Singular Surfaces and Meromorphic Functions

Mario Bonk

In the nineteenth century the theory of analytic functions grew into a major field of mathematical research. Mostly, mathematicians were interested in properties of specific classes of analytic functions such as elliptic functions. It was relatively late that general analytic functions became the focus of investigation. One of the early results in this respect is Picard's famous theorem of 1879, sometimes called "Picard's little theorem": *Every nonconstant meromorphic function in the complex plane* $\mathbb{C}$ *attains every value with at most two exceptions.*

Two exceptional values can really occur, as the exponential function shows; it omits 0 and ∞. Picard's theorem can easily be proved by resorting to a holomorphic universal covering map λ from the unit disk $\mathbb{D}$ onto the Riemann sphere $\hat{\mathbb{C}}$ punctured at 0, 1, and ∞. Namely, if a meromorphic function f omits three distinct values $a, b, c \in \hat{\mathbb{C}}$, then after composing f with a Möbius transformation, we may assume that $\{a, b, c\} = \{0, 1, \infty\}$. Then the map f has a holomorphic lift $\tilde{f}\colon \mathbb{C} \to \mathbb{D}$ such that $f = \lambda \circ \tilde{f}$. Now $\tilde{f}$ maps into the unit disk and is thus a bounded analytic function. According to Liouville's theorem the function $\tilde{f}$, and hence also f, is constant, proving the claim.

Mario Bonk is professor of mathematics at the University of Michigan, Ann Arbor. His e-mail address is `mbonk@math.lsa.umich.edu`.

The author would like to thank A. Eremenko, D. Drasin, J. Heinonen, P. Koskela, U. Lang, P. Poggi-Corradini, O. Schramm, and the editor for many useful comments.

Of course, Picard could not invoke these by now standard facts from the theory of covering spaces that ensure the existence of λ and the lift $\tilde{f}$. But at the time the function λ was explicitly known from the theory of elliptic modular functions. Picard defined $\tilde{f}$ locally as $\lambda^{-1} \circ f$ and then noted that there are no obstructions to analytic continuation of this analytic element along any path in the plane. Since the plane is simply connected, analytic continuation of $\tilde{f}$ along two paths with the same end points will lead to the same analytic element. In this way, we can define a single-valued holomorphic function $\tilde{f}$ on the entire complex plane.

According to a remark by Littlewood, this theorem and its proof could have been the shortest Ph.D. thesis in mathematics conceivable. Picard's theorem is a striking instance of a result that leaves one unsatisfied, because its short proof does not reveal the deeper reasons behind the statement. Indeed, after its discovery Picard's theorem stood in isolation for almost two decades. It was only in the late 1890s that the result was put into a broader context by work of E. Borel and J. Hadamard. This development culminated in the 1920s with R. Nevanlinna's value distribution theory for meromorphic functions, which provides far-reaching generalizations of Picard's theorem.

An important intermediate step in these investigations was the desire to obtain an *elementary* proof of Picard's theorem. Here "elementary" means avoiding the use of the elliptic modular function λ. After several such proofs had been given, in 1924 A. Bloch based his proof of Picard's theorem on a fundamental covering property of holomorphic functions that had not been noticed before [3]: *Every holomorphic function f in the unit disk $\mathbb{D}$ normalized by $|f'(0)| = 1$ univalently covers a disk of radius c_0, where $c_0 > 0$ is a numerical constant independent of f.*

Here we say that a disk (or a more general region) D is *univalently covered* by f, if there exists a subregion $U \subset \mathbb{D}$ such that the restriction $f|U$ is a one-to-one conformal map of U onto D. In other words, D is univalently covered if we can find a holomorphic branch of the inverse function f^{-1} defined on all of D.

Bloch's theorem says in particular that the image $f(\mathbb{D})$ of *every* holomorphic function f in $\mathbb{D}$ contains a disk of a fixed size given that $|f'(0)| = 1$. The best constant c_0 in Bloch's theorem (more precisely, the supremum of all constants c_0) is now called Bloch's constant B. Its precise value is still not known, but one has the estimates

$$.433 < B < .472$$

(see [8] for the most recent developments). An immediate corollary of Bloch's theorem is the following theorem of G. Valiron: *A nonconstant entire function univalently covers arbitrarily large disks.* It seems that this result was first stated explicitly by Bloch, who extracted it from work of Valiron.

In order to derive Picard's theorem from Valiron's theorem, let us assume, as we may, that a meromorphic function f in the plane omits the values $-1, 1, \infty$ from its range. Now one can use f to form other functions with an even larger set of omitted values. For example, just as in Picard's proof, the branches of $\arccos(f)$ lead to holomorphic functions in $\mathbb{C}$. Let $g = \frac{1}{\pi}\arccos(f)$ for one such branch. Then g is holomorphic and omits all the integers, in particular -1, 1, ∞. Hence we can repeat this procedure once more and form $h = \arccos(g)$. This will lead to an entire function omitting the values

$$k\pi \pm i\cosh^{-1}(m),\ k \in \mathbb{Z},\ m \in \mathbb{N}.$$

It is easy to see that these values form a set which does not contain arbitrarily large disks in its complement. Hence h is constant by Valiron's theorem, and so are g and f.

The preceding discussion shows that the value distribution theory of meromorphic functions (as exemplified by Picard's theorem) is connected to their covering properties. The main goal of this article is to present some recent developments in this area.

Covering Properties of Meromorphic Functions

The connection between covering properties and value distribution of analytic functions was systematically explored by L. Ahlfors in the 1930s. He was apparently motivated by Bloch's paper [4] and set out to put some of Bloch's speculations on a firm basis. Ahlfors's most beautiful result in this respect is his "Five-Islands Theorem": *Let f be a nonconstant meromorphic function in the plane, and let $\Omega_1, \dots, \Omega_5$ be five Jordan regions on the Riemann sphere $\hat{\mathbb{C}}$ with pairwise disjoint closures. Then one of the regions is covered univalently by f.* This theorem is a special case of a more general "Scheibensatz" proved in the paper [2] that earned Ahlfors one of the first two Fields Medals in 1936. A very elegant new proof of the Five-Islands Theorem has recently been found by W. Bergweiler.

An immediate consequence of the Five-Islands Theorem is the following Valiron-type covering theorem for meromorphic functions and the spherical metric [1]: *Every nonconstant meromorphic function in the plane univalently covers spherical disks of radii arbitrarily close to $\pi/4 = 45°$.*

Here we metrically identify the Riemann sphere $\hat{\mathbb{C}}$ with the unit sphere in $\mathbb{R}^3$ and use the intrinsic spherical metric to measure distances. So for example, a hemisphere can be considered as a disk of radius $\pi/2 = 90°$. In the usual model $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the spherical metric is given by the length element

$$\frac{2|dz|}{1 + |z|^2}.$$

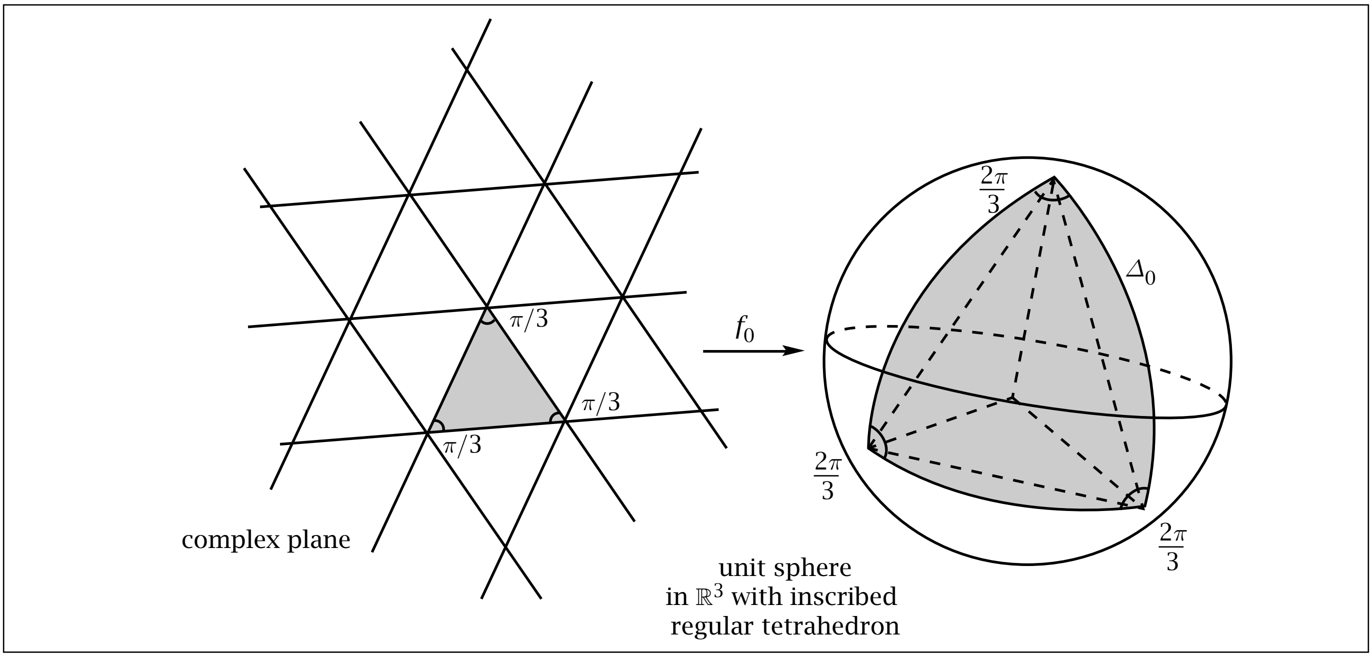
In order to derive the above covering property from the Five-Islands Theorem, choose five spherical disks of equal size as the regions Ω_i. Note that one can find four disjoint open disks of radius $45°$ with centers on a great circle, say the equator. Then there is space for another open disk with radius $45°$ disjoint from the others, if we center it at the north pole. Actually, there would even be space for a sixth such disk centered at the south pole, but this extra space cannot be used to increase the radii of the first five disks while keeping them disjoint. The five disks will touch at boundary points. In order to apply the Five-Islands Theorem, where we require *disjoint closures* of the regions, we shrink the radii of the disks by an arbitrarily small amount $\epsilon > 0$, and the result follows.

The covering theorem for meromorphic functions seems to be of a more fundamental nature than Bloch's covering theorem, because a fixed sized spherical disk is covered by *every* nonconstant meromorphic function independent of any normalization for the derivative of the function.

The precise value of Bloch's constant is not known. In contrast, A. Eremenko and the author recently proved the following sharp constant result [5].

Theorem 1. *Every nonconstant meromorphic function in the complex plane univalently covers spherical disks of radii arbitrarily close to*

$$\arctan(\sqrt{8}) \approx 70°32'.$$



The map f_0 maps an equilateral Euclidean triangle onto an equilateral spherical triangle with angles $2\pi/3$. The triangle Δ_o is the projection of a face of a regular tetrahedron inscribed in the unit sphere.

The number $\arctan(\sqrt{8})$ *is best possible.*

It is a nice exercise to give yet another proof of Picard's theorem by using this statement (take a cube root of a function omitting $\{0, 1, \infty\}$ and use the geometric interpretation of $\arctan(\sqrt{8})$ discussed below). Indeed, one can even derive the Five-Islands Theorem by using some standard machinery of quasiconformal mappings, but we will not go into this. It is an open problem whether the phrase "arbitrarily close" in Theorem 1 is essential, i.e., whether every nonconstant meromorphic function in the complex plane univalently covers a spherical disk of radius $\arctan(\sqrt{8})$.

The following example shows why the number $\arctan(\sqrt{8})$ in Theorem 1 cannot be replaced by any larger number. We consider a conformal map f_0 of an equilateral Euclidean triangle onto an equilateral spherical triangle Δ_0 with angles $2\pi/3$.

The Riemann mapping theorem guarantees that we can choose f_0 so that this map sends vertices to vertices. By repeated application of the Schwarz reflection principle, we can find an analytic continuation of the function f_0 as a meromorphic function in $\mathbb{C}$. At the vertices of our original Euclidean triangle and at all points corresponding to these vertices under successive reflections, the map f_0 doubles angles. Hence these points are critical points of f_0, i.e., zeros of the derivative. There are no other critical points. So the critical points of the function f_0 form a regular hexagonal lattice, and its critical values, the images of the critical points, correspond to the four vertices of a regular tetrahedron inscribed in the unit sphere.

If we place one of the vertices of the tetrahedron at the point corresponding to $\infty \in \hat{\mathbb{C}}$ and normalize the map by $z^2 f_0(z) \to 1$ as $z \to 0$, then f_0 becomes a Weierstrass $\wp$-function with a hexagonal lattice of periods. It satisfies the differential equation

$$(\wp')^2 = 4(\wp - e_1)(\wp - e_2)(\wp - e_3),$$

where the numbers e_j correspond to the three remaining vertices of the tetrahedron.

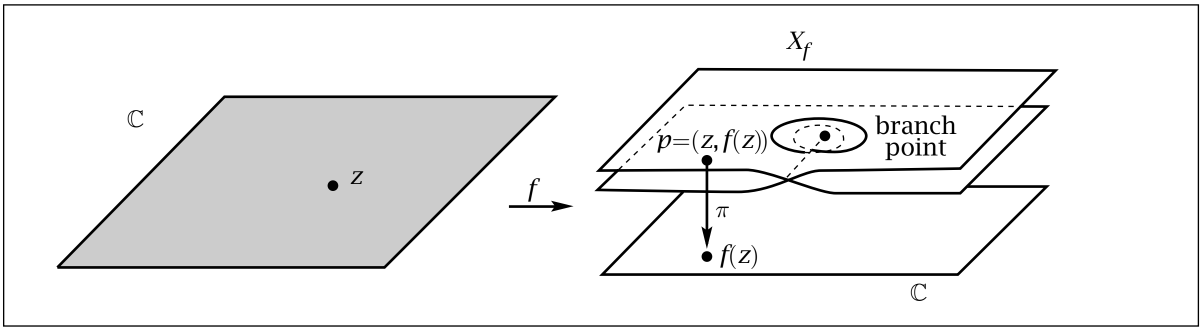
For a spherical disk to be covered univalently by f_0, it has to avoid the critical values of f_0, since near these points no branch of the inverse is locally holomorphic. Indeed, these points, the four vertices of our tetrahedron, are the only obstructions. So we obtain a largest spherical disk covered univalently by f_0 if we place its center at the center of Δ_0 and let its radius be equal to the circumscribed radius b_0 of Δ_0. Computation yields

$$b_0 = \arctan(\sqrt{8}).$$

The remainder of this article will be used to explain some of the ideas that go into the proof of Theorem 1. The main notions here are Riemann surfaces of analytic functions, the classical type problem, and curvature properties of surfaces with singularities. We hope to convey to the reader some of the appeal that the ensuing questions about surface geometry exhibit.

The Problem of Type and Riemann Surfaces of Analytic Functions

The classical uniformization theorem implies that every open simply connected Riemann surface X is biholomorphic either to the unit disk $\mathbb{D}$ or to the complex plane $\mathbb{C}$. According to these two cases, one says that X has *hyperbolic* or *parabolic conformal*



Schematic drawing of a multisheeted branched covering surface arising from a holomorphic function in the plane.

type. As with all classification theorems, it is important to know criteria for deciding in which class a given object belongs. In the case of the uniformization theorem we are lead to the *type problem*: how to decide the type of an open simply connected Riemann surface, if it is given in some explicit geometric way, for example as a branched covering surface over $\mathbb{C}$ or $\hat{\mathbb{C}}$. This question has been extensively studied since the 1930s.

If $f\colon \Omega \to \mathbb{C}$ is a holomorphic function defined in some region $\Omega \subset \mathbb{C}$, then we can associate with f the branched covering surface $X_f = \{(z, f(z)) : z \in \Omega\}$ over $\mathbb{C}$. The way to think of X_f becomes more intuitive if we use the following picture. A point $p = (z, f(z)) \in X_f$ should be thought of as sitting "above" the image point $w = f(z)$ of z. In this way, over each point $w \in \mathbb{C}$ there will be a fiber $f^{-1}(w)$ of several points z corresponding to its preimages under f. These fibers are merged to give the surface X_f which has the same topology as Ω. So if two points z_1 and z_2 are close, then they remain close after they are moved to the location "above" their image points $w_1 = f(z_1)$ and $w_2 = f(z_2)$.

We equip X_f with the metric coming from the pull-back of the Euclidean length element in the w-plane by the projection map $\pi\colon X_f \mapsto \mathbb{C}$, $(z, f(z)) \mapsto w = f(z)$. Then π will be a local isometry away from the discrete set of *branch points* of f, i.e., the points corresponding to the critical points of f. As a metric space, X_f is isometric to Ω equipped with the length element

$$|f'(z)|\,|dz|.$$

The map $z \mapsto (z, f(z))$ is a biholomorphism from Ω onto X_f. If Ω is simply connected, then X_f has the same property and it makes sense to speak of the type of X_f. In this case, the type of X_f is the same as the type of Ω, i.e., it is parabolic if $\Omega = \mathbb{C}$ and hyperbolic otherwise.

In this picture of the Riemann surface X_f as a branched covering surface of $\mathbb{C}$, it is easy to see when a disk $D \subset \mathbb{C}$ is covered univalently by f. This happens precisely if we can find a *schlicht disk* $U \subset X_f$ on the surface X_f, avoiding the "boundary" and the branch points of X_f, so that the projection map $\pi|U$ is a bijection (and hence an isometry) from U onto D.

If $f\colon \Omega \to \hat{\mathbb{C}}$ is a meromorphic function, then its Riemann surface Z_f is defined in a similar way. We also equip Z_f with a metric, this time by pulling-back the spherical length element. As a metric space, Z_f is isometric to Ω equipped with the length element

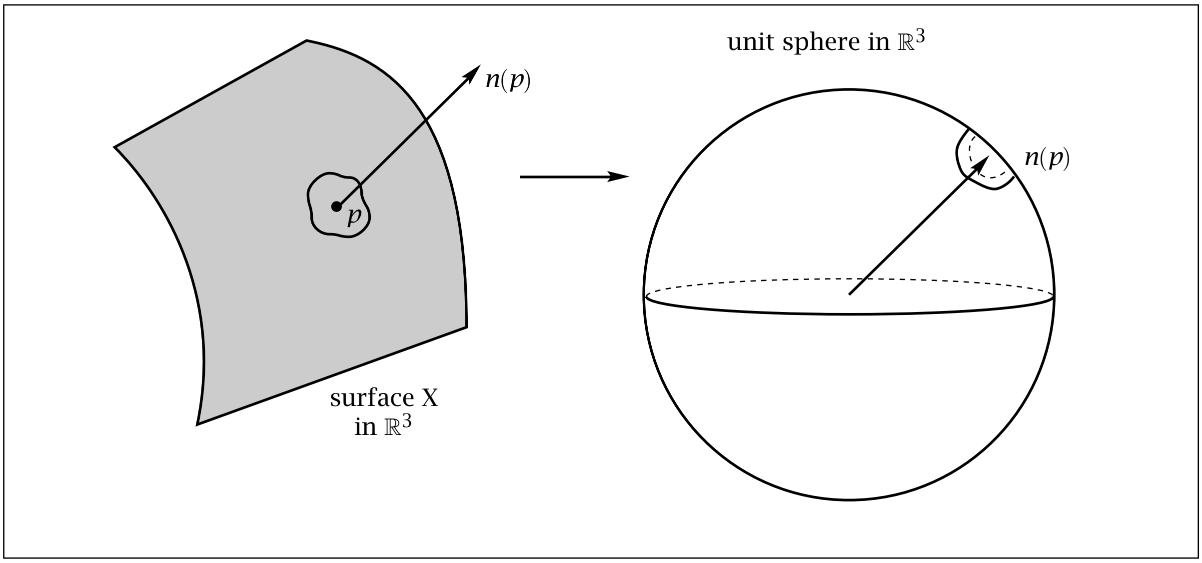
$$\frac{2|f'(z)|}{1+|f(z)|^2}\,|dz|.$$

Again the type of Z_f is the same as that of Ω, if Ω is simply connected.

One can derive covering properties of meromophic functions from type criteria as follows. Suppose we have some geometric condition that forces hyperbolic type of an open simply connected Riemann surface. If f is a meromorphic function in the plane, then its Riemann surface Z_f is of parabolic type and so it cannot satisfy the geometric condition in question. For example, we will see that if the branch points are distributed very "densely" on a surface, then the surface is of hyperbolic type. This implies that the branch points of the Riemann surface of a meromorphic function in the plane cannot be too dense. We will see that ideas like this lead to the proof of Theorem 1.

Polyhedral Surfaces

Let us have a closer look at the geometry of the Riemann surface X_f of a holomorphic function f and at curvature properties of this surface. Away from the set of branch points B_f, the surface is smooth and each point has a neighborhood isometric to a subset of $\mathbb{C}$. Hence $X_f \setminus B_f$ is "flat". Suppose z_0 is a critical point of f. Then one can find local biholomorphisms ϕ and ψ mapping z_0 and $w_0 = f(z_0)$ to the origin, respectively, so that the map $g := \psi \circ f \circ \phi^{-1}$ has the form $g(u) = u^k$ near the origin, where $k \in \mathbb{N}$, $k \geq 2$. It follows that f is k-to-1 near z_0. The number k is called the local degree of f at z_0. The local geometry of X_f near $p_0 = (z_0, f(z_0))$ can be described as follows: There is a neighborhood of p_0 which is isometric to a truncated Euclidean cone $C_a(\epsilon)$, where $a = k$. For arbitrary $a > 0$ we let $C_a(\epsilon)$ be the



The Gauss map.

disk $D(0, \epsilon)$ in the complex plane equipped with the length element

$$(0.1) \qquad |z|^{a-1}|dz|.$$

For $0 < a < 1$ the metric space $C_a(\epsilon)$ really looks like it should, namely like an ice-cream cone. For $a = 1$ we just get a flat Euclidean disk.

In our case, where $a = k$, we can obtain $C_k(\epsilon)$ also as follows. Take k copies of the open Euclidean disk of radius $\epsilon' = \epsilon^k/k$ centered at the origin. Make a slit into each of the disks along the radius $[0, \epsilon')$. These slits produce two "banks" in each disk; call them the upper and lower banks. Now glue the slit disks along the banks in cyclic order, so that the lower bank of a disk is glued to the upper bank of the slit of its successor. In this way, $C_k(\epsilon)$ can be visualized as a spiral staircase (with inevitable self-intersection if we want to realize this space in $\mathbb{R}^3$ in the way described).

If $a \neq 1$, then the vertex o of $C_a(\epsilon)$ (corresponding to the center of $D(0, \epsilon)$) is a singularity, where the surface ceases to be smooth. Nevertheless, this *conical singularity* is of a benign nature. Surfaces that are flat everywhere except at some isolated conical singularities of the above type are called *Euclidean polyhedral surfaces*. A singularity p of a Euclidean polyhedral surface can be detected from the *total angle* σ at p. This is the total angular variation of the directions issuing from p, if one goes around p once. So for $o \in C_a(\epsilon)$ the total angle σ is equal to $2\pi a$ and equal to 2π at all other points of $C_a(\epsilon)$. In this way the conical singularities of a Euclidean polyhedral surface can be described as the set of points where the total angle differs from 2π.

There is a well-developed theory of surfaces, created by A. D. Aleksandrov and his students in the 1950s, that includes both smooth and polyhedral surfaces and provides us with the usual amenities of differential geometry such as area, length, and angles. These surfaces are called *surfaces of bounded curvature in the sense of Aleksandrov*, or *Aleksandrov surfaces* for short (see [10] for a survey). For our purposes it suffices to say that an Aleksandrov surface is a surface where the length element can locally be expressed by complex coordinates z in the form

$$(0.2) \qquad \rho(z)\,|dz|,$$

where

$$(0.3) \qquad \rho(z) = \exp(u(z))$$

and u is the difference of two subharmonic functions. In addition, we require that the length element (0.2) is locally integrable along analytic curves. Note that, in general, u and ρ are nonsmooth functions. Complex coordinates for which the length element on the surface has the local representation (0.2) are called *isothermal* coordinates.

If the length element on a surface can be expressed locally as in (0.1), where $a > 0$, then it is locally integrable along analytic curves, and we have $u(z) = (a-1)\log|z|$. The function u is subharmonic or superharmonic depending on whether $a \geq 1$ or $a \leq 1$. This shows that Euclidean polyhedral surfaces and, in particular, Riemann surfaces of holomorphic functions are Aleksandrov surfaces.

On each Aleksandrov surface X one can define the *integral curvature*, denoted by μ (the surface is understood), as a measure on Borel sets of the surface.

The integral curvature of a (Borel) subset E of a smooth embedded surface X in $\mathbb{R}^3$ is the area of the image of E under the Gauss map. The Gauss map on X (for a fixed orientation of the surface) is the map into the unit sphere that associates with every point $p \in X$ the unit normal vector $n(p)$ to the surface X at p, where $n(p)$ is considered as a point in the unit sphere.

According to Gauss's Theorema Egregium, the integral curvature μ is completely determined by the inner geometry of the surface and does not depend on the particular embedding. Indeed, if we represent the surface locally by isothermal coordinates as in (0.2) and (0.3) which is always possible (and if we disregard the distinction between the set E and its image set under a coordinate map), then μ can be expressed as

$$\mu(E) = -\int_E \Delta \log \rho = -\int_E \Delta u,$$

where integration is with respect to Euclidean area in the z-plane. This formula remains valid on every Aleksandrov surface, if we interpret the Laplacian in a distributional sense.

Let us come back to our case of Euclidean polyhedral surfaces. Since they are Aleksandrov surfaces, it makes sense to attribute an integral curvature $\mu(p) := \mu(\{p\})$ to a conical singularity p. It is equal to $2\pi - \sigma$, where σ is the total angle at p. In the case of a Riemann surface X_f of a holomorphic function we have $\sigma = 2\pi k$, $k \in \mathbb{N}$, and so $\mu(p) = 2\pi(1-k)$ for $p \in X_f$. Since $k \geq 2$ for every branch point, we see that each branch point carries some definite mass of negative curvature. As X_f is flat everywhere else, the Riemann surface of a holomorphic function is a Euclidean polyhedral surface that is nonpositively curved.

Uniformly Hyperbolic Surfaces

Every Aleksandrov surface carries an essentially unique Riemann surface structure determined by isothermal coordinates. In particular, we can ask for its conformal type if the surface X is open and simply connected.

Suppose E is a continuum in such a surface X. Then X is of parabolic type if and only if for every $\epsilon > 0$ there exists a Borel measurable density $\lambda \colon X \to [0, \infty]$ such that

$$\int_X \lambda^2 \, dA < \epsilon,$$

where integration is with respect to area A on the surface, and

$$\int_\gamma \lambda \, ds \geq 1$$

for every locally rectifiable path γ connecting E to "infinity", i.e., γ starts at E and eventually leaves every compact subset of the surface. In the last integral, integration is with respect to arc length. If this condition is true for *some* continuum $E \subset X$, then it is actually satisfied for *all* continua $E \subset X$. In plain words, a surface is parabolic if a continuum E can be blocked from infinity at arbitrarily small "cost" and hyperbolic otherwise. The reader might find it a worthwhile exercise to check the parabolicity of the complex plane by using this criterion.

As we have seen, the Riemann surfaces of holomorphic functions are always nonpositively curved. In view of this, it is natural to start the investigation of the type question for surfaces that resemble these Riemann surfaces, namely nonpositively curved Aleksandrov surfaces. Roughly speaking, a dominant presence of negative curvature tends to make the surface hyperbolic, while the surface is parabolic if the negative curvature is diluted over the surface. So one should expect hyperbolic type if the negative curvature is somehow uniformly distributed, say if every disk of a fixed radius R_0 contains a definite amount of negative curvature. Indeed, the surfaces of this type admit various equivalent characterizations given in Theorem 2 below [6]. It is interesting that this result links several important notions that have found applications in other areas of mathematics. The new concepts in conditions (ii)–(iv) of this theorem will be briefly explained after the statement, but a full understanding of them is not really necessary for the rest of the article.

Theorem 2. *Let X be an open simply connected nonpositively curved Aleksandrov surface. Then the following conditions are equivalent:*

(i) *There exist $R_0 > 0$ and $\epsilon > 0$ such that every relatively compact open disk $D(a, R_0) \subset X$ has integral curvature less than $-\epsilon$.*

(ii) *X is hyperbolic in the sense of Gromov.*

(iii) *X satisfies a linear isoperimetric inequality.*

(iv) *The (differential) Kobayashi metric on X is uniformly bounded from below.*

Moreover, if X satisfies any of these conditions, then X is necessarily of hyperbolic type.

The *Gromov hyperbolicity* of a metric space (X, d) means that there exists a constant $\delta \geq 0$ such that for all points $x, y, z, w \in X$

$$(x, z)_w \geq \min\{(x, y)_w, (y, z)_w\} - \delta, \tag{0.4}$$

where $(u, v)_w := (1/2)\{d(u, w) + d(v, w) - d(u, v)\}$. A more intuitive definition can be given for complete surfaces X; in this case, the metric space X is *geodesic*, which means that every two points $x, y \in X$ can be joined by a *geodesic segment* $[x, y]$, i.e., a curve whose length is equal to the distance of x and y. Then the Gromov hyperbolicity of a geodesic metric space X is equivalent to the existence of a constant $\delta' \geq 0$ such that

$$\operatorname{dist}(u, [x_1, x_3] \cup [x_3, x_2]) \leq \delta',$$

whenever $x_1, x_2, x_3 \in X$ and $u \in [x_1, x_2]$. In other words, geodesic triangles $[x_1, x_2] \cup [x_2, x_3] \cup [x_3, x_1]$ are rather "thin". The concept of Gromov hyperbolicity encodes one of the basic properties of negatively curved spaces. It was originally utilized in the theory of discrete groups but has recently found applications in other areas (see [9] for general background and [7] for some recent developments).

The surface X satisfies a *linear isoperimetric inequality* if there exists a constant $C > 0$ such that all Jordan regions $\Omega \subset X$ satisfy

$$A(\Omega) \le C\ell(\partial\Omega),$$

where A stands for area and ℓ for length. Again, a linear isoperimetric inequality is a distinctive feature of negatively curved spaces. Ahlfors [2] was one of the first to recognize the importance of this concept. The equivalence of Gromov hyperbolicity and a linear isoperimetric inequality (if appropriately defined) is true in much greater generality.

Finally, the *Kobayashi metric* is one of several invariant metrics that occur in the theory of complex manifolds. Its uniform boundedness from below means in our case that there exists a constant $C > 0$ such that for all holomorphic maps $f: \mathbb{D} \to X$ we have $\|f'(z)\| \le C$ for $z \in \mathbb{D}$, where the norm $\|f'(z)\|$ of the derivative is the ratio at z of the pull-back of the length element on X to the length element

$$\frac{2|dz|}{1-|z|^2}$$

of the Poincaré metric of constant negative curvature -1 on $\mathbb{D}$.

We call surfaces satisfying the hypotheses and any of the equivalent conditions (i)–(iv) in Theorem 2 *uniformly hyperbolic*. We remark that as a mere type criterion this theorem is a rather weak statement. As one would expect, much better sufficient conditions for hyperbolic type are available.

As an illustration of how to apply Theorem 2, we will derive Valiron's theorem as an easy corollary. So suppose $f: \mathbb{C} \to \mathbb{C}$ is a nonconstant holomorphic function. Then the Riemann surface X_f of f is of parabolic type. Suppose that f does not univalently cover arbitrarily large disks. Let us assume that no disk of radius $R_0 > 0$ is covered univalently. So if a disk $D = D(a, R_0)$ is compactly contained in X_f, then D must contain a branch point p_0 of X_f. For otherwise, this disk would be a schlicht disk over the disk $D' = D(\pi(a), R_0) \subset \mathbb{C}$, and so D' would be covered univalently by f. Hence the integral curvature of D is bounded above by the integral curvature at p_0 which is $\le -2\pi$. Therefore, X_f satisfies condition (i) in Theorem 2, which forces the hyperbolicity of the surface. This is a contradiction.

Two other corollaries of Theorem 2 are worth pointing out. The first one deals with a very explicit case of polyhedral surfaces which can be obtained as follows: Take a triangulation of the plane by topological triangles, and assign a Euclidean triangle to each triangle of the triangulation. Assume that the lengths of the sides of two Euclidean triangles are the same if the sides correspond to the same edge in the triangulation. Then we can glue the Euclidean triangles together according to the combinatorial pattern of the triangulation and obtain a Euclidean polyhedral surface. We emphasize that the reader should visualize this surface as an abstract metric space not tied to any particular embedding into a Euclidean space. Theorem 2 now gives a sufficient condition for a surface of this type to be uniformly hyperbolic.

Corollary 1. *Let X be a Euclidean polyhedral surface homeomorphic to the plane which is equipped with a triangulation T. Suppose that each triangle $\Delta \in T$ is isometric to a Euclidean triangle and that there exist constants $\epsilon, C > 0$ such that:*

(i) $\mathrm{diam}(\Delta) \le C$ *for each triangle $\Delta \in T$.*

(ii) *The total angle at each vertex of T is $\ge 2\pi + \epsilon$.*

Then X is uniformly hyperbolic (and hence of hyperbolic type).

A holomorphic function $f: \mathbb{D} \to \mathbb{C}$ is called a *Bloch function* if there exists a constant $C \ge 0$ such that

$$|f'(z)| \le \frac{C}{1-|z|^2} \text{ for } z \in \mathbb{D}.$$

In other words, f is Bloch if it is a holomorphic Lipschitz map from the unit disk equipped with the Poincaré metric into the plane equipped with the Euclidean metric. Bloch functions were first considered in connection with Bloch's constant B defined above. They play an important role in the study of the boundary behavior of conformal maps.

Corollary 2. *Suppose $f: \mathbb{D} \to \mathbb{C}$ is holomorphic. Then f is a Bloch function if and only if its Riemann surface X_f is Gromov hyperbolic.*

Spherical Polyhedral Surfaces

In the previous section we have seen that Theorem 2 implies Valiron's theorem. This is an example of how results on surface geometry can be used to obtain results on covering properties of holomorphic functions.

In order to prove Theorem 1, which is about *meromorphic* functions one has to consider surfaces with polyhedral metric, but in this case the surfaces will be piecewise spherical instead of piecewise Euclidean.

A *spherical polyhedral surface Z* is an Aleksandrov surface with intrinsic metric so that each point has a neighborhood isometric to a truncated spherical cone, i.e., to the (Euclidean) disk $D(0,\epsilon)$ equipped with the length element

$$\frac{2a|z|^{a-1}|dz|}{1+|z|^{2a}}.$$

If $a=1$, a truncated cone is isometric to a spherical cap; for $a=k\in\mathbb{N}$, one can obtain the space as in the Euclidean case by gluing together k slit spherical caps.

Apart from isolated points on Z we have $a=1$. Near these regular points the local geometry coincides with the geometry on the unit sphere; in particular, the integral curvature of a set on the surface not containing any singularities is equal to the area of the set. At the singular points where $a\neq 1$, the total angle is $2\pi a$, and so the integral curvature has a Dirac mass $2\pi(1-a)$.

Now suppose that Z_f is the Riemann surface of a meromorphic function f. Then Z_f is a spherical polyhedral surface. The singular points of this surface are its branch points and correspond to the critical points of f. If the order of the critical point is $k\in\mathbb{N}$, $k\geq 2$, then $a=k$. In particular, each critical point of order k of f gives rise to a singular point carrying negative curvature of mass $2\pi(1-k)\leq -2\pi$. So the negative curvature of the surface Z_f is sitting in the branch points, and the other parts of the surface contribute to the positive curvature according to their areas.

The following theorem is analogous to Corollary 1.

Theorem 3. *Let Z be a spherical polyhedral surface homeomorphic to the plane which is equipped with a triangulation T. Suppose that each triangle $\Delta\in T$ is isometric to a spherical triangle and that there exists a constant $\epsilon>0$ such that:*

(i) *The circumscribed radius of each triangle Δ of T is $\leq \arctan(\sqrt{8})-\epsilon$.*

(ii) *The total angle at each vertex of T is $\geq 4\pi$.*

Then Z is of hyperbolic type.

Let us give a heuristic argument that explains the occurrence of the number $\arctan(\sqrt{8})$ in this theorem. Indeed, as we remarked above, this number is the circumscribed radius of an equilateral spherical triangle Δ_0 with all angles equal to $2\pi/3$. It is easy to believe and not hard to see that this triangle has largest area among all spherical triangles with circumscribed radius at most $\arctan(\sqrt{8})$. Since for our surface Z the positive part of the integral curvature of any set is equal to its area, each triangle in the triangulation T carries total positive curvature at most

$$A(\Delta)\leq A(\Delta_0)=\pi.$$

Let α,β,γ be the angles at the vertices of Δ. Then

$$A(\Delta)=\alpha+\beta+\gamma-\pi\leq\pi,$$

and so

$$\alpha+\beta+\gamma\leq 2\pi.$$

Now consider the negative curvature that one should attribute to the triangle Δ. The negative curvature sits in the vertices of the triangulation. We know that at each vertex v we have curvature $\mu(v)=2\pi-\sigma$, where σ is the total angle at v. In our case $\sigma\geq 4\pi$, and so we have for the negative part of the curvature

$$\mu^-(v)=-\mu(v)\geq\sigma/2. \tag{0.5}$$

Now let us "distribute" this curvature at v to the triangles in T which have v as a vertex. One way to do this is to let each triangle have a share of the curvature at v proportional to the angle of the triangle at this vertex. Then according to (0.5), each triangle will acquire negative curvature μ^- at least half as large as the sum of its angles. So the total curvature balance for Δ is

$$\begin{aligned}\mu(\Delta)&\leq-\frac{1}{2}(\alpha+\beta+\gamma)+A(\Delta)\\&=\frac{1}{2}(\alpha+\beta+\gamma)-\pi\leq 0.\end{aligned}$$

This means that the negative curvature on Z at least compensates for the positive curvature. The additional ϵ in the theorem forces the area of the triangles to be a definite amount smaller than the area of Δ_0. This tips the curvature balance in favor of the negative curvature, and we may have reason to expect hyperbolicity of our surface as Theorem 3 claims.

Is there an analog of Theorem 2 for spherical polyhedral surfaces? The answer is no. Indeed, there exist spherical polyhedral surfaces Z such that every disk of some fixed radius R_0 carries some definite negative curvature $\leq -\epsilon$, but the surface is still of parabolic type. As was pointed out by O. Schramm, an example of such a surface can be obtained as follows: Take a square grid in the plane. For each square let there be a corresponding hemisphere with four equally spaced points on its boundary. These points decompose the boundary of the hemisphere into four arcs of equal length corresponding to the sides of the square. Now glue the hemispheres according to the incidence pattern of the squares. We obtain a spherical polyhedral surface Z which happens to be isometric to the Riemann surface of the Weierstrass $\wp$-function satisfying the differential equation

$$(\wp')^2=4\wp(\wp^2-1).$$

In particular, the surface is of parabolic type. We leave it as an exercise for the reader to find explicit numbers $R_0>0$ and $\epsilon>0$ with the required properties.

It seems that in order to get a hyperbolicity criterion for a spherical polyhedral surface one has to impose a very tight control of the negative curvature over the positive curvature.

It is a natural question whether Theorem 3 holds for other appropriate values in conditions (i) and (ii): *Suppose X is as in Theorem 3 and suppose that the total angle at each vertex is at least $2\pi q$ with $q \in (1, 3]$ and that the supremum of the circumscribed radii of the triangles is at most*

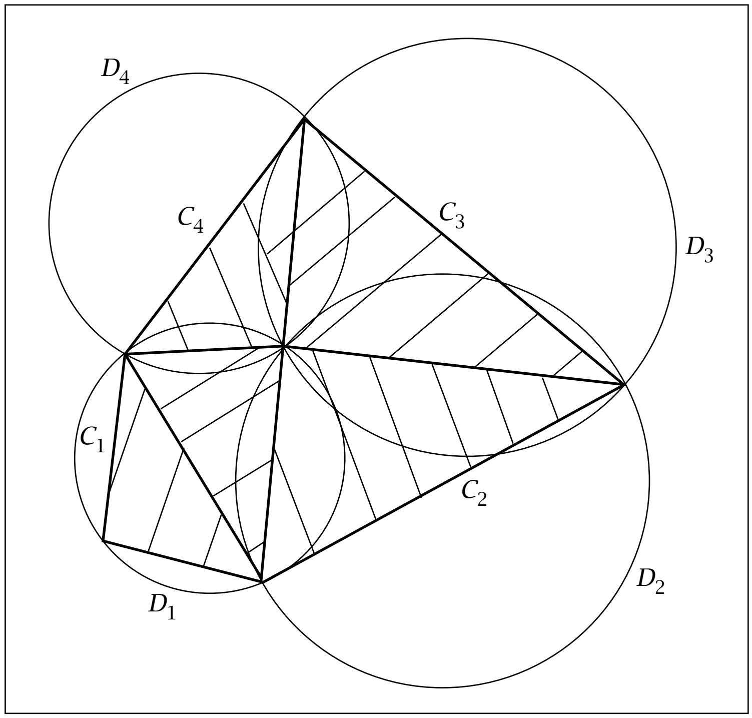
$$\arctan\sqrt{\frac{-\cos(\pi q/2)}{\cos^3(\pi q/6)}} - \epsilon. \tag{0.6}$$

Is X of hyperbolic type?

The number in (0.6) (where $\epsilon = 0$) is the circumscribed radius of the equilateral spherical triangle with angles $\pi q/3$. Theorem 3 corresponds to the case $q = 2$, and Corollary 1 can be considered as a limiting case $q \to 1$. In the case $q = 3$, the answer to this question is also positive [5].

Before we comment on the proof of Theorem 3, let us briefly indicate how this theorem can be used to prove Theorem 1. Consider the Riemann surface Z_f of a meromorphic function in the plane. Let us assume that the surface Z_f is complete. This will not be true in general, due to possible *asymptotic values* of f, but the general case can essentially be reduced to this one by invoking some techniques of quasiconformal surgery. We know that spherical disks D' covered univalently by the function correspond to schlicht disks D contained in Z_f. If Z_f is complete, then a disk $D \subset Z_f$ is schlicht if it does not contain any branch points of Z_f. In order to get a contradiction, let us assume that no schlicht disk has a radius exceeding $\arctan(\sqrt{8}) - \epsilon$, where $\epsilon > 0$. In this case one can construct a triangulation of Z_f (to be precise, it is rather a tiling of Z_f by triangles, but let us ignore this subtlety) satisfying the conditions of Theorem 3. The basic idea to obtain this triangulation is very simple: Consider all schlicht spherical disks D in Z_f containing at least three branch points in the boundary ∂D. Consider the (spherical) convex hull $C \subset \bar{D}$ of the branch points in ∂D. Then it can be shown that these convex sets C do not overlap and that they tile Z_f. If we draw diagonals in C, we obtain a triangulation of Z_f by triangles with vertices in the set of branch points. Moreover, each triangle is contained in a disk of radius $\le \arctan(\sqrt{8}) - \epsilon$, and so its circumscribed radius is bounded by the same number.

So we arrive at the set-up of Theorem 3, and we conclude that our surface should be of hyperbolic type. On the other hand, Z_f is the Riemann surface of a function meromorphic in the plane, and so it is of parabolic type. This is the desired contradiction.



The schlict disks in the surface containing at least three singular points in their boundaries can be used to construct a tiling of the surface by triangles. By the same construction, we can tile the plane by triangles with vertices in a given discrete set A which has the property that each point in the plane has uniformly bounded distance to A. Part of such a tiling is shown in the figure.

Bi-Lipschitz Maps

The idea of how to prove Theorem 3 is to reduce it to Corollary 1 by an auxiliary bi-Lipschitz map. A bijection f between two metric spaces (X, d_X) and (Y, d_Y) is called *bi-Lipschitz* if there exists a constant $L \ge 1$ (a bi-Lipschitz constant for f) such that

$$\frac{1}{L} d_X(x, y) \le d_Y(f(x), f(y)) \le L d_X(x, y) \quad \text{for } x, y \in X.$$

If there exists a bi-Lipschitz map between two metric spaces X and Y with bi-Lipschitz constant L, then the spaces are called *L-bi-Lipschitz equivalent*. If the constant L does not matter, we call the spaces *bi-Lipschitz equivalent*.

It is well known that bi-Lipschitz equivalent surfaces are of the same conformal type. Actually, there is an even larger class of type-preserving mappings, namely quasiconformal mappings, but it seems that this fact does not help in the proof of Theorem 3.

Let us now proceed to some more details: Suppose we somehow associate with each spherical triangle Δ of the triangulation T of our surface Z in Theorem 3 a corresponding Euclidean triangle $\tilde{\Delta}$. We want to glue the triangles $\tilde{\Delta}$ together according to the combinatorics of T to obtain a

Euclidean polyhedral surface $\tilde{Z}$. For the gluing to work, we need that if two triangles in the triangulation are adjacent, then the corresponding sides of the Euclidean triangles $\tilde{\Delta}$ have equal length. A way to guarantee this condition is to specify a function F so that if a, b, c are the side lengths of a spherical triangle $\Delta \in T$, then the side lengths of the corresponding Euclidean triangle are $F(a), F(b), F(c)$. In other words, the length of a side of a Euclidean triangle should depend only on the *length* of the side of the corresponding spherical triangle. We can choose F quite arbitrarily subject to some obvious conditions, e.g., we need that if a, b, c are the side lengths of a spherical triangle, then there should *exist* a Euclidean triangle with the side lengths $F(a), F(b), F(c)$. So these numbers have to satisfy the triangle inequalities. Let us denote the Euclidean triangle $\tilde{\Delta}$ associated to Δ by means of the function F by Δ_F.

A map from Z to $\tilde{Z}$ can be obtained by mapping each triangle Δ to the corresponding triangle $\tilde{\Delta} = \Delta_F$ using, say, barycentric coordinates. The length distortion of this map will be controlled by the values of the ratio $F(x)/x$. In general, the condition $F'(0) < \infty$ will be enough to ensure that the map is bi-Lipschitz.

A second condition on F comes from the following consideration. If we want to apply Corollary 1, then we have to make sure that the total angle at each vertex of the surface $\tilde{Z}$ exceeds 2π by a definite amount. Since the total angle at each vertex of Z is at least 4π, this will be true if the transition from Δ to $\tilde{\Delta}$ squeezes each angle by a factor bounded away from $1/2$. Let us state this more formally. Let α, β, γ be the angles of Δ and $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}$ be the corresponding angles of $\tilde{\Delta}$. Define the angle distortion of Δ with respect to F by

$$D(\Delta, F) := \min\{\tilde{\alpha}/\alpha, \tilde{\beta}/\beta, \tilde{\gamma}/\gamma\}.$$

Then we need a function F such that

$$D(\Delta, F) \geq 1/2 + \delta(\epsilon) > 1/2 \qquad (0.7)$$

for every spherical triangle Δ of circumscribed radius $\leq \arctan(\sqrt{8}) - \epsilon$. It is instructive to look at the extremal case of the equilateral spherical triangle Δ_0 with angles $2\pi/3$. Whatever F is, the corresponding Euclidean triangle $\tilde{\Delta}_0$ is equilateral and has angles $\pi/3$. Hence $D(\Delta_0, F) = 1/2$.

It can be shown that the function

$$F(t) = \min\left\{2k\sin(t/2), \sqrt{2\sin(t/2)}\right\}$$

satisfies $F'(0) < \infty$ and (0.7), where $k = k(\epsilon)$. The proof is elementary but quite cumbersome. Using this function F and the above construction, Theorem 3 can be reduced to Corollary 1.

One can summarize the whole proof of Theorem 3 by saying that the positive part of the curvature of the surface Z is eliminated by a carefully chosen bi-Lipschitz map so that on the target surface $\tilde{Z}$ enough negative curvature remains to ensure that $\tilde{Z}$ is uniformly hyperbolic.

Related techniques also work in other situations. U. Lang and the author have recently proved the following theorem.

Theorem 4. *Suppose X is an Aleksandrov surface which is complete, open, and simply connected. If $\mu^+(X) < 2\pi$ and $\mu^-(X) < \infty$, then X is bi-Lipschitz equivalent to $\mathbb{C}$.*

A similar statement had previously been established by J. Fu under the assumption $|\mu|(X) < \epsilon_0$, where $0 < \epsilon_0 < 2\pi$ is some numerical constant.

The constant 2π in Theorem 4 is best possible. To see this, let the surface X be a one-sided infinite cylinder, closed off by a hemisphere on its finite end. Then $\mu^+(X) = 2\pi$ and $\mu^-(X) = 0$, but it is an easy exercise to see that X is not bi-Lipschitz equivalent to $\mathbb{C}$.

The proof of Theorem 4 uses an approximation argument to reduce to the case of a Euclidean polyhedral surface. Then again we apply a "curvature elimination technique": We find appropriate pairwise disjoint flat sectors on the Euclidean polyhedral surface which go from a conical singularity to "infinity". By a bi-Lipschitz map these sectors are opened up or squeezed down according to whether the integral curvature at the singularity is positive or negative. In other words, the sectors are adjusted to compensate for the defect or excess of the total angle at the singularity with respect to 2π. In this way all the curvature on the target surface $\tilde{X}$ of the bi-Lipschitz map can be eliminated, and one obtains a flat, complete, open, and simply connected surface. Hence $\tilde{X}$ is isometric to $\mathbb{C}$, and X is bi-Lipschitz equivalent to $\mathbb{C}$.

The sectors can be chosen to give control on the bi-Lipschitz constant of the bi-Lipschitz equivalence. One obtains the stronger statement that each surface as in Theorem 4 is L-bi-Lipschitz equivalent to $\mathbb{C}$ where L depends only on $2\pi - \mu^+(X)$ and $\mu^-(X)$.

Concluding Remarks

It should have become clear from the preceding discussion that it is useful to know when two surfaces are bi-Lipschitz equivalent. At present, there are few techniques to decide this question and many basic problems remain open. For example, the fundamental question behind Theorem 4 is: *How to characterize $\mathbb{R}^2$ up to bi-Lipschitz equivalence?* This problem has additional interest, since it is equivalent to the problem of characterizing Jacobians of planar quasiconformal mappings up to multiplicative constants.

Of course, the characterization of other standard surfaces or even higher dimensional spaces up to bi-Lipschitz equivalence also merits attention; see [11] for information on this and related topics.

References

[1] L. V. AHLFORS, Sur les fonctions inverses des fonctions méromorphes, *C. R. Acad. Sci. Paris* **194** (1932), 1145–7.

[2] ____ Zur Theorie der Überlagerungsflächen, *Acta Math.* **65** (1935), 157–94.

[3] A. BLOCH, Les théorèmes de M. Valiron sur les fonctions entières, et la théorie de l'uniformisation, *C. R. Acad. Sci. Paris* **178** (1924), 2051–2.

[4] ____ La conception actuelle de la théorie des fonctions entières et méromorphes, *Enseign. Math.* **25** (1926), 83–103.

[5] M. BONK and A. EREMENKO, Covering properties of meromorphic functions, negative curvature and spherical geometry, *Ann. of Math.* (2) **152** (2000), 551–92.

[6] ____ Uniformly hyperbolic surfaces, *Indiana Univ. Math. J.* **49** (2000), 61–80.

[7] M. BONK, J. HEINONEN, and P. KOSKELA, Uniformizing Gromov hyperbolic spaces, *Astérisque* (2001), no. 270.

[8] H. CHEN, On the Bloch constant, *Approximation, Complex Analysis, and Potential Theory (Montréal, QC, 2000)* (N. Arakelian and P. M. Gauthier, eds.), Kluwer, Dordrecht, 2001, pp. 129–61.

[9] É. GHYS and P. DE LA HARPE (eds.), *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Birkhäuser, Boston, MA, 1990.

[10] YU. G. RESHETNYAK, Two-dimensional manifolds of bounded curvature, *Geometry IV, Nonregular Geometry* (Yu. G. Reshetnyak, ed.), Encylopaedia of Mathematical Sciences, vol. 70, Springer-Verlag, Berlin, 1993, pp. 3–163.

[11] S. SEMMES, On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights, *Rev. Mat. Iberoamericana* **12** (1996), 337–410.