . \end{aligned}$$

It is easy to check that

$$\big(\mathcal{O}(E')/\mathcal{O}(E)\big)_x = \begin{cases} \mathbf{C}, & x = p, \\ 0, & x \neq p, \end{cases}$$

and, moreover,

$$H^q(X, \mathcal{O}(E'))/\mathcal{O}(E)) = 0, \ \ q \geq 1,$$

since $\mathcal{O}(E')/\mathcal{O}(E)$ is a fine sheaf (a sheaf which admits a partition of unity; see the references for sheaf cohomology). Also, since we assume X is connected, $H^0(X, \mathcal{O}(E')/\mathcal{O}(E)) \cong \mathbf{C}$.

Thus the long exact sequence above becomes the exact sequence of five vector spaces

$$\begin{aligned} 0 \to H^0(X, \mathcal{O}(E)) \to H^0(X, \mathcal{O}(E')) \to H^0(\mathcal{O}(E')/\mathcal{O}(E)) \to \\ H^1(X, \mathcal{O}(E)) \to \mathrm{H}^1(X, \mathcal{O}(E')) \to 0. \end{aligned}$$

By elementary linear algebra we obtain an alternating sequence of dimensions of the form

$$\begin{aligned} \dim H^0(X, \mathcal{O}(E)) - \dim H^0(X, \mathcal{O}(E')) + 1 \\ - \dim H^1(X, \mathcal{O}(E)) + \dim H^1(X, \mathcal{O}(E') = 0, \end{aligned}$$

and this gives immediately

$$R(D) = \dim H^0(X, \mathcal{O}(E')) - \dim H^1(S, \mathcal{O}(E')) - \deg(D) - 1.$$

But $\deg(D') = \deg(D) + 1$, and thus $R(D)$ is a constant.

To complete Serre's proof of the Riemann–Roch theorem, we simply have to evaluate $R(D)$ for the special case of $D = 0$. We see that

$$R(0) = \dim H^0(X, \mathcal{O}) - \dim H^1(X, \mathcal{O}) - 0.$$

But, by Serre duality (14.65),

$$\dim H^1(X, \mathcal{O}) = \dim H^0(X, \Omega^1) = g,$$

the genus of X. Namely, the space $H^0(X, \Omega^1)$ is the space of holomorphic one-forms on X which we know by Hodge theory is the genus of X. This is also equivalent to Riemann's space of Abelian integrals of the first kind, which he showed in his original paper from 1857 [202] is the same as the genus (see our discussion of this in Chap. 10).

We need one more ingredient to completely transform the Riemann–Roch theorem to its modern form. We recall the Chern class of a holomorphic line bundle $c_1(E) \in H^2(X, \mathbf{Z})$, and that this Chern class can be represented, by de Rham's theorem, as a closed two-form on X. If $D = \sum_l n_l p_l$ is a divisor on X, we have that $\deg(D)$ is the numerical sum of these points with multiplicities, i.e., $\deg(D) = \sum_l n_l$. The Chern class of the line bundle allows us to compute the degree in the following manner:

$$\deg(D) = \int_X c_1(E), \tag{14.84}$$

where the integration refers to the integration of the two-form representing the Chern class $c_1(E)$.

Chern's fundamental paper from 1946 [44] introduced and proved many properties of what are now called *Chern classes* of Hermitian vector bundles on compact differentiable manifolds (he called them "characteristic classes," but they have been called Chern classes ever since). We described and used the concept of the first Chern class $c_1(E)$ for a holomorphic line bundle E in the earlier sections of this chapter. Now we need to use the higher-order Chern classes as well. Our primary interest is complex manifolds, so we restrict our attention to this category of manifolds, although the theory of Chern classes is much more general than this.

Let E be a Hermitian holomorphic vector bundle of rank r on a compact complex manifold of dimension n. Then there is a countable set of Chern classes of the form

$$c_j(E) \in H^{2j}(X, \mathbf{Z}),\ j = 0, \dots, k, \dots \ \text{where } c_j(E) = 0, \ \text{for ; } j > 2r,$$

and where we have set $c_0(E) = 1$.

Chern classes can, in fact, be defined axiomatically in terms of their fundamental properties, and we leave it to the references for the constructions and properties of these characteristic classes of vector bundles, which play such an important role in modern complex geometry (see, e.g., Hirzebruch [104], Chern [45], Griffiths and Harris [91], or Wells [239]).

In summary form, the construction of these classes can be obtained by defining the Chern classes for the universal bundles on Grassmannian manifolds, and then showing that any such holomorphic vector bundle is a pullback from the universal bundle on a suitable Grassmannian manifold, and defining the Chern classes as the pullbacks of specifically defined Chern classes for the universal bundle of a Grassmannian manifold. This shows, in particular, that the Chern classes are integral cohomology classes.

Chern also constructs the Chern classes in terms of curvature forms defined (using certain homogeneous terms of an expansion of a determinant expression involving curvature). This differential-geometric construction (which we saw in the case of a holomorphic line bundle) yields Chern forms which are closed $2j$-forms on X, which define $c_j(E)$ as an element of $H^{2j}(X, \mathbf{R})$. It takes a separate argument to show that these are integral cohomology classes, as can be shown from the Grassmannian definition, for instance, but there are other methods which yield the integrality as well.

We now define the *total Chern class* to be:

$$c(E) = 1 + c_1(E) + c_2(E) + \cdots c_r(E) \in H^*(X, \mathbf{Z}),$$

where $H^*(X, \mathbf{Z})$ is the cohomology ring of the manifold X. The multiplication here can be defined by using the de Rham representation of cohomology and using the exterior multiplication of differential forms.

A fundamental theorem of Chern is his higher-dimensional version of the Gauss–Bonnet theorem [46], which has the following form in terms of Chern classes. Letting $\chi(X)$ be the Euler characteristic of the manifold X, i.e., the alternating sum of Betti numbers,

$$\chi(X) = \sum_{k=0}^{2n} (-1)^k b_k,$$

then the Chern–Gauss–Bonnet theorem has the form:

$$\chi(X) = \int_X c_n(T(X)).$$

For dim $X = 1$, with genus g, this is simply

$$\int c_1(T(X)) = 2 - 2g.$$

Now we can rewrite the Riemann–Roch theorem for Riemann surfaces in terms of Chern classes, and from (14.83) we obtain

$$\dim H^0(X, \mathcal{O}(E)) - \dim H^1(X, \mathcal{O}(E)) = \int_X c_1(E) + \frac{1}{2}\left[c_1(T(X))\right]. \quad (14.85)$$

This is now the version of Riemann–Roch that generalizes to higher dimensions. In 1954 Hirzebruch announced [103] what is now called the *Hirzebruch–Riemann–Roch theorem* and published his results in a monograph in 1955 [105]. This was republished as a considerably expanded English monograph in 1966 [104]. We now have the language to formulate his theorem, which we now proceed to do.

Let now X be a compact complex manifold of dimension n, and let E be a holomorphic vector bundle of rank r on X. Hirzebruch defines the Euler characteristic for a holomorphic vector bundle E on X to be:

$$\chi(X, E) := \sum_{q=0}^{n} (-1)^q H^q(X, \mathcal{O}(E)), \quad (14.86)$$

and then has his version of the Riemann–Roch theorem in arbitrary dimensions:

Theorem 14.4 (Hirzebruch–Riemann–Roch [104]) *Let E be a holomorphic vector bundle on a projective algebraic manifold X, then there exists a homogeneous polynomial $p(X_1, \ldots, X_r, Y_1, \ldots, Y_n)$ with rational coefficients such that*

$$\chi(X, E) = \int p(c_1(E), \ldots, c_r(E), c_1(T(X)), \ldots, c_n(T(X))). \quad (14.87)$$

Hirzebruch has a very explicit formula for the polynomial p in this theorem in terms of Todd polynomials for the tangent bundle and Chern classes of the holomorphic vector bundle E (see Hirzebruch's monograph [104]).

We give a couple of examples to illustrate the theorem. First, we have the case of the classical Riemann–Roch theorem as formulated in Eq. (14.85), where here $p(X_1, Y_1) = X_1 + \frac{1}{2}Y_1$. Second, we consider the case of a holomorphic line bundle L on a projective algebraic manifold X of dimension two, and Hirzebruch's theorem simplifies to

$$\chi(X, L) = \int_X \frac{1}{2}[c_1(L)^2 + c_1(L)c_1(T(X))] + \frac{1}{12}[c_1(T(X))^2 + c_2(T(X))]. \quad (14.88)$$

We remind the reader that the products are in the cohomology ring $H^*(X, \mathbf{Q})$. A number of special cases of the Hirzebruch–Riemann–Roch theorem had been previously proved in the late nineteenth and early twentieth century. See Hirzebruch's monograph [104] for a description of these earlier results. One important fact which comes out in all of these theorems is that the right-hand side is *a priori* only a rational number, but the left-hand side is necessarily an integer, thus proving the integrality of the right-hand side, which is not at all easy to prove directly.

In 1963 Michael F. Atiyah (1929–) and Isadore M. Singer (1924–) announced their index theorem for elliptic operators on compact differentiable manifolds [11], and this result is now universally known as the *Atiyah–Singer index theorem*. A complete proof was published by Richard Palais in 1965 based on a seminar at the Institute for Advance Study [177], and Atiyah and Singer published a new and somewhat different proof in 1968 [12], which was followed up by a series of papers dealing with a number of different variations and generalization of this important theorem.

The fundamental result of Atiyah and Singer can be summarized as follows.

Theorem 14.5 (Atiyah–Singer index theorem) *Let E and F be differentiable complex vector bundles on a compact oriented differentiable manifold M of dimension m, and let D be a linear elliptic differential operator*

$$D : \mathcal{E}(E) \to \mathcal{E}(F),$$

where $\mathcal{E}(*)$ denotes C^∞ sections of the vector bundles. Then there is a cohomology class $\alpha \in H^m(X, \mathbf{Q})$ such that

$$\text{index}(D) := \dim \ker D - \dim\, \text{coker}\, D = \int_M \alpha.$$

Here $\text{coker}\, D = \mathcal{E}(F)/\text{im}\, \mathcal{E}(E)$. The cohomology class α in this theorem is defined explicitly by Atiyah and Singer in terms of the symbol of the differential operator (derived from the highest order derivative expression of D in local coordinates) as well as the Chern classes of the vector bundles E, F and $T(X)$.

A very important special case of the Atiyah–Singer index theorem is that the Hirzebruch–Riemann–Roch theorem, Theorem 14.4, is valid for arbitrary compact complex manifolds, not just for projective algebraic manifolds. With this result of Hirzebruch being generalized to arbitrary complex manifolds, we conclude our summary of Riemann–Roch theorems in the twentieth century.

Chapter 15
Noncompact Complex Manifolds

15.1 Introduction

In the previous chapter, we saw how Kodaira used a powerful combination of the tools from sheaf theory and the theory of harmonic differential forms to give a characterization of complex submanifolds of complex projective space. In this chapter we will discuss a similar characterization of complex submanifolds of complex Euclidean space.

The theory of several complex variables developed over the first half of the twentieth century, and a major theme concerned itself with numerous aspects of complex analysis that arose from Hartogs's discovery in 1906 that there are domains in $\mathbf{C}^n$, for $n > 1$, which admit simultaneous analytic continuation to larger domains. In particular, this included the characterizations of domains of holomorphy which are domains in $\mathbf{C}^n$ that do not admit simultaneous analytic continuation to a larger domain. In the following section we outline a brief history of this theory, which leads, in particular, to the notion of an abstract Stein manifold.

In 1951 Karl Stein (1913–2000) gave a definition of a complex manifold which was a generalization of a domain of holomorphy in $\mathbf{C}^n$. It turns out that a Stein manifold is precisely the notion of an abstract complex manifold which characterizes a closed submanifold of complex Euclidean space. This can be formulated as an embedding theorem for Stein manifolds, which was announced by Reinhold Remmert in 1957 and proved by Narasimhan and Bishop a couple of years later. We outline Bishop's proof of this theorem in Sects. 15.3–15.5.

One aspect of the theory of several complex variables was the formulation and eventual solution of the Levi problem. This problem is discussed in Sect. 15.2. It concerns itself with a characterization of domains of holomorphy (and generalizations thereof) in terms of a local differential-geometric condition on the boundary of such a domain (when there is a smooth boundary). In 1958 Hans Grauert gave a solution to the Levi problem in the context of a complex manifold setting, and was able to use this to prove that a real-analytic manifold admits a real-analytic embedding into

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2_15

real Euclidean space [86]. We outline the fundamental ideas in Grauert's paper in the final two sections of this chapter (Sects. 15.6 and 15.7).

15.2 Several Complex Variables

As we have mentioned earlier, Hartogs proved in 1906 [96] that there are no isolated singularities for holomorphic functions of n complex variables[1] for $n \geq 2$, but in fact he proved much more than this. Let's make this somewhat more precise. We consider a simple example of his theorem. Let

$$\Delta_0 = \{(z, w) \in \mathbf{C}^2 : \frac{1}{2} < |z| < 1, |w| < 1 \text{ or } |z| < 1, \frac{1}{2} < |w| < 1\},$$

and let

$$\Delta_1 = \{(z, w) : |z| < 1, |w| < 1\},$$

the unit bidisc of radius 1. We note that $\Delta_{\frac{1}{2}} := \Delta_1 - \overline{\Delta_0}$ is the open bidisc of radius $\frac{1}{2}$. Then Hartogs's theorem in this context says that if $f \in \mathcal{O}(\Delta_0)$, then there is a unique holomorphic function $\tilde{f} \in \mathcal{O}(\Delta_1)$ such that $\tilde{f}_{|\Delta_0} = f$. In other words, all holomorphic functions on Δ_0 admit simultaneous analytic continuation across the closed set $\overline{\Delta_{\frac{1}{2}}}$. The proof of this theorem uses the Cauchy integral formula in a very specific manner.

This type of simultaneous analytic continuation, which is only present in more than one complex dimension, is often referred to as a *Hartogs's phenomenon*.

A *domain of holomorphy* $D \subset \mathbf{C}^n$ is defined to be a domain such that there is no such simultaneous analytic continuation to a larger domain. In Sect. 9.6 we discussed Mittag-Leffler's generalization of the Weierstrass factorization theorem (a theorem on the complex plane $\mathbf{C}$) to a version which is valid for any domain $D \subset \mathbf{C}$. This shows immediately that any domain in the plane is a domain of holomorphy (see the discussion of this in Sect. 9.6). The example in the preceding paragraph due to Hartogs shows that not all domains in $\mathbf{C}^n, n > 1$, are domains of holomorphy. It became a major task in the first half of the twentieth century to characterize domains of holomorphy, both in terms of the geometry of the boundary of such a domain, and in terms of the behavior of holomorphic functions in the interior of the domain.

Let now D be a domain in $\mathbf{C}^n$. If K is a compact subset of D, then we define the *holomorphic hull* of K in D to be

$$\hat{K}_D := \{z \in D : |f(z)| \leq \sup_{\zeta \in K} |f(\zeta)|, \text{ for all } f \in \mathcal{O}(D)\}.$$

[1] Jacques Hadamard had remarked in his booklet *La Series de Taylor* [94], Chap. IX, that there are no isolated singularities of a holomorphic function of more than one variable using an analysis of power series near a point.

We say that a domain $D \subset \mathbf{C}^n$ is *holomorphically convex* if

$$K \text{ is compact in } D \Rightarrow \hat{K}_D \text{ is compact in } D. \tag{15.1}$$

This is a generalization of ordinary convexity, and it is not difficult to show that a convex set in $\mathbf{C}^n$ is indeed holomorphically convex. This yields immediately a rich family of holomorphically convex domains.

In 1932 Henri Cartan and Peter Thullen (1907–1996) proved an important result which linked the concept of domain of holomorphy with that of holomorphic convexity [222]:

Theorem 15.1 *A domain $D \subset \mathbf{C}^n$ is a domain of holomorphy if and only if it is holomorphically convex.*

In 1934 Heinrich Behnke (1898–1979) and Peter Thullen published an important monograph, *Theorie der Funktionen mehrerer komplexer Veränderlichen*[2] [13], which set the stage for the study of function theory of more than one complex variable for the next three decades. Behnke was based in Münster in the western part of Germany, and he created a school of several complex variables which played an important role in the development of this theory, culminating in major results in the 1950s, as we shall see later in this chapter. His students included Stein, Thullen, Grauert, Remmert, Hirzebruch and many others, all of whom played a major role in postwar mathematics in Germany.

In the 1930s there was a strong collaboration with the French school of several complex variables which was primarily led by Henri Cartan. This collaboration continued, almost unabated, after the second world war.

In 1929 Kiyoshi Oka (1901–1978) came to Paris to study several complex variables and returned to Japan three years later. In the course of two decades, Oka published ten fundamental papers solving a number of major problems, many of which had been formulated in the book by Behnke and Thullen. These papers all had the same title "Sur les fonctions analytiques de plusieurs variables",[3] with differing subtitles labeled I–X, e.g., the first paper had the subtitle "I. Domaines convexes par rapport aux fonctions rationnelles".[4] In 1961 Oka published a book [175] which contained the first nine of these papers. The tenth and final paper in the series appeared in 1962 [176].

This French–German collaboration before and after the second world war, along with the very important contributions made independently by Oka, formed the basis of what became the theory of several complex variables as it was understood at the beginning of the 1960s. There have been a number of books which appeared after 1960 which survey different aspects of the theory of several complex variables. We mention only four here for reference.

[2] *Theory of Functions of Several Complex Variables.*

[3] "Concerning analytic functions of several variables".

[4] "Rationally convex domains".

Gunning and Rossi's book *Analytic Functions of Several Complex Variables* [93] developed the theory of complex manifolds and complex spaces (which are locally subvarieties[5] of complex Euclidean space, which could have singularities). The treatment of singularities required the use of ring theory (including ideals, modules, etc.), and this needed to be extended to the sheaf-theoretic setting as well. The key work of Oka on coherent analytic sheaves was developed, and the culminating Theorems A and B of Cartan from his *Seminaire Cartan* from 1951–1952 were formulated and proved. This subsumed a good deal of the developments of the theory that had evolved since the book by Behnke and Thullen roughly thirty years earlier. We will introduce these theorems in Sect. 15.6, where they will be used in Grauert's solution of the Levi problem. We will discuss the Levi problem, which concerns a geometric characterization of domains of holomorphy, in the following paragraphs.

Lipman Bers published his very lucid lecture notes *Introduction to Several Complex Variables* in 1965 [17]. This covered much of the material in Gunning and Rossi, but it split the singularity theory and algebraic topology into a smaller aspect of the book and concentrated on the analysis on domains in $\mathbf{C}^n$ in the first part of the notes, making it somewhat easier for the reader to get the general picture.

The book by Lars Hörmander [112] brought a completely new point of view to the theory of several complex variables, showing how many of the problems that had been treated since the beginning of the theory at the time of Hartogs could be solved by solving suitable systems of over-determined partial differential equations, in particular, the inhomogeneous Cauchy–Riemann equations. His L^2-methods proved to be very powerful in this regard. However, they only worked at the time in the context of smooth complex manifolds.

Finally, the book by Stephen Krantz [130] brings an update of the theory and includes more material concerning various generalizations of the classical Cauchy integral formula from the nineteenth century, which were able to play important roles in solving a number of different problems.

We will refer to these references as we proceed with our study of the embedding theorems for noncompact complex manifolds and real-analytic manifolds in this chapter.

In 1910 Eugenio Ella Levi (1883–1917) discovered a very interesting local geometric criterion for the boundary of a domain of holomorphy in the case where the domain has a smooth boundary [143]. Suppose that $\varphi(z)$ is a real-valued differentiable function defined on a domain $U \subset \mathbf{C}^n$, then we say that φ is *plurisubharmonic* if the Hermitian symmetric matrix

$$H(\varphi) := \left(\frac{\partial^2 \varphi}{\partial z_i \partial \overline{z}_j}\right) \tag{15.2}$$

[5] A *subvariety* V of a complex manifold X is a closed subset $V \subset X$ which is locally defined near any point $x \in V$ to be the zero set of a finite number of holomorphic functions defined in a neighborhood of the point x.

is positive-semidefinite on U and is *strictly plurisubharmonic* on U if $H(\varphi)$ is positive-definite on U.

Let us now consider a domain $D \subset \mathbf{C}^n$ with a smooth boundary. Let $p \in \partial D$, where ∂D denotes the boundary of D. We say that the boundary of D is *pseudoconvex* (*strongly pseudoconvex*) near p if there is a plurisubharmonic function (strictly plurisubharmonic function) φ defined in a neighborhood U of p, and

$$D \cap U = \{z \in U : \varphi(z) < 0\}.$$

Levi showed in [143] that a domain of holomorphy with a smooth boundary must be pseudoconvex near each of its boundary points. More precisely, if $f \in \mathcal{O}(D \cap U)$, then f cannot be analytically continued beyond the point p. He also showed that

Lemma 15.1 Levi [143] *For any point p of the boundary of a strongly pseudoconvex domain D, there is a neighborhood U of p which is a domain of holomorphy and such that $U \cap D$ is also a domain of holomorphy.*

The *Levi problem* became the problem of showing that a pseudoconvex domain is itself a domain of holomorphy. The Levi problem became an important unsolved problem for several decades until it was resolved for domains in $\mathbf{C}^2$ by Oka in 1942 [173] and in $\mathbf{C}^n$ in 1953 by Oka [174] and by Hans Bremermann [26] and Francois Norguet [172] in 1954. It was successfully resolved in the general context of a complex manifold by Hans Grauert (1930–2011) in 1958 [86]. We will discuss this further in Sect. 15.7, as the solution to the Levi problem is a key ingredient in Grauert's proof of the real-analytic embedding theorem.

15.3 Stein Manifolds

In Chap. 14 we discussed Kodaira's embedding theorem for compact complex manifolds. Now we turn to an analogous theorem for noncompact complex manifolds which we call the Stein manifold embedding theorem. This theorem characterizes all complex manifolds which can be embedded as closed complex submanifolds of $\mathbf{C}^N$ for $N \geq 1$. We need to formulate first several concepts that are the basis for this characterization.

Let X be an n-dimensional complex manifold (as always, with a countable basis), and let $\mathcal{O}(X)$ be the ring of holomorphic functions on X. We can equip $\mathcal{O}(X)$ with the topology of uniform convergence on compact subsets of X, which makes $\mathcal{O}(X)$ into a Frchet space, and, moreover, $\mathcal{O}(X)$ is a complete metric space with this topology.

We say that X is *holomorphically separable* if, for any two distinct points x_1 and x_2 of X, there exists a function $f \in \mathcal{O}(X)$ such that $f(x_1) \neq f(x_2)$.[6] We say that X has *global local coordinates* if, for any point $x \in X$, there are n functions $f_1, \ldots, f_n \in$

[6]This is often expressed as saying that the globally defined holomorphic functions on X separate points on X.

$\mathcal{O}(X)$ such that the mapping $F = (f_1, \ldots, f_n)$ has rank n at x (these n functions, globally defined, could be used as a local coordinate system in a neighborhood of x). Finally, we say that X is *holomorphically convex* if, for any compact subset $K \subset X$, the *holomorphic hull* of K,

$$\hat{K}_X := \{x \in X : |f(x)| \leq \sup_{y \in K} |f(y|, \text{ for all } f \in \mathcal{O}(X)\} \tag{15.3}$$

is compact in X.

A complex manifold X is said to be a *Stein manifold* if it is holomorphically separable, has global local coordinates, and is holomorphically convex. This concept of a complex manifold satisfying these axioms was first formulated by Stein in 1951 [220], and Henri Cartan coined the name " Stein manifold" in his lectures of 1952 [34]. We can now formulate the *Stein manifold proper embedding theorem*.

Theorem 15.2 *Let X be a Stein manifold of dimension n, then there exists a proper holomorphic embedding*

$$F : X \to \mathbf{C}^{2n+1}.$$

Since the mapping is proper, the image of F in this theorem is a closed complex submanifold.

It is not difficult to show that any closed complex submanifold of complex Euclidean space is, indeed, a Stein manifold, and we will show this in the following paragraphs. Thus this theorem provides a characterization of which complex manifolds can be holomorphically embedded in complex Euclidean space as closed submanifolds.

First, we note that complex Euclidean space is clearly holomorphically convex. Namely, if K is compact in $\mathbf{C}^N$, then, letting $(z_1, \ldots, z_N)$ be coordinates on $\mathbf{C}^N$, we can set

$$M_j = \sup_K |z_j|,$$

and it follows that

$$\hat{K}_{\mathbf{C}^N} \subset \{(z_1, \ldots, z_N) : |z_j| \leq M_j, \ j = 1, \ldots N\},$$

which is compact.

Now, let X be a complex submanifold of $\mathbf{C}^N$, and suppose that K is a compact subset of X, then K has a holomorphic hull $\hat{K}_{\mathbf{C}^N}$ which is compact in $\mathbf{C}^N$, and hence its restriction to X is compact. But the holomorphic functions on $\mathbf{C}^N$ restricted to X is a subset of the holomorphic functions on X. It follows that

$$\hat{K}_{\mathbf{C}^N} \supset \hat{K}_X,$$

and hence $\hat{K}_X$ is compact in X. Moreover, X is clearly holomorphically separable and has global local coordinates, and hence it is a Stein manifold.

The Stein manifold proper embedding theorem was announced by Reinhold Remmert (1930–2016) in 1956 [193] with some indications of the nature of the proof (we note that he used the term "holomorphically complete" to describe a Stein manifold at that time). However, he never published a proof. Raghavan Narasimhan (1937–2015) in 1960 [163] and Errett Bishop (1928–1983) in 1961 [20] gave independent proofs a few years later. Both of these authors proved a more general result for complex spaces, which requires the proofs to be somewhat more technical. Hörmander's monograph on several complex variables from 1966 [112] gave a simplified proof (using the ideas of Bishop) in the case of complex manifolds, which is how we have formulated the theorem here. In our exposition we follow Hörmander's formulation of Bishop's proof.

The proof of this embedding theorem splits into two parts. The first part shows that there is an abundance of holomorphic embeddings of a complex manifold X which is holomorphically separable and has global local coordinates into $\mathbf{C}^{2n+1}$. Here we use some of the basic ideas that Whitney used in his proof of the differentiable embedding theorem (Chap. 12). This we discuss in the following Sect. 15.4.

To prove that there are proper holomorphic embeddings is considerably more difficult, and we outline the proof in Sect. 15.5. This proof uses basic ideas concerning analytic polyhedra as formulated and used by Bishop [20]. We cite here an important approximation theorem originally formulated and proved in Stein's paper [220] and which plays an important role in Bishop's proof of the Stein manifold proper embedding theorem. It is a generalization to this context of the classical Runge approximation theorem from the nineteenth century that we discussed in Sect. 9.6.

Theorem 15.3 *Let K be a compact subset of a Stein manifold X such that $\hat{K}_X = K$, then if f is a holomorphic function defined in a neighborhood of K, then f can be approximated uniformly on K by functions holomorphic on X.*

We will often say that a compact set $K \subset X$ is *holomorphically convex* if it satisfies the condition $\hat{K}_X = K$, as is hypothesized in this theorem.

15.4 Generic Embeddings for a Class of Complex Manifolds

Suppose that D is a domain in $\mathbf{C}^n$. Then clearly D is holomorphically separable and has global local coordinates. The following theorem shows that any complex manifold with these properties can be embedded into a complex Euclidean space (not necessarily a proper embedding, however). Here embedding means simply a regular one-to-one mapping, and we phrase it as such in this theorem.

Theorem 15.4 *Let X be an n-dimensional complex manifold which is holomorphically separable and which has global local coordinates, then the set of one-to-one regular mappings*

$$F : X \to \mathbf{C}^{2n+1}$$

is dense in $\mathcal{O}(X)^{2n+1}$.

As we said above, the hypotheses of this theorem are satisfied by any domain D in $\mathbf{C}^n$. In particular, there is no assumption of holomorphic convexity, as in the case of Stein manifolds. But the image of a one-to-one regular mapping of such a domain D need not be a closed submanifold of $\mathbf{C}^{2n+1}$. For instance, if Δ is the unit disc in $\mathbf{C}$, then the mapping

$$z \mapsto (z, 0, 0)$$

is a one-to-one regular mapping of Δ into $\mathbf{C}^3$, and its image (a copy of the unit disc) is clearly not a closed subset of $\mathbf{C}^3$.

The proof of this theorem can be summarized in two lemmas concerning the behavior of the mappings on compact subsets, which we discuss first, and then these are applied to obtain a proof of the theorem.

Lemma 15.2 *Let K be a compact subset of X, then there is an integer N and a mapping*

$$F : X \to \mathbf{C}^N, \tag{15.4}$$

which is one-to-one and regular on K.

We outline the proof of this lemma. Since X has global local coordinates, there is a finite covering of K by open sets U_α with functions $f_1^\alpha, \ldots, f_n^\alpha \in \mathcal{O}(X)$ such that the mappings $f^\alpha = (f_1^\alpha, \ldots, f_n^\alpha)$ are regular on U_α. Using all of these functions yields a finite set of functions $f_j \in \mathcal{O}(X)$, $j = 1, \ldots, M$, such that the mapping $(f_1, \ldots, f_M)$ is regular at each point of K. By this regularity, the mapping is one-to-one in the neighborhood of each point of K. Thus, there is a neighborhood W of the diagonal of $K \times K$ in $X \times X$ such that if $(x_1, x_2) \in W$, then $f_j(x_1) = f_j(x_2)$, $j = 1, \ldots, M$, implies that $x_1 = x_2$. By using the property of holomorphic separation we can find a finite number of functions $f_j \in \mathcal{O}(X)$, $j = M + 1, \ldots N$, such that if $(x_1, x_2) \in K \times K - W$, then $f_j(x_1) = f(x_2)$, $j = M + 1, \ldots, N$ implies that $x_1 = x_2$. Then the mapping

$$F = (f_1, \ldots, f_N) : X \to \mathbf{C}^N$$

will be a mapping which is one-to-one and regular on K.

This lemma insures that we have a suitable mapping on K into a possibly high-dimensional Euclidean space $\mathbf{C}^N$. The next lemma shows that we can successively perturb the mapping F in (15.4) to mappings into lower-dimensional spaces until we reach the desired embedding dimension of $2n + 1$.

Lemma 15.3 *Suppose that $N > 2n + 1$, and suppose that*

$$F = (f_1, \ldots, f_{N+1}) : X \to \mathbf{C}^{N+1}$$

is regular and one-to-one on a compact set $K \subset X$, then there is a neighborhood U of $0 \in \mathbf{C}^N$ such that for almost all $a \in U$, the perturbed mapping

$$\tilde{F} = (f_1 - a_1 f_{N+1}, f_2 - a_2 f_{N+1}, \dots, f_N - a_N f_{N+1}) \tag{15.5}$$

is regular and one-to-one on K.

We recall one of the key ingredients in the perturbations involved in the proof of the Whitney embedding theorem. Namely, if we have a smooth mapping f (C^1 is sufficient), defined on an open subset U of $\mathbf{R}^m$ which maps to a higher-dimensional Euclidean space $\mathbf{R}^N$, where $N > m$, then the image $f(U)$ has measure zero in $\mathbf{R}^N$ (see the discussion in Sect. 12.1). We note that a holomorphic mapping of the same type, namely,

$$f : U \subset \mathbf{C}^m \to \mathbf{C}^N,$$

where $N > m$, is a special case of the real-analysis result for C^1-mappings, and hence in this case we also have

$$\text{meas}\, f(U) = 0.$$

To prove Lemma 15.3 we need to successively choose two perturbations a in a neighborhood of $0 \in \mathbf{C}^N$. First we will need a perturbation to insure that $\tilde{F}$ is regular on K.

We assume, for simplicity, that K is contained in a coordinate neighborhood, and we can compute the Jacobian matrix of $\tilde{F}$ at a point $x \in K$,

$$\frac{\partial f_j}{\partial x_k} - a_j \frac{\partial f_{N+1}}{\partial x_k}, \ k = 1, \dots, n, \ j = 1, \dots, N. \tag{15.6}$$

We consider this as n vectors in $\mathbf{C}^N$, and we want to show that for appropriate a these are linearly independent. To see this we consider the sum, for $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbf{C}^n$,

$$\sum_{k=1}^{n} \lambda_k \left(\frac{\partial f_j}{\partial x_k} - a_j \frac{\partial f_{N+1}}{\partial x_k} \right) = 0, \tag{15.7}$$

and we need to show that, for suitable a, $\lambda = 0$.

We now set $a_{N+1} = 1$, and set

$$\mu = \sum_k \lambda_k \frac{\partial f_{N+1}}{\partial z_k},$$

then (15.7) becomes

$$\sum_{k=1}^{n} \lambda_k \frac{\partial f_j}{\partial z_k} = \mu a_j, \ \ j = 1, \dots, N+1. \tag{15.8}$$

The matrix

$$\frac{\partial f_j}{\partial z_k}, \ j = 1, \ldots, N+1, k = 1, \ldots, n,$$

has maximal rank, by hypothesis. Consider now the mapping

$$\Lambda : \mathbf{C}^n \times K \to \mathbf{C}^{N+1},$$

given by

$$\Lambda(\lambda, x) = \left\{ \sum_k \lambda_k \frac{\partial f_j}{\partial z_k} \right\}_{j=1}^{N+1} \in \mathbf{C}^{N+1}.$$

We see here that the domain of Λ has dimension $2n$, and the dimension of the image space is $N + 1 > 2n + 2$ and hence

$$\text{meas}\,(\Lambda(\mathbf{C}^n \times K)) = 0.$$

Since the matrix in the definition of the mapping Λ has, for fixed $x \in K$, maximal rank, then if $\lambda = (\lambda_1, \ldots \lambda_n) \neq 0$, in (15.8), then $\Lambda(\lambda, x) \neq 0$ in $\mathbf{C}^{N+1}$. Thus we have

$$\Lambda(\lambda, x) = \mu(a, 1) \neq 0,$$

and hence $\mu \neq 0$. By rescaling, and letting $\tilde{\lambda} = \frac{1}{\mu}\lambda$, we have

$$\Lambda(\tilde{\lambda}, x) = (a, 1) \in \Lambda(\mathbf{C}^n \times K).$$

But we know that $\text{meas}\,(\Lambda(\mathbf{C}^n \times K)) = 0$, so we can choose an arbitrarily small $a \in \mathbf{C}^N$ such that

$$(a, 1) \notin \Lambda(\mathbf{C}^n \times K).$$

If $(a, 1)$ is chosen in this manner, then it follows that $\lambda = 0$, which proves the desired linear independence of the vectors in the perturbed Jacobian matrix in (15.6), as desired.

To show that there is a choice of an arbitrarily small perturbation parameter $a \in \mathbf{C}^n$ so that $\tilde{F}$ is one-to-one, we proceed in a similar manner. We have, by hypothesis, that: if $x_1, x_2 \in K$, and

$$f_j(x_1) - f_j(x_2) = 0, \ j = 1, \ldots, N+1,$$

then $x_1 = x_2$. Suppose now that

$$f_j(x_1) - a_j f_{N+1}(x_1) - (f_j(x_2) - a_j f_{N+1}(x_2)) = 0, \ j = 1, \ldots, N,$$

for some small $a \in \mathbf{C}^n$, then these differences can be rewritten as ($a_{N+1} = 1$, as before):

$$f_j(x_1) - f_j(x_2) = a_j\lambda, \ \ j = 1, \dots, N+1,$$

where we have set $\lambda = f_{N+1}(x_1) - f_{N+1}(x_2)$. If $\lambda = 0$ for our choice of a, then it follows from the one-to-one nature of the mapping F, that $x_1 = x_2$. Suppose that $\lambda \neq 0$, then, letting $\mu = 1/\lambda$, we can define the mapping

$$\tilde{\Lambda} : \mathbf{C} \times K \times K \to \mathbf{C}^{N+1}$$

by

$$\tilde{\Lambda}(\mu, x_1, x_2) = \mu(f_1(x_1) - f_1(x_2), \dots, f_{N+1}(x_1) - f_{N+1}(x_2)).$$

Now we note, as we did in the previous case, that

$$\text{meas}\,(\tilde{\Lambda}(\mathbf{C} \times K \times K)) = 0,$$

since $2n + 1 < N + 1$. By rescaling as we did before, we can now choose small $a \in \mathbf{C}^n$ such that $(a, 1)$ is not in the range of $\tilde{\Lambda}$. Thus, for this choice of a, we must have $\lambda = 0$, and it follows that the perturbed mapping $\tilde{F}$ is one-to-one.

By choosing a so that $(a, 1)$ is not in the range of either Λ or $\tilde{\Lambda}$, then we can conclude that the perturbed mapping $\tilde{F}$ is both regular and one-to-one on K, and this concludes the outline of the proof of Lemma 15.3.

Using these two lemmas, it will be relatively easy to show why Theorem 15.4 is valid. We will use the language of Baire category theory for a metric space for this. Recall that the vector spaces $\mathcal{O}(X)$ or $\mathcal{O}(X)^N$, for some $N \geq 1$, are all Fréchet spaces, and, as such, they are complete metric spaces with the corresponding metric coming from the Fréchet structure. A subset of a complete metric space is said to be of *first category* if it is the union of closed subsets, each of which has no interior points, and a subset is said to be of *second category* if it is the complement of a set of first category. It follows that a set in a complete metric space of second category is dense in the metric space.

We now assume that $N \geq 2n + 1$, as hypothesized in Theorem 15.4. We let K be a compact set in X, and we let

$$S_K = \{F \in \mathcal{O}(X)^N : F \text{ is not regular or one-to-one on } K\}.$$

The complement of S_K, which we denote by T_K, is the set of mappings which are regular and one-to-one on K. First, we need to know that T_K is nonempty. By Lemma 15.2 there is a mapping $G = (g_1, \dots, g_M) \in \mathcal{O}(X)^M$, for some M, such that G is regular and one-to-one on K. Necessarily, $M \geq N$, but it might be considerably larger. If $M > N$, we can use Lemma 15.3 to successively find (by a sequence of perturbations of F each of which lowers the embedding dimension) a suitable mapping $\tilde{F} \in \mathcal{O}(X)^N$ which is regular and one-to-one on K. If $M < N$, then the mapping

$$\tilde{F} := (g_1, \dots, g_M, 0, \dots, 0) \in \mathcal{O}(X)^N$$

will be regular and one-to-one, as well, where we have added $N - M$ zeros to the mapping. Thus T_K is nonempty. It is easy to see using continuity considerations that the set T_K is open. Thus, the set S_K is closed. Now we need to show that S_K has no interior points.

We will show that in the neighborhood of any $F \in \mathcal{O}(X)^N$ (in particular, any $F \in S_K$), there is a mapping $\tilde{F}$ which is regular and one-to-one at all points of K. By Lemma 15.2 there is a mapping $G = (g_1, \ldots, g_M)$, which we used above, such that G is regular and one-to-one on K. Let now $F \in \mathcal{O}(X)^N$ be any mapping from X to $\mathbf{C}^N$. Consider the mapping

$$H := (f_1, \ldots, f_N, g_1, \ldots, g_M).$$

By successively applying Lemma 15.3 to H we can find a perturbation of F of the form

$$\tilde{f}_j = f_j + \sum_k a_{jk} g_k, \ \ j = 1, \ldots, N,$$

for suitably small $a_{jk} \in \mathbf{C}$, such that $\tilde{F} = (\tilde{f}_1, \ldots, \tilde{f}_N)$ is regular and one-to-one on K. Thus S_K has no interior points.

We can now find a countable set of compact sets K_α which cover X. Then, let T be the set of mappings in $f \in \mathcal{O}(X)^N$ such that F is regular and one-to-one on X. We see that, if we let S be the complement of T, then

$$S = \bigcup_\alpha S_{K_\alpha},$$

and hence is of first category, and thus T is of second category and is a dense subset of $\mathcal{O}(X)^N$, as desired.

Remark We note here that the ideas used in the proofs of this theorem and the two lemmas use the perturbation techniques that were used as part of the proof of Whitney's embedding theorem as in Sect. 12.3. In fact, this outline given here would have proved, by using differentiable functions on an n-dimensional manifold X, that the set of mappings in $\mathcal{E}(X)^{2n+1}$ mapping X to $\mathbf{R}^{2n+1}$ which are regular and one-to-one on X is a dense set. This proof of this fact using category-theoretic arguments would not have used cut-off functions as we did in Sect. 12.3, where we followed closely the steps outlined in the original paper of Whitney [245]. The final step in Whitney's proof was to find a proper mapping $\varphi(x)$ of X into $\mathbf{R}$, and this did use cut-off functions. It is this step, where we can't use cut-off functions, which is much more difficult in this holomorphic setting, and we will see that in much greater detail in the next section.

15.5 A Proper Embedding Theorem for Stein Manifolds

In this section we will see how holomorphic convexity of a Stein manifold helps to give a proper holomorphic embedding into complex Euclidean space. We've already seen in the previous section how the two hypotheses of holomorphic separability and having global local coordinates give a generic embedding.

Let now X be an n-dimensional complex manifold. Suppose that $f_1, \dots, f_N$ are N holomorphic functions on X. We define an *analytic polyhedron P of order N* on X to be a union of components of the open set

$$P_0 := \{x \in X : |f_j(x)| < 1, \; j = 1, \dots, N\}. \tag{15.9}$$

This concept generalizes to a manifold the notion of a polydisc in $\mathbf{C}^N$ defined by

$$P = \{z \in \mathbf{C}^N : |z_j| < 1, \; j = 1, \dots, N\},$$

where the coordinate functions z_j play the role of the f_j in the definition of an analytic polyhedron.[7]

We now formulate two lemmas which play a critical role in the proof of the proper embedding theorem. We assume for these lemmas that X is now a Stein manifold (of dimension n). The first one uses the holomorphic convexity explicitly and has a fairly straightforward proof.

Lemma 15.4 *Let K be a compact subset of X such that $K = \hat{K}_X$, then, for any neighborhood U of K in X, there exists an analytic polyhedron P such that*

$$K \subset P \subset U.$$

To see this, assume that U is relatively compact[8] in X. For each $x \in \partial U$, since $K = \hat{K}_X$, we can find a holomorphic function $f \in \mathcal{O}(X)$ such that $|f| < 1$ on K and $|f(x) > 1$. By the compactness of ∂U, there are a finite number of holomorphic functions $f_j \in \mathcal{O}(X)$, such that

$$P := \{z \in U : |f_j(x)| < 1\}$$

[7]An analytic polyhedron is often defined to be simply a set of the form P_0 in (15.9), but this slightly more subtle notion we are using was introduced by Bishop [20], and it allows different open components of the set P_0 to play a role in the proof of a theorem, as is the case in Bishop's proof of the proper embedding theorem, which we are following here. Consider the example of X being a two-sheeted covering of the punctured complex plane defined by $w = \sqrt{z}$, and let $\pi : X \to \mathbf{C} - \{0\}$ be the covering mapping, then if $z_0 \in \mathbf{C}$ satisfies, for instance, $|z_0| > 2$, then letting $f(x) := \pi(x) - z_0$, we see that $P_0 = \{x \in X : |f(x)) < 1\}$ consists of two copies of the open unit disc on the two sheets of X centered at the two points $\pi^{-1}(z_0)$, so an analytic polyhedron in this case could be one or both of these components.

[8]We recall that a set U is *relatively compact* in X if the closure $\overline{U}$ is compact in X. We will denote this by $U \subset\subset X$.

is an analytic polyhedron with $K \subset P \subset U$, as desired.

The second lemma, due to Bishop [20], is quite a bit more difficult, and it is the key to the proper embedding proof. Namely, we can reduce by a perturbation the order of a given analytic polyhedron if its order is greater than $2n$.

Lemma 15.5 *Let K be compact in X, and let P be an analytic polyhedron of order $N+1$ in X with $K \subset P$. If $N \geq 2n$, then there exists an analytic polyhedron P' of order N such that*

$$K \subset P' \subset P \subset X.$$

We will outline the construction of P' here. Suppose that P is defined by

$$P = \{x \in X : |f_j(x)| < 1, \text{ for } f_j \in \mathcal{O}(X), \ j = 1, \dots, N+1\}.$$

Let

$$c_0 < c_1 < c_2 < c_3 < 1 \tag{15.10}$$

be such that

$$|f_j(x)| < c_0, \ x \in K, \ j = 1, \dots, N+1.$$

We now want to find suitable approximations f_j' of the functions f_j. Using Lemma 15.3, and setting $f_{N+1}' = f_{N+1}$, we can find approximations $f_j'(x)$ so that

$$\begin{aligned}
&(f_1'/f_{N+1}, \dots f_N'/f_{N+1}) \text{ has rank } n \text{ on } \{x \in \overline{P} : |f_{N+1}(x)| \geq c_2\},\\
&|f_j'(x)| < c_0 \text{ in } K, \ j = 1, \dots, N,\\
&U := \{x \in P : |f_j'(x)| < c_3, j = 1, \dots, N+1\} \subset\subset P.
\end{aligned}$$

Here we have used the hypothesis that $N \geq 2n$.

So far we have utilized three of the intermediate constants in (15.10). The next step in our construction will use the remaining constant c_1. We set

$$\Delta_\nu := \{x \in X : |f_j'(x)^\nu - f_{N+1}(x)^\nu\} < c_1^\nu, \ j = 1, \dots, N\},$$

and let P_ν' be the union of the components of Δ_c which intersect K. We can then take our desired P' to be P_ν' for some sufficiently large ν. The proof that such a P' satisfies the conclusion of the lemma involves a delicate sequence of estimates which will be omitted here (see Hörmander [111], pp. 122–124).

We can now outline the proof of Theorem 15.2, using these lemmas on analytic polyhedra. Let X be an n-dimensional Stein manifold, and let

$$g : X \to \mathbf{C}^{2n+1}$$

be a one-to-one regular mapping (an embedding, but not necessarily proper) given by Theorem 15.4. Then suppose we can find a holomorphic mapping

$$f : X \to \mathbf{C}^{2n+1}$$

such that

$$\{x \in X : |f(x)| \leq k + |g(x)|\} \subset\subset X, \tag{15.11}$$

for all positive integers k (k will denote positive integers in the following arguments as well). Here we have set

$$|f(x)| = \max_j |f_j(x)|,$$

and similarly for $|g(x)|$. It will then follow that a suitable perturbation $\tilde{f}$ of the mapping f will be a proper holomorphic embedding of X into $\mathbf{C}^{2n+1}$.

Let us show how to find such a perturbation and that it provides a proper embedding, and then we will return to the question of the existence of a mapping f that satisfies (15.11). Namely, given the mappings f and g, we can combine them to form the mapping

$$(f, g) : X \to \mathbf{C}^{2(2n+1)},$$

and this is an embedding, since g is. By Lemma 15.3 we can find constants a_{jl} sufficiently small so that the perturbation of f given by

$$\tilde{f}_j(x) = f_j(x) + \sum_{l=1}^{2n+1} a_{jl} g_l$$

is an embedding. Supposing that $|a_{jl}| \leq 1$, then

$$\{x \in X : |\tilde{f}(x)| \leq k\} \subset \{x \in X : |f(x)| \leq k + |g(x)|\} \subset\subset X,$$

for all k, and it follows that $\tilde{f}$ is a proper holomorphic embedding of X into $\mathbf{C}^{2n+1}$, as desired.

To construct a mapping f which satisfies (15.11), we utilize an exhaustion of X by analytic polyhedra. By the holomorphic convexity of X it follows that there exists a sequence of compact sets K_k with K_k contained in the interior of K_{k+1} and satisfying $\hat{K}_k = K_k$ and $X = \bigcup_k K_k$. Now, using Lemma 15.5, we can find analytic polyhedra P_k of order $2n$ such that

$$K_k \subset P_k \subset K_{k+1},$$

and we can define the constants M_k by

$$M_k := \sup_{P_k} |g|.$$

If we can find a mapping $f : X \to \mathbf{C}^{2n+1}$ such that

$$|f(x)| \geq k + M_{k+1} \text{ in } P_{k+1} - P_k, \text{ for all } k, \tag{15.12}$$

then f will satisfy our condition (15.11). Namely, (15.12) implies that

$$|f(x)| \geq k + |g(x)| \text{ for } x \in P_{k+1} - P_k,$$

and hence

$$|f(x)| \geq k + |g(x)| \text{ for } x \in \bigcup_{l=k}^{\infty}(P_{l+1} - P_l) = X - P_k.$$

But this then implies that

$$\{x \in X : |f(x)| \leq k + |g(x)|\}$$

is necessarily a subset of P_k, which is relatively compact in X, and thus condition (15.11) is satisfied.

To find a mapping $f = (f_1, \ldots, f_{2n}, f_{2n+1})$ that satisfies (15.12), we proceed in two steps: first we find $(f_1, \ldots, f_{2n})$ that satisfy the inequality in (15.12) on the boundary of P_k (and some of the interior points of $P_{k+1} - P_k$), and then we construct a final function f_{2n+1} which satisfies this inequality at the remaining points of the set $P_{k+1} - P_k$ in (15.12). The first step uses specifically the properties of the analytic polyhedra P_k, and the second step uses the holomorphic approximation theorem for holomorphically convex compact sets in X (Theorem 15.3). In both cases the functions f_j are defined as the sum of infinite series of terms which are defined inductively in terms of the analytic polyhedra P_k.

We start with the construction of the functions $f_1, \ldots, f_{2n}$. Let

$$h_1^k(x), \ldots, h_{2n}^k(x)$$

be holomorphic functions on X defining the analytic polyhedron of order $2n$,

$$P_k = \{x \in X : |h_j^k(x)| < 1\}.$$

Thus we have

$$\max_{1 \leq j \leq 2n} |h_j^k| < 1 \text{ in } \overline{P}_{k-1},$$

and

$$\max_{1 \leq j \leq 2n} |h_j^k| = 1 \text{ on } \partial P_k.$$

We now let

$$f(x)_j^k = (a_k h(x)_j^k)^{m_k},$$

where the a_k are constants slightly larger than 1, and the m_k are integers sufficiently large so that

$$\max_{1\leq j\leq 2n} |f_j^k| \leq 2^{-k} \text{ in } P_{k-1},$$

and

$$\max_{1\leq j\leq 2n} |f_j^k| > M_{k+1} + k + 1 + \max_{1\leq j\leq 2n} |\sum_{l=1}^{k-1} f_j^l| \text{ on } \partial P_k.$$

Then it is easy to verify that the sum

$$f_j := \sum_{k=1}^{\infty} f_j^k, \ \ j = 1, \ldots, 2n$$

will converge on X, and will satisfy[9]

$$\max_{1\leq j 2n} |f_j(x)| > k + M_{k+1} \text{ on } \partial P_k. \tag{15.13}$$

This estimate for $f_1, \ldots, f_{2n}$ is part of the estimate we need in (15.12) (namely on a neighborhood of ∂P_k), but not on all of $P_{k+1} - P_k$. We need to construct a function f_{2n+1} which will provide the needed lower bound on the remainder of $P_{k+1} - P_k$. Thus we set

$$G_k = \{x \in P_{k+1} - P_k : \max_{1\leq j\leq 2n} |f_j(x)| \leq k + M_{k+1}\}.$$

This is the set of points where we need f_{2n+1} to be sufficiently large. We also define

$$H_k = \{x \in P_k : \max_{1\leq j\leq 2n} |f_j(x)| \leq k + M_{k+1}\}.$$

It is clear from (15.13) that these two compact sets G_k and H_k are disjoint. The holomorphic hull of the set $G_k \cup H_k$ is a compact subset of K_{k+2} and it has the form

$$G_k \cup H_k \cup H_k',$$

where

$$H_k' \subset X - P_{k+1}.$$

Here we note that each of the sets G_k and H_k are the closures of polynomial polyhedra, and hence are each holomorphically convex themselves. Also one can prove easily that $H' = \emptyset$ by comparing the values of the functions h_j^k on $G_k \cup H_k$ and on any possible point of H_k'. It now follows that if we define the locally constant function

---

[9]Here one uses the classical inequality $|A + B| \geq ||A| - |B||$, for complex numbers A and B.

$$\varphi := \begin{cases} 1 \text{ on } G_k, \\ 0 \text{ on } H_k, \end{cases}$$

which is a holomorphic function defined on a neighborhood of $G_k \cup H_k$, then we can approximate φ by a holomorphic function on X, using Theorem 15.3. We choose such holomorphic approximations $h_k \in \mathcal{O}(X)$ successively, such that

$$\begin{aligned} &|h_k(x)| < 2^{-k} \text{ on } H_k, \\ &h_k \geq 1 + M_k + k + 1 + |\textstyle\sum_{l=1}^{k-1} h_l(x)| \text{ on } G_k. \end{aligned}$$

Then, just as before, we define the limit

$$\sum_{k=0}^{\infty} h_k(x),$$

which is a well-defined holomorphic function on X satisfying

$$|f_{2n+1} \geq k + M_{k+1}, \text{ for } x \in G_k.$$

Thus the mapping $(f_1, \ldots, f_{2n+1})$ satisfies (15.12), and this concludes our outline of the proper holomorphic embedding theorem for Stein manifolds.

Remark This was the status in 1961, when this theorem was first proved by Narasimhan [163] and Bishop [20]. The question remained: could the embedding dimension $2n + 1$ be lowered? Otto Forster showed in 1970 [73] that the embedding dimension must be at least $\left[\frac{3n}{2}\right] + 1$, and he was able to lower the embedding dimension to $\left[\frac{5n}{2}\right] + 1$; and in 1992, Eliashberg and Gromov showed that [60], in fact, the embedding dimension could be lowered to the best possible $\left[\frac{3n}{2}\right] + 1$.

15.6 Grauert's Solution to the Levi Problem

Let X be a complex manifold, and suppose that $D \subset\subset X$ is a domain with a strongly pseudoconvex boundary. More precisely, we assume that there is a neighborhood W of ∂D and a strictly plurisubharmonic function $\varphi \in C^\infty(W)$ such that

$$D \cap W = \{x \in W : \varphi(x) < 0\}.$$

We formulate the Levi problem in this context to ask if such a domain D is holomorphically convex. We note that to ask if D were Stein would be asking too much as the following example shows.

Let B be in the unit ball in $X = \mathbf{C}^n$, $n \geq 2$, and let $\tilde{X} = Q_0 X$ be the quadratic transform of X at the origin, and let $\tilde{B} = Q_0 B$ be the corresponding quadratic transform of the domain B. It is clear then that $\tilde{B}$ is a strongly pseudoconvex domain in $\tilde{X}$ (the boundary of the ball or its quadratic transform hasn't been modified at all). However, $\tilde{B}$ is not Stein, since it is not holomorphically separable, as all holomorphic functions on $\tilde{B}$ are constant on

$$\pi^{-1}(0) \cong \mathbf{P}_{n-1},$$

where

$$\pi : Q_0X \to X$$

is the projection mapping of the quadratic transform. However, $\tilde{B}$ is holomorphically convex, as one can check directly.

That this example is holomorphically convex is also a consequence of the following quite general theorem of Grauert, which is a solution of the Levi problem in this complex manifold context.

Theorem 15.5 (Grauert [86]) *Let D be a relatively compact strongly pseudoconvex domain in a complex manifold X, then D is holomorphically convex.*

Grauert's proof of this theorem consists of two fundamental steps. The first step we formulate as a lemma (although it has the strong merit of being a theorem in its own right).

Lemma 15.6 *The cohomology groups*

$$H^q(D, \mathcal{O}), \ q \geq 1,$$

are all finite-dimensional.

The second step (which depends on the first) is to show that for each boundary point, there is a holomorphic function which doesn't analytically continue beyond that point. More precisely we have a second lemma.

Lemma 15.7 *For each point $p \in \partial D$, there exists a holomorphic function $f \in \mathcal{O}(D)$ such that*

$$\lim_{x \to p} |f(x| = \infty.$$

It is clear from this lemma that D must be holomorphically convex, and that concludes the proof of the theorem, assuming these two lemmas.

We now outline the proofs of these two lemmas. The proof of Lemma 15.6 is a generalization of Henri Cartan's proof of the finite-dimensionality of cohomology groups for compact complex manifolds:

Theorem 15.6 (H. Cartan [35]) *Let $\mathcal{F}$ be a coherent analytic sheaf on a compact complex manifold X, then the cohomology groups*

$$H^q(X, \mathcal{F}), \ q \geq 1,$$

are all finite-dimensional.

This theorem of Henri Cartan is, in turn, a generalization of Kodaira's finite-dimensionality theorem for cohomology groups with coefficients in a locally free sheaf which used a version of Dolbeault's theorem and the theory of harmonic differential forms for its proof (Hodge theory; see Sect. 14.4). Grauert's version of this theorem (Lemma 15.6) is for a compact complex manifold *with boundary* for the simplest locally free sheaf, namely the sheaf of holomorphic functions $\mathcal{O}$. Cartan's proof and Grauert's variation of that proof use the Čech-theoretic representation of cohomology groups in terms of coverings. We will discuss this shortly, but first we discuss a result from complex analysis which plays an important role in these finite-dimensionality theorems.

Let U be any open set in X, then, as we noted at the beginning of Sect. 15.3, we can equip the vector space $\mathcal{O}(U)$ with the structure of a Fréchet space with the topology of uniform convergence on compact subsets of U. If F_1 and F_2 are two Fréchet spaces, and u is a linear mapping

$$u : F_1 \to F_2,$$

then u is said to be *completely continuous*[10] if there is a neighborhood $\mathcal{U}$ of $0 \in F_1$ such that its image $u(\mathcal{U})$ in F_2 is relatively compact. We now have an important lemma that will be quite useful for us.

Lemma 15.8 *Let V and U be two open subsets of X with $V \subset\subset U$, then the natural restriction mapping*

$$r : \mathcal{O}(U) \to \mathcal{O}(V)$$

is a continuous linear mapping of Fréchet spaces, and moreover, it is a completely continuous mapping.

This is a generalization of theorems of Montel and Vitali from classical function theory to this more abstract setting. See, for instance, a proof of this in Gunning and Rossi's monograph [93] (Proposition 1, Chap. VIII, pp. 234–235).

We summarize briefly the Čech representation of cohomology here, which we need for our purposes. For a more formal presentation of this theory we refer to the monograph by Hirzebruch [104] or the summary in Wells [239].

Let $\mathcal{F}$ be a sheaf on a topological space X, and let $\mathfrak{U} = \{U_i\}$ be an open covering of X, then we define a *q-simplex*

$$\sigma = (U_0, \ldots, U_q)$$

to be an ordered collection of $U_i \in \mathfrak{U}$, and let the intersection

$$|\sigma| := \bigcap_{U_i \in \sigma} U_i$$

[10] Completely continuous mappings or operators are now referred to as compact linear mappings or compact operators.

be the *support* of the simplex σ. A q-*cochain* with coefficients in $\mathcal{F}$ is an assignment

$$c(\sigma) \in \mathcal{F}(|\sigma|),$$

(i.e., a section of the sheaf on the intersection) which is also alternating, that is, the sign of $c(\sigma)$ changes if two indices in σ are permuted.

By using pointwise addition of sections of the sheaf $\mathcal{F}$ on intersecting open sets, we can define

$$C^q(\mathfrak{U}, \mathcal{F})$$

to be the Abelian group of q-cochains. We let the coboundary operator

$$\delta : C^q(\mathfrak{U}, \mathcal{F}) \to C^{q+1}(\mathfrak{U}, \mathcal{F})$$

be defined by

$$\delta c(\sigma) = \sum_{i=1}^{q+1} (-1)^i r_{|\sigma|}^{|\sigma_i|} c(\sigma_i),$$

where

$$\sigma_i = (U_0, \ldots, U_{i-1}, U_{i+1}, \ldots, U_{q+1}),$$

and

$$r_{|\sigma|}^{|\sigma_i|} : \mathcal{F}(|\sigma_i|) \to \mathcal{F}(|\sigma|)$$

is the sheaf restriction mapping.

It follows easily that $\delta^2 = 0$, and we can form cocycles, coboundaries, and cohomology groups in the usual manner. Namely, we set

$$\begin{aligned} Z^q(\mathfrak{U}, \mathcal{F}) &:= \ker \delta : C^q(\mathfrak{U}, \mathcal{F}) \to C^{q+1}(\mathfrak{U}, \mathcal{F}), \\ B^q(\mathfrak{U}, \mathcal{F}) &:= \operatorname{im} \delta : C^{q-1}(\mathfrak{U}, \mathcal{F}) \to C^q(\mathfrak{U}, \mathcal{F}), \\ H^q(\mathfrak{U}, \mathcal{F}) &:= Z^q(\mathfrak{U}, \mathcal{F}) / B^q(\mathfrak{U}, \mathcal{F}). \end{aligned}$$

We call $H^q(\mathfrak{U}, \mathcal{F})$, $q \geq 0$, the Čech *cohomology groups of* X *with coefficients in* $\mathcal{F}$, *with respect to the covering* $\mathfrak{U}$. By using a limiting process with respect to refinements of coverings we can define the direct limit,

$$H^q(X, \mathcal{F}) := \lim_{\mathfrak{U}} H^q(\mathfrak{U}, \mathcal{F}),$$

and it is a theorem that this cohomology agrees with the sheaf cohomology introduced in Sect. 14.3.

Leray showed in 1950 that one can compute this cohomology in a certain family of special cases without taking a limit.

Theorem 15.7 (Leray [142]) *If the simplices of the covering $\mathfrak{U}$ have the property that*

$$H^q(|\sigma|, \mathcal{F}) = 0,$$

then

$$H^q(X, \mathcal{F}) \cong H^q(\mathfrak{U}, \mathcal{F}), \ q \geq 0.$$

This theorem is proved[11] and used in Henri Cartan's lectures in 1954 [35] on the finite-dimensionality of cohomology groups on compact complex manifolds, as mentioned above, and we will see below how Grauert used it in the same way in the following paragraphs. The key to using this theorem is to find an open covering which satisfies the hypotheses of vanishing cohomology for the intersections of the elements of the covering, and we see an example of this in our discussion below.

We now formulate two fundamental theorems on Stein manifolds for coherent analytic sheaves due to Henri Cartan from his lectures from 1952, which we mentioned earlier.

Theorem 15.8 (Henri Cartan [34]) *Let X be a Stein manifold, and let $\mathcal{F}$ be a coherent analytic sheaf on X, then*

A *For each $x \in X$, the stalk $\mathcal{F}_x$ can be generated over the ring $\mathcal{O}_x$ by a finite number of global sections of $\mathcal{F}$.*
B $H^q(X, \mathcal{F}) = 0, \ q \geq 1.$

The two parts of this theorem are referred to in the literature as Theorems A and B of Cartan (as they were labeled in Cartan's original paper [34]). They represent the culmination of almost 50 years of research in the field of several complex variables by numerous researchers. Proofs can be found in the references referred to earlier, including, in particular, the original paper by Cartan.

For instance, if $\mathcal{O}$ is the sheaf of holomorphic functions on a domain of holomorphy $X \subset \mathbf{C}^n$, then Theorem B asserts that $H^1(X, \mathcal{O}) = 0$, and this is equivalent to the solution to Cousin's first problem on such a domain (see, e.g., Gunning and Rossi [93] or the other references for a discussion of the two classical Cousin problems). In the case of any domain $D \subset \mathbf{C}$, the assertion that $H^1(D, \mathcal{O}) = 0$ is equivalent to the classical Mittag-Leffler theorem for the domain D (see Sect. 9.6, where we discuss the Mittag-Leffler theorem for a domain in the plane).

If $\mathcal{I}$ is the ideal sheaf of a subvariety of a Stein manifold X, then Henri Cartan showed in 1950 [33] that this ideal sheaf is a coherent analytic sheaf. Let's assume for simplicity that V is a submanifold of X, and we let $\mathcal{O}(V)$ denote the holomorphic functions on the submanifold V. From the exact sequence of sheaves

$$0 \to \mathcal{I} \to \mathcal{O} \to \mathcal{O}/\mathcal{I} \to 0,$$

[11] One can find proofs in any number of more modern references, e.g., Gunning and Rossi [93], as well as other references which use spectral sequences for a proof of this theorem. Spectral sequences were a part of Leray's original formalism; see, for instance, Roger Godement's classical reference on sheaf theory [83].

which we used in Kodaira's proof of his vanishing theorem in Sect. 14.6 for the case of V being a point, we obtain the exact sequence

$$H^0(X, \mathcal{O}) \to H^0(X, \mathcal{O}/\mathcal{I}) \to H^1(X, \mathcal{I}).$$

By Theorem 15.8 (Cartan's Theorem B) we have $H^1(X, \mathcal{I}) = 0$, and hence the mapping

$$\mathcal{O}(X) = H^0(X, \mathcal{O}) \to H^0(X, \mathcal{O}/\mathcal{I}) \cong \mathcal{O}(V)$$

is surjective. This implies that any holomorphic function on V is the restriction to V of a holomorphic function on X, an important consequence of Cartan's work.

We now return to our outline of the proof of Grauert's finite-dimensionality theorem for strongly pseudoconvex domains (Lemma 15.6). We consider a strongly pseudoconvex domain $D \subset X$, as hypothesized, and let $p \in \partial D$. We want to consider a small perturbation of the boundary near the point p, leaving the remainder of the boundary fixed. Let U be a suitably small neighborhood of p, and suppose that χ is a smooth nonnegative function defined on X such that, for ε small,

$$\begin{gathered} 0 < \chi(x) < \varepsilon, \;\; x \in U, \\ \chi(x) \equiv 0, \;\; x \in X - U. \end{gathered}$$

We also choose χ to have sufficiently small first and second derivatives. We let

$$\tilde{\varphi} := \varphi - \chi,$$

where φ is the defining function for the boundary of D, as formulated above, and we define the perturbed domain

$$D_1 := \{x \in X : \tilde{\varphi} < 0\},$$

and thus $D_1 \supset D$, for ε sufficiently small, and is still strongly pseudoconvex. We have "bumped" the domain D outward near the point p.

Now, by choosing the neighborhood U and the bumping function χ sufficiently small, we can find a finite Stein open covering $\{U_0, \ldots, U_N\}$ of D_1 with the property that $U \cap U_i = \emptyset, i = 1, \ldots N$, and $U \subset U_0$. By using Levi's theorem that there are Stein neighborhoods of boundary points of a strongly pseudoconvex domain whose intersection with the domain is Stein (Lemma 15.1), we can assume that the open sets

$$U_i \cap D, U_i \cap D_1$$

are all Stein. Moreover, it is easy to prove the fact that if $W, \tilde{W}$ are open Stein submanifolds of a complex manifold, then their intersection $W \cap \tilde{W}$ is also Stein.

We can now define coverings of the two domains D and D_1, using this data. Namely, let

$$\mathfrak{U}_D = \{U_i \cap D\}, \mathfrak{U}_{D_1} = \{U_i \cap D_1\}$$

be coverings of D and D_1 being defined by the restrictions of the covering $\mathfrak{U}$ to each of these domains.

Since the domains in these two coverings and their intersections are all Stein, if follows from Leray's theorem (Theorem 15.7) that we have

$$H^q(D_1, \mathcal{O}) \cong H^q(\mathfrak{U}_{D_1}, \mathcal{O}), q \geq 1,$$
$$H^q(D, \mathcal{O}) \cong H^q(\mathfrak{U}_D, \mathcal{O}), q \geq 1.$$

We now claim that the restriction mapping

$$r^* : H^q(D_1, \mathcal{O}) \to H^q(D, \mathcal{O}), \text{ is surjective for } q \geq 1. \tag{15.14}$$

Namely, we have

$$Z^q(\mathfrak{U}_{D_1}, \mathcal{O}) = Z^q(\mathfrak{U}_D, \mathcal{O}), \ q \geq 1. \tag{15.15}$$

This follows since the support of the cochains for these two cocyle groups, for $q \geq 1$, have the form

$$U_{i_0} \cap \cdots U_{i_q},$$

intersecting either D_1 or D. But these intersections do not intersect the original perturbation neighborhood U, and thus we obtain

$$U_{i_0} \cap \cdots U_{i_q} \cap D_1 = U_{i_0} \cap \cdots U_{i_q} \cap D.$$

Therefore we have (15.15), which then yields immediately (15.14).

By applying this process on successive bumps, which after a finite number of steps can be chosen to cover ∂D, we obtain a perturbation D' of the domain D with the property that

$$D \subset\subset D',$$

D' is strongly pseudoconvex, and the restriction mapping

$$r^* : H^q(D', \mathcal{O}) \to H^q(D, \mathcal{O}) \text{ is surjective, for all } q \geq 1. \tag{15.16}$$

Grauert now uses a variation of Cartan's proof of finite-dimensionality for compact complex manifolds extended to this strongly pseudoconvex setting (a compact manifold with boundary), and we outline this now.

Consider two finite Stein coverings of $\overline{D}$, $\{U_0, \ldots, U_K\}$ and $\{U_0', \ldots, U_K'\}$, with $U_i \subset\subset U_i'$, and $U_i \cap D$ and $U_i' \cap D$ being Stein manifolds. Let $W_i = U_i \cap D$, and let

$$\mathfrak{W} = \{W_i\}$$

be a well-defined Stein covering of D. We now choose a strongly pseudoconvex perturbation D' as above such that

$$\begin{aligned} D &\subset\subset D', \\ D' &\subset \textstyle\bigcup_i U_i', \end{aligned}$$

and where (15.16) is satisfied. We let

$$\mathfrak{W}' = \{W_i' = U_i' \cap D'\}$$

be a Stein covering of D' (here we have used that a perturbation of the boundary in Levi's theorem (Lemma 15.1) also yields a domain of holomorphy).

For these fixed coverings we have the vector spaces of q-cochains and q-cocycles

$$C^q(\mathfrak{W}, \mathcal{O}), Z^q(\mathfrak{W}, \mathcal{O}), C^q(\mathfrak{W}', \mathcal{O}), Z^q(\mathfrak{W}', \mathcal{O}),$$

each of which can be given the structure of a Fréchet space. The set of holomorphic functions on a typical intersection, say $W_{i_0} \cap \cdots \cap W_{i_q}$, is a Fréchet space (with the topology of uniform convergence on compact subsets), and we have a finite direct sum of such Fréchet spaces for the cochains and cocycles listed above.

We now set

$$\begin{aligned} V_1 &= Z^q(\mathfrak{W}', \mathcal{O}) \oplus C^{q-1}(\mathfrak{W}, \Omega), \\ V_2 &= Z^q(\mathfrak{W}, \Omega), \end{aligned}$$

for $q \geq 1$, and these are also then two Fréchet spaces. We define linear mappings

$$\begin{aligned} u &: V_1 \to V_2, \\ v &: V_1 \to V_2, \end{aligned}$$

by setting

$$\begin{aligned} u(z \oplus c) &= r(z) + \delta(c), \\ v(z \oplus c) &= = r(z), \end{aligned}$$

where r is the restriction mapping, and δ is the coboundary operator. These mappings are continuous linear mappings of the Fréchet spaces involved. Since

$$\begin{aligned} H^q(D', \mathcal{O}) &\cong H^q(\mathfrak{W}', \mathcal{O}), \\ H^q(D, \mathcal{O}) &\cong H^q(\mathfrak{W}, \mathcal{O}), \end{aligned}$$

it follows from (15.16) that the mapping u is surjective. The mapping v is completely continuous, using the generalization of Montel's theorem, Lemma 15.8.

In 1953 Laurent Schwartz (1915–2002) proved the following important theorem concerning Fréchet spaces, which was used by both Cartan and Grauert in their finite-dimensionality theorems.

Theorem 15.9 (Schwartz [211]) *Let F_1 and F_2 be Fréchet spaces, and let*

$$u : F_1 \to F_2$$

be a surjective continuous linear mapping and

$$v : F_1 \to F_2,$$

be a completely continuous mapping, then the vector space

$$F_2/(u(F_1) + v(F_1))$$

is finite-dimensional.

Schwartz devotes two pages to his proof, using the relatively new (at that time) formalism of locally convex topological vector spaces. A more detailed proof can be found in Gunning and Rossi ([93], pp. 294–295). See also the exposition by Serre in the Cartan seminar from 1954 [214].

In our case above we have

$$\begin{aligned} V_2/(u(V_1) + v(V_1)) &= Z^q(\mathfrak{W}, \mathcal{O})/(r(Z^q(\mathfrak{W}', \mathcal{O})) + \delta(C^{q-1}(\mathfrak{W}, \mathcal{O})) \\ &\qquad - r(Z^q(\mathfrak{W}', \mathcal{O}))) \\ &= Z^q(\mathfrak{W}, \mathcal{O})/\delta(C^{q-1}(\mathfrak{W}, \mathcal{O})) \\ &= H^q(D, \Omega), \end{aligned}$$

and, hence,

$$\dim H^q(D, \mathcal{O}) < \infty,$$

as desired.

Now we turn to the proof of Lemma 15.7. Let D be, as before, a strongly pseudoconvex domain in a complex manifold X. Let $p \in \partial D$. By Levi's result [143] that we used earlier, there is a Stein neighborhood U of p such that $D \cap U$ is Stein. Part of Levi's proof consists of showing that there is a holomorphic function $f \in \mathcal{O}(U)$ such that

$$\{x \in U : f(x) = 0\} \cap \partial D = p.$$

Thus if we consider $g = 1/f$, then $g \in \mathcal{O}(D \cap U)$, and

$$\lim_{x \to p} |g(x)| = \infty.$$

What Grauert does is to use this function f defined only in a neighborhood of the point p to find a similar function $\tilde{f}$ defined in a perturbation D' of D.

Let now D' be a strongly pseudoconvex perturbation of D such that D' coincides with D in $X - U$, and such that $U \cap D'$ is Stein. Let

$$S = \{x \in U \cap D' : f(x) = 0\},$$

and thus $S \cap \partial D = p$, as above. Let $[S]$ be the (positive) divisor on D' defined by S in U, and let F be the holomorphic line bundle determined by $[S]$. We let $\Gamma(D', \mathcal{O}(F))$ denote the holomorphic sections of the sheaf $\mathcal{O}(F)$ on the domain D', and let $h \in \Gamma(D', \mathcal{O}(F))$ denote the canonical holomorphic section that vanishes to first order on S.[12]

Consider the exact sequence

$$0 \to \mathcal{O}(F^k) \xrightarrow{\alpha} \mathcal{O}(F^{k+1}) \to \mathcal{O}(F^{k+1})/\alpha\mathcal{O}(F^k) \to 0,$$

defined by

$$\alpha(s) = h \cdot s.$$

We note that the quotient sheaf on the right in this exact sequence is 0 for points $x \in D' - S$ and has the form

$$\left(\mathcal{O}(F^{k+1})/\alpha\mathcal{O}(F^k)\right)_x \cong \mathcal{O}_S(F_S^{k+1})_x, \; x \in S,$$

where we let

$$F_S^{k+1} := F^{k+1}|_S$$

be the restriction of the line bundle F^{k+1} to the submanifold S, and $\mathcal{O}_S$ denotes the sheaf of holomorphic functions on S.

Since S is a complex submanifold of $D' \cap U$, which is Stein, it follows that S itself is Stein, which is easy to verify. We let $\mathcal{F}^{k+1}$ be the trivial extension of $\mathcal{O}_S(F_S^{k+1})$ to D' (i.e., this extension is 0 at points not on S), and the exact sequence above becomes

$$0 \to \mathcal{O}(F_k) \xrightarrow{\alpha} \mathcal{O}(F^{k+1}) \xrightarrow{\beta} \mathcal{F}^{k+1} \to 0, \tag{15.17}$$

for all $k \in \mathbf{Z}$.

For simplicity of notation we let

$$H^q(Y, E) := H^q(Y, \mathcal{O}(E)),$$

where Y is a complex manifold, and $\mathcal{O}(E)$ is the sheaf of holomorphic sections of a holomorphic vector bundle E.

---

[12]This can be made quite explicit in this special case. Let $U_0 = U \cap D'$ and $U_1 = D$ be a covering of D'. The functions

$$f_0 = f \text{ in } U_0,$$
$$f_1 = 1 \text{ in } U_1,$$

define the transition functions $g_{01} = f_0/f_1$ on $U_0 \cap U_1$ for the line bundle F associated to the divisor $[S]$, and the canonical section h has the (same) form

$$h_0 = f_0 \text{ in } U_0,$$
$$h_1 = 1 \text{ in } U_1.$$

From the short exact sequence (15.17) we have the long exact sequence of cohomology groups

$$\begin{aligned} 0 \to H^0(D', F^k) \to H^0(D', F^{k+1}) \to H^0(D', \mathcal{F}^{k+1}) \to \\ H^1(D', F^k) \to H^1(D', F^{k+1}) \to H^1(D', \mathcal{F}^{k+1}) \to \cdots . \end{aligned}$$

But

$$H^q(D', \mathcal{F}^{k+1}) \cong H^q(S, F_S^{k+1}) = 0,$$

for $q \geq 1$, by Theorem B of Cartan (Theorem 15.8B), since S is Stein.

This yields the exact sequence

$$H^0(D', F^{k+1}) \xrightarrow{\beta^*} H^0(S, F_S^{k+1}) \to H^1(D', F^k) \xrightarrow{\alpha^*} H^1(D', F^{k+1}) \to 0. \quad (15.18)$$

Thus the linear mapping α^* in this sequence is surjective, and it follows that

$$\dim H^1(D', F^k) \geq \dim H^1(D', F^{k+1}),$$

for $k \geq 0$. Letting

$$d_k = \dim H^1(D', F^k),$$

it follows that these dimensions are all finite since $d_0 < \infty$ by Lemma 15.6.

Since $d_k < \infty$, and $d_k \geq d_{k+1}$ for $k \geq 0$, it follows that there is a $k_0 \geq 0$ such that

$$d_k = d_k + 1, \; k > k_0.$$

Choose now any $k > k_0$, and it follows that $\ker \alpha^* = 0$ in (15.18), and hence the mapping

$$H^0(D', F^{k+1}) \xrightarrow{\beta^*} H^0(S, F_S^{k+1}) \quad (15.19)$$

in (15.18) is surjective.

Let now s_0 be a holomorphic section of F_S^{k+1} defined near p with $s_0(p) \neq 0$. By Cartan's Theorem A (Theorem 15.8A), there is a holomorphic section s of F_S^{k+1} on S with $s(p) \neq 0$, using again the fact that S is Stein. By the surjectivity of the mapping (15.19) it follows that there is a section $\tilde{s}$ of F^{k+1} on D' whose restriction to S is s, and hence $\tilde{s}(p) \neq 0$, as well.

We can now form the quotient

$$\tilde{f} := \frac{\tilde{s}}{h^{k+1}},$$

which is a meromorphic function on D'. Moreover, $\tilde{f}$ is holomorphic on D, since $h \neq 0$ on D, and since $h^{k+1}(p) = 0$ and $\tilde{s}(p) \neq 0$, it follows that

$$\lim_{x \to p} |\tilde{f}(x)| = \infty.$$

This completes our outline of the proof of Lemma 15.7 and hence completes our discussion of the proof of Grauert's Theorem 15.5 which asserts that a strongly pseudoconvex domain is indeed holomorphically convex.

As the example at the beginning of this section shows, a manifold can be holomorphically convex, but not necessarily Stein. We need additional information to conclude that certain holomorphically convex manifolds are Stein. Grauert showed in [85] that a holomorphically convex manifold X which is K-complete is a Stein manifold (and conversely). Here K-complete means that for any point x_0 of X, there are a finite number of holomorphic functions $f_1, \ldots, f_K$ on X such that

$$\{x \in X : f_j(x) = 0, j = 1, \ldots, K\}$$

consists of isolated points in X.

If φ is a strictly plurisubharmonic function on a complex manifold X, then it follows by a maximum principle argument that X has no compact subvarieties of positive dimension. On the other hand, a holomorphically convex manifold X is K-complete if and only if X has no compact subvarieties of dimension greater than zero. These results are relatively straightforward and are proved in Grauert's 1958 paper [86]. From these remarks, we have the following important result of Grauert, providing the final solution of the Levi problem in this context.

Theorem 15.10 (Grauert [86]) *Let D be a strongly pseudoconvex domain which has a strictly plurisubharmonic function φ defined on D, then D is a Stein manifold.*

15.7 The Grauert Real-Analytic Embedding Theorem

As we described in Chap. 12, Whitney proved in 1936 the first of the major embedding theorems for differentiable manifolds. A significant problem that remained unsolved at that time was whether a real-analytic manifold could be real-analytically embedded in real Euclidean space of some dimension. One year later, in 1937, Bochner showed [21] that if a *compact* real-analytic n-dimensional manifold admits a real-analytic Riemannian metric g_{ij}, then there is a real-analytic embedding of M into $\mathbf{R}^{2n+1}$. Note that Bochner's work did not obtain an isometric embedding; that came only much later with the work of Nash, as described in Chap. 13.

We now outline briefly Bochner's approach to this problem. Given the real-analytic metric, Bochner considered the Laplacian operator defined on functions

$$\Delta = d^* d$$

acting on smooth functions defined on M. The eigenfunctions of this Laplacian, solutions of

$$\Delta\varphi_j = \lambda_j\varphi,$$

form a discrete basis for the Hilbert space $\mathcal{H}$ consisting of all L^2-functions on M, and these are real-analytic eigenfunctions.

The embedding functions

$$(f_1, \ldots, f_{2n+1}); M \to \mathbf{R}^{2n+1},$$

given by Whitney, are C^∞ functions on M, and hence, of course, are L^2-functions and elements of the Hilbert space $\mathcal{H}$. Thus each such function can be approximated in the L^2-norm by a finite series of the basis functions, which yields L^2-norm approximations of the functions $f_1, \ldots, f_{2n+1}$ by real-analytic functions $r_1, \ldots, r_{2n+1}$. The major contribution of Bochner's paper, using delicate real-analysis estimates, is to exploit this L^2-approximation by these real-analytic functions and to show that they approximate the embedding functions in the C^1-norm, thus yielding a real-analytic embedding of M into $\mathbf{R}^{2n+1}$.

In 1958 two papers appeared, each of which provided real-analytic embeddings for compact real-analytic manifolds, improving on Bochner's result, by not requiring a real-analytic metric, and solving, in the case of compact manifolds, Whitney's problem. The first of these was by Charles B. Morrey, Jr. (1907–1984) [160], who was able to find a real-analytic immersion of M into a Euclidean space, which provided a Riemannian metric induced from the immersion, and he was able to then use Bochner's result to obtain an embedding. Morrey used important tools from the theory of elliptic differential equations in his proof.

The second result, due to Grauert, is a simple application of Grauert's solution to the Levi problem that we described in the previous section, which used powerful methods from the theory of several complex variables. We will describe Grauert's proof in the following paragraphs. Moreover, Grauert was able to refine his proof to obtain an embedding theorem for noncompact real-analytic manifolds, as well, thus providing the complete solution to the problem posed by Whitney in 1936. We formulate this result as our final embedding theorem of this book.

Theorem 15.11 (Grauert [86]) *Let M be a real-analytic n-dimensional manifold, then there is a proper real-analytic embedding*

$$f : M \to \mathbf{R}^{2(2n+1)}.$$

In Fig. 15.1 we show the first page of Grauert's paper in which he both solved the Levi problem on complex manifolds and applied this result to obtain his general real-analytic embedding theorem.

We will first outline Grauert's proof for the case where M is compact, and then we will indicate the refinements used to extend the proof to the noncompact case.

The first idea is to embed M as a real-analytic submanifold of a complex manifold which mimics the standard embedding of $\mathbf{R}^n$ in $\mathbf{C}^n$ given by

ANNALS OF MATHEMATICS
Vol. 68. No. 2, September, 1958
Printed in Japan

ON LEVI'S PROBLEM AND THE IMBEDDING OF REAL-ANALYTIC MANIFOLDS

BY HANS GRAUERT

(Received March 19, 1958)

Introduction

In 1911 E. E. Levi [18] showed that the boundary of a domain of holomorphy is not arbitrary. It satisfies certain condition of convexity and therefore is called pseudoconvex.[1] The pseudoconvexity is a local property. To prove that a domain with twice differentiable boundary is pseudoconvex it is only necessary to verify that some differential inequalities are satisfied (see [3]).

For more than forty years it was an open problem of the theory of several complex variables whether the Levi conditions are sufficient for the domains of holomorphy. At first the problem was solved for special domains. After refuting a counter-example of Blumenthal, H. Behnke proved that Levi's conjecture is true for (complex) 2-dimensional circular domains [1]. The first general result, however, was not obtained until 1942 by K. Oka [22]. Oka showed : *each pseudoconvex domain G of the* 2-*dimensional complex number space C^2 is a domain of holomorphy.* The problem was solved for the case of dimension $n > 2$ by K. Oka [23], H. Bremermann [5] and F. Norguet [21] in 1954. K. Oka [22] even proved that every unbranched (not necessarily schlicht) pseudoconvex domain over the n-dimensional complex number space C^n is a (holomorphically convex) domain of holomorphy. For branched domains and domains in the closed C^n, that is, in the n-dimensional complex projective space P^n, the problem is still unsolved.[2]

Using plurisubharmonic functions it can easily be proved that each pseudoconvex domain $G \subset C^n$ can be exhausted by strongly pseudoconvex domains $G_\nu \subset\subset G$ (see [23]). By a theorem of H. Behnke and K. Stein [2], the limit of domains of holomorphy is again a domain of holomorphy. So Levi's conjecture has to be verified only for strongly pseudoconvex domains (for definitions see § 1).

In this short paper relatively compact, *strongly* pseudoconvex subdomains G of complex manifolds $\mathfrak{M}$ are considered (without any assump-

[1] There are several definitions of pseudoconvexity. See [17].

[2] K. Oka has announced that the answer is also " yes " for these cases.

Fig. 15.1 First page from Grauert's 1958 paper on the solution to the Levi problem and the embedding of real-analytic manifolds [86]. *Reprinted with the permission of the Annals of Mathematics*

$$\mathbf{R}^n = \{z = (z_1, \ldots, z_n) \in \mathbf{C}^n : \operatorname{Im} z_j = 0,\ \ j = 1, \ldots, n\}. \tag{15.20}$$

This has been carried out in a very nice manner by Whitney and Francois Georges René Bruhat (1929–2007) in 1959 [243], which was used by both Morrey and Grauert in their separate proofs, and allows both authors to use complex variable methods to solve this real-analytic problem, an essential element in their solutions.

We will formulate this as a theorem, but first we introduce Whitney and Bruhat's definition of a complexification of a real-analytic manifold. A *complexification* of a real-analytic manifold M is a complex manifold X and a real-analytic embedding $\tau : M \to X$, such that for any coordinate chart neighborhood U of a point $p \in X$, there is a biholomorphic mapping

$$h : U \to U' \subset \mathbf{C}^n,$$

so that

$$h(U \cap \tau(M)) = U' \cap \mathbf{R}^n,$$

where $\mathbf{R}^n$ is embedded in $\mathbf{C}^n$ as in (15.20).

Theorem 15.12 (Whitney–Bruhat [243]) *Let M be a real-analytic manifold, then there exists a complexification (X, τ), and for any two such complexifications (X_1, τ_1), (X_2, τ_2), there is a biholomorphic extension of the real-analytic isomorphism*

$$\tau_2 \circ \tau_1^{-1} : \tau_1(M) \to \tau_x(M)$$

to a neighborhood of $\tau_1(M)$ in X.

The latter part of the theorem simply asserts that complexifications are unique near the embedded version of M in any such complexification.

We now return to Grauert's proof of Theorem 15.11. Let X be a complexification of the given real-analytic manifold M, and let $\{U_\alpha\}$ be a covering of M by open sets with biholomorphic mappings

$$h_\alpha : U_\alpha \to U'_\alpha \subset \mathbf{C}^n,$$

where each U'_α is an open ball in $\mathbf{C}^n$, and such that

$$h_\alpha(U_\alpha \cap M) = U'_\alpha \cap \mathbf{R}^n,$$

where

$$\mathbf{R}^n = \{z = (x_1 + iy_1, \ldots, x_n + iy_n) \in \mathbf{C}^n : y_j = 0, j = 1, \ldots, n\}.$$

We assume that the covering is a countable covering, and, in the case where M is compact, we can assume that this is a finite covering.

In a given U'_α we can define the function

$$\tilde{\varphi}_\alpha(z) := \sum_{j=1}^{n} y_j^2 = -\frac{1}{4}\sum_{j=1}^{n}(z_j - \overline{z}_j)^2.$$

This function is strictly plurisubharmonic, as is easy to compute. We recall that strictly plurisubharmonic means the Hermitian matrix

$$\frac{\partial^2 \tilde{\varphi}_\alpha}{\partial z_i \partial \overline{z}_j}$$

is positive-definite at points of U'_α, and, in particular, at points of $U'_\alpha \cap \mathbf{R}^n$, which will be important below. Also,

$$\begin{aligned} \tilde{\varphi}_\alpha &= 0 \text{ on } \mathbf{R}^n \cap U'_\alpha, \\ d\tilde{\varphi}_\alpha(z) &= 0 \text{ on } \mathbf{R}^n \cap U'_\alpha. \end{aligned}$$

If we let

$$\varphi_\alpha = \tilde{\varphi}_\alpha \circ h_\alpha,$$

then φ_α satisfies on U_α:

$$\begin{aligned} &\varphi_\alpha \text{ is strictly plurisubharmonic on } U_\alpha \cap M, \\ &\varphi_\alpha = 0 \text{ on } U_\alpha \cap M, \\ &d\varphi_\alpha = 0 \text{ on } U_\alpha \cap M. \end{aligned} \tag{15.21}$$

We call a C^∞ function satisfying the three properties in (15.21) a *p-function*.

Suppose that on any open set W in X with $W \cap M \neq \emptyset$, φ_1 and φ_2 are two p-functions in W, then, for any smooth function $\chi > 0$ on W,

$$\chi\varphi_1, \text{ and } \varphi_1 + \varphi_2,$$

are also p-functions on W. This is easy to verify, since, in suitable local coordinates, for instance,

$$\frac{\partial^2(\chi\varphi_1)}{\partial z_i \partial \overline{z}_j}(z) = \chi(z)\frac{\partial^2 \varphi_1}{\partial z_i \partial \overline{z}_j}(z) + R(z),$$

where $R(z)$ vanishes on $W \cap M$, since φ_1 and $d\varphi_1$ vanish on $W \cap M$. Thus we can use a partition of unity with respect to the covering U_α to obtain a p-function φ defined in a neighborhood U of M in X, such that

$$M = \{x \in U : \varphi(x) = 0\}.$$

Suppose now that M is compact, then there is an $\varepsilon > 0$ such that

$$T := \{x \in U : \varphi(x) < \varepsilon\}$$

is a strongly pseudoconvex domain in X, which has a strictly plurisubharmonic function φ defined on T. It now follows from Grauert's theorem (solution to the Levi problem), Theorem 15.10, that T is a Stein manifold. By the proper holomorphic embedding theorem for Stein manifolds (Theorem 15.2), there is a proper holomorphic embedding

$$f : T \to \mathbf{C}^{2n+1},$$

and hence the restriction of F to M, which we denote by g, gives us the desired proper real-analytic embedding

$$g : M \to \mathbf{R}^{2(2n+1)}.$$

In fact, here we only need to use Theorem 15.4 to obtain a holomorphic embedding defined in a neighborhood of M in T, since T being Stein implies that T is holomorphically separable and has global local coordinates, and M is a compact subset of T. Thus, in this case, we don't need the full power of the proper holomorphic embedding theorem.

In the case where M is noncompact, Grauert constructs a similar tube domain T (open set containing M, equipped with a strictly plurisubharmonic function φ defined on T), such that T is the union of relatively compact strongly pseudoconvex domains T_ρ, where ρ is a real parameter. The sets T_ρ are constructed in a similar manner to the strongly pseudoconvex domain used in the proof above for the case where M is compact. Moreover, they provide an exhaustion of the form

$$T_\rho \subset T_{\rho'}, \rho < \rho',$$

and

$$T = \bigcup_\rho T_\rho.$$

Each of the domains T_ρ is a Stein manifold, as in the earlier discussion, and by a theorem of Ferdinand Docquier and Grauert [57], the union T is also a Stein manifold. Now, using the proper embedding theorem for Stein manifolds (Theorem 15.2), it follows that there is a proper real-analytic embedding

$$g : M \to \mathbf{R}^{2(2n+1)}.$$

Bibliography

1. Abel, N.H.: Mémoire sur les équations algébriques ou l'on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré. In: Sylow, L., Lie, S. (eds.) Oeuvres complètes de Niels Henrik Abel, pp. 28–33. Grøndahl & Søn, Christiania (Oslo) (1881). Originally published as a printed brochure by Grendahl, Christiania 1824
2. Abel, N.H.: Mémoire sur une propriété générale d'une classe trés étendue de fonctions transcendantes. In: Sylow, L., Lie, S. (eds.) Oeuvres complètes de Niels Henrik Abel, pp. 145–211. Grøndahl & Søn, Christiania (Oslo) (1881). Presented to the Academy of Sciences in Paris 30 October 1826
3. Abel, N.H.: Recherches sur les fonctions elliptiques. In: Sylow, L., Lie, S. (eds.) Oeuvres complètes de Niels Henrik Abel, pp. 263–388. Grøndahl & Søn, Christiania (Oslo) (1881). Originally published in Journal für die reine und angewandte Mathematik, vols. 2, 3, 1827, 1828
4. Abel, N.H.: Remarques sur quelques propriétés générales d'une certaine sorte de fonctions transcendantes. In: Sylow, L., Lie, S. (eds.) Oeuvres complètes de Niels Henrik Abel, pp. 263–388. Grøndahl & Søn, Christiania (Oslo) (1881). Originally published in Journal für die reine und angewandte Mathematik, vols. 2, 3, 1827, 1828
5. Ahlfors, L.: Complex Analysis. McGraw-Hill, New York (1953)
6. Alexandroff, P.: Über die Strukture der bikompakten topologischen Räume. Math. Ann. **92**, 267–274 (1924)
7. Andersen, K.: Wessel's work on complex numbers and its place in history. In: Branner, B., Lützen, J. (eds.) On the Analytical Representation of Direction, pp. 65–98. The Royal Danish Academy of Science and Letters, Copenhagen (1999)
8. Andrews, B.: Notes on the isometric embedding problem and the Nash–Moser implicit function theorem. In: Surveys in Analysis and Operator Theory, pp. 157–208. Australia National University, Canberra (2002). Proceedings of the Centre for Mathematics and its Applications, Australia National University, 40
9. Argand, J.-R.: Essai sur un manière de représenter les quantités imaginaires dans les constructions géométriques. Gauthier-Villiars, Paris (1874)
10. Arnfinn, O., Piene, R. (eds.): The Legacy of Abel in Algebraic Geometry. Springer, Berlin (2004)
11. Atiyah, M.F., Singer, I.M.: The index of elliptic operators on compact manifolds. Bull. Am. Math. Soc. **69**, 422–433 (1963)
12. Atiyah, M.F., Singer, I.M.: The index of elliptic operators: I. Ann. Math. **87**, 484–530 (1968)
13. Behnke, H., Thullen, P.: Theorie der Funktionen mehrerer komplexer Veränderlichen. Springer, Heidelberg (1934)
14. Behnke, H., Stein, K.: Entwicklung analytischer Funktionen auf Riemannschen Flächen. Math. Ann. **120**, 430–461 (1947)

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2

15. Bernoulli, J.: Solution d'un problème concernant le calcul intégral. In: Opera Omni, Johannis Bernoulli, vol. 1, pp. 393–400. Carolum Jombert, Laussane & Geneva (1742)
16. Bers, L.: Riemann Surfaces. New York University, : Lecture notes on Riemann surfaces, 1957–58, notes by E. Rodlitz and R, Pollack (1958)
17. Bers, L.: Introduction to Several Complex Variables. New York University, New York (1965). Lectures given during the 1962–1963 academic year at NYU
18. Bers, L.: Uniformization, moduli, and Kleinian groups. Bull. Lond. Math. Soc. **4**(3), 257–300 (1972)
19. Betti, E.: Sopra gli spazi di un numero qualunque di dimensioni. Annali di Matematica Pura ed Applicata **4**, 140–158 (1871)
20. Bishop, E.: Mappings of partially analytic spaces. Am. J. Math. **83**, 209–242 (1961)
21. Bochner, S.: Analytic mapping of compact Riemann spaces into Euclidean space. Duke Math. J. **3**, 334–338 (1937)
22. Bochner, S.: Curvature and betti numbers. Ann. Math. **49**, 379–390 (1948)
23. Boehm, K.: Elliptische Funktionen. G.J. Göschen'sche Verlagshandlung, Leipzig (1908)
24. Bosse, A.: Maniere universelle de M. Desargues pour praticquer la Perspective par petit pied Comme le Geometral. A. Bosse, Paris (1647). A. Bosse was the author, printer and engraver
25. Boyer, C.B.: A History of Mathematics. Wiley, New York (1968)
26. Bremermann, H.: Über die Äquivalenz der pseudoconvexen Gebiete und der Holomorphiegebiete im Raum von n komplexen Veränderlichen. Math. Ann. **128**, 63–91 (1954)
27. Brianchon, C.-J.: Mémoire sur les surfaces courbes du second degré. Journal de l'Ecole Polytechnique **VI**, 297–311 (1806)
28. Brianchon, C.-J.: Mémoire sure les Lignes du second Ordre. Bachelier, Paris (1817)
29. Calabi, E., Rosenlicht, M.: Complex analytic manifolds without countable base. Proc. Am. Math. Soc. **4**, 335–340 (1953)
30. Cantor, G.: In: Zermelo, E. (ed.) Gesammelte Abhandlungen. Springer, Berlin (1932)
31. Carnot, L.N.M.: Géométie de Position. J.B.M. Duprat, Paris (1803)
32. Cartan, E.: Sur la possibilité de plonger un espace riemannien donné dans un espace euclidean. Annals de la sociéte polonaise de mathématique **6**, 1–7 (1927)
33. Cartan, H.: Idéaux et modules de fonctions analytiques de variables complexes. Bull. Soc. Math. Fr. **78**, 29–64 (1950). http://www.numdam.org/item?id=BSMF_1950__78__29_0
34. Cartan, H.: Théorie de la convexité (II), exp. No. 9. Séminaire Henri Cartan, tome 4 (1951–1952), pp. 1–12. École Normale Supérieure, Paris (1952). http://www.numdam.org/item?id=SHC_1951-1952__4__A9_0
35. Cartan, H.: Un théorème de finitude. Séminaire Henri Cartan, tome 6 (1953–1954), pp. 1–11. École Normale Supérieure, Paris (1955). http://www.numdam.org/item?id=SHC_1953-1954__6__A17_0
36. Cauchy, A.-L.: Mémoire sur les intégrales définies, prises entre des limites imaginaires. Chez de Bure Frères, Libraires du Roi et de la Bibliothèque du Roi (1825). Reprinted Bull. sciences mathématiques et astronomiques, 7, 1874, pp. 265–304, not contained in Oeurvres d'Cauchy
37. Cauchy, A.-L.: Mémoire sur les intégrales définies. Oeuvres complètes d'Augustin Cauchy Séries, pp. 320–475. Gauthier-Villars, Paris, : First published in Mémoires présentés par divers savants à l'Academie royale des Sciences de l'Institut de France. Sciences mathématiques et physiques. Tome **1**, 1827 (1882)
38. Cauchy, A.-L.: Exercices de mathématiques, 4 vols. Oeuvres complètes d'Augustin Cauchy, Séries 2, Tome 6–9. Gauthier-Villars, Paris, : First appeared as 4 volumes in 1826–1829, published by Debúre Frères. Libraires du Roi et de Bibiliotèque du Roi, Paris (1891)
39. Cauchy, A.-L.: Considérations nouvelles sur les intégrales définies qui s'étendent à tous les points d'une courbe fermée, et sur celles qui sont prises entre des limites imaginaires. Oeuvres complètes d'Augustin Cauchy, Séries 1, Tome 10, pp. 153–168. Gauthier-Villars, Paris (1897). First appeared as a paper in Comptes Rendu de l'Académie des Sciences, T. XXIII, p. 689, 1846

40. Cauchy, A.-L.: Sur les intégrales qui s'étendent à tous des points d'une courbe fermè. Oeuvres complètes d'Augustin Cauchy, Séries 1, Tome 10, pp. 70–74. Gauthier-Villars, Paris (1897). First appeared as a paper in Comptes Rendu de l'Académie des Sciences, T. XXIII, p. 251, 1846
41. Cauchy, A.-L.: Mémoire sur les fonctions irrationelles. Oeuvres complètes d'Augustin Cauchy, Séries 1, Tome 11, pp. 292–300. Gauthier-Villars, Paris (1899). First appeared as a paper in Comptes Rendu de l'Académie des sciences, T. XXXII, p. 68, 1851; see also other papers in remainder of this volume of Cauchy's Oeuvres
42. Cauchy, A.-L.: Exercices d'analyse et de physique mathématique. Oeuvres complètes d'Augustin Cauchy, Séries 2, Tome 12, pp. 8–469. Gauthier-Villars, Paris (1916). First appeared as a book published by Bachelier, Paris, 1841
43. Chasles, M.: Aperçu Historique des Méthodes en Géométrie. M. Hayez, Bruxelles (1837)
44. Chern, S.S.: Characteristic classes of Hermitian manifolds. Ann. Math. **47**, 85–121 (1946)
45. Chern, S.S.: Complex Manifolds Without Potential Theory. Van Nostrand Reinhold Co., New York (1967)
46. Chern, S.S.: On the curatura intega in a riemannian manifold. Ann. Math. **46**, 674–684 (1945)
47. Chow, W.-L.: On compact complex analytic varieties. Am. J. Math. **71**, 893–914 (1949)
48. Clairaut, A.C.: Recherches sur les courbes à double courbure. Nyons (1731)
49. Clebsch, A., Gordan, P.: Theorie der Abelschen Functionen. B.G. Teubner, Leipzig (1866)
50. Clifford, W.K.: Elements of Dynamic. Macmillan, London (1878)
51. Davis, D.: Table of immersions and embeddings of real projective space (2011). This is a table of results and references on the Internet: http://www.lehigh.edu/~dmd1/immtable
52. de Rham, G.: Sur l'analysis situs des variétés à n dimensions. Journal de mathématiques pures et appliquées **10**, 115–200 (1931)
53. de Rham, G.: Variétés Différentiables. Hermann, Paris (1960)
54. Desargues, G.: De Brouillon Projet d'une atteinte aux événements des rencontres du Cone avec un Plan. Melchior Tavernier, Paris (1642)
55. Descartes, R.: La Géométrie. Hermann, Paris (1886). Cornell University Library Digital Collections, first published in 1637
56. Dieudonne, J.: Une generalisation des espaces compacts. Journal de Mathématiques Pures et Appliquées **23**, 65–76 (1944)
57. Docquier, F., Grauert, H.: Levisches Problem und Rungescher Satz für Steinscher Mannigfaltigkeiten. Math. Ann. **140**, 94–123 (1960)
58. Dolbeault, P.: Sur la cohomologie des variétés analytiques complexes. C. R. Acad. Sci. Paris **236**, 175–177 (1953)
59. Eilenberg, S., Steenrod, N.: Foundations of Algebraic Topology. Princeton University Press, Princeton (1952)
60. Eliashberg, Y., Gromov, M.: Embedding of Stein manifolds of dimension n into the affine space of dimension $3n/2 + 1$. Ann. Math. **136**, 123–135 (1992)
61. Euler, L.: Solutio problematis ad geometriam situs. Commentarii academiae scientiarum Petropolitanae **8**, 128–140 (1741). Presented to the St. Petersburg Academy 1735. Opera Omnia; Series 1, vol. 7, pp. 1–10, Euler Archive Publication E053, see also Wikopedia: Seven Bridges of Königsberg
62. Euler, L.: Introductio in Analysin Infinitorum. St. Petersburg Royal Academy of Science (1748)
63. Euler, L.: Elementa doctrinae solidorum. Novi Comment. Acad. Sci. Imp. Petropol. **4**, 109–140 (1752). This is E230 in the Euler archive; David Epstein has an interesting overview of Euler's formula for polyhedra: http://www.ics.uci.edu/~eppstein/junkyard/euler/
64. Euler, L.: De integratione aequationius differentialis $\frac{mdx}{\sqrt{1-x^4}} = \frac{ndy}{\sqrt{1-y^4}}$. Novi Comment. Acad. Sci. Petrpol. **6**, 37–57 (1761). Opera Omnia: Series 1, vol. 20, pp. 58–79; English Translation by Stacy G. Langton, Euler 251. www.eulerarchiv.maa.org
65. Euler, L.: Recherches sur la courbure des surfaces. Memoires de l'academie des sciences de Berlin **16**, 119–143 (1767)

66. Euler, L.: Integratio aeqationis $dx/\sqrt{A + Bx + Cx^2 + Dx^3 + Ex^4} = dy/\sqrt{A + By + Cy^2 + Dy^3 + Ey^4}$. Novi Comment. Acad. Sci. Petropol. **12**, 3–16 (1768). Also in: Opera Omnia: Series 1, vol. 20, pp. 302–317; Euler Archive e345; see also the German translation from Latin e345de
67. Euler, L.: Introduction à l'Analyse Infinitésimale. Barrois, Paris (1796)
68. Euler, L.: Institutionum Calculi Integrali, vol. 1. Royal Academy of Sciences, St. Petersburg (1824). 3rd Edition, 1st Edition published 1768, E342 in www.eulerarchive.com with translation by Ian Bruce
69. Euler, L.: Part 2, chapter VI: comparatione quantitatum transcendentium contentarum in forma $\int \frac{Pdx}{\sqrt{a+2bx+cxx+2dx^3+ex^4}}$. Institutionum Calculi Integrali, vol. 1, pp. 365–388. Royal Academy of Sciences, S. Petersburg (1826)
70. Euler, L.: Part 2, chapter V: comparatione quantitatum transcendentium in forma $\int \frac{Pdx}{\sqrt{a+2bx+cxx}}$ contentarum. Institutionum Calculi Integrali, vol. 1, pp. 365–388. Royal Academy of Sciences, S. Petersburg (1826). 3rd Edtion, 1st Edition published 1768; E342 in www.eulerarchive.com with English Translation by Ian Bruce, Concerning the comparison of transcendental quantities contained in the form $\int \frac{Pdx}{\sqrt{A+2B+Cxx}}$
71. Galois, É.: OEuvres mathématematiques d'Évariste Galois. Journal des mathématiques pures et appliquées **11**, 381–444 (1846). This was published posthumously after his early death in 1832
72. Farkas, H.M., Kra, I.: Riemann Surfaces. Springer, New York (1992)
73. Forster, O.: Plongement de variétés de Stein. Comm. Math. Helv. **45**, 170–184 (1970)
74. Fréchet, M.: Sur quelques point du calcul fonctionnel. Rend. Circ. Matem. Palermo **22**, 1–72 (1906)
75. Frenet, F.: Sur les courbes à double courbure. J. Math. Pures Appl. **17**, 437–447 (1852)
76. Fröhlicher, A.: Relations between the cohomology groups of Dolbeault and topological invariants. Proc. Natl. Acad. Sci. USA **41**, 641–644 (1955)
77. Gauss, C.F.: Disquisitiones arithmeticae. Gauss Werke, vol. 1, pp. 1–478. Königlichen Gesellschaft der Wissenschaften, Göttingen (1863). First published in 1801
78. Gauss, C.F.: Theoria residuorum biquadraticorum, commentatio secunda. Gauss Werke, vol. 2, pp. 169–178. Königlichen Gesellschaft der Wissenschaften, Göttingen, : In German (Latin title), originally published in Göttingsche Gelehrte Anzeigen, 1831, V. 1. Apr. **23**, 625–638 (1863)
79. Gauss, C.F.: Auflösung, Allgemeine, der Aufgabe die Theile einer gegebenen Fläche so abzubilden, dass die Abgebildeten in den kleinsten Theilen ähnlich wird. Gauss Werke, vol. 4, pp. 189–216. Königlichen Gesellschaft der Wissenschaften, Göttingen (1873). First published in 1825, Astronomische Abhandlungen, edited by H.C. Schumacher, Altona, 1825
80. Gauss, C.F.: Anzeige: Disquisitiones generales circa superficies curvas. Gauss Werke, vol. 4, pp. 341–352. Königlichen Gesellschaft der Wissenschaften, Göttingen (1873)
81. Gauss, C.F.: Disquisitiones generales circa superficies curvas. Gauss Werke, vol. 4, pp. 217–258. Königlichen Gesellschaft der Wissenschaften, Göttingen, : Originally published Commentationes Societatis Regiae Scientiarum Gottingensis recentiores, vol. VI **1828**, 99–146 (1873)
82. Gauss, C.F.: Letter from Gauss to Bessel. Gauss Werke, vol. 8, pp. 90–91. Teubner, Göttingen (1900)
83. Godement, R.: Théorie des Faisceaux. Hermann, Paris (1965)
84. Grassmann, H.: Die Ausdehnungslehre. Otto Wigand, Leipzig (1844)
85. Grauert, H.: Charakterisieriung der holomorph vollständigen komplexen Räume. Math. Ann. **129**, 233–259 (1955)
86. Grauert, H.: On Levi's problem and the imbedding of real-analytic manifolds. Ann. Math. **68**, 460–472 (1958)
87. Grauert, H.: Über Modifikationen und exzeptionelle analytische Mengen. Math. Ann. **146**, 331–368 (1962)

88. Gray, J.: On the history of the Riemann mapping theorem. Rendiconti del Circolo Matematico di Palermo, Serie II. supplemento N. 34, pp. 47–94 (1994)
89. Green, G.: Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism. T. Wheelhouse, Nottingham (1828). Reprinted in three parts in Journal für die reine und angewandte Mathematik vol. 39, 1 (1850) pp. 73–89; vol. 44, 4 (1852) pp. 356–74; and vol. 47, 3 (1854) pp. 161–221
90. Green, G.: On the determination of the exterior and interior attractions of ellipsoids of variable densities. Trans. Camb. Philos. Soc. 5(3), 395–430: Reprinted in Mathematical Papers of the Late George Green. Macmillan, London **1871**, 187–222 (1835)
91. Griffiths, P., Harris, J.: Principles of Algebraic Geometry. Wiley, New York (1978)
92. Gunning, R.C.: Lectures on Riemann Surfaces. Princeton University Press, Princeton (1966)
93. Gunning, R.C., Rossi, H.: Analytic Functions of Several Complex Variables. Prentice-Hall, Englewood Cliffs (1965)
94. Hadamard, J.: La Série de Taylor, vol. 12. Scientia, Paris (1901)
95. Han, Q., Hong, J.-X.: Isometric Embedding of Riemannian Manifolds in Euclidean Spaces. American Mathematical Society, Providence (2006)
96. Hartogs, F.: Einige Folgerungen aus der Cauchyschen Integralformel bei Funktionen mehrerer Veränderlichen. Sitzungsberichte der Königlich Bayerischen Akademie der Wissenschaften zu München, Mathematisch-Physikalische Klasse **36**, 223–242 (1906)
97. Hausdorff, F.: Grundzüge der Mengenlehre. Veit, Leipzig (1914). Later editions: 1927, 1935, de Gruyter; English translation editions (Set Theory): 1957, 1962, Chelsea
98. Heath, T.L.: Apollonius. Treatise on Conic Sections. Cambridge University Press, Cambridge (1896)
99. Helgason, S.: Differential Geometry. Lie Groups and Symmetric Spaces. Academic Press, New York (1978)
100. Hensel, K., Landsberg, G.: Theorie der algebraischen Funktionen einer Variabeln. B.G. Teubner, Leipzig (1902)
101. Hilbert, D.: Ueber Flächen von constanter Gausscher Krümmung. Trans. Am. Math. Soc. **2**, 87–99 (1901)
102. Hilbert, D.: Über das Dirichletsche Prinzip. Math. Ann. **59**, pp. 161–186 (1904). Also in Hilbert's Gesammelte Abhandlungen, vol. III, pp. 15–37, Springer, 1970
103. Hirzebruch, F.: Arithmetic genera and the theorem of Riemann-Roch for algebraic varieties. Proc. Natl. Acad. Sci. USA **40**, 110–114 (1954)
104. Hirzebruch, F.: Topological Methods in Algebraic Geometry. Springer, New York (1966). Grundlehren der Mathematischen Wissenschaften, vol. 131, 3rd enlarged edition, translation from the German 2nd edition with new appendix by R.L.E. Schwarzenberger; additional section by A. Borel
105. Hirzebruch, F.: Neue Topologische Methoden in der Algebraischen Geometrie. Springer, Berlin (1962). Ergebnisse der Mathematik series: first edition 1955; Second edition with appendix 1962
106. Hodge, W.V.D.: A Dirichlet problem for harmonic functionals, with applications to analytic varieties. Proc. Lond. Math. Soc. **36**, 257–303 (1934)
107. Hodge, W.V.D.: Harmonic functionals in a Riemannian space. Proc. Lond. Math. Soc. **38**, 72–95 (1935)
108. Hodge, W.V.D.: The existence theorem for harmonic integrals. Proc. Lond. Math. Soc. **41**, 483–496 (1936)
109. Hodge, W.V.D.: The Theory and Applications of Harmonic Integrals. Cambridge University Press, Cambridge (1941). Second Edition published in 1952
110. Hodge, W.V.D.: The topological invariants of algebraic varieties. In: Proceedings International Congress of Mathematicians 1950, vol. 1, pp. 182–191. American Mathematical Society (1952)
111. Hörmander, L.: Linear Partial Differential Operators. Springer, New York (1963)
112. Hörmander, L.: An Intoduction to Complex Analysis in Several Variables. Princeton, D. Van Nostrtand Company (1966)

113. Hurwitz, A., Courant, R.: Vorlesungen über Funktionentheorie und Elliptische Funktionen. Springer, Berlin (1964). Fourth Edition with an Anhang by H. Röhrl
114. Huygens, C.: Horologium Oscillatorium. Muget, Paris (1673). Translation into English online by Ian Bruce at www.17centurymaths.com
115. Jacobi, C.G.J.: Demonstratio theorematis ad theoriam functionum ellipticarum spectantis. In: Borchardt, C.W. (ed.) Gesammelte Werke, vol. 1, pp. 39–48. G. Reimer (1881). First published in 1827, Schumacher Astronomische Nachrichten, vol. 6, pp. 37–48
116. Jacobi, C.G.J.: Fundamenta nova theoriae functionum ellipticarum. In: Borchardt, C.W. (ed.) Gesammelte Werke, vol. 1, pp. 49–239. G. Reimer (1881). First published in 1829, Regiomonti. Sumptibus fratrum Bornträger
117. Janet, M.: Sur la possibilité de plonger un espace riemannien donné dans un espace euclidean. Annals de la sociéte polonaise de mathématique **65**, 38–443 (1926)
118. Kähler, E.: Über eine bemerkenswerte Hermitesche Metrik. Abhand. Math. Sem. Hamburg **9**, 173–186 (1933)
119. Kleiman, S.L.: What is Abel's theorem anyway? In: Laudal, O.A., Piene, R. (eds.) The Legacy of Niels Henrik Abel, pp. 395–440. Springer, Berlin (2004)
120. Klein, F.: Zur Theorie der Liniencomplexe des ersten und zweiten Grades. Math. Ann. **2**, 189–226 (1870)
121. Klein, F.: Vergleichende Betrachtungen über neuere geometrische Forschungen. Andreas Deichart, Erlangen (1872). Program zum Eintritt in die philosophische Facultät und den Senat der k. Friederich-Alexanders-Universität zu Erlangen
122. Klein, F.: Gesammelte Mathematische Abhandlungen, Band 1–3. Springer, Berlin (1921). Klein's Collected works were published in three volumes from 1921–1923
123. Klein, F., : In: Rosemann, W. (ed.) Vorlesungen über Nicht-Euklidische Geometrie. Julius Springer, Berlin (1928)
124. Klein, F., Lie, S.: Ueber diejenigen ebenen Curven, welche durch ein gesehlossenes System von einfach unendlich vielen vertauschbaren linearen Transformationen in sich übergehen. Math. Ann. **4**, 50–84 (1871)
125. Kline, M.: Mathematical Thought from Ancient to Modern Times. Oxford University Press, New York (1972)
126. Kodaira, K.: A differential-geometric method in the theory of analytic stacks. Proc. Natl. Acad. Sci. USA **39**, 1268–1273 (1953)
127. Kodaira, K.: On cohomology groups of compact analytic varieties with coefficients in some analytic faisceaux. Proc. Natl. Acad. Sci. USA **39**, 865–868 (1953)
128. Kodaira, K.: On Kähler varieties of restricted type (an intrinsic characterization of algebraic varieties). Ann. Math. **60**, 28–48 (1954)
129. Koebe, P.: Über die Uniformisierung beliebiger analytischer Kurven (2). Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen **633–669**, 1907 (1907)
130. Krantz, S.G.: Function Theory of Several Complex Variables. Wiley, New York (1982)
131. Kuratowski, C.: Sur l'opération $\overline{A}$ de l'analysis situs. Fundamenta Mathematicae **3**, 182–199 (1922)
132. Lancret, M.-A.: Sur les développoïdes des courbes planes, des courbes à double courbure et des surfaces développables. Mémoires de Mathématiques et de Physique, Présenté à l'Institut national de France par divers savans **2**, 416–454 (1806)
133. Lang, S.: Fonctions implicites et plongements riemanniens. Séminaire N. Bourbaki 1961–1962, 245–254 (1962)
134. Laurent, P.-A.: Extension du théorème de M. Cauchy relatif à la convergence du dévelopement d'une fonction suivantes puissances ascendantes de la variable. Comptes Rendu de l'Academie des Sciences **17**, 348–349 (1843)
135. Laurent, P.-A.: Mémoire sur la théorie des imaginaires, sur l'équilibre des températures et sur l'équilibre d'élasticité. Journal de l'École Impériale Polytechnique **23**, 75–204 (1863)
136. Lebesgue, H.: Intégrale, longuer, aire. Annali di Matematica Pura ed Applicata **7**, 231–359 (1902)

137. Lebesgue, H.: Leçons sur l'Integration et la Recherche des Fonctions Primitives. Gauthiers-Villars, Paris (1904)
138. Legendre, A.-M.: Exercices de Calcul Intégral, vol. 1. Courcier, Paris (1811)
139. Legendre, A.-M.: Exercices de Calcul Intégral, vol. 3. Courcier, Paris (1816)
140. Legendre, A.-M.: Exercices de Calcul Intégral, vol. 2. Courcier, Paris (1817)
141. Leibniz, G.W.: Leibniz an Huygens, A Hannover ce 8 de Sept. 1679. In: Gerhardt, C.J. (ed.) Der Briefwechsel von Gottfried Wilhelm Leibniz mit Mathematikern, pp. 567–575. Mauer & Müller (1899)
142. Leray, J.: L'anneau spectral et l'anneau filtré d'homologie d'un espace localement compact et d'une application continue. Journl de Math. Pures et Appl. **29**, 1–139 (1950)
143. Levi, E.E.: Studii sui punti singolari essentiali delle funzioni analitiche di due o più variabili. Ann. di Mat. **17**, 61–88 (1910)
144. Lie, S.: Über die Reziprozitätsverhältnisse des Reyeschen Komplexes. Göttingen Nachrichten **4**, 53–66 (1870)
145. Lie, S.: Theorie der Transformationsgruppen I. Math. Ann. **16**, 441–528 (1880)
146. Lie, S.: Gesammelte Abhandlungen, Band 1–6, mit Anmerkungen. Teubner, Leipzig (1922). Volumes 1–6 with notes (Anmerkungen) in separate volumes were published between 1922 and 1937
147. Liouville, J.: Remarques. Comptes Rendu de l'Académie des Sciences **19**, 1261–1263 (1844)
148. Liouville, J.: Leçons sur les fonctions doublement périodiques faites en 1847. Journal für die reine und angewandte Mathematik **88**, 277–310 (1880)
149. Listing, J.B.: Vorstudien zur Topologie. Vandenhoeck und Ruprecht, Göttingen (1848)
150. Mahoney, M.S.: The mathematical career of Pierre de Fermat, 1601–1665. Princeton University Press, Princeton (1994)
151. Markushevich, A.I.: Introduction to the Classical Theory of Abelian Functions. American Mathematical Society, Providence (1992)
152. Miller, H.R.: Leray in Oflag XXVIIA. Gazette des Mathématiciens **84 suppl**(2000), 17–34 (2000). This volume is entitled Jean Leray (1906–1998)
153. Milne, J.J.: An Elementary Treatise on Cross-Ratio Geometry. Cambridge University Press, Cambridge (1911)
154. Milnor, J.: Differential Topology (1958). Lecture notes at Princeton University, 1958, notes by James Munkres
155. Mittag-Leffler, G.: Om den analytiska framställningen' af en funktion af rationel karakter med en godtyckligt vald *gränspunkt*. Öfversigt af Kongl. Vetenskaps-Akademiens Förhandlingar, pp. 33–43 (1877)
156. Mittag-Leffler, G.: Ytterligare om den analytiska framställningen *af funktioner utaf rationel karakter*. Öfversigt af Kongl. Vetenskaps-Akademiens Förhandlingar, pp. 17–32 (1877)
157. Mittag-Leffler, G.: Sur la represéntation analytique des fonctions monogènes uniformes d'une variable indépendante. Acta Math. **4**, 1–70 (1884)
158. Möbius, A.F.: Der barycentrische Calcul. Johann Ambrosius Barth, Leipzig (1827)
159. Monge, E.G.: Géométrie Descriptive. Bachelier, Paris (1827)
160. Morrey, Jr., C.B.: The analytic embedding of abstract real-analytic manifolds. Ann. Math. **68**, 159–201 (1958)
161. Moser, J.: A new technique for the construction of solutions of nonlinear partial differential equations. Proc. Natl. Acad. Sci. USA **47**, 1824–1831 (1961)
162. Nakano, S.: On complex analytic vector bundles. J. Math. Soc. Jpn. **7**, 1–12 (1955)
163. Narasimhan, R.: Imbedding of holomorphically complete complex spaces. Am. J. Math. **84**, 917–934 (1960)
164. Nash, J.: C^1 isometric imbeddings. Ann. Math. **60**, 383–396 (1954)
165. Nash, J.: The imbedding problem for Riemannian manifolds. Ann. Math. **63**, 20–63 (1956)
166. Newton, I.: Enumeratio linearum tertii ordinis. Opticks, pp. 138–167. Royal Society, London (1704)
167. Newton, I.: Opticks. Royal Society, London (1704)

168. Newton, I.: Tractatus de Quadratura Curvarum. Opticks, pp. 165–211. Royal Society, London (1704)
169. Newton, I.: The Method of Fluxions and Infinite Series. Henry Woodfall, London (1736). Translated by John Colson from unpublished Latin ms of Newton (Methodus Fluxionum et Serierum Infinitarum) from 1671; download available from http://www.archive.org/details/methodoffluxions00newt
170. Newton, I.: Isaac Neswton's Philosophiae Naturalis Principia Mathematica, The Third Edition (1726) with Variant Readings. Cambridge University Press, Cambridge (1972). Edited and annotated by Alexandre Koyré and I. Bernhard Cohen
171. Nirenberg, L.: The Weyl and Minkowski problems in differential geometry in the large. Commun. Pure Appl. Math. **6**, 337–394 (1953)
172. Norguet, F.: Sur les domaines d'holomorphie des fonctions uniformes de plusieurs variables complexes. Bull. Soc. Math. Fr. **82**, 137–159 (1954)
173. Oka, K.: Sur les fonctions analytique de plusieurs variables. VI. Domaines pseudoconvexes. Tôhoku Math. J. **49**, 19–52 (1942)
174. Oka, K.: Sur les fonctions analytique de plusieurs variables. IX. Domains finis sans point critique interior. Jpn. J. Math. **23**, 97–155 (1953)
175. Oka, K.: Sur les Fonctions Analytiques de Plusieurs Variables. Iwanami Shoten, Tokyo (1961)
176. Oka, K.: Sur les fonctions analytiques de plusieurs variables: X une mode nouvelle engendrant les domains pseudoconvex. Jpn. J. Math. **32**, 1–12 (1962)
177. Palais, R.: Seminar on the Atiyah-Singer Index Theorem. Princeton University Press, Princeton (1965). Annals of Mathematics Studies, No. 57
178. Pascal, B.: Oeuvres de Blaise Pascal, vol. 5. Lefèvre, Paris (1819)
179. Peterson, F.P.: Some non-embedding problems. Bol. Soc. Mat. Mexicana **2**(2), 9–15 (1957)
180. Plücker, J.: Analytisch-Geometrische Entwicklungen. Erster Band. G.D. Baedeker, Essen (1828)
181. Plücker, J.: Über eine neue Art, in der analyischen Geometrie durch Puncte und Curven durch Gleichungen darzustellen. Journal für die reine und angewandte Mathematik **1830**(6), 107–146 (1830)
182. Plücker, J.: Analytisch-Geometrische Entwicklungen. Zweiter Band. G.D. Baedeker, Essen (1831)
183. Plücker, J.: On a new geometry of space. Phil. Trans. R. Soc. Lond. **155**, 725–791 (1865)
184. Plücker, J.: Neue Geometrie des Raumes. B.G. Teubner, Leipzig (1868)
185. Poincaré, H.: Analysis situs. Journal de l'École Polytechnique **1**, 1–121 (1895)
186. Poincaré, H.: Sur l'uniformisation des fonctions analytiques. Acta Math. **31**, 1–63 (1907)
187. Poincaré, H.: Oeuvres de Henri Poincaré. Tome VI. Gauthiers-Villars, Paris (1953)
188. Poincaré, H.: Papers on Topology; Analysis Situs and its Five Supplements. American Mathematical Society, Providence (2010). Translation by John Stillwell
189. Poncelet, J.-V.: Applications d'Analyse et de Géométrie. Mallet-Bachelier, Paris (1862)
190. Poncelet, J.-V.: Traité des Propriétés Projectives des Figures, Tome 1, 2nd edn. Gauthiers-Villars, Paris (1865)
191. Poncelet, J.-V.: Traité des Propriétés Projectives des Figures, Tome 2, 2nd edn. Gauthiers-Villars, Paris (1866)
192. Radó, T.: Über den Begriff der Riemannschen Fläche. Acta. Univ. Szeged. **2**(II), 101–121 (1925)
193. Remmert, R.: Sur les espaces analytiques holomorphiquement séparables et holomorphiquement convexes. C. R. Acad. Sci. Paris **243**, 118–121 (1956)
194. Resnikoff, H.L., Wells, Jr., R.O.: Mathematics in Civilization, 3rd edn. Dover Publications, Inc. New York (2015)
195. Riemann, B.: Theorie der Abel'schen Funktionen. Journal für reine und angewandte Mathematik **54**, 115–155 (1857)
196. Riemann, B.: Allgemeine Voraussetzungen und Hülfsmitel für die Untersuchung von Funktionen unbeschränkt veränderlicher Grössen. Journal für reine und angewandte Mathematik **54**, 101–104 (1857)

197. Riemann, B.: Bestimmung einer Function einer veränderlichen complexen Grösse durch Grenz- und Unstetigkeitsbedingungen. Journal für reine und angewandte Mathematik **54**, 111–114 (1857)
198. Riemann, B.: Lehrsätze aus der analysis situs für die Theorie der Integrale von zweigliedrigen vollständigen Differentialien. Journal für reine und angewandte Mathematik **54**, 105–110 (1857)
199. Riemann, B.: Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse. In: Weber, H. (ed.) Gesammelte Mathematische Werke, pp. 3–47. Teubner, Leipzig (1876). Inaugural Dissertation, Göttingen, 1851
200. Riemann, B.: In: Weber, H. (ed.) Gesammelte Mathematische Werke. B.G. Teubner, Leipzig (1876). A 1990 edition with additional notes from numerous commentators on Riemann's work was published by Springer and edited by R. Narasihman
201. Riemann, B.: Ueber die Hypothesen, welche der Geometrie zu Grunde liegen. Gesammelte Mathematische Werke, pp. 254–269. B.G. Teubner, Leipzig (1876)
202. Riemann, B.: Theorie der Abel'schen Functionen. In: Weber, H. (ed.) Gesammelte Mathematische Werke, pp. 81–135. Teubner, Leipzig (1876). This paper in Riemann's Werke combines four papers from the Journal für reine und angewandte Mathematik, vol. 54, 1857
203. Risi, V.: Geometry and Monadology: Leibniz's Analysis Situs and Philosophy of Space. Verlag C.H. Beck im Internet. www.beck.de (2007)
204. Royden, H.L.: Real Analysis, 2nd edn. Collier-Macmillan Ltd., London (1968)
205. Runge, C.: Zur theorie der eindeutigen analytischen functionen. Acta Math. **6**, 229–244 (1885)
206. Russell, B.: History of Western Philosophy. George Allen and Unwin Ltd., London (1946)
207. Sanderson, B.J.: Immersions and embeddings of projective space. Proc. Lond. Math. Soc. **3**(14), 137–153 (1964)
208. Schaefli, L.: Nota alla Memoria del sig. Beltrami, Sugli spazii di curvatura costante. Ann. di Mat. 2nd Ser. **5**, 170–193 (1871)
209. Schwartz, J.T.: On Nash's implicit function theorem. Commun. Pure Appl. Math. **13**, 509–530 (1960)
210. Schwartz, J.T.: Nonlinear Functional Analysis. New York University, New York (1965). Lectures at NYU in 1963–1964
211. Schwartz, L.: Homomorphismes et applications complètement continues. C. R. Acad. Sci. **236**, 2472–2473 (1953)
212. Schwarz, H.A.: Über die Integration der partiellen Differentialgleichung $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ unter vorgeschriebenen Grenz- und Unstetigkeitsbedingungen. Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin **1870**, 767–795 (1870). Published in 1871 by the Berlin Academy of Sciences
213. Serfati, M.: A note on the geometry and descartes's mathematical work. Not. AMS **55**, 50–53 (2008)
214. Serre, J.-P.: Deux théorèmes sur les aplications complètement continue. Seminaire Henri Cartan, tome 6, 1953–1954, exp. 16 **6**, 1–7 (1954)
215. Serre, J.-P.: Un théorème de dualité. Comment. Math. Helv. **29**, 9–26 (1955)
216. Serret, J.A.: Sur quelques formules relatives à la théorie des courbes à double courbure. J. Math. Pures Appl. **16**, 193–207 (1851)
217. Spivak, M.: A Comprehensive Introduction to Differential Geometry, vol. 1, 3rd edn. Publish or Perish, Houston (1999)
218. Spivak, M.: A Comprehensive Introduction to Differential Geometry, vol. 2, 3rd edn. Publish or Perish, Houston (1999)
219. Steenrod, N.: The Topology of Fibre Bundles. Princeton University Press, Princeton (1951)
220. Stein, K.: Anaytische Funktionen mehrerer komplexer Veränderlichen zu vorgegebenen Periodizitätsmoduln und das zweite Cousinsche Problem. Math. Ann. **123**, 201–222 (1951)
221. Thomson, W.: Theorems with reference to the solution of certain partial differential equations. Camb. Dublin Math. J. **3**, 84–87 (1848)
222. Thullen, P., Cartan, H.: Zur Theorie der Singularitäten der Funktionen mehrerer komplexen Veränderlichen. Regularitäts- und Konvergenzberieche. Math. Ann. **106**, 617–647 (1932)

223. van der Waerden, B.L.: Science Awakening. Wiley, New York (1963)
224. Vandermonde, M.: Mémoire sur la résolution des équations. Histoire de l'Académie Royale des Sciences, Paris, pp. 365–416 (1771)
225. Veblen, O., Whitehead, J.H.C.: A set of axioms for differential geometry. Proc. Natl. Acad. Sci. **17**, 551–660 (1931)
226. Veblen, O., Whitehead, J.H.C.: The Foundations of Differential Geometry. Cambridge University Press, Cambridge (1932)
227. Walker, R.J.: Algebraic Curves. Dover Publications, New York (1962)
228. Weierstrass, K.: Zur Theorie der eindeutigen analytischen Functionen. In: Abhandlungen aus der Functionenlehre, pp. 1–52. Verlag von Julius Springer, Berlin (1886)
229. Weierstrass, K.: Mathematische Werke, vol. 1. Mayer & Müller, Berlin (1894). Abhandlungen I
230. Weierstrass, K.: Über continuierliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen. Mathematische Werke von Karl Weierstrass, vol. 2, pp. 71–74. Mayer & Müller, Berlin (1895). Gelesen in der Königl. Akademie der Wissenschaften am 18. Juli 1872
231. Weierstrass, K.: Über das sogenannte Dirichletsche Prinzip. Mathematische Werke, vol. 2, pp. 49–54. Mayer & Müller, Berlin (1895). This paper was read to the Berlin Academy of Sciences in 1870
232. Weierstrass, K.: Mathematische Werke, vol. 2. Mayer & Müller, Berlin (1895). Abhandlungen II
233. Weierstrass, K.: Mathematische Werke, vol. 4. Mayer & Müller, Berlin (1902). Vorlesungen über die Theorie der Abelschen Transcendenten
234. Weierstrass, K.: Mathematische Werke, vol. 3. Mayer & Müller, Berlin (1903). Abhandlungen III
235. Weierstrass, K.: Mathematische Werke, vol. 5. Mayer & Müller, Berlin (1915). Vorlesungen über die Theorie der elliptischen Functionen
236. Weierstrass, K.: Mathematische Werke, vol. 6. Mayer & Müller, Berlin (1915). Vorlesungen über Anwendungen der ellilptischen Functionen
237. Weil, A.: On Picard varieties. Am. J. Math. **74**, 865–894 (1952)
238. Weil, A.: Introduction àl'Etude des Variétés Kählériennes. Hermann, Paris (1958)
239. Wells, Jr., R.O.: Differential Analysis on Complex Manifolds, 3rd edn. Springer, New York (2008)
240. Wessel, C.: On the Analytical Representation of Direction. The Royal Danish Academy of Sciences and Letters, Copenhagen (1999). First published in Danish by the same academy in 1797
241. Weyl, H.: Die Idee der Riemannschen Fläche. B.G. Teubner, Leipzig (1913)
242. Weyl, H.: Die Idee der Riemannschen Fläche. 2nd edn. B.G. Teubner, Stuttgart (1955)
243. Whitney, H., Bruhat, F.: Quelques propriétés fondamentales des ensemble analytiques-réel. Comment. Math. Helv. **33**, 132–160 (1959)
244. Whitney, H.: Analytic extensions of differentiable functions defined in closed sets. Trans. Am. Math. Soc. **36**, 63–89 (1934)
245. Whitney, H.: Differentiable manifolds. Ann. Math. **37**, 645–680 (1936)
246. Whitney, H.: The self-intersections of a smooth n-manifold in $2n$-space. Ann. Math. **45**, 220–246 (1944)
247. Whittaker, E.T., Watson, G.N.: A Course of Modern Analysis. Cambridge University Press, Cambridge (1962)

Index

An index entry which is *italicized* refers to the principal occurrence of a reference to a person or the primary definition or explanation of a mathematical term in the text.

R. O. Wells, Jr., *Differential and Complex Geometry: Origins, Abstractions and Embeddings*, DOI 10.1007/978-3-319-58184-2