Contents

The Authors

Aurélien Alvarez
Université d'Orléans, UFR Sciences, Bâtiment de mathématiques - Route de Chartres, B.P. 6759, 45067 Orléans Cedex 2.

Christophe Bavard
Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex.

François Béguin
Université Paris 13 Nord, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse.

Nicolas Bergeron
Institut de Mathématiques de Jussieu, UMR CNRS 7586, Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris Cedex 05.

Maxime Bourrigan
Département de Mathématiques et Applications, École Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05.

Bertrand Deroin
DMA ENS Paris CNRS UMR 8553, 45 rue d'Ulm, 75005 Paris.

Sorin Dumitrescu
Université Nice Sophia Antipolis, Laboratoire J.-A. Dieudonné UMR 7351 CNRS, Parc Valrose, 06108 Nice Cedex.

Charles Frances
IRMA Université de Strasbourg, 7 rue René Descartes, 67084 Strasbourg Cedex.

Étienne Ghys
Unité de Mathématiques Pures et Appliquées, UMR 5669 CNRS-ENS de Lyon, 46 allée d'Italie, 69007 Lyon.

Antonin Guilloux
Institut de Mathématiques de Jussieu, UMR CNRS 7586, Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris Cedex 05.

Frank Loray
Institut de Recherche Mathématique de Rennes, Campus de Beaulieu, 35042 Rennes Cedex.

Patrick Popescu-Pampu
Université Lille 1, UFR de Maths., Bâtiment M2, Cité Scientifique, 59655, Villeneuve d'Ascq Cedex.

Pierre Py
IRMA, Université de Strasbourg & CNRS, 67084 Strasbourg.

Bruno Sévennec
Unité de Mathématiques Pures et Appliquées, UMR 5669 CNRS-ENS de Lyon, 46 allée d'Italie, 69007 Lyon.

Jean-Claude Sikorav
Unité de Mathématiques Pures et Appliquées, UMR 5669 CNRS-ENS de Lyon, 46 allée d'Italie, 69007 Lyon.

Foreword

Mathematicians have ambiguous relations with the history of their discipline. They experience pride in describing how important new concepts emerged gradually or suddenly, but sometimes tend to prettify the history, carried away with imaginings of how ideas might have developed in harmonious and coherent fashion. This tendency has sometimes irritated professional historians of science, well aware that the development has often been much more tortuous.

It is our implicit belief that the uniformization theorem is one of the major results of 19th century mathematics. In modern terminology its formulation is simple:

Every simply connected Riemann surface is isomorphic to the complex plane, the open unit disc, or the Riemann sphere.

And one can even find proofs in the recent literature establishing it by means of not very complicated argumentation in just a few pages (see e.g. [Hub2006]). Yet it required a whole century before anyone managed to formulate the theorem and for a convincing proof to be given in 1907. The present book considers this maturation process from several angles.

But why is this theorem interesting? In the introduction to his celebrated 1900 article [Hil1900b] listing his 23 most significant open problems, David Hilbert proposed certain "criteria of quality" characterizing a good problem. The first of these requires that the problem be easy to state, and the uniformization theorem certainly satisfies this condition since its statement occupies only two lines! The second requirement — that the proof be beautiful — we leave to the reader to check. Finally, and perhaps most importantly, it should generate connections between different areas and lead to new developments. The reader will see how the uniformization theorem evolved in parallel with the emergence of modern algebraic geometry, the creation of complex analysis, the stirrings of functional analysis, the blossoming of the theory of linear differential equations, and the birth of topology. It is one of the guiding principles of 19th century mathematics. And furthermore Hilbert's twenty-second problem was directly concerned with uniformization.

We should give the reader fair warning that this book represents a rather modest contribution. Its authors are not historians — many of them can't even read German! They are mathematicians wishing to cast a stealthy glance at the past of this so fundamental theorem in the hope of bringing to light some of the beautiful — and potentially useful — ideas lying hidden in long-forgotten papers. Furthermore, the authors cannot claim to belong to the first rank of specialists in modern aspects of the uniformization theorem. Thus the present work is not a complete treatise on the subject, and we are aware of the gaps we should have plugged if only we had had the time.

Our exposition is perhaps somewhat unusual. We don't so much describe the history of a result as re-examine the old proofs with the eyes of modern mathematicians, querying their validity and attempting to complete them where they fail, first as far as possible within the context of the background knowledge of the period in question, or, if that turns out to be too difficult, then by means of modern mathematical tools not available at the time. Although the proofs we arrive at as a result are not necessarily more economical than modern ones, it seems to us that they are superior in terms of ease of comprehension. The reader should not be surprised to find many anachronisms in the text — for instance calling on Sobolev to rescue Riemann! Nor should he be surprised that results are often stated in a much weaker form than their modern-day versions — for example, we present the theorem on isothermal coordinates, established by Ahlfors and Bers under the general assumption of measurability, only in the analytic case dealt with by Gauss. Gauss' idea seems to us so limpid as to be well worth presenting in his original context.

We hope that this book will be of use to today's mathematicians wishing to glance back at the history of their subject. But we also believe that it can be used to provide masters-level students with an illuminating approach to concepts of great importance in contemporary research.

The book was conceived as follows: In 2007 fifteen mathematicians foregathered at a country house in *Saint-Germain-la-Forêt*, Sologne, to spend a week expounding to one another fifteen different episodes from the history of the uniformization theorem, given its first complete proof in 1907. It was thus a week commemorating a mathematical centenary! Back home, the fifteen edited their individual contributions, which were then amalgamated. A second retreat in the same rural setting one year later was devoted to intensive collective rewriting, from which there emerged a single work in manuscript form. After multiple further rewriting sessions, this time in small subsets of the fifteen, the present book ultimately materialized.

We are grateful for the financial support provided by the grant ANR Symplexe BLAN06-3-137237, which made the absorbing work of producing this book feasible.

We wish to thank also Mark Baker, Daniel Bennequin, Catherine Goldstein, Alain Hénaut, Christian Houzel, Frédéric Le Roux, Pierre Mounoud, and Ahmed Sebbar for useful conversations, François Poincaré for translating the Klein–Poincaré correspondence, Arnaud Chéritat and Jos Leys for producing the diagrams, and Marc Deroin and Karim Noui for the electrostatic photographs.

Translator's Note. I am grateful to Edwin Beschler, Manfred Karbe, Norm Purzitsky, Tom Salisbury, Walter Tholen, and especially Frank Loray for various things to do with this translation, carried out at York University, Toronto, and The University of Queensland, Brisbane.

General introduction: The uniformization theorem

The study of plane curves is one of the chief preoccupations of mathematicians. The ancient Greeks investigated in detail straight lines, circles, as well as the conic sections and certain more exotic curves such as Archimedean spirals. A systematic study of general curves became possible only with the introduction of Cartesian coordinates by Fermat and Descartes during the first half of the 17th century [Fer1636, Desc1637], marking the beginning of algebraic geometry. For the prehistory of algebraic geometry the reader may consult [BrKn1981, Cha1837, Die1974, Weil1981].

Two ways of representing a curve

A plane curve can be modelled mathematically in two — in some sense dual — ways:

- by an *implicit equation* $F(x, y) = 0$, where $F : \mathbb{R}^2 \to \mathbb{R}$ is a real function of two real variables;
- as a parametrized curve, that is, as the image of a map $\gamma : \mathbb{R} \to \mathbb{R}^2$.

We shall see that the *uniformization theorem* allows one to pass from the first of these representations to the second. If F is a polynomial, the curve is said to be *algebraic* (formerly such curves were called "geometric"), otherwise *transcendental* (formerly "mechanical"). A significant part of this book is concerned with algebraic curves but, as we shall see, the uniformization theorem in its final version provides an *entrée* into the investigation of (almost) all curves.

Among transcendental curves we find various kinds of spirals and catenaries, the brachistochrone and other tautochrones, which played a fundamental role in the development of mathematics in the 17th century.

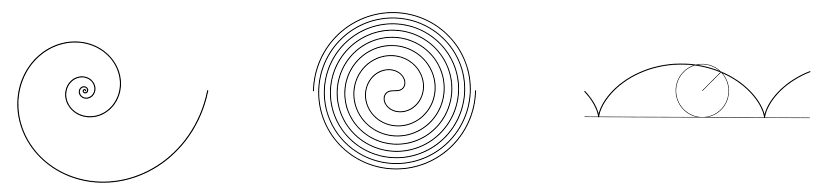


Figure 1: Some transcendental curves

Formerly the study of algebraic curves consisted in a case-by-case examination of a large number of examples of curves with complicated names (lemniscates, cardioids, folia, strophoids, cissoids, etc.) which used to be found among the exercises in undergraduate textbooks and which continue to give pleasure to amateur mathematicians[1].

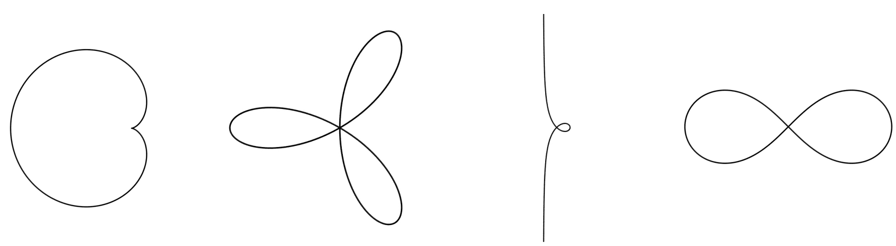


Figure 2: Some algebraic curves

The first invariant that suggests itself for an algebraic curve is the *degree* of the polynomial F, readily seen to be independent of the system of plane (Cartesian) coordinates to which the curve is referred. It is clear that straight lines are precisely the curves of degree 1, and it is not difficult to show that the venerable conic sections of the ancient Greeks are just the curves of degree 2. In a celebrated work [New1704] Newton took up the task of producing a "qualitative" classification of

[1]See e.g. http://www.mathcurve.com/ or http://www.2dcurves.com/.

the curves of degree 3, concluding that there are 72 different types[2]. Evidently it would be difficult if not impossible to continue in this fashion since the number of possible "types" increases very rapidly with the degree and the situation soon becomes impenetrable.

Three innovations

Three major mathematical innovations led to significant clarification of the situation. First came the understanding that the projection from a point in 3-dimensional space of one plane onto another, both situated in that space and neither containing the point of projection, transforms an algebraic curve in one plane into an algebraic curve on the other, moreover of the same degree, said to be *projectively equivalent* to the first. For example, every non-degenerate conic section is the image of a circle under a suitable such projection; hence from the projective point of view the distinction between ellipses, parabolas, and hyperbolas disappears: there now exists just a single equivalence class of non-degenerate conic sections. Similarly, after having defined a *diverging parabola* to be a curve given by an equation of the form $y^2 = ax^3 + bx^2 + cx + d$, Newton states that:

> Just as the circle lit by a point-source of light yields by its shadow all curves of the second degree, so also do the shadows of diverging parabolas give all curves of the third degree.

Here we are at the beginning of *projective geometry*, initiated by Girard Desargues [Desa1639]. Rather than considering a curve $F(x, y) = 0$ in the plane coordinatized by pairs (x, y), one considers it in the *projective plane*, coordinatized by means of homogeneous coordinates $[X : Y : Z]$, where now the curve is given by a homogeneous polynomial $\bar{F}(X, Y, Z) = 0$. Each set of points of the projective plane with $Z \neq 0$ and fixed values for the ratios X/Z and Y/Z corresponds to the point of the affine plane with coordinates $x = X/Z$ and $y = Y/Z$, so that the projective plane is in effect the ordinary affine plane with a *line at infinity* adjoined, each of whose points corresponds to a line through $[0 : 0 : 0]$ in XYZ-space with $Z \equiv 0$ and with $[0 : 0 : 0]$ omitted. It follows that the two branches of a hyperbola in the affine plane join up at two points on the line at infinity, namely the points of that line determined by its two asymptotes, while a parabola is actually tangential to the line at infinity. Thus utilization of projective

[2] Note however that he "missed" 6, his definition of "type" in this context was criticized by Euler, and Plücker, using a different criterion, distinguished 219 types.

geometry simplifies the geometrical picture in a significant way, reducing apparently distinct cases down to an examination of the relative positions of a *projective algebraic curve* and a (projective) line. The adjunction of the line at infinity has other advantages: for instance, every pair of distinct projective lines intersects in a point, which will be at infinity precisely when the corresponding affine lines are parallel.

The second major innovation, dating back to the turn of the 19th century, was *the systematic use of complex numbers in geometry*, leading to the need to consider the complex points of the algebraic curve under investigation, that is, the complex solutions of the equation $F(x, y) = 0$, where furthermore the polynomial $F(x, y)$ is, naturally, now allowed to have complex coefficients. The fact that the field of complex numbers is algebraically closed — awareness of which grew gradually until it was finally established in the 19th century — entails a substantial consolidation of geometrical statements. Clearly projective geometry and complex geometry represent natural enlargements of the original context of the study of plane curves, and indeed until relatively recently were together taken as providing the most natural framework for algebraic geometry.

To take a simple example, the straight line $y = 0$ now meets every "parabola" $y = ax^2 + bx + c$ (with not all of a, b, c zero) in two points. The sign of the discriminant is no longer of any importance — indeed it no longer really has a sign! — but if it vanishes then the two roots merge into one. If $a = 0, b \neq 0$ one of the points is at infinity and if $a = b = 0, c \neq 0$ there is a "double root at infinity"[3]. (The case $a = b = c = 0$ is exceptional.) Thus does one see the unifying power of complex projective algebraic geometry. An even more compelling example concerns the *cyclic points*, which are both imaginary and on the line at infinity. These are just the points $[1 : i : 0]$ and $[1 : -i : 0]$. It is not difficult to see that a conic section in the Euclidean plane is a circle if and only if, considered as a conic in the complex projective plane, it passes through the cyclic points. From this fact many of the properties of circles can be inferred, since they in fact reduce to the question of the position of a conic relative to two points.

Even if we study complex algebraic curves only up to projective coordinate changes, a systematic classification still eludes us except in small degrees. To see this it suffices to note, as Cramer did in 1750, that the vector space of algebraic curves of degree d has dimension $d(d + 3)/2$, while the group of projective transformations has "only" dimension 8 [Cra1750].

The third major innovation, due to Poncelet, Plücker, and Steiner [Ponc1822, Plü1831, Ste1832] among others, rested on the discovery that one can investigate curves by means of *non-linear coordinate changes*. Among such coordinate trans-

[3] To see this rewrite the equation in terms of homogeneous coordinates.

formations *inversion* plays an important role. (Up until the 1960s many chapters of high school geometry textbooks used to be devoted to inversion.) One very simple algebraic version is the transformation (to which the name De Jonquières is attached) sending each point with affine coordinates (x, y) to the point $(1/x, 1/y)$, or, in its "homogenized" variant, mapping the point with projective coordinates $[X : Y : Z]$ to $[YZ : XZ : XY]$. This prompts two remarks. First, this "transformation" σ is not everywhere defined. When two of the homogeneous coordinates are zero — the three such points forming the vertices of a triangle with one side on the line at infinity — the image is not defined (since $[0 : 0 : 0]$ does not correspond to a point of the projective plane). Secondly, the transformation is not injective: the line at infinity $Z = 0$ is sent entirely to the point $x = y = 0$. However, apart from such "details", which hardly bothered our predecessors, this transformation may be regarded as a legitimate change of variables. It is "almost" bijective in view of the fact that it is involutory: if σ is defined both at a point p and its image $\sigma(p)$, then $(\sigma \circ \sigma)(p) = p$. On transforming an algebraic curve via σ we obtain another algebraic curve *but of different degree*. For example, the image of the straight line $x + y = 1$ is the conic $1/x + 1/y = 1$, or, to be precise, a conic with certain points removed.

The non-linear transformations we have in mind form a group (named after Cremona) which is much larger that the projective group, so that one can hope for a precise and at the same time tractable classification of algebraic curves up to such a non-linear transformation. Here we have the beginnings of *birational geometry*, one of Riemann's great ideas. We say that two projective algebraic curves $\bar{F}(X,Y,Z) = 0$ and $\bar{G}(X,Y,Z) = 0$ are *birationally equivalent* if there is a (possibly non-linear) transformation of the form $(X,Y,Z) \mapsto (p(X,Y,Z), q(X,Y,Z), r(X,Y,Z))$ where the coordinates p, q, r are homogeneous polynomials of the same degree, which maps the first curve "bijectively" to the second. Here the quotation marks are meant to indicate that, as in the above example, the transformation may not be defined everywhere. One insists only that each of the two curves has a finite set of points such that the transformation sends the complement of the finite subset of the first curve bijectively to the complement of the subset of the second.

A signal virtue of birational transformations is that they allow us to avoid the problem of *singular points*. Early geometers were soon confronted with the need to study double points, cusp points, etc. In the real domain the theory of such points is relatively simple, at least in its topological aspects. Every point of a real algebraic curve has a neighbourhood in which the curve is made up of an *even* number of arcs. Such a curve cannot have an end-point, for instance.

On the other hand for complex algebraic curves, local analysis of their singular points has established that they can have an extraordinarily intricate structure: investigations begun by Newton and continued by Puiseux [New1671, Pui1850, Pui1851] show that their topological structure is connected with the theory of knots, which theory does not, however, come within the compass of the present book. For us it suffices to know that every algebraic curve is birationally equivalent to a curve possessing only especially simple singular points, namely ordinary double points (Noether, Bertini [Noe1873, Bert1882]) — in other words, points in a neighbourhood of which the curve consists of two smooth arcs with distinct tangents.

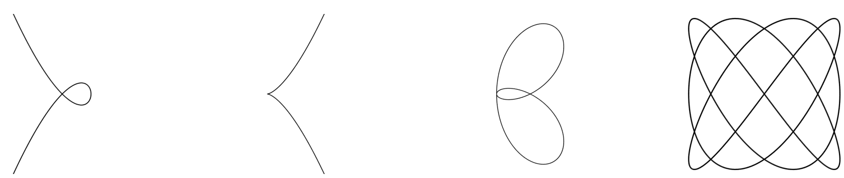


Figure 3: Some types of singular points

To summarize, geometers have progressively reduced the study of algebraic plane curves to that of algebraic curves which, to within a birational transformation, have only ordinary double points.

Rational curves

The introduction of complex numbers had consequences far beyond projective geometry: the beginning of the 19th century also witnessed the advent of the geometric theory of holomorphic functions, which are at one and the same time functions of a single complex variable and of two real variables. Gauss not only knew that it is useful to coordinatize the plane with the complex numbers, but understood equally well that any surface in space can be coordinatized by the complex numbers *conformally* (see Chapter I). Thus a surface is locally determined by a single number. The step had been taken: *a real surface can be considered a complex curve.* Some thirty years later Riemann understood that there is, reciprocally, some advantage in regarding a complex curve as a real surface (see Chapter II).

We are now in a position to address the question of *parametrized curves*. A curve is called *rational* if it is birationally equivalent to a straight line. (Formerly such curves were called *unicursal*, meaning that they could be "traced out with a single stroke of the pen".) More concretely, a curve $F(x, y)$ is rational if it can be parametrized by means of rational functions

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

where p, q and r are polynomials in a single (complex) variable t, and the parametrization is a bijection outside a finite subset of values of t. Here are some simple examples.

Non-degenerate conics are rational. To see this it suffices to take a point m on the conic C and a projective line D not passing through m (see Figure 4). Then for each point t on D, the line determined by m and t meets the conic in two points, of which one is of course m. Denoting the other point by $\gamma(t)$, one readily checks that the map $\gamma : D \to C$ is a birational equivalence.

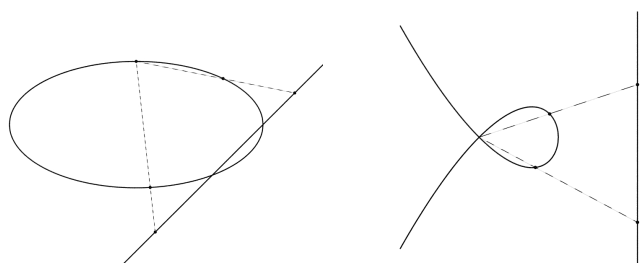


Figure 4: Parametrization of a conic and a singular cubic

A cubic curve with a double point is also a rational curve. For this it suffices to choose a straight line not passing through the singular point, and consider for each point p of that line the line through p and the singular point (see Figure 4). Each such line meets the conic in three points, two of which coincide with the double point of the cubic. The third point of intersection then determines a birational equivalence between the initially chosen line and the given cubic. For example, the origin is a double point of the curve $y^2 = x^2(1 - x)$. We choose $x = 2$ as our parametrizing line. The line passing through the origin and the point $(2,t)$ has equation $y = tx/2$, so intersects the given cubic where $t^2x^2/4 = x^2(1 - x)$, which has the expected double root $x = 0$ and the third solution $x = 1 - t^2/4$, yielding the desired rational parametrization $y = t(1 - t^2/4)/2$ of the curve.

Although rational curves are of considerable interest, they represent just a small proportion of all algebraic curves. We do not know exactly when mathematicians became fully aware of this, that is, of the fact that most algebraic curves are not rational. There are several elementary means of convincing oneself of it, and later on we shall give a topological argument rendering it "obvious". Or one can argue as follows. Note first that a curve given in the form $x = p(t)/r(t), y = q(t)/r(t)$ is of degree d where d is the largest of the degrees of the polynomials p, q, r: one can see this by counting the number of points of intersection with a generic straight line, which points will be given as the solutions of an equation of degree d. The vector space of triples of polynomials of degree d has dimension $3(d + 1)$. However multiples of p, q, r by any non-zero scalar yield the same curve, and replacement of t by a suitable rational function of t (depending on at least three parameters) will also leave the curve unchanged. Thus the space of rational curves of degree d depends on at most $3d - 1$ parameters. As noted earlier, a count of the number of coefficients of a polynomial of degree d in two variables yields $d(d + 3)/2$ for the number of parameters. Since $d(d + 3)/2 > 3d - 1$ for $d \geq 3$, we conclude that *in general algebraic curves of degree at least three are not rational curves.*

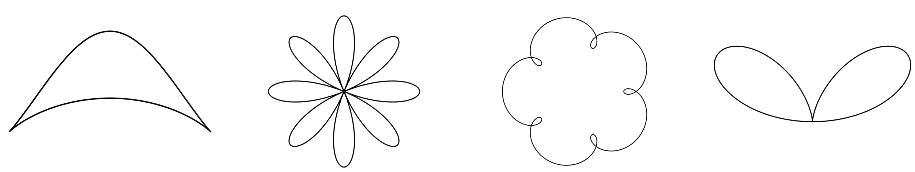


Figure 5: Some rational curves

Elliptic curves

It is completely natural that effort should first be concentrated on the cubics. As we have seen, Newton himself produced an initial classification which was neither projective nor complex, even though he found hints of certain features of projectivity and the complex numbers. His aim was to understand in some fashion the possible topological dispositions of cubic curves in the plane: the positions of asymptotes, singular points, etc. We saw above that singular cubics are rational. However non-singular cubics are never rational; we recommend that the reader attempt to prove this by elementary means.

We limit ourselves here to a brief overview of the principal results. First, every smooth cubic curve is projectively equivalent to a curve with equation in the following normal form (named for Weierstrass although it should properly be attributed to Newton):

$$y^2 = x^3 + ax + b,$$

where a, b are complex numbers. If $4a^3 + 27b^2 \neq 0$, this cubic is smooth. From the inception of the theory mathematicians struggled to evaluate integrals of the form

$$f(x) = \int \frac{dx}{y} = \int \frac{dx}{\sqrt{x^3 + ax + b}}.$$

They called such integrals "elliptic" since evaluation of the length of an arc of an ellipse leads to such a formula. Difficulties arise when one tries to make sense of such integrals with x and y allowed to be complex. The first problem is that the value of the integral depends on which square root one chooses for the denominator of the integrand. The second, linked to the first, consists in the dependence of the integral on the path of integration. Faced with these difficulties, one is forced to the conclusion that one must resign oneself to regarding f as a "many-valued[4] function", that is, that each point x may have several images, all denoted by $f(x)$ however — a situation somewhat distasteful to present-day mathematicians, brought up as they are on the modern set-theoretic definition of a function.

Gauss, Abel, and Jacobi conceived the ingenious idea (to be expounded in Chapter I) that it is not so much the function f that is of interest but its *inverse*. They were perhaps led to this by the analogy with the circle

$$x^2 + y^2 = 1$$

(which is certainly a rational curve) and the integral

$$\int \frac{dx}{y} = \int \frac{dx}{\sqrt{1 - x^2}} = \arcsin x.$$

The "function" arcsin as so defined is multivalued, but it is the inverse function of sin, a function in the strict sense of the word — each point x has a uniquely defined image $\sin x$. The many-valuedness of arcsin arises from the *periodicity* of the sine function, and in like manner the inverse $\wp$ of f is a "genuine" meromorphic

[4]The French term is "multiforme".

function (to emphasize that the modern interpretation of the word "function" is the one intended, the adjective *single-valued*[5] is sometimes used), and the fact that f is many-valued is explained by the periodicity of the single-valued function $\wp$.

It is important to stress this periodicity. While the sine function is periodic of period 2π, the periodicity of the meromorphic function $\wp$ is much richer: it has *two* linearly independent periods. In more precise terms, there is a subgroup Λ of $\mathbb{C}$ of rank 2 (depending on a and b) such that

$$\forall \omega \in \Lambda, \ \ \wp(z) = \wp(z + \omega).$$

(In fact the elements of Λ are just the integrals of dx/y around closed curves in the x-plane.) It follows that we may regard $\wp$ as defined on the quotient of $\mathbb{C}$ by the lattice Λ. Topologically, the quotient space $\mathbb{C}/\Lambda$ is a 2-dimensional torus. Locally, each point of the torus is associated with a complex number in such a way that it inherits the structure of a holomorphic manifold of complex dimension 1, an example of a *Riemann surface* (see Chapter II).

Since $\wp$ is periodic, its derivative $\wp' = d\wp/dz$ is also, and we then obtain a map $(\wp, \wp')$ from the Riemann surface $\mathbb{C}/\Lambda$ with the poles of $\wp$ and $\wp'$ removed, to $\mathbb{C}^2$. It is not difficult to prove[6] that this map extends from $\mathbb{C}/\Lambda$ to the original cubic curve in the complex projective plane (with the three excluded points restored, now sent to three points at infinity). In this way one obtains an identification of the projective cubic curve and the torus $\mathbb{C}/\Lambda$.

A few remarks are apropos. First, it now becomes *topologically* clear why such cubics are not rational: a complex projective line is homeomorphic to a sphere (the Riemann sphere) and the removal of finitely many points will not make it homeomorphic to a torus.

We also see from the above that every smooth cubic, considered as a real surface (in the complex projective plane), is homeomorphic to the torus. On the other hand, considered as Riemann surfaces, these tori are not holomorphically equivalent to one another: given two distinct lattices Λ_1, Λ_2 in $\mathbb{C}$, there is in general no holomorphic bijection between $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$. (There is such a bijection if and only if $\Lambda_2 = k\Lambda_1$ for some non-zero k.) Hence in contrast with rational curves, which are all parametrized by the complex projective line (the Riemann sphere), smooth cubics are not all parametrized by the same complex torus $\mathbb{C}/\Lambda$: each of them is parametrized by a complex parameter (determined by a lattice in $\mathbb{C}$ defined to within a homothety[7]), called a *modulus*.

---

[5]The French term is "uniforme". "Uniformization" is thus the process of representing many-valued functions by single-valued ones. *Translator*

[6]Using the fact that $(\wp')^2 = \wp^3 + a\wp + b$. *Trans*

[7]That is, defined to within a similarity.

The reader may now take the measure of the progress achieved since Newton's attempt at classification: the birational equivalence classes of smooth cubic curves also depend on a single complex parameter.

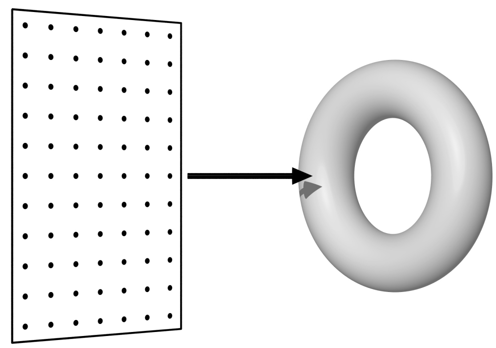


Figure 6: Uniformization of an elliptic curve

Even though the domain $\mathbb{C}/\Lambda$ of the parametrization of a smooth cubic depends on the cubic, it should be noted that the universal cover of $\mathbb{C}/\Lambda$, the complex line $\mathbb{C}$ considered as a Riemann surface, is in fact independent of the cubic. We shall now elaborate on this point — at the expense of perpetrating an anachronism since the concept of the universal cover evolved only gradually in the course of the 19th century, and reached final form only in the 20th. (In this connection one should also mention that some of the motivation for the development of topology came from the study of curves.) A topological space X is said to be *simply connected* if every loop $c : \mathbb{R}/\mathbb{Z} \to X$ can be contracted to a point, that is, if there is a continuous family of loops $c_t, t \in [0,1]$, with $c_0 = c$ and c_1 a constant loop. It can be shown that provided X is a "reasonably well-behaved space" — which is certainly the case for manifolds — there exists a simply-connected space $\tilde{X}$ and a projection map $\pi : \tilde{X} \to X$ whose fibres are the orbits of a discrete group Γ acting on $\tilde{X}$ (fixed point) freely and properly[8]. The space $\tilde{X}$ is then called *the universal covering space* of X, and Γ *the fundamental group* of X. In the case where X is the torus $\mathbb{C}/\Lambda$, it is obvious from its very construction that its universal cover is $\mathbb{C}$ and its fundamental group is Λ, which is isomorphic to the group $\mathbb{Z}^2$. When X is endowed with the additional structure of a Riemann surface, such a structure

[8]That is, with the map $\Gamma \times \tilde{X} \to \tilde{X} \times \tilde{X}$ given by $(g, x) \mapsto (gx, x)$ proper, meaning that complete inverse images of compact sets are compact.

is naturally induced on its universal cover, most often non-compact, so from the above it follows that the universal cover of every non-singular cubic curve is isomorphic to the complex line $\mathbb{C}$. Thus even though the isomorphism classes of smooth cubic curves depend on a modulus, their universal covers are all isomorphic. We summarize this, bringing in for the first time the term “uniformization”:

Every smooth cubic curve C in the complex projective plane has a holomorphic uniformization $\pi : \mathbb{C} \to C$ which parametrizes the curve in the sense that two points of $\mathbb{C}$ have the same image under π if and only if their difference belongs to a certain lattice Λ of $\mathbb{C}$.

And the converse rounds off the theory into a harmonious whole: corresponding to each lattice Λ of $\mathbb{C}$, there exists a smooth cubic curve that is holomorphically isomorphic to $\mathbb{C}/\Lambda$.

Beyond elliptic curves

Our Chapter II constitutes an invitation to read the papers of Riemann devoted to algebraic functions and their integrals. These texts, so important for the history of mathematics, are difficult of access, and it took a considerable time for them to be finally assimilated. Although there are historical articles commenting on these, our approach is quite different, in particular in not at all attempting to be exhaustive. Riemann’s great contribution was to turn Gauss’ idea on its head: although it is useful to think of real surfaces as complex curves, it turns out to be more fruitful to think of a complex curve — with equation $P(x, y) = 0$, say — as a real surface. It is on this that Riemann bases his theory of surfaces, in which one-dimensional and two-dimensional notions become associated with one another. For example, he makes no bones about cutting a surface along a real curve, thereby introducing topological methods into algebraic geometry. Regarding an algebraic curve — that is, an object of one complex dimension situated in the complex projective plane — as a real two-dimensional surface presents no difficulties if the given curve is smooth, since then the real surface is also smooth. However, as we have already seen, this is far from representing the general situation since singular points arise frequently. In this case, however, one can to within a birational equivalence assume that the singularities are ordinary double points, and then it is not difficult to make the surface smooth: for this it suffices to regard the double point as actually two distinct points, on separate branches, and one constructs in this way a smooth surface associated with the original algebraic curve. This is how Riemann associates with each given algebraic curve a Riemann surface, that is, a holomorphic manifold of dimension 1, or, to put it another way, a real manifold of dimension 2 endowed with a complex structure. (We shall revisit this theme throughout the book.) Riemann went on to (almost) prove the following two statements:

— *Two algebraic curves are birationally equivalent if and only if their associated Riemann surfaces are holomorphically isomorphic.*

— *Every "abstract" compact Riemann surface is holomorphically isomorphic to the Riemann surface of some algebraic curve.*

Thus the algebraic problem of describing algebraic curves is transformed into the transcendental one of describing Riemann surfaces. The first invariant derived by Riemann was a purely topological one (and had a major impact on the development of topology since, among other things, it was in attempting to generalize it that Poincaré was led to the modern form of that discipline). It is well known that every compact orientable surface is homeomorphic to a sphere with a certain number of handles attached, which number is nowadays termed the *genus* of the surface. It follows that every algebraic curve has a specific associated genus which is invariant with respect to birational equivalence and so of much greater significance than the degree.

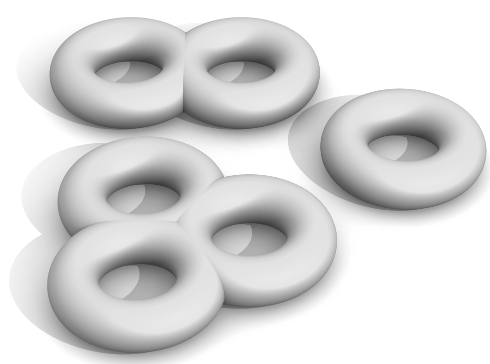


Figure 7: Topological surfaces of genus 1, 2, and 3

Here are some of the results concerning the genus that we shall encounter later on.

Having genus zero means that the curve's associated Riemann surface is homeomorphic to the 2-sphere. It does not then follow immediately that it is holomorphically isomorphic to the Riemann sphere. This fact was established in two different ways by Alfred Clebsch (see Chapter II) and Hermann Schwarz (see Chapter IV): every Riemannian metric on the sphere is globally conformally equivalent to that of the standard sphere. In other words (closer to those of Schwarz) every Riemann surface homeomorphic to the sphere is holomorphically equivalent to the Riemann sphere. In yet other words:

The algebraic curves of genus zero are precisely the rational curves.

This represents a further stage on the way to general uniformization: a single topological datum about a curve determines whether or not it has a rational parametrization.

Having genus 1 signifies that the Riemann surface is homeomorphic to a torus of two real dimensions. It follows, although not obviously — Clebsch proved it in 1865 — that it is holomorphically isomorphic to a quotient of the form $\mathbb{C}/\Lambda$ (see Chapter II). Thus:

The algebraic curves of genus 1 are precisely those birationally equivalent to smooth cubics (the so-called "elliptic" curves).

The case of genus greater than or equal to 2 is more complicated, and it is to this case that the present book is devoted. Before summarizing the situation, we clarify the connection between genus and degree: it can be shown that if C is a curve of degree d with k singular points, all ordinary double points, the genus is given by the formula

$$g = \frac{(d-1)(d-2)}{2} - k.$$

It is then immediate that straight lines and conics have genus zero, smooth cubics genus 1, singular cubics genus zero, and smooth quartics genus 3.

Riemann demonstrates great mastery by the manner in which he generalizes from the case of elliptic curves. For instance, for each fixed value $g \geq 2$ of the genus, he seeks to describe the space of moduli of the curves of that genus — that is, the space of algebraic curves of genus g considered to within a birational transformation — showing it has complex dimension $3g - 3$. Among other results of Riemann, we should also mention the celebrated one asserting that every non-empty simply connected open subset of $\mathbb{C}$ is biholomorphically equivalent to the open unit disc — a result of fundamental importance, although Riemann's proposed proof leaves a little to be desired (see Chapter II). It sometimes happens that this result, albeit an important special case, is confused with the "great" uniformization theorem forming the theme of the present book, which has to do not just with open sets of $\mathbb{C}$ but, much more impressively, with *all* Riemann surfaces.

Riemann's work in this field exerted a considerable influence on his immediate successors. In Chapter IV we describe Schwarz's attempts to establish explicitly certain particular cases of the conformal representation theorem while skirting the technical difficulties on which Riemann's proof founders.

Among the best expositions of Riemann's ideas, that of Felix Klein, another hero of the present work, stands out. In 1881 he wrote up what he believed to be the idea behind Riemann's intuition, even though Riemann's actual articles make no mention of it. We will never know if Klein was right in this, but the resulting new approach, via Riemannian metrics, seems to us especially illuminating. It relies on an electrostatic or perhaps hydrodynamic interpretation, making it particularly accessible to the intuition. We describe this way of looking at the subject and its modern developments in Chapter III.

Uniformizing algebraic curves of genus greater than 2

The question of parametrizing curves of general genus g remained open, or, more precisely, no one suspected that every algebraic curve might be parametrizable by *single-valued* holomorphic functions. However, following Riemann's work, evidence for this began to accumulate from the examination of certain remarkable examples.

In a marvellous article Klein studied the curve C given by the homogeneous equation $x^3y + y^3x + z^3x = 0$ as a Riemann surface, showing that it is isomorphic to the quotient of the upper half-plane by an explicit group of holomorphic transformations. In other words, he constructed a (single-valued) holomorphic function π with domain the upper half-plane $\mathbb{H}$ and with fibres the orbits of a group Γ of holomorphic transformations acting freely and properly. The analogy with the situation of elliptic curves was striking: the half-plane replaces the complex line and the group Γ of Möbius transformations replaces the group Λ acting via translations. *Thus is Klein's quartic uniformized by* π.

Even though this remarkable specimen was actually the first example of uniformization in higher genus, it was nonetheless taken at the time for an unparalleled gem, as it were, incapable of generalization like the regular polyhedra. As such it marked an interlude prior to attempts at establishing general uniformization. We shall expound Klein's example in Chapter V.

Motivated by quite different considerations arising in the theory of linear differential equations, Poincaré was led to the systematic investigation of the discrete subgroups Γ of the group $\mathrm{PSL}(2,\mathbb{R})$, which he called *Fuchsian*, and the quotients $\mathbb{H}/\Gamma$ obtained from them. He saw that among such quotients there are compact Riemann surfaces of genus at least 2. He showed that there is some latitude in the choice of the group, depending on certain parameters (see Chapter VI).

In light of Poincaré's results, Klein realized that the algebraic curves uniformized by $\mathbb{H}$ are in fact not isolated examples as he had thought, but form continuous families depending on parameters to be determined. Almost simultaneously Klein and Poincaré saw that the latter's constructions might be of sufficient flexibility for *all* compact Riemann surfaces to be uniformizable by $\mathbb{H}$. A dimensional count quickly showed that the space of Poincaré's Fuchsian groups, considered up to conjugation, yielding a surface of genus g depends on $6g-6$ real parameters — highly suggestive given Riemann's result that Riemann surfaces of genus g depend on $3g-3$ complex moduli. The race was on between Klein and Poincaré to prove the theorem. We encourage the reader to read the impassioned correspondence on this topic between our two heros reproduced at the end of the book. Here Klein and Poincaré introduce a new method of proof, namely *by continuity*.

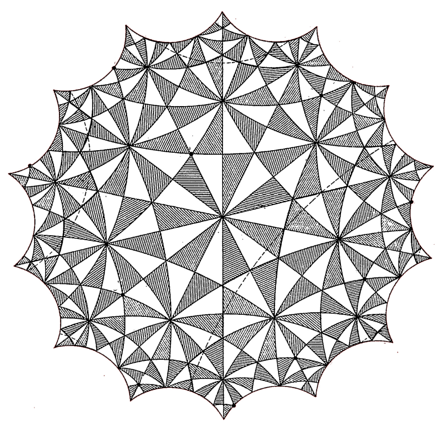


Figure 8: Klein's Fuchsian group (shown here as a group of automorphisms of the unit disc rather than the upper half-plane).

To us neither Klein's proof nor Poincaré's is totally convincing. In Chapter VII we try to resurrect Klein's proof[9]; to obtain a rigorous proof we had to use modern tools derived from quasiconformal techniques, which Klein and Poincaré certainly did not have at their disposal. Then in Chapter VIII we make an attempt to resuscitate — at least in part — Poincaré's approach, which was not motivated

[9]The matter is actually more complex; in fact some parts of the proof given in Chapter VII are closer to certain of Poincaré's arguments than to those of Klein.

by uniformization but rather by the desire to solve linear differential equations. The reader will observe there the emergence for the first time of a great number of concepts familiar to modern mathematicians. Chapter IX is devoted to the explicit investigation of some examples of uniformization of surfaces of higher genus.

By 1882 Klein and Poincaré had become fully convinced of the truth of the following uniformization theorem:

Theorem. *Let X be any compact Riemann surface of genus* ≥ 2. *There exists a discrete subgroup* Γ *of* $PSL(2,\mathbb{R})$ *acting freely and properly on* $\mathbb{H}$ *such that X is isomorphic to the quotient* $\mathbb{H}/\Gamma$. *In other words, the universal cover of X is holomorphically isomorphic to* $\mathbb{H}$.

To summarize, Klein and Poincaré had now effectively solved one of the main problems handed down by the founders of algebraic geometry: to parametrise an algebraic curve $F(x, y) = 0$ (of genus at least 2) by single-valued meromorphic functions $x, y : \mathbb{H} \to \mathbb{C}$. This magnificent result rounded out the theory dealing with the particular cases of rational and elliptic curves. Thus *Fuchsian functions* were now seen to be the appropriate generalizations of elliptic functions. Of course, as in the case of the elliptic functions, it was now necessary to admit new transcendental functions into the menagerie of basic mathematical objects, find their (convergent) series representations, etc. In fact Poincaré subsequently devoted a number of papers to such questions.

Beyond algebraic curves

But why should we confine ourselves to algebraic curves? What is the situation with "transcendental" curves? Spurred by his success with algebraic curves, Poincaré went on to address the problem of non-compact Riemann surfaces, which *a priori* have no relation to algebraic geometry. Although the method of continuity could no longer be applied, nonetheless already by 1883 Poincaré had managed to show that every Riemann surface admitting a non-constant meromorphic function can be uniformized in a certain weakened sense of the word "uniformize": one has now to allow parametrisations that may not be locally injective, that is, with ramification points. This result is the subject of Chapter XI. The question of the uniformization of non-algebraic surfaces seems to have stagnated for a while thereafter, until, in 1900, Hilbert stressed the incomplete nature of Poincaré's result, and encouraged mathematicians to re-apply themselves to it; this was Hilbert's twenty-second problem. At last, in 1907, Poincaré and Koebe arrived independently at the general uniformization theorem:

Theorem. *Every simply connected Riemann surface is holomorphically isomorphic to the Riemann sphere* $\bar{\mathbb{C}}$, *the complex plane* $\mathbb{C}$, *or the upper half-plane* $\mathbb{H}$.

Koebe's and Poincaré's approaches to this theorem are described in Chapters XII and XIII.

Of course, this classification of simply connected Riemann surfaces yields immediately a characterisation of all Riemann surfaces, since every Riemann surface is a quotient of its universal cover by a group acting holomorphically, freely, and properly. Thus by the theorem of Koebe and Poincaré every Riemann surface is identical with either the Riemann sphere or a quotient of $\mathbb{C}$ by a discrete group of translations, or a quotient of the half-plane $\mathbb{H}$ by a Fuchsian group. The work of Poincaré and Koebe, occupying Part C, allowed a new page to be turned in potential theory, and represents the end of an important epoch in the history of mathematics.

Meanwhile, over the decade 1890–1900, Picard and Poincaré worked out a new proof of the uniformization theorem based on a suggestion by Schwarz, valid in the compact case at least, and depending of the solution of the equation $\Delta u = e^u$. We present this in Chapter X.

The uniformization theorem was at the centre of the evolution of mathematics in the 19th century. In the diversity of its algebraic, geometric, analytic, topological, and even number-theoretic aspects it is in some sense symbolic of the mathematics of that century.

Our book ends in 1907, even though the story of the uniformization theorem continues. Among later developments, one might mention Teichmüller's work on moduli spaces, or those of Ahlfors and Bers in the 1960s relating to the concept of quasiconformal mappings (see for example [Hub2006]). There is also the progress in higher dimensions, in particular Kodaira's classification of complex surfaces, that is, of 2 complex dimensions. But that's another story!

Part A

Riemann surfaces

Chapter I

Antecedent works

Any account of the evolution of the uniformization theorem must begin with a description of the methods of Riemann and his immediate successors[1]. The aim of this first part is to provide such a description.

Of all mathematicians of the middle of the 19th century, it was without doubt Riemann who left the deepest imprint on the theory of algebraic curves. Here, for example, are the first few sentences of Hermite's preface to Riemann's complete works:

> Bernhard Riemann's *oeuvre* is the greatest and the most beautiful of our era: it has received unanimous acclaim and will have a permanent influence on scientific development. The work of present-day geometers is inspired by his ideas whose significance and fruitfulness are reconfirmed every day in their discoveries.

In this preliminary chapter we give a succinct exposition of two topics which in Riemann's time (around 1851) had emerged quite recently and which may well have served as "detonators" for his work on algebraic curves:

— Gauss' application of complex numbers to *cartography*, and the "local" uniformization theorem allowing local parametrization of any surface by a "conformal map".

— the rise of the theory of *elliptic functions*, initiated by Euler and reaching maturity with the work of Abel and Jacobi just prior to Riemann's thesis.

However before discussing cartography and elliptic functions, we consider very briefly the birth of the geometric interpretation of the complex numbers as points of the plane.

[1]Even though the correspondence between Klein and Poincaré reproduced at the end of the present book shows clearly that when Poincaré began his investigations of Fuchsian functions he had not read Riemann!

I.1. On the development of the complex numbers

The story of the complex numbers is a rather involved one and there are many detailed histories devoted to them, such as, for instance [Mar1996, Neue1981]. Our present aim is certainly not to recount their history, but rather to recall just a few of the more important stages in their development in order for the reader to appreciate more fully the innovatory character of the works of Gauss, Abel, and Jacobi described below. (For additional details we refer the reader to [Mar1996], pp. 121–132.)

Although in 1777 Euler had indeed coordinatized the points of the plane with complex numbers $x + iy$, this geometric interpretation received its full formalization only at the turn of the 19th century (by Wessel in 1799, and Argand and Buée in 1806), and it took some time before the geometric point of view was taken for granted.

Of course, Gauss understood many things before anyone else... His first "proof" of the Fundamental Theorem of Algebra, in 1799, cannot be understood without an appreciation of the geometric and topological way of viewing the complex numbers[2]. According to [Mar1996] it was only following the publication of Gauss's 1831 article "Theoria residuorum biquadraticorum" that the notion of a complex number as a point of the plane gained universal recognition.

The theory of analytic, or holomorphic, functions also took a long time to crystallize out, at least in its geometric aspect. Here the great instigator in the development of the theory was Cauchy. According to [Mar 1996], the path he followed was long and tortuous. In 1821 he was still talking of imaginary expressions: "An imaginary equation is merely a symbolic representation of two equations in two variables." It took till 1847 for him to largely shed such terminology, speaking instead of "geometric quantities", and reach the point of conceiving a function visually as we do today, that is, as transforming a variable point in the input plane to another variable point in the output plane.

The concept of the complex integral $\int f(z)dz$ along a path, the dependence of the integral on the path, and residue theory: all of these familiar results also underwent a long period of gestation, primarily at the hands of Cauchy. His first paper on these questions appeared in 1814 but the theory of residues dates from 1826–1829.

Here also Gauss was ahead of his time, but refrained from publishing his ideas. A letter from him to Bessel dating from 1811 shows that he had a clear idea of the integral along a path and that he had already grasped the concept of the residues at the poles of the functions being integrated.

[2]In order to show that a non-constant polynomial P vanishes somewhere in the complex plane, he studies the behavior towards infinity of the curves $\operatorname{Re} P = 0$ and $\operatorname{Im} P = 0$, deducing that they must of necessity cross.

In sum, in 1851 Riemann has at his disposal a geometric theory of holomorphic functions just recently created. By introducing the concept of a Riemann surface, he will now liberate holomorphic functions from the coordinates x and y, and the theory will assume a fundamentally geometric form. By contrast, twenty-five years earlier, Abel and Jacobi had none of the basic concepts of complex function theory — such as for example Cauchy's residue formula — at their disposal.

I.2. Cartography

The pursuit of the science of cartography, both terrestrial and celestial, led scholars of antiquity to pose the question as to how a portion of a sphere might be represented by a planar map. Ptolemy's *Geography* contains several possible solutions. It soon became clear that distortions are inevitable, whether of shapes, distances, areas, etc.

In 1569 Mercator proposed a projection which he used to produce a map of the world with properties especially convenient for navigation. Although his method of drawing the map was empirical, the underlying idea nonetheless paved the way for the application of mathematical analysis to cartography. It was in the 18th century that these two disciplines came together in a series of works by Johann Heinrich Lambert, Leonhard Euler, and Joseph Louis Lagrange. Lambert's work, published in 1772, heralded the birth of modern mathematical cartography. According to Lagrange, Lambert was the first to formulate the basic problems associated with the representation of a region of the sphere on a plane in terms of certain partial differential equations.

In 1822, inspired by cartographical problems and methods, the Royal Society of Copenhagen set as the subject of a prize essay the problem of "representing parts of a given surface on another surface in such a way that the representation be similar to the original in infinitely small regions". This was a prime opportunity for Gauss, greatly interested as he was in both the theory and practice of cartography, to prove the existence of a locally conformal representation of any real analytic surface on the Euclidean plane, the first step towards uniformization. The main goal of the present section is to expound this theorem.

I.2.1. From practice to theory

First constructions. — Written by Ptolemy in the 2nd century AD, the famous geographical treatise *Geography* maintained its authority till the Renaissance. It describes (and applies) several methods of representing the then known world as precisely as possible on a planar map. Of course, the geometers and astronomers of antiquity were aware that it is impossible to represent a part of a sphere on a flat surface so as to preserve all pertinent geometrical information (distances, angles, areas, etc.) — that is, isometrically.

This impossibility is due to the curvature of the sphere, in modern terminology — that is, in the precise sense of "curvature of a surface" as defined by Gauss. Of course, ancient astronomers had no such sophisticated mathematical artillery at their disposal, but they must surely have been aware of simple manifestations of that curvature. For example, a geodesic triangle forming an exact eighth part of the sphere, with its angles all right angles, readily occurs to the imagination, and shows clearly that not all spherical triangles can be faithfully represented on a plane (see Figure I.1).

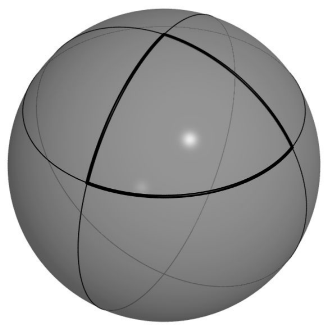


Figure I.1: A spherical triangle

We might also mention that although Ptolemy and his forerunners (Eratosthenes in the 3rd century BC, Hipparchus in the following one) did indeed take the Earth to be spherical in their model of the world, the attempt to obtain a useful planar representation of the celestial sphere of fixed stars presents in any case the same difficulty independently of the question of the shape of the Earth.

Of course, constraints on the planar representation of large parts of a sphere will depend on the intended use of the map. A sovereign exacting taxes proportional to areas of land under cultivation, a sailor navigating with compass and astrolabe, or an astronomer observing the heavens — these all have different requirements. Leaving aside (important) questions of aesthetics, it would seem relevant to demand, for example, one or more of the following:

— that areas be preserved (or, rather, be in a fixed proportion to the originals); in this case one calls the map *equivalent*;

— that angles be preserved (a *conformal* map);

— that the distances from a particular reference point be preserved (an *equidistant* map;

— that certain distinguished curves be mapped onto straight line segments. In this connection it is natural to think of geodesics (*geodesic* maps), but a sailor would naturally tend to give priority to routes of constant heading (*loxodromic maps*).

There are obviously many other constraints that one might impose on the planar map, yielding as many different problems to solve or to be shown incapable of solution. The book [Sny1993] is a good introduction to such aspects of the history of cartography.

Box I.1: Conformal mappings

In this book we shall often have occasion to revisit the concept of conformal maps so it may be appropriate to give the precise terminology. For a linear operator L on a Euclidean vector space $(E, \|\cdot\|)$ the following properties are equivalent:

— L preserves angles;

— L is a similarity, that is, there exists a positive constant c such that $\|L(v)\| = c\|v\|$ for every vector v of E.

The word "similarity" conveys preservation of shape; in German one finds the adjective *winkeltreu*, directly conveying the preservation of angles.

A diffeomorphism between two open sets of the Euclidean plane is said to be *conformal* if its differential map has the above two properties at every point. The expression "similar in infinitesimally small regions" also used to be current in both French and German. Later on we shall see that once the plane has been made over into the complex plane $\mathbb{C}$, one is in a position to speak of a given diffeomorphism as being holomorphic or not holomorphic. Note also the analogous meanings of the Greek and Latin roots *morph* and *form*, and likewise how the prefixes *holo-* and *con-* both convey the sense of preservation.

Even before Ptolemy various projections had been used by ancient Greek scholars. An intermediate step, crucial both theoretically and practically, was the introduction of the idea of latitude and longitude, already familiar to Hipparchus. This provided in effect a means of pinpointing the positions of two distant towns, say, using a single system of coordinates. Astronomical criteria — most notably observations of the stars — could be used to ascertain position. A fairly typical example: if two towns A and B in the northern hemisphere are such that a particular star is visible the whole night through by an observer at A, while an observer at B sees it rise and set, one can infer that the town A is more northerly than B.

One method of drawing a map is therefore to impose constraints on the images on the map of the circles of constant latitude — called *parallels* — and the great half-circles of fixed longitudes — the *meridians*. Cartographers call the network of images of lines of latitude and longitude a *graticule*. Thus *rectangular* maps are those where the parallels and meridians are represented by horizontal and vertical straight lines respectively: here the graticule is made up of rectangles. Among these we find the *equirectangular* map (often known also by the French name *plate carrée* (squared flat [projection]), dating from before Ptolemy's *Geography*, generally signifying a constant spacing of the equal jumps in latitude and longitude. In this case, therefore, the graticule is a grid of squares of fixed size (Figure I.2).

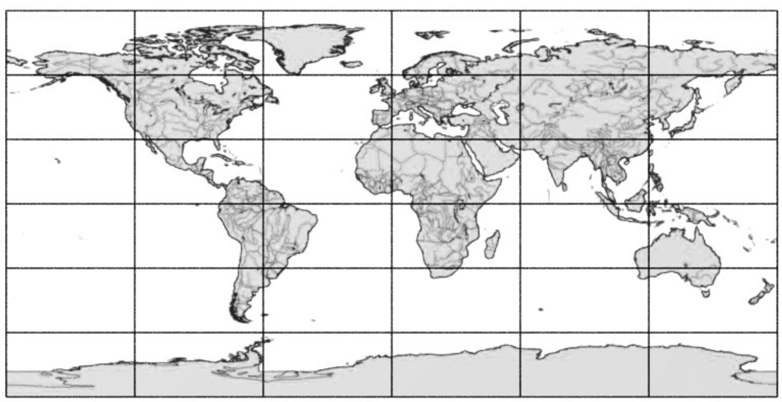


Figure I.2: The *plate carrée*

Another natural method of producing maps is to apply certain simple geometric operations to the space in which the sphere is situated in order to obtain a planar image: one might apply an orthogonal projection onto a suitably positioned plane, or a projection from some point, or indeed map the sphere onto a surface such as a cone or cylinder, and then develop the resultant image onto a flat surface. One very old such method is stereographic projection, known to Hipparchus and probably earlier. This involves projecting the Earth's surface onto the tangent plane to one of its points (the South Pole, for instance) from the antipodal point (Figure I.3). This procedure yields a planar map representing the whole of the Earth's surface except for the point of projection. Clearly, the distortion occasioned by this method increases with the distance from the point of tangency.

An essential property of this map is its conformality: angles drawn on the sphere remain the same on the map. This property, as important for celestial maps as for terrestrial or maritime ones, seems to have been noticed and proved for the first time by the famous English astronomer Edmond Halley towards the close of the 17th century [Hal1695]. The book [HiCo1932] contains an elegant proof of this fact.

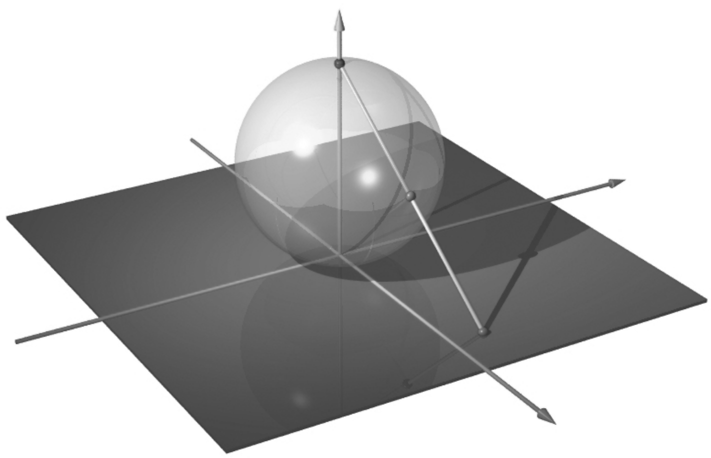


Figure I.3: The stereographic projection

Although the stereographic and equirectangular projections are still in use (the first in producing maps of the celestial sphere and the second as affording the simplest means of sketching a map by computer from knowledge of the latitude and longitude of certain towns), they have largely yielded in importance to other projections. The most familiar projection is that invented by Mercator in 1569.

Mercator's aim was to produce a rectangular map, like the *plate carrée*, with the difference that now routes of constant heading[3] on the sphere are represented on the map by straight lines, making the map suitable for maritime navigation. However, this constraint entails a wider and wider spacing of the images of the parallels of latitude as one approaches the poles, resulting in the familiar distortion of areas. Mercator actually constructed a model of his map, probably by calculating graphically the necessary spacing between pairs of parallels differing

[3]That is, constant bearing relative to true north.

by ten degrees latitude. Mercator's is the second conformal projection after the stereographic projection.

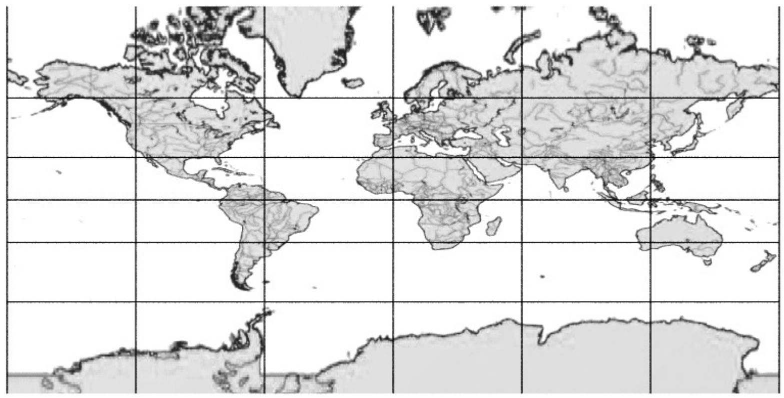


Figure I.4: Mercator's projection

Introduction of the differential calculus. From a mathematical point of view the 18th century marked a renascence in both the conception and study of geographical map-making through the application of the differential calculus. The pioneer in this development was Johann Heinrich Lambert.

Equally well known for his work in the physical sciences (the *Law of Beer–Lambert* describes the absorption of light by a chemical solution as a function of its concentration) and especially for giving the first proof of the irrationality of π, in his work *Beyträge zum Gebrauche der Mathematik und deren Anwendung*[4], and especially *Anmerken und Zusätze zur Entwerfung der Land und Himmelscharten*[5] [Lam1772], written between 1765 and 1772, Lambert described numerous methods of obtaining cartographical representations and opened the way to a systematic analytic study of the various constraints, notably equivalence and conformality. On the practical side, it is to him that we are indebted for *Lambert's conformal conical projection.*, the present-day official projection used for the maps of France, but he also gave the first analytic proofs of the conformality of the stereographic and Mercator projections, re-proved by Euler in 1777 in [Eul1777].

Inspired by Lambert's translation of cartographical questions into mathematical language, Lagrange [Lag1779] saw that the subject suggests more general

[4] Contributions to the utilization of mathematics and its application.

[5] Notes and comments on the construction of terrestrial and celestial maps.

questions than just those associated with the production of conformal maps and the verification of their properties. The problem occurred to him of determining all conformal maps that one can make of the Earth's surface, but with a refinement of the model of the Earth commonly used: he assumed a "spheroidal" shape for the Earth — more precisely, that it is a surface "generated by the revolution of some curve about a fixed axis".

In summarizing the history of cartography, Lagrange observes, without citing Mercator explicitly, that the possibility of producing conformal maps other than by direct projection of the terrestrial sphere onto a tangential cone or cylinder leads one to a more general and fruitful perspective on the problem, allowing its transformation from a purely practical question into a mathematical one:

> This investigation [of conformal maps], as interesting for the analytical techniques it requires as for its potential application to the drawing of geographical maps, seems to me a topic worthy of the attention of geometers and appropriate subject-matter for a memoir.

Thus he proposes determining all conformal planar representations of a surface of revolution. His idea is to imitate Mercator's projection in the sense of identifying the constraints on the spacing of parallels ensuring conformality.

We first introduce appropriate notation: the surface in question is obtained by revolving a planar arc about the axis joining its end-points, the *poles* of the surface. Each point of the surface is then naturally coordinatized by the longitude φ and the length s of the arc of the generating curve from the point to one of the poles. (In the case of a sphere of radius 1, the coordinate s is $\pi/2$ minus the latitude.) Each point (φ, s) of the surface lies on a horizontal circle (representing a parallel) of radius $q(s)$, say. (In the case of the unit sphere this radius is $\sin s$, or, equivalently, the cosine of the latitude.)

In this notation the Riemannian metric[6] — also called "the first fundamental form" — of the surface is easily seen to be $ds^2 + q(s)^2 d\varphi^2$.[7] Representing the surface conformally on the plane then comes down to expressing the rectangular Cartesian coordinates x and y as functions of s and φ in such a way that the elements of distance computed in terms of x, y on the one hand and φ, s on the other satisfy the proportionality relation

$$dx^2 + dy^2 = n(\varphi, s)^2(ds^2 + q(s)^2 d\varphi^2),$$

[6] We don't hesitate to call this metric "Riemannian" even though it considerably predates Riemann.

[7] This is the square of an infinitesimal element of length on the surface, considered embedded in Euclidean space. Thus the length of a smooth arc $(\varphi(t), s(t))$, $a \le t \le b$, is $\int_a^b \sqrt{(ds/dt)^2 + q(s(t))^2 (d\varphi/dt)^2} dt$. *Trans*

where n is a non-vanishing function representing the dilation factor of distances at each point.

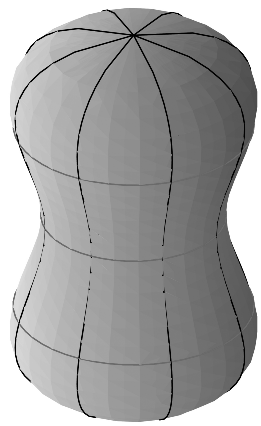


Figure I.5: A surface of revolution

Lagrange finds a system of coordinates u, v solving this equation for x and y, and representing a generalization, within the limits of his investigation, of Mercator's projection. The functions u, v in question are given by

$$u(s) = \int_0^s \frac{d\sigma}{q(\sigma)}, \quad v = \varphi,$$

which satisfy

$$du^2 + dv^2 = \frac{ds^2}{q(s)^2} + d\varphi^2 = \frac{1}{q(s)^2}(ds^2 + q(s)^2 d\varphi^2),$$

and therefore define (locally, away from the poles) a conformal coordinate system for the surface of revolution.

Having found one conformal coordinate system, Lagrange goes on in his memoir to consider the problem of determining the other possible such systems, in particular, for practical reasons in the case where the graticule — the network of images of the parallels and meridians — is made up of circles. In the evolution of cartography this theoretical result represents the first occasion where conformal coordinates are found for a relatively general class of surfaces.

I.2.2. Gauss's view of conformal representation

In 1822 the Royal Danish Academy of Sciences and Letters in Copenhagen proposed as a prize problem that of representing portions of a given surface on another in such a way that the representation be "similar to the original in infinitesimally small regions." In 1825 Gauss published in Schumacher's *Astronomische Abhandlungen* his famous memoir on the topic [Gau1825], also to be found in his collected works [Gau1863].

The term "conformal representation" was introduced by Gauss only in 1844 in Section I of the first part of his memoir on higher geodesy. This work largely bypasses the particular theme of geographical maps, playing in the theory of functions a role analogous to that of his *Disquisitiones generales circa superficies curvas* in the theory of surfaces.

To return to Gauss's result of 1825: he shows that *every (analytic) surface is locally conformally equivalent to the Euclidean plane* (whence it is immediate that any two analytic surfaces are locally conformally equivalent)[8]. A local system of coordinates $(x, y) \in \mathbb{R}^2$ on a surface is called *conformal* if in terms of x, y the metric has the form $m(x, y)(dx^2 + dy^2)$. Gauss's theorem then states that:

Theorem I.2.1 (Gauss). *Let g be a real analytic Riemannian metric defined in a neighborhood of a point p of an analytic surface. Then there exists a conformal map $V \to \mathbb{R}^2$ from some open neighborhood V of p to the Euclidean plane.*

We shall now sketch Gauss's marvellous proof of this theorem.

We first choose coordinates in some neighborhood of p; expressed in terms of these coordinates the metric on the surface may be considered as defining an analytic metric g in an open neighborhood U of the origin in $\mathbb{R}^2$.

To ease understanding we first prove the exact analogue of Gauss's theorem in the case where the open set U is endowed with a *Lorentzian* metric g. This means that at each point of U there is given a quadratic form of signature $(+, -)$, and we wish to show that this Lorentzian metric is conformal to the standard Lorentzian metric $dx^2 - dy^2$ on $\mathbb{R}^2$ — in the sense of the obvious extension of conformality to the Lorentzian situation. One proceeds as follows. At each point of U the metric g determines two directions where it vanishes — the two "isotropic" directions of the metric. Hence locally one obtains two non-singular vector fields determined by these directions, and on integrating them one obtains two families of isotropic curves intersecting transversely. For example, in the case of the standard Lorentzian metric $dx^2 - dy^2$ these curves will clearly be just the lines of slopes ± 1.

[8]Note that for Gauss the surfaces in question are embedded in Euclidean space, from which they inherit their metric.

We now choose the origin O as base point in U and denote by P_0 any particular point of $\mathbb{R}^2$. Denote by C_1 and C_2 the two isotropic curves through O of the Lorentzian metric g and by $\mathcal{D}_1$ and $\mathcal{D}_2$ the isotropic lines through P_0 of the standard Lorentzian metric on $\mathbb{R}^2$. Let $f_1 : U \to \mathbb{R}^2$ be any diffeomorphism sending C_1 onto $\mathcal{D}_1$ and f_2 a diffeomorphism sending C_2 onto $\mathcal{D}_2$. Now let m be an arbitrary point of U close to the origin, and let $\tilde{C}_1$ and $\tilde{C}_2$ be the two isotropic curves of the metric g passing through it. By replacing U by a smaller open set V if necessary, we may assume that $\tilde{C}_1$ intersects C_2 in a single point p_1 and likewise that $\tilde{C}_2$ intersects C_1 in just one point p_2.

The map ψ that we seek is then that sending each such point m of V[9] to the point of intersection $M = \psi(m) \in \mathbb{R}^2$ of the isotropic lines of $\mathbb{R}^2$ through the points $P_1 = f_1(p_1)$ and $P_2 = f_2(p_2)$. The map ψ so defined sends g-isotropic directions in V to those of the standard Lorentzian metric on $\mathbb{R}^2$.

We now appeal to the crucial, and easily seen, fact that two quadratic forms of signature $(+,-)$ on a real vector space of dimension 2 are proportional if and only if they have the same isotropic directions. We must therefore have $\psi_* g = m(x,y)(dx^2 - dy^2)$ for some non-vanishing function $m(x,y)$. In other words, ψ is a conformal map, and Gauss's theorem is thus established *in the Lorentzian case* — and moreover without the assumption of analyticity.

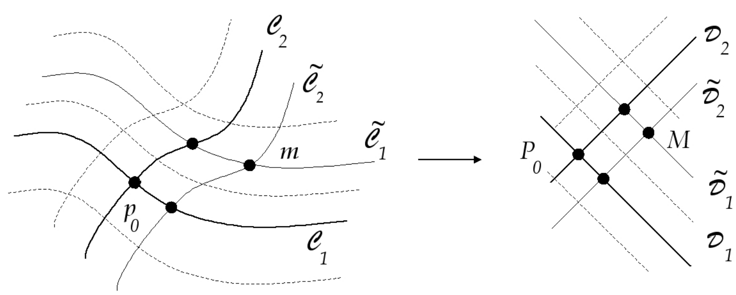


Figure I.6: The Lorentzian version of Gauss's theorem

In the case where g is a real analytic *Riemannian* metric, although certainly one no longer has isotropic directions to play with, nevertheless the same underlying idea can be made to work given sufficient imagination.

[9]Where now V is an open subset contained in U all of whose points have the property pertaining to the point m. *Trans*

We first express the basic ideas of the argument in modern terminology. That argument begins with the complexification of the open set U into an open set $\hat{U} \subset \mathbb{C}^2$; this is just an open neighborhood of U considered as a subset of $\mathbb{C}^2$. We write $\hat{g}_0 = dx^2 + dy^2$ for the standard "complex Riemannian" metric on $\mathbb{C}^2$, with x and y now the standard (complex) coordinates of $\mathbb{C}^2$. (Strictly speaking this is not a Riemannian metric since the underlying quadratic form takes on complex values.) Since by assumption the coefficients of the metric g are real analytic functions we can, by restricting $\hat{U}$ if need be, extend g uniquely to a complex analytic — that is, holomorphic — metric $\hat{g}$ on the open set $\hat{U}$. Furthermore, since the coefficients of g are real the metric $\hat{g}$ will be invariant under complex conjugation $(x, y) \mapsto (\bar{x}, \bar{y})$. Now in $\mathbb{C}^2$ one does have two transverse families of isotropic complex lines of the metric $\hat{g}_0$, with equations of the form $y = \pm ix + \text{const.}$, while on $\hat{U}$ the metric $\hat{g}$ likewise gives rise to two families of holomorphic vector fields, which one integrates to obtain two families of holomorphic curves intersecting transversely. (Note that these holomorphic complex curves in $\mathbb{C}^2$ correspond to surfaces in $\mathbb{R}^4$.)

Next one maps the origin O of U to an arbitrary real point P_0 of $\mathbb{R}^2 \subset \mathbb{C}^2$. Through O there passes a complex isotropic curve C_1 and the complex curve C_2 obtained by complex conjugation of the curve C_1. By means of these curves one defines, exactly as in the Lorentzian situation, a mapping $\hat{\psi}$ of a suitable neighborhood $\hat{V}$ of O contained in $\hat{U}$, with image in $\mathbb{C}^2$. The diffeomorphism $\hat{\psi}$ has the additional property of being invariant under complex conjugation, so that it induces a diffeomorphism ψ from $V = \hat{V} \cap \mathbb{R}^2$ to its image $\hat{\psi}(\hat{V}) \cap \mathbb{R}^2$. The fact that the complexification of the diffeomorphism ψ preserves the isotropic directions of the complexification of the metric g means precisely that the map is conformal. This completes the proof of Gauss's theorem. □

Gauss does not set out his proof exactly as above, although his method is essentially the same.

First he writes g out as

$$g = a(x, y)dx^2 + 2b(x, y)dxdy + c(x, y)dy^2, \; ac > b^2.$$

Then he factors the quadratic form as a product of two conjugate linear forms (defining the isotropic directions):

$$\begin{aligned} g &= \frac{1}{a}\left(adx + (b + i\sqrt{ac - b^2})dy\right)\left(adx + (b - i\sqrt{ac - b^2})dy\right) \\ &= \frac{1}{a}\omega\bar{\omega}. \end{aligned}$$

Here ω is what is now called a "holomorphic 1-form" in the complex variables x, y. The equation $\omega = 0$ may be regarded as a differential equation whose solutions are locally of the form $f(x, y) = $ const. where f is defined in some neighborhood of the origin — in other words, ω has the form hdf for some function h.[10] Resolving f into its real and imaginary parts, we have

$$\omega = h(du + idv) \text{ (whence also } \bar{\omega} = \bar{h}(du - idv)),$$

whence, finally,

$$g = \frac{h\bar{h}}{a}(du^2 + dv^2).$$

Here the coordinates u, v are real by construction and the similarity coefficient $m := \frac{h\bar{h}}{a}$ is obviously a real analytic function of x, y, so Theorem I.2.1 is proved. □

Conformal maps ψ are certainly not unique, but of course any two of them differ by a conformal self-map of the Euclidean plane. Thus in order to classify all locally conformal maps one needs to ascertain those coordinate transformations $(x, y) \mapsto (X, Y)$ between open sets of $\mathbb{R}^2$ that are conformal, that is, for which $dX^2 + dY^2 = m(x, y)(dx^2 + dy^2)$ for an appropriate function m — in other words, it is necessary and sufficient that the differential map determine a similarity at each point. If one assumes in addition that orientation is preserved — that the similarity is "direct" — then the condition is expressed by the formulae

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

familiar as the so-called "Cauchy–Riemann" equations expressing the holomorphicity of $X + iY$ as a function of $x + iy$. In fact this way of expressing the conformality of a map in terms of $dx + idy$ was known to Euler as long ago as 1777!

Here then in modern terminology is what Gauss showed:

Every (oriented and analytic) surface can be represented by a map to the Euclidean plane (identified with the complex plane) that is locally conformal and orientation-preserving. Any two such maps differ by a holomorphic change of coordinates.

It follows from this theorem that any surface endowed with a (real analytic) Riemannian metric is a "Riemann surface", as defined in Chapter II below.

Gauss's theorem, established here only in the situation of a real analytic metric, remains true under the weaker assumption that the metric is C^∞ or even just measurable, but the proof is then much more difficult. The C^∞ case was proved by Korn in 1914 and Lichtenstein in 1916, and, finally, in 1960, Ahlfors and Bers established the theorem in the measurable case (see [Ahl2006]).

[10] This is because ω and df both vanish exactly on vectors tangent to the curves $f(x, y) = $ const., thus must be proportional. *Trans*

Gauss does not rest content with merely proving his theorem, but illustrates it with many examples: he begins by showing how to represent the surface of a certain solid on the plane, and then a cone and a sphere. He does not lose sight of the particular question prompting the Danish Academy's choice of problem, and ends his memoir with a treatment of the case of an ellipsoid of revolution. The determination of conformal maps of a more general ellipsoid requires the use of elliptic integrals, which form the theme of the next section.

I.3. Overview of the development of elliptic functions

By the end of the 19th century elliptic functions were at the center of mathematics. They turned up everywhere: in geometry, algebra, number theory, analysis, and even mechanics, and assumed the status of an indispensable accessory of mathematical culture.

Elliptic functions proved useful in allowing certain algebraic curves (those of genus 1) to be uniformized, and they are therefore important in relation to the theme of this book. However, they played a more important role in providing a source of inspiration for Riemann, Klein, and Poincaré — among others — in their investigations of general algebraic curves. Poincaré, for example, presented his theory of Fuchsian functions as a "simple" generalization of that of elliptic functions, and for this reason we now describe the latter theory and its development.

There are many excellent books on elliptic functions, including those taking a historical tack. Among those we prefer, the reader may consult for example [McKMo1997, Bos1992, Hou1978]. In view of the treatments in such works as these, rather than going into the detailed history we shall confine ourselves here to describing just the main developments, concentrating on just those aspects we shall be needing in the sequel.

At the beginning of the 19th century analysts had essentially just a small number of types of *elementary functions* at their disposal[11]: polynomials and rational functions, of course, *algebraic functions* $y(x)$, that is, satisfying a polynomial equation $F(x, y) = 0$ (even if many-valued), and also the exponential and trigonometric functions. Early attempts to "find new transcendental functions with which to enrich analysis" consisted in studying the anti-derivatives of functions already at hand. This method had already proven itself in connection with the "discovery"

[11]Although in the 18th century Euler had introduced the zeta and gamma functions, for instance, as well as the idea of a general function as being defined by a power series. *Trans*

of the natural logarithm

$$\log x = \int \frac{dx}{x}.$$

Euler, Gauss, Legendre, Abel, and Jacobi, among others, began the general investigation of *Abelian integrals*, as Jacobi called them, that is, integrals of the form

$$\int R(x, y)dx,$$

where R is a rational function of x and y with y an algebraic function of x. We present here their respective contributions to this subject.

I.3.1. Euler

The first step consisted in a somewhat "magical" calculation performed by Euler in commenting on an article by Fagnano. This concerned one of the very simplest of anti-derivatives not expressible in terms of the known elementary functions, namely

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

which conforms to the preceding definition of an Abelian integral with $y^2 = 1-x^4$ and $R(x, y) = 1/y$. This integral arises in the attempt to evaluate the length of an arc of the lemniscate with equation, in polar coordinates, $r^2 = \cos 2\theta$ (the last of the curves depicted in Figure 2 in the General Introduction).[12]

In 1752 Euler proved the following "addition theorem":

$$\int_0^x \frac{dt}{\sqrt{1-t^4}} + \int_0^y \frac{dt}{\sqrt{1-t^4}} = \int_0^z \frac{dt}{\sqrt{1-t^4}},$$

where

$$z = \frac{x\sqrt{1-y^4} + y\sqrt{1-x^4}}{1+x^2y^2}.$$

He was doubtless led to this by the analogy with the integral

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

[12]Arc length in polar coordinates is calculated by means of the integral $\int \sqrt{(dr/d\theta)^2 + r^2}d\theta$, which reduces to $\int \frac{1}{\sqrt{\cos 2\theta}}d\theta$ for the lemniscate (so that the length of a loop of the lemniscate is obtained by evaluating this integral from $\theta = -\pi/4$ to $\theta = \pi/4$). The substitution $x := \tan\theta$ then yields the above indefinite integral. *Trans*

for which the formula

$$\sin(a+b) = \sin a \cos b + \sin b \cos a$$

yields the addition formula

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

where

$$z = x\sqrt{1-y^2} + y\sqrt{1-x^2}.$$

It should be observed that at this stage in the development these identities are considered as holding for x, y in the interval $[0,1]$. For values of x and y outside this interval the problem of choice of square root arises. Note also that Euler makes no explicit use of complex variables in this work.

I.3.2. Gauss

Although during his lifetime Gauss published nothing on this topic, his letters show that he had a clear understanding of the issue as early as 1796. His first idea was to invert the function

$$a = \int_0^x \frac{dt}{\sqrt{1-t^4}}$$

and consider x as a function of a, which he denotes by $x = \sin\text{lemn}\, a$. The analogy with the circular functions doubtless again played a role: the sine and cosine are convenient for parametrizing the circle by arc length. He translates Euler's addition formula into an addition formula for $\sin\text{lemn}\,(a+b)$, but does not stop there. Even though he is still, at that early stage of the game, hesitant about letting x be complex in the above integral, he is tempted to choose x purely imaginary, of the form iy, and to consider the integral

$$\int_0^y \frac{i\,dt}{\sqrt{1-t^4}}.$$

This leads him to conclude that $\sin\text{lemn}\,(ib) = i \sin\text{lemn}\, b$, and this in turn, in view of the addition formula, allows him to define $\sin\text{lemn}\,(a+ib)$ in terms of $\sin\text{lemn}\, a$ and $\sin\text{lemn}\, b$. *Thus is the elliptic function* $\sin\text{lemn}$ *of a complex variable* $a+ib$ *born.*

Gauss continues his investigation of this function using the analogy with the sine function. Starting from the addition formula, now conveniently extended to all of $\mathbb{R}$, he shows that the function sin lemn is periodic of period

$$2\varpi = 4\int_0^1 \frac{dt}{\sqrt{1-t^4}}.$$

By the same means he finds a second period equal to $2i\varpi$.[13] Thus the function sin lemn has *two* linearly independent periods, subsequently the *defining property* of elliptic functions. The adjective "elliptic" originates from the fact that these new transcendental functions arise not only in attempting to calculate arc length of a lemniscate but also that of an ellipse.

Although the rest of Gauss's work on this theme is of equal significance, it would take us too far out of our way to discuss it. However, we cannot but mention such marvellous expressions for sin lemn z involving doubly infinite products, as

$$\sin\operatorname{lemn} z = z\frac{\prod'_{m,n}(1-\frac{z}{\alpha_{m,n}})}{\prod'_{m,n}(1-\frac{z}{\beta_{m,n}})},$$

where $\prod'$ denotes the product over all pairs $(m,n) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, $\alpha_{m,n} = (m+in)\varpi$, and $\beta_{m,n} = ((2m-1)+i(2n-1))\varpi/2$. Note here the appearance of the famous "Gaussian integers".

I.3.3. Abel and Jacobi

We mentioned above that Gauss never published his discoveries on this theme. Twenty-five years later Abel and Jacobi, in ignorance of Gauss's work, retraced his steps, until around 1827 they began to go well beyond him, in part independently and in part mutually stimulated by a relatively protracted rivalry. On this subject there has survived a lively correspondence between the young Jacobi and an aging Legendre sometimes assuming the role of intermediary [LeJa1875].

The mention of Legendre's name affords an opportunity to note that he also must be considered one of the precursors of the theory, having dedicated forty years of his life to it, beginning in 1786. His labors culminated in the publication in 1830 of the three volumes of his *Traité des fonctions elliptiques*. In this connection one should mention, however, that Legendre's elliptic functions are functions

[13]The notation is explained by the fact that the quantity 2ϖ is the length of the lemniscate in question.

of a single real variable, and that one of his chief motivations was to establish numerical tables, with a view to applications. And moreover he never penetrated to the double periodicity of the *inverse* of the anti-derivative of $1/\sqrt{1-t^4}$.

We remind the reader that anti-derivatives of the form $\int p(x)/\sqrt{q(x)}dx$, where p is a polynomial of any degree and q one of degree at most 2, can be explicitly evaluated in terms of logarithms and rational functions. Geometrically, this comes down to the fact that the curve defined by the equation $y^2 = q(x)$ is a conic, which therefore admits a rational parametrization, by means of which the problem is reduced to that of anti-differentiating a rational function, and for these there is the well known standard procedure involving logarithms arising as integrals of expressions cognate to $1/x$.[14]

One of Legendre's contributions was a systematic classification of integrals of the form $\int \left(p(x)/\sqrt{q(x)}\right) dx$ when the degree of q is 3 or 4. He shows that in this case the calculation reduces to three precise types of anti-derivative playing in some sense a logarithm-like role, whose values he tabulates.

Be that as it may, Abel and Jacobi investigated integrals of the form

$$u = \int_0^x \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}},$$

in connection with which both hit on the good idea of considering x as a function of u — unaware that Gauss had had the same idea earlier. The parameter k is called the "modulus", and since it is a parameter not varying within the integral, they denoted the inverse function simply by $x = \sin\operatorname{am} u$. They "showed", more or less, that x is a single-valued meromorphic doubly-periodic function of u satisfying a certain addition formula, and they went on to obtain a great number of series expansions of such functions.

A central theme of their investigations concerned certain "transformations" — rather magical-seeming formulae relating values of $\sin\operatorname{am} u$ for different values of the parameter k, some of which had been found earlier by Euler. This marked the début of the theory of modular equations, which, however, we shall not broach here, even though they will turn up in the course of our discussion of Klein's quartic.

I.3.4. Jacobi and the ϑ-functions

In 1835–36 Jacobi developed extremely powerful tools for constructing elliptic functions as ratios of holomorphic functions. These are the so-called "ϑ-functions".

[14]And also trigonometric functions. *Trans*

They are relatively simply defined: taking ω to be a complex parameter, and writing $p := \exp(i\pi z)$, $q := \exp(i\pi\omega)$, we may define them as follows:

$$\begin{aligned}
\vartheta_1(z) = \vartheta_1(z|\omega) &= i \sum_{n=-\infty}^{\infty} (-1)^n p^{2n-1} q^{(n-1/2)^2}, \\
\vartheta_2(z) = \vartheta_2(z|\omega) &= i \sum_{n=-\infty}^{\infty} p^{2n-1} q^{(n-1/2)^2}, \\
\vartheta_3(z) = \vartheta_3(z|\omega) &= i \sum_{n=-\infty}^{\infty} p^{2n} q^{n^2}, \\
\vartheta_4(z) = \vartheta_4(z|\omega) &= i \sum_{n=-\infty}^{\infty} (-1)^n p^{2n} q^{n^2}.
\end{aligned}$$

For $\operatorname{Im}\omega > 0$ these series converge and define 1-periodic functions of z. Although they themselves are not elliptic functions — having only the single basic period 1 — all the same $\vartheta_i(z+\omega)$ can be expressed very simply in terms of $\vartheta_i(z)$. For example,

$$\vartheta_1(z+\omega) = -p^{-2}q^{-1}\vartheta_1(z).$$

The point is then that ratios of two ϑ-functions may be doubly periodic. Thus ϑ_1/ϑ_4 is an elliptic function with periods 1 and ω. The ϑ-functions satisfy a tremendous number of identities each more astonishing than the one before, and their applications — notably in number theory — continue to prove their worth.

To learn *much* more on this theme, one may consult for example [McKMo1997, Mum1983, Mum1999].

I.3.5. Bringing the theory into final form: Eisenstein, Liouville, and Weierstrass

From 1840 onwards the theory of *elliptic functions* stabilized, taking on the form familiar to us today. From that time on an elliptic function is defined as any meromorphic function f of the complex plane admitting two independent periods ω_1, ω_2:

$$f(z + m\omega_1 + n\omega_2) = f(z)$$

for all $z \in \mathbb{C}$ and all integers m, n.

The functions obtained by Abel and Jacobi as inverses of anti-derivatives of $1/\sqrt{(1-t^2)(1-k^2t^2)}$ are examples of such functions, but are there any others? Is there an elliptic function for every choice of the two periods? Here again we must limit ourselves to merely stating the main results, obtained independently by

Eisenstein, Liouville, and Weierstrass — results one may consider known when Riemann began working on his thesis.

Given two complex numbers ω_1 and ω_2, linearly independent over $\mathbb{R}$, the lattice Λ they generate is the set of points of the form $m\omega_1 + n\omega_2 \in \mathbb{C}$, $m, n \in \mathbb{Z}$. These form a discrete subgroup of $\mathbb{C}$, and the fact that a function has periods ω_1 and ω_2 means that it is in fact defined on the quotient torus $\mathbb{C}/\Lambda$, which is, as we shall soon see, a basic example of a Riemann surface.

The *Weierstrass ℘-functions* are elliptic functions with prescribed periods, defined by

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \setminus \{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

It can be shown that each such series converges where defined and defines a meromorphic function with lattice of periods precisely Λ. It has a pole of order 2 at the origin of $\mathbb{C}/\Lambda$ and is holomorphic everywhere else.[15]

His next step was to show that this function satisfies a differential equation, namely

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3,$$

where g_2 and g_3 are the *Eisenstein series*

$$\begin{aligned} g_2(\Lambda) &= 60 \sum_{\omega \in \Lambda \setminus \{0\}} \omega^{-4}, \\ g_3(\Lambda) &= 140 \sum_{\omega \in \Lambda \setminus \{0\}} \omega^{-6}. \end{aligned}$$

He establishes this equation using a method due to Liouville: the difference between the two sides represents a (meromorphic) elliptic function, and one then chooses the coefficients in such a way as to eliminate the pole, thus obtaining a holomorphic function. Since the only holomorphic elliptic functions are constants (in view of the compactness of $\mathbb{C}/\Lambda$ and Liouville's theorem), we have the result.

It follows that the projective algebraic curve C with affine equation

$$y^2 = 4x^3 - g_2x - g_3$$

is *uniformized* by the torus $\mathbb{C}/\Lambda$ via the parametrization

$$z \in \mathbb{C}/\Lambda \mapsto (\wp(z), \wp'(z)) \in C.$$

[15]It can be shown that every meromorphic doubly periodic function with basic periods ω_1 and ω_2 is a rational expression in $\wp, \wp'$ with coefficients in $\mathbb{C}$, so that the totality of such functions is the algebraic function field $\mathbb{C}(\wp, \wp')$. *Trans*

It remains to show that, conversely, for any g_2 and g_3 such that the curve C is non-singular, that is, satisfying $g_2^3 - 27g_3^2 \neq 0$, there is a lattice Λ with Eisenstein invariants equal to g_2 and g_3. This can be achieved in several ways, the simplest of which is to consider the integral

$$\int \frac{dz}{\sqrt{4z^3 - g_2 z - g_3}},$$

and imitate Gauss, Abel, and Jacobi by inverting it. The periods of the elliptic function thus obtained are then the required ones.

Furthermore, one can also show that every smooth curve of degree 3 in the complex projective plane is projectively equivalent to a curve of the above form (named after Weierstrass, although it was Newton who originally discovered this): this is done effectively by projecting to infinity a tangent line at a point of inflection of the given cubic curve.

The upshot of our discussion is thus that:

Every smooth curve of degree 3 in the complex projective plane is isomorphic to a torus of the form $\mathbb{C}/\Lambda$, and furthermore by means of an isomorphism determined by an elliptic function.

A final point to end this preliminary chapter: since $\mathbb{C}/\Lambda$ is an Abelian group, the same group structure is induced on the smooth cubic curve it parametrizes. The addition formula discovered by Euler reflects this. It turns out that the rule of addition on the cubic is extremely simple. First one chooses a point of inflection to represent the identity (or zero) element, and then one declares that the three points of intersection of the curve with any straight line have sum zero. This defines the rule of addition completely. The proof that this geometric construction does indeed yield an addition defining a group is an interesting exercise in projective geometry (see for example [McKMo1997]).

It may be of interest to remark that the simple projective definition of this group structure appears to have been unknown to the heroes of this chapter. From [Cat2004, Scha1991] it appears that perhaps even Poincaré had no clear idea that the rational points of a cubic curve defined over $\mathbb{Q}$ form an Abelian group (even though he spoke of it as having "finite rank").

Chapter II

Riemann

In this chapter we examine two of Riemann's memoirs: his doctoral thesis [Rie1851] defended in Göttingen in 1851, where he develops the theory of holomorphic functions and proves the "Riemann mapping theorem", and his article on Abelian functions [Rie1857] published in Crelle's journal six years later. In the latter work Riemann applies the techniques developed in his thesis to the construction of a general theory of algebraic functions and their associated Abelian integrals. Recall that a function $s(z)$ is called *algebraic* if it satisfies a polynomial equation $P(s(z), z) = 0$, and that an *Abelian integral* is one of the form $\int F(s(z), z)dz$ where F is a rational function of two variables.

Subsequently the paper [Rie1857] came to be considered as initiating major directions of mathematical research, including the topology and analytical geometry of compact Riemann surfaces, their moduli spaces, the Riemann–Roch theorem, birational geometry, the theory of general theta-functions and Abelian varieties, the Dirichlet problem, Hodge theory, etc. Over just the 25 years following the publication of this article, we see its results geometrized by Clebsch, and then by Brill and Noether, then arithmetized by Dedekind and Weber — and a start made by Clebsch and Noether on extending the results to algebraic surfaces.

It has been an absorbing task to bring to light the seeds of all of these developments contained in this single article.

II.1. Preliminaries: holomorphic functions and Riemann surfaces

II.1.1. Holomorphic functions

We begin by explicating Riemann's work on the uniformization of simply connected open sets of the plane, contained in his thesis [Rie1851] published in 1851.

We describe first of all how Riemann defines the concept of a holomorphic function in the very first section of this memoir. He considers a complex-valued function $w(z) = u(z) + iv(z)$ of a quantity $z = x + iy$ varying over an open set U of the complex plane, and studies the differential quotient:

$$\frac{dw}{dz} = \lim_{z' \to z} \frac{w(z) - w(z')}{z - z'},$$

observing that:

> When the dependence of the magnitude w on z is chosen arbitrarily, the quotient $\frac{du+idv}{dx+idy}$ will generally vary with the values of dx and dy.

This may be unpacked as follows: if we denote by

$$dw_z : \mathbb{C} \simeq \mathbb{R}^2 \to \mathbb{C} \simeq \mathbb{R}^2$$

the differential map at a point z of the function w considered as a real differentiable function $\mathbb{R}^2 \to \mathbb{R}^2$, and consider an infinitesimal increment $dz = \varepsilon e^{i\varphi}$ of the variable z, then we have[1]

$$\begin{aligned}\frac{dw_z(\varepsilon e^{i\varphi})}{\varepsilon e^{i\varphi}} = \frac{1}{2}\left(\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\right) \\ + \frac{1}{2}\left(\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right) e^{-2i\varphi}.\end{aligned}$$

If the term

$$\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)$$

does not vanish at the point z, the quantity $\frac{dw_z(\varepsilon e^{i\varphi})}{\varepsilon e^{i\varphi}}$ will vary with $e^{i\varphi}$. However, as Riemann observes, for all functions w obtained from z by means of "elementary computational operations", the quantity $\frac{dw_z(\varepsilon e^{i\varphi})}{\varepsilon e^{i\varphi}}$ does not depend on $dz = \varepsilon e^{i\varphi}$.

[1] This can be seen as follows. Suppose for simplicity that $z = 0$ and $w(z) = w(0) = 0$. The differentiability of $w(z)$ considered as a function of the two real variables x, y means that there exist numbers α, β — in fact these are just $\partial w/\partial x$ and $\partial w/\partial y$ at $z = 0$ — such that $w(z) = \alpha x + \beta y + \eta(z) z$ where $\eta(z) \to 0$ as $z \to 0$. Rewriting this as $w(z) = \left(\frac{\alpha - i\beta}{2}\right) z + \left(\frac{\alpha + i\beta}{2}\right) \bar{z} + \eta(z) z$, it follows that $\frac{w(z)}{z} = \left(\frac{\alpha - i\beta}{2}\right) + \left(\frac{\alpha + i\beta}{2}\right) \frac{\bar{z}}{z} + \eta(z)$. Taking the limit as $z \to 0$, that is, setting $z = dz = \varepsilon e^{i\varphi}$ (and noting that $d\bar{z}/dz = e^{-2i\varphi}$), we obtain the formula that follows. *Trans*

He therefore proposes taking the vanishing of this term as the defining condition of what he calls a *function of a complex variable*:

> A variable complex quantity w is called a function of another variable complex quantity z when it varies with z in such a way that the value of the derivative $\frac{dw}{dz}$ is independent of the value of the differential dz.

In other words, for Riemann the term "function of a complex variable" always means a holomorphic function. Thus such functions are by definition just those satisfying the Cauchy–Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \tag{II.1}$$

which is equivalent to the closure of the complex differential 1-form

$$w(z)dz = (u + iv)(dx + idy).$$

It can be shown that then the function $w'(z) := \frac{dw}{dz}$ is well-defined and again holomorphic, so that w is in fact infinitely differentiable.[2]

If a function $w = u + iv$ is holomorphic, it follows from the Cauchy–Riemann equations and the fact that it is twice differentiable that the functions u and v satisfy

$$\Delta u = \Delta v = 0,$$

where $\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the *Laplacian* associated with the complex coordinate z. Functions of two variables annihilated by the Laplacian are said to be *harmonic*. Thus the real and imaginary parts of a holomorphic function are harmonic.

Conversely, given a function u defined and harmonic on a simply connected open set $U \subset \mathbb{C}$, there exists a holomorphic function $f_u : U \to \mathbb{C}$, uniquely defined to within a purely imaginary additive constant, such that $u = \mathrm{Re}(f_u)$. The function f_u is in fact simply a primitive of the holomorphic 1-form

$$du - i du \circ i.$$

The function $u^* = \mathrm{Im}(f_u)$, defined only up to an additive constant, is called the *conjugate function* of u.

This close affinity between holomorphic and harmonic functions is central to the methods used in [Rie1851, Rie1857], since his proofs of the conformal

[2]Even more, a function holomorphic in a neighborhood of a point z_0 is *analytic*, that is, equal to its Taylor series expansion, in some neighborhood of z_0. *Trans*

representation theorem (or "Riemann mapping theorem") and the existence of certain Abelian integrals are based on a close study of harmonic functions, and especially "Dirichlet's principle".

We recall also the following "mean-value" property of harmonic functions. If $D(z_0,r) = \{z \in \mathbb{C} | |z - z_0| \leq r\}$ is a disc contained in U, then

$$\begin{aligned} u(z_0) &= \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{i\theta})d\theta \\ &= \frac{1}{\pi r^2}\int_{D(z_0,r)} u(x,y)dxdy. \end{aligned}$$

In fact this property characterizes harmonic functions: a continuous (or even just measurable) function is harmonic if and only if it has the above mean value property on closed discs in U. It follows that harmonic functions satisfy the *maximum principle*: if a harmonic function u has a local extremum at a point z_0 of U, then it must be constant in some neighborhood of z_0. Another consequence is that a function $v : U \to \mathbb{R}$ that is a uniform limit of harmonic functions defined on compact subsets of U is itself harmonic.

It is noteworthy that, in contrast with Abel, whose approach is essentially algebraic, consisting of manipulations of functions of several variables and of algebraic and differential equations, Riemann works with functions independently of specific formulae, basing his argumentation on their defining properties, as he explains in the introduction to [Rie1857]:

> I shall consider as a function of $x + yi$ any quantity w that varies with the first quantity in such a way as to satisfy the equation
>
> $$i\,\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y},$$
>
> without resorting to an expression for w in terms of x and y.[3]

This desire to avoid starting out with particular expressions for his functions is taken up again a little further on:

> By a known theorem, mentioned earlier, the property of a function of being single-valued comes down to the possibility of developing it by means of positive or negative integer powers of increments of the variables, while the many-valuedness of a function reduces to the impossibility of doing so. However it does not appear to be useful to express properties independent of the mode of representation by means of symbols based on an explicit and determinate form of expression for the function.

[3] The reader will recognize here an alternative formulation of the Cauchy–Riemann equations.

In order to read Riemann's article the following must be kept in mind: by a "function of x and y" he means a function without implicitly understood properties; but by a "*function of* $x+iy$" he means a holomorphic function, in both cases allowing the function to be many-valued or even discontinuous. The following excerpt from his thesis [Rie1851, §5] clarifies the type of discontinuities he had in mind, and is also interesting for the light it sheds on the meaning he gives the phrase "*in a general manner*":

> A variable quantity which, in a general manner, that is, without excluding exceptional isolated points or lines, at every point O of a surface T takes on a definite value varying in a continuous way with the position of the point, can clearly be regarded as a function of x, y, and henceforth whenever functions of x, y are being discussed, this definition is to be understood.

II.1.2. Riemann surfaces

The modern definition. — Nowadays a Riemann surface is defined as a complex manifold of dimension 1:

Definition II.1.1 (**Riemann surface**). A *Riemann surface* is a (connected, Hausdorff) topological space X endowed with an *atlas* $\{(U_\lambda, \phi_\lambda)\}_{\lambda\in\Lambda}$ where $(U_\lambda)_{\lambda\in\Lambda}$ is an open cover of X and the maps $\phi_\lambda : U_\lambda \to V_\lambda$ are homeomorphisms to open sets of $\mathbb{C}$ (the *charts* of the atlas), such that the composite maps[4]

$$\phi_\lambda \circ \phi_\mu^{-1} : \phi_\mu(U_\lambda \cap U_\mu) \to \phi_\lambda(U_\lambda \cap U_\mu)$$

are biholomorphic transformations (that is, holomorphic bijections).

Furnished with this definition one can immediately extend local properties and objects from $\mathbb{C}$ to any Riemann surface; in particular the concepts of a holomorphic or meromorphic[5] function or form on a Riemann surface, and holomorphic and biholomorphic mappings (isomorphisms) between such surfaces now acquire meaning.

Gauss's theorem of Chapter I now provides us with a plentiful supply of examples: we can re-interpret that theorem as asserting that every analytic real Riemannian metric on an analytic surface furnishes it with the structure of a Riemann surface. The interplay between this structure and the geometry arising from the metric will play a leading role in the work of Klein considered in the next chapter.

[4]Coordinate changes. *Trans*

[5]A *meromorphic* function is a function that is locally the quotient of two holomorphic functions. It can be interpreted as a holomorphic function taking its values in $\overline{\mathbb{C}}$. *Trans*

The Riemann sphere. — Apart from $\mathbb{C}$ and its open subsets the first examples of Riemann surfaces that come to mind are the tori $\mathbb{C}/\Lambda$ met with in the introduction, and the *Riemann sphere*: indeed, one can cover the unit sphere

$$\mathbb{S}^2 := \{(X,Y,Z) \in \mathbb{R}^3 \mid X^2 + Y^2 + Z^2 = 1\}$$

by the two open sets $\mathbb{S}^2 \setminus N$ and $\mathbb{S}^2 \setminus S$ (where $S = (0,0,-1)$ and $N = (0,0,1)$ are the south and north poles), on which one defines the *stereographic projections*

$$\begin{array}{cccc} \varphi_N & : & \mathbb{S}^2 \setminus N & \to & \mathbb{R}^2 \simeq \mathbb{C} \\ & & P = (X,Y,Z) & \mapsto & \frac{X+iY}{1-Z} \end{array}$$

and

$$\begin{array}{cccc} \varphi_S & : & \mathbb{S}^2 \setminus S & \to & \mathbb{R}^2 \simeq \mathbb{C} \\ & & P = (X,Y,Z) & \mapsto & \frac{X-iY}{1+Z}. \end{array}$$

For a point P of the sphere other than the poles, one checks that $\varphi_N(P) = 1/\varphi_S(P)$; since $z \mapsto 1/z$ is a holomorphic function on $\mathbb{C}^*$, this furnishes the sphere with the structure of a Riemann surface, denoted by $\overline{\mathbb{C}}$, which can be thought of as the natural compactification of $\mathbb{C}$ by a point at infinity, or, equivalently, as the complex projective line $\mathbb{CP}^1$. These two notations for the Riemann sphere will recur throughout the book.

As recounted in [Cho2007, p. 98], the construction of the Riemann sphere by means of stereographic projections appeared first in print in [Neum1865], the first textbook devoted to the theory of Riemann surfaces. In the introduction to his book Neumann mentions that Riemann taught the above construction, which was then handed down only orally.

The disc, the plane, the sphere, and their automorphisms. — It follows from the uniformization theorem that the disc $\mathbb{D}$, the plane $\mathbb{C}$, and, lastly, the Riemann sphere $\overline{\mathbb{C}}$ are, up to isomorphism, the only simply connected Riemann surfaces. We now describe the automorphism groups of these three surfaces.

Firstly, taking

$$\mathbb{D} := \{z \in \mathbb{C} \mid |z| < 1\},$$

the map

$$z \mapsto w = i\frac{1+z}{1-z} \qquad \text{(II.2)}$$

is a holomorphic isomorphism from $\mathbb{D}$ onto the *upper half-plane*

$$\mathbb{H} := \{w \in \mathbb{C} \mid \operatorname{Im} w > 0\}.$$

It follows that the automorphism groups of $\mathbb{D}$ and $\mathbb{H}$ are isomorphic. Thus by means of conjugation by the transformation (II.2) one can pass from the action of an automorphism on $\mathbb{D}$ to that of the corresponding automorphism on $\mathbb{H}$. The model $\mathbb{H}$ has the advantage that one can easily see that its group of automorphisms is isomorphic to $\mathrm{PSL}(2,\mathbb{R}) := \mathrm{SL}(2,\mathbb{R})/\{\pm I\}$. The precise action is as follows: an element $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$ acts on $\mathbb{H}$ on the left according to the rule

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot w = \frac{aw+b}{cw+d}.$$

The automorphism group of $\mathbb{C}$ is simply the group $\mathrm{Aff}(\mathbb{C})$ of complex affine transformations of $\mathbb{C}$:

$$(a,b)\cdot z = az+b,$$

where $a \in \mathbb{C}^*$ and $b \in \mathbb{C}$.

In the case of $\overline{\mathbb{C}}$, the automorphism group is $\mathrm{PSL}(2,\mathbb{C})$, acting on the left by the rule

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

The latter transformations are called *homographies*.[6]

These three automorphism groups are variously transitive:

1. $\mathrm{Aut}(\mathbb{D})$ is 1-transitive and each of its elements is completely determined by its action on an arbitrary point of $\mathbb{D}$ and an arbitrary point of the boundary $\partial\mathbb{D}$ (to which the group action extends by continuity).

2. $\mathrm{Aut}(\mathbb{C})$ is 2-transitive and each of its elements is completely determined by its action on any two distinct points of $\mathbb{C}$.

3. $\mathrm{Aut}(\overline{\mathbb{C}})$ is 3-transitive and each of its elements is completely determined by its action on any three distinct points of $\overline{\mathbb{C}}$.

Many-valued functions and Riemann surfaces. — Our definition (above) of a Riemann surface is anachronistic: for Riemann these surfaces arose as a means for handling *many-valued functions*. Starting from a holomorphic function defined on an open subset of the plane, he sought to extend its domain of definition by means of analytic continuation. The first sentence of the following quotation announces the procedure of analytic continuation and the second explains how one may by such means be confronted with the problem of many-valuedness. It is precisely this situation that justifies the introduction of the term "many-valued function", which is really not a function at all in the modern set-theoretic sense.

[6]Or "Möbius transformations" or "linear-fractional transformations". *Trans*

> A function of $x + yi$ given on a part of the (x, y)-plane can be extended continuously beyond [that region] in just one way. [. . .] Now, depending on the nature of the function being extended, either it will or will not always take on the same value at a single z-value independently of the path along which the continuation was performed.
>
> In the first case I call the function *single-valued*: such a function is then precisely determined for each value of z and it never becomes discontinuous along any line. In the second case, where we call the function *many-valued*, one must first of all, in order to grasp how it develops, pay attention to certain points of the z-plane around which the function extends onto another [plane]. Such a point is, for example, the point a for the function $\log(z-a)$.

The points at which the value of the function varies with the path along which analytic continuation is carried out are so important in the sequel that Riemann gives them a name:

> We will call the various extensions of a single function over the same region of the z-plane the *branches*[7] of the function, and a point near which one branch extends onto another a *branch point* of the function. Wherever there is no branching the function is to be called *monodrome* or *single-valued*.

After explaining the types of functions he will be considering, he introduces the surfaces now bearing his name, repeating a construction appearing in his thesis [Rie1851]. What's novel here are the intuitive pictures he proposes, of an "*infinitely thin body*" and of a " *helicoid*" of "*infinitely narrow thread*":

> Imagine a surface extended above the (x, y)-plane and coincident with it (or if one likes a body infinitesimally thin [spread] over the plane), which extends exactly as far as the function is given. When the function is extended, this surface is to be continued equally far. In a region of the plane where the function has two or several continuations this surface will be double or multiple. It is thus made up of two or more sheets each of which corresponds to a branch of the function. Near a branch point of the function one sheet of the surface extends onto another in such a way that in a neighborhood of this point the surface can be considered as a helicoid with infinitely narrow thread and with axis perpendicular to the (x, y)-plane at that point. However, if the function, after z has traced several turns about the branch point, should again take on its initial value (just as, for example, $(z-a)^{\frac{m}{n}}$, m, n relatively prime, does after z has executed n turns about a), one must assume that the uppermost sheet reconnects with the bottom sheet, passing through the others.

[7] Or "ramifications". *Trans*

The last few lines show that Riemann did indeed picture the surface situated in the 3-dimensional space of common intuition. Or did he use such language merely to facilitate the understanding of his readers, while himself conceiving the surface as an abstract manifold? Whatever the case may be, as is mentioned in [Cho2007, p. 59], Hensel and Landsberg [HeLa1902, p. 91] continue describing the situation in a manner close to that of Riemann:

> Imagine n coordinate planes placed one above the other at an infinitesimally small distance [...] in such a way that their origins and axes are superimposed [...]

Riemann's description of the surface as situated in a space of dimension 3 forces him to talk of sheets which cross each other, which has historically been a source of difficulty for those trying to learn his theory. The fact that these intersections need not and should not be considered is implied by the following property of such a surface:

> A many-valued function admits at each point of a surface which so represents its mode of branching, a *single* determinate value, and can therefore be regarded as a function uniquely determined at the place (of a point) on that surface.

From this it is clear that the surface associated with a many-valued function is considered as a means of resolving the problem of its many-valuedness.

Box II.1: The Riemann surface of a germ of a function

We explain here how one nowadays constructs a Riemann surface associated with a germ of a holomorphic function $f : (\mathbb{C}, x) \to \mathbb{C}$.[a] Let $\mathcal{G} = \left\{\text{germs of holomorphic functions } (\mathbb{C}, x) \to \mathbb{C} \,\middle|\, x \in \mathbb{C}\right\}$. We first define a Hausdorff topology on this set. For each open set U of $\mathbb{C}$ and each holomorphic function $f : U \to \mathbb{C}$, we define

$$\mathcal{U}(U, f) = \left\{\text{germs } f_x : (\mathbb{C}, x) \to \mathbb{C} \,\middle|\, x \in U\right\}$$

and we endow $\mathcal{G}$ with the topology generated by the $\mathcal{U}(U, f)$.

[a] That is, the set (equivalence class) of all holomorphic functions g agreeing with f on some open neighborhood (depending on g) of x. *Trans*

It is then immediate that the map

$$\pi : \begin{array}{ccc} \mathcal{G} & \to & \mathbb{C} \\ (f_x : (\mathbb{C}, x) \to \mathbb{C}) & \mapsto & x \end{array}$$

is continuous and that its restriction to each open set $\mathcal{U}(U, f)$ defines a local homeomorphism. These local homeomorphisms then allow us to endow $\mathcal{G}$ with the structure of a one-dimensional complex manifold (leaving aside for the moment the requirement of a countable open basis).

This topology is Hausdorff. To see this, note first that two germs based at distinct points are already separated by the continuous function π. Consider two germs $f_x : (\mathbb{C}, x) \to \mathbb{C}$ and $g_x : (\mathbb{C}, x) \to \mathbb{C}$ at x, and let U be a connected open set of $\mathbb{C}$ containing x such that f_x and g_x are the germs of $f, g : U \to \mathbb{C}$. If there were a germ $h_y : (\mathbb{C}, y) \to \mathbb{C}$ in the intersection $\mathcal{U}(U, f) \cap \mathcal{U}(U, g)$, the functions f and g would coincide on an open set contained in the domain of h, whence $f_x = g_x$. Thus if $f_x \neq g_x$, the open sets $\mathcal{U}(U, f)$ and $\mathcal{U}(U, g)$ must be disjoint, and so serve to separate the two germs.

Now let $f_x : (\mathbb{C}, x) \to \mathbb{C}$ be a germ of a holomorphic function. The *Riemann surface of* f_x is then defined to be the connected component $\mathcal{S}(f_x)$ of $\mathcal{G}$ containing f_x. The germs $g_y : (\mathbb{C}, y) \to \mathbb{C}$ in $\mathcal{S}(f)$ are obtained by analytic continuation of $(f_x : (\mathbb{C}, x) \to \mathbb{C})$ along a path joining x to y. In particular, if $f_x : (\mathbb{C}, x) \to \mathbb{C}$ is a germ of a function f that is many-valued in a neighborhood of x, the surface $\mathcal{S}(f_x)$ will contain a point "above" x (that is, in $\pi^{-1}\{x\}$) for each value of f at x. Thus the surface $\mathcal{S}(f_x)$ comes with a (single-valued) holomorphic map $\hat{f} : \mathcal{S}(f) \to \mathbb{C}$ determining f.

The Poincaré–Volterra theorem guarantees that the surface $\mathcal{S}(f)$ has a countable open basis (see Box XI.1).

One can imitate the above construction of $\mathcal{S}(f)$ for other regularity classes of germs. For example, one may construct in the same way the maximal meromorphic continuation. More generally, this procedure can be extended to a sheaf on a topological space[a] and what was called around 1950 the associated "étale space" of the sheaf.

[a] A "sheaf" on a topological space X is a structure associating with each open set U of X an Abelian group or ring (usually of functions defined on U), equipped with a restriction operation satisfying certain conditions. *Trans*

The Riemann surface associated with an algebraic function. — In this section we consider the Riemann surface associated with an algebraic function $s(z)$. The graph of s, in $\overline{\mathbb{C}}_z \times \overline{\mathbb{C}}_s$, is determined by an irreducible polynomial equation $F(z, s) = 0$; such an equation defines an irreducible *algebraic curve* in $\overline{\mathbb{C}}_z \times \overline{\mathbb{C}}_s$.

Although Riemann never used such geometrical terminology in his article, he must have been aware of the geometric interpretation, as Klein explains in [Kle1928].

> From the beginning Riemann recognized the importance of his theory for algebraic geometry. However, in his courses he went into detail only in the case of quartics. This came out only much later from an examination of his lecture notes. It required a much more extroverted nature to establish his results on a broader basis and introduce them to a wider readership. It was Clebsch who understood this.

Riemann proposed “determining the mode of branching of the function s or of the surface T representing it”. Initially T is the Riemann surface of the regular part of the function, that is, the maximal analytic continuation of any of its regular (single-valued and holomorphic) germs. Next he shows that there exists a unique smooth compactification of T obtained as follows. He first defines the simplest possible branch points on the surface T:

> A point of the surface T where just two branches of a function join in such a way that near this point the first branch continues into the second and the second into the first, I will call a *simple branch point*.

(We recognize here a branching like that of the two-valued function $\sqrt{z}$ at the origin.) Every other branch point is regarded as the limit of simple branch points:

> A point around which the surface turns about itself $(\mu+1)$ times can then be considered as consisting of μ coincident (or infinitely close) simple branch points.

He then introduces *local parameters* (or, as we also call them nowadays, *local uniformizing parameters*) in a neighborhood of every point of the closed surface T, choosing them explicitly as functions of z. Thus in a neighborhood of a point $z = a$ where the surface T does not branch, he chooses $z - a$, and then:

> For a point where the surface T turns about itself μ times, when z is equal to a finite value a, [we choose] $(z-a)^{\frac{1}{\mu}}$ [...]; but at $z = \infty$, it is $(\frac{1}{z})^{\frac{1}{\mu}}$, which becomes infinitely small to the first order.

He next explains how to use such a local parameter to develop in series “*the functions we shall be dealing with here*”, which is to say meromorphic functions and their integrals.

Here we see that Riemann *desingularizes* the curve defined by the equation $F(z, s) = 0$ using only local monodromy[8] of the values taken on by s as z varies around each branch point a: a set of μ branches around a are given simultaneously by one and the same meromorphic function of $(z - a)^{\frac{1}{\mu}}$. Each irreducible local component of the curve is thus parametrized by a disc, namely the image of $\{|z - a| < \varepsilon\}$ under the map $(z - a)^{\frac{1}{\mu}}$. Of course, a point has been added to T above a in order to compactify the irreducible local component. In this way Riemann by-passes the so-called algorithm of Newton–Puiseux (which, moreover, he does not refer to). The uniqueness of the resulting compactification is immediate from another theorem of Riemann, namely that on removable singularities, according to which a holomorphic function bounded on a punctured disc can be extended holomorphically to the missing point.

From all this it follows that the Riemann surfaces associated with two birationally equivalent algebraic curves (see Subsection II.3.1) are isomorphic: after the removal of a finite number of points on each of them, the given birational map will define an isomorphism between the punctured surfaces, which then extends automatically to an isomorphism between compact surfaces.

Thus does Riemann open the way to the modern abstract notion of a Riemann surface, with all local parameters obtained from each other by means of biholomorphisms considered equivalent.

Algebraicity of compact Riemann surfaces. — We saw in Box II.1 that every germ of a holomorphic function f can be associated in a natural way with a Riemann surface $\mathcal{S}(f)$. When the function is algebraic, this surface compactifies into a compact Riemann surface — its maximal meromorphic analytic continuation. We now wish to consider the converse: if the maximal meromorphic analytic continuation of f is compact, then f is algebraic.

In anticipation of the Riemann–Roch theorem (see Section II.2.4, Corollary II.2.13) we remark that every (abstract) Riemann surface carries enough meromorphic functions to separate its points. This allows one to prove the following theorem.

Theorem II.1.2. — *Every compact Riemann surface T is isomorphic to the Riemann surface of an algebraic function.*

Proof. — Let f_1 be a non-constant meromorphic function on T, and consider f_1 as defining a branched covering of $\overline{\mathbb{C}}$ of degree d.[9] Let $\{P_1, \ldots, P_d\}$ be a generic fiber of the covering, and f_2 a meromorphic function separating these d points.

[8]"Monodromy" is a general term for the change in an appropriate mathematical object with variation around a singularity. *Trans*

[9]This is a consequence of the compactness of T, which allows one to trace the various branches of the inverse. *Trans*

The image of S under (f_1, f_2) is an analytic curve C in $\overline{\mathbb{C}}_z \times \overline{\mathbb{C}}_w$. We wish to show that this curve is algebraic.

Since the non-constant function $f_1 : T \to \overline{\mathbb{C}}_z$ defines a branched covering, the fiber $f_1^{-1}(z)$ over z always consists of the same number d of points of T except for a finite number of points $z_1, \ldots, z_k$ of $\overline{\mathbb{C}}_z$. For $z \in \overline{\mathbb{C}}_z - \{z_1, \ldots, z_k\}$, we write $f_1^{-1}(z) = \{P_1(z), \ldots, P_d(z)\}$. It is important to observe that the $P_i(z)$ are many-valued: the set $\{P_1(z), \ldots, P_d(z)\}$ is well-defined but it is not possible to arrange the preimages so as to obtain d holomorphic functions globally defined on $\overline{\mathbb{C}}_z - \{z_1, \ldots, z_k\}$. The ordinates of the d points where the line $\{z\} \times \overline{\mathbb{C}}_w$ meets C are $w_i(z) = f_2(P_i(z))$, i varying from 1 to d. Once again we obtain d "functions" w_i, many-valued on $\overline{\mathbb{C}}_z - \{z_1, \ldots, z_k\}$. We now consider the basic symmetric expressions in the $w_i(z)$:

$$\begin{aligned} S_1(z) &= w_1(z) + \cdots + w_d(z), \\ S_2(z) &= w_1(z)w_2(z) + \cdots + w_{d-1}(z)w_d(z), \\ &\vdots \\ S_d(z) &= w_1(z) \cdots w_d(z). \end{aligned}$$

These functions are meromorphic on $\overline{\mathbb{C}}_z$, whence they are rational functions[10] of the variable z. The polynomial $F(z, w)$ obtained from $w^d - S_1(z)w^{d-1} + \cdots + (-1)^d S_d(z)$ by multiplying by a suitable polynomial in z to cancel the denominators, vanishes precisely on the curve C. The Riemann surface T is then just the Riemann surface of any germ at which the above polynomial vanishes: these surfaces are both compact and coincide except for a finite number of points. □

Observe that we have shown here that the field $\mathbb{C}(f_1, f_2)$ has degree precisely d over the subfield $\mathbb{C}(f_1)$. The same argument shows that for every meromorphic function g, the field $\mathbb{C}(f_1, g)$ has degree at most d over the same subfield. It follows by the primitive-element theorem[11], that the field generated by f_1, f_2 and g is the same as that generated by f_1 and f_2. We conclude, finally, that the field of meromorphic functions on T is precisely $\mathbb{C}(f_1, f_2)$.

One infers from this that if we choose two other functions f_1' and f_2' as in the above proof, then the resulting curve C' is birationally equivalent to C — effectively since f_1' and f_2' can be expressed as rational functions of f_1 and f_2. We conclude that two isomorphic compact Riemann surfaces give rise to birationally

[10] A *meromorphic* function from the Riemann sphere to itself is a holomorphic function with a finite number of poles, possibly including ∞, and so, by an easy extension of Liouville's theorem that a bounded holomorphic function is constant, must be a rational function. *Trans*

[11] Asserting that if $K \supseteq F$ is a finite field extension and (in particular) the characteristic is zero then there is an element $a \in K$ such that $K = F(a) (= F[a])$. *Trans*

equivalent algebraic curves. It is this that Riemann expounds in Sections XI and XII of [Rie1857], the point of departure for his investigation of moduli.

Theorem II.1.2 can be made more precise:

Theorem II.1.3. — *Every compact Riemann surface S can be immersed*[12] *in $\mathbb{CP}^2$, injectively apart from a finite number of points, and with image an algebraic curve C having as singularities only double points at which the two tangents are distinct.*

To see this one first embeds the given Riemann surface in some projective space $\mathbb{CP}^n$. Such an embedding is given, in projective coordinates, by

$$z \mapsto (1 : f_1(z) : f_2(z) : \cdots : f_n(z))$$

where we have supplemented the earlier functions f_1 and f_2 with further meromorphic functions f_i on S in order to ensure injectivity:

- — if all of the f_i have the same value at some two points of S, one adds another function taking distinct values at those points;
- — if all the f_i have a common critical point on S, one adds a function regular at that point.

One may construct such functions directly from f_1 and f_2 (working in the field they generate), or, better yet, by appealing to the Riemann–Roch theorem. This achieved, a suitable projection $\mathbb{CP}^n \to \mathbb{CP}^2$ affords us the desired immersion.

In fact the Riemann–Roch theorem provides a privileged representation of a compact Riemann surface as an algebraic curve in a projective space. In genus $p \geq 2$, the dimension of this space is $p - 1$ for all non-hyperelliptic curves (compare for example [GrHa1978]).

II.1.3. Theorems of "Analysis Situs"

There remains the major problem of actually defining meromorphic functions and forms on a given Riemann surface. This will be the main goal of Section II.2 below.

Riemann bases his construction of meromorphic functions and forms on what he calls "Dirichlet's principle", which plays a role also in his thesis [Rie1851]. To that end he needs to integrate a closed form $X dx + Y dy$ (that is, for which $\frac{\partial X}{\partial y} = \frac{\partial Y}{\partial x}$), the integration to be carried out along paths on a surface above the (x, y)-plane. He begins this section by declaring that he will be needing results from *Analysis Situs* (that is, topology). What is at issue here is nothing less than a major conceptual leap: to investigate the construction of an algebraic function on a surface using topological methods in relation to the surface. We now expound these ideas.

By means of a special case of Stokes' theorem, Riemann first of all shows that:

> [...] the integral $\int (X\,dx + Y\,dy)$, evaluated along two different paths joining two fixed points, yields the same value when the the union of these two paths forms the complete boundary of part of the surface T.

In modern terminology, the integral of a closed form along a path with fixed end-points depends only on the homology class of the path.

Riemann next introduces a measure of the connectivity of a surface, giving the extent of its departure from simple-connectedness. His definition is the forerunner of that of the Betti numbers with integer coefficients. Here he implicitly assumes his surfaces are compact and connected with non-empty boundary. Faced with a surface without boundary, he begins by removing a disc.

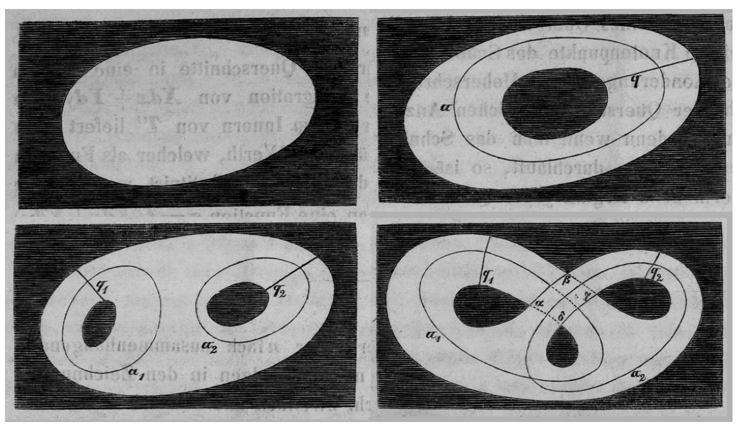


Figure II.1: "Riemann surfaces" (pp. 99, 100 of [Rie1857])

For him a simply connected surface (homeomorphic to a disc) is one with degree of connectivity 1. When a surface is not simply-connected, he performs cuts along sections of it until it becomes simply connected:

[12]An *immersion* of a manifold M in a manifold N is a differentiable map $f : M \to N$ such that the induced map between tangent spaces (the derivative as linear transformation) is injective at each point. Such a map f is an *embedding* if it is also injective. *Trans*

> A multiply connected surface can be transformed, by means of cuts, into a simply connected surface. [...]
> When one can trace n closed curves $a_1, a_2, \ldots, a_n$ on a surface F which, whether one considers them separately or united, do not form the complete boundary of a portion of the surface, but which, when supplemented by any other closed curve do form the complete boundary of a portion of the surface, the surface is said to be $(n+1)$ times connected.

Riemann provides four diagrams intended to aid comprehension of the notions of a multiply connected surface and its *degree of connectivity*. These are the only diagrams in the article [Rie1857]!

How does all this apply to the integration of a closed form of degree 1? After having cut along certain curves of section[13] he obtains a surface representing a simply connected region of the original, so that the closed form is now exact on that region, that is, is the differential of a single-valued function. In passing across each curve of section this function undergoes constant jumps of discontinuity, which Riemann calls *moduli of periodicity*. Nowadays one talks rather of *periods* as the integrals of the closed form around loops, representing, therefore, a concept dual to Riemann's moduli of periodicity. Thus in Figure II.2 the modulus of periodicity corresponding to the transverse section XX' is equal to the period taken along the dual loop l_X $(X = A, B)$.

Box II.2: Simple connectedness

Note how the terminology has evolved: today a surface is called *simply connected* if every loop on the surface is homotopic to a constant loop. However the definition Riemann used is different:

> This gives rise to a distinction among surfaces into simply connected ones, where every closed curve completely bounds a portion of the surface [...] and multiply connected ones, where this is not the case.

A modern reader will see here a homological definition: a surface is simply connected if every loop bounds a subsurface. In higher dimensions this definition (which is just that $H_1(X,\mathbb{Z}) = 0$) is weaker than that given above (equivalent to $\pi_1(X) = 0$ — weaker since we know that $H_1(X,\mathbb{Z})$ is always the Abelianization of $\pi_1(X)$). However for a surface the two definitions are equivalent.

[13]That is, simple closed paths on the surface. *Trans*

An important consequence of simple connectedness is the vanishing of the first homology group: every closed 1-form on a connected and simply connected surface is exact.

To conclude this parenthetical terminological discussion, we note that in 1905 Poincaré was still not using the term "simply connected" in its modern sense. For him a compact manifold of dimension 3 is "simply connected" if it is homeomorphic to a ball. Thus, as he stated it, the famous *Poincaré conjecture* sounds rather odd to a modern ear:

> Is it possible for the fundamental group of V to reduce to the identity substitution, and yet V not be simply connected?

Here is what Riemann says:

> When the surface T [...] is n-connected, one can decompose it into a simply connected surface T' by means of n transverse sections. [...] one obtains a function of x, y, $z = \int (X dx + Y dy)$ completely determined at every point of T' and varying continuously throughout the interior of T', but which in crossing one of the transverse sections varies in general by a finite amount all along the line leading from one vertex of the network of sections to the next.

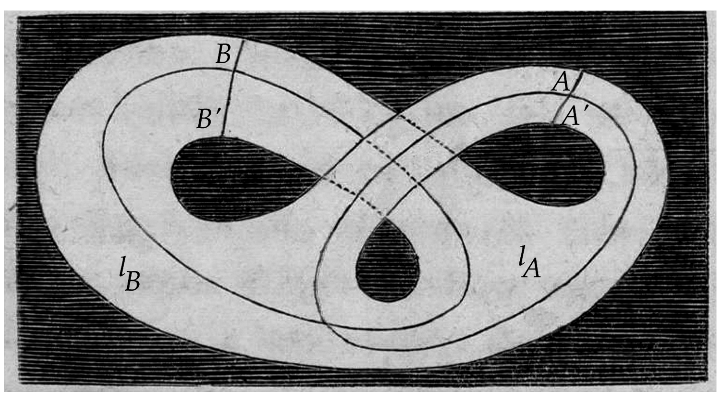


Figure II.2: Moduli/periods

Here we are confronted with a second method (the first being that described above of constructing the associated Riemann surface) for passing from a many-valued function to a single-valued one, namely, just that of making a choice of a particular value of the function on each sub-region of the full domain of definition. Both methods are used throughout the article, the first in dealing with an algebraic function and the second in dealing with an Abelian integral.

One might well ask what prevents Riemann from applying the first method in the context of Abelian integrals. In this case he would have had to describe a branched covering of the complex plane of infinite degree, which situation he might have illustrated with the example of $\log(z-a)$, used in Section 2.1 to explain the phenomenon of many-valuedness. However when he explains how one should think about branched coverings, his illustrative examples are just the coverings of finite degree associated with expressions of the form $(z-a)^{\frac{m}{n}}$. Was he in some sense wary of infinite-degree coverings?

Note also that in the same section he considers *differentials* $X dx + Y dy$ as real objects, while never throughout the article talking of the analogous complex objects (that is, of holomorphic or meromorphic forms). Further on, when he turns to Abelian integrals, the problem is, in modern language, that of finding (many-valued) primitives of meromorphic forms on the surface in question.

For the particular case of a *closed surface* — that is, compact, connected and without boundary — Riemann introduces the topological invariant that today we call the *genus*.

> Let us imagine [...] we have decomposed that surface into a simply connected surface T'. Since the boundary curve of a simply connected surface is uniquely determined, whereas a closed surface has, as the result of an odd number of sections, an even number of bounded regions, and as a result of an even number of sections an odd number of bounded regions, in order to effect this decomposition of the surface it is necessary to execute an even number of sections. Let $2p$ be the number of such transverse sections.

Box II.3: Degree of connectivity, genus, and Euler characteristic

For any compact, connected, orientable (topological) surface S there are two particular topological invariants available: the *Euler characteristic* $\chi(S)$ and the *genus* $g(S) \geq 0$.

When S is without boundary these two invariants are linked by the formula

$$\chi(S) = 2 - 2g(S).$$

The genus $g(S)$ has the following interpretations:

- as equal to $\frac{1}{2}\text{rank}_{\mathbb{Z}} H_1(S)$;[a]
- as the largest number of homologically independent, pairwise disjoint, simple closed curves that can be drawn on S.

If the boundary of S is non-empty, the genus $g(S)$ is defined as the latter number after S has been modified as follows: for each component of the boundary ∂S of S, attach a disc with boundary identified with that component. One then has the formula

$$\chi(S) = 2 - 2g(S) - b(S),$$

where $b(S)$ denotes the number of components of ∂S.

Using this formula one can show that, if $c(S)$ denotes the degree of connectivity of S introduced by Riemann, then

$$c(S) = 2 + 2g(S) - b(S).$$

Figure II.3 shows the $c(S) - 1$ steps of the surgery yielding a disc in Riemann's fourth example (in Figure II.1 above), whence one sees via the latter formula that S is a surface of genus 1.

Many details on the evolution of the notion of genus may be found in [Pop2012].

[a] The rank of the Abelian group $H_1(S)$. *Trans*

Here Riemann is tacitly assuming that the surfaces he considers are all *orientable*. In fact in the case of a non-orientable surface, if one cuts along a simple closed curve along which the orientation reverses, one obtains just one "part bounded by the curve". Such a curve has a neighborhood that is a Möbius band, which surface was described explicitly only many years later, in [Möb1886] (see also [Pont1974, p. 108]). In Riemann's article the opposition *orientable/non-orientable* (or *two-sided/one-sided* as it came to be called for a certain time) is never mentioned.

Note the use of the letter p, still largely in use nowadays to denote various notions of genus arising in geometry and algebraic geometry (mainly in the form of arithmetic and geometric genera of curves and surfaces). Riemann himself does not name this invariant; it seems to have been Clebsch who introduced the term "genus" in [Cle1865a].

We now return to the surface T associated with an algebraic function $w(z)$ defined by an irreducible polynomial equation $F(z,w) = 0$. Riemann now proposes calculating its genus g, to which end he establishes the special case of what we now call the Riemann–Hurwitz theorem (see Box II.4) where the target surface of the function is the Riemann sphere. He shows that if the irreducible algebraic curve of bi-degree m,n defined by $F(z,w) = 0$ in $\overline{\mathbb{C}}_z \times \overline{\mathbb{C}}_w$ has as singularities only r double points with distinct tangents, then T has genus $g = (n-1)(m-1)-r$.

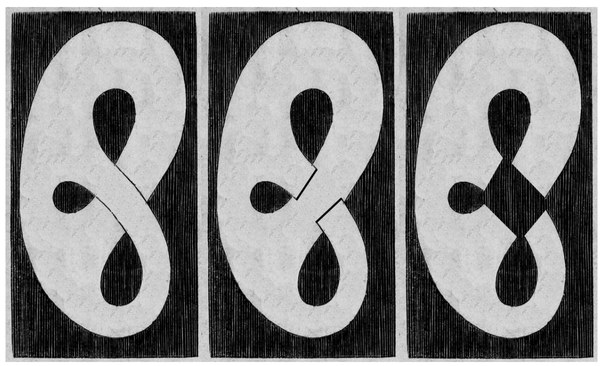


Figure II.3: “Cuts” (or “sections”)

It follows in much the same way that a curve in $\mathbb{CP}^2$ defined by a polynomial of degree n having as its only singularities r double points at which the two tangents are distinct, has genus

$$g = \frac{(n-1)(n-2)}{2} - r. \qquad \text{(II.3)}$$

Box II.4: The Riemann–Hurwitz theorem

Let S and S' be two compact, connected Riemann surfaces and f a holomorphic mapping from S to S'. A point $s \in S$ at which $df = 0$ is called a *critical point* of f, and the image of such a point under f a *branch point* of f.

With each point $s \in S$ we associate its *ramification index* $\nu(s) \geq 1$, defined as the local degree[a] of f in a neighborhood of s. There then exist local coordinates z in some neighborhood of s and w in some neighborhood of $f(s)$,

[a]That is, the number of preimages of individual points in $f(U)$, U some small neighborhood of s. *Trans*

where f takes the form $w = z^{\nu(s)}$. The critical points of S are then just those points of ramification index at least 2.

If f has no critical points (in which case the covering defined by f is unramified) the genera g of S and g' of S' and the global degree d of f are linked by the simple formula $2 - 2g = d(2 - 2g')$.

The *Riemann–Hurwitz theorem* generalizes this to the situation where there are critical points (finite in number in view of the compactness of S):

$$2 - 2g = d(2 - 2g') - \sum_{s \in S} (\nu(s) - 1).$$

A straightforward proof starts from a triangulation of S' whose vertices include all the ramification points of f. One then lifts this to a triangulation of S, and shows that the Euler characteristic of the latter triangulation equals the right-hand side of the above formula. We have already seen that it coincides with the left-hand side expression.

II.2. Dirichlet's principle and its consequences

II.2.1. Dirichlet's problem

Given an open set $U \subset \mathbb{C}$ and a function $\underline{u} : \partial U \to \mathbb{R}$ — continuous, for example — the *Dirichlet problem* is that of finding a harmonic function $u : \overline{U} \to \mathbb{R}$ defined throughout U and continuously extending $\underline{u}$.

We begin with a basic construction using the fact that the imaginary part Im w of a holomorphic function w is harmonic and identically zero on the real axis. By means of the biholomorphism $\mathbb{H} \to \mathbb{D}$; $w \mapsto z = \frac{w-i}{w+i}$, which maps the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} | \text{Im } z > 0\}$ onto the unit disc $\mathbb{D} = D(0,1)$, we obtain a harmonic function defined on the disc, namely $z \mapsto \frac{1-|z|^2}{|1-z|^2}$, which is thus automatically harmonic and extends onto $\partial\mathbb{D} \setminus \{1\}$ as the zero function there. By the mean-value property of harmonic functions we must have, for any disc $D(0,r)$, $0 < r < 1$,

$$\frac{1}{2\pi r} \int_{\partial D(0,r)} f = f(0) = 1.$$

Our function f is thus harmonic on the open unit disc and seems to extend continuously to the zero function on the boundary with the point 1 deleted, while at the same time having integral 1 over that boundary. We have here a "point charge" or "Dirac mass" at 1.

This observation together with the linearity of the Dirichlet problem allows us to retrieve Poisson's formula, providing the solution to the Dirichlet problem for the unit disc.[14] Indeed, given a continuous function $\underline{u} : \partial\mathbb{D} \to \mathbb{R}$, an harmonic extension u of $\underline{u}$ to the whole of $\mathbb{D}$ is:

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-|z|^2}{|z-e^{i\theta}|^2}\underline{u}(e^{i\theta})\, d\theta. \tag{II.4}$$

And when the boundary function $\underline{u}$ is not continuous but merely integrable, the extension u will still satisfy, for radial limits,

$$\lim_{r\to 1} u(re^{i\theta}) = \underline{u}(e^{i\theta})$$

for almost all angles θ (in the sense of Lebesgue measure).

Our present aim is to solve the Dirichlet problem for a simply connected open set U of the plane with boundary ∂U a smooth Jordan curve, and for any continuous function $\underline{u} : \partial U \to \mathbb{C}$.

We begin by remarking that there exists at most one solution. For if u_1 and u_2 are two solutions, the function $u_1 - u_2 : \overline{U} \to \mathbb{R}$ is bounded and harmonic on U. Let z_0 be a point of $\overline{U}$ satisfying:

$$|u_1(z_0) - u_2(z_0)| = \max_{\overline{U}} |u_1 - u_2|.$$

If $z_0 \in U$, $u_1 - u_2$ must be constant by the maximum principle for harmonic functions (see earlier), so equal to zero since it vanishes on the boundary of U. If z_0 is on the boundary of U, then $u_1(z_0) = u_2(z_0)$, and it is immediate from the maximality property of $|u_1(z_0) - u_2(z_0)|$ that $u_1 = u_2$ on U.

From Section 16 to Section 18 of [Rie1851], Riemann explains how to solve Dirichlet's problem by minimizing a certain functional. He starts with a smooth function $\alpha : \overline{U} \to \mathbb{C}$ satisfying $\alpha = \underline{u}$ on the boundary of U. Then he adds a function λ vanishing on the boundary and seeks to arrange that $\alpha + \lambda$ be harmonic. Such a function λ will minimize the integral

$$\Omega(\alpha+\lambda) = \int_U \left(\frac{\partial\alpha}{\partial x} + \frac{\partial\lambda}{\partial x}\right)^2 + \left(\frac{\partial\alpha}{\partial y} + \frac{\partial\lambda}{\partial y}\right)^2 dxdy.$$

Thus the problem arises as to whether the functional $\lambda \mapsto \Omega(\alpha+\lambda)$ has a minimum. Write $L = \int_U (\frac{\partial\lambda}{\partial x})^2 + (\frac{\partial\lambda}{\partial y})^2 dxdy$. We quote from Riemann's text where

[14]Poisson's formula seems to have been unknown to Riemann and his immediate successors. Schwarz presents the formula as if new in [Schw1870a]. According to [Die1978], Green was the first to show, in 1828, that a continuous function of the points of any (simple, closed) curve extends to a harmonic function in the interior. In the case of a sphere Poisson gave an explicit formula "in 1820". Prym, in his 1871 commentary on Riemann's works mentions that the only known method of extending a function harmonically into the interior of a circle is by developing it in a Fourier series, even though the convergence of the series is not guaranteed by continuity alone [Pry1871].

he justifies the existence of such a minimum of the functional Ω:

> For each function λ, Ω takes on a finite value tending to infinity with L and varying in a continuous manner with the form of λ, but is bounded below by 0; hence for at least one value of the function $\alpha + \lambda$, the integral Ω attains a minimum value.

It is the purported existence of a function realizing this minimum that Riemann calls the "Dirichlet principle". We must stress here the conceptual leap that this form of the principle represents: a function is considered implicitly as a particular point of an infinite-dimensional space.

Riemann next shows that for every function λ_0 minimizing the integral $\Omega(\alpha + \lambda)$, the function $\alpha' = \alpha + \lambda_0$ is harmonic. He thinks he has thus solved the Dirichlet problem.

Riemann's argument concerning the existence of a minimum is, however, not rigorous — and not only in the eyes of a 21st century reader: Weierstrass criticized the argument already in [Weie1870]. The reader may also consult Volume II of the *Traité d'analyse* [Pic1893d, p. 38] where Picard revisits Weierstrass's criticisms, as well as giving the counter-example appearing in [Weie1870] of a functional which does not attain its greatest lower bound. As Picard says ([Pic1893d], p. 39):

> One cannot be certain *a priori* that there exists a function u satisfying continuity conditions, at which the integral actually attains its lower limit. This is a serious objection and M. Weierstrass has shown by means of a very simple example the danger of this kind of reasoning.

Here is Weierstrass's counter-example. He considers the space of functions $y(x)$ of class C^1 on the interval $[-1, 1]$ with values at the end-points equal to a and b (with $a \neq b$), and introduces the functional defined by

$$J(y) = \int_{-1}^{1} x^2 \left(\frac{dy}{dx}\right)^2 dx.$$

It is not difficult to check that, for the family of functions

$$y_\varepsilon(x) = \frac{a+b}{2} + \frac{(b-a)\arctan\left(\frac{x}{\varepsilon}\right)}{2\arctan\left(\frac{1}{\varepsilon}\right)},$$

one has that $J(y_\varepsilon)$ tends to 0 with ε. The greatest lower bound of J is thus 0, which is not attained at any function in the given space since $a \neq b$. This is made possible by the fact that the space $C^1([-1, 1])$ is not complete. Note that the space of functions with which Riemann is working here — consisting of the functions continuous on $\overline{U}$ and smooth in the interior — is likewise not complete.

The modern method of skirting this obstacle, conceived in a famous 1900 paper of Hilbert [Hil1900a], is to work in a larger space of functions which *is* complete; see for example [Jos2002].

II.2.2. The Riemann mapping theorem (or conformal representation theorem)

We first quote Riemann's own statement of the conformal representation theorem:

> Any two given simply connected, planar surfaces can always be mapped one to the other in such a way that to each point of one there corresponds a unique point of the other whose position varies in a continuous manner with that of the first, and such that the smallest corresponding portions of the surfaces are similar; furthermore, for a point of the interior and for a point of the boundary of one surface, the corresponding points of the other surface may be given arbitrarily; but then this determines the correspondence for all points.

The modern statement of this theorem is more general since it incorporates regularity conditions on the boundaries. Recall that a *Jordan curve* is any continuous embedding of the circle in the plane.

Theorem II.2.1. — *Let U be any simply connected open set in the plane, not equal to the whole plane. Then there exists a biholomorphic mapping $f : U \to \mathbb{D}$. Furthermore if the boundary of U is a Jordan curve, then f extends to a homeomorphism from the closure of U onto the closed unit disc.*

Note that Riemann implicitly assumes that the boundary — which he calls the "frame" of the surface — is a Jordan curve since he defines the images of the boundary points. In the present subsection we give a proof of the conformal representation theorem (or "Riemann mapping theorem") directly inspired by Riemann's ideas. We shall always assume the boundaries to be Jordan curves — in fact even smooth Jordan curves. (The methods of proof when the boundary is not a Jordan curve are different; see for example [Rud1987, Chapter 4].)

Proof of the first statement of the Riemann mapping theorem (assuming we know how to solve the Dirichlet problem): Let $U \subset \mathbb{C}$ be open and simply connected and let z_0 be any particular point of U. We begin with a definition.

Definition II.2.2.[15] — A *Green's function* for U relative to the point z_0 is a function $u : U \setminus \{z_0\} \to \mathbb{R}$ with the following properties:

1. u is harmonic on the open set $U \setminus \{z_0\}$;
2. the function $z \in U\setminus\{z_0\} \mapsto u(z)+\log|z-z_0|$ extends to a function harmonic at z_0;
3. $u(z)$ tends to 0 as z approaches the boundary of U.

Note that there exists at most one such function: this follows in much the same way as the uniqueness of a solution of Dirichlet's problem. A similar argument shows also that a Green's function u must be strictly positive on U. Indeed, if u assumed a non-positive value at a point z_1 of U, then by invoking the facts that $\lim_{z\to z_0} u(z) = +\infty$ and $\lim_{z\to\partial U} u(z) = 0$, we could infer that u attained its minimum on $U \setminus \{z_0\}$, whence, by virtue of its harmonicity u would be constant, which is absurd in view of the fact that it has a logarithmic pole at z_0.

We now show how to construct a Green's function relative to a point $z_0 \in U$ under the assumption that we know how to solve the Dirichlet problem.

Consider the function $\underline{v} : \partial U \to \mathbb{R}$ defined by $\underline{v}(z) = \log|z - z_0|$. Since the Dirichlet problem is assumed to have a solution on U, we have an harmonic extension $v : \overline{U} \to \mathbb{R}$. Write $u_U(z,z_0) = v(z) - \log|z - z_0|$ for z in $U \setminus \{z_0\}$. Then since v is continuous on $\overline{U}$, the function $u_U(\cdot, z_0)$ approaches 0 on the boundary of U. Hence $u_U(\cdot, z_0)$ is the Green's function of U relative to the point z_0.

Example II.2.3. — For the unit disc $\mathbb{D} = \{z \in \mathbb{C} | |z| < 1\}$, the Green's function relative to the point $z_0 = 0$ is

$$u_{\mathbb{D}}(z,0) = -\log|z|.$$

Resuming our proof, we denote by v^* a harmonic conjugate[16] of v, and consider the holomorphic function on U defined by

$$\phi(z) = (z - z_0)\mathrm{e}^{-(v(z)+iv^*(z))}.$$

Since $(z - z_0) = e^{|z-z_0|+i\arg(z-z_0)}$, this may be rewritten "formally" as

$$\phi(z) = \mathrm{e}^{-(u_U(z,z_0)+iu_U(z,z_0)^*)},$$

where $u_U(z,z_0)^* := v^* - \arg(z - z_0)$. Since, as we have already observed, the Green's function is strictly positive, the function ϕ takes its values in the unit disc. It can be inferred from the condition

$$u_U(z,z_0) \longrightarrow 0 \text{ as } z \to \partial U,$$

[15] One may find in [Taz2001] many details on the historical development of this notion.

[16] That is, such that $v + iv^*$ is holomorphic. See earlier. *Trans*

that ϕ is proper.[17] Hence it is surjective (its image being both open and closed) and the cardinality (including multiplicities) of its fibres is constant. Since the fibre above 0 is just $\{z_0\}$, of cardinality 1, the map φ is injective, and we have the desired biholomorphism between U and $\mathbb{D}$.

Remark II.2.4. — The second part of the theorem was proved by Carathéodory in 1916 (see [Coh1967] for example). It is interesting that his proof yields the solution of the Dirichlet problem in the case where the boundary of U is a Jordan curve. The Riemann mapping theorem and Carathéodory's theorem are thus equivalent to the solvability of the Dirichlet problem on U. □

II.2.3. Abelian integrals

Recall that our aim is to construct meromorphic functions on a given surface. Riemann seeks such functions as primitives of meromorphic forms.

We have seen above how to associate a given Riemann surface T with an algebraic function.

> A similar system of algebraic functions with the same ramifications and integrals of the functions will be first of all the object of our study.

In other words, once the surface T has been constructed, one investigates the space of meromorphic forms on T (having the same ramifications as the equation used to construct T) and their primitives.

Here is the title chosen by Riemann to present his vision of the construction of such functions:

> Determination of a function of a variable complex quantity by the conditions it satisfies relative to the boundary and discontinuities.

Thus the functions in question should be determined by their values on the boundary and by their behavior in the neighborhood of discontinuities. The holomorphicity of the function being sought renders all other data superfluous. And it is once again the "Dirichlet principle" which allows one to construct the desired functions starting from a "system of independent conditions among them".

[17]That is, that the preimage of every compact set in $\mathbb{D}$ is compact in U. It follows that φ is a closed map; it is also open since holomorphic. *Trans*

Here is the theorem on which Riemann bases the whole of his theory of functions of a complex variable — a theorem already present in his thesis [Rie1851]:

> If on a connected surface T, decomposed by means of transverse sections into a simply connected surface T', one gives a complex function $\alpha + \beta i$ of x, y for which the integral
>
> $$\int \left[\left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right] dT, \tag{II.5}$$
>
> evaluated over the whole surface, has a finite value, this function can always, and in a *unique manner*, be transformed into a function of $x + yi$ by the subtraction of a function $\mu + \nu i$ of x, y satisfying the following conditions:
>
> 1. On the contour $\mu = 0$, or at least differs from zero only at isolated points; at *one point* ν is given arbitrarily.
> 2. The variation of μ on T and of ν on T' is discontinuous only at isolated points, and then only in such a way that the integrals
>
> $$\int \left[\left(\frac{\partial \mu}{\partial x} \right)^2 + \left(\frac{\partial \mu}{\partial y} \right)^2 \right] dT \text{ and } \int \left[\left(\frac{\partial \nu}{\partial x} \right)^2 + \left(\frac{\partial \nu}{\partial y} \right)^2 \right] dT \tag{II.6}$$
>
> over the whole surface, remain finite; furthermore the variations of ν along a transverse section should be equal on the two sides.

The steps in the proof sketched by Riemann have been explicated in intrinsically modern terminology by Ahlfors in [Ahl1953]. These involve harmonic analysis, and we shall return to them in Section III.1. We now explain briefly the above passage.

To ease our explanation of Riemann's text we denote by the lower-case Roman letters a, b, m, n the differentials (closed but not exact) of the discontinuous functions α, β, μ, ν introduced by Riemann.

If T is a Riemann surface (ultimately with boundary), its real tangent bundle is furnished with an operator J of square -1: multiplication by i. If a is a real differential form of degree 1 on T, its conjugate differential is defined by $*a = -a \circ J$. A form is said to be *co-closed* if its conjugate is closed. We write also $D[a] = \int_T a \wedge *a$, the norm (also called the Dirichlet energy) of a. One then sees that the three integrals in the above excerpt from Riemann's paper are respectively $D[a + *b], D[m], D[n]$.

Given a real closed differential form a on T with its periods, (isolated) singularities, and prescribed values on the boundary, we assume there exists a closed differential form b such that the energy $D[a + *b]$ is finite. One then chooses an exact form m whose restriction to the boundary is zero and whose distance from $a + *b$ in terms of the norm is least.

Here one encounters difficulties in proving the existence of such a form m analogous to those we described in Section II.2.1 in connection with the "proof" of Dirichlet's principle via minimization. As there, so also here in order to circumvent these difficulties one needs to work in a more appropriate function space not available to Riemann.

The existence of m is equivalent to the existence of an orthogonal decomposition $a + *b = m + *n$ with m exact and n closed, whence it follows that $a - m = *n - *b$ is both closed and co-closed, therefore harmonic.

From this the existence of a holomorphic form on T follows: the harmonic form $a - m$ has the prescribed periods, singularities, and values on the boundary (those of a), and writing $u := a - m$ and $v := b - n$, we see that $u + iv$ is a holomorphic form whose integrals yield the "functions of $x + yi$" on T in Riemann's statement.

Riemann now uses his result on the existence of harmonic forms on a given surface to construct meromorphic forms. Starting with a closed Riemann surface T he takes a finite number of points $P_1, \dots, P_m$ of T and in a neighborhood of each of these he takes as *principal part* a finite sum expressed in terms of local parameters z_i:

$$(A_i z_i^{-1} + B_i z_i^{-2} + C_i z_i^{-3} + \cdots) dz_i \tag{II.7}$$

He chooses $2g$ cuts (not passing through any of the P_i) yielding a simply connected surface, and then establishes the existence theorem, which, in modern terminology, is as follows:

Theorem II.2.5 (The existence of meromorphic 1-forms on a surface). — *Assuming the sum of the residues A_i is zero, for each choice of $2g$ real numbers there exists a unique meromorphic form on T with poles at just the points P_i and the given principal parts, and with periods evaluated along the $2g$ cuts having as real parts those prescribed $2g$ numbers.*

The importance of this theorem was recognized well before a perfectly rigorous proof was given. It had a great influence on Riemann's successors, in the forefront of whom were Hermann Schwarz and Felix Klein, whose work will be considered in the following chapters. A modern proof in the spirit of Riemann may be found in [Coh1967], and we shall give another (inspired by [Spr1957]) in Subsection III.2.1.

Certain of the forms figuring in this theorem were destined to play a special role, the so-called forms of the first, second, and third kind. Nowadays a form is said to be *of the first kind* if it is holomorphic, *of the second kind* if it is meromorphic with all residues zero, and, finally, *of the third kind* if it is meromorphic and has only simple poles. The simplest forms of the second kind are then those with just one pole P on T: their many-valued primitives are what Riemann called *integrals of the second kind*, denoted by t_P. The simplest forms of the third kind are those with just two simple poles P_1, P_2, and in this case Riemann called their primitives *integrals of the third kind* ϖ_{P_1,P_2}. His motivation in using such terms derives from their use by Legendre in his classification of elliptic integrals.

We now turn to Riemann's proof. He first shows using Theorem II.2.5 that the complex vector space of integrals of the first kind has dimension $g+1$ (1 more than that of the space of holomorphic forms on account of the constant of integration). This affords an analytic interpretation of the genus g, originally defined topologically. He also shows that such an integral is uniquely determined to within an additive constant by the real parts of the moduli of periodicity relative to a system of transverse sections rendering the surface simply connected.

Similarly, an integral of the third kind is uniquely determined to within an additive constant by the data of the poles, the residues of its differential at these poles, and the real parts of its moduli of periodicity relative to the transverse sections (chosen so as to avoid the poles).

The existence of meromorphic 1-forms on an algebraic curve. — It took until the beginning of the 20th century before Dirichlet's principle and the "proof" imagined by Riemann of the existence of meromorphic 1-forms with prescribed poles on the surfaces bearing his name were given a rigorous foundation. However, then the question had become that of defining such forms on *abstract* Riemann surfaces. In actuality, following on the work of Abel and Jacobi, 19th century mathematicians knew how to construct meromorphic 1-forms explicitly (or rather their many-valued integrals — Abelian integrals) on Riemann surfaces defined as algebraic curves; we will now explain how they did this.

We begin with a compact Riemann surface T. By Theorem II.1.3, T can be immersed in $\mathbb{CP}^2$ as an algebraic curve C with all of its singular points double with distinct tangents. For a suitable choice of affine coordinates we may arrange that the curve C is transverse to the line at infinity, and that in a neighborhood of each double point the first projection $x : C \to \mathbb{CP}^1$ is a coordinate on each branch of the curve.

First we construct holomorphic 1-forms on T. Denote by E the vector space of polynomials $P \in \mathbb{C}[x, y]$ of degree at most $d - 3$[18] which vanish at the double points of C. For each point $P \in E$ we write ω_P for the lift onto T of the Abelian differential

$$P(x, y)\frac{dx}{F'_y}, \tag{II.8}$$

where $F'_y = \frac{\partial F}{\partial y}$.[19]

[18] Where d is the degree of C. *Trans*

[19] Here $F(x, y) = 0$ is the polynomial equation (of degree d) defining the algebraic curve C. *Trans*

Proposition II.2.6. — *1. For every polynomial P in E, the form ω_P is holomorphic on T.*

2. *The map $P \mapsto \omega_P$ from E to the space $\Omega^1(T)$ of holomorphic 1-forms on T is linear and injective.*

3. *The dimension of E is greater than or equal to g, the genus of T.*

Proof. — 1. The formula (II.8) defines *a priori* a holomorphic 1-form on C from which have been removed:

— the points where the first projection $x : C \to \mathbb{CP}^1$ does not define a holomorphic local coordinate, that is, the points of intersection of C with the line at infinity, and the branch points of $x : C \to \mathbb{CP}^1$;

— the points where F'_y vanishes, that is, the double points of C and the branch points of $x : C \to \mathbb{CP}^1$;

— the points where $F(x, y)$ becomes infinite, that is, the points of intersection of C with the line at infinity.

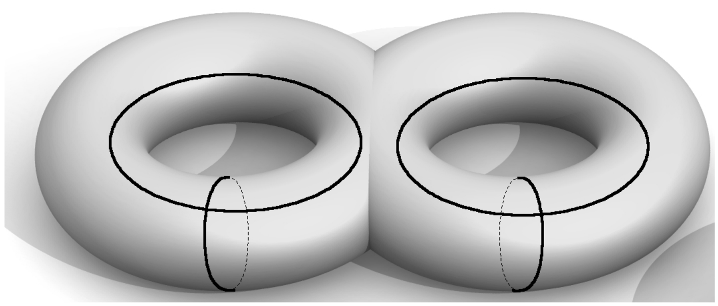


Figure II.4: A symplectic basis for the homology

Next observe that the 1-form given by (II.8) extends holomorphically to the ramification points of $x : C \to \mathbb{CP}^1$; indeed it follows from the identity $F'_x dx + F'_y dy = 0$ that (II.8) can be rewritten as

$$\omega = -\frac{P(x, y)}{F'_x} dy$$

(where this makes sense), and this expression defines a holomorphic 1-form in the neighborhood of every ramification point of x. Then since at each double point of C the polynomial F'_y has a zero of order 1 and the polynomial $P(x, y)$ also vanishes, the lift of the 1-form defined by (II.8) extends to the double points of T. Finally, by means of the change of variables $X = \frac{1}{x}$ and $Y = \frac{1}{y}$, one sees that the 1-form defined by (II.8) extends holomorphically to the points of intersection of C with the line at infinity since the polynomial P has degree at most $d - 3$ (using here the fact that C is transverse to the line at infinity).

2. This statement is immediate.

3. We count dimensions. The polynomials in x, y of degree at most $d-3$ form a vector space of dimension $\frac{(d-1)(d-2)}{2}$. In order for such a polynomial to vanish at all of the r, say, double points of C, its coefficients must satisfy r linear equations. Hence the dimension of the space E is at least

$$\frac{(d-2)(d-1)}{2} - r,$$

which by (II.3) is equal to the genus of T. □

We shall now show that in fact the dimension of E is precisely g. As does Riemann, we fix on $2g$ simple closed paths on T and cut along them so as to obtain a simply connected surface. These are loops representing homology classes on T.

Reverting to modern terminology, we consider the intersection product defined by these loops:

$$H_1(T,\mathbb{Z}) \times H_1(T,\mathbb{Z}) \to \mathbb{Z}.$$

Being bilinear and antisymmetric, it defines a symplectic form. Moreover $H_1(T,\mathbb{Z})$ has a basis which is *symplectic* relative to the intersection product, that is, a basis $(a_1, \ldots, a_g, b_1, \ldots, b_g)$ such that for $i, j = 1, \ldots, g$:

$$a_i \cdot a_j = 0, \quad b_i \cdot b_j = 0, \quad a_i \cdot b_j = \delta_{ij}$$

(see Figure II.4). Each such basis corresponds to a dissection of T into a $4g$-sided polygon. Riemann next shows — with the aid of Stokes's theorem — that for every symplectic basis $(a_1, \ldots, a_g, b_1, \ldots, b_g)$ of $H_1(T,\mathbb{Z})$ and two given closed 1-forms η and η' on T, one has

$$\int_T \eta \wedge \eta' = \sum_{i=1}^{g} \left(\int_{a_i} \eta \int_{b_i} \eta' - \int_{a_i} \eta' \int_{b_i} \eta \right). \tag{II.9}$$

It follows from this that the linear map

$$\left\{ \begin{array}{ccc} \Omega^1(T) & \longrightarrow & \mathbb{C}^g \\ \omega & \longmapsto & \left(\int_{a_i} \omega\right)_{i=1\ldots g} \end{array} \right. \tag{II.10}$$

is injective. It follows in particular from (II.9) that a non-zero holomorphic 1-form ω satisfies

$$i \int_T \omega \wedge \overline{\omega} > 0.$$

That is essentially the proof given by Riemann in Section 20 of [Rie1857]; and it is also the first half of the proof establishing the *bilinear relations of Riemann* (see [Bos1992] for further details).

Instead of Ψ we might have considered the linear map

$$\Phi : \begin{cases} \Omega^1(T) & \longrightarrow & (\mathbb{R}\times\mathbb{R})^g \\ \omega & \longmapsto & \left(\mathrm{Re}\left(\int_{a_i}\omega\right), \mathrm{Re}\left(\int_{b_i}\omega\right)\right)_{i=1\dots g} . \end{cases}$$

The injectivity of this map follows in much the same way, giving the uniqueness assertion of Theorem II.2.5. As far as holomorphic forms are concerned, the existence claim — problematic for Riemann — is made good by Proposition II.2.6 and its proof. The linear map Ψ is thus an isomorphism and we have Theorem II.2.5 in the case of holomorphic forms:

Proposition II.2.7. — *For each g-tuple $\mathbf{n} = (n_1, \dots, n_g)$ of complex numbers, there exists a unique holomorphic 1-form $\omega_{\mathbf{n}}$ on T whose integral along the loop α_i is equal to n_i for $i = 1, \dots, g$.*

Furthermore the 1-form $\omega_{\mathbf{n}}$ depends linearly (so certainly holomorphically) on the g-tuple $\mathbf{n} = (n_1, \dots, n_g)$.

It remains to construct *meromorphic* forms on T. We choose fixed loops $\alpha_1, \dots, \alpha_g$ representing the classes $a_1, \dots, a_g$, and denote their union by A. We shall now prove Theorem II.2.5 for forms of the second and third kind.

We first consider the case of meromorphic 1-forms having only simple poles. Such a form can always be expressed as a linear combination of meromorphic 1-forms each with precisely two simple poles at which the residues are $+1$ and -1. Moreover by adding suitable holomorphic 1-forms if necessary, in view of Proposition II.2.7 we can assume without loss of generality that the integrals of these 1-forms around the loops $\alpha_1, \dots, \alpha_g$ are all zero. We are thus left to prove the following result:

Proposition II.2.8 — *Corresponding to any two distinct points $p, q \in T \setminus A$, there exists a unique meromorphic* 1*-form $\omega_{p,q}$ on T, having simple poles at p and q with residues respectively* $+1$ *and* -1*, and without any other poles, and such that the integral around each of the loops $\alpha_1, \dots, \alpha_g$ is zero.*

Proof. — Consider the vector space $\Omega_{p,q}$ of meromorphic 1-forms on T with simple poles at p and q and no other poles. Write $\Theta : \Omega_{p,q} \to \mathbb{C}^{g+1}$ for the linear map associating with each element of $\Omega_{p,q}$ its integrals around the loops $\alpha_1, \dots, \alpha_g$ and its residue at p (the residue at q being the negative of that at p).

Proving the above proposition is then equivalent to showing that Θ is bijective. We know it's injective since any two elements in the kernel differ by a holomorphic 1-form whose integral around each of the loops $\alpha_1, \ldots, \alpha_g$ is zero. It thus suffices to prove that the dimension of the vector space $\Omega_{p,q}$ is at least $g + 1$.

The proof of this is similar to that of Proposition II.2.6: as in the proof of that proposition one constructs the desired forms on the curve C of degree d of Theorem II.1.3 — the image under an immersion of T in $\mathbb{CP}^2$. We may assume that the images of the points p and q in C do not coincide with any singular point and do not lie on the line at infinity; we continue denoting them by p and q.

Let D denote the line of $\mathbb{CP}^2$ determined by p and q. We choose an equation $(ax + by + c = 0)$ for D and consider those elements of $\Omega_{p,q}$ expressible in the form

$$\omega = \frac{P(x, y)}{(ax + by + c)F'_y} dx, \tag{II.11}$$

for some polynomial $P(x, y)$. The line D intersects the curve C in d points, counted according to their multiplicities; to simplify the argument we shall assume these points pairwise distinct and off the line at infinity. The formula (II.11) defines *a priori* a holomorphic 1-form on the curve C from which the ramification points of the map $x : C \to \mathbb{CP}^1$, the points of intersection of C with the line at infinity, the double points of C, and the points of intersection of C with the line D have been removed. The same reasoning as in the proof of Proposition II.2.6 then shows that the formula (II.11) lifts to an element of $\Omega_{p,q}$ if and only if:

- the polynomial P has degree at most $d - 2$;
- the polynomial P vanishes at each double point of C; and
- the polynomial P vanishes at each of the $d - 2$ points of intersection of C with D distinct from p and q.

The polynomials in the variables x, y of degree at most $d - 2$ form a vector space of dimension $\frac{d(d-1)}{2}$. The vanishing of a polynomial at the r double points of C and the $d - 2$ points of $C \cap D$ distinct from p and q, yields $r + (d - 2)$ linear equations in its coefficients. Hence the dimension of the space $\Omega_{p,q}$ is greater than or equal to

$$\frac{d(d-1)}{2} - r - (d-2) = \frac{(d-1)(d-2)}{2} - r + 1,$$

which by (II.3) is equal to $g + 1$.

A similar (but trickier to expound) count of dimensions yields the same outcome in the case where D has multiple intersection points with C. □

The 1-form $\omega_{p,q}$ given by Proposition II.2.8 "depends holomorphically on the points p, q". A precise meaning can be given to this assertion as follows. Choose an open set $U \subset T$ on which the coordinate x defines an injective map. Then on U the form $\omega_{p,q}$ can be expressed as:

$$\omega_{p,q}(r) = \left(\frac{1}{x_r - x_p} - \frac{1}{x_r - x_q} + G^x_{p,q}(r)\right) dx_r,$$

where x_p, x_q, x_r are the values of the coordinate x at the points p, q, r. Then for every pair $(p,q) \in (U\backslash A)^2$ of distinct points, the function $r \mapsto G^x_{p,q}(r)$ is holomorphic on U. In fact:

Proposition II.2.9. — *The function $(p,q,r) \mapsto G_{p,q}(r)$ with $p \neq q$ is holomorphic in the three variables as a map to*

$$\{(p,q,r) \in (U\backslash A) \times (U\backslash A) \times U | p \neq q\}.$$

Furthermore it extends holomorphically to the diagonal $p = q$.

Proof. — We first repeat the construction of the 1-form $\omega_{p,q}$ in the proof of the previous proposition: that 1-form was given in terms of the x-coordinate by

$$\omega_{p,q} = \frac{P_{p,q}(x,y)}{(ax + by + c)F'_y} dx,$$

where $(ax + by + c = 0)$ was an equation of the line determined by the points p, q, and $P_{p,q}(x,y)$ was a polynomial of degree at most $d - 2$. The coefficients of this polynomial satisfy a system of affine equations made up of the following: $d(d-3)/2 - (g-1) + (d-2)$ linear equations deriving from the fact that $\omega_{p,q}$ belongs to the space $\Omega_{p,q}$, then g linear equations expressing the condition that the integral of $\omega_{p,q}$ around each of the loops $\alpha_1, \dots, \alpha_g$ is zero, and finally one equation from the condition that the residue of $\omega_{p,q}$ at p should be 1. Clearly the coefficients in these affine equations depend holomorphically on p and q. It follows via the uniqueness of $\omega_{p,q}$ and therefore of the polynomial $P_{p,q}$ that this system has maximal rank[20], whence it follows in turn that the polynomial $P_{p,q}$ itself depends holomorphically on the points p and q. The first assertion of the proposition is then immediate.

[20]Since there are altogether $d(d-1)/2$ equations in the same number of coefficients of $P_{p,q}$. *Trans*

We now show that the function $(p,q,r) \mapsto G_{p,q}(r)$ extends holomorphically to the diagonal $p = q$. For pairwise distinct $p,q,t \in U \setminus A$, the uniqueness assertion of Proposition II.2.8 implies that

$$\omega_{p,q} = \omega_{p,t} + \omega_{t,q} \quad \text{and} \quad G_{p,q}(r) = G_{p,t}(r) + G_{t,q}(r).$$

As the points p and q are allowed to merge into a single point (different from t), the quantity $G_{p,t}(r) + G_{t,q}(r)$ extends holomorphically; hence the same is true of $G_{p,q}(r)$. □

The same sort of arguments as above allow the construction of meromorphic 1-forms with poles of orders greater than or equal to 2. Not pretending to exhaustiveness, we merely state a typical result in this direction:

Proposition II.2.10. — *Given a point $p \in T \setminus A$, there exists a unique meromorphic* 1*-form on T having a pole of order* 2 *at p, with principal part $\frac{1}{(x-x_p)^2}$, and with no other poles, and whose integral around each of the loops $\alpha_1,\dots,\alpha_g$ is zero.*

Proof. — This follows as in the proof of Proposition II.2.8, except that the role of the line D is now played by the tangent to the curve C at p. □

Remark II.2.11. — In the statements of Propositions II.2.7, II.2.8, and II.2.9, one may — as in the statement of Theorem II.2.5 — replace the condition "whose integral around each of the loops $\alpha_1,\dots,\alpha_g$ is zero" by the condition "whose integral around each of the loops $\alpha_1,\dots,\alpha_g,\beta_1,\dots,\beta_g$ is purely imaginary. To see this, it suffices to consider the map Φ defined earlier in place of the map Ψ.

II.2.4. The Riemann–Roch theorem

In Section V of his memoir, Riemann begins his investigation of the space of meromorphic functions on a given compact surface T. He proposes determining the functions by means of their poles: this is the *Riemann–Roch* problem. (According to Gray [Gra1998], this name was bestowed by Brill and Noether in [BrNo1874].)

Riemann first considers a given set $\{P_1,\dots,P_m\}$ of points, candidates for simple poles, the case of poles of greater order to be dealt with subsequently by passing to a limit where several poles merge. This procedure is used several times by Riemann, and it is not always easy to make it work formally, even if it is clear enough intuitively.

The set of meromorphic functions having at most simple poles at the points $P_1,\dots,P_m$ is obviously a complex vector space. Riemann quickly shows that it has finite dimension, and obtains an upper bound for the dimension by considering such meromorphic functions as particular cases of *many-valued* functions of a special kind (in the following quote we have taken the liberty of changing some of the notation):

> The general expression of a function s, which becomes infinitely large of the first order at m points $P_1, P_2, \ldots, P_m$ of the surface T is, by virtue of the above,
>
> $$s = \beta_1 t_1 + \beta_2 t_2 + \cdots + \beta_m t_m + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_g w_g + \text{const.},$$
>
> where t_i is any function t_{P_i} and where the quantities α and β are constants.

If we wish to avoid using many-valued functions on T, we may instead couch our argument in terms of the differentials of the functions in question: if f is one of the functions under study, then its differential df is a linear combination of differentials of the second kind associated with each point P_i (the dt_i) and of differentials of the first kind (the dw_j).

The existence of these forms and the fact that those of the first kind constitute a space of dimension *exactly* g constitute a special case of Theorem II.2.5. In taking their (many-valued) primitives, one should not forget to add 1 to the dimension on account of the constant of integration.

Next one needs to distinguish the differentials without periods, that is, those that integrate to yield meromorphic functions, the object of the investigation. By considering a basis for the first homology group, one obtains $2g$ conditions for the vanishing of the periods, which one interprets as $2g$ linear conditions on the space of forms in question. It follows that altogether their dimension is at most $2g$, whence the following:

Theorem II.2.12 (Riemann's inequality). — *Let T be a compact Riemann surface of genus g. The vector space of meromorphic functions having at most simple poles at the points $P_1, \ldots, P_m$ has dimension at least $m - g + 1$.*

By varying the set of poles imposed, we infer the following corollary:

Corollary II.2.13 (Riemann). — *A compact Riemann surface admits infinitely many meromorphic functions linearly independent over $\mathbb{C}$.*

It was Gustav Roch, a student of Riemann — deceased, alas, very young, in the same year as his supervisor — who subsequently succeeded, in [Roc1865], in interpreting the difference between the dimension sought and the quantity $m-g+1$.

Here is the full statement, embracing also the case of multiple poles:

Theorem II.2.14 (Riemann–Roch). — *Let T be a compact Riemann surface of genus g, and let $P_1, \ldots, P_m$ be points with which are associated "multiplicities"*

$n_1, \ldots, n_m$ from $\mathbb{N}^*$, and write $m = \sum n_i$, the sum of these multiplicities. Then the difference between the dimension of the vector space of functions having a pole of order at most n_i at the point P_i and $(m - g + 1)$ is equal to the dimension of the vector space of holomorphic forms having a zero of order at least n_i at the point P_i.

Application to the uniformization of curves of genus 0 and 1. — It is difficult to overestimate the importance of the Riemann–Roch theorem for the modern approach to the theory of algebraic curves. In particular, it is this theorem that is regularly invoked in order to prove that every compact, simply connected Riemann surface is isomorphic to the Riemann sphere.

Theorem II.2.15. — *A compact Riemann surface of genus zero is biholomorphic to the Riemann sphere.*

Proof. — It follows directly from the Riemann–Roch theorem that such a surface S admits a meromorphic function with just one, simple, pole, that is, there exists a holomorphic mapping $S \to \bar{\mathbb{C}}$ of degree 1. Since S has genus zero, by the Riemann–Hurwitz theorem this mapping can have no ramification points, so that it is an isomorphism. □

Despite the simplicity of this proof, this result was most probably not thought of by Riemann or Roch, whose interest, it must be recognized, was not centered on the genus zero case. In Chapter IV we shall give an analytic proof of this theorem due to Schwarz, and a little further on in the present section a proof due to Clebsch using birational geometry.

In much the same way, the Riemann–Roch theorem allows one to uniformize curves of genus 1.

Theorem II.2.16. — *A compact Riemann surface of genus 1 is biholomorphic to a quotient of $\mathbb{C}$ by a lattice of translations.*

Proof. — Applied to the case $m = 0$, $g = 1$, the Riemann–Roch theorem tells us that on a surface S of genus 1 there exists a nowhere vanishing holomorphic form ω. Consider now the vector field dual to ω, that is, the non-singular vector field X such that $\omega(X) = 1$. Integration of this field affords an action of $\mathbb{C}$ on the surface S. Since X is non-singular, every (complex) integral curve of X — the orbits of the action, in other words — are open. Then since the complement of an orbit is a union of orbits, these must also be closed. Moreover since S is connected, the action is transitive, so S can be identified with $\mathbb{C}/\Lambda$, where Λ is the stabilizer of a point, a closed subgroup of $\mathbb{C}$. Since S is compact and has the same dimension as $\mathbb{C}$, Λ must necessarily be a lattice in $\mathbb{C}$. □

Once again it seems that Riemann never wrote this result down explicitly, even if, as seems likely, he had at some time conceived it.

Box II.5: The Riemann–Roch theorem and Serre duality

A word on how the Riemann–Roch theorem is formulated nowadays. The system of multiplicities n_i attached to the points P_i of the surface T is called a *divisor* $D := \sum n_i P_i$. The sum $\sum n_i$ is then defined to be the degree $\deg(D)$ of D.

The (germs of) functions having at most a pole of order n_i at the point P_i form a *sheaf*, denoted by $\mathcal{O}(D)$. The two cohomology groups $H^0(\mathcal{O}(D))$ and $H^1(\mathcal{O}(D))$ of this sheaf are naturally endowed with the structure of *finite-dimensional* complex vector spaces, with dimensions denoted respectively by $h^0(\mathcal{O}(D))$ and $h^1(\mathcal{O}(D))$. In the literature the dimension $h^0(\mathcal{O}(D))$ is also often denoted by $l(D)$.

The first vector space $H^0(\mathcal{O}(D))$ can be interpreted as that consisting of the meromorphic functions in question with poles of order *at most* D and *defined globally* on T. The second can be interpreted globally only via *the Serre duality theorem*, affirming that there is a *canonical* isomorphism

$$H^1(\mathcal{O}(D)) \simeq (H^0(\Omega(-D)))^*,$$

where $\Omega(-D)$ is the sheaf of holomorphic forms vanishing at least to the order D. If K is the divisor of a global holomorphic (or meromorphic) differential form, then the sheaf $\Omega(-D)$ becomes identified with the sheaf $\mathcal{O}(K - D)$, whence the following version of the Riemann–Roch theorem (for curves):

$$l(D) - l(K - D) = \deg(D) - g + 1.$$

The *Euler characteristic* $\chi(\mathcal{O}(D))$ of the sheaf $\mathcal{O}(D)$ is by definition the difference $h^0(\mathcal{O}(D)) - h^1(\mathcal{O}(D))$. Thus the Riemann–Roch theorem may also be stated in the form

$$\chi(\mathcal{O}(D)) = \deg(D) - g + 1.$$

Thus, via Serre duality, one retrieves the version II.2.14 of the theorem. Viewed this way, the above modern version might seem to be merely a tautological reformulation. However, the significance of this reformulation derives from the fact that it allows the statement to be extended to higher dimensions, as was shown by Kodaira, Hirzebruch, Serre, and Grothendieck in the 1950s: the Euler characteristic $\chi(\mathcal{F})$ of a sheaf $\mathcal{F}$ of sections of an algebraic fibre bundle over a compact algebraic variety or of a holomorphic bundle over a compact analytic manifold is expressed uniquely in terms of topological invariants of the bundle in question and the tangent bundle of the manifold; and the vector spaces $H^i(\mathcal{F})$ entering into the definition of $\chi(\mathcal{F})$ are naturally isomorphic to $(H^{n-i}(\Omega(\mathcal{F}^*)))^*$, where n is the dimension of the variety or manifold.

We shall now prove the preceding two results using an idea due to Clebsch [Cle1865a, Cle1865b]. This method has the advantage of being completely algebraic in the sense that it uses no analysis (unlike the proof of Riemann–Roch via the Dirichlet principle). On the other hand, it has the shortcoming that it deals only with Riemann surfaces assumed *a priori* to be algebraic.

By Theorem II.1.3 the surface S can be so immersed in $\mathbb{CP}^2$ that its image is an algebraic curve C having as its only singularities double points with distinct tangents.

Let n be the degree of C. Recall from the formula (II.3) that the genus of S is equal to $\frac{(n-1)(n-2)}{2} - k$, where k is the number of double points.

Curves of genus zero. — Suppose S has genus zero. Then the curve C has $N = (n-1)(n-2)/2$ double points $x_1, x_2, \ldots, x_N$, say. Choose any particular $n-3$ other points $y_1, \ldots, y_{n-3}$ on C. Recall that the projective space of curves of a given degree d has dimension $d(d+3)/2$, so that the projective space E of curves of degree $n-2$ passing through the N points x_i and the $n-3$ points y_i has dimension at least

$$\frac{(n-2)(n+1)}{2} - N - (n-3) = 1.$$

Let z_1 and z_2 be any two distinct points of C. Through each of these there passes at least one curve from E. By replacing E by the line in E determined by a curve through z_1 and a curve through z_2 (considered as points of the projective space E), we may suppose that E has dimension precisely 1.

By Bézout's theorem, each curve of degree $n-2$ meets C in $n(n-2)$ points, counted according to their multiplicities. Thus apart from the x_i and the y_i, the curves in E meet C in

$$n(n-2) - 2N - (n-3) = 1$$

points. This affords us a rational map from E to C. It is not constant since z_1 and z_2 are distinct points in the image. It is even birational since the preimage of a point is a proper projective subspace of the one-dimensional space E, so consists of a single point.

Curves of genus 1. — Now suppose the genus of S is 1. In this case the curve C has $N = (n-1)(n-2)/2 - 1$ double points $x_1, x_2, \ldots, x_N$. Choose any particular $n-3$ other points $y_1, \ldots, y_{n-3}$ on C. The space E consisting of curves of degree $n-2$ passing through the N points x_i and through the y_i has projective dimension at least:

$$\frac{(n-2)(n+1)}{2} - N - (n-3) = 2.$$

Analogously to the case of genus 0, if E happens to have dimension greater than 2, we replace it by a generic subspace of dimension precisely 2. By Bézout's theorem, apart from the x_i and y_i, the curves from E each meet C in

$$n(n-2) - 2N - (n-3) = 3$$

points. It follows that corresponding to each generic point x of C, there exists exactly one curve from E tangent to C at x (and passing through the x_i and y_i). This defines a *rational* map from the curve C to the projective plane E. We shall now show that the image of this map is a cubic curve.

To this end we consider the pencil of curves in the projective space E determined by two of its elements, and find the condition that a curve in this pencil be tangent to C. The condition takes the form of an equation of degree 3 since it involves the vanishing of the discriminant of a polynomial of degree 3. The cubic thus obtained must be non-singular since we know that a singular cubic has genus zero. We have thus established a birational equivalence between the curve C and a smooth cubic, which may now in turn be projectively transformed into Weierstrass normal form.

II.3. The Jacobi variety and moduli spaces

After having investigated his surfaces individually, Riemann seeks to comprehend them collectively. This represents the birth of the "space" of moduli. Difficulties in defining this space notwithstanding, this opens the way to a topological approach to the uniformization theorem: the method of continuity, forming the theme of the second part of the present book.

II.3.1. Moduli spaces of Riemann surfaces

Birational equivalence. — At the beginning of his investigation, Riemann considers the surface T to be associated with an algebraic function $s(z)$ as a branched covering of the sphere $\overline{\mathbb{C}}_z$ associated with the plane of the complex variable z.[21]

[21]More precisely, above that plane. The use of the Riemann sphere is made explicit subsequently in work of Neumann [Neum1865].

However he next envisages changing the variable z employed to represent T:

> A function z_1 of z, ramified like T, which becomes infinite to the first order at n_1 points of that surface [...], takes each of its values at n_1 points of the surface T. Consequently, when one imagines each point of T represented by a point of a plane representing geometrically the value of z_1 at that point, the totality of these points forms a surface T_1 everywhere covering the z_1-plane n_1 times, a surface which is, one understands, a representation, similar to it in its smallest parts, of the surface T. To each point of either of these surfaces there then corresponds a *unique* point of the other.

Mathematicians later learned to say that T and T_1 are *isomorphic* as Riemann surfaces, and, in particular, homeomorphic. However, in order to begin using such language, it would be necessary to come to the recognition that various sorts of mathematical objects have internal structures defining their *form*, and it was to this realization that in fact the work of Riemann contributed in no small measure.

After having represented T in a new way with the aid of a meromorphic function z_1, one can go on to consider the representation one obtains by means of a further meromorphic function:

> If one denotes by s_1 any other function whatever of z, ramified like T [...], then (§ V) s_1 and z_1 will be linked by an equation of the form $F_1(s_1, z_1) = 0$, where F_1 is a power of an irreducible entire function of s_1, z_1, and when this power is the first, one can express every function of z_1 ramified like T rationally in terms of s_1 and z_1, and, consequently, all rational functions of s and z (§VIII). The equation $F(s,z) = 0$ can thus, by means of a rational transformation, be transformed into $F_1(s_1, z_1) = 0$ and *vice versa*.

The equivalence relation that he introduces stemming from such considerations represents the point of departure of *birational geometry* (see Klein [Kle1928, Chapter VII]):

> We now consider as forming part of the same class, all irreducible algebraic equations in two variable quantities that can be transformed one to the other by means of rational substitutions [...].

The choice of an equation $F(s,z) = 0$ in such a class, and of one of the two variables s say, as representing, via this equation, an algebraic function of the other variable z, allows us to define "a system of identically ramified algebraic functions", or, in modern terminology, *a finite extension of the field* $\mathbb{C}(z)$, that is, the field of rational functions over the curve defined by the equation $F(s,z) = 0$

(which may also be thought of as the field of meromorphic functions on the associated Riemann surface). One thus arrives at the present definition: two algebraic curves are *birationally equivalent* if their fields of rational functions are isomorphic as field extensions of $\mathbb{C}$. And in fact two non-singular curves are birationally equivalent if and only if they are biholomorphic.

Counting moduli. — At this point Riemann introduces the *moduli* problem for Riemann surfaces of genus g — the problem of studying the birational equivalence classes for each fixed topological type of Riemann surface, that is, for each fixed value of the genus.

Riemann explains that, for $g \geq 2$:[22]

> [...] a class of systems of functions identically ramified and $(2p + 1)$-connected and the class of algebraic equations belonging to it, depend on $3p - 3$ quantities varying in a continuous manner, which will be called the *moduli of the class.*

Nowadays we speak of the *moduli space*, but here we see that Riemann refers only to the number of parameters needed to determine the points of the space, that is, its complex dimension, without any mention of the possibility of a global construction of such a "space". Nonetheless he has thought of this possibility, as is shown by the following excerpt from his habilitation address [Rie1854, pp. 282–283], delivered three years earlier:

> Concepts of size are possible only where there exists a general concept allowing different modes of determination. According to whether it is or is not possible to pass from one of these modes of determination to another in a continuous manner, they form a continuous or discrete manifold [...] the occasions giving rise to concepts whose modes of determination form a continuous manifold are so rare in everyday life that the positions of sensible objects and their colours are practically the only simple concepts whose modes of determination form a manifold of several dimensions. It is only in higher mathematics that occasions for the formation and development of such concepts become more common.
>
> Such investigations have become necessary in many areas of mathematics, notably in the study of analytic many-valued functions, and it is primarily on account of their imperfection that Abel's celebrated theorem, as well as works of Lagrange, Pfaff, and Jacobi on the general theory of differential equations, have remained sterile for so long.

[22] In the following quotations the genus is denoted by p.

Riemann proposes two methods for calculating the number of moduli, the first valid only for $g > 1$, and the second for $g \geq 1$. (We saw in the preceding chapter that a Riemann surface of genus $g = 0$ is isomorphic to $\overline{\mathbb{C}}_z \simeq \mathbb{CP}^1$.)

First method. — Here for each $\mu > 2g$ Riemann considers the set of meromorphic functions on T with exactly μ poles (counted according to their multiplicities). In other words he considers the space of holomorphic maps of degree μ from T to $\mathbb{CP}^1$. It follows from the Riemann–Roch theorem (Theorem II.2.14 and Box II.5) that this space has dimension $2\mu - g + 1$.

By the Riemann–Hurwitz theorem (see Box II.4) a function from T to $\mathbb{CP}^1$ with μ poles has $2(\mu + g - 1)$ ramification values, that is, the set of images of its critical points is a finite subset of points of the Riemann sphere of this cardinality. By allowing the function to vary (by varying the "arbitrary constants" on which it depends), this finite set can be varied. And then:

> These constants can be given values in such a way that the $2\mu - p + 1$ ramification points take on any prescribed values provided the functions determined by these constants are independent, which can be achieved in only finitely many ways since the equations expressing this condition are algebraic.

Riemann now asserts that the condition that the functions be independent is satisfied provided $g > 1$. In this case, by choosing the meromorphic function on T so that the $2\mu - g + 1$ "ramification points take on any prescribed values", there remain $3g - 3$ unused ramification values, which therefore afford a complete system of parameters for the moduli of T.

Second method. — Rather than considering, as in the above approach, properties of meromorphic functions on T, the second method exploits properties of integrals w of holomorphic forms ("integrals of the first kind") — or, more precisely, of their periodicity moduli relative to a fixed system of sections transforming T into a simply connected surface T' and their values at the zeros of the associated holomorphic form, that is, the critical values of $w|_{T'}$.

The calculation yielding the desired $3g - 3$ moduli of the surface T is then as follows:

> [...] we can, in the quantity $w = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + c$, treated as an independent variable, determine both the quantities α, where of $2p$ periodicity moduli p can be given prescribed values, and the constant c, provided $p > 1$, and in such a way that one of the $2p - 2$ ramification values of the periodic functions of w take on a prescribed value. In this way w is completely determined, and consequently the remaining $3p - 3$

> quantities on which depend the mode of ramification and the periodicity of these functions of w likewise [. . .].

The question that suggests itself next is whether the set of isomorphism classes of Riemann surfaces of a fixed genus g can be naturally endowed with supplementary structures. Is there, for example, a topology on that set with respect to which the parameters considered by Riemann in the above two approaches become continuous functions? It is only when one has imposed on the set of isomorphism classes some sort of structural concepts of a geometrical nature that one can speak of the *space* of moduli. The problem of moduli, as it arose following Riemann, is that of defining such structures reflecting the properties of the objects under examination.

For example, if we consider compact Riemann surfaces as complex algebraic curves, we may ask if the space of moduli can itself be regarded as a complex algebraic variety. Contemporary research has shown that this is indeed the case (see the book [HaMo1998]).

Proposition II.3.1. — *There exists an irreducible quasi-projective complex variety $\mathcal{M}_g$ (hence connected) that is a moduli space for the compact complex smooth algebraic curves of genus g.*

We now elucidate the meaning of this statement. It is easy to define the concept of an *algebraic family* of curves of genus g: such an object is given by an algebraic morphism $X \xrightarrow{\pi} B$ with fibres $\pi^{-1}(b)$ curves of genus g. Thus we have a family of curves "parametrized" by the base B. Our space $\mathcal{M}_g$ is characterized by the property that for each family of this sort, there exists a *unique* algebraic map $B \xrightarrow{\gamma} \mathcal{M}_g$ such that for each $b \in B$ the curve $\pi^{-1}(b)$ belongs to the isomorphism class represented by the point $\gamma(b) \in \mathcal{M}_g$. In particular, therefore, there is a canonical bijection between the points of $\mathcal{M}_g$ and the isomorphism classes of curves of genus g, by means of which the algebraic structure of $\mathcal{M}_g$ induces a geometric structure on the set of moduli.

An important point here is that $\mathcal{M}_g$ itself is not the base of any algebraic morphism $X \xrightarrow{\pi} \mathcal{M}_g$ for which for every $b \in \mathcal{M}_b$, the fibre $\pi^{-1}(b)$ is in the isomorphism class represented by b. For this reason one says that $\mathcal{M}_g$ is only a *coarse* moduli space.

In the above two methods of Riemann one is in effect considering Riemann surfaces endowed with certain supplementary structures: a meromorphic function defined on the surface together with an enumeration of its critical values, or again a basis for its homology. The question of the existence of moduli spaces for such "enriched" Riemann surfaces turns out to be an important one. The advantage of such an approach is that by enriching the additional structure sufficiently one obtains objects *without nontrivial automorphisms*, which facilitates study of the moduli problem. For example, this allows one to show that $\mathcal{M}_g$ is in fact the quotient of a *smooth* algebraic variety by a finite group action.

II.3.2. The "Abelian" uniformization of Jacobi and Riemann

A further important contribution of [Rie1857] was the solution of the "inversion problem" left open by Abel and Jacobi. In order to better explain Riemann's contribution, we go back to Abel, who — around 1829 — managed to generalize Euler's addition theorem to Abelian integrals. He starts with an integral

$$\int_{x_0}^{x} y\,dx$$

where $y(x)$ is an algebraic function defined by an irreducible polynomial equation $F(x, y) = 0$, and he shows that there exists an integer $\mu \geq 0$ such that for any given $\mu + 1$ complex numbers $x_1, x_2, \ldots, x_{\mu+1}$ one can find μ complex numbers $x'_1, \ldots, x'_\mu$ — uniquely determined up to order — depending rationally on $x_1, x_2, \ldots, x_{\mu+1}$ such that

$$\int_{x_0}^{x_1} y\,dx + \int_{x_0}^{x_2} y\,dx + \cdots + \int_{x_0}^{x_{\mu+1}} y\,dx = \int_{x_0}^{x'_1} y\,dx + \cdots + \int_{x_0}^{x'_\mu} y\,dx,$$

to within a period of $\int y\,dx$. One should think of this as merely a formal equality between sums of anti-derivatives. For example, for the integral of the form of the second kind dx/x, that is, the complex logarithm, one has

$$\int_1^a \frac{dx}{x} + \int_1^b \frac{dx}{x} = \int_1^{ab} \frac{dx}{x}.$$

By applying Abel's theorem several times one sees that it leads to an "addition" of μ-tuples of points. More precisely:

For any given μ-tuples $(x_1, \ldots, x_\mu)$ and $(x'_1, \ldots, x'_\mu)$, defined up to order, there is a μ-tuple $(x''_1, \ldots, x''_\mu)$, uniquely determined up to order, depending rationally on $(x_1, \ldots, x_\mu)$ and $(x'_1, \ldots, x'_\mu)$, such that

$$\sum_1^\mu \int_{x_0}^{x_i} y\,dx + \sum_1^\mu \int_{x_0}^{x'_i} y\,dx = \sum_1^\mu \int_{x_0}^{x''_i} y\,dx.$$

Thus whereas Euler and Gauss found an addition rule for the points of a lemniscate (the case $\mu = 1$), Abel found such a rule for sets of size μ.

The situation remained for some time in this rather mysterious — and moreover not quite valid — form. In particular, the significance of the integer μ remained hidden from view. It had to wait for the work of Jacobi and especially Riemann before it was understood that when the Abelian integral is of the first kind, μ is equal to the genus g of the Riemann surface associated with y, and when the integral is of the second kind — as in the case of the logarithm — μ is $g + 1$. We should not forget that at the time of Abel and Jacobi no one thought of an algebraic curve as a surface endowed with a topology.

For an exposition of Abel's theorem from his point of view, the reader may consult [Cat2004, Kleim2004], where it will be seen that Abel considered several rather different versions of his theorem.

Hyperelliptic functions and Jacobi's inversion problem. — One of the first families of Abelian integrals beyond elliptic integrals consists of those of the following form:

$$u = \int_0^x \frac{(\alpha + \beta x)\, dx}{\sqrt{P}},$$

where P is a polynomial of degree 6. This corresponds to the curve C with equation $y^2 = P(x)$, a Riemann surface S of genus 2 to which $(\alpha + \beta x)\, dx/y$ lifts as a holomorphic differential. The integral therefore has precise meaning provided one specifies the homotopy class of the path of integration joining the two limits of integration. We note once again that the concepts expressed in this sentence were not available to Jacobi.

Thus the "function" u lifts to a many-valued function on S. Recall that in the case of a polynomial P of degree 3 or 4 the analogue of the map u (where P is assumed to have degree 6) has a doubly periodic inverse. In the present case the study of the inverse of u encounters two major difficulties.

The first difficulty arises from the vanishing of the form $(\alpha + \beta x)\, dx/y$ at two points on the surface ($x = -\alpha/\beta$ corresponds to the two points of the surface arising from the two values taken by $\sqrt{P}$ at that x). Hence u has critical points, whence its "inverse" — assuming it existed — would have branch points and so not be single-valued! This difficulty did not arise in the elliptic case because the form $dx/\sqrt{P}$ (with P of degree 3 or 4) does not vanish on the corresponding elliptic curve.

The second difficulty arises from the fact that $(\alpha + \beta x)\, dx/y$ affords *four* periods, given by integrals around four loops encircling pairs of roots of P. If it existed, the inverse function would thus have four independent periods. Jacobi established the fact, clear to a modern mathematician, that a subgroup of rank 4 of $\mathbb{C}$ cannot be discrete, and so cannot serve as the group of periods of a non-constant meromorphic function.

Following on the appearance of Abel's article on the laws of addition of μ-tuples of points, Jacobi had two brilliant new ideas for finding a way out of the impasse.

The first of these consisted in using two holomorphic forms simultaneously on C. Given any loop γ on C, one can integrate $\frac{dx}{\sqrt{P}}$ and $\frac{x\,dx}{\sqrt{P}}$, obtaining a pair of periods $(\omega_1(\gamma), \omega_2(\gamma)) \in \mathbb{C}^2$. As γ ranges over all loops on C, these pairs of periods range over a subgroup Λ of rank 4 of $\mathbb{C}^2$ and no longer of $\mathbb{C}$ as previously. It is therefore possible that Λ is a discrete subgroup of $\mathbb{C}^2$ and indeed this turns out to be the case. Thus one now has available a holomorphic map — called the "Abel–Jacobi map" — utilizing two forms, namely:

$$x \in C \mapsto \left(\int_0^x \frac{dx}{\sqrt{P}}, \int_0^x \frac{x\,dx}{\sqrt{P}} \right) \in \mathbb{C}^2/\Lambda.$$

The complex torus $\mathbb{C}^2/\Lambda$ is today called the *Jacobian* of the curve C. However, uniformization has not been achieved here since the torus $\mathbb{C}^2/\Lambda$ has dimension 2, so could not possibly parametrize the curve C, of dimension 1.

Jacobi's second idea was to use pairs of points, that is, to use the map

$$(x_1, x_2) \in C^2 \mapsto \left(\int_0^{x_1} \frac{dx}{\sqrt{P}} + \int_0^{x_2} \frac{dx}{\sqrt{P}}, \int_0^{x_1} \frac{x\,dx}{\sqrt{P}} + \int_0^{x_2} \frac{x\,dx}{\sqrt{P}} \right) \in \mathbb{C}^2/\Lambda.$$

The domain and codomain of this map have the same dimension, but the map is not bijective since it sends (x_1, x_2) and (x_2, x_1) to the same image. One gets around this by working instead with the "symmetric square" $C^{(2)}$ of C, the quotient of C^2 by the involution switching the two factors; the elements of $C^{(2)}$ are therefore essentially just the unordered pairs of not necessarily distinct points of C. By means of elementary symmetric functions one then endows $C^{(2)}$ with the structure of a smooth algebraic variety of dimension 2. Thus one now has at one's disposal a holomorphic map from $C^{(2)}$ to $\mathbb{C}^2/\Lambda$, and it is this map that Jacobi seeks to invert. The question of its surjectivity is the "Jacobi inversion problem", which he himself failed to solve. This particular problem was solved around that time by Adolph Göpel and Georg Rosenhain in the special case of hyperelliptic curves that we have been expounding here. But it is to Riemann that we owe the complete solution of the problem.

Riemann and the Jacobi inversion problem. — Riemann begins by generalizing the construction to any surface S whatever, not necessarily hyperelliptic. Recall that by Theorem II.2.5 the space of holomorphic forms on S has dimension equal to the genus g of S. By integrating g such forms comprising a basis for that space over all loops on S we obtain a subgroup Λ of $\mathbb{C}^g$. Riemann proves that this subgroup is a lattice, that is, that it is discrete with compact quotient $\mathbb{C}^g/\Lambda$. Much as above one constructs an Abel–Jacobi map from C to $\mathbb{C}^g/\Lambda$. By taking the sum of the images in $\mathbb{C}^g/\Lambda$, one then obtains a map from the symmetric power $C^{(g)}$ to $\mathbb{C}^g/\Lambda$. Riemann now establishes the following two fundamental theorems:

Theorem II.3.2. — *The Jacobian $\mathbb{C}^g/\Lambda$ of an algebraic curve is an algebraic variety, that is, it embeds holomorphically in a projective space of sufficiently high dimension as an algebraic subvariety.*

Theorem II.3.3. — *The Abel–Jacobi map $C^{(g)} \to \mathbb{C}^g/\Lambda$ is birational.*

This does not mean that this map is an isomorphism — in fact a little topological reasoning shows that for $g \geq 2$ these two spaces are not even homeomorphic. However, by way of compensation we get the existence of a rational map $\mathbb{C}^g/\Lambda \to C^{(g)}$ that is inverse to the Abel–Jacobi map where defined.

To give the proofs of Theorems II.3.2 and II.3.3 would take us too far afield. We limit ourselves to sketching briefly a proof of the surjectivity of the map in Theorem II.3.3.

Proof. — We shall show that the Abel–Jacobi map has non-zero "topological degree". Recall (see [Mil1965]) that the *topological degree* of a C^∞ map between two compact orientable manifolds is the sum of the signs of the Jacobian determinants over the preimages of a regular value. Hence a map of non-zero topological degree must be surjective.

The Abel–Jacobi map is holomorphic and is therefore orientation-preserving. Thus it suffices to prove that its image contains a regular value. We shall show that there exists a g-tuple $l \in C^{(g)}$ where the derivative is invertible; this will suffice since then by the (local) Inverse Function Theorem the image of our map must contain a non-trivial open set and hence at least one regular point.

Observe that the derivative of the Abel–Jacobi map fails to be invertible at a g-tuple $l = (x_1, \ldots, x_g) \in C^{(g)}$ if and only if there exists a form ω of the first kind on C that vanishes at all of the x_i. It therefore suffices to find a g-tuple l at which no form vanishes. To this end we consider the projective space $\mathbb{P}(\Omega^1(C))$, of dimension $g-1$. The subset A of $\mathbb{P}(\Omega^1(C)) \times C^{(g)}$ consisting of all pairs $(\bar{\omega}, l)$ where $\bar{\omega}$ is the complex line defined by a differential form ω vanishing on the g-tuple l, is an analytic subset of dimension $g-1$, so that its projection on the factor $C^{(g)}$ cannot be surjective. □

With a little more work it can be shown that the topological degree of this map is exactly 1. This implies that there exist dense open sets in $C^{(g)}$ and in $\mathbb{C}^g/\Lambda$ that are biholomorphic to one another, so that the Abel–Jacobi map is birational.

One can find a proof of the first of Riemann's theorems above (giving a necessary and sufficient condition — in terms of the "Riemann bilinearity conditions" — for a torus $\mathbb{C}^g/\Lambda$, which is automatically a holomorphic manifold, to be in fact an algebraic variety) in [Bos1992]. This proof exploits a higher-dimensional generalization of the Jacobi ϑ-functions, called since then the "Riemann theta-function".

While the Jacobian naturally carries the structure of an Abelian group, Theorem II.3.2 endows it besides with the structure of an algebraic variety. This can be phrased more concisely by saying that "the Jacobian is an Abelian variety", which serves to join in a single statement the names of the two protagonists whose rivalry should not be allowed to obscure the similarity of their mathematical visions[23].

These theorems afford a new global perspective on the theories of the integrals of algebraic functions and of algebraic curves:

— the birational identification of the first theorem renders transparent Abel's theorem on the law of addition of g-tuples of points: a mysterious addition formula becomes a simple consequence of the group operation on the Jacobian;

— without yielding a parametrization of the points of the curve, Riemann's theorems provide a simple algebraic model for g-tuples of points by referring them to a given algebraic model.

However, in the course of time, it transpired that one could do much better than this "Abelian uniformization". Twenty-five years later Klein and Poincaré showed that the points of C itself (and not just of $C^{(g)}$) can be uniformized.

[23]In a letter addressed to Legendre which has remained famous, dated July 2, 1830, Jacobi writes: "But M. Poisson should not have reproduced in his report the not very clever statement of the late M. Fourier, by which the latter reproaches us, Abel and me, for not having occupied ourselves instead with the motion of heat. It is true that M. Fourier was of the opinion that the principal aim of mathematics is public utility and the explanation of natural phenomena; but a philosopher like him should have known that the only aim of science is to honor the human spirit, and that under this banner a question about numbers is worth just as much as a question about the system of the world."

Chapter III

Riemann surfaces and Riemannian surfaces

In 1881 Felix Klein gave a course [Kle1882] on Riemann's work, in which he tried to make the theory of Riemann surfaces more intuitive. By then, of course, a considerable length of time had elapsed since the appearance of Riemann's memoirs of 1851 and 1857 [Rie1851, Rie1857]; this reworking of Riemann's results was contemporaneous with the first announcements of the uniformization theorem, which we shall be considering in Part B and of which Klein was one of the major heroes.

Especially notable was his reinterpretation of Riemann's Theorem II.2.5 on the existence of meromorphic forms on a Riemann surface, in terms of fluid flow on the surface. The better to grasp his idea, we reconsider the Riemann Mapping Theorem from this point of view. Thus consider a bounded, simply connected (open) region of the complex plane, and imagine that its boundary is a perfectly conducting wire. If we attach one terminal of a battery to a point inside the region and the other to a point on the boundary, we obtain a flow of charge in the region, following the flow lines of the gradient of the potential. One sees that this potential has a logarithmic singularity at the point in the interior to which the battery is attached, and is constant on the boundary in view of the assumption that the boundary wire is perfectly conducting. Thus we have a Green's function on the open set, and have thereby "proved" the Riemann Mapping Theorem by experimental means.

In his course, Klein aimed at illustrating Riemann's theory by extending this physical intuition to an arbitrary compact surface.

In order to describe Klein's physical illustration in mathematical terms, one needs to introduce a Riemannian metric on the surface under consideration. Even though this is far removed from Klein's actual preoccupations and techniques, it will show how this new structure allows one to look at Riemann's theory from a more modern point of view. In particular, it sheds new light on Theorem II.2.5 on the existence of meromorphic forms on a given surface and also on the moduli problem.

III.1. Felix Klein and his illustration of Riemann's theory

As was explained in the preceding chapter, with each algebraic function of a complex variable z Riemann associates a surface covering the z-plane several times. For most of his exposition of the theory, Riemann uses the parameter z of the plane to describe objects that today we would consider as living on the surface. Getting to the point of rendering unto the surface that which belongs to it has been a long and difficult process. Here, for instance, is how Klein talks of the matter in the preface to his course [Kle1882c], taught in 1881:

> I am not sure if I'd have been able to develop a coherent conception of the current subject as a whole if, many years ago now (1874), during an opportune conversation, M. Prym had not said something to me that has assumed more and more importance in the course of my subsequent reflections. He said that "Riemann's surfaces are, fundamentally, not necessarily surfaces of several sheets above the plane, but on the contrary, complex-valued functions of position that can be studied on arbitrarily given curved surfaces in exactly the same way as on surfaces above the plane".

In [Kle1882c] Klein proposes expounding the theory of meromorphic forms and functions living on a compact Riemann surface in an intrinsic language no longer employing projections on the plane. And, even more important, he wants to teach his students to *think* in physical terms, since:

> [...] there are certain physical considerations that have been developed subsequently [...]. I have not hesitated in taking these physical conceptions as my point of departure. As we know, Riemann used Dirichlet's principle instead. However, I have no doubt that he started from precisely these physical problems and then, in order to lend the support of mathematical reasoning to what was obvious from a physical point of view, he replaced them with Dirichlet's principle.

The path from the "physically obvious" to mathematical rigour is thus strewn with pitfalls. And in progressing towards rigour, one risks losing all intuition. According to Klein that is what happened in this particular case, and what motivated him to to design his course:

> We are familiar with all the tortuous and difficult considerations by means of which, over the last several years, some, at least, of the theorems of Riemann that we deal with here have been given reliable proofs. Such considerations play a completely negligible part in what follows and I thus renounce the use of anything except intuitive foundations for the theorems stated. These proofs should not in any way be subsequently mixed up with the ideas that I have tried to preserve [. . .]. However they should obviously follow them [. . .].

We are unable to resist quoting the following excerpt from the review by Young [You1924] of the third volume of Klein's collected works:

> A topic that will interest the reader of Volume III is Klein's attitude to Riemann. Although Klein never saw Riemann, they can be freely compared to Plato and Socrates. Many philologists maintain that the Platonic Socrates is unhistoric. I would put this otherwise. What Plato tells us of Socrates is what he thought he saw in his master, and in order to see [what he did see] a "formidable mind" such as Plato's was necessary. What Klein tells us about Riemann is what he thought he saw of the master in his writings, and, I dare say, this intuition gave Klein access to points of view of Riemann that none of the latter's disciples had suspected. One has only to look at Riemann's portrait to see that he was modest. I am prepared to believe that he had many latent ideas of which he himself was not conscious.
>
> One should read what Klein relates on p. 479 on the subject of his paper "Algebraische Funktionen und ihre Integrale" (1882), where he claims to have revealed the actual basis of the ideas underlying Riemann's conception of his theory of functions, an essentially concrete and physical basis for abstract and metaphysical notions. Just as the real values of an algebraic function were then represented by points on a curve, so Riemann introduced flat surfaces consisting of several superimposed sheets meeting only at their branch points, in order to separate the complex values of an algebraic function $f(x+iy)$. Klein claimed that it was only by reflecting on physical phenomena that Riemann arrived at this idea, and that Riemann's original surface was not so very abstract and complicated but a completely natural curved surface realizable in space, such as the torus.
>
> On such a surface the phenomena of stable flow of heat or electricity is represented mathematically by a function, the potential, satisfying the fundamental differential equation $\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2 = 0$ of the theory of complex functions $f(x+iy)$. In his paper Klein develops this idea in a very satisfying way, and shows that from this point of view most of the theorems of function theory become intuitive. According to Klein, Riemann must have introduced the surfaces bearing his name only later on, in order to elucidate his arithmetized exposition. In this connection Klein cites the statement of Prym, a student of Riemann, that "the surfaces of Riemann

> were originally not necessarily surfaces of several sheets superimposed on the plane. One can study complex functions of position on any curved surface just as well as on planar ones".
>
> However, Klein realized that he had interpreted Prym's thought incorrectly. The latter issued a formal denial (April 8, 1882) of having said that Riemann had conceived the idea of separating the values of a complex function on a curved surface as Klein does in his paper.
>
> The above remarks are a response to the reproach made against Klein of lacking mathematical rigour in the notions forming the basis of his paper, as also, incidentally, in other places in his writings. Klein defends himself with the principle of intuitive methods that he makes use of.
>
> "I seek", he says, "to arrive by means of reflections on physical nature at a true understanding of the fundamental ideas of the Riemannian theory. I would wish for like procedures be used more often, since the usual style of mathematical publication involves a habitual relegation to the background of the important question as to the means by which one was led to formulate certain problems or to make certain deductions. I am of the opinion that the fact that most mathematicians pass over in total silence their intuitive reflections, publishing only proofs (certainly necessary) in rigorous form and for the most part mathematized, is a fault. They seem to be held back by a certain fear of not appearing scientific enough to their colleagues. Or is the reason, in some cases, that of not wishing to reveal the source of their own ideas to the competition?" He also says: "It is as physician that I wrote my note on Riemann, and furthermore in this I met with the approval of several other physicians."

In the following sections we first of all explain the intuitions developed by Klein about meromorphic forms and functions on a Riemann surface. Then we give a modern proof of Theorem II.2.5 on the existence of meromorphic forms on a given surface: being much less physically intuitive, this nicely illustrates the above statements of Klein.

Klein's physical explanations are based on the idea of considering on the given Riemann surface a Riemannian metric compatible with its complex structure. Such a metric allows us to regard, via duality, real forms as vector fields. When a real form is the real part of a meromorphic 1-form, the associated dual field inherits particular dynamical properties which can be formulated in the terminology of Riemannian geometry and interpreted in hydrodynamic or electrostatic terms. (The very name "electric current" bears witness to the analogies between different branches of physics observed in the 19th century.)

Further details of these physical interpretations and their history may be found in [Coh1967].

III.1.1. Compatible metrics on a Riemann surface

Klein uses the following fact. Given a Riemann surface S, there is always a Riemannian metric $g = \langle \cdot, \cdot \rangle$ on S that is *compatible with the complex structure*, meaning that it determines the same angle measure. Such a metric has the following form, expressed in terms of a holomorphic local coordinate system $z = x + iy$: $g = e^{u(x,y)} \sqrt{dx^2 + dy^2}$, where u is a smooth function. It is very easy to construct such a metric by modern means. It is enough to cover S by open sets U_i endowed with holomorphic maps $z_i : U_i \to \mathbb{C}$, and consider a partition of unity (ρ_i) subordinate to the open cover by the (U_i).[1] One may then use the metric

$$g = \sum_i \rho_i \cdot z_i^*(\sqrt{dx^2 + dy^2}).$$

Remark III.1.1. — If we assume the Riemann surface S embedded in some projective space $\mathbb{CP}^N$, we can construct globally a real analytic Riemannian metric compatible with the complex structure on S; it suffices to restrict to S the Fubini–Study metric on the projective Fubini–Study space (see for example [GrHa1978] for the definition of this metric).

The complex structure on S also induces an *orientation* of S, determined via the holomorphic charts by the standard orientation of $\mathbb{C}$. Indeed the coordinate changes on the overlaps of charts are biholomorphisms between open subsets of $\mathbb{C}$, so preserve the standard orientation.

It can be shown that, conversely, a given *oriented* surface (S, g) endowed with a smooth Riemannian metric admits a *unique* compatible Riemann-surface structure (see Section I.2.2). This *local uniformization theorem* is much more difficult to prove than Gauss's theorem I.2.1 (which is the particular case of this local uniformization theorem for real analytic metrics).

In summary: defining the structure of a Riemann surface on a given differentiable surface S is the same thing as choosing an orientation and a conformal class of Riemannian metrics.

The unique Riemann-surface structure on S allows us to define an associated *almost complex* operator[2] $J : TS \to TS$, which, from a geometric point of view, is just rotation through the angle $\pi/2$ in the positive sense. In fact the existence of such an operator (satisfying the equation $J^2 = -I$) is equivalent to the specification of an orientation and a conformal class of metrics, and therefore of a

[1]Thus each ρ_i is a continuous map $S \to [0,1]$ with support contained in U_i, and for each $s \in S$ all but finitely many of the ρ_i vanish in some neighborhood of s and $\sum_i \rho_i(s) = 1$. *Trans*

[2]Here TS denotes the tangent bundle over S. *Trans*

Riemann-surface structure on S. By means of J one can rotate both tangent vectors and real differential forms[3]:

$$\begin{aligned} *\vec{v} &:= J(\vec{v}), & \text{for } \vec{v} \in TS, \\ *\alpha &:= -\alpha \circ J, & \text{for } \alpha \in T^*S. \end{aligned} \tag{III.1}$$

Once we have fixed on a metric g compatible with the complex structure, we can associate with each real-valued differential 1-form α on S the vector field $\vec{v}_\alpha$ dual to it relative to g:

$$\alpha(\cdot) = \langle \vec{v}_\alpha, \cdot \rangle.$$

This then allows us to define pointwise a scalar product of two 1-forms as that of the respective dual vector fields. Denoting by vol the area form determined by g and the fixed orientation of S, we have the following formulae:

$$\begin{cases} \mathrm{vol}(*\vec{v}_1, \vec{v}_2) = -\langle \vec{v}_1, \vec{v}_2 \rangle, & \forall\, \vec{v}_1, \vec{v}_2 \in TS \\ \langle \alpha_1, \alpha_2 \rangle \mathrm{vol} = \alpha_1 \wedge *\alpha_2, & \forall \alpha_1, \alpha_2 \in T^*S, \end{cases} \tag{III.2}$$

easily proved by calculating in terms of an orthonormal basis.

With the aid of the duality between forms and vectors one can also define the concepts of the *curl* (or *rotation*) and *divergence* of a vector field (see Box III.1).

Box III.1: The curl and divergence

Let (S, g) be an oriented surface endowed with a smooth Riemannian metric. Denote by vol the associated area form. Let $\vec{v}$ be a smooth vector field on S and $\alpha = \langle \vec{v}, \cdot \rangle$ the form dual to $\vec{v}$. The 2-form $d\alpha$ is then the product of the area form by a smooth function called the *curl* of $\vec{v}$:

$$d\alpha = \mathrm{curl}\vec{v} \cdot \mathrm{vol}.$$

By Stokes' theorem, for every region U of S with smooth boundary ∂U, one has

$$\int_{\partial U} \langle \vec{v}, \vec{t} \rangle \, dl = \int_U \mathrm{curl}\vec{v} \cdot \mathrm{vol},$$

where $\vec{t}$ is the unit tangent vector field to ∂U and dl the element of length on ∂U. The left-hand side of this equation is called the *circulation* of the field $\vec{v}$ around the curve ∂U. The field $\vec{v}$ is said to be *irrotational* if its curl is identically zero, or, equivalently, if the 1-form α is closed.

[3]Note the (usual) sign convention here.

Now consider the 1-form $*\alpha$. The 2-form $d(*\alpha)$ is the product of the area form vol by a smooth function called the *divergence* of $\vec{v}$:

$$d(*\alpha) = \mathrm{div}\vec{v} \cdot \mathrm{vol}.$$

By means of Stokes' theorem this equality translates into the following standard form: for every region U of S with smooth boundary,

$$\int_{\partial U} \langle \vec{v}, \vec{n} \rangle \, dl = \int_U \mathrm{div}\vec{v} \cdot \mathrm{vol},$$

where $\vec{n}$ is the vector field normal to the boundary ∂U. The left-hand side quantity here is called the *flux* of the field $\vec{v}$ across the curve ∂U. If $\vec{v}$ models a fluid flow, this measures the infinitesimal change in the amount of fluid contained in U. The flow is called *incompressible* if the divergence of v is everywhere zero, or, equivalently, if the form $*\alpha$ is closed.

III.1.2. Meromorphic forms and vector fields

Suppose now that the field $\vec{v}$ is irrotational. The dual 1-form α is then closed, and therefore locally exact. Thus in a neighborhood of each point of S there exists a function u such that $du = \langle \vec{v}, \cdot \rangle$; in other words $\vec{v}$ is the gradient of the function $u : \vec{v} = \mathrm{grad} u$. This is often expressed the other way around by saying that the function u is a *potential* from which $\vec{v}$ is derived.

If $\vec{v}$ is both incompressible and irrotational, then u is a *harmonic function*. This follows from the relation

$$\Delta u := \mathrm{div}\ \mathrm{grad} u.$$

(Note that even though the definition of the Laplacian depends on the metric, the concept of a harmonic function depends only on the associated conformal structure.) It follows in particular that the function u and the field $\vec{v}$ are automatically analytic. Conversely, every harmonic function defines via its gradient an incompressible and irrotational flow.

Consider next the field $*\vec{v}$. The following equations hold:

$$\mathrm{curl}(*\vec{v}) = \mathrm{div}\vec{v} \quad \text{and} \quad \mathrm{div}(*\vec{v}) = -\mathrm{curl}\vec{v}.$$

Hence if $\vec{v}$ is incompressible then $*\vec{v}$ is irrotational, and *vice versa*. It follows that if the field $\vec{v}$ is both incompressible and irrotational, then so is the field $*\vec{v}$, whence, in particular, it is the derivative of a potential function u^*. Like u, the function u^* is only defined locally and up to an additive constant. The complex-valued 1-form $\eta = du + idu^*$ is, however, well-defined on the whole surface S.

Lemma III.1.2. — *The 1-form η is holomorphic.*

Proof. — Consider an open set U of S on which the field $\vec{v}$ does not vanish. Since the gradients of u and u^* are orthogonal and of the same norm, the map $u + iu^* : U \to \mathbb{C}$ is holomorphic. Observe that, since this map is a local diffeomorphism, the functions u and u^* give conformal local coordinates on U. Another way of saying this is that the 1-form η is holomorphic on the set consisting of S with the zeros of $\vec{v}$ removed. However, since that form is defined on the whole of S (and the zeros of $\vec{v}$ are isolated), η must then in fact be a holomorphic 1-form on S. □

Conversely, given a holomorphic 1-form η on S, the dual field of the real part of η, that is, the field $\vec{v}$ defined by

$$\operatorname{Re} \eta = \langle \vec{v}, \cdot \rangle,$$

is both incompressible and irrotational. This allows us to gain an understanding of the local properties of the critical points of incompressible and irrotational vector fields. In a neighborhood of such a point we have $\eta = df$ for some holomorphic function f. Hence there exists a local holomorphic coordinate z and a non-negative integer n such that $f(z) = z^n$, whence $\eta = nz^{n-1}dz$ and $\langle \vec{v}, \cdot \rangle = \operatorname{Re}\,(nz^{n-1}dz)$. In the case $n = 3$ the field lines are as in Figure III.1.

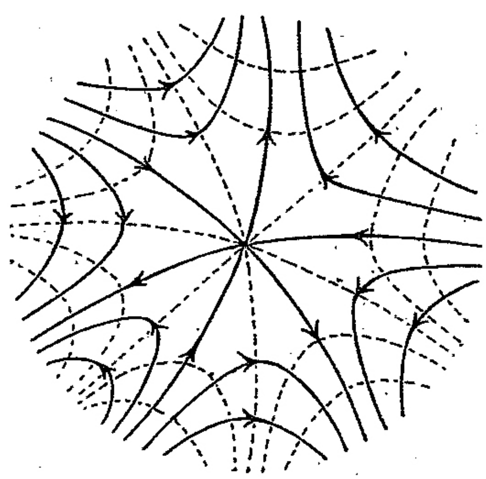


Figure III.1: A figure taken from Klein's book [Kle1882c]: the neighborhood of a critical point

In his course Klein also considered the case where the form η has poles. The field $\vec{v}$ is then defined only on the surface with the poles of η removed, which are, of course, finite in number. We shall now examine qualitatively the behaviour of the flow lines of $\vec{v}$ in the vicinity of the poles. In a neighborhood of a pole of η one can always find (see Box III.2) a holomorphic local coordinate w such that

$$\eta = \left(\frac{\lambda}{w} + \frac{1}{w^\nu}\right) dw,$$

where $\lambda \in \mathbb{C}$. Hence the field $\vec{v}$ dual to the real form Re η decomposes as a superposition of the fields dual to the forms $\lambda dw/w$ and dw/w^ν. Consider first the case $\nu = 1$, where we have $\eta = \mu dw/w$, with $\mu = \lambda + 1$. We now further decompose the field dual to Re $(\mu dw/w)$ as the superposition of a field with μ real and another with μ purely imaginary. For real μ one finds that the potential of $\vec{v}$ is, to within an additive constant, the function $u = \mu \log r$, where $w = re^{i\theta}$. In this case the field lines are orthogonal to the concentric circles about the point $w = 0$, which is then either a positive *source* or a negative *sink*, depending on the sign of μ (see Figure III.2).

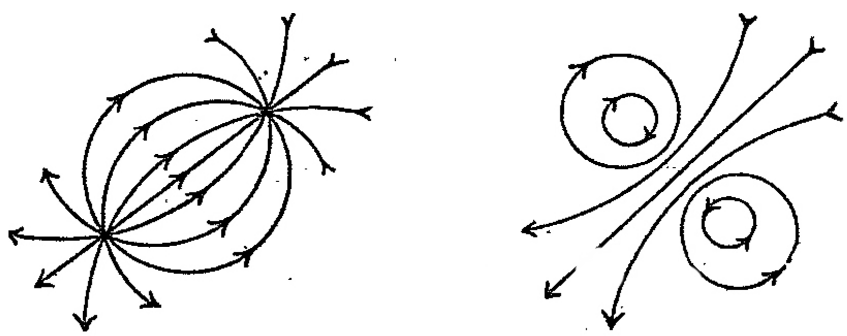


Figure III.2: Taken from Klein's book [Kle1882c]: sink/source and vortices

When μ is purely imaginary the potential is, to within an additive constant, the function $u = i\mu\theta$, and the flow lines are then concentric circles about $w = 0$, traced out with speed $|\mu|$. We have here the case of a *vortex* (see Figure III.2).

The case of the field dual to the 1-form dw/w^2 is dealt with by first observing that

$$\frac{1}{2\varepsilon}\left(\frac{dw}{w-\varepsilon} - \frac{dw}{w+\varepsilon}\right) \longrightarrow \frac{dw}{w^2} \text{ as } \varepsilon \to 0.$$

For real ε this represents the superposition of a source "of flow ε" and a sink "of flow $-\varepsilon$" positioned at points p and p' a distance 2ε apart (see Figure III.3).

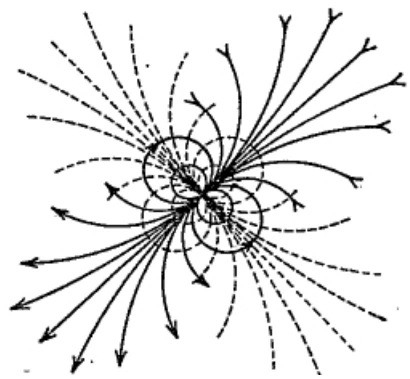


Figure III.3: Taken from Klein's book [Kle1882c]: a dipole

The same procedure can be used to investigate the fields dual to the form Re (dw/w^{ν}), ν any integer ≥ 2. Thus one arranges for ν points representing sources, sinks, or vortices to approach the same limit.

We next consider how to interpret the periods of the holomorphic 1-form η in terms of the vector field $\vec{v}$. Let $a \in H_1(S,\mathbb{Z})$ be any particular homology class. The real part of the period of η on the class a is then by definition

$$[\mathrm{Re}\ \eta](a) := \mathrm{Re}\ \left(\int_{\gamma} \eta\right),$$

where γ is an oriented (multi-)curve representing a. As in Box III.1, we denote by $\vec{t}$ the unit vector field tangent to this curve and by $\vec{n}$ the unit normal vector field to the curve, chosen so that $(\vec{t},\vec{n})$ furnishes an indirect basis for the tangent space[4] at each point of the curve. As before, $\vec{v}$ denotes the field dual to Re η; thus Re $\eta = \langle \vec{v}, \cdot \rangle$. Observing then that $\langle \vec{v},\vec{t}\rangle = -\langle *\vec{v},\vec{n}\rangle$, we deduce the following equality (where dl is the element of length along γ):

$$[\mathrm{Re}\ \eta](a) = \int_{\gamma} \langle -(*\vec{v}),\vec{n}\rangle dl\,.$$

Thus the period $[\mathrm{Re}\ \eta](a)$ is equal to the flux of the field $-(*\vec{v})$ across the curve γ.

Furthermore, by means of the first of the relations (III.2) we may rewrite the equality Re $\eta = \langle \vec{v}, \cdot\rangle$ in the form Re $\eta = -\mathrm{vol}(*\vec{v},\cdot)$, whence

$$[\mathrm{Re}\ \eta](a) = -\int_{\gamma} \mathrm{vol}(*\vec{v},\cdot).$$

[4]That is, one not agreeing with the chosen orientation of the surface. *Trans*

The period $[\mathrm{Re}\,\eta](a)$ is therefore also equal to the surface area of the infinitesimal cylinder obtained by displacing the curve γ by the flow of $-(*\vec{v})$.

Box III.2: The local normal form of a meromorphic form

Consider a meromorphic form η in a neighborhood of one of its poles. We explain here how one finds a local coordinate w in such a neighborhood in terms of which η assumes the normal form $\eta = \left(\frac{\lambda}{w} + \frac{1}{w^\nu}\right) dw$.

In terms of an arbitrary fixed holomorphic local coordinate z, the form η can be written as

$$\eta = \frac{\lambda}{z} dz + d\left(\frac{h(z)}{z^{\nu-1}}\right),$$

where h is a holomorphic function, $\lambda \in \mathbb{C}$ is the residue of the form η at 0, and $\nu \geq 2$ is an integer. We seek a coordinate change of the form $w(z) = z \cdot u(z)$ where u is holomorphic and $u(0) = 1$, such that

$$\frac{\lambda dz}{z} + d\left(\frac{h}{z^{\nu-1}}\right) = \frac{dw}{w^\nu} + \lambda \frac{dw}{w}.$$

In view of the assumed expression for w, this simplifies to

$$\frac{dw}{w^\nu} + \lambda \frac{du}{u} = d\left(\frac{h}{z^{\nu-1}}\right).$$

Integrating, we obtain

$$\frac{-1}{(\nu-1)w^{\nu-1}} + \lambda \log u - \frac{h}{z^{\nu-1}} = C,$$

where $C \in \mathbb{C}$ is a constant. Multiplication by $z^{\nu-1}$ then yields:

$$\frac{-1}{(\nu-1)u^{\nu-1}} + \lambda z^{\nu-1} \log u - h - C z^{\nu-1} = 0.$$

Denote by $\Phi(u,z)$ the left-hand side of this equation. In a neighborhood of the point $(1,0)$, $\Phi(u,z)$ is a holomorphic function of the two variables u and z. Its derivative with respect to u at the point $(u,z) = (1,0)$ is 1, so the holomorphic version of the Implicit Function Theorem may be applied, yielding the holomorphic local function $u(z)$ required for the coordinate change from z to $w(z) = z \cdot u(z)$.

The above considerations thus allow us to interpret Theorem II.2.5 (on the existence of meromorphic forms on a given surface) in terms of vector fields (or, in more physical terminology, "flows"). Finding a meromorphic 1-form η with poles at prescribed points then becomes equivalent to constructing an incompressible and irrotational vector field $\vec{v}$ having singularities at the poles of η (sinks, sources, or vortices). Similarly, prescribing the periods amounts to pre-determining the flux across curves forming a basis for the homology of the surface.

III.1.3. Experimental proof of Theorem II.2.5

In his course [Kle1882c] Klein described electrostatic or hydrodynamical experiments yielding incompressible and irrotational vector fields. His idea was to exhibit a stationary flow with prescribed singularities and flux across certain curves.

To that end we imagine the given surface to be made of a perfectly conducting, infinitely thin material. If one attaches the two terminals of a battery to the surface, a flow of charge will occur with a source and a sink at the points where the terminals are attached. The physics of the motion of electric charges assures us that this flow will be incompressible and irrotational away from the singularities. We can further imagine that we can arrange by means of suitable electromotive forces[5] for the flux across certain curves to be as prescribed. As we have seen, this amounts to fixing the real parts of the periods.

By way of example, we investigate how one might generate a prescribed flow in the plane. We first recall a few relevant facts. The curl of a given vector field Y, denoted by curl Y, in Euclidean $\mathbb{R}^3$ is the vector field defined by the following equality of differential 2-forms:

$$d(\langle Y, \cdot\rangle) = \mathrm{vol}(\mathrm{curl}\, Y, \cdot, \cdot),$$

where $\langle\cdot,\cdot\rangle$ and vol are respectively the Euclidean scalar product and the usual volume form in $\mathbb{R}^3$.

In $\mathbb{R}^3$ the electric field $\overrightarrow{E}$ and magnetic field $\overrightarrow{B}$ of a stationary electromagnetic field are given by the following special case of *Maxwell's equations*:

$$\mathrm{div}\overrightarrow{E} = \frac{\rho}{\varepsilon_0}, \quad \mathrm{curl}\overrightarrow{E} = 0,$$
$$\mathrm{div}\overrightarrow{B} = 0, \quad \mathrm{curl}\overrightarrow{B} = \mu_0 j,$$

[5]The practical realization of such constraints poses a problem. Klein found himself obliged to imagine surfaces made up of pieces kept at different temperatures. Our treatment will remain at the level of a thought experiment.

where ε_0 (the dielectric permittivity of the vacuum) and μ_0 (the magnetic permeability of the vacuum) are constants, ρ is the charge density, and j is a vector field representing current density. Note that where the quantities ρ and j vanish, both the electrostatic and magnetic fields are incompressible and irrotational.

This allows one to produce examples of incompressible and irrotational flows on the plane. Thus consider finitely many points $P_1, \ldots, P_n$ of the plane[6] $\mathbb{R}^2$, and arrange a sequence of charges of uniform density on each of the vertical lines $\{P_i\} \times \mathbb{R} \subset \mathbb{R}^3$. Since this configuration is invariant under the symmetry $(x, y, z) \mapsto (x, y, -z)$, the electrostatic field must be tangential to horizontal planes, and therefore determines an incompressible and irrotational flow on the plane $\mathbb{R}^2 \times \{0\}$ away from the singularities, which will be sources or sinks depending on the charge densities one has chosen (see Figure III.4).

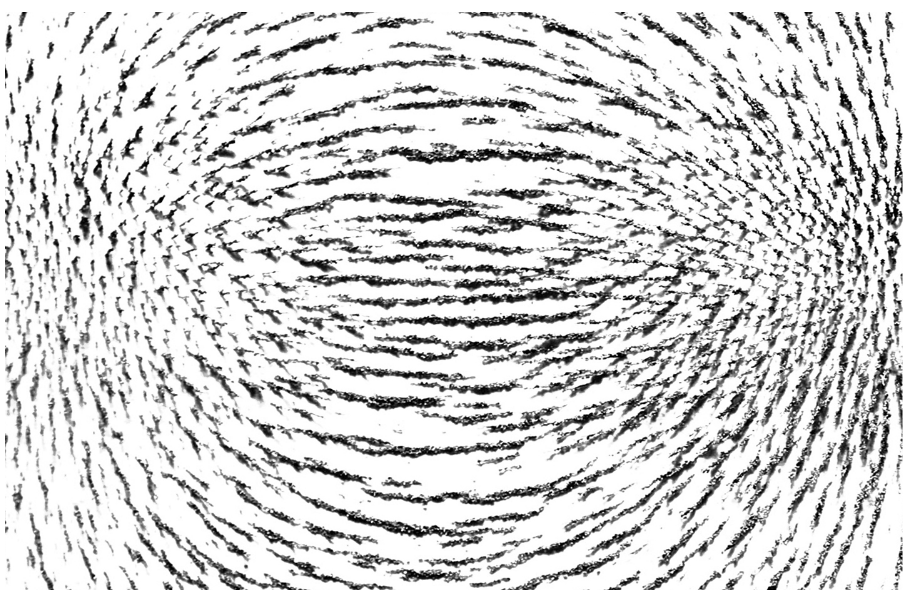


Figure III.4: A planar electrostatic field

Similarly, one may have constant electric currents passing through the verticals $\{P_i\} \times \mathbb{R}$, in which case the magnetic field $\overrightarrow{B}$ in the stationary regime is again tangential to horizontal planes and induces an incompressible and irrotational flow on the horizontal (x, y)-plane away from finitely many points where there are vortices.

By superimposing these two sorts of fields, one obtains experimentally all the types of poles possible for meromorphic 1-forms on $\mathbb{C}$.

[6]Considered as the (x, y)-plane of (x, y, z)-space $\mathbb{R}^3$. *Trans*

III.2. Revisiting Riemann's theory in modern terms

III.2.1. Hodge theory and a modern proof of the existence of meromorphic forms on a given surface

Klein's idea of endowing a given Riemann surface with a Riemannian metric compatible with the complex structure underlies modern proofs of the existence of meromorphic forms on the surface (Theorem II.2.5). The proof we give here was inspired by Springer's book [Spr1957].

Once again we consider a compact Riemann surface S endowed with a compatible Riemannian metric. Let ω be a smooth 1-form on S. Recall that such a form ω is called *co-closed* if $*\omega$ is closed, and *harmonic* if it is both closed and co-closed.

Our proof of Theorem II.2.5 relies on the following particular result of *Hodge theory*, a theory applying in all dimensions, and developed precisely in order to generalize the two-dimensional situation:

Theorem III.2.1. — *Every real, smooth 1-form ω decomposes uniquely as a sum of three 1-forms:*

$$\omega = \omega_h + dF + *dG,$$

where ω_h is a smooth harmonic form and F and G are real-valued, smooth functions defined globally on S.

Of course the uniqueness applies to the 1-forms ω_h, dF, $*dG$, and not to the functions F and G, which are defined only to within additive constants. We first explain why this result implies Theorem II.2.5.

Derivation of Theorem II.2.5 from Theorem III.2.1.

We begin by proving the theorem in question for holomorphic forms. Since a holomorphic 1-form is completely determined by its real part and since harmonic 1-forms are just the real parts of holomorphic 1-forms, it suffices to prove that there is a unique harmonic 1-form with the prescribed periods.

We shall use, without justification, a very elementary version of de Rham's theorem — a result, needless to say, very substantially postdating Klein: firstly, on a compact orientable surface of genus g one can define a closed 1-form with its $2g$ periods prescribed, and, secondly, such a form is exact if and only if these periods are all zero.

Hence for the *existence* of the desired harmonic 1-form, consider a real-valued closed 1-form ω with the prescribed periods (as guaranteed by de Rham's theorem), and apply Theorem III.2.1 to it. Since ω is closed its co-exact part $*dG$ is also closed. The following proposition shows that in fact we must then have

$*dG = 0$, so that ω is in the same cohomology class as its harmonic component ω_h, which thus provides a solution to the problem.

Proposition III.2.2. — *For every smooth function $G : S \to \mathbb{R}$ one has*

$$\int_S |*dG|^2 \operatorname{vol} = -\int_S G \cdot d(*dG).$$

Proof. — Since the sum

$$G \cdot d(*dG) + dG \wedge *dG$$

is exact (being equal to the differential of the 1-form $G \cdot *dG$) we have

$$\int_S G \cdot d(*dG) = -\int_S dG \wedge *dG,$$

and since here the left-hand side is zero ($*G$ being closed), so is the right-hand side. On the other hand since the forms dG and $*dG$ are orthogonal and have the same norm, we have $dG \wedge *dG = |*dG|^2 \operatorname{vol}$, whence the desired conclusion. □

For the *uniqueness* it suffices to prove that a harmonic 1-form ω with all its periods zero must itself be zero. Now such a form is exact (by de Rham's theorem — see above) and its primitive is a harmonic function. The maximum principle for harmonic functions then implies that such a function is constant, so that ω is zero.

We now turn to the case of meromorphic forms. Thus let $P_1, \ldots, P_m$ be any prescribed points of the surface, $A_1, \ldots, A_m$ any complex numbers summing to zero, and choose any real numbers for the prescribed real parts of the periods. We seek a meromorphic form on S with its poles at the points P_i with principal parts given by the A_i, and with the real parts of its periods as chosen.

Let α_0 be a smooth real 1-form on $S \setminus \{P_1, \ldots, P_m\}$ satisfying

$$\alpha_0 = \operatorname{Re}\left((A_i z_i^{-1} + B_i z_i^{-2} + C_i z_i^{-3} + \cdots)dz_i\right)$$

in a neighborhood of each point P_i. Since α_0 is harmonic in a neighborhood of each of the points P_i, the 2-form $d\alpha_0$ is zero there, and therefore extends to a smooth form on all of S.

Lemma III.2.3. — *We have* $\int_S d\alpha_0 = 0$.

Proof. — For each $i = 1, \ldots, m$, let D_i be a disc centered at P_i on which $d\alpha_0 = 0$. We then have by Stokes' theorem

$$\int_S d\alpha_0 = \int_{S \setminus \cup_i D_i} d\alpha_0 = -\sum_i \int_{\partial D_i} \alpha_0.$$

The last sum is equal to

$$\operatorname{Re}\left(2i\pi\sum_i A_i\right),$$

which is zero by the assumption on the prescribed residues of the principal parts. □

Since $d\alpha_0$ is a smooth form with zero integral over S, it admits a primitive ω, say. Consider the closed 1-form $\alpha_1 = \alpha_0 - \omega$ on $S\setminus\{P_1, \dots, P_m\}$. As earlier, the 2-form $d(*\alpha_1)$ extends to a smooth 2-form on S with zero integral over S. Taking β to be a smooth primitive of the 2-form $d(*\alpha_1)$ on S and applying Theorem III.2.1 to it, we obtain $\beta = \beta_h + dF + *dG$. From the equality $d\beta = d(*\alpha_1)$ we then infer that

$$d(*dG) = d(*\alpha_1).$$

The 1-form $\alpha_2 = \alpha_1 - dG$ is closed (since α_1 is closed) and co-closed by the preceding equation. It is therefore harmonic away from its poles. We have thus found the real part of the desired meromorphic form, and the proof of Theorem II.2.5 is complete. □

Proof of Theorem III.2.1. — We introduce the space $\Omega^1_{L^2}(S)$ of differentiable real 1-forms on S whose coefficients are measurable square integrable functions. This is a Hilbert space once endowed with the scalar product

$$\langle\omega_1, \omega_2\rangle_{L^2} := \int_S \langle\omega_1, \omega_2\rangle_P \operatorname{vol}_P = \int_S \omega_1 \wedge *\omega_2,$$

where the second equality comes from (III.2).

We denote by E the closure of the set of exact smooth forms in $\Omega^1_{L^2}(S)$, and by E^* the closure in $\Omega^1_{L^2}(S)$ of the set of smooth co-exact forms — that is, those expressible as $*dF$ for a smooth function F. For any smooth 1-form ω and smooth function F, one has

$$\int_S \langle dF, \omega\rangle \operatorname{vol} = -\int_S dF \wedge *\omega = \int_S F d(*\omega), \tag{III.3}$$

where the first equality follows from (III.2), and the second from Leibniz's formula $d(F * \omega) = dF \wedge *\omega + F d(*\omega)$ and Stokes' theorem.

By substituting a co-exact form $\omega = *dG$, G a smooth function, in these equations, one infers that $\langle dF, *dG\rangle_{L^2} = 0$. Thus the spaces E and E^* are orthogonal. Writing H for the orthogonal complement of $E \oplus E^*$, one then has the following decomposition of our original Hilbert space as a direct sum:

$$\Omega^1_{L^2}(S) = H \oplus E \oplus E^*.$$

It remains to show that H is actually the space of smooth harmonic forms. For this it suffices to prove that every element of H is locally a smooth harmonic form.

Since S comes with locally conformal coordinates, by choosing the metric to be locally the standard Euclidean one (which is permitted since the above function spaces depend only on the conformal structure of S), the problem is reduced to the following local lemma:

Lemma III.2.4 (Weyl's lemma). — *Consider the unit disc $\mathbb{D}$ endowed with the Euclidean metric $dx^2 + dy^2$. If a measurable square integrable 1-form on $\mathbb{D}$ is orthogonal to all exact or co-exact forms with compact support, then it is harmonic.*

Proof. — Let ω be a square integrable 1-form on $\mathbb{D}$ orthogonal to every exact or co-exact form on $\mathbb{D}$ with compact support. We first prove that if ω is also assumed smooth, then it is harmonic. For this we need to show that ω is both closed and co-closed. Since $*\omega$ has the same basic properties as ω, it suffices to show, for instance, that $*\omega$ is closed. Now since by assumption ω is orthogonal to all exact forms with compact support, by (III.3) the form $d(*\omega)$ is orthogonal to all functions with compact support, so must in fact be zero.

The idea is now to regularize ω by convolution. If the convolution kernel is chosen to be rotation-invariant, then, as we shall show, the form ω will be equal to its convolution, by virtue of the mean-value property of a harmonic function according to which its value at a point is equal to its average on any circle centered at the point. Thus we now prove that ω coincides with its convolute, whence its smoothness.

The details are as follows. For each $\rho \in (0,1)$, we denote by D_ρ the closed disc of radius ρ centered at 0, and choose a *regularizing kernel* $(K_\rho)_{\rho\in(0,1)}$, with, for each $\rho \in (0,1)$, the following properties:

1. K_ρ is a smooth non-negative function on $\mathbb{D}$ with support D_ρ and integral equal to 1;
2. $K_\rho(x,y)$ depends only on $x^2 + y^2$.

Then for every integrable function $f : \mathbb{D} \to \mathbb{R}$ and every $\rho \in (0,1)$, we consider the function $M_\rho f$ defined by

$$M_\rho f(x,y) = \int_{\mathbb{D}} K_\rho(x'-x, y'-y) f(x',y') dx'dy' .$$

Similarly, for every 1-form $\omega = \omega_x dx + \omega_y dy$ on $\mathbb{D}$ and every $\rho \in (0,1)$, we consider the 1-form $M_\rho\omega$ defined by

$$M_\rho\omega = (M_\rho\omega_x)dx + (M_\rho\omega_y)dy.$$

Such functions and forms have the following properties:

(i) For every integrable function f, the function $M_\rho f$ is defined and smooth on $D_{1-\rho}$.

(ii) For every integrable function f, one has $M_\rho df = d\left(M_\rho f\right)$ on $D_{1-\rho}$. Similarly, for every 1-form ω, one has $M_\rho(*\omega) = *(M_\rho\omega)$ on $D_{1-\rho}$.

(iii) For every two 1-forms ω_1 and ω_2 where the support of ω_1 is contained in $D_{1-\rho}$, one has $\langle M_\rho\omega_1, \omega_2\rangle = \langle\omega_1, M_\rho\omega_2\rangle$.

(iv) For every integrable function f, and all $\rho, \rho' \in (0,1)$, one has $M_\rho M_{\rho'} f = M_{\rho'} M_\rho f$ on $D_{1-\rho-\rho'}$.

(v) For every $0 < r < 1$ and function f in $L^2(D_r)$ the functions $M_\rho f$ tend to f as ρ tends to 0.

(vi) If u is a harmonic function on $\mathbb{D}$ then $M_\rho u = u$ on $D_{1-\rho}$.

The last point is crucial, and follows from the mean-value property of harmonic functions and the choice of K_ρ as rotation-invariant.

Properties (ii) and (iii) show that if ω is orthogonal to all smooth exact or co-exact forms with compact support, then $M_\rho\omega$ is orthogonal to all smooth exact or co-exact forms with support contained in $D_{1-\rho}$. Then since $M_\rho\omega$ is smooth on $D_{1-\rho}$ by property (i), it follows, as we have already seen, that it is in fact harmonic in $D_{1-\rho}$.

It remains to prove that $M_\rho\omega$ is almost everywhere equal to ω on $D_{1-\rho}$. To this end, note that properties (ii) and (vi) imply that for every ρ, ρ' with $0 < \rho, \rho' < 1$, one has

$$M_{\rho'} M_\rho\omega = M_{\rho'}(du) = d\left(M_{\rho'} u\right) = du = M_\rho\omega$$

on the disc $D_{1-\rho-\rho'}$, where u is the potential of $M_\rho\omega$. It then follows from (v) that

$$M_\rho\omega = M_{\rho'} M_\rho\omega = M_\rho M_{\rho'}\omega = M_{\rho'}\omega$$

on the disc $D_{1-\rho-\rho'}$. Hence for each r, $0 < r < 1$, the family $M_\rho\omega$ is constant for $0 < \rho < 1 - r$ and approaches ω in $\Omega^1_{L^2}(B(0,r))$ as ρ approaches 0.

We have thus proved that ω is almost everywhere equal to a smooth harmonic form, and the lemma follows. □

To conclude the proof of Theorem III.2.1, we need a second local lemma.

Lemma III.2.5. — *Let $\mathbb{D}$ be the unit disc endowed with the Euclidean metric $dx^2 + dy^2$, and let ω be a smooth 1-form on $\mathbb{D}$. There then exist smooth functions F and G on the disc such that*

$$\omega = dF + *dG.$$

Remark III.2.6. — This lemma might seem to indicate that in the statement of Theorem III.2.1 the harmonic form ω_h is always zero. However, this is certainly not always the case: Lemma III.2.5 is a local result applying specifically to the disc.

Proof. — If ω is closed then it is exact on $\mathbb{D}$. Thus we seek a function G such that $\omega - *dG$ is closed. Write $d\omega = \phi dx \wedge dy$, the measure of the departure of ω from being closed. We may assume ω is defined throughout the plane by extending it smoothly to the whole of $\mathbb{R}^2$.

It follows from (III.3) that

$$d(\omega - *dG) = (\phi - \Delta G)dx \wedge dy\,.$$

Thus it remains to solve the equation $\Delta G = \phi$ on $\mathbb{D}$. If ϕ is a Dirac point-mass at 0, the Green's function $G_0(re^{i\theta}) = -\log(r)$ is a solution. In the general case it suffices by linearity to take the convolution of ϕ with G_0. Thus one checks that the following function works:

$$G(x, y) = -\frac{1}{2\pi}\int_{\mathbb{D}} \log \sqrt{(x' - x)^2 + (y' - y)^2}\varphi(x', y')dx'dy'\,.$$

□

We are now within reach of the completion of the proof of Theorem III.2.1. Let ω be a smooth 1-form on S. We have already shown that

$$\omega = \omega_h + a + b,$$

where ω_h is harmonic, a belongs to E and b belongs to E^*. We first show that a and b are smooth. It suffices, of course, to prove this locally, so we consider a disc D in S, sufficiently small to admit conformal coordinates (x, y). By Lemma III.2.5 there exist smooth functions F and G such that

$$\omega_h + a - dF = *dG - b.$$

Here the left-hand side form is orthogonal to co-exact forms with compact support on D, while the right-hand side form is orthogonal to exact forms with compact support on D. This differential form is therefore harmonic and smooth by Weyl's lemma (Lemma III.2.4), forcing the regularity of the forms a and b.

Finally, we need to show that a and b are exact. This is done by showing, for instance by means of Ascoli's theorem, that a smooth form that is a limit of exact smooth forms in the L^2-topology, is itself exact. □

III.2.2. The continuity of the dependence of the moduli on the metric

As before, our context is that of a compact Riemann surface S endowed with a compatible Riemannian metric h. We saw in Section II.3.1 that for $g > 1$ Riemann defined certain complex local parameters, or *moduli*, for what is now called the "moduli space" $\mathcal{M}_g$ of complex curves of genus g. Here we are interested in those of the second type, namely the "periods" of a holomorphic 1-form on S.[7]

One may ask whether these moduli depend continuously on the chosen Riemannian metric. The space $\mathrm{Met}(S)$ of Riemannian metrics on S is naturally endowed with a topology[8]. As we have seen, each such metric h determines the structure of a Riemann surface on S, which we shall denote by $S^{\mathbb{C}}(h)$. We may therefore ask how the moduli vary with h. To be precise, the aim of the present section is to indicate why the map

$$\begin{array}{ccc} \mathrm{Met}(S) & \to & \mathcal{M}_g \\ h & \mapsto & S^{\mathbb{C}}(h) \end{array}$$

is continuous with respect to the topology "defined" by Riemann on the space $\mathcal{M}_g$ of (isomorphism classes of) Riemann surfaces of genus g.

We briefly expound the ideas allowing one to prove that this map is indeed continuous. With each Riemannian metric h one can associate a subspace of dimension $2g$ of the space of real differential 1-forms, namely the space $\mathrm{Harm}^1_h(S,\mathbb{R})$ of harmonic forms. One can view this space as the kernel of the Laplacian $\Delta_h = dd^* + d^*d$[9] associated with h. This is an elliptic operator varying continuously with h. Fredholm theory allows one to establish the following theorem (see [Hod1941]):

Theorem III.2.7 (Hodge). — *Let (S,h) be a Riemannian surface (compact, oriented, and without boundary) of genus g. Then in the space $\Omega^1(S,\mathbb{R})$ of C^∞ 1-forms on S, the subspace $\mathrm{Harm}^1_h(S,\mathbb{R})$ of harmonic forms (that is, both closed and co-closed) has dimension $2g$ and varies continuously with the metric h.*

To see that $S^{\mathbb{C}}(h)$ depends continuously on h, one first observes that the Hodge star operator $*_h$ defines a complex structure on $\mathrm{Harm}^1_h(S,\mathbb{R})$ since $*_h^2 = -\mathrm{Id}$ in this degree. Since moreover the Hodge star commutes with the Laplacian, the eigenspace of *holomorphic* forms

$$H^{1,0}(h) = \ker(*_h + i\,\mathrm{Id}) \subset \mathrm{Harm}^1_h(S,\mathbb{C})$$

also varies continuously with h. (Here $\mathrm{Harm}^1_h(S,\mathbb{C})$ stands for the space of harmonic complex-valued 1-forms on S.)

---

[7] Recall that for this a basis for the homology of S needs to be distinguished.

[8] The C^∞ topology.

[9] Here d^* is the adjoint of d with respect to the scalar product on L^2 defined above, now extended to all forms.

Let $(A_i, B_i)_{1 \leq i \leq g}$ be a fixed basis for the homology $H_1(S)$. The forms ω_i defined by

$$\int_{A_j} \omega_k(h) = \delta_{jk},\ 1 \leq j, k \leq g\,,$$

then comprise the intersection of the space of holomorphic forms (which depends continuously on h) with some affine subspace (independent of h), and so vary continuously with h.

In particular, the periods $\int_{B_j} \omega_k(h) = \Pi_{jk}(h)$ are continuous, as well as the zeros $P_i(h)$ of the $\omega_j(h)$, and so linear combinations of them (with constant coefficients), whence also the integrals $\int_{P_1(h)}^{P_i(h)} \omega(h)$ between two such zeros.

Note that we know today that in fact the Π_{jk} determine the complex curve (Torelli's theorem).

Chapter IV

Schwarz's contribution

In this chapter we expound the relevant works of Hermann Schwarz. Around 1870 he undertook to prove the Riemann Mapping Theorem in particular cases and to find expressions for the uniformizing functions.

In the introduction to the article [Schw1869], Schwarz tells of how, when he was attending Riemann's course on the theory of analytic functions during the winter of 1863–1864, Franz Mertens remarked to him that it was extraordinary that, although the existence of a biholomorphic mapping between the disc and, say, a triangle was "established" by Riemann's theorem, it was not at all clear how one might go about determining such a mapping explicitly. It is primarily to the problem of explicit uniformization of certain simply-connected regions of the plane that Schwarz addresses himself in [Schw1869]. We shall see that in large measure he succeeds: he obtains a necessary expression for the biholomorphism in each of the cases he considers. However this expression depends on certain constants — *accessory parameters* — which he is able to determine explicitly only in the case of a "triangle" with sides arcs of circles. Schwarz's method — marking the first connection between the *Schwarzian derivative* and the uniformization problem — is the direct forerunner of Poincaré's approach via differential equations. When, following on the publication of Poincaré's papers, Schwarz eventually discovered this, he added a note to his paper as it appeared in his complete works (see Chapter IX).

Interest in the existence of biholomorphisms between the disc and certain regions of the plane was quickened at the time by Weierstrass's objections to Riemann's proof. In his next article [Schw1870] Schwarz, abandoning the search for explicit formulae, gives a different proof of the Riemann Mapping Theorem for compact regions with analytic boundary. His method is constructive, proceeding via successive approximations. This represents a decisive stage on the way to the uniformization theorem; it will later be taken up and elaborated on by Poincaré under the name "the scanning method". We shall expound Schwarz's results and examine their influence in detail in Chapter IX; here we confine ourselves to describing the method in a simple case yielding nonetheless the first result in the direction of the uniformization of abstract compact surfaces, namely the uniqueness of the conformal structure on the sphere.

IV.1. Explicit cases of conformal representation

One of the first directions taken by Schwarz's work was the explicit determination of certain conformal maps. The point of departure of the paper [Schw1869] is the *symmetry principle*, to which his name was thenceforth attached.

Theorem IV.1.1 (Symmetry principle). — *Let U be an open set of the upper half-plane $\mathbb{H}$, with closure intersecting the real axis in an interval I, and let $z \mapsto f(z)$ be a holomorphic function on U. We assume that f extends via continuity to the union $U \cup I$ and that I is sent by this extension to (an arc of) a circle C. Then, denoting by U' the image of U obtained by reflecting in the real axis and by σ the Möbius inversion relative to the circle C, one can extend f to a function holomorphic on $U \cup I \cup U'$ by means of the formula $f(\overline{z}) = \sigma \circ f(z)$.*

With the aid of this principle, Schwarz was able to infer the form of uniformizing functions first for polygonal regions and then for such regions bounded by circular arcs. He was unable to deduce directly the existence of the uniformizing functions — except in the case of triangular regions — since there remained accessory parameters to be determined. We shall expound this work here.

IV.1.1. Uniformization of polygonal regions with straight sides

Here we begin with a simply connected polygonal region P of the plane[1]. The boundary of P is made up of a finite number of line segments meeting in vertices $w_1, \dots, w_n$. The interior angle at w_i will be denoted by $\lambda_i \pi$, $0 < \lambda_i < 2$ (with $\lambda_i \neq 1$ at vertices where the boundary of the polygon is not "flat"). The assumption of simple connectedness implies that $\sum(1 - \lambda_i) = 2$. The problem Schwarz sets himself is that of finding a biholomorphism s from $\mathbb{H}$ onto the interior of the region P, extending to a homeomorphism on the boundary (where here the boundary of $\mathbb{H}$ is understood to be its boundary in the Riemann sphere, which is a circle).

[1] In [Schw1869] Schwarz considers to begin with the case of a square; however his approach is completely general.

We begin with the following local problem. Suppose a is a point on the real axis and that some function s sends the half-disc

$$V = \{z = a + re^{i\varphi} \mid 0 \leq \varphi \leq \pi,\ 0 < r < r_0\}$$

to an angular sector

$$S = \{z = re^{i\theta} \mid 0 \leq \theta \leq \lambda\pi,\ 0 < r < r_0\},$$

where $0 < \lambda < 2$. We shall assume that s is a homeomorphism from V to its image and that it is holomorphic on $V \cap \mathbb{H}$. We further suppose that the intersection of V with the real axis is sent to the union of the two half-lines delimiting S.

Consider the mapping $s(z)^{\frac{1}{\lambda}}$. This is defined and continuous on the intersection of $\overline{\mathbb{H}}$, the closure of $\mathbb{H}$, with a small disc $D(a, \epsilon)$ centred at a; furthermore, it is holomorphic on $D(a, \epsilon) \cap \mathbb{H}$ and is real-valued at real values of its argument. Schwarz now applies his symmetry principle to infer that $s(z)^{\frac{1}{\lambda}}$ extends to a function holomorphic on $D(a, \epsilon)$. One may therefore write, in a neighborhood of a:

$$s(z) = (z - a)^{\lambda} H(z), \tag{IV.1}$$

where $z \mapsto H(z)$ is holomorphic in that neighborhood, and does not vanish at a (since otherwise the injectivity of s would be contradicted). Moreover since $H(z)$ is real-valued at real values of its argument, the coefficients of the Taylor series for H about a, that is, of the expansion as a series in powers of $(z - a)$, are themselves real.

When a is the point at infinity, the above analysis, applied now to the function $s_1(z) = s\left(-\frac{1}{z}\right)$, gives

$$s_1(z) = z^{-\lambda} H\left(\frac{1}{z}\right) \tag{IV.2}$$

where H is a function holomorphic in a neighborhood of 0, whose Taylor series expansion about 0 has real coefficients.

Returning to our initial problem, we assume there exists a homeomorphism $s : \overline{\mathbb{H}} \to P$ holomorphic on $\mathbb{H}$. In order to determine the form of such an s we show that it must satisfy a certain natural differential equation. Note that the problem is invariant under the action (on the codomain) of the affine group. Denote by $a_1, \ldots, a_n$ the (putative) preimages of the vertices $w_1, \ldots, w_n$ under s. Modulo the application of a suitable Möbius transformation to the domain, we may assume that a_n is the point at infinity.

Thus we seek a quantity that is invariant under the affine group. The function $z \mapsto \frac{d}{dz} \log \frac{ds}{dz}(z)$ will do: it is holomorphic on $\mathbb{H}$ and invariant under affine transformations following s. The symmetry principle together with the above local investigation then allows us to prove the following lemma.

Lemma IV.1.2. — *Let $s : \mathbb{H} \to P$ be a uniformizing function which extends to a homeomorphism on the boundary. Then one has*

$$\frac{d}{dz} \log \frac{ds}{dz}(z) = \sum_{i=1}^{n-1} \frac{\lambda_i - 1}{z - a_i} .$$

Proof. — For each open interval (a_i, a_{i+1}), one can apply the symmetry principle to extend s to a function s_i holomorphic on

$$\mathbb{H} \cup (a_i, a_{i+1}) \cup \mathbb{H}^-$$

where $\mathbb{H}^-$ denotes the lower half-plane. This extension satisfies the condition $s_i(\overline{z}) = h_i \circ s_i(z)$ where h_i is the reflection in the edge (w_i, w_{i+1}). It follows, in particular, that s_i is injective, so that its derivative does not vanish. Hence the function $z \mapsto \frac{d}{dz} \log \frac{ds}{dz}(z)$ extends by continuity to $\overline{\mathbb{H}}$ with the points $a_1, \dots, a_n$ removed.

Since s maps each segment (a_i, a_{i+1}) onto the segment (w_i, w_{i+1}), there exist complex numbers A_i and B_i such that $\hat{s} = A_i s + B_i$ maps (a_i, a_{i+1}) onto an open interval of the real axis. We thus conclude, invoking the affine invariance, that $\frac{d}{dz} \log \frac{d\hat{s}}{dz} = \frac{d}{dz} \log \frac{ds}{dz}$ is real-valued on (a_i, a_{i+1}) (for each $i = 1, \dots, n$). By examining the situation locally near each of the points a_i we will be able to identify the function $\frac{d}{dz} \log \frac{ds}{dz}(z)$.

Indeed, from the local formula (IV.1), considered at each $a_i, i = 1, \dots, n-1$, we infer that

$$\frac{d}{dz} \log \frac{ds}{dz}(z) = \frac{\lambda_i - 1}{z - a_i} + d_1 + d_2(z - a_i) + d_3(z - a_i)^2 + \cdots , \tag{IV.3}$$

where the coefficients d_j are real. Hence the map defined by the difference

$$z \mapsto \frac{d}{dz} \log \frac{ds}{dz}(z) - \sum_{i=1}^{n-1} \frac{\lambda_i - 1}{z - a_i}$$

is holomorphic on $\mathbb{H}$, extends to a continuous map on $\overline{\mathbb{H}} \setminus \{\infty\}$, and is moreover real-valued at real values of its argument. We shall now show that this difference is in fact zero.

To prove this, we once again apply the symmetry principle to extend this function to an entire function. This done, in a neighborhood of the point at infinity the formula (IV.2) gives $s(z) = w_n + (z)^{-\lambda_n} H(\frac{1}{z})$, from which one infers that

$$\lim_{z \to \infty} \frac{d}{dz} \log \frac{ds}{dz}(z) = 0 .$$

Thus the entire function $\frac{d}{dz}\log\frac{ds}{dz}(z) - \sum_{i=1}^{n-1}\frac{\lambda_i - 1}{z - a_i}$ approaches 0 out to infinity, so must be identically zero. □

One now has only to integrate twice to obtain:

Proposition IV.1.3 (The Schwarz–Christoffel formula). — *Let P be a simply connected polygonal region with vertices $w_1, \dots, w_n$ and interior angles $\lambda_1\pi, \dots, \lambda_n\pi$. Let $s : \mathbb{H} \to P$ be a uniformizing function which extends to a homeomorphism on the boundary and which maps the point at infinity to w_n. Then there exist $n-1$ real numbers $a_1, \dots, a_{n-1}$ such that*

$$s(z) = C\int_{z_0}^{z}\frac{dw}{(w-a_1)^{1-\lambda_1}\dots(w-a_{n-1})^{1-\lambda_{n-1}}}. \tag{IV.4}$$

This expression is called the *Schwarz–Christoffel formula.* Such formulae had in fact been introduced independently by Christoffel [Chr1867][2]. Returning to the above argument, we see that we have shown that *if* there is a transformation s sending the upper half-plane onto the region P biholomorphically and extending to a homeomorphism on the boundary, then the composite of s with an appropriate Möbius transformation applied to the domain, is given by the formula (IV.4) for an appropriate choice of the real constants $a_1, \dots, a_{n-1}$. On the other hand, if the polygon P is fixed from the start, one is unable in general to determine the corresponding real numbers a_i, so that this approach does not provide a complete proof of Riemann's theorem for polygonal regions.

IV.1.2. Uniformization of polygonal regions with sides in the form of circular arcs

Schwarz also considers the more general case of a polygonal region P with vertices $w_1, \dots, w_n$ with sides arcs of circles or straight-line segments. It is assumed that the vertices w_i are so indexed that as one traverses $[w_i, w_{i+1}]$ from w_i to w_{i+1}, the interior of P lies to the left, that the interior angle at each vertex w_i is $\lambda_i\pi$, $0 < \lambda_i < 2$, and, once again, that we have a map s sending the upper half-plane $\mathbb{H}$ biholomorphically onto the interior of the polygon P and extending to a homeomorphism on the boundary. By composing, if necessary, with an appropriate Möbius transformation acting on the domain, we may assume that the point at infinity is not mapped to a vertex of P, and denote by $a_1 < a_2 < \cdots < a_n$ the preimages of the vertices under s.

Here one has that the problem is invariant under the action on the codomain of the group of complex Möbius transformations, that is, the group consisting of all

[2] A classic reference for conformal representation of planar regions is [Neh1952]; for constructive aspects one may consult [DrTr2002].

transformations of the form $z \mapsto \frac{az+b}{cz+d}$, $a, b, c, d \in \mathbb{C}$, $ad-bc \neq 0$. Schwarz seeks a new differential expression in s, invariant under complex Möbius transformations applied to s, playing a role analogous to that of $\frac{d}{dz} \log \frac{ds}{dz}$ in the previous section. This leads him to consider the *Schwarzian derivative*

$$\{s, z\} = \frac{d^2}{dz^2} \log \frac{ds}{dz} - \frac{1}{2}\left(\frac{d}{dz} \log \frac{ds}{dz}\right)^2.$$

Box IV.1: The Schwarzian derivative

The *cross ratio* of four points x, y, z, t of the projective line $\mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}$ is defined to be

$$[x, y, z, t] = \frac{(x-z)(y-t)}{(x-y)(z-t)}.$$

This is a "projective invariant", that is, it is unchanged by any Möbius transformation applied to x, y, z, t.

The *Schwarzian derivative* is a "local" projective invariant measuring the "departure from infinitesimal projectivity" of a local biholomorphism w of $\mathbb{CP}^1$. It can be defined in several ways — for instance by means of a comparison of the cross ratios of the four points $x, y = x + \varepsilon, z = x + 2\varepsilon, t = x + 3\varepsilon$ with that of their images under w (assuming x in the domain of w and ε sufficiently small). An elementary calculation shows that

$$[w(x), w(y), w(z), w(t)] = [x, y, z, t] - 2\{w, x\}\varepsilon^2 + o(\varepsilon^2),$$

where

$$\{w, x\} := \frac{d^2}{dx^2}\left(\log \frac{dw}{dx}\right) - \frac{1}{2}\left(\frac{d}{dx}\left(\log \frac{dw}{dx}\right)\right)^2.$$

This defines the Schwarzian derivative $\{w, x\}$ and furnishes its intuitive interpretation as a sort of projectively invariant derivative. It was actually first introduced by Lagrange in connection with his investigations into the drawing of geographical maps (see [Lag1779, p. 652], where it has the form $\frac{\Phi''(z)}{\Phi(z)}$ with $\Phi = \frac{1}{\sqrt{F'(z)}}$, according to [OvTa2009]). It was later called the "Schwarzian derivative" by Cayley, unaware of Lagrange's work.

The third-order differential equation $\{w, x\} = 0$ has the Möbius transformations as general solutions. From its very definition it is clear that the Schwarzian derivative is invariant under Möbius transformations, that is, that $\{h \circ w, x\} = \{w, x\}$ for all such transformations h.

The Schwarzian derivative of a composite of any two local biholomorphisms is calculated via the formula

$$\{f \circ g, x\} = \left(\frac{dg}{dx}\right)^2 \{f, x\} \circ g + \{g, x\}. \tag{IV.5}$$

Here the presence of the term involving $\left(\frac{dg}{dx}\right)^2$ indicates that it may be useful to interpret the Schwarzian derivative as a quadratic differential $\{f, x\}dx^2$. This allows one to interpret the preceding formula as a "cocycle"

$$\{f \circ g, x\}dx^2 = g^*(\{f, x\}dx^2) + \{g, x\}dx^2.$$

More generally, if U is an open set of a Riemann surface furnished with a coordinate $x : U \to \mathbb{C}$, one can still define the Schwarzian derivative $\{w, x\}$ of a local biholomorphism $w : U \to \mathbb{CP}^1$. Given another coordinate y on U, one has the following transformation rule:

$$\{w, x\}^2 = \{w, y\}dy^2 + \{y, x\}dx^2. \tag{IV.6}$$

This gives in particular a verification of the fact that the the quadratic differential $\{w, x\}dx^2$ is invariant under projective coordinate changes, that is, $\{y, x\} = 0$.

Let $q(x)dx^2$ be a local holomorphic quadratic differential. The third-order differential equation $\{f, x\} = q$ then admits local solutions any two of which differ by a Möbius transformation (acting on the codomain).

Here is yet another way of looking at the Schwarzian derivative. If f is a local biholomorphism between two open sets of $\mathbb{C} \cup \{\infty\}$, at each point x of its domain one can determine a unique Möbius transformation $m(x)$ in $\mathrm{PSL}(2, \mathbb{C})$ coinciding with f up to order two in x. In this way one obtains a curve in $\mathrm{PSL}(2, \mathbb{C})$ (a "Frenet frame" *à la* Darboux) whose derivative again measures the deviation of f from a Möbius transformation. This derivative is defined to be $m^{-1}dm$, regarded as an element of the Lie algebra of $\mathrm{PSL}(2, \mathbb{C})$, consisting of the zero-trace matrices. A simple calculation then yields

$$m(x)^{-1}dm(x) = -\frac{\{f, x\}}{2} \begin{pmatrix} x & x^2 \\ 1 & -x \end{pmatrix} dx.$$

Note that for the above reasons the Schwarzian derivative is also a basic tool in real projective geometry in connection with the study of the diffeomorphisms of the circle; see the book [OvTa2005].

The approach in this case is thus very similar to the earlier one, and yields as upshot:

Proposition IV.1.4. (The Schwarz equation). — *Let P be a simply connected polygonal region with sides circular arcs and with vertices $w_1, \ldots, w_n$ and interior angles $\lambda_1\pi, \ldots, \lambda_n\pi$. Let $s : \mathbb{H} \to P$ be a uniformizing function extending to a homeomorphism on the boundary. Then there exist $2n$ real numbers $a_1, \ldots, a_n$ and $\beta_1, \ldots, \beta_n$ such that*

$$\{s, z\} = \sum_{i=1}^{n} \frac{1}{2}\frac{1-\lambda_i^2}{(z-a_i)^2} + \frac{\beta_i}{z-a_i}\,. \tag{IV.7}$$

Furthermore, the λ_i, a_i, and β_i are linked by the following relations:

(i) $\sum_{i=1}^{n} \beta_i = 0$;

(ii) $\sum_{i=1}^{n} \frac{1-\lambda_i^2}{2} + \beta_i a_i = 0$;

(iii) $\sum_{i=1}^{n} a_i(1-\lambda_i^2) + \beta_i a_i^2 = 0.$

Proof. — For each $i = 1, \ldots, n$, one can by means of the symmetry principle extend s to a function s_i holomorphic on $\mathbb{H} \cup (a_i, a_{i+1}) \cup \mathbb{H}^-$, satisfying $s_i(\overline{z}) = h_i \circ s_i(z)$, where here h_i denotes inversion in the circle having (w_i, w_{i+1}) as an arc. In particular, therefore, these extensions are injective, so their derivatives are non-vanishing. It follows that $\{s, z\}$, which is holomorphic on $\mathbb{H}$, extends by continuity to $\overline{\mathbb{H}}$ with the points $a_1, \ldots, a_n$ removed. Now for each $i = 1, \ldots, n$, there is an appropriate choice of complex numbers A_i, B_i, C_i, D_i, with the property that the transformation $\hat{s} = \frac{A_i s + B_i}{C_i s + D_i}$ sends a_i to 0 and the two segments $[a_{i-1}, a_i]$, $[a_i, a_{i+1}]$ onto two straight-line segments meeting at 0 and with the angle between them equal to $\lambda_i\pi$ at 0. Applying the formula (IV.1) to $\hat{s}$, one then obtains, locally,

$$\hat{s}(z) = (z-a_i)^{\lambda_i} H_i(z),$$

with H_i holomorphic in a neighborhood of a_i. Moreover the coefficients of the power series expansion of each H_i in a neighborhood of a_i are real.

The invariance of the Schwarzian derivative under complex Möbius transformations then yields

$$\{s, z\} = \{s', z\} = \frac{1}{2}\frac{1-\lambda_i^2}{(z-a_i)^2} + \frac{\beta_i}{z-a_i} + d_{2i} + d_{3i}(z-a_i) + \cdots \tag{IV.8}$$

The coefficients of this series are real, determined by the power series expansions of the H_i.

One infers, setting

$$F(z) = \sum_{i=1}^{n} \frac{1}{2} \frac{1-\lambda_i^2}{(z-a_i)^2} + \frac{\beta_i}{z-a_i},$$

that the function $z \mapsto \{s,z\} - F(z)$, which is holomorphic on $\mathbb{H}$, extends to a continuous function on $\overline{\mathbb{H}}$. Since it is real-valued at real values of the argument, we can once again apply the symmetry principle to extend it to an entire function. It then remains to investigate the behaviour of this function at infinity in order to deduce that it is zero.

Thus we apply the symmetry principle relative to the segment $[a_n, a_1]$: the map s extends holomorphically to a neighborhood of infinity, so that we must have $s(z) = b_0 + \frac{b_1}{z} + \frac{b_2}{z^2} + \cdots$ for z of large modulus in $\mathbb{H}$. It follows that

$$\frac{ds}{dz} = -\frac{b_1}{z^2} - \frac{2b_2}{z^3} - \cdots,$$
$$\frac{d^2s}{dz^2} = \frac{2b_1}{z^3} + \frac{6b_2}{z^4} + \cdots,$$

and hence that

$$\frac{d}{dz}\left(\log \frac{ds}{dz}\right) = \frac{-2}{z}(1 + \frac{c_1}{z} + \cdots).$$

This allows us to calculate the Schwarzian up to order 3:

$$\{s,z\} = (\frac{2}{z^2} + \frac{4c_1}{z^3} + \cdots) - \frac{1}{2}(\frac{4}{z^2} + \frac{8c_1}{z^3} + \cdots).$$

Thus the Schwarzian $\{s,z\}$ has a zero of order at least 4 at infinity. Hence the function $\{s,z\} - F(z)$ is entire and vanishes at infinity. It is therefore identically zero, which establishes the first part of the proposition.

We now establish the interdependence of the a_i, λ_i and β_i. In the expansion in powers of $\frac{1}{z}$ of the rational function $\sum_{i=1}^n \frac{1}{2}\frac{1-\lambda_i^2}{(z-a_i)^2} + \frac{\beta_i}{z-a_i}$, the coefficient of $\frac{1}{z}$ is $\sum_{i=1}^n \beta_i$, the coefficient of $\frac{1}{z^2}$ is $\sum_{i=1}^n \frac{1-\lambda_i^2}{2} + \beta_i a_i$, and the coefficient of $\frac{1}{z^3}$ is $\sum_{i=1}^n a_i(1-\lambda_i^2) + \beta_i a_i^2$.

The vanishing of $\{s,z\}$ to the order 4 at infinity then yields the desired conditions:

(i) $\sum_{i=1}^n \beta_i = 0$;

(ii) $\sum_{i=1}^n \frac{1-\lambda_i^2}{2} + \beta_i a_i = 0$;

(iii) $\sum_{i=1}^n a_i(1-\lambda_i^2) + \beta_i a_i^2 = 0$. □

IV.1.3. The special case of a triangle

In Proposition IV.1.4, the constants a_i and β_i are in general impossible to determine for a given polygon. However, there is an important case where they can be found: that of a triangle with sides circular arcs. Since this case will be of central significance in Chapter IX, we provide the details here. Thus we choose a fixed such triangle in the plane, and denote by s a uniformizing function extending continuously onto the boundary.

Let w_1, w_2, and w_3 denote the vertices and $\lambda\pi$, $\mu\pi$, and $\nu\pi$ the interior angles of the triangle, and let a, b, and c on the real axis be the preimages of the vertices under s.

The equations (i), (ii), (iii) of Proposition IV.1.4 constitute a linear system in the β_i as unknowns. On solving this system one finds that the Schwarzian $\{s, z\}$ must satisfy

$$\{s,z\} = \frac{1}{(z-a)(z-b)(z-c)}\left[\frac{1-\lambda^2}{2}\frac{(a-b)(a-c)}{(z-a)} + \frac{1-\mu^2}{2}\frac{(b-a)(b-c)}{z-b} + \frac{1-\nu^2}{2}\frac{(c-a)(c-b)}{z-c}\right].$$

The crucial point now is that the parameters a, b, and c can be determined completely: by composing with a suitable Möbius transformation acting on the domain, we can arrange that $a = 0$, $b = \infty$, and $c = 1$. This done, after a little reorganization of the terms, the preceding formula becomes

$$\{s,z\} = \frac{1-\lambda^2}{2z^2} + \frac{1-\nu^2}{2(1-z)^2} - \frac{\lambda^2-\mu^2+\nu^2-1}{2z(1-z)}. \tag{IV.9}$$

We have thus found a differential equation with rational coefficients defined on $\overline{\mathbb{H}} \setminus \{0, 1, \infty\}$. In the case where the triangle is convex, every solution of this differential equation will be a uniformizing function for a triangle with appropriate angles. Then by mapping its vertices to w_1, w_2, and w_3 by a suitable Möbius transformation applied to the codomain, we obtain a uniformizing function for the triangle we began with. Thus we have the following result:

Theorem IV.1.5 (Uniformization of triangles). — *Let T be a triangle with vertices w_1, w_2 and w_3 and with angles respectively $\lambda_1\pi, \lambda_2\pi$ and $\lambda_3\pi$, where $\lambda_i \in (0, 1)$. Then the solution of equation (IV.9) sending 0, ∞ and 1 to w_1, w_2 and w_3, sends the upper half-plane biholomorphically onto the triangle T. Moreover, this solution extends to $\overline{\mathbb{H}}$ and sends $\partial\overline{\mathbb{H}}$ homeomorphically onto the boundary of the triangle.*

We conclude by examining the means employed by Schwarz to uniformize all polygonal regions with circular arcs as sides, thereby settling the question of the accessory parameters. Here is the key result proved by Schwarz in [Schw1870a]; it goes significantly further than the the situation of plane polygons. Before enunciating it we define one of our terms precisely: an open set U of a Riemann surface is *uniformizable right to the boundary* by the unit disc if there exists a biholomorphism from U onto the disc which extends to a homeomorphism between the boundary of U and the unit circle.

Theorem IV.1.6. — *Let S be a Riemann surface and U, V two open sets of S uniformizable right to the boundary by the unit disc. If the intersection $U \cap V$ is homeomorphic to a disc, then the union $U \cup V$ is uniformizable right to the boundary by the unit disc.*

This theorem can then be used in the following way to prove the existence of a uniformizing function for any polygon with circular arcs as sides: one decomposes an arbitrary quadrilateral into a union of two triangles (which we know how to uniformize) with intersection a simply connected region. Theorem IV.1.6 then assures us that we can uniformize the quadrilateral. Then one can inductively increase the number of sides.

Schwarz's proof of Theorem IV.1.6 uses his "alternating method" (the term he himself employs in [Schw1870b]), proceeding by successive approximations. We refer the reader to [Cho2007, p. 123 *et seqq.*] for another exposition of this method, and also of the very similar one of Neumann [Neum1884]. Schwarz, and then Poincaré and Koebe, extended this strategy very much further (see part C). Thus ultimately this led to a complete proof of the uniformization theorem. We ourselves shall adapt the method to prove, in Corollary XI.1.6, that any simply connected region with analytic boundary of a Riemann surface is uniformizable right to the boundary.

IV.2. The conformal structure of the sphere

We propose now to explain how Schwarz uses his alternating method in [Schw1870a] to uniformize spheres. The precise result is as follows:

Theorem IV.2.1. — *Every compact, simply connected Riemann surface S is biholomorphic to the Riemann sphere $\overline{\mathbb{C}}$.*

A few preliminary remarks are in order. First, although Schwarz's proof of this theorem in [Schw1870a] is not complete, he does not hesitate to enunciate it. Here is his original statement (where by "surface of a circle" he means a disc and by "surface of a ball" a sphere) with an approximate translation.

> Dem von Riemann ausgesprochenen Satze, dass es stets möglich sei, einen einfach zusammenhängenden Bereich zusammenhängend und in den kle-

> insten Theilen ähnlich auf die Fläche eines Kreises abzubilden, kann der folgende Satz zur Seite gestellt werden :
>
> *Es ist stets möglich, einen einfach zusammenhängenden und geschlossenen Bereich zusammenhängend und in den kleisten Theilen ähnlich auf die Fläche einer Kugel abzubilden und zwar nur auf eine Weise so, dass drei beliebig vorgeschriebenen Punkten jenes Bereiches drei ebenfalls vorgeschriebene Punkte der Kugelfläche entsprechen.*

Our translation is as follows:

> The theorem announced by Riemann that it is always possible to represent a simply connected region [of the plane] on the surface of a circle continuously and in such a way that similarity is preserved in infinitely small parts allows one to establish the following theorem:
>
> *It is always possible to represent a simply connected and closed region on the surface of a ball continuously and in such a way that similarity is preserved in infinitely small regions, and furthermore uniquely if three prescribed points of the region are to correspond to three prescribed points of the surface of the ball.*

He does not, however, give a convincing proof except in the case of polyhedral surfaces, which case does not strictly speaking fall within the purview of the theorem since the latter concerns smooth surfaces. Schwarz shows in effect that every finite, simply connected polyhedron can be mapped homeomorphically onto the Riemann sphere in such a way that the mapping is conformal on the faces. Yet in 1881 in his note [Kle1881], Klein mentions the uniqueness of the conformal structure for surfaces of genus 0 and attributes the result to Schwarz.

In order for Schwarz's argument to genuinely yield the statement of Theorem IV.2.1, it is necessary to know beforehand how to uniformize a simply connected region with analytic boundary of a Riemann surface. This result is not, however, fully established in [Schw1870a], although it is possible to obtain a complete proof via an elaboration of the alternating method. This is in fact what we undertake to do in Chapter IX. In the meantime we give a proof of Theorem IV.2.1 assuming ahead of time the uniformization of simply connected regions with boundary.

Before starting the proof we note that by 1860 the *topological* classification of compact surfaces was considered as achieved (even if a long interval of time had to elapse before it was finally established rigorously — see the box in the introduction to Part C — so that, in particular, it was considered "known" that a connected, simply connected surface without boundary is homeomorphic to the

2-dimensional sphere. It was also considered "clear" that the sphere with a disc removed is homeomorphic to a disc.

Proof. — Thus to prove the theorem we need to construct a meromorphic function f on S with a single pole of order 1 on S. We achieve this by constructing a harmonic function u, which will then serve as the real part of the desired function on S with the pole removed.

Choose two distinct points n and s of S (the north and south poles[3]), and consider a holomorphic chart $\varphi_n : U \to \mathbb{D}$ defined on an open set U containing n whose closure avoids the south pole s, and extending to a homeomorphism from the boundary of U to the unit circle. Write V for the open set $S \setminus \varphi_n^{-1}(\overline{\mathbb{D}_{1/4}})$. Since we are presupposing the uniformization theorem for simply connected regions with analytic boundary (see Corollary XI.1.6), we may assume there exists a conformal map $\varphi_s : V \to \mathbb{D}$ sending the point s to 0. The argument mainly involves the disc U. The following diagram depicts the situation.

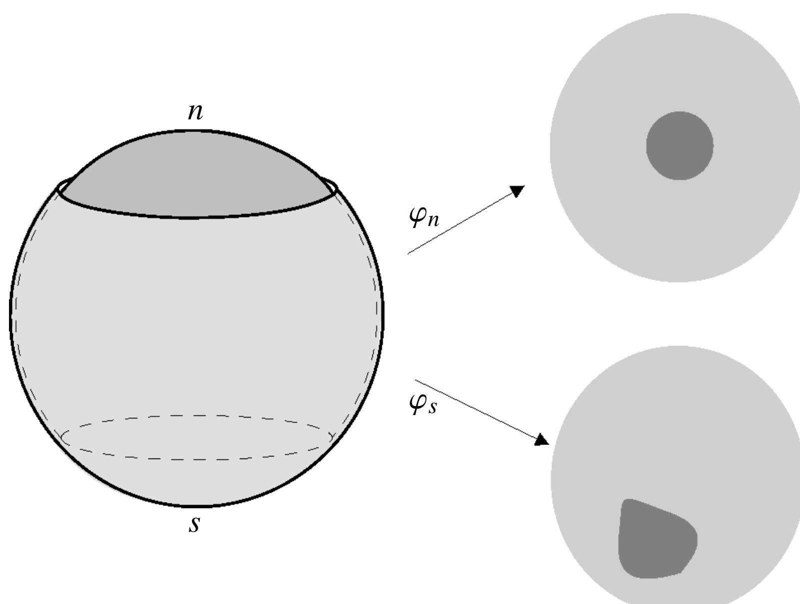


Figure IV.1: The sphere S and Schwarz's charts

The mapping φ_n (resp. φ_s) allows us to solve the Dirichlet problem for the disc U (resp. V) with any continuous boundary value. In fact it suffices to consider the

[3]Note that in view of the discussion preceding this proof, it is assumed that S is a topological 2-sphere. *Trans*

problem on the unit disc, whither it has been transferred by the map φ_n (resp. φ_s). Recall (from §II.2) that in order to solve the Dirichlet problem on the unit disc with a continuous boundary value $\underline{u}$, we use Poisson's formula, assuring us that the following function is harmonic on the disc and extends continuously via $\underline{u}$ to the boundary (here $z = re^{i\theta}$):

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} \underline{u}(e^{it})\frac{1-r^2}{1-2r\cos(\theta-t)+r^2}dt\ .$$

This formula (a variant of the formula (II.4)) allows one to obtain an upper bound for the modulus of u when the mean of $\underline{u}$ is zero on the unit circle. The value of u at the origin is clearly zero and we shall show that its values in a disc of radius $1/4$ are uniformly bounded in terms of the supremum of the modulus of u on the boundary. The precise result is as follows:

Lemma IV.2.2. — *Let $\underline{u}$ be a continuous function of mean zero on the unit circle and let u be its harmonic extension to the whole of the unit disc. Then for every point z of modulus less than $\frac{1}{4}$, one has $|u(z)| \leq \frac{2}{3}\|u\|_{\overline{\mathbb{D}}}$, where* $\|u\|_{\overline{\mathbb{D}}} = \max_{z\in\overline{\mathbb{D}}} |u(z)|$).

Proof. — This bound is established by means of a straightforward calculation using Poisson's formula. Thus for all θ one has

$$\begin{aligned} 2\pi u(re^{i\theta}) &= \int_0^{2\pi} \underline{u}(e^{it})\frac{1-2r\cos(\theta-t)+r^2+2r\cos(\theta-t)-2r^2}{1-2r\cos(\theta-t)+r^2}dt \\ &= \int_0^{2\pi} \underline{u}(e^{it})dt + \int_0^{2\pi} \underline{u}(e^{it})\frac{2r\cos(\theta-t)-2r^2}{1-2r\cos(\theta-t)+r^2}dt \\ &= 2r\int_0^{2\pi} \underline{u}(e^{it})\frac{\cos(\theta-t)-r}{1-2r\cos(\theta-t)+r^2}dt. \end{aligned}$$

Examination of the function $a \mapsto \frac{a-r}{1-2ar+r^2}$ on $[-1,1]$ quickly shows that the modulus $|u(re^{i\theta})|$ is bounded above by $\frac{2r}{1-r}\|u\|_{\overline{\mathbb{D}}}$. Hence in particular for $r < \frac{1}{4}$, the modulus $|u(re^{i\theta})|$ is bounded above by $\frac{2}{3}\|u\|_{\overline{\mathbb{D}}}$. □

It is now not difficult to construct the function f with a pole at n. We set $u_{-1}(z) = \frac{1}{r}\cos\theta$ where $\varphi_n(z) = re^{i\theta}$, and write $\hat{u}_{-1}$ for the harmonic function on the disc U satisfying the boundary condition $(\hat{u}_{-1})_{|\partial U} = (u_{-1})_{|\partial U}$. Write $u_0 = u_{-1} - \hat{u}_{-1}$; the function u_0 then vanishes on the circle ∂U.

We now begin the procedure of successive approximation, constructing inductively two sequences of functions u_k and v_k satisfying:

- u_0 is as just defined;
- for every $k \geq 0$, the function v_k is the harmonic function on the disc V satisfying the boundary condition $(v_k)_{|\partial V} = (u_k)_{|\partial V}$;

- for every $k \geq 1$, the function u_k is the harmonic function on the disc U satisfying the boundary condition $(u_k)_{|\partial U} = (v_{k-1})_{|\partial U}$.

We wish next to bound these two sequences in order to establish the convergence of the corresponding series. To this end, we note first that the maximum-modulus principle affords us the following bounds:

- for every $k \geq 1$, the modulus of v_k is bounded above by its maximum on the circle ∂V, and hence by the maximum value of $|u_k|$ on the circle ∂V. It follows that $|v_k|$ is bounded above by the maximum value of $|u_k|$;
- similarly, for every $k \geq 1$, $|u_k|$ is bounded above by the maximum value of $|v_{k-1}|$.

Thus the sequences of the $\|u_k\|_{\overline{U}}$ and $\|v_k\|_{\overline{V}}$ are decreasing. However, that does not of course suffice for the corresponding series to converge. We shall use Lemma IV.2.2 to show that for every $k \geq 1$, the function $|v_k|$ is in fact bounded above by $\frac{2}{3}\|u_k\|_{\overline{U}}$. For this purpose we need to be sure that the mean of the functions $u_k \circ \varphi_n^{-1}$ on the unit circle is zero:

Lemma IV.2.3. — *The means of all of the functions $u_k \circ \varphi_n^{-1}$ and $v_k \circ \varphi_n^{-1}$ on the circles of radius* 1 *and* $\frac{1}{4}$ *are zero.*

Proof. — Let A denote the annulus bounded by the circles of radii $\frac{1}{4}$ and 1. A function holomorphic on A will have equal integrals around the bounding circles in view of the residue formula. Recall also that a harmonic function on a simply connected region is the real part of a holomorphic function. Hence considering U and V in turn, we infer that for $k \geq 1$ the functions u_k and v_k are the real parts of holomorphic functions on $\varphi_n^{-1}(A)$. This is also true of u_0 by construction. Hence the means of the functions $u_k \circ \varphi_n^{-1}$ and $v_k \circ \varphi_n^{-1}$ on the two bounding circles are equal.

Since the function $u_0 \circ \varphi_n^{-1}$ has zero mean on the unit circle, it now follows that this is likewise true for the circle of radius $\frac{1}{4}$. Since on the one hand $u_k \circ \varphi_n^{-1}$ and $v_k \circ \varphi_n^{-1}$ coincide on the circle of radius $\frac{1}{4}$, and on the other hand $u_{k+1} \circ \varphi_n^{-1}$ and $v_k \circ \varphi_n^{-1}$ coincide on the circle of radius 1, this property propagates itself inductively to all of the $u_k \circ \varphi_n^{-1}$ and $v_k \circ \varphi_n^{-1}$, that is, they all have mean zero on the circles of radii $\frac{1}{4}$ and 1. □

Now Lemma IV.2.2 assures us that for every $k \geq 1$ the function $u_k \circ \varphi_n^{-1}$ has $\frac{2}{3}\|u_k \circ \varphi_n^{-1}\|_{\overline{\mathbb{D}}}$ as an upper bound on the circle of radius $\frac{1}{4}$. Hence the function u_k is bounded above by $\frac{2}{3}\|u_k\|_{\overline{U}}$ on ∂V. Now v_k was defined as the function solving the Dirichlet problem on V with boundary condition given by the restriction of u_k to ∂V. Hence a further application of the maximum-modulus principle yields the upper bound $\frac{2}{3}\|u_k\|_{\overline{U}}$ for $|v_k|$ on $\overline{V}$.

We thence obtain, again by induction, the following bounds (for all $k \geq 1$):

$$\|u_{k+1}\|_{\overline{U}} \leq \|v_k\|_{\overline{V}} \leq \left(\frac{2}{3}\right)^{k+1} \|u_1\|_{\overline{U}}\,.$$

Hence the two series with terms u_k and v_k respectively, $k = 1, 2, \ldots$, are dominated in modulus by the geometric series with common ratio $\frac{2}{3}$, and therefore converge respectively to a function $\hat{u}$ defined and continuous on $\overline{U}$ and harmonic on U, and a function $\hat{v}$ defined and continuous on $\overline{V}$ and harmonic on V. Writing $u = \hat{u} + u_0$ and $v = \hat{v} + v_0$, we have by construction that the functions u and v coincide on the boundaries ∂U and ∂V. Hence the function $u - v$ is harmonic in the interior of $U \cap V$ and zero on the boundary. Then by the maximum-modulus principle it must vanish also in the interior, so that $u = v$ on the whole annulus. The function defined on S as equal to u on $U \setminus \{n\}$ and to v on V is then well-defined and harmonic on the whole of $S \setminus \{n\}$. It is therefore the real part of a meromorphic function, and with a single pole at n of order 1 since $u_0 \circ \varphi_n^{-1} = \mathrm{Re}(\frac{1}{z})$. We have thus found the function sought.

Finally, such a function with prescribed images of any three distinct points will be unique since the only conformal bijection of the Riemann sphere to itself fixing three arbitrarily chosen points is the identity map. This completes the proof of the theorem. □

The above proof is a good illustration, though in a rather simple situation, of Schwarz's alternating method. Recall that that method is interesting chiefly for its use in proving Theorem IV.1.6, which we have for the time being assumed. Thus we shall be returning to the method in Chapter IX.

Intermezzo

Chapter V

The Klein quartic

The theory of elliptic integrals, so intensively developed over the course of the 19th century, gave rise to new functions. With each elliptic integral there is associated the *marked lattice* of its periods, that is, a given discrete subgroup of rank 2 of the additive group $(\mathbb{C},+)$ and a basis $(\omega_1,\omega_2) \in \mathbb{C}^* \times \mathbb{C}^*$ for this lattice satisfying $\mathrm{Im}(\omega_1/\omega_2) > 0$. It is therefore natural to introduce the set

$$\mathcal{M} = \{(\omega_1,\omega_2) \in \mathbb{C}^* \times \mathbb{C}^* | \omega_1/\omega_2 \in \mathbb{H}\}$$

of marked lattices. Observe that $\mathcal{M}$ is invariant under the natural action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{C}^2$ (and that $\mathrm{SL}(2,\mathbb{Z})\backslash\mathcal{M}$ may be identified with the set of lattices of $\mathbb{C}$ — see [Ser1970]). This action induces an action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{H}$ by Möbius transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

We recall in Section V.1 below the proof that there exists a function

$$j : \mathbb{H} \to \mathbb{C},$$

invariant under the action of $\mathrm{SL}(2,\mathbb{Z})$, such that two lattices Λ_1 and Λ_2 of $\mathbb{C}$ are homothetic[1] if and only if $j([\Lambda_1]) = j([\Lambda_2])$, where $[\Lambda_i]$ denotes the homothety class of the lattice Λ_i. Following on foundational work, notably of Gauss, Legendre, Abel, and Jacobi, a basic problem became that of linking $j(\tau)$ and $j'(\tau) = j(N\tau)$ for $\tau \in \mathbb{H}$ and N an integer ≥ 2. It can be shown[2] (see Subsection V.1.3 below) that there exists a polynomial $\Phi_N \in \mathbb{C}[X,Y]$ such that

$$\Phi_N(j', j) = 0. \tag{V.1}$$

[1]Or *similar*, meaning that there exists a $\lambda \in \mathbb{C}^*$ such that $\lambda\Lambda_1 = \Lambda_2$, which is equivalent to the statement that $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ are isomorphic.

[2]Recall, by way of analogy, that the trigonometric functions $\cos(x)$ and $\cos(Nx)$ are linked by an algebraic equation

$$\cos(Nx) = T_N(\cos x),$$

where T_N is the Nth Chebyshev polynomial.

When Φ_N is minimal, this is called the *modular equation*[3] associated with transformations of order N. The modular equation associated with transformations of order 7 (that is, the case $N = 7$) is the main subject of the article [Kle1878c]. There Klein produces a remarkable geometric model of the surface $X(7)$ obtained by compactifying the quotient of $\mathbb{H}$ by the action of the subgroup

$$\Gamma(7) = \{\alpha \in \mathrm{SL}(2,\mathbb{Z}) | \alpha \equiv I_2 \pmod 7\},$$

where I_2 denotes the identity matrix. More generally, one defines the *principal congruence subgroup of level N* by

$$\Gamma(N) = \{\alpha \in \mathrm{SL}(2,\mathbb{Z}) | \alpha \equiv I_2 \pmod N\}.$$

We show in Section V.1 below that the quotient $\Gamma(N)\backslash\mathbb{H}$ can be compactified to form a Riemann surface, denoted by $X(N)$. The group $\Gamma(1)$ coincides, of course, with $\mathrm{SL}(2,\mathbb{Z})$ and we shall see that the surface $X(1)$ is isomorphic to the Riemann sphere $\mathbb{CP}^1$.

Klein shows that the surface $X(7)$ is isomorphic to the smooth plane quartic C_4 with equation[4] $x^3y+y^3z+z^3x = 0$, invariant under the action of a group G isomorphic to $\mathrm{PSL}(2,\mathbb{F}_7)$ (the automorphism group of $X(7)$ — see Proposition V.1.1 below). In this projective model the natural morphism from $X(7)$ onto $X(1) \simeq \mathbb{CP}^1$ is made concrete as the projection of C_4 on $G\backslash C_4$ (identified with $\mathbb{CP}^1$); this is a Galois covering whose generic fibre is considered by Klein as "the Galois resolvent"[5] of the modular equation of level 7. Relying on numerous geometric properties of his quartic and on his investigations of the equation $\Phi_7(\cdot, j) = 0$, Klein arrived at a description of the fibre over a given value $j(\tau)$ in terms of quotients of explicit modular forms defined on the half-plane $\mathbb{H}$, the domain of the variable τ (the ratio of periods). This result was the most significant one of the article [Kle1878c]. In addition to that he presents another novelty: the explicit parametrization of a curve of genus > 1 (to within a finite number of points) by means of a uniform complex variable. He proves the following theorem:

Theorem V.0.6. — *The Riemann surface associated with the plane quartic C_4 defined by the equation*

$$x^3y + y^3z + z^3x = 0$$

[3] An earlier version, Jacobi's modular equation, concerns the modulus $\lambda = k^2$.

[4] In terms of homogeneous coordinates for 2-dimensional complex projective space. *Trans*

[5] Meaning that the function field of C_4 is the splitting field of this equation over $\mathbb{C}(j)$.

with its 24 points of inflection removed, is uniformized by the variable $\tau \in \mathbb{H}$ via the formulae

$$\frac{x}{z} = q^{-1/7} \frac{\sum_{m\in\mathbb{Z}}(-1)^m[q^{\frac{1}{2}(21m^2+37m+16)} - q^{\frac{1}{2}(21m^2+19m+4)}]}{\sum_{n\in\mathbb{Z}}(-1)^{n+1}q^{\frac{1}{2}(21n^2+7n)}}, \tag{V.2}$$

$$\frac{y}{x} = q^{-4/7} \frac{\sum_{m\in\mathbb{Z}}(-1)^m[q^{\frac{1}{2}(21m^2+25m+8)} - q^{\frac{1}{2}(21m^2+31m+12)}]}{\sum_{n\in\mathbb{Z}}(-1)^{n+1}q^{\frac{1}{2}(21n^2+7n)}}, \tag{V.3}$$

$$\frac{z}{y} = q^{-2/7} \frac{\sum_{m\in\mathbb{Z}}(-1)^m[q^{\frac{1}{2}(21m^2+m)} + q^{\frac{1}{2}(21m^2+13m+2)}]}{\sum_{n\in\mathbb{Z}}(-1)^{n+1}q^{\frac{1}{2}(21n^2+7n)}}, \tag{V.4}$$

where $q = e^{2i\pi\tau}$.

In other words, the above formulae give a concrete universal covering map $\mathbb{H} \to C_4 \setminus \mathcal{I}_4$, where $\mathcal{I}_4$ is the set of points of inflection of C_4. Over the two year period 1878–1879, Klein published a series of papers on modular equations, in particular [Kle1878b, Kle1878c, Kle1879b], devoted respectively to transformations of order $p = 5$, 7 and 11. In each case he constructs by geometric means a Galois resolvent, gives its roots explicitly — using modular forms — and shows how to find the modular equation itself (of degree $p + 1$) as well as a resolvent of degree p for each of these particular values of p. For $p = 5$, the geometric model of $X(5)$ he uses is the regular icosahedron[6], the resolvent of degree 5 being linked, as had been shown by Hermite [Her1858], to the general quintic equation. Just as the sphere has a regular tiling induced from the faces of an inscribed icosahedron, so also does the modular surface $X(7)$ admit a regular tiling by triangles. This tiling is inherited combinatorially from a tiling of $\mathbb{H}$ of type $(2,3,\infty)$[7], and its triangles are of type $(2,3,7)$. This is described in [Kle1878c] (or see pp. 125–127 of his complete works) and depends on elementary geometric properties of C_4. Arithmetic, algebraic, geometric, and combinatorial facets are tightly imbricated in this work of Klein, revealing the quartic C_4 to be a central and fascinating mathematical object. The reader may also consult [Levy1999a], and especially

[6]Klein shows that the morphism $X(5) \to X(1)$ is isomorphic to that taking the quotient of the unit sphere in $\mathbb{R}^3$ by the action of the symmetry group of the regular icosahedron.

[7]A tiling of $\mathbb{H}$ by triangles is said to be of type (a,b,c) if it is realized by hyperbolic triangles (a,b,c), that is, with angles $(\frac{2\pi}{a}, \frac{2\pi}{b}, \frac{2\pi}{c})$.

[Elk1999] who places C_4 in the context of modern number theory. We ourselves, following [Kle1878c], will be concentrating here on a particular result, namely the parametrization of the Klein quartic, representing an important stage in the explicit uniformization of algebraic curves. In particular, the above formulae of Theorem V.0.6 will be derived in the final Section V.2.5.

V.1. V.1. Modular forms, the invariant map j

V.1.1. Modular surfaces

It has been known since the time of Gauss (see Box V.1 and Figure V.1) that the set

$$\mathcal{D}(1) = \{\tau \in \mathbb{H} | |\tau| \geq 1, \ |\mathrm{Re}(\tau)| \leq 1/2\}$$

is a fundamental region of the action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{H}$, meaning that every orbit of the action of $\mathrm{SL}(2,\mathbb{Z})$ meets $\mathcal{D}(1)$ and that the translates of $\mathrm{int}(\mathcal{D}(1))$ by $\mathrm{SL}(2,\mathbb{Z})$ are pairwise disjoint. We shall be returning to this topic in Chapter VI.

Proposition V.1.1. — *Let* Γ *be a subgroup of finite index of* $\mathrm{SL}(2,\mathbb{Z})$*. The quotient*

$$Y_\Gamma = \Gamma \backslash \mathbb{H}$$

admits the structure of a noncompact Riemann surface, biholomorphic to a compact Riemann surface X_Γ *with a finite number of points removed.*

Box V.1: Gauss's reduction theory

The theory of the reduction of quadratic forms consists in the study of the orbits of the group $\mathrm{SL}(n,\mathbb{Z})$ acting on the vector space of quadratic forms in n variables according to the rule[a] $(A \cdot q)(x) = q(A^T x)$. This action is natural from the point of view of number theory: two quadratic forms in the same orbit have the same integer values. It is of interest, in particular, to look for a fundamental region of the action of $\mathrm{SL}(n,\mathbb{Z})$ on the set X of positive definite quadratic forms.

In his *Disquisitiones Arithmeticae* Gauss considers the case $n = 2$. Every such positive definite quadratic form q can be factored uniquely as

$$q(x, y) = a(\tau x + y)(\overline{\tau} x + y) \tag{V.5}$$

where $a > 0$ and τ lies in the half-plane $\mathbb{H}$.

[a] Here x stands for the column vector with components $x_1, \ldots, x_n$.

The action of positive scalars on X commutes with that of $\mathrm{SL}(2,\mathbb{Z})$, and it follows from (V.5) that the quotient $X/\mathbb{R}_+^*$ is isomorphic to the half-plane $\mathbb{H}$. The group $\mathrm{SL}(2,\mathbb{Z})$ acts via Möbius transformations on $\mathbb{H}$. Gauss proves the following celebrated result concerning a fundamental region of this action (see [Ser1970]):

Theorem V.1.2 (Gauss). — *The subset*

$$\mathcal{D}(1) = \{\tau \in \mathbb{H} \mid |\tau| \geq 1 \text{ and } |\mathrm{Re}(\tau)| \leq 1/2\}$$

of $\mathbb{H}$ *is a fundamental region for the action of* $\mathrm{SL}(2,\mathbb{Z})$ *on* $\mathbb{H}$.

Proof. — We begin with the case $\Gamma = \mathrm{SL}(2,\mathbb{Z})$. It is easy, starting from the fundamental region $\mathcal{D}(1)$, to equip the quotient $Y(1) = Y_\Gamma$ with the structure of a non-compact Riemann surface of genus 0 and with one end; this construction is carried out in the most general case of an arbitrary Fuchsian group in Chapter VI. The *horoballs*[8] centered at infinity, given by

$$B_a = \{\tau \in \mathbb{H} \mid \mathrm{Im}\tau \geq a\} \ (a > 0)$$

become in the passage to the quotient punctured discs forming neighborhoods of the ends of Y_Γ. Setting $q = e^{2i\pi\tau}$, understood as defining a chart, one obtains thereby a compact surface X_Γ representing a completion of the open Riemann surface Y_Γ.

In the general case of a subgroup $\Gamma \subset \mathrm{SL}(2,\mathbb{Z})$ of finite index, the quotient Y_Γ is a branched covering of $Y(1)$. It compactifies uniquely to a branched covering X_Γ of $X(1)$. The projective action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{QP}^1$ is transitive; the set $X_\Gamma \setminus Y_\Gamma$ is finite, in one-to-one correspondence with the classes $\Gamma\backslash\mathbb{QP}^1$, whose elements are still called *cusps* of X_Γ (or of Γ). Let $x = \rho(\infty)$ be a representative of a cusp ($\rho \in \mathrm{SL}(2,\mathbb{Z})$) and denote by Γ_x the stabiliser of x in Γ. The group $\rho^{-1}\Gamma_x\rho$ is, independently of the choice of representative of the cusp and of ρ, generated by $\gamma(z) = z + m$ for some integer $m \geq 1$. Setting $q = e^{2i\pi\tau}$, one takes as chart $\rho(w)$ with

$$w = e^{2i\pi\tau/m} = q^{1/m}. \qquad \text{(V.6)}$$

Note finally that each inclusion $\Gamma_1 \subset \Gamma_2$ of finite-index subgroups of $\mathrm{SL}(2,\mathbb{Z})$ induces a holomorphic map from X_{Γ_1} onto X_{Γ_2}. □

[8] In hyperbolic n-space a *horoball* is the limit of an increasing sequence of balls sharing a tangent hyperplane and its point of tangency. *Trans*

For $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ one has $Y_\Gamma \simeq \mathbb{C}$ and $X_\Gamma \simeq \overline{\mathbb{C}} \simeq \mathbb{CP}^1$ (the case of a single cusp). When Γ is one of the principal congruence subgroups $\Gamma(N)$, the numerical invariants (genus, the number of cusps) of the associated Riemann surface are known (see [Shi1971, pp. 20–23]). In particular, the surface $X(7)$ has genus 3 and 24 cusps (see Section V.2.1).

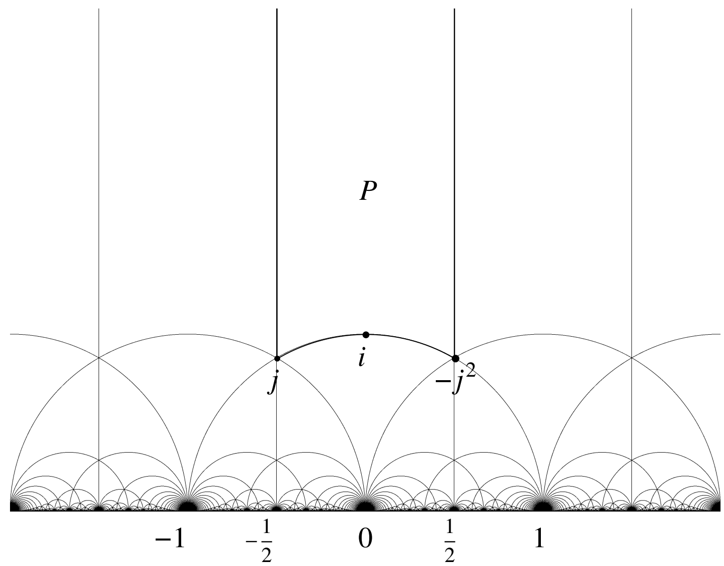


Figure V.1: A tiling for $\mathrm{PSL}(2,\mathbb{Z})$

V.1.2. Modular forms

For greater detail concerning the contents of this section, one may consult [Ser1970]. As above we consider a subgroup Γ of finite index of $\mathrm{SL}(2,\mathbb{Z})$. Recall that the set $\mathcal{M}$ of marked lattices is stable under the natural action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{C}^2$.

An *automorphic form of weight* k on $\mathbb{H}$ relative to Γ is defined to be any function $f : \mathbb{H} \to \mathbb{C}$ for which

$$f(\tau) = \hat{f}(\tau, 1)$$

where $\hat{f} : \mathcal{M} \to \mathbb{C}$ is a homogeneous function of degree $-2k$, invariant under Γ and such that $\hat{f}(\tau, 1)$ is meromorphic on $\mathbb{H}$ and also at the cusps coordinatized by the variable w defined by (V.6). In particular, the function f satisfies

$$f\left(\frac{a\tau+b}{c\tau+d}\right)(c\tau+d)^{-2k} = f(\tau) \qquad \left(\tau \in \mathbb{H}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma\right). \tag{V.7}$$

Among the automorphic forms it is useful to distinguish certain subsets. First, we denote by $K(\Gamma)$ the set of such forms of weight $k = 0$, which may be identified with the field of meromorphic functions on X_Γ. Next there is the set $M_k(\Gamma)$ of forms of weight k holomorphic on $\mathbb{H}$ and holomorphic in the variable w at each cusp of Γ: the *modular forms*. Under multiplication, the direct sum $M(\Gamma) = \bigoplus_{k\in\mathbb{Z}} M_k(\Gamma)$ is a graded $\mathbb{C}$-algebra.

Consider now the case $\Gamma(1) = \mathrm{SL}(2,\mathbb{Z})$, and assume $k > 2$. For each $(\omega_1,\omega_2) \in \mathcal{M}$ we write

$$G_k(\omega_1,\omega_2) = \sum_{\lambda\in\Lambda}{}' \frac{1}{\lambda^{2k}}, \tag{V.8}$$

where Σ' designates summation over the non-zero vectors of the lattice $\Lambda = \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2$, convergence being assured by the assumption $k > 2$. By construction, $G_k(\omega_1,\omega_2)$ is homogeneous of degree $-2k$ and $\mathrm{SL}(2,\mathbb{Z})$-invariant; an argument using normal convergence — in a fundamental region of $\mathrm{SL}(2,\mathbb{Z})$ — shows that $G_k(\tau,1)$ is holomorphic on $\mathbb{H}$ and also at the point ∞ (see [Ser1970, Chapter VII]). It is also known that the algebra of modular forms for $\mathrm{SL}(2,\mathbb{Z})$ is a polynomial algebra, generated by $g_2 = 60G_2$ and $g_3 = 140G_3$ of respective weights 2 and 3: $M(\mathrm{SL}(2,\mathbb{Z})) = \mathbb{C}[g_2, g_3] \simeq \mathbb{C}[X,Y]$.

In order to construct a meromorphic function on $\mathbb{H}$ that is $\mathrm{SL}(2,\mathbb{Z})$-invariant and non-constant, one considers the first homogeneous summand of $M(\mathrm{SL}(2,\mathbb{Z}))$ of dimension at least 2, with a view to forming the quotient of two linearly independent modular forms of the same weight. One shows (see [Ser1970, Chapter VII]) that this first summand is in fact $M_6(\mathrm{SL}(2,\mathbb{Z}))$, which contains the form $\Delta = g_2^3 - 27g_3^2$, not vanishing on $\mathbb{H}$. Hence it is natural to define

$$J = g_2^3/\Delta, \quad \text{and} \quad j = (12)^3 J. \tag{V.9}$$

The function j, called a *modular invariant*, is holomorphic on $\mathbb{H}$ with a simple pole at infinity of residue 1. Via passage to the quotient, it induces an isomorphism between $X(1)$ and $\mathbb{C}P^1$.

By considerations of symmetry, one obtains $g_3(i) = 0$ and $g_2(\rho) = 0$ for $\rho = (1 + i\sqrt{3})/2$ (see equation (V.8)), whence the particular values

$$j(i) = 12^3 = 1728 \quad \text{and} \quad j(\rho) = 0. \tag{V.10}$$

Finally, the field of meromorphic $\mathrm{SL}(2,\mathbb{Z})$-invariant functions coincides with $\mathbb{C}(j)$, which is isomorphic to the field of rational functions over $\mathbb{C}$ in a single variable. Hence for every finite-index subgroup Γ of $\mathrm{SL}(2,\mathbb{Z})$, the field of functions $K(\Gamma)$ is a finite extension of $\mathbb{C}(j)$, Galois if and only if Γ is a normal subgroup of $\Gamma(1)$, which is the case for the principal congruence subgroup $\Gamma(N)$, whose degree is equal to that of the branched covering $X_\Gamma \to X(1)$ [Rey1989, p. 60].

V.1.3. Modular equations

Given an integer $N \geq 2$, we seek an equation linking $j(\tau)$ and $j'(\tau) = j(N\tau)$ for $\tau \in \mathbb{H}$. It is easy to check that j' is left invariant by the group

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) | c \equiv 0 \ (\mathrm{mod}\ N) \right\}, \tag{V.11}$$

which is, in fact, precisely the stabiliser of j'.

On the other hand j' is meromorphic at the cusps of $\Gamma_0(N)$. Indeed, by means of the action of $\Gamma(1) = \mathrm{SL}(2,\mathbb{Z})$, one reduces the situation to the cusp ∞ and to a function of the form $j \circ \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix} \right)$ with a, b and d integers; for sufficiently large k the product of the latter with $q^{k/m}$ (see equation (V.6)) is bounded in a neighborhood of $q^{1/m} = 0$. The extension $K(\Gamma_0(N))/\mathbb{C}(j)$ being finite, this implies the existence of an algebraic relation between j and j'. In order to exhibit such a relation, one considers the transforms of j' by the elements of $\Gamma(1)$, that is, the $j \circ \alpha$ with α ranging over the orbit O_N of the point

$$p_N = \Gamma(1) \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma(1)\backslash\Delta_N,$$

under the action of $\Gamma(1)$ on the right; here Δ_N denotes the set of integer matrices of determinant N. One readily checks that the stabiliser of the point p_N in $\Gamma(1)$ is $\Gamma_0(N)$, so that the orbit O_N may be identified with the quotient $\Gamma_0(N)\backslash\Gamma(1)$. Write d_N for the index of $\Gamma_0(N)$ in $\Gamma(1)$ and $\alpha_k \in \Delta_N$ ($k = 1, \ldots d_N$) for a system of representatives of the orbit O_N. Then the coefficients of the polynomial $\prod_{k=1}^{d_N} (X - j \circ \alpha_k)$ are invariant under $\Gamma(1)$, holomorphic on $\mathbb{H}$ and (by the same argument as above) meromorphic at the cusp ∞. We have thus found a polynomial $\Phi_N \in \mathbb{C}[X,Y]$ of degree d_N in X such that

$$\Phi_N(j', j) = 0. \tag{V.12}$$

This is the *modular equation* associated with transformations of order N. The stabiliser of $j \circ \alpha_k$ is conjugate to $\Gamma_0(N)$ (the stabiliser of j'), whence the subgroup fixing all the $j \circ \alpha_k$ coincides with $\Gamma(N) = \bigcap_{\gamma \in \Gamma(1)} \gamma \Gamma_0(N) \gamma^{-1}$. It follows that the splitting field of $\Phi_N \in \mathbb{C}[j][X]$ is $K(\Gamma(N))$. Moreover $\Gamma(1)$ acts as a set of automorphisms of $K(\Gamma(N))$ in permuting transitively the roots of this polynomial, which is therefore irreducible, whence, in particular, $K(\Gamma_0(N)) = \mathbb{C}(j, j')$ (see also [Shi1971, p. 34]). When $N = p$, a prime, one readily sees that the matrices $\left(\begin{smallmatrix} 1 & k \\ 0 & p \end{smallmatrix}\right)$ $(0 \le k < p)$ and $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ form a system of representatives of $O_p = \Gamma(1)\backslash\Delta_p$; the index of $\Gamma_0(p)$ is thus $d_p = p + 1$.

An elementary calculation shows that $\Gamma_0(N)$ is normalized by the matrix

$$\begin{pmatrix} 0 & N^{-1/2} \\ -N^{1/2} & 0 \end{pmatrix} \tag{V.13}$$

which induces an involutary automorphism of the surface $X_0(N) = X_{\Gamma_0(N)}$ and of its function field: this is the *Fricke involution* interchanging j and j'. One infers from its existence that $\Phi_N \in \mathbb{C}[X, Y]$ is symmetric. Klein relies on this symmetry in his investigation of the modular equation for $N = 2, 3, 4, 5, 7$ and 13 [Kle1878b, §II]. For these values of N the surface $X_0(N)$ is of genus 0 and there exists $\xi \in K(\Gamma_0(N))$ such that $K(\Gamma_0(N)) = \mathbb{C}(j, j') = \mathbb{C}(\xi)$; one then has $j = F(\xi)$ and $j' = F(\xi')$ with $F \in \mathbb{C}(Z)$, the function ξ' being linked to ξ by a Möbius transformation (the Fricke involution). In each of these cases Klein describes a fundamental region for the action of $\Gamma_0(N)$ on the half-plane[9], then deduces from ramification data an expression for F and gives the relation between ξ (in its alternative guise as a function of q) and ξ'. Note that for $N \in \{2, 3, 4, 5\}$, the surface $X(N)$ is also of genus 0, with respective automorphism groups (leaving the set of cusps globally fixed) the dihedral group, the tetrahedral group A_4, the group S_4 of the cube and the octahedron, and the group A_5 of the dodecahedron and icosahedron.

V.1.4. The surface $X_0(7)$

We shall now expound in detail the case $N = 7$. Our first task is to determine a fundamental region for the action of the group $\Gamma_0(7)$. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ from $\mathrm{SL}(2, \mathbb{R})$, there is the following well-known formula:

$$\mathrm{Im}\gamma(z) = \frac{\mathrm{Im} z}{|cz + d|^2} \quad (z \in \mathbb{H}). \tag{V.14}$$

Since for each fixed z there are only finitely many pairs $(c, d) \in \mathbb{Z}^2$ for which the modulus $|cz + d|$ is less than a given number, it follows that each orbit of the action of $\Gamma_0(7)$ on $\mathbb{H}$ contains a point $z \in \mathbb{H}$ whose imaginary part is greatest, that

[9] He uses the $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2, \mathbb{Z})$ with $b \equiv 0 \pmod N$, which comes to the same thing.

is, such that $|cz + d| \geq 1$ for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(7)$. Hence every orbit of $\Gamma_0(7)$ meets the set

$$\mathcal{D}' = \bigcap_{d \notin 7\mathbb{Z}} \{|z + d/7| \geq 1/7\} \cap \{|\mathrm{Re} z| \leq 1/2\}.$$

Here the inequality $|\mathrm{Re} z| \leq 1/2$ is a consequence of the fact that the translation

$$t : \tau \mapsto \tau + 1$$

is an element of $\Gamma_0(7)$. By applying the rotations $r_1 = \left(\begin{smallmatrix} 2 & -1 \\ 7 & -3 \end{smallmatrix}\right)$ and $r_2 = \left(\begin{smallmatrix} -2 & -1 \\ 7 & 3 \end{smallmatrix}\right)$, the set $\mathcal{D}'$ is transformed into the fundamental region $\mathcal{D}$ of Figure V.2.

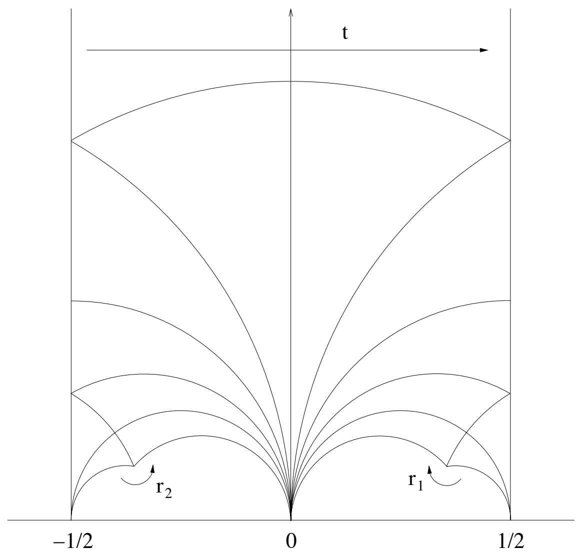


Figure V.2: A fundamental region for $\Gamma_0(7)$

There we see a tiling by eight translates of $\mathcal{D}(1)$ by $\Gamma(1)$, and since the image $\overline{\Gamma}_0(7)$ of $\Gamma_0(7)$ in $\mathrm{PSL}(2,\mathbb{Z})$ is a subgroup of index 8, we conclude that $\mathcal{D}$ and $\mathcal{D}'$ are fundamental regions for $\Gamma_0(7)$. Finally, we see in Figure V.2 that $X_0(7)$ is of genus 0 with two cusps (0 and ∞) and that $\overline{\Gamma}_0(7)$ is generated by r_1, r_2 and the translation t realizing the identifications used to obtain $X_0(7)$.

Proposition V.1.3. — *The expression*

$$\xi = \left(\frac{\Delta(\tau)}{\Delta(7\tau)}\right)^{1/6} = \frac{1}{q} \prod_{n=1}^{\infty} \left(\frac{1 - q^n}{1 - q^{7n}}\right)^4 \tag{V.15}$$

affords a rational coordinate on $X_0(7)$.

Proof. — The second equality is a consequence of the following one (see [Ser1970]):

$$\Delta = (2\pi)^{12}\, q \prod_{n=1}^{\infty} (1-q^n)^{24} \ .$$

One next verifies that, as for $j(7\tau)$, $\Delta(7\tau)$ is modular for $\Gamma_0(7)$, whence $\xi^6 \in M(\Gamma_0(7))$. It then follows that $\xi \circ \overline{\gamma} = \chi(\overline{\gamma})\xi$ for all $\overline{\gamma} \in \overline{\Gamma}_0(7)$, where χ is a character of the group $\overline{\Gamma}_0(7)$ with values sixth roots of unity. Recall that $\overline{\Gamma}_0(7)$ is generated by t, r_1 and r_2. One has $\xi \circ t = \xi$ (since ξ is expressible uniquely in terms of integer powers of q) and $\chi(r_1) = \chi(r_2) = 1$ because these rotations have a fixed point in $\mathbb{H}$. Hence the character χ is trivial. Finally, since Δ is holomorphic and does not vanish on $\mathbb{H}$ (see Section V.1.2), ξ can take the values 0 and ∞ only at cusps. In view of the fact that $X_0(7)$ has only two cusps, the function ξ is thus necessarily of degree 1. Hence, in particular, the subgroup of $\Gamma(1)$ leaving ξ invariant is $\Gamma_0(7)$. □

V.1.5. The modular invariant as a function on $X_0(7)$

We can now determine j as a function of ξ following the method employed by Klein in [Kle1878b II §14]. The result is as follows:

Proposition V.1.4. — *We have*

$$j = \frac{1}{\xi^7}(\xi^2 + 13\xi + 49)(\xi^2 + 245\xi + 2401)^3.$$

Proof. — Write $j = \phi(\xi)/\psi(\xi)$, a rational function of degree 8 in ξ. The equation $j = \infty$ has a simple root $\xi = \infty$ — corresponding to $q = 0$ — and a root of multiplicity 7 at $\xi = 0$ (see equation (V.15) and Figure V.2); we may therefore take $\psi(\xi) = \xi^7$. Similarly, ϕ has two triple roots and two simple ones, and $\phi - 1728\psi$ has four double roots (equation (V.10) and Figure V.2). Furthermore ψ must be monic since $j(q)$ has a simple pole of residue 1 at $q = 0$. These conditions serve to determine ϕ uniquely. Indeed, we must have $\phi = UV^3$ and $\phi - 1728\psi = W^2$ where U, V and W are monic polynomials of degrees 2, 2 and 4 respectively; and moreover U, V, W and ξ are pairwise relatively prime. The "functional determinant" $\phi'\psi - \psi'\phi$ is therefore a monic polynomial of degree 14 divisible by both ξ^6V^2 and W, whence $\phi'\psi - \psi'\phi = \xi^6V^2W$. This relation gives W in terms of the coefficients of U and V; carrying this over to $UV^3 - 1728\xi^7 = W^2$ then yields the expression claimed for j. □

The Fricke involution sends ξ to $\xi'(\tau) = \xi(-1/(7\tau))$. It induces an involutary automorphism σ of the surface $X_0(7)$, so that ξ' is an involutary homographic function of ξ. One has $j = F(\xi)$ and $j' = F(\xi')$. The two cusps of $X_0(7)$ make up the fibre common to j and j' above infinity, and are interchanged by σ, whence $\xi\xi' = C$ for some constant C. In similar fashion σ interchanges the simple roots of $F = 0$ (the images of the centres of rotation of r_1 and r_2 — see Figure V.2) linked by the equation $z_1 z_2 = -1/7$), at which $j = j' = 0$; hence $\xi^2 + 13\xi + 49 = 0$ implies that also $\xi'^2 + 13\xi' + 49 = 0$, whence $C = 49$. Finally, therefore, the Fricke involution is given by

$$\xi\xi' = 49. \tag{V.16}$$

It then follows from Proposition V.1.4 that $j = (\xi'^2+13\xi'+49)(\xi'^2+5\xi'+1)^3/\xi'$. This is actually the expression obtained by Klein in [Kle1878b] for $J = j/1728$.

The description of the fibre of $j : X_0(7) \to \mathbb{C}P^1$ will be important in the sequel as an essential intermediate step towards the parametrization of C_4. From the above expression for J as a function of ξ', we infer the following expression for J' as a function of ξ:

$$J' - 1 = \frac{1}{12^3\xi}(\xi^4 + 14\xi^3 + 63\xi^2 + 70\xi - 7)^2. \tag{V.17}$$

At the points $-1/7\tau$ and $(\tau + k)/7$ ($k = 0, \ldots, 6$), the function J' takes the same value $J(\tau)$. Hence $\xi_\infty = \xi' = \xi(-1/7\tau)$ and $\xi_k = \xi(\tau/7 + k/7)$, $k = 0, \ldots, 6$, are roots of the equation

$$(z^4 + 14z^3 + 63z^2 + 70z - 7)^2 - 12^3(J - 1)z = 0. \tag{V.18}$$

Since the functions $-1/7\tau$ and $(\tau + k)/7$ ($k = 0, \ldots, 6$) are distinct modulo $\Gamma_0(7)$, the same holds for the ξ_k ($k = \infty, 0, \ldots, 6$) as functions on $\mathbb{H}$. Changing the point τ in the fibre above J, induces a permutation of the ξ_k ($k = \infty, 0, \ldots, 6$) as roots of the equation (V.18) — the permutation can be made explicit using (V.15). We now set $q^{1/2} = e^{i\pi\tau}$, $\Delta^{1/2} = (2\pi)^6 q^{1/2} \prod_{n=1}^{\infty}(1 - q^n)^{12}$. Since $J - 1 = 27g_3^2/\Delta$ (see equation (V.18)), the square roots of the solutions of (V.18) are up to sign solutions of

$$w^8 + 14w^6 + 63w^4 + 70w^2 - 6^3 g_3 w/\Delta^{1/2} - 7 = 0. \tag{V.19}$$

They can be expressed in terms of $\pm\xi^{1/2}$, $\xi^{1/2} = q^{-1/2} \prod_{n=1}^{\infty}(1 - q^n)^2(1 - q^{7n})^{-2}$. The sign is determined by the behaviour of the leading term of (V.19) as q tends to 0. Using $\lim_{q\to 0}(g_3) = 280\zeta(6)$ [Ser1970, Chapter VII, §2.3], one finds that $6^3 g_2/\Delta^{1/2} = q^{-1/2}(1 + o(q))$. Hence the roots of equation (V.19) are

$w_\infty = -\xi^{1/2}(-1/7\tau)$ and $w_k = (-1)^k \xi^{1/2}(\tau/7 + k/7)$ ($k = 0,\ldots,6$), or, setting $\gamma = e^{2i\pi/7}$,

$$w_\infty = -7q^{1/2}\prod_{n=1}^{\infty}\left(\frac{1-q^{7n}}{1-q^n}\right)^2, \tag{V.20}$$

$$w_k = \gamma^{-4k}q^{-1/14}\prod_{n=1}^{\infty}\left(\frac{1-\gamma^{nk}q^{n/7}}{1-q^n}\right)^2 \qquad (k = 0,\ldots,6). \tag{V.21}$$

V.2. How Klein parametrized his quartic

V.2.1. The group $\mathrm{PSL}(2,\mathbb{F}_7)$ and the surface $X(7)$

As we have seen above (§V.1.3), the function field $K(\Gamma(7))$ of the surface $X(7)$ is the splitting field of the polynomial $\Phi_7 \in \mathbb{C}(j)[X]$ associated with transformations of order 7 (equation (V.1)). We first need to examine the action of "homographic substitutions modulo 7"[10] [Kle1878c, §§1–2] on $X(7)$. Let $\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$ be the field of seven elements. Since $\mathrm{SL}(2,\mathbb{F}_7)$ is generated by $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$, the reduction morphism modulo 7 from $\mathrm{SL}(2,\mathbb{Z})$ to $\mathrm{SL}(2,\mathbb{F}_7)$ is surjective, whence the exact sequence

$$1 \to \overline{\Gamma}(7) \to \mathrm{PSL}(2,\mathbb{Z}) \to \mathrm{PSL}(2,\mathbb{F}_7) \to 1. \tag{V.22}$$

In particular, the quotient $G = \mathrm{PSL}(2,\mathbb{Z})/\overline{\Gamma}(7)$ is isomorphic to $\mathrm{PSL}(2,\mathbb{F}_7)$, a simple group of order 168 (see Remark V.2.2 below). The group G acts on $X(7)$ via automorphisms and $G\backslash X(7)$ can be identified with $X(1)$. Thus the fibres of the projection $X(7) \to X(1)$ are the orbits of the action of G on $X(7)$. There are therefore three singular fibres corresponding to the values $J = \infty, 0$ and 1 (recall that $J = j/1728$), whose elements are called A-points, B-points and C-points in Klein's terminology, with stabilisers of orders respectively 7, 3 and 2. These fibres have cardinality 24, 56 and 84; all others have 168 elements. By the Riemann–Hurwitz formula, the genus g of $X(7)$ satisfies the relation

$$2 - 2g = 2\cdot 168 - 6\cdot 24 - 2\cdot 56 - 84, \tag{V.23}$$

whence $g = 3$.

[10] That is, Möbius transformations modulo 7. *Trans*

Remark V.2.1. — The automorphism group of a compact Riemann surface of genus $g \geq 2$ is finite of cardinality at most $84(g-1)$; this is the *Hurwitz bound*. Thus the surface $X(7)$ attains this bound.[11]

On lifting to $X(7)$ the decomposition of $X(1) \simeq \mathbb{C}P^1$ into two triangles with vertices $(1, 0, \infty)$, one obtains a polyhedral triangular structure of type $(2,3,7)$ on $X(7)$. The surface is tiled by 336 triangles which can be grouped to obtain either a tiling by 24 heptagons (centered at the A-points) with 84 edges centered at the C-points and 56 vertices (the B-points), or the dual tiling comprised of 56 triangles centered at the B-points, 84 edges centered at the C-points, and 24 vertices, the A-points (see [Kle1878c, 1921a] and Figure V.3.).

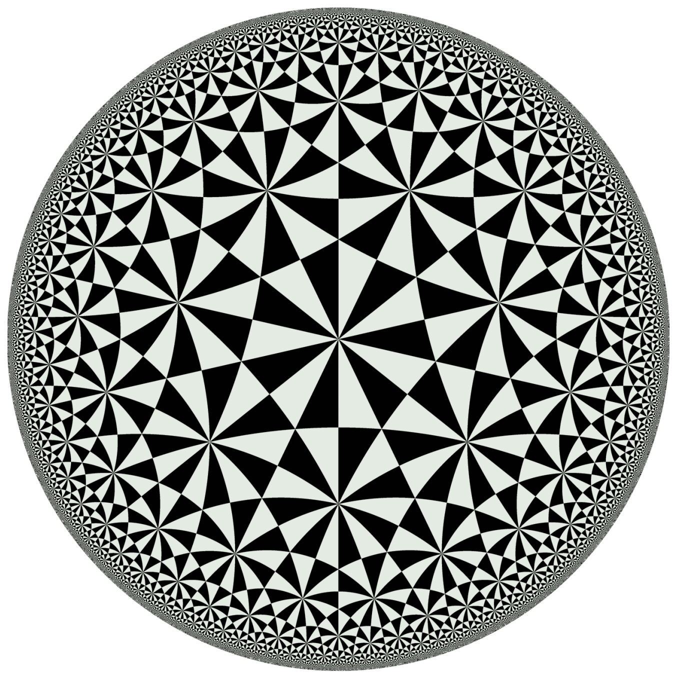


Figure V.3: Polyhedral triangular structure of $X(7)$

Each $g \in G$ lifts to a matrix $\gamma \in \mathrm{SL}(2, \mathbb{F}_7)$ uniquely defined up to sign. The order of g is therefore related to the trace of γ. For example, if g has

[11]This is a consequence of the fact that the triangle $(2,3,7)$ has least hyperbolic area among all hyperbolic triangles of type (a,b,c). By the Gauss–Bonnet theorem, this derives from the fact that the largest value less than 1 of the sum $\frac{1}{a} + \frac{1}{b} + \frac{1}{c}$ is attained uniquely by $(2,3,7)$.

order 4, the minimal polynomial of γ is one of the factors of the polynomial $x^4 + 1 = (x^2 - 3x + 1)(x^2 - 4x + 1)$ and one has $\mathrm{tr}\gamma = \pm 3$; conversely, this condition clearly implies that g has order 4. One thus obtains that the elements of orders 2, 3, 4 and 7 are characterized respectively by $\pm\mathrm{tr}\gamma = 0, 1, 3$ and 2 (with $\gamma \neq \pm I_2$). An easy count then shows that G has respectively 21, 56, 42 and 48 elements of these orders.

We now make an inventory of the cyclic subgroups of G. First one observes that the involutions in G form a single conjugacy class (their lifts to $\mathrm{SL}(2, \mathbb{F}_7)$ all have $x^2 + 1$ as minimal polynomial) so must have fixed points; in fact each of the 21 involutions must fix four C-points. These involutions are in one-to-one correspondence with the subgroups of order 4, also all conjugate. Each element $g \in G$ of order 4 is fixed-point free and acts by means of a bi-transposition on the 4 fixed points of the involution g^2. For topological reasons (a surface of genus 3 cannot be an unramified triple covering), the elements of order 3 all have fixed points. Thus they form 28 subgroups each fixing a pair of B-points, and all conjugate since G is transitive on the B-points. In similar fashion, one sees that the 48 elements of order 7 make up altogether 8 subgroups forming a single conjugacy class, and each fixing three A-points.

A few additional remarks will facilitate the determination of a geometric model of $X(7)$. Since the action of G by conjugation on the set of 8 subgroups of order 7 is transitive, the normalizer of each such subgroup has order 21, and is non-Abelian since G has no elements of order 21. Denote by G'_{21} (in the notation of [Kle1878c]) any one of these normalizers; its structure is necessarily that of the semi-direct product generated by two elements h and r satisfying (up to interchanging r and r^{-1}) the relations

$$h^7 = r^3 = 1 \quad \text{and} \quad rhr^{-1} = h^4. \tag{V.24}$$

The three A-points fixed by h, which have cyclic stabilizer, are permuted cyclically by r. Analogous reasoning shows that the centralizer of r is a subgroup of G'_6 isomorphic to the symmetric group S_3, generated by r and an involution s permuting the two B-points fixed by r.

Remark V.2.2. — Klein makes no mention of the simplicity of G, although this may be deduced using the elementary argument given by him in [Kle1884, p. 19] for $\mathrm{PSL}(2, \mathbb{F}_5)$. From the knowledge that all cyclic subgroups of the same order are conjugate — in fact all elements of the same order are conjugate — it follows that the cardinality of a normal subgroup H of G is of the form $1 + 21\alpha_1 + 56\alpha_2 + 42\alpha_3 + 48\alpha_4$ with $\alpha_k = 0$ or 1 ($k = 1, \ldots, 4$). The only possibilities yielding a divisor of 168 are then quickly seen to be $\alpha_k = 0$ for all k (H trivial) and $\alpha_k = 1$ for all k ($H = G$).

V.2.2. The quartic C_4

In this section we determine an explicit algebraic equation for the Riemann surface $X(7)$. Recall that, given a compact Riemann surface X and the dual $V = \Omega(X)^*$ of the space of holomorphic 1-forms on X, there is a natural embedding of X in the projective space $P(V)$. This map sends each point x of X to the projectification of the space of holomorphic 1-forms vanishing at x. One shows that this space is always a vector hyperplane of $\Omega(X)$, so may be identified with a point of $P(V)$. It is further known that the map ϕ from X to the projective space $P(V)$ so defined is a holomorphic embedding except for the case when X is hyperelliptic[12] (see [Rey1989, p. 102]).

In order to exclude the latter possibility in the case of $X(7)$, Klein uses a specific known model for plane hyperelliptic curves of genus 3 (see [Levy1999b, p. 295]). The fact that $X(7)$ is not hyperelliptic may also be inferred from the behaviour of its involutions[13]: they each have four fixed points (§V.2.1), whereas a hyperelliptic involution in genus 3 must have eight. When $g = 3$ the projective space $P(V)$ has dimension 2; hence the image C_4 of the embedding of $X(7)$ is a plane smooth quartic.

The group G of automorphisms of $X(7)$ acts linearly on the space $V = \Omega(X(7))^*$ as follows:

$$(g \cdot \xi)(\omega) = \xi(g^*\omega) \qquad (g \in G,\, \xi \in V,\, \omega \in \Omega(X)). \tag{V.25}$$

This projective action of G is essential to the geometric investigation of this quartic. Observe that the representation (V.25) takes its values in $\mathrm{SL}(V) \simeq \mathrm{SL}(3,\mathbb{C})$ since G has no nontrivial quotient (see Remark V.2.2).

Recall that every plane curve possesses special points, notably: inflections, points of contact with bitangents[14], and points where the curve admits a superosculating conic, that is, having contact with the curve of order at least 6. Following Cayley, these last are called *sectactic*; for example, if the curve has an axis of symmetry, its intersections with that axis will be sectactic by symmetry. For a smooth projective curve with equation $f = 0$ of degree at least 4, each of these three sets of special points are obtained as the intersection with another curve associated with f (for C_4 see equation (V.31) of §V.2.3 below), for example with $\det \mathrm{Hess}\, f = 0$ in the case of inflections, where $\mathrm{Hess}\, f$ is the Hessian of f. Smooth curves of degree 4 have 24 inflections, 56 contacts with bitangents, and 84 sectactics points (counting multiplicities).

[12] Recall that a Riemann surface X is said to be *hyperelliptic* if there exists a branched covering of degree 2 of $\mathbb{CP}^1$ by X. The unique nontrivial covering automorphism is then called a *hyperelliptic involution*.

[13] Or from the simplicity of G, since a hyperelliptic involution is always central.

[14] That is, lines tangent at two points of the curve. *Trans*

Henceforth we shall identify G with the subgroup of $\mathrm{SL}(V)$ leaving C_4 invariant under the projective action (V.25). Being projectively invariant, each family of special points of C_4 is the union of orbits of the action of G. This leads to the following single possibility: the inflections correspond to the A-points, the points of contact with bitangents to the B-points, and the sectactic points to the C-points (the other orbits each having 168 elements). Since the inflections are simple, each tangent line to such a point $P \in C_4$ meets C_4 in another point P'. As P ranges over the set $\mathcal{I}$ of inflection points, the corresponding points P' range over a orbit consisting of 24 elements, which therefore coincides with $\mathcal{I}$ itself by uniqueness. Furthermore, an element of G fixing P must also fix P'. Knowing that the stabilizers of A-points each fix a triplet of them, we obtain a decomposition of the set $\mathcal{I}$ into 8 cycles of length 3. Hence the tangents to inflection points subdivide into 8 *inflection triangles* (in Klein's terminology) which will have an important role to play subsequently. For instance, they allow us to prove the following:

Theorem V.2.3. — *An equation for C_4, invariant under G, is*

$$x^3y + y^3z + z^3x = 0. \tag{V.26}$$

Proof. — Let $f = 0$ be an equation for C_4, invariant under the action of G. The simplicity of the group G imposes an additional constraint. Indeed, since each character of G with values in $\mathbb{C}^*$ is trivial, every projectively G-invariant polynomial is G-invariant; in particular *f is G-invariant*. Since the linear projective group is 3-transitive on the plane, one can choose coordinates $[x, y, z]$ so that the axes form an inflection triangle of C_4 and the tangent to the point $[1, 0, 0]$ is $y = 0$. In terms of these coordinates the polynomial f is then of the form

$$f = ax^3y + by^3z + cz^3x + xyz(ux + vy + wz). \tag{V.27}$$

We also know from §V.2.1 that there exists an element of order 3 of G which permutes the three base points of the coordinate system cyclically. Such an element is necessarily conjugate by a diagonal matrix from $\mathrm{GL}(3, \mathbb{C})$ to the matrix r introduced above (see equation (V.28)). By applying a diagonal coordinate change (necessarily preserving the form of f), we may suppose that r acts as a cyclic permutation of x, y and z. The invariance of f under r then entails $a = b = c$ and $u = v = w$. On the other hand, each of the base points is fixed by a (diagonal) element $h \in G$ of order 7. The diagonal entries of h must have the form γ^k, γ^l and γ^m where k, l, m are integers and $\gamma = \exp(2i\pi/7)$. From the invariance of f under h it follows that $l = 4k$, $m = 2k$, whence $u = 0$, yielding the desired equation. □

One can in fact characterize the homogeneous coordinate systems relative to which C_4 has equation (V.27). They are those for which $xyz = 0$ defines an inflection triangle and x, y, z are permuted cyclically by an element of order 3 of G.

From now on f will denote the polynomial $x^3y + y^3z + z^3x$. One also infers from the above proof that the stabilizer of the inflection triangle $xyz = 0$ is the non-Abelian subgroup G'_{21} of order 21 of G (see §V.2.1) generated by the following two matrices:

$$r = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \gamma^4 & 0 \\ 0 & 0 & \gamma^2 \end{pmatrix}. \tag{V.28}$$

In order to complete the description of G, it remains to find an involution s normalizing r. By means of a change of coordinates, Klein obtains

$$s = \begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix} \quad \text{with} \quad \begin{cases} a = i(\gamma^2 - \gamma^5)/\sqrt{7} = -\frac{2}{\sqrt{7}}\sin(\frac{3\pi}{7}), \\ b = i(\gamma^4 - \gamma^3)/\sqrt{7} = \frac{2}{\sqrt{7}}\sin(\frac{\pi}{7}), \\ c = i(\gamma - \gamma^6)/\sqrt{7} \ = -\frac{2}{\sqrt{7}}\sin(\frac{2\pi}{7}). \end{cases} \tag{V.29}$$

Here is how one can reconstruct this result. First, it follows from the relation $srs^{-1} = r^{-1}$ that s has the form shown in the left-hand matrix in (V.29). Next, the relation $s^2 = 1$ yields $ab + bc + ca = 0$ and $a^2 + b^2 + c^2 = 1$. This implies that the conic C_2 with equation $xy + yz + zx = 0$ is also stabilized by s. The intersection of C_2 with C_4 is comprised of the points $[1, \alpha, \alpha^2]$, $[1, \alpha^2, \alpha]$ ($\alpha = e^{2i\pi/3}$), the 3 base points $[1, 0, 0]$, $[0, 1, 0]$, $[0, 0, 1]$ and the 3 points $[a, b, c]$, $[b, c, a]$, $[c, a, b]$ with a, b, c defined as in (V.29). Up to replacing s by rs or r^2s (which amounts to permuting a, b, c cyclically), one sees in this way that s is as in (V.29). Note also that one has $a + b + c = -1$ and $abc = 1/7$.

The normalizer G'_6 of r in G, generated by r and s, is isomorphic to the symmetric group S_3 (see §V.2.1). It acts on the intersection $C_4 \cap C_2$ as follows: the two inflection triangles are each permuted cyclically by r and interchanged by s; the points of tangency $B = [1, \alpha, \alpha^2]$ and $B' = [1, \alpha^2, \alpha]$ with the straight line $x + y + z = 0$ are fixed by r and interchanged by s, which therefore stabilizes this bitangent line. Since the polynomial $x^2 + y^2 + z^2$ is invariant under s, each conic of the pencil

$$u(xy + yz + zx) + v(x^2 + y^2 + z^2) = 0, \quad ([u, v] \in \mathbb{CP}^1) \tag{V.30}$$

is stable under the action of G'_6 (this is the pencil of conics bitangent at B and B'). A single conic of this pencil passes through each point $P \in C_4$, and its intersection with C_4 contains the orbit of P under the action of G'_6, generally consisting of 6 points.

By means of an *ad hoc* use of various coordinate systems, Klein locates a subgroup G''_{24} isomorphic to the symmetric group S_4, whose existence had been intimated in his investigation of G [Kle1878c §1]. This subgroup realizes every permutation of the 4 bitangents and stabilizes the conic C of the pencil (V.30) corresponding to $u/v = (-1 + i\sqrt{7})/2$ [*ibid.*, §§4–5]. The action of G on C defines a family of 7 conics using which Klein then infers a "resolvent of degree 7" [*ibid.*, §10].

V.2.3. Invariant polynomials

After having constructed a projective model of $X(7)$, Klein returns to the fundamental problem of describing the modular invariant in this context, that is, the function from C_4 to $\mathbb{CP}^1$ denoted by J whose fibres are the orbits of the group G. To this end, he determines all polynomials left invariant by G — in any case of use in the sequel — and then deduces the expression for J.

In order to find new invariant polynomials (other than $f = x^3y + y^3z + z^3x$), Klein uses his knowledge of "covariants". Denote by $S_d(\mathbb{C}^3)$ the subspace of $\mathbb{C}[x, y, z]$ of homogeneous polynomials of degree d. A *covariant* is a polynomial map $\Phi : S_d(\mathbb{C}^3) \to S_{d'}(\mathbb{C}^3)$ equivariant under the action of the special linear group; for example, the Hessian is a covariant with $d' = 3(d-2)$. If $P \in S_d(\mathbb{C}^3)$ is G-invariant, then so also is $\Phi(P)$. Klein introduces three G-invariant polynomials:

$$\nabla = \frac{1}{54}\begin{vmatrix} f''_{x^2} & f''_{xy} & f''_{xz} \\ f''_{yx} & f''_{y^2} & f''_{yz} \\ f''_{zx} & f''_{zy} & f''_{z^2} \end{vmatrix}, \qquad C = \frac{1}{9}\begin{vmatrix} f''_{x^2} & f''_{xy} & f''_{xz} & \nabla'_x \\ f''_{yx} & f''_{y^2} & f''_{yz} & \nabla'_y \\ f''_{zx} & f''_{zy} & f''_{z^2} & \nabla'_z \\ \nabla'_x & \nabla'_y & \nabla'_z & 0 \end{vmatrix}$$

$$\text{and } K = \frac{1}{14}\begin{vmatrix} f'_x & \nabla'_x & C'_x \\ f'_y & \nabla'_y & C'_y \\ f'_z & \nabla'_z & C'_z \end{vmatrix}, \tag{V.31}$$

of respective degrees 6, 14 and 21, with $\nabla = 5x^2y^2z^2 - (xy^5 + x^5z + z^5y)$. To verify the invariance, for any three polynomials P, Q, R from $\mathbb{C}[x, y, z]$, denote by $\nabla(P)$, $C(P, Q)$, $K(P, Q, R)$ the polynomials obtained by replacing (f, ∇, C) by (P, Q, R)

in (V.31). Writing $P \cdot u = P \circ u$, $u \in \mathrm{GL}(3,\mathbb{C})$, one then sees that

$$\nabla(P \cdot u) = (\det u)^2\, \nabla(P) \cdot u, \tag{V.32}$$

$$C(P \cdot u, Q \cdot u) = (\det u)^2\, C(P,Q) \cdot u, \tag{V.33}$$

$$K(P \cdot u, Q \cdot u, R \cdot u) = (\det u)\, K(P,Q,R) \cdot u. \tag{V.34}$$

Proposition V.2.4. — *The algebra of G-invariant polynomials is generated by f, ∇, C and K.*

Proof. — The intersections of the quartic C_4 with the curves defined by ∇, C and K are unions of orbits of the group G; by Bézout's theorem they have respectively 24, 56 and 84 points. Thus they are comprised of the inflection points ($\nabla = 0$), the points of contact with bitangents ($C = 0$), and the sectactic points ($K = 0$). Furthermore, the quotient C^3/∇^7 defines a nonconstant meromorphic function, hence surjective, from C_4 to $\mathbb{CP}^1$. Since the degree of both ∇^7 and C^3 is 42, the intersection of a curve of the form $\lambda\nabla^7 + \mu C^3 = 0$ with C_4 can have at most 168 points. On the other hand, since C^3/∇^7 is G-invariant, the fibres are unions of orbits of G, so in fact have size exactly 168 (counting multiplicities). Since the cardinality of a union of orbits of G must be of the form $24\alpha + 56\beta + 84\gamma + 168\zeta$, where $\alpha, \beta, \gamma = 0$ or 1, the only possibility is for just one of $\alpha, \beta, \gamma, \zeta$ to be nonzero; in other words each fibre consists of a single orbit.

Thus each orbit of the action of G on C_4 is given by a curve of the pencil $\lambda\nabla^7 + \mu C^3 = 0$ ($[\lambda, \mu] \in \mathbb{CP}^1$). Hence if P is a G-invariant polynomial not proportional to f, the intersection $\{P = 0\} \cap C_4$ must be a finite union of G-orbits, and there exists $Q \in \mathbb{C}[\nabla, C, K]$ such that $\{P = 0\} \cap C_4$ coincides with the intersection $\{Q = 0\} \cap C_4$ (with equality of multiplicities). Thus P/Q defines a holomorphic function on C_4, implying the existence of a constant $\lambda \in \mathbb{C}$ such that $P - \lambda Q = 0$ on C_4. It follows (choosing an affine chart) that $P - \lambda Q$ belongs to the ideal generated by f and therefore P belongs to $\mathbb{C}[f, \nabla, C, K]$. □

Hence in particular f, ∇ and K are the only homogeneous invariant polynomials (up to multiplication by a constant) of degrees 4, 6 and 21. By the uniqueness, one sees that K is necessarily the product of 21 degree-one factors corresponding to the straight lines made up of points fixed by the involutions of G (for example, $(a+1)x + by + cz = 0$ in the case of s — see equation (V.29)), each meeting C_4 in 4 sectactic points; the group G permutes these 21 straight lines, whence the invariant polynomial of degree 21.

The orbit of the sectactic points (counted twice) is the intersection of C_4 with some curve with equation $\lambda\nabla^7 + \mu C^3 = 0$ and is clearly also the intersection with $K^2 = 0$. Hence there is a relation of the form $\lambda\nabla^7 + \mu C^3 + \nu K^2 = 0$ modulo f.

Evaluating this at the point $[1,0,0]$, and using the fact that $C = x^{14}+y^{14}+z^{14}+\cdots$ and $K = -(x^{21} + y^{21} + z^{21}) + \cdots$, one infers that $\mu = -\nu$. In order to obtain another such relation, Klein evaluates the above equation at the points of contact with bitangents, a calculation most simply carried out in terms of the following coordinates $[y_1, y_2, y_3]$, introduced by Klein in connection with his investigation of the involutions [Kle1878c, §5]:

$$-i\sqrt{3}\sqrt[3]{7}x = y_1 + \beta y_2 + \beta' y_3 \tag{V.35}$$

$$-i\sqrt{3}\sqrt[3]{7}y = y_1 + \alpha^2\beta y_2 + \alpha\beta' y_3 \tag{V.36}$$

$$-i\sqrt{3}\sqrt[3]{7}z = y_1 + \alpha\beta y_2 + \alpha^2\beta' y_3 \tag{V.37}$$

where $\alpha = e^{2i\pi/3}$, $\beta^3 = 7(3\alpha^2 + 1)$ and $\beta\beta' = 7$. In terms of these coordinates, the bitangent $x + y + z = 0$ has equation $y_1 = 0$ and the points of contact become $[0,1,0]$ and $[0,0,1]$. The polynomial $f(x,y,z)$ becomes

$$F = 3^{-1}7^{-4/3}(y_1^4 + 21y_1^2y_2y_3 - 147y_2^2y_3^2 + 49y_1(y_2^3 + y_3^3)). \tag{V.38}$$

Since (V.35) *et seqq.* defines an element of $\mathrm{SL}(3,\mathbb{C})$, one can apply (V.32) *et seqq.* to the calculation of the transforms of ∇, C and K directly from F, yielding $\nabla = 7^2y_3^6/3^3 + \cdots$ and $K = -2^37^7y_3^{21}/3^9 + \cdots$, whence $\lambda + 12^3\nu = 0$ and, finally, the desired relation

$$12^3\nabla^7 + C^3 - K^2 = 0 \pmod{f}. \tag{V.39}$$

One may consult [Adl1999, p. 262] for a relation linking ∇, C, K and f. The determination of the function J is now at hand. One observes that the function $J\nabla^7/C^3$ is holomorphic on C_4 and so constant; its value can be obtained by evaluating the relation (V.39) at the sectactic points (using the fact that J takes the value 1 at those points), whence one obtains

$$J = -\frac{C^3}{12^3\nabla^7} \quad \text{and} \quad j = -\frac{C^3}{\nabla^7}. \tag{V.40}$$

V.2.4. Inflection triangles and a resolvent of degree 8

The following step is essential [Kle1878c, §8]: in his systematic study of the action of G on C_4, especially noteworthy are his investigation of the action of G on the inflection triangles and the reappearance of a degree-8 equation already solved in his earlier article [Kle1878b, §18]; the equation in question is (V.19) above, now linked to $X_0(7)$ (see §V.1.5). We now supply the details.

Recall that the stabilizer of the inflection triangle $xyz = 0$ is the subgroup G'_{21} of G, generated by r and h (see equations (V.24) and (V.28)). Let s be the involution defined by (V.29). The right cosets of G'_{21} are G'_{21} itself and $G'_{21}sh^k$ $(k = 0,\dots,6)$. The action of G on $\delta_\infty = -7xyz$ yields the following polynomials:

$$\delta_k = xyz + \gamma^{-k}(x^2y - z^3) + \gamma^{-4k}(y^2z - x^3) + \gamma^{-2k}(z^2x - y^3) + 2\gamma^k z^2y + 2\gamma^{4k}x^2z + 2\gamma^{2k}y^2x \quad (k = 0,\dots,6)$$

which determine the 8 inflection triangles of C_4 (with $\gamma = e^{2i\pi/7}$). It follows that the coefficients of $P = (\delta - \delta_\infty)\prod_{k=0}^{6}(\delta - \delta_k)$, considered as a polynomial in δ, are G-invariant polynomials. By taking account of degrees, one sees that

$$P = \delta^8 + a_6\nabla\delta^6 + a_4\nabla^2\delta^4 + a_2\nabla^3\delta^2 + a_1K\delta + a_0\nabla^4 \pmod f \tag{V.41}$$

where the a_j are constants.

Klein indicates that one may determine the coefficients a_j by identification. However, to ease the calculation we shall find them by evaluating (V.41) at judiciously chosen points. The coefficients of P correspond to symmetric functions in the polynomials δ_∞, δ_k, so the a_j are real; the coefficient of δ^6 is proportional to ∇ (up to a constant factor the unique invariant polynomial of degree 6), but the others are as shown in (V.41) only modulo f. Setting (x,y,z) equal to $(1,1,1)$ and then to $(1,\alpha,\alpha^2)$ (with $\alpha = e^{2i\pi/3}$), one obtains $a_6 = -14$ and $a_0 = -7$. Substitution of the point of inflection $(1,0,0)$ and $\delta = \delta_0$ yields immediately $a_1 = -1$ since $K(1,0,0) = \delta_0(1,0,0) = -1$. Evaluating the polynomials δ_∞, δ_k at the points of inflection yields $\delta^8 + \delta = 0$. Next, the fact that the a_j are real gives $a_2 = -70$ and $a_4 = 63$ on substituting $\delta = \delta_\infty$ and $(x,y,z) = (1,\alpha,\alpha^2)$ (here the value $K(1,\alpha,\alpha^2)$ is calculated using the coordinate change of (V.35) *et seqq* of §V.2.3. Finally, at a point $[x,y,z] \in C_4$, the polynomials δ_∞ and δ_k $(k = 0,\dots,6)$ are roots of

$$\delta^8 - 14\nabla\delta^6 + 63\nabla^2\delta^4 - 70\nabla^3\delta^2 - K\delta - 7\nabla^4 = 0. \tag{V.42}$$

Compare this equation with (V.19). Since $12^3(J-1) = -K^2/\nabla^7$ on C_4, it follows that if δ is a root of (V.42), then $\delta/\sqrt{-\nabla}$ is a root of (V.19) (for a appropriate square root) and its square $-\delta^2/\nabla$ is a root of (V.18) describing the fibre of $X_0(7) \to X(1)$ (see §V.1.5). In fact Klein begins by showing that the solutions of the "modular equation" (V.18) can be expressed as rational functions of a point of the curve C_4 (namely $-\delta_\infty^2/\nabla$ and $-\delta_k^2/\nabla$, $k = 0,\dots,6$), and then infers the relation (V.42). He remarks that $-\delta_\infty^2/\nabla$ induces an isomorphism between $G'_{21}\backslash C_4 \simeq X_0(7)$ and $\mathbb{CP}^1$.

V.2.5. Conclusion

The modular equation (V.42) — or degree-8 resolvent — has a remarkable property, discovered by Jacobi [Jac1828, p. 308]: the square roots of its 8 solutions, chosen appropriately, depend linearly on 4 parameters.

It is easy to bring this property to light starting from the expressions for δ_∞ and δ_k. Following Klein, we write, for $[x,y,z] \in C_4$,

$$A_0 = \sqrt{xyz},\ A_1 = \sqrt{-y^3 - z^2x},\ A_2 = \sqrt{-z^3 - x^2y} \text{ and } A_3 = \sqrt{-x^3 - y^2z}. \tag{V.43}$$

Then, to within a change of sign of A_1, A_2 or A_3, we have $A_0A_1 = x^2y$, $A_0A_2 = y^2z$ and $A_0A_3 = z^2x$. Hence $A_1A_2 = xy^2$, $A_2A_3 = yz^2$, $A_3A_1 = zx^2$, and the equations of the inflection triangles (see the expressions δ_∞, δ_k above) take the form[15]

$$\sqrt{\delta_\infty} = \sqrt{-7}A_0, \tag{V.44}$$

$$\sqrt{\delta_k} = A_0 + \gamma^{-k}A_1 + \gamma^{-4k}A_2 + \gamma^{-2k}A_3, \quad (k = 0,\dots,6). \tag{V.45}$$

Recall that $\delta_\infty/\sqrt{-\nabla}$ and $\delta_k/\sqrt{-\nabla}$ are solutions of equation (V.19). It remains only to express the ratios $A_1/A_0 = x/z$, $A_2/A_0 = y/x$ and $A_3/A_0 = z/y$ as functions of the ratios $\sqrt{\delta_k}/\sqrt{\delta_\infty}$, that is, in terms of the solutions w_∞ and w_k $(k = 0,\dots,6)$ of equation (V.19), in order to obtain a parametrization of C_4 by the single variable q. The appropriate choice of signs for the square roots can be determined by eliminating the A_j in the equations (V.44) and (V.45). We take (see equations (V.20), (V.21))

$$\sqrt{w_\infty}/\sqrt{-7} = -q^{1/4}\prod_{n=1}^{\infty}\frac{1-q^{7n}}{1-q^n}, \tag{V.46}$$

$$\sqrt{w_k} = \gamma^{-2k}q^{-1/28}\prod_{n=1}^{\infty}\frac{1-\gamma^{nk}q^{n/7}}{1-q^n} \quad (k = 0,\dots,6). \tag{V.47}$$

Clearly, it should be possible to express the A_j/A_0 $(j = 1,2,3)$ directly in terms of the δ_k/δ_∞ $(k = 0,\dots,6)$. The approach via the square roots allows the use of Euler's "pentagonal identity" (see, for example, [McKMo1997, p. 143]):

$$\prod_{n=1}^{\infty}(1-q^n) = \sum_{n\in\mathbb{Z}}(-1)^n q^{n(3n+1)/2}, \tag{V.48}$$

[15] Klein also gives an interpretation of these formulae in terms of "cubics of contact" containing the inflection triangles [Kle1878c, §9].

while expanding $\prod_{n\geq 1}(1-q^n)^2$ is not easy. From the relations (V.44) and (V.45) one infers that $7A_1 = \sum_{0\leq k\leq 6}\gamma^k\sqrt{\delta_k}$ and $A_0 = \sqrt{\delta_\infty}/\sqrt{-7}$, whence, taking account of (V.46), (V.47) and (V.48), one obtains:

$$\begin{aligned} 7q^{2/7}\prod_{n=1}^{\infty}(1-q^{7n})\frac{A_1}{A_0} &= -\sum_{k=0}^{6}\gamma^{-k}\prod_{n=1}^{\infty}(1-\gamma^{nk}q^{n/7}) \\ &= -\sum_{n\in\mathbb{Z}}(-1)^n q^{\frac{n(3n+1)}{14}}\sum_{k=0}^{6}(\gamma^{\frac{n(3n+1)}{2}-1})^k. \qquad \text{(V.49)} \end{aligned}$$

Here the sum over k is zero except when $n(3n+1)/2-1$ is divisible by 7 (in which case the sum is 7). This occurs if and only if n is of the form $7m+3$ or $7m+6$, whence $n(3n+1)/14$ is $1/7+(21m^2+19m+4)/2$ or $1/7+(21m^2+37m+16)/2$. By means of a similar calculation for A_j/A_0 ($j = 2,3$), one finally obtains the formulae of Theorem V.0.6.

To within a permutation of the variables x, y, z and replacing $q^{1/2}$ by $q = e^{i\pi\tau}$, these are the formulae given by Klein[16] in [Kle1878c, §9].

It is also of interest to consult [Elk1999, p. 84] where a direct parametrization is described (via the canonical embedding — see §V.2.2) in terms of 1-forms on $X(7)$:

$$x, y, z = \epsilon q^{a/7}\prod_{n=1}^{\infty}(1-q^n)^3(1-q^{7n})\prod_{n>0,\, n\equiv\pm n_0 \bmod 7}(1-q^n)$$

where $q = e^{2i\pi\tau}$ and the triple (ϵ, a, n_0) is $(-1,4,1)$ for x, $(1,2,2)$ for y, and $(1,1,4)$ for z. In this version, the G-invariant polynomials define parabolic modular forms for $\mathrm{SL}(2,\mathbb{Z})$ and may therefore easily be linked to the variable q. Thus $x^3y + y^3z + z^3x = 0$ or $\nabla(x,y,z)$ is proportional to the discriminant Δ.

[16] He revisits the question again in [Kle1880b] in order to obtain more pleasing formulae, expressed in terms of the average of partial values of theta functions and also allowing of a uniform treatment for transformations of orders 5, 7, and 11.

Part B

The method of continuity

The years 1880–1882 are crucial to our theme. It was then that Klein and Poincaré announced and then "proved" that *all algebraic curves of genus at least 2 can be uniformized by the disc.* This came as a great surprise to the mathematicians of the time. Examples were known — we saw some of them earlier — but that the result held in such generality seemed incredible. Even today it has the status of a major and highly nontrivial fact about the geometry of algebraic curves — to such an extent, indeed, that many mathematicians profess to know it "so well" that they forget that it is so highly nontrivial and all too often confuse it with one or another of two theorems which, although certainly important, are much older (and much simpler): the *Riemann Mapping Theorem* (the first convincing proof of which, as we shall see, was given by Osgood) according to which a nontrivial simply connected open subset of the plane is conformally equivalent to the disc, and *Gauss's theorem* (often wrongly attributed to Riemann) stating that a (real analytic) surface is locally conformal to an open subset of the plane.

Even though the present work is not a history book, a brief introduction to the protagonists may nonetheless be useful.

In 1880 Poincaré was a 26-year-old assistant professor[17]. He had defended his thesis two years earlier on the research topic of differential equations. It is undeniable that differential equations were at the root of almost all of his subsequent discoveries. The Paris Academy had proposed in 1878, as the theme of its competition for the Grand Prize in the mathematical sciences to be awarded in 1880, the following problem: "To bring to perfection in some significant aspect the theory of linear differential equations in a single independent variable". Since he had founded the qualitative theory of dynamical systems a few months earlier[18], Poincaré now began to investigate differential equations in a single variable. In March 1880 he submitted a first memoir on the real theory, and then withdrew it in June of the same year. In the meantime he had become aware — in May 1880 — of an article by Fuchs on second order linear differential equations with algebraic coefficients. The memoir that he finally submitted to the Academy — in June 1880 — contains reflections inspired by Fuchs's article, reproduced in [Poin1951, Tome I, pp. 336–373]. In the work that so stimulated Poincaré, Fuchs sought to generalize Jacobi's inversion. He considered in particular the inverse function of the quotient of two independent solutions of a second-order differential equation and gave a necessary and sufficient condition for this function to be meromorphic. Since Fuchs's theory is essentially only local, Poincaré was struck by the result but found it unconvincing. He understood that Fuchs's result was an (excessively

[17] Actually a *maître de conférences*, equivalent to senior lecturer in a British university or assistant professor in North America. *Trans*

[18] At this time Poincaré was also feeling the need to develop an autonomous topological theory (which, as we know, he subsequently realized).

strong) version of uniformization[19]. Be that as it may, at that time Poincaré was in the midst of an attempt to understand second-order linear differential equations with algebraic coefficients *via* Fuchs's theory — and it was in connection with this aim that he created the theory of Fuchsian groups. Details of this first stage in this engaging story[20] are omitted from the present book. Happily, the existence of [Poin1997] excuses us somewhat. We might summarize these first months by saying simply that Poincaré immersed himself, with all his genius but also with a certain "naïveté", in the new theory. His correspondence with Klein shows, for instance, that at that time he had not read Riemann!

Klein was six years older than Poincaré. He had by that time been a professor already for ten years, and, possessed of an immense mathematical culture, was probably the most prominent mathematician of the era. He was certainly one of the finest connoisseurs of Riemann's works and knew the theory of elliptic functions thoroughly. He was one of the most influential propagators of the group concept in mathematics: his "*Erlanger Programm*" of 1872, announced on the occasion of his nomination to a professorship (at the age of 23), shows astonishing perspicacity. He had at that time already published major articles on the uniformization of certain particular algebraic curves arising in number theory. He had also established the projective character of (real) non-Euclidean geometry. When Professor Klein learns of Poincaré's first notes on Fuchsian groups (dating from February 1881) he is astounded both at the generality of the latter's constructions and his ignorance of the literature — in particular German — on the topic. On June 12, 1881 he begins a correspondence with his young colleague on the other side of the Rhine, destined to continue till September 22, 1882.

We reproduce this celebrated correspondence in an appendix, and strongly recommend it to the reader. One sees there a (scientific!) confrontation between a beginner and an established professor, tinged with oblique political references. Also in evidence is the increase in mutual respect over the course of the correspondence. But best of all one sees there the genesis of the uniformization theorem, gaining in precision of formulation almost day by day. It should also be mentioned how Poincaré's genius compels Klein's respect — respect he gladly acknowledges subsequently.

---

[19]The day after the submission of his memoir to the Academy, Poincaré also sent to Fuchs the first of a series of letters in which the young assistant professor tried — without success — to explain to Professor Fuchs that a local diffeomorphism need not necessarily be a covering. Note in this connection that, throughout his work on uniformization, the ease with which Poincaré deals with what is not yet explicitly covering-space theory is certainly one of his essential assets — to such an extent that some have wished to see in the construction of the universal covering space Poincaré's main contribution to the problem. However, as we will see, the latter contention is largely an exaggeration.

[20]That is, the part consisting not only of the memoir submitted to the Academy but also the three supplements brought to light by Gray in 1979 and published in [Poin1997].

The first Fuchsian functions Poincaré constructs (in a note of May 23, 1881, [Poin1951, T. II, pp. 12–15]) uniformize surfaces obtained by removing a finite number of *real* points from a sphere (Poincaré also allows what are now called "orbifold singularities"[21]). He arrives independently at functions introduced earlier, as Klein points out to him, by Schwarz (see Chapter IV). Poincaré's method is quite different, however. He considers (Fuchsian) groups generated by reflections in the sides of ideal n-sided hyperbolic polygons. These groups depend on $n-3$ real parameters $1 < x_1 < \ldots < x_{n-3}$ and he identifies the space of these groups with the space of moduli of spheres with n real points removed. This represents the first appearance of the *method of continuity*[22]. It is clear that "from the beginning Poincaré has a lead that Klein can no longer make up" [Freu1955]. On August 8, 1881 Poincaré makes the following announcement [Poin1951, Tome II, pp. 29–31]:

> We conclude from this that:
>
> 1. Every linear differential differential equation with algebraic coefficients is integrable by means of zeta Fuchsian functions;
> 2. The coordinates of the points of every algebraic curve can be expressed by means of functions of an auxiliary variable.

This represents the very first enunciation of the uniformization theorem. It is, however, always necessary to moderate the enthusiasm of the young Poincaré a little. What he had actually proved (but completely rigorously) was appreciably weaker: *every algebraic curve can be "uniformized" by means of a function from the disc to the curve except for at most a finite number of points*. For Poincaré, motivated as he was by the integration of differential equations by means of functions given explicitly by series, excepting a finite number of points was not a problem. Moreover the proof of his result was in fact especially simple and elegant: given an algebraic curve branched over the sphere, up to removing the branch points one obtains a covering of the sphere with a finite number of points removed. It then only remains to show that up to the removal of finitely many more points, one has a covering of the sphere with finitely many *real* points removed. (This last step is an elementary exercise which we recommend to the reader.)

It is in fact Klein to whom the honor belongs of enunciating the uniformization theorem for algebraic curves as we now understand it. Klein, less interested in differential equations, in effect prefers finite polygons. Moreover his intimate

---

[21] An "orbifold" is a certain generalization of a manifold with singularities. *Trans*

[22] The method of continuity, as conceived by Poincaré, is explicitly described in Chapter IX in the case of spheres with 4 points removed. We leave to the reader as an exercise the verification that the method becomes considerably simpler when the 4 points are real.

knowledge of Riemann's work allows him to identify the number of moduli of curves of a fixed genus with the number of parameters on which Poincaré's polygons of the same genus depend. He is thus more naturally inclined to produce the "correct formulation" (see Freudenthal [Freu1955] and Scholz [Schol1980]); "this is the only essential point in which Klein, in his research on automorphic functions, surpassed Poincaré" [Freu1955]. The great principle is still the method of continuity, however implementing it in the needed generality is difficult. The correspondence between Klein and Poincaré shows very clearly just how each interprets it according to his own point of view.

Thus Klein observes that Poincaré's construction of Fuchsian groups produces uniformizable algebraic curves, and that these depend on parameters equal in number to those of the moduli space of curves of fixed genus. He notes also that if a Riemann surface can be uniformized, then this is possible in one way only. Thus the problem reduces to showing that the space of uniformizable curves is both open and closed. The question of the connectedness of the moduli space is mentioned by Klein as established in his book [Kle1882c], which we have already described[23].

Poincaré, on the other hand, was interested in second-order linear differential equations on an algebraic curve and showed that their description depended on a "monodromy" representation of the fundamental group (which he had then not as yet "invented") in $\mathrm{SL}(2,\mathbb{C})$. When a differential equation on a fixed algebraic curve is allowed to vary, so also does this representation vary. In his examples of uniformizable curves (given by Fuchsian groups) one of these differential equations is privileged and has real monodromy group: Poincaré calls this equation *Fuchsian*. He asserts that every algebraic curve possesses a Fuchsian equation and that this allows him to show that his construction of Fuchsian groups is flexible enough to yield a description of all algebraic curves. The "proof" that he proposes also contains a component devoted to openness and another to closure. His attachment to ideal polygons allows him to more easily identify the difficulties associated with closure; see [Schol1980].

Both Klein and Poincaré later published descriptions of this period in their lives. Poincaré's text on "mathematical invention", dating from 1908, is famous [Poin1908]. There he describes his discovery of the link between differential equations and hyperbolic geometry as pre-dating his first epistolary contact with Klein.

> At that time I left Caen, where I was then living, in order to take part in a geology course undertaken by the School of Mines. The hazards of the trip caused me to forget my mathematical labors; when we arrived in Coutances

[23]His "proof" is hardly convincing.

> we climbed into an omnibus to go I knew not whither. At the instant I placed my foot on the step, the idea came to me, seemingly without anything in my mind having prepared me for it earlier, that the transformations I had used to define Fuchsian functions were identical to those of non-Euclidean geometry. I carried out no verification of this, I wouldn't have had the time since scarcely had I entered the omnibus when I resumed an earlier conversation; nonetheless I immediately felt complete certitude. Once back in Caen, I checked the result at leisure to satisfy my conscience.

It is indisputable that Poincaré had grasped the essentials of the theory before beginning his correspondence with Klein. In his third supplement to the memoir for the prize of the Academy, submitted on December 20, 1880, he "conjectures" that Fuchsian functions allow one to solve *all* linear differential equations with algebraic coefficients [Poin1997]:

> I have no doubt, moreover, that the many equations envisaged by M. Fuchs in his memoir in Volume 71 of Crelle's journal... will furnish an infinity of transcendentals... and that these new functions will allow the integration of all linear differential equations with algebraic coefficients.

One observes here, however, the absence of any formulation of the situation in terms of the uniformization of algebraic curves.

As for Klein, in his book on the development of 19th century mathematics [Kle1928] he explains:

> During the last night of my journey, that from March 22 to March 23 [1882], which I spent sitting on a couch on account of an attack of asthma, suddenly, towards 3:30, the central theorem dawned on me as if it had been sketched in the figure of the 14-sided polygon. Next morning, in the coach which at that time travelled between Norden and Emden, I thought about what I had found, going over all the details once more. I knew then that I had found an important theorem. Once arrived in Düsseldorf, I wrote up the memoir, dated March 27, sent it off to Teubner, and had copies sent to Poincaré and Schwarz, and also to Hurwitz.

In [Kle1921a, Vol. 3, pp. 577–586], there is an addendum to the effect that he considered that neither he nor Poincaré had a complete proof and that the proof using the method of continuity had been firmly established only by Koebe in 1912 [Koe1912]. He also describes that episode in his life as marking "the end of his productive period". He fell ill in the autumn of 1882[24].

---

[24]"Leipzig seemed to be a superb outpost for building the kind of school he now had in mind:

Unfortunately the second part of his book makes only a superficial contribution to the description of this mathematical adventure. Freudenthal's fine article [Freu1955] served us as a point of departure. Klein's book [Kle1928] is an essential reference for the history of 19th century mathematics, written by one of the heroes of the present work. By way of complementing these, the reader may also consult the relevant chapter of the historical book by J. Gray [Gra1986], the remarkable analysis [Die1982] by J. Dieudonné, the introduction [Poin1997] to Poincaré's three supplements to his memoir on the discovery of Fuchsian functions, J. Stillwell's commentary to his translation into English of Poincaré's articles on Fuchsian functions [Poin1985], the relevant chapter of the impressive thesis by Chorlay [Cho2007], the commentaries attached to the French version of the Klein–Poincaré correspondence [Poin1989], or Fricke's article [Fric1901] in the *Encyklopädie der mathematischen Wissenschaften*. Finally, there is the article by Abikoff [Abi1981], from which, while interesting also mathematically speaking, we quote, for our present historical purposes, only his version of the reception by Hurwitz, Schwarz, and Poincaré of the latter's proof of the uniformization theorem:

> – Hurwitz: I accept it without reservation.
>
> – Schwarz: It's false.
>
> – Poincaré: It's true. I knew it and I have a better way of looking at the problem.

Chapter VI is an introduction to Fuchsian groups. The reader will find there, for instance, the construction of the Fuchsian group associated with a fundamental polygon, and also the construction of automorphic forms and Fuchsian functions invariant under the action of a given Fuchsian group. As current references for Fuchsian groups we might mention the books [Kat1992] and [Bea1983], the latter dealing also with their generalization to higher dimensions: discrete groups of isometries of hyperbolic space, notably the Kleinian groups in dimension 3. For Kleinian groups one may also consult [Dal2007] and [Mas1988]. The paper [Mas1971] gives the first complete and correct proof of Poincaré's polygon theorem (Theorem VI.1.10 below).

---

one that would draw heavily on the abundant riches offered by Riemann's geometric approach to function theory. But unforeseen events and his always delicate health conspired against this plan. [In him were] two souls [. . .] one longing for the tranquil scholar's life, the other for the active life of an editor, teacher, and scientific organiser. [. . .] It was during the autumn of 1882 that the first of these two worlds came crashing down upon him [. . .] his health collapsed completely, and throughout the years 1883–1884 he was plagued by depression" [Row1989].

Chapter VII is a variation on Klein's approach, and there we make no attempt to pronounce on the validity of the proofs proposed by him[25]. We propose a "re-constitution" of what purported to be a proof of the theorem on the uniformization of algebraic curves along the lines of the method of continuity as viewed by Klein. This proof uses tools developed later, but in a weak form. In sum, Chapter VII is in some sense the article Klein might have written if he had more tools at his disposal. In the space of twenty years the literature on the representations of surface groups has grown enormously. The article [GolW1988] is an important reference for the questions evoked in this chapter. For a presentation adhering more closely to the ideas of Klein, the reader may consult the classic book [FrKl1897].

Chapter VIII is an introduction to the approach of Poincaré. We first explain there how uniformization theory can be expressed in terms of second-order linear differential equations, and then give a proof of the openness of the space of uniformizable curves. In this connection it is necessary to complete some of Poincaré's arguments, but in a relatively light-handed manner. However, as far as his approach to closure is concerned, we do not expound it because it fails to convince us, and also because we cannot see our way to "repairing" it without in fact using the arguments of Chapter VII.

Finally, in Chapter IX we put Poincaré's approach to work in the analysis of special cases, and also describe the subsequent life of these methods. In particular, we expound there the explicit examples of uniformization obtained by Schwarz in his investigation of the hypergeometric equation.

As we have already mentioned, the uniformization theorem is not confined to algebraic curves. Emboldened by this "special case" (yet an already enormously general one), Poincaré went on to attempt to generalize it to all simply-connected Riemann surfaces not necessarily universal covering spaces of compact surfaces. Here he can no longer resort to finite-dimensional moduli spaces or monodromy groups. Koebe and Poincaré succeeded in 1907, and we shall explain how in Part C.

---

[25]Except to say that his approach to closure does not appear convincing to us.

Chapter VI

Fuchsian groups

In his articles of 1882–1886 in *Acta Mathematica*, Poincaré proposes new "transcendentals" on the model of elliptic functions (Chapter I). His initial motivation was to develop in power series global solutions of linear differential equations with algebraic coefficients. The then recent work of Fuchs on singular points of linear differential equations [Fuc1880, Fuc1881] showed that the solutions can be expressed as analytic functions of a finitely ramified variable, $z^{1/q}$, or an infinitely ramified one, $\log(z)$ (corresponding to the case "$q = \infty$"). Poincaré seeks a global analogue, at first in the form of the universal cover of $\mathbb{CP}^1$ with finitely many points removed, including the singular points of the differential equation; then, in conjunction with the successive appearances of notes in the *Comptes rendus de l'Académie des sciences de Paris* between February 1881 and April 1882, one sees the statement become progressively more precise, culminating in the universal covering space of a compact Riemann surface with an "orbifold structure". First, however, we expound his construction of Fuchsian groups.

VI.1. Fuchsian groups, the fundamental polygon, and hyperbolic tilings

Recall that $\mathbb{H}$ denotes the Poincaré half-plane

$$\mathbb{H} = \{x + \sqrt{-1}y = z \in \mathbb{C} \mid y > 0\}$$

endowed with the hyperbolic metric $y^{-2}dzd\overline{z}$. Here the geodesics are semicircles centered on the real axis $y = 0$ together with the vertical half-lines.

Sometimes the disc model of the hyperbolic plane is used instead, that is, the disc $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ endowed with the hyperbolic metric $4(1 - |z|^2)^{-2}dzd\overline{z}$, where the geodesics are the arcs of circles and segments orthogonal to the unit circle $\partial\mathbb{D}$.

VI.1.1. The isometries of the Poincaré half-plane

The group $\mathrm{PSL}(2,\mathbb{R})$ acts on $\mathbb{H}$ by Möbius transformations: $z \mapsto \frac{az+b}{cz+d}$ where a, b, c and d are real numbers satisfying $ad - bc = 1$. This action is isometric with respect to the hyperbolic metric. We remind the reader that, even more, $\mathrm{PSL}(2,\mathbb{R})$ coincides with the group of holomorphic self-diffeomorphisms of $\mathbb{H}$.

There are three types of elements in $\mathrm{PSL}(2,\mathbb{R})$, characterized by their respective fixed points in $\overline{\mathbb{H}} = \mathbb{H} \cup \partial\mathbb{H}$. (Here we are thinking of $\mathbb{H}$ as a disc in the Riemann sphere, so that $\partial\mathbb{H}$ contains the point at infinity.) The first type consists of *elliptic* transformations, which are just those for which the inequality $|a+d| < 2$ holds; each such transformation has just one fixed point in $\overline{\mathbb{H}}$, which is in fact in $\mathbb{H}$. Every elliptic transformation is conjugate to a unique transformation of the form $z \mapsto \frac{\cos\theta\ z+\sin\theta}{-\sin\theta\ z+\cos\theta}$ for some $\theta \in \mathbb{R}$. (For these transformations it is more convenient to use the disc model, where every elliptic transformation is conjugate to a rotation $z \mapsto e^{i\theta} z$.) Note that an elliptic transformation generates a relatively compact subgroup of $\mathrm{PSL}(2,\mathbb{R})$.

The second type of isometry is characterized by the inequality $|a + d| > 2$. These are the *hyperbolic* transformations, and are each conjugate in $\mathrm{PSL}(2,\mathbb{R})$ to a unique transformation of the form $z \mapsto \lambda z$, with $\lambda > 0$ and not equal to 1. A transformation φ of this type has exactly two (distinct) fixed points in $\overline{\mathbb{H}}$, both on the boundary $\partial\mathbb{H}$. One of these points, which we denote by p^+, is *attractive*, and the other, p^-, is *repulsive*, in the following sense: if $z \in \overline{\mathbb{H}}$ is different from p^-, then $\varphi^n(z)$ tends to p^+ as n tends to $+\infty$, and, analogously, if $z \in \overline{\mathbb{H}}$ is different from p^+, then $\varphi^n(z)$ tends to p^- as n tends to $-\infty$. The hyperbolic geodesic connecting p^+ and p^- is called the *axis* of φ.

Lastly, the elements of $\mathrm{PSL}(2,\mathbb{R})$ different from the identity and satisfying $|a + d| = 2$ are called *parabolic*. A parabolic transformation φ is conjugate in $\mathrm{PSL}(2,\mathbb{R})$ to one of the two transformations $z \mapsto z \pm 1$, and has just one fixed point p in $\overline{\mathbb{H}}$, which is in fact located on the boundary $\partial\mathbb{H}$. For every $z \in \overline{\mathbb{H}}$, $\varphi^n(z)$ tends to p as n tends to ∞.

VI.1.2. Fuchsian groups

A *Fuchsian group* is defined to be a discrete subgroup Γ of $\mathrm{PSL}(2,\mathbb{R})$.

Proposition VI.1.1. — *A subgroup Γ of $\mathrm{PSL}(2,\mathbb{R})$ is discrete if and only if it acts discontinuously on $\mathbb{H}$, that is, if and only if each of its orbits is discrete.*

Proof. — The group $\mathrm{PSL}(2,\mathbb{R})$ acts freely and transitively on the unit tangent bundle $\mathrm{U}\mathbb{H}$. Hence Γ is discrete if and only if it acts discontinuously on $\mathrm{U}\mathbb{H}$. Then since the fibres of the projection $\mathrm{U}\mathbb{H} \to \mathbb{H}$ are compact, the orbits are discrete in $\mathrm{U}\mathbb{H}$ if and only if the orbits of $\mathbb{H}$ under the action of Γ are discrete. □

VI.1.3. A fundamental polygon and its associated tiling

A *polygon* $P \subset \mathbb{H}$ is defined to be a closed, convex subset with piecewise geodesic boundary, where the number of bounding geodesic arcs is locally finite in $\mathbb{H}$. A *side* of P is then a maximal geodesic arc contained in the boundary of P. Thus the intersection of two sides of P is either empty or consists of a single point, in which case we call this point a *vertex* of P. We say that P is *finite* if it has only finitely many sides.

Given a Fuchsian group Γ, a polygon $P \subset \mathbb{H}$ is called a *fundamental polygon* for Γ if each orbit $\Gamma(z)$, $z \in \mathbb{H}$, intersects P in at least one point and intersects the interior $\mathring{P}$ in at most one point. It follows that the set of translates $\varphi(P)$, $\varphi \in \Gamma$, defines a *tiling* of the hyperbolic plane:

$$\bigcup_{\varphi \in \Gamma} \varphi(P) = \mathbb{H} \quad \text{and} \quad \varphi(\mathring{P}) \cap \psi(\mathring{P}) = \emptyset \quad \text{for all} \quad \varphi \neq \psi.$$

The set of tiles $\{\varphi(P) \mid \varphi \in \Gamma\}$ is thus in one-to-one correspondence with Γ. We saw a classic example of a tiling for $\mathrm{PSL}(2,\mathbb{Z})$ in the preceding chapter (see Figure V.1). Two more are given in Figure VI.1 below.

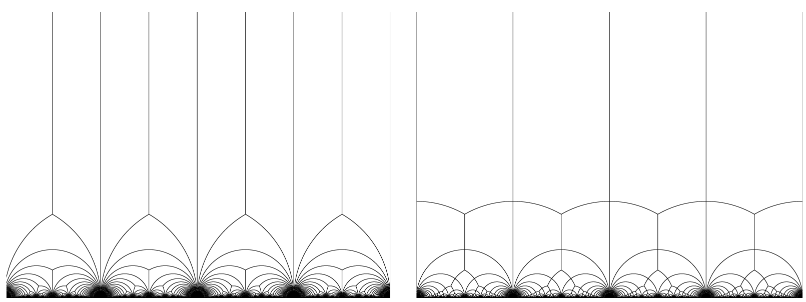


Figure VI.1: Two more tilings for $\mathrm{PSL}(2,\mathbb{Z})$ (variants)

Theorem VI.1.2. — *Let* $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$ *be a Fuchsian group and* z_0 *a point of* $\mathbb{H}$ *not fixed by any nontrivial element of* Γ. *Then the set*

$$P = \left\{ z \in \mathbb{H} \mid d_{\mathrm{hyp}}(z, z_0) = d_{\mathrm{hyp}}(z, \Gamma(z_0)) \right\}$$

of points $z \in \mathbb{H}$ *that are closer to* z_0 *than to any other point of the orbit* $\Gamma(z_0)$ *is a (convex) fundamental polygon for* Γ. *Furthermore* P *is finite if* Γ *is finitely generated.*

Poincaré considers only finitely generated Fuchsian groups, showing that for such a group there exists a fundamental region bounded by a finite number of curvilinear arcs, and explaining how to modify such a region to make it polygonal. However, it seems that here he glosses over a genuine difficulty. There do in fact exist finitely generated discrete subgroups of $\mathrm{PGL}(2,\mathbb{C})$ having no finite fundamental polyhedron in the hyperbolic space $\mathbb{H}^3$ (see [BoOt1988]). We describe here the construction of the polygon P, the *Dirichlet polygon*, for an arbitrary Fuchsian group. It is in fact finite if the Fuchsian group is finitely generated, but we refer the reader to [Dal2007] for a proof of this additional fact.

Proof without the finiteness assertion. — Let z_0 be a point of $\mathbb{H}$ not fixed by any nontrivial element of Γ and consider the set P of points $z \in \mathbb{H}$ closer to z_0 than to all other points of the orbit $\Gamma(z_0)$:

$$P = \left\{ z \in \mathbb{H} \mid d_{\mathrm{hyp}}(z, z_0) = d_{\mathrm{hyp}}(z, \Gamma(z_0)) \right\}.$$

First of all, since Γ is Fuchsian, the orbit $\Gamma(z_0)$ is discrete, so P contains a neighborhood of z_0. Next observe that P is the intersection of the "half-planes"

$$P_i = \left\{ z \in \mathbb{H} \mid d_{\mathrm{hyp}}(z, z_0) \le d_{\mathrm{hyp}}(z, z_i) \right\}, \quad z_i \in \Gamma(z_0) \setminus \{z_0\},$$

and since each P_i is convex (with respect to the hyperbolic metric) their intersection P must also be convex, whence, in particular, connected and simply connected. The part of P contained in a hyperbolic disc of radius $r > 0$ coincides with the intersection of finitely many of the P_i, namely those corresponding to points $z_i \in \Gamma(z_0)$ contained in the disc of radius $2r$. Hence P is a polygon with piecewise geodesic boundary.

Now consider any point $z \in \mathbb{H}$. Its distance to $\Gamma(z_0)$ is attained by some point $z_1 \in \Gamma(z_0)$: indeed, since $\Gamma(z_0)$ is discrete it must be closed. Let $\varphi_1 \in \Gamma$ be an element sending z_0 to z_1; since φ_1 is an isometry, the point $z_0 \in \Gamma(z_0)$ minimises the distance to $z_1' := \varphi_1^{-1}(z)$ in $\Gamma(z_0)$, so that z_1' must belong to P. Hence $\Gamma(z) \cap P \neq \emptyset$. Finally, if $\Gamma(z)$ intersected P in at least two distinct points, say z_1' and $z_2' = \varphi_2^{-1}(z)$, then z would be equidistant from z_1 and $z_2 = \varphi_2(z_0)$, so that z_1' would be equidistant from z_0 and $\varphi_1^{-1}\varphi_2(z_0)$. Hence z_1' and z_2' must in fact both lie on the boundary ∂P. □

VI.1.4. Finite polygons

In what follows we consider only Dirichlet polygons, that is, those constructed as in Theorem VI.1.2 above.

Proposition VI.1.3. — *Let* $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$ *be a Fuchsian group and* P *a Dirichlet fundamental polygon. If* P *is finite then* Γ *is finitely generated. More precisely,* *there exists a decomposition of the oriented boundary P into an even number of oriented geodesic arcs* $\delta_1, \ldots, \delta_{2n}$*, together with a fixed-point free involutary permutation* σ *of the set* $\{1, \ldots, 2n\}$*, and generators* $\varphi_1, \ldots, \varphi_{2n}$ *for* Γ*, satisfying*

$$\varphi_i(\delta_i) = \delta_{\sigma(i)}^{-1} \quad \text{and} \quad \varphi_{\sigma(i)} = \varphi_i^{-1}, \quad i = 1, \ldots, 2n.$$

A fundamental polygon with a specified such even decomposition of its boundary is said to be *adapted* to the group Γ. Note that the assumption that P be Dirichlet is significant here: in [Bea1983, pp. 210–213] an example is given of a Fuchsian group with a convex fundamental polygon with just 5 sides, one of which is not associated with the others and is the limit of sides of infinitely many translates $\varphi(P)$, $\varphi \in \Gamma$. Thus in this example the tiling is not locally finite.

Proof. — Let $\delta_1, \ldots, \delta_p$ denote the sides of P, with the orientation inherited from that chosen for the boundary of P. Observe first that for every point $z \in \partial P$ there is at least one nontrivial transformation $\varphi \in \Gamma$ sending z to a point $z' \in \partial P$ (possibly the same). To see this, let z_0 be a point used to define P as a Dirichlet polygon (as in the statement of Theorem VI.1.2); then the distance of z from the orbit $\Gamma(z_0)$ is attained at z_0 and at least one other point $z_1 \neq z_0$ of the orbit. It then suffices to choose a transformation $\varphi \in \Gamma$ sending z_1 to z_0: the distance from $z' = \varphi(z)$ to $\Gamma(z_0)$ is then attained at both $z_0 = \varphi(z_1)$ and $\varphi(z_0)$, so that z' is on the boundary of P.

Now in the case where z and z' are "smooth" points of ∂P (that is, not vertices), φ is determined uniquely and must send germs (δ_i, z) of segments of the side containing z to germs (δ_j^{-1}, z') of the side containing z', since otherwise there would be points arbitrarily close to z each of whose orbits intersected the interior of P at least twice. The equivalence class[1] of a point $z \in \partial P$ is finite, since the sum of the corresponding angles must be $\leq 2\pi$ and the angles of ∂P are all strictly positive. (Here we are of course considering only points in $\mathbb{H}$). By regarding the points of the equivalence classes of non-smooth points of ∂P as new vertices, we obtain a new decomposition of ∂P into geodesic arcs $\delta_1, \ldots, \delta_p$, with the property that for each $i = 1, \ldots, p$ there exist a unique j and $\varphi_i \in \Gamma$ such that $\varphi_i(\delta_i) = \delta_j^{-1}$; and moreover the orbit of a point $z \in \delta_i$ other than its end-points intersects ∂P in just the two points z and $\varphi_i(z)$. It is possible that $i = j$, in which case φ_i has a fixed point in the middle of the arc δ_i; if this should occur, we add this point to the set of vertices and subdivide δ_i into two arcs to fulfil the claim of the theorem (see Figure VI.2).

To verify that the elements $\varphi_1, \ldots, \varphi_{2n} \in \Gamma$ thus constructed do indeed generate Γ, we argue in terms of the tiling as follows. Note first that the tile adjacent to P along the side δ_i is $\varphi_i^{-1}(P)$. Thus if $\varphi(P)$ and $\varphi'(P)$ (with $\varphi, \varphi' \in \Gamma$) are

[1] That is, the set $\Gamma(z) \cap P$. *Trans*

two tiles sharing the side $\varphi'(\delta_i)$, we must have $\varphi'^{-1} \circ \varphi(P) = \varphi_i^{-1}(P)$, whence $\varphi' = \varphi \circ \varphi_i$. Now given any element $\varphi \in \Gamma$, we choose a path γ joining P to the tile $\varphi(P)$ avoiding all the vertices of the tiling (that is, all the translates by Γ of the vertices of P). Suppose the path γ successively traverses the tiles

$$P_0 = P,\ P_1,\ P_2,\ \dots\ ,P_N = \varphi(P).$$

If it enters P_k across the side δ_{i_k} (that is, the side of P_k sent to δ_{i_k} by a (unique) element of Γ), $k = 1, \dots, N$, then it follows immediately that $P_k = \varphi_{i_1} \circ \varphi_{i_2} \circ \cdots \circ \varphi_{i_k}(P)$, whence $\varphi = \varphi_{i_1} \circ \varphi_{i_2} \circ \cdots \circ \varphi_{i_N}$. □

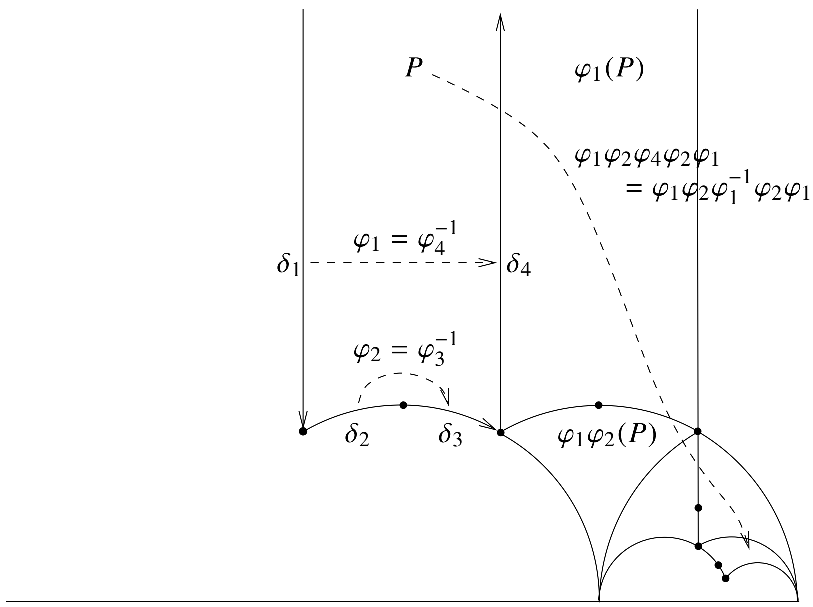


Figure VI.2: A generating system for PSL$(2,\mathbb{Z})$

From this proof we see that a shortest word in the generators $\varphi_1, \dots, \varphi_{2n}$ representing a given element φ of Γ corresponds to a path from P to $\varphi(P)$ meeting the least possible number of tiles.

Henceforth we shall consider only adapted fundamental polygons with sides understood to be the δ_i of the above proposition.

VI.1.5. The angle of an elliptic cycle and relations

The group Γ determines an equivalence relation on the boundary ∂P of the fundamental polygon, and, more particularly, on the set of its vertices. We shall call the equivalence class of a vertex under this relation a *cycle*, and define the *angle of a cycle* to be the sum of the angles at the vertices of the cycle in question. We shall ignore for the moment sides that go to infinity and do not intersect in $\mathbb{H}$: we are concerned for the time being only with the vertices (that is, at finite distances). We index the sides δ_i and vertices s_i of P cyclically, so that $s_i = \delta_i \cap \delta_{i+1}$ (assuming always that the polygon is connected and simply connected). For non-compact P, the set of indices of the vertices is a proper subset of $\{1,\ldots,2n\}$. We denote by $\tilde{\sigma} \in \mathrm{Perm}\{1,\ldots,2n\}$ the permutation defined by $\tilde{\sigma}(i) := \sigma(i) - 1$ for $i = 1,\ldots,2n$ (with the convention $0 \equiv 2n$) where σ is the permutation given in Proposition VI.1.3. It is then immediate that the cycle of a vertex s_i is

$$s_i, \quad s_{\tilde{\sigma}(i)}, \quad s_{\tilde{\sigma}^2(i)}, \quad \ldots \quad , s_{\tilde{\sigma}^l(i)}$$

where $l \in \mathbb{N}$ is the least number such that $\tilde{\sigma}^{l+1}(i) = i$.

Proposition VI.1.4. — *The angle of each cycle is an integer fraction of 2π. Furthermore, in the notation introduced above, if the angle of the cycle of the vertex s_i is $\frac{2\pi}{q}$, $q \in \mathbb{N}^*$, then the following relation holds:*

$$(\varphi_{\tilde{\sigma}^l(i)} \circ \cdots \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i)^q = \mathrm{id}\,.$$

Every relation between the generators $\varphi_1,\ldots,\varphi_{2n}$ follows from these relations (with s_i ranging over a system of representatives of cycles) together with the relations $\varphi_{\sigma(i)} = \varphi_i^{-1}$.

We have thus obtained an explicit *presentation* of the group Γ in terms of n generators and r relations, where $2n$ is the number of sides of the fundamental polygon P and r the number of finite-distance cycles.

Proof. — Let s_i be a vertex of P, and consider the tiling of $\mathbb{H}$ by the $\varphi(P)$, $\varphi \in \Gamma$. The transformation φ_i sends the side δ_i to its conjugate $\delta^{-1}_{\sigma(i)}$, and in particular s_i to $s_{\sigma(i)-1}$: the tile $P_1 := \varphi_i^{-1}(P)$ is the one encountered on leaving the polygon P across the side δ_i. In turning about s_i, we successively encounter

$$P_0 = P,$$

$$P_1 = \varphi_i^{-1}(P),$$

$$P_2 = (\varphi_{\tilde{\sigma}(i)} \circ \varphi_i)^{-1}(P),$$

$$P_3 = (\varphi_{\tilde{\sigma}^2(i)} \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i)^{-1}(P), \quad \ldots .$$

Writing $\varphi := \varphi_{\tilde{\sigma}^l(i)} \circ \cdots \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i$, we have $\varphi(s_i) = s_i$. Hence φ is an elliptic transformation, and since Γ is discrete, there exists an integer q for which $\varphi^q(P) = P$, so that $\varphi^q = \mathrm{id}$. Writing $Q = P_0 \cup P_1 \cup P_2 \cup \ldots \cup P_l$, we see that the polygons $Q, \varphi(Q), \ldots, \varphi^{q-1}(Q)$ have pairwise disjoint interiors and cover a neighborhood of s_i. The angle at the vertex s_i of the polygon Q is therefore of size $\frac{2\pi}{q}$, and is equal to the angle of the cycle containing s_i. Moreover Q and $\varphi^{-1}(Q)$ are, by construction, adjacent along $\varphi^{-1}(\delta_{i+1})$, so the angle of the rotation φ is $\frac{2\pi}{q}$. Note that the element φ generates the stabilizer of s_i in Γ.

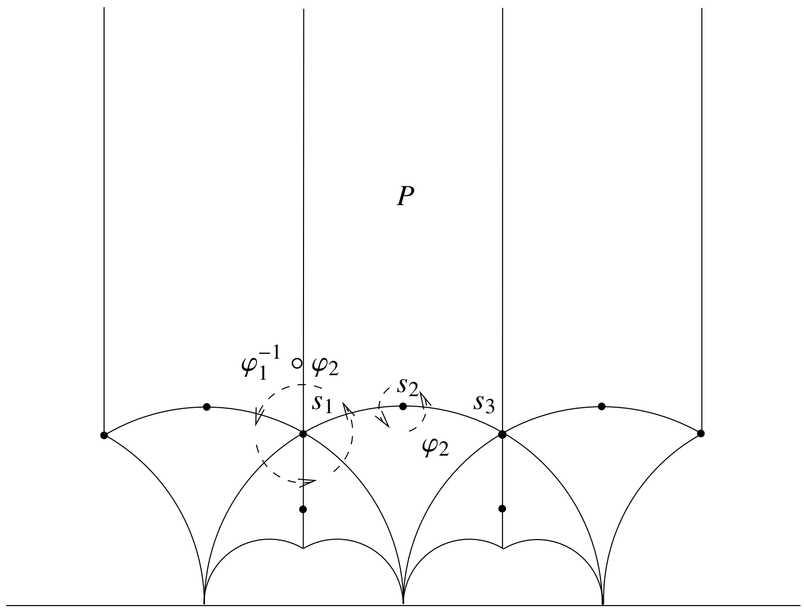


Figure VI.3: Two cycles of angles $\frac{2\pi}{3}$ and $\frac{2\pi}{2}$ for PSL$(2,\mathbb{Z})$

It remains to show that every relation among the generators φ_i is a consequence of the relations given by the cycles, just discussed, together with the relations $\varphi_i^{-1} = \varphi_{\sigma(i)}$. To this end, we denote by G the group generated by $2n$ generators $a_1,\dots,a_{2n}$, say, subject to the relations $a_{\sigma(i)} = a_i^{-1}$, $i = 1,\dots,2n$, together with others corresponding to the cycles, and define a group morphism $\rho : G \to \Gamma$ by setting $a_i \mapsto \varphi_i$, with kernel denoted by N. We need to prove that in fact N is trivial.

With this in view, we introduce the space $\mathcal{H} = P \times G$ endowed with the product topology, with the topology on G taken to be discrete, and consider the equivalence relation on this space generated by

$$(z,g) \sim (z',g') \quad \text{whenever there exists } 1 \le i \le 2n,\ z' = \varphi_i(z) \text{ and } g = g' \cdot a_i.$$

Denote by $\mathcal{H}^*$ the topological quotient space of $\mathcal{H}$ by this equivalence and by π

the natural projection of $\mathcal{H}$ onto $\mathcal{H}^*$. One easily checks that this map is proper[2].

The group G acts on $\mathcal{H}$ in the obvious way: each element $g \in G$ determines a homeomorphism

$$\tau_g : (z, g') \in \mathcal{H} \mapsto (z, g \cdot g') \in \mathcal{H}.$$

This action is proper[3] and free, and has $P \times \{\mathrm{id}\}$ as a fundamental region. This action of G on $\mathcal{H}$ induces an action of G on the quotient $\mathcal{H}^*$, also proper since the projection $\pi : \mathcal{H} \to \mathcal{H}^*$ is proper. On the other hand, if $(z, g) \in \mathcal{H}$ is equivalent to (z, g'), where $g \neq g'$, then z must of necessity be a vertex s_i, and we then have $g' = g a_{\tilde{\sigma}^k(i)}, \ldots, a_{\tilde{\sigma}(i)} a_i$, with $0 \leq k < (l_{s_i} + 1) q_{s_i}$, where $\frac{2\pi}{q_{s_i}}$ is the angle of the cycle containing s_i and $l_{s_i} + 1$ the number of vertices in this cycle. Since for $0 \leq k < (l_{s_i} + 1) q_{s_i}$ we have $\rho(a_{\tilde{\sigma}^k(i)} \ldots a_{\tilde{\sigma}(i)} a_i) \neq \mathrm{id}$, we infer that for all nontrivial $\nu \in N$, (z, g) is not equivalent to $(z, \nu g)$. Hence the action of N on $\mathcal{H}^*$ is free (as well as proper).

We now introduce the map $p : \mathcal{H} \to \mathbb{H}$ defined by $p(z, g) = \rho(g)(z)$. This map preserves the equivalence relation $\sim$ and therefore induces a continuous map $p^* : \mathcal{H}^* \to \mathbb{H}$. Note that p^* is a local homeomorphism. To see this, let s be a vertex of ∂P, $l_s + 1$ the size of the cycle to which s belongs, and $\frac{2\pi}{q_s}$ the angle of this cycle. Then on taking the union of P with the $\varphi_i(P)$, $i = 1, \cdots, 2n$, and the $\varphi_{\tilde{\sigma}^k(i)} \circ \ldots \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i(P)$, $0 \leq k < (l_{s_i} + 1) q_{s_i}$, where s_i ranges over the vertices of P, we obtain a neighborhood of P. Hence the projection on $\mathcal{H}^*$ of $P \times \{\mathrm{id}\}$, together with the $P \times \{a_i\}$, $i = 1, \cdots, 2n$, and the $P \times \{a_{\tilde{\sigma}^k(i)} \ldots a_{\tilde{\sigma}(i)} a_i\}$, $0 \leq k < (l_{s_i} + 1) q_{s_i}$ yields a neighborhood W of $\pi(P \times \{\mathrm{id}\})$ and p^* is a homeomorphism from W to its image. It follows further that for all $g \in G$, gW is a neighborhood of $\pi(P \times \{g\})$, and p^* is a homeomorphism from gW to its image. Finally, observe that since the φ_i generate Γ and $\bigcup_{\varphi \in \Gamma} \varphi(P) = \mathbb{H}$, the map p^* is surjective.

We shall now show that the fibres of p^* are precisely the orbits of the action of N on $\mathcal{H}^*$. Clearly, those orbits are contained in the fibres. Now if $p(z, g) = p(z', g')$, then $\rho(g)(z) = \rho(g')(z')$, implying in turn that $(z, g) \sim (z', ga)$ where $a = g^{-1} g'$. Hence $\rho(ga)(z') = \rho(g')(z')$. If z' is not a vertex of P, we obtain directly the existence of a $\nu \in N$ such that $ga = \nu g'$. In other words, $\pi(z, g) = \pi(z', \nu g')$, so $\pi(z, g)$ is certainly in the same N-orbit as $\pi(z', g')$. If z' is a vertex, s_i, say, then there will exist $0 \leq k < q_{s_i}$ such that $ga = \nu g' (a_{\tilde{\sigma}^{l_{s_i}}(i)} \ldots a_{\tilde{\sigma}(i)} a_i)^k$. Since (z', g') is equivalent to $(z', g' (a_{\tilde{\sigma}^{l_{s_i}}(i)} \ldots a_{\tilde{\sigma}(i)} a_i)^k)$, we have that $\pi(z, g)$ and $\pi(z', g')$ are in the same N-orbit. Thus we conclude that $\mathcal{H}^*/N$ is homeomorphic to $\mathbb{H}$. Then since $\mathbb{H}$ is simply connected, N must be trivial, and in fact G is isomorphic to Γ. □

---

[2] That is, the preimage of every compact set is compact.

[3] An action $a : G \times X \to X$ is *proper* if the map $a \times \mathrm{pr}_2 : G \times X \to X \times X$ is proper.

Remark VI.1.5. — We mention by the way the fact that if Γ contains a nontrivial element φ fixing a point $z_0 \in \mathbb{H}$, then z_0 is in the orbit of a cycle of P and $\varphi^q = \mathrm{id}$ where $\frac{2\pi}{q}$ is the angle of the cycle. In particular, φ is conjugate in Γ to an element (of order q) of the isotropy group of z_0. Hence the cycles of angle $< 2\pi$ are in one-to-one correspondence with the conjugacy classes of maximal elliptical subgroups of Γ.

VI.1.6. Cycles at infinity

We shall now consider the intersection of the closure of P with *the circle at infinity* $\partial\mathbb{H}$. We denote by $\overline{P}$ the closure of P in $\overline{\mathbb{H}} = \mathbb{H} \cup \partial\mathbb{H}$. The boundary of P at infinity $\overline{P} \cap \partial\mathbb{H}$ decomposes into a finite number (possibly zero) of points and closed intervals in $\partial\mathbb{H}$; we call the isolated such points and the endpoints of the intervals contained in $\overline{P} \cap \partial\mathbb{H}$ the *vertices at infinity* of P. Recall that we are assuming P to be convex, so that its closure $\overline{P}$ is also. Extending the convention adopted above, we denote by s_i the connected component of the boundary of P at infinity connecting the side δ_i to the side δ_{i+1}. Once again the group Γ induces an equivalence relation on the boundary of P at infinity, or, more specifically, on the set of connected components of that boundary. The *cycles at infinity* of P are then the latter equivalence classes. Such a cycle will be called *parabolic* if it consists only of isolated points, and *hyperbolic* if it also contains intervals of positive length. As for finite-distance cycles, one may consider the isotropy subgroup of a vertex $x = s_i$ contained in a parabolic cycle. A nontrivial element of this group is then $\varphi := \varphi_{\tilde{\sigma}^l(i)} \circ \cdots \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i$ where $l \in \mathbb{N}$ is the smallest natural number for which $\tilde{\sigma}^{l+1}(i) = i$.

Proposition VI.1.6. — *If the vertex $x \in \overline{P} \cap \partial\mathbb{H}$ belongs to a parabolic cycle, then the element $\varphi \in \Gamma$ indicated above is parabolic and generates the isotropy group of x.*

Proof. — By suitably dissecting and reassembling the fundamental polygon we can arrange for the various ends corresponding to the cycle in question to coincide; in this way we are reduced to the case of a fundamental polygon where the parabolic cycle of interest consists of a single point x. The two sides of P adjacent to x are then sent one to the other by φ so that of course $\varphi(x) = x$.

To see that φ is parabolic, we consider the two geodesics γ_0 and γ_1 bounding P in a neighborhood of x and such that $\varphi(\gamma_0) = \gamma_1$. Supposing φ hyperbolic, consider the position of its second fixed point y relative to these two geodesics. If y, or (what comes to the same thing) the geodesic γ joining x to y, is between γ_0 and γ_1, one readily sees (by considering the germ of P at x along γ) that the intersection $P \cap \varphi(P)$ must have non-empty interior, contradicting the fact that P

is a fundamental region. Hence γ cannot lie between γ_0 and γ_1. By replacing φ by φ^{-1} if necessary, we may suppose that x is repulsive and y attractive in order that the sequence of geodesics $\gamma_n := \varphi^n(\gamma_0)$ tend to γ. The side of P along γ_0 will then be sent by φ^n to a side of $P_n = \varphi^n(P)$ along "larger and larger" portions of γ_n, that is, tending to the whole of γ. Hence the tiling cannot be locally finite in a neighborhood of γ, giving a contradiction.

The preceding argument shows in fact that the isotropy group contains only parabolic elements. That isotropy group is therefore contained in a one-parameter group $\{\psi^t\}$ of parabolic elements, where $\varphi = \psi^1$, say. If the isotropy group contained an element not a power of φ, then on multiplying it by φ, we would obtain an element $\tilde{\varphi} = \psi^{t_0}$, $0 < t_0 < 1$, sending γ_0 to a geodesic lying strictly between γ_0 and γ_1, from which it would follow that $P \cap \tilde{\varphi}(P)$ had nonempty interior. This contradiction shows that in fact the isotropy group of x is generated by φ. □

VI.1.7. Orbifolds and Riemann surfaces

If a discrete group Γ acts properly on a Riemann surface S, the quotient space S/Γ can be endowed with the structure of a Riemann surface in such a way that the projection $S \to S/\Gamma$ is holomorphic. If we are given that the action is free and S simply connected, then Γ and S and the action of Γ on S can actually be retrieved from knowledge of S/Γ; in fact S is just the universal covering space of S/Γ and Γ the fundamental group of S/Γ acting via covering transformations. However, if the action is not free, this is no longer the case. For example, as we have seen in Chapter V, the quotient of $\mathbb{H}$ by the action of $\mathrm{PSL}(2,\mathbb{Z})$ is a Riemann surface isomorphic to $\mathbb{C}$, which is simply connected. In this case, in order to recover Γ and its action on S from S/Γ, one needs supplementary information about the latter Riemann surface. This leads to the concept of a 2-dimensional "orbifold", originally introduced in all dimensions by Satake [Sat1956] under the name of "V-manifold", and popularized by Thurston [Thu1980, Chapter 13] under the name orbifold[4]. Here we shall rest content with a naive approach, close to that of Poincaré.

Thus for us an orbifold is a Riemann surface X on which there is specified a family (x_i) of isolated points assigned integer weights $n_i \geq 2$. One sometimes hears the points x_i called *ramification points* of the orbifold and the n_i their *multiplicities*. When a group Γ acts properly on a Riemann surface S (preserving orientation), the quotient S/Γ naturally acquires the structure of an orbifold. Each point of S with nontrivial (finite cyclic) stabilizer defines a ramification point of S/Γ whose multiplicity is the order of the stabilizer. The quotient S/Γ is then called a *quotient orbifold*.

With two exceptions, described below, every orbifold X determines a unique

[4]The corresponding object of algebraic geometry bears the name "stack".

proper action of a group Γ on a simply connected Riemann surface S such that X is isomorphic to S/Γ. The Riemann surface S is called the universal cover of the orbifold X and Γ its fundamental group. The exceptions referred to above are just the sphere with either one ramification point or two ramification points of different multiplicities. Here are examples that we have already encountered (see Chapter V): the quotient of $\mathbb{H}$ by PSL$(2,\mathbb{Z})$ is $\mathbb{C}$ with two ramification points of multiplicities 2 and 3; the quotient of $\mathbb{CP}^1$ by the icosahedral group is $\mathbb{CP}^1$ with three ramification points of multiplicities 2, 3 and 5; and the quotient of $\mathbb{H}$ by $\Gamma_0(7)$ is $\mathbb{C}^*$, with two ramification points of multiplicity 3 (see §V.1).

Thus orbifolds represent a generalization of Riemann surfaces. One can define the concepts of a holomorphic mapping between two orbifolds, a covering of orbifolds, and so on. If X is a compact orbifold, with ramification points x_i of multiplicities $n_i \geq 2$, its orbifold Euler–Poincaré characteristic χ_{orb} is defined to be $\chi_{\text{orb}}(X) = \chi(X) + \sum(1/n_i - 1)$. This definition is dictated mainly by the fact that if $X_1 \to X_2$ is a covering map of degree d, then $\chi_{\text{orb}}(X_1) = d\chi_{\text{orb}}(X_2)$, a version of the Riemann–Hurwitz relations[5].

VI.1.8. Quotients viewed as Riemann surfaces, then as orbifolds

Consider now the quotient

$$\pi : \mathbb{H} \to S := \mathbb{H}/\Gamma.$$

The structure of S as Riemann surface can be described as follows. As topological space S is homeomorphic to the quotient of the fundamental polygon P by the relation identifying each side δ_i with its conjugate $\delta^{-1}_{\sigma(i)}$; thus in particular each finite-distance cycle corresponds to a single point of S. It is also quite natural to consider the compactification $\overline{S}$ obtained by making the same identifications of $\overline{P}$.[6] A cycle at infinity then projects to a point or a circle according as it is parabolic or hyperbolic. The map $\pi : \mathbb{H} \to S$ is a ramified covering; or, more precisely, it is totally ramified above each cycle of angle $\frac{2\pi}{q}$ with $q > 1$, and, at each point of the fibre π can be expressed in the form $z \mapsto z^q$ in terms of local complex coordinates; and elsewhere the covering π is regular.

Proposition VI.1.7. — *The genus g of the Riemann surface $S = \mathbb{H}/\Gamma$ is equal to $\frac{n+1-c}{2}$ where $2n$ is the number of sides of the polygon P and c the number of cycles, both infinite and finite.*

---

[5]Another interpretation involves understanding an "ordinary" point as having characteristic 1 and an orbifold point characteristic $\frac{1}{n}$.

[6]One must take into account here that $\overline{\mathbb{H}}/\Gamma$, endowed with the quotient topology, is not Hausdorff if Γ is infinite.

Proof. — The genus of S is by definition that of the compact surface without boundary $S' = \overline{S} \cup \bigcup D_i$, where the D_i are discs attached to the connected components $\partial_i \overline{S}$ of the boundary of $\overline{S}$ via identifications between the $\partial_i \overline{S}$ and the ∂D_i.

Denote by $c_{\rm hyp}$ (resp. $c_{\rm par}$, $c_{\rm ell}$) the number of hyperbolic (resp. parabolic, elliptic) cycles of Γ. We then have

$$2 - 2g = \chi(S') = \chi(\overline{S}) + c_{\rm hyp} = \chi(S) + c_{\rm par} + c_{\rm hyp},$$

and also $\chi(S) = c_{\rm ell} - n + 1$, by considering the images of the vertices and sides of P in its quotient S. Hence $2 - 2g = 1 - n + c$ as claimed. □

Since Γ acts by conformal transformations, one can endow S with a complex structure in such a way that π is holomorphic; a conformal local coordinate w at a point $\pi(z_0)$ is given by the formula $w = (z - z_0)^q$, where q is the order of the isotropy subgroup of z_0 in Γ. At a parabolic end, for instance $z_0 = \infty$ with isotropy group generated by $\varphi(z) = z + 1$, the function $w = \exp(2i\pi z)$ projects to a conformal local coordinate in a neighborhood of the corresponding point of the compactification $\overline{S}$. Hence in the case when the group Γ has no hyperbolic cycles, the Riemann surface $\overline{S}$ so defined is compact without boundary; the surface S is then obtained by removing from $\overline{S}$ a finite number of points, one for each parabolic cycle.

When the group Γ has one or more hyperbolic cycles, the compactification $\overline{S}$ is naturally endowed with the structure of a Riemann surface *with boundary*.

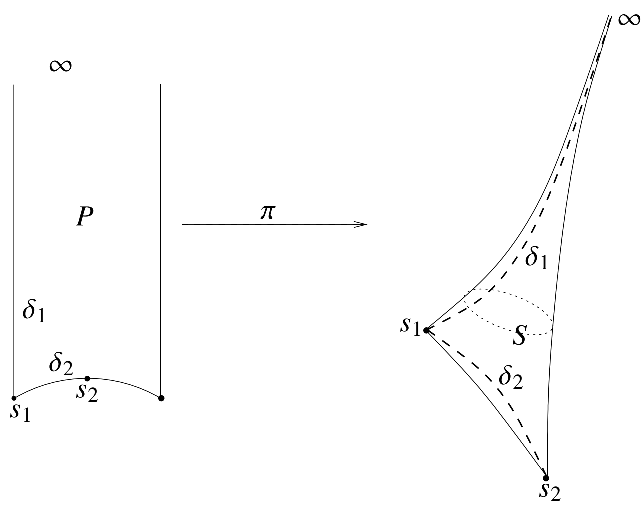


Figure VI.4: The quotient orbifold $\mathbb{H}/\mathrm{PSL}(2,\mathbb{Z})$

The hyperbolic metric on $\mathbb{H}$ is invariant under the action of Γ and therefore π induces a metric of constant curvature -1 with orbifold singularities on the quotient S: corresponding to each elliptical cycle one has an orbifold (or conical) singularity with angle that of the cycle. In this way one may view $\pi : \mathbb{H} \to S$ as providing an "orbifold uniformization" of S. As far as parabolic ends are concerned, it is natural to consider them as orbifold points of angle zero in $\overline{S}$.

The area of a hyperbolic triangle with angles measuring α, β and γ is

$$\pi - (\alpha + \beta + \gamma),$$

even if one or more of the angles is zero. Hence, in particular, parabolic ends have finite area, or, to be more precise, by dissecting the polygon into triangles, and calculating directly, one obtains:

Proposition VI.1.8. — *If the polygon P is without hyperbolic ends, then its hyperbolic area is finite, given by*

$$\text{area}(P) = (2n - 2)\pi - \alpha,$$

where $2n$ is the number of sides of P and $\alpha = \sum_i \alpha_i$, the sum of the angles of P over all its vertices.

It follows that the orbifold Euler–Poincaré characteristic of the quotient S is given by $\chi_{\text{orb}} = -\text{area}(P)/2\pi$ — a particular case of the Gauss–Bonnet theorem. Since the polygon P has area > 0, we conclude that the Euler–Poincaré characteristic of S is necessarily < 0.

If the polygon P has a hyperbolic end its area is infinite. To give a fundamental polygon for the action of a group Γ, is to give a "tessellation" of, or piecewise geodesic graph on, the quotient orbifold S whose vertices include all the orbifold points and whose complement is connected and simply connected. Hence one may construct new fundamental polygons by modifying or deforming this graph. For instance, one can always modify the graph in such a way that a given orbifold point becomes a vertex on a single side; the corresponding cycle of the new fundamental polygon is thus reduced to a single vertex. In the particular case of a cycle of angle 2π, it becomes an interior point of the new fundamental polygon. One cannot, however, use this device to eliminate all cycles of angle 2π — for example, in the case of a compact surface (without orbifold points). All the same, they can be moved about relatively freely in the course of deforming the fundamental polygon.

Remark VI.1.9. — Given any Fuchsian group Γ (not necessarily finitely generated), one can form the quotient $\mathbb{H}/\Gamma$ and endow this with an orbifold structure rendering the projection a universal orbifold covering map. The number of orbifold points or ends may then be infinite. In fact it follows from the topological classification of connected oriented surfaces[7] that the (orbifold) fundamental group of such a surface is finitely generated if and only if it has a finite number of orbifold points and ends. Thus the quotient of $\mathbb{H}$ by a finitely generated subgroup Γ of $\mathrm{PSL}(2,\mathbb{R})$ must be geometrically finite; it is then easy, starting from a geodesic triangulation, to infer the existence of a finite fundamental polygon.

VI.1.9. The polygon theorem

Hitherto we have always started with a Fuchsian group Γ and constructed from it a fundamental polygon. We now reverse this point of view and look for conditions on a polygon for it to be the fundamental polygon of a Fuchsian group.

Theorem VI.1.10. — *Let $P \subset \mathbb{H}$ be a connected and simply connected polygon with boundary consisting of an even number of geodesic arcs $\delta_1, \ldots, \delta_{2n}$ in cyclic order. Suppose given the following:*

— *an involutary fixed-point free permutation σ of $\{1, \ldots, 2n\}$ allowing the pairwise identification of the arcs;*

— *for each i, a transformation $\varphi_i \in \mathrm{PSL}(2,\mathbb{R})$ sending δ_i onto $\delta_{\sigma(i)}^{-1}$ and satisfying $\varphi_{\sigma(i)} = \varphi_i^{-1}$.*

Suppose in addition that:

— *the angle of each finite-distance cycle is an integer fraction of 2π;*

— *for each parabolic cycle at infinity the corresponding "return" transformation φ defined in §VI.1.6 is parabolic.*

Then the group Γ generated by $\varphi_1, \ldots, \varphi_{2n}$ is Fuchsian and P is a fundamental polygon for Γ.

Remark VI.1.11. — Given two oriented geodesic arcs δ and δ', there exists an element $\varphi \in \mathrm{PSL}(2,\mathbb{R})$ sending δ onto δ' (and matching the orientations) if and only if one of the following conditions holds:

— δ and δ' are of the same finite length;

— δ and δ' are either both future half-geodesics or both past half-geodesics;

— δ and δ' are both complete geodesics.

[7] This classification came later than the work of Poincaré we are concerned with here; see the introduction to the final part (Part C).

The transformation φ is unique except in the third case where it is defined only modulo the action of a one-parameter group. Consequently, if the polygon of the above theorem has no "doubly infinite" sides, then, once the involution σ is given, it determines the group Γ uniquely. Hence for instance in the case of a compact polygon, the theorem can be stated without reference to the φ_i, requiring only that for each i the sides δ_i and $\delta_{\sigma(i)}$ have the same length. In the non-compact case the parabolicity condition has to be appropriately translated. As Proposition VI.1.6 shows, this condition is necessary. For example, if one chooses to identify the two sides of the polygon $\{(x, y) \in \mathbb{H} \mid \frac{1}{2} \leq x \leq 1\}$ using the hyperbolic transformation $\varphi(z) = \frac{z}{2}$, then the associated tiling covers only a quarter of the plane $\{x, y > 0\}$. In his first notes Poincaré omits this needed assumption.

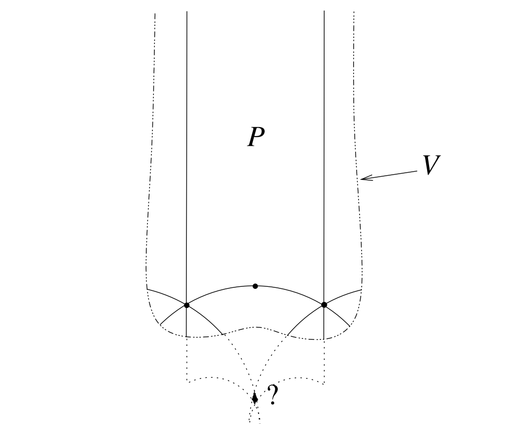


Figure VI.5: The tiled neighborhood V

Proof of Theorem VI.1.10. — It suffices to show that the polygons $\varphi(P)$, $\varphi \in \Gamma$ constitute a tiling of the half-plane $\mathbb{H}$, that is:

— the union of the $\varphi(P)$ covers $\mathbb{H}$;

— the intersection of two translates $\varphi(P)$ and $\psi(P)$ is empty or a union of sides (or else vertices) or $\varphi(P) = \psi(P)$ in which case $\varphi = \psi$.

Observe first that the condition on the elliptic cycles at least allows one to tile a neighborhood of P in $\mathbb{H}$. This is done by first attaching the germ of $P_i = \varphi_i^{-1}(P)$ to the side δ_i. Next one fills in the region around each (finite-distance) vertex s_i using a sequence

$$P_{i,j} = \varphi_{i,j} \circ \cdots \circ \varphi_{i,1}(P), \quad j = 1, \cdots, k_i$$

of tiles, as in Proposition VI.1.4. (Here one considers the germ of each $P_{i,j}$ in a neighborhood of the vertex s_i, except for $j = 1$ and $j = k_i$ where one considers germs neighbouring the sides δ_i and δ_{i+1}). Let V be a neighborhood of P for which the tiling so constructed is well defined. In order to appreciate the difficulty and thus the force of this theorem, note that *a priori* this neighborhood of P is tiled only by germs (see Figure VI.5).

We shall now construct a global tiling of a surface covering $\mathbb{H}$. To that end we again employ the construction used in the proof of Proposition VI.1.4. The group G, the morphism $\rho : G \to \Gamma$, the spaces $\mathcal{H}$ and $\mathcal{H}^*$, and the maps $p : \mathcal{H} \to \mathbb{H}$ and $p^* : \mathcal{H}^* \to \mathbb{H}$ are all as in that proof.

The action of G on $\mathcal{H}^*$ is discrete and $\pi(P \times \{\mathrm{id}\})$ is a fundamental region for that action. To obtain the polygon theorem we need to show that p^* is a homeomorphism between $\mathcal{H}^*$ and $\mathbb{H}$. Since we do not know *a priori* that the $\varphi(P)$, $\varphi \in \Gamma$, tile $\mathbb{H}$, it is no longer clear that p^* is surjective. Nonetheless it is still true that p^* defines a local homeomorphism from $\mathcal{H}^*$ to $\mathbb{H}$. We verify this first in a neighborhood of $\pi(P \times \{\mathrm{id}\})$. Just as we constructed above the tiled neighborhood V of P in $\mathbb{H}$, we now construct the analogous neighborhood W of $P \times \{\mathrm{id}\}$ in $\mathcal{H}$ by adjoining the corresponding germs $P \times \{g\}$ obtained by replacing the φ_i by the a_i. Write $U = \pi(W)$, the projection of W on $\mathcal{H}^*$. The map $p^*|_U : U \to V$ is a homeomorphism by construction. Noting that $p^* \circ \tau_g = \rho(g) \circ p^*$ for all $g \in G$, we see that $\tau_g(U)$ defines a neighborhood of $\pi(P \times \{g\})$ and p^* restricts to a homeomorphism from $\tau_g(U)$ onto $\rho(g)(V)$.

In order to show that the local homeomorphism $p^* : \mathcal{H}^* \to \mathbb{H}$ is in fact a global homeomorphism, it is enough to prove that p^* is a covering map from $\mathcal{H}^*$ onto $\mathbb{H}$. We establish this by showing that p^* has the homotopy lifting property for paths. It is in this connection that the condition on parabolic cycles comes into play. Thus let s_i be a vertex of $\overline{P}$ belonging to a parabolic cycle with $n_i + 1$ vertices at infinity. By gluing together germs at s_i of the $\varphi_{\tilde{\sigma}^k(i)} \circ \cdots \circ \varphi_{\tilde{\sigma}(i)} \circ \varphi_i(P)$, $0 \leq k \leq n_i$, one obtains a tiling of an angular sector C_i lying between two geodesic arcs α and β issuing from s_i. The return map φ, which generates the isotropy group of s_i, is a parabolic transformation sending α onto β. By considering the union of the $\varphi^m(C_i)$ as m ranges over $\mathbb{Z}$, one can tile the whole of the interior of a horosphere based at s_i. This procedure is repeated for each parabolic vertex of $\overline{P}$, and the interiors of the horospheres thus tiled are adjoined to the region V, yielding a new neighborhood V' of P. The analogous construction on $\mathcal{H}$ with the φ_i replaced by the a_i, yields a neighborhood W' of $P \times \{\mathrm{id}\}$, and then by projecting to $\mathcal{H}^*$, we obtain a new neighborhood U' of $\pi(P \times \{\mathrm{id}\})$ on which $p^* : \mathcal{H}^* \longrightarrow \mathbb{H}$ is still injective.

Now for the neighborhood V', there exists a $\varepsilon > 0$ such that for all $z \in P$, the hyperbolic disc with centre z and radius ε is contained in V'. This implies that the open set $\bigcup_{\varphi \in \Gamma} \varphi(P)$ coincides with its ε-neighborhood in $\mathbb{H}$, or, in other words, that the map p^* is surjective. Let h^* be the metric on $\mathcal{H}^*$ obtained by lifting the

Poincaré metric via p^*. Since p^* is a homeomorphism between U' and V', we have that for each $a \in \pi(P \times \{\mathrm{id}\})$, p^* determines an isometry between $D^*(a, \epsilon)$, the disc centred at a and of radius ε with respect to the metric h^*, and its image $D(p^*(a), \varepsilon)$. Since G acts isometrically on $\mathcal{H}^*$, we conclude that every h^*-disc centred at a point of $\mathcal{H}^*$ and of radius ε is sent isometrically by p^* onto its image. It is then easy to show that every path in $\mathbb{H}$ lifts to $\mathcal{H}^*$, so that p^* is indeed a covering map. □

VI.2. Examples

We now apply the preceding theorem to the construction of certain Fuchsian groups and thence of uniformizable orbifolds.

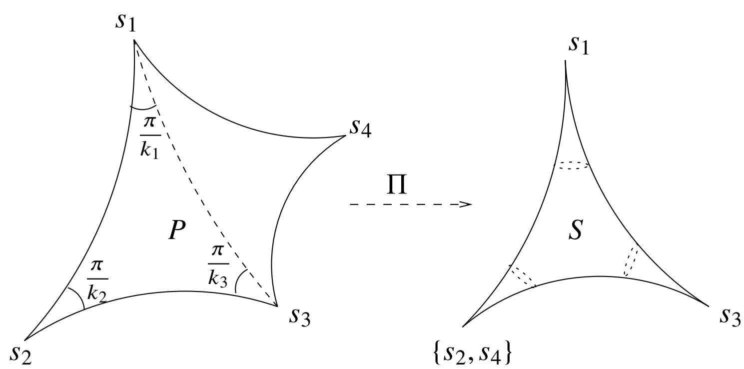


Figure VI.6: Hyperbolic triangles and orbifold spheres

VI.2.1. The sphere with 3 orbifold points

Consider a hyperbolic triangle T with vertices s_1, s_2 and s_3, and angles

$$\alpha_1 = \frac{\pi}{k_1}, \quad \alpha_2 = \frac{\pi}{k_2} \quad \text{and} \quad \alpha_3 = \frac{\pi}{k_3},$$

with $k_i \in \mathbb{N}^* \cup \{\infty\}$. If $k_i = \infty$, then $s_i \in \partial\mathbb{H}$. We denote by σ_i the reflection in the side $s_j s_k$, $\{i, j, k\} = \{1, 2, 3\}$. These three reflections generate a discrete group of isometries of $\mathbb{H}$. The subgroup Γ generated by

$$\varphi_i = \sigma_{i+1} \circ \sigma_{i+2}, \quad i = 1, 2, 3 \bmod 3$$

has index 2: it is the subgroup of orientation-preserving transformations. The group Γ is Fuchsian, with fundamental region $P := T \cup \sigma_2(T)$, for example. The vertices of P are then s_1, s_2, s_3 and $s_4 = \sigma_2(s_2)$. Writing δ_i for the geodesic arc $s_i s_{i+1}$, we have

$$\varphi_1(\delta_1) = \delta_4, \quad \varphi_3(\delta_3) = \delta_2 \quad \text{and} \quad \varphi_1 \circ \varphi_2 \circ \varphi_3 = \text{id}.$$

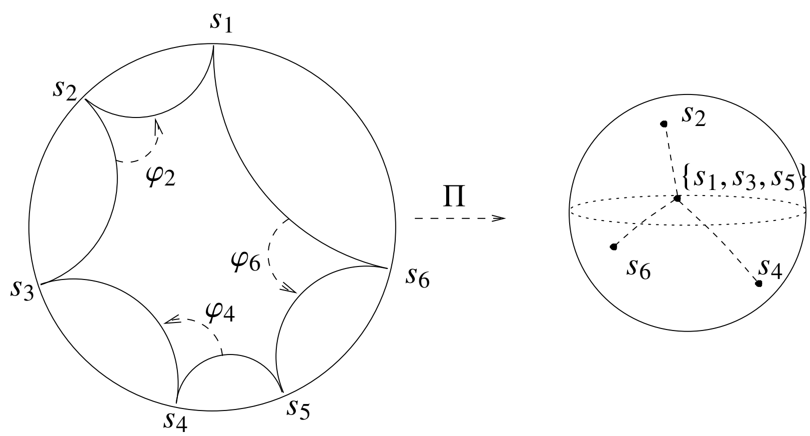


Figure VI.7: The sphere with 4 points removed

The cycles are

$$\{s_1\}, \quad \{s_2, s_4\}, \quad \{s_3\}$$

with angles respectively

$$\frac{2\pi}{k_1}, \quad \frac{2\pi}{k_2}, \quad \frac{2\pi}{k_3},$$

and with isotropy groups generated respectively by

$$\varphi_1, \quad \varphi_2 = (\varphi_3\varphi_1)^{-1}, \quad \varphi_3.$$

The group relations are just

$$\varphi_i^{k_i} = \text{id}$$

for each $i = 1, 2, 3$ for which k_i is finite. When $k_i = \infty$, one verifies that the parabolicity condition holds for the cycle associated with the vertex s_i, $i = 1, 2, 3$. The surface $\overline{S}$ is compact, of genus 0, and has 3 orbifold points (possibly of angle zero). Every sphere with 3 orbifold points is obtained in this way provided $\frac{1}{k_1} + \frac{1}{k_2} + \frac{1}{k_3} < 1$, that is,

$$\operatorname{area}(T) = \pi - \left(\frac{2\pi}{k_1} + \frac{2\pi}{k_2} + \frac{2\pi}{k_3}\right) > 0\,,$$

which is just the necessary and sufficient condition for a hyperbolic triangle with these angles to exist.

VI.2.2. The Riemann sphere with $n+1$ points removed

Next we consider an n-sided polygon P in the hyperbolic disc $\mathbb{D}$ with all of its vertices $s_1, \dots, s_{2n}$ on the boundary $\partial\mathbb{D}$, and cyclically ordered. We denote by δ_{i+1} the side $s_i s_{i+1}$ and by φ_{2k} the parabolic transformation fixing s_{2k} and sending δ_{2k+1} to δ_{2k}. The cycles are then

$$\{s_2\}, \quad \{s_4\}, \quad \dots, \quad \{s_{2n}\} \quad \text{and} \quad \{s_1, s_3, \dots, s_{2n-1}\}.$$

The corresponding isotropy groups are generated by

$$\varphi_2, \quad \varphi_4, \quad \dots, \quad \varphi_{2n} \quad \text{and} \quad \varphi := \varphi_2 \circ \varphi_4 \circ \dots \circ \varphi_{2n}.$$

The transformation φ is parabolic if and only if

$$\frac{(s_1 - s_3)(s_3 - s_5)\cdots(s_{2n-1} - s_1)}{(s_2 - s_4)(s_4 - s_6)\cdots(s_{2n} - s_2)} = -1.$$

If this condition is satisfied, the group Γ generated by the φ_i is Fuchsian. The surface $\overline{S}$ is compact, smooth, of genus 0, and has $n+1$ orbifold points of zero angle. Once endowed with the complex structure defined in §VI.1.8, $\overline{S}$ is just $\mathbb{CP}^1$, while $S = \mathbb{H}/\Gamma$ is $\mathbb{CP}^1$ with $n+1$ points removed. Modulo the action of $\mathrm{PSL}(2,\mathbb{R})$ one can fix three vertices, say s_2, s_4 and s_6; there then remain $2n-3$ parameters subject to the cyclic inequalities and the parabolicity condition. Hence the set of such polygons forms a real semi-algebraic subset of $\mathbb{R}^{2n-3}$ of dimension $2n-4$. Given two $(n+1)$-tuples E_1 and E_2 of points of $\mathbb{CP}^1$, the Riemann surfaces $\mathbb{CP}^1 \setminus E_1$ and $\mathbb{CP}^1 \setminus E_2$ are biholomorphically equivalent if there exists a transformation from $\mathrm{PSL}(2,\mathbb{C})$ sending E_1 onto E_2. Since the action of $\mathrm{PSL}(2,\mathbb{C})$ is transitive on triples of points, a structure of this form is completely determined by a given $n-2$ distinct points of $\mathbb{CP}^1 \setminus \{0, 1, \infty\}$, again yielding the real dimension $2n-4$.

VI.2.3. The surface of genus $g > 1$

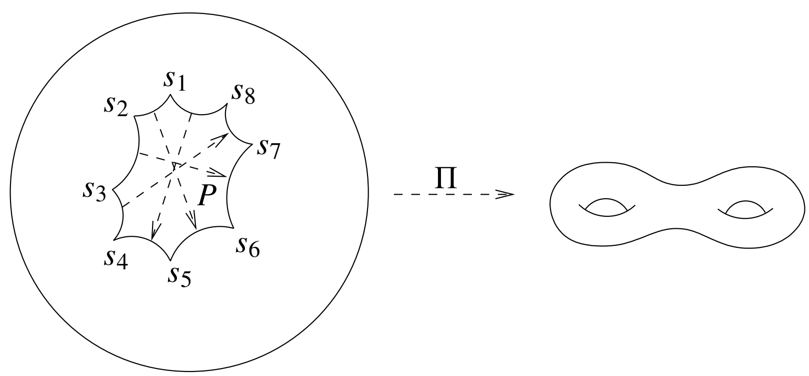


Figure VI.8: The surface of genus 2

Lastly, we consider a $4g$-sided polygon P in $\mathbb{H}$ with cyclically ordered vertices $s_1, \ldots, s_{4g} \in \mathbb{H}$, writing as before $\delta_i = s_i s_{i+1}$. We assume also that δ_i has the same length as δ_{i+2g} and denote by φ_i the transformation sending the first of these onto the second and reversing the orientation induced from P. The vertices then form a single cycle; we shall further assume that the sum of their angles is exactly 2π. The quotient by the Fuchsian group generated by the φ_i is then a compact Riemann surface of genus g. Modulo $\mathrm{PSL}(2, \mathbb{R})$, one needs $8g - 3$ real parameters to determine such a polygon. These parameters are subject to $2g + 1$ equations, of which $2g$ arise from the condition that opposite sides be congruent and one from the condition on the sum of the angles. This leaves $6g - 4$ parameters, which is two more than the dimension of the space of possible complex structures on the compact orientable surface of genus g. These two dimensions are accounted for by the circumstance that the same Riemann surface can be represented by a 2-parameter family of such polygons: the cycle (of angle 2π) plays no role in the surface and may be changed at will; one can also continuously deform the geodesic graph on the surface. In other words, one can deform the polygon without changing either the surface or the group Γ. Hence one obtains finally $6g - 6$ essential real parameters.

Remark VI.2.1. — These dimensional calculations allow us to glimpse the conceptual leap involved in introducing Fuchsian groups; while earlier uniformized Riemann surfaces seemed to represent only exceptional cases — consider for example the case of the Klein quartic — one now saw entire open sets of complex structures uniformized by means of Fuchsian groups. We shall provide a more rigorous treatment of this aspect in the next chapter.

VI.3. Algebraisation according to Poincaré

VI.3.1. Automorphic forms

Recall (see §V.1.2) that an *automorphic form* of weight $\nu \in \mathbb{N}$ for a subgroup Γ of $\mathrm{PSL}(2,\mathbb{R})$ is a (holomorphic or meromorphic) differential form on $\mathbb{H}$ of degree ν

$$\Theta = \theta(z)(dz)^\nu$$

invariant under Γ, that is, satisfying

$$\theta \circ \varphi(z) \cdot (\varphi'(z))^\nu = \theta(z), \quad \forall z \in \mathbb{H},\ \varphi \in \Gamma.$$

This is only of interest if Γ is discrete. Abusing terminology, we also call the function θ an automorphic form of weight ν.

Theorem VI.3.1. — *Let Γ be a Fuchsian group, $\nu \geq 2$, and f a rational function such that the differential form $\Theta_0 = f(z)(dz)^\nu$ has no poles on $\partial\mathbb{H}$ (and the function f vanishes to the order 2ν at $z = \infty$). Then the series*

$$\theta(z) := \sum_{\varphi \in \Gamma} f \circ \varphi(z) \cdot (\varphi'(z))^\nu$$

converges uniformly to a meromorphic automorphic form on every compact subset of $\mathbb{H}$.

Proof. — To establish the convergence we go over to the disc model of the hyperbolic plane

$$\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$$

via the identification $\psi : \mathbb{D} \to \mathbb{H}$, given, for example, by $\psi(z) = i\frac{1-z}{z+1}$. We have to prove the convergence in mean of the differential form $\psi^*\Theta_0 = f \circ \psi(z)(\psi'(z))^\nu(dz)^\nu$ under the action of the group $\psi \circ \Gamma \circ \psi^{-1}$. To avoid undue notational complexity, we shall henceforth understand the group Γ as acting on $\mathbb{D}$ and the differential form $\Theta_0 = f(z)(dz)^\nu$ as given by a rational function f without poles on $\partial\mathbb{D}$.

Now choose a point z_0 of the disc not fixed by any element of $\Gamma \setminus \{\mathrm{Id}\}$, and denote by $D = \overline{\mathbb{D}(z_0, \varepsilon)}$ a closed disc centred at z_0 of radius ε (in the Euclidean metric) contained in $\mathbb{D}$. We first establish the uniform convergence of the series

$$\sum_{\varphi \in \Gamma} |\varphi'(z)|^2$$

on D for sufficiently small ε. For ε small enough the $\varphi(D)$, $\varphi \in \Gamma$ will be pairwise disjoint, whence the total (Euclidean) area $\sum_{\varphi \in \Gamma} \mathrm{area}(\varphi(D))$ will be finite.

We claim (see below) that there exists a constant $K(\varepsilon) > 0$ such that for every $\varphi \in \Gamma$ one has

$$\mathrm{Max}_{z\in D}|\varphi'(z)|^2 < K^2 \frac{\mathrm{area}(\varphi(D))}{\mathrm{area}(D)}.$$

It follows from this that

$$\sum_{\varphi\in\Gamma} |\varphi'(z)|^2 \le \frac{K^2}{\mathrm{area}(D)} \sum_{\varphi\in\Gamma} \mathrm{area}(\varphi(D)) < \infty,$$

whence the normal convergence of the series of interest on D. In particular, the quantities $|\varphi'(z)|$ are uniformly bounded on D by a constant $C > 0$. Hence for every $\nu > 2$, the quantities $|\varphi'(z)|^{\nu-2}$ are bounded above by $C^{\nu-2}$, whence the normal convergence on D of the series

$$\sum_{\varphi\in\Gamma} |\varphi'(z)|^\nu \le C^{\nu-2} \sum_{\varphi\in\Gamma} |\varphi'(z)|^2 < \infty.$$

Now if $f(z)$ is a rational function, or even just meromorphic in a neighborhood of the closed disc $\overline{\mathbb{D}}$, without poles on the boundary $\partial\mathbb{D}$, then it will be uniformly bounded on all the $\varphi(D)$ possibly except for a finite number of them on which it is meromorphic, whence the convergence of the series

$$\sum_{\varphi\in\Gamma} f \circ \varphi \cdot (\varphi'(z))^\nu$$

on D for all $\nu \ge 2$.

It remains to establish the above upper bound for $|\varphi'|$ on D. If $\varphi(z) = \frac{az+b}{cz+d}$, $ad - bc = 1$, then

$$|\varphi'(z)| = \frac{1}{|cz+d|^2} = \frac{1}{|c^2|} \cdot \frac{1}{\mathrm{dist}(z, -\frac{d}{c})^2},$$

where the only variable quantity on D is the Euclidean distance from z to the point $\varphi^{-1}(\infty) = -\frac{d}{c}$. Note that the equality $\varphi(\infty) = \infty$ holds only for finitely many elements of the group. Leaving aside the terms corresponding to these elements, we may assume $c \neq 0$. Let M_φ and m_φ be the largest and smallest values of $|\varphi'(z)|$ on D. Then

$$\frac{M_\varphi}{m_\varphi} \le \left(\frac{\mathrm{dist}(D, \varphi^{-1}(\infty)) + 2\epsilon}{\mathrm{dist}(D, \varphi^{-1}(\infty))}\right)^2 < \left(1 + \frac{2\epsilon}{\mathrm{dist}(D, \partial\mathbb{D})}\right)^2.$$

Writing K for the right-hand majorant here, A for the Euclidean area of D, and A_φ for the Euclidean area of $\varphi(D)$, we then have

$$A_\varphi > m_\varphi^2 \cdot A > \frac{M_\varphi^2}{K^2} \cdot A,$$

which gives directly

$$\mathrm{Max}_{z\in D}|\varphi'(z)|^2 < K^2 \frac{\mathrm{area}(\varphi(D))}{\mathrm{area}(D)}.$$

□

The number of poles of $\theta(z)\cdot(dz)^\nu$ in the fundamental region P is (neglecting simplifying reductions) equal to the number of poles of f in the disc. In the notation of §VI.1.8, the quotient $S = \mathbb{H}/\Gamma$ is a Riemann surface of finite type. Its compactification $\overline{S}$ obtained by adjoining the cycles at infinity is a Riemann surface with boundary.

Lemma VI.3.2. — *The automorphic form $\theta(z)(dz)^\nu$ of Theorem VI.3.1 defines a meromorphic differential form Θ of degree ν on the compact Riemann surface $\overline{S}$.*

Proof. — Since $\theta(z)(dz)^\nu$ is meromorphic on $\mathbb{H}$ and automorphic (Γ-invariant), it projects to a meromorphic differential form of degree ν on

$$S^* = S - \text{elliptic cycles}.$$

Thus we now need to examine the behaviour of Θ in a neighborhood of each type of cycle.

We begin with elliptic cycles. We again work in the disc model, and assume that 0 belongs to an elliptic cycle of angle $2\pi/q$, so that the isotropy subgroup of 0 in Γ is generated by the elliptic transformation $\varphi(z) = e^{2i\pi/q}z$. The form Θ is invariant under φ, in particular, and can be written as

$$\Theta = \sum_{k\geq k_0} a_{kq} z^{kq} \left(\frac{dz}{z}\right)^\nu = \sum_{k\geq k_0} \frac{a_{kq}}{q^\nu} w^k \left(\frac{dw}{w}\right)^\nu,$$

where $w = z^q$ is a local coordinate on S in a neighborhood of the corresponding cycle. Hence the form Θ is meromorphic on S.

Note that even if the rational function f of Theorem VI.3.1 is holomorphic in a neighborhood of the orbit of 0, the form Θ will still have a pole at the corresponding point of S. Indeed, if $f(z) = \sum_{n=0}^\infty a_n z^n$, then averaging over the isotropy subgroup one obtains

$$\sum_{l=1}^{q} f\circ\varphi^l \cdot (d\varphi^l)^\nu = q \sum_{k\geq\frac{\nu}{q}}^{\infty} a_{kq-\nu} z^{kq} \left(\frac{dz}{z}\right)^\nu = q \sum_{k\geq\frac{\nu}{q}}^{\infty} \frac{a_{kq-\nu}}{q^\nu} w^k \left(\frac{dw}{w}\right)^\nu.$$

Hence the form Θ will in general have order $k-\nu$, where k is the smallest integer $\geq \frac{\nu}{q}$. Since $\nu, q \geq 2$, as in our case, one has $k-\nu < 0$.

We next consider parabolic cycles. We return to $\mathbb{H}$ and assume that the point ∞ belongs to a parabolic cycle, with isotropy group generated by $\varphi(z) = z+1$. To that cycle there corresponds a point s on the surface $\overline{S}$ at which a local coordinate is given by $w = \mathrm{e}^{2i\pi z}$; a neighborhood base is given by the family of horospheres $\mathbb{H}_M := \{\mathrm{Im}(z) > M\}$, $M > 0$. By modifying the fundamental polygon P if necessary, one may assume that

$$P_M := P \cap \mathbb{H}_M = \left\{-\frac{1}{2} \leq \mathrm{Re}(z) \leq \frac{1}{2}\right\} \cap \mathbb{H}_M$$

for $M \gg 0$. By construction, the form $\theta(z)(dz)^\nu$ has only a finite number of poles in P and is therefore holomorphic on P_M for $M \gg 0$. On the other hand, $\theta(z)(dz)^\nu$ is φ-invariant, so projects to a form

$$\Theta = \tilde{\theta}(w)\left(\frac{dw}{w}\right)^\nu := \frac{1}{(2i\pi)^\nu}\theta\left(\frac{\log(w)}{2i\pi}\right)\left(\frac{dw}{w}\right)^\nu .$$

The whole difficulty consists in showing that $\tilde{\theta}(w)$, which is defined holomorphically *a priori* only on a punctured neighborhood, grows at a moderate rate and so extends to $w = 0$. We first prove this for the series

$$\theta_0(z) := \sum_{k=-\infty}^{+\infty} f(z+k).$$

Recall that the rational form $f(z)(dz)^\nu$ is, by assumption, holomorphic in a neighborhood of the limit set of Γ; since dz has a pole of order 2 at ∞, it follows that the function f vanishes to the order 2ν at ∞. Hence for $M \gg 0$ one has

$$|f(z)| \leq \frac{C}{|z|^{2\nu}} \quad \forall z \in \mathbb{H}_M,$$

where $C > 0$ is a constant. Hence for all $z_0 \in P_M$

$$|\theta_0(z_0)| \leq \sum_{k=-\infty}^{+\infty} \frac{C}{|z_0+k|^{2\nu}} \leq \sum_{k=0}^{+\infty} \frac{2C}{(|z_0|^2+k^2)^\nu}$$

$$\leq \frac{2C}{|z_0|^{2\nu}} \sum_{k=0}^{+\infty} \frac{1}{\left(1+\left(\frac{k}{|z_0|}\right)^2\right)^\nu},$$

and since there are at most $|z_0| + 1$ integers $k \in \mathbb{N}$ such that $n \leq \frac{k}{|z_0|} < n+1$, it follows that

$$|\theta_0(z_0)| \leq \frac{2C}{|z_0|^{2\nu}}(|z_0|+1)\sum_{n=0}^{+\infty} \frac{1}{(1+n^2)^\nu} \leq \frac{C'}{|z_0|^{2\nu-1}}$$

for some constant $C' > 0$. Therefore

$$\theta_0(z)(dz)^\nu = \tilde{\theta}_0(w)\left(\frac{dw}{w}\right)^\nu$$

with $\tilde{\theta}_0(w)$ holomorphic in a neighborhood of $w = 0$ and vanishing at $w = 0$.

The general case now follows readily. One chooses from each right coset of Γ modulo $\langle\varphi\rangle$ a representative φ_i, $i \in I$, and rearranges the series

$$\theta(z) = \sum_{i\in I}\sum_{k=-\infty}^{+\infty} f_i(z+k)\cdot(dz)^\nu = \sum_{i\in I}\theta_i(z)(dz)^\nu,$$

where $f_i := f \circ \varphi_i \cdot (\varphi_i')^\nu$, $i \in I$. By Theorem VI.3.1 this series converges uniformly on every compact subset of $\mathbb{H}$, so in particular in the annulus $\{r \le |w| \le r'\}$ defined about $w = 0$ by $P_M \setminus P_{M+1}$. Each function f_i is rational and vanishes to the order 2ν at ∞; hence each series $\tilde{\theta}_i(w)$ is holomorphic and vanishes at the point $w = 0$. Thus on the annulus $\{r \le |w| \le r'\}$ the series $\tilde{\theta}(w)$ is a uniform limit of holomorphic functions on the disc $\{|w| \le r'\}$ vanishing at $w = 0$. The limit is therefore holomorphic, vanishing at $w = 0$, and the differential form $\Theta = \tilde{\theta}(w)\left(\frac{dw}{w}\right)^\nu$ is meromorphic of order $\le \nu - 1$.

It remains to show that Θ extends meromorphically to the compact Riemann surface $\overline{S}$ in the case where Γ has hyperbolic cycles. To this end, we return to the disc $\mathbb{D}$ and consider the fundamental region P' symmetric relative to $\partial\mathbb{D}$, obtained by applying a Schwarz reflection to P. The convergence of the series

$$\sum_{\varphi\in\Gamma}|\varphi'(z)|^\nu$$

holds on every compact set $D \subset \mathbb{C}$ not approaching the orbit $\Gamma(\infty)$, once established at a point $z_0 \in \mathbb{D}$. To see this it suffices to observe that, for $z \in D$, one has

$$\frac{|\varphi'(z)|}{|\varphi'(z_0)|} = \left(\frac{\operatorname{dist}(z_0,\varphi^{-1}(\infty))}{\operatorname{dist}(z,\varphi^{-1}(\infty))}\right)^2 < \left(\frac{\operatorname{dist}(z_0,\varphi^{-1}(\infty))}{\operatorname{dist}(D,\Gamma(\infty))}\right)^2;$$

since Γ acts discretely on $\mathbb{CP}^1 \setminus \overline{\mathbb{D}}$, one can uniformly bound $\operatorname{dist}(z_0,\varphi^{-1}(\infty))$ on all nontrivial elements of Γ, thus establishing the convergence of the series on D. In fact, provided the compact set D contains no limit point of $\Gamma(\infty)$, then possibly after omitting a finite number of terms containing a pole, the series still converges. Hence the form Θ extends meromorphically to $\overline{S}$. □

When there are no hyperbolic cycles (and the differential form Θ of degree ν is not identically zero on the compact surface $\overline{S}$) one has

$$\text{number of zeros} - \text{number of poles} = 2\nu(g-1)$$

where g is the genus of the surface $\overline{S}$, or, equivalently, by Proposition VI.1.7:

$$\text{number of zeros} - \text{number of poles} = \nu(n - c - 1)$$

where $2n$ is the number of sides of P and c the number of distinct cycles (elliptic or parabolic). In order to use Theorem VI.3.1 to construct a form Θ that is not identically zero, it is enough to ensure that it has a pole. Note however that several of the poles of f may belong to the same orbit of Γ and simplify in the series θ. To avoid this eventuality, one might for instance choose the function $f(z)$ with all of its poles in the interior $\mathring{P}$ of the fundamental polygon: these poles will then persist (with the same orders) in the form Θ.

VI.3.2. Fuchsian functions and the algebraisation of the Riemann surface

We define a *Fuchsian function* for the group Γ to be any meromorphic function $f(z)$ on the disc $\mathbb{D}$ left invariant by Γ:

$$f \circ \varphi(z) = f(z), \quad \forall z \in \mathbb{D},\ \varphi \in \Gamma$$

(in other words, an automorphic form of weight $\nu = 0$). One constructs such functions by taking the quotient of two automorphic forms of the same weight; in order to ensure nontriviality, it suffices to make an appropriate choice of poles of the rational functions used in Theorem VI.3.1. These are the "*new transcendentals*" proposed by Poincaré.

Proposition VI.3.3. — *Suppose the polygon P has no hyperbolic cycles. Then the field of Fuchsian functions is generated by just two of them, that is, has the form* $\mathbb{C}(x, y)$, *where* $x = x(z)$ *and* $y = y(z)$ *satisfy an algebraic relation* $F(x, y) = 0$, $F \in \mathbb{C}[X, Y]$. *The map*

$$\mathbb{H} \to X = \{F(x, y) = 0\}\ ;\ z \mapsto (x(z), y(z))$$

identifies the compact quotient $\overline{S}$ *with a compactification/desingularization of the algebraic curve* $X = \{F(x, y) = 0\} \subset \mathbb{C}^2$. *The genus of this curve is as calculated above.*

Remark VI.3.4. — Here the curve is considered up to birational equivalence (that is, to within an isomorphism of the function fields). When the genus g is ≥ 3, it is possible that the curve obtained be non-smooth. It may have "apparent" singularities depending on the choice of the generators x and y. We can nevertheless talk of the underlying Riemann surface by passing to its desingularization, or, what amounts to the same thing, to an embedding in some projective space $\mathbb{P}^N$. The genus of the curve is thus well defined.

Proof. — First of all, it is easy to construct a non-constant Fuchsian function on the quotient S by forming the quotient of two automorphic forms of the same weight: here one should choose the poles of the second so that they do not cancel those of the first. Let $x(z)$ be the function so constructed. We shall now show that the field K of meromorphic functions on S is a finite extension of $k = \mathbb{C}(x)$. With this in view, we first prove that every element $y(z) \in K$ is algebraic over k. Since $x(z)$ effectively defines a ramified covering of the Riemann sphere of degree d, say, its inverse has at a generic point x exactly d values $z_i(x)$, $i = 1, \dots, d$. The elementary symmetric functions $\sigma_k(x)$ of the d local values of $y(z_i(x))$ are then well defined and meromorphic on the Riemann sphere, whence rational in x. Finally, $y(z)$ satisfies the polynomial equation

$$y^d - \sigma_1(x)y^{d-1} + \cdots + (-1)^{d-1}\sigma_{d-1}(x)y + (-1)^d\sigma_d(x) = 0.$$

Hence the degree over k of every element of K is bounded by d, and it follows (from the primitive-element theorem) that K is a finite extension of k and that $K = \mathbb{C}(x(z), y(z))$ for some $y \in K$. □

VI.3.3. The dependence of Fuchsian functions on the group Γ

Consider a fundamental polygon P_0 of a Fuchsian group Γ_0, with generators $\varphi_1, \dots, \varphi_n$ as given by Proposition VI.1.3. Now imagine a deformation

$$t \mapsto P_t \subset \mathbb{H}$$

of the polygon P_0 where the finite and infinite vertices vary continuously with the parameter t without collisions, and with the assumptions of Theorem VI.1.10 preserved identically throughout. Then the generators deform continuously:

$$t \mapsto \varphi_1^t, \dots, \varphi_n^t,$$

in such a way that the groups Γ_t they generate are all Fuchsian. For example, in the compact case, it is enough for the angles of the cycles to be kept constant and for the sides conjugated by the φ_i^t to remain conjugate — that is, of the same length.

In this case the family of automorphic forms

$$\theta_t(z) := \sum_{\varphi \in \Gamma_t} f \circ \varphi(z) \cdot (\varphi'(z))^\nu$$

constructed from a given rational function f will also depend continuously on the parameter t; to see this it suffices to note that in the proof of Theorem VI.3.1, all of the constants involved in bounding the series

$$\sum_{\varphi \in \Gamma_t} |\varphi'(z)|^2$$

on D depend continuously on t. This continuous dependence automatically extends to parabolic cycles of the surface $\overline{S}_t$. In taking the quotient of two such forms, one obtains a meromorphic function $x_t(z)$ again depending continuously on t. In particular the degree of the meromorphic function $x_t : \overline{S} \to \mathbb{CP}^1$ must be constant. Hence the function $y_t(z)$ constructed from x_t as in the proof of Proposition VI.3.3 will depend continuously on t, and, finally, so also will the curve $X_t = \{F_t(X,Y) = 0\}$.

VI.4. Appendix

We conclude this chapter with two technical lemmas to be used in Chapter IX in connection with appreciating how Poincaré uniformized the complex structures obtained by removing 4 points from the Riemann sphere. The reader may wish to omit this section at first reading.

Consider the unit disc $\mathbb{D}$ furnished with the hyperbolic metric. We will denote distance and area with respect to this metric by “dist” and “area” respectively. For example, the open disc centred at 0 and of radius $R > 0$ (in the hyperbolic metric) is

$$D(0,R) := \{z \in \mathbb{D} \mid \operatorname{dist}(0,z) < R\} \quad \left(= \mathbb{D}\left(0, \frac{e^R - 1}{e^R + 1}\right)\right)$$

and its (hyperbolic) area grows exponentially with R:

$$\operatorname{area}(D(0,R)) = 2\pi(\cosh(R) - 1), \quad \cosh(R) = \frac{e^R + e^{-R}}{2}.$$

Suppose now that Γ is a Fuchsian group acting on the disc $\mathbb{D}$.

Lemma VI.4.1. — *For every radius $R > 0$, there exists $k \in \mathbb{N}$ and $\varepsilon > 0$ such that for all $z_0 \in D(0,R)$, the disc $D(z_0,\varepsilon)$ contains at most k other points of the orbit $\Gamma(z_0)$. If Γ is finitely generated then k may be chosen independently of R provided one takes $\varepsilon(R) = ce^{-c'R}$, $c, c' > 0$.*

Proof. — Consider first the case where Γ has no elliptic cycles in some neighborhood of $\overline{D(0,R)}$: in such a neighborhood no nontrivial element of Γ will have a fixed point. Hence there is an $\varepsilon > 0$ separating every two distinct points of $D(0,R)$ in the same orbit under Γ — if this were not the case, then there would have to exist a sequence of points $z_n \in D(0,R)$ and elements $\varphi_n \in \Gamma$ such that z_n

and $\varphi_n(z_n)$ converged to the same point $z_0 \in \overline{D(0,R)}$, since this set is closed, and then we could find a transformation φ fixing z_0 in the closure of the φ_n. However since Γ is discrete, such a sequence of transformations would eventually become constant, whence $\varphi \in \Gamma$, a contradiction. Thus in this case we have $k = 0$.

If on the other hand Γ has an elliptic fixed point $z_0 \in \overline{D(0,R)}$, of order l, say, then in the orbit of a point arbitrarily close to z_0 we can find l points arbitrarily close to one another. Thus by choosing $k+1$ as an upper bound for the highest order of an elliptic point in $\overline{D(0,R)}$, we can apply the above argument to obtain an appropriate ε.

If Γ is finitely generated, then for R sufficiently large the complement of $D(0,R)$ will contain only finitely many parabolic and hyperbolic cycles of a suitable fundamental polygon. In the case of a parabolic end the distance between any two points in one and the same orbit will not have a positive lower bound: if φ is parabolic, direct calculation shows that $\operatorname{dist}(z,\varphi(z)) \le ce^{-c' \operatorname{dist}(0,z)}$ where $c' > 1$ can be chosen arbitrarily close to 1 for suitable choice of $c > 0$. □

For each $n \in \mathbb{N}$, we denote by C_n the annulus

$$C_n := D(0,n+1) \setminus D(0,n),$$

and write the series of Theorem VI.3.1 as the sum $\theta(z) = \sum_{n\in\mathbb{N}} \theta_n(z)$ where

$$\theta_n(z) := \sum_{\varphi\in\Gamma,\ \varphi(z)\in C_n} f\circ\varphi(z)\cdot(\varphi'(z))^\nu.$$

Lemma VI.4.2. — *Given a Fuchsian group* Γ *and* $\nu > 1$, $R > 0$, *there exists a constant* $K > 0$ *depending only on* ν, *R, and the constants* ε *and k of the preceding lemma, satisfying*

$$\sum_{\varphi\in\Gamma,\ \varphi(z)\in C_n} |\varphi'(z)|^\nu \le Ke^{(1-\nu)n}.$$

Proof. — Given a value $R > 0$ for the radius, the preceding lemma tells us there exist constants $\varepsilon > 0$ and $k \in \mathbb{N}$ with the property that for all $z_0 \in D(0,R)$, each point of the disc $\mathbb{D}$ is contained in $\varphi(D(z_0,\varepsilon))$ for at most k distinct elements $\varphi \in \Gamma$. Hence the number of points of the orbit of z_0 contained in the disc $D(0,n)$, $n \in \mathbb{N}^*$, is bounded above as follows:

$$|\Gamma(z_0)\cap D(0,n)| \le k\frac{\operatorname{area}(D(0,n))}{\operatorname{area}(D(z_0,\varepsilon))} \le k\frac{2\pi(\cosh(n+\varepsilon)-1)}{2\pi(\cosh(\varepsilon)-1)} \le k\frac{e^n+\varepsilon}{\varepsilon^2}.$$

Certainly, therefore, the number of such points in the annulus C_n also satisfies

$$|\Gamma(z_0)\cap C_n| \le k\frac{e^{n+1}+\varepsilon}{\varepsilon^2}.$$

One also readily verifies that for every automorphism φ of the disc $\mathbb{D}$, one has

$$|\varphi'(z)| = \frac{1 - |\varphi(z)|^2}{1 - |z|^2}.$$

And since for $\varphi(z_0) \in C_n$, one also has

$$|\varphi(z_0)|^2 \geq \frac{e^n - 1}{e^n + 1},$$

it follows that

$$1 - |\varphi(z_0)| \leq \frac{2}{\cosh(n) + 1} \leq \frac{4}{e^n},$$

whence

$$|\varphi'(z_0)| \leq \frac{4}{(1 - R^2)e^n}.$$

Thus

$$\sum_{\varphi(z_0) \in C_n} |\varphi'(z_0)|^\nu \leq K e^{(1-\nu)n} \quad \text{where} \quad K = k \frac{e^{1+\varepsilon}}{\varepsilon^2} \left(\frac{4}{1 - R^2} \right)^\nu .$$

□

It follows from these lemmas that

$$\sum_{n \geq N} \sup\{|\theta_n(z)| \mid z \in D(0,R)\} \;\leq\; K \frac{e^{(1-\nu)N}}{(\nu - 1)N} \sup\{|f(z)| \mid z \in D(0,R)\}.$$

Hence in the situation of a subgroup Γ' of Γ with the property that the Γ-orbits of the points of $D(0,R)$ coincide with the Γ'-orbits on restricting to a sufficiently large disc $D(0,N)$, the corresponding series θ and θ' will be close to one another on $D(0,R)$. This property will be useful to us in understanding the behaviour of the map $P_t \mapsto X_t$ constructed in §VI.3.3 above, in the case where the family of polygons P_t approaches the boundary of the moduli space.

Chapter VII

The "method of continuity"

The aim of this chapter is to establish the uniformization theorem for compact Riemann surfaces in the spirit of the "method of continuity" developed in parallel by Klein and Poincaré. This method consists in showing that the space of uniformizable Riemann surfaces is both open and closed in the space of all Riemann surfaces. The proof we present here is more along the lines of Klein's approach, at least insofar as the "closure" is concerned; Poincaré's approach will be considered in the next chapter. What we actually show in the present chapter is that every compact Riemannian surface of negative Euler–Poincaré characteristic is conformally equivalent to a quotient of the hyperbolic plane. Since every Riemann surface admits a Riemannian metric compatible with its complex structure, this then certainly shows that every compact Riemann surface of negative Euler–Poincaré characteristic is uniformizable by the hyperbolic plane. We describe the set of uniformizable metrics as the continuous image of a space of Fuchsian groups (modulo conjugation) in the space $\mathcal{T}_g$ of metrics (modulo conformal equivalence). Then, noting that $\mathcal{T}_g$ is a real connected manifold of dimension $6g-6$ ($g \geq 2$), we prove that the set of uniformizable metrics is both open and closed in that space.

VII.1. Preliminaries

VII.1.1. Introduction

We now give the definitions of these various objects. Let S be an oriented, closed, connected surface of genus $g \geq 2$ endowed with a smooth structure, and let Met_S be the space of Riemannian metrics on S endowed in turn with the uniform convergence topology. We shall say that two metrics m_1 and m_2 are equivalent ($m_1 \sim m_2$) if there exists $\varphi \in \mathrm{diff}^0(S)$ (the group of diffeomorphisms isotopic to the identity map) such that m_1 is conformal with $\varphi^* m_2$. We denote by

$$\mathcal{T}_g = \mathrm{Met}_S / \sim$$

the quotient by this relation. By the local theorem of Gauss on the existence of locally conformal coordinates — extended to the smooth case by Korn and Lichtenstein — the space $\mathcal{T}_g$ is isomorphic to the *Teichmüller space* of complex structures on the surface S, up to an isotopy.[1]

On the other hand the space of Fuchsian groups is described in terms of representations. Denoting by Γ the fundamental group of the surface S, we write $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ for the set of discrete faithful representations of Γ in $\mathrm{SL}(2,\mathbb{R})$ (with the topology induced from the product topology on $\mathrm{SL}(2,\mathbb{R})^{\Gamma}$) and consider the quotient with respect to conjugation

$$\mathcal{R}_{\mathbb{R}}^{fd}(g) = \mathrm{Rep}_{\mathbb{R}}^{fd}(g)/\mathrm{SL}(2,\mathbb{R}).$$

We shall show that $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is a manifold of dimension $6g-6$; in fact it is a union of connected components of the manifold of irreducible representations, considered up to conjugation (see §§VII.2, VII.3).

VII.1.2. From representations to metrics

We wish to construct a continuous map from the manifold of representations $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ to the Teichmüller space $\mathcal{T}_g$. Henceforth we fix on a connected component $\mathcal{X}$ of $\mathcal{R}_{\mathbb{R}}^{fd}(g)$, and denote by X the component of $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ above $\mathcal{X}$. The action of the group Γ on $X \times \mathbb{H}$, where $\mathbb{H}$ is the Poincaré half-plane, is given by

$$\gamma \cdot (\rho, z) = (\rho, \rho(\gamma) \cdot z), \quad (\rho, z) \in X \times \mathbb{H}.$$

The projection of $X \times \mathbb{H}$ on X extends to the quotient as a submersion of the space $E = \Gamma \backslash (X \times \mathbb{H})$ onto X; it is a C^∞ fibration, locally trivial by Ehresmann's theorem. Furthermore, the fibre S_ρ above each $\rho \in X$ is a compact surface naturally

[1]The fact that this space is a manifold (in fact a smooth manifold) of real dimension $6g-6$ can be proved using Riemann's methods. The space Riemann considered implicitly, namely the modular space $\mathcal{M}_g$ of complex structures modulo diffeomorphisms, is singular (an orbifold) at the points corresponding to surfaces having nontrivial automorphisms (such as Klein's surface!), but one can get around this problem by means of "level structures" (one takes a fixed basis for $H_1(S,\mathbb{Z}/n\mathbb{Z})$ for $n \geq 3$); see the discussion following Proposition II.3.1. One obtains in this way a space intermediate between $\mathcal{T}_g$ and $\mathcal{M}_g$ and Riemann's methods allow one to prove that it is in fact a smooth complex manifold of dimension $3g-3$, and therefore *a fortiori* a smooth real manifold of dimension $6g-6$. For further details on this matter, consult for example [HaMo1998]. Otherwise, there are many expositions available of the theory of Teichmüller spaces themselves, from which it follows directly that the space in question is diffeomorphic to $\mathbb{R}^{6g-6}$. One may consult for instance [ImTa1992], which also contains substantial historical asides.

endowed with a hyperbolic metric m_ρ. As our surface of reference, we choose the fibre $S = S_{\rho_0}$ above a fixed base point $\rho_0 \in X$. For each $\rho \in X$, denote by $c(t) = \rho_t$ ($t \in [0,1]$) a piecewise-smooth path from ρ_0 to ρ. The fibration c^*E is trivializable, and every trivialization $F : [0,1] \times S \to c^*E$ determines a continuous family of diffeomorphisms $f_t = F(t,\cdot) \in \mathrm{diff}(S, S_{\rho_t})$. We may always assume that f_0 is the identity map on S.

Lemma VII.1.1. — *The isotopy class of the metric $f_1^* m_\rho \in \mathrm{Met}_S$ is independent of the choice of trivialization of c^*E (normalized as above) and the choice of the path c from ρ_0 to ρ.*

Proof. — By construction, the group Γ acts via ρ as a group of automorphisms of the universal covering $\mathbb{H}$ of S_ρ (for every $\rho \in X$). The above trivialization F lifts to $\tilde{F} : [0,1] \times \mathbb{H} \to [0,1] \times \mathbb{H}$ (between universal covers), whence we obtain a continuous family $\tilde{f}_t : \tilde{S} \to \tilde{S}_{\rho_t}$ of lifts; here we may assume that $\tilde{f}_0$ is the identity map of $\tilde{S}$. Each $\tilde{f}_t$ defines an automorphism θ_t of the group Γ via the equality $\tilde{f}_t \circ \rho_0(\gamma) = \rho_t(\theta_t(\gamma)) \circ \tilde{f}_t$, where $\gamma \in \Gamma$. However, since θ_t depends continuously on t, we must have $\theta_t = \theta_0 = \mathrm{Id}_\Gamma$ for all $t \in [0,1]$. In other words $\tilde{f}_t$ is Γ-equivariant, whence, in particular,

$$\tilde{f}_1 \circ \rho_0(\gamma) = \rho(\gamma) \circ \tilde{f}_1 \quad (\gamma \in \Gamma). \tag{VII.1}$$

Now consider another path σ_t from ρ_0 to ρ (possibly equal to ρ_t) in order to deal with a change of trivialization. Suppose $g_1 \in \mathrm{diff}(S, S_\rho)$ is obtained via a normalized trivialization. Its lift $\tilde{g}_1$, constructed as above, also satisfies (VII.1). Hence for all $\gamma \in \Gamma$ we have

$$\tilde{f}_1^{-1} \circ \tilde{g}_1 \circ \rho_0(\gamma) = \rho_0(\gamma) \circ \tilde{f}_1^{-1} \circ \tilde{g}_1.$$

The diffeomorphism $\varphi = f_1^{-1} \circ g_1$ clearly satisfies $\varphi^*(f_1^* m_\rho) = g_1^* m_\rho$. Moreover the above equation shows that the outer automorphism of the group $\Gamma = \mathrm{Aut}_S \tilde{S}$ determined by φ is trivial. It follows that φ is isotopic to the identity map (see [ZVC1970, 5.13]). □

Now consider two conjugate representations $\rho, \sigma \in X$, joined by paths ρ_t and σ_t to the base point ρ_0 ($t \in [0,1]$). Choose continuous families $f_t \in \mathrm{diff}(S, S_{\rho_t})$ and $g_t \in \mathrm{diff}(S, S_{\sigma_t})$ as in the above proof, with $f_0 = g_0 = \mathrm{Id}_S$. Let $A \in \mathrm{SL}(2,\mathbb{R})$ be such that $\rho = A\sigma A^{-1}$ and let A_t be a smooth path from I to A in $\mathrm{SL}(2,\mathbb{R})$ ($t \in [0,1]$). Each element A_t induces a diffeomorphism h_t between S_σ and $S_{A_t \sigma A_t^{-1}}$, with $h_0 = \mathrm{Id}_{S_\sigma}$ and $h_1^* m_\rho = m_\sigma$. Hence g_t ($t \in [0,1]$) followed by $h_t \circ g_1$ ($t \in [0,1]$) yields a continuous family of diffeomorphisms above a path from ρ_0 to ρ. By Lemma VII.1.1 (and invoking $g_0 = \mathrm{Id}_S$) we have that the isotopy classes of the metrics $f_1^* m_\rho$ and $(h_1 \circ g_1)^* m_\rho = g_1^* m_\sigma$ coincide.

Write $[\rho] \in \mathcal{X}$ for the conjugacy class of $\rho \in X$ and $[m] \in \mathcal{T}_g$ for the class of a metric $m \in \mathrm{Met}_S$. In view of the foregoing, we can define a map Φ from $\mathcal{X}$ to $\mathcal{T}_g$ by setting

$$\Phi([\rho]) = [f_1^* m_\rho] \in \mathcal{T}_g \quad ([\rho] \in \mathcal{X}).$$

It is then immediate that Φ is a continuous trivialization of the fibration E above a contractible open set containing ρ and ρ_0. The uniformization theorem *à la* Klein is then subsumed in the following result.

Theorem VII.1.2. — *Let $\mathcal{X}$ be a connected component of $\mathcal{R}_{\mathbb{R}}^{fd}(g)$. Then the map $\Phi : \mathcal{X} \to \mathcal{T}_g$ is a homeomorphism.*

VII.2. Representations of surface groups

VII.2.1. The manifold of representations

Let Γ be the fundamental group of a closed, connected, orientable surface of genus $g \geq 2$ (with a distinguished base point), and let $(\gamma_i)_{i=1,\ldots,2g}$ be a set of generators of Γ satisfying $\prod_{i=1}^{g}[\gamma_i, \gamma_{i+g}] = 1$, which we shall henceforth call a *standard generating set*. The set of representations of Γ in $\mathrm{SL}(2,\mathbb{C})$ is then identifiable with the subset $\mathrm{Rep}_{\mathbb{C}}(g)$ consisting of the $2g$-tuples $(A_1, \ldots, A_{2g}) \in \mathrm{SL}(2,\mathbb{C})^{2g}$ satisfying

$$\prod_{i=1}^{g}[A_i, A_{i+g}] = I.$$

The subset $\mathrm{Rep}_{\mathbb{C}}(g)$ is an affine algebraic subvariety of $M_2(\mathbb{C})^{2g}$. Write $\mathrm{Rep}_{\mathbb{C}}^*(g)$ for the subset of $\mathrm{Rep}_{\mathbb{C}}(g)$ consisting of the irreducible representations ρ over $\mathbb{C}$ (that is, such that the only invariant subspaces of $\rho(\Gamma)$ are $\mathbb{C}^2$ and $\{0\}$). By §VI.2.3 this is a non-empty subset of $\mathrm{Rep}_{\mathbb{C}}(g)$ since a faithful discrete representation is necessarily irreducible, as one may readily verify[2]. It is moreover an open subset (even in the Zariski topology) since the set

$$\{(\rho, D) \in \mathrm{Rep}_{\mathbb{C}}(g) \times \mathbb{C}\mathrm{P}^1 \mid \rho(\Gamma) D \subset D\}$$

of pairs made up of a representation ρ and a nontrivial ρ-invariant subspace D of $\mathbb{C}^2$ is a closed set whose first projection (closed since $\mathbb{C}\mathrm{P}^1$ is compact) is $\mathrm{Rep}_{\mathbb{C}}(g) \setminus \mathrm{Rep}_{\mathbb{C}}^*(g)$.

[2]If the group $\rho(\Gamma)$ had a proper, nontrivial invariant subspace of $\mathbb{C}^2$ (in other words an invariant line) $\rho(\Gamma)$ would then be solvable since contained in a conjugate of the subgroup of upper-triangular matrices of $\mathrm{SL}(2,\mathbb{C})$.

For each representation ρ of Γ in $\mathrm{SL}(2,\mathbb{C})$, one may view the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ as a Γ-module via the adjoint action defined by $\gamma\cdot\xi = \mathrm{Ad}\rho(\gamma)(\xi) = \rho(\gamma)\xi\rho(\gamma)^{-1}$; we shall denote this module by $\mathfrak{sl}(2,\mathbb{C})_\rho$. Recall that a 1-*cocycle* is a map c from Γ to $\mathfrak{sl}(2,\mathbb{C})_\rho$ such that for all $\gamma,\gamma'\in\Gamma$

$$c(\gamma\gamma') = c(\gamma) + \gamma\cdot c(\gamma'),$$

and that a 1-*cobordism* is a 1-cocycle of the form $c_\xi(\gamma) = \xi - \gamma\cdot\xi$ with ξ in $\mathfrak{sl}(2,\mathbb{C})_\rho$. We write $Z^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$ for the space of 1-cocycles and $B^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$ for the subspace of 1-cobordisms, and define

$$H^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho) = Z^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)/B^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho).$$

The tangent space to $\mathrm{SL}(2,\mathbb{C})$ at any point σ is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$; the map $\xi\mapsto\exp(\xi)\sigma$ affords via $\mathfrak{sl}(2,\mathbb{C})$ a local chart on $\mathrm{SL}(2,\mathbb{C})$ in a neighborhood of σ.

Proposition VII.2.1. — *The space* $\mathrm{Rep}^*_{\mathbb{C}}(g)$ *is a complex submanifold of dimension* $6g-3$ *of* $SL(2,\mathbb{C})^{2g}$. *For every* $\rho\in\mathrm{Rep}^*_{\mathbb{C}}(g)$ *the map associating with each* $c\in Z^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$ *the element* $(c(\gamma_i))_{1\le i\le 2g}\in\mathfrak{sl}(2,\mathbb{C})^{2g}$, *induces an isomorphism from* $Z^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$ *to* $T_\rho\mathrm{Rep}^*_{\mathbb{C}}(g)$.

Proof. — We follow the argument given by Hubbard in [Hub1981]. Consider the map $f:\mathrm{SL}(2,\mathbb{C})^{2g}\to\mathrm{SL}(2,\mathbb{C})$ defined by

$$f(\sigma_1,\dots,\sigma_{2g}) = \prod_{i=1}^{g}[\sigma_i,\sigma_{i+g}].$$

The set $\mathrm{Rep}_{\mathbb{C}}(g)$ has $f = I$ as analytic equation. A straightforward calculation beginning with

$$\begin{aligned}[\mathrm{e}^{\xi_1}\sigma_1,\mathrm{e}^{\xi_{1+g}}\sigma_{1+g}] &= \mathrm{e}^{\xi_1}(\sigma_1\mathrm{e}^{\xi_{1+g}}\sigma_1^{-1})(\sigma_1\sigma_{1+g}\sigma_1^{-1}\mathrm{e}^{-\xi_1}\sigma_1\sigma_{1+g}^{-1}\sigma_1^{-1})[\sigma_1,\sigma_{1+g}]\mathrm{e}^{-\xi_{1+g}}\\ &= \mathrm{e}^{\xi_1}\mathrm{e}^{\mathrm{Ad}\sigma_1\cdot\xi_{1+g}}\mathrm{e}^{-\mathrm{Ad}(\sigma_1\sigma_{1+g}\sigma_1^{-1})\cdot\xi_1}\mathrm{e}^{-\mathrm{Ad}[\sigma_1,\sigma_{1+g}]\cdot\xi_{1+g}}[\sigma_1,\sigma_{1+g}],\end{aligned}$$

and ending by invoking $\mathrm{e}^{\chi_1}\mathrm{e}^{\chi_2} = \mathrm{e}^{\chi_1+\chi_2} + O(|\chi_1|^2+|\chi_2|^2)$, implies that the differential of f at $\sigma = (\sigma_1,\dots,\sigma_{2g})\in\mathrm{SL}(2,\mathbb{C})^{2g}$ in the direction $\xi = (\xi_1,\dots,\xi_{2g})\in\mathfrak{sl}(2,\mathbb{C})^{2g}$ is

$$\sum_{i=1}^{g}\prod_{j=1}^{i-1}[\sigma_j,\sigma_{j+g}]\cdot\Big((1-\sigma_i\sigma_{i+g}\sigma_i^{-1})\cdot\xi_i + (\sigma_i - [\sigma_i,\sigma_{i+g}])\cdot\xi_{i+g}\Big).$$

Observe that a similar calculation shows that the map $\gamma_i\mapsto\xi_i$ ($i = 1\dots 2g$) extends to a (necessarily unique) 1-cocycle $\Gamma\to\mathfrak{sl}(2,\mathbb{C})_\rho$ if and only if $d_\sigma f(\xi_1,\dots,\xi_{2g}) = 0$, where $\sigma = (\rho(\gamma_i))_{i=1,\dots,2g}$.

Thus it suffices to verify that the map $d_\sigma f : \mathfrak{sl}(2,\mathbb{C})_\rho^{2g} \to \mathfrak{sl}(2,\mathbb{C})$ is surjective if ρ is irreducible. We deduce this by means of two applications of the following lemma:

Lemma VII.2.2. — *If $\sigma_1,\sigma_2 \in SL(2,\mathbb{C})$ do not commute, then the map $\mathfrak{sl}(2,\mathbb{C})_\rho \times \mathfrak{sl}(2,\mathbb{C})_\rho \to \mathfrak{sl}(2,\mathbb{C})_\rho$ sending (ξ_1,ξ_2) to $(1-\sigma_1)\cdot\xi_1 + (1-\sigma_2)\cdot\xi_2$ is surjective.*

Resuming the proof of Proposition VII.2.1, we first assume that for some integer i, $1 \le i \le g$, the elements σ_i and σ_{i+g} do not commute. Then, since

$$(1-\sigma_i\sigma_{i+g})\cdot\xi_i + (\sigma_i - [\sigma_i,\sigma_{i+g}])\cdot\xi_{i+g} = \sigma_i\left((1-\sigma_{i+g}\sigma_i)\sigma_i^{-1}\cdot\xi_i + (1-\sigma_{i+g}\sigma_i^{-1}\sigma_{i+g}^{-1})\xi_{i+g}\right),$$

we infer from Lemma VII.2.2 that the restriction of the map $d_\sigma f$ to ξ_i and ξ_{i+g} is surjective. If we now assume that each σ_i commutes with σ_{i+g}, then the differential of f takes the form

$$d_\sigma f(\xi) = \sum_{i=1}^{g}\left((1-\sigma_{i+g})\cdot\xi_i + (\sigma_i - 1)\cdot\xi_{i+g}\right)$$

and Lemma VII.2.2 can be applied again since, by virtue of the irreducibility of ρ there exist at least two indices i and j in $[1,2g]$ such that σ_i and σ_j do not commute. □

Proof of Lemma VII.2.2. — Let $\sigma \in SL(2,\mathbb{C})$, $\sigma \neq \pm I$. The endomorphism $f_\sigma(\xi) = (1-\sigma)\cdot\xi$ has rank 2 and $\ker f_\sigma = \mathbb{C}(2\sigma - \mathrm{tr}\sigma I)$. One verifies directly that $\ker f_\sigma$ is the orthogonal complement of $\mathrm{im} f_\sigma$ with respect to the nondegenerate bilinear form on $\mathfrak{sl}(2,\mathbb{C})$ defined by $b(\xi,\xi') = \mathrm{tr}(\xi\xi')$. Therefore, since σ_1 and σ_2 do not commute the images of f_{σ_1} and f_{σ_2} must be distinct . □

The action of the group $SL(2,\mathbb{C})$ on $\mathrm{Rep}_\mathbb{C}(g)$ by conjugation preserves $\mathrm{Rep}^*_\mathbb{C}(g)$, and restricted to this subspace this action is locally free. To see this, let $\mathbb{C}^2_\rho$ be the simple $\mathbb{C}[\Gamma]$-module defined by ρ. The ring $\mathrm{End}_{\mathbb{C}[\Gamma]}(\mathbb{C}^2_\rho)$ reduces to homotheties (in particular since each of its elements is either null or invertible, and has a single eigenvalue), so that the centralizer of ρ in $SL(2,\mathbb{C})$ is just $\{\pm I\}$. We write

$$\mathcal{R}^*_\mathbb{C}(g) = \mathrm{Rep}^*_\mathbb{C}(g)/SL(2,\mathbb{C})$$

for the quotient by this action. For every $\rho \in \mathrm{Rep}^*_\mathbb{C}(g)$ the differential at the identity of the inclusion map of $SL(2,\mathbb{C})/\{\pm 1\}$ in $\mathrm{Rep}^*_\mathbb{C}(g)$ defined by $\sigma \mapsto \sigma\circ\rho\circ\sigma^{-1}$, is the map from $\mathfrak{sl}(2,\mathbb{C})_\rho$ to $Z^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$ given by $\xi \mapsto c_\xi$. The following theorem now follows readily from Proposition VII.2.1.

Theorem VII.2.3. — *The space $\mathcal{R}^*_\mathbb{C}(g)$ is naturally endowed with the structure of a complex manifold of dimension $6g-6$. Its tangent space at a point ρ is canonically isomorphic to $H^1(\Gamma,\mathfrak{sl}(2,\mathbb{C})_\rho)$.*

Proof. — It remains only to verify that the action of $\mathrm{SL}(2,\mathbb{C})$ on $\mathrm{Rep}^*_{\mathbb{C}}(g)$ is proper, that is, that the set

$$E_K = \{\sigma \in \mathrm{SL}(2,\mathbb{C}) \mid \sigma K \sigma^{-1} \cap K \neq \emptyset\}$$

is compact in $\mathrm{SL}(2,\mathbb{C})$ for all compact K of $\mathrm{Rep}^*_{\mathbb{C}}(g)$. To this end, observe that the set F of all pairs $(\rho,\rho') \in \mathrm{Rep}^*_{\mathbb{C}}(g)^2$ for which the linear equation

$$X\rho - \rho' X = 0$$

admits a non-zero solution $X \in M_2(\mathbb{C})$ is closed (since the projection $P(M_2(\mathbb{C})) \times (\mathrm{Rep}^*_{\mathbb{C}}(g))^2 \longrightarrow \mathrm{Rep}^*_{\mathbb{C}}(g)^2$ is proper) and non-empty; furthermore, in F the solution space represents a line $D(\rho,\rho') = \mathrm{Hom}_{\mathbb{C}[\Gamma]}(\mathbb{C}^2_\rho, \mathbb{C}^2_{\rho'})$ — generated by an invertible element — depending continuously on (ρ,ρ'). For any given sequence (σ_k) of elements of E_K, there exists a pair $(\rho_k,\rho'_k) \in (K \times K) \cap F$ such that $\sigma_k\rho_k = \rho'_k\sigma_k$, the sequence (ρ_k,ρ'_k) being assumed convergent to $(\rho_\infty,\rho'_\infty)$, say, by compactness. Hence we have $D(\rho_k,\rho'_k) = \mathbb{C}\tau_k$ where τ_k is invertible ($k = 0,\dots,\infty$), $\lim \tau_k = \tau_\infty$ (as k tends to ∞), and we can normalize so as to ensure $\tau_k \in \mathrm{SL}(2,\mathbb{C})$. It follows that $\sigma_k = \pm\tau_k$ ($k < \infty$) and, by choosing a suitable subsequence if necessary, we have $\lim \sigma_k = \pm\tau_\infty$. The set E_K is therefore compact. □

We now turn to the set $\mathrm{Rep}^*_{\mathbb{R}}(g)$ of $\mathbb{C}$-irreducible representations of Γ in $\mathrm{SL}(2,\mathbb{R})$, that is, the intersection of the set of real points $\mathrm{Rep}_{\mathbb{R}}(g)$ of $\mathrm{Rep}_{\mathbb{C}}(g)$ with $\mathrm{Rep}^*_{\mathbb{C}}(g)$. Note that as before this is a non-empty open set of $\mathrm{Rep}_{\mathbb{R}}(g)$.

Corollary VII.2.4. — *The space* $\mathcal{R}^*_{\mathbb{R}}(g) = \mathrm{Rep}^*_{\mathbb{R}}(g)/SL(2,\mathbb{R})$ *is a real manifold of dimension* $6g - 6$.

Proof. — The points of $\mathrm{Rep}^*_{\mathbb{R}}(g)$ are smooth points of $\mathrm{Rep}_{\mathbb{R}}(g)$ since by Proposition VII.2.1 they are smooth points of $\mathrm{Rep}_{\mathbb{C}}(g)$ and this manifold is defined over $\mathbb{R}$. As earlier, one verifies that the action of $\mathrm{SL}(2,\mathbb{R})$ on $\mathrm{Rep}^*_{\mathbb{R}}(g)$ is locally free (for $\rho \in \mathrm{Rep}^*_{\mathbb{R}}(g)$ the module $\mathbb{C}^2_\rho$ is simple). The argument showing that this action is proper is then analogous to that of the proof of Theorem VII.2.3. □

Remark VII.2.5. — More generally, let S be a compact surface of genus g with M points removed (where $M \geq 1$ and $2g + M > 2$). The fundamental group Γ of S (which is free on $2g + M - 1$ generators) is generated by $2g + M$ elements $\sigma_1,\dots\sigma_{2g}, c_1,\dots c_M$ satisfying the single relation $\prod_{i=1}^{g}[\sigma_i,\sigma_{i+g}]\prod_{j=1}^{M} c_j = 1$. Write $R_{\mathbb{C}}(g,M)$ for the space of representations ρ of Γ in $\mathrm{SL}(2,\mathbb{C})$ satisfying the supplementary constraint $\mathrm{tr}\rho(c_j) = -2$ ($1 \leq j \leq M$). One can then show as above that the spaces $R^*_{\mathbb{C}}(g,M)$ of irreducible representations, and their quotients $\mathcal{R}^0_{\mathbb{C}}(g,M) = R^*_{\mathbb{C}}(g,M)/\mathrm{SL}(2,\mathbb{C})$, are complex manifolds of dimensions $6g - 3 + 2M$ and $6g - 6 + 2M$ respectively. Moreover these results continue to hold when $\mathbb{C}$ is replaced by $\mathbb{R}$.

VII.2.2. Characters and the fundamental invariants

Let Γ be a finitely presented group and ρ a representation of Γ in $\mathrm{SL}(N,\mathbb{C})$. The *character* of ρ is then the function χ_ρ from Γ to $\mathbb{C}$ defined by $\chi_\rho(\gamma) = \mathrm{tr}(\rho(\gamma))$. Note that conjugate representations in $\mathrm{SL}(N,\mathbb{C})$ have the same character.

Poincaré was interested in the reciprocal relation when he considered, in the first section of his memoir [Poin1884b], a monodromy representation ρ_E arising from a differential equation on a surface. In this situation the group Γ is the fundamental group of the surface and the $\chi_{\rho_E}(\gamma)$, for $\gamma \in \Gamma$, are invariants that he associates with the substitutions $\rho_E(\gamma) \in \mathrm{SL}(2,\mathbb{C})$. He states that:

> If one knows the invariants of all the substitutions $\rho_E(\gamma)$, the group Γ will be completely determined, since we do not consider it as distinct from its transforms $\sigma^{-1}\Gamma\sigma$. But it is not necessary to know all of these invariants, it suffices to know a certain number of them which we will call *fundamental invariants* and of which all the others are functions.

We shall now give a proof of this assertion (which Poincaré states without proof). It will also be used in Chapter VIII. The following proposition captures the first statement of the above quotation.

Proposition VII.2.6. — *Let ρ and ρ' be two representations of Γ in $SL(N,\mathbb{C})$. If $\chi_\rho = \chi_{\rho'}$ and if ρ and ρ' are irreducible, then ρ and ρ' are conjugate in $SL(N,\mathbb{C})$.*

The proof we give is due to Selberg [Sel1960], and reduces to the following two lemmas.

Lemma VII.2.7. — *Let ρ be an irreducible representation of Γ in $SL(N,\mathbb{C})$. Then there exist N^2 elements $\gamma_1,\dots,\gamma_{N^2}$ of Γ such that the family $(\rho(\gamma_j))_{j=1,\dots,N^2}$ spans the complex vector space $M_N(\mathbb{C})$ of square matrices.*

Proof. — It suffices to prove that the vector subspace R spanned by the $\rho(\gamma)$, with γ ranging over Γ, is the space $M_N(\mathbb{C})$. This is just Burnside's lemma: see for instance [Lan2002, XVII, Corollary 3.4]; in fact R is a subalgebra of $M_N(\mathbb{C})$ and the space $\mathbb{C}^N$ is a simple R-module. □

Lemma VII.2.8. — *Every algebra automorphism $\psi : M_N(\mathbb{C}) \to M_N(\mathbb{C})$ is inner.*

Proof. — Consider the basic matrices E_{ij}. The images $p_i = \psi(E_{ii})$ satisfy $p_i^2 = p_i$, $p_i p_j = 0$ if $i \neq j$, and $\sum p_i = Id$, so they represent projections on n independent lines. Hence $p_i = E_{ii}$ to within a conjugation. It follows that $E_{kk}\psi(E_{ij}) = 0$ if $k \neq i$ and $\psi(E_{ij})E_{kk} = 0$ if $k \neq j$, whence $\psi(E_{ij}) = a_{ij}E_{ij}$ ($a_{ij} \in \mathbb{C}$). We have $a_{ij}a_{jk} = a_{ik}$, whence $a_{ij} = b_i/b_j$, and ψ is conjugation by the matrix $(\delta_{ij}b_i)$. □

Proof of Proposition VII.2.6. — Let $\gamma_j \in \Gamma$ $(j = 1, \ldots, N^2)$ be as in Lemma VII.2.7. If $\sum_{j=1}^{N^2} \lambda_j \rho(\gamma_j) = 0$, then $\lambda_j = 0$ for every j (by choice of the γ_j), whence in turn $\sum_{j=1}^{N^2} \lambda_j \rho'(\gamma_j) = 0$. The endomorphism of $M_N(\mathbb{C})$ sending a linear combination $\sum_{j=1}^{N^2} \lambda_j \rho(\gamma_j)$ to the matrix $\sum_{j=1}^{N^2} \lambda_j \rho'(\gamma_j) \in M_N(\mathbb{C})$ is therefore well-defined. We shall now show that if a linear combination $\sum \lambda_\gamma \rho(\gamma)$ $(\lambda_\gamma \in \mathbb{C})$ is zero in $M_N(\mathbb{C})$, then also $\sum \lambda_\gamma \rho'(\gamma) = 0$. For all $\gamma_0 \in \Gamma$, we have

$$\begin{aligned}
\operatorname{tr}\left(\sum \lambda_\gamma \rho'(\gamma)\rho'(\gamma_0)\right) &= \sum \lambda_\gamma \operatorname{tr}(\rho'(\gamma\gamma_0)) \\
&= \sum \lambda_\gamma \operatorname{tr}(\rho(\gamma\gamma_0)) \\
&= \operatorname{tr}\left(\sum \lambda_\gamma \rho(\gamma)\rho(\gamma_0)\right) = 0.
\end{aligned}$$

It follows from Lemma VII.2.7 (and from the fact that the trace defines a non-degenerate bilinear form on $M_N(\mathbb{C})$) that $\sum \lambda_\gamma \rho'(\gamma) = 0$. Hence, finally, our endomorphism is an algebra morphism, and we can invoke Lemma VII.2.8 to complete the proof. □

In the second statement of the above quotation from Poincaré's memoir, he claims that it is in fact enough to know the invariants of only a finite number of substitutions, his "fundamental invariants".

Proposition VII.2.9. — *Let* Γ *be the fundamental group of a closed surface of genus* g*, and for each* $\gamma \in \Gamma$ *consider the function* $\tau_\gamma : \mathcal{R}_{\mathbb{C}}(g) \to \mathbb{C}$ *defined by* $\tau_\gamma(\rho) = \operatorname{tr}\rho(\gamma) = \chi_\rho(\gamma)$*. The ring* T *generated by the functions* τ_γ $(\gamma \in \Gamma)$ *is finitely generated.*

Proof. — Here $N = 2$. The proof depends on the identity

$$\operatorname{tr}(A)\operatorname{tr}(B) = \operatorname{tr}(AB) + \operatorname{tr}(AB^{-1}) \quad (A, B) \in \mathrm{SL}_2(\mathbb{C})^2,$$

(an immediate consequence of the Cayley–Hamilton theorem) and on the fact that Γ is finitely generated. (As before we denote standard generators by γ_i, $1 \le i \le N$.) We recall the argument (see [Hor1972] or [CuSh1983, p. 116]): Let T_0 be the ring generated by the τ_γ with $\gamma = \gamma_{i_1} \cdots \gamma_{i_k}$, and the indices $i_1, \ldots, i_k$ all distinct (so that in particular $k \le N$). Consider a general element

$$\delta = \gamma_{i_1}^{m_1} \cdots \gamma_{i_r}^{m_r} \in \Gamma$$

with $m_j \neq 0$ $(j = 1, \ldots, r)$. Assuming to begin with that the indices $i_1, \ldots, i_r$ are distinct, we verify that in this case $\tau_\delta \in T_0$. We proceed by induction on $q = \sum_{j=1}^{r} \max(m_j - 1, -m_j)$. When $q = 0$, we have $\tau_\delta \in T_0$ by definition. For

$q > 0$, up to replacing δ by a conjugate we may assume that $m_r \neq 1$. If $m_r < 0$, then by the above trace identity we have

$$\tau_\delta = \tau_{\delta\gamma_{i_r}} \tau_{\gamma_{i_r}^{-1}} - \tau_{\delta\gamma_{i_r}^2} \in T_0$$

since $\tau_{\delta\gamma_{i_r}}, \tau_{\delta\gamma_{i_r}^2} \in T_0$ (by the inductive hypothesis) and $\tau_{\gamma_{i_r}^{-1}} = \tau_{\gamma_{i_r}} \in T_0$ by definition. If $m_r \geq 2$ we write τ_δ as above with γ_{i_r} replaced by $\gamma_{i_r}^{-1}$. We dispose of the general case by means of a second induction on r. In view of what we have just proved, we may suppose that $r \geq 2$, and then, by conjugating if need be, that there exists an index $j < r$ such that $i_j = i_r$. It now suffices to define $\alpha = \gamma_{i_1}^{m_1} \cdots \gamma_{i_j}^{m_j}$, $\beta = \gamma_{i_{j+1}}^{m_{j+1}} \cdots \gamma_{i_r}^{m_r}$ and express τ_δ in the form $\tau_\delta = \tau_{\alpha\beta} = \tau_\alpha \tau_\beta - \tau_{\alpha\beta^{-1}}$. □

Let $(\alpha_1, \dots, \alpha_m)$ be a fixed finite sequence of elements of Γ such that the functions $\tau_{\alpha_1}, \dots, \tau_{\alpha_m}$ generate the ring T. One then defines a mapping t from $\mathcal{R}_{\mathbb{C}}(g)$ to $\mathbb{C}^m$ by

$$t(\rho) = (\tau_{\alpha_1}(\rho), \dots, \tau_{\alpha_m}(\rho)). \tag{VII.2}$$

The numbers $\tau_{\alpha_i}(\rho)$, $i = 1, \dots, m$ are then exactly what Poincaré called *fundamental invariants* of the group $\rho(\Gamma)$.

VII.3. Real faithful and discrete representations

VII.3.1. The faithful and discrete representations form an open set

As in §VII.2.1, let Γ be the fundamental group of a connected, closed surface S of genus $g \geq 2$. In what follows we will understand the group Γ to have a fixed action as a group of automorphisms of the universal cover $\tilde{S}$. Consider the set $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ of faithful, discrete representations of Γ in $\mathrm{SL}(2,\mathbb{R})$ (a nonempty set — see §VII.2.3). These representations are $\mathbb{C}$-irreducible — see §VII.2.1. For $\rho \in \mathrm{Rep}_{\mathbb{R}}^{fd}(g)$, the action of $\rho(\Gamma)$ on the half-plane $\mathbb{H}$ is faithful — since Γ has no elements of order 2 — and the surface $\rho(\Gamma)\backslash\mathbb{H}$ is diffeomorphic to S.

We denote by $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ the quotient of $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ by the conjugations. The first thing to note is that, by virtue of the following proposition, $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is a manifold of dimension $6g - 6$.

PropositionVII.3.1. — *The set $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is an open subset of the manifold $\mathcal{R}_{\mathbb{R}}^{*}(g)$ of irreducible representations.*

Proof. — This reduces to showing that $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ is an open subset of $\mathrm{Rep}_{\mathbb{R}}^{*}(g)$. Consider $\rho_0 \in \mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ and let $\delta_0 : \tilde{S} \to \mathbb{H}$ be a smooth and $(\Gamma, \rho_0(\Gamma))$-equivariant diffeomorphism. Choose a compact set K of $\tilde{S}$ such that $\bigcup_{\gamma\in\Gamma} \gamma(K) = \tilde{S}$. For any given $\rho \in \mathrm{Rep}_{\mathbb{R}}(g)$, we shall show that there is a map $\delta_\rho : \tilde{S} \to \mathbb{H}$ that is smooth, $(\Gamma, \rho(\Gamma))$-equivariant, and C^1-close to δ_0 on K when ρ is close to ρ_0. For the moment we assume this. Then if ρ is sufficiently close to ρ_0, the map δ_ρ is an immersion in a neighborhood of K (immersions forming an open set in the C^1-topology), and therefore by equivariance an immersion on $\tilde{S}$. By considering the inverse image of the hyperbolic metric of $\mathbb{H}$ under δ_ρ, we see that $\tilde{S}$ inherits a Γ-invariant metric. Then since the action of Γ on $\tilde{S}$ is cocompact, it follows from the Hopf–Rinow theorem that this metric is complete. Hence the local isometry δ_ρ is a covering map for $\mathbb{H}$, and therefore a diffeomorphism. The $(\Gamma, \rho(\Gamma))$-equivariance then ensures that ρ is faithful and discrete.

It remains to show how to construct δ_ρ. Choose a fixed open cover of K by open sets $U_1^1, \ldots, U_s^1$, together with successive refinements $U_1^l, \ldots, U_s^l$ $(2 \le l \le s)$ of this cover. More precisely, the refinements are to be chosen so that $\overline{U}_j^{l+1}$ is contained in U_j^l $(l \le s-1,\ j \le s)$ while maintaining $K \subset \bigcup_{j=1}^{s} U_j^l$ for $1 \le l \le s$, and the initial U_j^1 are chosen so as to satisfy $\gamma(U_j^1) \cap \gamma'(U_j^1) = \emptyset$ for $\gamma \ne \gamma'$ in Γ. Now set $V_j^l = \bigcup_{\gamma\in\Gamma} \gamma(U_1^l \cup \ldots \cup U_j^l)$ for all $1 \le j, l \le s$. The map δ_ρ we are seeking is now constructed by means of successive "restriction-extensions" from V_l^l to V_{l+1}^{l+1}. We first define δ_ρ^1 as the $(\Gamma, \rho(\Gamma))$-equivariant map from V_1^1 to $\mathbb{H}$ coinciding with δ_0 on U_1^1. We then suppose that for some $l \in \{1, \ldots, s-1\}$, we have constructed a smooth, $(\Gamma, \rho(\Gamma))$-equivariant map δ_ρ^l from V_l^l to $\mathbb{H}$. The open set $U_{l+1}^{l+1} \cap V_l^{l+1}$ has compact closure in V_l^l, so that, given a smooth map $f : V_l^l \to \mathbb{H}$, it is possible, by means of suitable plateau functions, to extend the restriction of f to $U_{l+1}^{l+1} \cap V_l^{l+1}$ to a smooth map $\overline{f} : U_{l+1}^{l+1} \to \mathbb{H}$. Furthermore this extension procedure may be arranged so as to be continuous in the C^1-topology, so that if f is C^1-close to δ_0 on $V_l^l \cap K$, then $\overline{f}$ is C^1-close to δ_0 on $U_{l+1}^{l+1} \cap K$. It is therefore possible by means of this process to extend the restriction of δ_ρ^l to V_l^{l+1} to a smooth map from $U_{l+1}^{l+1} \cup V_l^{l+1}$ to $\mathbb{H}$; one completes via $(\Gamma, \rho(\Gamma))$-equivariance to obtain $\delta_\rho^{l+1} : V_{l+1}^{l+1} \to \mathbb{H}$. After s steps one arrives at a smooth and $(\Gamma, \rho(\Gamma))$-equivariant map $\delta_\rho = \delta_\rho^s$ from $V_s^s = \tilde{S}$ to $\mathbb{H}$. Now if a sequence $(\rho_k)_{k\ge1}$ converges to ρ_0, then for each $\gamma \in \Gamma$ the sequence $(\rho_k(\gamma))$ converges to $\rho_0(\gamma)$ on the compact set K, in the sense of the C^1-topology; moreover, for $1 \le j, l \le s$, the set of those $\gamma \in \Gamma$ for which $\gamma(U_j^l) \cap K \ne \emptyset$ is finite. This explains why, if ρ is chosen sufficiently close to ρ_0, each of the maps δ_ρ^l $(1 \le l \le s)$ will be C^1-close to δ_0 on $V_l^l \cap K$. □

VII.3.2. Closure of the faithful and discrete representations

We now show that $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is closed in $\mathcal{R}_{\mathbb{R}}^{*}(g)$, a result due in this form to Wielenberg [Wiel1977].

Proposition VII.3.2. — *A limit of discrete, faithful representations of* Γ *in* $SL(2,\mathbb{R})$ *is again faithful and discrete.*

A few preliminaries are in order. Denote by $||\cdot||$ the algebra norm on $M_2(\mathbb{R})$ determined by the usual Euclidean norm on $\mathbb{R}^2$. Given two matrices $A, B \in \mathrm{SL}(2,\mathbb{R})$, and writing $\alpha = A - I$ and $\beta = B - I$, we have:

$$[A,B] - I = (AB - BA)A^{-1}B^{-1} = (\alpha\beta - \beta\alpha)A^{-1}B^{-1}.$$

If $||\alpha||, ||\beta|| < 1$, then $||A^{-1}|| = ||\sum_{n=0}^{\infty}(-\alpha)^n|| \le (1 - ||\alpha||)^{-1}$ and also $||B^{-1}|| \le (1 - ||\beta||)^{-1}$. Hence

$$||[A,B] - I|| \le \frac{2||\alpha||||\beta||}{(1 - ||\alpha||)(1 - ||\beta||)}. \qquad \text{(VII.3)}$$

In particular, if $||A - I||$ and $||B - I|| < 2 - \sqrt{3}$, then $||[A,B] - I|| < ||B - I||$.

The following classical lemma, which in a much more general form is due to Zassenhaus, is also known as Margulis's lemma by reason of the latter's non-linear generalization of it (see for example [Kap2001,§4.12]).

Lemma VII.3.3. — *Let A and B be two elements of* $SL(2,\mathbb{R})$ *with* $||A - I||$ *and* $||B - I||$ *strictly less than* $2 - \sqrt{3}$. *If A and B generate a discrete subgroup of* $SL(2,\mathbb{R})$*, then they commute.*

Proof. — Since the group Π generated by A and B is discrete, there exists an element $C \in \Pi - \{I\}$ such that the norm $||C - I||$ is least. Hence by the inequality (VII.3) we have

$$[A,C] = [B,C] = I,$$

so that C is a nontrivial element in the centre of Π. This element is hyperbolic, parabolic, or elliptic. In the first case, the group Π preserves the axis of C and is therefore Abelian. If C is parabolic, Π fixes the fixed point of C at infinity, and is therefore a subgroup of the group of similarities of $\mathbb{R}$. However since C cannot commute with a strictly contracting or dilating similarity, the group Π must again be Abelian. Lastly, if C is elliptic, then Π fixes the unique fixed point of C and is therefore once again Abelian. □

Proof of Proposition VII.3.2. — Recall first of all that if a representation $\rho : \Gamma \to \mathrm{SL}(2,\mathbb{R})$ is discrete and faithful, then all elements of $\rho(\Gamma)$ must be hyperbolic by the compactness of $\rho(\Gamma)\backslash\mathbb{H}$ — see for example [ImTa1992, p. 46] —

and two elements of $\rho(\Gamma)$ commute if and only if they have a fixed point in common (since their commutator would then necessarily be parabolic). Note also that if $A \in \mathrm{SL}(2,\mathbb{R})$ is hyperbolic, then the only finite subsets of $\mathbb{H} \cup \partial\mathbb{H}$ left invariant by A are the subsets of the set comprised of the two fixed points. Hence if A and B are hyperbolic and satisfy $[A, BAB^{-1}] = I$, then A and BAB^{-1} have the same fixed points, whence one infers that A and B have the same fixed points.

Now let (ρ_k) be a sequence of faithful and discrete representations of Γ in $\mathrm{SL}(2,\mathbb{R})$ convergent to ρ. As before, let $(\gamma_i)_{i=1,\dots,2g}$ be a standard generating family for Γ. We first show that ρ is faithful. Suppose we have an element $\gamma \neq 1$ of Γ such that $\rho(\gamma) = I$. Then of course $\rho(\gamma_i\gamma\gamma_i^{-1}) = I$ for all $i = 1,\dots,2g$, whence by Lemma VII.3.3 for sufficiently large k we must have that $\rho_k(\gamma)$ and $\rho_k(\gamma_i\gamma\gamma_i^{-1})$ commute, and therefore in turn (see above) that $\rho_k(\gamma)$ and $\rho_k(\gamma_i)$ commute. However this means that $\rho_k(\gamma)$ is central in $\rho_k(\Gamma)$, which is absurd.

Finally we show that ρ is discrete. Let Ω be the set of those $A \in \mathrm{SL}(2,\mathbb{R})$ for which $||A - I|| < 2 - \sqrt{3}$, and $\Omega' \subset \Omega$ an open neighborhood of the identity such that $\rho_k(\gamma_i)\Omega'\rho_k(\gamma_i)^{-1} \subset \Omega$ for all k and i. If ρ were not discrete we should have for sufficiently large k that there existed an element $\delta \neq 1$ of Γ such that $\rho_k(\delta)$ belonged to Ω'. This would then entail, as before, that $\rho_k(\delta)$ and $\rho_k(\gamma_i)$ commute, once again yielding a contradiction. □

Thus by Propositions VII.3.1 and VII.3.2, we have that $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is a union of connected components of the manifold $\mathcal{R}_{\mathbb{R}}^{*}(g)$ of irreducible representations. Let G be a covering of $\mathrm{PSL}(2,\mathbb{R})$. The connected components of the space $\mathrm{Hom}(\Gamma, G)/G$ have been described by Goldman in [GolW1988]. First, for $G = \mathrm{PSL}(2,\mathbb{R})$, they coincide with the fibres of the *Euler class*[3]

$$\mathrm{eu} : \mathrm{Hom}(\Gamma, \mathrm{PSL}(2,\mathbb{R}))/\mathrm{PSL}(2,\mathbb{R}) \to \mathbb{Z}.$$

The faithful and discrete representations constitute the two connected components associated with the maximal value $|\mathrm{eu}| = 2g - 2$ — each homeomorphic to a ball. Next, the projection of $\mathrm{SL}(2,\mathbb{R})$ onto $\mathrm{PSL}(2,\mathbb{R})$ determines a covering of degree 2^{2g} by $\mathrm{Hom}(\Gamma, \mathrm{SL}(2,\mathbb{R}))/\mathrm{SL}(2,\mathbb{R})$ of the components of even Euler class of $\mathrm{Hom}(\Gamma, \mathrm{PSL}(2,\mathbb{R}))/\mathrm{PSL}(2,\mathbb{R})$. Hence $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ has at most 2^{2g+1} components. However for $[\rho] \in \mathcal{R}_{\mathbb{R}}^{fd}(g)$, the signs of the $\mathrm{tr}\rho(\gamma_i)$, $i = 1,\dots,2g$ (for a standard generating family) and that of $\mathrm{eu}(\rho)$ are continuous and may be changed

[3]Here is the definition of $\mathrm{eu}(\rho)$. For each i one chooses a lift $\widetilde{\rho(\gamma_i)}$ of $\rho(\gamma_i)$ in the universal cover $\widetilde{\mathrm{PSL}}(2,\mathbb{R}) \to \mathrm{PSL}(2,\mathbb{R})$, whose kernel, canonically isomorphic to $\mathbb{Z}$, is the centre of $\widetilde{\mathrm{PSL}}(2,\mathbb{R})$. One then has $\mathrm{eu}(\rho) = \prod_{i=1}^{g}[\widetilde{\rho(\gamma_i)}, \widetilde{\rho(\gamma_{i+g})}]$, which is independent of both the choice of lifts and the standard generators γ_i, provided the homology classes $[\gamma_i] \in H_1(S,\mathbb{Z})$ form a symplectic basis, that is, satisfy $[\gamma_i]\cdot[\gamma_{j+g}] = \delta_{ij}$, $[\gamma_i]\cdot[\gamma_j] = [\gamma_{i+g}]\cdot[\gamma_{j+g}] = 0$, $1 \le i,j \le g$.

arbitrarily by replacing $\rho(\gamma_i)$ by $-\rho(\gamma_i)$ for any of the i, or by inverting the relation $\prod_{i=1}^{g}[\rho(\gamma_i),\rho(\gamma_{i+g})] = I$ (which changes the Euler class into its opposite, and comes down to changing the orientation of the surface S). Hence the manifold $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ has exactly 2^{2g+1} connected components.

VII.3.3. The map determined by the “fundamental invariants” is proper

Consider fundamental invariants as defined by the formula (VII.2) of §VII.2.2. Since they are conjugation-invariant, they induce (via restriction to the faithful and discrete representations) a map $\bar{t}$ from $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ to $\mathbb{R}^m$. We denote by $\mathcal{R}_{\mathbb{R}}^{fd,+}(g) \subset \mathcal{R}_{\mathbb{R}}^{fd}(g)$ the submanifold of the representations of positive Euler class (in fact necessarily equal to $2g-2$).

Proposition VII.3.4. — *The map* $\bar{t} : \mathcal{R}_{\mathbb{R}}^{fd,+}(g) \to \mathbb{R}^m$ *is injective and proper.*

Proof. — Let $[\rho],[\rho'] \in \mathcal{R}_{\mathbb{R}}^{fd}(g)$ be such that $\bar{t}([\rho]) = \bar{t}([\rho'])$. Since the fundamental invariants determine the character, we have $\chi_\rho = \chi'_\rho$ (Proposition VII.2.9). Hence there exists $A \in \mathrm{SL}(2,\mathbb{C})$ conjugating ρ to ρ' (Proposition VII.2.6), or, in other words, a nonzero complex solution $X = A$ of the system of linear equations *with real coefficients* $X\rho(\gamma) = \rho'(\gamma)X$, $\gamma \in \Gamma$, whose solution space is then the line $\mathbb{C}A$. This immediately gives the existence of a non-zero real solution, necessarily invertible since ρ and ρ' are irreducible. Hence we may choose $A \in \mathrm{GL}(2,\mathbb{R})$, $\det A = \pm 1$, and in fact $\det A = 1$ provided $\mathrm{eu}(\rho)$ and $\mathrm{eu}(\rho')$ have the same sign. Thus we have $[\rho] = [\rho']$.

We now verify that $\bar{t}$ is proper. Let $(\gamma_i)_{1\le i\le 2g}$ be a standard generating family of the fundamental group Γ of the surface S such that the intersection of γ_1 and γ_2 in the homology $H_1(S,\mathbb{Z})$ is ± 1. For $\rho \in \mathrm{Rep}_{\mathbb{R}}^{fd}(g)$, the 1-dimensional homology classes of the surface $\rho(\Gamma)\backslash\mathbb{H}$ are represented by closed geodesics, projections of the axes of the hyperbolic elements $\rho(\gamma)$ ($\gamma \in \Gamma$). We therefore see that the axes of $\rho(\gamma_1)$ and $\rho(\gamma_2)$ meet[4]. Let $(\rho_k)_{k\in\mathbb{N}}$ be a sequence of representations with $t(\rho_k)$ bounded. For every $\gamma \in \Gamma$ the sequence $\mathrm{tr}\rho_k(\gamma)$ is then bounded (see Proposition VII.2.9). For each k, to within a conjugation we may assume that $\rho_k(\gamma_1)$ and $\rho_k(\gamma_2)$ are of the form

$$A_k = \rho_k(\gamma_1) = \begin{pmatrix} u_k & 0 \\ 0 & 1/u_k \end{pmatrix} \text{ and } B_k = \rho_k(\gamma_2) = \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix}.$$

Furthermore we can conjugate again by a diagonal matrix (which does not change A_k) in order to arrange that $|b_k| = |c_k|$. The condition that the axes of A_k

[4]One may also easily construct an example of such a representation ρ in each connected component of $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$; see the conclusion of §VII.3.2.

and B_k intersect translates into the condition $b_k c_k > 0$. From the equality

$$\mathrm{tr}^2 B_k - 4 = (a_k - d_k)^2 + 4 b_k c_k,$$

we infer that the $a_k - d_k$ and $b_k c_k$ are bounded, and then that B_k is bounded. An elementary calculation yields

$$\mathrm{tr}[A_k, B_k] - 2 = -b_k c_k (\mathrm{tr}^2 A_k - 4), \tag{VII.4}$$

whence $\mathrm{tr}[A_k, B_k] < -2$ (since $b_k c_k > 0$), that is, $b_k c_k (\mathrm{tr}^2 A_k - 4) > 4$. Since $b_k c_k$ and $\mathrm{tr} A_k$ are bounded, it follows that $\inf_k (b_k c_k)$ and $\inf_k (\mathrm{tr}^2 A_k - 4)$ are positive.

Now consider any $\gamma \in \Gamma$ and denote by a'_k, b'_k, c'_k, d'_k the entries of $B'_k = \rho_k(\gamma)$. Equation (VII.4), with B'_k in place of B_k, implies that the $b'_k c'_k$ are bounded; therefore, since $a'_k + d'_k = \mathrm{tr} B'_k$ and $a'_k d'_k = b'_k c'_k + 1$ are bounded, so are the diagonal entries a'_k and d'_k. Similarly, by considering $\rho_k(\gamma_2 \gamma)$, we see that the $a_k a'_k + b_k c'_k$ and $d_k d'_k + c_k b'_k$ are bounded, and then, since $\inf_k |b_k| = \inf_k |c_k|$ is positive, that the b'_k and c'_k are bounded. Thus all of the generators $\rho_k(\gamma_i)$ are bounded, for $i = 1, \ldots, 2g$, and we can find a sub-sequence of the ρ_k that converges, necessarily to a faithful and discrete representation (Proposition VII.3.2). □

Note that the analogous result holds for the submanifold $\mathcal{R}_{\mathbb{R}}^{fd,-}(g)$ of classes of discrete and faithful representations with negative Euler class.

VII.4. Proof of uniformization

VII.4.1. The set of uniformizable surfaces is open

Consider the map $\Phi : \mathcal{X} \to \mathcal{T}_g$ defined in §VII.1.2. Recall that $\mathcal{X}$ is an arbitrary connected component of the manifold $\mathcal{R}_{\mathbb{R}}^{fd}(g)$.

Proposition VII.4.1. — *The map $\Phi : \mathcal{X} \to \mathcal{T}_g$ is injective and open.*

Proof. — For the injectivity, it is useful to first re-examine the definition of Φ. This is based on a natural fibration E into hyperbolic surfaces above the component X of $\mathrm{Rep}_{\mathbb{R}}^{fd}(g)$ corresponding to $\mathcal{X}$. The metrics m_ρ on the fibres $S_\rho = \rho(\Gamma)\backslash\mathbb{H}$ ($\rho \in X$) are pulled back to a reference fibre $S = S_{\rho_0}$ by trivializing the fibration E above paths. Thus for $\rho, \sigma \in X$, we defined $\Phi([\rho]) = [f_1^* m_\rho]$ and so also $\Phi([\sigma]) = [g_1^* m_\sigma]$, with f_1 and g_1 obtained via trivialization (see §VII.1.2 for the details). Now suppose there exists $\varphi \in \mathrm{diff}_0(S)$ such that $\varphi^*(f_1^* m_\rho)$ is conformal with $g_1^* m_\sigma$. We know that f_1 and g_1 lift equivariantly to universal covers (see the equality (VII.1)). This holds also for φ, which is isotopic to the identity. Hence the diffeomorphism $\psi = f_1 \circ \varphi \circ g_1^{-1}$ admits a lifting that conjugates the representation σ to ρ. However ψ is directly conformal from S_σ to S_ρ. Hence by Schwarz's lemma (or rather its corollary $\mathrm{Aut}^+(\mathbb{D}) = \mathrm{PSL}_2(\mathbb{R})$), such a diffeomorphism lifts to a Möbius transformation of the half-plane. Thus $[\rho] = [\sigma]$.

The space $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ is a manifold of dimension $6g - 6$ (Theorem VII.2.3). Allowing — as was made precise in the introduction §VII.1.1 — that $\mathcal{T}_g$ is also a manifold of dimension $6g - 6$, one concludes that Φ is open by the (admittedly much later) theorem of Brouwer on the invariance of the domain. □

VII.4.2. The set of uniformizable surfaces is closed

Proposition VII.4.2. — *The map $\Phi : \mathcal{X} \to \mathcal{T}_g$ is proper.*

Let $[\rho_k]$ be a sequence of points of $\mathcal{R}_{\mathbb{R}}^{fd}(g)$ such that $\Phi([\rho_k])$ converges in $\mathcal{T}_g$. Taking, as always, Klein's point of view (see §III.1), we note that this convergence means that there exist Riemannian metrics ds_k^2 on S, $k = 1, \ldots, \infty$, such that ds_k^2 converges to ds_∞^2, each metric ds_k^2 ($k \in \mathbb{N}$) being conformally equivalent to the hyperbolic metric on S associated with'ρ_k (well-defined to within an isotopy — see §VII.1.2). We need to show that, up to extracting a subsequence, the sequence $([\rho_k])$ converges to a limit $[\rho_\infty]$. If this is the case, then the representation ρ_∞ will be faithful and discrete (Proposition VII.3.2) and the metric ds_∞^2 (or the associated complex structure) will be uniformized by $\rho_\infty(\Gamma)$.

We know (see §VII.2.2) that there exists a finite family $(\alpha_j)_{1 \le j \le m}$ of nontrivial free homotopy classes of simple closed curves on S such that each $[\rho_k]$ is determined by the lengths $(\ell_{\rho_k}(\alpha_j))_{1 \le j \le m}$ of the classes α_j relative to the hyperbolic metric associated with ρ_k — lengths corresponding to Poincaré's fundamental invariants[5]. By virtue of the fact that the "fundamental invariants" map is proper (Proposition VII.3.4), the existence of a convergent subsequence of $([\rho_k])$ is a consequence of the following proposition.

Proposition VII.4.3. — *Let α be a free homotopy class of simple closed curves on S. There exists a constant $C_\alpha < +\infty$ such that $\ell_{\rho_k}(\alpha) \le C_\alpha$ for all $k \in \mathbb{N}$.*

Proof. — The proof rests on a simple argument around "extremal length"[6]. Consider first a Riemannian metric ds^2 on S. For every positive function φ on S, we set

$$L_\varphi(\alpha, ds^2) = \inf_{c \in \alpha} \int_c \varphi ds \quad \text{and} \quad A_\varphi(S, ds^2) = \int_S \varphi^2 dA,$$

[5]The traces are certainly determined by the lengths since their signs are fixed in the component $\mathcal{X}$.

[6]In the spirit of the work of Ahlfors and Beurling [Ahl1973].

where dA is the measure of area in the metric ds^2; these quantities are respectively the length of α and the area S with respect to the metric $\varphi^2 ds^2$. The *extremal length* of α is then defined to be

$$E_{ds^2}(\alpha) = \sup_{\varphi>0} \frac{L_\varphi^2(\alpha, ds^2)}{A_\varphi(S, ds^2)}.$$

This quantity is a conformal invariant associated with the class of ds^2 in $\mathcal{T}_g$ (for a fixed α). Furthermore if ds^2 is uniformized by a representation ρ of Γ in $\mathrm{SL}(2,\mathbb{R})$, one has, by definition, $\ell_\rho^2(\alpha) \leq 4\pi(g-1)E_{ds^2}(\alpha)$. The following lemma on semi-continuity then gives us the desired conclusion. □

Lemma VII.4.4. — *Let α be a free homotopy class of simple closed curves on S and (ds_k^2) a sequence of Riemannian metrics on S converging to ds_∞^2 as k tends to infinity. We then have the inequality*

$$\overline{\lim}_k E_{ds_k^2}(\alpha) \leq E_{ds_\infty^2}(\alpha).$$

Proof. — In the above definition of the extremal length we may confine ourselves to functions satisfying $\varphi \leq 1$ since S is compact and φ continuous. Under this condition the sequence $\varphi^2 ds_k^2$ converges to $\varphi^2 ds_\infty^2$ uniformly with respect to φ. Let L be strictly greater than the length of the class α in the metric ds_∞^2 and let $\epsilon > 0$. For every curve $c \in \alpha$ of length $\mathrm{Length}(c, ds_\infty^2)$ less than L, there exists a k_0 independent of $\varphi \leq 1$ and c such that

$$\frac{L_\varphi^2(\alpha, ds_k^2)}{A_\varphi(S, ds_k^2)} \leq \frac{\mathrm{Length}^2(c, \varphi^2 ds_k^2)}{A_\varphi(S, ds_k^2)} \leq \frac{\mathrm{Length}^2(c, \varphi^2 ds_\infty^2)}{A_\varphi(S, ds_\infty^2)} + \epsilon \quad (k \geq k_0).$$

By taking the infimum over the curves c and then the supremum over the functions φ, we infer that $E_{ds_k^2}(\alpha) \leq E_{ds_\infty^2}(\alpha) + \epsilon$ for $k \geq k_0$, whence the desired conclusion. □

Proof of Theorem VII.1.2. — This is now an immediate consequence of Propositions VII.4.1 and VII.4.2, and the fact that, as a quotient of Met_S, $\mathcal{T}_g$ is connected. □

Chapter VIII

Differential equations and uniformization

The aim of this chapter is to examine the route to the uniformization theorem taken by Poincaré, who was interested above all in the solution of linear differential equations. Uniformization was not his initial goal, and only emerged incidentally as a byproduct of his results.

It seemed to us useful to precede the main part of this chapter with a preliminary section summarizing the various features of algebraic differential equations that were undoubtedly present to Poincaré's mind when he began his investigations.

VIII.1. Preliminaries: certain aspects of first-order algebraic differential equations

The Riccati differential equation. — This is any equation of the form

$$\frac{dy}{dx} = a(x)y^2 + b(x)y + c(x),$$

where a, b, c are rational functions of a complex variable x (which may also vary over a more general algebraic curve).

It was this family of equations that led Poincaré to uniformization. Here we recall their basic properties, long ago become classical.

These equations are well known as "disguised" linear differential equations. Starting from a first-order linear differential equation in *two unknowns*

$$\frac{d}{dx}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = A(x)\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{VIII.1}$$

where $A(x)$ is a 2×2 matrix depending rationally on x, one infers that the quotient $y = u_1/u_2$ satisfies a Riccati equation, and conversely every Riccati equation derives from a linear equation of this form.

In fact the Riccati equation can be transformed into a scalar linear equation, but *of the second order*. More precisely, the change of variable given by the formula $y(x) = -\frac{w'(x)}{a(x)w(x)}$ yields the second-order linear equation

$$\frac{d^2w}{dx} + p(x)\frac{dw}{dx} + q(x)w(x) = 0,$$

for w, where $p = -a'/a - b$ and $q = ac$. And conversely, if w is a solution of this second-order linear equation, then the function $y = -w'/w$ satisfies the Riccati equation

$$\frac{dy}{dx} = y(x)^2 - p(x)y(x) + q(x).$$

We recall also that a second-order linear equation in a single unknown w reduces to a first-order linear equation in the two unknowns (w, w').

An important (elementary) property of linear differential equations consists in the fact that the domain of definition of the solutions is the same as that of the equation. Considering for example the above equation in two unknowns associated with a matrix $A(x)$, one may continue a local solution in the neighborhood of a point along any path whatever avoiding the poles of A. Of course, in doing this one may encounter the phenomenon of *monodromy* lying at the heart of this chapter, but the solutions of the associated Riccati equation present only poles as singularities — apart from the poles of the coefficients a, b, c. What is perhaps of greatest interest here is the fact that this property characterizes them. Here is a result in this direction.

Proposition VIII.1.1 — *Let Ω be a simply connected open set of $\mathbb{C}$ and $F : \Omega \times \mathbb{C} \to \mathbb{C}$ a holomorphic function. Consider the first-order differential equation $\frac{dy}{dx} = F(x, y)$. The following two statements are equivalent:*

(a) For every initial condition $(x_0, y_0) \in \Omega \times \mathbb{C}$, there exists a meromorphic *solution y defined on Ω and satisfying $y(x_0) = y_0$;*

(b) There exist holomorphic functions a, b and c defined on Ω such that $F(x, y) = a(x)y^2 + b(x)y + c(x)$.

Proof — We first show that (b) implies (a). The above change of variable shows that the solutions of the Riccati differential equation can be expressed as quotients of two solutions of a linear equation, which are therefore defined (and holomorphic on Ω). The solutions of the Riccati equation are thus meromorphic on Ω.

Here is another proof. The graphs of the solutions of our differential equation are integral curves of the vector field $\frac{\partial}{\partial x} + F(x,y)\frac{\partial}{\partial y}$ defined on $\Omega \times \mathbb{C}$. For each fixed x, the component $F(x,y)\frac{\partial}{\partial y}$ of this vector field is (assuming (b)) quadratic in y and therefore extends to a well-defined holomorphic vector field on $\mathbb{CP}^1$. In fact, if we set $Y = 1/y$, our equation becomes

$$\begin{aligned}\frac{dY}{dx} &= -\frac{1}{y^2}\frac{dy}{dx} = Y^2\frac{dy}{dx}\\ &= -Y^2\left[a(x)\frac{1}{Y^2} + \frac{b(x)}{Y} + c(x)\right]\\ &= -c(x)Y^2 - b(x)Y - a(x).\end{aligned}$$

We can thus *compactify* our equation and, provided we admit meromorphic solutions, seek graphs of solutions in the form of integral curves of a well-defined vector field on $\Omega \times \mathbb{CP}^1$. Since the fibre $\mathbb{CP}^1$ is compact, this means that the meromorphic solutions of our equation are defined along every differentiable curve c contained in Ω, whence, in particular, such solutions are meromorphic on Ω.

We now prove that (a) implies (b). The graphs of the solutions of our differential equation are curves transverse to the fibres $\mathbb{CP}^1$ of the product $\Omega \times \mathbb{CP}^1$. Since these solutions are assumed single-valued and defined on Ω, the projection on the factor Ω determines a diffeomorphism from the graph of each solution to its domain of definition Ω.

Consider the graph of the solution y taking the value $y_0 \in \mathbb{CP}^1$ at the point $x_0 \in \Omega$. This graph meets the fibre above $x \in \Omega$ in a point $y \in \mathbb{CP}^1$. The map sending an initial point y_0 to the point y is a biholomorphism between two fibres, each isomorphic to $\mathbb{CP}^1$. It must therefore be a Möbius transformation, so that

$$y(x) = \frac{\alpha(x)y_0 + \beta(x)}{\gamma(x)y_0 + \delta(x)}.$$

We recover the vector field $F(x,y)$ by taking the derivative at x_0 of $y(x)$, that is, $-\gamma'(x_0)y_0^2 + (\alpha'(x_0) - \delta'(x_0))y + \beta'(x_0)$, which is in fact quadratic in y. (This ultimately derives from the fact that the Lie algebra of PSL(2,$\mathbb{C}$) is identifiable with the polynomials of degree two.) □

We also need to describe briefly the *work of Fuchs* on linear differential equations, which, although it directly inspired Poincaré's work in that direction, is not considered elsewhere in the present book. Fortunately, the existence of Hille's book [Hil1976] somewhat excuses our brevity. We note also the excellent works [Forsy1902, Gra1986, Inc1944, IKSY1991, Val1945].

Thus consider once again a linear equation of type (VIII.1). Fuchs seeks conditions on the matrix A (rationally dependent on x) guaranteeing that local solutions in a neighborhood of the poles of A can be expressed as power series in $(x - x_i)^\lambda$ and $\log(x - x_i)$ for certain λ. He shows that this is the case if and only if A has only simple poles. He also shows how to calculate the exponents λ simply as roots of an equation called "radicial", easily made explicit.

Fuchs's theory is essentially *local*, and he was therefore led to study linear equations of a similar type — called *Fuchsian equations* — where now A is an algebraic function of x, or, in other words, a meromorphic function on a certain compact Riemann surface extended over the x-plane. We will revisit some of Fuchs's work in detail in §IX.1.

Revisiting differential equations and elliptic functions. — We turn once again to elliptic functions, only touched on in Chapter I. The point of departure for that theory was the investigation of integrals of the form

$$x = \int \frac{dy}{\sqrt{(y-\alpha)(y-\beta)(y-\gamma)}}$$

with α, β, γ distinct complex numbers.

We first need to justify the idea of Gauss, Abel, and Jacobi to the effect that y is a single-valued (and periodic) function of x.

The differential form

$$\omega = \frac{dy}{\sqrt{(y-\alpha)(y-\beta)(y-\gamma)}}$$

is well-defined and *nonsingular* on the double cover of the projective line $\mathbb{CP}^1$, ramified over the points $\alpha, \beta, \gamma, \infty$. In other words, the smooth projective cubic C with affine equation $z^2 = (y-\alpha)(y-\beta)(y-\gamma)$ inherits a nonsingular holomorphic volume form. In fact the local coordinate v of the cubic C in a neighborhood of $y = \alpha$ is such that $y - \alpha = v^2$, whence $dy = 2v dv$ and $\omega \simeq 2dv$. A similar calculation in a neighborhood of infinity (where the local variable v satisfies the equation $1/y = v^2$) shows that ω is also holomorphic and nonsingular at infinity.

Dually, C inherits a nonsingular holomorphic vector field X, defined by $\omega(X) = 1$. Since C is compact, the (complex) flow determined by X is complete and its (transitive) action parametrizes C as the quotient of $\mathbb{C}$ by the stabilizer Λ of a point. The smooth cubic is thus uniformized by $\mathbb{C}$. The equation $\omega(X) = 1$ shows that the parametrization of a given orbit of X (the uniformizing map) is the inverse of the corresponding integral x. This inverse its therefore a Λ-periodic elliptic function satisfying, by construction, the differential equation

$$\left(\frac{dy}{dx}\right)^2 = (y-\alpha)(y-\beta)(y-\gamma),$$

the solutions of which are the Weierstrass $\wp$-function and its translates $y = \wp(x + \mathit{const})$.

One may instead base the development of the theory of elliptic functions on the following differential equation:

Theorem VIII.1.2. — *The nontrivial solutions of the equation*

$$\left(\frac{dy}{dx}\right)^2 = (y-\alpha)(y-\beta)(y-\gamma)$$

are elliptic functions which uniformize the smooth projective curve C with affine equation $z^2 = (y-\alpha)(y-\beta)(y-\gamma)$.

Proof. — The proof uses a geometric method invented by Lie (see [PaSe2004]). Denote the quantity $\frac{dy}{dx}$ by z; then $dy = zdx$. We seek solutions in the form of maps $f : \mathbb{C} \to \mathbb{C}^2$, $f(x) = (y(x), z(x))$, with graphs tangent to the contact field $dy - zdx = 0$ and contained in $\mathbb{C} \times C$. The intersection of the tangent space to $\mathbb{C} \times C$ with the contact field allows us to define a nonsingular line field. In fact if we set $F(y) = (y-\alpha)(y-\beta)(y-\gamma)$, this intersection coincides with the kernel of the holomorphic differential 2-form

$$(dy - zdx) \wedge (2zdz - F'(y)dy) = z\Omega,$$

where Ω is a non-vanishing holomorphic 2-form. The kernel of Ω is a line field (defined even at $z = 0$) and one may verify that it extends to a line field $\mathcal{F}$ defined also at the point at infinity of C (and therefore on the whole of $\mathbb{C} \times C$). Furthermore, $\mathcal{F}$ is transverse, that is, its projection on the coordinate x is an isomorphism.

The solutions of our differential equation are curves tangential to $\mathcal{F}$. Let $c : [0,1] \to \mathbb{C}$ be a differentiable curve and (y_0, z_0) a point on C above $c(0)$. There then exists a unique differentiable curve $\tilde{c} : [0,1] \to \mathbb{C} \times C$ that is a lift of c and satisfies $\tilde{c}(0) = (0, y_0, z_0)$. In view of the compactness of the fibre C, the curve $\tilde{c}$ is well-defined on the whole interval $[0,1]$.

This shows that corresponding to every initial condition and every curve c in $\mathbb{C}$, there exists a meromorphic solution defined along that curve. Hence the nontrivial solutions of our equation are defined on $\mathbb{C}$.

Denote by Λ the group of periods of a nontrivial solution y. Since y is completely determined by its initial value, one also has

$$\Lambda = \{\lambda \in \mathbb{C} | y(\lambda) = y(0)\}.$$

Now since the zeros of an analytic function are isolated, Λ is a discrete subgroup of $\mathbb{C}$, so that we have an injective map $f : \mathbb{C}/\Lambda \to C$. One readily verifies that f sends the canonical 1-form dx to the form $\omega = dy/z$ on C: one has, indeed,

$$f^*(\omega) = f^*\left(\frac{dy}{z}\right) = \frac{y'(x)dx}{y'(x)} = dx.$$

In particular, f sends the real volume form $dx \wedge \overline{dx}$ to $vol \wedge \overline{vol}$.

Suppose by way of obtaining a contradiction that Λ is not a lattice in $\mathbb{C}$. Then f maps the manifold of infinite volume $\mathbb{C}/\Lambda$ injectively to the finite-volume cubic C, an absurdity. Since the image of $\mathbb{C}/\Lambda$ is both open and closed (being compact) it must coincide with C. □

Non-linear differential equations. — In the remarkable article [Poin1885b], which nevertheless went unnoticed at the time, Poincaré succeeded in delineating Riccati or elliptic differential equations within the jungle of algebraic differential equations: they are just those without "mobile singularities".

The definition is as follows. Poincaré says of a differential equation

$$R\left(x, y, \frac{dy}{dx}\right) = 0,$$

that it has no *mobile singularities* if one can find a finite number of points $x_1, \ldots, x_n$ (*singular* for the given differential equation) such that for every $x_0 \neq x_1, \ldots, x_n$, every $y_0 \in \mathbb{C} \cup \{\infty\}$, and every path $\gamma : [0,1] \to \mathbb{C} \setminus \{x_1, \ldots, x_n\}$ starting at x_0, there exists a meromorphic solution y of the equation along γ such that $y(x_0) = y_0$. Thus the solutions may be many-valued, but apart from the "fixed singularities" $x_1, \ldots, x_n$ they can only have poles as singularities.

Here are some examples of equations admitting mobile singularities.

The rational equation $\frac{dy}{dx} + \frac{1}{2}y^3 = 0$, with general solution $y(x) = 1/\sqrt{x-c}$ has mobile singularities of *algebraic type*. Observe that the singularity $x = c$ does indeed depend on the initial condition.

The equation (not rational this time) $\frac{dy}{dx} + \exp(y) = 0$ affords another example since the solutions $y(x) = -\log(x-c)$ admit mobile singularities of logarithmic type.

A further example is the equation $\frac{dy}{dx} + y \log^2 y = 0$ whose solutions of the form $y(x) = \exp(1/(x-c))$ present essential mobile singularities.

Theorem VIII.1.3. — *Let $R\left(x, y, \frac{dy}{dx}\right) = 0$ be a differential equation without mobile singularities, where R is polynomial in $y, \frac{dy}{dx}$ and analytic in x. There are then three possibilities: the equation "derives" from a Riccati equation, or its general solution is expressible in terms of elliptic functions, or the general solution is an algebraic function of the coefficients of y and $\frac{dy}{dx}$ in R.*

Proof. — Again we seek graphs of solutions of the differential equation $R\left(x, y, \frac{dy}{dx}\right) = 0$ in the form of curves situated naturally on the surface with equation $R(x, y, z) = 0$ (or more precisely on the projective compactification with respect to the variables (y, z)) and tangential to the plane field defined by $dy - z dx = 0$. This surface is generated by the family of curves obtained by assigning x some fixed value. In view of the assumption that there are no mobile singularities, we can, for every path $\gamma : [0,1] \to \mathbb{C} \setminus \{x_1, \ldots, x_n\}$ and every initial condition x_0, y_0 (with $x_0 = \gamma(0)$), find a solution along γ. The end-point $\gamma(1)$ of this solution allows us to *identify holomorphically* the curves $R(\gamma(0), y, z) = 0$ and $R(\gamma(1), y, z) = 0$ (of course it must be shown that such identifications extend to the associated compact surfaces). In other words, all surfaces $R(x, y, z) = 0$ with x given a fixed value (other than $x_1, \ldots, x_n$) are birationally equivalent.

Thus every path γ avoiding the x_i defines a "monodromy" isomorphism. We distinguish three cases.

If the genus of these curves is 0, that is, if they are copies of the Riemann sphere, the monodromies are projective transformations and the monodromy group is a subgroup of $\mathrm{PSL}(2, \mathbb{C})$. As we saw in the proof of Proposition VIII.1, this characterizes equations of Riccati type or linear.

If the genus is 1 then, as we have seen in connection with Theorem VIII.1.2, the theory of elliptic curves allows us to identify these curves with the quotient of $\mathbb{C}$ by a lattice, and the solutions can be parametrized by means of the corresponding elliptic functions.

Finally, *if the genus is greater than or equal to 2*, then a theorem due to Klein affirms that the group of holomorphic automorphisms of such a curve is finite. It follows that the monodromy group is finite or the general solutions take on only finitely many values for each value of x (different from the x_i). From this it is not difficult to conclude that the general solution is algebraic. (For many further details concerning this proof, the reader may consult [PaSe2004].) □

Later developments. — Thus the preceding theorem of Poincaré shows that *the quest for new transcendentals via first-order algebraic differential equations should be concentrated on Riccati differential equations, or, equivalently, second-order linear differential equations.* Here we have the principal motivation behind the articles of Poincaré of interest to us here.

Of course one might also attempt an investigation of *nonlinear higher-order* algebraic differential equations, which lie outside the scope of the theorem of Poincaré we have just expounded. It was Painlevé who, following the work of Poincaré we are concerned with here, made major contributions to the topic. There is perhaps some point in mentioning two of them. The first shows that for first-order algebraic equations the mobile singularities are of a limited sort.

Theorem VIII.1.4. — *An equation of the form $\frac{dy}{dx} = F(x, y)$ where F is a rational function in y with coefficients algebraic functions of the variable x, can only have mobile singularities of algebraic type.*

Proof. — The graphs of solutions of such an equation must be contained in the surface given by the equation $R(x, y, z) = z - F(x, y) = 0$ and be tangent to the plane field given by $dy - zdx = 0$. The trace of this plane field on the surface $R(x, y, z) = 0$ defines a holomorphic singular line field $\mathcal{F}$ (with isolated points as singularities). Moreover this line field compactifies into a well-defined singular line field on the compactification of the curve $R(x, y, z) = 0$ in $\mathbb{CP}^1 \times \mathbb{CP}^2$. Note that $\mathcal{F}$ may become vertical (contained in $dx = 0$) at infinity.

The singular points $x_1, \ldots, x_n$ of the equation will then be the projections of the singular points of the foliation on the projective line $\mathbb{CP}^1$, the domain of the variable x.

Consider a differentiable curve c contained in the plane of the variable x. We have the following two possibilities: If the curve c lifts to a curve $\tilde{c}$ tangential to $\mathcal{F}$ (along which the foliation $\mathcal{F}$ is not vertical), then there exists a meromorphic solution of the equation, differentiable along the curve c.

If the tangent to a lift $\tilde{c}$ of c becomes vertical above a point $c(t_0)$, then the analytic continuation of a local solution along c will not be meromorphic, but have an algebraic singularity at $c(t_0)$. In order to see that that singularity is indeed algebraic, one reverses the roles of the variables x and y and observes that at the point in question x is then a holomorphic function of y with derivative zero since $\frac{dx}{dy} = F(x, y) = 0$. If that derivative vanishes to the order $k \in \mathbb{N}^*$, then y represents an algebraic mobile singularity given by a series in $(x - c(t_0))^{\frac{1}{k}}$. □

The other major contribution of Painlevé is his systematic investigation of second-order nonlinear equations without mobile singularities with the aim of discovering the "Painlevé transcendentals". He succeeded in classifying equations of the form $y'' = F(x, y, y')$ without mobile singularities, where F is an analytic function in x and rational in y, y'. He shows that only six classes of such equations (the simplest being of the form $y'' = 6y^2 + x$) effectively yield new transcendentals, that is, functions not expressible in terms of algebraic functions and known transcendental functions.

VIII.2. Poincaré's approach

We now return to Poincaré, sustained as he is by elliptic functions and convinced by his theorem that linear equations provide a suitable framework for research on new transcendentals. He considers a Riccati equation on which he imposes the conditions that it be Fuchsian and that the λ arising in connection with the poles be reciprocals of integers. The equations he obtains, which he calls "normal", have all the properties necessary for the inverses of their solutions, *à la* Jacobi, to be *locally* single-valued (the local exponents becoming integers). But are the inverses of these solutions *globally* single-valued? Will the elliptic miracle occur again? *Nothing of the kind*, but nonetheless Poincaré makes a remarkable discovery. Among all the normal differential equations defined on a single Riemann surface S, there exists a *unique* one[1] the inverses of whose solutions are single-valued in a disc D. This yields a parametrization of S by D. Just as an elliptic curve is the quotient of $\mathbb{C}$ by a lattice acting as group via translations, the Riemann surface S is the quotient of the disc by a discrete group of holomorphic automorphisms. *The Fuchsian functions were born.*

Poincaré then shows that Fuchsian functions allow the solution not only of this privileged Fuchsian equation but of all normal equations on S. Thus did he discover new transcendentals, single-valued on a disc, and show that they allow the solution of all normal equations. Mission accomplished![2]

As it were incidentally — this was not Poincaré's research goal — he had uniformized all surfaces of genus at least 2. These are all isomorphic to the quotient of a disc by the holomorphic action of a discrete group. However, this major result — which so surprised Klein — was only of secondary importance to Poincaré.

It is perhaps best to quote relevant excerpts from Poincaré's announcement [Poin1921].

> I was thus led to examine linear equations with rational and algebraic coefficients.
>
> [. . .]
>
> This close study of the nature of integral functions cannot be achieved without the introduction of new transcendentals, about which I shall now say a few words. These transcendentals have great analogies with elliptic functions, and one should not be astonished at this, since if I conceived these new functions, it was in order to do for linear differential equations what had been done by means of the elliptic and Abelian ϑ-series for the integrals of algebraic differentials.

[1]Poincaré calls this equation *Fuchsian*. We prefer the terminology *uniformizing equation*.

[2]We quote the testimony of Lecornu, classmate at l'École Polytechnique and l'École des Mines, as reported by Appell in [App1925]:

> I remember that, invited by me to dine with my parents on December 31, 1879, he spent the evening walking up and down, not hearing what we said to him or else replying in monosyllables, and forgetting the time to such an extent that just after midnight I took it upon myself to remind him gently that we were now in 1880. He seemed at that moment to return to earth, and decided to take leave of us. Some days afterwards, when I met him on the quay of the port of Caen, he told me carelessly: I know how to integrate all differential equations. Fuchsian functions have just been born. And I guessed then what he had been thinking about in going from 1879 to 1880.

This quote, mentioned in [MiPo1999] does not fully gibe with the chronology of the discovery of Fuchsian functions as reported by Poincaré himself in [HaPo1993]; but at least it shows clearly what motivated him.

It is thus the analogy with elliptic functions that has served me as guide in all my investigations. The elliptic functions are single-valued functions that remain unchanged when one increases the variable by certain periods. This idea is so useful in Mathematical Analysis, that all geometers must long ago have thought how convenient it would be to generalize it by seeking single-valued functions of a single variable x which remain unchanged when one applies certain transformations to that variable; however such transformations cannot be chosen in any way whatever.

[...]

It is easy to see what particular kind of discrete groups it is appropriate to introduce. Recall how the elliptic functions arise: one considers certain integrals said to be of the first kind, and then, by means of a process known under the name of inversion, one regards the variable x as a function of the integral; the function so defined is single-valued and doubly periodic.

In the same way, we take a second-order linear equation and, by means of a sort of inversion, we regard the variable as a function, no longer of the integral, but of the ratio z of two integrals of our equation. In certain cases, the function so defined will be single-valued, and then it will remain unaltered by an infinity of linear substitutions, changing z into $\frac{\alpha z+\beta}{\gamma z+\delta}$.

[...]

The results so obtained as yet give only a very incomplete solution to the problem I set myself, that is, the integration of linear differential equations. The equations I have called Fuchsian, and which can be integrated by means of a simple inversion, are just very special cases of second-order linear equations. One should not be surprised at this if one reflects a little on the analogy with elliptic functions. The inversion process only allows the calculation of integrals of the first kind. For integrals of the second and third kinds, it is necessary to proceed otherwise.

Consider for example the integral of the second kind

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

To evaluate it we consider as auxiliary equation that giving the integral of the first kind

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

where by inversion $x = \operatorname{sn} z$. Replacing x by $\operatorname{sn} z$, we find that u is equal to a single-valued function of z, $Z(z)$, which increases by a constant amount when z increases by a period. We are thus led to use an analogous procedure: given a linear differential equation E of any order, with coefficients algebraic in x, we use an auxiliary second-order equation E', and this auxiliary equation must be chosen in such a way that x is a Fuchsian function of the ratio z of two integrals of E' and that the integrals of E are single-valued functions of z.

Is it always possible to make this choice so as to satisfy all these conditions? Such is the question that naturally arises. This comes down moreover to wondering if, among linear equations satisfying certain conditions that it is pointless to state here, there is always a Fuchsian equation. I have managed to prove that one can answer this question affirmatively. I cannot explain here in what consists the method we, M. Klein and I, have employed in studying diverse particular examples; how M. Klein has sought to apply the method in the general case; nor how I filled in the gaps which persisted in the proof of the German geometer in introducing a theory having the most profound analogies with that of the reduction of quadratic forms.

[...]

Thus is it possible to express the integrals of linear equations with algebraic coefficients in terms of new transcendentals, in the same way as one expresses, in terms of Abelian functions, the integrals of algebraic differentials. Furthermore the latter integrals are themselves susceptible of being obtained by means of Fuchsian functions, and one then arrives at a new expression, entirely different from that involving ϑ-series in several variables.

VIII.3. Second-order linear differential equations, normal equations and uniformizing equations

The idea of using second-order linear differential equations to uniformize Riemann surfaces arose essentially from the following two observations. On the one hand, if S is a Riemann surface uniformizable by the half-plane and $w : \widetilde{S} \to \mathbb{H}$ is a biholomorphism, then in a neighborhood of every point w can be expressed as the quotient of two independent solutions of a certain second-order linear differential equation on an open set of S. On the other hand, if E is a second-order linear differential equation on an open set U of S, then the quotient w of two independent solutions of E is always a local biholomorphism from $\widetilde{U}$ to $\mathbb{CP}^1$. The aim of this section is to justify these two assertions and introduce the definitions needed to formulate the uniformization question "*à la* Poincaré", that is, as the problem of the existence of a "*uniformizing*" differential equation.

VIII.3.1. Second-order linear differential equations

Let U be a connected open set of a Riemann surface S and $x : U \to \mathbb{C}$ a coordinatizing holomorphism defined on this open set[3]. For us, a *second-order linear differential equation on U in the coordinate x* will be a differential equation of the form

$$\frac{d^2v}{dx^2} + f\frac{dv}{dx} + gv = 0 \tag{E}$$

where $f, g : U \to \mathbb{C}$ are given *holomorphic* functions (we stress the fact that the functions f and g are not permitted to have poles in U) and $v : U \to \mathbb{C}$ is the unknown function[4].

A second-order linear differential equation on U in the coordinate x will be said to be *reduced* if it has no order-one term, that is, if it has the form

$$\frac{d^2v}{dx^2} + hv = 0 \tag{E'}$$

where h is a holomorphic function on U.

The notion of a second-order linear differential equation is certainly stable under coordinate changes: if we rewrite the equation E above in terms of a coordinate y, the differential equation obtained will still be second-order linear. This is why one may speak of a second-order linear differential equation without specifying the coordinate. On the other hand the concept of a reduced equation is *not* invariant under coordinate changes: if the equation E' is rewritten in terms of a new coordinate y, one obtains in general an equation that is no longer reduced.

If E is a second-order linear differential equation on an open set U, then its solutions are holomorphic functions $v : U \to \mathbb{C}$ that are, in general, *many-valued*. In other words, it is more appropriate to view such solutions as (genuine) functions defined on the universal cover $\pi : \widetilde{U} \to U$, that is, as solutions of the differential equation on $\widetilde{U}$ induced from E:

$$\frac{d^2\widetilde{v}}{d\widetilde{x}^2} + (f \circ \pi)\frac{d\widetilde{v}}{d\widetilde{x}} + (g \circ \pi)\widetilde{v} = 0,$$

where $\widetilde{x} = x \circ \pi$. In practice it is convenient to hold both points of view: we will understand the solutions of E to be functions on $\widetilde{U}$, but it will sometimes be practical to use the language of many-valued functions.

---

[3]By this we mean that x is a local biholomorphism from U onto an open set $\mathbb{C}$. Typically we shall be considering a ramified covering $x : S \to \mathbb{CP}^1$ and U will be the surface S with $x^{-1}(\infty)$ and the ramification points of x removed.

[4]Here $\frac{dv}{dx}$ denotes the "ratio" of dv and dx as sections of the cotangent bundle.

VIII.3.2. Quotients of solutions and projective equivalence

Let U be a connected open set of a Riemann surface S and $\widetilde{U}$ its universal cover. We are interested in functions arising as quotients of two solutions of a second-order linear differential equation on U. We begin with a few elementary remarks concerning such functions.

Proposition VIII.3.1. — *Consider a second-order linear differential equation E on U and two independent solutions v_1, v_2 of that equation. Write w for the quotient v_1/v_2. Then w is a local biholomorphism*[5] *from $\widetilde{U}$ to $\mathbb{CP}^1$ and for each automorphism γ of the universal cover $\widetilde{U}$ there exists a Möbius transformation $\rho(\gamma)$ such that $w \circ \gamma = \rho(\gamma) \circ w$.*[6] *Furthermore, a function is expressible as the quotient of two independent solutions of E if and only if it is the composite of w with a Möbius transformation.*

Proof. — The solutions v_1, v_2 are holomorphic many-valued functions on U, which is to say that they are holomorphic functions on $\widetilde{U}$. Their quotient w is meromorphic on $\widetilde{U}$ and its derivative is given (to within a non-vanishing factor) by the Wronskian $\frac{dv_1}{dx}v_2 - v_1\frac{dv_2}{dx}$, which is non-vanishing in view of the independence of the solutions v_1 and v_2. This shows that $w : \widetilde{U} \to \mathbb{CP}^1$ is étale.

Now let γ be an automorphism of the universal cover $\widetilde{U}$ and v a solution (that is, a solution of the equation on $\widetilde{U}$ induced from E — see §VIII.3.1). Since the coordinate of $\widetilde{U}$ (induced from that on U) is invariant under γ, it follows that $v \circ \gamma$ is also a solution. Thus the pair $(v_1 \circ \gamma, v_2 \circ \gamma)$ of solutions is obtained from the pair (v_1, v_2) of independent solutions by means of an element of $\mathrm{GL}_2(\mathbb{C})$. Hence there exists a Möbius transformation $\rho(\gamma)$ such that $w \circ \gamma = \rho(\gamma) \circ w$.

The final assertion of the theorem follows from the argument involving comparison of bases of a vector space. □

We now introduce an equivalence relation on the set of second-order linear differential equations taking into account the fact that our interest lies not so much with the equations' solutions in themselves as with quotients of pairs of independent solutions.

Proposition VIII.3.2. — *Given two second-order linear differential equations E and E' on an open set U, the following three conditions are equivalent:*

(i) *the set of functions expressible as the quotient of two solutions of E coincides with the set of functions expressible as the quotient of two solutions of E';*

[5] One also says that w is *étale*.

[6] In other words, w is a many-valued meromorphic function on U, each of whose local determinations (branches) is étale from U to $\mathbb{CP}^1$, with passage from one determination to another achieved by composing with a Möbius transformation.

(ii) *the quotient of any two independent solutions whatever of E' can be obtained by composing the quotient of any two solutions of E with some Möbius transformation;*

(iii) *there exists a non-vanishing holomorphic many-valued function k such that the solutions of E' are obtained by multiplying those of E by k.*

Proof. — The equivalence (i) $\Leftrightarrow$ (ii) follows immediately from the final assertion of Proposition VIII.3.1. The implication (iii) $\Rightarrow$ (i) is obvious. Thus it remains to prove the implication (i) $\Rightarrow$ (iii).

Suppose (i) holds and consider two independent solutions v_1', v_2' of E'. By assumption there exist two independent solutions v_1, v_2 of E such that $v_1/v_2 = v_1'/v_2'$. Writing $k = v_1'/v_1$, we have $v_1' = kv_1$ and $v_2' = kv_2$. Since (v_1', v_2') is a basis for the solutions of E', it follows that every solution of E' is obtained by multiplying some solution of E by k. The function k is meromorphic *a priori*; but it is easy to see that in fact it has neither zeros nor poles. For example, if k vanished at a point x_0 of U, then the solutions v_1' and v_2' would have to vanish simultaneously, which is impossible since they are assumed independent. One shows similarly that k has no poles by reversing the roles of the solutions. Hence k is a non-vanishing holomorphic function and (iii) holds. □

Definition VIII.3.3. — Under the conditions of the above proposition the equations E and E' will be called *projectively equivalent*.

Proposition VIII.3.4 below allows us to replace the abstract space of projective-equivalence classes of second-order linear differential equations on U by a more concrete space: that of second-order linear differential equations that are reduced relative to a fixed coordinate.

Proposition VIII.3.4. — *Let $x : U \to \mathbb{C}$ be a holomorphic local coordinate. Every second-order linear differential equation on U is projectively equivalent to a unique equation reduced in the coordinate x.*

Proof. — Consider a second-order linear differential equation E on U in the variable x:

$$\frac{d^2v}{dx^2} + f\frac{dv}{dx} + gv = 0. \qquad (E)$$

By condition (iii) of Proposition VIII.3.2, an equation E' is projectively equivalent to E if and only if it can be obtained from E by means of a change of unknown of the form $v = k(x)v'$. Applying such a change to E, we obtain the equation

$$\frac{d^2v'}{dx^2} + \left(f + \frac{2}{k}\frac{dk}{dx}\right)\frac{dv'}{dx} + \left(g + \frac{f}{k}\frac{dk}{dx} + \frac{1}{k}\frac{d^2k}{dx^2}\right)v' = 0.$$

For this equation to be reduced, k must satisfy $\frac{dk}{dx} = -\frac{1}{2} f k$, in which case the latter equation becomes

$$\frac{d^2 v'}{dx^2} + \left(g - \frac{1}{2}\frac{df}{dx} - \frac{1}{4} f^2 \right) v' = 0. \qquad (E')$$

This proves the proposition. Observe that although the function k may *a priori* be many-valued, the functions appearing in the final equation E' are single-valued. Its form implies the uniqueness of the reduced equation in the coordinate x. □

We turn now to the following problem: *what (many-valued) functions appear as quotients of two independent solutions of a second-order linear differential equation?*

By the preceding discussion such a function w must be étale and its branches interchanged by Möbius transformations. We shall now show, by means of the Schwarzian derivative (see Box IV.1), that these two conditions suffice to characterize the functions in question. This elementary but fundamental fact highlights the connection between uniformization and differential equations (see Corollary VII.3.7).

Proposition VIII.3.5. — *Let $w : U \to \mathbb{CP}^1$ be a many-valued meromorphic function that is étale and whose branches are interchanged by Möbius transformations. Let $x : U \to \mathbb{C}$ be a coordinate on U. Then w is the quotient of two independent solutions of the following second-order linear equation:*

$$\frac{d^2 v}{dx^2} + \frac{1}{2}\{w, x\} v = 0. \qquad (E')$$

Proof. — Since w is étale, its derivative with respect to x is non-vanishing and $\{w, x\}$ is holomorphic. Furthermore, in view of the projective invariance (of the Schwarzian derivative of w), *the Schwarzian $\{w, x\}$ is single-valued on U*. Hence the equation E' is a second-order linear differential equation in our restricted sense. Set

$$v_1 = w \Big/ \sqrt{\frac{dw}{dx}} \quad \text{and} \quad v_2 = 1 \Big/ \sqrt{\frac{dw}{dx}}.$$

Then obviously $w = v_1/v_2$. We also have

$$\{w, x\} = -2\sqrt{\frac{dw}{dx}} \frac{d^2}{dx^2} \left(\sqrt{dw} dx \right)^{-1},$$

and by using the above equation E' with $k(x) = \left(\sqrt{dw} dx \right)^{-1}$ (k as in Propositions VIII.3.2 and VIII.3.4) we obtain an equation of which 1 and w are obvious solutions. The functions v_1 and v_2 are therefore solutions of the equation E'. □

Corollary VIII.3.6. — *Let w be the quotient of two independent solutions of a second-order linear differential equation E on U. Then the equation E' above is the unique reduced equation projectively equivalent to E.*

Proof. — This is immediate from Proposition VIII.3.2. □

Corollary VIII.3.7. — *Let S be a Riemann surface uniformizable by the half-plane $\mathbb{H}$, $\pi : \widetilde{S} \to S$ the universal cover of S, and $\varphi : \widetilde{S} \to \mathbb{H}$ a biholomorphism. Then for every open set U of S furnished with a coordinate x, the restriction w of φ to $\pi^{-1}(U)$ is the quotient of two independent solutions of the differential equation E' on U.*

Proof. — The conjugates by φ of the automorphisms of the universal cover $\widetilde{S}$ are biholomorphisms of the half-plane $\mathbb{H}$, that is, Möbius transformations with real coefficients. It follows that the function w, considered as a many-valued function on U (being actually defined on a covering of U), satisfies the assumptions of Proposition VIII.3.5. □

We now need to consider the problem of changing coordinates in second-order linear equations. Suppose we are given a second-order linear equation on U, reduced in a coordinate x. If we rewrite this equation in terms of another coordinate y, we will in general obtain a non-reduced equation. According to Proposition VIII.3.4, however, the equation in terms of y admits a unique reduced projectively equivalent equation. The precise result is as follows:

Proposition VIII.3.8. — *Let x and y be two coordinates on U, and consider a second-order linear differential equation reduced in the coordinate x:*

$$\frac{d^2v}{dx^2} + hv = 0. \qquad (E_x)$$

Then the unique equation projectively equivalent to E_x and reduced in the coordinate y has the form

$$\frac{d^2v}{dy^2} + Hv = 0, \quad \text{with } h = \left(\frac{dy}{dx}\right)^2 H + \frac{1}{2}\{y, x\}. \qquad (E_y)$$

Proof. — The set of quotients of independent solutions is common to all equations projectively equivalent to E_x. Let w be such a quotient and E equivalent to E_x. We know that the unique reduced equation equivalent to E is defined by the Schwarzian derivative of w in the coordinate of E (Corollary VIII.3.6). Hence the equation we seek is just the formula (IV.6) for the transformation of the Schwarzian derivative under a change of coordinate (see Box IV.1). □

VIII.3.3. Globalizable equations

Let S be a Riemann surface. So far we have been considering only differential equations defined on an open set U of S. Except in very exceptional circumstances, the open set U cannot be taken equal to the whole of the surface S since in general there does not exist a holomorphic coordinate x defined on the whole of S. Recall, however, that our aim is to use differential equations to solve a global problem: the uniformization of S. Hence we need to consider second-order linear differential equations with the property that the quotient of two independent solutions extends to a many-valued function defined globally on the surface S (in other words a function defined on $\widetilde{S}$). This is equivalent to considering second-order linear differential equations which "extend to the whole surface S to within projective equivalence". Or, more formally:

Definition VIII.3.9. — Let U_0 be an open set of a Riemann surface S and E_0 a second-order linear differential equation on U_0. We shall say that the equation E_0 is *globalizable* if there exist

- — open sets $U_1, \ldots, U_n$ of S such that $S = U_0 \cup \cdots \cup U_n$,
- — second-order linear differential equations $E_1, \ldots, E_n$ on the open sets $U_1, \ldots, U_n$, such that, for every pair $(i, j) \in \{1, \ldots, n\}^2$, the equations E_i and E_j are projectively equivalent when restricted to $U_i \cap U_j$.

Alert geometers among our readers will certainly have perceived that the notion of a globalizable second-order linear differential equation is closely allied to the more classical idea (for us today) of a *complex projective structure*. Recall that a *projective structure* — here assumed compatible with the complex structure — on a Riemann surface S is given by a holomorphic atlas whose charts take their values in $\mathbb{CP}^1$ and with coordinate changes on overlaps locally projective (restrictions of Möbius transformations). With each projective structure on S one associates a local biholomorphism $w : \widetilde{S} \to \mathbb{CP}^1$, termed *structure developing*, obtained by analytic continuation of a germ of a fixed chart. This map clearly depends on the initial chart: two developments differ by a Möbius transformation (acting on the codomain). Two projective structures whose developments differ by a Möbius transformation are said to be *equivalent*.

Now let U_0 be an open set of S, furnished with a coordinate. Then every globalizable second-order linear differential equation on U_0 defines a projective structure on S and conversely. One passes from the equation to the projective structure by considering quotients of solutions on small open sets (see Proposition VIII.3.2) and from the projective structure to the equation by means of a development (see Proposition VIII.3.5). Furthermore, two second-order linear differential equations on U_0 are projectively equivalent if and only if they correspond to equivalent projective structures.

Proposition VIII.3.10. — *Let S be a Riemann surface with universal cover $\pi : \widetilde{S} \to S$. Let E_0 be a second-order linear differential equation on a connected open set U_0 of S and $w_0 : \widetilde{U_0} \to \mathbb{CP}^1$ the quotient of two independent solutions of E_0. Then the equation E_0 is globalizable if and only if both of the following conditions hold:*

(i) *w_0 extends*[7] *to a local biholomorphism $w : \widetilde{S} \to \mathbb{CP}^1$;*

(ii) *for every $\gamma \in \pi_1(S)$ (viewed as an automorphism of the covering $\widetilde{S}$) there exists $\rho(\gamma) \in PSL(2,\mathbb{C})$ such that $w \circ \gamma = \rho(\gamma) \circ w$.*

Proof. — If E_0 is globalizable, then one constructs w by gluing step by step on $\widetilde{S}$ quotients w_i of solutions of the equations E_i; the compatibility of the equations E_i ensures that one can find w_is which can be glued in this way. Conversely, if w_0 satisfies (i) and (ii), then one can cover S by finitely many open sets U_i furnished with coordinates x_i, and the equations E_i globalizing E_0 are then obtained via the Schwarzian derivative $\{w, x_i\}$ of w on the open sets U_i (see Proposition VIII.3.5). □

Remark VIII.3.11. — Since condition (ii) is automatically satisfied for all $\gamma \in \pi_1(U_0)$, it becomes superfluous in the case that $\pi_1(U_0)$ maps naturally onto $\pi_1(S)$. This occurs when $S \setminus U_0$ is finite — for example if E_0 is meromorphic on S and U_0 is the complement of the set of poles of E_0.

The property of a differential equation of being globalizable is clearly invariant under projective equivalence. Thus Proposition VIII.3.4 allows us to restrict attention to equations that are reduced in a given coordinate.

Proposition VIII.3.12. — *Let U_0 be a connected open set of a Riemann surface S, x_0 a coordinate on U_0, and $h_0 : U_0 \to \mathbb{C}$ a holomorphic map. Then the reduced second-order linear differential equation*

$$\frac{d^2 v}{dx_0^2} + h_0 v = 0$$

is globalizable if and only if the following conditions hold:

— *there exist open sets $U_1, \ldots, U_n$ of S such that $S = U_0 \cup \cdots \cup U_n$,*

[7] The use of the verb "extends" here is not quite accurate. In fact $\widetilde{U_0}$ does not in general embed in $\widetilde{S}$. In order to be able to "extend" w_0 to a function defined on $\widetilde{S}$, a prior condition would be that w_0 pass via the appropriate quotient to a function defined on $\pi^{-1}(U_0) \subset \widetilde{S}$.

— *there exist holomorphic coordinates* $x_1, \ldots, x_n$ *and holomorphic functions* $h_1, \ldots, h_n$ *on the open sets* $U_1, \ldots, U_n$*, such that on each* $U_i \cap U_j$ *(*$(i,j) \in \{1, \ldots, n\}^2$*) one has*

$$h_i = \left(\frac{dx_j}{dx_i}\right)^2 h_j + \frac{1}{2}\{x_j, x_i\}.$$

Proof. — It suffices to express the equations E_i of Definition VIII.3.9 in the form $\frac{d^2v}{dx_i^2} + h_i v = 0$ and apply Proposition VIII.3.8. □

VIII.3.4. Normal equations on algebraic curves

In his article [Poin1884b] Poincaré never considers abstract Riemann surfaces; he confines himself rather to Riemann surfaces defined explicitly as algebraic curves in $\mathbb{CP}^2$ and from now on we will follow suit. Note that he is perfectly well aware that as far as uniformization is concerned only the structure of the Riemann surface in the abstract counts. It is simply that he needs to have the surfaces defined by a polynomial equation in order to "calculate" certain objects on them. Today we can of course re-derive the whole theory intrinsically — by regarding second-order linear differential equations as connections and the quotients of solutions of these equations as sections of a certain vector bundle; showing the existence of globalizable equations then reduces to establishing the triviality of a certain Čech cohomology group, this being equivalent to the condition of Proposition VIII.3.12 (see for example [Gun1967, p. 75]). However the aim of the present chapter is to resuscitate Poincaré's point of view, which, although more concrete than the modern approach, contains the seed from which it developed.

Recall that, from the point of view of abstract structures of Riemann surfaces, confining oneself to algebraic curves in $\mathbb{CP}^2$ is equivalent to restricting one's attention to compact Riemann surfaces since in fact every compact Riemann surface can be immersed in $\mathbb{CP}^2$ with image an algebraic curve (see Theorem II.1.3).

The projective plane $\mathbb{CP}^2$ is obtained from the affine plane $\mathbb{C}^2$ with coordinates (x, y) by adjoining a line at infinity. Thus we shall consider projective algebraic curves X, assumed reduced and *irreducible* (but possibly singular), given in the form

$$X := \overline{\{(x, y) \in \mathbb{C}^2 \mid F(x, y) = 0\}}$$

where $F(x, y)$ is a polynomial in the variables x and y, the closure being taken in $\mathbb{CP}^2$. The Riemann surface associated with X will be denoted by S.

Definition VIII.3.13. — Let $X \subset \mathbb{CP}^2$ be an irreducible algebraic curve and U_x the open set of X on which the first projection $x : X \to \mathbb{CP}^1$ is a holomorphic coordinate[8]. A *normal equation* on X is a second-order linear differential equation on an open set $U \subset U_x$, reduced in the coordinate x and globalizable.

Notation VIII.3.14. — We denote by $\mathcal{E}(X)$ the space of normal equations on an algebraic curve $X \subset \mathbb{CP}^2$ endowed with the following topology: equations $\frac{d^2v}{dx^2} + h_1 v = 0$ and $\frac{d^2v}{dx^2} + h_2 v = 0$ are *close* if the rational maps h_1 and h_2 are close.

VIII.3.5. Uniformizing equations

The following result provides the main motivation for studying normal equations in connection with uniformization of surfaces.

Proposition VIII.3.15. — *Suppose that the Riemann surface S associated with X is uniformizable by the half-plane. Then every global biholomorphism $w : \widetilde{S} \to \mathbb{H}$ is the quotient of two solutions of a normal equation on X.*

Proof. — By Corollary VIII.3.7, the restriction of w to U_x is the quotient of two solutions of a second-order linear differential equation on U_x reduced in the coordinate x. Proposition VIII.3.10 then shows that this equation is globalizable. It is therefore normal. □

We are thus led naturally to the following definition:

Definition VIII.3.16. — A normal equation on X will be called *uniformizing*[9] if there exist two solutions v_1, v_2 of E such that the quotient $w := v_1/v_2$ extends to a global biholomorphism between $\widetilde{S}$ and $\mathbb{H}$.

Note that a uniformizing equation is automatically globalizable. Poincaré lent great importance to the following fact:

Proposition VIII.3.17. — *There exists at most a single normal uniformizing equation on S.*

Proof. — Let E and E' be two uniformizing normal equations. There then exist two solutions v_1, v_2 of E and v_1', v_2' of E' such that the quotients $w := v_1/v_2$ and $w' := v_1'/v_2'$ extend to global biholomorphisms from $\widetilde{S}$ to $\mathbb{H}$. It follows that $w' = h \circ w$ where h is an automorphism of $\mathbb{H}$ and therefore a Möbius transformation. The equations E and E' are thus projectively equivalent and the uniqueness condition of Proposition VIII.3.4 then implies that $E = E'$. □

We are now in a position to formulate the uniformization problem for the surface S, *à la* Poincaré, in terms of linear differential equations:

To show that, among all normal equations on X, there is one that is uniformizing. If possible, to find this equation.

[8]In other words, U_x is the surface X with $x^{-1}(\infty)$ and the ramification points of the covering $x : X \to \mathbb{CP}^1$ removed.

[9]These are in fact just the equations Poincaré calls *Fuchsian equations.*

VIII.4. The set of normal equations on a fixed curve

VIII.4.1. The existence of such an equation on a given curve

Our aim in this section is to understand the structure of the set of normal equations on a given plane algebraic curve of genus $g \geq 2$.

Surprisingly (for us) Poincaré seems to consider as self-evident that every algebraic curve should support at least one normal equation. The "modern" proof of this fact consists in seeing that the obstruction given by Proposition VIII.3.12 actually has its being in a cohomology group which can be shown to vanish by means of Serre duality (see [Gun1967, p. 75]). Here we present a proof which could have been given by Poincaré — even though the reader will perhaps perceive the cohomology groups lurking between the lines!

Thus let X be a reduced, irreducible, plane algebraic curve, with affine equation $F(x, y) = 0$. We seek a normal equation on X of the form

$$\frac{d^2v}{dx^2} + h(x, y)v = 0. \qquad (E_0)$$

We may always assume that to within a birational transformation of the projective plane X is a *nodal* curve (that is, that its only singularities are ordinary double points), that $1/x$ furnishes a local coordinate at every point at infinity (so that, in particular, X is transverse to the line at infinity), and that the singularities of the projection on the x-axis are quadratic.

Proposition VIII.3.12 gives a necessary and sufficient condition for the equation to be normal. The surface S (associated with X) inherits a covering by open sets U_0, U_1, U_2 each equipped with a holomorphic coordinate (in the sense of VIII.3.10), namely x, $1/x$ and y respectively. Denote by R the divisor of S determined by the critical points of the projection on the x-axis. We need to find holomorphic functions h_i on the U_i such that

$$h_i = \left(\frac{dx_j}{dx_i}\right)^2 h_j + \frac{1}{2}\{x_j, x_i\},$$

where x_i is equal to x, $1/x$ or y as the case may be.

The compatibility condition on coordinates between U_0 and U_1 implies that $h_0 dx^2$ extends to a holomorphic quadratic differential in a neighborhood of the points of X at infinity. The compatibility condition between the coordinates of U_0 and U_2 is expressed by

$$h_0 dx^2 = h_2 dy^2 + \frac{1}{2}\{y, x\}dx^2,$$

which implies that at each point of R the differential $h_0 dx^2$ is meromorphic in the local coordinate y and that its polar part coincides with that of $\frac{1}{2}\{y,x\}dx^2$, of order 2 in y. Conversely, if a meromorphic quadratic differential on S fulfills these conditions, then its local expression $h_1 d(1/x)^2$ on U_1 also satisfies the appropriate gluing condition on $U_1 \cap U_2$ in view of the formula (IV.6) for the transformation of the Schwarzian derivative under coordinate changes (see Box IV.1).

To summarize: *a normal equation is determined by a meromorphic quadratic differential with poles and polar parts — of order 2 — prescribed (by initial choice of projective model).*

Let K be the canonical divisor of S and m a natural number. The dimension $l(2K+mR)$ of the space $\mathcal{L}(2K+mR)$ of meromorphic quadratic differentials with poles on R and of order at most m is given by the Riemann–Roch theorem (see Box II.5):

$$\begin{aligned} l(2K+mR) &= l(K-2K-mR) + \deg(2K+mR) + 1 - g \\ &= 3g - 3 + m\deg(R) \end{aligned}$$

where we have set $l(-K-mR) = 0$ (in view of the fact that $\deg(K+mR) > 0$). Hence in particular the dimension $l(2K)$ of the space of *holomorphic* quadratic differentials on S is $3(g-1)$. One has, moreover, $l(2K+2R) - l(2K) = 2\deg(R)$, whence it follows that the polar parts of an element of $\mathcal{L}(2K+2R)$ may be imposed arbitrarily. This completes the proof of the existence of a normal equation on every algebraic curve.

Proposition VIII.3.12 shows that *the normal equations naturally form an affine space associated with the vector space (of dimension* $3g-3$*) of* holomorphic *quadratic differentials on the surface.* For an algebraic curve this fact can be expressed in the following concrete form:

Proposition VIII.4.1. — *Let X be a reduced and irreducible nodal curve in $\mathbb{CP}^2$, of degree d and with affine equation $F(x,y) = 0$. Suppose further that X is transverse to the line at infinity and that the branches of its double points are transverse to the fibres of the coordinate x.*

If E_0 is a normal equation on X:

$$\frac{d^2v}{dx^2} + h_0 v = 0, \qquad (E_0)$$

then the normal equations on X are exactly those of the form

$$\frac{d^2v}{dx^2} + \left(h_0 + \frac{P}{F_y'^2}\right)v = 0,$$

where $P(x, y)$ is a polynomial of degree less than $2d - 6$ which, together with its first partial derivatives, vanishes at the double points of X (along with $F'_y = \frac{\partial F}{\partial y}$). Note furthermore that the polynomials P need be considered only modulo F.

Proof. — We need to show that every holomorphic quadratic differential on X can be expressed in terms of the coordinate x in the form

$$Q = \frac{P}{F_y'^2}\, dx^2, \tag{VIII.2}$$

where P is a polynomial satisfying the assumptions of the proposition. (The reader may also like to examine the proof of Proposition II.2.8, where the argument is similar.)

Let Q be a meromorphic quadratic differential on X. We may write it in the form (VIII.2) with P a rational function. Observe that the form $\omega = \left(\frac{\partial F}{\partial y}\right)^{-1} dx = -\left(\frac{\partial F}{\partial x}\right)^{-1} dy$ is holomorphic at the ramification points of x since these are smooth. In order for Q to be holomorphic on the affine part Y of X, it is necessary and sufficient that P be regular on Y (that is, that it be a polynomial) and vanish to the appropriate order at the double points of X. At those points the function $\left(\frac{\partial F}{\partial y}\right)^2$ vanishes to the order 2; the desired condition is thus that P, together with first derivatives, should vanish at the double points of X. Finally, since the form ω vanishes to the order $d-3$ on the line at infinity and X is transverse to that line, the form Q is holomorphic at infinity (given that we know it is holomorphic on Y) if and only if the degree of P is less than $2d - 6$. ◻

VIII.4.2. The space of normal equations on curves

Let $g \geq 0$ and $d \geq 1$ be integers. Setting $N = d(d + 3)/2$, we consider the set $\mathcal{S}_{g,d} \subset \mathbb{CP}^N$ of reduced and irreducible nodal curves of degree d and genus g. In fact $\mathcal{S}_{g,d}$ is a *smooth* manifold of dimension $3d + g - 1 = N - \delta$, where δ is the number of double points of a curve of $\mathcal{S}_{g,d}$, given by $g + \delta = (d-1)(d-2)/2$ (see Box VIII.1 below). In what follows we shall consider only curves of genus $g \geq 2$.

We shall say that a curve $X \in \mathcal{S}_{g,d}$ is *in general position* with respect to an affine coordinate system (x, y) if X is transverse to the line at infinity and the singularities of the coordinate x are of quadratic type, distinct from the double points. These conditions clearly define an open set of the manifold $\mathcal{S}_{g,d}$. The corresponding space of polynomials in (x, y) will be denoted by $\mathcal{P}_{g,d}$. Every curve $X \in \mathcal{S}_{g,d}$ admits such a system of affine coordinates. Hence as far as local questions on $\mathcal{S}_{g,d}$ are concerned, one may confine attention to curves X_F defined by polynomials $F \in \mathcal{P}_{g,d}$.

Recall (Theorem II.1.3) that every compact Riemann surface S admits a holomorphic immersion in the projective plane with image a nodal curve, which moreover one may choose to be in general position relative to an appropriate affine coordinate system; in other words, S always admits an algebraic model of the form X_F with $F \in \mathcal{P}_{g,d}$. In line with the spirit of the present chapter we tackle the uniformization of compact Riemann surfaces via algebraic curves, or, more particularly, in the framework — dear to Poincaré — of differential equations on those curves.

Notation VIII.4.2. — For a given two integers $g \geq 2$ and $d \geq 4$, we denote by $\mathcal{E}_{g,d}$ the space of pairs (X, E) with $X \in \mathcal{S}_{g,d}$ and $E \in \mathcal{E}(X)$ (see Notation VIII.3.14).

One then has, of course, the natural projection from $\mathcal{E}_{g,d}$ onto $\mathcal{S}_{g,d}$. By Proposition VIII.4.1, the fibres of this projection are complex affine spaces of dimension $3g - 3$. In fact one has the following proposition:

Proposition VIII.4.3. — *Take any $g \geq 2$ and $d \geq 4$. Then the space $\mathcal{E}_{g,d}$ is a fibration of affine spaces over $\mathcal{S}_{g,d}$. In particular $\mathcal{E}_{g,d}$ is a smooth manifold.*

Proof. — Consider the curves X_F with $F \in \mathcal{P}_{g,d}$. The ramification divisor R_F of the coordinate x is the (transverse) intersection of X_F with $F'_y = \frac{\partial F}{\partial y} = 0$, of degree $d(d-1)$. The meromorphic quadratic differentials on X_F with poles of order at most 2 at the points of R_F are of the form

$$Q = \frac{P}{F_y'^4}\, dx^2,$$

where P belongs to the space $\mathcal{P}_F$ of polynomials of degree at most $4d-8$ vanishing to the fourth order at the double points of X_F (here it is enough to adapt the proof of Proposition VIII.4.1). The polynomials P need to be considered only modulo F. Choose (locally) a complement Q_F of $\mathbb{C}[x, y]F \cap \mathcal{P}_F$ in $\mathcal{P}_F$; we know that this space has dimension $3g - 3 + 2d(d - 1)$ independently of F (see Proposition VIII.4.1). We therefore have a holomorphic vector bundle over $\mathcal{P}_{g,d}$.

The restrictions on the polar parts at the points of R_F characterizing the normal equations are affine. They depend holomorphically on F (considering $\{y, x\}$ explicitly as a rational function of the partial derivatives of F) and determine an affine subspace of Q_F of dimension $3g - 3$ independently of F (see Proposition VIII.4.1). This establishes the proposition. □

Box VIII.1: The manifold of nodal curves

Let $d \geq 1$ be an integer and $\mathbb{CP}^N$ the projectification of the space of homogeneous polynomials of degree d (so that $N = d(d + 3)/2$). Each curve of degree d in the projective plane $\mathbb{CP}^2$ can be identified with a point of $\mathbb{CP}^N$ (the equation of the curve in the usual homogeneous coordinates). Furthermore the projective coordinate changes of $\mathbb{CP}^2$ correspond to projective transformations of $\mathbb{CP}^N$.

Proposition VIII.4.4 (Severi). — *Let $g \geq 0$ and $d \geq 1$ be integers and $\mathcal{S}_{g,d} \subset \mathbb{CP}^N$ be the set of reduced and irreducible nodal curves of degree d and genus g, assumed non-empty (which is equivalent to requiring that $(d-1)(d-2) \leq 2g$ — see for example* [Loe1988, Cor. 2.2]*). Then $\mathcal{S}_{g,d}$ is a smooth manifold (a "Severi manifold") of dimension*

$$3d + g - 1 = N - \delta,$$

where δ is the number of double points of a curve from $\mathcal{S}_{g,d}$ (given by Clebsch's formula $g + \delta = (d-1)(d-2)/2$).

Proof (see [HaMo1998], p. 30]*).* — Set

$$\Sigma = \{(X,p) | X \in \mathcal{S}_{g,d}, p \in X_{sing}\} \subset \mathbb{CP}^N \times \mathbb{CP}^2,$$

where X_{sing} denotes the singular locus of X. We shall show that Σ is smooth at (X_0, p_0) provided p_0 is an ordinary double point of X_0. Choose affine coordinates (x,y) so that $p_0 = (0,0)$, and let $F_0(x,y) = 0$ be an affine equation for X_0. The condition for a point $p = (a,b)$ to belong to the singular locus of a curve X of degree d with affine equation $F(x,y) = 0$ is

$$\Phi(F,a,b) := (F(a,b), F'_x(a,b), F'_y(a,b)) = 0 \in \mathbb{C}^3.$$

Now the Jacobian matrix of Φ at the point $(F_0,0,0)$ with respect to the variables $(F(0,0),a,b)$ is

$$\begin{pmatrix} 1 & 0 \\ 0 & HF_0(0,0) \end{pmatrix},$$

where HF_0 is the Hessian matrix of F, invertible at $(0,0)$ since p_0 is an ordinary double point of X_0, whence we infer that Σ is a smooth submanifold of codimension 3 in a neighborhood of (X_0,p_0). Moreover at this point the projection $\mathbb{CP}^N \times \mathbb{CP}^2 \to \mathbb{CP}^N$ induces a local immersion from Σ to a germ of the smooth hypersurface $\mathcal{H}(p_0)$ whose tangent hyperplane corresponds to the space of polynomials vanishing at p^0.

Returning to the set $\mathcal{S}_{g,d}$, we see that for $\delta = 0$ the proposition is obvious since the smooth curves form an open set in $\mathbb{CP}^N$. Suppose now that $\delta \geq 1$.

Consider $X \in \mathcal{S}_{g,d}$ and let $p_1, \ldots, p_\delta$ be the double points of X. By the above, every curve in $\mathcal{S}_{g,d}$ sufficiently close to X belongs to the intersection of the $\mathcal{H}(p_k)$ $(k = 1, \ldots, \delta)$. Conversely, every curve $X' \in \bigcap_{k=1}^{\delta} \mathcal{H}(p_k)$ sufficiently close to X must belong to $\mathcal{S}_{g,d}$ (in particular the only singularities of X' are its δ double points close to the p_k). Hence in a neighborhood of X the set $\mathcal{S}_{g,d}$ coincides with the intersection of the hypersurfaces $\mathcal{H}(p_k)$. To complete the proof it remains to show that these are in general position, that is, that the space of polynomials vanishing at the points $p_1, \ldots, p_\delta$ has codimension δ.

Let $X \in \mathcal{S}_{g,d}$ be as before and (x, y) an affine coordinate system in general position with respect to X. Write P_m for the space of polynomials of degree $\leq m$ in (x, y). It follows from Clebsch's formula $g + \delta = (d-1)(d-2)/2$ (inferred, for instance, from the Riemann–Hurwitz formula in the case of a generic projection on $\mathbb{CP}^1$) and the characterization of holomorphic differentials on X in terms of polynomials of degree $\leq d-3$ (see the proof of Proposition II.2.8), that the conditions $P(p_k) = 0$ $(k = 1, \ldots, \delta)$ are independent on P_{d-3}. In other words, setting $\varphi_m(P) = (P(p_1), \ldots, P(p_\delta)) \in \mathbb{C}^\delta$ for $P \in P_m$, we have that the map φ_{d-3} is surjective. Hence φ_m is surjective for every $m \geq d-3$ — so in particular for $m = d$ — since its restriction to P_{d-3} is already surjective. This completes the proof. □

VIII.5. Monodromy of normal equations and uniformization of algebraic curves

In this section we explain why the set of algebraic curves supporting a uniformizing normal equation is open (from which it follows that the set of uniformizable algebraic curves is open).

VIII.5.1. The monodromy representation

Our main tool for detecting uniformizing equations will be the concept of *monodromy*. With each normal equation E on an algebraic curve $X \in \mathcal{S}_{g,d}$, we shall relate a conjugacy class of representations in $\mathrm{PSL}(2,\mathbb{C})$ of the fundamental group of the associated Riemann surface S.

For any Riemann surface S, we denote by $\mathcal{R}_{\mathbb{C}}(S)$ the space of conjugacy classes of representations of the fundamental group $\pi_1(S)$ in $\mathrm{PSL}(2,\mathbb{C})$[10]. We

[10]N. B. We use the same notation as was used in Chapter VII in connection with representations in $\mathrm{SL}(2,\mathbb{C})$ although now they are in $\mathrm{PSL}(2,\mathbb{C})$.

denote by $\mathcal{R}_{\mathbb{R}}(S)$ the subset of $\mathcal{R}_{\mathbb{C}}(S)$ consisting of the conjugacy classes of representations containing a representation in $\mathrm{PSL}(2,\mathbb{R})$. We saw earlier (Theorem VII.2.3) that $\mathcal{R}_{\mathbb{C}}(S)$ is a complex analytic manifold of complex dimension $6g-6$; similar arguments show that $\mathcal{R}_{\mathbb{R}}(S)$ is a real analytic submanifold of $\mathcal{R}_{\mathbb{C}}(S)$, of real dimension $6g-6$ (see Corollary VII.2.4).

Let X be an algebraic curve in $\mathbb{CP}^2$, S the associated Riemann surface, and E a normal equation on X. To begin with consider two independent solutions of E and denote by w the ratio of these solutions. Since the equation E is globalizable, the function w (defined *a priori* on the universal cover of an open set of S) can be analytically continued to a function defined on the universal cover of S (Proposition VIII.3.10). By Proposition VIII.3.1, for every $\gamma \in \pi_1(S)$ there then exists a Möbius transformation $\rho(\gamma) \in \mathrm{PSL}(2,\mathbb{C})$ such that $w(\gamma.z) = \rho(\gamma) \circ w(z)$. This defines a representation

$$\rho : \pi_1(S) \to \mathrm{PSL}(2,\mathbb{C}),$$

the *monodromy representation*. Consider now two other independent solutions of E, and denote by w' their quotient and by ρ' the monodromy representation of w'. By Proposition VIII.3.1, there exists a Möbius transformation $m \in \mathrm{PSL}(2,\mathbb{C})$ such that $w' = m \circ w$. We have therefore $\rho' = m \circ \rho \circ m^{-1}$; thus in particular the monodromy representations of w and w' are conjugate. This justifies the following definition: we call the single conjugacy class of monodromy representations of quotients of pairs of independent solutions of the equation E the *monodromy of the equation E*, denoted by

$$\mathrm{Mon}_X(E) \in \mathcal{R}_{\mathbb{C}}(S).$$

Remark VIII.5.1 (fundamental). — *If an equation E uniformizes S by means of the half-plane $\mathbb{H}$, then its monodromy is real.* Indeed, under this assumption there exist two solutions of E whose quotient defines a global biholomorphism from $\widetilde{S}$ to $\mathbb{H}$. Thus the associated monodromy representation takes its values in the automorphism group of $\mathbb{H}$, that is, in $\mathrm{PSL}(2,\mathbb{R})$, and the monodromy $\mathrm{Mon}_X(E)$ belongs to the submanifold $\mathcal{R}_{\mathbb{R}}(S)$.

Note that the converse is "almost true": if X is close to a uniformizable curve and if there exists a normal equation E on X with monodromy $\mathrm{Mon}_X(E)$ belonging to the submanifold $\mathcal{R}_{\mathbb{R}}(S)$, then E is uniformizing for X; in particular, X is uniformizable[11].

Thus for each algebraic curve $X \in \mathcal{S}_{g,d}$ we now have a monodromy map defined on the space of normal equations on X. But this is not enough! We

[11]However there are normal equations with monodromy in $\mathcal{R}_{\mathbb{R}}(S)$ (and even with Fuchsian monodromy) that are not uniformizing; see [GolW1987].

need a monodromy map defined on the fibration $\mathcal{E}_{g,d}$ of normal equations over all algebraic curves, or at least defined in a neighborhood of a given fibre.

Fix on integers $g \geq 2$ and $d \geq 4$. Recall that $\mathcal{P}_{g,d}$ denotes the set of polynomials $F \in \mathbb{C}[x, y]$ of degree d such that the projective curve X_F defined by F is nodal and in general position with respect to the coordinates (x, y) (see §VIII.4.2). For any $F_0 \in \mathcal{P}_{g,d}$, we have by the Tubular Neighborhood Theorem (for immersed submanifolds) that there exists a neighborhood $\mathcal{U}_0$ of F_0 in $\mathcal{P}_{g,d}$ and a smooth map $\Phi : \mathcal{U}_0 \times X_{F_0} \to \mathbb{CP}^2$ such that for every $F \in \mathcal{U}_0$, the map $\Phi(F, .)$ is a diffeomorphism from X_{F_0} to X_F. Although the map Φ is not itself unique, the homotopy class of maps $\Phi(F, .)$ is well-defined and for F sufficiently close to F_0 affords an identification of the fundamental group of the associated surface S_F with that of the surface S_{F_0}.

For $F \in \mathcal{U}_0$ and $E \in \mathcal{E}(X_F)$ (see Notation VIII.4.2), one may therefore consider the monodromy $\mathrm{Mon}_{X_F}(E)$ as an element of the manifold $\mathcal{R}_{\mathbb{C}}(S_{F_0})$. Writing $\mathcal{E}_{\mathcal{U}_0} := \{(F, E) \mid F \in \mathcal{U}_0, E \in \mathcal{E}(X_F)\}$, we therefore have a map

$$\begin{aligned} \mathrm{Mon} : \ \mathcal{E}_{\mathcal{U}_0} &\longrightarrow \mathcal{R}_{\mathbb{C}}(S_{F_0}) \\ (F, E) &\longmapsto \mathrm{Mon}_{X_F}(E). \end{aligned} \tag{VIII.3}$$

Recall that $\mathcal{E}_{\mathcal{U}_0}$ is a fibre bundle of affine spaces over the open set $\mathcal{U}_0$ (Proposition VIII.4.3).

Proposition VIII.5.2. — *The map Mon*: $\mathcal{E}_{\mathcal{U}_0} \longrightarrow \mathcal{R}_{\mathbb{C}}(S_{F_0})$ *is holomorphic.*

Proof. — This follows from the theorem concerning the holomorphic dependence of the solutions of a linear differential equation on the coefficients of that equation. □

VIII.5.2. The set of uniformizable Riemann surfaces is open

The aim of the long memoir [Poin1884b], which Poincaré published in 1884, is to show, by means of the method of continuity that he had conceived simultaneously with Klein, that algebraic curves are all uniformizable.

Theorem VIII.5.3. — *Every compact Riemann surface of genus g greater than or equal to 2 is uniformized by the upper half-plane.*

Recall that for each fixed integer $g \geq 2$ the symbol $\mathcal{M}_g$ denotes the moduli space of compact Riemann surfaces of genus g (see Proposition II.3.1). Recall also that the method of continuity consists in the following:

— observing that $\mathcal{M}_g$ is connected and that there exists at least one point of $\mathcal{M}_g$ corresponding to a uniformizable Riemann surface;

— showing that the set of points of $\mathcal{M}_g$ corresponding to uniformizable Riemann surfaces is both open and closed in $\mathcal{M}_g$.

This clearly suffices to establish that every compact Riemann surface of genus g is uniformizable.

Poincaré seems to consider it evident that the moduli space $\mathcal{M}_g$ is arcwise connected; on this point the reader may consult Chapter II, in particular Proposition II.3.1. The existence of at least one uniformizable Riemann surface of genus g follows easily from Poincaré's work on Fuchsian groups, expounded in §VI.2.3.

We shall not enter here into the arguments Poincaré uses to establish closure, except to remark that he fully appreciated the difficulties involved in this, and that in some sense one may regard the proof given in Chapter VII as a "putting in order" of his attempt at a proof. The particular case of the sphere with four points removed is instructive: here he produced a perfectly rigorous proof — apparently as the result of concentrated effort. (We give the details in the next chapter.) However that may be, we would say that the creation of the tools necessary for a correct proof of closure in general lay far in the future of the researchers of that era, even such a one as Poincaré.

Thus here we rest content with a treatment of the openness (within the framework of algebraic curves) closely following Poincaré's method, which requires modifications only of minor points to make it fully rigorous.

Proposition VIII.5.4. — *Let $\mathcal{S}_{g,d}$ be the manifold of curves of genus $g \geq 2$ and degree $d \geq 4$. Then the set of curves $X \in \mathcal{S}_{g,d}$ uniformizable by the upper half-plane is open in $\mathcal{S}_{g,d}$.*

Proof. — Let $F_0 \in \mathcal{P}_{g,d}$ be such that the Riemann surface S_{F_0} is uniformizable by the upper half-plane. We need to find a neighborhood $\mathcal{U}$ of F_0 in $\mathcal{P}_{g,d}$ such that the Riemann surface S_F is uniformizable for every $F \in \mathcal{U}$. In §VIII.5.1 we defined a certain neighborhood $\mathcal{U}_0$ of F_0 in $\mathcal{P}_{g,d}$ for which the space

$$\mathcal{E}_{\mathcal{U}_0} = \{(F,E) \mid F \in \mathcal{U}_0 \ , \ E \in \mathcal{E}(X_F)\}$$

is a fibre bundle of affine spaces over the open set $\mathcal{U}_0$, and the "monodromy" map (see (VIII.3)) from $\mathcal{E}_{\mathcal{U}_0}$ to $\mathcal{R}_{\mathbb{C}}(S_{F_0})$ is holomorphic.

Writing E_0 for the uniformizing equation of the surface S_{F_0}, we choose two solutions of E_0 whose quotient w_0 defines a biholomorphism from $\widetilde{S}_{F_0}$ to $\mathbb{H}$, and denote by $\rho_0 : \pi_1(S_{F_0}) \to \mathrm{PSL}(2,\mathbb{R})$ the monodromy representation of w_0. We then have $\mathrm{Mon}_{X_{F_0}}(E_0) = [\rho_0] \in \mathcal{R}_{\mathbb{R}}(S_{F_0})$ (see Remark VIII.5.1).

Recall that $\mathcal{R}_{\mathbb{C}}(S_{F_0})$ is a smooth complex manifold of complex dimension $6g-6$ (or real dimension $12g-12$) in a neighborhood of ρ_0, and that $\mathcal{R}_{\mathbb{R}}(S_{F_0})$

is a smooth real submanifold of real dimension $6g - 6$. The following point is crucial[12].

Lemma VIII.5.5. — *The map* $Mon_{X_{F_0}} : \mathcal{E}(X_{F_0}) \to \mathcal{R}_{\mathbb{C}}(S_{F_0})$ *is transverse at* E_0 *to the subvariety* $\mathcal{R}_{\mathbb{R}}(S_{F_0})$.

We shall assume this result for the moment and use it to complete the proof of Proposition VIII.5.4. Thus by this lemma, the inverse image of $\mathcal{R}_{\mathbb{R}}(S_{F_0})$ under the map Mon defines a germ of a real smooth manifold Σ passing through E_0 and of (real) codimension $6g-6$. This submanifold is transverse to the fibre $\mathcal{E}(X_{F_0})$ since its codimension is equal to the dimension of $\mathcal{E}(X_{F_0})$. Hence for F sufficiently close to F_0 there exists a normal equation E_F (depending smoothly on F) on the curve X_F with monodromy in $\mathcal{R}_{\mathbb{R}}(S_{F_0})$. We now show that E_F is uniformizing.

Let ξ_F be a germ of a projective chart with values in $\mathbb{CP}^1$ and defined on an open set of X_F by a quotient of solutions of E_F (see §VIII.3.3). We may choose ξ_{F_0} to correspond to w_0 and assume that ξ_F depends smoothly on F (in view of the dependence of solutions on the parameters of the equation). In what follows we shall suppose the polynomial F to be as close to F_0 as need be. The many-valued analytic continuation of ξ_F will then be C^0-close to ξ_{F_0} on compact sets. The single-valued version of this affirms that ξ_F defines a local biholomorphism $w_F : \widetilde{S}_F \to \mathbb{CP}^1$, and we can find a compact fundamental region D_F of the universal cover $\widetilde{S}_F \to S_F$ such that $w_F(D_F)$ is Hausdorff close to $w_0(D_{F_0})$. Since $w_0(D_{F_0})$ is a compact set contained in the half-plane $\mathbb{H}$, we also have $w_F(D_F)$ contained in $\mathbb{H}$. In view of the equivariance of w_F (Proposition VIII.3.1) and the fact that the monodromy of E_F is real, this entails that $w_F(\widetilde{S}_F)$ *is contained in* $\mathbb{H}$.

Now let h denote the usual hyperbolic metric on the half-plane $\mathbb{H}$. Since w_F takes its values in $\mathbb{H}$, we can pull h back to a metric $\tilde{g} = w_F^* h$ on $\widetilde{S}_F$, which by construction then induces a (hyperbolic) metric g on S_F. Recall that if M and N are Riemannian manifolds, with N complete, then a local isometry from M to N is a covering map if and only if M is also complete. Here the metric g is complete since S_F is compact, whence we infer that $\tilde{g}$ is also complete, so that w_F is a covering map. Hence w_F is a biholomorphism from $\widetilde{S}_F$ to $\mathbb{H}$ and E_F is uniformizing. □

In order to complete the proof of Proposition VIII.5.4, it remains to prove the transversality lemma.

[12]In connection with this lemma we should mention Klein's and Poincaré's claim, made without proof, that a certain "functional determinant" is non-zero, signifying transversality.

Commentary. — A crucial point in the following proof concerns the equality of the (real) codimension of $\mathcal{R}_{\mathbb{R}}(g)$ in $\mathcal{R}_{\mathbb{C}}(g)$ and the dimension of the space of quadratic differentials. It is of course not surprising that the dimension of $\mathcal{R}_{\mathbb{C}}(g)$ is twice that of $\mathcal{R}_{\mathbb{R}}(g)$ since the first space is just the complexification of the second. Thus it is a question of understanding why the space of quadratic differentials and the space $\mathcal{R}_{\mathbb{R}}(g)$ have the same dimension. The first of these dimensions is calculated using the Riemann–Roch theorem and the second by counting generators and relations — two independent calculations yielding $6g - 6$ real dimensions but leaving it a mystery as to the reason. (To put it in other terms, the moduli space of curves of genus g has the same dimension as the space of quadratic differentials on a given curve.) Poincaré seems not to have been surprised by this coincidence. A modern way of making it "clear" is as follows: A holomorphic quadratic differential is a section of the double of the canonical divisor K. According to a general principle, an infinitesimal deformation of the complex structure on a curve S is parametrized by an element of the first cohomology group of S with values in the sheaf of holomorphic vector fields, which is to say $-K$. These two spaces are then dual by Serre duality.

VIII.5.3. Proof of the transversality Lemma VIII.5.5

Poincaré does not prove this lemma (and in any case the concept of tranversality had of course as of then not yet settled out)! On the other hand he proves in detail the above lemma according to which a given algebraic curve possesses at most one uniformizing equation, whence, in particular, it follows that in a neighborhood of the uniformizing equation E_0 the image of the map $\mathrm{Mon}_{X_{F_0}} : \mathcal{E}(X_{F_0}) \to \mathcal{R}_{\mathbb{C}}(S_{F_0})$ meets the submanifold $\mathcal{R}_{\mathbb{R}}(S_{F_0})$ in a single point. Of course this does not immediately entail transversality at this point of intersection, yet Poincaré seems to make the leap without hesitation[13]. We now propose giving a proof of transversality by means of methods that Poincaré might have used (as it seems to us).

We begin with a preliminary calculation. Let w be a biholomorphism between two open sets of the upper half-plane and 2φ the logarithm of the Jacobian of w in the hyperbolic metric. We first establish the equation

$$\Delta_{\mathbb{H}}(\varphi) = \exp(2\varphi) - 1, \tag{VIII.4}$$

where $\Delta_{\mathbb{H}}$ denotes the hyperbolic Laplacian in $\mathbb{H}$.

Writing s and t for the real and imaginary parts of a point z of $\mathbb{H}$, we have that the Laplacian for the hyperbolic metric $(ds^2 + dt^2)/t^2$ is given by $\Delta_{\mathbb{H}} = t^2\Delta$,

[13] "Is it because the functional determinant of the coordinates of μ with respect to those of δ vanish? But this can never happen since the lemma of Section VII shows that to every point μ there can correspond only a single point δ" [Poin1884b, p. 370].

where Δ is the Euclidean Laplacian. Setting $v = \mathrm{Im}(w)$, we have the Jacobian of w in the hyperbolic metric in the form $|w'(z)|^2 t^2/v^2$. Its logarithm 2φ is therefore given by

$$\varphi = \log|w'(z)| + \log t - \log v.$$

Here the first term is harmonic (as the real part of a holomorphic function). Furthermore

$$\Delta(\log v) = \frac{1}{v}\Delta(v) - \frac{1}{v^2}|\mathrm{grad} v|^2 = -\frac{1}{v^2}|\mathrm{grad} v|^2 = \frac{1}{v^2}|w'(z)|^2$$

since v is also harmonic. Finally, we obtain the desired equation (VIII.4): $\Delta_{\mathbb{H}}(\varphi) = t^2(|w'(z)|^2/v^2 - 1/t^2) = \exp(2\varphi) - 1$.

Remark VIII.5.6. — In order to motivate what follows, we make a few remarks on matters Riemannian. The equation (VIII.4) is a special case of the formula linking the curvatures K_1 and K_2 of two conformal metrics g_1 and g_2 on a surface, namely, if $g_2 = \exp(2\varphi)g_1$, then

$$K_2 = \exp(-2\varphi)(K_1 - \Delta_{g_1}(\varphi)),$$

where Δ_{g_1} is the Laplacian of the metric g_1. Then in order to obtain (VIII.4), one takes the usual hyperbolic metric and its inverse image under w, both of curvature -1. Even if this line of thought was not actually present to the mind of Poincaré, we shall see in Chapter X that he was very familiar with such formulae, at least in the case where g_1 and g_2 are of constant curvature.

The question of the uniqueness of the uniformizing equation on an algebraic curve X can be reformulated in terms of Riemannian metrics on the surface S associated with X. Thus two uniformizing equations furnish biholomorphisms w_i from $\widetilde{S}$ to $\mathbb{H}$ and hyperbolic metrics g_i on S, $i = 1, 2$ (stemming, as above, from the hyperbolic metric of $\mathbb{H}$). The metrics g_1 and g_2 determine the complex structure of S, so belong to the same conformal class. The projective equivalence of the uniformizing equations translates [14] as $g_1 = g_2$. From the point of view of metrics on S, the uniqueness is thus equivalent to the fact that there exists at most one metric of curvature -1 in a given conformal class, that is, that there is no nonzero function φ on a compact surface that solves the preceding equation. This last point can be argued directly as follows: Such a function φ must change sign since the integral of a Laplacian is zero; at a point where φ attained its maximum, the Laplacian is zero or negative while the second term is positive, an absurdity.

We are now in a position to conclude the proof of the transversality lemma.

[14] Via the identity $\mathrm{Isom}^+(\mathbb{H}) = \mathrm{PSL}(2,\mathbb{R})$, inferred from $\mathrm{PSL}(2,\mathbb{R}) = \mathrm{Aut}(\mathbb{H})$, which clearly already settles the question of uniqueness!

Thus consider the uniformizing normal equation for the curve X_{F_0}:

$$\frac{d^2v}{dx^2} + h_0 v = 0, \tag{E_0}$$

and a vector tangent to $\mathcal{E}(X_{F_0})$ at the point E_0, which is to say tangent at $\varepsilon = 0$ to a certain curve in the space of normal differential equations of the form

$$\frac{d^2v}{dx^2} + (h_0 + \varepsilon q)\, v = 0, \tag{E_ε}$$

where $q(x, y)dx^2$ defines a holomorphic quadratic differential on S_{F_0}. Via the monodromy map one obtains a curve in $\mathcal{R}_{\mathbb{C}}(S_{F_0})$ parametrized by ε. We wish to show that, if this curve is tangent to the real submanifold $\mathcal{R}_{\mathbb{R}}(S_{F_0})$ at $\varepsilon = 0$, then q is identically zero.

Since E_0 is uniformizing, we have an identification between $\widetilde{S}$ and the upper half-plane $\mathbb{H}$ and between the fundamental group of S and a discrete group Γ of isometries of $\mathbb{H}$. Let z be the usual coordinate on $\mathbb{C}$, that is, for which $\mathbb{H}$ is defined by $\mathrm{Im}(z) > 0$. As always, the quotient of two solutions of the differential equation E_ε furnishes a local biholomorphism w_ε from $\widetilde{S} = \mathbb{H}$ to $\mathbb{CP}^1$, well-defined up to a projective transformation of the codomain. We may assume that w_0 is the identity map.

The quadratic differential $\{w_\varepsilon, z\}dz^2$ is equal to $\varepsilon q dx^2$. To see this, observe first that by Proposition VIII.3.5 the Schwarzian derivative of w with respect to the coordinate x is equal to $h_0+\varepsilon q$. But then, by the same proposition, the Schwarzian derivative $\{z, x\}$ is equal to h_0 since the identity function $w_0(z) = z$ is the quotient of two solutions of the equation E_0. The equality $\{w_\varepsilon, z\}dz^2 = \varepsilon q dx^2$ now follows from the formula (IV.6). Thus in order to show that q vanishes identically, it suffices to prove that the derivative of $\{w_\varepsilon, z\}$ with respect to the parameter ε is zero at $\varepsilon = 0$.

By definition of the monodromy representation, for each element $\gamma \in \Gamma$, the fundamental group of S, there is a Möbius transformation depending on ε such that:

$$w_\varepsilon\,(\gamma(z)) = \frac{a(\gamma, \varepsilon)w_\varepsilon(z) + b(\gamma, \varepsilon)}{c(\gamma, \varepsilon)w_\varepsilon(z) + d(\gamma, \varepsilon)}.$$

For $\varepsilon = 0$ the numbers $a(\gamma, \varepsilon), b(\gamma, \varepsilon), c(\gamma, \varepsilon), d(\gamma, \varepsilon)$ are real. By assumption we may choose the w_ε in such a way that the derivatives of these numbers at $\varepsilon = 0$ are also real. To see this it suffices to recall that w_ε is defined only up to a Möbius transformation acting on the codomain, and that the monodromy representation is defined only up to conjugation.

Of course w_ε does not necessarily preserve $\mathbb{H}$; however, for every compact subset K of the half-plane the compact subset $w_\varepsilon(K)$ is contained in $\mathbb{H}$ for ε sufficiently small. We may therefore consider the functions $2\varphi_\varepsilon$ — logarithms of the hyperbolic Jacobians of w_ε — these being defined in a neighborhood of a given point for sufficiently small ε. By the equation (VIII.4) we have $\Delta_{\mathbb{H}}(\varphi_\varepsilon) = \exp(2\varphi_\varepsilon) - 1$. Setting $\psi = \frac{d}{d\varepsilon}_{|\varepsilon=0}\varphi_\varepsilon : \mathbb{H} \to \mathbb{R}$, and differentiating the preceding equation, we obtain

$$\Delta_{\mathbb{H}}(\psi) = 2\psi.$$

We claim that *ψ is invariant under the action of* Γ. To see this, note first that since the numbers $a(\gamma,\varepsilon), b(\gamma,\varepsilon), c(\gamma,\varepsilon), d(\gamma,\varepsilon)$ are real to the first order, the hyperbolic Jacobian of the corresponding Möbius transformation is equal to 1 to the first order. It follows that

$$\varphi_\varepsilon(\gamma(z)) = \varphi_\varepsilon(z) + O(\varepsilon^2)$$

for every γ, moreover uniformly on every compact subset of $\mathbb{H}$. Differentiating with respect to ε at 0, we obtain the desired Γ-invariance of ψ. Hence ψ induces a function, which we also denote by ψ, on the compact surface S, and then as earlier (see the conclusion of Remark VIII.5.6), by examining the sign of the Laplacian at the extrema of ψ, one sees that ψ must vanish identically. In other words, we have established that w_ε preserves the hyperbolic metric up to the order $O(\varepsilon^2)$, uniformly on every compact subset.

It remains to show that *the Schwarzian derivative of w_ε is also of order $O(\varepsilon^2)$ uniformly on every compact subset.* To this end it is convenient to go over to the unit disc model of the hyperbolic plane: $\mathbb{D} = \{|z| < 1\}$. Consider any point $z_0 \in \mathbb{H}$, and choose a compact neighborhood K of z_0 such that $w_\varepsilon(K) \subset \mathbb{H}$ (with ε sufficiently small), and Möbius transformations $f, g_\varepsilon : \mathbb{D} \to \mathbb{H}$ such that $f(0) = z_0$, $g_\varepsilon(0) = w_\varepsilon(z_0)$ with g_ε a smooth function of ε. We take $K = f(\mathbb{D}_{1/2})$ where $\mathbb{D}_{1/2} = \{|z| \le 1/2\}$. We now replace w_ε by the function $g_\varepsilon^{-1} \circ w_\varepsilon \circ f$, while continuing to denote it by w_ε. This (new) function w_ε fixes 0, and the quadratic differential $\{w_\varepsilon, x\}dx^2$ is left unchanged (since the functions f and g_ε represent projective coordinate changes). Furthermore, f and g_ε are isometries for the hyperbolic metrics on $\mathbb{H}$ and $\mathbb{D}$, so that w_ε preserves the hyperbolic metric up to the order $O(\varepsilon^2)$, uniformly on compact subsets, so in particular on the disc $\mathbb{D}_{1/2}$.

We now claim that *this entails that, always within the disc $\mathbb{D}_{1/2}$, the distance between w_ε and a certain rotation about 0 as centre (depending on ε) is of order $O(\varepsilon^2)$*. To see this, observe first that the image under w_ε of a radius joining the origin to a point of the circle $C_{1/2} = \partial\mathbb{D}_{1/2}$ has hyperbolic length different from its Euclidean length only by an amount $O(\varepsilon^2)$. It follows that the image $w_\varepsilon(\mathbb{D}_{1/2})$ is contained in a disc of radius $1/2 + O(\varepsilon^2)$. By noting that for small ε

the restriction of w_ε to $\mathbb{D}_{1/2}$ is a diffeomorphism onto its image, and applying the same reasoning as before to the inverse of w_ε, we infer that $w_\varepsilon(\mathbb{D}_{1/2})$ is contained in the annulus between two discs of radii $1/2 - O(\varepsilon^2)$ and $1/2 + O(\varepsilon^2)$. Schwarz's classical lemma then implies that $w'_\varepsilon(0)$ has modulus $1 + O(\varepsilon^2)$.

Now consider the restriction of w_ε to the circle $C_{1/2}$. Its image is a curve contained in an annulus of width $O(\varepsilon^2)$ around $C_{1/2}$. Hence the radial projection on the circle $C_{1/2}$ furnishes, for sufficiently small ε, a diffeomorphism of the circle with derivative majorized by $1 + O(\varepsilon^2)$. This diffeomorphism therefore differs from a rotation by $O(\varepsilon^2)$. We have thus shown that w_ε differs from a rotation by an amount $O(\varepsilon^2)$ on the boundary of the disc $\mathbb{D}_{1/2}$, whence this holds also throughout the disc by the maximum principle. Cauchy's formula then shows that the second and third derivatives of w_ε at the origin are of order $O(\varepsilon^2)$, so that the Schwarzian derivative $\{w_\varepsilon, z\}$ at the origin is of order $O(\varepsilon^3)$.

Returning to the upper half-plane, we see that the derivative with respect to ε, at $\varepsilon = 0$, of the Schwarzian derivative $\{w_\varepsilon, z\}$ is zero at every point z_0 of the half-plane. As we have seen, this is equivalent to the fact that *the holomorphic quadratic differential q vanishes identically*, which we wished to prove. □

Chapter IX

Examples and further developments

We begin the chapter with a detailed exposition of Schwarz's work [Schw1873] on the hypergeometric equation — work which led him to the famous list of values of the parameters for which the solutions are algebraic. This precursory work contains the seeds of many of the ideas later developed by Klein and Poincaré. Then we examine in detail normal equations on certain particular algebraic curves; this leads us to revisit certain classical families of differential equations. These depend on certain "accessory parameters" of which the values yielding uniformizing equations are known only in certain exceptional cases. As Schwarz himself realized in light of the subsequent work of Klein and Poincaré, it follows incidentally that the general solution of the hypergeometric equation allows one to uniformize many algebraic curves. Next, following Poincaré's lead, we apply the method of continuity in a direct and elementary manner to the case of the sphere with four points removed. The chapter concludes by evoking some of the consequences of the method of continuity. For more on this theme the reader may like to consult [Gra1986].

IX.1. Fuchs's local theory

As we have noted on several occasions, if one is interested only in uniformizing smooth compact curves, one is obliged, if only for computational reasons, to consider differential equations with poles, corresponding to singular projective structures. This is the case, for instance, when one seeks to express the uniformizing equation of a curve with equation $F(x, y) = 0$ in terms of the variable x. Otherwise one has to resort to the following result of Fuchs, which inpressed Poincaré sufficiently for him to feel the terms "Fuchsian functions" and "Fuchsian groups" justified. (The result in question was reproved slightly later by Schwarz.)

A linear differential equation

$$\frac{d^2v}{dx^2} + f\frac{dv}{dx} + gv = 0 \tag{E}$$

with meromorphic coefficients is called *Fuchsian* at a point $x = x_0$ if at that point f and g have at worst only poles of orders 1 and 2 respectively. This is equivalent to requiring that the associated reduced equation

$$\frac{d^2v}{dx^2} + hv = 0, \qquad (E'),$$

where $h = g - \frac{1}{2}\frac{df}{dx} - \frac{1}{4}f^2$, has at worst a double pole at x_0. We then say that the projective structure induced in a punctured neighborhood of x_0 possesses a *Fuchsian singularity* at x_0. Note that (E) and (E') are projectively equivalent, as defined in §VIII.3.2, only on such a punctured neighborhood, since neither (E) nor (E') has a solution at x_0. According to a well-known result of Fuchs, equations with Fuchsian singularities are characterized among meromorphic second-order equations by the fact that their solutions have moderate growth in a neighborhood of their singular points (on sectors). However, it is a different result of Fuchs that interests us here.

We wish to describe the type of singularities presented by the charts w of the projective structure induced around x_0, as well as their monodromies around x_0. Recall that by Proposition VIII.3.5 such a chart is determined by the quotient $w = v_1/v_2$ of two independent solutions of (E) around x_0, or, equivalently, as a solution of the Schwarzian equation $\{w, x\} = 2h$, where h is the coefficient of the associated reduced equation (E'). Thus the problem we are faced with is that of solving the Schwarzian equation $\{w, x\} = 2h$ in a neighborhood of a double pole x_0 of h.

If $y(x)$ is another coordinate, sending the point x_0 to the point $y_0 = y(x_0)$, then the new Schwarzian equation $\{w, y\} = 2H$, as given by the change-of-coordinate formula (see Box IV.1), will still present a double pole at y_0. Furthermore the dominant coefficient λ, defined by

$$\{w, x\} = \frac{\lambda}{(x - x_0)^2} + \frac{\mu}{x - x_0} + O(1),$$

remains unchanged:

$$\{w, y\} = \frac{\lambda}{(y - y_0)^2} + \frac{\tilde{\mu}}{y - y_0} + O(1).$$

This is the residue of the projective structure at the singular point. It is calculated in terms of the differential equation

$$\frac{d^2v}{dx^2} + \left(\frac{\lambda_1}{x - x_0} + O(1)\right)\frac{dv}{dx} + \frac{1}{2}\left(\frac{\lambda_2}{(x - x_0)^2} + \frac{\mu}{x - x_0} + O(1)\right)v = 0 \tag{E}$$

using the formula

$$\lambda = \lambda_2 + \frac{1-(\lambda_1-1)^2}{2}.$$

One then defines the *index* θ up to sign by

$$\lambda = \frac{1-\theta^2}{2}.$$

We remark in passing that when the coefficient g in the equation (E) has only a simple pole, the index $\theta = \lambda_1 - 1$ is given directly as the usual residue of f. The Fuchs–Schwarz result is then:

Theorem IX.1.1. — *The Schwarzian equation*

$$\{w,x\} = \frac{1-\theta^2}{2(x-x_0)^2} + \frac{\mu}{x-x_0} + O(1)$$

has as a particular solution around x_0

— *either* $w(x) = y^\theta$,

— *or* $w(x) = \frac{1}{y^n} + \log y$, *in which case* $\theta = \pm n$ *(*$n \in \mathbb{N}$*),*

where $y(x)$ *is a local coordinate at* x_0, $y(x_0) = 0$.

When θ is not an integer we are in the first case and every other solution of the Schwarzian equation is then clearly of the form $w(x) = \frac{ay^\theta+b}{cy^\theta+d}$, $ad - bc \neq 0$ (see Box IV.1). In particular, $w(x) = y^{-\theta}$ is also a solution, which is consistent with the fact that θ is determined by the equation only up to sign. The monodromy around x_0 is given by multiplication by $e^{2i\pi\theta}$. A basis for the solutions of the reduced equation (E') is afforded by $v(x) = y^{\frac{1\pm\theta}{2}}$.

On the other hand, when θ is an integer, say $\theta = n \in \mathbb{N}$, then there exists a local coordinate $y(x)$ at x_0 in terms of which the projective structure is defined

— either by the chart $w = y^n$ (or $w = \frac{1}{y^n}$), in which case the monodromy is trivial,

— or by the chart $w(x) = \frac{1}{y^n} + \log y$, in which case the monodromy is a non-trivial translation of the form $w(e^{2i\pi}y) = w(y) + 2i\pi$.

The first, exceptional, case is characterized as in the following proposition:

Proposition IX.1.2. — *Under the assumptions of Theorem IX.1.1 and with $\theta = n$ (an integer), the following statements are equivalent:*

— *there exists a local coordinate y for which $w = y^n$ is a solution of the Schwarzian equation;*

— *every solution $w(x)$ of the Schwarzian equation is single-valued on a punctured neighborhood of x_0 — in other words the monodromy is trivial;*

— *there exists a local coordinate y for which the equation (E) is projectively equivalent to*

$$\frac{d^2v}{dy^2} + \frac{1-n^2}{4y^2}v = 0;$$

— *there exists a local coordinate y in terms of which*

$$\frac{dy}{dx}\{w,x\} + \frac{1}{2}\{y,x\} = \frac{1-n^2}{2x^2};$$

— *there exists a local coordinate y in terms of which*

$$\frac{dy}{dx}\{w,x\} + \frac{1}{2}\{y,x\} = \frac{1-n^2}{2x^2} + O(x^{n-1}).$$

We shall say that the singularity is *apparent* in the last case, otherwise *logarithmic*.

A direct formal calculation using the above characterization yields the following list of possibilities up to $n = 4$ for the singularity of the Schwarzian equation

$$\{w,x\} = \frac{1-n^2}{2(x-x_0)^2} + \frac{\mu}{x-x_0} + \mu_0 + \mu_1(x-x_0) + \mu_2(x-x_0)^2 + O((x-x_0)^3) :$$

— $n = 0$, in which case it is always logarithmic;

— $n = 1$, in which case it is apparent if and only if $\mu = 0$ (that is, if and only if it is holomorphic since then $\lambda = 0$);

— $n = 2$, in which case it is apparent if and only if $\mu^2 + 2\mu_0 = 0$;

— $n = 3$, in which case it is apparent if and only if $\mu^3 + 8\mu\mu_0 + 16\mu_1 = 0$;

— $n = 4$, in which case it is apparent if and only if $\mu^4 + 20\mu^2\mu_0 + 36\mu_0^2 + 96\mu\mu_1 + 288\mu_2 = 0$.

The proofs of Theorem IX.1.1 and Proposition IX.1.2 consist in first finding a formal coordinate change $y(x)$ yielding one of the solutions mentioned in the statement of the theorem and then establishing its convergence by means of dominant series. (Alternatively, one might simply cite the earlier theorem of Briot–Bouquet, proved by similar means.) We propose here to simplify our task by using without proof the fact of the moderate growth of the solutions of (E) (also established by Fuchs), today considered classical (see for example [Hil1976]).

Idea of the proof. — Since the equation (E) is Fuchsian, its (many-valued) solutions v near x_0 have moderate growth at x_0, that is, satisfy

$$|v(x-x_0)| \leq C|x-x_0|^M$$

for constants $C, M > 0$ provided certain restrictions are observed — such as, for instance, confining attention to a sector $\{-\alpha < \arg(x-x_0) < \alpha\}$. The situation will be similar for every projective chart $w = v_1/v_2$: its monodromy around x_0 will be given by a Möbius transformation. To within a change to another projective chart, we may assume the monodromy to have the form

$$w(e^{2i\pi}x) = aw(x) \quad \text{or} \quad w(x)+b$$

(where we suppose $x_0 = 0$ for the sake of simplicity).

First case: $w(e^{2i\pi}x) = aw(x)$. The differential form $\frac{dw}{w}$ is also of moderate growth at 0, and is well-defined (single-valued) so extends meromorphically to 0. In fact it must have a simple pole there since otherwise w would have exponential growth at 0. We may therefore write

$$\frac{dw}{w} = \theta\frac{dx}{x} + df,$$

where θ is the residue of $\frac{dw}{w}$ and $f(x)$ is holomorphic at 0. Integrating, we obtain $w = x^\theta \exp f = y^\theta$ with $y = x\exp\frac{f}{\theta}$. Hence, finally, $\{w, y\} = \frac{1-\theta^2}{2y^2}$. (The case $\theta = 0$ needs to be considered separately.)

Second case: $w(e^{2i\pi}x) = w(x) + b$. Here the differential form dw is meromorphic at 0 and we may write

$$dw = f\frac{dx}{x^{n+1}}$$

with f holomorphic at 0. There is here an important point to consider: a logarithm will appear on integrating if and only if the coefficient of order n of f is non-zero; this yields non-trivial monodromy. A simple calculation shows that this condition is determined by the first n coefficients of $\{w, x\}$. One then has $w = \frac{u}{x^n} + \beta\log(x)$, with u holomorphic and non-zero, which can be rewritten as $\tilde{w} = \frac{w}{\beta} = \frac{1}{y^n} + \log y$, and then by setting $y = xv(x)$, one arrives at the non-zero holomorphic function $v(x)$ via the Implicit Function Theorem. □

IX.2. Gauss's hypergeometric equation and Schwarz's list

The hypergeometric equation of Gauss, namely

$$x(x-1)\frac{d^2v}{dx^2} + ((\alpha+\beta+1)x-\gamma)\,\frac{dv}{dx} + \alpha\beta v = 0, \tag{IX.1}$$

is a family of Fuchsian equations on $\mathbb{CP}^1$ in the three parameters α, β, γ (real or complex) with poles at 0, 1 and ∞. By "Fuchsian at infinity" one understands that with respect to the variable $\tilde{x} = \frac{1}{x}$ it extends meromorphically to $\tilde{x} = 0$ with the singularity Fuchsian. The indices $\pm\theta_i$ at the points $i = 0, 1, \infty$ are given by

$$\theta_0 = \gamma - 1, \;\; \theta_1 = \alpha + \beta - \gamma \quad \text{and} \quad \theta_\infty = \alpha - \beta. \tag{IX.2}$$

Every Fuchsian equation on $\mathbb{CP}^1$ with poles at 0, 1 and ∞ is projectively equivalent to an equation of this family. To see this, one verifies that the associated reduced equation must have the form

$$\frac{d^2v}{dx^2} + \left(\frac{\lambda_0}{x^2} + \frac{\lambda_1}{(x-1)^2} + \frac{\lambda_\infty - \lambda_0 - \lambda_1}{x(x-1)}\right) v = 0 \tag{IX.3}$$

where $\lambda_i = \frac{1-\theta_i^2}{4}$ is the residue[1] at the point $i = 0, 1, \infty$. In other words, a Fuchsian projective structure on $\mathbb{CP}^1$ with three singularities is completely determined by its singular points and their residues (or indices).

In [Schw1873], Schwarz revisits his earlier work [Schw1869] on conformal representation (see §IV.1.2) in order to answer the following question of Gauss: *for which values of the parameters* (α,β,γ) *is the hypergeometric series*[2] $F(\alpha,\beta,\gamma,x)$ *an algebraic function of its argument?*

In [Schw1873] Schwarz gives a complete answer to this question, determining the triples (α,β,γ) for which equation (IX.1) admits at least one algebraic integral. His exhaustive solution involves many technical details, but the most interesting part of Schwarz's work, at least from a geometrical point of view, is

[1]It should be noted that here we are considering the residues of the coefficient of the reduced equation, which differs by a factor 2 from the second term of the associated Schwarzian equation considered earlier in §IX.1.

[2]For $\gamma \notin \mathbb{Z}^-$, the hypergeometric equation (IX.1) has as a particular solution the hypergeometric series introduced by Gauss:

$$F(\alpha,\beta,\gamma,x) = \sum_{n=0}^{+\infty} \frac{(\alpha)_n(\beta)_n}{n!(\gamma)_n} x^n,$$

with the convention $(x)_0 = 1$ and $(x)_n := x(x+1)\dots(x+n-1)$. When either α or β is zero or a negative integer, $F(\alpha,\beta,\gamma,x)$ is a polynomial in x. Otherwise the series defining $F(\alpha,\beta,\gamma,x)$ has radius of convergence 1.

his answer to the simpler question as to when *all* the solutions of the hypergeometric equation are algebraic functions of their argument. By taking a fair degree of liberty with Schwarz's article, which is, to put it mildly, rather elliptical, we will explain just how he arrived at his famous list. In the course of doing this, he described the projective monodromy of the hypergeometric equation for all real triples (α, β, γ) of parameters.

IX.2.1. The algebraicity of solutions and the monodromy of the equation

We first note, as does Schwarz, that the algebraicity of solutions of the hypergeometric equation is directly linked to the finitude of the monodromy group.

Proposition IX.2.1. — *Consider a Fuchsian equation*

$$\frac{d^2v}{dx^2} + f\frac{dv}{dx} + gv = 0 \tag{E}$$

such that $g(x)$ has only simple poles. Then the following statements are equivalent:

1. *all solutions of (E) are algebraic;*
2. *the quotient $w = v_1/v_2$ of some two independent solutions of (E) is algebraic;*
3. *the projective monodromy of (E) takes its values in a finite subgroup of $PSL(2,\mathbb{C})$.*

In this case all singular points of the equation have rational indices θ, and are non-logarithmic for integer indices.

This proposition applies directly to the hypergeometric equation (IX.1). Note that the indices are rational if and only if the parameters α, β and γ are rational. We shall see in the next section that when one of the indices is an integer the corresponding singularity of the equation (IX.1) is of logarithmic type and the monodromy is then infinite.

Proof. — Let v_1 and v_2 be two linearly independent solutions of equation (IX.1). If v_1 and v_2 are algebraic then of course $w = \frac{v_1}{v_2}$ is also. It is then immediate that w has only a finite number of determinations under analytic continuation, so that, since the projective monodromy of the equation is just the monodromy of w, that also is finite.

For the converse, suppose the equation (E) has finite projective monodromy. Then, in particular, the local projective monodromy around each singularity of the equation is finite, which by §IX.1 is equivalent to the rationality of the index θ of each singularity, with those singularities with integral θ not being of logarithmic type. Furthermore, the quotient $w = v_1/v_2$ of two solutions will have a finite number of determinations and admit an algebroid continuation (that is, of the form $w = \varphi^\theta$, $\theta \in \mathbb{Q}$) at each singular point (see §IX.1). Hence in view of Riemann's work $w(x)$ is an algebraic function of x.

Finally, if w is algebraic, then this will also be the case for

$$\frac{dw}{dx} = \frac{v_2\frac{dv_1}{dx} - v_1\frac{dv_2}{dx}}{v_2^2}.$$

Now the Wronskian of two solutions v_1 and v_2 is given by:

$$v_2\frac{dv_1}{dx} - v_1\frac{dv_2}{dx} = Ce^{-\int f\,dx},$$

whence it follows that

$$v_2^2 = C\left(\frac{dw}{dx}\right)^{-1} e^{-\int f\,dx}.$$

Since the equation (E) is Fuchsian (even at infinity), f must have the form

$$f(x) = \sum_i \frac{\lambda_i}{x - x_i}\ .$$

Moreover since g has only simple poles, the residue of f at every singularity x_i is the rational number $\lambda_i = 1 + \theta_i$ (see §IX.1). Hence $e^{-\int f\,dx}$ is algebraic, whence in turn also v_2^2, and thence also v_2 and $v_1 = wv_2$. □

IX.2.2. Revisiting the article [Schw1873]

In terms of the preparatory work of Schwarz on conformal representation (see §IV.1), we may paraphrase Theorem IV.1.5 as follows:

Theorem IX.2.2. — *If* $0 \le \theta_0, \theta_1, \theta_\infty \le 2$ *are such that there exists a triangle with sides arcs of circles and angles* $\pi\theta_i$*, then the quotient* $w = v_1/v_2$ *of two particular solutions of the associated hypergeometric equation (IX.1) maps the half-plane* $\mathbb{H}$ *conformally onto the triangle (sending* 0*,* 1 *and* ∞ *to its vertices).*

In this situation, the projective monodromy of the hypergeometric equation (IX.1) coincides with the monodromy of the projective coordinate w. Recall (from §IV.1) that the group thus generated is the subgroup of index 2 of the group of (anti-)conformal transformations generated by the reflections in the sides of the triangle.

Recall that Theorem IX.2.2 is established by means of a direct approach when $0 \leq \theta_i < 1$ and indirectly as a consequence of the Riemann Mapping Theorem for $0 \leq \theta_i \leq 2$. In order to reduce the general case of real parameters θ_i to that of the statement of the theorem, we employ a group of symmetries become classical in view of later work of Schlesinger.

IX.2.3. Symmetries

We note once and for all that several hypergeometric equations may yield one and the same reduced equation, and therefore define the same projective structure. To be precise, the coefficients α, β and γ of the hypergeometric equation are, via the formulae (IX.2), mutually uniquely determined by the indices θ_i, but the latter are defined only up to sign by the projective structure. For example, the equation with parameters

$$(\alpha', \beta', \gamma') = (1 + \alpha - \gamma, 1 + \beta - \gamma, 2 - \gamma)$$

is projectively equivalent to the one given by (α, β, γ); it is arrived at via (IX.2) by means of the change of indices given by

$$(\theta_0', \theta_1', \theta_\infty') = (-\theta_0, \theta_1, \theta_\infty).$$

Thus it will be advantageous in the sequel to work in terms of the indices θ_i as parameters rather than the classical (α, β, γ). In this connection note also the natural action of the group $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ on the space of triples $(\theta_0, \theta_1, \theta_\infty) \in \mathbb{C}^3$ of parameters — which is to say on the space of hypergeometric equations — with quotient the space of projective structures.

The permutation group S_3 also acts, via the coordinate changes given by

$$x' = 1 - x \quad \text{and} \quad x' = \frac{1}{x},$$

on $\mathbb{CP}^1$, inducing the parameter changes

$$(\theta_0', \theta_1', \theta_\infty') = (\theta_1, \theta_0, \theta_\infty) \quad \text{and} \quad (\theta_\infty, \theta_1, \theta_0)$$

respectively. Combining these actions, we obtain a linear action of a group of order 48 on $\mathbb{C}^3$.

Lastly, recall that on setting $dy = \frac{dv}{v}$ one obtains a Riccati equation (see the beginning of Chapter VIII), with monodromy the same as the projective monodromy of equation (IX.1). Then on applying the birational transformation $y' = -\frac{\alpha\beta}{x(x-1)y}$, one obtains a new Riccati equation with the same monodromy since the change of unknown is regular away from the three poles of the equation. Next, direct calculation shows that on setting $dy' = \frac{dv'}{v'}$ we retrieve the hypergeometric equation, though now in terms of the parameters

$$(\alpha', \beta', \gamma') = (-\alpha, -\beta, 1-\gamma),$$

whose projective monodromy must be the same as for the parameters (α, β, γ). This corresponds to the transformation

$$(\theta_0', \theta_1', \theta_\infty') = (-\theta_0 - 1, -\theta_1 - 1, -\theta_\infty).$$

(It is important to note, however, that the projective structure has changed.) By combining this transformation with the earlier ones, one easily sees that the group generated is an infinite group Γ of affine transformations of the space of parameters, with a normal subgroup of index 6 given by

$$\Gamma' = \{(\pm\theta_0 + n_0, \pm\theta_1 + n_1, \pm\theta_\infty + n_\infty) \mid (n_0, n_1, n_\infty) \in \mathbb{Z}^3,\ n_0 + n_1 + n_\infty \in 2\mathbb{Z}\}.$$

The quotient Γ/Γ' is the symmetric group S_3 of degree 3.

Proposition IX.2.3. — *Two hypergeometric equations have the same projective monodromy representation if and only if one can be obtained from the other by the action of the symmetry group* Γ'.

Proof. — It suffices to verify that two hypergeometric equations having the same projective monodromy are sent one to the other by an element of Γ'. The projective monodromy of equation (IX.1) is given, in terms of the standard system of generators of the automorphism group of $\mathbb{CP}^1 \setminus \{0, 1, \infty\}$, by a triple $(\varphi_0, \varphi_1, \varphi_\infty) \in \mathrm{PSL}(2, \mathbb{C})$ satisfying $\varphi_0\varphi_1\varphi_\infty = \mathrm{id}$. The transformation φ_i represents the local monodromy of the projective structure around the respective pole $i = 0, 1, \infty$ and is conjugate in $\mathrm{PSL}(2, \mathbb{C})$ to an affine transformation of the form $w \mapsto e^{2i\pi\theta_i} w + b$, $b \in \mathbb{C}$. Each transformation φ_i has two preimages $\pm M_i \in \mathrm{SL}(2, \mathbb{C})$; we choose M_i so that $\mathrm{tr} M_i = 2\cos(\pi\theta_i)$. We claim that then the relation $\varphi_0\varphi_1\varphi_\infty = \mathrm{Id}$ lifts to $M_0 M_1 M_\infty = -I$.

To see this, note first that we must have $M_0 M_1 M_\infty = \pm I$ and the sign of the right-hand side depends continuously on the parameters θ_i of the equation; it must therefore be constant on the space $\mathbb{C}^3$ of parameters, so that it suffices to determine it in a particular case. By applying Theorem IX.2.2 to $w = z$, for example, the conformal representation of $\mathbb{H}$ viewed as a triangle with all its angles equal to π, one finds that $\theta_i = 1$ and $M_i = -I$ for $i = 0, 1, \infty$.

If another hypergeometric equation, with parameters θ_i', has the same projective monodromy, then up to conjugation we shall have $M_i' = \pm M_i$ with $M_0' M_1' M_\infty' = -I$. Hence in particular $\mathrm{tr}(M_i') = \pm\mathrm{tr}(M_i)$, which is equivalent to

$\cos(\pi\theta_i') = \pm\cos(\pi\theta_i)$, in turn equivalent to $\theta_i' = \pm\theta_i + n_i$, $n_i \in \mathbb{Z}$; the condition $M_0'M_1'M_\infty' = -I$ now follows since $n_0 + n_1 + n_\infty$ is even. □

By Proposition IX.1.2, when θ_0 is an integer, say $\theta_0 = n \in \mathbb{N}$, in order to determine if the singularity is logarithmic, we need to consider the first n terms of the Laurent series of the coefficient in the reduced equation (IX.3). A remarkable consequence of the use of the symmetry group is that there are no such non-logarithmic singularities:

Proposition IX.2.4. — *If for any of the singular points $i = 0, 1, \infty$ of the hypergeometric equation the index $\theta_i \in \mathbb{Z}$ is an integer, then the singularity is of logarithmic type.*

Proof. — By Proposition IX.1.2, the nature of the singularity can be read off the monodromy; it is therefore invariant under the action of the group Γ, and we may, in particular, assume $i = 0$ and $\theta_0 = 0$. However in this case the reduced equation (IX.3) has a double pole at 0 ($\lambda_0 = \frac{1}{4}$), so that the singularity is logarithmic, again by Proposition IX.1.2. □

IX.2.4. Triangles and geometries

In order to understand the structure of the group generated by the reflections in the sides of a triangle, Schwarz was naturally led to consider — though without putting it in these terms — three geometries. Recall that the Riemann sphere $\mathbb{CP}^1$ may be identified with the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ via the stereographic projection. One then defines a *circle* on $\mathbb{CP}^1$ as the image of a plane intersecting $\mathbb{S}^2$, as long as this intersection is neither empty nor consists of just a single point; these correspond to the circles and straight lines of $\mathbb{C}$ in the Euclidean metric. The group $\mathrm{PSL}(2,\mathbb{C})$ acts transitively on the circles of $\mathbb{CP}^1$.

Spherical geometry. The Euclidean metric $\mathbb{R}^3$ induces a metric of constant curvature $+1$ on the sphere $\mathbb{S}^2$, with isometry group generated by the antipodal involution $\sigma(z) = -\frac{1}{\bar{z}}$ together with the rotation group

$$\mathrm{PSU}(2,\mathbb{C}) = \{\varphi \in \mathrm{PSL}(2,\mathbb{C}) \mid \varphi \circ \sigma = \sigma \circ \varphi\}.$$

The geodesics are the great circles (the intersections of $\mathbb{S}^2$ with planes $\Pi \subset \mathbb{R}^3$ passing through the origin). These are just the circles intersecting the equator $\mathbb{RP}^1$, for instance, in a pair of antipodal points.

Euclidean geometry. The Euclidean metric of $\mathbb{C} \subset \mathbb{CP}^1$ has as its geodesics the straight lines on $\mathbb{C}$, that is, the circles of $\mathbb{CP}^1$ passing through the point at infinity.

Hyperbolic geometry. The Poincaré metric of constant curvature -1 on $\mathbb{H}$ has as geodesics arcs of circles of $\mathbb{CP}^1$ orthogonal to the equator.

In what follows, by "triangle" we shall understand a simply connected region of $\mathbb{CP}^1$ with boundary made up of three circular arcs forming a Jordan curve with distinct vertices. We shall denote the vertices by w_0, w_1 and w_∞, by A_i the circular arc opposite the vertex w_i, and by C_i the complete circle of which A_i is an arc. Schwarz observes the following trichotomy:

Proposition IX.2.5. — *Let C_0, C_1 and C_∞ be three circles on the Riemann sphere $\mathbb{CP}^1$ intersecting pairwise in one or two points. We have the following three possibilities:*

— *C_∞ intersects $C_0 \cap C_1$, or*

— *C_∞ separates $C_0 \cap C_1$ (that is, $C_0 \cap C_1$ intersects both components of $\mathbb{CP}^1 \setminus C_\infty$ but not C_∞), or*

— *C_∞ isolates $C_0 \cap C_1$ (that is, $C_0 \cap C_1$ is wholly contained in one of the connected components of $\mathbb{CP}^1 \setminus C_\infty$).*

— *In each case there exists an element of $PSL(2,\mathbb{C})$ mapping the three circles C_i simultaneously onto Euclidean, spherical, or hyperbolic geodesics respectively.*

Proof. — In the first case it suffices to send any point of $C_0 \cap C_1 \cap C_\infty$ to the point ∞. (Here we do not exclude the possibility that the circles become combined.) In the second case, we first map the two points of $C_0 \cap C_1$ to 0 and ∞, and then, by means of appropriate homotheties $\varphi(z) = az$, send C_∞ onto a great circle. In the third case, assuming that two of the circles, say C_0 and C_1, intersect in two distinct points, then once again we map these points to 0 and ∞ and then juggle homotheties $\varphi(z) = az$ to map the third circle C_∞ to one also orthogonal to the equator; thus all three will have been sent to geodesics relative to the hyperbolic metric on the unit disc. Finally, if the three circles are pairwise tangential to each other, then the circle C through the three points of tangency will be orthogonal to them, and on mapping C to $\mathbb{RP}^1$, the circles C_i will become geodesics in the hyperbolic metric on $\mathbb{H}$. □

In the sequel we shall call:

— a *spherical triangle* any triangle on the Riemann sphere $\mathbb{CP}^1$ bounded by geodesics relative to the spherical metric;

— a *Euclidean triangle* any triangle[3] on $\mathbb{C}$ bounded by straight lines, geodesics for the Euclidean metric;

[3]We include here the possibility that one of the vertices is at infinity, with the restriction that then the two adjacent sides should be parallel. In other words, unbounded Euclidean triangles are allowed, with angles $(0, \pi\theta, \pi(1-\theta))$, $0 \leq \theta \leq 1$, with the case $(0, 0, \pi)$ degenerating to a strip.

— a *hyperbolic triangle* any triangle in the half-plane $\mathbb{H}$ bounded by geodesics relative to the hyperbolic metric.

Even though Proposition IX.2.5 tells us that every triangle is equivalent modulo the action of the group $\mathrm{PSL}(2,\mathbb{C})$ to a triangle with geodesic sides relative to one of the three geometries, one should be careful to note that that triangle itself may not be any of the three preceding types. To take the hyperbolic case, for instance, it may happen that several triangles are bounded by arcs of the same circles C_0, C_1 and C_∞, only one of which is contained in $\mathbb{H}$. The same sort of possibility occurs in the Euclidean case. However, the transformation group generated by the symmetries in the three circles depends only on those circles and not on the triangle chosen. Schwarz actually shows that the triangle minimizing the angle sum $\theta_0 + \theta_1 + \theta_\infty$ is indeed hyperbolic, Euclidean or spherical. Note that he does not even ask for which $0 \leq \theta_i \leq 2$ there exists a triangle with angles $\pi\theta_i$. We shall circumvent Schwarz's arguments and the associated technical difficulties by appealing to the group of symmetries, as in the following proposition.

Proposition IX.2.6. — *Let $0 \leq \theta_0 \leq \theta_1 \leq \theta_\infty \leq 2$. There exists a triangle (with sides arcs of circles) with angles $\pi\theta_i$ if and only if:*

$$2\theta_\infty - 1 < \theta_0 + \theta_1 + \theta_\infty < 2\theta_0 + 3, \tag{IX.4}$$

and in this case the triangle is unique modulo the action of $PSL(2,\mathbb{C})$. Furthermore, when $\theta_0 + \theta_1 + \theta_\infty < 2\theta_0 + 1$, these conditions are satisfied and the triangle is equivalent modulo the action of $PSL(2,\mathbb{C})$ to one of the following:

— *a hyperbolic triangle if $\theta_0 + \theta_1 + \theta_\infty < 1$;*

— *a Euclidean triangle if $\theta_0 + \theta_1 + \theta_\infty = 1$; and*

— *a spherical triangle if $\theta_0 + \theta_1 + \theta_\infty > 1$.*

Proof. — Consider a triangle with angles $\pi\theta_i$ where

$$0 \leq \theta_0, \theta_1, \theta_\infty \leq 2. \tag{IX.5}$$

Modulo the action of $\mathrm{PSL}(2,\mathbb{C})$ we may suppose the vertices are 0, 1 and ∞. Denote by A_{ij} the circular arc forming the side of the triangle joining the vertices i and j. We choose an orientation so that the oriented side A_{01} goes from 0 to 1 with the interior of the triangle to the left. (Note that A_{01} is a circular arc — which in the degenerate case coincides with the interval $[0,1]$ — while the two other sides $A_{1\infty}$ and $A_{\infty 0}$ are straight lines, that is, circles passing through the point at infinity).

We now introduce the parameters

$$-1 < \delta_0, \delta_1, \delta_\infty < 1 \tag{IX.6}$$

defined as follows: for $(i, j, k) = (0,1,\infty), (1,\infty,0), (\infty,0,1)$, we take $\pi\delta_i$ to be the angle at the point j made by the arc A_{jk} with the real interval I_{jk}; the sign is chosen here so that $\delta_i > 0$ if the arc A_{jk} lies in the interior of $\mathbb{H}$. We have excluded the possibility $\delta_i = 1$ since in this case the arc A_{jk} would be incident with the vertex i, and the boundary of the triangle would no longer be a Jordan curve. Hence the angles of the triangle are given by

$$\begin{cases} \theta_0 &= 1 - \delta_1 - \delta_\infty \\ \theta_1 &= 1 - \delta_\infty - \delta_0 \\ \theta_\infty &= 1 - \delta_0 - \delta_1. \end{cases} \tag{IX.7}$$

Since $0 \leq \pi\theta_i \leq 2\pi$, the parameters δ_i are subject to the constraints

$$-1 \leq \delta_0 + \delta_1, \delta_1 + \delta_\infty,\ \delta_0 + \delta_\infty \leq 1. \tag{IX.8}$$

Conversely, every triple $(\delta_0, \delta_1, \delta_\infty)$ of real numbers satisfying the conditions (IX.6) and (IX.8) corresponds to a triangle with angles given by (IX.7).
Inverting the system (IX.7), we obtain

$$\begin{cases} \delta_0 &= 1 + \theta_0 - \theta_1 - \theta_\infty/2 \\ \delta_1 &= 1 + \theta_1 - \theta_0 - \theta_\infty/2 \\ \delta_\infty &= 1 + \theta_\infty - \theta_0 - \theta_1/2 \end{cases} \tag{IX.9}$$

so that the triangle is determined, modulo the action of $\mathrm{PSL}(2,\mathbb{C})$, by its angles. It remains to express the constraints (IX.6) in terms of the angles of the triangle:

$$2\theta_0 - 1, 2\theta_1 - 1, 2\theta_\infty - 1 \leq \theta_0 + \theta_1 + \theta_\infty \leq 2\theta_0 + 3, 2\theta_1 + 3, 2\theta_\infty + 3. \tag{IX.10}$$

When $\theta_0 \leq \theta_1 \leq \theta_\infty$, these constraints reduce to those of (IX.4).

We now turn to the second part of the theorem. The set of triples $(\theta_0, \theta_1, \theta_\infty)$ of parameters defined by the inequalities (IX.5) and (IX.10) is a convex region T of $\mathbb{R}^3$, regarded as the space of all triangles. By Proposition IX.2.5 we can partition T in accordance with the configuration of the three circles bounding each triangle. The set E of Euclidean configurations, characterized by the property that these three circles have a common point of intersection, is closed and separates the (open) components made up of the hyperbolic and spherical types. We now give the equations for E. If the common point of all three circles lies on the boundary of the triangle, then it must be a vertex, and this occurs precisely when one of the δ_i vanishes, that is, when

$$\theta_0 + \theta_1 + \theta_\infty = 2\theta_0 + 1, \quad 2\theta_1 + 1 \quad \text{or} \quad 2\theta_\infty + 1.$$

If this is not the case then of course the common point of the three circles is either in the interior or exterior of the triangle. In the second case we have a Euclidean triangle with $\theta_0 + \theta_1 + \theta_\infty = 1$, and in the first case the complement of the triangle (with angles $2\pi - \pi\theta_i$) is Euclidean whence $\theta_0 + \theta_1 + \theta_\infty = 5$. It is classical and straightforward (so we omit the details) that the condition $\theta_0 + \theta_1 + \theta_\infty = 1$ characterizes Euclidean triangles. On the other hand the open component $\theta_0 + \theta_1 + \theta_\infty < 1$ corresponds to hyperbolic triangles. Indeed, if a triangle is hyperbolic, then its hyperbolic area is given by $\pi(1 - \theta_0 - \theta_1 - \theta_\infty) > 0$; by continuity, therefore, every other triangle satisfying this inequality is hyperbolic. If one now confines oneself to the parametric region T^+ determined by the inequalities $\theta_0 \leq \theta_1 \leq \theta_\infty$, then the other component bordering on the plane $\theta_0 + \theta_1 + \theta_\infty = 1$ is defined by

$$1 < \theta_0 + \theta_1 + \theta_\infty < 2\theta_0 + 1,$$

in which case the triangles are spherical. □

IX.2.5. Monodromy

The following crucial lemma allows us to bring every hypergeometric equation into the form of an equation uniformizing a hyperbolic, Euclidean, or spherical triangle.

Lemma IX.2.7. — *Let* $(\theta_0, \theta_1, \theta_\infty) \in \mathbb{R}^3$ *be a triple of reals. Its orbit under the symmetry group* Γ *contains a unique positive ordered triple* $(\theta'_0, \theta'_1, \theta'_\infty)$*:*

$$0 \leq \theta'_0 \leq \theta'_1 \leq \theta'_\infty$$

minimizing the sum

$$\theta'_0 + \theta'_1 + \theta'_\infty,$$

which we call a ***reduced triple****. It also satisfies*

$$\theta'_0 + \theta'_1 + \theta'_\infty \leq 1 + 2\theta'_0$$

with equality only if $\theta'_0 = 0$*.*

Proof. — By applying first of all appropriate changes of signs and "even" translations $(n_0, n_1, n_\infty) \in (2\mathbb{Z})^3$, one obtains a unique triple $(\theta_0, \theta_1, \theta_\infty)$ for which $0 \leq \theta_i < 1$. Then by further applying transformations of the forms

$$(1-\theta_0, 1-\theta_1, \theta_\infty), \quad (1-\theta_0, \theta_1, 1-\theta_\infty) \quad \text{and} \quad (\theta_0, 1-\theta_1, 1-\theta_\infty),$$

one can minimize the sum $\theta_0 + \theta_1 + \theta_\infty$. Note that if one of these triples, the first, say, has the same sum as $(\theta_0, \theta_1, \theta_\infty)$, then $\theta_0 + \theta_1 = 1$, so that these two triples are permutations of one another. Finally, by applying the symmetric group S_3 appropriately one can order the θ_i so that $0 \leq \theta_0 \leq \theta_1 \leq \theta_\infty$; the triple thus obtained is then unique. Now if we had

$$\theta_0 + \theta_1 + \theta_\infty \geq 1 + 2\theta_0,$$

then the triple $(\theta_0', \theta_1', \theta_\infty') = (\theta_0, 1-\theta_1, 1-\theta_\infty)$ would satisfy

$$\theta_0' + \theta_1' + \theta_\infty' \leq 1$$

whence, by the minimality of the sum, the above two inequalities become equalities and $\theta_0 = 0$. □

Corollary IX.2.8. — *The projective monodromy of a hypergeometric equation with real coefficients is given by the group generated by the reflections in the sides of a hyperbolic, parabolic, or elliptic triangle.*

More precisely, if G denotes the transformation group generated by the three reflections, then the image of the projective monodromy representation is the subgroup of index 2 consisting of the orientation-preserving transformations in G.

Proof. — Consider the hypergeometric equation (IX.1) with the coefficients α, β and γ assumed real, so that the indices θ_i defined by (IX.2) are also real. By Proposition IX.2.3, every other hypergeometric equation obtained via the action of the group Γ' has the same monodromy. By the above proposition, we can, for example, arrange for the indices to satisfy $0 \leq \theta_0 \leq \theta_1 \leq \theta_\infty \leq 1 + \theta_0 - \theta_1$. In particular, Theorem IX.2.2 tells us that the monodromy is generated by the reflections in the sides of a triangle with angles $\pi\theta_i$ and Proposition IX.2.6 that that triangle is hyperbolic, Euclidean, or spherical. □

IX.2.6. The hyperbolic case

When the indices of the hypergeometric equation (IX.1) satisfy

$$\theta_0 + \theta_1 + \theta_\infty < 1,$$

the monodromy is generated, modulo conjugation in $\mathrm{PSL}(2,\mathbb{C})$, by the reflections in the sides of a hyperbolic triangle. In this case the monodromy group is infinite. This occurs, for example, when there are vertices of the triangle located on the boundary of $\mathbb{H}$, since on composing the reflections in the two sides incident with such a vertex, one obtains a parabolic element (of infinite order) of $\mathrm{PSL}(2,\mathbb{C})$ (see §VI.1.6). On the other hand, when the triangle is compact in $\mathbb{H}$, the images of the triangle obtained by reflecting in the sides can approach arbitrarily closely to the boundary of $\mathbb{H}$, without ever reaching it. After a finite number of successive reflections the vertices of the image triangles will always be at non-zero distance from the boundary of $\mathbb{H}$ and the procedure of repeated application of symmetries can be carried on indefinitely. *The projective coordinate w is a transcendental function, since it is an infinitely multi-valued function of x.*

The diagram reproduced below as Figure IX.1, appears in [Schw1873] as an illustration of the case $(\theta_0, \theta_1, \theta_\infty) = \left(\frac{1}{5}, \frac{1}{4}, \frac{1}{2}\right)$, yielding a tiling of the disc by triangles.

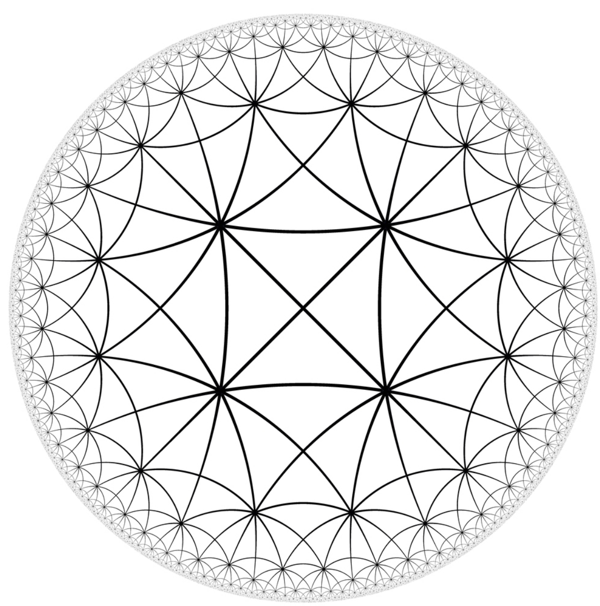


Figure IX.1: A tiling by the triangles $(\frac{\pi}{5}, \frac{\pi}{4}, \frac{\pi}{2})$

Returning to the example of §VI.2.1, we see that the hypergeometric equation (IX.1) with parameters (as defined in (IX.2)) of the form

$$\theta_i = \frac{1}{k_i}, \quad k_i \in \mathbb{N}^*, \quad \frac{1}{k_0} + \frac{1}{k_1} + \frac{1}{k_\infty} < 1,$$

must be the uniformizing equation for the sphere with three conical points of angles $\frac{2\pi}{k_i}$. In this case the variable x is a single-valued function of w. Schwarz notes that this is the only case where this circumstance arises, but without providing any proof. Recall (see §VI.2.1) that this represents a particular case of Poincaré's Theorem VI.1.10. This later prompted the following remark of Poincaré in a letter to Mittag-Leffler:

> In his memoir M. Schwarz has thus stated a result of the greatest importance, namely the one I quoted. He gives no proof. In the proof of this result there is a very delicate point, a difficulty of a special kind; I don't know how M. Schwarz overcame it.

Unfortunately, however, Schwarz dwells no further on the hyperbolic case. He does revisit it later on, after the relevant works of Klein and those of Poincaré on Fuchsian functions appeared. In the second volume of Schwarz's complete works, there is an addendum to [Schw1873] in which he reformulates the different cases investigated earlier in terms of hyperbolic, Euclidean, and spherical geometry, and then evokes, by means of several examples, the major fact that had eluded him in [Schw1873], namely the property of the "transcendental" functions $w(x)$ associated with the parameters $\theta_i = \frac{1}{k_i}$ of uniformizing a great number of algebraic curves (see §IX.3). We mention by the way that the hypergeometric equation

$$x(x-1)\frac{d^2v}{dx^2} + (2x-1)\frac{dv}{dx} + \frac{v}{4} = 0 \tag{IX.11}$$

uniformizes $\mathbb{CP}^1 \setminus \{0, 1, \infty\}$.

IX.2.7. The Euclidean case

When the indices of the hypergeometric equation (IX.1) satisfy

$$\theta_0 + \theta_1 + \theta_\infty = 1,$$

the monodromy is generated, modulo conjugation in $\mathrm{PSL}(2, \mathbb{C})$, by the reflections in the sides of a Euclidean triangle, including the case $\theta_\infty = 0$ where the corresponding vertex is at the point at infinity and the two adjacent sides are parallel half-lines. Like Schwarz, we shall not linger over this case. The monodromy is again infinite for reasons similar to those in the hyperbolic case. Note that once again the function $x(w)$ is single-valued if and only if the triangle tiles the plane, which is the case precisely for the triples of indices

$$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),\quad \left(\frac{1}{2},\frac{1}{4},\frac{1}{4}\right),\quad \left(\frac{1}{2},\frac{1}{3},\frac{1}{6}\right)\quad \text{and}\quad \left(\frac{1}{2},\frac{1}{2},0\right).$$

IX.2.8. The spherical case

When the indices of the hypergeometric equation (IX.1) satisfy

$$0 < \theta_0 \leq \theta_1 \leq \theta_\infty \leq 1$$

with

$$1 < \theta_0 + \theta_1 + \theta_\infty < 2\theta_0 + 1,$$

the monodromy is generated, modulo conjugation in $\mathrm{PSL}(2,\mathbb{C})$, by the reflections in the sides of a spherical triangle with angles $\pi\theta_i$. In this case the projective monodromy group of the equation is a subgroup of $\mathrm{SO}(3)$, the rotation group of the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Its finite subgroups are well known to be as follows:

The finite cyclic groups. — For each $n \in \mathbb{N}^*$ the rotation $w \mapsto e^{\frac{2i\pi}{n}} w$ generates a cyclic subgroup of order n with quotient $x(w) = w^n$, viewed as a metric space, a sphere with two conical points both of angle $\frac{2\pi}{n}$.

The dihedral groups. — The group D_n ($n \in \mathbb{N}, n \geq 2$) generated by the involution $w \mapsto \frac{1}{w}$ and the rotation $w \mapsto e^{\frac{2i\pi}{n}} w$, has order $2n$, and is isomorphic as abstract group to the semi-direct product $\mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$. It is a subgroup of index 2 of the group generated by the reflections in the sides of the spherical triangle with angles $\left(\frac{\pi}{2},\frac{\pi}{2},\frac{\pi}{n}\right)$. The quotient, given by $x(w) = \frac{(1-w^n)^2}{4w^n}$, is a sphere with 3 conical points of angles $\left(\pi,\pi,\frac{2\pi}{n}\right)$. The inverse function $w(x)$ is the quotient of two solutions of the hypergeometric equation with indices $\left(\frac{1}{2},\frac{1}{2},\frac{1}{n}\right)$.

The tetrahedral group. — When one tiles the sphere with 4 triangles with all their angles equal to $\frac{2\pi}{3}$, one obtains a spherical tetrahedron. The group of rotations preserving this tiling has order 12, and is isomorphic as abstract group to the alternating group A_4. If one adds the reflections, one obtains a group of order 24 with fundamental region the triangle with angles $\left(\frac{\pi}{2},\frac{\pi}{3},\frac{\pi}{3}\right)$, which defines a subtiling of the above tiling. The group A_4 may thus be viewed as a subgroup of index 2 of the group generated by the reflections in the sides of the latter triangle. Passage to the quotient by A_4 is given, for example[4], by

$$x(w) = -12i\sqrt{3}\frac{w^2(w^4-1)^2}{(w^4-2i\sqrt{3}w^2+1)^3}.$$

[4]The formulae given by Klein for the quotient by this group and the following two correspond to $\tilde{x} = 1 - x(w)$.

The inverse function $w(x)$ is the quotient of two solutions of the hypergeometric equation with indices

$$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{3}\right).$$

The octahedral group. — This is the group of rotations of order 24, isomorphic to the symmetric group S_4, leaving invariant the octahedral tiling of the sphere by 8 triangles with all their angles equal to $\frac{\pi}{2}$. One may also view it as the group fixing the cubic tiling by 6 quadrilaterals of angles $\frac{2\pi}{3}$. It has index 2 in the group generated by the reflections in the sides of the triangle with angles $\left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4}\right)$. The tiling determined by the latter triangle contains as subtilings both the octahedral and cubic ones, situated dually with respect to each other[5]. The passage to the quotient is given by

$$x(w) = -\frac{1}{108} \frac{(w^{12} - 33w^8 - 33w^4 + 1)^2}{w^4(w^4 - 1)^4}.$$

The inverse function $w(x)$ is the quotient of two solutions of the hypergeometric equation with indices

$$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right).$$

The icosahedral group. — This is the group of rotations of order 60, isomorphic to the alternating group A_5, and leaving invariant the icosahedral tiling of the sphere by 20 triangles with all their angles equal to $\frac{2\pi}{5}$. It is also the group fixing the dodecahedral tiling of the sphere by 12 regular pentagons of angle $\frac{2\pi}{3}$. It has index 2 in the group generated by the reflections in the sides of the triangle with angles $\left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5}\right)$. The passage to the quotient is given by

$$x(w) = \frac{1}{1728} \frac{(w^{30} + 522w^{25} - 10005w^{20} - 10005w^{10} - 522w^5 + 1)^2}{w^5(w^{10} + 11w^5 - 1)^5}.$$

The inverse function $w(x)$ is the quotient by two solutions of the hypergeometric equation with indices

$$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5}\right).$$

[5]That is, with a vertex of the first at the center of each face of the second and *vice versa*.

IX.2.9. Schwarz's list

The main result obtained by Schwarz in [Schw1873] is the following:

Theorem IX.2.9. — *A hypergeometric equation has all its solutions algebraic if and only if its parameters are equivalent, via the symmetry group* Γ*, to one of the triples in the following table:*

Group	Reduced triples $(\theta_0, \theta_1, \theta_\infty)$
D_n dihedral	$\left(\frac{1}{2}, \frac{1}{2}, \frac{k}{n}\right)$, with $k = 1, \ldots, n-1$
A_4 tetrahedral	$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{3}\right)$ or $\left(\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$
S_4 octahedral	$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right)$ or $\left(\frac{2}{3}, \frac{1}{4}, \frac{1}{4}\right)$
A_5 icosahedral	$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5}\right)$, $\left(\frac{2}{5}, \frac{1}{3}, \frac{1}{3}\right)$, $\left(\frac{2}{3}, \frac{1}{5}, \frac{1}{5}\right)$, $\left(\frac{1}{2}, \frac{2}{5}, \frac{1}{5}\right)$, $\left(\frac{3}{5}, \frac{1}{3}, \frac{1}{5}\right)$, $\left(\frac{2}{5}, \frac{2}{5}, \frac{2}{5}\right)$, $\left(\frac{2}{3}, \frac{1}{3}, \frac{1}{5}\right)$, $\left(\frac{4}{5}, \frac{1}{5}, \frac{1}{5}\right)$, $\left(\frac{1}{2}, \frac{2}{5}, \frac{1}{3}\right)$ or $\left(\frac{3}{5}, \frac{2}{5}, \frac{1}{3}\right)$.

Idea of the proof. — The method Schwarz proposes (without giving the details) is as follows. Suppose that the monodromy of a hypergeometric equation is finite and that its defining triple $(\theta_0, \theta_1, \theta_\infty)$ of indices is reduced (see Lemma IX.2.7). The monodromy group G of the equation is then of index 2 in the group $G^\pm$ generated by the reflections in the sides of the spherical triangle with angles $\pi\theta_i$ (since, as we have seen, the triangle cannot be hyperbolic or Euclidean). Hence the group G may be identified with one of the finite rotation groups of the sphere $\mathbb{S}^2$ described in the preceding subsection.

We take the case of the tetrahedral group A_4 by way of illustration, and denote by $A_4^\pm$ the reflection group of the triangle T_0 with angles $\left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{3}\right)$. Thus here $G^\pm = A_4^\pm$. In fact $G^\pm$ is always generated by G together with the reflection σ in any of the sides of the triangle T with angles $\pi\theta_i$. Since in the present case $G = A_4$ is a normal subgroup, σ defines, via passage to the quotient, an anti-holomorphic transformation $\bar{\sigma}$ on $\mathbb{CP}^1/A_4$; furthermore, since σ is an isometry, $\bar{\sigma}$ will preserve the metric structure of the quotient and hence fix the 3 conical points. It follows that $\bar{\sigma}$ is the reflection in the circle passing through those 3 points — the generator of the action of $A_4^\pm$ on the quotient $\mathbb{CP}^1/A_4$.

Here we are exploiting the fact that the quotients have 3 conical points, which is the case for all the finite groups except the cyclic ones, which need to be eliminated by means of a different argument.

Thus the sides of the triangle T furnishing the generators of $G^{\pm}$ must be contained among the geodesics of the tiling determined by $A_4^{\pm}$, so that T is tiled by a finite number of copies of T_0. The angles of T are of the form $\frac{\pi}{2}$, $\frac{\pi}{3}$, and $\frac{2\pi}{3}$ (they are $< \pi$ and are formed from the angles of T_0 by successive reflections). These constraints give us 3 reduced triples, namely

$$\left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}\right), \quad \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{3}\right) \quad \text{and} \quad \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

The last is Euclidean but the first two, of areas $\frac{\pi}{3}$ and $\frac{\pi}{6}$ respectively, conform: the first is made up of two copies of T_0 and the second is T_0 itself.

In the case of A_5, one finds 15 reduced triples of the indices $\frac{1}{2}$, $\frac{1}{3}$, $\frac{2}{3}$, $\frac{1}{5}$, $\frac{2}{5}$, $\frac{3}{5}$ or $\frac{4}{5}$ corresponding to spherical triangles. The triple $\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5}\right)$ corresponds to the triangle of minimal area $\frac{\pi}{30}$: it is a fundamental region for the action of $A_5^{\pm}$. After a tedious case-by-case analysis, one discovers that of the 15 possibilities, only 10 triangles are tiled by means of the fundamental triangle for $A_5^{\pm}$. For example, in order to eliminate the triple $\left(\frac{2}{5}, \frac{2}{5}, \frac{3}{5}\right)$, which corresponds to the triangle of maximal area $\frac{12\pi}{30}$, it suffices to note that the corresponding hypergeometric equation can be obtained by lifting, via the branched covering map $x \mapsto (2x-1)^2$, the hypergeometric equation corresponding to the reduced triple $\left(\frac{1}{2}, \frac{2}{5}, \frac{3}{15}\right)$. The monodromy group of the former has at most index 2 in that of the latter, which cannot be finite since it contains an element of order 15. □

In [Kle1884], Klein reconsiders Schwarz's work, bringing to it the clarification lent by Galois theory. He also re-derives Schwarz's list by means of a different approach to finite monodromy via Fuchsian equations:

Theorem IX.2.10. — *Consider a Fuchsian equation*

$$\frac{d^2v}{dx^2} + f\frac{dv}{dx} + gv = 0 \qquad (E)$$

on $\mathbb{CP}^1$*, and suppose its projective monodromy group is finite non-Abelian. Then (E) is projectively equivalent to the lift via a rational map* $\phi(x)$ *of a hypergeometric equation with indices*

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{n}\right), \quad \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{3}\right), \quad \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right) \quad \text{or} \quad \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5}\right).$$

The proof is remarkable for its simplicity.

Proof. — Consider the quotient $w = v_1/v_2$ of two independent solutions of the equation (E): this is a many-valued local biholomorphism

$$\mathbb{CP}^1 \setminus \mathrm{Sing}(E) \to \mathbb{CP}^1$$

whose monodromy coincides with one of the finite groups D_n, A_4, S_4 or A_5 described above. (Here $\mathrm{Sing}(E)$ denotes the singular locus of the equation (E).) Composing the map w with the map $\mathbb{CP}^1 \to \mathbb{CP}^1/G$ determined by passage to the quotient by the corresponding group G, we obtain a single-valued local biholomorphism

$$\phi : \mathbb{CP}^1 \setminus \mathrm{Sing}(E) \to \mathbb{CP}^1/G,$$

which, by the local investigation of singularities undertaken in §IX.1, extends continuously to $\mathrm{Sing}(E)$; we thus obtain a rational map $\phi : \mathbb{CP}^1 \to \mathbb{CP}^1/G \simeq \mathbb{CP}^1$. The projective structure induced on $\mathbb{CP}^1 \setminus \mathrm{Sing}(E)$ by the equation (E) is the lift via $w(x)$ of the standard projective structure on $\mathbb{CP}^1$ and consequently the lift via $\phi(x)$ of the orbifold projective structure on the quotient $\mathbb{CP}^1/G$; for each of the finite groups listed above, the quotient structure has precisely 3 conical points and is defined by a hypergeometric equation with indices as given in §IX.2.8. □

This statement obviously remains valid (with the same proof) when (E) is a globalizable Fuchsian equation on any curve. Moreover it allows us to give a very different proof of Schwarz's theorem using the techniques of branched coverings.

IX.3. Examples of families of normal equations

In the following subsections we give other examples of normal equations, notably in the smooth (non-orbifold) case. On each occasion where the curve's symmetries allow the accessory parameters to be determined, we observe that the equation in fact reduces to hypergeometric form.

IX.3.1. Heun's equation and the sphere with 4 points removed

As in the case of the hypergeometric equation, one readily verifies by passing to the reduced form that every Fuchsian equation on $\mathbb{CP}^1$ with poles at 0, 1, λ, and ∞ is projectively equivalent to Heun's equation

$$\frac{d^2v}{dx^2} + \left(\frac{\alpha}{x} + \frac{\beta}{x-1} + \frac{\gamma}{x-\lambda}\right)\frac{dv}{dx} + \frac{\delta x + c}{x(x-1)(x-\lambda)}v = 0,$$

which has indices at its singular points respectively

$$\alpha - 1, \quad \beta - 1, \quad \gamma - 1 \quad \text{and} \quad \sqrt{(\alpha+\beta+\gamma-1)^2 - 4\delta}.$$

The normal equation associated with the uniformization of $\mathbb{CP}^1 \setminus \{0,1,\lambda,\infty\}$ is thus projectively equivalent to an equation of the form

$$\frac{d^2v}{dx^2} + \left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{x-\lambda}\right)\frac{dv}{dx} + \frac{x+c}{x(x-1)(x-\lambda)}v = 0.$$

Here c is what is called an "accessory parameter" of the equation. The uniformizing function $w(x)$ is given as the quotient $w = v_1/v_2$ of two independent solutions of the equation for a single value of c. In fact, since in the smooth case (see Proposition VIII.3.17) two uniformizations $x : \mathbb{H} \to \mathbb{CP}^1 \setminus \{0,1,\lambda,\infty\}$ will induce the same projective structure on $\mathbb{CP}^1 \setminus \{0,1,\lambda,\infty\}$ and therefore yield the same reduced equation

$$\frac{d^2v}{dx^2} + \left\{\frac{1}{4x^2} + \frac{1}{4(x-1)^2} + \frac{1}{4(x-\lambda)^2} + \frac{2c+\lambda+1-x}{2x(x-1)(x-\lambda)}\right\}v = 0, \quad \text{(IX.12)}$$

it follows that two distinct values of c yield two distinct reduced equations. It is not known which c yield a uniformizing equation except in special cases. For example, when $\lambda = -1$, the Möbius transformation $\varphi(x) = -x$ permutes the 4 singular points, and since the uniformizing equation must be left invariant, we infer that $c = 0$; thus the equation

$$\frac{d^2v}{dx^2} + \left\{\frac{1}{4x^2} + \frac{1}{4(x-1)^2} + \frac{1}{4(x+1)^2} + \frac{x}{2x(x-1)(x+1)}\right\}v = 0$$

uniformizes $\mathbb{CP}^1 \setminus \{-1,0,1,\infty\}$. Sure enough, this equation corresponds to the hypergeometric equation (IX.11) via the (unbranched) double cover

$$x \in \mathbb{CP}^1 \setminus \{-1,0,1,\infty\} \mapsto x^2 \in \mathbb{CP}^1 \setminus \{0,1,\infty\}.$$

Similarly, when an affine transformation φ of order 3 permutes the 3 singular points 0, 1 and λ, then $\lambda^2 - \lambda + 1 = 0$ and $c = -\frac{\lambda+1}{3}$. Once again the uniformizing equation passes to the quotient under the action of $\varphi(x) = 1 - \frac{x}{\lambda}$ and one retrieves the hypergeometric equation uniformizing the orbifold sphere with indices $\left(\frac{1}{3},0,0\right)$. These are the only two cases where one can determine c by means of symmetries. For example, although the Möbius transformation $\varphi(x) = \frac{\lambda}{x}$ also permutes the 4 singular points — whatever the value of λ — one observes that every equation (IX.12) is left invariant by φ, so that c cannot be determined.

IX.3.2. The sphere with $r + 1$ points removed

We now choose $r + 1$ distinct points of $\mathbb{CP}^1$. We assume one of these points is the point at infinity and denote the others by $a_1, \ldots, a_r$. A calculation analogous to that of the preceding example shows that the uniformizing equation for $\mathbb{CP}^1 \setminus \{a_1, \ldots, a_r, \infty\}$ has the form

$$\frac{d^2v}{dx^2} + \left\{ \frac{1}{4} \sum_{j=1}^{r} \frac{1}{(x - a_j)^2} - \frac{Q(x)}{\prod_{j=1}^{r}(x - a_j)} \right\} v = 0, \tag{IX.13}$$

where Q is a polynomial of degree $r - 2$ with leading term $\frac{r-1}{4} x^{r-2}$. All the other coefficients of Q are "accessory parameters".

When the points $a_1, \ldots, a_r$ are permuted by an affine rotation φ of order r, say $\varphi(x) = \mu x$ and $a_i = \mu^i$ with μ a primitive r^{th} root of unity, then the invariance of the uniformizing equation under φ yields $Q(x) = \frac{(r-1)x^{r-2}}{4}$. The equation can also be obtained by lifting the hypergeometric equation with indices $\left(\frac{1}{r}, 0, 0\right)$ via the branched covering map $x \mapsto x^r$.

IX.3.3. Lamé's equation and the torus with a point removed

In the case of a curve X of genus 1, given say in Legendre's form by

$$y^2 = x(x-1)(x-\lambda), \quad \lambda \in \mathbb{C} \setminus \{0, 1\}, \tag{IX.14}$$

the uniformizing equation of a projective structure with a single orbifold singularity at the point $x = \infty$ is projectively equivalent to Lamé's equation

$$\frac{d^2v}{dx^2} + \frac{1}{2}\left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{x-\lambda}\right)\frac{dv}{dx} + \frac{c - \frac{n(n+1)}{4}x}{x(x-1)(x-\lambda)} v = 0.$$

The index of the equation at the singular point (on the curve X) is $2n+1$. This is a special case of Heun's equation, except that we are not considering it on $\mathbb{CP}^1$, but rather on its elliptic double cover X. In other words, every projective structure on X with a Fuchsian singularity of index $2n + 1$ at the point at infinity, derives, via the double cover

$$X \to \mathbb{CP}^1 \; ; \; (x, y) \mapsto x,$$

from a projective structure on $\mathbb{CP}^1$ singular at 0, 1, λ and ∞, with respective indices

$$\frac{1}{2}, \quad \frac{1}{2}, \quad \frac{1}{2} \quad \text{and} \quad n + \frac{1}{2}.$$

The equation becomes non-singular at infinity precisely when $n = 0$ (or $n = -1$, by the symmetry of the equation under $n \mapsto -n-1$). For $n = -\frac{1}{2}$, we obtain the uniformizing equation for the affine curve $X \setminus \{\infty\}$ (that is, for the projective curve X with the point at infinity removed). We are able to determine the accessory parameter c in the same two cases as were described in §IX.3.1 for $\mathbb{CP}^1$ with 4 points removed.

When $\lambda = -1$, the Möbius transformation $\varphi(x) = -x$ affords a symmetry of the curve and thence a uniformizing equation: one finds that $c = 0$. The map $(x, y) \mapsto x^2$ induces a branched covering $X \to \mathbb{CP}^1$ of degree 4 and we see that our Lamé equation is just the lift of the hypergeometric equation with indices $\left(\frac{1}{4}, \frac{1}{2}, 0\right)$.

Similarly, when $\lambda^2 - \lambda + 1 = 0$, the transformation $\varphi(x, y) = \left(1 - \frac{x}{\lambda}, y\right)$ defines an automorphism of the curve X of order 3, which allows one to determine that $c = \frac{\lambda+1}{48}$. Passing to the quotient by the group of order 6 generated by φ and the elliptic involution, one obtains the hypergeometric equation with indices $\left(\frac{1}{3}, \frac{1}{2}, 0\right)$.

IX.3.4. Hyperelliptic curves

A normal equation without singular points on the hyperelliptic curve of genus g

$$y^2 = P(x), \quad P(x) = \prod_{j=1}^{2g+1} (x - a_j)$$

is projectively equivalent to a unique equation of the form

$$\frac{d^2 v}{dx^2} + \frac{1}{2} \frac{P'(x)}{P(x)} \frac{dv}{dx} + \frac{A(x)y + B(x)}{P(x)} v = 0 \tag{IX.15}$$

where A and B are polynomials of degrees satisfying $\deg(A) \leq g - 3$ and $\deg(B) = 2g - 1$, with B having leading term $\frac{g(g-1)}{4} x^{2g-1}$. The absence of a term of the form $\frac{y}{(x-a_j)^2}$ in the coefficient of v is a necessary and sufficient condition for there to be no logarithmic singularity on the curve at the branch points. One also verifies that the set of normal equations has dimension $3g - 3$ (and that for $g \leq 2$ they depend only on the variable x).

For the highly symmetric pair of curves

$$y^2 = x^{2g+1} - x \quad \text{and} \quad y^2 = x^{2g+1} - 1,$$

the uniformizing equation is given by

$$A = 0 \quad \text{and} \quad B(x) = \frac{g-1}{8(2g+1)} \frac{d^2}{dx^2} \left(\sum_{j=1}^{2g+1} (x - a_j) \right).$$

Once more we have here an avatar of the hypergeometric equation.

In the case

$$y^2 = x^{2g+1} - 1,$$

the projection $p : x \mapsto x^{2g+1}$ induces a branched covering (of degree $4g + 2$) by the hyperelliptic curve of genus g of the sphere with three conical points at 0, 1 and ∞, of angles $\frac{2\pi}{2g+1}$, π and $\frac{\pi}{2g+1}$ respectively. Figure IX.2 below represents the case of genus 2. The only uniformizing equation of type (IX.15) on the curve

$$y^2 = x^{2g+1} - 1$$

is obtained by lifting the corresponding hypergeometric equation.

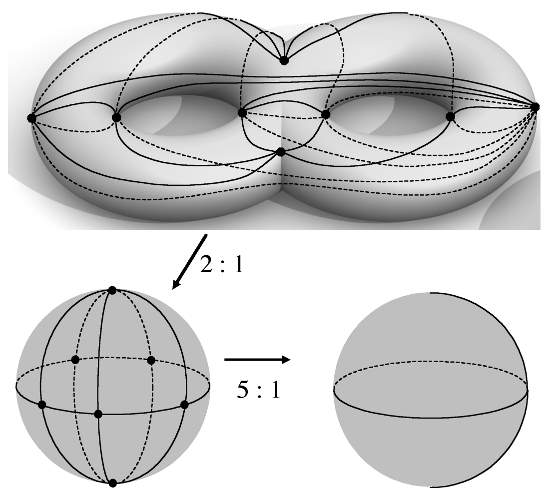


Figure IX.2: A branched covering of degree 10

In the case

$$y^2 = x^{2g+1} - x,$$

the projection $p : x \mapsto x^{2g}$ induces a branched covering (of degree $4g$) by the hyperelliptic curve of genus g of the sphere with three conical points at 0, 1 and ∞ of angles $\frac{\pi}{2g}$, $\frac{\pi}{g}$ and $\frac{\pi}{2g}$ respectively. The only uniformizing equation of type (IX.15) on the curve

$$y^2 = x^{2g+1} - x$$

is obtained by lifting the corresponding hypergeometric equation.

IX.3.5. Curves of arbitrary genus

As we mentioned earlier, in the second volume of the complete works of Schwarz, there is an addendum to [Schw1873], in which he revisits several points in that article. These were added after the works of Klein and those of Poincaré on Fuchsian functions had appeared. Schwarz begins by reformulating the different cases he had investigated earlier in terms of hyperbolic and spherical geometry. Then he evokes by means of several examples the major fact that had escaped him when he wrote the paper [Schw1873], namely the property of hypergeometric equations with parameters $\theta_i = \frac{1}{k_i}$ of allowing many algebraic curves to be uniformized. In fact, provided a curve X admits a covering $\pi : X \to \mathbb{CP}^1$ branched precisely over 0, 1 and ∞ whose fibres are totally ramified of order k_0, k_1 and k_∞ respectively, then the uniformizing equation of the curve X is obtained by lifting the corresponding hypergeometric equation via π. By way of example Schwarz gives the following family of curves, of which the above examples are all special cases, namely the family of curves X with equation

$$y^m = x^n(1 - x^p)^q \tag{IX.16}$$

where $m, n, p, q \in \mathbb{N}^*$. Such a curve X is irreducible if and only if $(m, n, q) = 1$ (that is, m, n and q are relatively prime).

To see this, note that the projection $(x, y) \mapsto x$ induces a branched covering[6] $\pi : X \to \mathbb{CP}^1$ of degree m; the monodromy around $x = 0$ and $x = 1$, given respectively by $y \mapsto e^{2i\pi\frac{n}{m}}$ and $y \mapsto e^{2i\pi\frac{q}{m}}$, acts transitively on the fibre provided (n, q) is relatively prime to m. The smooth part of the curve X is thus connected, whence X is irreducible.

Composing π with $\pi' : x \mapsto x^p$, we obtain a branched covering $\Pi : X \to \mathbb{CP}^1$ of degree mp ramified precisely over the points 0, 1 and ∞. Above $x = 0$, the curve X has exactly (m, n) branches with a local parametrization given by

$$t \mapsto \left(t^{\frac{m}{(m,n)}}, t^{\frac{n}{(m,n)}} u(t)\right)$$

where $u(0)$ is a (m, n)th root of unity (which depends on the branch chosen). On each branch Π is given by $\Pi : t \mapsto t^{\frac{mp}{(m,n)}}$. Thus the fibre of Π above 0 is totally ramified to the order $\frac{mp}{(m,n)}$. Above $x = 1$ the curve has (m, q) branches parametrized by $t \mapsto (1 + t^{\frac{m}{(m,q)}}, t^{\frac{q}{(m,q)}} u(t))$ on which $\Pi(t) = (1 + t^{\frac{m}{(m,q)}})^p$ is ramified to the order $\frac{m}{(m,q)}$. The calculation is similar when x ranges over the other pth roots of unity, and the fibre of Π above 1 is totally ramified to the order $\frac{m}{(m,q)}$. An analogous calculation shows that the fibre of Π above ∞ is totally ramified to

[6] Abusing notation, we understand X as denoting the disjoint union of the Riemann surfaces associated with the irreducible components of the singular curve.

the order $\frac{mp}{(m,n+pq)}$ and the uniformizing equation of the curve X is the lift via Π of the hypergeometric equation with indices

$$(\theta_0, \theta_1, \theta_\infty) = \left(\frac{(m,n)}{pm}, \frac{(m,q)}{m}, \frac{(m,n+pq)}{pm} \right).$$

The Riemann–Hurwitz formula gives us the genus of the curve X:

$$g(X) = 1 + \frac{pm - (m,n) - p(m,q) - (m,n+pq)}{2},$$

and when $m = 2k+1$ and $n = p = q = 1$, this gives $g(X) = k$. In this way one obtains explicit uniformizations of curves of any genus.

IX.3.6. Revisiting Klein's quartic

We now return to Klein's quartic (see Chapter V), given by the equation

$$X^3Y + Y^3Z + Z^3X = 0$$

in $\mathbb{CP}^2$. The projection

$$(X : Y : Z) \mapsto (X^3Y : Y^3Z : Z^3X)$$

induces a cyclic covering of order 7 by the curve over

$$\mathbb{CP}^1 = \{(a : b : c) \in \mathbb{CP}^2 \mid a + b + c = 0\}.$$

We have $(Y/Z)^7 = ab^2/c^3$ so that a point $(X : Y : Z)$ is completely determined by the point $(a : b : c)$ and the choice of a 7th root Y/Z of ab^2/c^3. Setting $y = Y/Z$ and $x = -b/c = -Y^3/Z^2X$, we see that Klein's quartic is birationally equivalent to the curve with equation

$$F(x,y) = y^7 - x^2(x-1) = 0,$$

a particular case of the family of curves considered in the preceding subsection. The uniformizing equation of this curve is projectively equivalent to the equation

$$\frac{d^2v}{dx^2} + \frac{12}{49}\frac{x^2 - x + 1}{x^2(x-1)^2}v = 0.$$

Despite a century of effort most of the known cases of explicit uniformization generally speaking reduce to the hypergeometric equation. A notable exception will be considered at the end of the present chapter. The difficulty of the problem is doubtless to be found in the real analytic nature of the uniformizing section $S \mapsto E_0(S)$ (see §IX.5.1).

IX.4. Uniformization of spheres with 4 points removed

Klein's quartic was the first example of a Riemann surface of genus at least 2 demonstrating the uniformization theorem. In Chapter VI we explained how Fuchsian groups allow one to uniformize a whole open set of the moduli space of curves (with orbifold singularities) of a given genus g. The first example of a moduli space for which one could prove that all its curves are uniformizable was that of spheres with 4 points removed. In this case Poincaré was able to put the method of continuity to work in completely rigorous fashion. In the present section we shall follow Poincaré in proving a particular case of the uniformization theorem while staying at a relatively elementary level.

IX.4.1. A space of polygons

We return to the polygon considered in Example VI.2.2 in the case $n = 3$: here there are the 4 cycles

$$\{0\}, \quad \{1\}, \quad \{\infty\} \quad \text{and} \quad \{x, y, z\}$$

(see Figure IX.3 below). The generators of the group Γ are given by

$$\varphi_0(w) = \frac{w}{1+rw}, \quad \varphi_1(w) = \frac{(1+s)w - s}{sw + (1-s)} \quad \text{and} \quad \varphi_\infty(w) = w - t,$$

where

$$r = \frac{1}{x} - \frac{1}{z}, \quad s = \frac{1}{z-1} - \frac{1}{y-1} \quad \text{and} \quad t = x - y.$$

Thus in this case the group Γ is determined by the polygon. The isotropic subgroup of a point x is generated by $\varphi_0 \circ \varphi_1 \circ \varphi_\infty$; this transformation is parabolic if and only if

$$x(1-z) = z(1-y). \qquad \text{(IX.17)}$$

The parameters x, y and z are real and subject to the conditions

$$x < 0 < z < 1 < y.$$

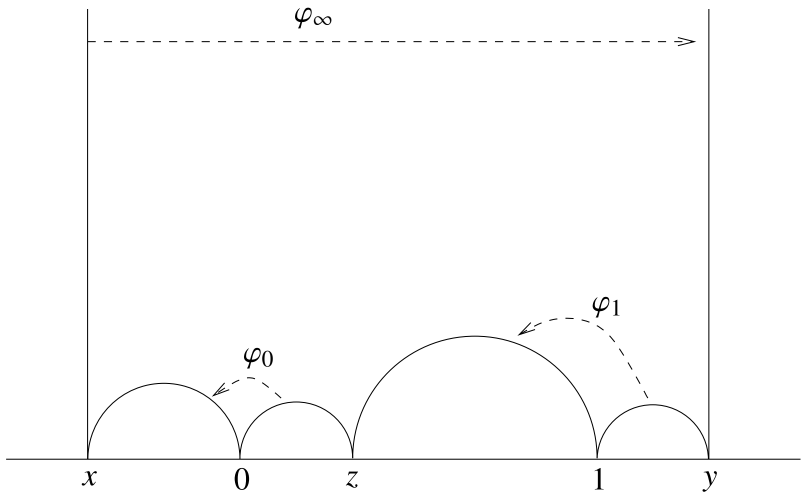


Figure IX.3: A fundamental polygon

To see this, note that for every point (x, y) belonging to the parameter space

$$T := \{(x, y) \mid x < 0 \text{ and } 1 < y\},$$

the point $z = \frac{x}{1+x-y}$ (defined by the condition (IX.17)) is immediately seen to satisfy $0 < z < 1$: thus the region T of the plane is precisely the space of polygons of the above form (with 0, 1 and ∞ fixed).

The quotient of $\mathbb{H}$ by the group Γ is the Riemann sphere with 4 points removed, namely the images of the above cycles, say $0,\ 1,\ \infty$ and λ.

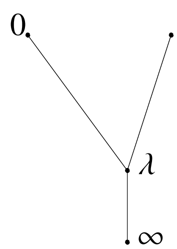


Figure IX.4: The quotient (with cuts)

The space of parameters for the quotient is therefore

$$\mathbb{CP}^1 \setminus \{0, 1, \infty\} \quad \ni \quad \lambda.$$

Following in Poincaré's footsteps, we shall prove the following theorem by elementary means.

Theorem IX.4.1 (Uniformization of spheres with 4 points removed)

The map

$$\Pi : T \to \mathbb{CP}^1 \setminus \{0, 1, \infty\} \ ; \ (x, y) \mapsto \lambda$$

is surjective.

We know already that the map Π is continuous (see §VI.3.3). Thus it suffices to prove that it is both open and closed.

IX.4.2. Openness

We have already proved the openness in §VIII.5.2 in the case of a smooth and complete curve of genus $g > 1$. But rather than adapting that proof to the non-compact case of interest to us here, we use Poincaré's argument.

The price we have to pay for this is the use of the theorem on the invariance of the domain, which, although proved by Brouwer[7] only a considerable time after the appearance of the works of Poincaré we are considering here, seems to have been regarded by Poincaré himself as something obvious: *If* Π *is a continuous map between two manifolds of the same dimension that is locally injective, then it is also locally surjective and therefore open.*

Now local injectivity can be proved using the same argument as was used in the proof of Proposition VIII.3.17. If two points (x_1, y_1) and (x_2, y_2) of T have the same image λ, then the corresponding polygons P_1 and P_2 are the fundamental regions of two Fuchsian uniformizations

$$\pi_1, \pi_2 : \mathbb{H} \to \mathbb{CP}^1 \setminus \{0, 1, \lambda, \infty\}.$$

Hence $\pi_2 = \pi_1 \circ \varphi$ for some automorphism φ of $\mathbb{H}$ and the two groups are conjugate. In other words, P_1 and P_2 are, to within a conjugation (by φ), fundamental regions for the same Fuchsian group. Hence if P_1 and P_2 are close, then φ will be close to the identity, whence in fact $P_1 = P_2$.

[7] We quote in this connection words of Freudenthal published in the book marking Poincaré's centenary:

> The principle of continuity and the concept of a topological manifold attracted the attention of Brouwer, who was then able to create by these means his proof of the invariance of the domain the indispensable and fundamental methods that Topology has used from that time till this.

IX.4.3. Closure

For the proof of Theorem IX.4.1 it remains to show that the map Π is closed. To this end we consider a sequence of points λ_n in the image converging to $\lambda_\infty \in \mathbb{CP}^1 \setminus \{0, 1, \infty\}$, with the aim of showing that the latter point is also in the image of the map. By assumption, there exists a sequence of polygons $P_n \in T$ such that $\Pi(P_n) = \lambda_n$. If the sequence P_n has a cluster point P_∞ in the interior of T, then it follows immediately, by the continuity of Π, that there is a subsequence of the sequence λ_n converging to $\Pi(P_\infty)$; thus in this case $\lambda_\infty = \Pi(P_\infty)$ is in the image. We therefore need only consider the case where the sequence P_n approaches the boundary of T.

Suppose for instance that the corresponding coordinate x tends to 0; then z will also tend to 0 or else y will tend to 1 (these two cases not being exclusive). We consider in detail the first case. Suppose for the moment that y has a closure point satisfying $1 < \hat{y} < \infty$. Then the polygon P will approach a simpler polygon $\hat{P}$ having only the vertices 0, 1, $\hat{y}$ and ∞. The transformations φ_1 and φ_∞ will tend to the transformations

$$\hat{\varphi}_1(w) = \frac{(1+\hat{s})w - \hat{s}}{\hat{s}w + (1-\hat{s})} \quad \text{and} \quad \hat{\varphi}_\infty(w) = w - \hat{t}$$

with

$$\hat{s} = \frac{\hat{y}}{\hat{y}-1} \quad \text{and} \quad t = -\hat{y}$$

respectively. On the other hand $r = \frac{y}{x} - 1$ will diverge and also φ_0, so we may ignore this.

Now in order for the group $\hat{\Gamma}$ generated by $\hat{\varphi}_1$ and $\hat{\varphi}_\infty$ to be Fuchsian, it is necessary and sufficient that

$$\hat{\varphi}_1 \circ \hat{\varphi}_\infty(w) = \frac{w}{(\hat{y}-1)^2 - \hat{y}w}$$

be parabolic, that is, that $\hat{y} = 2$, or else

$$\frac{x}{z} \to -1.$$

In this case we can argue as follows.

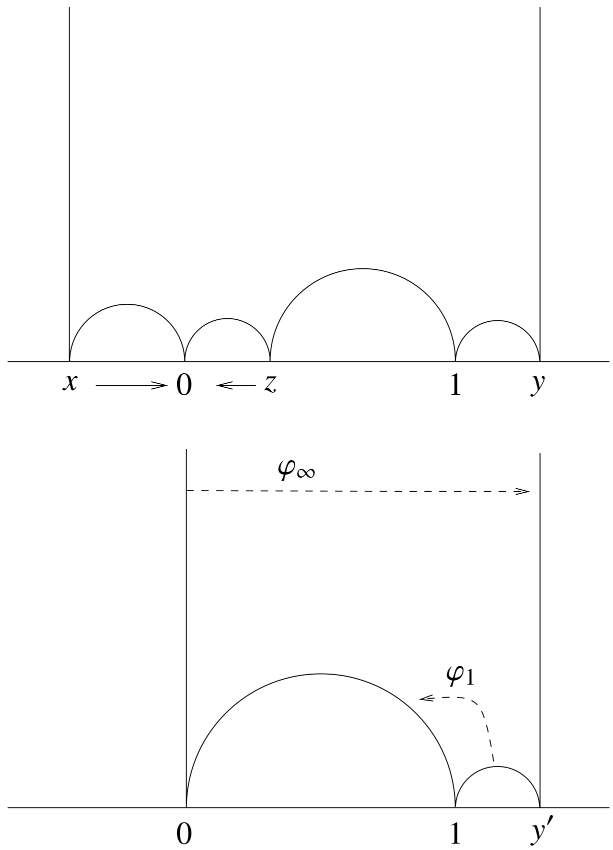


Figure IX.5: The limit polygon

As the polygon P approaches the polygon $\hat{P}$, the tiling determined by P will, on a disc $D(z_0, N)$ of ever-increasing size, coincide with the tiling determined by the subgroup $\hat{\Gamma}$ generated by $\hat{\varphi}_1$ and $\hat{\varphi}_\infty$. Hence for a fixed rational function f the difference between the corresponding Poincaré series

$$\theta(z) = \sum_{\varphi\in\Gamma} f\circ\varphi(z)\cdot(\varphi'(z))^\nu \quad\text{and}\quad \hat{\theta}(z) = \sum_{\varphi\in\hat{\Gamma}} f\circ\varphi(z)\cdot(\varphi'(z))^\nu$$

will by Lemma VI.4.2 become negligible since each approaches the series $\hat{\theta}(z)$ corresponding to the limit group $\hat{\Gamma}$ generated by $\hat{\varphi}_1$ and $\hat{\varphi}_\infty$. One then infers the continuity of the curve and the Fuchsian equation from the continuity of Fuchsian functions: as P tends to $\hat{P}$, the invariant $\lambda = \Pi(P)$ tends to 0, giving a contradiction.

In the general case, we can avoid having the sequence of polygons P_n converging to the point $(x, y) = (0, 2)$ of the boundary of T as x tends to 0, by exploiting the non-uniqueness of the sequence P_n. One may indeed easily construct a sequence of polygons by means of successive modifications of a given polygon P_0, with the property that it converges artificially to the boundary of T while the image under Π remains constant. In order to avoid this sort of situation, Poincaré proposes choosing the sequence P_n (preimages of the λ_n) as follows. One sets up an exhaustion function on T, such as, for instance,

$$F(x, y) = -\frac{x(y-1)}{(1-x)y(y-1-x)}.$$

It is easy to check that F is positive on T, tends to 0 as one approaches the boundary of T, and has $(x, y) = (-1, 2)$ as its only maximum point, where it takes the value $\frac{1}{8}$. Now for each λ_n one chooses a point $P_n \in \Pi^{-1}(\lambda_n)$ maximizing F while remaining as far as possible from the boundary. We shall now verify that if nonetheless the sequence P_n should approach the boundary of T as x approaches 0, then the coordinate y must approach 2, yielding a contradiction.

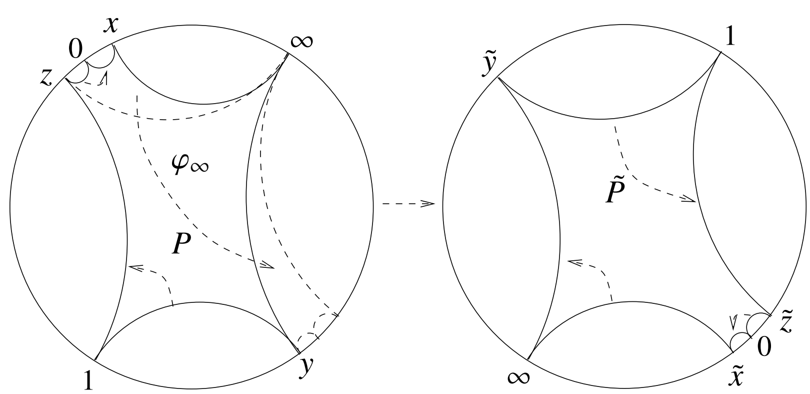


Figure IX.6: Modification of the polygon

We first consider how one can modify a polygon P without changing its image λ. Thus if we cut the polygon P along the geodesic joining z to ∞ and then translate the left-hand portion by means of φ_∞, we obtain a new polygon with vertices z, 1, y, $y-x$, $y-x+z$ and ∞. We then apply the Möbius transformation $w \mapsto \frac{w-y+x}{w-1}$ to bring the new cycles $\{y-x\}$, $\{\infty\}$ and $\{1\}$ to $\{0\}$, $\{1\}$ and $\{\infty\}$, respectively. We thus obtain a new polygon $\tilde{P}$ and denote by Φ the transformation induced on T by the modification $P \mapsto \tilde{P}$:

$$\Phi : T \to T\ ;\ (x, y) \mapsto \left(\frac{x}{y-1}, \frac{z-y+x}{z-1}, \frac{z}{y-x+z-1} \right).$$

In fact Π is not Φ-invariant but rather Φ^2-invariant: we have permuted the cycles and if $\Pi(P) = \lambda$, then $\Pi(\Phi(P)) = \frac{\lambda}{\lambda-1}$. However, this does not constitute a difficulty in connection with our present preoccupation since if we know how to uniformize the curve for λ, then we can also do it for $\frac{\lambda}{\lambda-1}$.

Proposition IX.4.2 — *The above-defined transformation Φ is of infinite order and without a periodic orbit. A fundamental region for it is given by*

$$D = \{x + 2 \leq y \leq -x + 2\}.$$

On each orbit F is maximal precisely on D.

Proof. — Consider the function $G(x, y) = \frac{2-y}{x}$ giving, up to sign, the slope of the straight line determined by $(x, y) \in T$ and the point $(0, 2)$ of the boundary. Then the difference

$$G \circ \Phi - G = -\frac{x^2 + y^2 - 2xy + 2x - 4y + 4}{x}$$

is positive and bounded below by 2, this bound being attained precisely on the straight line $G(x, y) = -1$. It follows immediately that Φ has no fixed points and traverses the region D defined by $-1 \leq G(x, y) \leq 1$ exactly once (except for the boundary points). A straightforward though tedious calculation now shows that $F \circ \Phi - F$ has the same sign on T as $x + 2 - y$, so that indeed the fundamental region D maximizes F on each of the orbits of Φ. □

Thus if a sequence P_n, a lift of λ_n maximizing F, is such that the coordinate x approaches 0, then the sequence necessarily tends to the point $(x, y) = (0, 2)$ of the boundary of T, whence one obtains a contradiction as before. More generally, if the sequence P_n approaches the boundary of T, then at least two of the vertices approach each other arbitrarily closely and after permuting the vertices suitably, one obtains a contradiction as before.

IX.4.4. The action of the modular group

At the end of his text Poincaré returns to this example in order to fill in the details, in particular to give a complete description of the covering T, and we shall now follow suit. For the sake of precision we introduce two further modifications, namely Schwarz's reflection in the geodesic joining z to ∞ on the one hand, and on the other the reflection in the geodesic joining y to 0.

From this we obtain the following two involutions acting on the space of polygons:

$$\sigma_1 : T \to T \; ; \; (x, y, z) \mapsto (1 - y, 1 - x, 1 - z)$$

and

$$\sigma_2 : T \to T \; ; \; (x, y, z) \mapsto \left(\frac{z}{z-1}, \frac{y}{y-1}, \frac{x}{x-1} \right).$$

The locus of fixed points of σ_1 is the half-line $x + y = 1$. The Schwarz reflection conjugating P to $\sigma(P)$ commutes with the corresponding Fuchsian groups and induces a conjugation of the quotients. Since it fixes ∞ and permutes 0 and 1, this will be the reflection

$$\tilde{\sigma}_1 : \mathbb{CP}^1 \setminus \{0, 1, \infty\} \to \mathbb{CP}^1 \setminus \{0, 1, \infty\} \; ; \; \lambda \mapsto 1 - \overline{\lambda}$$

in the line $\left\{ \mathrm{Re}(\lambda) = \frac{1}{2} \right\}$. In other words, the map Π "semi-conjugates" σ_1 to $\tilde{\sigma}_1$. In particular, the fixed points σ_1 are sent to the fixed points of $\tilde{\sigma}_1$.

Similarly, the involution σ_2 fixes the half-line $y = 2$ pointwise and is semi-conjugated, via Π, to the reflection

$$\tilde{\sigma}_2 : \mathbb{CP}^1 \setminus \{0, 1, \infty\} \to \mathbb{CP}^1 \setminus \{0, 1, \infty\} \; ; \; \lambda \mapsto \frac{\overline{\lambda}}{\overline{\lambda} - 1}$$

in the circle $\{|\lambda - 1| = 1\}$. The map Π also sends the half-line $y = 2$ into this circle. Finally, the map

$$\sigma_3 := \sigma_2 \circ \Phi : T \to T \; ; \; (x, y, z) \mapsto \left(\frac{z}{1 + x - y}, \frac{x - y + z}{1 + x - y}, \frac{x}{1 + x - y} \right)$$

is also an involution, one fixing the segment $y = x + 2$ pointwise, and semi-conjugate to the reflection in the real axis

$$\tilde{\sigma}_3 : \mathbb{CP}^1 \setminus \{0, 1, \infty\} \to \mathbb{CP}^1 \setminus \{0, 1, \infty\} \; ; \; \lambda \mapsto \overline{\lambda}.$$

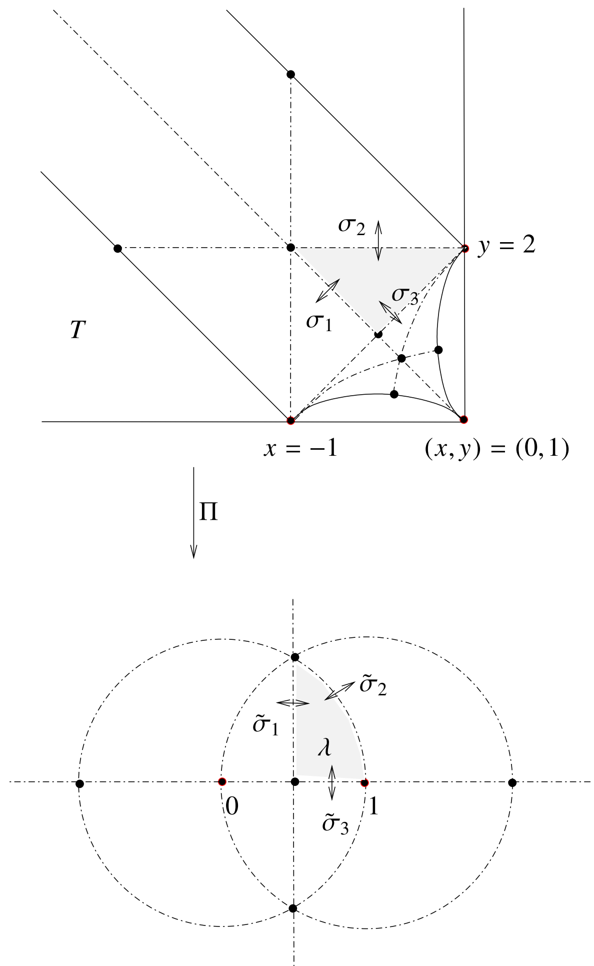


Figure IX.7: The action of the modular group

The action on T of the group $H := \langle\sigma_1,\sigma_2,\sigma_3\rangle$ generated by these involutions is in fact conjugate by means of a homeomorphism to the action on $\mathbb{H}$ of the reflection group of the hyperbolic triangle with angles 0, $\frac{\pi}{2}$ and $\frac{\pi}{3}$. The map Π semi-conjugates this action to that of the reflection group $\tilde{H} := \langle\tilde\sigma_1,\tilde\sigma_2,\tilde\sigma_3\rangle$ of order 12 on $\mathbb{CP}^1\setminus\{0,1,\infty\}$. The orientation-preserving subgroup $H' = \langle\sigma_1\sigma_2,\sigma_2\sigma_3,\sigma_1\sigma_3\rangle$ of index 2 in H is conjugate to $\mathrm{PSL}(2,\mathbb{Z})$; it satisfies $(\sigma_1\sigma_2)^2 = (\sigma_1\sigma_3)^3 = \mathrm{Id}$ and $\Phi = \sigma_2\sigma_3$. The subgroup G of index 12 in H leaving Π (or λ) invariant corresponds to the action on T of the modular group of the sphere with 4 distinguished points (0, 1, ∞ and λ). It is generated by Φ^2 and Ψ^2, where $\Psi = (\sigma_1\sigma_2)(\sigma_1\sigma_3)$, and is conjugate to the subgroup $\Gamma(2) \subset \mathrm{PSL}(2,\mathbb{Z})$ of index 6 consisting of the matrices congruent to I modulo 2. Figure IX.7 shows a fundamental region for G as well as the subtiling induced by H.

The vertices of the tiling belong to polygons possessing symmetries; they are sent to the symmetric locations of λ. One thus infers the following explicit uniformizations (each of which reduces to triangle groups via finite coverings):

$$T \ : \ \begin{cases} \left(-1, 2, \frac{1}{2}\right) & \mapsto \quad \frac{1+i\sqrt{3}}{2} \\ \left(-\frac{1}{2}, \frac{3}{2}, \frac{1}{2}\right) & \mapsto \quad \frac{1}{2}. \end{cases}$$

In fact the upper shaded tile in Figure IX.7, which has these two points as vertices, is sent to the lower shaded tile. To see this note first that the fixed points of σ_i are sent to those of $\tilde{\sigma}_i$, $i = 1, 2, 3$; it therefore suffices to verify that as one approaches the third vertex $(0, 2, 0) \in \partial T$ of the upper tile, the image under Π approaches 1 in the lower diagram. This is precisely the argument we used earlier, and we infer from it that the map $\Pi : T \to \mathbb{CP}^1 \setminus \{0, 1, \infty\}$ is a Galois covering of the group G.

IX.5. Posterity

IX.5.1. Uniformization of complex algebraic varieties

By reviewing Poincaré's version of the uniformization theorem but considering this time analytic families of algebraic curves and their associated families of normal equations, Griffiths [Gri1971] was able to prove the following beautiful theorem, a little-known[8] generalization of a weak version of the uniformization theorem to complex algebraic varieties of any dimension.

Theorem IX.5.1. — *Let V be a complex, quasi-projective, smooth and irreducible algebraic variety of dimension n. For any given point $x \in V$, there exists a neighborhood U of x open in the Zariski topology, such that the universal cover $\widetilde{U}$ of U is homeomorphic to a ball and biholomorphic to a bounded region of $\mathbb{C}^n$.*

[8]We thank J.-B. Bost for bringing this reference to our attention.

We quote Griffiths:

> ... all known complete proofs [of the uniformization theorem] seem to be potential-theoretic and offer very little insight into just how to explicitly locate the Fuchsian D.E.

and he goes on to propose considering an analytic family $\pi : U \to B$ of algebraic curves (that is, such that each preimage $\pi^{-1}(b)$ is an algebraic curve C_b). We can then define an affine holomorphic fibration with base B

$$\mathbf{E} \to B$$

with fibre $\mathbf{E}_b$ the space $E(C_b)$ of normal equations on the curve C_b. The Fuchsian equation then determines a section

$$f : B \to \mathbf{E}$$

of the fibration. This section is defined locally as the preimage under the holomorphic map $\mathrm{Mon}_S : E \to \mathcal{R}_{\mathbb{C}}(S)$ of the real analytic submanifold $\mathcal{R}_{\mathbb{R}}(S)$ of the complex manifold $\mathcal{R}_{\mathbb{C}}(S)$, the space of conjugacy classes of representations of the fundamental group $\pi_1(S)$ of the underlying topological surface of C_b (see §VIII.5). It follows that the map f is real analytic and Griffiths suggests the problem of characterizing f as a solution of an explicit differential equation. Hitchin [Hit1987] has found an apparently satisfying solution to this problem.

For similar reasons Poincaré's take on uniformization stimulated wider interest in abstract algebraic geometry, especially from the 1970s and the appearance of the early work of Ihara [Iha1974] on p-adic uniformization up till the recent work of Mochizuki [Moc1999]. In conclusion, we return to Poincaré's treatment but now over an arbitrary field.

IX.5.2. Algebraic uniformization

Let k be a field of characteristic zero and X an irreducible algebraic curve over k. Write $k(X)$ for the function field of X and let x be any particular non-constant function in $k(X)$. Denote by $R = k(X)[d/dx]$ the ring of differential operators on X; we shall use the notation $D = d/dx$ and $v' = dv/dx$. A *differential equation* on X is defined to be an equation of the form $Lv = 0$ with $L \in R$; we shall in what follows denote such an equation simply by L.

An equivalent formulation consists in regarding a differential equation as a left R-module M generated by an element v: one sets $M = R/RL$ and $v = 1 + RL \in M$. One can then form the tensor product $(M \otimes_{k(X)} N, v \otimes w)$ of two differential equations (M,v) and (N,w). The action of $D = d/dx$ on $(M \otimes_{k(X)} N, v \otimes w)$ is given by $D(p \otimes q) = (Dp) \otimes q + p \otimes (Dq)$.

By means of these definitions one can reformulate Poincaré's treatment of uniformization over the ground field k. A differential equation $D^n + f_1 D^{n-1} + \cdots + f_n$ is called *Fuchsian* if it satisfies the following two conditions:

1. At every point $p \in X$ that is neither a pole nor a critical point of x, the function f_i has at worst a pole of order i;
2. When x is replaced by another function, the same holds for those p at which that new function has neither pole nor critical point.

(Thus the definition does not depend on the choice of x.)

It follows from the proof of Fuchs's theorem that the class of Fuchsian equations is closed under tensor products. Finally, two differential equations (M,v) and (N,w) are called *projectively equivalent* if there exists a first-order differential equation (A,a) such that

$$(M,v) \cong (N,w) \otimes (A,a).$$

Let $L = D^2 + fD + g$ be a second-order differential equation. For every $a \in k(X)$, there exists a unique differential equation of the form $D^2 + aD + b$ projectively equivalent to L. For $a = 0$, one obtains $D^2 + (g - f'/2 - f^2/4)$ (see Proposition VIII.3.4). The uniformization theorem may now be formulated as follows:

Theorem IX.5.2. — *For every curve X over $\mathbb{C}$, there exists a unique function $h \in \mathbb{C}(X)$ such that the quotient of two solutions of the differential equation*

$$v'' + hv = 0 \tag{IX.18}$$

is the developing map of a hyperbolic structure on X. The equation (IX.18) is Fuchsian.

We then call the equation (IX.18) the *uniformizing equation* of the curve X. The following natural question still seems to be open.

Question: Which curves X defined over $\overline{\mathbb{Q}}$ have uniformizing equation also defined over $\overline{\mathbb{Q}}$?

The answer to this question is known for a particular class of algebraic curves already uniformized, namely arithmetic surfaces. We recall briefly the construction of those among such surfaces called *rational*, introduced by Poincaré for the first time in [Poin1887], and whose associated Fuchsian groups are historically the first such groups to have been constructed.

Let A be a quaternion algebra over $\mathbb{Q}$ and suppose that $A \otimes \mathbb{R} = \mathcal{M}_2(\mathbb{R})$, the algebra of 2×2 real matrices. Let $\mathcal{O}$ be an order in A and $\mathcal{O}^1$ the set of its elements of (reduced) norm 1. Then $\mathcal{O}^1$ is a Fuchsian group in $\mathrm{SL}(2,\mathbb{R})$ and every rational arithmetic Fuchsian group in $\mathrm{SL}(2,\mathbb{R})$ is commensurable in the wide sense to such a group. (We recall that two subgroups of a group are called *commensurable* if their intersection has finite index in both and *commensurable in the wide sense* if one is commensurable to a conjugate of the other.) Two rational arithmetic lattices in $\mathrm{SL}(2,\mathbb{R})$ associated with quaternion algebras A_1 and A_2 are commensurable in the wide sense if and only if A_1 and A_2 are isomorphic [MaRe2003].

The following theorem is a corollary of a theorem of Ihara [Iha1974, Theorem A].

Theorem IX.5.3. — *Let $\Gamma \subset SL(2,\mathbb{R})$ be a rational arithmetic Fuchsian group. Then the curve $X = \Gamma\backslash\mathbb{H}$ and its uniformizing equation are both defined over $\overline{\mathbb{Q}}$.*

Proof. — Let $\mathcal{O} \subset \mathcal{M}_2(\mathbb{R})$ be a maximal order in the quaternion algebra A associated with the group Γ. By a result of Shimura [Shi1959], $\mathcal{O}^1\backslash\mathbb{H}$ is the moduli space of Abelian surfaces with multiplication by $\mathcal{O}$; in particular it is defined over $\mathbb{Q}$. Hence the curve X is defined over $\overline{\mathbb{Q}}$. It remains to investigate the uniformizing equation.

Let g be a rational element of norm 1 in A. The groups Γ and $g\Gamma g^{-1}$ are commensurable, and we write $X_0 = (g\Gamma g^{-1} \cap \Gamma)\backslash\mathbb{H}$, $X_1 = X$, $X_2 = g\Gamma g^{-1}\backslash\mathbb{H}$, and consider the following diagram of finite Galois coverings:

$$\begin{array}{ccc} & X_0 & \\ p_1 \swarrow & & \searrow p_2 \\ X_1 & & X_2 \end{array} \qquad \text{(IX.19)}$$

Let L_i $(i = 0,1,2)$ be the uniformizing equation of X_i. The equation L_0 is obtained by lifting the equation L_i via p_i $(i = 1,2)$.

Since the rational elements of A are dense in A and the only Lie subgroup of $\mathrm{SL}(2,\mathbb{R})$ containing Γ is $\mathrm{SL}(2,\mathbb{R})$ itself, we may so choose g that the group generated by Γ together with g is dense in $\mathrm{SL}(2,\mathbb{R})$. We may then characterize the uniformizing equation L_0 as the only Fuchsian equation on X_0 invariant under the groups of the coverings p_1 and p_2 to within projective equivalence. Indeed, a Fuchsian equation on X_0 invariant under the groups of the coverings p_1 and p_2 will differ from the uniformizing equation by a holomorphic quadratic differential $q(z)dz^2$ invariant under a dense subgroup of $\mathrm{SL}(2,\mathbb{R})$, which immediately forces $q \equiv 0$.

The coverings p_1 and p_2 obviously being defined over $\overline{\mathbb{Q}}$, we infer immediately that the uniformizing equation of X_0 (and therefore that of X) is defined over $\overline{\mathbb{Q}}$. □

All examples where one has been able to explicitly determine the uniformizing equation of a curve appertain to the case of the hypergeometric equation and therefore to a Fuchsian group commensurable with a triangle group. In the following subsection we discuss briefly a different example. In connection with both this and the above question we mention the following conjecture of Krammer [Kram1996], which if settled would furnish an answer to that question.

Conjecture IX.5.4. — *Let $X = \Gamma\backslash\mathbb{H}$ be a curve and L its uniformizing equation. Then X and L can both be defined over $\overline{\mathbb{Q}}$ if and only if Γ is arithmetic or commensurable with a triangle group.*

IX.5.3. A final example

We conclude the chapter by considering Krammer's use [Kram1996] of the idea of the proof of Theorem IX.5.3 in order to calculate explicitly the uniformizing equation of a particular arithmetic curve. We present here several points of his remarkable construction by way of comparison with the approaches of Poincaré and Ihara.

Krammer considers a rational arithmetic cocompact Fuchsian group $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$ for which the curve $X = \Gamma\backslash\mathbb{H}$ is isomorphic to $\mathbb{CP}^1$ with 4 conical points of orders 2, 2, 2 and 6. Note that the double cover branched over these 4 points is an elliptic orbifold curve with a single conical point of order 3.

Thus we start with $\mathbb{CP}^1$ with three conical points 0, 1 and $a \in \mathbb{C} \setminus \{0,1\}$ of order 2 and the conical point ∞ of order e. The uniformizing equation is projectively equivalent to the Lamé equation (with $P(x) = x(x-1)(x-a)$):

$$P(x)v'' + \frac{1}{2}P'(x)v' + \left(C - \frac{n(n+1)}{4}x\right)v = 0, \qquad \text{(IX.20)}$$

where $n = \frac{1}{e} - \frac{1}{2}$ and C is an accessory parameter to be determined. In our case $e = 6$, we need to determine the point a from the group (or find the algebraic equation of X) and determine the value of C for which the equation is the uniformizing equation.

In the spirit of the proof of Theorem IX.5.3, Krammer considers a rational element g of norm 1 in A, and the corresponding diagram (IX.19). He manages to choose g so that:

1. the maps p_1 and p_2 are of degree 3;
2. the curve X_0 (like $X_1 = X$ or X_2) has genus 0;

3. the element g acts on X_0 as an involution; and

4. there exist points $x_1, \ldots, x_8 \in X_0$, $y_1, \ldots, y_4 \in X_1$ such that p_1 ramifies only above the points y_i, and g permutes the x_i other than x_2 and x_4 according to the following table, where n_i denotes the ramification index of p_1 at x_i.

i	1	2	3	4	5	6	7	8
$g(x_i)$	x_8	?	x_5	?	x_3	x_7	x_6	x_1
n_i	1	2	1	2	1	1	1	3
$p_1(x_i)$	y_1	y_1	y_2	y_2	y_3	y_3	y_3	y_4

The combinatorics of this table suffice for the algebraic equations of the maps $p_1 : X_0 \to X_1$ and $g : X_0 \to X_0$ to be determined. We may take $X_0 = X_1 = \mathbb{CP}^1$ and assume that $x_2 = y_1 = 0$, $x_4 = y_2 = 1$ and $x_8 = y_4 = \infty$. We begin by checking that this completely determines p_1. Since p_1 has degree 3 and fixes the point 0 with ramification of degree 2 and ∞ with degree-3 ramification, we have $p_1(x) = ax^2 + bx^3$. Then since $p_1(1) = 1$ and $p_1'(1) = 0$, we obtain the values of a and b, whence

$$p_1(x) = 3x^2 - 2x^3.$$

We therefore have $x_1 = 3/2$, $x_3 = -1/2$ and g interchanges $3/2$ and ∞, $-1/2$ and x_5, and x_6 and x_7. After translating through $3/2$, we obtain an involution of $\mathbb{CP}^1$ interchanging 0 and ∞, -2 and $x_5 - 3/2$, and $x_6 - 3/2$ and $x_7 - 3/2$. This involution must then be of the form $x \mapsto p/x$, whence

$$-2\left(x_5 - \frac{3}{2}\right) = \left(x_6 - \frac{3}{2}\right)\left(x_7 - \frac{3}{2}\right). \qquad \text{(IX.21)}$$

Furthermore, since $p_1(x_5) = p_1(x_6) = p_1(x_7)$, the numbers x_6 and x_7 are solutions in t, different from x_5, of the equation $p_1(t) = p_1(x_5)$, that is,

$$3t^2 - 2t^3 = 3x_5^2 - 2x_5^3 \Leftrightarrow (t - x_5)(3(t + x_5) - 2(t^2 + tx_5 + x_5^2)) = 0$$
$$\Leftrightarrow t^2 + (t + x_5)\left(x_5 - \frac{3}{2}\right) = 0.$$

Hence

$$(t - x_6)(t - x_7) = t^2 + (t + x_5)\left(x_5 - \frac{3}{2}\right)$$

for all t. Setting $t = 3/2$ we have in view of (IX.21) that

$$-2\left(x_5 - \frac{3}{2}\right) = \left(\frac{3}{2} - x_6\right)\left(\frac{3}{2} - x_7\right) = \left(\frac{3}{2}\right)^2 + \left(x_5 + \frac{3}{2}\right)\left(x_5 - \frac{3}{2}\right) = x_5^2$$

whence

$$x_5^2 + 2x_5 - 3 = 0.$$

Then since $x_5 \neq x_4 = 1$, we have $x_5 = -3$, $y_3 = p_1(-3) = 81$, $p = -2(x_5 - 3/2)$ and x_6, x_7 are roots of $2t^2 - 9t + 27 = 0$. We assemble these results in the following table.

i	1	2	3	4	5	6,7	8
x_i	$\frac{3}{2}$	0	$-\frac{1}{2}$	1	-3	$\frac{3}{4}(3 \pm \sqrt{15})$	∞
$p_1(x_i)$	0	0	1	1	81	81	∞

We now return to equation (IX.20). We have $e = 6$ and $a = 81$, and it remains to determine C. It is at this juncture that the additional symmetry given by g becomes crucial. One demands that the equations obtained by lifting equation (IX.20) via the maps p_1 and $p_2 = g \circ p_1$ are projectively equivalent. With the help of a computer Krammer verifies that this forces $C = -1/2$.

Thus Krammer obtains the following theorem, the first example, as far as we know, not an avatar of the hypergeometric equation. Ihara's method has been applied by Elkies in a systematic way to other examples of a similar type (rational Shimura curves with 4 conical points) in [Elk1998]. These are, we believe, the only examples where one knows how to determine the uniformizing equation explicitly.

Theorem IX.5.5. — *The Fuchsian differential equation*

$$P(x)v'' + \frac{1}{2}P'(x)v' + \frac{x-9}{18}v = 0,$$

where $P(x) = x(x-1)(x-81)$*, is the uniformizing equation of the orbifold elliptic curve* $y^2 = P(x)$ *with a single conical point of order* 3 *(at infinity). Its group is not commensurable with any triangle group.*

Proof. — It remains to show only that the group of the equation we have obtained is not commensurable with any triangle group. This follows from the classification of the arithmetic triangle groups by Takeuchi [Tak1977], according to which their quaternion algebras are all different from that considered by Krammer. □

Intermezzo

Chapter X

Uniformization of surfaces and the equation $\Delta_g u = 2e^u - \varphi$

Before proceeding to the general uniformization theorem, we wish to present an alternative approach to the uniformization of algebraic Riemann surfaces, probably originating with Schwarz. It would seem, indeed, to have been at Schwarz's initiative that the Göttingen Royal Society of Sciences decided to draw mathematicians' attention to the connections between uniformization of surfaces and the solution of the equation $\Delta u = ke^u$. Here is an approximate translation of the topic of the competition proposed by this learned society in 1890 [Got1890, Page IX]:

> The problem of representing a region of the plane (that is, a region of the complex plane or of a Riemann surface extended over the plane) conformally on a portion of a curved surface of constant curvature k is related to the problem of integrating the partial differential equation
>
> $$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2k\, e^u$$
>
> with prescribed singularities and boundary values.
>
> For this problem one must in the first place be concerned with the boundary values and the singularities as specified by Riemann in his theory of Abelian functions.
>
> The Royal Society wishes for a complete answer to the following question: is it possible to integrate the above differential equation on a given region, with prescribed boundary values and singularities of a certain type, under the assumption that the constant k has a negative value [?]
>
> In particular, the Royal Society would like to see the above question answered in the case where the region of the plane under consideration is a

> closed Riemann surface of several sheets, and where the function u admits only logarithmic singularities[1].

Let us explain this a little. Given a biholomorphism f from a region S of the complex plane to a region of the unit disc $\mathbb{D}$, one can lift the Poincaré metric of the disc to S, obtaining thereby a metric on S conformally equivalent to the Euclidean metric, that is, of the form $g = e^u dz d\bar{z}$. Furthermore, g is of constant curvature -1, which is equivalent to the function u satisfying the equation $\Delta u = 2e^u$. This generalizes to the case where S is a Riemann surface extended over the plane and f is no longer a global biholomorphism, but the multivalued inverse of a Fuchsian function. The only difference is that the function u will then have singularities at the branch points of S. To summarize: if a Riemann surface S extended over the plane is uniformized by the disc, then this surface supports a solution of the equation $\Delta u = 2e^u$ with singularities of a certain type at the branch points of S. The converse of this statement is equally valid, so that one can appreciate the interest for the uniformization of surfaces in solving the equation $\Delta u = 2e^u$.

The problem set by the Göttingen Royal Society was very quickly solved: Picard published a solution in 1890 [Pic1890]. However, in this memoir he only shows the existence of solutions of the equation $\Delta u = ke^u$ (with singularities of a certain type) in the case of a bounded region of $\mathbb{C}$.[2] Moreover he later felt the need to return several times to his proof in order to clarify certain points: in 1893 [Pic1893c, Pic1893b, Pic1893a], in 1898 [Pic1898], and in 1905 [Pic1900]. It

[1]*Die Aufgabe der conformen Abbildung eines ebenen Bereiches auf ein Stück einer krummen Fläche, deren Krümmungmass überall den constanten Werth k besitz, hängt zusammen mit der Aufgabe, die partielle Differentialgleichung*

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2k\, e^u$$

vorgeschriebenen Grenz- und Unstetigkeitsbedingungen gemäss zu integriren.

Für diese Aufgabe kommen zunächst die von Riemann in seiner Theorie der Abelschen Functionen angegebenen Grenz- und Unstetigkeitsbedingungen in Betracht.

Die Königliche Gesellschaft wünscht die Frage, ob es möglich ist, die angebene partielle Differentialgleichung für einen gegebenen Bereich unter vorgeschriebenen Grenz- und Unstetigkeitsbedingungen der angegebenen Art zu integriren, vorausgesetzt, dass der Konstanten k negative Werthe beigelegt werden, vollständig beanwortet zu sehen.

Insbesondere wünscht die Königliche Gesellschaft den Fall der angeführten Aufgabe behandelt zu sehen, in welchen der betrachtete eben Bereich, eine geschlossene mehrfach zusammenhängende Riemannsche Fläche ist, während die Function u keine anderen als logarithmische Unstetigkeiten annehmen soll.

[2]Picard asserts that the case of a closed Riemann surface presents no additional difficulties, but — as he himself later recognizes — this is a very significant underestimation of those difficulties.

seems to us that the last of these articles does indeed contain a complete and rigorous proof of the existence and uniqueness of solutions of the equation $\Delta u = ke^u$, with prescribed singularities, on any closed Riemann surface. For a complete and very clear exposition of this proof, see [Pic1931, Chapter 4].

In 1898, Poincaré in turn published a memoir in response to the question posed by the Göttingen Royal Society [Poin1898][3]. That memoir is entitled *Fuchsian functions and the equation* $\Delta u = e^u$; however, contrary to what this title might lead one to expect, the memoir does not actually deal with the equation $\Delta u = e^u$ — and this is in fact one of the main reasons for its being of interest! For the equation $\Delta u = ke^u$ only makes sense on a region of $\overline{\mathbb{C}}$, or on a Riemann surface S extended over $\overline{\mathbb{C}}$. In order to link this equation with the problem of the existence of a uniformizing parameter for a surface S, one therefore needs to choose a meromorphic coordinate on S (thus allowing S to be viewed as extended over $\overline{\mathbb{C}}$), and look for solutions of the equation with singularities at the branch points and on the fibre at infinity. In his memoir, Poincaré replaces this equation by another one which has the effect of limiting the situation to objects defined in an intrinsic manner on S, namely the equation $\Delta_g u = 2e^u - \varphi$.

In the present chapter, we shall follow Poincaré's memoir step by step. Thus, in particular, we shall have yet another opportunity of coming face to face with his genius. Earlier we admired Poincaré's creative power, capable of building *ex nihilo* a new mathematical universe populated with Fuchsian groups, automorphic functions, and projective structures, as well as his almost incredible ability to manipulate concepts of the most extreme abstraction. But what delights the reader of the 1898 memoir is, on the contrary, Poincaré's ability to solve a difficult analytic problem using only perfectly elementary arguments. In the way he looks for solutions in the form of series, bounds their terms, employs normal convergence, etc., one sometimes has the impression one is reading the solution to a problem from the entrance examination to a large engineering school! Yet by the end of his memoir, Poincaré has obtained a result — the existence of solutions of the equation $\Delta_g u = \theta e^u - \varphi$ and the existence of metrics of prescribed curvature on compact surfaces — that remains nontrivial a century later! However, if the techniques Poincaré uses are highly classical, on the other hand the fact that he deals with the intrinsic equation $\Delta_g u = 2e^u - \varphi$ rather than the equation $\Delta u = 2e^u$ is in contrast very modern.

Remark X.0.6. — Although it may seem like a historical slight, we have chosen not to expound the work of Picard, preferring to concentrate on that of Poincaré. There are several reasons for this. First of all, as we have already mentioned, Picard expounded his method of solution of the equation $\Delta u = ke^u$ very well in

[3] In this memoir Poincaré states clearly that the question has already been answered by Picard in 1890 — but then proceeds as if Picard's solution did not exist.

his book [Pic1931, Chapter 4]; it would be pointless to reproduce here that classic work. Secondly, Picard's proof depends on a judicious use of *Schwarz's alternating procedure*, and this technique plays a major role in Chapters IV and XI; we have therefore opted to "diversify pleasures" by giving Poincaré's proof, which employs somewhat different arguments. Finally, it seemed to us that Poincaré's article is more innovative than Picard's. For, on the one hand Poincaré goes over to an intrinsic variant of the equation $\Delta u = ke^u$ while Picard continues working "in terms of a meromorphic coordinate", and on the other hand Poincaré's proof is, unlike that of Picard, purely global in nature[4].

X.1. Uniformization of surfaces and the equation $\Delta_g u = 2e^u - \varphi$

We begin by describing in detail the various ways in which the equation $\Delta_g u = 2e^u - \varphi$ is related to the uniformization of surfaces.

X.1.1. From the existence of a uniformizing parameter to the existence of a solution of the equation $\Delta_g u = 2e^u - \varphi_g$

Let S be a Riemann surface, and assume S is uniformized by the disc; in other words, assume there exists a biholomorphism $F : \mathbb{D} \to \widetilde{S}$ (where $\widetilde{S}$ is the universal cover of S). Denote by $f : \mathbb{D} \to S$ the map induced by F.

The automorphisms of the covering $f : \mathbb{D} \to S$ are biholomorphisms of the disc $\mathbb{D}$, that is, elements of PSL(2,$\mathbb{R}$). Now the elements of PSL(2,$\mathbb{R}$) are not just automorphisms of the holomorphic structure of $\mathbb{D}$, but also isometries with respect to the standard hyperbolic metric on $\mathbb{D}$. Hence f defines a Riemannian metric[5] on the surface S, induced from the standard hyperbolic metric on $\mathbb{D}$.

In terms of a single coordinate. — Consider a holomorphic local coordinate $z : U_z \to \mathbb{C}$ defined on an open set of S. (Note that the coordinate z need not be injective on U_z, but simply a holomorphic immersion; typically one chooses a

[4]In this connection, we quote from Poincaré's commentary on his own work made at the request of Mittag-Leffler [Poin1921]:

M. Picard was the first to integrate it [Poincaré is talking here of the equation $\Delta u = ke^u$]. The method I have proposed is entirely different [. . .]. What characterizes my method and distinguishes it from M. Picard's, is that it immediately embraces the whole of the Riemann surface, whereas M. Picard first considers a restricted region, and then extends his results more and more widely until they are established for the whole surface.

[5]In fact, this Riemannian metric, viewed as an intrinsic object on S, does not appear explicitly in Poincaré's article. Instead, Poincaré chooses holomorphic local coordinates on S with domains of definition covering S, and considers in terms of each of these coordinates the formula for the metric induced from the hyperbolic metric via f.

meromorphic function $z : S \to \mathbb{C} \cup \{\infty\}$, and restricts to an open set U_z where z is finite-valued with non-vanishing derivative.) On this chart, the pull-back by f^{-1} of the hyperbolic metric of $\mathbb{D}$ can be written as

$$g_{\text{hyp}} = 4\frac{\left|\frac{df^{-1}}{dz}\right|^2}{(1-|f^{-1}|^2)^2}dzd\overline{z}.$$

(The inverse f^{-1} of f is of course multivalued, but the above expression is independent of the particular choice of preimage.) We see that, in particular, the metric g_{hyp} is conformally equivalent to the metric $dzd\overline{z}$. We denote by e^{u_z} the factor of conformality linking these two metrics. In other words, we consider the function $u_z : U_z \to \mathbb{R}$ defined by

$$e^{u_z} = 4\frac{\left|\frac{df^{-1}}{dz}\right|^2}{\left(1-|f^{-1}|^2\right)^2} .$$

We then have

$$u_z = \log 4 + \log\frac{df^{-1}}{dz} + \log\frac{d\overline{f^{-1}}}{d\overline{z}} - 2\log\left(1 - f^{-1}\overline{f^{-1}}\right),$$

whence, taking into account the fact that f^{-1} is holomorphic and $\overline{f^{-1}}$ anti-holomorphic, we obtain

$$\frac{\partial^2 u_z}{\partial z\partial\overline{z}} = \frac{2\frac{df^{-1}}{dz}\frac{d\overline{f^{-1}}}{d\overline{z}}}{\left(1-f^{-1}\overline{f^{-1}}\right)^2} = \frac{1}{2}e^{u_z}.$$

Thus the function $u_z : U_z \to \mathbb{R}$ is a solution of the equation

$$\Delta_z u = 2e^u, \qquad \text{(X.1)}$$

where $\Delta_z = 4\frac{\partial^2}{\partial z\partial\overline{z}}$ is the Laplace operator associated with the coordinate z.

The intrinsic point of view. — The most important contribution of Poincaré's memoir when compared with Picard's articles is the introduction of an intrinsic point of view. To achieve this, he endows the surface S with a Riemannian metric g compatible with the complex structure[6] (see §III.1.1). Thus we shall

[6]Poincaré calls this an "*isotropic metric*". Note that at that time the existence of such metrics was far from obvious (the notion of a partition of unity not having yet been invented!); Poincaré says that this follows from work of Schwarz and Klein, and that "the proof is fairly long". He also believes he may use *anisotropic* metrics, but makes a significant mistake at that juncture. (His expression for the Laplace–Beltrami operator in terms of a local coordinate is valid only for a metric compatible with the complex structure.)

be comparing two objects defined globally on S, namely, the Riemannian metrics g_{hyp} (induced from the hyperbolic metric on the disc via f) and g.

For every holomorphic local coordinate z defined on an open set U_z of S, we consider the function $\sigma_z : U_z \to \mathbb{R}$ satisfying

$$dz d\overline{z} = e^{\sigma_z} g.$$

(Thus e^{σ_z} is the conformality factor linking the metrics $dz d\overline{z}$ and g).

If z_1 and z_2 are two holomorphic local charts, then on their region of overlap we have

$$\Delta_{z_2} = \left| \frac{dz_1}{dz_2} \right|^2 \Delta_{z_1}, \tag{X.2}$$

whence

$$e^{\sigma_{z_1}} \Delta_{z_1} = e^{\sigma_{z_2}} \Delta_{z_2}.$$

Hence there exists an operator $\Delta_g : C^2(S, \mathbb{R}) \to C^0(S, \mathbb{R})$ such that, on the domain of every holomorphic local chart z of S, we have

$$\Delta_g = e^{\sigma_z} \Delta_z.$$

The operator Δ_g is, of course, the Laplace–Beltrami operator associated with the Riemannian metric g.

We also have that the Riemannian metrics g_{hyp} and g are conformally equivalent. In other words, there exists a function $u_g : S \to \mathbb{R}$ satisfying

$$g_{\text{hyp}} = e^{u_g} g.$$

Thus for every holomorphic local coordinate z, we have the two equations $g_{\text{hyp}} = e^{u_z} dz d\overline{z} = e^{u_g} g$ and $dz d\overline{z} = e^{\sigma_z} g$, whence

$$u_g = u_z + \sigma_z$$

on the domain of definition of the coordinate z.

Hence for every holomorphic local coordinate z, we have:

$$\begin{aligned}\Delta_g u_g &= \Delta_g u_z + \Delta_g \sigma_z \\ &= e^{\sigma_z} \Delta_z u_z + \Delta_g \sigma_z \\ &= 2e^{\sigma_z} e^{u_z} + \Delta_g \sigma_z \\ &= 2e^{u_g} + \Delta_g \sigma_z .\end{aligned}$$

This calculation shows that the quantity $\Delta_g \sigma_z$ is independent of the choice of holomorphic local coordinate z. In other words, there exists a function $\varphi_g : S \to \mathbb{R}$ such that, on the domain of every holomorphic local chart z, we have:

$$\varphi_g = -\frac{1}{2} \Delta_g \sigma_z = -\frac{1}{2} e^{\sigma_z} \Delta_z \sigma_z .$$

(The reason for the factor $\frac{1}{2}$ will be apparent later on.) Most significantly, the above calculation shows that the function $u_g : S \to \mathbb{R}$ is a solution of the equation

$$\Delta_g u = 2e^u - 2\varphi_g . \tag{X.3}$$

To summarize, Poincaré has to this point established the following

Proposition X.1.1. — *Let S be a Riemann surface endowed with a Riemannian metric g compatible with its complex structure, and consider the function $\varphi_g : S \to \mathbb{R}$ such that for every holomorphic local coordinate z, one has $dzd\overline{z} = e^{\sigma_z} g$ and $\varphi_g = -\frac{1}{2}\Delta_g \sigma_z$ on the domain of definition of z. Then if the surface S is uniformized by the disc, equation (X.3) admits a solution $u : S \to \mathbb{R}$.*

X.1.2. How to obtain a uniformizing parameter from a solution of the equation $\Delta_g u = 2e^u - 2\varphi_g$

Consider again a Riemann surface S endowed with a Riemannian metric g compatible with its complex structure. *Supposing now that this surface is uniformized by the disc*, let $f : \mathbb{D} \to S$ be a uniformization. (Recall that f is unique modulo composition with an element of $\mathrm{PSL}(2,\mathbb{R})$.) We also assume that the equation $\Delta_g u = 2e^u - 2\varphi_g$ has a unique solution $u_0 : S \to \mathbb{R}$. The aim of this section is to explain — *à la* Poincaré — how one can retrieve the uniformizing parametrization f from u_0 and g.

Choose a meromorphic function $z_0 : S \to \mathbb{C} \cup \{\infty\}$, and denote by U_{z_0} the region of S where z_0 is finite and a local diffeomorphism. Write $\sigma_{z_0} : U_{z_0} \to \mathbb{R}$ for the function defined by $dz_0 d\overline{z}_0 = e^{\sigma_{z_0}} g$. We know (from Corollary VIII.3.7) that the (multivalued) inverse of the function f is expressible as a quotient

$$f^{-1} = \frac{v_2}{v_1} .$$

Here v_1 and v_2 are two multivalued functions from S to $\mathbb{C}$ forming a basis for the solutions of a Fuchsian equation, which, in terms of the coordinate z_0, has the form

$$\frac{d^2v}{dz_0^2} = \sigma v, \tag{X.4}$$

where the function $\sigma : S \to \mathbb{C} \cup \{\infty\}$ is determined uniquely. (It is the Schwarzian derivative of f — but this will not be of particular import here.) To within composition with an element of $\mathrm{PSL}(2,\mathbb{R})$, the function f is independent of the choice of the basis (v_1, v_2) of solutions. Thus finding f reduces to finding the function σ.

In terms of each holomorphic local coordinate z defined on an open set U_z of S, we consider functions u_z, σ_z and u_g defined as in §X.1.1. We showed above that the function u_g is a solution of the equation (X.3). Since by assumption u_0 is the unique solution of that equation, we have $u_0 = u_g$. Hence

$$e^{u_0 - \sigma_{z_0}} = 4\frac{\left|\frac{df^{-1}}{dz_0}\right|^2}{(1 - |f^{-1}|^2)^2}.$$

In view of the equality $f^{-1} = v_2/v_1$, it then follows that

$$\begin{aligned} e^{-\frac{1}{2}(u_0 - \sigma_{z_0})} &= \frac{1}{2}\left|\frac{df^{-1}}{dz_0}\right|^{-1}\left(1 - |f^{-1}|^2\right) \\ &= \frac{1}{2}\left|\frac{\frac{dv_2}{dz_0}v_1 - \frac{dv_1}{dz_0}v_2}{v_1^2}\right|^{-1}\left(1 - \left|\frac{v_2}{v_1}\right|^2\right) \\ &= \frac{(|v_1|^2 - |v_2|^2)}{2\left(\frac{dv_2}{dz_0}v_1 - \frac{dv_1}{dz_0}v_2\right)}. \end{aligned}$$

The denominator of the last expression is twice the Wronskian of the basis (v_1, v_2) of solutions. This Wronskian is constant (since the equation (X.4) has no term in $\frac{dv}{dz_0}$) and, since the basis (v_1, v_2) for the solutions can be chosen arbitrarily, we may suppose it equal to $\frac{1}{2}$. This assumed, we have

$$e^{-\frac{1}{2}(u_0 - \sigma_{z_0})} = |v_1|^2 - |v_2|^2 = v_1\overline{v_1} - v_2\overline{v_2}.$$

Since $\overline{v_1}$ and $\overline{v_2}$ are anti-holomorphic, it follows that

$$\frac{d^2 e^{-\frac{1}{2}(u_0 - \sigma_{z_0})}}{dz_0^2} = \frac{d^2v_1}{dz_0^2}\overline{v_1} - \frac{d^2v_2}{dz_0^2}\overline{v_2},$$

and then equation (X.4) yields

$$\frac{d^2 e^{-\frac{1}{2}(u_0-\sigma_{z_0})}}{dz_0^2} = \sigma e^{-\frac{1}{2}(u_0-\sigma_{z_0})},$$

whence

$$\sigma = -\frac{1}{2}\frac{d^2(u_0-\sigma_{z_0})}{dz_0^2} + \frac{1}{4}\left(\frac{d(u_0-\sigma_{z_0})}{dz_0}\right)^2.$$

Thus if we know the unique solution u_0 of equation (X.3) and the conformal factor σ_{z_0} distinguishing the metric g from the metric $dz_0 d\overline{z}_0$, we can find the Fuchsian equation giving the uniformizing map f.

In conclusion, if one knows how to solve the equation (X.3) on the surface S, then one can find the unique function admissible as a candidate for a uniformizing map for S — but we do not at this stage know how to *prove* that this unique candidate function does actually uniformize S.

Of course, the precise meaning of the phrase "find the unique function admissible as a candidate for a uniformizing map for S" depends on the sense one gives the phrase "solve the equation (X.3)". Note however, that calculation by any means whatever of a numerical approximation to the unique solution of the equation (X.3) would allow us to infer, via the above formulae, a corresponding approximation to the uniformizing map.

X.1.3. Why the existence of a solution of the equation $\Delta_g u = 2e^u - 2\varphi_g$ implies the uniformization theorem

We shall now explain why, for a given compact Riemann surface S, the existence of a solution of equation (X.3) for an appropriate metric g on S entails the uniformizability of S by the disc. There is no explicit mention of this aspect of the problem in the announcement of the competition by the Göttingen Royal Society of Sciences, nor in Picard's memoirs[7]. On the other hand, although Poincaré seems to be fully conscious of the need to establish this connection[8], he nevertheless does not trouble himself to do so. It may be that in fact he is less interested in such a general abstract result as the existence of uniformizing maps, than in the possibility of constructing them explicitly. One should not forget that Poincaré considers that he and Klein had proved fifteen years before, by means

[7]Picard simply asserts repeatedly that the solution of $\Delta u = e^u$ is of capital importance for the theory of Fuchsian functions.

[8]In the introduction to his memoir, he writes:

The integration of this equation would indeed lead directly to a solution of the problem of interest to us [that of establishing that there always exists a Fuchsian equation, and therefore a uniformizing map].

of the method of continuity, that algebraic Riemann surfaces are uniformizable[9]; however, he emphasizes that the method of continuity is "*extremely complicated and of an indirect character*", never leading to the construction of an explicit uniformizing map for a particular given surface.

We begin by interpreting equation (X.3) in terms of curvature. As before, let S be a Riemann surface, g a Riemannian metric on S compatible with its complex structure, and $\varphi_g : S \to \mathbb{R}$ the function defined at the end of §X.1.1.

Note first that, if D is an open region of $\mathbb{C}$ and $u : D \to \mathbb{R}$ a function of class C^2, then the Gaussian curvature of the Riemannian metric $e^u dzd\overline{z}$ on D is given by

$$-\frac{1}{2}e^{-u}\Delta_z u.$$

(This follows by means of direct calculation; see for example [Jos2002].) It follows that, if z is a holomorphic local coordinate defined on an open set U_z of S, and $u_z : U_z \to \mathbb{R}$ a function of class C^2, then the function u_z is a solution of the equation

$$\Delta_z u = 2e^u$$

if and only if the Gaussian curvature relative to the metric $e^{u_z} dzd\overline{z}$ has the constant value -1. On the open set U_z we may write $g = e^{-\sigma_z} dzd\overline{z}$, whence we have $-\varphi_g = \frac{1}{2}e^{\sigma_z}\Delta_z\sigma_z$ (see §X.1.1). This formula shows that $-\varphi_g$ is none other than the Gaussian curvature with respect to the metric g.

Now let $u : S \to \mathbb{R}$ be a function of class C^2. For every holomorphic local coordinate z defined on an open set U_z of S, we denote by $u_z : U_z \to \mathbb{R}$ the function defined by $u_z = u - \sigma_z$. By repeating the calculations in §X.1.1, one easily verifies that u is a solution of the equation (X.3) if and only if, for every holomorphic local chart z, the function u_z is a solution of the equation (X.1). Furthermore, for every holomorphic local chart z, we have, on the domain of that chart, $e^u g = e^{u_z} dzd\overline{z}$.

This argument establishes the following

Proposition X.1.2 — *A function $u : S \to \mathbb{R}$ of class C^2 is a solution of the equation (X.3) if and only if the Gaussian curvature with respect to the metric $e^u g$ has the constant value* -1.

This interpretation in terms of curvature allows us to show that the uniformization theorem for algebraic Riemann surfaces follows from the existence of solutions of equation (X.3):

[9]Excerpt from the introduction to [Poin1898]:

The first proof that was given was based on what is called *the method of continuity*. M. Klein and I arrived at this method independently. In my Memoir on *groups of linear equations* [. . .] I gave a complete exposition of the method; I have nothing further to add to it. I feel I have brought that method to perfectly rigorous form.

Proposition X.1.3. — *Let S be a compact Riemann surface of negative Euler characteristic, endowed with a Riemannian metric g compatible with its complex structure. Then if the equation*

$$\Delta_g u = 2e^u - \varphi$$

has a solution for every function $\varphi : S \to \mathbb{R}$ with positive integral, the universal covering space of S is biholomorphic to the disc.

Proof. — Let $\varphi_g : S \to \mathbb{R}$ denote the negative of the Gaussian curvature with respect to the metric g. By the Gauss–Bonnet theorem, since S is assumed to have negative Euler characteristic, the integral of φ_g will be positive. The equation $\Delta_g u = 2e^u - 2\varphi_g$ therefore has a solution $u : S \to \mathbb{R}$. By Proposition X.1.2, the Gaussian curvature with respect to the metric $e^u g$ will then have constant value -1. By the Hopf–Rinow theorem, this metric is complete. The universal cover of S endowed with the lift of the Riemannian metric $e^u g$ is thus a simply connected surface equipped with a Riemannian metric of constant curvature -1. However, to within an isometry there is only one such surface: the disc $\mathbb{D}$ with the standard hyperbolic metric. Thus the universal cover of S endowed with the lift of the metric $e^u g$ is isometric to the disc with the hyperbolic metric, and we conclude that the universal cover of S is biholomorphic to the disc $\mathbb{D}$ (since the metric $e^u g$ is compatible with the complex structure of S). □

X.2. How Poincaré solved the equation $\Delta_g u = \theta e^u - \varphi$

In what follows, we consider a compact algebraic Riemann surface[10] S. Given a Riemannian metric g on S compatible with its complex structure, we wish to solve the equation

$$\Delta_g u = \theta e^u - \varphi, \tag{X.5}$$

where $\theta : S \to \mathbb{R}$ is a given positive function (for example, we may take θ to be of constant value 2 if our goal is just to uniformize the surface S) and $\varphi : S \to \mathbb{R}$ is a given function with positive integral.

Poincaré's strategy for solving the equation (X.5) may be summarized as follows. One begins with an equation that one can integrate explicitly. Then one wins territory by attempting to integrate "neighboring" equations by successive developments in series, always approaching little by little the equation (X.5). Here in greater detail is how he proceeds:

[10] Recall that every compact Riemann surface is algebraic. The algebraicity of S allows one to construct explicitly meromorphic 1-forms with prescribed poles on S, and Poincaré makes use of this in his proof.

1. First of all one considers the equation $\Delta_g u = -\varphi$.

   (a) The surface S being assumed algebraic, one knows how to find (explicitly) meromorphic functions with prescribed poles on S. The real parts of such functions then furnish harmonic functions with prescribed singularities on S.

   (b) One next shows that one can solve the Poisson equation $\Delta_g u = -\varphi$ for every function φ with integral zero. In order to do this one constructs a Green's function using the harmonic functions found earlier; the solutions of the equation $\Delta_g u = -\varphi$ are then given in the form of an integral involving φ and that Green's function.

2. Next one considers the equation $\Delta_g u = \eta u - \varphi$.

   (a) Here one shows first that the equation $\Delta_g u = \lambda \eta u - \varphi$ can be integrated for all given functions η and φ provided the real number λ is sufficiently small. This is done by constructing a putative solution of the equation formally as a series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$, then showing that the u_i are solutions of Poisson equations (which one knows how to solve from Step 1), and, finally, proving the series converges for λ sufficiently small.

   (b) Next one shows (via a series development again) that if one can integrate the equation $\Delta_g u = \lambda_0 \eta u - \varphi$ for some λ_0, then one can also integrate the equation $\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi$ provided $\lambda < \lambda_0$.

   (c) From the two preceding steps , one easily infers that one can integrate the equation $\Delta_g u = \lambda \eta u - \varphi$ for all $\lambda > 0$. Thus, in particular, one can integrate the equation $\Delta_g u = \eta u - \varphi$.

3. Finally, one considers the equation $\Delta_g u = \theta e^u - \varphi$.

   (a) One first observes that the equation $\Delta_g u = \theta e^u - \varphi$ has an obvious solution (namely, a constant function) if φ is proportional to θ.

   (b) Then by means of yet another series development one shows that if the equation $\Delta_g u = \theta e^u - \varphi_0$ can be integrated for some function φ_0, then the equation $\Delta_g u = \theta e^u - (\varphi_0 + \lambda\psi)$ can be integrated for every function ψ provided only that λ is sufficiently small.

   (c) From the preceding two steps, one infers that one can integrate the equation $\Delta_g u = \theta e^u - \varphi$ provided the function φ is everywhere positive.

(d) It now remains only to resort to an elementary trick to pass from the case where the function φ is everywhere positive, to that where it is positive only on average.

That may all appear rather laborious, but one should not forget that Poincaré did not have available to him the beautiful modern apparatus of distributions, Sobolev injections, weak compactness, elliptic regularity, etc. In fact, for each equation he examines, Poincaré exhibits a solution, be it in the form of a convolution product with a Green's function or a convergent series each of whose terms is a solution of a "simpler" equation — that is, one that he already knows how to integrate. Poincaré's article having been forgotten, it took till 1971 before finally a proof was published, by M. S. Berger, of the existence of a solution of the equation $\Delta_g u = \theta e^u - \varphi$, in which a large amount modern machinery was brought to bear, namely the aforementioned distributions, Sobolev injections, weak compactness, and elliptic regularity! (See the Box at the end of this chapter.)

Remark X.2.1. — It is perhaps of some interest to note the similarity between the strategy Poincaré adopts here, and that he employed 15 years earlier in his 1884 article on the uniformization of algebraic surfaces using the method of continuity (see Chapter VIII). Here, in order to integrate the equation $\Delta_g u = \theta e^u - \varphi$, Poincaré starts from a partial differential equation for which he has an explicit solution and, deforming it, proceeds via equations that he can integrate until he at last reaches the equation $\Delta_g u = \theta e^u - \varphi$. Similarly, in his article using the method of continuity, starting from an algebraic surface that he can uniformize explicitly, he traces a path through the appropriate moduli space of algebraic surfaces, proceeding via surfaces that he successively discovers how to uniformize, until he finally arrives at the algebraic surface of primary interest.

We shall now trace the sequence of steps whereby Poincaré constructs the unique solution of the equation $\Delta_g u = \theta e^u - \varphi$.

X.2.1. The solution of the equation $\Delta_g u = -\varphi$

The first step in Poincaré's construction consists in the integration of the Poisson equation

$$\Delta_g u = -\varphi, \tag{X.6}$$

where $\varphi : S \to \mathbb{R}$ is a given function of class C^1 and $u : S \to \mathbb{R}$ is the unknown function. Poincaré begins with the remark that, since the surface S is closed, Stokes' theorem implies that the integral of $\Delta_g u$ over S is zero for every function $u : S \to \mathbb{R}$ (of class C^2); hence for the equation (X.6) to have a solution, it is necessary that the integral of the function φ be zero. His next step is to show that this necessary condition is also sufficient.

Proposition X.2.2. — *For every function $\varphi : S \to \mathbb{R}$ of class C^1 and vanishing integral over S, one can construct a solution of class C^2 of the equation*

$$\Delta_g u = -\varphi.$$

To construct such a solution Poincaré's tactic is to "introduce a function that will play the same role as a Green's function in potential theory". In other words, by analogy with the situation in Euclidean space, one seeks a Green's function $G : S \times S \to \mathbb{R}$ such that, for every function $\varphi : S \to \mathbb{R}$ of class C^1 and zero integral, the function $u : S \to \mathbb{R}$ given by the formula

$$u(p) = \int_S G(p,q)\varphi(q)\, dv_g(q)$$

is a solution of the equation (X.6).

In 2-dimensional Euclidean space, the solutions of the Poisson equation are given by this formula with $G(p,q) = \log(\|p-q\|)$. On the surface S the most natural approach would therefore seem to be to use a function G with the property that for all $q \in S$, the map $p \mapsto G(p,q)$ is harmonic on $S \setminus \{q\}$ with a logarithmic singularity at q. Unfortunately, it has long been known — at least since Riemann — that a compact surface does not support a harmonic function with a single logarithmic singularity. One is therefore compelled to work with a function G whose partial function $p \mapsto G(p,q)$ has two logarithmic singularities: one at the point q and the other at a different (fixed) point q_0.

Proof of Proposition X.2.2. — We assume we have a fixed function $\varphi : S \to \mathbb{R}$ of class C^1 and vanishing integral.

Let p_0 be a particular point of S. We will construct a function $u_{p_0} : S \to \mathbb{R}$ representing a solution of the Poisson equation (X.6) vanishing at the point p_0. One should keep in mind that all other solutions of equation (X.6) are obtained from u_0 by adding an arbitrary constant, since the difference between any two (putative) solutions of (X.6) is a harmonic function on S, and the only harmonic functions defined on a closed surface are constant functions.

Let q_0 be a point of S distinct from p_0. For every point $q \in S \setminus \{q_0\}$, consider the unique meromorphic 1-form $\omega_{q_0,q}$ on S satisfying the following conditions (see Proposition II.2.8):

— firstly, $\omega_{q_0,q}$ should have two simple poles at the points q_0 et q, of residues -1 and $+1$ respectively, and no other poles;

— secondly, the real part of every period of $\omega_{q_0,q}$ should vanish, or, in other words, the real part of the integral of $\omega_{q_0,q}$ along any closed curve on S should vanish.

We then consider the function G_{p_0,q_0} defined on the set of pairs $(p,q) \in (S\backslash\{q_0\})^2$ with $p \neq q$, by the formula

$$G_{p_0,q_0}(p,q) = \frac{1}{2\pi}\,\mathrm{Re}\left(\int_{\gamma_{p_0,p}} \omega_{q_0,q}\right),$$

where $\gamma_{p_0,p}$ is any path from the point p_0 to the point p avoiding the points q_0 and q. Note that the quantity $G_{p_0,q_0}(p,q)$ is independent of the choice of the path $\gamma_{p_0,p}$ in view of the condition that the real parts of the periods of $\omega_{q_0,q}$ are zero, and the fact that the integral of a 1-form around a small closed curve encircling a simple pole with real residue is purely imaginary. Proposition II.2.9 implies that this quantity has the following properties:

— the function $(p,q) \mapsto G_{p_0,q_0}(p,q)$ is analytic in both variables and harmonic in the variable p on the set $\{(p,q) \mid p \neq q_0\,,\, p \neq q\}$;

— for every open set U of $S\backslash\{q_0\}$ with compact closure and every holomorphic local coordinate z on U, one can express $G_{p_0,q_0}(p,q)$ at each $(p,q) \in U^2$, $p \neq q$, in terms of the coordinate z, in the form

$$G_{p_0,q_0}(p,q) = H(p,q) + \frac{1}{2\pi}\log|p-q|, \tag{X.7}$$

where H is a function defined on $U \times U$ (including the diagonal), analytic in both variables and harmonic in the variable p.

We can now define a function $u_{p_0} : S \setminus \{q_0\} \to \mathbb{R}$, a candidate for a solution of the equation (X.6), by

$$u_{p_0}(p) := \int_S G_{p_0,q_0}(p,q)\varphi(q)\; dv_g(q). \tag{X.8}$$

That this integral converges is immediate in view of the following facts: the surface S is compact, the function $q \mapsto \varphi(q)$ is of class C^1 on S, the function $q \mapsto G_{p_0,q_0}(p,q)$ is continuous on $S \setminus \{p\}$, and the singularity of this function at p is logarithmic.

The integral in (X.8) above makes no sense when $p = q_0$; however, we shall show later on that the function u_{p_0} defined in (X.8) extends to q_0. Moreover the extension is independent of the choice of the point q_0.

We now need to show that the function u_{p_0} so defined satisfies equation (X.6). To this end, we consider an open set U of $S \setminus \{q_0\}$ with compact closure and sufficiently small for there to exist a holomorphic local coordinate z on it. Then for each $p \in U$, we decompose $u_{p_0}(p)$ as a sum of three terms:

$$u_{p_0}(p) = \int_{S\setminus U} G_{p_0,q_0}(p,q)\varphi(q)\,dv_g(q) + \int_U H(p,q)\varphi(q)\,dv_g(q) + \frac{1}{2\pi}\int_U \log|p-q|\varphi(q)\,dv_g(q),$$

where H is as in equation (X.7). The first term here is an analytic and harmonic function of the variable p; this follows from the fact that the function $(p,q) \mapsto G_{p_0,q_0}(p,q)$ is analytic in both variables and harmonic in the variable p on $U \times (S \setminus U)$. Hence the function $p \mapsto \int_{S\setminus U} G_{p_0,q_0}(p,q)\varphi(q)\,dv_g(q)$ is analytic on U, and for every $p \in U$, we have

$$\Delta_g\left(\int_{S\setminus U} G_{p_0,q_0}(p,q)\varphi(q)\,dv_g(q)\right) = \int_{S\setminus U} \Delta_g G_{p_0,q_0}(p,q)\varphi(q)\,dv_g(q) = 0.$$

Similarly, the second term in the above decomposition of $u_0(p)$ is an analytic and harmonic function of the variable p (since the map $(p,q) \mapsto H(p,q)$ is analytic in both variables and harmonic in the variable p on $U \times U$). It remains to examine the third term of the decomposition of $u_0(p)$. A standard calculation in $\mathbb{R}^2$ shows that, for φ of class C^1, the function $p \mapsto \int_U \log|p-q|\varphi(q)\;dv_g(q)$ is of class C^2 and satisfies

$$\Delta_z\left(\frac{1}{2\pi}\int_U \log|p-q|\varphi(q)\;dz d\bar{z}(q)\right) = -\varphi(p)$$

for all $p \in U$, where $\Delta_z = \frac{\partial^2}{\partial z \partial \bar{z}} = \frac{dv_g}{dz d\bar{z}}\Delta_g$. It follows that for all $p \in D$, we have

$$\Delta_g\left(\frac{1}{2\pi}\int_U \log|p-q|\varphi(q)\;dv_g(q)\right) = -\varphi(p).$$

We have thus shown that u_{p_0} is of class C^2 on U and satisfies the equation

$$\Delta_g u_{p_0}(p) = -\varphi(p) \tag{X.9}$$

for all $p \in U$. In view of the condition that U have compact closure in $S \setminus \{q_0\}$ and be sufficiently small, we infer that u_{p_0} is of class C^2 on $S \setminus \{q_0\}$ and satisfies equation (X.9) for all $p \in S \setminus \{q_0\}$.

It remains to show that the function u_{p_0} defined by (X.8) extends continuously to q_0, and that the resulting extended function has class C^2 (in which case it will automatically satisfy equation (X.6) also at q_0). For this, it suffices to repeat the construction of the function u_{p_0} but now with the point q_0 replaced by a different point $\hat{q}_0$ (still distinct from p_0). The key observation here is the following one: for every point $q \in S \setminus \{q_0, \hat{q}_0\}$, the uniqueness of the 1-form $\omega_{\hat{q}_0,q}$ (see Proposition II.2.8) implies that

$$\omega_{\hat{q}_0,q} = \omega_{q_0,q} + \omega_{\hat{q}_0,q_0}.$$

Hence for every pair of distinct points $p, q \in S \setminus \{q_0, \hat{q}_0\}$, we shall have

$$G_{p_0,\hat{q}_0}(p,q) = G_{p_0,q_0}(p,q) + G_{p_0,\hat{q}_0}(p,q_0),$$

and then, finally, for all $p \in S \setminus \{q_0, \hat{q}_0\}$, we obtain

$$\begin{aligned}
&\int_{S\setminus\{p\}} G_{p_0,\hat{q}_0}(p,q)\varphi(q)\ dv_g(q) \\
&\qquad = \int_{S\setminus\{p\}} G_{p_0,q_0}(p,q)\varphi(q)\ dv_g(q) + G_{p_0,\hat{q}_0}(p,q_0) \int_{S\setminus\{p\}} \varphi(q)\ dv_g(q) \\
&\qquad = \int_{S\setminus\{p\}} G_{p_0,q_0}(p,q)\varphi(q)\ dv_g(q),
\end{aligned}$$

where in the last equality we have used the fact that φ has vanishing integral. We thus see that if we replace the point q_0 by any other point $\hat{q}_0$ in the definition of u_{p_0}, we again obtain a function defined and of class C^2 on $S \setminus \{\hat{q}_0\}$ coinciding with u_{p_0} on $S \setminus \{q_0, \hat{q}_0\}$. Hence the function u_{p_0} defined by (X.8) extends by continuity to q_0, is of class C^2, and satisfies equation (X.6). □

Remark X.2.3. — If the function φ is of class C^k (with $k \geq 1$), then every solution of equation (X.6) will be of class C^{k+1}. To see this, it suffices to imitate the proof of Proposition X.2.2 above, having noted beforehand that, if φ has class C^k, then the function $p \mapsto \int_U \log|p - q|\varphi(q)\ dv_g(q)$ will have class C^{k+1}.

It is also useful to note that if the function φ is merely bounded, then the function $u_{p_0} : p \mapsto \int_S G_{p_0,q_0}(p,q)\varphi(q)\ dv_g(q)$ will be well-defined and of class C^1.

From the integral formula (X.8) for the solution of equation (X.6) vanishing at p_0, we can infer an upper bound on the norm of that solution in terms of the norm of the function φ:

Addendum X.2.4. — *Let $p_0 \in S$ be any particular point. There exists a constant $\beta > 0$ such that, for every function $\varphi : S \to \mathbb{R}$ of class C^1 and vanishing integral, one has*

$$\|u_{p_0,\varphi}\|_\infty \leq \beta\|\varphi\|_\infty \quad \textit{and} \quad \|\vec{\nabla}_g u_{p_0,\varphi}\|_\infty \leq \beta\|\varphi\|_\infty,$$

where $u_{p_0,\varphi}$ is the unique solution of equation (X.6) vanishing at p_0.

Proof. — Choose a point $q_0 \in S$ different from p_0, and a neighborhood U_0 of q_0. By the proof of Proposition X.2.2, whatever the function φ, for all $p \neq q_0$ one has

$$u_{p_0,\varphi}(p) := \int_S G_{p_0,q_0}(p,q)\varphi(q)\ dv_g(q).$$

Hence for all $p \neq q_0$ one has

$$|u_{p_0,\varphi}(p)| \leq \left(\int_S \left|G_{p_0,q_0}(p,q)\right|\ dv_g(q)\right)\ \|\varphi\|_\infty.$$

The integral $\int_S \left|G_{p_0,q_0}(p,q)\right|\ dv_g(q)$ is finite for every $p \in S \setminus \{q_0\}$ and depends continuously on p. Hence the quantity

$$\beta_0 := \sup_{p \in S\setminus U_0} \int_S \left|G_{p_0,q_0}(p,q)\right|\ dv_g(q)$$

is finite, and moreover one has, for all $p \in S \setminus U_0$,

$$|u_{p_0,\varphi}(p)| \leq \beta_0 \|\varphi\|_\infty.$$

Recall from the proof of Proposition X.2.2 that the function u_{p_0} is independent of the choice of the point q_0. Thus if one considers a point $\hat{q}_0 \in S$ different from q_0 and p_0, and a corresponding neighborhood $\widehat{U}_0$ of $\hat{q}_0$, one will have, for all $p \in S \setminus \widehat{U}_0$,

$$|u_{p_0,\varphi}(p)| \leq \widehat{\beta}_0 \|\varphi\|_\infty$$

where

$$\widehat{\beta}_0 := \sup_{p \in S\setminus \widehat{U}_0} \int_S \left|G_{p_0,\hat{q}_0}(p,q)\right|\ dv_g(q).$$

It now remains only to choose the neighborhoods U_0 and $\widehat{U}_0$ sufficiently small for $(S \setminus U_0) \cup (S \setminus \widehat{U}_0) = S$ to hold. It will then follow that

$$\|u_{p_0,\varphi}\|_\infty = \sup_{p \in (S\setminus U_0)\cup(S\setminus \widehat{U}_0)} |u_{p_0,\varphi}(p)| \leq \beta \|\varphi\|_\infty,$$

with $\beta := \max(\beta_0, \widehat{\beta}_0)$. The upper bound for $\|\vec{\nabla}_g u_{p_0,\varphi}\|_\infty$ is then obtained by replacing the function G_{p_0,q_0} (resp. $G_{p_0,\hat{q}_0}$) by its gradient with respect to the variable p in the foregoing argument. □

Remark X.2.5. — If we denote by $C^1_0(S)$ the set of functions $\varphi : S \to \mathbb{R}$ of class C^1 and vanishing integral, then we can interpret Proposition X.2.2 as showing that the linear operator $\Delta_g : C^2(S) \to C^1_0(S)$ is surjective and providing a formula for the inverse of this operator. Addendum X.2.4 then shows that the inverse of the linear operator $\Delta_g : C^2(S) \to C^1_0(S)$ is continuous with respect to the C^1-norm on the space $C^2(S)$.

X.2.2. The solution of the equation $\Delta_g u = \eta u - \varphi$

The second major stage in Poincaré's strategy consists in integrating the equation

$$\Delta_g u = \eta u - \varphi, \tag{X.10}$$

where $\eta : S \to \mathbb{R}$ is a given positive class-C^1 function and $\varphi : S \to \mathbb{R}$ a given class-C^1 function. Starting from the fact that now he can integrate the Poisson equation $\Delta_g u = \varphi$, Poincaré deduces to begin with that he can also integrate any equation of the form $\Delta_g u = \lambda\eta u - \varphi$ with $\lambda > 0$ sufficiently small. Then he shows that if one can integrate the equation $\Delta_g u = \lambda_0 \eta u - \varphi$ for a particular $\lambda_0 > 0$, then one can integrate any equation of the form $\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi$ with λ satisfying $0 < \lambda < \lambda_0$. From these two results he infers immediately that one can integrate any equation of the form $\Delta_g u = \lambda\eta u - \varphi$ with $\lambda > 0$. Taking $\lambda = 1$ in particular, it follows that one can integrate equation (X.10).

Uniqueness of solutions. — Rather early on in his article, Poincaré asserts that the "*fundamental property of the expression* $\Delta_g u$" is the following one:

Fact X.2.6. — *If* $u : S \to \mathbb{R}$ *is a twice differentiable function, then at a point where* u *has a maximum,* $\Delta_g u$ *is non-positive, and at a point where* u *has a minimum,* $\Delta_g u$ *is non-negative.*

Proof. — It is enough to recall that Δ_g is proportional toă $\Delta_z = \frac{\partial^2}{\partial z \partial \bar{z}}$ for every holomorphic local coordinate z. □

Poincaré gives a very convincing physical interpretation of this fact: *a point where the temperature has a maximum can yield heat to neighboring points, but not receive heat from them.* He immediately infers a "maximum principle" for the equation of present interest:

Proposition X.2.7. — *For any given positive* λ*, any positive function* $\eta : S \to \mathbb{R}$*, and any function* $\varphi : S \to \mathbb{R}$*, the partial differential equation*

$$\Delta_g u = \lambda\eta u - \varphi$$

has at most one solution $u : S \to \mathbb{R}$*.*

Proof. — Suppose we have two solutions $u, v : S \to \mathbb{R}$ of the equation $\Delta_g u = \lambda\eta u - \varphi$. Since S is compact, there will be points $p_-, p_+ \in S$ such that the function $u - v$ attains its least value at p_- and its greatest value at p_+. By Fact X.2.6, we shall then have

$$\Delta_g(u - v)(p_-) \geq 0 \quad \text{and} \quad \Delta_g(u - v)(p_+) \leq 0.$$

Furthermore, the function $u - v$ is a solution of $\Delta_g(u - v) = \lambda\eta(u - v)$, and since the function $\lambda\eta$ is positive by assumption, it follows that

$$(u - v)(p_-) \geq 0 \quad \text{and} \quad (u - v)(p_+) \leq 0.$$

However, since the function $u-v$ is greatest at p_+ and least at p_-, these inequalities imply that it must in fact be identically zero. □

Bounding the solutions above a priori. — By exploiting Fact X.2.6 again, Poincaré establishes an *a priori* upper bound for the norm of a putative solution of an equation of type $\Delta_g u = \lambda\eta u - \varphi$:

Proposition X.2.8. — *Let λ be any positive real number, $\eta : S \to \mathbb{R}$ a positive function, and $\varphi : S \to \mathbb{R}$ any function. If $u : S \to \mathbb{R}$ is a solution of the equation $\Delta_g u = \lambda\eta u - \varphi$, then the following inequality holds:*

$$\|u\|_\infty \leq \frac{1}{\lambda}\left\|\frac{\varphi}{\eta}\right\|_\infty.$$

Proof. — Let $p_-, p_+ \in S$ be points where the function u attains its least and greatest values respectively on S. We then have the inequalities $\Delta_g u(p_-) \geq 0$ and $\Delta_g u(p_+) \leq 0$. From the equation $\Delta_g u = \lambda\eta u - \varphi$ and the positivity of η, we obtain in turn the inequalities

$$\frac{\varphi(p_-)}{\lambda\eta(p_-)} \leq u(p_-) \quad \text{and} \quad u(p_+) \leq \frac{\varphi(p_+)}{\lambda\eta(p_+)}.$$

Since u is least at p_- and greatest at p_+, these imply that

$$\frac{\varphi(p_-)}{\lambda\eta(p_-)} \leq u(p) \leq \frac{\varphi(p_+)}{\lambda\eta(p_+)}$$

for all $p \in S$, from which the desired bound for $\|u\|_\infty$ is immediate. □

The solution of the equation $\Delta_g u = \lambda\eta u - \varphi$ for λ sufficiently small. — Poincaré seeks a solution of the equation $\Delta_g u = \lambda\eta u - \varphi$ in the form of a power series in λ, whose convergence he establishes for sufficiently small λ:

Proposition X.2.9. — *Let β be the constant given by Addendum X.2.4. One can find a solution of class C^2 of the equation*

$$\Delta_g u = \lambda\eta u - \varphi \tag{X.11}$$

for every positive class-C^1 function $\eta : S \to \mathbb{R}$, every C^1-function $\varphi : S \to \mathbb{R}$, and every real number $\lambda > 0$ satisfying $2\lambda\beta\|\eta\|_\infty < 1$.

Proof of Proposition X.2.9. — Fix on a positive class-C^1 function $\eta : S \to \mathbb{R}$ and a class-C^1 function $\varphi : S \to \mathbb{R}$. Choose also a particular point $p_0 \in S$.

First step: looking for a series solution.

We write φ in the form $\varphi = \varphi_0 + \lambda\varphi_1$ where φ_0 and φ_1 are functions of class C^1, the first with vanishing integral. (One could, for example, choose φ_1 constant.) We look for a solution u of equation (X.11) in the form of a series

$$u = (u_0 + c_0) + \lambda(u_1 + c_1) + \lambda^2(u_2 + c_2) + \cdots,$$

the u_i being functions vanishing at p_0, and the c_i constants. If one substitutes this series for u in equation (X.11), and groups the terms in like powers of λ (working purely formally), one obtains the following sequence of equalities:

$$\begin{aligned}\Delta_g u_0 &= -\varphi_0\\ \Delta_g u_1 &= \eta(u_0 + c_0) - \varphi_1\\ \Delta_g u_2 &= \eta(u_1 + c_1)\\ \Delta_g u_3 &= \eta(u_2 + c_2)\\ &\cdots\cdots\cdots\end{aligned}$$

We see at once that these yield

$$0 = \int \eta(u_0 + c_0) - \varphi_1 = \int \eta(u_1 + c_1) = \int \eta(u_2 + c_2) = \cdots .$$

The equation $\Delta_g u_0 = -\varphi_0$ is a Poisson equation in the unknown u_0. Since φ_0 is of class C^1 and vanishing integral, this equation has a unique solution u_0 of class C^2 vanishing at p_0 (for which Proposition X.2.2 affords us an integral expression). Once we have u_0, we can choose the constant c_0 so that the function $\eta(u_0+c_0)-\varphi_1$ has vanishing integral. Then, knowing the function u_0 and the constant c_0, the equation $\Delta_g u_1 = \eta(u_0 + c_0) - \varphi_1$ may be viewed as a Poisson equation in the unknown function u_1. Since $\eta(u_0 + c_0) - \varphi_1$ has class C^1 and vanishing integral, this equation has a unique solution u_1 of class C^2 (see Remark X.2.3) vanishing at p_0. We may therefore choose the constant c_1 so that the function $\eta(u_1 + c_1)$ has vanishing integral. Then, $\eta(u_1 + c_1)$ being known and of class C^1, the equation $\Delta_g u_2 = \eta(u_1+c_1)$ has a unique solution u_2 of class C^2 vanishing at p_0. Continuing in this way (inductively), we find functions u_i and constants c_i satisfying the above sequence of equalities. Moreover, the functions u_i and the constants c_i are unique, and the functions u_i are of class C^2.

Second step: the convergence of the series $(u_0+c_0)+\lambda(u_1+c_1)+\lambda^2(u_2+c_2)+\cdots$. Our aim now is to determine the values of the parameter λ for which the series $(u_0 + c_0) + \lambda(u_1 + c_1) + \lambda^2(u_2 + c_2) + \cdots$ converges. To this end, we avail ourselves of Addendum X.2.4, which furnishes us estimates of the functions u_i and the constants c_i. Note first that, the function $\eta(u_i + c_i)$ having zero integral, we have, for all $i \geq 1$,

$$|c_i| \leq \|u_i\|_\infty.$$

Next one observes that from the equation $\Delta_g u_{i+1} = \eta(u_i + c_i)$ together with Addendum X.2.4, it follows that, for $i \geq 1$,

$$\|u_{i+1}\|_\infty \leq \beta\|\eta\|_\infty\,(\|u_i\|_\infty + c_i) \leq 2\beta\|\eta\|_\infty\|u_i\|_\infty.$$

From this we infer the existence of a constant K with the property that, for all $i \geq 2$, we have the upper bound

$$\|u_i + c_i\|_\infty \leq 2\|u_i\|_\infty \leq 2K\,(2\beta\|\eta\|_\infty)^{i-1}\,.$$

From this it is immediate that the series of functions

$$(u_0 + c_0) + \lambda(u_1 + c_1) + \lambda^2(u_2 + c_2) + \cdots$$

converges normally to a function u for values of the parameter λ satisfying

$$2\beta\|\eta\|_\infty\lambda < 1.$$

It is appropriate to observe in this connection that for these values of λ, the function u is automatically of class C^1. Indeed, for $i \geq 1$, Addendum X.2.4 and the equation $\Delta_g u_{i+1} = \eta(u_i + c_i)$ together imply that

$$\left\|\vec{\nabla} u_{i+1}\right\|_\infty \leq \beta\|\eta\|_\infty \left\|\vec{\nabla} u_i\right\|_\infty .$$

Thus we conclude that the series $\vec{\nabla} u_0 + \lambda\vec{\nabla} u_1 + \lambda^2\vec{\nabla} u_2 + \cdots$ of derivatives converges normally provided $\beta\|\eta\|_\infty\lambda < 1$, and therefore that the function u is certainly of class C^1 for every λ satisfying $2\beta\|\eta\|_\infty\lambda < 1$.

Third step: verifying that the function u is indeed a solution of equation (X.11).

It remains to show that the function $u = (u_0 + c_0) + \lambda(u_1 + c_1) + \lambda^2(u_2 + c_2) + \cdots$ is a solution of the equation (X.11). To this end, we consider a function $v : S \to \mathbb{R}$ of class C^2 satisfying

$$\Delta_g v = \lambda\eta u - \varphi.$$

The existence of such a function is a consequence of Proposition X.2.2. To see this, note that the function $\lambda\eta u - \varphi$ is of class C^1 (this being the case for each of u, η and φ), and has vanishing integral since

$$\lambda \int \eta u = \sum_{i \geq 0} \lambda^{i+1} \int \eta(u_i + c_i) = \lambda \int \eta(u_0 + c_0) = \lambda \int \varphi_1 = \int \varphi.$$

We now consider, for each $n \geq 0$, functions R_n and S_n defined as follows:

$$R_n := u - \big((u_0 + c_0) + \lambda(u_1 + c_1) + \cdots + \lambda^n(u_n + c_n)\big),$$
$$S_n := v - \big((u_0 + c_0) + \lambda(u_1 + c_1) + \cdots + \lambda^{n+1}(u_{n+1} + c_{n+1})\big).$$

For each $n \geq 0$ we then have

$$u - v = R_n - S_n - \lambda^{n+1}(u_{n+1} + c_{n+1}).$$

The terms $\lambda^{n+1}(u_{n+1} + c_{n+1})$ and R_n tend uniformly to 0 as $n \to \infty$ (since the series $(u_0 + c_0) + \lambda(u_1 + c_1) + \lambda^2(u_2 + c_2) + \cdots$ converges uniformly to u). Furthermore, from the equations satisfied by v and the u_i, we see that, for all $n \geq 0$,

$$\Delta_g S_n = \lambda\eta R_n.$$

These equations, together with Addendum X.2.4, imply that, for all $n \geq 0$, one has

$$\|S_n\|_\infty \leq \lambda\beta\|\eta\|_\infty\|R_n\|_\infty,$$

whence it follows that S_n also tends uniformly to 0 as $n \to \infty$. Hence by letting n tend to infinity in the equality $u - v = R_n - S_n - \lambda^{n+1}(u_{n+1} + c_{n+1})$, we obtain the desired equality $u = v$. From this we conclude that the function u is also of class C^2, and then, since v satisfies the equation $\Delta_g v = \lambda\eta u - \varphi$, that the function u satisfies equation (X.11). Finally, by Proposition X.2.7, equation (X.11) has no other solution. □

From the equation $\Delta_g u = \lambda_0\eta u - \varphi$ to the equation $\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi$. — Having shown that one can integrate the equation $\Delta_g u = \lambda_0\eta u - \varphi$ for certain values of λ_0, Poincaré goes on to perturb this equation in order to widen the field of equations he knows how to solve:

Proposition X.2.10. — *Let $\eta : S \to \mathbb{R}$ be a positive function of class C^1, and λ_0 any positive real number. If one can find a class-C^2 solution of the equation*

$$\Delta_g u = \lambda_0\eta u - \varphi \tag{X.12}$$

for every class-C^1 function $\varphi : S \to \mathbb{R}$, then one can find a class-C^2 solution of the equation

$$\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi \tag{X.13}$$

for every class-C^1 function $\varphi : S \to \mathbb{R}$ and every positive real λ such that $\lambda < \lambda_0$.

Observe that the functions η and φ enjoy different status in the above statement: the function η is fixed once and for all, while, by contrast, the assumptions of the theorem concern crucially *all* functions φ of class C^1. Thus in order to integrate the equation $\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi_0$ for a certain function φ_0, one needs to be able to integrate the equation $\Delta_g u = \lambda_0\eta u - \varphi$ for an infinity of appropriate functions φ.

Proof of Proposition X.2.10. — The argument is very similar to that establishing Proposition X.2.9.

First step: looking for a series solution.

We seek a solution $u : S \to \mathbb{R}$ of equation (X.13) in the form

$$u = u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$$

Substituting this in equation (X.13) and grouping terms in the same powers of λ, we obtain the following sequence of equations:

$$
\begin{aligned}
\Delta_g u_0 &= \lambda_0 \eta u_0 - \varphi \\
\Delta_g u_1 &= \lambda_0 \eta u_1 + \eta u_0 \\
\Delta_g u_2 &= \lambda_0 \eta u_2 + \eta u_1 \\
&\cdots\cdots\cdots
\end{aligned}
$$

Here the first equation is of type (X.12), and by assumption this can be integrated. Once we have the unique solution u_0 of the first equation, the second becomes an equation of type (X.12) in the unknown function u_1 (with $\varphi = -\eta u_0$), and once again we can, by hypothesis, integrate it. Once we have the unique solution u_1 of that second equation, the third equation becomes an equation of type (X.12) in the unknown function u_2, which again, by hypothesis, we can integrate. Continuing in this way (inductively), we find successively the solutions $u_0, u_1, u_2, \ldots$ (of class C^2 and unique) of all of the equations of the above sequence.

Second step: the convergence of the series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$.

By virtue of the *a priori* upper bound obtained in Proposition X.2.8, we have, for all $i \geq 0$,

$$\|u_{i+1}\|_\infty \leq \left\| \frac{\eta u_i}{\lambda_0 \eta} \right\|_\infty \leq \frac{1}{\lambda_0} \|u_i\|_\infty.$$

Hence there exists a constant K such that $\|u_i\| \leq K\lambda_0^{-i}$ for all $i \geq 0$, so that the series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ of functions converges normally to a function u (*a priori* only continuous) for all $\lambda < \lambda_0$.

Third step: showing that the function u is a solution of equation (X.13).

Thus we need to show that u has class C^2 and satisfies (X.13).

To this end, Poincaré considers the function $(p, q) \mapsto G_{p_0,q_0}(p, q)$ introduced in the proof of Proposition X.2.2, in terms of which he defines a sequence of functions $v_0, v_1, v_2, \ldots : S \to \mathbb{R}$ by setting, for all $p \in S \setminus \{q_0\}$,

$$
\begin{aligned}
v_0(p) &:= \int_S G_{p_0,q_0}(p, q)(\lambda_0 \eta u_0 - \varphi)(q) dv_g(q) \\
v_1(p) &:= \int_S G_{p_0,q_0}(p, q)(\lambda_0 \eta u_1 + \eta u_0)(q) dv_g(q) \\
v_2(p) &:= \int_S G_{p_0,q_0}(p, q)(\lambda_0 \eta u_2 + \eta u_1)(q) dv_g(q) \\
&\cdots\cdots\cdots
\end{aligned}
$$

As in the proof of Proposition X.2.2 one shows that the functions $v_0, v_1, v_2, \dots$ extend to q_0 and are of class C^2. Since the function series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ converges normally to the function u, the function series $(\lambda_0 \eta u_0 - \varphi) + \lambda(\lambda_0 \eta u_1 + \eta u_0) + \lambda^2(\lambda_0 \eta u_2 + \eta u_1) + \cdots$ converges normally to the continuous function $(\lambda_0 + \lambda)\eta u - \varphi$. From this fact it follows that the function series $v_0 + \lambda v_1 + \lambda^2 v_2 + \cdots$ converges normally to the function

$$\begin{aligned} v &= \int_S G_{p_0,q_0}(p,q)\,((\lambda_0 \eta u_0 - \varphi) + \lambda(\lambda_0 \eta u_1 + \eta u_0) + \cdots)\,(q)\; dv_g(q) \\ &= \int_S G_{p_0,q_0}(p,q)\,((\lambda_0 + \lambda)\eta u - \varphi)\,(q)\; dv_g(q). \end{aligned}$$

Hence by Remark X.2.3 the function v is of class C^1.

Since the functions $\varphi, u_0, u_1, \dots$ are of class C^1, it follows from the equations defining the functions $v_0, v_1, \dots$ that

$$\begin{aligned} \Delta_g v_0 &= \lambda_0 \eta u_0 - \varphi &&= \Delta_g u_0 \\ \Delta_g v_1 &= \lambda_0 \eta u_1 + \eta u_0 &&= \Delta_g u_1 \\ \Delta_g v_2 &= \lambda_0 \eta u_2 + \eta u_1 &&= \Delta_g u_2 \\ &\cdots\cdots\cdots \end{aligned}$$

Hence for each $i \geq 0$, the functions u_i and v_i differ by an additive constant: there exist $c_i \in \mathbb{R}$ such that $v_i = u_i + c_i$. Since the series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ and $v_0 + \lambda v_1 + \lambda^2 v_2 + \cdots$ converge normally, the series $c_0 + \lambda c_1 + \lambda^2 c_2 + \cdots$ converges absolutely. Hence the functions u and v differ by a constant: we have $v = u + c$ where $c = c_0 + \lambda c_1 + \lambda^2 c_2 + \cdots$.

Since v is of class C^1, and v differs from u by a constant, the function u is of course also of class C^1, and then this holds also for the function $\eta u - \varphi$. It follows that the function

$$v = \int_S G_{p_0,q_0}(p,q)((\lambda_0 + \lambda)\eta u - \varphi)(q)\; dv_g(q)$$

is actually of class C^2 (see Remark X.2.3), and satisfies

$$\Delta_g v = (\lambda_0 + \lambda)\eta u - \varphi.$$

Then once again in view of the fact that u and v differ by a constant, we conclude that the function u is also of class C^2, and

$$\Delta_g u = \Delta_g v.$$

From the preceding two equations we infer, finally, that u satisfies the equation $\Delta_g u = (\lambda_0 + \lambda)\eta u - \varphi$. □

Solution of the equation $\Delta_g u = \eta u - \varphi$. — It now only remains to harvest the fruits of the above labors:

Proposition X.2.11. — *For all positive class-C^1 functions $\eta : S \to \mathbb{R}$ and arbitrary class-C^1 functions $\varphi : S \to \mathbb{R}$, one can find a solution of class C^2 of the equation*

$$\Delta_g u = \eta u - \varphi.$$

Proof of Proposition X.2.11. — Let $\eta : S \to \mathbb{R}$ be any fixed positive class-C^1 function. By Proposition X.2.9, we can find a solution of class C^2 of the equation

$$\Delta_g u = \lambda \eta u - \varphi \tag{X.14}$$

for every function $\varphi : S \to \mathbb{R}$ of class C^1 and every real positive number $\lambda < (2\beta\|\eta\|_\infty)^{-1}$. From Proposition X.2.10 we then infer that we can find a solution of class C^2 of the equation (X.14) for every class-C^1 function φ and every $\lambda < 2(2\beta\|\eta\|_\infty)^{-1}$. Then by means of a second application of Proposition X.2.10, we infer in turn that we can find a class-C^2 solution of equation (X.14) for every class-C^1 function φ and every $\lambda < 3(2\beta\|\eta\|_\infty)^{-1}$. By induction we conclude that we can find a class-C^2 solution of equation (X.14) for every class-C^1 function φ and all $\lambda > 0$. To get the desired conclusion we now take $\lambda = 1$. □

X.2.3. The solution of the equation $\Delta_g u = \theta e^u - \varphi$

The third major stage in Poincaré's strategy is concerned with the equation

$$\Delta_g u = \theta e^u - \varphi, \tag{X.15}$$

where $\theta : S \to \mathbb{R}$ is a given positive function of class C^1 and $\varphi : S \to \mathbb{R}$ is a given function of class C^1 and positive integral. Poincaré notes first of all that this equation has an obvious solution when the functions φ and θ are proportional. He then goes on to show — using the technique of series developments that the reader must by now be accustomed to — that if one can integrate the equation $\Delta_g u = \theta e^u - \varphi_0$ for some function φ_0, then one can also integrate the equation $\Delta_g u = 2e^u - (\varphi_0 + \lambda\psi)$ provided λ is sufficiently small.

The uniqueness of the solution. — The same sort of reasoning as in the proof of Proposition X.2.7 shows that equation (X.15) has at most one solution:

Proposition X.2.12. — *Given a positive function $\theta : S \to \mathbb{R}$ and any function $\varphi : S \to \mathbb{R}$, the equation*

$$\Delta_g u = \theta e^u - \varphi$$

has at most one solution $u : S \to \mathbb{R}$.

Proof. — Let $u, v : S \to \mathbb{R}$ be two solutions of $\Delta_g u = \theta e^u - \varphi$. It then follows that

$$\Delta_g(u - v) = \theta e^v (e^{(u-v)} - 1).$$

Since S is compact, there are points $p_-, p_+ \in S$ where the function $u - v$ attains its least and greatest values respectively. By Fact X.2.6, we then have

$$\Delta_g(u - v)(p_-) \geq 0 \quad \text{and} \quad \Delta_g(u - v)(p_+) \leq 0.$$

The equation u and v satisfy and the positivity of the function θ, then together yield the inequalities

$$e^{(u-v)}(p_-) \geq 1 \quad \text{and} \quad e^{(u-v)}(p_+) \leq 1,$$

which in turn imply that the function $u - v$ is identically zero. □

The case where the functions φ and θ are proportional. — The following proposition, though trivial, is fundamental to Poincaré's strategy:

Proposition X.2.13. — *For every positive function $\theta : S \to \mathbb{R}$ and positive real number α, the equation*

$$\Delta_g u = \theta e^u - \alpha\theta$$

has a constant solution.

Proof. — The constant function $u = \log \alpha$ is a solution of the equation, and, by Proposition X.2.12, it is the only solution. □

Passing from the equation $\Delta_g u = \theta e^u - \varphi_0$ to the equation $\Delta_g u = \theta e^u - (\varphi_0 + \lambda\psi)$. — Poincaré establishes a final result of the type "if one can integrate such and such an equation then one can also integrate …".

Proposition X.2.14. — *Let $\theta : S \to \mathbb{R}$ and $\varphi_0 : S \to \mathbb{R}$ be positive class-C^1 functions and $\psi : S \to \mathbb{R}$ any class-C^1 function. If one can find a solution of class C^2 of the equation*

$$\Delta_g u = \theta e^u - \varphi_0, \tag{X.16}$$

then one can also find a solution of class C^2 ofă

$$\Delta_g u = \theta e^u - (\varphi_0 + \lambda\psi) \tag{X.17}$$

for all positive real λ satisfying

$$\lambda \left\| \frac{\psi}{\theta} \right\|_\infty \left\| \frac{\theta}{\varphi_0} \right\|_\infty < 2 \log 2 - 1.$$

Proof. — The argument is similar to that of Propositions X.2.9 and X.2.10; any differences are due simply to certain technical complications arising from the non-linear nature of equation (X.17).

First step: the search for a series solution.
We look for a possible solution u of equation (X.17) in the form of a series

$$u = u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots .$$

Assuming u of that form, it follows that the function e^u has the form

$$e^u = e^{u_0}(1 + \lambda u_1 + \lambda^2(u_2 + w_2) + \lambda^3(u_3 + w_3) + \lambda^4(u_4 + w_4) + \cdots),$$

where $w_2 = \frac{u_1^2}{2}$, $w_3 = \frac{u_1^3}{6} + u_1 u_2$, $w_4 = \frac{u_1^4}{24} + u_1 u_3 + \frac{u_2^2}{2} + \frac{u_1^2 u_2}{2}$, and, more generally,

$$w_i = P_i(u_1, u_2, \ldots, u_{i-1}),$$

where P_i is a polynomial in $i - 1$ variables and *with positive coefficients*. Substituting these series for u and e^u in equation (X.17) and grouping terms in like powers of λ, we obtain the following sequence of equations:

$$\begin{aligned}
\Delta_g u_0 &= \theta e^{u_0} - \varphi_0 \\
\Delta_g u_1 &= \theta e^{u_0} u_1 - \psi \\
\Delta_g u_2 &= \theta e^{u_0}(u_2 + w_2) \\
\Delta_g u_3 &= \theta e^{u_0}(u_3 + w_3) \\
&\cdots\cdots\cdots
\end{aligned}$$

By assumption we can integrate the first equation in this list; thus we may suppose we have a function $u_0 : S \to \mathbb{R}$ satisfying that equation. Then, u_0 being known, the second equation becomes one of type (X.10) (in the unknown function u_1). By Proposition X.2.11, we can integrate this equation, obtaining thereby the function u_1, and thence $w_2 = P_2(u_1)$. Then, u_0 and w_2 being known, the third equation becomes one of type (X.10) (in the unknown function u_2), and, again by Proposition X.2.11 we can integrate that equation to obtain u_2. The functions u_1 and u_2 being known, we can then calculate $w_3 = P_3(u_1, u_2)$. Continuing in this way (inductively) we can integrate all equations in the above list, finding one after the other the functions $u_0, u_1, u_2, \ldots$.

Second step: the convergence of the series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$.
We now need to determine values (if any) of the parameter λ for which the series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ converges.

We begin by finding an upper bound for the uniform norm of u_1. The equation $\Delta_g u_1 = \theta e^{u_0} u_1 - \psi$ is an equation (in the unknown u_1) of type (X.10); hence by

Proposition X.2.8 we have

$$\|u_1\|_\infty \le \left\| \frac{\psi}{\theta e^{u_0}} \right\|_\infty \le \left\| \frac{\psi}{\theta} \right\|_\infty \left\| \frac{1}{e^{u_0}} \right\|_\infty .$$

Our problem is thus reduced to finding an upper bound for the uniform norm of the function $1/e^{u_0}$. Let p be a point of S where u_0 attains its least value. Then the function $1/e^{u_0}$ attains its greatest value at p. We also have $\Delta_g u_0(p) \ge 0$, whence $(\theta e^{u_0} - \varphi)(p) \ge 0$. Hence

$$\left\| \frac{1}{e^{u_0}} \right\|_\infty \le \frac{1}{e^{u_0(p)}} \le \frac{\varphi(p)}{\theta(p)} \le \left\| \frac{\varphi_0}{\theta} \right\|_\infty .$$

From this and the earlier inequality we obtain the following upper bound for the norm of u_1:

$$\|u_1\|_\infty \le \left\| \frac{\psi}{\theta} \right\|_\infty \left\| \frac{\varphi_0}{\theta} \right\|_\infty . \tag{X.18}$$

We now seek to bound the uniform norm of u_i, for $i \ge 2$. The equation $\Delta_g u_i = \theta e^{u_0}(u_i + w_i)$ is of type (X.10), so that, by Proposition X.2.8, we have, for all $i \ge 2$,

$$\|u_i\|_\infty \le \left\| \frac{\theta e^{u_0}.w_i}{\theta e^{u_0}} \right\|_\infty = \|w_i\|_\infty = \|P_i(u_1, \ldots, u_{i-1})\|_\infty .$$

From this and the fact that the coefficients of the polynomial P_i are all positive, we obtain, for all $i \ge 2$,

$$\|u_i\|_\infty \le P_i\left(\|u_1\|_\infty, \ldots, \|u_{i-1}\|_\infty\right) . \tag{X.19}$$

Consider now the sequence of positive real numbers $(a_i)_{i \ge 1}$ defined inductively as follows:

$$\begin{cases} a_1 & := \left\| \frac{\psi}{\theta} \right\|_\infty \left\| \frac{\varphi_0}{\theta} \right\|_\infty \\ a_{i+1} & := P_i(a_1, \ldots, a_i) \quad \text{for all } i \ge 1. \end{cases}$$

From the inequalities (X.18) and (X.19) above it is immediate that $\|u_i\|_\infty \le a_i$ for all $i \ge 1$. Hence in order to find values of λ for which the function series

$u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ converges normally, it suffices to examine the convergence of the series with real positive coefficients given by $\lambda a_1 + \lambda^2 a_2 + \lambda^3 a_3 + \cdots$.

Thus our task is now to determine values (if any) of the parameter λ for which the series $\lambda a_1 + \lambda^2 a_2 + \lambda^3 a_3 + \cdots$ converges. Supposing it converges, let a be its sum. We shall then have

$$\begin{aligned} e^a &= 1 + \lambda a_1 + \lambda^2(a_2 + P_2(a_1)) + \lambda^3(a_3 + P_3(a_1, a_2)) + \cdots \\ &= 1 + \lambda a_1 + 2(\lambda^2 a_2 + \lambda^3 a_3 + \cdots) \\ &= 2a + 1 - \lambda a_1. \end{aligned}$$

Hence, in particular, the equation $e^x = 2x + 1 - \lambda a_1$ will have a solution. By examining the function $x \mapsto e^x = 2x + 1 - \lambda a_1$, one sees that for this to be possible, it is necessary that $\lambda a_1 \leq 2 \log 2$. Conversely, if $0 < \lambda a_1 \leq 2 \log 2$ then the equation $e^x = 2x + 1 - \lambda a_1$ has two solutions (both positive); let a be the smaller of these. For $N \geq 1$, write

$$A_N := \lambda a_1 + \lambda^2 a_2 + \cdots + \lambda^N a_N,$$

and consider the function $f : x \mapsto e^x - x - 1 + \lambda a_1$. By developing e^{A_N} as a series in powers of λ, we see that, for all $N \geq 1$,

$$\begin{aligned} f(A_N) &= A_{N+1} + \lambda^{N+2} P_{N+2}(a_1, \ldots, a_N, 0) + \cdots \\ &\geq A_{N+1}. \end{aligned}$$

We also have, since f is increasing on $[0, +\infty)$, that $f(x) < a$ for all $x \in (0, a)$. It follows that for all $N \geq 1$, A_N is bounded above by a, so that the series $\lambda a_1 + \lambda^2 a_2 + \lambda^3 a_3 + \cdots$ converges. To summarize, we have shown that the function series $u_0 + \lambda u_1 + \lambda^2 u_2 + \cdots$ converges normally to a function u (*a priori* only continuous) provided that $\lambda a_1 \leq 2 \log 2$, that is, provided that

$$\lambda \left\| \frac{\psi}{\theta} \right\|_\infty \left\| \frac{\varphi_0}{\theta} \right\|_\infty < 2 \log 2 - 1.$$

Third step: showing that the function u is a solution of equation (X.17).

As in the proof of Proposition X.2.10, one defines functions $v_0, v_1, v_2, \cdots : S \to \mathbb{R}$

by setting, for all $p \in S \setminus \{q_0\}$:

$$v_0(p) := \int_S G_{p_0,q_0}(p,q)(\theta e^{u_0} - \varphi)(q)dv_g(q)$$

$$v_1(p) := \int_S G_{p_0,q_0}(p,q)(\theta e^{u_0}.u_1 - \psi)(q)dv_g(q)$$

$$v_2(p) := \int_S G_{p_0,q_0}(p,q)(\theta e^{u_0}(u_2 + w_2))(q)dv_g(q)$$

$$v_3(p) := \int_S G_{p_0,q_0}(p,q)(\theta e^{u_0}(u_3 + w_3))(q)dv_g(q)$$

$$\cdots\cdots\cdots$$

where $(p,q) \mapsto G_{p_0,q_0}(p,q)$ is the function used in the proof of Proposition X.2.2. By means of the same arguments as were employed in the proof of Proposition X.2.10, one infers that the functions $v_0, v_1, v_2 \ldots$ then extend to q_0, are of class C^2, and that the function series $v_0 + \lambda v_1 + \lambda^2 v_2 + \cdots$ converges normally to

$$v = \int_S G_{p_0,q_0}(p,q)(\theta e^{u_0}(1 + \lambda u_1 + \cdots) - \varphi_0 - \lambda\psi)(q)dv_g(q)$$

$$= \int_S G_{p_0,q_0}(p,q)(\theta e^{u} - \varphi_0 - \lambda\psi)(q)dv_g(q).$$

By Remark X.2.3, the function v is of class C^1, and one then shows as in the proof of Proposition X.2.10 that the functions u and v differ by an additive constant c, so that u also is of class C^1, whence in turn the function v is of class C^2 and satisfies

$$\Delta_g v = \theta e^u - \varphi_0 - \lambda\psi.$$

Thence, finally, it follows that u is of class C^2, and satisfies $\Delta_g u = \Delta_g v$, whence u satisfies equation (X.17). □

The case where φ is positive. — By means of Propositions X.2.13 and X.2.14 we can now show how to integrate the equation $\Delta_g u = \theta e^u - \varphi$ in the case where the function φ is everywhere positive.

Proposition X.2.15. — *For all positive class-C^1 functions $\theta : S \to \mathbb{R}$ and $\varphi : S \to \mathbb{R}$, one can find a solution of class C^2 of the equation*

$$\Delta_g u = \theta e^u - \varphi.$$

Proof. — We seek to reduce the situation to that where Propositions X.2.13 and X.2.14 apply. To that end, we write the function φ in the form $\varphi = \alpha\theta + \psi$ where α is a positive real number and $\psi : S \to \mathbb{R}$ is a positive function. (One could, for example, take $\alpha < \min(\varphi/\theta)$.) We shall show that one can integrate the equation

$$\Delta_g u = \theta e^u - (\alpha\theta + \lambda\psi) \tag{X.20}$$

for all $\lambda > 0$. Set

$$\lambda_0 = (2\log 2 - 1)\alpha \left\| \frac{\psi}{\theta} \right\|_\infty^{-1}.$$

By Proposition X.2.13, since the constant α is positive, one can find a solution of class C^2 of equation (X.20) for $\lambda = 0$. Then from Proposition X.2.14 (with $\varphi_0 = \alpha\theta$), we infer that one can find a solution of class C^2 of equation (X.20) for $0 < \lambda < \lambda_0$. Applying yet again Proposition X.2.14 (this time with $\varphi_0 = \alpha\theta + \lambda_0\psi$), we infer in turn that one can find a solution of class C^2 of equation (X.20) for $0 < \lambda < 2\lambda_0$. By iterating this argument, we conclude that one can find a solution of class C^2 of equation (X.20) for $0 < \lambda < n\lambda_0$ for every $n \in \mathbb{N}$. To conclude, it only remains to observe that equation (X.20) reduces to the equation $\Delta_g u = \theta e^u - \varphi$ when $\lambda = 1$. □

The solution of the equation $\Delta_g u = 2e^u - \varphi$. — We finally arrive at the desired result:

Theorem X.2.16. — *For every positive class-C^1 function $\theta : S \to \mathbb{R}$ and class-C^1 function $\varphi : S \to \mathbb{R}$ with positive integral over S, one can find a solution of class C^2 of the equation*

$$\Delta_g u = \theta e^u - \varphi.$$

Proof. — Let $\varphi : S \to \mathbb{R}$ be a function of class C^1 and positive integral. Denote that integral by c and write $\varphi_0 = \varphi - c$. Since φ_0 is a function of class C^1 and vanishing integral, by Proposition X.2.2 we can find a unique function $v : S \to \mathbb{R}$ of class C^2 satisfying

$$\Delta_g v = -\varphi_0.$$

Since the function θe^v is then of class C^1 and positive, and also c is positive, by Proposition X.2.15 one can find a function $w : S \to \mathbb{R}$ of class C^2 satisfying

$$\Delta_g w = \theta e^v e^w - c.$$

We now have immediately that the function $u = v + w$ satisfies the equation $\Delta_g u = \theta e^u - \varphi$. □

X.3. Conclusion: uniformization of algebraic Riemann surfaces, prescription of the curvature, and the calculus of variations

Putting Theorem X.2.16 and Proposition X.1.3 together, we obtain the following uniformization theorem:

Theorem X.3.1. — *The universal cover of every compact algebraic Riemann surface of negative Euler characteristic is biholomorphic to the disc.*

This was Poincaré's second proof of this theorem — and, since the first was far from satisfactory by the criteria of rigor even of that time, we can hardly reproach him for reproving it.

And all the more since, while in the above proof the uniformization of surfaces depends on the existence of a solution of the equation $\Delta_g u = 2e^u - \varphi$, Poincaré actually establishes the existence and uniqueness of a solution of the somewhat more general equation $\Delta_g u = \theta e^u - \varphi$, where θ is any positive function. This equation can be interpreted naturally in terms of the existence of metrics of prescribed curvature. To see this, consider a compact surface S of negative Euler characteristic, endowed with a Riemannian metric g. The metric g induces the structure of a Riemann surface on S and this Riemann surface is automatically algebraic in view of its compactness. Write $-\varphi_g$ for the curvature with respect to g. The formula given in §X.1.3 then generalizes as follows: for every function $u : S \to \mathbb{R}$ (of class C^2), the curvature of the metric $g' = e^u g$ is equal to $-\varphi_{g'}$ where $\Delta_g u = 2\varphi_{g'} e^u - 2\varphi_g$. Theorem X.2.16 shows that the equation

$$\Delta_g u = \theta e^u - \varphi_g \tag{X.21}$$

has a solution for every positive function $\theta : S \to \mathbb{R}$ (of class C^2). Thus Poincaré's work implies the following result[11]:

Theorem X.3.2. — *Let S be a compact surface of negative Euler characteristic, g a Riemannian metric on S, and $\theta : S \to \mathbb{R}$ a positive function of class C^1. Then there exists a Riemannian metric in the conformal class of g relative to which the Gaussian curvature is equal to $-\theta$.*

In the box below we sketch a "modern" proof of Theorem X.2.16 — and thus also of Theorem X.3.2 — using the concepts of *sub-solution* and *super-solution*. This proof, suggested to us by H. Brezis, is perfectly elementary, and seems to us very close in spirit to Poincaré's.

[11]Note however, that the interpretation of equation (X.21) in terms of curvature is absent from Poincaré's memoir.

Box X.1: The method of sub-and super-solutions.

Let S, g, θ and φ be as in the statement of Theorem X.2.16. Consider a partial differential equation of the form

$$\Delta_g u = F(x,u), \tag{X.22}$$

where $F : S \times \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function[a].
A *sub-solution* of equation (X.22) is a function $u_- : S \to \mathbb{R}$ of class C^2, such that

$$\Delta_g u_- \leq F(x,u_-),$$

and a *super-solution* is a function $u_+ : S \to \mathbb{R}$ of class C^2, such that

$$\Delta_g u_+ \geq F(x,u_+).$$

The following result demonstrates the significance of these concepts:

Theorem X.3.3. — *If equation (X.22) possesses a sub-solution u_- and a super-solution u_+ with $u_- \leq u_+$, then equation (X.22) has a solution u satisfying $u_- \leq u \leq u_+$.*

Theorem X.2.16 follows easily from this result: the equation

$$\Delta_g u = \theta e^u - \varphi$$

certainly has the form of (X.22), and for sufficiently large C the constant functions $u_- = -C$ and $u_+ = C$ will be respectively a sub-solution and super-solution of that equation.

The proof of Theorem X.3.3 is based on the construction of a sequence of functions very similar to those introduced by Poincaré in his memoir. One begins by choosing a constant ρ sufficiently large for the function $F(x,s) + \rho.s$ to be increasing with respect to s for s in the interval $[\min_{x \in S} u_-(x), \max u_+(x)]$ (and for any x); such a constant ρ must exist in view of the continuous differentiability of F and the compactness of the surface S. Now equation (X.22) is certainly equivalent to the following one:

$$\Delta_g u + \rho u = F(x,u) + \rho u. \tag{X.23}$$

[a] What follows remains valid with Δ_g replaced by a linear elliptic differential operator of order two with arbitrary continuous coefficients.

In order to solve the latter, one constructs inductively a sequence $(u_k)_{k\geq 0}$ of functions as follows: One first sets $u_0 = u_-$, then, assuming u_k known, defines u_{k+1} as the unique solution of the Poisson equation

$$\Delta_g u_{k+1} + \rho u_{k+1} = F(x, u_k) + \rho u_k. \tag{X.24}$$

From the fact that the function $s \mapsto F(x,s) + \rho s$ is increasing, together with the maximum principle, one easily infers that the sequence $(u_k)_{k\geq 0}$ is increasing and bounded above by u_+. The fact that the u_k are bounded above by u_+ allows one to find, via equation (X.24), an upper bound for the second derivative of u_{k+1} independent of k. Hence the sequence $(u_k)_{k\geq 0}$ converges uniformly to a function u of class C^2, and one then verifies easily that u is a solution of equation (X.23). (For more details on this theme we refer the reader to [Ku1959].)

Another proof of Theorem X.3.2, more involved than the one presented in Box X.1, was given by M. S. Berger in [Berg1971]. The idea behind that proof was apparent already in Poincaré's work. Indeed, in the middle of his memoir, Poincaré makes a pause:

> However, before proving by rigorous means the integrability of this equation [the equation $\Delta_g u = \theta . e^u - \varphi$], I would first like to present it in terms of an insight based on the calculus of variations which is sometimes employed in mathematical physics.

Poincaré then constructs a functional whose critical points of class C^2 are solutions of the equation $\Delta_g u = \theta . e^u - \varphi$. He goes on to show that, provided θ is positive and the integral of φ is positive, the functional so constructed is bounded below. The above quote shows that he understood perfectly well that this does not *a priori* imply the existence of a smooth function u at which the functional attains its minimum. It is, roughly speaking, the existence of this minimum that M. Berger[12] establishes in [Berg1971]. We sketch Berger's proof in the following box.

Box X.2: Variational solution of the equation $\Delta_g u = \theta e^u - \varphi_g$

Let S, g, θ be as in the statement of Theorem X.3.2, and $\varphi_g : S \to \mathbb{R}$ the negative of the Gaussian curvature with respect to g.

[12] In fact the functional considered by Berger is slightly different from that considered by Poincaré.

Consider the functional

$$F : u \mapsto \int_S \left(\frac{1}{2}|du|^2 - \varphi_g u\right) dv_g$$

on the submanifold V of the Sobolev space $H^1(S)$, defined by the Gauss–Bonnet constraint:

$$V = \left\{u \in H^1(S) \mid \int_S \theta e^u \, dv_g = -2\pi\chi(S)\right\}.$$

Since S is compact, we have in dimension 2 the Sobolev inclusions $H^1(S) \hookrightarrow L^p(S)$ for all $p < \infty$ (but not at $p = \infty$). These inclusions are moreover *compact*. A further ingredient in the proof is the following inequality, a consequence of the *Trudinger inequalities*:

$$\int_S e^{|u|} \, dv_g \leq C \exp\left(C\|u\|^2_{H^1}\right).$$

This implies that the functional F is of class C^1 on $H^1(S)$, and that V is indeed a submanifold of $H^1(S)$. Inequalities of the same type are then used to show that $u \mapsto e^u$ defines a continuous map from $H^1(S)$ endowed with the weak topology, to L^1, so that V is closed *in the weak topology* in $H^1(S)$.

An easy calculation shows that the critical points of the functional F on S are solutions of the equation in question, that is, of $\Delta_g u = \theta e^u - \varphi_g$.

Since the functional F is convex and continuous (with respect to the strong topology), it is lower semi-continuous in the weak topology (Hahn-Banach). Furthermore, it is bounded below on V, which fact follows by means of (among other things) *Poincaré's inequality* on S:

$$\int_S |u_0|^2 \, dv_g \leq C \int_S |du_0|^2 \, dv_g, \quad \text{if } u_0 \in H^1(S) \text{ and } \int_S u_0 \, dv_g = 0,$$

together with the assumption $\theta > 0$.

From this point the strategy is the standard one of the calculus of variations: one considers a minimizing sequence (u_n) in V, that is, such that $F(u_n) \to \inf_V F$. It follows that (u_n) is bounded in $H^1(S)$, whence, modulo choosing a subsequence, that it converges *weakly* in $H^1(S)$, to a limit u_∞. Since V is closed, this limit belongs to V and $F(u_\infty) = \inf_V F$ by weak semi-continuity. Classical arguments concerning elliptic regularity then show that this is a smooth solution of the problem.

Berger's article was followed by a series of papers on the existence of metrics of prescribed curvature on non-compact surfaces (see especially [KaWa1974] and, for a more complete bibliography, the survey paper [HuTr1992]), and then on the existence of metrics of constant scalar curvature on manifolds of dimension at least 3 (the so-called *Yamabe problem*; see for example [Aub1998]). Note that in his memoir, Poincaré also deals with the case of compact surfaces with a finite number of points removed. Alas, we have not succeeded in the present context in interpreting Poincaré's results in terms of the existence of metrics of prescribed curvature and with prescribed behavior near the removed points... but we cannot exclude the possibility that a closer reading of his memoir would yield further beautiful surprises!

Part C

Towards the general uniformization theorem

The last part of this book is devoted to describing the path that, from 1882 to 1907, led from the uniformization of algebraic Riemann surfaces by the method of continuity to the general uniformization theorem as we know it today. Gray has written a very detailed study [Gra1994] devoted to the Riemann Mapping Theorem [Gra1994] to which we may refer the reader. We recommend also earlier entries in the *Encyklopädie der mathematischen Wissenschaften*: [OsgW1901, Bie19210].

In 1882, Klein and Poincaré became convinced that every algebraic Riemann surface could be uniformized by the sphere, the plane, or the unit disc. Although some of the details of the proof of this marvellous result remained to be filled in, Poincaré, never lacking in mathematical audacity, was already launched on the conquest of much wider territory, attempting to uniformize Riemann surfaces associated with arbitrary, so not necessarily algebraic, *germs of analytic functions*.

The memoir [Poin1883b] Poincaré published in 1883 begins with a statement of *the theorem of uniformization of functions* that he proposes to prove:

> Let y be any analytic function of x, not single-valued. One can always find a variable z such that x and y are single-valued functions of z.

What is the missing link between this statement and what we call today the uniformization theorem for Riemann surfaces? In his memoir, Poincaré recalls how to construct from a "non-single-valued analytic function y of the variable x" an abstract Riemann surface extended over the plane of the variable x, on which y is naturally defined as a single-valued analytic function. In modern terminology, given a germ of an analytic function y of a variable x, one constructs the maximal Riemann surface on which one can extend the germ y to a (single-valued) analytic function (see Box II.1): this is the Riemann surface associated with the germ y. Finding a variable z such that x and y are single-valued functions of z comes down to uniformizing the Riemann surface associated with the germ y, that is, to parametrizing this surface with a single complex variable z. In 1883, Poincaré did not succeed in obtaining a parametrization that is a local biholomorphism at every point, and was forced to allow for branch points. His precise result was as follows:

Theorem. — *Let S a Riemann surface admitting a non-constant meromorphic function. Then there exists a branched covering map $\pi : U \to S$, where U is a bounded open subset of $\mathbb{C}$.*

The uniformization theorem for functions announced by Poincaré follows immediately from this result: if S is the Riemann surface associated with a germ of an analytic function y of a complex variable x, and if U is the open set in $\mathbb{C}$ given by the above theorem, then x and y may be viewed as single-valued functions on the surface S, and therefore as single-valued functions of the coordinate z of the complex plane containing U.

The concept of the universal cover of a Riemann surface plays an important role in Poincaré's memoir. As far as we know, it is in this memoir that there appears for the very first time a definition of the universal cover of the Riemann surface associated with a germ of a function (or with a finite family of germs of functions; see Box XI.2 below). In 1898, Osgood reckoned this definition a crucial feature (and perhaps the most important contribution) of Poincaré's memoir ([OsgW1898]). To establish the above theorem, Poincaré shows the existence of a Riemann surface Σ that is a branched covering space of S and is such that its universal covering space $\widetilde{\Sigma}$ is biholomorphic to a bounded open subset of $\mathbb{C}$. To achieve this, it suffices — as Riemann had observed — to find a Riemann surface Σ that is a branched covering of S such that $\widetilde{\Sigma}$ admits a positive harmonic function with a logarithmic pole.

The basic tool in Poincaré's proof is the following result, which he attributes to Schwarz, and which does indeed follow immediately from techniques invented by the latter in [Schw1870a] (even if it would seem that Schwarz himself was unaware in 1870 of having effectively established such a general result):

Theorem. — *Let Ω be a region of compact closure, with analytic or polygonal boundary, of a Riemann surface. Then Ω admits a Green's function*[13]*. It follows that if Ω is simply connected, then it is biholomorphic to the unit disc in $\mathbb{C}$.*

Poincaré considers an exhaustion of a simply connected Riemann surface $\widetilde{\Sigma}$ by means of simply connected regions with compact closure (however without justifying its existence); he applies Schwarz's theorem to each of these regions, obtaining thereby a sequence of Green's functions; if this sequence converges, then the limit will automatically be a positive harmonic function defined on $\widetilde{\Sigma}$ with a logarithmic pole, and $\widetilde{\Sigma}$ will therefore be biholomorphic to an open subset of the unit disc. However, in general one does not obtain a convergent sequence of Green's functions, and this is why, instead of considering the universal cover $\widetilde{S}$ of the Riemann surface S of interest, Poincaré has to resort to the universal cover $\widetilde{\Sigma}$ of a branched covering space Σ of S.

The result Poincaré obtained in 1883 represents an exceptional advance from the point of view of analytic functions, but is much less satisfactory if one is interested in Riemann surfaces for their own sake, and not merely as a simple tool to be used to investigate analytic functions.

Recall that Klein and Poincaré had shown (or at least believed they had shown) that the universal cover $\widetilde{S}$ of an algebraic Riemann surface S is always biholo-

[13]Recall that a Green's function on Ω is a positive, harmonic function with a logarithmic pole, that tends to zero in the neighborhood of the boundary of Ω.

morphic to the sphere, complex plane, or unit disc, and that therefore S can be identified with the quotient of one of these surfaces by the action of a group of automorphisms. In the case where S is not algebraic, Poincaré managed "only" to prove in 1883 that S has a branched covering space $\widetilde{\Sigma}$ biholomorphic to a bounded simply connected region U of $\mathbb{C}$. The primary drawback in this result consists in the fact that one has no control over the region U, which *a priori* depends on the surface S. (Note that at that time the Riemann Mapping Theorem had been proved rigorously only in special cases.) And even if one knew how to identify the region U, the presence of branch points makes for a considerably weaker result: indeed, for a fixed Riemann surface S and region U of $\mathbb{C}$, there exist in general infinitely many branched coverings $\pi : U \to S$ not obtained from one another by composing with biholomorphisms of U. And lastly, one is hard put to content oneself with Poincaré's result when one reflects that it yields a "uniformization" of the complex plane by means of an open subset of the unit disc![14]

In his address to the International Congress of Mathematicians in 1900 [Hil1900b], Hilbert praises Poincaré's work on algebraic Riemann surfaces and also his uniformization theorem for analytic functions, but also emphasizes the imperfections of the latter result. In view of the importance of the question, he reckons it essential to try to obtain a result for general Riemann surfaces as satisfying as that obtained by Klein and Poincaré for algebraic surfaces. This constitutes his 22nd problem.

An initial advance was made on the problem in 1900 by W. Osgood, in proving the following result:

Theorem. — *Every simply connected region of the complex plane that admits a positive harmonic function with a logarithmic pole (for example, every bounded simply connected region) is biholomorphic to the unit disc.*

Thus at this stage it was known that every Riemann surface has a branched covering biholomorphic to the unit disc in $\mathbb{C}$. It took another seven years before the uniformization theorem as we know it today was proved ...

Over the first several years of the 20th century there were various unsuccessful attempts to solve Hilbert's 22nd problem. We mention, in particular, Johansson ([Joh1906a, Joh1906b]). Then at the meeting of May 11, 1907 of the Göttingen Scientific Society, Klein presented a note by P. Koebe [Koe1907b] announcing that he had proved the general uniformization theorem:

Theorem. — *Every simply connected Riemann surface (supporting a non-constant meromorphic function*[15]*) is biholomorphic to the Riemann sphere, the complex plane, or the unit disc.*

---

[14]It is interesting to read Osgood's presentation of Poincaré's result and its inadequacies in a series of talks given in Cambridge in 1898 [OsgW1898].

[15]At that time Riemann surfaces were always conceived as extended over the plane. However,

The case of compact simply connected Riemann surfaces (homeomorphic to the sphere $\mathbb{S}^2$) had been dealt with already in papers by Schwarz and Neumann: they are all biholomorphic to the Riemann sphere. Thus there remained only the case of non-compact simply connected Riemann surfaces. Given such a Riemann surface S, Koebe considers an exhaustion of it by means of an increasing sequence $(D_n)_{n\geq 0}$ of simply connected regions with compact closure and with polygonal boundaries, and chooses a fixed point $p_0 \in D_0$. Schwarz had shown the existence, for each n, of a biholomorphism φ_n from D_n onto the unit disc of $\mathbb{C}$, sending the prescribed point p_0 to the origin. If the sequence of moduli of the derivatives of the φ_n at p_0 could be shown to be bounded, then from work of Harnack and Osgood it would follow that the surface S is uniformized by the unit disc. Thus the whole of Koebe's paper is devoted to showing that, if the sequence of derivatives of the φ_n at p_0 should diverge, one can nonetheless construct from the sequence $(\varphi_n)_{n\geq 0}$ a different sequence $(\psi_n)_{n\geq 0}$ of biholomorphisms that converges to a biholomorphism between S and the complex plane. The key argument involved in constructing the sequence $(\psi_n)_{n\geq 0}$ is very subtle, and contains in embryo a version of the so-called *Koebe's Quarter Lemma*. But even if it is difficult to grasp[16], Koebe's proof is nevertheless perfectly rigorous.

Six months later, an article by Poincaré [Poin1907] appeared in *Acta Mathematica* in which he also proposed a proof of the general uniformization theorem, one very different from Koebe's[17]. For a given non-compact, simply connected Riemann surface S, Poincaré considers the region A obtained by removing a small disc. He notes that the surface S will be biholomorphic to the complex plane or the unit disc provided A admits a *Green's majorant*, that is, a positive harmonic function with at least one logarithmic pole. It then remains to construct such a function. To this end, Poincaré again generalizes the *alternating procedure* invented by Schwarz, and gives a physical interpretation of the procedure he defines, which he calls the "*sweeping method*".[18] Suppose one wishes to construct on a surface A a function u with a logarithmic pole at a point p_0, harmonic on $A \setminus \{p_0\}$, and tending to zero at infinity. Such a function may be thought of as given by the electric potential associated with a negative point charge situated at

---

Koebe's proof works for abstract Riemann surfaces.

[16]It is appropriate to mention that the article [Koe1907a] was in the form of a communication to the Göttingen Scientific Society, and that the details suppressed in such communications were often intended for publication in a "real" mathematics journal. In fact Koebe continued for the rest of his life to reprise different proofs of the theorem in order to make it more accessible and more general, and improve its presentation. See, for instance, [Koe1907a, Koe1907b, Koe1908a, Koe1909a, Koe1909b, Koe1909c, Koe1909d, Koe1910b, Koe1911].

[17]Poincaré did not know of Koebe's proof when he was preparing his article, submitted in March 1907.

[18]Usually translated into English as the "scanning method". *Trans*

the point p_0. To construct it, Poincaré imagines the following:

— starting with an arbitrary function $u_0 : A \to \mathbb{R}$ with a logarithmic pole at p_0 and tending to zero at infinity, visualized as the potential associated with a distribution of charge $\rho_0 := \Delta u_0$;

— letting each small region of the surface gradually become more and more "conducting" in order to be able to "sweep" the charges (except for that at p_0, which is to be maintained artificially in place) towards the boundary of each of these regions. One hopes that at the end of this process, all charges (save that at p_0) will have been "swept to infinity"; the associated potential will then give the desired function. In mathematical terms, one covers A by holomorphic discs, and constructs a sequence $(u_n)_{n \geq 0}$ of continuous functions, with the property that u_{n+1} is the same as u_n everywhere except on one of the discs, on which it is harmonic.

Of course, the bulk of the work is involved in showing that the sequence $(u_n)_{n \geq 0}$ converges. As so often, Poincaré's proof, although not a model of rigor, contains luminous intuitions. In particular, he uses a physical argument (the conservation of the total electric charge when a disc in the Riemann surface is "made conducting") difficult to justify mathematically without using the theory of distributions.

Poincaré's memoir appeared at the beginning of November 1907. At the end of that same month, Koebe, who had read Poincaré's memoir avidly, submitted a new note to the Göttingen Scientific Society [Koe1907b], containing a proof of the general uniformization theorem largely inspired by Poincaré's proof. In fact Koebe reprises the global strategy of Poincaré's proof, but with the "sweeping method" replaced by a much more direct construction based on an exhaustion of A by regions of compact closure, thus gaining in simplicity (and rigor) what had been lost in physical intuition.

In the introduction to Part B, we explained how in 1881 Klein was an established professor who soon found himself outmatched by the young Poincaré. In 1907 it was Poincaré who was the established one and who must have felt a little hustled by the young Koebe, only 25 years old. The following anecdote shows clearly the difference in status between the two rivals: at the International Congress in Rome in 1908, both Koebe and Poincaré gave addresses. Koebe's was entitled "On the uniformization problem…", while Poincaré's was "On the future of mathematics"!

In sum, in 1883 Poincaré shows that every Riemann surface (on which a meromorphic function can be defined) admits a branched covering space biholomorphic to a bounded simply connected open subset of the plane. His proof depends on ideas of Schwarz allowing the uniformization of every relatively compact, simply connected region with polygonal boundary of a Riemann surface, and uses an exhaustion of his arbitrary non-compact simply connected Riemann surface by means of such regions, that is, by a sequence of relatively compact, simply connected regions with polygonal boundaries. The existence of such an exhaustion — which Poincaré does not prove — is not difficult to establish in the case of a Riemann surface extended over the plane[19] (see §XI.2). Then in 1900, Osgood shows that every bounded simply connected open subset of the plane is biholomorphic to the disc. Thus then it becomes known that every Riemann surface has a branched covering space biholomorphic to the disc. In his address to the International Congress of Mathematicians of that year, Hilbert emphasizes the inadequacy of this result, and urges mathematicians to try to prove a "true" uniformization theorem for non-algebraic Riemann surfaces. In May 1907, Koebe publishes the first proof of the general uniformization theorem, also based on work of Schwarz and on the existence of exhaustions by means of relatively compact, simply connected regions with polygonal boundary. (This proof seems to us to be perfectly correct and rigorous.) Just prior to the publication of Koebe's memoir, Poincaré also completes and prepares for publication a proof of the general uniformization theorem, which appears in early November 1907. (This proof, based on physical intuition, seems a very natural one to us; however, it cannot be made rigorous without recourse to the theory of distributions.) At the end of November 1907, Koebe publishes a "simplified" version of Poincaré's proof, with the "sweeping method" replaced by an appeal to work of Schwarz and the use of an exhaustion of a simply connected Riemann surface with a small disc removed, by means of relatively compact annuli. This "cleaned up" version of Poincaré's proof is, although certainly less intuitive than its original, especially brief, and seems to us rigorous as it stands.

Thus by the close of the year 1907, the uniformization theorem was firmly established. Of course, the process of assimilation of the result was far from complete, and it would take another fifteen years before the proofs began to appear that one finds in today's books (in this connection, see our annotated bibliography). Early on, mathematicians switched predominantly to the search for results beyond the uniformization theorem: Koebe was already beginning to think about uniformizing non-simply-connected Riemann surfaces [Koe1910b], and Hilbert was already inviting mathematicians to investigate the uniformizability of complex manifolds of higher dimensions. . . but our book stops in 1907.

Chapter XI is devoted to Schwarz's theorem on the uniformization of simply connected regions with compact closure, Poincaré's results of 1883 on the uniformization of functions, and Osgood's theorem. Then, in Chapter XII, we expound the first of Koebe's proofs of the general uniformization theorem from

---

[19]Prior to the work of Weyl in the 1910s, a Riemann surface was *by definition* extended over the plane.

his note [Koe1907a]. Finally, Chapter XIII is devoted to Poincaré's proof of the same theorem in [Poin1907], and also to the simplification of that proof proposed by Koebe in [Koe1907b].

Box: The classification of surfaces

In a completely natural way the theory of Riemann surfaces — the veritable topic of this book — evolved in parallel with the *topological* theory of surfaces, that is, of 2-dimensional manifolds not endowed *a priori* with a complex structure. While the history of these developments would furnish enough material for another book, we thought it nonetheless apropos to indicate here some of the most important milestones. In their progress towards the general uniformization theorem, Poincaré and Koebe used, proved, or quite simply anticipated the main results of the topology of surfaces. Often the borrowings from topology are completely implicit. Yet again do we find the situation somewhat confused.

The topological classification of compact surfaces took place gradually, progressively gaining in rigour and generality. The very concept of surface was some time in maturing, from the idea of a surface as embedded in 3-dimensional space to the conception of an abstract surface. Moreover two surfaces embedded in 3-space could be homeomorphic without there existing any homeomorphism of the ambient space sending one to the other: thus, for instance, a torus might be knotted in 3-space. And then it became necessary to distinguish degrees of regularity of surfaces under investigation, which might be smooth or merely topological. Fractal sets, arising at the same time as Kleinian groups, furnish many examples of curves that are not differentiable and whose local properties are such as to make one despair of any topological classification.

The main theorem, become classical, may be stated as follows:

Theorem. — *Every compact connected orientable surface is homeomorphic to a sphere or to a connected sum of tori.*

This theorem was "known" — and used — by B. Riemann, with no attempt made to justify it. For its history one may consult [Pont1974]: the most important names in this connection are F. A. Möbius [Möb1863], C. Jordan [Jor1866] and W. von Dyck [Dyc1888]. The first proofs meeting (almost) today's standards of rigour date from the 1860s and use two different kinds of ideas. They assume implicitly that the surfaces are smooth.

First Möbius produced a remarkable proof in the case of compact surfaces embedded in space (which is *a posteriori* equivalent to orientability).

His proof involves choosing a real-valued function on the surface and investigating the nature of its level curves. By means of successive modifications of the function he simplifies it so as to step by step eliminate critical points and reduce it to "standard" form. (One sees here the germ of what will much later be called Morse theory.) He also proves that the surface may be cut into two planar surfaces, that these are characterized to within a homeomorphism in terms of the number n of components of their boundary, and that this number is the only invariant of the initial closed surface. He observes also that $n - 1$ is the greatest number of disjoint closed curves on the surface that do not disconnect it, thus recovering Riemann's definition of genus. These ideas were then elaborated on and consolidated by, among others, J.C. Maxwell [Max1870] and C. Jordan [Jor1872].

Jordan takes a different approach, in some sense reprising Riemann's method, which consists in cutting the surface along disjoint simple closed curves. His surfaces are compact, smooth, and without boundary, but not necessarily embedded in 3-space: curves of self-intersection are allowed, so that in fact he allows his surfaces to be immersed in 3-space.

The classification of nonorientable surfaces was also carried out progressively. In 1861, J.B. Listing (to whom, incidentally, we owe the word "topology" [Lis1847]) appears to have been the first to describe the nonorientable surface with boundary that today we call the Möbius strip [Lis1861], and in 1882 Klein described the "bottle" bearing his name in an article discussed earlier [Kle1882c]. In 1886, Möbius clearly defines the concept of orientability [Möb1886], and then Dyck obtains the classification of arbitrary compact, smooth surfaces, possibly with boundary, possibly nonorientable [Dyc1888]. Volume 6 of Poincaré's collected works includes a glossary allowing one to pass from the topological terminology of 1950 back to that of Poincaré. For example, opposite "Möbius strip" one finds Poincaré's term "the one-sided surface that everyone knows".

This was all made precise in an article by Dehn and Heegaard in 1907 [DeHe1907]. Here the surfaces are triangulated, and are allowed to be nonorientable and have non-empty boundary. The classification is combinatorial in nature, and the arguments are convincing. Klein comments on this article that it is "written in a rather abstract style. . . . It begins by formulating the concepts and facts fundamental to topology. Then the rest is deduced in a purely logical manner. This contrasts completely with the inductive presentation that I have always recommended. To be understood plainly, [let me say that] this article presupposes of the reader that he has already pondered the topic deeply in the inductive manner" [Kle1925].

The characterization of *topological* surfaces to within a homeomorphism will take more time, as we shall now see.

The Jordan Curve Theorem and the Osgood–Schoenflies Theorem

Theorem. — *The complement of a simple closed curve in the plane has exactly two connected components.*

This theorem was stated by Jordan in 1887 [Jor1887]. The "proof" he proposed did not seem convincing to those who commented on it [Veb1905, Ale1920, Schoe1906, DoTi1978], and it should be noted that it assumed the statement to be obvious in the case of a polygonal (or smooth) curve ...
A proof for a polygonal curve was in fact first given by Schoenflies in 1896 [Schoe1896]. The first complete proof of the full theorem seems to be that given by Veblen in 1905 [Veb1905].

Consider these dates in relation to the period of relevant activity of the protagonists of this part of our book — Poincaré and Koebe — from 1883 to 1907. Since their interest lay with Riemann surfaces, which are necessarily smooth, all the theorems on the classification of surfaces were at their disposal, and indeed they exploited them to the full, though sometimes without mention.

The following theorem, especially delicate in the case of non-smooth curves, progressively makes its appearance during the same period.

Theorem. — *Every simple closed curve in the plane can be mapped onto a circle by means of a global homeomorphism of the plane.*

Here are some comments on the history of this result, traditionally called "Schoenflies' Theorem", drawn largely from a recent publication of Siebenmann [Sieb2005].

Even though the arguments Jordan used in his attempt to prove the Jordan Curve Theorem [Jor1887] were not convincing, they still showed essentially that the bounded component of the complement of a curve is homeomorphic to an open disc. This fact was explicitly established using conformal methods in 1900 by Osgood in an article we will be discussing later on [OsgW1900].

It was in 1902 that Osgood stated Schoenflies' Theorem [OsgW1902]; however it would take another ten years or so before the first complete proofs appeared, again using conformal methods [Car1913a, Car1913b, Car1913c, Koe1913a, Koe1913b, Koe1915, OsTa1913, Stu1913]. Schoenflies stated "his" theorem clearly enough in 1906 [Schoe1906]. His proof, fully correct in the case of a polygonal or smooth curve, was, however, lacking in the general case.

The first correct proof, using only topological arguments (and not conformal ones) seems to be due to Tietze in 1914 [Tie1913, Tie1914] or to Antoine in 1921 [Ant1921]. The name "Schoenflies' Theorem" was given to the result by Wilder in 1949 [Wil1949].

The *topological* classification of surfaces

We emphasize once again that since the interest of Poincaré and Koebe was concentrated on Riemann surfaces, and these are automatically smooth, the question of the structure of topological surfaces was of no direct interest to them at that time. It seemed to us nonetheless useful to give a quick description of the later developments concerning topological surfaces.

Schoenflies' theorem would be the key allowing Radó to prove in 1925 that every topological surface countable at infinity is triangulable, and thence to obtain a classification in the compact case [Rad1925].

The outstanding case of noncompact surfaces was dealt with thanks to the introduction of the idea of end compactification by Freudenthal, Kerékjártó and Schoenflies. The complete classification in the noncompact case was obtained by Kerékjártó in 1923 [Ker1923], and fully rigorized by Richards in 1963 and Goldman in 1971 [Ric1963, GolM1971].

We mention in conclusion a particular case that will be needed in the proof of Lemma XI.2.1: a noncompact, simply connected surface is homeomorphic to the plane.

Chapter XI

Uniformization of functions

As mentioned in the introduction to this, the last part of the book, in the 19th century the uniformization problem for Riemann surfaces was above all a problem in function theory: given an non-single-valued analytic function y of a variable x, to find a variable z such that y and x are single-valued functions of z. This comes down to parametrizing the Riemann surface defined by the many-valued function y by means of a variable z ranging over an open subset of $\mathbb{C}$. In his 1883 memoir, Poincaré achieves a double *tour de force*: on the one hand, he uniformizes *all* (analytic) functions, with the caveat that the parametrization $z \mapsto (x(z), y(z))$ may have branch points, and on the other hand, he shows that the variable z can be constrained to range over a bounded simply connected open subset of the plane. One has just reflect for a moment on the diversity and complexity of many-valued functions — solutions of algebraic differential equations provide good examples and surely formed one of Poincaré's main motivations — to be convinced of the revolutionary character of this assertion.

It took till 1900 and the work of Osgood before a parametrization by the disc was obtained. In fact, it follows from Osgood's main result [OsgW1900] that every (non-empty) bounded simply connected open subset of $\mathbb{C}$ is biholomorphic to the unit disc. For the material of the present chapter, one may also consult [Cho2007, Gra1994].

XI.1. Uniformization of relatively compact regions with boundary

We begin by returning for a moment to the important stage represented by Schwarz's work at the turn of the 1870s. We have seen in Chapter IV how Schwarz obtained, in [Schw1869], "explicit" formulae for biholomorphisms from the disc onto polygonal regions of the plane. The logical path he took in achieving this is as follows: he assumed *a priori* the existence of a biholomorphism between the region in question and the unit disc, and then determined a formula for it. Thus, in order to complete the argument, it remained for him to show that every simply connected polygonal region is indeed biholomorphic to the unit disc — for example, by proving rigorously that such regions admit Green's functions (see below). It was to this task that he addressed himself in [Schw1870a], introducing a very elegant method now known as *the alternating method*. This method was taken up and generalized by his contemporaries — for example, by C. Neumann in the last chapter of [Neum1884] — and it is still used today to obtain solutions of certain partial differential equations. Poincaré elaborated on it in his 1890 article [Poin1890] under the name of the "sweeping method" and also used it crucially in his 1907 proof of the general uniformization theorem. Schwarz's techniques are also of use in going well beyond the case of planar polygonal regions to that of relatively compact regions of an arbitrary Riemann surface, with boundary analytic or polygonal. (We shall give precise meaning below to the word "polygonal" as applied to a region of a Riemann surface.)

We first of all clarify the connection between uniformization and the construction of Green's functions. We have already defined (in §II.2.2) what a Green's function is for an open subset of the plane. We now generalize that definition to an arbitrary Riemann surface:

Definition XI.1.1. — (Of a simple logarithmic singularity): Let S be a Riemann surface. We shall say that a function $u : S \to \mathbb{R}$ presents a *simple logarithmic singularity* at a point p_0 of S if, given a local holomorphic coordinate z in a neighborhood of p_0 on S, the function

$$p \mapsto u(p) + \log|z(p) - z(p_0)|$$

is bounded in a neighborhood of p_0.

Definition XI.1.2. — (Of a Green's function): Let S be a (connected) noncompact Riemann surface. A function $u : S \to \mathbb{R}$ is called a *Green's function* on S if there exists a point p_0 of S such that:

(i) u is harmonic on $S \setminus \{p_0\}$;

(ii) u presents a simple logarithmic singularity at p_0;

(iii) $u(p)$ tends to 0 as p leaves every compact subset of S.

Remark XI.1.3. — It follows from the maximum principle that a Green's function must be positive and that a given surface S can have at most one Green's function with the logarithmic singularity at a prescribed point.

The relevance of Green's functions to the uniformization of Riemann surfaces is made clear by the following theorem.

Theorem XI.1.4. — *If a simply connected Riemann surface S admits a Green's function, then it is biholomorphic to the unit disc $\mathbb{D}$.*

Proof. — The proof is very similar to that in §II.2.2 dealing with the case where S is a simply connected open subset of the plane.

Let u be a Green's function on S with its logarithmic singularity at a point p_0 of S. Let α be the 1-form defined on $S \setminus \{p_0\}$ by

$$\alpha(\xi) = -du(i\xi).$$

The harmonicity of u ensures that the form α is closed. Let γ be a small loop turning once around p_0. A calculation within a chart containing the given singularity yields

$$\int_\gamma \alpha = -2\pi.$$

Write u^* for a primitive function of α; this will be a many-valued function (since the integral of α around γ is nonzero) but the function $F := e^{-(u+iu^*)}$ is nonetheless well-defined (that is, single-valued): for, the surface S being simply connected, every loop is homologous to an integer multiple of γ and therefore the integral of α along any loop will be an integer multiple of 2π. Since $u > 0$, the values taken by F will lie in the unit disc $\mathbb{D}$. As p leaves every compact subset, we have $u(p) \to 0$ (by definition of a Green's function), whence $|F(p)| \to 1$. Thus the map F is proper, so that its image is closed. This image is also open since this is always the case for a non-constant holomorphic map. Hence F is surjective. It remains to verify that F is injective. Since $F : S \to \mathbb{D}$ is proper and holomorphic, the fibre $F^{-1}(w)$ over a point w of $\mathbb{D}$ is of finite cardinality, and this cardinality is locally constant provided we count multiplicities. Since $F^{-1}(0) = \{p_0\}$ and $F'(p_0) \neq 0$, it follows that each fibre is a singleton, so the map F is injective. □

We now state the result yielded by the techniques developed by Schwarz in [Schw1870a].

Theorem XI.1.5. — *Let S be a Riemann surface and Ω a region with compact closure, and with boundary analytic or polygonal. Then for every continuous function $u : \partial\Omega \to \mathbb{R}$ we have:*

1. *there exists a unique continuous extension $\tilde{u} : \overline{\Omega} \to \mathbb{R}$, harmonic in the interior of Ω;*

2. *if p_0 is a point in the interior of Ω, then there exists a unique continuous extension $\overline{u} : \overline{\Omega} \setminus \{p_0\} \to \mathbb{R}$, harmonic on $\Omega \setminus \{p_0\}$ and presenting a simple logarithmic singularity at p_0.*

Hence, in particular, we have:

Corollary XI.1.6. — *Let S be a Riemann surface and Ω a simply connected region of S with compact closure and with boundary analytic or polygonal. Then for every point p_0 of Ω, there exists a Green's function on Ω with its logarithmic singularity situated at p_0. It follows that Ω is biholomorphic to the unit disc.*

As already mentioned, this corollary together with Theorem IV.1.6, allow one to show that every compact, simply connected Riemann surface is biholomorphic to the Riemann sphere (see Chapter IV).

We now give the precise definition of the phrase *region with analytic boundary* appearing in the statements of Theorem XI.1.5 and Corollary XI.1.6: this is an open subset with boundary a non-empty real analytic submanifold of dimension one[1]. And by *region with polygonal boundary* we mean an open subset with boundary a one-dimensional topological submanifold with the further property that there exist a holomorphic atlas of the Riemann surface such that, on each chart meeting the boundary of the region, the intersection appears as a straight-line segment or as a "corner" (two segments issuing from a single point). We should point out that Theorem XI.1.5 is not exactly as Schwarz stated it in [Schw1870a], although his contemporaries did explicitly attribute it to him in this form. Thus the techniques that he used were capable of being adapted by sometimes "reading between the lines" so as to obtain the degree of generality exhibited in the above statement of Theorem XI.1.5.

Proof of Corollary XI.1.6. — In view of the second conclusion of Theorem XI.1.5, applied to the zero function on the boundary $\partial\Omega$, the open set Ω admits a Green's function. The corollary now follows by applying Theorem XI.1.4 to that open set. □

Proof of Theorem XI.1.5. — We begin by determining, for each point p of $\overline{\Omega}$, a "privileged" neighborhood $\overline{D_p}$ of p in $\overline{\Omega}$. If p is an interior point of Ω, we take as the neighborhood D_p a small Euclidean disc in a local holomorphic coordinate such that $\overline{D_p}$ is contained in Ω. If p belongs to the smooth part of the boundary of Ω, then there exists, by the assumption concerning $\partial\Omega$, a holomorphic chart $z_p : U_p \to \mathbb{C}$ with U_p a neighborhood of p in S, such that $z_p(U_p \cap \partial\Omega)$ is a line-segment. In this case we consider a Euclidean triangle T contained in $z_p(U_p \cap \overline{\Omega})$, such that one of its sides is contained in $z_p(U_p \cap \partial\Omega)$ and $z_p(p)$ is not a vertex of T, and we take $D_p = z_p^{-1}(\mathrm{Int}(T))$. Finally, if $\partial\Omega$ is polygonal and p is a "vertex" of $\partial\Omega$, then we choose a holomorphic chart $z_p : U_p \to \mathbb{C}$ with the property that in a neighborhood of p, the chart z_p maps $\partial\Omega$ onto the boundary of a Euclidean sector with vertex $z_p(p)$. We then complete this sector to a triangle and, as in the

[1]Thus, in particular, a compact Riemann surface and a compact Riemann surface with a point removed are not considered regions with analytic boundary.

preceding case, define the desired neighborhood D_p to be the pre-image of the interior of this triangle under z_p.[2]

For each p in $\overline{\Omega}$, the neighborhood $\overline{D_p}$ so constructed has the important property that there exists a biholomorphism $z : D_p \to \mathbb{D}$ from D_p onto the unit disc $\mathbb{D}$ that extends to a homeomorphism from $\overline{D_p}$ to $\overline{\mathbb{D}}$. This is clear in the case when p is an interior point of Ω, and when p is on the boundary of Ω it follows from Schwarz's work in [Schw1869]. (See §IV.1.3: for every Euclidean triangle T, there in fact exists a function of the form

$$z \mapsto A + \int_0^z \frac{B}{(w-a)^{1-\alpha}(w-b)^{1-\beta}} dw$$

sending the upper half-plane biholomorphically onto T and extending to a homeomorphism on the boundary.) In what follows we will consider a finite cover $\mathcal{R}$ of $\overline{\Omega}$ by the neighborhoods just defined.

If D is a region in $\mathcal{R}$, then for a continuous (or possibly just piecewise continuous) function $u : \partial D \to \mathbb{R}$ we define *the harmonic extension of u to D* as follows: By identifying $\overline{D}$ with the unit disc $\overline{\mathbb{D}}$ by means of a biholomorphism, we may regard u as a function defined on the boundary of $\mathbb{D}$. There exists a harmonic function $\tilde{u} : \mathbb{D} \to \mathbb{R}$ such that $\tilde{u}(p)$ tends to $u(\zeta)$ as an arbitrary point p of the disc tends to a point ζ of the circle $\partial\mathbb{D}$ where u is continuous; this function $\tilde{u}$ is unique and is given by Poisson's formula:

$$\tilde{u}(z) = \int_0^1 \frac{1-|z|^2}{|1-e^{-2i\pi\theta}z|^2} u(e^{2i\pi\theta}) d\theta.$$

We now regard $\tilde{u}$ as being defined on $\overline{D}$: this is then the harmonic extension of u to D.

If D is as before a region of $\mathcal{R}$ and $u : \overline{\Omega} \to \mathbb{R}$ is a function for which $u_{|\partial D}$ is piecewise continuous, we can make it harmonic on D by forming the function $B(u, D)$ that coincides with u on $\Omega \setminus D$ and with the harmonic extension of $u_{|\partial D}$ on D. Note that if one begins with a function u continuous on $\overline{\Omega}$, then $B(u, D)$ will remain continuous.

1. The existence of a harmonic extension. — Consider any continuous function $u : \partial\Omega \to \mathbb{R}$. We start by extending u to a function $u_0 : \overline{\Omega} \to \mathbb{R}$ coinciding with u on $\partial\Omega$ and taking the value $m = \inf_{\partial\Omega} u$ on Ω. We shall progressively render u_0 harmonic on an infinite sequence (D_n) of discs of the open cover $\mathcal{R}$. We begin by choosing $D_1 \in \mathcal{R}$ such that $\overline{D_1} \cap \partial\Omega$ has non-empty interior in $\partial\Omega$. Next we choose a finite sequence $D_1, \ldots, D_r$ of members of $\mathcal{R}$ so that $\partial D_{j+1} \cap D_j$ is non-empty for

[2] Note that if p is on the boundary of Ω, then we shall not have $p \in D_p$, although we will have $p \in \overline{D}_p$.

each $j = 1, \ldots, r-1$, and every D of $\mathcal{R}$ is equal to some D_j (possibly several); the sequence (D_n) is then defined to be $D_1, \ldots D_r, D_1, \ldots D_r, \ldots$, in terms of which we define a sequence of functions $u_0, u_1, \ldots$ by the formula $u_{n+1} = B(u_n, D_{n+1})$.

The reader will be able to check that this sequence is well-defined and that the functions u_n are continuous in the interior of Ω. Although they need not all be continuous on $\overline{\Omega}$, we do have u_n continuous on $\overline{\Omega}$ for $n > r$. To see this, let p be a point on the boundary of Ω. There exists a smallest index $k < r$ such that p is in the interior of $\overline{D_k} \cap \partial\Omega$. The function u_{k-1} restricted to a neighborhood of p in ∂D_k is continuous, and its harmonic extension to D_k defined by the Poisson formula is also continuous at p. One then proves by induction that all of the functions u_n with $n \geq k$ are continuous at p. For, suppose this is not the case for some such u_n. Then if ∂D_{n+1} does not contain p, we shall have $u_{n+1} = u_n$ in a neighborhood of p in $\overline{\Omega}$, which implies that u_{n+1} is continuous at p. For otherwise the point p would have to be in ∂D_{n+1}, and since the restriction of u_n to ∂D_{n+1} is continuous at p, its harmonic extension to the interior of D_{n+1} must also be continuous at p. Hence u_{n+1} is indeed continuous at p and it follows inductively that u_n is continuous at p for all $n > r$.

We next show that the sequence (u_n) is increasing. To this end, consider a function $w : \overline{\Omega} \to \mathbb{R}$ with the property that for all $D \in \mathcal{R}$ one has $B(w, D) \geq w$. This is the case for u_0, for instance, by the maximum principle. Let D' be a disc in $\mathcal{R}$ and write $w' = B(w, D')$. We claim that then w' again satisfies $B(w', D) \geq w'$ for every disc D of $\mathcal{R}$. To see this, note first that $B(w', D) = w'$ outside D. Also, $w' \geq w$ by the assumption on w. In $D \setminus D'$ we have $w' = w$, whence the desired inequality since $B(w', D) \geq B(w, D) \geq w$. It remains to verify the inequality in $D \cap D'$. Now the function $B(w', D) - w'$ vanishes on ∂D and, on $\partial D'$, we have

$$B(w', D) - w \geq B(w, D) - w \geq 0.$$

Hence by the maximum principle we have $B(w', D) \geq w'$ on $D \cap D'$. Applying this result first with $w = u_0$, then with $w = u_1$, and so on, we infer that the sequence (u_n) is indeed increasing.

For each positive integer n, we define $v_n := u_{n+2r} - u_{n+r}$, whence $v_{n+1} = B(v_n, D_{n+1})$. We shall need the following lemma.

Lemma XI.1.7. — *Let v_0 be a continuous function on $\overline{\Omega}$, vanishing on the boundary of Ω, and (v_n) the sequence defined inductively by $v_{n+1} = B(v_n, D_{n+1})$. Then there exists a constant q, $0 \leq q < 1$, such that, for every positive integer k,*

$$\|v_{(k+1)r}\|_\infty \leq q \|v_{kr}\|_\infty. \tag{XI.1}$$

Proof. — Consider the sequence of functions $f_0, f_1, \dots, f_r$ on $\overline{\Omega}$ defined as follows: The function f_0 is defined to be zero $\partial\Omega$ and identically equal to 1 in the interior of Ω. One then defines inductively $f_{k+1} = B(f_k, D_{k+1})$ for $k = 0, \dots, r-1$, and sets $q := \sup_{\overline{\Omega}} f_r$. Since $v_{kr} \leq f_0 \|v_{kr}\|_\infty$, it follows that

$$v_{kr+1} = B(v_{kr}, D_1) \leq B(f_0\|v_{kr}\|_\infty, D_1) = f_1\|v_{kr}\|_\infty.$$

Hence we have inductively that

$$v_{(k+1)r} \leq f_r\|v_{kr}\|_\infty \leq q\|v_{kr}\|_\infty.$$

To establish the lemma it remains to show that $q < 1$. It follows as proved above for the functions u_n, $n \geq r$, that the function f_r is continuous on $\overline{\Omega}$. Hence the function f_r attains its greatest value at a point p_0 of $\overline{\Omega}$. Since — as the reader can easily check — f_r is not identically zero, the point p_0 must lie in the interior of Ω, and therefore in the interior of some disc D_j. Note now that if a piecewise continuous function lies between 0 and 1 on the boundary of a region $D \in \mathcal{R}$ and is less than 1 on a non-empty open subset of that boundary, then its harmonic extension, obtained using the Poisson formula, is less than 1 in the interior of that region. Hence the function f_1 is less than 1 in the interior of D_1 and therefore on a non-empty open subset of ∂D_2. Iterating the argument, we infer that f_2 is less than 1 in the interior of D_2, and therefore on a non-empty open subset of ∂D_3. It follows by induction that all f_j are less than 1 in the interior of D_j. Since the sequence (f_k) is decreasing (this follows by an argument similar to that used above to establish that the sequence (u_n) was increasing), it follows that $q = f_r(p_0) \leq f_j(p_0) < 1$. This completes the proof of the lemma. □

This lemma now gives, for all $i = 1, \dots, r$ and every positive integer k, that

$$\|v_{kr+i}\|_\infty \leq \|f_i\|_\infty\|v_{kr}\|_\infty \leq \|v_{kr}\|_\infty \leq q^k\|v_0\|_\infty.$$

Hence the series

$$u_{i+r} + (u_{i+2r} - u_{i+r}) + (u_{i+3r} - u_{i+2r}) + \cdots$$

converges uniformly on $\overline{\Omega}$ to a continuous function $U_i : \overline{\Omega} \to \mathbb{R}$ equal to u on $\partial\Omega$ and harmonic on D_i (to see which, one observes that each term of the above series is harmonic on D_i, whence by the dominated convergence theorem U_i has the mean-value property). To conclude the proof, observe first that $U_i \leq U_{i+1}$ since $U_i = \lim_{k\to\infty} u_{i+kr}$, $U_{i+1} = \lim_{k\to\infty} u_{i+1+kr}$ and $u_{i+kr} \leq u_{i+1+kr}$. Then since $U_{i+r} = U_i$, it follows that $U_{i+1} = U_i$ for all i. The function U_i is therefore harmonic in the interior of Ω, continuous on $\overline{\Omega}$, and coincides with u on $\partial\Omega$. The Dirichlet problem is solved.

2. The existence of a harmonic extension with a simple logarithmic singularity. — Choose a point p_0 in the interior of Ω. We may assume, up to modifying the open cover $\mathcal{R}$, that, on the one hand, no disc in $\mathcal{R}$ has p_0 on its boundary, and, on the other hand, p_0 belongs to just the one disc D_{i_0} of the sequence $D_1, \ldots, D_r$ and has coordinate $z_{i_0}(p_0) = 0$. We now modify the definition of the sequence (u_n) in the following way: the first term of the sequence is to be u_{i_0}, defined by $u_{i_0}(p) = \log\left(\frac{1}{|z_{i_0}(p)|}\right)$ in the interior of D_{i_0} and by $u_{i_0} = 0$ outside D_{i_0}. Then we define u_{n+1} inductively in terms of u_n by $u_{n+1} = B(u_n, D_{n+1})$ if $n+1$ is not congruent to i_0 modulo r, and by $u_{n+1} = u_{i_0} + B(u_n - u_{i_0}, D_{n+1})$ if $n+1$ is congruent to i_0 modulo r. It should now be clear that for $n \geq i_0$ each function $u_n : \overline{\Omega} \setminus \{p_0\} \to \mathbb{R}$ is continuous. One shows as before that $u_n \leq u_{n+1}$.

For each positive integer $n \geq i_0$, we set $v_n = u_{n+r} - u_n$. We then have $v_{n+1} = B(v_n, D_{n+1})$ and the restrictions of the v_n to the boundary of Ω are identically zero. Furthermore v_0 is continuous, so Lemma XI.1.7 applies, and the inequality (XI.1) holds. Hence if $\tilde{u}$ denotes the harmonic extension of u to Ω (obtained as in Part 1 of the present proof) then an argument similar to that of Part 1 shows that for $i = i_0, \ldots, i_0 + r - 1$, the series

$$U_i = \tilde{u} + u_{i+r} + (u_{i+2r} - u_{i+r}) + \cdots$$

converge uniformly to one and the same limit function, which is then the harmonic extension of u to Ω with a simple logarithmic singularity at p_0. □

XI.2. Exhaustion by means of relatively compact, simply connected regions

Suppose we wish to uniformize a relatively compact, simply connected Riemann surface S. In view of Corollary XI.1.6, a natural strategy would be to choose an exhaustion $D_0 \subset D_1 \subset \ldots$ of the surface by simply connected regions D_k with compact closure, each admitting a Green's function g_k with a logarithmic singularity at a point p_0 of D_0, and then attempt to understand the behaviour of the sequence (g_k). This, very briefly, represents the starting point of the approach to the problem taken by Poincaré in 1883, and then by Koebe in 1907. We will return to this later on. First, however, we need to show that an exhaustion $D_0 \subset D_1 \subset \cdots$ such as just described does indeed exist.

Lemma XI.2.1. — *Every noncompact, simply connected Riemann surface S admits an exhaustion by simply connected regions with compact closures and polygonal boundaries.*

Proof. — We first observe that, leaving aside the question of boundaries, it is easy to prove the existence of an exhaustion by a family of simply connected regions with compact closures if one knows that S is homeomorphic to the plane — and this could have been inferred immediately from the classification of surfaces were it not for the very pertinent fact that the classification of non-compact surfaces was not achieved till 1923, so Poincaré and Koebe could not avail themselves of it! That did nothing to prevent them, however, from considering the existence of such an exhaustion obvious: for example, Koebe writes in [Koe1907a]: *"Die Konstruktion einer allen angeführten Bedingungen genügenden Folge von Bereichen bietet keine prinzipiellen Schwierigkeiten dar."*[3] But in fact in higher dimensions the situation is more complicated; for example, Whitehead constructed examples of *contractible* 3-dimensional manifolds not homeomorphic to Euclidean space that do not admit an exhaustion by means of topological balls.

However, by appealing to the classification of smooth surfaces with boundary, which had been established by the time of interest here, it is not difficult to prove the existence of an exhaustion $\tilde{D}_0 \subset \tilde{D}_1 \subset \cdots$ by means of topological discs with piecewise smooth boundaries. To do this, one first covers S by countably many small closed discs (see Box XI.1 below). It may always be assumed that the boundaries of these discs intersect transversely in pairs, so that the union of a finite number of the discs is a compact surface Σ embedded in S with boundary made up of a finite number of (arcs of) pairwise disjoint topological circles embedded in S. One then observes that the proof of the Jordan Curve Theorem generalizes readily to a curve on S (since in that proof only the simple-connectedness is used), so that each component of the boundary of Σ separates S into two components, one of which is relatively compact. One of these relatively compact components $\hat{\Sigma}$ will contain all the others, so that Σ is contained in $\hat{\Sigma}$, a surface with just one boundary component; thus in going from Σ to $\hat{\Sigma}$ we have effectively "plugged the holes". The surface $\hat{\Sigma}$ is necessarily of genus zero for "homological" reasons, since in the contrary case one could find two curves in $\hat{\Sigma}$ (and therefore on S) intersecting transversely in a single point — an impossibility on S since it would contradict the Jordan Curve Theorem. Thus we have constructed a closed disc containing an arbitrary finite union of closed discs, from which the existence of an exhaustion of S by a sequence $(\tilde{D}_k)$ of closed topological discs follows.

We now look to the boundaries of these discs. We begin by showing that one may arrange for the members of the exhaustion to have polygonal boundaries if there exists a non-constant meromorphic function f on our surface S (which is, we remind the reader, assumed by Koebe and Poincaré since they do not have the concept of an abstract Riemann surface at their disposal). Now, modulo modifying the $\tilde{D}_k$ slightly, we may suppose that their boundaries $\partial\tilde{D}_k$ do not pass through any branch points of f, nor through any of its poles. Then for every

[3] "The construction of a sequence of regions satisfying all the conditions itemized presents no fundamental difficulty."

positive integer k, there exists a neighborhood U_k of $\partial\tilde{D}_k$ on which f is a local biholomorphism. Hence $\gamma_k = f(\partial\tilde{D}_k)$ is a closed smooth curve in $\mathbb{C}$ (possibly with multiple points). By choosing U_k small enough, we can approximate γ_k arbitrarily closely in terms of the Hausdorff metric by means of a closed Euclidean polygonal curve ν_k that lifts via f to a simple closed curve $\tilde{\nu}_k$ in U_k. Write D_k for the connected component of compact closure of $S \setminus \tilde{\nu}_k$. The region D_k is then simply connected, of compact closure in S and with polygonal boundary by construction. If, for each k, we choose $\tilde{\nu}_k$ sufficiently close to $\partial\tilde{D}_k$, then we shall have $\overline{D_k} \subset D_{k+1}$ and (D_k) will be the desired exhaustion of S.

If one should wish to avoid the assumption that there exists a non-constant meromorphic function on S, it becomes necessary to resort to much more sophisticated arguments which the reader might prefer to omit on a first reading. Note, for example, that it is not clear that on an abstract Riemann surface one can always join two given points by means of a real analytic curve. (To prove this one needs to resort to methods and results established much later than those available at the turn of the 19th century — such as, for example, the fact that an open Riemann surface is a *Stein* surface, which then entails the existence of a non-constant holomorphic function.)

But here is an argument that would doubtless have convinced Poincaré. We begin by investigating two germs of real analytic curves in the plane, meeting transversely at a point p at an angle α, and try to see if it is possible to find a germ of a diffeomorphism, holomorphic in a neighborhood of p, transforming the two curves into straight lines. To this end, we observe that a germ of a real analytic curve is the locus of fixed points of a unique anti-holomorphic involution: the Schwarz symmetry associated with the curve. Thus on composing the two anti-holomorphic involutions associated with the two germs of curves in question, we obtain a germ of a holomorphic diffeomorphism φ fixing the point p and therefore with derivative $e^{2i\alpha}$ at p. Transforming the two curves by means of a germ h of a diffeomorphism amounts to conjugating ϕ by h. Thus the problem of rectifying the two curves by means of a germ of a diffeomorphism is equivalent to that of conjugating the germ ϕ to its linear part. The problem of linearization of germs of holomorphic functions in a neighborhood of a chosen point constitutes a chapter, become classical, of holomorphic dynamics. Poincaré showed that this linearization is possible if the derivative at the chosen point does not have modulus 1. However, in our present situation the derivative at the origin *is* of modulus 1, and even apart from that circumstance the situation is rather delicate. One had to wait till 1942 before Siegel proved that linearization is possible in our situation if α is irrational and satisfies a certain diophantine condition [Sieg1942]. What is significant here is that the set of these diophantine numbers is in fact dense in $\mathbb{R}$. (For the history of this problem of holomorphic "rectification" of germs of pairs of real analytic curves, one may consult [Kas1913, Pfe1917].)

Thus one starts as before with an exhaustion $\tilde{D}_0 \subset \tilde{D}_1 \subset \dots$ by means of topological discs with smooth boundaries. For each k, one can choose a topological disc $D_k \subset \tilde{D}_k$ with boundary piecewise analytic and such that ∂D_k and $\partial \tilde{D}_k$ are arbitrarily close in terms of the Hausdorff metric. Thus we now have a new exhaustion (D_k). To ensure that each D_k admits a Green's function one now needs to have Part 2 of the proof of Theorem XI.1.5 go through for the regions D_k. This will be possible if the boundaries of the D_k are polygonal and not just piecewise analytic. In view of the above-mentioned denseness of the "diophantine irrationals" one can in fact choose the open sets D_k in such a way that the angles at the vertices of the ∂D_k are "diophantine", so that by the argument sketched above all their corners are rectifiable and the lemma goes through. □

We have the following consequence:

Corollary XI.2.2. — *Every non-compact, simply connected Riemann surface S admits an exhaustion by means of simply connected regions with compact closure and analytic boundaries.*

Proof. — By the preceding lemma, there exists an exhaustion of S by means of simply connected regions $\tilde{D}_k$ of compact closure and with polygonal boundaries. By Corollary XI.1.6, for each positive integer k there exists a biholomorphism $F_k : \tilde{D}_k \to \mathbb{D}$. Choose a sequence (n_k) of positive integers tending to infinity, and write γ_k for the complete inverse image under F_k of the circle centred at 0 and of radius $1 - \frac{1}{n_k}$. Then if D_k is the connected component of $\tilde{D}_k \setminus \gamma_k$ whose closure in $\tilde{D}_k$ is compact, then, provided (n_k) tends sufficiently quickly to infinity, the sequence $D_0 \subset D_1 \subset \dots$ will be an exhaustion of S by means of simply connected regions with compact closures and analytic boundaries. □

Box XI.1: Remarks concerning countability

Every germ of a holomorphic function of a variable continues analytically to a maximal many-valued function, or, equivalently, defines a Riemann surface S furnished with a holomorphic function. In his 1883 article, Poincaré assumes (implicitly) that such a many-valued function takes at most countably many values at each point, or, in other words, that S has a countable base of open sets. In an 1888 note, Vivanti remarks that this limits *a priori* the domain of applicability of Poincaré's theorem:"Dunque la dimonstrazione di Poincaré vale solo per le funzioni aventi la 1 a potenza" [Viv1888a]. It is not surprising that Poincaré responded immediately that this note "interests me greatly and has prompted [in me] diverse reflections".

These "diverse reflections" went into an article appearing in the same volume as Vivanti's note [Poin1888], in which he proved in effect that such a many-valued function always takes on at most countably many values at each point. This volume also contains a second article by Vivanti [Viv1888b] giving a proof — unfortunately false — of the same result.... Another proof (correct and more detailed than Poincaré's) was published independently and almost simultaneously by Volterra [Vol1888] of what we know today as the "Poincaré–Volterra theorem".

The history of this theorem and later developments is described in a very interesting article by Ullrich [Ull2000]. One learns from it, in particular, that Cantor and Weierstrass had proved this result a little earlier. Today the proof does not seem very complicated: one can, for example, consider analytic continuations of $f_x : (\mathbb{C}, x) \to \mathbb{C}$ along piecewise polygonal paths with rational vertices, and then simply observe that the set of such paths is countable and every analytic continuation can be inferred from one along such a path. Subsequently, the Poincaré–Volterra theorem took on a purely topological character: a connected space extended over a separable space with a countable base of open sets itself has a countable open base.

It is appropriate to mention also that in 1925 Radó proved that *all* (connected) Riemann surfaces have a countable open base, independently of any *a priori* assumption about the existence of a holomorphic function [Rad1925]. For a modern proof, see [Forst1977].

XI.3. Parametrization by a simply connected open subset of the disc

In this section, we prove the following result of Poincaré from [Poin1882c]:

Theorem XI.3.1. — *The Riemann surface $S(y)$ of a germ of a meromorphic function $x \in U \subset \mathbb{CP}^1 \longmapsto y(x) \in \mathbb{CP}^1$ can be fully parametrized*[4] *by a simply connected open subset of the unit disc.*

Proof. — We take a non-constant Fuchsian function $F : \mathbb{D} \to \mathbb{CP}^1$ obtained by considering a *cocompact* lattice Γ of $\mathrm{Aut}(\mathbb{D})$ and defining $F(z)$ as the quotient of two Fuchsian series of the form

$$\sum_{\gamma \in \Gamma} R(\gamma(z))\gamma'(z)^2,$$

[4] That is, with critical points allowed.

where R is a rational function with no poles on the unit circle. Recall that F is invariant under Γ and can therefore be obtained as the composite of the covering map $\mathbb{D} \to \Gamma\backslash\mathbb{D}$ with a non-constant meromorphic function on $\Gamma\backslash\mathbb{D}$; this follows from the surjectivity of F and the fact that it has the path-lifting property: for every continuous path $\gamma : [0,1] \to \mathbb{CP}^1$ with initial point $\gamma(0) = x_0$ on the Riemann sphere and every point $z_0 \in F^{-1}(\gamma(0))$, there exists a continuous path $\beta : [0,1] \to \mathbb{D}$ such that $\beta(0) = z_0$ and $F \circ \beta = \gamma$.[5]

For the rest of the proof we shall not in fact need to assume that $S(y)$ is the Riemann surface of a function, but simply that it admits a non-constant meromorphic function; in the situation of the surface $S(y)$, the function x will do, and, moreover, it is not branched. Thus in order to simplify notation, we may instead consider a Riemann surface S that we wish to show can be parametrized by a simply connected open subset of the disc, and a non-constant function $x : S \to \mathbb{CP}^1$ which we may assume to be without critical points[6].

Poincaré's idea is to use the function F to construct a Riemann surface that will be at once "above" both S and $\mathbb{D}$. More precisely, he constructs a Riemann surface Σ and holomorphic maps $f : \Sigma \to S$ and $h : \Sigma \to \mathbb{D}$ such that $x \circ f = F \circ h$. To do this he considers the product $S \times \mathbb{D}$ and a nonempty connected component Σ of the holomorphic curve defined by the equation $x \circ \mathrm{pr}_S = F \circ \mathrm{pr}_{\mathbb{D}}$, where pr_S and $\mathrm{pr}_{\mathbb{D}}$ are the projections of $S \times \mathbb{D}$ on S and $\mathbb{D}$ respectively. Note that then Σ is non-singular since x has no branch points. We set $f = \mathrm{pr}_{S|\Sigma}$ and $h = \mathrm{pr}_{\mathbb{D}|\Sigma}$.

Note next that $f : \Sigma \to S$ is surjective. To see this, choose a point $(p_0, z_0) \in \Sigma$ and write $x_0 = x(p_0) = F(z_0)$. Let p be an arbitrary point of S and $\alpha : [0,1] \to S$ a continuous path joining $p_0 = \alpha(0)$ to $p = \alpha(1)$. We then write $\gamma := x \circ \alpha$ and consider a path $\beta : [0,1] \to \mathbb{D}$ for which $\beta(0) = z_0$ and $F \circ \beta = \gamma$. By the connectivity of Σ, the point $(\alpha(t), \beta(t))$ of $S \times \mathbb{D}$ belongs to Σ for all $0 \leq t \leq 1$. In particular, $p = \mathrm{pr}_S(p, \beta(1)) = f(\alpha(1), \beta(1))$ is in the image of f, so that f is indeed surjective.

To conclude the proof of the theorem, it only remains to show that the universal covering space $\widetilde{\Sigma}$ of Σ is biholomorphic to a simply connected open subset of the unit disc. Poincaré has first to *construct* this universal covering space, or, in other words, to demonstrate the existence of a simply connected Riemann surface $\widetilde{\Sigma}$ that covers Σ. He does this in less than a page, and in a very natural way; we give his construction in Box XI.2 below. Then in order to show that $\widetilde{\Sigma}$ is biholomorphic to a simply connected open subset of the unit disc, he uses the map $h : \Sigma \to \mathbb{D}$, a *non-constant* holomorphic mapping. He is able to view it as a function defined on $\widetilde{\Sigma}$, modulo composing it on the right with the covering map $\widetilde{\Sigma} \to \Sigma$.

[5]Note however that the uniqueness property may not hold here since F may have branch points.

[6]Critical points would appear if, instead of the Riemann surface of the function, we added to it points corresponding to "algebraic" singularities. The proof goes through in this case also, but we prefer to indulge ourselves this small simplification.

We may suppose, by composing it with an automorphism of the unit disc if necessary, that it vanishes precisely to the order 1 at some point p_0 of $\widetilde{\Sigma}$. We now introduce the function $t = -\log|h|$ defined on $\widetilde{\Sigma}$. This function has a singularity at p_0, as well as other singularities $\{p_i\}_{i\geq 1}$ forming a discrete subset of $\widetilde{\Sigma}$. These singularities are all of "logarithmic type": if z is a holomorphic coordinate in a some neighborhood of p_i, there exists an integer n_i such that $t(p) + n_i \log|z(p) - z(p_i)|$ is bounded. It is important to observe that $n_0 = 1$, that is, that the point p_0 is a simple logarithmic singularity (see Definition XI.1.1). Away from these singularities the function t is harmonic and positive.

By Lemma XI.2.1, there exists an exhaustion of $\widetilde{\Sigma}$ by means of simply connected, compact regions $D_0 \subset D_1 \subset \cdots$ with polygonal boundaries. We may also choose the D_k so that the ∂D_k contain no poles of t and $p_0 \in D_0$. By Theorem XI.1.6, for all positive integers k the interior of D_k supports a Green's function g_k with its logarithmic singularity at p_0.

Note next that the function sequence (g_k) is increasing, that is, g_{k+1} is greater than g_k on D_k. This is immediate from the fact that the difference $g_{k+1} - g_k$ is harmonic in the interior of D_k and positive on its boundary.

Furthermore, each of the functions g_k is bounded above by t. To see this, note that each difference $t - g_k$ is positive on the boundary of D_k, harmonic in the interior of D_k except for a finite number of points at which it has a logarithmic singularity, so that, by the minimum principle, $t - g_k$ attains its minimum on the boundary of D_k.

Thus the sequence $(g_k(p))$ has an upper bound independent of k provide p is not a pole of t. However, if p is a pole of t other than p_0, we still have such a bound, since $g_k(p)$ is bounded above independently of k by the maximum value of t on a small circle around p. Hence (g_k) converges simply to a function g on $\widetilde{\Sigma} \setminus \{p_0\}$ that is locally bounded. The theorem of dominated convergence, used in the Poisson formula expressing the values of g_k on a small disc as a function of those on the boundary, implies that g is harmonic on $\widetilde{\Sigma} \setminus \{p_0\}$ and that (g_k) converges uniformly to g relative to the C^∞ topology on the compact subsets of $\widetilde{\Sigma} \setminus \{p_0\}$. Since $g_k - g_0$ is harmonic on a small neighborhood V of p_0 with compact closure, we have that $|g_k - g_0|$ is bounded above on V by $\sup_{\partial V} |g_k - g_0|$. The latter quantity is in turn bounded above independently of k since on compact sets of $\widetilde{\Sigma} \setminus \{p_0\}$ convergence is uniform. Hence $g - g_0$ is bounded in a neighborhood of p_0, and g has a simple logarithmic singularity at p_0.

Choose a point $p_1 \in D_0 \setminus \{p_0\}$, and for each positive integer k consider the harmonic conjugate g_k^* of g_k on $D_k \setminus \{p_0\}$, defined by

$$g_k^*(p) = \int_{p_1}^{p} *dg_k.$$

Although this is a many-valued function, as we saw in the proof of Theorem XI.1.4, the function $G_k = e^{-(g_k + i g_k^*)}$ is a single-valued holomorphic function on $D_k \setminus \{p_0\}$ which extends holomorphically to p_0 and induces a biholomorphism from D_k to the unit disc. Since, relative to the C^1 topology, (g_k) converges to g uniformly on compact sets, the sequence (G_k) converges uniformly on compact sets to the holomorphic function $G = e^{-(g+ig^*)}$. This function is not constant in view of the fact that g has a simple logarithmic singularity at p_0. Hurwitz's theorem implies that, the G_k being injective, G also is injective. We have thus proved that G is a biholomorphism from $\widetilde{\Sigma}$ to a simply connected open subset of the unit disc. □

It is at this point that Poincaré calls a halt to his 1883 memoir. He does not know *a priori* what simply connected open subset of the disc it is that parametrizes the Riemann surface of the function $x \mapsto y(x)$ and furthermore he is careful not to say that it is biholomorphic to the unit disc. This is the case, however, although he had to wait till 1900 and Osgood's article [OsgW1900] for a proof.

Box XI.2: The universal covering space of a Riemann surface

The concept of the universal covering space is, at least implicitly, at the heart of the work of Klein and Poincaré on the uniformization of algebraic curves. However, as far as we know, Poincaré's memoir on the uniformization of functions contains the first explicit *definition* of the universal cover of a Riemann surface. Thus, for given germs of analytic functions $y_1, \dots, y_m$ of a variable x, Poincaré constructs the universal cover $\widetilde{S}$ of the Riemann surface associated with the germs $y_1, \dots, y_m$. He begins by asserting that $\widetilde{S}$ is to be a Riemann surface extended over the x-plane, and that $\widetilde{S}$ will be completely defined if one knows, for every loop C in the x-plane, under what condition the two ends of a lift of C will lie on the same sheet of $\widetilde{S}$. (Here it is implicit that this condition should depend only on C, and not on the chosen lift.) Poincaré goes on to distinguish two sorts of loops C:

1. those with the property that the continuation of at least one of the germs $y_1, \dots, y_m$ does not return to its initial value when the variable x traverses the loop C;
2. those for which the continuation of every germ $y_1, \dots, y_m$ returns to its initial value when the variable x traverses C.

He subdivides the loops of the second sort into two types:

1. loops of the first type are those that one can deform continuously to a point in such a way that throughout the deformation the loop remains one of the second sort;
2. the other loops of the second sort are then of the second type.

It only remains to note that the initial and final points of a lift of a loop C will lie on the same sheet of $\widetilde{S}$ if and only if C is of the second sort and the first type. This completely defines the surface $\widetilde{S}$.

R. Chorlay's thesis contains an interesting analysis of this construction; see [Cho2007, pp. 187–190].

XI.4. Osgood's theorem

The function t used in the proof of Theorem XI.3.1 resembles a Green's function on $\widetilde{\Sigma}$ although it does not tend to 0 at infinity and potentially admits an infinitude of poles (rather than just one). We owe it to Osgood to have noticed that the existence of such a function implies the existence of a "genuine" Green's function, and therefore, for a simply connected surface, of a biholomorphism to the unit disc ([OsgW1900]).

Definition XI.4.1. — Let S be a Riemann surface. A *Green's majorant* is a positive function $f : S \to \mathbb{R}$ with a discrete set of singularities, harmonic outside those singularities, tending to $+\infty$ in a neighborhood of each singularity, and such that at least one of its singularities, p_0 say, is of simple logarithmic type (that is, in terms of a holomorphic coordinate z in a neighborhood of p_0, the expression $f(p) + \log|z(p) - z(p_0)|$ is bounded).

Example XI.4.2. — A nonempty open set U contained in the unit disc possesses a Green's majorant. To see this, simply observe that, modulo composition with an automorphism of the disc, one may assume that U contains the point 0 and then $z \mapsto \log \frac{1}{|z|}$ is a Green's majorant on U.

Example XI.4.3. — If a Riemann surface admits a Green's majorant, then its universal covering space does also: it suffices to compose with the covering map. Hence, in view of the preceding example, the universal cover of a nonempty open set contained in the unit disc admits a Green's majorant.

Note that in general a Riemann surface may admit Green's majorants without admitting a genuine Green's function.

Example XI.4.4. — Let p be a point of the unit disc and consider the surface $S = \mathbb{D}\setminus\{p\}$. If there were a Green's function on S with logarithmic singularity at 0, then it would extend to a Green's function on $\mathbb{D}$ vanishing at p, contradicting the maximum principle; hence S does not admit a Green's function. In compensation, however, the restriction to S of the function $z \mapsto \log \frac{1}{|z|}$ is a Green's majorant on S.

We are now ready to state Osgood's theorem:

Theorem XI.4.5. — *Let S be a Riemann surface. If S admits a Green's majorant then its universal cover is biholomorphic to the unit disc.*

Hence in particular we have the

Corollary XI.4.6. — *Every simply connected open set of* $\mathbb{C}$ *admitting a Green's majorant is biholomorphic to the unit disc.*

Only this corollary appears explicitly in Osgood's article [OsgW1900]. To prove it, it suffices to obtain a parametrization by the unit disc in the proof of Theorem XI.3.1. However, Osgood's arguments establish the more general statement XI.4.5, and it was this version that Poincaré used in his proof of the general uniformization theorem in 1907.

Proof of Theorem XI.4.5. — Write t for the Green's majorant on S. By taking into account that the composite of t with the universal covering map is a Green's majorant on the universal cover, we may assume that S is simply connected.

We now reprise the notation and strategy of the proof of XI.3.1: we construct an exhaustion of S by means of simply connected regions D_k with compact closure and polygonal boundaries. We ensure that the D_k are such that their boundaries ∂D_k contain none of the singularities of t while D_0 contains a simple logarithmic singularity p_0 of t. We then consider the Green's functions g_k on the D_k, all with their logarithmic singularity at p_0, whose existence is guaranteed by Theorem XI.1.5.

We saw in the proof of XI.3.1 that the function sequence (g_k) tends uniformly on compact subsets of $S \setminus \{p_0\}$ to a positive harmonic function g with a simple logarithmic singularity at p_0. We shall prove that in fact g is a Green's function on S.

For each positive integer k and every Green's majorant t' having p_0 as a simple logarithmic singularity, the function $t' - g_k$ is positive on ∂D_k, and harmonic in the interior of D_k, apart from finitely many singularities in neighborhoods of which it tends to infinity. The maximum principle ensures that, for each k, $t' - g_k$ attains its minimum on the boundary of D_k. In other words, g_k is less than *every* Green's majorant on S, whence p_0 is a simple logarithmic singularity. And then the same holds for the function g.

Suppose that g does not tend to 0 at infinity. Then there exists a sequence (q_n) of points of S tending to infinity for which there is a number $\epsilon > 0$ such that $g(q_n) \geq \epsilon$. Consider the function $G = e^{-(g+ig^*)}$, the limit of the functions $G_k = e^{-(g_k+ig_k^*)}$. We saw in the preceding section that G is injective. The sequence $G(q_n)$ satisfies $|G(q_n)| \leq e^{-\epsilon} < 1$. Modulo choosing a suitable subsequence, we may therefore suppose that $G(q_n)$ converges to a point q of $\mathbb{D}$.

This point cannot belong to the image of G. For, if it did, then we could choose a preimage p of q under G and a relatively compact neighborhood U of p such that $G(U)$ is a neighborhood of q (since the non-constant holomorphic map G is open). Then almost all terms of the sequence $(G(q_n))$ would belong to $G(U)$ and the injectivity of G would guarantee that almost all the (q_n) were in U, contradicting the fact that the sequence (q_n) leaves every compact subset.

We now consider the universal cover of the punctured disc[7] $\pi : \mathbb{D} \to \mathbb{D} \setminus \{q\}$. Since S is simply connected, the map $G : S \to G(S) \subset \mathbb{D} \setminus \{q\}$ lifts to a map $\widetilde{G} : S \to \mathbb{D}$ such that $G = \pi \circ \widetilde{G}$. Then since the sequence $\widetilde{G}(q_n)$ leaves every compact set, we have $|\widetilde{G}(q_n)| \to 1$.

If F is the Green's function on the disc with logarithmic singularity at $\widetilde{G}(p_0)$, then the map $\widetilde{F} = F \circ \widetilde{G}$ affords a new Green's majorant on S with a simple logarithmic singularity at p_0 and satisfying $\widetilde{F}(q_n) \to 0$. However, since g is less than every Green's majorant on S, this contradicts the assumption $g(q_n) \geq \epsilon$ and completes the proof of the theorem. □

Note that Osgood's theorem implies in particular the Riemann Mapping Theorem.

Corollary XI.4.7. — *Every simply connected region D of the plane different from the whole plane, is biholomorphic to the unit disc.*

Proof. — By assumption, the boundary of D in $\mathbb{C}$ is nonempty and, since D is simply connected, this boundary contains at least two points. Modulo replacing D by its image under an affine transformation, we may assume that 0 and 1 are on the boundary of D. Now we know that there exists a holomorphic covering map $\pi : \mathbb{D} \to \mathbb{C} \setminus \{0, 1\}$: such a covering arises from the action on the upper half-plane of the Fuchsian group generated by $z \mapsto \frac{z}{2z+1}$ and $z \mapsto z + 2$. The inclusion ι of D in $\mathbb{C} \setminus \{0, 1\}$ then lifts to an injection $\tilde{\iota} : D \to \mathbb{D}$. We can then define a Green's majorant on D by inducing one from $\mathbb{D}$ via $\tilde{\iota}$, and apply Corollary XI.4.6 to conclude that D is biholomorphic to the unit disc. □

[7] It suffices to compose the universal covering map

$$\begin{array}{ccc} \mathbb{D} & \to & \mathbb{D} \setminus \{0\} \\ z & \mapsto & \exp\left(\frac{z+1}{z-1}\right) \end{array}$$

with an automorphism of the disc.

XI.5. The problem of branch points

By examining the proof of Theorem XI.3.1, one sees that the parametrization of the the surface determined by a function $x \mapsto y(x)$ admits critical points (branch points). The set E of such points is the preimage under x of the set of critical points of F. The set E is discrete in S. The surface $S \setminus E$ is parametrized, without branch points, by an open subset of the unit disc. This is how Poincaré formulates his result in his 1883 memoir. Theorem XI.3.1 is not entirely satisfactory since we would like to obtain a parametrization of the *whole* surface S. Osgood emphasized this point in the presentation of Poincaré's theorem that he made in the course of a series of talks he gave in 1898 ([OsgW1898]), and made suggestions as to what a "truly satisfactory" statement about uniformization might be. Hilbert also stressed, at the 1900 International Congress of Mathematicians, the problem presented by the presence of branch points in Poincaré's result. Here is how he formulated his 22nd problem:

> Wie Poincaré zuerst bewiesen hat, gelingt die Uniformisirung einer beliebigen algebraischen Beziehung zwischen zwei Variabeln stets durch automorphe Functionen einer Variabeln; d. h. wenn eine beliebige algebraische Gleichung zwischen zwei Variabeln vorgelegt ist, so lassen sich für dieselben stets solche eindeutigen automorphen Functionen einer Variabeln finden, nach deren Einsetzung die algebraische Gleichung identisch in dieser Variabeln erfüllt ist. Die Verallgemeinerung dieses fundamentalen Satzes auf nicht algebraische, sondern beliebige analytische Beziehungen zwischen zwei Variabeln hat Poincaré { Bulletin de la Société Mathématique de France XI. 1883 } ebenfalls mit Erfolg in Angriff genommen und zwar auf einem völlig anderen Wege als derjenige war, der ihn bei dem anfangs genannten speziellen Probleme zum Ziele führte. Aus Poincarés Beweis für die Möglichkeit der Uniformisirung einer beliebigen analytischen Beziehung zwischen zwei Variabeln geht jedoch noch nicht hervor, ob es möglich ist, die eindeutigen Functionen der neuen Variabeln so zu wählen, da, während diese Variabele das reguläre Gebiet jener Functionen durchläuft, auch wirklich die Gesamtheit aller regulären Stellen des vorgelegten, analytischen Gebildes zur Darstellung gelangt. Vielmehr scheinen in Poincarés Untersuchungen, abgesehen von den Verzweigungspunkten, noch gewisse andere im Allgemeinen unendlichviele diskrete Stellen vorgelegten analytischen Gebildes ausgenommen zu sein, zu denen man nur gelangt, indem man die neue Variable gewissen Grenzstellen der Functionen nähert. Eine Klärung und Lösung dieser Schwierigkeit scheint mir in Anbetracht der fundamentalen Bedeutung der Poincaréschen Fragestellung äußerst wünschenswert.

Here is an approximate translation:

> As Poincaré was the first to prove, it is always possible to uniformize any algebraic relation between two variables by means of automorphic functions of a single variable. That is, given an algebraic equation in two variables, one can always express the latter as functions of a third variable so that, after substituting, the algebraic relation holds identically. Poincaré also tackled with success the generalization of this fundamental theorem to any relations between two variables that are not algebraic but just analytic, and this by using methods completely different from those he brought to the solution of the first problem mentioned. From Poincaré's proof of the possibility of uniformizing an arbitrary analytic relation between two variables, it still does not always follow that it is possible to choose the single-valued functions of the new variable so that as said variable ranges over its domain of definition, the totality of regular points of the analytic surface under consideration are obtained. On the contrary, there appear in Poincaré's investigations, apart from the branch points, certain other points, which in general comprise an infinite discrete part of the surface in question, and which cannot be reached except by letting the new variable tend to certain points on the boundary of its domain of definition. In view of the fundamental importance attaching to Poincaré's problem, it seems to me that a clarification and a resolution of this difficulty would be most desirable.

Chapter XII

Koebe's proof of the uniformization theorem

When he produced his masterful solution of Hilbert's twenty-second problem, Koebe was a 25-year-young mathematician who just two years before had defended a thesis under the supervision of Schwarz and Schottky. The result in question, nowadays called *The Uniformization Theorem*, was presented to the Göttingen Scientific Society by Klein on May 11, 1907 [Koe1907a]:

Theorem XII.0.1. — *Every simply connected Riemann surface is biholomorphic to the Riemann sphere, the complex plane, or the unit disc in* $\mathbb{C}$.

Remark XII.0.2. — As always at the time in question, the Riemann surfaces considered by Koebe were assumed extended over the plane. However, his proof of the uniformization theorem goes through without change for an abstract Riemann surface.

XII.1. Idea of the proof

Let S be a simply connected Riemann surface. We wish to show that S is biholomorphic to the Riemann sphere, to $\mathbb{C}$, or to the unit disc. Since Schwarz had already proved that a *compact* simply connected Riemann surface is biholomorphic to the Riemann sphere, Koebe could assume that S is not compact. Consider an exhaustion $(D_k)_{k\in\mathbb{N}}$ of S by means of simply connected regions with compact closures and polygonal boundaries, and such that for every k, the inclusion $\overline{D}_k \subset D_{k+1}$ holds (see Lemma XI.2.1 for the existence of such an exhaustion). Choose a point p_0 in the interior of D_0, and denote by (g_k) the sequence of Green's functions associated with (D_k), each with its logarithmic singularity at p_0, the existence of this sequence of functions being guaranteed by Theorem XI.1.5. Now fix on another point p_1 in D_0 and consider "the" harmonic conjugate g_k^* of g_k, defined on $D_k \setminus \{p_0\}$ by

$$g_k^*(p) = \int_{p_1}^{p} * \, dg_k .$$

Then as we saw in the proof of Theorem XI.1.4 the function $G_k := e^{-(g_k + i g_k^*)}$ is single-valued and realizes a biholomorphism from D_k to the unit disc. Koebe's approach consists in a (more or less) direct investigation of the convergence of the G_k, following possible renormalization. One of the most interesting ideas of his proof is that of regulating the size of the image of the unit disc under a holomorphic injection in terms of the modulus of the derivative at 0. This forms the essential content of Lemma XII.3.2, which is a slightly more primitive version of what is now called the "Koebe Quarter Theorem" (see for example [Pom1975]):

Lemma XII.1.1. — *If $f : \mathbb{D} \to \mathbb{C}$ is a holomorphic injection of the unit disc fixing the origin, then the image $f(\mathbb{D})$ contains the disc of radius $|f'(0)|/4$ centered at the origin.*

Consider a holomorphic chart $z : U_0 \to \mathbb{C}$ defined in a neighborhood U_0 of p_0 and centered at p_0 (that is, $z(p_0) = 0$). Each function g_k has a simple logarithmic singularity at p_0 (see Definition XI.1.1); in other words, there is a neighborhood of p_0 in which it has a development of the form

$$g_k(z) = \log \frac{1}{|z|} + c_k + o(1).$$

For all k the function $g_{k+1} - g_k$ is harmonic on D_k, extends continuously to $\overline{D}_k$, and is positive on ∂D_k. Hence by the maximum principle it is positive on D_k. It follows in particular that its value at p_0, namely $c_{k+1} - c_k$, is positive, so that the sequence (c_k) is strictly increasing. Moreover, one verifies that on the chart U_0 one has $|G_k'(p_0)| = e^{-c_k}$. Koebe now splits his proof into two cases: that where the sequence (c_k) converges to a finite limit c_∞, and that where c_k tends to $+\infty$.

XII.2. The case where the sequence (c_k) is bounded

We are assuming here that c_k tends to a real number c_∞. We have seen above that on every relatively compact open set D of S, the sequence (g_k) is, from some k on, defined, harmonic, and strictly increasing on $D \setminus \{p_0\}$. On our open set with chart U_0 about p_0, we consider the sequence of functions $u_k := g_k - \log\left(\frac{1}{|z|}\right)$. This is an increasing sequence of harmonic functions for which $u_k(p_0) = c_k$ forms a convergent sequence. We may therefore apply Harnack's principle, whose statement and proof we now recall.

Theorem XII.2.1 (Harnack's principle). — *Let Ω be a connected open subset of a Riemann surface and (u_k) an increasing sequence of harmonic functions defined on Ω. Then (u_k) converges uniformly on every compact subset of Ω, either to $+\infty$ or to a harmonic function.*

Proof. — Since every compact subset of Ω can be covered by open subsets biholomorphic to $\mathbb{D}$ by means of biholomorphisms extending to homeomorphisms on the boundary, we need consider only the case where (u_k) is an increasing sequence of functions continuous on $\overline{\mathbb{D}}$ and harmonic on $\mathbb{D}$.

Harnack's principle is then a consequence of the following result:

Proposition XII.2.2 (Harnack's inequality). — *Corresponding to each compact subset K of $\mathbb{D}$ there is a constant $C_K > 0$ such that for every continuous, positive function $u : \overline{\mathbb{D}} \to \mathbb{R}$ harmonic on $\mathbb{D}$, one has*

$$\forall (x,y) \in K^2 \qquad u(x) \leq C_K\, u(y). \tag{XII.1}$$

Proof. — Observe first that there exists a constant $C_K > 0$ such that (XII.1) holds for every function u of the form

$$P_\theta : z \mapsto \frac{1-|z|^2}{|1-e^{-i\theta}z|^2},$$

where $\theta \in [0, 2\pi]$. To see this, it suffices to note that the map

$$(x,y,\theta) \mapsto \frac{P_\theta(x)}{P_\theta(y)}$$

is bounded on $K^2 \times [0,2\pi]$. Now let $u : \overline{\mathbb{D}} \to \mathbb{R}$ be continuous, positive, and harmonic in $\mathbb{D}$. Poisson's formula gives

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} P_\theta(z)\, u(e^{i\theta}) d\theta.$$

Hence for every $(x,y) \in K^2$ we have

$$u(x) = \frac{1}{2\pi}\int_0^{2\pi} P_\theta(x)\, u(e^{i\theta}) d\theta \leq C_K\, \frac{1}{2\pi}\int_0^{2\pi} P_\theta(y)\, u(e^{i\theta}) d\theta = C_K\, u(y).$$

□

To deduce Harnack's principle from his inequality, one begins by reducing to the case where the functions u_k are positive: to this end it suffices to consider instead the sequence $(u_k - u_0 + 1)$, for example. This assumed, since the sequence (u_k) is increasing, it tends to a Borel function $u : \overline{\mathbb{D}} \to \mathbb{R} \cup \{+\infty\}$. We then have a dichotomy according to whether $u(0) = +\infty$ or $u(0) \in \mathbb{R}$.

In the first case we have, for every compact subset K of $\mathbb{D}$ containing 0 and for every natural number k,

$$\min_K u_k \geq u_k(0)/C_K,$$

so that the sequence (u_k) tends to $+\infty$ uniformly on every compact subset. In the second case, we have, for every compact subset K of $\mathbb{D}$ containing 0 and every k,

$$\max_K u_k \leq C_K\, u_k(0) \leq C_K\, u(0),$$

showing that the function u is bounded on every compact subset of $\mathbb{D}$. Hence by the dominated convergence theorem u has the mean-value property:

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} u(z + re^{i\theta}) d\theta$$

for all z in $\mathbb{D}$ and all r satisfying $0 < r < 1-|z|$. It follows that u is harmonic on $\mathbb{D}$. To see this it suffices to use a regularizing kernel ρ_n: the regularized function $\rho_n \star u$ is smooth and has the mean-value property, and is therefore harmonic. But then $\rho_n \star u = u$ by the mean-value property. Thus in particular the function u is continuous, and we conclude by Dini's theorem that (u_k) converges uniformly to u on every compact subset of $\mathbb{D}$. □

Returning to the case of the uniformization theorem we are at present concerned with, we note that by Harnack's principle, the function sequence (u_k) converges uniformly on every compact subset of $U_0 \setminus \{p_0\}$. In particular, for any p in U_0 other than p_0, the sequence $(g_k(p))$ is bounded. Hence in this case we may apply Harnack's principle to (g_k) to infer that (g_k) converges uniformly on every compact subset $S \setminus \{p_0\}$ to a function g that is harmonic, positive, and has a simple logarithmic singularity at p_0. The function g is therefore a Green majorant on S, and Theorem XI.4.5 then tells us that S is biholomorphic to the unit disc. □

XII.3. The case where the sequence (c_k) tends to infinity

The sequence $\left(g_k - \log \frac{1}{|z|}\right)$ is an increasing sequence of harmonic functions defined on the open set with chart U_0, and with kth term taking the value c_k at p_0. We are now assuming that c_k tends to $+\infty$. In this case, Harnack's inequality implies that the sequence $\left(g_k - \log \frac{1}{|z|}\right)$ diverges to infinity on a neighborhood of p_0, so that this is *a fortiori* the case for (g_k). Hence by Harnack's principle g_k tends uniformly to infinity on the compact subsets of S, or, what amounts to the same thing, G_k tends uniformly to 0 on compact subsets of S since $|G_k| = e^{-g_k}$.

It therefore natural to consider the sequence of functions $F_k = e^{c_k} G_k$ with derivatives at p_0 equal to 1 on our chart $z : U_0 \to \mathbb{C}$. Koebe proves the following proposition:

Proposition XII.3.1. — *For all $k, m \geq 0$, the image of D_k under F_{m+k} contains a disc centered at* 0 *and of radius Ce^{c_k} where C is a positive universal constant.*

Proof. — Koebe's idea is that a holomorphic embedding of the unit disc whose derivative at 0 has large modulus should contain large discs. More precisely:

Lemma XII.3.2. — *Let $\varphi : \mathbb{D} \to \mathbb{C}$ be a holomorphic injection fixing* 0*. If R denotes the largest real such that the disc of center* 0 *and radius R is contained in $\varphi(\mathbb{D})$ then*

$$R \geq C|\varphi'(0)|,$$

where C is a positive constant independent of φ.

Proof. — One chooses a point $Re^{i\theta}$ on the boundary of $\varphi(\mathbb{D})$ and considers the function

$$\psi(w) := \frac{\varphi(w)}{Re^{i\theta}}.$$

The open set $\psi(\mathbb{D})$ contains the unit disc but not the point 1. Write r for holomorphic determination of the square root restricted to the disc of radius 1 centered at -1, taking the value i at -1. For every complex number w in the image of r, its negative $-w$ is not in that image. Hence there is a neighborhood U of $-i$ on the Riemann sphere that does not intersect the image of the function r and for which there is a biholomorphism $k : \widehat{\mathbb{C}} \setminus U \to \mathbb{D}$ satisfying $k(i) = 0$. Write $h(w) = r(\psi(w) - 1)$. The function $k \circ h$ is then a holomorphic map of $\mathbb{D}$ to itself, sending 0 to 0. It follows from Schwarz's lemma that

$$|k'(i)||h'(0)| \leq 1.$$

Furthermore, from $h^2 = \psi - 1$ one infers that

$$|h'(0)| = \frac{|\psi'(0)|}{2} = \frac{|\varphi'(0)|}{2R},$$

whence $R \geq |k'(i)||\varphi'(0)|$. (Note that $|k'(i)|$ depends only on r and the choice of U, so is independent of φ). This establishes the claim. □

One calculates without difficulty that

$$|(F_{k+m} \circ G_k^{-1})'(0)| = e^{c_k}.$$

The proposition now follows directly from the preceding lemma since $F_{k+m} \circ G_k^{-1}$ is a holomorphic embedding of $\mathbb{D}$ into $\mathbb{C}$ fixing 0 and with image precisely $F_{k+m}(D_k)$. □

Using this proposition one would like to show directly that the sequence (F_k) converges to a biholomorphism from S to $\mathbb{C}$. However, Koebe is unable to make this direct approach work, so he considers instead the sequence of functions

$$K_k = e^{-c_k}\left(\frac{1}{G_k} - G_k\right),$$

which he investigates utilizing what he knows about the functions F_k. The transformation $w \mapsto \frac{1}{w} - w$ sends the unit disc biholomorphically onto the complement, in $\mathbb{CP}^1$, of the segment $[-2i, 2i] \subset i\mathbb{R}$. Hence K_k sends D_k onto the complement in $\mathbb{CP}^1$ of the segment $[-2ie^{-c_k}, 2ie^{-c_k}]$. On the chart $z : U_0 \to \mathbb{C}$, the function K_k admits a series development of the form

$$K_k = \frac{1}{z} + O(1).$$

Koebe's idea is now to use Proposition XII.3.1 not to establish the convergence of the sequence (K_k) directly, but rather to show the convergence of the real parts $U_k = \mathrm{Re}(K_k)$ and to show that this suffices to yield the uniformization theorem. Thus suppose the sequence of functions $U_k : D_k \setminus \{p_0\} \to \mathbb{R}$ does indeed converge uniformly on compact subsets of $S \setminus \{p_0\}$. Poisson's formula then shows that the partial derivatives of the U_k also converge uniformly on compact subsets of $S \setminus \{p_0\}$. Hence, in particular, for any point p_1 of D_0 other than p_0, the single-valued functions

$$U_k^* : p \mapsto \int_{p_1}^{p} *dU_k$$

on $D_k \setminus \{p_0\}$ converge uniformly on compact subsets of S. By construction, $U_k + iU_k^*$ is equal to $K_k + a_k$ for a certain purely imaginary number a_k and, with this choice of a_k, the sequence of functions $(K_k + a_k)$ converges uniformly on compact subsets of S to a holomorphic function $K : S \to \mathbb{CP}^1$ with a simple pole at p_0. Hence K is certainly not constant, whence, by Hurwitz's theorem, it is injective. The open subset $K(S)$ of $\mathbb{CP}^1$ is simply connected. Koebe now argues that the the boundary of $K(S)$ consists of a single point, so that S must in fact be biholomorphic to the complex plane. In fact, in view of the simple connectedness of $K(S)$, if its boundary contained more than one point it would have to contain infinitely many, and in this case, by Corollary XI.4.7, $K(S)$, and therefore S, would be biholomorphic to the unit disc. This leads to a contradiction, however, since then S would admit a Green majorant g with a simple logarithmic singularity at p_0, with the consequence that the c_k are bounded.

It now only remains to show how Koebe establishes the convergence of the sequence (U_k). The functions U_k are characterized by their harmonicity on

$D_k \setminus \{p_0\}$, their convergence to 0 on the boundary of D_k, and their having a development in a neighborhood of p_0 of the form $U_k(z) = \mathrm{Re}\left(\frac{1}{z}\right) + O(1)$. The uniform convergence of the sequence (U_k) on compact subsets of $S \setminus \{p_0\}$, is established by deducing from these properties that it is uniformly Cauchy.

Lemma XII.3.3. — *For all $k, m \geq 0$, the following inequality holds on the region D_k:*

$$|U_{k+m} - U_k| \leq \left(1 + \frac{1}{C}\right) e^{-c_k}. \tag{XII.2}$$

Proof. — On $F_{k+m}(D_{k+m}) = D(0, e^{c_{k+m}})$, the function

$$V_{k+m}(w) := U_{k+m} \circ F_{k+m}^{-1}(w) - \mathrm{Re}\left(\frac{1}{w}\right)$$

is harmonic and extends to a continuous function on $\overline{D}(0, e^{c_{k+m}})$. The function

$$V_k(w) := U_k \circ F_{k+m}^{-1}(w) - \mathrm{Re}\left(\frac{1}{w}\right)$$

is harmonic on $F_{k+m}(D_k)$ and extends continuously to the boundary $\overline{F_{k+m}(D_k)}$. On the boundary of $D(0, e^{c_{k+m}})$, one has

$$V_{k+m}(w) = -\mathrm{Re}\left(\frac{1}{w}\right), \quad \text{whence} \quad |V_{k+m}| \leq e^{-c_{m+k}} \leq e^{-c_k}.$$

By the maximum principle this inequality is then valid on $D(0, e^{c_{k+m}})$, and so, in particular, on $F_{k+m}(D_k)$. Furthermore, by Proposition XII.3.1, $F_{k+m}(D_k)$ contains a disc centered at 0 of radius Ce^{c_k}. Hence on the boundary of $F_{k+m}(D_k)$, where $V_k(w) = -\mathrm{Re}\left(\frac{1}{w}\right)$, one has

$$|V_k| \leq \frac{1}{C} e^{-c_k},$$

and one concludes finally that, on $F_{k+m}(D_k)$,

$$|V_{m+k} - V_k| \leq 2e^{-c_k},$$

yielding the desired inequality (XII.2). □

The sequence (U_k) is thus uniformly Cauchy on compact subsets of $S \setminus \{p_0\}$. As explained earlier, this implies that the surface S is biholomorphic to the complex plane, and the proof of Theorem XII.0.1 is complete.

Remark XII.3.4. — The dichotomy $(c_k \to c_\infty)$ versus $(c_k \to +\infty)$ allows one to distinguish between surfaces S biholomorphic to the unit disc and those biholomorphic to the complex plane. Note also that, instead of introducing — admittedly very cleverly — the functions K_k, it would have been preferable to establish the desired result by working directly with the renormalized maps F_k, if possible. It turns out that this can indeed be done using Lemma XII.3.2., which tells us that for all $k, m \geq 0$, $F_{k+m} \circ F_k^{-1}(\mathbb{D})$ contains the unit disc. In fact, in the proof of that lemma, we saw in effect that, setting $h_m(z) = r(F_{k+m} \circ F_k^{-1} - 1)$, we have that the complement of $h_m(D_k)$ contains a neighborhood U of $-i$ independent of m and k. It follows that the sequence $\left(\frac{1}{h_m+i}\right)_{m\in\mathbb{N}}$ is bounded in modulus and therefore constitutes a normal family. Hence there is a subsequence of (h_m) that converges uniformly on compact subsets of D_k. It follows, finally, that (F_k) possesses a subsequence converging uniformly on compact subsets of S to a holomorphic function $F : S \to \mathbb{C}$. This function cannot be constant since its derivative at p_0 is 1, so by Hurwitz's theorem it is injective. Then one concludes as before, via Corollary XI.4.7, that $F(S) = \mathbb{C}$.

Note that it was in June 1907, so almost simultaneously with Koebe's note, that P. Montel defended his thesis entitled *On infinite sequences of functions* [Mon1907], in which he defined the notion of a normal family.

Chapter XIII

Poincaré's proof of the uniformization theorem

This last chapter is based essentially on the memoir [Poin1907], written at the beginning of 1907 and published in November of that year, in which Poincaré gives his own proof of the uniformization theorem XII.0.1, one based on a generalization of Schwarz's alternating procedure which he calls the "sweeping-out method".[1] The first two sections of this chapter are devoted to an exposition of this proof. In the third section, we present the contents of the note [Koe1907b], in which Koebe revisits Poincaré's proof, simplifying it significantly.

Throughout the chapter we will be concerned with a simply connected Riemann surface S. Theorem XII.0.1 states that S is biholomorphic to the Riemann sphere, the complex plane, or the unit disc, so this is what we have to prove. The compact case was dealt with earlier: according to Theorem IV.2.1, if S is compact then it is biholomorphic to the Riemann sphere. Thus it will be assumed henceforth in this chapter that S is non-compact (and the aim will then be to show that S is biholomorphic to the complex plane or the unit disc). For any subset X of S, we shall, as earlier, denote by $\overline{X}$ the closure of X in S and by ∂X the boundary of X in S.

XIII.1. Strategy of the proof

Poincaré's strategy for proving the uniformization theorem rests on Osgood's theorem XI.4.5, affirming that the universal cover of a non-compact Riemann surface admitting a Green's majorant is biholomorphic to the unit disc. He attempts to construct a Green's majorant on the Riemann surface A obtained by removing a

[1]Sometimes translated into English as the "scanning method". The original French word "balayage" is also used in English texts. (See also e.g. Chapter XI.) *Trans*

small disc $\overline{\Delta}$ from S. His procedure is, more precisely, as follows. He first chooses a holomorphic chart $z : U \to \mathbb{C}$ defined on an open subset U of S. Up to following z by a suitable automorphism of $\mathbb{C}$, one may suppose $z(U)$ contains the unit disc in $\mathbb{C}$. Poincaré then chooses a real number $r \in (0,1)$ and sets

$$\Delta := \{p \in S \mid |z(p)| < r\} \qquad \text{and} \qquad A := S \setminus \overline{\Delta}.$$

Observe that A is an open subset of the surface S homeomorphic to an annulus, and that the boundary of A in S is the curve $\partial A = \partial\Delta = \{p \in S \mid |z(p)| = r\}$. The construction of a Green's majorant on A will form the subject-matter of section XIII.2; here we shall explain why the existence of such a Green's majorant implies that S is biholomorphic to the complex plane $\mathbb{C}$ or the unit disc $\mathbb{D}$.

Proposition XIII.1.1. — *If A admits a Green's majorant, then it is biholomorphic to the complex plane or the unit disc.*

Proof. — We begin by showing that the fundamental group of A is isomorphic to $\mathbb{Z}$.[2] Choose $r' \in (r,1)$ and write

$$\Delta' := \{p \in S \mid |z(p)| < r'\}.$$

The existence of an exhaustion of S by means of topological discs with smooth boundaries (see the proof of Lemma XI.2.1) implies that the inclusion of $\partial\Delta'$ in A induces a surjection from $\pi_1(\partial\Delta')$ onto $\pi_1(A)$, which is therefore cyclic. Furthermore, by Theorem XI.4.5 the universal cover of A is biholomorphic to the unit disc $\mathbb{D}$, whence A is biholomorphic to the quotient of the disc $\mathbb{D}$ by a cyclic subgroup Γ of $\mathrm{Aut}(\mathbb{D})$. If Γ were trivial, the surface A would be a disc and S would be the union of two relatively compact subsets so compact. Hence the group Γ is generated by an element γ acting fixed-point free on the disc, so necessarily parabolic or hyperbolic and of infinite order. (For the classification of the automorphisms of the disc, see §VI.1.1.)

It is now easy to see that there is a biholomorphism h sending A to a plane annulus of the form $\mathbb{A}(r_1,r_2) = \{w \in \mathbb{C} \mid r_1 < |w| < r_2\}$ where $0 \le r_1 < r_2 < \infty$ (more precisely $0 = r_1 < r_2 < \infty$ if γ is parabolic and $0 < r_1 < r_2 < \infty$ if γ is hyperbolic). Write $A' := S \setminus \overline{\Delta'}$. For sufficiently small $\varepsilon > 0$, A' contains either the annulus $h^{-1}(\mathbb{A}(r_1, r_1+\varepsilon))$ or the annulus $h^{-1}(\mathbb{A}(r_2-\varepsilon, r_2))$. We distinguish three cases.

First case: A' contains $h^{-1}(\mathbb{A}(r_2-\varepsilon, r_2))$. If one attaches the annulus $\mathbb{A}(r_2-\varepsilon, r_2+\varepsilon)$ to the surface S by identifying the subset $\mathbb{A}(r_2-\varepsilon, r_2)$ of it with $h^{-1}(\mathbb{A}(r_2-\varepsilon, r_2))$, one obtains a Riemann surface $\widehat{S}$ of which S is a simply

[2] As mentioned in the introduction to Part C, Poincaré and Koebe consider it obvious that S is homeomorphic to the plane, so that $A = S \setminus \Delta$ is homeomorphic to an annulus.

connected region with compact closure and analytic boundary. Hence by Corollary XI.1.6, one concludes in this case that S is biholomorphic to the unit disc.

Second case: A' contains $h^{-1}(\mathbb{A}(r_1, r_1+\varepsilon))$ and $r_1 > 0$. If one attaches the annulus $\mathbb{A}(r_1-\varepsilon, r_1+\varepsilon)$ to the surface S by identifying its subset $\mathbb{A}(r_1, r_1+\varepsilon)$ with $h^{-1}(\mathbb{A}(r_1, r_1+\varepsilon))$, one obtains a Riemann surface $\widehat{S}$ in which S is again a simply connected region with compact closure and analytic boundary. As in the first case we conclude via Corollary XI.1.6 that S is biholomorphic to the unit disc.

Third case: A' contains $h^{-1}(\mathbb{A}(r_1, r_1+\varepsilon))$ and $r_1 = 0$. If one attaches the disc $\mathbb{D}(0, r_1+\varepsilon) = \{w \in \mathbb{C} \mid |w| < r_1+\varepsilon\}$ to the surface S by identifying its subset $\mathbb{A}(r_1, r_1+\varepsilon)$ with the annulus $h^{-1}(\mathbb{A}(r_1, r_1+\varepsilon))$, one obtains a Riemann surface $\widehat{S}$ homeomorphic to the sphere[3]. By Theorem IV.2.1, $\widehat{S}$ is then biholomorphic to the Riemann sphere. The surface S is the complement of a point in $\widehat{S}$, so is biholomorphic to the Riemann sphere with a point removed, that is, to the complex plane. □

In view of this proposition, the proof of the uniformization theorem XII.0.1 reduces to establishing the existence of a Green's majorant on the annulus A.

XIII.2. The existence of a Green's majorant on the annulus A

We shall now explain how Poincaré proves the existence of a Green's majorant on A. His proof uses a generalization of *Schwarz's alternating procedure* called the *sweeping-out method*, for which he provides an electrostatic interpretation. The implementation of this method, especially the justification of its convergence, presents problems of analysis whose difficulty Poincaré fully appreciates but which were only solved a half-century later with the advent of the theory of distributions. Lacking these tools, Poincaré relies on physical arguments. In the proof we give here we expound the ideas from [Poin1907], occasionally supplementing them with anachronistic arguments in order to render them rigorous.

XIII.2.1. Presentation of the sweeping-out method

This method rests on an electrostatic analogy: given a distribution of charges on a surface covered by discs, one lets the discs successively become conducting, which has the effect of "sweeping out" the charges, that is, causing them to move to the boundary of their disc without altering the potential except in the interior

[3]Actually setting $r_1 = 0$ may make this clearer. *Trans*

of the disc, the total charge being conserved. One hopes that by iterating this procedure, the charges will be ultimately "dispersed" to infinity, leaving a perfectly harmonic potential on the Riemann surface.

In order to construct a Green's function, one begins with a charge distribution consisting of a single positive point charge and some distribution of negative charges. The procedure is then modified so as to preserve the positive point charge. In other words, only the negative charges are subject to sweeping out.

Here is what the method amounts to mathematically speaking. One considers a Riemann surface S and a point p_0 of S. Each point p of S has a neighborhood D for which there is a biholomorphism from D to the open unit disc $\mathbb{D}$ of $\mathbb{C}$, extending to a homeomorphism from $\overline{D}$ to $\overline{\mathbb{D}}$ (so that, in particular, D has compact closure). Every continuous function $v : \partial D \to \mathbb{R}$ has an *harmonic extension* to $\overline{D}$, that is, a function $\bar{v} : \overline{D} \to \mathbb{R}$ continuous on $\overline{D}$ and harmonic in D. (This function is obtained via convolution of v with the Poisson kernel after sending $\overline{D}$ to the closed unit disc.) Furthermore, if p_0 is in D, then there is a Green's function on D with a simple logarithmic singularity at p_0.

Let u be a continuous real-valued function defined on $S \setminus \{p_0\}$ with a simple logarithmic singularity at p_0. The *scan*[4] on D of the function u is the function $B(u, D)$ defined as follows: $B(u, D)$ is equal to u outside D; if D does not contain p_0, then $B(u, D)_{|D}$ is the harmonic extension of $u_{|\partial D}$; and finally if D contains p_0, then $B(u, D)_{|D}$ is to be the sum of the harmonic extension of $u_{|\partial D}$ and the Green's function on D with a simple logarithmic singularity at p_0. Thus the possible logarithmic singularity of u is preserved.

The *sweeping-out process* consists in repeating this operation infinitely often on a family $\mathcal{R}$ of discs covering the surface S and not containing the point p_0 on their boundaries. It is important that each disc be swept out infinitely often. If $\mathcal{R} = \{D_1, D_2 \dots\}$, one might, for instance, sweep out the discs in the order D_1, D_2, D_1, D_2, D_3, D_1, D_2, D_3, $D_4 \dots$.

Starting with a positive continuous function $u_0 : S \setminus \{p_0\} \to \mathbb{R}$ with a simple logarithmic singularity at p_0, the scanning process determines a sequence of functions $(u_n)_{n \geq 0}$ via the recurrence relation

$$u_{n+1} := B(u_n, D_{n+1}).$$

Box XIII.1: Electrostatic interpretation of sweeping-out

We begin by recalling certain facts from electrostatics. A planar charge

[4] Or "sweeping-out" or "balayage". *Trans*

distribution q of compact support gives rise to an electric field $\vec{E}$ given by:

$$\vec{E}(x) := \int_{\mathbb{R}^2} \frac{\overrightarrow{xy}}{\|\overrightarrow{xy}\|^2} dq(y). \tag{XIII.1}$$

The flux of the field $\vec{E}$ across an arbitrary simple closed curve C is equal to the total charge contained in the region D bounded by C:

$$\int_C \langle \vec{E}, \vec{n} \rangle = \int_D q, \tag{XIII.2}$$

where $\vec{n}$ is the outward-directed unit normal vector field along C. Green's theorem yields another expression for the left-hand integral:

$$\int_C \langle \vec{E}, \vec{n} \rangle = \int_D \operatorname{div} \vec{E}. \tag{XIII.3}$$

Furthermore, the field $\vec{E}$ can be considered as arising from a potential u, that is, a real-valued function for which $E = -\vec{\nabla} u$. The potential is *a priori* well defined only up to an additive constant. The equations (XIII.2) and (XIII.3) imply that the potential satisfies

$$\Delta u = \operatorname{div} \vec{\nabla} u = -\operatorname{div} \vec{E} = -q. \tag{XIII.4}$$

The elementary potential function $u(z) = -\log|z|$ yields for every (small) circle $C = C(0, r)$:

$$\begin{aligned} \int_C \langle \vec{E}, \vec{n} \rangle &= r \int_0^{2\pi} \tfrac{\partial u}{\partial n}(re^{i\theta}) d\theta \\ &= r \int_0^{2\pi} -\tfrac{d\theta}{r} \\ &= -2\pi. \end{aligned}$$

The Laplacian Δu corresponds intuitively to $-2\pi\delta_0$, where δ_0 is a Dirac mass at 0. This represents a relatively recent analytic formulation of a much older intuition concerning a positive point charge.

Imagine now a charge distribution including a positive point charge at p_0. Such a distribution is described by a potential u with a logarithmic singularity at p_0. We now carry out the following thought experiment: we allow a small disc D in the plane to "become conducting".

If the disc D contains p_0, we only allow $D \setminus \{p_0\}$ to become conducting. This has the effect of "sweeping" the electric charges contained in D (or in $D \setminus \{p_0\}$, as the case may be) out towards the boundary of D. After having allowed the charge distribution to evolve to equilibrium, we obtain a new distribution q', arising from a potential u'. The potential u' is harmonic in D (resp. $D \setminus \{p_0\}$) since there are no longer any charges in D (resp. $D \setminus \{p_0\}$). On the complement of $\overline{D}$, we will have $\Delta u' = \Delta u$ since the charges outside $\overline{D}$ have not moved, the complement of $\overline{D}$ being supposed to have remainednon-conducting. Since u and u' both tend to 0 at infinity, it follows that u' is precisely the scan[a] of u on D, that is, $u' = B(u, D)$.

Thus a sweeping-out process on a Riemann surface S "corresponds" to the following electrostatic experiment (which is clearly only a "thought experiment" if S is not an open subset of the plane): Starting from a charge distribution including a positive point charge at p_0, we let the members of a family of discs covering the surface become conducting one after the other.

[a] Or "sweeping-out" or "balayage". *Trans*

In order to see more clearly the importance of sweeping-out for the construction of Green's majorants, suppose for a moment that the sequence (u_n) converges uniformly on compact subsets of $S \setminus \{p_0\}$ to a function u. Then $(u_n - u_0)$ converges uniformly on compact subsets of $S \setminus \{p_0\}$ to a function v, say. Letting D denote any disc in $\mathcal{R}$ and $u_{\varphi(n)}$ be terms of a subsequence satisfying $u_{\varphi(n)} = B(u_{\varphi(n)-1}, D)$, we then have that each function $u_{\varphi(n)} - u_0$ is harmonic on D, whence, since uniform passage to the limit preserves the mean value property, we shall have that v is also harmonic. Thus u is harmonic on $S \setminus \{p_0\}$ with a simple logarithmic singularity at p_0. It is readily verified that if we start with a positive function u_0, all the terms u_n will likewise be positive. Hence the function u is indeed a Green's majorant on $S \setminus \{p_0\}$.

XIII.2.2. Growth in the sweeping-out process and convergence criteria

Our aim is now clear: we have to make some sweeping-out process converge. A appropriate means to this end involves beginning the process with a *subharmonic* function.

Box XIII.2: Subharmonic functions

Let U be an open subset of $\mathbb{C}$. A continuous function $u : U \to \mathbb{R}$ is called *subharmonic* if the value of u at each point of U is less than or equal to the average of u on every circle centered at that point, that is, for all $x_0 \in U$ and $r > 0$,

$$u(x_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(x_0 + re^{i\theta}) d\theta. \tag{XIII.5}$$

Using Green's theorem, one infers that, provided the function u is smooth, this property is equivalent to the condition that the Laplacian Δu of u be non-negative. In fact, it is more natural to introduce the operator $d^c = \star d$ acting on functions and to consider the 2-form $dd^c u$, which is invariant under holomorphic coordinate changes. Recall that the Hodge operator $\star$, acting on 1-forms via the formula $\star\alpha(\xi) = -\alpha(i\xi)$, is a conformal invariant. This is also the case for dd^c, since on a chart $x = a + ib$, one may write

$$dd^c u = \Delta u \; da \wedge db.$$

Furthermore, this 2-form has an electrostatic interpretation: given a negative charge distribution defined by a 2-form μ, the potential associated with this charge distribution is a function u satisfying $-dd^c u = \mu$. From this it follows that a smooth function u is subharmonic if and only if $dd^c u$ is a positive 2-form.

However, in the sweeping-out process of interest to us, we shall be considering subharmonic functions that are not differentiable. An example typical of those we shall encounter is the maximum of two harmonic functions; for such functions we would like to define the "corresponding repartition of charges". Thus we seek to define $dd^c u$ for any continuous function u, and for this purpose we shall need to have recourse to the theory of distributions.

Let $C_c^\infty(U)$ denote the space of smooth functions on U with compact support. For any continuous function u on U, one defines *$dd^c u$ in the distributional sense* to be the linear form on $C_c^\infty(U)$ given by

$$\langle dd^c u, \varphi \rangle := \int_U u \; dd^c \varphi.$$

If u is smooth, then $dd^c u$ in the distributional sense coincides with $dd^c u$ in the usual sense, interpreted as a linear form on $C_c^\infty(U)$ via integration.

Proposition XIII.2.1. — *A continuous function $u : U \to \mathbb{R}$ is subharmonic if and only if $dd^c u$ is a positive linear form, that is, a positive measure.*

Proof. — Let $u : U \to \mathbb{R}$ be a continuous function, and consider a rotation-invariant, positive regularizing kernel $(\rho_\varepsilon)_{\varepsilon > 0}$ with compact support. For each

$\varepsilon > 0$, write $u_\varepsilon := u * \rho_\varepsilon$; this is a smooth function tending uniformly to u on every compact subset of U as ε tends to 0.

Assume first that $dd^c u$ is, in the distributional sense, a positive linear form, and consider $\varphi \in C_c^\infty(U)$ with $\varphi \geq 0$. Then for ε sufficiently small, we have

$$\begin{aligned}\langle dd^c u_\varepsilon, \varphi\rangle &= \int_U u_\varepsilon \, dd^c \varphi \\ &= \int_U u \, dd^c(\varphi * \rho_\varepsilon) \\ &= \langle dd^c u, \varphi * \rho_\varepsilon\rangle \geq 0.\end{aligned}$$

This shows that $dd^c u_\varepsilon$ is (in the distributional sense) also a positive linear form. Since u_ε is smooth, $dd^c u_\varepsilon$ is well defined in the standard sense and (as a 2-form) positive. Hence u_ε satisfies the inequality (XIII.5). Then since u_ε tends to u uniformly on every compact subset as ε tends to 0, it follows that u also satisfies (XIII.5), and so is subharmonic.

Conversely, suppose u is subharmonic, that is, that it satisfies the inequality (XIII.5). From the rotational invariance of ρ_ε it follows that u_ε also satisfies that inequality. Since u_ε is smooth, we infer in turn that $dd^c u_\varepsilon$ is positive. By invoking the uniform convergence of u_ε to u, one then has, for every function $\varphi \in C_c^\infty(U)$ with $\varphi \geq 0$,

$$\begin{aligned}\langle dd^c u, \varphi\rangle &= \int_U u dd^c \varphi \\ &= \lim_{\varepsilon\to 0} \int_U u_\varepsilon dd^c \varphi \\ &= \lim_{\varepsilon\to 0} \int_U \varphi dd^c u_\varepsilon \geq 0.\end{aligned}$$

This shows that $dd^c u$ is, in the distributional sense, a positive linear form. □

As mentioned earlier, if u is a smooth function, then $dd^c u$ is invariant under holomorphic coordinate changes. Hence $dd^c u$ has a well defined sense when u is a smooth function defined on a region of a Riemann surface. One sees immediately that this remains true (in the distributional sense) for a function that is merely continuous. For, if u is a continuous function on a Riemann surface S, then by Proposition XIII.2.1 the following properties are equivalent:

(i) the linear form $dd^c u$ is a positive measure;

(ii) in terms of any holomorphic coordinate on S, for every x_0 and every positive r,

$$u(x_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(x_0 + re^{i\theta}) d\theta.$$

We call the function u *subharmonic* if it has these two properties.

Note that since the function defined on $\mathbb{C} \setminus \{0\}$ by $u(z) = -\log|z|$ is integrable over compact subsets of $\mathbb{C}$, it has, in the distributional sense, a Laplacian on $\mathbb{C}$ defined by

$$\langle dd^c u, \varphi \rangle := \int_{\mathbb{C}} u \, dd^c \varphi$$

for every smooth function φ with compact support. Green's theorem gives $dd^c u = -2\pi\delta_0$. To see this, one considers real numbers $0 < \varepsilon < r$ such that φ vanishes outside the disc of radius r. One then has

$$\int_{\varepsilon \leq |z| \leq r} u \, dd^c \varphi = -\int_{|z|=\varepsilon} \frac{\varphi}{\varepsilon} + \int_{|z|=\varepsilon} \log \varepsilon \frac{\partial \varphi}{\partial n},$$

from which one obtains $\langle dd^c u, \varphi \rangle = -2\pi\varphi(0)$ on letting ε tend to 0. Hence for a Riemann surface S, every subharmonic function on $S \setminus \{p_0\}$ with a simple logarithmic singularity at p_0 admits a Laplacian in the distributional sense defined on the whole surface S, namely a signed measure with mass -2π at p_0 and positive elsewhere.

We now consider a Riemann surface S and a subharmonic function $u_0 : S \to \mathbb{R}$ on S, with a simple logarithmic singularity. One way of constructing such functions is to choose a disc D with compact closure in S, for which there exists a biholomorphism $\varphi : D \to \mathbb{D}$; it then suffices to define u_0 by $u_0(p) := -\log|\varphi(p)|$ for $\varphi(p)$ in $\mathbb{D} \setminus \{0\}$ and $u_0(p) = 0$ otherwise. The logarithmic singularity is then located at the point $p_0 := \varphi^{-1}(0)$. Denote by $(u_n)_{n \geq 0}$ the sequence of functions generated by a sweeping-out process on S starting with u_0.

Proposition XIII.2.2. — *For all n the function u_n is subharmonic on $S \setminus \{p_0\}$ with a simple logarithmic singularity at p_0. Moreover, the sequence $(u_n)_{n \geq 0}$ is increasing.*

Proof. — Assuming inductively that u_n is subharmonic on $S \setminus \{p_0\}$, let D be the open disc such that u_{n+1} has been obtained from u_n by sweeping-out on D. Assume D contains p_0. Recall that u_{n+1} is equal to u_n on $S \setminus D$, and that $u_{n+1|D}$ is the sum of the harmonic extension of $u_{n|\partial D}$ and the Green's function on D with a logarithmic singularity at p_0. It follows that $u_n - u_{n+1}$ is subharmonic on D

and vanishes on ∂D. This implies, via the maximum principle for subharmonic functions, that $u_n \leq u_{n+1}$.

We now show that u_{n+1} is subharmonic. This is clearly the case on $D \setminus \{p_0\}$ and on $S \setminus \overline{D}$. If, finally, $x_0 \in \partial D$, if $z : U_0 \to \mathbb{C}$ is a holomorphic chart on a neighborhood of x_0 with $z(x_0) = 0$, and if $r > 0$ is such that $\mathbb{D}(0,r) \subset z(U_0)$, then $u_{n+1}(x_0) = u_n(x_0)$ and

$$u_n(x_0) \leq \frac{1}{2\pi}\int_0^{2\pi} u_n(z^{-1}(re^{i\theta}))d\theta \leq \frac{1}{2\pi}\int_0^{2\pi} u_{n+1}(z^{-1}(re^{i\theta}))d\theta.$$

This shows that u_{n+1} is indeed subharmonic on $S\setminus\{p_0\}$. In the case that the disc D does not contain p_0, the proof is similar. □

Proposition XIII.2.3. — *If, for a point p of $S \setminus \{p_0\}$, the sequence $(u_n(p))_{n\geq 0}$ is bounded above, then the function sequence (u_n) converges uniformly on compact subsets of $S \setminus \{p_0\}$ to a Green's majorant $u : S \to \mathbb{R}$ possessing a unique logarithmic singularity, simple and located at p_0.*

Proof. — This follows directly from Harnack's principle (Theorem XII.2.1) and Proposition XIII.2.2. □

The following proposition is crucial to Poincaré's proof. It is this that allows the sweeping-out process to be controlled. It represents the mathematical interpretation of the following physical intuition: *throughout the sweeping-out procedure, the total electric charge does not change*. Poincaré considers the result physically obvious and gives no proof.

Proposition XIII.2.4. — *The mass $\int_S dd^c u_n$ is independent of n.*

Proof. — Choose any n and let D denote the disc of S such that u_{n+1} is obtained by sweeping-out u_n on D. If $\varphi \in C_c^\infty(S)$ is equal to 1 on D, then

$$\langle dd^c u_n, \varphi\rangle = \int_S u_n \; dd^c\varphi \; = \int_S u_{n+1} \; dd^c\varphi \; = \langle dd^c u_{n+1}, \varphi\rangle.$$

The middle equality follows from the facts that $dd^c\varphi$ vanishes on D and u_n is equal to u_{n+1} outside D. If one lets φ approach the function taking the constant value 1 on D and 0 elsewhere, one obtains in the limit the equality of the total mass $dd^c u_n$ with that of $dd^c u_{n+1}$. □

XIII.2.3. The convergence of the sweeping-out process on A

We shall now prove that any sweeping-out process on the annulus A starting from a subharmonic function with a simple logarithmic singularity, converges. This establishes the existence of a Green's majorant on A.

We recall the construction of the annulus A. We chose a holomorphic chart z defined on an open subset U of S such that the image of z contains the unit disc of $\mathbb{C}$; we also chose a real number $r \in (0,1)$ and set $\Delta := \{p \in S \mid |z(p)| < r\}$ and $A = S \setminus \overline{\Delta}$. We now choose a real number $r' \in (r,1)$ and write

$$\Delta' := \{p \in S \mid |z(p)| < r'\}.$$

Fix on a point $p_0 \in A$ such that $p_0 \notin \Delta'$, and consider a subharmonic function $u_0 : S \to \mathbb{R}$ on S with a simple logarithmic singularity at p_0 and with support a compact subset of A. Let $(u_n)_{n\geq 0}$ be the sequence of functions generated by a sweeping-out process on A starting with u_0. By Proposition XIII.2.2, this is an increasing sequence of subharmonic functions on $A \setminus \{p_0\}$, all with a simple logarithmic singularity at p_0. Observe that furthermore these functions all have compact support in A. By Proposition XIII.2.3, in order to establish the existence of a Green's majorant on A it suffices to find a point $p \in A$ for which the sequence $(u_n(p))_{n\geq 0}$ is bounded. For each $n \geq 0$, we consider the function $\bar{u}_n : S \to \mathbb{R}$ coinciding with u_n on A and identically zero on

$$S \setminus A = \overline{\Delta} = \{p \in S \mid |z(p)| \leq r\}.$$

We shall need the following lemma.

Lemma XIII.2.5. — *Let r and r' be such that $0 < r < r' < 1$. For each s, denote by $\mathbb{D}(0,s)$ the open disc of radius s and center the origin in $\mathbb{C}$. If $u : \mathbb{D}(0,1) \to \mathbb{R}$ is a continuous subharmonic function vanishing on the disc $\mathbb{D}(0,r)$, then*

$$\frac{1}{2\pi}\int_0^{2\pi} u(r'e^{i\theta})d\theta \leq \log\frac{r'}{r}\int_{\mathbb{D}(0,r')} dd^c u. \tag{XIII.6}$$

Proof. — For each s such that $0 < s < 1$, write $J(s) = \frac{1}{2\pi}\int_0^{2\pi} u(se^{i\theta})d\theta$. We first consider the case that u is a smooth function; we shall in this case bound $J(r')$ by integrating its derivative. For $0 \leq s < 1$, we have

$$sJ'(s) = \int_0^{2\pi} \frac{du}{ds}(se^{i\theta})s\frac{d\theta}{2\pi}.$$

By Green's theorem the right-hand side of this equation can be interpreted as the integral of the Laplacian of u over the disc $\mathbb{D}(0,s)$: we therefore have, for $0 < s < r'$,

$$J'(s) = \frac{1}{s}\int_{\mathbb{D}(0,s)} dd^c u \ \leq \frac{1}{s}\int_{\mathbb{D}(0,r')} dd^c u.$$

Integration of this inequality yields

$$J(r') - J(r) \leq \log\frac{r'}{r}\int_{\mathbb{D}(0,r')} dd^c u.$$

Then since $J(r)$ is zero, we obtain the desired inequality (XIII.6).

In the case that u is not smooth, we consider, as in the proof of Weyl's lemma III.2.4 or that of Proposition XIII.2.4, a positive, rotation-invariant regularizing kernel $(\rho_\varepsilon)_{\varepsilon>0}$ with compact support. For $0 < \varepsilon < 1-r_1$, the function $u_\varepsilon := u * \rho_\varepsilon$ is subharmonic and smooth, so the inequality (XIII.6) holds for u_ε. As ε approaches 0, the function u_ε tends uniformly to u and the measure $dd^c u_\varepsilon$ converges weakly to $dd^c u$ (see the proof of Proposition XIII.2.1). The inequality (XIII.6) for u therefore follows from that same inequality for u_ε. □

From Lemma XIII.2.5 it follows that for each $n \geq 0$

$$\int_{\partial\Delta'} u_n = \int_{\partial\Delta'} \bar{u}_n \leq \log\frac{r'}{r} \int_{\Delta'} dd^c \bar{u}_n = \log\frac{r'}{r} \int_{\Delta'\setminus\overline{\Delta}} dd^c u_n.$$

Note that $dd^c u_0$ restricted to $A \setminus \{p_0\}$ is a measure of finite mass. Recall also that by Proposition XIII.2.4 "the total electric charge does not change in the course of the sweeping-out process", or, in other words, the integral $\int_{A\setminus\{p_0\}} dd^c u_n$ is independent of n and consequently $\int_{\Delta'} dd^c u_n$ is bounded independently of n. It follows from this that $\int_{\partial\Delta'} u_n$ is also bounded above independently of n, by a constant C, say.

This last bound suffices to control the functions u_n at a point. Consider the sequence of functions u_n restricted to $\partial\Delta'$. We have just shown that they are all of mean at most C. Moreover, the sequence (u_n) is increasing. It follows from the theorem on nested compact sets that there exists a point p of $\partial\Delta' \subset A \setminus \{p_0\}$ for which the sequence $(u_n(p))$ is bounded above by C. The sequence $(u_n)_{n\geq 0}$ (viewed as a sequence of functions on $A \setminus \{p_0\}$, rather than on $S \setminus \{p_0\}$) therefore converges uniformly on compact subsets of $A \setminus \{p_0\}$ to a Green's majorant $u : A \setminus \{p_0\} \to \mathbb{R}$ with a simple logarithmic singularity at p_0. By Proposition XIII.1.1, the uniformization theorem XII.0.1 now follows. □

XIII.3. The more direct proof of Koebe

We now expound the proof of the existence of a Green's majorant on the annulus A given by Koebe in his note [Koe1907b]. Koebe explains that this proof was inspired by his reading of Poincaré's memoir [Poin1907]. In fact it is not a question of a really new proof, but rather of a "radical tidying up" of the one in [Poin1907]. Apart from being significantly shorter than Poincaré's, Koebe's proof has a further advantage: it is no longer necessary to appeal to the theory of distributions to render it rigorous.

We shall use the notation of §XIII.1. Koebe's proof utilises an exhaustion of $A := S \setminus \overline{\Delta}$ by means of relatively compact annuli. One begins by choosing an exhaustion $D_0 \subset D_1 \subset \cdots$ of the surface S by means of an increasing sequence of relatively compact, simply connected regions, with analytic boundaries (see Corollary XI.2.2), with D_0 containing $\overline{\Delta}$. For each $n \geq 0$, we write $A_n := D_n \setminus \overline{\Delta}$. Thus we now have an exhaustion $A_0 \subset A_1 \subset \cdots$ of the annulus A by means of an increasing sequence of relatively compact topological annuli in S. For each n, the boundary of A_n in S is made up of the two components $\partial\Delta$ and ∂D_n.

Choose a point p_0 in A_0. Work of Schwarz shows that, for each n, the annulus A_n admits a Green's function u_n with its pole at p_0 (Corollary XI.1.6). Recall that this means that u_n is a well defined harmonic function on $A_n \setminus \{p_0\}$ with a simple logarithmic singularity at p_0 (see Definition XI.1.1), and tends to zero on leaving every compact subset of A_n. The function u_n is extended to the annulus $\overline{A}$ by setting $u_n = 0$ on $\overline{A} \setminus A_n$; the extended function so obtained is continuous.

Koebe's proof consists in showing that the function sequence $(u_n)_{n\geq 0}$ converges (uniformly on every compact subset of $A \setminus \{p_0\}$) to a Green's majorant on A. For each $n \geq 0$, the function u_n is harmonic on $A_n \setminus \{p_0\}$, zero on $A \setminus A_n$, and tends to $+\infty$ near p_0; hence by the maximum principle u_n is non-negative. This implies that the function $u_{n+1} - u_n$ is non-negative on $A \setminus A_n$ (since u_{n+1} is non-negative and u_n zero on $A \setminus A_n$); then since $u_{n+1} - u_n$ is harmonic on A_n (including p_0), we infer, once again from the maximum principle, that $u_{n+1} - u_n$ is non-negative on the whole of A. Hence $(u_n)_{n\geq 0}$ is an increasing sequence of non-negative functions. By Harnack's principle XII.2.1, in order to show that the sequence $(u_n)_{n\geq 0}$ converges uniformly on every compact subset of A to a function $u : A \setminus \{p_0\} \to \mathbb{R}$ (which will then automatically serve as a Green's majorant), it suffices to find a point $p \in A \setminus \{p_0\}$ such that the sequence $(u_n(p))_{n\geq 0}$ is bounded.

The key argument here concerns the bounding of the integral along ∂A of the partial derivative of u_n in a direction normal to ∂A. For sufficiently small $\varepsilon > 0$, denote by B_ε the open disc of radius ε centered at p_0, referred to a local holomorphic coordinate system defined in a neighborhood of p_0, and write $A_{n,\varepsilon} := A_n \setminus B_\varepsilon$. The boundary of $A_{n,\varepsilon}$ (in S) is comprised of the three components ∂A, ∂D_n and ∂B_ε. For any point p of the boundary of $A_{n,\varepsilon}$, write $\frac{\partial u_n}{\partial \nu}(p)$ for the partial derivative of the function u_n in the direction of the normal directed into the interior of $A_{n,\varepsilon}$, evaluated at the point p. Green's theorem gives:

$$\int_{\partial A} \frac{\partial u_n}{\partial \nu} = -\int_{A_{n,\varepsilon}} \Delta u_n - \int_{\partial B_\varepsilon} \frac{\partial u_n}{\partial \nu} - \int_{\partial D_n} \frac{\partial u_n}{\partial \nu}. \qquad \text{(XIII.7)}$$

Since the function u_n is harmonic on $A_n \setminus \{p_0\}$, the first term on the right-hand side is zero. Secondly, if w is a holomorphic coordinate given on a neighborhood of p_0, the function $u_n(p)$ behaves like $-\log|w(p) - w(p_0)|$ to within a bounded quantity, so that the second integral on the right-hand side approaches -2π as ε tends to 0. Finally, since, as we saw above, the function u_n is non-negative on $\overline{A_n}$

and zero on ∂A_n, the derivative $\frac{\partial u_n}{\partial \nu}(p)$ must be non-negative at every point p of ∂A_n. Hence the third integral on the right-hand side of (XIII.7) is non-negative. We have thus established the following bound:

$$\int_{\partial A} \frac{\partial u_n}{\partial \nu} \leq 2\pi. \tag{XIII.8}$$

Remark XIII.3.1. — Just like Poincaré, Koebe ignores, purely and simply, problems concerning the regularity of the functions he considers! For $n \geq 0$, the Green's function u_n is analytic on $A_n \setminus \{p_0\}$ but *a priori* only continuous on $\overline{A_n} \setminus \{p_0\}$; thus talk of the normal derivative of u_n along ∂A_n and applying Green's theorem makes no sense *a priori*. In order to resolve this sort of problem in Poincaré's proof, we had to appeal to the theory of distributions. However, in the present proof the difficulties turn out to be only apparent: one can show that the function u_n can be continued to an analytic function on a neighborhood of $\overline{A_n} \setminus \{p_0\}$. Here is a proof. Consider two copies $\overline{A_n^1}$ and $\overline{A_n^2}$ of the closed annulus $\overline{A_n}$. By gluing $\overline{A_n^1}$ and $\overline{A_n^2}$ together along their boundaries, one obtains a Riemann surface Σ_n, the double of the annulus A_n. The surface Σ_n has the topology of the torus $\mathbb{T}^2$, so it is an elliptic curve. Denote by σ the involution of Σ_n interchanging $\overline{A_n^1}$ and $\overline{A_n^2}$. By §II.2.4, there exists a unique function $v_n : \Sigma_n \to \mathbb{R}$ that is harmonic save at the points p_0^1 and p_0^2, with a singularity of the form $-\log|w-w(p_0^1)|$ at p_0^1 and of the form $\log|w - w(p_0^2)|$ at p_0^2. The function $v_n \circ \sigma$ is therefore harmonic save at the points $p_0^2 = \sigma(p_0^1)$ and $p_0^1 = \sigma(p_0^2)$, with a singularity of the form $-\log|w-w(p_0^2)|$ at p_0^2 and of the form $\log|w-w(p_0^1)|$ at p_0^1. By uniqueness, one must have $-v_n = v_n \circ \sigma$. Since the points of $\partial A_n^1 = \partial A_n^2$ are fixed points under σ, it follows that v_n vanishes on $\partial A_n^1 = \partial A_n^2$. Hence the restriction of v_n to $\overline{A_n^1}$ is a Green's function on $\overline{A_n^1} \simeq \overline{A_n}$ with a pole of the form $-\log|z - z(p_0)|$ at p_0, so by uniqueness again it must be the function u_n! This shows that u_n may be regarded as the restriction of an analytic function defined on a Riemann surface containing $\overline{A_n}$; there is therefore no need to scruple in talking of the normal derivative of u_n along ∂A_n and applying Green's theorem.

It remains to show that the inequality (XIII.8) suffices for bounding the sequence $(u_n(p))_{n\in\mathbb{N}}$ at a point $p \in A$. Recall that the annulus A is the complement in S of the closed disc $\overline{\Delta} := \{p \in S \mid |z(p)| \leq r\}$ $(r < 1)$ defined in terms of a holomorphic chart z whose image contains the unit disc of $\mathbb{C}$. Choose $r' \in (r, 1)$ and write $\Delta' := \{p \in S \mid |z(p)| < r'\}$. By taking r' sufficiently close to r, we may assume the point p_0 is not in $\overline{\Delta'}$. For each $n \geq 0$, denote by m_n the least value of the function u_n on the circle $\partial\Delta' = \{p \in S \mid |z(p)| = r'\}$ and consider the function h_n defined on the open subset U by

$$h_n(q) := m_n \frac{\log \frac{|z(q)|}{r}}{\log \frac{r'}{r}}.$$

On the circle $\partial\Delta = \partial A = \{p \in S \mid |z(p)| = r\}$ the functions u_n and h_n are both zero, and on the circle $\partial\Delta' = \{p \in S \mid |z(p)| = r'\}$ they satisfy

$$u_n \geq h_n = m_n.$$

Since the functions u_n and h_n are both harmonic on $\Delta' \setminus \overline{\Delta}$, it follows that u_n majorizes h_n on $\overline{\Delta'} \setminus \Delta$. Since u_n and h_n coincide on $\partial\Delta$, this implies that

$$\frac{\partial u_n}{\partial \nu}(q) \geq \frac{\partial h_n}{\partial \nu}(q)$$

at every point q of $\partial\Delta = \partial A$. In view of the bound on the integral of $\frac{\partial u_n}{\partial \nu}$ obtained above, we therefore have

$$\int_{\partial A} \frac{\partial h_n}{\partial \nu} \leq 2\pi.$$

On the other hand, a direct calculation using the formula defining h_n yields

$$\int_{\partial A} \frac{\partial h_n}{\partial \nu} = 2\pi r m_n.$$

It follows that the real sequence $(m_n)_{n\geq 0}$ is bounded. Choose any point p on $\partial\Delta'$. Harnack's inequality (XII.1) now implies the existence of a constant K such that

$$u_n(p) \leq K m_n$$

for all n. We therefore conclude, as desired, that the sequence $(u_n(p))_{n\geq 0}$ is bounded, and, as explained earlier, this entails that the sequence $(u_n)_{n\geq 0}$ converges (uniformly on every compact subset of $A \setminus \{p_0\}$) to a Green's majorant on the annulus A. □

Remark XIII.3.2. — The key arguments in Koebe's proof are very similar to those of Poincaré's proof: in both cases one controls the sequence $(u_n(p))_{n\in\mathbb{N}}$ at a certain point p of A by bounding the integral over a (or possibly several) circle(s) centered at p, of the derivative of u_n in a direction transverse to this (these) circle(s) (this bound being obtained via Green's theorem). An important difference is that in Poincaré's proof the measure $dd^c u_n$ has its support in a compact subset of A, while in that of Koebe it is concentrated on $\partial A_n \cup \{p_0\}$. (In Koebe's proof, if one extends the function u_n to S by setting $u_n = 0$ on $S \setminus \overline{A}$, then $\overline{u_n}$ becomes a subharmonic function of S with a simple logarithmic singularity at p_0; on the other hand, the measure $dd^c \overline{u_n}$ is concentrated at p_0 and on ∂A_n, and the linear density of $dd^c\overline{u_n}$ along ∂A_n is nothing more or less than the function $\frac{\partial u_n}{\partial \nu}$.) Thus Koebe's contribution in his note [Koe1907b] was to show:

— first, that Poincaré's memoir contained a fundamentally new argument, namely, that the conservation of electric charge during the sweeping-out process implies that the integral along the boundary of A of the normal derivative of the functions arising in that process is bounded;

— second, that this argument of Poincaré suffices by itself for the desired conclusion, provided the sequence of functions arising in the sweeping-out process is replaced by a much simpler sequence of functions — for example, the sequence of Green's functions associated with an exhaustion of A.

Epilogue

The uniformization theorem from 1907 to 2007

By the end of 1907 the uniformization theorem was definitely proved. Koebe's proof and that of Poincaré (as revised by Koebe) seem to us rigorous in the modern sense of the word. However, it took some time for these proofs to be tidied up and the "simple and natural" versions arrived at that one finds in modern works. The theorem itself has ceased to be a research topic and has become a tool that mathematicians hold in their minds to be endlessly polished.

In 1909 Koebe was invited by Poincaré to republish his proofs of the uniformization theorem in his Notes in the *Comptes rendus de l'Académie des sciences de Paris* [Koe1909b, Koe1909c]. In the note [Koe1909b], Koebe states that the uniformization problem separates into two sub-problems: one of Analysis Situs (given a Riemann surface, find a not necessarily simply connected covering homeomorphic to an open subset of the Riemann sphere), and a conformal problem (replace "homeomorphic" by "biholomorphic"). According to Weyl [Wey1955]:

> From then on, Koebe spent his whole scientific life in studying the problem of uniformization thoroughly from all sides, and with the most varied methods. To him above all we owe it that today the theory of uniformization, which certainly may claim a central role in complex function theory, stands before us as a mathematical structure of a particular harmony and grandeur.

The list of Koebe's publications on uniformization is indeed impressive. He used the very latest techniques to improve the proof — for instance, ideas of Hilbert — or to produce more general theorems [Koe1908a, Koe1908b, Koe1910a, Koe1910b, Koe1910c]. He also published articles on the case of algebraic surfaces, examining the connections between different aspects of uniformization (linear differential equations and projective structures, Fuchsian groups and their deformations and fundamental polygons, Fuchsian functions, the method of continuity, uniformization *à la* Schottky, explicit examples of uniformization, ...) [Koe1909e, Koe1910a, Koe1912, Koe1914].

In his 1976 article reviewing the status of Hilbert's twenty-second problem (uniformization), Bers points up his theme with the following quotation from Goethe:

> Was du ererbt von deinen Vätern hast, erwirb es, um es zu besitzen.[1]

According to Bers [Bers1976]

> Each generation of mathematicians, obedient to Goethe's advice, rethinks and reworks the solutions discovered by their predecessors and places them within the framework of the concepts and notation of their epoch.

It is not our aim to describe here the development of the theory of Riemann surfaces in the course of the 20th century. We shall, however, provide a reading list that will allow the reader to retrace the process by means of which the various proofs of the theorem have been progressively simplified between 1907 and today.

Apart from their complexity, Koebe's and Poincaré's 1907 proofs possess other weak points. A Riemann surface as they understood the concept, is endowed *a priori* with a globally defined meromorphic function. It is true, however, that Poincaré and Koebe would not have considered this a weakness since it was built into their definition of Riemann surface. The construction of Green's functions rested on Schwarz's alternating method in Koebe's proof, and on "sweeping-out" in Poincaré's. Now although Schwarz's alternating method is certainly solid enough, it is not very natural since it depends on a choice of triangulation, whereas an abstract Riemann surface does not possess a canonical triangulation. And the sweeping-out method requires certain analytic refinements not actually available to Poincaré.

Improvements were made in quick succession. In June 1907, Montel published his first article [Mon1907] on *normal families*, work which lent itself to significantly clarifying and simplifying the last part of Koebe's and Poincaré's proofs, dealing with the convergence of the partial uniformizations defined on relatively compact regions.

In 1899, Hilbert resuscitated Dirichlet's principle using new methods allowing him to show that certain functionals in effect admit a minimum (see [Hil1900a, Hil1904, Hil1905]). Then in 1909, he finally announced that his results sufficed to establish the existence of a Green's majorant in a very general framework

[1] In order to possess what you have inherited from your parents, you need to earn it.

[Hil1909], and this served to liberate the proofs of the uniformization theorem from the need to use Schwarz's alternating procedure. Koebe was not slow to publish, in 1909 and 1910, articles in which he improved his proof by means of Hilbert's methods [Koe1909a, Koe1910b]. Of course, the scope of Hilbert's ideas went beyond this since they opened the way to Hodge theory.

Let us summarise the situation as of 1909: for a Riemann surface endowed with a meromorphic function, one has a simple proof of the uniformization theorem. The topological part (the existence of a countable base of open subsets) is given by the Volterra–Poincaré theorem while the conformal part is given by Hilbert's results.

During the Göttingen Winter semester of 1911–1912, Weyl expounded the theory of Riemann surfaces. His book *Die Idee der Riemannschen Fläche* appeared in 1913 and was revised several times right up to 1955 [Wey1913, Wey1955]. It played a fundamental role. One of the contributions of this book was to give for the first time the modern definition of an abstract manifold and hence of a Riemann surface not *a priori* endowed with a meromorphic function. Weyl notes that this "abstract" approach goes back essentially to Klein, who himself talked in this regard of the influence of Prym, however.[2] The proof of the uniformization theorem offered by Weyl uses the ideas of Hilbert, at that time fresh, concerning Dirichlet's principle. This proof does not easily extend to abstract Riemann surfaces, however, since it was not then clear that these possess a countable open base. The first edition also used triangulated Riemann surfaces.

As far as we know, Weyl's book remained the principal complete reference for the uniformization theorem till the 1950s. Of course, various facets of the theorem were considered in other books. We may mention the two volumes of Appell–Goursat and Fatou [ApGo1929, Fat1930], and Ford's book [Ford1929]. There is also the overview by Fricke and Klein, essential, but not easy, reading.

In 1925, Radó showed that every abstract Riemann surface is triangulable and has a countable open base [Rad1925], which eliminated the need for these assumptions noted above in connection with Weyl's book. Thus the theorem was now completely proved for abstract surfaces.

In 1941, van der Waerden proposed a very simple topological argument — not using the topological classification of surfaces — showing that an open, simply connected Riemann surface is an increasing union of compact, simply connected regions with polygonal boundaries [Wae1941]. This simplified a portion of the proofs by Koebe and Poincaré that had not been clear in detail in their articles, and allowed Carathéodory to give a simplified proof of the uniformization theo-

[2] *Thus we are faced with the unique situation that the less restrictive idea of a Riemann surface is due to some unknown comment made by Prym and misunderstood by Klein*, according to the review of Weyl's book by Sario.

rem in 1952, once again using Schwarz's method for polygons [Car1932, Second Edition].

Another important approach was opened up in the 1920s. The use of subharmonic functions and "Perron families" led to a considerable simplification of the construction of Green's functions. The method was established in [Per1923] and simplified in [Wien1924a, Wien1924b, Wien1925] and then in [Bre1939]. The Dirichlet principle thus attained its simplest formulation and proof at this time. One may consult [Han1979, Gard1979] for the history of this principle and successive approaches to it.

In 1949, Heins observed that Perron's methods require no topological assumption and so allow one to bypass the results of Radó mentioned above [Hei1949].

Thus by the beginning of the 1950s all the requisite tools were in place: normal families, harmonic and subharmonic functions, and the elements of topology. This was also the epoch when the theory of sheafs made its appearance, allowing the limpid formulation of a great number of concepts often somewhat unclear before: divisors, cohomology, the Riemann–Roch theorem, Serre duality, etc. The time was ripe for "optimized" presentations of the theory of Riemann surfaces, and in particular the uniformization theorem, using little topology and little analysis.

Here are a few reference works published since 1950, all containing a proof and in which the successive improvements are clear from the order of the details.

1. Nevanlinna, *Uniformisierung* [Nev1953]. Existence theorems are proved with the aid of Schwarz's alternating method. One finds here a new proof of the fact that a Riemann surface has a countable open base.

2. Springer, *Introduction to Riemann surfaces* [Spr1957]. Uses the orthogonal projection method *à la* Hilbert and follows Koebe's proof [Koe1910b]. This book presents physical intuitions well. However, it is less rigorous topologically (in connection with the existence of a triangulation).

3. Ahlfors and Sario, *Riemann surfaces* [AhSa1960]. This book gives a detailed exposition of the topological classification of surfaces and Radó's theorem on the existence of a triangulation. It also presents Dirichlet's principle and analytical methods (Hilbert's method, and capacities).

4. Ahlfors, *Conformal invariants* [Ahl1973]. Here Perron's method is used systematically to construct Green's functions.

5. Farkas and Kra, *Riemann surfaces* [FaKr1980]. Contains a complete presentation of the uniformization theorem based on Perron's method, occupying around 80 pages.

6. Beardon, *A primer on Riemann surfaces* [Bea1984]. Gives a complete proof of the uniformization theorem based on Perron's method.

7. Reyssat, *Quelques aspects des surfaces de Riemann* [Rey1989]. Here twenty pages suffice for a clear presentation of the theorem. The main argument in the proof involves Perron's method.

8. Forster, *Lectures on Riemann surfaces* [Forst1977]. Surprisingly, this book talks of "Riemann's theorem" rather than the uniformization theorem.

9. Jones, *Rudiments of Riemann surfaces* [Jon1971]. Provides a detailed proof of uniformization using Perron's method.

10. Abikoff, *The uniformization theorem* [Abi1981]. In less that twenty (effective and instructive) pages the author presents a proof of the uniformization theorem inspired by Perron's method.

11. Hubbard, *Teichmüller theory and applications to geometry, topology, and dynamics* [Hub2006], provides a crystal-clear proof that "goes by itself". However, behind it looms a two-hundred-year-long mathematical heritage — even if one looks in vain for a reference to "Vätern"!

12. Donaldson, *Riemann surfaces* [Don2011]. This last book on our list bases the proof of the uniformization theorem on the global surjectivity of the Laplacian under certain constraints. It avoids the use of Perron's method and does not give historical information.

To conclude, we should mention again that the special case of compact surfaces has recently undergone a resurgence of interest as a result of the emergence of methods of Hamilton–Perelman type used in proving the Poincaré conjecture. It was of course natural to attempt to prove that every Riemannian metric on a compact surface is conformally equivalent to a metric of constant curvature by tracking the "Ricci flow". The proof that this procedure functions effectively is unfortunately not as elementary as one might have hoped (it dates from 2006 [CLT2006]).

Appendices

The correspondence between Klein and Poincaré

We give here the correspondence between Klein and Poincaré from the period 1881–1882. A translation of Klein's letters (originally in German) into French by François Poincaré, Henri Poincaré's grandson, was published in the *Cahiers du Séminaire d'Histoire des Mathématiques* [Poin1989].[3] These letters have also been published, in untranslated form, in *Acta Mathematica* [KlePoi1923]. These two references are extensively annotated.

By way of introduction, we confine ourselves to a quotation from Freudenthal [Freu1955]:

> Twenty-six letters were exchanged between Klein and Poincaré on the topic of automorphic functions. Klein wrote first following the appearance of Poincaré's third Note. In this correspondence, Poincaré is the pupil who asks the questions and Klein the master who, in all sincerity and fidelity guides his pupil and forces him to make good the huge gaps in his mathematical knowledge. Only one thing was in dispute: Klein disapproved of the name Fuchsian functions which Poincaré had chosen unaware of the merits of the mathematicians of Riemann's school, but Poincaré persisted with it. In talking of automorphic functions one is accepting Klein's viewpoint.
>
> Who can measure the feelings provoked in Klein by the tremendous and instantaneous advance made by Poincaré along the path where they, Klein and his students, had progressed by slow steps? The more one considers this situation, the more one has to admire Klein's irreproachable attitude.

[3]The English translation of Klein's correspondence appearing below is based on the original German, with François Poincaré's French version used as a crib. *Trans*

I

Leipzig, June 12, 1881

Dear Sir!

Your three Notes in the *Comptes rendus*: "On Fuchsian functions" [Poin1881b], which I first became acquainted with yesterday and then only fleetingly, are so closely related to the reflections and endeavors that have occupied me these last years, that I feel obliged to write to you. First of all I would like to tell you of various works of mine on elliptic functions published in Volumes XIV [Kle1878b, Kle1878a, Kle1878c], XV [Kle1879a, Kle1879b], and XVII [Kle1880a] of the *Mathematische Annalen*. As far as modular elliptic functions are concerned, I dealt with only a special case of the independence relation that you consider; but a closer examination will show you that I did indeed have a general point of view. In this regard, I draw your attention to certain particular points:

— p. 128 of Volume XIV [Kle1878b] deals with general functions representable by modular functions, independently of being connected with doubly periodic functions. Then follows, first in a special case, the important theorem on the fundamental polygon;

— pp. 159–160 of Volume XIV [Kle1878b] where I expound the result that every hypergeometric series can be represented by single-valued functions of suitable modular functions;

— p. 428 *et seqq.* of Volume XIV [Kle1878b] contain a table illustrating the mutual disposition of triangles with sides circular arcs and angles $\frac{\pi}{7}, \frac{\pi}{3}, \frac{\pi}{2}$ (which is also an example from the classes of special functions studied by Halphen), apropos of which I should mention by the way that Mr. Schwarz has elaborated the case $\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{4}$ in Volume LXXV of Crelle's Journal.

Starting from p. 62 of Volume XVII [Kle1880a], I present a rapid overview of the more mature conceptions of the theory of elliptic modular functions which in the meantime had occurred to me. I have not published anything on these; however I did present them during the summer of 1879 in a course at the Munich Polytechnic. My line of thought, which comes very close to your exposition at many points, was as follows:

1. Periodic and doubly periodic functions are just examples of single-valued functions admitting linear transformations.[4] It is the task of modern analysis to determine all these functions.

[4]The naming of these function became the subject of a lively debate between Klein and Poincaré, who called them Fuchsian functions. Today they are called automorphic functions.

2. The number of such transformations may be finite; these give the equations of the icosahedron, octahedron, ... that I investigated earlier (Math. Annalen IX [Kle1875], XII [Kle1877a, Kle1877b]) and served as the point of departure for all of these sets of ideas.

3. Groups of infinitely many linear transformations, giving rise to useful groups (discontinuous groups, in your terminology), are obtained *for example* when one starts with a polygon with circular arcs as sides, such that these circles cut a fixed circle orthogonally and have angles equal to exact simple fractions of π.

4. One should investigate all such functions (as indeed you have already begun doing); however, in order to attain concrete goals, let us confine ourselves to triangles of circular arcs, and, in particular, to elliptic modular functions.

Since then I have been much occupied with these questions, also in discussions with other mathematicians, but except to say that I have not yet obtained any definitive result, this is really not the place for them. I would like to limit myself to what I have published or lectured on. Perhaps I should have made contact with you sooner, or with one of your friends, such as Mr. Picard. (When the occasion presents itself, would you draw Mr. Picard's attention to *Annalen*, XIV, p. 122, §8! [Kle1878b].) For, the line of development of the set of ideas that have engaged you for the last 2 or 3 years, is, in actuality, very close to mine. I would also be very happy if this first letter led to a continuing correspondence. It is true that at the moment other business takes me from these questions, but I am the more encouraged to take them up again in that next Winter I am to give a course on differential equations.

Please convey my compliments to Mr. Hermite. I have often thought of starting a correspondence with him, and would have done so long ago — doubtless to my great profit — if it were not for the language problem. As you may perhaps know, I was in Paris long enough to learn to speak and write in French; however, in the meantime this ability has faded through disuse.

With the greatest respect,

Prof. Dr. F. Klein

Address: Leipzig, Sophienstraße 10/II.

II

[Caen] June 15 [1881]

Monsieur,

Your letter proves that you anticipated me in some of the results I have obtained in the theory of Fuchsian functions. I am not at all surprised for I know how well versed you are in non-Euclidean geometry, which is the true key to the problem now occupying us.

I will do justice to you in that regard when I publish my results; I hope to be able to obtain here Volumes 14, 15, and 17 of *Mathematische Annalen*, which are not in the University of Caen library. Regarding the talk you gave at the Munich Polytechnic, I would ask you if you might give me some details on that subject so that I may add to my memoir a note rendering full justice to you; for I will doubtless not be able to get direct access to your work.

Since without doubt I won't *immediately* be able to get hold of *Mathematischen Annalen*, I would beg you also to kindly give me explanations of some points in your letter. You speak of *the elliptic modular functions**.[5]

Why the plural? If the modular function is the square of the modulus expressed as a function of the ratio of the periods, then there is only one such function; thus the expression *modular functions** must mean something else.

What do you mean by algebraic functions capable of being represented by modular functions? Also, what is the *theory of the fundamental polygon**?

I would also ask you to clarify for me the following points: Have you found all *polygons of circular arcs** giving rise to a discontinuous group?

Have you proved the existence of functions corresponding to all discontinuous groups?

I have written to Mr. Picard to communicate your remark.

I am pleased, Monsieur, to have the opportunity of making contact with you. I have taken the liberty of writing to you in French since you tell me you know that language.

Please be assured, Monsieur, of my respectful regard.

Poincaré

III

Leipzig, June 19, 1881

Dear Sir!

On receiving your welcome letter yesterday, I immediately sent you those offprints I still have of works relating to our topic. Allow me to add today a few lines of explanation pertaining to them. Of course, the question will not be settled by a single letter; rather must we correspond more extensively in order to establish mutually close contact. I would like today to emphasize the following points:

[5]In Poincaré's letters, an italicised expression marked with an asterisk indicates that that expression was in German in the original letters.

1. Of the papers sent to you the three most important from Volume XIV of the *Annalen* [Kle1878a, Kle1878b, Kle1878c] are missing, as also my investigations of the icosahedron in Volumes IX and XII [Kle1875, Kle1877b], and my second memoir on linear differential equations (which seems also to be unknown to Mr. Picard) in Volume XII [Kle1877a]. I entreat you to procure them somehow. I have been sending various offprints to Paris, for instance to Hermite.

2. The work of my students Dyck and Gierster complements my own. I am asking both of them to send you their offprints. Mr. Hurwitz's doctoral dissertation, relating to these same theories, is soon to be published and you will get a copy within a few weeks.

3. A compatriot of yours whose name you must surely know since he studied under Picard and Appell, namely Mr. Brunel (at the address Liebigstraße 38/2), has been here since last Autumn. Perhaps you would be interested in starting a correspondence also with him; he could better than I tell you about the organization of our seminar and the role played there by single-valued functions invariant under linear transformations.

4. I have had Mr. Gierster write up a set of notes from the course I gave in the summer of 1879. For the time being it is on loan, but I should get it back in a few days and will go through it with Mr. Brunel before giving you an account of it.

5. I reject the name "Fuchsian functions" although I understand that you were led to these ideas via work of Fuchs. In essence, all these investigations are based on those of Riemann. My own evolution in this regard was strongly influenced by Schwarz's deliberations, closely linked to Riemann's and of great significance, appearing in Volume 75 of *Borchardt's Journal* [Schw1873] (and which I can strongly recommend if you are unaware of them). Mr. Dedekind's memoir on elliptic modular functions appeared only in Volume 83 of *Borchardt's Journal* [Ded1877], when the geometric representation of modular functions was already clear to me (by the autumn of 1877). In their ungeometric form, Fuchs's memoirs stand in deliberate opposition to that of Dedekind. I don't deny the great service that Mr. Fuchs has rendered other parts of the theory of differential equations, but his work here leaves so much to be desired that on the only occasion when, in a letter to Hermite, he expatiated on elliptic modular functions, he made a fundamental mistake, which Dedekind criticizes only lightly in the abovementioned memoir.

6. One may, in particular, define a function invariant under linear transformations by the property that it maps the *half-plane* onto a given polygon with sides circular arcs. This in fact represents only a special case of the general situation (I don't know yet if you have been limiting yourself to just this particular case). The corresponding group of linear transformations is then characterized by the fact that it is contained in a group of operations twice as large, which, in addition

to linear transformations, contains reflections (inversions). In this case the existence of the function was rigorously established in much earlier work of Schwarz and Weierstrass, if only one would prefer not to appeal to general principles of Riemann. See Volume 70 of *Borchardt*: "Mapping the half-plane on polygons of circular arcs" [Schw1869].

7. Even in this special case I haven't yet been able to find all "discontinuous groups"; I have only established that there are many for which there is no determined fundamental circle, so that the analogy with non-Euclidean geometry (with which, by the way, I am very familiar) does not hold. If you take, for example, any polygon with sides any tangentially incident circles *whatever*, then generation via symmetry will always yield a *discontinuous group*.[6]

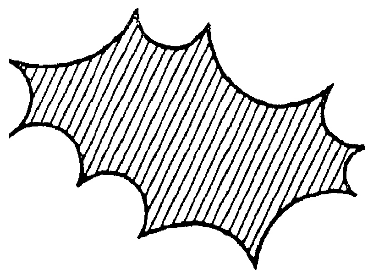


Fig. 1.

Figure 1: A polygon of circular arcs giving rise to a discontinuous group

8. You will doubtless find answers to the other questions you pose in your letter in the articles I am sending you, in particular to those concerning the plural form of "modular functions" and, especially, "fundamental polygons".

In the hope of hearing from you again soon

your very devoted

F. Klein

IV

Caen, June 22, 1881

Monsieur,

I have not yet received the items you mentioned but I shall doubtless not long await their arrival. However, I did not want to put off thanking you for your undertaking, and for your letter, which I read with the greatest interest. Immediately

[6]See Figure 1.

upon receiving it, I ran to the library to request Volume 70 of *Borchardt*; unfortunately that volume was on loan so I was unable to read Mr. Schwarz's memoir. However, I expect I'll be able to reconstruct it from what you have told me and to recognize there certain results I had found without thinking they had been the object of earlier research. I therefore expect to understand that the Fuchsian functions determined by the investigations of Mr. Schwarz and yourself are none other than those with which I have been occupied, in particular in my note of May 23 [Poin1881c]. The particular group you speak of in your last letter seems to me of great interest and I beg your permission to cite that passage of your letter in a communication that I will shortly be delivering to the Académie in which I will try to generalize your result.

As far as the name Fuchsian functions is concerned, I will not change it. The regard in which I hold Mr. Fuchs forbids changing it. Furthermore, even if it is true that the viewpoint of the scholarly Heidelberg geometer is completely at odds with yours and mine, it is nonetheless certain that it was his work that served as the point of departure of all that has been done since in that theory. It is thus only fair that his name remain attached to the functions that play such an important role in it.

Please be assured, Monsieur, of my respectful regard.

Poincaré

V

Leipzig, June 25, 1881

Dear Sir!

Please send me at once a postcard informing me if my package of offprints has not yet arrived; I myself took it to the post office eight days ago. You will express yourself differently about F.[7] when you become familiar with the relevant literature. The theory concerning maps of polygons of circular arcs is completely independent of his paper in Volume 66 [Fuc1866]; the only thing they have in common is to have been inspired by Riemann.

With the greatest respect

Prof. Dr. F. Klein

[7]Fuchs.

VI

Caen, June 27, 1881

Monsieur,

At the very moment when I received your postcard, I was on the point of writing to thank you for your parcel and announce its arrival. If it was held up it was through a mistake on the part of the post office, which sent it first to the Sorbonne and then to the Collège de France, even though it was addressed perfectly correctly.

As far as Mr. Fuchs and the naming of Fuchsian functions is concerned, clearly I would have chosen another name if I had been aware of Mr. Schwarz's work; however, I came to know of that only from your letter and therefore after the publication of my results, so that I cannot now change the name I gave those functions without demonstrating a lack of respect for Mr. Fuchs. I have begun reading those of your papers of liveliest interest to me, mainly that entitled "Über elliptische Modulfunktionen" [Kle1880a]. It is concerning that paper that I would like to put certain questions to you.

1. Have you determined the *fundamental polygons** of all the *subgroups** that you call *congruence groups**, and, in particular, of the following one:

$$\alpha = \delta = 1, \ \beta = \gamma = 0 \bmod n.$$

2. In my memoir on Fuchsian functions, I classified Fuchsian groups according to various principles, among others by means of a number I called their genus. You similarly classify *subgroups** by means of a number you call their *Geschlecht*[8]. Are the *genus* (as I understand it) and the *Geschlecht* one and the same number? I haven't been able to find out since I do not know what the *Geschlecht im Sinne der Analysis Situs*[9] is. All I see is that these numbers cancel each other out. Would you please, therefore, do me the favor of telling me what *Geschlecht im Sinne der Analysis Situs* means, or, if the definition is too long to be included in a letter, in which work I might find it? In your last letter you ask me if I have confined myself to the special case where *"the group of linear transformations is characterized by the fact that it is contained in a group of operations twice as large, which, besides linear transformations, also contains reflections"**. I have in fact not limited myself to this case, but I have assumed that all the linear transformations preserve a certain fundamental circle. Furthermore, I believe I can address the more general case by means of a similar method.

[8] That is, "genus". We leave the word untranslated since it is the object of a discussion between Klein and Poincaré.

[9] Genus in the sense of Analysis Situs.

Apropos of that, it seems to me that all the *subgroups** relevant to modular functions do not fall into this special case.

Concerning the discontinuous group you spoke of, obtained via reflections and by reproduction via symmetries of a polygon bounded by arcs of circles tangent in pairs, it seems to me that there is a supplementary condition that you failed to mention although it doubtless had not escaped you: no two of the circular arcs should intersect. Would it abuse your complaisance if I asked you yet another question?

You write *"in this case, the existence of the function was established in much earlier work of Schwarz and Weierstrass"**, and you add *"if only one would prefer not to appeal to general principles of Riemann"**. What do you mean by this?

I recently wrote to Mr. Hermite; I communicated succinctly the content of your letters, and conveyed to him your compliments as requested.

Please be assured, Monsieur, of my consideration and respect.

Poincaré

VII

Leipzig, July 2, 1881

Dear Sir!

Allow me to answer at once the various questions you pose in your welcome letter of June 27.

1. In Volume 14 [Kle1878b, Kle1878c] I describe in detail the congruence groups $\alpha \equiv \delta \equiv 1$, $\beta \equiv \gamma \equiv 0$ mod n for $n = 5$ (where by means of simultaneous deformation of the edges one obtains the icosahedron) and for $n = 7$. The general case $n =$ a prime number is the topic of a memoir of Dyck, at present in press. I have not yet completed the investigation of the case where n is composite.

2. A "Geschlecht im Sinne der Analysis Situs" is associated with every closed surface. It is equal to the largest number of closed curves that one can draw on the surface without dividing it in two. If one now considers the surface in question as the image of the values of the numbers w, z satisfying an algebraic equation $f(w, z) = 0$, then its genus is that of the equation. Your *genus* and my *Geschlecht* are thus *in fact the same number*; only my interpretation is presumably more clearly associated with the Riemann surface and the definition of p that it affords.

3. There do exist, however, in the group of modular functions, subgroups with asymmetric fundamental polygon, including in particular, as I have shown in Volume 14 [Kle1878a], those subgroups corresponding to the resolvents of the modular equation for $n = 7$ and $n = 11$.

4. I did of course know that in the case of a polygon, the circles should not intersect when extended towards the exterior if one wishes to have a single-valued function. In my opinion, one should concentrate precisely on this point in

order to prove that the coordinates w, z of the points of any algebraic curve can be represented by a single-valued function invariant under linear transformations. I will now indicate to you how far I have advanced with this question. By work of Schwarz and Weierstrass, one can always map the half-plane onto a polygon of circular arcs in such a way that the points I, II, III, IV, V corresponding to the points 1, 2, 3, 4, 5 of the boundary of the half-plane are positioned arbitrarily.[10]

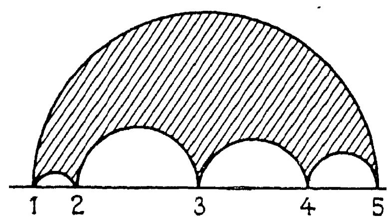


Figure 2: A polygon in the half-plane

Suppose now that I, II, III, IV, V, . . . are the branch points of an algebraic function $w(z)$ and that this algebraic function has *no other* branch points. Then obviously w and z are single-valued functions of the desired sort of auxiliary variables in the plane of the indicated polygon. *If, therefore, all of the branch points of an algebraic function $w(z)$ lie on a circle in the z-plane, then the answer to the question is immediately in the affirmative.* What if, however, that is not the case? Then I arrive, in fact, at polygons of the type noted last time. If the figure has no symmetry , I obtain (by establishing associated differential equations of the form $\frac{\eta'''}{\eta'} - \frac{3}{2}\left(\frac{\eta''}{\eta'}\right)^2 = R(z)$, which I have dealt with before) a fundamental region in similar fashion, where the edges meet tangentially, and which, moreover, are grouped together in pairs by means of certain linear substitutions. *But I cannot prove that this fundamental region together with its iterates covers only a part of the complex plane.* And this difficulty has held me up for a long time.

5. Furthermore, one obtains other remarkable examples of discontinuous groups if one takes an arbitrary number of pairwise disjoint circles and reflects them onto one another by inversions. For greater clarity I have shaded the part of the plane exterior to all the circles, representing the fundamental half-polygon.[11] These groups have been investigated occasionally by Schottky (*Journal de Borchardt*, Volume 83, pp. 300–351 [Schot1877]) without bringing their fundamental significance to light.

[10]See Figure 2.
[11]See Figure 3.

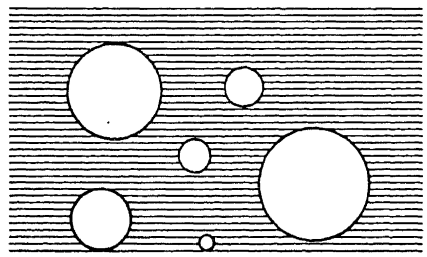


Figure 3: A "Schottky" group

6. Riemann's principles don't at first yield any means for actually constructing a function whose existence has been established. One is therefore inclined to consider them as uncertain even though the results following from them are correct. On the other hand, in connection with the abovementioned mapping of a polygon of circular arcs, Weierstrass and Schwarz have effectively determined the relevant constants by means of convergent processes. If one is prepared to use Riemann's principles, then one can prove the following very general theorem. Suppose given a polygon with one or more separate boundary components. The polygon may have several sheets, joined at branch points. Each boundary component is to be made up of several pieces, each of which can be transformed into another via a prescribed linear substitution. One can then always construct a function with arbitrarily prescribed discontinuities in the interior of the polygon and with real part taking on certain prescribed periodicity moduli as one passes from one piece of the boundary to the corresponding one by traversing the interior of the polygon. These functions include, in particular, those that are single-valued everywhere in the interior of the polygon and take the same value at each pair of corresponding points of the boundary. The proof is along precisely the same lines as that given by Riemann in §12 of the first part of his "Abelian Functions" [Rie1857] in the case of the special polygon made up of p parallelograms arranged one above the other and joined by means of $2p - 2$ branch points. This theorem, which by the way I have only fully established in the last few days, includes, it seems to me, all the existence proofs you mention in your notes as special cases or easy inferences. Incidentally, my theorem, like many that I write down these days, is not yet formulated precisely; otherwise I would have had to go into much more detail; you will easily get my meaning.

7. Allow me to make a remark concerning another of your publications [Poin-1881a]. You say that the θ-functions resulting from inversion of algebraic integrals on curves of genus p are not θ-functions of the most general sort. You could not know that in Germany precisely these ideas are common knowledge: a large number of young mathematicians are engaged in finding the conditions distinguishing so-called Riemann θ-functions from general θ-functions. On the other hand, I am surprised that you give the number of moduli of Riemann θ-functions as $4p + 2$, whereas in fact it should be $3p - 3$. Haven't you read the relevant explanations of Riemann? And aren't you aware of the whole discussion that Brill and Noether settled in Volume 7 of *Math. Annalen*, pp. 300–307 [BrNo1874]?

Hoping to hear from you soon, very respectfully yours,

F. Klein

VIII

Caen, July 5, 1881

Monsieur,

I have received you letter and read it with the greatest interest. I offer a thousand apologies for the question I asked you concerning "Geschlecht im Sinne der Analysis Situs". I could have saved you the trouble of responding since I found the explanation on the next page of your memoir. You will doubtless recall that in one of my recent letters I asked your permission to quote a sentence in a communication in which I proposed generalizing your results. You have not responded on this matter, so I am taking silence as acquiescence. The communication was made twice, at the meetings of June 27 and July 4 [Poin1881d, Poin1881e].

You will find that we have overlapped on some points. However, I think you will find the citing of your sentence sufficient guarantee.

Please allow me, Monsieur, another question: where can I find the works of Messrs. Schwarz and Weierstrass that you speak of, first concerning the theorem that: *one can always map the half-plane onto a polygon of circular arcs in such a way that the points I, II, III, IV, V corresponding to the points 1, 2, 3, 4, 5 of the boundary of the half-plane are positioned arbitrarily**? This theorem was not unknown to me, since I myself gave a proof of it in my communication of May 23 [Poin1881c]. But where can I find the works of my anticipators? In Volume 70 of *Crelle*? Where also can I find the developments of which you speak in the following sentence: *On the other hand, in connection with the above-mentioned problem of mapping a polygon of circular arcs, Weierstrass and Schwarz have effectively determined the relevant constants by means of convergent processes**?

The theorem you say you have discovered interests me a great deal. It is clear that, as you say, your result includes as special cases *all my existence proofs**. However, it comes later.

I now come to your remark about Abelian functions. When I spoke of $4p + 2$ constants, it wasn't a question of the number of moduli. What I actually said was this [Poin1881a]: an algebraic equation of genus p can always be reduced to degree $p + 1$. An equation of degree $p + 1$ depends on $4p + 2$ parameters; for, a *general* equation of degree $p + 1$ depends on

$$\frac{(p+1)(p+4)}{2}$$

parameters. But there are

$$\frac{p(p-1)}{2} - p$$

double points. Therefore there remain $4p + 2$ independent parameters. I thus obtain, not the number of moduli, but an upper bound for that number, and this was sufficient for my aim.

Please be assured, Monsieur, of my respectful regard.

Poincaré

IX

Leipzig, July 9, 1881

Dear Sir!

By way of a quick reply to your letter, I have more-or-less the following to say.

1. As far as I'm concerned, it is completely correct for you to have quoted that passage of my letter. Thus far I only have your first Note of June 27 [Poin1881d]. As to the name you have given that class of functions, here I was quite surprised; but then I myself had no more than noticed the existence of these groups. For my part, I will use neither "Fuchsian" nor "Kleinian" but stay with my "functions invariant under linear transformations".

2. What I said about the value of Riemann's principles was not precise enough. There can be no doubt that Dirichlet's principle must be abandoned as not at all conclusive. However, it can be completely replaced by more rigorous methods of proof. You will find this expounded in more detail in a work by Schwarz that I have just recently seen (in connection with my course) and in which you will find information on the determination of the constants, which was only indicated in *Borchardt's Journal* [Schw1873] (you must in any case examine the memoirs published in Volumes 70, 74, and 75 of *Borchardt's Journal*); the work of Schwarz in question is in the *Berliner Monatsberichten* 1870, pp. 767–795 [Schw1870a].

3. The general existence proof I mentioned last time remains valid, naturally, for groups made up of arbitrary analytic (not necessarily linear) substitutions. It is

remarkable that in this sense every group of operations defines functions remaining unchanged by them. “Discontinuous groups” have the advantage that they have associated *single-valued* functions, a very fundamental property, moreover. Might one be able to master higher cases by means of *single-valued* functions of *several* variables as was the custom in connection with the case treated by Riemann in §12 [Rie1857] relating to the Jacobi inversion problem?

Enough for today. In the meantime, with Mr. Brunel I have looked over my works, notably the lecture notes from 1877–78 and 78–79 (which I had reworked back then), and he will shortly write to you about these.

With the greatest respect, your devoted

Prof. Dr. F. Klein.

X

Leipzig, December 4, 1881
Sophienstraße 10/II

Dear Sir!

After having long reflected casually on the the problems of mutual interest to us, this morning I took the opportunity of reading together the different communications that you have published successively in the *Comptes rendus*. I see that now you have definitely proved (as of August 8): *that every linear differential equation with algebraic coefficients is integrable by means of zeta Fuchsian functions and that the coordinates of any algebraic curve whatever can be expressed via Fuchsian functions of an auxiliary variable*[12]. While congratulating you on the results you have obtained, I would also like to put a proposition to you respecting both your interests and mine equally. I would ask you to to send me, for *Mathematische Annalen*, a more or less long article, or, if you don’t have the time to write up such a work, then a *letter*, expounding, in broad strokes, your points of view and results. I would then write an accompanying note in which I would describe how I view the whole matter and how at this juncture the research program that you are now pursuing has served as a fundamental guiding principle for my work on modular functions. Of course, I would submit this note to you for your approval prior to publication. Such a publication would have a double effect: first, it would definitely draw the attention of the readers of *Math. Annalen* to your work, doubtless a desirable outcome for you; and second, your work would be presented to a large general readership, at the same time demonstrating the connections with my work as they actually are. As I know from what you wrote to me, you intend to

[12]In French in Klein’s letter.

analyse these equations in a detailed memoir; but writing it will take time, and I would like an announcement to be made in the *Annalen* also.

For my part, I have meanwhile written up a a little treatment of "Riemann's theory" [Kle1882c] which may be of interest to you since it presents a version of the concept of a Riemann surface that I believe R. himself actually worked on. Perhaps Mr. Brunel has told you of this. I have also been busy lately with different existence proofs designed to replace Dirichlet's principle, and I am convinced that the methods expounded by Schwarz in the *Berliner Monatsberichten*, 1870, p. 767 *et seqq.* [Schw1870a] in any case suffice completely to yield, for example, the general theorem I wrote to you about once or twice this past summer.

With the greatest respect,

F. Klein

XI

December 8, 1881
Paris, rue Gay-Lussac 66

Monsieur,

I am infinitely grateful for the obliging offer you make me and am fully prepared to avail myself of it. I shall shortly send you the letter you request; I would ask you, however, how much room you are prepared to devote to it in the *Annales*. I know that your journal's clientele is numerous and that the space you can allow each article is necessarily limited and I would not wish to abuse your benevolence. As soon as I know what length I can make my letter, I shall write it for you at once.

I shall soon be honored to send you various notes relating to the general theory of functions, if you should wish to accept them.

I have recently read Schwarz's memoir in the *Monatsberichten* [Schw1870a] and his proofs appeared rigorous to me.

Please accept, Monsieur, my thanks and the expression of my great consideration.

Poincaré

XII

Leipzig, December 10, 1881

Dear Sir!

I am pleased that my proposition is agreeable to you: *voilà une loi de réciprocité.*[13] Concerning your question, I wish primarily to say that your article will be

[13]There we have a reciprocity law.

the more apropos the sooner I receive it. If I receive it by the 20th of this month, I will include it in Issue 4 of Volume 19 of the *Annalen*; it will then appear at the beginning of March at the latest. As to the size, I would say, if you're agreeable, about a printed sheet (16 pages). This is enough for you to be able to convey clearly the essentials, and not too long for a quick perusal. I would also ask you especially to give the necessary indications as to the *methods* of your proofs, that is, how you actually construct the functions in question, etc. But you can better assess all such matters than I can prescribe them.

One more question! Are you now permanently in Paris? And what is Picard's present address? I would be happy if I could also obtain a contribution to the *Annalen* from him.

With the greatest respect, your devoted

F. Klein

XIII

Paris, December 17, 1881
rue Gay-Lussac 66

Monsieur,

I have the honor of sending you the little work in question [Poin1882c]; I have not, as you asked me, described succinctly my methods. I was unable to do so without substantially exceeding the limits you imposed. I know that they are not at all rigid. However, on the other hand I don't believe that a proof can be summarized; one cannot subtract from it without depriving it of its rigor and a proof without rigor is no proof. I would prefer, therefore, to send you from time to time a series of short letters in which I would give successive proofs of the stated results or at least the main ones. You may do with these letters whatever you think fit. I do indeed live in Paris now; I am a lecturer in the Faculty of Science.

Here is Picard's address: Adjunct Professor in the Faculty of Science, rue Michelet 13, Paris.

Here also is Appell's: Lecturer at the École Normale Supérieure, rue Soufflot 22, Paris.

Please be assured, Monsieur, of my highest regard.

Poincaré

XIV

Leipzig, January 13, 1882

Dear Sir!

I have not yet thanked you personally for sending your article, for which I am exceedingly obliged to you. As things stand, it will go to press in a few days. You will be getting page proofs, which I ask you to return to Teubner Publishers in Leipzig. Would you, in particular, examine the short commentary that I have, along the lines mentioned earlier, appended to your article, and in which I protest, as strongly as I can, against the two names *Fuchsian* and *Kleinian*, citing Schottky with respect to the latter, and, incidentally, pointing to Riemann as the one who initiated all these investigations? I endeavored to preserve as moderate a tone as possible in this commentary, but I beg you to write to me immediately if you should require further amendments. I in no way wished to diminish your work with these comments. Furthermore, I have written another short article [Kle1882a] which will appear right after yours. It contains, also without proof, some results relating to the area in question, primarily the following one: *every algebraic equation* $f(w, z) = 0$ *can be solved in one and only one way by* $w = \varphi(\eta), z = \psi(\eta)$ *in terms of p independent reentrant cuts*[14] *on the corresponding Riemann surface, where* η *is a discontinuous group of the kind you spoke of regarding my letter.* This theorem is the more beautiful in that this group has exactly $3p - 3$ essential parameters, that is, the same as the number of moduli possessed by the equations of the given p. In this connection further considerations arise which seem to me to be of interest. In order to share this with you fully I have ordered the publisher to send you the page proofs of my article for you to have at your disposal.

As far as the *proofs* are concerned, they are difficult. I always operate with Riemann's ideas respecting "geometria situs". This is very difficult to get clear. I shall make every effort to achieve this in due time. Meanwhile I would like very much to correspond with you on that subject and also on your proofs. You can be sure that I will study with the greatest interest any letters on this subject that you may send me, and accord them a quick reply. If you should wish to publish them in one form or other, the *Annalen* are naturally at your disposal.

Very respectfully, your devoted

F. Klein

[14]Or "cuspidal" or "recurrent" cuts. In German, "Rückkehrschnitte". *Trans*

XV

[Paris, January 1882]

Monsieur,

I have received the page proofs from Teubner, and am returning them corrected. I have read your note and don't see any need to change anything there. However, I hope you will permit me to write you a few lines in an attempt to justify my terminology. I await impatiently the theorem you have told me of and which seems to me of the greatest interest.

Please be assured, Monsieur, of my greatest regard.

Poincaré

XVI

Paris, March 28, 1882

Monsieur,

You have adjoined to my article *On the single-valued functions that replicate themselves under linear substitutions* [Poin1882c] a note in which you list the reasons for your rejection of my terminology. You were good enough to send me the page proofs and to ask me if I wanted any changes made. I am grateful to you for the delicacy of your actions, and I was unable to abuse it by asking you to keep half your opinion to yourself.

You must understand, however, that I cannot leave the readers of the *Annales* with the impression that I have committed an injustice. That is why I wrote, if you recall, asking for no change to your note, but also permission to address a few lines to you justifying my terminology.

Thus here are those lines; perhaps you might judge them suitable for insertion. In turn I wonder if you would like me to make changes to this little note. I am ready to make all such changes on condition they don't alter the sense.

Please excuse my importunity and forgive this small piece of advocacy *pro domo*.

Please rest assured, Monsieur, of my greatest regard.

Poincaré

I would be obliged if you could tell me the address of Mr. Hurwitz, to whom I would like to pay homage by sending him a copy of my work.

I would be very grateful also, if you could indicate the general features of the proof by means of which you establish the theorem stated in your latest work: *"Über eindeutige Funktionen mit linearen Transformation in sich"* [Kle1882a].

On single-valued functions that replicate themselves under linear substitutions
(Excerpt from a letter addressed to Mr. F. Klein)
By H. Poincaré in Paris

. . . Recently you were good enough to have published in the *Mathematische Annalen* (Vol. XIX, pp. 553–564) [Poin1882c] an article of mine on single-valued functions that replicate themselves under linear substitutions, with a note appended in which you explain why you find the names I have given to these transcendental functions unsuitable. Allow me to address you a few lines in defense of my terminology, which I did not choose at random. If I believed I should bestow on these new functions the name of Mr. Fuchs, it was not out of disregard for the value of your work and that of Mr. Schwarz — on the contrary, I would be the first to appreciate its great importance. However, it was impossible for me to ignore the remarkable discoveries published by the Heidelberg professor in *Crelle's Journal*. They form the basis of the theory of linear equations and without them I would not have been able to begin my investigations of my transcendental functions, so directly linked to that theory. In his first articles, it is true, Mr. Fuchs takes a point of view a little different from mine and concerns himself neither with the discontinuity of the groups nor with the single-valuedness of the functions. However, neither is Mr. Schwarz, in his memoirs in Volumes 70 and 74 of *Crelle's Journal*, concerned with these things; he has a few words to say in that connection in a very special case in his memoir in Volume 75, which I cite in my note. It is only there that he finds himself in the domain of Fuchsian functions. In your beautiful research on modular functions, your manner of perceiving things differs little from mine, but you have more in mind the study of elliptic functions than linear equations. As for Mr. Fuchs, in his memoirs in Volumes 83 and 89 of *Crelle's Journal* he has risen to a new point of view and shed light on the close connection between the theory of differential equations and that of certain single-valued functions. It was the reading of these memoirs that set me on the path to my investigations.

As far as Kleinian functions are concerned, I should have felt I had committed an injustice had I given them a name other than yours. It was Mr. Schottky who discovered the figure you discuss in your letter, but it was you who *"underlined their fundamental importance"**, as you did at the conclusion of your learned work "Über eindeutige Funktionen mit linearen Transformationen in sich" [Kle1882a]. I cannot fully subscribe to what you say concerning Riemann. He was one of those geniuses who so change the face of science that they leave their imprint, not only on the works of their immediate students, but also on those of all their successors over many years. Riemann created a new theory of functions, and it will always be possible to find there the seed of everything that has been done and will be done after him in mathematical analysis. . . .

Paris, March 30, 1882

XVIII

Düsseldorf, April 3, 1882
Address: Bahnstraße 15

Dear Sir,

Your letter, which I received yesterday via Leipzig, arrived at the very moment when I was about to write you a few words apropos of my most recent note in the *Annalen* [Kle1882a], whose page proofs should already be in your hands. In the meantime I have obtained a copy of Prof. Fuchs's note published in the *Göttingen Nachrichten*. If I had to say two words concerning the latter, they would be to the effect that I judge it to be completely beside the point. I claim only that Fuchs has never published anything on "Fuchsian functions". It follows that the second article that he cites (which I must get hold of in order to examine it more closely) is objectless. The first may, however, be considered to be concerned with "Fuchsian functions" insofar as it deals with modular functions, but, lacking geometric intuition, Fuchs has not correctly recognized the proper character of the latter, which resides in the nature of singular lines, as Dedekind showed already in Volume 83 of *Borchardt* [Ded1877]. Finally, as for the insinuations at the end of his note to the effect that my own work has been essentially stimulated by his, this is quite simply historically false. My research began in 1874 with the determination of all finite groups of linear transformations in one variable [Kle1875]. Then in 1976 I showed that the problem raised at that time by Fuchs of determining all integrable second order linear differential equations was *eo ipso* solved [Kle1877a]. The situation is precisely the reverse of what Fuchs claims. It was not that I took his ideas, but rather that I showed that his topic should be treated using my ideas.

As you may well suspect, I am not in agreement with your presentation of the matter. If it were a matter of a general appreciation of Fuchs's *oeuvre*, I would willingly have his name bestowed on some *new* class of functions that no one had yet studied, or even, for instance, on the functions of several variables that he has put forward. (Are these *really* single-valued? All I understand is that over the whole of the range of values taken by them there is no branching. However, I may be mistaken in this.) However, the functions you have named after Fuchs already belonged to others before you suggested the name. I am also quite sure that you would have not proposed this name had you been then (at the beginning) familiar with the literature. You then offer me, so to speak in compensation, "Kleinian functions". To the extent that I perceive your friendly intention here, to just that extent is it impossible for me to accept the offer, as again perpetrating an historical untruth. If my memoir in Volume XIX [Kle1882a] might give the impression that I am now especially preoccupied with "Kleinian" functions, my more recent work in Volume XX [Kle1882b] shows that as before I continue to regard "Fuchsian" functions as my domain.

But enough on that theme. I immediately dispatched your note to the printer, appending only a remark to the effect that for my part I adhere to my previous opinion (and on this occasion drawing the public's attention to Mr. Fuchs's note). You will receive the page proofs very soon and I ask you to return them quickly to me here (where I am spending the Easter holidays), and I will then do what is necessary as far as publication is concerned. (Your note will appear directly after mine.) As far as the passage about Schottky is concerned, I would like to point out to you a posthumous article in Riemann's *Works*, p. 413, where exactly the same ideas are developed. I should say, however, that it would be difficult to determine the extent of a possible contribution by the editor, Prof. Weber. Riemann's *Works* appeared in 1876, and Schottky's dissertation was completed in 1875 and published in 1877 as a memoir in *Borchardt's Journal* [Schot1877]. However, the 1875 dissertation constitutes only a part of the 1877 memoir, and I cannot recall if the figure in question had already appeared in the 1875 text.

I should add that on my part I have no intention of prolonging our *terminological* disagreement (once I have added the above-mentioned footnote to your explanation). However, if I should be led to intervene in the matter anew then I would, it is true, give a very complete and frank account of it. Let us rather compete to see which of us is best equipped to advance the theory in question! On my side, I believe my new note represents a certain advance. A whole series of theorems on algebraic functions can be proved using the new η-function — for instance, the theorem that, in my book on Riemann, I initially indicated as only probably true, namely that a surface of genus $p > 0$ never admits infinitely many single-valued *discrete* transformations (since otherwise it would decompose into an infinite number of "equivalent fundamental polygons"). And also the theorem that various of Picard's results for the case $p = 0$ extend to general p, etc.

As for the methods I use to prove my theorems, I will write to you of them as soon as I have clarified them some more. In the meantime could you describe for me the ideas you are pursuing at the moment? I scarcely need add that we would be pleased to publish in the *Mathematische Annalen* any article you cared to send us. It would mean a great deal to me to remain in active contact with you. Lively contact with mathematicians aspiring to similar ends is always for me a prior condition of my own mathematical production.

Very respectfully, your devoted

F. Klein

Until further notice, Dr. Hurwitz's address is: *Hildesheim*, Langer Hagen.

XIX

Paris, April 4, 1882

Monsieur,

I have just received your letter and hasten to reply. You say you wish to cease our sterile debate for science's sake and I can only congratulate you on your resolution. I know it cannot have cost your too dear since it was you who, in the note you appended to my last letter, had the last word, but I am nonetheless grateful to you. As for me, I did not begin the dispute and entered into it only to voice once and for all my opinion that it was impossible for me to keep silent. It was not I who prolonged it, and I only spoke up because I was compelled to; and there are not too many things that could so compel me.

If I named Kleinian functions after you, it was for the reasons I gave and not, as you insinuate, *"by way of compensation**" since there is nothing to compensate you for; I will only recognize a property right prior to mine when you can show me that someone had earlier investigated the discontinuity of the groups and the single-valuedness of the functions in even just a slightly general case and had given the series expansions of these functions. I wish to respond to a query in a footnote to a page of your letter. Speaking of the functions defined by Mr. Fuchs in Volume 89 of *Crelle*, you ask: *"Are these really single-valued? All I understand is that over the whole of the range of values taken by them there is no branching**." Here is my response. The functions investigated by Mr. Fuchs subdivide into three large classes: those of the first two are indeed single-valued; those of the third are in general only *without branching**; they are only single-valued if one adds a condition to those given by Mr. Fuchs. These distinctions are not made in the first of Mr. Fuchs's works; they can be found in two additional notes, unfortunately too concise, one of which is in *Borchardt's Journal* and the other in the *Göttingen Nachrichten* of 1880 [Fuc1880, Fuc1881].

I thank you very much for the last note that you had the goodness to send me. The results you state interest me greatly, and here is why: I discovered them myself some time ago, but did not publish them since I wished to clarify the proof a little; that is why I would like to see your proof when you in turn have clarified it.

I hope that the battle, albeit fought with weapons of courtesy, over a name, will not alter our good relations. At every instance not at all wanting to take the offensive, I hope that you will not hold it against me that I went on the defensive. In any case, it would be ridiculous for us to continue falling out over a name. *Name ist Schall und Rauch*[15] and after all that, it's all the same to me: do as you wish, and for my part I will do as I wish.

Please be assured, Monsieur, of my greatest regard.

Poincaré

[15]"A name is just noise and smoke", from Goethe's *Faust*.

XX

Paris, April 7, 1882

Monsieur,

I have the honor of returning to you the corrected page proofs of my letter. Now that our little dispute is over — and I hope it never revives — allow me to thank you for the courtesy which you demonstrated throughout it.

Please be assured, Monsieur, of my highest regard.

Poincaré

XXI

Leipzig, May 7, 1882
Sophienstraße 10

Dear Sir!

A short time ago I read your note in the *Comptes rendus* of April 10 [Poin1882a]. It was all the more interesting to me in that I believe your present considerations, and also methods, are related to mine. I prove my theorems using *continuity*, relying on the the following two lemmas: 1) to every "discontinuous group" there belongs a Riemann surface, and 2) to each suitably cut Riemann surface there belongs just *one* such group (if any group at all belongs to it). Up till now I have not attempted to derive the series developments that you establish. How in fact do you prove the existence of the number m for which $\sum \frac{1}{(\gamma_i \eta + \delta_i)^m}$ converges absolutely? And do you have an *exact* lower bound or an approximate one?

As for me, I have in the meantime been able to give the theorems in question an even more general form, but I shall have to write to you about that later in view of the need to prepare a note for the *Annalen* for which I have insufficient time. In the case of my first theorem, the whole of the closed sphere η with infinitely many points removed is covered by the images of the fundamental region. In the case of the second theorem, the interior of a disc, but of *one only*, remains uncovered. I have now noted the existence of representations (which for each Riemann surface are always unique) for which the case of *infinitely many discs* is excluded. In this direction, I formulate here only the the simplest theorem (in which it is assumed always that one has an unbranched representation of the Riemann surface). Let $p = \mu_1 + \mu_2 + \cdots + \mu_m$, where none of the μ is equal to 1. Choose m points $O_1, \ldots, O_m$ on the Riemann surface, and starting from O_1 make, in the usual way, $2\mu_1$ cuts $A_1, B_1; A_2, B_2; \ldots; A_{\mu_1}, B_{\mu_1}$; from O_2 make $2\mu_2$ cuts, etc. On the other hand on the sphere η, one draws m pairwise disjoint circles and in the interior of the space bounded by the latter taken together a polygon bounded by $4\mu_1$ circular arcs orthogonal to the first fundamental circle, next one bounded by $4\mu_2$ circular arcs orthogonal to the second fundamental circle, etc. (so that each polygon of circular arcs is m-fold connected). The bounding circles are ordered in pairs in accordance with the familiar sequence $A_1, B_1, A_1^{-1}, B_1^{-1}, A_2, B_2, \ldots$, that is, by means of linear substitutions acting on η in each case leaving the fundamental circle invariant. Furthermore, assume the product of the linear transformations in question — for example $A_1 B_1 A_1^{-1} B_1^{-1} \ldots A_{\mu_1}^{-1} B_{\mu_1}^{-1}$ — is always the identity. *Then there is always one and only one analytic function mapping the sectioned Riemann surface onto one of the given polygons of circular arcs.* The case when one of the μ is equal to 1 differs only in that then the corresponding fundamental circle reduces to a point and the corresponding linear substitutions become "parabolic", and fix that point. So, enough for today. Would it be possible to obtain a complete collection of your relevant offprints? After Pentecost I will be giving a series of lectures in my seminar on single-valued functions admitting linear transformations, and would, if possible, like to make available such a collection to my auditors.

Very respectfully yours,

F. Klein

XXII

Paris, May 12, 1882

Monsieur,

I have been slow in replying and I beg your forgiveness; I had to be away for a short time. Like you I believe our methods are very similar and differ less in principle than in the details. As for the lemmas you speak of, the first I proved using ideas related to series expansions, while you, I conjecture, used the theorem of which you spoke in one of your letters of last year. The second lemma presents no difficulties, and it's likely that we both prove it in the same way. Once these two lemmas are established, then that is in fact the point from which I start, as you do. Like you I appeal to continuity, but there are many ways of doing so and it is possible that we differ in some of the details.

You ask me how I established the convergence of the series $\sum \frac{1}{(\gamma_i \eta + \delta_i)^{2m}}$. I have two proofs but they are both too long to be included in a letter; I shall be publishing them shortly. The first is based in principle on the fact that the surface of the fundamental circle is finite. The second requires the same assumption, but is based on non-Euclidean geometry. What is a lower bound of the number m? It's 2. Here if one assumes m to be an integer one has an exact bound. As far as series relating to *zeta Fuchsian* functions are concerned, in contrast I have only an approximate bound. What most interested me in your letter was what you had to say about functions admitting an infinity of circles as their lacunary spaces. I also have encountered such functions and have an example in one or two of my notes. However, I arrived at them by a completely different route from yours. It seems likely that your functions and mine are closely related; it is not at all obvious, however, that they are identical. I am willing to believe that your method and mine can be very extensively generalized so that both yield a large class of transcendental functions including as special cases those we have already met with.

You asked me about offprints of my articles. Do you mean my notes in the *Comptes rendus*? I did not order offprints and it will now be difficult to obtain them, at least for the earlier articles.

I will shortly send you, that is, as soon as I receive them, offprints of my two most recent articles; the first is *On curves defined by differential equations* [Poin1885a]. This concerns the geometric form of curves defined by first-order differential equations. Unfortunately, only the first part of this memoir has as yet appeared and contains only the preliminaries. The second is on the topic of cubic ternary forms, of which I wish to conduct an arithmetic investigation. I wanted first to recall certain algebraic results involved in the first part of the memoir. This first part has appeared only in Issue 50 of the *Journal de l'École Polytechnique*, and the remainder is to appear in Issue 51 [Poin1882b]. Thus the first part will not be of much interest to you. It does contain, however, a study of linear transformations and certain *continuous* groups contained in the linear group in 3 or 4 variables.

By the way, I don't remember if I sent you my dissertation, or earlier articles on differential equations and a work on functions related to lacunary spaces.

Please be assured, Monsieur, of my highest regard.

Poincaré

XXIII

Leipzig, May 14, 1882

Dear Sir!

In answer to your letter, just received, I would like to explain in a few words how I use "continuity". Only in principle, of course, since the detailed exposition, which cost me much effort to write up, can in any case be modified in many ways. I will confine myself completely to unbranched η-functions of the second kind, as I called them in my note. Here it is primarily a question of proving that the two manifolds to be compared — on the one hand the set of systems of substitutions considered and on the other the set of all existing Riemann surfaces — not only have the same dimension ($6p - 6$ real dimensions) but form *analytic* manifolds with *analytic* boundaries (in the sense of the terminology introduced by Weierstrass). By the first lemma stated in my previous letter, these two manifolds are related in a $(1 \longleftrightarrow x)$-valued manner, where, by the second lemma, for the various parts of the second manifold x can take only the values 0 or 1. Now this relation turns out to be *analytic* and, what follows from both lemmas, an analytic relation *whose functional determinant vanishes nowhere*. From this I deduce that x always has the value 1. For in fact if there existed a transition from a region with $x = 0$ to one with $x = 1$, then, in view of the analyticity of the correspondence, the first region would correspond to definite points (those actually attained) of the other region and then for these, contrary to what has already been noted, the functional determinant of the relation would have to vanish. Such is my proof. Mr. Schwarz communicated to me a quite different proof, also based on considerations of continuity, when I visited him recently (on April 11) in Göttingen. Although I haven't obtained his permission to do so, I thought I should tell you about it all the same. He imagines the Riemann surface cut appropriately and then covered infinitely often with the various sheets so connected along the cuts that a complete surface results corresponding to a set of polygons placed side by side in the plane. This complete surface, if one may so term an infinitely extended surface (which is something to be cleared up), is, in the case of an η-function of the second kind (the case first considered by Schwarz), *simply connected and with simple boundary curve*, and then it is a question of seeing whether such a simply connected surface with simple boundary can be mapped in the usual way onto the interior of a disc. In any case, Schwarz's chain of argument is very beautiful.

You mention offprints. I would, above all, not wish to trouble you about that, and the more so in that I can always get hold of all your works with the single exception of your dissertation. However, I would very much like to have as complete a collection as possible of your offprints. Thus I would be very happy if you could send me anything at all (I don't have any).

Have you perhaps had the chance to read Lie's theory of transformation groups? Lie always takes the parameter figuring in his groups to be complex-valued; it would be interesting to see how his results extend to the situation where one considers groups generated only by a *real* iteration of certain infinitely small operations.

Some time ago Hermite sent me a lithographed copy of his *Cours d'analyse*. Would it be possible (of course for appropriate payment) to obtain all such copies? That would be especially useful to me as far as my present aims in my seminar are concerned.

As always

your most devoted

F. Klein

XXIV

Paris, May 18, 1882

Monsieur,

I don't need to tell you how much your last letter interested me. I now see clearly that your proof and mine differ at most in the terminology and details: it is thus possible that we don't establish in the same way the analytic character of the relation between the two *manifolds** you speak of; I myself connect this fact to the convergence of my series, but it is clear that one can obtain the same result by other means.

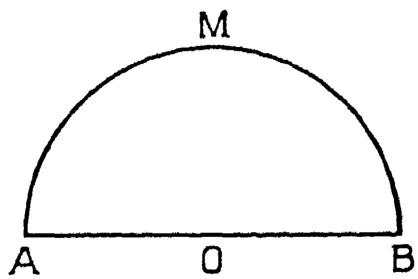


Figure 4: A half-disc

Mr. Schwarz's ideas are of much greater scope; it is clear that the general theorem in question, assuming he has proved it, will find application to the theory of a great many functions and in particular to that of functions defined by *non-linear* differential equations. It was in studying such equations that I myself came to ask if a Riemann surface of infinitely many sheets could be extended over a disc, and in this regard I was led to the following problem, which would allow one to prove that such an extension is indeed possible:

One begins with a partial differential equation

$$X_1 \frac{d^2u}{dx^2} + X_2 \frac{d^2u}{dxdy} + X_3 \frac{d^2u}{dy^2} + X_4 \frac{du}{dx} + X_5 \frac{du}{dy} = 0,$$

and a half-disc AMBO[16]. Here X_1, X_2, X_3, X_4, X_5 are given functions of x and y; these functions are analytic in the interior of the half-disc but not on its perimeter. Can one always find a function u of x and y satisfying the equation, analytic in the interior of the half-disc, and tending to 1 as the point x, y approaches the curved

[16] See Figure 4.

portion of the perimeter and to 0 as it approaches the diameter AOB? All my efforts in this direction have proved fruitless, but I hope that Mr. Schwarz, who has solved the problem in the simplest case, will be more fortunate than I.

I am sending you the offprints of my earlier articles, and I hope to be able to send you soon other more recent memoirs I mentioned to you of which offprints shouldn't be too long in arriving.

Mr. Hermite's lithographed course is published by Hermann, Librairie des Lycées, rue de la Sorbonne; the price of a subscription is 12 francs. I don't believe the editor sent Mr. Hermite any offprints.

Please be assured of my most devoted sentiments and my sincere esteem.

Poincaré

XXV

Leipzig, September 19, 1882
Sophienstraße 10/II

Dear Sir!

While on the point of completing a rather long paper on the new functions, I reread yet once more your article in Volume 19 of the *Annalen* [Poin1882c]. There is a point there that I do not understand. You mention, in two places (in the middle of p. 558 and at the bottom of p. 560), "Fuchsian functions" existing only in a space bounded by infinitely many circles all orthogonal to the fundamental circle. Now I am very familiar with functions (as I wrote to you three months ago) having as their natural boundary an infinitude of circles. However, the corresponding group always contains substitutions leaving invariant only a single limit circle, chosen at random. Now you define as "Fuchsian" those functions *all* of whose substitutions are real (p. 552), and this definition remains essentially unchanged by the generalization on p. 557 where the real axis is replaced by an arbitrary circle. Thus the functions I am familiar with do not fall under your definition of "Fuchsian". Is this a misunderstanding on my part or an imprecision in the formulation on yours? As far as my own work is concerned, I confine myself to expounding the geometric viewpoint, by virtue of which I consider I have defined the new functions in Riemann's sense. One finds, as is only natural, many points of contact with your geometric conception of the subject. The most general groups that I consider are generated by a certain number of "isolated" substitutions and a certain number of groups "with a fundamental circle" (which may be real or imaginary or reduced to a point) "nested in one another". The theorems of my two *Annalen* notes then become special cases of the following general theorem: *to every Riemann surface with arbitrarily prescribed branching and cuts there corresponds one and only one η-function of the type in question.*

I have heard from Mittag-Leffler that you also are at present busy with important work. I don't need to tell you how much I would be interested in knowing more about that work. If in a month's time you're in Paris, you will meet my friend S. Lie, who has been visiting me for a few days and who, although not yet a function-theoretician, is very interested in the progress made latterly in function theory.

With the greatest respect,

F. Klein

XXVI

Nancy, September 22, 1882

Monsieur,

Here are a few details concerning the functions I spoke of in my note in the *Annalen* with natural limit comprised of an infinitude of circles. To simplify the exposition, I shall take by way of example a very special case. Suppose we have four points a,b,c,d on the fundamental circle and four circles meeting it orthogonally: the first at a and b, the second at b and c, the third at c and d, and the fourth at d and a. Thus we have a curvilinear quadrilateral. Let us consider two substitutions (hyperbolic or parabolic), the first sending the circle ab to the circle ad, and the second sending the circle cb to the circle cd. *Iterates** of our quadrilateral will cover the fundamental disc or only part of that surface; however, in every case the group will obviously be discontinuous. One sees easily that the fundamental disc will be entirely covered only in a single case, namely when the four points $abcd$ are harmonic and the two substitutions (ab,ad) and (cb,cd) are parabolic. Thus here we have to do with the modular function. In every other case, one finds that the *iterates** in question cover only a region bounded by an infinity of circles. Now the whole plane can be *mapped** onto our quadrilateral in such a way that two corresponding points on the perimeter correspond to the same point of the plane. This *mapping** determines a function defined only on the region covered by the *iterates**. However, there is an important point to make here. The group generated by the two substitutions (ab,ad) and (cb,cd) may be considered as generated in another manner. Consider four circles C_1,C_2,C_3,C_4 meeting the fundamental circle orthogonally and such that none of them contains any other. Consider two substitutions interchanging C_1 with C_2 and C_3 with C_4; the group they generate is obviously discontinuous and, for a suitable choice of the four circles, it will be the same as the group considered above. The part of the plane exterior to the four circles is a sort of quadrilateral which can be *mapped** onto a Riemann surface of genus 2 and which therefore determines a function defined on the whole plane. So there you have one and the same group giving rise to two essentially different functions. One can ask in this connection a great many delicate questions which I shall not go into here.

In sum, you see that we can have functions defined only in a region bounded by an infinitude of circles and yet which are still "Fuchsian functions" since all the group's substitutions preserve the fundamental circle. Each of the circles of the boundary is preserved by a substitution from the group, which also preserves the fundamental circle. You know, of course, that every hyperbolic substitution preserves all circles passing through the two double points.

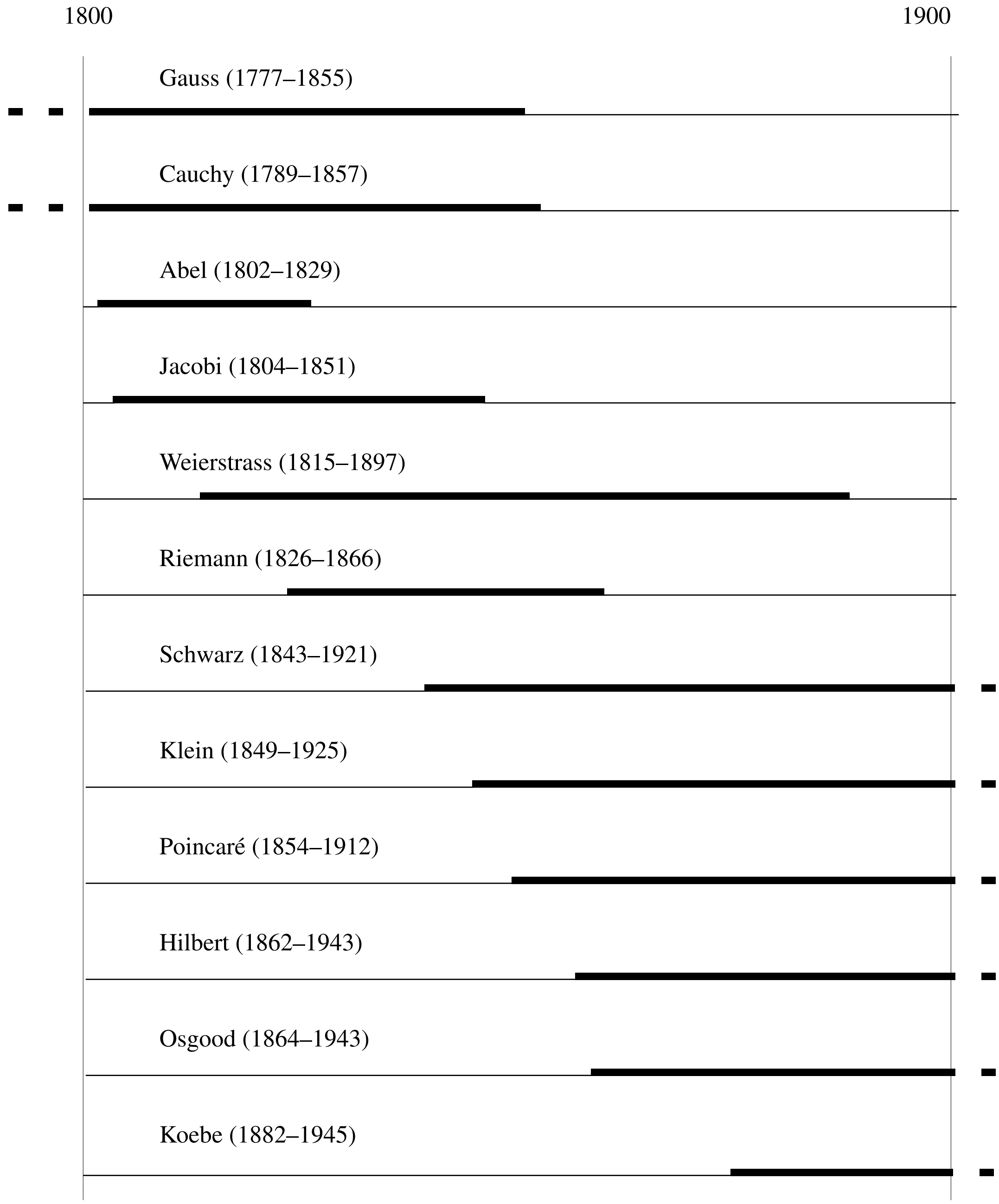
I learn with pleasure that you are preparing a substantial piece of work of the topic of common interest. I shall read it with the greatest pleasure. As Mr. Mittag-Leffler told you, I myself am writing a paper on that theme; but in view of its length I am dividing it up into five memoirs: the first, which will appear this year, on groups of real substitutions (those I called Fuchsian groups) [Poin1882c]; the second on Fuchsian functions (I will shortly complete writing this up) [Poin1882d]; and the third on the more general groups and functions that I call Kleinian [Poin1883a]. In the fourth, I will address a range of issues left aside in the second memoir — the proof of the existence of functions satisfying certain conditions, such as, for example, the proof of the fact that to every Riemann surface there corresponds an analogous function and the determination of the relevant constants [Poin1884b].

Finally, in the fifth, I will discuss zetafuchsian functions and the integration of linear equations [Poin1884b].

I have to return to Paris the day after tomorrow; thus I will be there when Mr. Lie is visiting. I would be sorry to miss the opportunity of seeing this famous geometer. You should have received the first part of my paper on curves defined by differential equations. I will send you the second part shortly and at the same time my memoir on cubic forms.

Please be assured, Monsieur, of my highest regard.

Poincaré


1800
1900
Gauss (1777–1855)
Cauchy (1789–1857)
Abel (1802–1829)
Jacobi (1804–1851)
Weierstrass (1815–1897)
Riemann (1826–1866)
Schwarz (1843–1921)
Klein (1849–1925)
Poincaré (1854–1912)
Hilbert (1862–1943)
Osgood (1864–1943)
Koebe (1882–1945)

Some historical reference points

In the following pages, we give some chronological reference points showing how the "uniformization adventure" fits into the general historical context. For each decade from 1800 to 1910, we list certain events of significance for the time in question, grouped under the following heads:

— the theme of this book: The Uniformization Theorem;

— mathematics in general;

— science;

— technology;

— the arts, philosophy, and the humanities;

— history (Franco–German, since it was in these two countries that the adventure's protagonists lived).

It goes without saying that the choice of such events is extremely subjective, especially with respect to the last headings. Despite this, we hope that these chronological glimpses will help the reader to better situate the uniformization theorem within its scientific, cultural, social, and historical contexts.

The "light" character of these few pages allows us to dispense with exact references for this rather disconnected inventory; they were inspired by the great number of friezes or historical chronologies one comes across here and there.

1800–1809

Uniformization

1806 : The *geometric representation of complex numbers*, sketched by Euler, is formalized by Argand and Buée.

Mathematics

1801 : Gauss publishes “Disquisitiones arithmeticae”. This book remains a model of rigor. Devoted to the theory of numbers, it above all contains such pleasures as *quadratic reciprocity* and criteria for the *construction of regular polygons by straight-edge and compasses*, prefiguring Galois theory.

1803 : L. Carnot publishes his book “The geometry of position”, in which he systematically gives signs to oriented geometric entities.

1806 : Legendre introduces the *method of least squares*, allowing the best fitting of a curve to possibly flawed experimental data. This method will play an important role in the development of the experimental sciences.

1806 : Poinsot discovers the last two regular stellated Kepler–Poinsot polyhedra.

1807 : Fourier submits his first important memoir “On the propagation of heat in solid bodies”, in which he decomposes a periodic function into an infinite sum of harmonics: *harmonic analysis* is born.

Sciences

1800–1805 : Cuvier publishes the five volumes of his “Lessons in comparative anatomy”.

1801 : In “On the theory of light and colours”, Young brings out the *wave character of light*.

1801 : Pinel publishes his “Medico-philosophic treatise on insanity”: the first classification of mental illnesses.

1802 : Publication of *Dalton’s law* on pressure in gases.

1804 : In his “Chemical investigations into plant life”, de Saussure proves that water is consumed in *photosynthesis*.

1807 : Davy isolates the elements *sodium* and *potassium*.

1808 : Dalton, in his “New system of chemical philosophy”, proposes an *atomic theory* according to which all matter is made up of a (small) number of elements.

1808 : Malus discovers the *polarization of light*.

1809 : Lamarck publishes “Zoological philosophy”, proposing his theory of “transformism”, the first attempt to give a materialistic and mechanistic explanation of the *evolution of living creatures*.

Technological Advances

1801 : Jacquard develops the *loom*.

1801 : Volta produces an *electric battery*.

1803 : Fulton brings into operation the first *steampowered boat* on the Seine.

1804 : Trevithick builds the first *locomotive*.

1805 : Appert develops the first effective technique for *conserving* foodstuffs.

1806 : *Morphine* is isolated by Setürner. It will be used for therapeutic purposes in 1827.

1806 : The British engineer William Murdoch installs gas lighting in a cotton spinning mill in Manchester, U.K. First gas plant opened at Salford.

The Arts, Philosophy, et Humanities

1801 : In “René”, Chateaubriand describes for the first time the “wave of passions”, destined to become a commonplace of Romanticism. The next year will see the appearance of the “Genius of Christianity”, in which he claims to prove that Christianity is no less favorable to art than the “fictions” of Antiquity.

1802 : The philologist Grotefend deciphers cuneiform script.

1804–1805 : Beethoven composes the “Appassionata”, his sonata for piano No. 23 in F minor, described as a “torrent of fire in a bed of granite” by the writer Romain Rolland.

1808 : Appearance of the first part of Goethe’s “Faust”.

1808 : Ingres paints “The Valpinçon Bather” (Fr: “La Grande Baigneuse”), the first picture in a series of depictions of bathing women that he will continue producing for the rest of his life.

1809 : First volume of the “Description of Epypt or compilation of observations and investigations carried out in Egypt during the French expedition”.

Franco–German history

1804 : Napoleon promulgates the French *code of civil law* (the so-called “Napoleonic code”), defining new rights and obligations of the French people.

1804 : Napoleon Bonaparte has himself proclaimed *emperor of the French* by the Senate in May, and then crowns himself in Notre-Dame Cathedral on December 2. The imperial symbolism is designed to recognize Napoleon I as heir of Charlemagne and the Roman Empire.

1806 : On August 6, the last emperor of the Germanic *Holy Roman Empire*, submissive to Napoleon’s ultimatum, renounces his crown. Prussia reacts by declaring war on Napoleon. In October, the battles of Jena and Auerstädt end in total victory for Napoleon’s “Grande Armée” over the Prussian army. Napoleon enters Berlin and parades his troops along Unter den Linden. The trauma resulting from this humiliation will trigger a violent *German nationalism*, developing over the course of the 19th century and leading to the *unification of the German nation*.

1810–1819

Uniformization

1811 : In a letter to Bessel, Gauss explains *integration of holomorphic functions along paths*.

1814 : First of Cauchy's articles on the *theory of residues*.

Mathematics

1810 : Gergonne publishes the first volume of the "Annales de mathématiques pures et appliquées", which will play an important role in the diffusion of mathematics.

1811 : Poisson publishes his "Treatise on mechanics", containing applications to electricity, magnetism, and of course mechanics.

1812 : The two volumes of Laplace's "Analytic theory of probabilities" establish probability at the center of mathematics.

1814 : Publication of "Barlow's tables", containing squares, cubes, square roots, logarithms, etc. of the integers from 1 to 10,000.

1815 : Pfaff publishes an article on *systems of partial differential equations*, which will later influence Jacobi and Lie.

1817 : Bolzano publishes "Rein analytischer Beweis ... ". One finds here, in particular, the infinite-simal-free definition of continuous functions and the Bolzano–Weierstrass theorem.

1817 : Bessel discovers a new class of transcendental functions, today called *Bessel functions*, satisfying a second-order algebraic differential equation.

Science

1811 : Avogadro proposes the idea that equal volumes of gas at the same temperature and pressure contain the same number of molecules.

1813 : Von Fraunhofer invents the *spectroscope*, permitting him to identify the lines of the solar spectrum (1814) and later to classify stars according to their light spectra (1822).

1815 : Prout states that the *atomic weights* of the elements are multiples of that of hydrogen.

1815 : Ampère establishes the difference between an *atom* and a *molecule*.

1816 : Fresnel shows that the phenomena of *interference* and *diffraction* can be explained in terms of the wave theory of light.

1817 : Pelletier and Caventou isolate *chlorophyll*.

Technological advances

1816 : Laennec invents the *stethoscope*.

1816 : The *Davy lamp* is a safety lamp allowing miners to work in the presence of inflammable gases such as firedamp (methane).

1817 : Drais invents the *draisine* (or *velocipede* (bicycle)).

1819 : The steamship Savannah crosses the Atlantic Ocean (although part of the voyage is by sail).

1819 : Pelletier isolates *quinine*.

The Arts, Philosophy, and the Humanities

1811–1816 : Four of Jane Austen's novels appear in succession: "Sense and Sensibility", "Pride and Prejudice", "Mansfield Park", and "Emma".

1812 : First publication of "Kinder und Hausmärchen" (Children's and household tales) by the brothers Grimm.

1814 : Goya paints "El tres de mayo", portraying the execution of Spanish prisoners by soldiers of Napoleon's army on May 3, 1808, following a revolt the day before. The painting was commissioned by the provisional government of Spain at Goya's suggestion.

1817 : Friedrich paints "Der Wanderer über dem Nebelmeer" (Wanderer above the sea of fog), emblematic of *German romantic art*.

1818–1819 : Géricault paints "Le Radeau de la Méduse" (The raft of the Medusa).

Franco–German history

1812 : The *Russian campaign*, marked by the burning of Moscow by the Russians the day after the entrance of the French into the city, ends with a catastrophic retreat of Napoleon's army through Poland and Germany. Cold, snow, and harassing attacks by Cossacks result in considerable losses; the "Grande Armée" is destroyed.

1813 : Napoleon succeeds in assembling an army of 200,000 men but the *Battle of Nations* at Leipzig results in a decisive defeat for the French army opposed to a coalition made up of almost all of Europe.

1814 : Napoleon is forced to abdicate and is exiled to the island of Elba. The French nobility is restored and the allies install Louis XVIII on the throne. The defunct Germanic *Holy Roman Empire* is replaced by a *German Confederation* of thirty-nine states under the nominal leadership of the Habsburgs.

1815 : Napoleon returns to France at the head of a small army and succeeds in resuming power. However, the allies do not accept his return and take up arms once again against France; Napoleon's army is finally defeated at the *Battle of Waterloo*. Napoleon is exiled to Saint Helena. He leaves behind a France bled white.

1820–1829

Uniformization

1825 : Gauss proves his *theorem on local uniformization*: every surface is locally conformally equivalent to the Euclidean plane. This result will later allow an arbitrary *Riemannian surface* to be viewed as a *Riemann surface.*

1827–1829 : Abel and Jacobi understand that the inverses of elliptic integrals are doubly periodic *single-valued* functions; the way is open to the uniformization of *elliptic curves.*

Mathematics

1821 : Cauchy publishes his “Cours d’analyse”, presenting mathematical analysis as rigorously as possible.

1824 : Abel proves that the general equation of degree 5 is not solvable by radicals.

1825 : Fermat’s last theorem is proved for $n = 5$ by Dirichlet and Legendre.

1826 : Lobachevsky announces that he has developed a geometry in which *Euclid’s fifth postulate* does not hold: through a point off a straight line there are infinitely many straight lines parallel to that line.

1826 : Founding of the “Journal für die reine und angewandte Mathematik” by Crelle.

1828 : In his “Disquisitiones generales circa superficies” Gauss proves his “Theorema Egregium”, which opens the way to the concept of the *curvature* of an *abstract* surface, independent of how it is embedded in space.

1829 : Proof of Sturm’s theorem, permitting the calculation of the number of distinct real roots of a polynomial contained in a given interval.

Science

1820 : Oersted observes that an electric current produces a *magnetic field*, thus initiating the study of *electromagnetism*.

1821 : Arago and Gay-Lussac invent the *electromagnet*.

1822 : By studying the Rosetta stone, Champollion manages to decipher Egyptian hieroglyphs.

1824 : Carnot publishes his “Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance” (Reflections on the motive power of fire and the machines appropriate to develop this power). This contains his theory of the Carnot cycle, which will be fundamental to the development of steam engines and more generally of *thermodynamics*.

1825–1828 : Publication of the three volumes of Legendre’s “Traité des fonctions elliptiques et intégrales eulériennes". This monumental work, a standard reference for the manipulation and calculation of elliptic integrals, is, however, superseded in its theoretical point of view by work of Gauss, Abel, and Jacobi.

1827 : The botanist Brown observe the erratic movements of particles of pollen in suspension in a liquid; this is *brownian motion*.

1829 : Lyell publishes his “Principles of Geology” in which he affirms, in particular, that the Earth’s surface changes very slowly but continuously; this will be an important idea for the theory of evolution.

Technological advances

1820 : Fabian von Bellingshausen discovers the *continent Antarctica*.

1821 : Mary Anning discovers the first complete skeleton of a plesiosaurus.

1821 : Faraday perfects the *electric motor*.

1822 : Mantell discovers the first fossils of *dinosaurs*: the teeth of an iguanodon.

1823 : Macintosh files a patent for his *waterproof cloth*.

1825 : Chevreul and Gay-Lussac manufacture candles made of soapstone (steatite).

1825 : Inauguration of the first *commercial railroad line* Stockton–Darlington; first passenger train drawn by the steam locomotive invented by Stephenson.

1826 : Moray builds the first *internal combustion engine*, fueled by a mixture of ethanol and turpentine.

1826 : The first *photograph*, "Point de vue du Gras" (View of le Gras), is taken by Niépce.

The Arts, Philosophy, and the Humanities

1823 : Pushkin begins writing his *chef d'œuvre* "Eugene Onegin".

1824 : Beethoven composes his ninth and last symphony.

1826 : Mendelssohn composes his overture to "A midsummer night's dream".

1827 : A year before his death, Schubert composes "Winterreise" (Winter voyage).

1829 : Chopin composes his first book of *Études* at the age of 19.

1829 : In May 1829, the young Gérard, not yet calling himself de Nerval, is called upon by Hugo to lend his support to Hugo's play "Hernani". The year before, he had published a translation of "Faust", serving as the inspiration of the composer Berlioz's opera "La Damnation de Faust". Like Galois he spent some of the year 1831 in Sainte-Pélagie prison.

1829 : "William Tell", Rossini's last opera, opens in Paris on August 3.

Franco–German history

1824 : On the death of Louis XVIII, his brother accedes to the throne as Charles X, at the age of 66. Charles X revives the tradition of the coronation at Reims on May 29, 1825, surrounded by pomp recalling the *Ancien Régime* at its height. His reign will be marked by the domination of the ultra-royalists, and he will alienate the public with his granting of indemnities to the *émigrés*, his law punishing by death anyone guilty of sacrilege, and his reintroduction of censorship.

1830–1839

Uniformization

1835 : Jacobi invents *ϑ-functions*, thereby providing a very general procedure for constructing elliptic functions.

Mathematics

1831 : Cauchy proves that holomorphic functions can be developed in *power series*.

1832 : Bolyai's work on *non-Euclidean geometry* is published.

1832 : Galois dies following a duel. During the night before his death, he sketches in haste the theory which today bears his name and will exert a profound influence on almost all of mathematics.

1834 : Hamilton publishes "On a general method in dynamics".

1835 : Quételet publishes "Sur l'homme et le développement de ses facultés, essai de physique sociale" (On man and the development of his faculties, an essay on social physics).

1837 : In his book "Recherches sur la probabilité des jugements" (Investigations into the probability of judgements), Poisson introduces the law of probability bearing his name and applies it to concrete problems related to the functioning of law courts.

1837 : Dirichlet publishes the first version of his *theorem on arithmetic progressions*, thereby founding analytic number theory.

Science

1833 : In his work "On a general method of expressing the paths of light and of the planets by the coefficients of a characteristic function" Hamilton founds *Hamiltonian dynamics*, centered on the concept of *action*. The principle according to which the action is *stationary* generalizes the principle of least action and remains today one of the pillars of science.

1834 : Payen and Persoz isolate the first *enzyme*: diastase.

1834 : Babbage conceives of a *mechanical programmable calculating machine* using punched cards.

1838 : Schleiden observes that plants are constituted of *cells* and shows the importance of their *nuclei*.

Technological advances

1831 : The discovery of *electromagnetic induction* by Faraday allows the production of alternating current. Pixii builds the first *alternator*, soon to be improved.

1833 : Gauss and Weber develop the *electromagnetic telegraph*.

1835 : Colt invents the *revolver*.

1835 : Invention of *Morse code.*

1836 : Madersperger invents the *sewing machine.*

The Arts, Philosophy, and the Humanities

1830 : Publication of the third and last version of the "Enzyklopädie der philosophischen Wissenschaften im Grundrisse" (Encyclopedia of the philosophical sciences) in which Hegel expounds his system of philosophy.

1830–1842 : Comte expounds the principles of *scientific positivism* in his "Cours de philosophie positive".

1831 : Delacroix's painting "La Liberté guidant le peuple" (Liberty guiding the people), depicting the July revolution, is exhibited at the Salon de Paris.

1831 : "La Peau de chagrin" (The magic skin), published in the series "Romans et contes philosoph-iques" (Philosophical novels and stories), is the first volume of what will become Balzac's "Comédie humaine".

1831 : On December 26, "Norma" is performed at La Scala, Milan.

1835 : Chopin publishes his "Ballad No. 1 in G minor", an *odyssey of the soul of Chopin*, as Liszt later describes it, mingling happiness, melancholy, sadness, and delight.

1837 : Büchner dies, leaving his "Woyzeck" unfinished.

1839 : Stendhal's "La Chartreuse de Parme" is published in two volumes..

Franco–German history

1830 : On May 11, using as pretext an altercation between the Dey of Algiers and the French consul, Charles X launches the *Algerian campaign*. This date marks the beginning of the acquisition of a *second French colonial empire* (the first having been entirely dissolved at the fall of the first Empire).

1830 : In July, an attempt by Charles X to dissolve the Chamber of deputies, change the electoral law, and abolish the freedom of the press, triggers a riot rapidly escalating into a *popular revolution*. Barricades are erected everywhere in Paris, the armed forces are caught in the middle, and Charles X is forced to flee. The liberal deputies, for the most part monarchists, succeed in saving the situation, and are able to preserve constitutional monarchy at the expense of a dynastic change. Louis-Philippe I is proclaimed "king of the French". He will be the "bourgeois king"; his reign will be characterized by rapid development and enrichment of the manufacturing and financial *bourgeoisie*.

1840–1849

Uniformization

1844 : In his memoir on *higher geodesy*, Gauss introduces the word *conformal*.

1847–51 : Publication of Eisenstein's articles on *elliptic functions*: the series he introduces allow the coefficients of the polynomial appearing in an elliptic integral to be related to the periods of the inverse function.

1847 : In Göttingen, Riemann attends Eisenstein's course on elliptic functions.

Mathematics

1841 : Quételet opens the first *bureau of statistics*. Statistics develops and becomes a mathematically-based field in its own right.

1843–1845 : The concept of a vector space of *n dimensions* is independently formulated by Cayley and Grassmann.

1843 : Hamilton discovers the *quaternions*.

1844 : Liouville describes the first *transcendental numbers*.

1845 : Cayley studies the composition of linear mappings in his "Theory of linear transformations".

1847 : Boole shows that one can *algebraicize* logic in "The Mathematical Analysis of Logic".

1847 : Kummer publishes "Über die Zerlegung der aus Wurzeln der Einheit gebildeten complexen Zahlen in ihre Primfactoren" (On the decomposition into prime factors of complex numbers formed from roots of unity). He succeeds in proving Fermat's last theorem for all integers *n* up to 100 except for 37, 59, and 67.

1847 : Von Staudt publishes "Geometrie der Lage" (Positional geometry), the first metric-free treatment of projective geometry.

Science

1843 : Joule and Mayer independently demonstrate the *equivalence between heat and mechanical energy*.

1843 : Schwabe discovers the cycle of *sunspots*.

1844 : Darwin writes, but does not publish, an essay prefiguring his *theory of evolution of species*.

1847 : Helmholz formulates the *law of conservation of energy* in "Über die Erhaltung der Kraft".

1848 : Kelvin proposes an *absolute temperature scale* which will be named after him.

1849 : Fizeau measures the *speed of light* using rotating cog wheels.

Technological advances

1841 : The *steam hammer*, an industrial forging tool, is improved by Bourdon, resulting in a leap forward in metallurgy, then a rapidly developing industry.

1842 : Long develops the process for inducing *anaesthesia* using ether. He convinces James Venable, one of his patients, that he can remove a tumor painlessly.

1844 : Manzetti suggests the feasibility of a *talking telegraph*, the future *telephone*, which will be gradually perfected.

1846 : Hoe invents the *rolling press*, allowing more rapid printing.

1847 : Krupp produces his first *steel cannons*.

1848 : Lambot and Monier invent *reinforced concrete*.

The Arts, Philosophy, and the Humanities

1840 : Appearance of the second volume of de Tocqueville's "Democracy in America".

1840 : Publication of Schopenhauer's "Über die Grundlage der Moral" (On the foundations of morality).

1844 : Heine, living in Paris since 1831, makes a last voyage to Germany, where his works are banned, and writes "Deutschland: Ein Wintermärchen" (Germany: A Winter's tale).

1844 : Turner paints "Rain, Steam and Speed — The Great Western Railway".

1845 : Schumann completes his "Concerto for piano in A minor op. 54".

Franco–German history

1848 : On July 23, incited by republicans, Paris rises. Louis-Philippe, refusing to begin an assault on the Parisians, is forced to abdicate. The revolutionaries install a republican provisional government, removing the July monarchy and proclaiming the *second republic* on February 25, 1848. In December, Louis-Napoleon Bonaparte, Napoleon's nephew, is elected *president of the French republic* by universal male suffrage.

1848 : In March 1848, at the news of the Parisian and Viennese revolutions, rebellion ignites also in Germany. The revolutionaries wish to create a united and democratic Germany. They confront the king of Prussia Frederick-William IV, who shortly thereafter convokes a constituent assembly but soon dissolves it when the balance of power goes over to him. The revolutions of 1848 are destined to fail in securing German unity, however the majority of German states concede constitutions mollifying the liberal bourgeoisie.

1850–1859

Uniformization

1851 : In his dissertation, Riemann develops a systematic theory of *holomorphic functions*; he defines the surfaces bearing his name, explains how to manipulate them and how to construct meromorphic functions on them; he "proves" the Riemann Mapping Theorem.

1854 : Weierstrass publishes an article expounding his viewpoint of elliptic functions developed in normally convergent series of functions.

1857 : In an article in Crelle's journal, Riemann creates a general theory of *algebraic functions* and *Abelian integrals*. There he explains, in particular, how to view algebraic curves as compact Riemann surfaces, and describes the topology and analytic geometry of these surfaces, their moduli spaces, etc.

Mathematics

1852 : Chasles introduces the concepts of the *cross ratio*, a *pencil* of conics, and *involutions* in his "Treatise on geometry".

1854 : In his Habilitation dissertation "Über die Hypothesen welche der Geometrie zu Grunde liegen", Riemann lays the foundations of *differential geometry*, introducing in particular the concept of an n-dimensional manifold.

1854 : Cayley defines the notion of an abstract *group* of permutations.

1858 : Dedekind proposes a rigorous way of constructing the real numbers by means of *Dedekind cuts*.

1859 : Riemann publishes a memoir entitled "Über die Anzahl der Primzahlen unter einer gegebenen Grösse", in which he uses, among other things, the *zeta function* to estimate the number of primes less than a given number.

1859 : Mannheim invents the modern *slide rule* with its sliding scale and cursor.

Science

1850 : Clausius introduces the concept of *entropy* and formulates the *second law of thermodynamics*, generalizing the Carnot cycle.

1851 : The *Foucault pendulum* affords an experimental demonstration of the rotation of the Earth.

1859 : Darwin publishes "On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life".

Technological advances

1850 : A *commercial undersea telegraphic cable* is laid between France and England, which will keep functioning for more than 40 years.

1852 : Giffard constructs the first *dirigible airship* which, by virtue of a steam engine installed in its gondola, can change direction relative to the wind.

1853 : Otis founds the Otis Elevator Company and sells the first *reliable elevators*, thus paving the way to the construction of skyscrapers.

1859 : Colonel Drake drills the first *oil wells* in Pennsylvania.

The Arts, Philosophy, and the Humanities

1850 : "Les Chouans" is the final volume of Balzac's "La Comédie humaine".

1850 : Melville begins writing "Moby Dick".

1851–1854 : Comte writes his "Système de politique positive": he has gone from *scientific positivism* to *religious positivism*.

1852–1853 : Liszt composes his "Sonata in B minor", dedicated to Schumann.

1856 : "Madame Bovary" is serialized in the "Revue de Paris". The following year the manager of the Revue, the publisher, and Flaubert will be tried for "contempt for public and religious morals and good ethics"; they will be acquitted.

1857 : Publication of "Les Fleurs du Mal", a collection including almost all of Baudelaire's poetic production since 1840.

Franco–German history

1850 : This year marks the beginning of a period of *industrial development* of great importance for Germany. This development, affecting all of Germany's states, culminates in an economic union around Prussia, foreshadowing the political union.

1851–1852 : The *coup d'état* of December 2, organized by Louis-Napoleon Bonaparte, puts an end to the Second Republic. The National Assembly is dissolved and the prince-president's mandate extended to ten years. Democratic-socialist and republican deputies are exiled in great numbers, and only the conservative press is authorized to publish. A few months later, the title "imperial highness" is reinstated by referendum. Louis-Napoléon Bonaparte becomes officially "Napoleon III, Emperor of the French" on December 2, 1852. From 1859 his regime will take a significantly more liberal turn.

1853 : Napoleon III appoints Haussmann Prefect of the Seine. Till 1870 he will undertake a radical transformation of Paris, symbolic of the era of economic development and explosive capitalism characterizing the Second Empire.

1860–1869

Uniformization

1865 : Clebsch proves that every compact Riemann surface of genus zero is biholomorphic to the Riemann sphere and every compact Riemann surface of genus 1 is biholomorphic to the quotient of the plane by a lattice.

1866 : Fuchs publishes his articles on second-order *linear differential equations* which will serve as a basis for those of Poincaré.

1869 : Schwarz publishes *explicit examples of uniformization* of open subsets of the plane with polygonal boundaries. The connection he discovers between the *Schwarzian derivative* and the uniformization problem is the direct precursor of Poincaré's approach to uniformization via differential equations.

Mathematics

1861 : Weierstrass constructs a continuous, nowhere differentiable curve.

1863 : Weierstrass's construction of the set of real numbers.

1865 : Plücker introduces the coordinates bearing his name identifying the lines of the (complex projective) space of dimension 3. This is one of the first occasions where a *space* is considered whose elements are not necessarily points.

1868 : Beltrami constructs the first *concrete model* of the non-Euclidean geometry of Lobachevsky and Bolyai.

Science

1861 : Bunsen and Kirchhoff lay the foundations of *spectrography*.

1861 : Pasteur refutes the theory of *spontaneous generation*.

1862 : Pasteur publishes his theory of germs: infections are caused by *micro-organisms* which reproduce themselves.

1862 : Foucault makes a precise measurement of the *speed of light*.

1865 : *Maxwell's equations*, appearing for the first time in "A Dynamical theory of the electromagnetic field", unite the electric and magnetic fields into a single *electromagnetic* field.

1866 : In "Versuche Über Pflantenhybriden", Mendel interprets *heredity* in terms of pairs of dominant or recessive traits.

1869 : Mendeleev and Meyer independently draw up the *periodic table of the elements*, in particular leaving gaps allowing for elements not yet discovered.

Technological advances

1860 : In London, Fowler contributes to the construction of the first *underground railway*, the "Metropolitan line", a short line built using the cut and cover method.

1862 : Gatling applies for a patent of his *machine gun*, employing from six to ten rotating barrels thus allowing the parallelization of the firing mechanism and time for each barrel to cool down without reducing the rate of fire.

1866 : The *Sholes and Glidden typewriter* (later known as the "Remington No. 1") is the first such to enjoy commercial success.

1867 : Nobel patents his invention of *dynamite*, the first powerful explosive that is both cheap and *stable*.

1867 : The engineer Eiffel sets up his workshops.

The Arts, Philosophy, and the Humanities

1862 : Publication of Hugo's "Les Misérables", an immediate great popular success.

1865–1869 : Tolstoy's "War and Peace" is published in "Russkiĭ Vestnik". The novel relates the history of the Napoleonic Wars as they affected Russia.

1866 : Courbet paints "L'Origine du monde" (The origin of the world).

1867 : Appearance of Book 1 of Marx's "Das Capital", devoted to the development of *capitalist production*. Books 2 and 3 will be published in 1885 and 1894 by Engels from Marx's drafts following on the latter's death.

1869 : Publication of the first part of Jules Verne's novel "Twenty thousand leagues under the sea". The second part will appear the following year.

Franco–Germanic history

1864 : In France, the "crime of association" is repealed, giving workers the right to force improvements in their working conditions by organizing strikes.

1867 : The *North German Confederation*, uniting the twenty-two German states north of the river Main, is formed at the initiative of the Prussian minister-president von Bismarck following the Prussian victory over Austria and the dissolution of the German Confederation. *De facto*, the Kingdom of Prussia simply annexes those states that had supported Austria. Although Bismarck refrained from including the south German states in order not to provoke Napoleon III, those states do sign treaties of military alliance with Prussia.

1870–1879

Uniformization

1870 : Schwarz invents his *alternating method*, allowing him to prove that open subsets of a Riemann surface that are simply connected, relatively compact and have polygonal boundaries are biholomorphic to the disc.

1878–79 : Klein publishes a series of articles on *modular equations*. These articles contain, in particular, the first example of explicit uniformization of a Riemann surface of higher genus (albeit with a finite number of points removed).

Mathematics

1871 : Dedekind introduces the notions of *field, ring, module, ideal*.

1872 : In his *Erlanger Programm*, Klein views a geometry as a set of structures left invariant by a group of transformations.

1872 : Founding of the *Mathematical Society of France*.

1872 : Sylow proves his theorems on finite groups in “Theorems on groups of substitutions”.

1873 : Hermite proves the *transcendence of the number e*.

1874 : Cantor lays the foundations of the *theory of sets*, introducing, in particular, the idea of countability.

1878 : Sylvester founds the “American Journal of Mathematics”.

Science

1872 : Bolzmann explains that the growth of entropy and the second law of thermodynamics can only be understood in terms of large populations of particles, and not in terms of their individual trajectories; this marks the beginning of *statistical mechanics*.

1873 : Schneider describes *chromosomes* undergoing mitosis.

1873 : In “Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires” (Experimental statics and liquids subject only to molecular forces), Plateau shows experimentally that, in the absence of external forces, the surface of a liquid minimizes its area.

1878 : Gibbs develops the general concept of *thermodynamic equilibrium* in “On the equilibrium of heterogeneous substances”.

Technological advances

1873 : The first mass-produced commercial *automobile*: “L’Obéissante” (The obedient one) of Amédée Bollée.

1877 : Edison invents the *phonograph* and tests it by recording “Mary had a little lamb”.

1879 : Edison presents his first *electric light bulb*. In the same year a 7 kW hydroelectric power plant is built at Saint-Moritz.

1879 : Beginning of *commercial exploitation of the telephone* in France.

The Arts, Philosophy, and the Humanities

1871 : The premiere of Verdi’s opera “Aïda” is held at the Khedival Opera House in Cairo.

1872 : Monet paints “Impression, soleil levant” (Impression, rising sun), a view of the former port of Le Havre. It is this picture that will lend its name to the *impressionist movement*.

1873 : Rimbaud writes “Une saison en enfer” (A season in hell).

1875 : The premiere of Bizet’s “Carmen” at the Paris Opéra-Comique on March 3.

1876 : The Bayreuth “Festspielhaus”, specially conceived by Wagner for the performance of his operas, opens its doors on the occasion of the performance of “The Ring of the Nibelung”.

1879–1880 : “The brothers Karamazov”, Dostoevsky’s last novel, is serialized in the magazine “Russkiĭ Vestnik”.

Franco–Germanic history

1870 : On July 19 France declares war on Prussia. The French lack of preparation leads to disaster: Napoleon III is taken prisoner and on September 19 the siege of Paris begins.

1871 : On January 18, the *German Empire* is proclaimed — very symbolically at Versailles. The constitution provides for the Reichstag to pass laws, but real power rests in the hands of William I and his chancellor Bismarck.

1871 : On January 28, France officially capitulates. Alsace and part of Lorraine become German (again). The National Guard and workers of Paris, refusing to accept defeat, take control of the capital and install an insurrectional government: this is the *Paris Commune*. With tacit Prussian agreement, they are crushed during *bloody week* by the Thiers government, removed to Versailles.

1880–1889

Uniformization

1881 : Klein presents his view of Riemann's works; he develops hydrodynamic and electrostatic interpretations of results on the existence of meromorphic forms.

1881 : Poincaré's investigations of second-order linear equations lead him to the "invention" of Fuchsian groups.

1882 : Following on Poincaré's articles on Fuchsian groups, Klein and Poincaré "prove" the uniformization theorem for algebraic Riemann surfaces using the *method of continuity*.

1883 : Poincaré proves his *uniformization theorem for functions*: the Riemann surface associated with a germ of any analytic function can be parametrized by an open subset of the plane (with possible branch points).

Mathematics

1882 : Lindemann proves the *transcendence of* π, hence showing the impossibility of squaring the circle by means of straight-edge and compasses.

1884 : Volterra publishes his work on integral equations.

1885 : In his memoir "On curves defined by differential equations", Poincaré creates the theory of *dynamical systems*.

1888 : Engel and Lie publish the first volume of "Theorie der Transformationsgruppen", laying the foundations of the *theory of Lie groups*.

1889 : In his memoir on the three-body problem, Poincaré discovers the possibility of *deterministic chaos*.

1890 : Peano exhibits a *continuous curve completely filling a square*.

Science

1882 : Tesla invents the *induction motor* using the idea of a rotating magnetic field.

1883 : Roux suggests that threads visible in the nuclei of cells are vectors of heredity.

1885 : Pasteur *vaccinates* the young Joseph Meister, bitten by a rabid dog.

1886 : Hertz demonstrates *radio waves* experimentally.

1887 : Michelson and Morley demonstrate the *non-existence of the ether* experimentally by measuring the speed of light in various directions.

Technological advances

1885 : Benz develops the "Benz Patent Motorwagen", installing a cylinder generating 560 watts on a tricycle.

1886 : Mergenthaler's *Linotype machine* is used to print the "New York Tribune": this is a machine employing an alphanumeric keyboard allowing the composition of a complete line of text to be cast as a single lead piece.

1887 : Berliner proposes a *zinc disc recording machine* with horizontal sound groove. (Edison's device used a cylinder with a vertical groove.)

1888 : The name "Kodak" appears for the first time with the launching of the first photo apparatuses using photographic film.

The Arts, Philosophy, and the Humanities

1880–1881 : Henry James's "The Portrait of a Lady" is serialized in the "Atlantic Monthly" and "Macmillan's Magazine".

1885 : Zola's novel "Germinal", describing the daily life, labor, and suffering of the miners, is a significant popular success.

1888 : Mahler begins composing his first two symphonies ("Titan" and "Resurrection") which he will complete in 1896 and 1894.

1889 : Publication of "Jenseits von Gut und Böse" (Beyond good and evil) in which Nietzsche proposes transcending "belief in the oppositions of values".

1889 : Van Gogh paints "Starry night".

Franco–German history

1880–1887 : William I, in a weakened state, tends more and more to abdicate power to Bismarck. In 1882, the latter signs the *Triple Alliance* with Austria-Hungary and the Kingdom of Italy.

1879–1885 : In France, Jules Ferry is the dominant member of the first republican governments. He establishes free compulsory and secular elementary education. He has laws passed ensuring the freedom of the press and freedom of association. He is also a very active supporter of French colonial expansion. In this he is opposed to Clemenceau, who regards colonialist adventures as distracting attention from the lost provinces of Alsace and Lorraine; for this reason Ferry earns the goodwill of Bismarck.

1888 : Following the death of William I and the very short reign of Frederick III, William II is proclaimed emperor (Kaiser Wilhelm II).

1890–1899

Uniformization

1890 : The Royal Scientific Society of Göttingen proposes a new approach to the uniformization theorem, via an investigation of the equation $\Delta u = k.e^u$. This suggestion soon leads Picard and Poincaré to a new proof of the uniformization theorem for compact Riemann surfaces.

Mathematics

1894 : In his dissertation, É. Cartan classifies *complex semi-simple Lie algebras*.

1895 : Poincaré publishes his first article on “Analysis Situs”, in which he founds *algebraic topology*.

1895 : Weber publishes his famous “Lehrbuch der Algebra”.

1895 : “Lessons from Stockholm”: Painlevé replaces Poincaré (on the advice of Mittag-Leffler) as the appropriate choice to expound before the Swedish king Oscar II the latest progress in the analytic theory of differential equations.

1896 : Hadamard and de la Vallée Poussin independently prove the *prime number theorem*, thereby giving an estimate of the number of primes less than a given number.

1896 : Frobenius makes a major contribution to group representations and the theory of finite fields in his publication “Über die Gruppencharactere”.

1897 : Burnside publishes “The Theory of Groups of Finite Order”.

1898 : Hadamard’s article on the geodesics of *negatively curved surfaces* lays the foundation of *symbolic dynamics*.

Science

1891 : Dubois discoves the first remains of *Homo erectus* at Trinil, east of the Indonesian island of Java. He names it *Pithecanthropus erectus* in an account published in 1894.

1895 : Röntgen discovers *X-rays*.

1896 : Becquerel discovers the *radioactivity of uranium*.

1897 : Thomson discovers the *electron*.

1898 : P. et M. Curie discover *radium* and *polonium* and show that radioactivity is an atomic property.

Technological advances

1890 : On board his aircraft *Éole*, Ader lifts off and shaves the ground for around fifty meters.

1895 : The brothers Lumière organize the first *cinema* projections.

1895 : Diesel perfects an *internal combustion engine with spontaneous compression-ignition* (rather than ignition by spark plugs).

1899 : Hoffmann patents *aspirin*.

The Arts, Philosophy, and the Humanities

1890 : Klimt's "Judith" is considered a *crime against artistic creation*.

1891 : Gauguin moves to Tahiti in the hope of escaping from western civilization. He paints sixty-six canvases in the course of a few months.

1893 : The definitive edition of Michelet's "Histoire de France".

1895 : In his "Rules of the sociological method" Durkheim describes the method he employs to guarantee the scientific rectitude of the discipline he founded.

1898 : Apollinaire begins writing his anthology "Alcools".

1899 : Conrad's novella "Heart of darkness" is serialized in the Scottish review "Blackwood's Magazine".

Franco–German history

1890 : Bismarck is forced to step down. *Realpolitik* gives way to *Weltpolitik*, whose goal is to find for Germany a "place in the sun" proportionate to her industrial power. Colonial politics is resumed and the construction of a navy to rival the Royal Navy is made a national priority.

1891 : Creation of the *pan-German League* which meets with rapid success and will last till 1939. Its aim is to unite all German speakers, Germanize all aliens living in the Reich, recover lost territories, and seize lands "necessary for the development of the German race".

1893 : The *Franco–Russian Alliance*, signed at the end of December, stipulates that those countries will provide mutual support if attacked by a member of the Triple Alliance.

1894 : Captain Dreyfus, accused of spying for Germany, is condemned to deportation to Guyana. Several individuals try in vain to prove his innocence. The Dreyfus affair really explodes in 1898, on the publication of Zola's article "J'accuse" (I accuse). France is split between dreyfusards, for whom truth must prevail whatever the consequences, and the antidreyfusards, for whom the prestige of the army — the instrument of revenge — demands that a judgement once passed should not be revised.

1900–1909

Uniformization

1900 : The uniformization of Riemann surfaces is one of the *twenty-three problems* put forward by Hilbert at the *International Congress of Mathematicians*.

1900 : Osgood shows that every bounded, simply connected, open subset of the plane is biholomorphic to the unit disc.

1907 : Poincaré and Koebe simultaneously prove the *uniformization theorem*.

Mathematics

1900 : Hilbert publishes “Grundlagen der Geometrie”, providing a complete *axiomatization* of Euclidean geometry.

1902 : In his dissertation, Lebesgue founds *integration theory*.

1906 : Markov introduces the *random processes* bearing his name, to become a basic tool in probability theory.

1907 : Brouwer’s dissertation on the foundations of mathematics marks the beginning of the *intuitionist school*.

Sciences

1900 : Freud publishes “Die Traumdeutung” (The interpretation of dreams). Ida Bauer begins a course of treatment with him: psychoanalysis is born!

1900 : Planck lays the foundations of *quantum mechanics* in “Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum”: his *quantum of action* provides an explanation of the experimental facts concerning black body radiation.

1904 : In “Elektromagnetische Vorgänge in einem Systeme…”, Lorentz formulates the *group of transformations* bearing his name, which will play a central role in the special theory of relativity.

1904–1905 : Weber publishes in the form of articles “Die protestantische Ethik und der Geist des Kapitalismus” (The protestant ethic and the spirit of capitalism). This work, centered on individual motives for action, will have a considerable influence on sociology.

1905 : In “Drei Abhandlungen zur Sexualtheorie”, Freud assembles his hypotheses on the place of sexuality and its future in the development of the personality.

1905 : Einstein’s four fundamental papers revolutionize physics and open a new era. They concern the *photoelectric effect, brownian motion, the special theory of relativity* and the *equivalence of mass and energy*.

1906 : De Saussure begins his investigations of the structural principles of *linguistics*.

Technological advances

1904 : Fleming, of the Marconi company, invents the *diode vacuum tube*.

1907 : World's first genuine takeoff of a *helicopter*, invented and piloted by Cornu.

1907 : *Bakelite* is one of the first plastic materials, developed by the chemist Baekeland.

1908 : Painlevé flies as a passenger with one of the Wright brothers in their biplane.

1910 : Claude invents *neon lighting*.

The Arts, Philosophy, and the Humanities

1901 : Chekhov's play "The three sisters" is performed at the Moscow Arts Theater.

1902 : On April 30, at the Paris Comic Opera, Messager directs "Pelléas et Mélisande", a lyrical drama in five acts by Debussy with libretto adapted from a play by Maeterlinck.

1906 : Appearance of Musil's novel "Die Verwirrungen des Zöglings Törless" (The confusions of young Törless). Musil's "Preliminary studies for a novel" prefigure "The man without qualities".

1907 : Proust begins writing "À la recherche du temps perdu", the various volumes of which will be published between 1913 and 1927.

1908–1911 : Schoenberg progressively invents the musical language "free atonality", liberated from tonal functions and hierarchies. His "String quartet No. 2" (1908), "Five orchestral pieces" (1909) and "Six little pieces for piano" (1911) mark important stages on the way to this new language.

1909 : Picasso paints "Les Demoiselles d'Avignon", considered to represent the point of departure of *cubism*.

Franco–German history

1902 : In France, legislative elections give power to an alliance of radicals and socialists: the *Left Bloc*. The government program of Émile Combes consists essentially in combatting the influence of the church. This culminates in the *law of separation of church and state* of 1905 and the nationalization of church property.

1904 : France and the United Kingdom sign a treaty, the *Entente Cordiale*, marking a diplomatic *rapprochement* between the two countries, faced with the Triple Alliance. The Entente Cordiale, the Franco–Russian Alliance, and the *Anglo–Russian Convention*, which will be signed in 1907, will together constitute the *Triple Entente*.

1905 : In Germany, von Schlieffen presents his plan for an offensive strategy on the French and Russian fronts.

Bibliography

[Abe1826] N. H. ABEL – "Mémoire sur une propriété générale d'une classe très étendue de fonctions transcendantes", presented to the Paris Academy of Sciences on Octobre 30, 1826. = [Abe1881], pp. 145–211.

[Abe1881] — , *Œuvres complètes de Niels Henrik Abel*, Second edition, Christiana [Oslo], 1881. Reprinted: Gabay, Sceaux, 1992.

[Abi1981] W. ABIKOFF – "The uniformization theorem", *Amer. Math. Monthly* **88** (1981), pp. 574–592.

[Adl1999] A. ADLER – "Hirzebruch's curves $F_1, F_2, F_4, F_{14}, F_{28}$ for $\mathbf{Q}(\sqrt{7})$", in [Levy1999b], pp. 221–285.

[Ahl1953] L. V. AHLFORS – "Development of the theory of conformal mapping and Riemann surfaces through a century", in *Contributions to the theory of Riemann Surfaces*, Ann. of Math. Studies 30, Princeton Univ. Press, 1953, pp. 3–13.

[Ahl1973] — , *Conformal invariants: topics in geometric function theory*, McGraw-Hill, New York, 1973.

[Ahl2006] — , *Lectures on quasiconformal mapping, second edition. With supplemental chapters by C. J. Earle, I. Kra, M. Shishikura and J. H. Hubbard*. Univ. Lecture Series 38, Amer. Math. Soc., Providence (RI), 2006.

[AhSa1960] L.V. AHLFORS & L. SARIO – *Riemann Surfaces*, Princeton Univ. Press, 1960.

[Ale1920] J. W. ALEXANDER – "A proof of Jordan's theorem about a simple closed curve", *Ann. of Math.* **21** (1920), No. 3, pp. 180–184.

[Ant1921] L. Antoine – “Sur l’homéomorphie de deux figures et de leurs voisinages. Thèses présentées à la Faculté des sciences de Strasbourg”, *J. Math. Pures Appl. (8^e série)* **4** (1921), pp. 221–325.

[App1925] P. Appell – *Henri Poincaré*, Plon, Paris, 1925.

[ApGo1929] P. Appell & É. Goursat – *Théorie des fonctions algébriques et de leurs intégrales. I Étude des fonctions analytiques sur une surface de Riemann*, Second edition, Gauthier-Villars, Paris, 1929. Volume II : [Fat1930].

[Arg1806] J.-R. Argand – *Essai sur une manière de représenter les quantités imaginaires par des constructions géométriques*, Veuve Blanc, Paris, 1806. Second edition: Gauthier-Villars, Paris, 1874.

[Aub1998] T. Aubin – *Some nonlinear problems in Riemannian geometry*, Monographs in Mathematics, Springer, New York, 1998.

[Bea1983] A. F. Beardon – *The geometry of discrete groups*, Grad. Texts in Math. 91, Springer, New York, 1983.

[Bea1984] — , *A primer on Riemann surfaces*, London Math. Soc. Lect. Note Series 78, Cambridge Univ. Press, Cambridge, 1984.

[BeSt1949] H. Behnke & K. Stein – “Entwicklung analytischer Funktionen auf Riemannschen Flächen”, *Math. Ann.* **120** (1949), pp. 430–461.

[Berg1971] M. S. Berger – “Riemannian structures of prescribed Gaussian curvature for compact 2-manifolds”, *J. Differential Geometry* **5** (1971), pp. 325–332.

[Ber1976] L. Bers – “On Hilbert’s 22nd problem”, in *Mathematical developments arising from Hilbert problems (Symposium of the Amer. Math. Soc., De Kalb, 1974)*, Proc. Sympos. Pure. Math., Vol. XXVIII, Amer. Math. Soc., Providence (RI), 1976.

[Bert1882] E. Bertini – “Sui sistemi lineari”, *Istit. Lombardo Accad. Sci. Lett. Rend. A Istituto (II)* **15** (1882), pp. 24–28.

[Bie1921] L. Bieberbach – “Neuere Untersuchungen über Funktionen von komplexen Variablen”, *Encyklopädie der mathematischen Wissenschaften*, Vol. 2, Part 3, 2nd half, Teubner, Berlin, 1921, pp. 379–532.

[Bol1909] O. Bolza – *Lectures on the Calculus of Variations*, Chelsea Publishing Company, New York, 1909.

[BoOt1988] F. Bonahon & J.-P. Otal – "Variétés hyperboliques à géodésiques arbitrairement courtes", *Bull. London Math. Soc.* **20** (1988), No. 3, pp. 255–261.

[Bos1992] J.-B. Bost – "Introduction to compact Riemann surfaces, Jacobians, and Abelian varieties", *From number theory to physics (Les Houches, 1989)*, Springer, Berlin, 1992, pp. 64–211.

[Bre1939] M. Brelot – "Familles de Perron et problème de Dirichlet", *Acta Sci. Math.* (Szeged) **IX** (1939), pp. 133–153.

[BrKn1981] E. Brieskorn & H. Knörrer – *Ebene Algebraische Kurven*, Birkhäuser, Basel, 1981. English translation: *Plane algebraic curves*, *ibid.*, 1986.

[BrNo1874] A. Brill & M. Nöther – "Über die algebraischen Functionen und ihre Anwendung in der Geometrie", *Math. Ann.* **7** (1874), pp. 269–316.

[BrNo1894] A. Brill & M. Noether – "Die Entwicklung der algebraischen Funktionen in älterer und neuerer Zeit", *Jahresber. Deutschen Math.-Vereinigung* **3** (1894), pp. 107–566.

[BrBo1859] C. Briot & J. C. Bouquet – *Théorie des fonctions doublement périodiques et, en particulier, des fonctions elliptiques*, Mallet-Bachelier, Paris, 1859.

[Bro1912] L. E. J. Brouwer – "Beweis der Invarianz des n-dimensionale Gebiets", *Math. Ann.* **71** (1912), pp. 305–313.

[Bué1806] A.-Q. Buée – "Mémoire sur les quantités imaginaires", *Philos. Trans. Roy. Soc. London* **96** (1806), pp. 23–88.

[Car1906] C. Caratheodory – "Über die starken Maxima und Minima bei einfachen Integralen", *Math. Ann.* **62** (1906), pp. 449–503.

[Car1912] — , "Untersuchungen über die konformen Abbildungen von festen und veränderlichen Gebieten", *Math. Ann.* **72** (1912), pp. 107–144.

[Car1913a] — , "Zur Ränderzuordnung bei konformer Abbildung", *Nachr. Ges. Wiss. Göttingen* (1913), pp. 509–518.

[Car1913b] — , "Über die gegenseitige Beziehung der Ränder bei der konformen Abbildung des Inneren einer Jordanschen Kurve auf einen Kreis", *Math. Ann.* **73** (1913), pp. 305–320.

[Car1913c] — , "Über die Begrenzung einfach zusammenhängender Gebiete", *Math. Ann.* **73** (1913), pp. 323–370.

[Car1914a] — (Ed.), *H. A. Schwarz-Festschrift. Mathematische Abhandlungen, Herrmann Amandus Schwarz zu seinem fünfzigjährigen Doktorjubiläum am 6. August 1914 gewidmet von Freunden und Schülern*, Springer, Berlin, 1914.

[Car1914b] — , "Elementarer Beweis für den Fundamentalsatz der konformen Abbildungen", in [Car1914a], pp. 19–40.

[Car1932] — , *Conformal Representation*, Cambridge Tracts in Mathematics and Mathematical Physics Nr. 28, Cambridge Univ. Press, Cambridge, 1932. Second edition 1952, reprinted by Dover, Mineola, 1998.

[Cat2004] F. Catanese – "From Abel's Heritage: Transcendental Objects in Algebraic Geometry and their Algebraization", in [LaPi2004], pp. 349–394.

[Cha1837] M. Chasles – *Aperçu historique sur l'origine et le développement des méthodes en géométrie : particulièrement de celles qui se rapportent à la géométrie moderne*, followed by *Mémoire de géométrie sur deux principes généraux de la science, la dualité et l'homographie*, Hayez, Bruxelles, 1837.

[CLT2006] X. Chen, P. Lu & G. Tian – "A note on uniformization of Riemann surfaces by Ricci flow" *Proc. Amer. Math. Soc.* **134** (2006), No. 11, pp. 3391–3393.

[Che1951] C. Chevalley – *Introduction to the Theory of Algebraic Functions of One Variable*, Mathematical Surveys VI, Amer. Math. Soc., New York, 1951.

[Cho2007] R. Chorlay – *L'émergence du couple local-global dans les théories géométriques, de Bernhard Riemann à la théorie des faisceaux*, Thesis on the history of science, Univ. Paris 7, 2007. Available at http://www.rehseis.cnrs.fr/spip/php?article35#these.

[Chr1867] E. B. CHRISTOFFEL – "Sul problema delle temperature stazionarie e la rappresentazione di una data superficie (trans. D. Formenti), *Ann. Mat. Pura Appl.* **1** (1867), pp. 89–103. *Gesammelte Abhandlungen*, Teubner, Leipzig, 1910.

[Cle1865a] A. CLEBSCH – "Über diejenigen ebenen Curven, deren Coordinaten rationale Functionen eines Parameters sind", *J. Reine Angew. Math.* **64** (1865), pp. 43–65.

[Cle1865b] — , "Ueber diejenigen curven, deren coordinaten sich als elliptische Functionen eines Parameters darstellen lassen", *J. Reine Angew. Math.* **64** (1865), pp. 210–270.

[Clif1877] W. K. CLIFFORD – "On the canonical form and dissection of a Riemann's surface", *Proc. of the London Math. Society* **VIII** No. 122, (1877), pp. 292–304. = [Clif1882, pp. 241–255].

[Clif1882] — , *Mathematical papers*, Macmillan, London, 1882. Reprinted by Chelsea, New York, 1968.

[Coh1967] H. COHN – *Conformal mapping on Riemann surfaces*, McGraw-Hill, New York, 1967. Reprinted by Dover, New York, 1980.

[Cra1750] G. CRAMER – *Introduction à l'analyse des lignes courbes algébriques*, Frères Cramer et Cl. Philibert, Genève, 1750.

[CuSh1983] M. CULLER & P. SHALEN – "Varieties of group representations and splittings of 3-manifolds", *Ann. of Math.* **117** (1983), pp. 109–146.

[Dal2007] F. DAL'BO – *Trajectoires géodésiques et horocycliques*, EDP Sciences & CNRS éditions, Paris, 2007.

[Ded1877] R. DEDEKIND – "Über die Theorie der elliptischen Modul-Functionen ", *J. Reine Angew. Math.* **83** (1877), pp. 265–292.

[DeHe1907] M. DEHN & P. HEEGAARD – "Analysis Situs", *Encyklopädie der mathematischen Wissenschaften [Vol. 3, Part 1, 1st half]*, 1907, pp. 153–250.

[Desa1639] G. DESARGUES – *Brouillon-project d'une atteinte aux evenemens des rencontres du cone avec un plan*, Paris, 1639. *Œuvres de Desargues*, Leiber, Paris, 1861, pp. 103–181.

[Desc1637] R. Descartes – *La Géométrie*, appendix to *Discours de la méthode pour bien conduire sa raison ...* , Jan Maire, Leyde, 1637. Reprinted: Gabay, Sceaux, 1991.

[Die1974] J. Dieudonné – *Cours de géométrie algébrique. 1 Aperçu historique sur le développement de la géométrie algébrique*, Presses Univ. France, Paris, 1974.

[Die1978] — (Ed.), *Abrégé d'histoire des mathématiques 1700–1900*, Hermann, Paris, 1978.

[Die1982] — , "La découverte des fonctions fuchsiennes", *Actes du sixième congrès des mathématiciens d'expression latine*, Gauthier-Villars, Paris, 1982, pp. 3–23.

[Die1989] — , *A history of algebraic and differential topology, 1900–1960*, Birkhäuser, Boston, 1989.

[Don2011] S. Donaldson – *Riemann surfaces.* Oxford University Press, 2011.

[DoTi1978] M. Dostal & R. Tindell – "The Jordan curve theorem revisited", *Jahresber. Deutschen Math.-Vereinigung* **80** (1978), pp. 11–128.

[DoDo1977] R. & A. Douady – *Algèbre et théories galoisiennes*, CEDIC, Paris, 1977. New enlarged edition, Cassini, Paris, 2005.

[DrTr2002] T. A. Driscoll & L. N. Trefethen – *Schwarz–Christoffel mapping*, Cambridge Univ. Press, 2002.

[Dyc1888] W. Dyck – "Beyträge zur Analysis Situs", *Math. Ann.* **32** (1888), pp. 457–512.

[Elk1998] N. Elkies – "Shimura curve computations", *Proceedings of the 3rd Algorithmic Number Theory Symposium*, Lect. Notes in Comput. Sci. **1423** (1998), pp. 1–47.

[Elk1999] — , "The Klein quartic in number theory", in [Levy1999b], pp. 51–101.

[Eul1777] L. Euler – "De repraesentatione superficiei sphaericae super plano / De projectione geographica superficiei sphaericae / De projectione geographica De Lisliana in mappa generali imperii russi usitata", *Acta Acad. Sci. Imper. Petropol.* 1777. German translation: “Drei Abhandlungen über Kartenprojection”, Engelmann, Leipzig, 1898.

[FaKr1980] H. M. Farkas, I. Kra – *Riemann surfaces*, Graduate Texts in Mathematics 71, Springer, New York, 1980.

[Fat1930] P. Fatou – *Théorie des fonctions algébriques et de leurs intégrales. II Fonctions automorphes*, Gauthier-Villars, Paris, 1930. Volume I : [ApGo1929].

[Fer1636] P. de Fermat – “Isagoge ad locos planos et solidos”, 1636. *Œuvres de Fermat*, Vol. 1, Gauthier-Villars, Paris, 1891, pp. 91–102.

[Ford1929] L. A. Ford – *Automorphic functions*, McGraw-Hill, New York, 1929.

[Forst1977] O. Forster – *Riemannsche Flächen*, Springer, Berlin, 1977. English translation: *Lectures on Riemann surfaces*, Grad. Texts in Math. 81, Springer, New York, 1981.

[Forsy1902] A. R. Forsyth – *Theory of differential equations, Part III, ordinary linear equations, Vol. IV*, Cambridge Univ. Press, 1902.

[Freu1955] H. Freudenthal – “Poincaré et les fonctions automorphes”, in *Le livre du centenaire de la naissance de Henri Poincaré*, Gauthier-Villars, Paris 1955, pp. 212–219. Also in [Poin1951], Vol. XI.

[Fric1901] R. Fricke – “Automorphe Funktionen und Modulfunktionen”, *Encyklopädie der mathematischen Wissenschaften [Vol. 2, Part 2, 2nd half]*, Teubner, Leipzig, 1901, pp. 349–470.

[FrKl1897] R. Fricke & F. Klein – *Vorlesungen über die Theorie der automorphen Funktionen*, 2 volumes (the second in 3 parts), Teubner, Leipzig, 1897 to 1912.

[Fub1906] G. Fubini – “Sul principio di Dirichlet (Da una lettera al Prof. Beppo Levi)”, *Rend. Circ. Mat. Palermo* **22** (1906), pp. 383–386.

[Fub1907a] — , “Il problema di Dirichlet considerato come limite di un ordinario problema di minimo”, *Atti R. Accad. Lincei, Rend.* Ser. 5, **16** (1907), 162–167.

[Fub1907b] — , “Il principio di minimo”, *Rend. Circ. Mat. Palermo* **23** (1907), pp. 300–301.

[Fub1907c] — , “Il principio di minimo e i teoremi di esistenza per i problemi al contorno relativi alle equazioni alle derivate parziali di ordine pari”, *Rend. Circ. Mat. Palermo* **23** (1907), 53–84.

[Fub1907d] — , “Nuove applicazioni del principio di minimo”, *Annali di Mat.* **14** (1907), pp. 113–141.

[Fuc1866] L. Fuchs – “Zur Theorie der linearen Differentialgleichungen mit veranderlichen Coeffizienten”, *J. Reine Angew. Math* **66** (1866), pp. 121–160.

[Fuc1880] — , “Über die Functionen, welche durch Umkehrung der Integrale von Lösungen der linearen Differentialgleichungen entstehen”, *Nachr. Ges. Wiss. Göttingen* (1880), pp. 445–453.

[Fuc1881] — , “Auszug aus einem Schreiben des Herrn L. Fuchs an C.W. Borchardt”, *J. Reine Angew. Math.* **90** (1881), pp. 71–73.

[Gard1979] L. Gårding – “The Dirichlet problem ”, *Math. Intelligencer* **2** (1979/80), pp. 43–53.

[Gars1961] A. M. Garsia – “An imbedding of closed Riemann surfaces in Euclidean space”, *Comment. Math. Hel.* **35** (1961), pp. 93–110.

[Gars1962] — , “On the conformal types of algebraic surfaces of euclidean space ”, *Comment. Math. Hel.* **36** (1962), pp. 49–60.

[Gau1825] C. F. Gauss – “Allgemeine Auflösung der Aufgabe die Theile einer gegebenen Fläche auf einer anderen gegebenen Fläche so auszubilden . . . ”, in *Astronomische Abhandlungen [Vol. 3]*, edited by H.C. Schumacher, Altona, 1825. = [Gau1863], vol. 4, pp. 189–216. Translated into French as: “Solution générale de ce problème : représenter les parties d’une surface donnée sur une autre surface donnée”, Hermann, Paris, 1915.

[Gau1831] — , “Theoria residuorum biquadraticorum, commentatio secunda”. *Göttingische gelehrte anzeigen* **8** (1831). = [Gau1863], Vol. 4, pp. 169–178.

[Gau1845] — , “Untersuchungen über der höhere Geodäsie. Erste Abhandlung ”, *Abhand. Kön. Ges. Wiss. Göttingen. Math. Kl.* **2** (1845), pp. 3–45. = [Gau1863], Vol. 4, pp. 259–300.

[Gau1863] — , *Werke*, 9 volumes, Universität-Drückerei, Göttingen, 1863–1906.

[Glu2008] A. Glutsyuk – "Simple proofs of uniformization theorem ", in *Holomorphic Dynamics and Renormalization*, Fields Instit. Comm., **53**, Amer. Math. Soc., Providence, 2008, pp. 125–143.

[GolM1971] M. E. Goldman – "An algebraic classification of noncompact 2-surfaces", *Trans. Amer. Math. Soc.* **156**, 1971, pp. 241–258.

[GolW1987] W. M. Goldman – "Projective structures with Fuchsian holonomy", *J. Diff. Geom.* **25** (1987), pp. 297–326.

[GolW1988] — , "Topological components of spaces of representations", *Invent. Math.* **93** (1988), pp. 557–607.

[Got1890] *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen* **36** (1890).

[Gra1986] J. Gray – *Linear differential equations and group theory from Riemann to Poincaré*, Birkhäuser, Boston, 1986. Second edition, 2000.

[Gra1994] — , "On the history of the Riemann mapping theorem,", *Rend. Mat. Circ. Palermo ser. II* **34** (1994), pp. 47–94.

[Gra1998] — , "The Riemann–Roch theorem and geometry, 1854–1914", *Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998)*. Doc. Math. 1998, Extra Vol. III, 1998, Documenta Mathematica, Bielefeld, pp. 811–822.

[Gre1828] G. Green – *An essay on the application of mathematical analysis to the theories of electricity and magnetism*, Nottingham, 1828. *Mathematical papers of the late George Green*, Macmillan, London, 1871, pp. 1–82.

[Gri1971] P. Griffiths – "Complex-analytic properties of certain Zariski open sets on algebraic varieties", *Ann. of Math.* **94** (1971), pp. 21–51.

[GrHa1978] P. Griffiths & J. Harris – *Principles of algebraic geometry*, Wiley, New York, 1978.

[Gun1967] R. C. Gunning – "Special coordinate coverings of Riemann surfaces", *Math. Ann.* **170** (1967), pP. 67–86.

[Gun1972] — , *Lectures on Riemann surfaces, Jacobi varieties*, Math. Notes 12, Princeton Univ. Press, 1972.

[Had1906] J. Hadamard – "Sur le principe de Dirichlet", *Bull. Soc. Math. France* **34** (1906), pp. 135–138.

[Had1999] — , *Non-Euclidean geometry in the theory of automorphic functions*, Amer. Math. Soc., Providence (RI), 1999. Translation of a 1951 Russian version of the original French version lost in the 1920s, with an historical introduction by J. Gray.

[HaPo1993] J. Hadamard & H. Poincaré – *Essai sur la psychologie de l'invention dans le domaine mathématique [Hadamard]. L'invention mathématique [Poincaré]*. Reprinting of the 1975 translation of Hadamard's essay and of Poincaré's 1908 text, Gabay, Sceaux, 1993.

[Hal1695] E. Halley – "An Easie Demonstration of the Analogy of the Logarithmick Tangents to the Meridian Line or sum of the Secants : with various Methods for computing the same to the utmost Exactness", *Philos. Trans.* **19** (1695–1697), pp. 202–214.

[Han1979] W. Hansen – "The Dirichlet problem" *Equadiff IV. Proc. Czechoslovak Conf. Differential Equations and their Applications, Prague, (1977)*, Lecture Notes in Math. 703, Springer, Berlin, 1979.

[HaMo1998] J. Harris & I. Morrison – *Moduli of curves*, Grad. Texts in Math. 187, Springer, New York, 1998.

[Hat2002] A. Hatcher – *Algebraic topology*, Cambridge Univ. Press, 2002.

[Haw2000] T. Hawkins – *Emergence of the theory of Lie groups, An essay in the History of Mathematics, 1869–1926*, Springer, 2000.

[Hei1949] M. Heins – "The conformal mapping of simply-connected Riemann surfaces", *Ann. of Math.* **50** (1949). pp. 686–690.

[HeLa1902] K. Hensel & G. Landsberg – *Theorie der algebraischen Funktionen einer Variabeln und ihre Anwendung auf algebraische Kurven und Abelsche Integrale*, Teubner, Leipzig, 1902. Reprinted by Chelsea, New York, 1965.

[Her1858] C. Hermite – "Sur la résolution de l'équation du cinquième degré", *C. R. Acad. Sci. Paris* **46** (1858), pp. 508–515.

[Hil1899] D. Hilbert – "Über das Dirichlet'schen Prinzip", *Jahresber. Deutschen Math.-Vereinigung* **8** (1899), pp. 184–188.

[Hil1900a] — , “Über das Dirichletsche Prinzip”, Deutsche Math. Ver. **81** (1900), pp. 184–188. Translated into French by L. Laugel: “Sur le principe de Dirichlet ”, *Nouv. Ann. (3rd series)* **19** (1900), pp. 337–544. Reprised in [Hil1905].

[Hil1900b] — , “Sur les problèmes futurs des mathématiques”, Gauthier-Villars, Paris, 1902. Reissued by Gabay, Sceaux, 1990.

[Hil1904] — , “Über das Dirichletsche Prinzip”, *Math. Ann.* **59** (1904), pp. 161–186.

[Hil1905] — , “Über das Dirichletsche Prinzip”, J. Reine Angew. Math. **129** (1905), pp. 63–67. Reprise of [Hil1900a], with commentaries on intermediate works.

[Hil1909] — , “Zur Theorie der konformen Abbildung”, *Nachr. Ges. Wiss. Göttingen* (1909), pp. 314–323.

[HiCo1932] D. Hilbert & S. Cohn-Vossen – *Anschauliche Geometrie*, Springer, Berlin, 1932. English translation: *Geometry and the imagination*, Chelsea, New York, 1952. Reprinted by Amer. Math. Soc., Providence (RI), 1998.

[Hil1976] E. Hille – *Ordinary differential equations in the complex domain*, Wiley, New York, 1976. Reprinted by Dover, Mineola (NY), 1997.

[Hit1987] N. Hitchin – “The self-duality equations on a Riemann surface”, *Proc. London Math. Soc.(3)* **55** (1987), pp. 59–126.

[Hod1941] W. V. D. Hodge – *The Theory and Applications of Harmonic Integrals*, Cambridge Univ. Press, 1941.

[Hor1972] R. D. Horowitz – “Characters of free groups represented in the two-dimensional special linear group”, *Comm. Pure Appl. Math.* **25** (1972), pp. 635–649.

[Hou1978] C. Houzel – *Fonctions elliptiques et intégrales abéliennes*, in [Die1978], pp. 1–113.

[Hou2002] — , *La géométrie algébrique – recherches historiques*, Blanchard, Paris, 2002.

[Hub1981] J. Hubbard – "The monodromy of projective structures", *Riemann surfaces and related topics. Proceedings of the 1978 Stony Brook conference*, Ann. of Math. Studies 97, Princeton Univ. Press, 1981, pp. 257–275.

[Hub2006] — , *Teichmüller theory, vol. 1*, Matrix Editions, Ithaca NY, 2006.

[HuTr1992] D. Hulin & M. Troyanov – "Prescribing curvature on open surfaces", *Math. Ann.* **293** (1992), pp 277–315.

[Iha1974] Y. Ihara – "Schwarzian equations", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **21** (1974), pp. 97–118.

[ImTa1992] Y. Imayoshi & M. Taniguchi – *An introduction to Teichmüller spaces*, Springer, Tokyo, 1992.

[Inc1944] E. L. Ince – *Ordinary Differential Equations*, Dover, New York, 1944.

[IKSY1991] K. Iwasaki, H. Kimura, S. Shimomura & M. Yoshida – *From Gauss to Painlevé. A modern theory of special functions*, Aspects of Mathematics, Vieweg, Braunschweig 1991.

[Jac1828] C. G. Jacobi – "Suite des notices sur les fonctions elliptiques", *J. Reine Angew. Math.* **3** (1828), pp. 303–310; in [Jac1969], Vol. I, pp. 255–263.

[Jac1969] — , *Gesammelte Werke*, Chelsea, New York, 1969.

[Joh1906a] S. Johansson – "Ein Satz über die konforme Abbildung einfach zusammenhängender Riemannscher Flächen auf den Einheitskreis.", *Math. Ann.* **62** (1906), pp. 177–183.

[Joh1906b] — , "Beweis der Existenz linear-polymorpher Funktionen vom Grenzkreistypus auf Riemannschen Flächen", *Math. Ann.* **62** (1906), pp. 184–193.

[Jon1971] B. F. Jones – *Rudiments of Riemann surfaces*, Lecture Notes in Mathematics, **2**, Rice University, Houston ,Texas, 1971.

[Jor1866] C. Jordan – "Sur la déformation des surfaces ", *J. Math. Pures Appl. (2nd series)* **11** (1866), pp. 105–109.

[Jor1872] — , "Sur les lignes de faîte et de thalweg", *C. R. Acad. Sci. Paris* **74** (1872), pp. 1457–1459, and **75** (1872), pp. 625–627 and 1023–1025.

[Jor1887] — , *Cours d'analyse de l'École polytechnique. Tome III, Calcul intégral*, Gauthier-Villars, Paris, 1887.

[Jos2002] J. Jost – *Compact Riemann surfaces, An introduction to Contemporary mathematics, Second edition*, Springer, New York, 2002.

[Kap2001] M. Kapovich – *Hyperbolic manifolds and discrete groups*, Progr. in Math. 183, Birkhäuser, Boston, 2001.

[Kas1913] E. Kasner – "Conformal geometry", in *Proceedings of the fifth international congress of mathematicians (Cambridge 1912), Vol. 2.*, Cambridge Univ. Press, Cambridge, 1913, pp. 81–87.

[Kat1992] S. Katok – *Fuchsian Groups*, The Univ. of Chicago Press, 1992.

[KaWa1974] J. L. Kazdan & F. W. Warner – "Curvature functions for compact 2-manifolds", *Ann. of Math.* **99** (1974), pp. 14–47.

[Ker1923] B. Kerékjártó – *Vorlesungen über Topologie*, Springer, Berlin, 1923.

[Kleim2004] S. L. Kleiman – "What is Abel's theorem anyway?", in [LaPi2004], pp. 395–440.

[Kle1875] F. Klein – "Über binäre Formen mit linearen Transformationen in sich selbst", *Math. Ann.* **9** (1875), pp. 183–208. = [Kle1921a], Vol. II, pp. 275–301.

[Kle1877a] — , "Über lineare Differentialgleichungen", *Math. Ann.* **12** (1877), pp. 167–179. = [Kle1921a], Vol. II, pp. 307–320.

[Kle1877b] — , "Weitere Untersuchungen über das Ikosaeder", *Math. Ann.* **12** (1877), pp. 503–560. = [Kle1921a], Vol. II, pp. 321–384.

[Kle1878a] — , "Über die Erniedrigung der Modulargleichungen", *Math. Ann.* **14** (1878), pp. 417–427. = [Kle1921a], Vol. III, pp. 76–89.

[Kle1878b] — , "Über die Transformation der elliptischen Funktionen und die Auflösung der Gleichungen fünften Grades", *Math. Ann.* **14** (1878), pp. 111–172. = [Kle1921a], Vol. III, pp. 13–75.

[Kle1878c] — , "Über die Transformation siebenter Ordnung der elliptischen Functionen", *Math. Ann.* **14** (1878), pp. 428–471. = [Kle1921a], Vol. III, pp. 90–135. English translation: "On the Order-Seven Transformations of Elliptic Functions", in [Levy1999b], pp. 287–331.

[Kle1879a] — , "Über Multiplicatorgleichungen", *Math. Ann.* **15** (1879), pp. 86–88. = [Kle1921a], Vol. III, pp. 137–139.

[Kle1879b] — , "Über die Transformation elfter Ordnung der elliptischen Functionen", *Math. Ann.* **15** (1879), pp. 533–555. = [Kle1921a], Vol. III, pp. 140–165.

[Kle1880a] — , "Zur Theorie der elliptischen Modulfunctionen", *Math. Ann.* **17** (1880), pp. 62–70. = [Kle1921a], Vol. III, pp. 169–178.

[Kle1880b] — , "Über gewisse Teilwerte der Θ-Function", *Math. Ann.* **17** (1880), pp. 565–574. = [Kle1921a], Vol. III, pp. 186–197.

[Kle1881] — , "Über die conforme Abbildung von Flächen", *Math. Ann.* **19** (1881), pp. 159–160.

[Kle1882a] — , "Über eindeutige Functionen mit linearen Transformationen in sich", *Math. Ann.* **19** (1882), pp. 565–569. = [Kle1921a], Vol. III, pp. 622–626.

[Kle1882b] — , "Über eindeutige Functionen mit linearen Transformationen in sich", *Math. Ann.* **20** (1882), pp. 49–51. = [Kle1921a], Vol. III, pp. 627–629.

[Kle1882c] — , *Über Riemanns Theorie der algebraischen Funktionen und ihrer Integrale*, Teubner, Leipzig, 1882. English translation (revised): *On Riemann's theory of algebraic functions and their integrals*, Dover, Mineola (NY), 2003.

[Kle1882d] — , "Neue Beiträge zur Riemann'schen Functionentheorie", *Math. Ann* **21**, (1882) pp. 141–218. = [Kle1921a], Vol. III, pp. 630–710.

[Kle1884] — , *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen vom fünften Grade*, Teubner, Leipzig, 1884. English translation (revised): *Lectures on the icosahedron and the solution of equations of the fifth degree*, Dover, Mineola (NY), 2003.

[Kle1921a] — , *Gesammelte mathematische Abhandlungen*, 3 volumes, Springer, Berlin, 1921–1923.

[Kle1921b] — , *Zur Vorgeschichte der automorphen Funktionen*, [Kle1921a], Vol. III, pp. 577–586.

[Kle1925] — , *Elementarmathematik vom höheren Standpunkte aus: II*, Verlag von Julius Springer, Berlin, 1925. English translation: *Elementary mathematics from an advanced standpoint*. Geometry, Dover, Mineola, 2004.

[Kle1928] — , *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Springer, Berlin, 1928. Reprinted by Chelsea, New York, 1979. English translation: *Lectures on the development of mathematics in the 19th century*, Math. Sci Press, Brookline (Mass.), 1979.

[KlePoi1923] F. Klein, H. Poincaré – "La correspondance d'Henri Poincaré et de Félix Klein", *Acta Math.* **39** (1923), pp. 94–132. = [Poin1951], Vol. XI, pp. 26–65.

[Koe1907a] P. Koebe – "Über die Uniformisierung beliebiger analytischer Kurven", *Nachr. Ges. Wiss. Göttingen* (1907), pp. 191–210.

[Koe1907b] — , "Über die Uniformisierung beliebiger analytischer Kurven (zweite Mitteilung)", *Nachr. Ges. Wiss. Göttingen* (1907), pp. 633–649.

[Koe1908a] — , "Über die Uniformisierung beliebiger analytischer Kurven (dritte Mitteilung)", *Nachr. Ges. Wiss. Göttingen* (1908), pp. 337–358.

[Koe1908b] — , "Konforme Abbildunge der Oberfläche einer von endlich vielen regulären analytischen Flächenstücken gebildeten körperlichen Ecke ", *Nachr. Ges. Wiss. Göttingen* (1908), pp. 359–360.

[Koe1909a] — , "Sur un principe général d'uniformisation", *C.R. Acad. Sci. Paris* **148** (1909), p. 824–828.

[Koe1909b] — , "Fonction potentielle et fonction analytique ayant un domaine d'existence donné à un nombre quelconque (fini ou infini) de feuillets", *C. R. Acad. Sci. Paris* **148** (1909), pp. 1446–1448.

[Koe1909c] — , "Über die Uniformisierung beliebiger analytischer Kurven (vierte Mitteilung)", *Nachr. Ges. Wiss. Göttingen* (1909), pp. 324–361.

[Koe1909d] — , "Über ein allegemeines Uniformisierungprinzip", *Atti del IV congresso internazionale dei matematici (Roma 1908)* (1909), pp. 25–30.

[Koe1909e] — , "Über die Uniformisierung der algebraischen Kurven I", *Math. Ann.* **67** (1909), pp. 145–224.

[Koe1910a] — , "Über die Uniformisierung der algebraischen Kurven II", *Math. Ann.* **69** (1910), pp. 1–81.

[Koe1910b] — , "Über die Uniformisierung beliebiger analytischer Kurven. Erster Teil : Das allgemeine Uniformisierungsprinzip", *J. Reine Angew. Math.* **136** (1910), pp. 192–253.

[Koe1910c] — , "Über die Hilbertische Uniformisierungmethode", *Nachr. Ges. Wiss. Göttingen* (1910), pp. 59–74.

[Koe1911] — , "Über die Uniformisierung beliebiger analytischer Kurven. Zweiter Teil : Die zentralen Uniformisierungsprobleme", *J. Reine Angew. Math.* **139** (1911), pp. 251–292.

[Koe1912] — , "Über die Uniformisierung der algebraischen Kurven III. (Erster Beweis der allgemeinen Kleinschen Fundamentaltheoreme)", *Math. Ann.* **72** (1912), pp. 437–516.

[Koe1913a] — , "Ränderzuordnung bei konformer Abbildung ", *Nachr. Ges. Wiss. Göttingen* (1913), pp. 286–288.

[Koe1913b] — , "Das Uniformisierungstheorem und seine Bedeutung für Funktionentheorie und nichteuklidische Geometrie) ", *Annali di Mat.* **21** (1913), No. 1, pp. 57–68.

[Koe1914] — , "Über die Uniformisierung der algebraischen Kurven IV. Zweiter Existenezbeweis der allgemeinen kanonischen uniformisierenden Variablen: Kontinuitätsmethode", *Math. Ann.* **75** (1914), pp. 42–129.

[Koe1915] — , "Abhandlungen zur Theorie der konformen Abbildung: I. Die Kreisabbildung des allgemeinsten einfach und zweifach zusammenhängenden schlichten Bereichs und die Ränderzuordnung bei konformer Abbildung", *J. Reine Angew. Math.* **145** (1915), pp. 177–225.

[Kor1914] A. Korn – "Zwei Anwendungen der Methode der sukzessive Annäherungen", in [Car1914a], pp. 215–229.

[Kram1996] D. Krammer – "An example of an arithmetic Fuchsian group", *J. Reine Angew. Math.* **473** (1996), pp. 69–85.

[Kran2006] S. G. KRANTZ – *Geometric function theory. Explorations in complex analysis*, Birkhäuser, Boston, 2006.

[Ku1959] A. KUTSCHER – *The Method of Super- and Sub-Solutions.* http://www.math.utsc.utoronto.ca/gpde/notes/subsup.pdf

[Lag1779] J.-L. LAGRANGE – "Sur la constructions des cartes géographiques", *Nouveau Mémoire de l'Académie de Berlin* (1779), pp. 161–210. *Oeuvres*, Gauthier-Villars, Paris, 1867–1892, Vol. 4, pp. 637–692.

[Lam1772] J. H. LAMBERT – "Anmerkungen und Zusätze zur Entwerfung der Land- und Himmelscharten", *Beyträge zum gebrauche der Mathematik und deren Anwendung [Vol. 3]*, Buchandlung der Realsschule, Berlin, 1772, 6th article. Reissued by A. Wangerin, W. Engelmann, Leipzig, 1894.

[Lan2002] S. LANG – *Algebra. Revised third edition*, Grad. Texts in Math. 211, Springer, New York, 2002.

[LaPi2004] O. A. LAUDAL & R. PIENE – *The legacy of Niels Henrik Abel. Papers from the bicentennial conference at the University of Oslo, June 3–8, 2002*, Springer, New York, 2004.

[Leb1907a] H. LEBESGUE – "Sur le problème de Dirichlet", *C. R. Acad. Sci. Paris* **144** (1907), pp. 316–318 and 622–623.

[Leb1907b] H. LEBESGUE – "Sur le problème de Dirichlet", *Rend. mat. Palermo* **24** (1907), pp. 371–402.

[LeJa1875] A.-M. LEGENDRE & C. G. JACOBI – "Correspondance", *J. Reine Angew. Math.* **80** (1875), pp. 205–279. = [Jac1969], Vol. 1, pp. 390–461.

[Levi1906a] B. LEVI – "Sul principio di Dirichlet ", *Rend. mat. Palermo* **22** (1906), pp. 293–360.

[Levi1906b] — , "Sul principio di Dirichlet (Da una lettera al *Prof. Guido Fubini*)", *Rend. mat. Palermo* **22** (1906), pp. 387–394.

[Levy1999a] S. LEVY – "On the Order-Seven Transformations of Elliptic Functions", translation of [Kle1878c], in [Levy1999b].

[Levy1999b] — (Ed.), *The eightfold way. The beauty of Klein's quartic curve*, Math. Sci. Res. Inst. Publ. 35, Cambridge Univ. Press, 1999. Chapters available at: http://www.msri.org/communications/books/Book35/index.html.

[Lic1916] L. Lichtenstein – "Zur Theorie der konformen Abbildungen. Konforme Abbildung nichtanalytischer, singularitätenfreier Flächenstücke auf eben Gebiete", *Bull. Int. Acad. Sci. Cracovie A* (1916), pp. 192–217.

[Lic1921] — , "Neuere Entwicklung der Potentialtheorie. Konforme Abbildungen", *Encyklopädie der mathematischen Wissenschaften (Vol. 2, Part 3, 1st half)*, Teubner, Berlin (1921), pp. 1–377.

[Lin1920] E. Lindelöf – "Sur la représentation conforme d'une aire simplement connexe sur l'aire d'un cercle", *Fourth congress of Scandinavian mathematicians (Stockholm 1916)*, Almkvist & Wiksells, Uppsala, 1920, pp. 59–90.

[Lis1847] J. B. Listing – "Vorstudien zur Topologie", *Göttinger Studien* (1847), pp. 811–875.

[Lis1861] — , "Der Census räumlicher Complexe oder Verallgemeinerungen des Eulerschen Satzes von den Polyedern", *Abhandlungen der Math. Classe der König. Ges. der Wiss. Gött.* **10** (1861), pp. 97–182.

[Loe1988] F. Loeser – "Déformations de courbes planes (d'après Severi et Harris)", *Séminaire Bourbaki, Vol. 1986/87*, Astérisque **152–153** (1988), pp. 187–205.

[MaRe2003] C. Maclachlan & A. W. Reid *The arithmetic of hyperbolic 3-manifolds*, Grad. Texts in Math. 219, Springer, New York, 2003.

[Mar1996] A. I. Markushevich – "Analytic function theory", *Mathematics of the 19th century. Geometry, analytic function theory*, Eds. A. N. Kolmogorov and A. P. Yushkevich, Birkhäuser, Basel, 1996. Translated from Russian in 1981.

[Mas1971] B. Maskit – "On Poincaré's theorem for fundamental polygons", *Adv. in Math.* **7** (1971), pp. 219–230.

[Mas1988] — , *Kleinian groups*, Grund. Math. Wiss. 287, Springer, New York, 1988.

[Max1870] J. C. Maxwell – "On Hills and Dales", *Philos. Mag. Series* **40** (1870), pp. 421–427. In *Scientific papers (Vol. 2)*, Cambridge Univ. Press, 1890, pp. 233–240.

[McKMo1997] H. McKean & V. Moll – *Elliptic curves. Function theory, geometry, arithmetic*, Cambridge Univ. Press, 1997.

[Mil1965] J. Milnor – *Topology from the differentiable viewpoint*, The Univ. Press of Virginia, Charlottesville, 1965.

[MiPo1999] *Correspondance entre Henri Poincaré et Gösta Mittag-Leffler*, edited by P. Nabonnand. Publications of the Henri-Poincaré Archive, Birkhäuser, Basel, 1999.

[Möb1861] A. F. Möbius – "Zur Theorie der Polyëder und der Elementarverwandtschaft", manuscript. Published in [Möb1886], pp. 519–559.

[Möb1863] — , "Theorie der elementaren Verwandschaft", *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math.-Phys. Kl.* **15** (1863), pp. 18–57. = [Möb1886], pp. 473–512.

[Möb1865] — , "Über die Bestimmung des Inhaltes eines Polyeders", *Ber. Verh. Kön. Sächs. Ges. Wiss., Math.-phys. Klasse* **17**, pp. 31–68. In [Möb1886], pp. 433–471.

[Möb1886] — , *Gesammelte Werke (Vol. 2)*, S. Hirzel, Leipzig, 1886.

[Moc1999] S. Mochizuki – *Foundations of p-adic Teichmüller theory*, AMS/IP Studies in Advanced Mathematics 11, Amer. Math. Soc., Providence (RI), 1999.

[Mon1907] P. Montel – "Sur les suites infinies de fonctions", *Ann. Sci. École Norm. Sup.* **24** (1907), pp. 233–334.

[Mum1975] D. Mumford – *Curves and their Jacobians*, The Univ. of Michigan Press, Ann Arbor, 1975.

[Mum1983] — , *Tata lectures on theta I*, Progr. in Mathematics 28, Birkhäuser, Basel, 1983.

[Mum1999] — , *The red book of varieties and schemes*, second edition. Contains [Mum1975]. Lecture Notes in Mathematics, 1358. Springer, Berlin, 1999.

[Neh1952] Z. NEHARI – *Conformal mapping*, McGraw-Hill, New York, 1952.

[Neue1981] E. NEUENSCHWANDER – "Studies in the history of complex function theory. II Interactions among the French school, Riemann, and Weierstrass", *Bull. Amer. Math. Soc. (N.S.)* **5** (1981), pp. 87–105.

[Neum1865] C. NEUMANN – *Vorlesungen über Riemanns Theorie der abelschen Integrale*, Teubner, Leipzig, 1865.

[Neum1870] — , "Zur Theorie des Potentiales", *Math. Ann.* **2** (1870), p. 514.

[Neum1877] — , *Untersuchungen über das logarithmische und Newton'sche Potential*, Teubner, Leipzig, 1877.

[Neum1884] — , *Vorlesungen über Riemanns Theorie der abelschen Integrale* (second edition), Teubner, Leipzig, 1884.

[Nev1953] R. NEVANLINNA – *Uniformisierung*, Grund. Math. Wiss. LXIV. Springer, 1953.

[New1671] I. NEWTON – *De methodis fluxionum et serierum infinitorum...*, 1671, not published during Newton's life (also called *Artis analytica specimina sive Geometria analytica*). English translation in *The mathematical work of Isaac Newton, vol. III : 1670–1673*, Cambridge Univ. Press, 1969.

[New1704] — , "Enumeratio linearum tertii ordinis", appendix to *Opticks or a Treatise of the reflexions, refractions, inflexions and colours of light*, Smith et Walford, London, 1704. English translation (of 1710) in *The mathematical work of Isaac Newton, Vol. I*, Johnson Reprint Corp., New York, 1969.

[Noe1873] M. NOETHER – "Ueber Flächen welche Schaaren rationaler Curven besitzen", *Math. Ann.* **3** (1873), pp. 161–227.

[OsgB1998] B. OSGOOD – "Old and new on the Schwarz derivative", in *Quasiconformal mappings and analysis (Ann Arbor, 1995)*, Springer, New York, 1998, pp. 275–308.

[OsgW1898] W. OSGOOD – "Selected topics in the general theory of functions", *Amer. Math. Soc. Bull.* (2) **5** (1898), pp. 59–87.

[OsgW1900] W. OSGOOD – "On the existence of the Green's function for the most general simply connected plane region", *Trans. Amer. Math. Soc.* **1** (1900), pp. 310–314.

[OsgW1901] — , “Analytische Funktionen komplexe Grössen, in *Encyklopädie der mathematischen Wissenschaften*, Vol. II, Part 2 (1901) pp. 1–114.

[OsgW1902] — , “On the transformation of the boundary in the case of conformal mapping”, *Bull. Amer. Math. Soc.* **9** (1902) pp. 233–235.

[OsTa1913] W. Osgood & E. Taylor – “Conformal transformations on the boundaries of their regions of definition”, *Trans. Amer. Math. Soc.* **14** (1913), pp. 277–298.

[OvTa2005] V. Ovsienko & S. Tabachnikov – *Projective geometry, old and new. From the Schwarzian derivative to the cohomology of diffeomorphism groups*, Cambridge Univ. Press, 2005.

[OvTa2009] — , “What is . . . the Schwarzian derivative?”, *Notices Amer. Math. Soc.* **56** (2009), pp. 2–4.

[PaSe2004] I. Pan & M. Sebastiani – *Les équations différentielles algébriques et les singularités mobiles*, Ensaios Matemáticos, vol. 8, Sociedade Brasileira de Matemática, Rio de Janeiro, 2004.

[Per1923] O. Perron – “Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$ ”, *Math. Zeit.* **18** (1923), pp. 42–54.

[Pfe1917] G. Pfeifer – “On the conformal mapping of curvilinear angles. The fundamental equation $\phi(f(x)) = a\phi(x)$”, *Trans. Amer. Math. Soc.* **18** (1917), pp. 185–198.

[Pic1890] É. Picard – “Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives”, *J. Math. Pures Appl. (4th series)* **6** (1890), pp. 145–210. = [Pic1979], pp. 384–450.

[Pic1893a] — , “De l’équation $\Delta u = ke^u$ sur une surface de Riemann fermée”, *J. Math. Pures Appl. (4th series)* **9** (1893), pp. 273–292. = [Pic1979], pp. 525–543.

[Pic1893b] — , “Sur l’équation $\Delta u = e^u$ ”, *C. R. Ac. Sci. Paris* **116** (1893), pp. 1015–1017. = [Pic1979], pp. 521–523.

[Pic1893c] — , “Sur une équation aux dérivées partielles”, *C. R. Ac. Sci.* **116** (1893), pp. 454–456. = [Pic1979], pp. 517–519.

[Pic1893d] — , *Traité d'analyse, Vol. II*, Gauthier-Villars, Paris, 1893. Reprinted: Gabay, Sceaux, 1991.

[Pic1898] — , "De l'équation $\Delta u = e^u$ ", *J. de Math. Pures et Appl. (5th series)* **4** (1898), pp. 313–316. = [Pic1979], pp. 601–604.

[Pic1900] — , "De l'intégration de l'équation $\Delta u = e^u$ sur une surface de Riemann fermée", *Bull. Sci. Math.* **24** (1900), pp. 196–210. = *J. Reine Angew. Math* **130** (1905), pp. 243–258. = [Pic1979], pp. 649–664.

[Pic1931] — , *Quelques applications analytiques de la théorie des courbes et des surfaces algébriques*, Gauthier-Villars, Paris, 1931.

[Pic1979] — , *Œuvres*, Vol. II, CNRS, Paris, 1979.

[Plü1831] J. Plücker – *Analytisch-geometrische Entwicklungen*, Vol. 2, Baedeker, Essen, 1831.

[Poin1881a] H. Poincaré – "Sur les fonctions abéliennes", *C. R. Acad. Sci. Paris* **92** (1881), pp. 958–959. = [Poin1951], Vol. IV, pp. 299–301.

[Poin1881b] — , "Sur les fonctions fuchsiennes", *C. R. Acad. Sci. Paris* **92** (1881), pp. 333–335 (February 14); pp. 395–396 (February 21); pp. 859–861 (April 4). = [Poin1951], Vol. II, pp. 1–10.

[Poin1881c] — , "Sur les fonctions fuchsiennes", *C. R. Acad. Sci. Paris* **92** (1881), pp. 1198–1200. = [Poin1951], Vol. II, pp. 19–22.

[Poin1881d] — , "Sur les fonctions fuchsiennes", *C. R. Acad. Sci. Paris* **92** (1881), pp. 1484–1487. = [Poin1951], Vol. II, pp. 12–15.

[Poin1881e] — , "Sur les groupes kleinéens", *C. R. Acad. Sci. Paris* **93** (1881), pp. 44–46. = [Poin1951], Vol. II, pp. 23–25.

[Poin1882a] — , "Mémoire sur les fonctions fuchsiennes", *Acta Math.* **1** (1882), pp. 193–294. = [Poin1951], Vol. II, pp. 169–257.

[Poin1882b] — , "Sur les fonctions fuchsiennes", *C. R. Acad. Sci. Paris* **94** (1882), pp. 1038–1040. = [Poin1951], Vol. II, pp. 41–43.

[Poin1882c] — , "Sur les fonctions uniformes qui se reproduisent par des substitutions linéaires", *Math. Ann.* **19** (1882), pp. 565–568. = [Poin1951], Vol. II, pp. 92–105.

[Poin1882d] — , "Sur les formes cubiques ternaires et quaternaires", *J. Ec. Polytech.* **50/51** (1882), pp. 199–253 and pp. 45–91. = [Poin1951], Vol. V, pp. 28–72 and 293–334.

[Poin1882e] — , "Théorie des groupes fuchsiens", *Acta Math.* **1** (1882), pp. 1–62. = [Poin1951], Vol. II, pp. 108–168.

[Poin1883a] — , "Mémoire sur les groupes kleinéens", *Acta Math.* **3** (1883), pp. 49–92. = [Poin1951], Vol. II, pp. 258–299.

[Poin1883b] — , "Sur un théorème de la théorie générale des fonctions", *Bull. Soc. Math. France* **11** (1883), pp. 112–125. = [Poin1951], Vol. IV, pp. 57–69.

[Poin1884a] — , "Mémoire sur les fonctions zétafuchsiennes", *Acta Math.* **5** (1884), pp. 209–278. = [Poin1951], Vol. II, pp. 402–462.

[Poin1884b] — , "Sur les groupes des équations linéaires", *Acta Math.* **4** (1884), pp. 201–312. = [Poin1951], Vol. II, pp. 300–401.

[Poin1885a] — , "Sur les courbes définies par les équations différentielles", *J. Math Pures Appl. (4th series)* **1** (1885), pp. 167–244. = [Poin1951], Vol. I, pp. 3–84.

[Poin1885b] — , "Sur un théorème de M. Fuchs", *Acta Math.* **7** (1885), pp. 1–32. = [Poin1951], Vol. III, pp. 4–32.

[Poin1887] — , "Les fonctions fuchsiennes et l'arithmétique", *J. Math. Pures Appl. (4th series)* **3** (1887), pp. 405–464. = [Poin1951], Vol. II, pp. 463–511.

[Poin1888] — , "Sur une propriété des fonctions analytiques", *Rend. Circ. Mat. Palermo* **2** (1888), pp. 197–200. = [Poin1951], Vol. IV, pp. 11–13.

[Poin1890] — , "Sur les équations aux dérivées partielles de la physique mathématique", *Amer. J. Math.* **12** (1890), No. 3, pp. 211–294. = [Poin1951], Vol. IX, pp. 28–113.

[Poin1898] — , "Les fonctions fuchsiennes et l'équation $\Delta u = e^u$ ", *J. Math. Pures Appl. (5th series)* **4** (1898), pp. 137–230. = [Poin1951], Vol. IX, pp. 512–591.

[Poin1907] — , "Sur l'uniformisation des fonctions analytiques", *Acta Math.* **31** (1907), pp. 1–64. = [Poin1951], Vol. IV, pp. 70–139.

[Poin1908] — , "L'invention mathématique", talk delivered to the Institute of Psychology on May 23, 1908. *Enseignement Math.* **10** (1908), pp. 357–371.

[Poin1921] — , "Analyse des travaux scientifiques de Henri Poincaré faite par lui-même", *Acta Math.* **38** (1921), pp. 1–153. = [Poin1951], Vol. I, pp. I–CXXIX.

[Poin1951] — , *Œuvres*, 11 volumes, Gauthier-Villars, Paris, 1951.

[Poin1985] — , *Papers on Fuchsian functions*, translated into English, with commentary, by J. Stillwell, Springer, 1985.

[Poin1989] — , "La correspondance avec des mathématiciens de J à Z", *Cahiers du séminaire d'histoire des mathématiques* **10** (1989), pp. 83–229.

[Poin1997] — , *Trois suppléments sur la découverte des fonctions fuchsiennes*, edited by J. Gray and S. A. Walter. Publications of the Henri Poincaré Archive, Blanchard, Paris, 1997.

[Pois1820] S. D. Poisson – "Mémoire sur la manière d'exprimer les fonctions par des séries de quantités périodiques ", *J. Éc. Polytechnique* **11** (1820) pp. 417–489.

[Pom1975] C. Pommerenke – *Univalent functions*, Vandenhoeck & Ruprecht, Göttingen, 1975.

[Ponc1822] J.-V. Poncelet – *Traité des propriétés projectives des figures*, Bachelier, Paris, 1822.

[Pont1974] J.-C. Pont – *La topologie algébrique des origines à Poincaré*, Presses Univ. France, Paris, 1974.

[Pop2012] P. Popescu-Pampu – Qu'est-ce que le genre ? In *Histoires de mathématiques* Actes des Journées X-UPS 2011, 55–198, Ed. Ecole Polytechnique, Palaiseau, 2012.

[Pry1871] F. E. Prym – "Zur Integration der Differentialgleichung $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$", *J. Reine Angew. Math.* **73**, (1871) pp. 340–364.

[Pui1850] V. Puiseux – "Recherches sur les fonctions algébriques", *J. Math. Pures Appl. (1st series)* **15** (1850), pp. 365–480.

[Pui1851] — , "Nouvelles recherches sur les fonctions algébriques", *J. Math. Pures Appl. (1st series)* **16** (1851), pp. 228–240.

[Rad1925] T. Radó – "Über den Begriffe der Riemannsche Fläche", *Acta Litt. Sci. Szeged* **2** (1925), pp. 101–121.

[Rey1989] É. Reyssat – *Quelques aspects des surfaces de Riemann*, Progr. Math. 77, Birkhäuser, Boston, 1989.

[Ric1963] I. Richards – "On the classification of noncompact surfaces", *Trans. Amer. Math. Soc.* **106** (1963), pp. 259–269.

[Rie1851] B. Riemann – "Grundlagen für eine allgemeine Theorie der Functionen einer veränderlicher komplexer Grösse", Inauguraldissertation, Göttingen, 1851. = [Rie1892], pp. 3–45. English translation: "Foundations for a general theory of functions of a complex variable", in [Rie2004], pp. 1–42.

[Rie1854] — , "Über die Hypothesen, welche der Geometrie in Grunde liegen", Habilitationschrift, Göttingen, 1854. = [Rie1892], pp. 213–251. English translation: "The hypotheses on which geometry is based", in [Rie2004], pp. 257–272.

[Rie1857] — , "Theorie der Abelschen Functionen", *J. Reine Angew. Math.* **54** (1857), pp. 115–155. = [Rie1892], pp. 81–135. English translation: "The Theory of Abelian Functions" [sic], in [Rie2004], pp. 79–134.

[Rie1892] — , *Gesammelte mathematische Abhandlungen*, edited by H. Weber and R. Dedekind, Teubner, Leipzig, 1892. English translation: [Rie2004]. Reprinted with important commentaries in English: Teubner, Leipzig, & Springer, Berlin, 1990.

[Rie2004] — , *Collected Papers*, Translated from the 1892 German edition by Roger Baker, Charles Christenson and Henry Orde. Kendrick Press, Heber City, UT, 2004.

[Rit1908] W. Ritz – "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik.", *J. für Math.* **135** (1908), pp. 1–61.

[Roc1865] G. Roch – "Über die Anzahl der willkürlichen Constanten in algebraischen Functionen", *J. Reine Angew. Math.* **64** (1865), pp. 372–376.

[Row1989] D. Rowe – "Klein, Hilbert, and the Göttingen mathematical tradition", *Osiris* (1989), pp. 186–213.

[Rud1987] W. Rudin – *Real and complex analysis*, second edition, McGraw-Hill, New York, 1987.

[Sat1956] I. Satake – "On a generalization of the notion of manifold", *Proc. Nat. Acad. Sci. U.S.A.* **42** (1956), pp. 359–363.

[Scha1991] N. Schappacher, "Développement de la loi de groupe sur une cubique", *Seminar on the theory of numbers, Paris 1988/89*, Progress in Mathematics 91, (Birkhaüser) 1991, pp. 159–184.

[Schoe1896] A. Schoenflies – "Über einen Satz auf der Analysis Situs", *Nachr. Ges. Wiss. Göttingen* (1899), pp. 79–89.

[Schoe1906] — , "Beiträge zur Theorie der Punktmengen. III", *Math. Ann.* **62** (1906), pp. 286–328.

[Schol1980] E. Scholz – *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré*, Birkhäuser, Boston, 1980.

[Schot1877] F. Schottky – "Über die konforme Abbildung mehrfach zusammenhängender ebener Flächen", *J. Reine Angew. Math* **83** (1877), pp. 300–351.

[Schw1869] H. Schwarz – "Über einige Abbildungsaufgaben", *J. Reine Angew. Math* **70** (1869), pp. 105–120.

[Schw1870a] — , "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 Unstetigkeitbedingungen", *Monatsber. der Königlichen Akademie der Wissenschaft zu Berlin* (1870), pp. 767–795.

[Schw1870b] — , "Ueber einen Grenzübergang durch alternirendes Verfahren", *Wolf J.* **XV** (1870), pp. 272–286. = *Gesammelte mathematische Abhandlungen (2ter Band)* pp. 133–143, Springer, Berlin, 1890.

[Schw1873] — , "Über diejenige Fälle, in welchen die Gaussische hypergeometrische Reihe eine algebraische Function ihres vierten Elementes darstellt", *J. Reine Angew. Math.* **75** (1873), pp. 292–335.

[Sel1960] A. Selberg – "On discontinuous groups in higher-dimensional symmetric spaces", *Contributions to function theory (Internat.*

Colloq. Bombay, 1960), Tata Inst. Fund. Res., Bombay, 1960, pp. 147–164.

[Ser1970] J.-P. SERRE – *Cours d'arithmétique*, Presses Univ. France, Paris, 1970.

[Shi1959] G. SHIMURA – "On the theory of automorphic functions", *Ann. of Math.* **70** (1959), pp. 101–144.

[Shi1971] — , *Introduction to the arithmetic theory of automorphic functions. Kanô Memorial Lectures, No. 1*, Publ. Math. Soc. Japan, No. 11. Iwanami Shoten, Tokyo, & Princeton Univ. Press, 1971.

[Sieb2005] L. SIEBENMANN – "Vozvrashchenie k teoreme Osguda-Schoenfliesa", *Uspehi Mat. Nauk.* **60** (2005), No. 4, pp. 67–96. English translation: "The Osgood-Schoenflies theorem revisited", *Russ. Math. Surv.* **60** (2005), No. 4, pp. 645–672.

[Sieg1942] C. L. SIEGEL – "Iteration of analytic functions", *Ann. of Math. (2)* **43** (1942), pp. 607–612.

[Sny1993] J. SNYDER – *Flattening the Earth : Two Thousand Years of Map Projections*, Univ. of Chicago Press, 1993.

[Spr1957] G. SPRINGER – *Introduction to Riemann surfaces*, Addison-Wesley, Reading (Mass.), 1957.

[Ste1832] J. STEINER – *Systematische Entwicklung der Abhängigkeit geometrischer Gestalten von einander (1st part)*, G. Fincke, Berlin, 1832.

[Sti2002] J. STILLWELL – *Mathematics and its history*, second edition, Undergrad. Texts in Math., Springer, New York, 2002.

[Stu1913] E. STUDY – *Vorlesungen über ausgewälte Gegenstänstende der Geometrie (2nd part). Konforme Abbildung einfach zusammenhängender Bereiche*, Teubner, Berlin, 1913.

[Tak1977] K. TAKEUCHI – "Arithmetic triangle groups", *J. Math. Soc. Japan* **29** (1977), pp. 91–106.

[Taz1994] R. TAZZIOLI – "Il teorema di rappresentazione di Riemann: critica e interpretazione di Schwarz", *Rend. Circ. Mat. Palermo II*, **34** (1994), pp. 95–132.

[Taz2001] — , Green's function in some contributions of 19th century mathematicians. *Historia Mathematica* **28** (2001), 232–252.

[Thu1980] W. P. Thurston – *The geometry and topology of three-manifolds*, typescript, Princeton University, 1980. Electronic version: http://www.msri.org/publications/books/gt3m, 2002.

[Tie1913] H. Tietze – "Sur les représentations continues des surfaces en elles-mêmes", *C. R. Acad. Sci. Paris* **157** (1913), pp. 509–512.

[Tie1914] — , "Über einige Abbildungen einer Quadratfläche", *Rend. Circ. Mat. Palermo* **38** (1914), pp. 247–304.

[Ull2000] P. Ullrich – "The Poincaré-Volterra theorem: from hyperelliptic integrals to manifolds with countable topology", *Archiv Hist. Exact Sci.* **54** (2000), pp. 375–402.

[Val1945] G. Valiron – *Équations Fonctionnelles. Applications*, Masson, Paris, 1945.

[Veb1905] O. Veblen – "Theory of plane curves in non-metrical Analysis Situs", *Trans. Amer. Math. Soc.* **6** (1905), No. 1, pp. 83–98.

[Viv1888a] G. Vivanti – "Sulle functioni ad infiniti valori", *Rend. Circ. Mat. Palermo* **2** (1888), pp. 135–138.

[Viv1888b] — , "Ancora sulle functioni ad infiniti valori", *Rend. Circ. Mat. Palermo* **2** (1888), pp. 150–151.

[Vol1888] V. Volterra – "Sulle funzioni analitiche polidrome", *Atti R. Accad. Lincei, Rend. Ser. 4*, **16** (1888), pp. 355–361. = [Vol1964], pp. 356–362.

[Vol1964] — , *Opere mathematiche. Memorie e Note, I*, Acc. Lincei, 1964.

[Wae1941] B. L. van der Waerden – "Topologie und Uniformisierung der Riemannschen Flächen. Paul Koebe zum 60. Geburstag am 15 Februar 1942 gewidmet", *Ber. Sächs. Akad. Wiss. Leipzig, math.-phys. Kl.* **93** (1941), pp. 147–160.

[Weie1870] K. Weierstrass – "Über das sogenannte Dirichlet'sche Princip", read on July 14, 1870 at the Göttingen Academy. In *Mathematische Werke*, Mayer & Müller, Berlin, 1895, pp. 49–54.

[Weil1981] A. Weil – "Sur les origines de la géométrie algébrique", *Compositio Math.* **44** (1981), pp. 395–406.

[Wes1799] C. Wessel – "Om directionens analytiske betegning", *Nye Samling af del Kongelige Danske Videnskabernes Selkabs Skrifter*, Femte Del., Kjöbenhavn, 1799. pp. 469–518. Annotated French translation: "Essai sur la représentation analytique de la direction", Andr.-Fred Host og son, Copenhagen, 1897.

[Wey1913] H. Weyl – *Die Idee der Riemannschen Fläche*, Teubner, Leipzig, 1913. Reprinted with English commentary, Teubner, Stuttgart, 1997.

[Wey1955] — , *Die Idee der Riemannschen Fläche* [3rd edition], Teubner, Stuttgart, 1955. English translation: *The concept of a Riemann surface*, Addison-Wesley, Reading (Mass.), 1964.

[Wiel1977] N. J. Wielenberg – "Discrete Moebius groups: fundamental polyhedra and convergence", *Amer. J. Math.* **99** (1977), pp. 861–877.

[Wien1924a] N. Wiener – "Certain notions in potential theory", *J. Math. and Phys. (MIT)* **3** (1924), pp. 24–51.

[Wien1924b] — , "The Dirichlet problem", *J. Math. and Phys. (MIT)* **3** (1924), pp. 127–146.

[Wien1925] — , "Note on a paper by O. Perron", *J. Math. and Phys. (MIT)* **4** (1925), pp. 21–32.

[Wil1949] R. L. Wilder – *Topology of Manifolds*, Amer. Math. Soc. Coll. Publ., Vol. 32, Amer. Math. Soc., New York, 1949.

[You1924] G. C. Young – "Compte-rendu de lecture du volume 3 des œuvres de Klein" [Kle1921a], *Enseignement Math.* **24** (1924–1925), pp. 349–352.

[ZVC1970] H. Zieschang, E. Vogt & H.-D. Coldewey – *Flächen und ebene diskontinuierliche Gruppen*. Lect. Notes in Math. 122, Springer, Berlin, 1970. Translated by J. Stillwell: *Surfaces and planar discontinuous groups*, Lect. Notes in Math. 835, *ibid.*, 1980.

Index