J. London Math. Soc. (2) 69 (2004) 153–168 © 2004 London Mathematical Society DOI: 10.1112/S0024610703004836

STRUCTURE THEOREMS FOR RIEMANN AND TOPOLOGICAL SURFACES

VENANCIO ÁLVAREZ AND JOSÉ M. RODRÍGUEZ

1. Introduction

The classification theorem of compact surfaces states that every topological orientable compact surface is homeomorphic to a sphere or to a 'torus' of genus g, with g = 1, 2, ... (see for example [12]).

We say that the closure of a three-holed sphere (which is a bordered topological compact surface whose border is the union of three pairwise disjoint simple closed curves) is a *Y*-piece or a pair of pants. A Y-piece can be visualized as tubing with the shape of the letter Y.

With this definition, the classification theorem of compact surfaces states that every topological orientable compact surface except for the sphere and the torus (of genus 1) can be obtained by gluing Y-pieces along their boundaries.

In this paper we obtain as a corollary of the main theorem the generalization of this result to non-compact surfaces. We only need one simple definition.

We say that a closed subset of a topological surface is a cylinder if it is homeomorphic to $S^1 \times [0, \infty)$, where S^1 denotes the one-dimensional sphere.

THEOREM 1.1. Every topological orientable surface except for the sphere, the plane and the torus is the union (with pairwise disjoint interiors) of Y-pieces and cylinders.

REMARK 1.1. In this paper we only consider surfaces that are connected and which have a topology with a countable basis.

We also have a similar result for bordered surfaces.

The main result is a geometric version of this theorem for complete surfaces with constant negative curvature. In this case we have more information about the basic blocks of the surface: we can decompose the surface in such a way that the boundary of the blocks is the union of at most three simple closed geodesics. Since the Riemannian structure is more restrictive than the topological one, we need an additional piece to achieve the decomposition, the half-disk.

We state now the main result. We refer to the next section for definitions and background.

Received 26 November 2000; final revision 16 April 2003.

²⁰⁰⁰ Mathematics Subject Classification 30F45.

The first author's research was partially supported by a grant from MCYT (BFM2001-1736), Spain. The second author's research was partially supported by a grant from MCYT (BFM2000-0022), Spain.

THEOREM 1.2. Every hyperbolic Riemann surface except for $\mathbf{D} \setminus \{0\}$ is the union (with pairwise disjoint interiors) of funnels, half-disks and a set G which can be exhausted by geodesic domains. Furthermore, if the surface is not \mathbf{D} or an annulus, the set G appears always in the decomposition.

We will see in Proposition 3.2 that G is a union of generalized Y-pieces (see Section 2) whose boundary is a union of simple closed geodesics.

If we exclude the case of the disk, it is not clear why we need half-disks in order to decompose a surface. The necessity of half-disks is in fact the most surprising and difficult part of the proof of this theorem.

Theorem 1.2 (and in particular its corollary (see Section 4)) is a useful result in the study of Riemann surfaces. It plays an important role in the proof of the following theorem of J. L. Fernández and M. V. Melián [9].

THEOREM A. Let S be a hyperbolic surface. There are three possibilities.

(i) S has finite area. Then for every $p \in S$ there is exactly a countable collection of directions in $\mathcal{E}(p)$.

(ii) S is transient. Then for every $p \in S$, $\mathcal{E}(p)$ has full measure.

(iii) S is recurrent and of infinite area. Then $\mathcal{E}(p)$ has zero length but its Hausdorff dimension is 1.

We call a surface transient (respectively recurrent) if Brownian motion of S is transient (respectively recurrent). Also, we define $\mathcal{E}(p)$ as the set of unitary directions v in the tangent plane of S at p such that the unit speed geodesic emanating from p in the direction of v escapes to infinity.

In the applications of Theorem 1.2, it is a crucial fact that the boundaries of the blocks are simple closed geodesics. There is a clear reason for this: it is very easy to cut and paste surfaces along simple closed geodesics.

One may think that perhaps in the decomposition of Theorem 1.2 we do not need half-disks. The example after the proof of Theorem 1.2 shows that we do need them.

REMARK 1.2. Theorem 1.1 is an easy consequence of the proof of Theorem 1.2 (see Section 4). There are unpublished proofs of Theorem 1.1 that do not use any geometries (which involve the 2-dimensional Dehn lemma).

REMARK 1.3. A result similar to Theorem 1.2 can be deduced from Nielsen's work on the convex cores of Fuchsian groups. However, the usual treatment of Nielsen's convex core only deals with surfaces whose fundamental group is finitely generated (see for example [8], [17] or [5]). Our approach to the subject is based on entirely different arguments.

The outline of the paper is as follows. Section 2 presents the definitions we need. We prove some technical results in Section 3. Section 4 is dedicated to the proofs of the theorems. Last, Section 5 describes some situations in which we can guarantee the existence of half-disks in the decomposition. NOTATION 1.1. We denote by $\Re z$ and $\Im z$ respectively the real and imaginary part of the complex number z. By $A \subset B$ we mean that the set A is strictly contained in B. A simple curve is always a non-closed simple curve.

2. Definitions and results

We collect here some definitions relating to Riemann surfaces.

A hyperbolic Riemann surface S is a Riemann surface whose universal covering space is the unit disk $\mathbf{D} = \{z \in \mathbf{C} : |z| < 1\}$, endowed with its Poincaré metric (also called the hyperbolic metric), that is, the metric obtained by projecting the Poincaré metric of the unit disk

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

With this metric, S is a complete Riemannian manifold with constant curvature -1. The only Riemann surfaces which are left out are the sphere, the plane, the punctured plane and the tori. Note that every Riemann surface is connected.

We remark that this definition of a hyperbolic Riemann surface is not universally accepted, since sometimes the word *hyperbolic* refers to the existence of Green's function.

Let S be a hyperbolic Riemann surface with a puncture r (an isolated point in the boundary of S in the case $S \subset \mathbf{C}$); a puncture end in a hyperbolic Riemann surface is also usually called *cusp*. A collar in S about r is a doubly connected domain in S 'bounded' by r and a Jordan curve (called the boundary curve of the collar) orthogonal to the pencil of geodesics emanating from r. It is well known that the length of the boundary curve is equal to the area of the collar (see for example [3]).

A collar in S about r of area α will be called an α -collar and it will be denoted by $C(r, \alpha)$. A theorem of Shimizu [15] gives that for every puncture in any hyperbolic Riemann surface, there exists an α -collar for every $0 < \alpha \leq 1$ (see also [11, pp. 60–61]). We also have the following result (see for example [3]).

LEMMA A. Let S be a hyperbolic Riemann surface with a puncture r. Then we have

$$C(r,1) \cap \gamma = \emptyset,$$

for any simple closed geodesic γ in S.

We say that a curve is homotopic to a puncture r if it is freely homotopic to $\partial C(r, \alpha)$ for some (and then for every) $0 < \alpha < 1$.

We say that S is a bordered hyperbolic Riemann surface if it is a bordered orientable Riemannian manifold of dimension 2 and its Riemannian metric has constant negative curvature -1; S must also be a complete metric space.

A half-disk is a bordered hyperbolic Riemann surface which is topologically a half-plane and the border of which is a non-closed simple geodesic. Every half-disk is conformally equivalent to the subset $\{z \in \mathbf{D} : \Re z \ge 0\}$ of the hyperbolic disk \mathbf{D} .

A funnel is a bordered hyperbolic Riemann surface which is topologically a cylinder and the border of which is a simple closed geodesic. Given a positive number a, there is a unique (up to conformal mapping) funnel such that its boundary curve has length a. Every funnel is conformally equivalent, for some $\beta > 1$, to the subset $\{z \in \mathbf{C} : 1 \leq |z| < \beta\}$ of the hyperbolic Riemann surface $\{z \in \mathbf{C} : 1/\beta < |z| < \beta\}$.

Every doubly connected end of a hyperbolic Riemann surface is a puncture (if there are homotopically non-trivial curves with arbitrary small length) or a funnel (otherwise).

An open connected set is called a *domain*. A geodesic domain in a Riemann surface S is a domain $G \subset S$ (which is not simply or doubly connected) that has finite area and is such that ∂G consists of finitely many simple closed geodesics. G does not have to be relatively compact since it may contain finitely many cusps. We can think of a puncture as a boundary geodesic of zero length. Recall that if γ is a simple closed curve in S, then there is a unique simple closed geodesic of minimal length in its free homotopy class, unless γ is homotopic to a point or to a puncture; in these cases it is not possible to find such a geodesic because there are curves in the homotopy class with arbitrary small length.

Geodesic domains play an important role in the study of the hyperbolic isoperimetric inequality of a Riemann surface. We say that a Riemann surface Ssatisfies a hyperbolic isoperimetric inequality if there is a positive constant h such that

$$A(D) \leqslant h \, L(\partial D) \tag{2.1}$$

holds for every relatively compact domain $D \subset S$ with smooth boundary, where A(D) and $L(\partial D)$ denote, respectively, the hyperbolic area of D and the hyperbolic length of ∂D in S. We denote by h(S) the infimum of the constants h verifying (2.1). In [10, Lemma 1.2], it was proved that if S verifies (2.1) for geodesic domains, then it satisfies a hyperbolic isoperimetric inequality. In fact, if $h_g(S)$ is the infimum of the constants h such that the inequality (2.1) is true for any geodesic domain, then we have

$$h(S) \leqslant h_g(S) + 2$$

There are interesting relations of the hyperbolic isoperimetric inequality with other conformal invariants of a Riemann surface (see for example [2; 6, p. 95; 7; 10; 13, p. 145; 16, p. 333]).

A Löbell Y-piece is a compact bordered hyperbolic Riemann surface which is topologically a Y-piece and the boundary curves of which are simple closed geodesics. Given three positive numbers a, b, c, there is a unique (up to conformal mapping) Löbell Y-piece such that its boundary curves have lengths a, b, c (see for example [14, p. 410]). They are a standard tool for constructing Riemann surfaces. A clear description of these Y-pieces and their use is given in [4, Chapter 1; 6, Chapter X.3].

A generalized Löbell Y-piece is a bordered or non-bordered hyperbolic Riemann surface which is topologically a sphere without n open disks and m points, with integers $n, m \ge 0$ and n + m = 3, so that the n boundary curves are simple closed geodesics and the m deleted points are punctures. Observe that a generalized Löbell Y-piece is topologically the union of a Y-piece and m cylinders, with $0 \le m \le 3$ (see Figure 1).



FIGURE 1.

It is clear that the interior of every generalized Löbell Y-piece is a geodesic domain. Furthermore, it is known that the closure of every geodesic domain is a finite union (with pairwise disjoint interiors) of generalized Löbell Y-pieces (see Proposition 3.2).

We say that the set A is exhausted by $\{A_n\}$ if $A_n \subseteq A_{n+1}$ for every n and $A = \bigcup_n A_n$.

We say that a bordered topological surface S is simple if the border of S is a (finite or infinite) union of pairwise disjoint simple closed curves. We have the following result.

THEOREM 2.1. Every simple bordered topological orientable surface except for the bordered disk and the cylinder with two boundary curves is the union (with pairwise disjoint interiors) of Y-pieces and cylinders.

We have a similar result for bordered hyperbolic Riemann surfaces. A bordered hyperbolic Riemann surface S is simple if the border of S is a (finite or infinite) union of pairwise disjoint simple closed geodesics. The closure of any geodesic domain is a simple bordered hyperbolic Riemann surface.

THEOREM 2.2. Every simple bordered hyperbolic Riemann surface is the union (with pairwise disjoint interiors) of funnels, half-disks and a set V which can be exhausted by the closures of geodesic domains.

REMARK 2.1. The proof of Theorem 1.2 gives the following recipe for constructing hyperbolic Riemman surfaces.

Join funnels and/or generalized Y-pieces by identifying simple closed geodesics of equal length obtaining a surface without border S_0 . If S_0 is complete, then it is a hyperbolic Riemann surface. If S_0 is not complete, then we can obtain a hyperbolic Riemann surface S by gluing to the metric completion of S_0 half-disks (half-disks are the unique blocks we can add to S_0 in order to obtain a hyperbolic Riemann surface).

Theorem 1.2 states that this method allows one to construct any hyperbolic Riemann surface except for $\mathbf{D} \setminus \{0\}$.

3. Geodesics and geodesic domains

In this section we include some technical results about geodesics and geodesic domains that we need in the proofs of the theorems.

LEMMA B [14, p. 405]. Let α and β be two disjoint simple closed curves not freely homotopic in the hyperbolic Riemann surface S. If α_0 and β_0 are respectively simple closed geodesics in the homotopy classes of α and β , then α_0 and β_0 are also disjoint.

A similar result is true if α is a non-closed simple geodesic.

LEMMA 3.1. Let S be a hyperbolic Riemann surface, γ_1 be a non-closed simple geodesic in S, σ be a simple closed curve in S and γ_2 be a simple closed geodesic freely homotopic to σ in S. If γ_1 and σ are disjoint, then γ_1 and γ_2 are also disjoint.

Proof. The proof follows the arguments in the proof of Lemma B. We include the details for the sake of completeness.

Let us consider a universal covering map $\pi : \mathbf{D} \longrightarrow S$. Without loss of generality, we can assume that π applies the interval (-1, 1) onto γ_1 , that is, that $\tilde{\gamma}_1 = (-1, 1)$ is a lifting of γ_1 . Let us consider liftings $\tilde{\sigma}$, $\tilde{\gamma}_2$ of σ , γ_2 , respectively, that have the same endpoints $A, B \in \partial \mathbf{D}$.

We only need to check that $\tilde{\gamma}_1 \cap \tilde{\gamma}_2 = \emptyset$. Assume that this is not the case. Then $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ intersect at a single point in **D**, since otherwise they would be equal, which is a contradiction. Consequently $\Im A \cdot \Im B < 0$. This implies that $\tilde{\gamma}_1 \cap \tilde{\sigma} \neq \emptyset$, since $\tilde{\sigma}$ also has the same endpoints $A, B \in \partial \mathbf{D}$, but this contradicts $\pi(\tilde{\gamma}_1) \cap \pi(\tilde{\sigma}) = \gamma_1 \cap \sigma = \emptyset$. This finishes the proof of Lemma 3.1.

It is well known that if a simple closed curve σ is not homotopic to a point or to a puncture in a hyperbolic Riemann surface S, then there is a unique simple closed geodesic freely homotopic to σ in S. The following result is not surprising and it is probably known. However we do not know any reference for it.

PROPOSITION 3.1. Let S be a bordered hyperbolic Riemann surface such that the border of S is a pairwise disjoint union of non-closed simple geodesics and/or simple closed geodesics. If a simple closed curve σ is not homotopic to a point or to a puncture in S, then there is a unique simple closed geodesic γ freely homotopic to σ in S. Furthermore, γ is contained in the interior of S if σ is not freely homotopic to a boundary geodesic in ∂S .

Proof. If σ is freely homotopic to a boundary geodesic in ∂S , the result is trivial. Otherwise, without loss of generality, we can assume that σ is contained in the interior of S, since if it is not the case, then we can take a simple closed curve σ_0 contained in the interior of S and freely homotopic to σ in S. Let us consider the Schottky double S_0 of S. Roughly speaking, S_0 is the union of S and S^* , the symmetric surface of S, identifying the symmetric points in ∂S and ∂S^* (see [1, p. 119] for details). We have σ is not homotopic to a point or to a puncture in S_0 ; then there is a unique simple closed geodesic γ freely homotopic to σ in S_0 . Lemmas B and 3.1 give that γ does not intersect the border of S, since we are assuming that σ is not freely homotopic to a boundary geodesic in ∂S . Consequently γ is contained in the interior of S. The following result is well known, although we do not know any reference for it. We include the proof for the sake of completeness.

PROPOSITION 3.2. The Riemann surface obtained by taking the closure of a geodesic domain is a finite union (with pairwise disjoint interiors) of generalized Löbell Y-pieces.

REMARK 3.1. This Riemann surface can be bordered or not. The fact that a compact Riemann surface is a finite union (with pairwise disjoint interiors) of Löbell Y-pieces can be found in many books (see for example [4, p. 94]). The heart of the proof of Proposition 3.2 is to associate, in an appropriate way, a compact Riemann surface to each geodesic domain.

Proof of Proposition 3.2. Let G be a geodesic domain and let us denote by S the closure of G. We have $\partial S = \partial G$ is the union of pairwise disjoint simple closed geodesics $\gamma_1, \ldots, \gamma_k$ (we take k = 0 if $\partial G = \emptyset$). For each $j = 1, \ldots, k$, let us consider a Y-piece Y_j with $\partial Y_j = \sigma_j^1 \cup \sigma_j^2 \cup \sigma_j^3$, $L(\sigma_j^1) = L(\gamma_j)$ and $L(\sigma_j^2) = L(\sigma_j^3)$. Let us denote by Z_j the bordered surface Y_j with σ_j^2 and σ_j^3 identified; then $\partial Z_j = \sigma_j^1$. Let us consider the non-bordered Riemann surface $S_1 := \bigcup_j Z_j \cup S$ obtained by gluing (identifying) for each $j = 1, \ldots, k$ the curves γ_j and σ_j^1 . If r_1, \ldots, r_m are the punctures in S_1 , we consider

$$S_2 := S_1 \setminus \bigcup_{i=1}^m \overline{C(r_i, 1/3)}.$$

Recall that Lemma A gives that every simple closed geodesic of S is contained in S_2 . There is a homeomorphism $f: S_1 \longrightarrow S_2$. We can take as f the identity map in $S_1 \setminus \bigcup_{i=1}^m C(r_i, 1/2)$ and for each $i = 1, \ldots, m$ any homeomorphism between $C(r_i, 1/2)$ and $C(r_i, 1/2) \setminus \overline{C(r_i, 1/3)}$ which fixes $\partial C(r_i, 1/2)$.

If we consider now the Riemann surface S_2 with its hyperbolic metric (for which it is a complete metric space), let us take the simple closed geodesics η_1, \ldots, η_m in S_2 such that each η_i is homotopic to r_i in S_1 for $i = 1, \ldots, m$, and the funnels F_i bounded by η_i . For each $i = 1, \ldots, m$, let us consider a Y-piece Y^i with $\partial Y^i = \tau_1^i \cup \tau_2^i \cup \tau_3^i$, $L(\tau_1^i) = L(\eta_i)$ and $L(\tau_2^i) = L(\tau_3^i)$. Let us denote by Z^i the bordered surface Y^i with τ_2^i and τ_3^i identified; we have $\partial Z^i = \tau_1^i$. Let us consider the hyperbolic compact Riemann surface $S_3 := (S_2 \setminus \bigcup_i F_i) \cup (\bigcup_i Z^i)$ obtained by gluing for each $i = 1, \ldots, m$ the curves η_i and τ_1^i . If η_{m+j} is the simple closed geodesic in S_2 (or in S_3) freely homotopic to γ_j for each $j = 1, \ldots, k$, then it is well known (see for example [4, p. 94]) that $m + k \leq 3g - 3$, where g is the genus of S_3 , and that there exist simple closed geodesics $\eta_{m+k+1}, \ldots, \eta_{3g-3}$ such that $\eta_1, \ldots, \eta_m, \eta_{m+1}, \ldots, \eta_{m+k}, \eta_{m+k+1}, \ldots, \eta_{3g-3}$ decompose S_3 into Löbell Y-pieces.

It is not difficult to see that we can 'pullback' this decomposition to S. The punctures r_1, \ldots, r_m correspond to the geodesics η_1, \ldots, η_m , and this shows the necessity of considering generalized Löbell Y-pieces instead of Löbell Y-pieces.

4. Proof of the results

Let us start with the proof of Theorem 1.2.

THEOREM 1.2. Every hyperbolic Riemann surface except for $\mathbf{D} \setminus \{0\}$ is the union (with pairwise disjoint interiors) of funnels, half-disks and a set G which can be exhausted by geodesic domains. Furthermore, if the surface is not \mathbf{D} or an annulus, the set G appears always in the decomposition.

Proof. Let us consider such a surface S. If S is simply connected, then $S = \mathbf{D}$, which is a union of two half-disks. If S is doubly connected, then it is an annulus (since $S \neq \mathbf{D} \setminus \{0\}$), which is a union of two funnels.

If S is of connectivity greater than 2, then there is at least one geodesic domain. In this case let us consider a fixed point $p \in S$ and any positive number t, and let us denote by B(t) the open ball in S with center p and radius t. The boundary of B(t) is a finite union of pairwise disjoint simple closed curves η_1, \ldots, η_k except for $t \in A$ with A a numerable set. In the following we only consider values of $t \notin A$. For $i = 1, \ldots, k$, we denote by γ_i the empty set if η_i is homotopic to a point or to a puncture, and the simple closed geodesic freely homotopic to η_i otherwise. Observe that $\gamma_1, \ldots, \gamma_k$ are pairwise disjoint by Lemma B, since η_1, \ldots, η_k are pairwise disjoint. We denote by G(t) the geodesic domain bounded by $\gamma_1, \ldots, \gamma_k$ 'corresponding' to B(t). There is a positive t_0 such that $G(t) = \emptyset$ if $t < t_0$ and $G(t) \neq \emptyset$ if $t > t_0$. We also have $G(t) \subseteq G(t')$ if $t \leq t'$.

In the following we need some results which appear in the following lemmas.

LEMMA 4.1. If there exists $t_1 > 0$ such that γ is a simple closed geodesic contained in $\partial G(t)$ for every $t \ge t_1$, then γ is the boundary of a funnel in S.

Proof. For $t \ge t_1$, let us consider the simple closed curve $\eta_t \subseteq \partial B(t)$ freely homotopic to γ . If $t_2 := \max\{\text{dist}(p, z) | z \in \gamma\}$ and $t > t_3 := \max\{t_1, t_2\}$, then we have $\gamma \cap \eta_t = \emptyset$. Let us denote by F_t the doubly connected closed set bounded by γ and η_t for $t > t_3$ and $F := \bigcup_{t > t_3} F_t$. We have F is a doubly connected end in Sbounded by the simple closed geodesic γ ; therefore F is a funnel in S.

LEMMA 4.2. (a) If γ is a simple closed geodesic contained in B(t), then γ is contained in the closure of G(t).

(b) If r is a puncture and $\partial C(r, \alpha)$ is contained in B(t) for some $0 < \alpha < 1$, then C(r, 1) is contained in G(t).

Proof. Assume that γ is not contained in the closure of G(t).

If $\gamma \cap G(t) = \emptyset$, let us consider U_i the open subset of $B(t) \setminus G(t)$ bounded by η_i and γ_i if $\gamma_i \neq \emptyset$. We have $\gamma \subset U_j$ for some j since the $\{U_i\}$ are pairwise disjoint. Since γ is not homotopically trivial we have U_j is doubly connected and $\gamma = \gamma_j \subset \overline{G(t)}$. This is a contradiction to $\gamma \cap \overline{G(t)} = \emptyset$.

Therefore $\gamma \cap G(t) \neq \emptyset$. Consequently, $\gamma \cap \partial G(t) \neq \emptyset$ since we are assuming that γ is not contained in the closure of G(t). Then we have $\gamma \cap \partial B(t) \neq \emptyset$ by Lemma B. This is a contradiction to $\gamma \subset B(t)$.

We now prove the second statement of the lemma. There is a simple closed curve $\eta_i \subseteq \partial B(t)$ contained in $C(r, \alpha)$, since $\partial C(r, \alpha) \subset B(t)$. We also have that η_i is homotopic to r. Then Lemma A gives that C(r, 1) is contained in G(t) since no simple closed geodesic can intersect C(r, 1).

We now continue the proof of Theorem 1.2.

Assume that there is t_1 such that $G(t) = G(t_1)$ for every $t \ge t_1$. If $\partial G(t_1) = \emptyset$, then $G(t_1) = S$ and the proof is finished with $G = G(t_1)$. Otherwise, the boundary of $G(t_1)$ is a finite union of simple closed geodesics $\gamma_1, \ldots, \gamma_n$. By Lemma 4.1, we have each γ_i is the boundary of a funnel in S, and this finishes the proof in this case taking also $G = G(t_1)$.

Assume now that there is an increasing sequence $\{t_n\}$ with limit ∞ such that $G_n := G(t_n)$ verifies $G_n \subset G_{n+1}$. By Lemmas 4.1 and 4.2, we can assume without loss of generality that $\partial G_n \cap \partial G_{n+1}$ is the empty set or a union of simple closed geodesics each of which is the boundary of a funnel in S.

REMARK 4.1. If γ_n, γ_{n+m} $(m, n \in \mathbf{Z}^+)$ are simple closed geodesics contained, respectively, in ∂G_n , ∂G_{n+m} , and η is a curve connecting γ_n with γ_{n+m} , then the closed curve $\beta := \eta + \gamma_{n+m} - \eta + \gamma_n$ cannot be homotopically trivial, since in this case γ_{n+m} would be freely homotopic to γ_n and this would imply $\gamma_{n+m} = \gamma_n$ and consequently m = 0. It can be homotopic to a puncture if m = 1. However, β cannot be homotopic to a puncture if $m \ge 2$, since there is at least one 'topological obstacle' between γ_n and γ_{n+1} .

Let us consider now the open set H_n obtained as the union of G_n and the funnels bounded by a curve in ∂G_n . Observe that we always have $\overline{H_n} \subset H_{n+1}$. Let us define $H := \bigcup_n H_n$ and $d_n := \text{dist}(p, \partial H_n)$. Observe that H is an open set. If $d_n \to \infty$ as $n \to \infty$, then S = H and we have finished the proof in this case with $G = \bigcup_n G_n$. If d_n is bounded, then $S \setminus H$ is a closed non-empty set.

We will finish the proof by showing that each connected component of $S \setminus H$ is a half-disk. First we will show that if $q \in \partial H$ and U is any simply connected neighbourhood of q, then $q \in \sigma \subseteq \partial H$ where $\sigma \cap U$ is a geodesic arc.

If $q \in \partial H$, then there are $q_n \in \gamma_n$ converging to q with γ_n a simple closed geodesic contained in ∂H_n . We want to see that the sequence of geodesics $\{\gamma_n\}$ converges to a geodesic arc σ in U. To see this it is enough to consider the lifting of U to the universal covering space **D**; the statement in **D** is now trivial, since the $\{\gamma_n\}$ are pairwise disjoint.

We will see that $\sigma \cap U$ is contained in ∂H . Let us consider any point $q' \in \sigma \cap U$. We have $q' \notin \operatorname{ext} H$, since it is the limit of points in $\{H_n\}$. Then, in order to see that $\sigma \cap U$ is contained in ∂H , it is enough to see that $q' \notin H$. Assume that $q' \in H$; then q' belongs to a neighbourhood $V \subset H_{n_0}$ for some n_0 . Consequently, $V \subset H_n$ for every $n \geq n_0$, but $V \cap \partial H_n \neq \emptyset$ for $n \geq n_1$, since q' is the limit of points in $\bigcup_n \partial H_n$, which is a contradiction.

We also have that if σ is a geodesic such that $\sigma \cap W \subseteq \partial H$ for some non-empty open set W, then $\sigma \subseteq \partial H$ (recall that $\sigma \cap U \subseteq \partial H$ for every simply connected open set U with $U \cap W \neq \emptyset$). We will prove that such a geodesic σ is a non-closed simple curve. Otherwise, we have σ is a simple closed geodesic or it autointersects nontangentially. If σ is a simple closed geodesic, then, by compactness, it is the limit of simple closed geodesics $\gamma_n \subset \partial H_n$; then γ_n is freely homotopic to σ for $n \ge n_2$, which is a contradiction, since in each free homotopy class there is at most one simple closed geodesic. If σ autointersects non-tangentially, then ∂H_n autointersects for $n \ge n_3$, which is another contradiction. This last argument also proves that if $\sigma_1, \sigma_2 \subset \partial H$ are geodesics with $\sigma_1 \neq \sigma_2$, then they are simple and disjoint. LEMMA 4.3. If σ is a non-closed simple geodesic contained in ∂H , $q \in \sigma$, and U is a simply connected neighbourhood of q such that $U \setminus \sigma$ has exactly two connected components U_1 and U_2 , then there is $i \in \{1, 2\}$ with $U_i \cap \overline{H} = \emptyset$.

Proof. Without loss of generality, we can assume that $\partial H_n \cap U_1 \neq \emptyset$ if $n \ge n_1$. We now prove that $U_2 \cap \overline{H} = \emptyset$. Assume that this is not true; then there is a point $h \in H \cap U_2$. In fact, there is $n_2 \ge n_1$ such that $h \in H_{n_2} \cap U_2$. This fact implies that $\partial H_{n_2} \cap U_2 \neq \emptyset$, since otherwise $\sigma \cap U$ would be contained in ∂H_{n_2} ; this is not possible, since the boundary of H_{n_2} can only contain simple closed geodesics and σ is non-closed.

Let us consider $n_3 \ge n_2 + 2$ and a compact curve $\eta \subset U$, starting in U_2 and finishing in U_1 , which connects a closed simple geodesic $\gamma_{n_2} \subset \partial H_{n_2}$ with a simple closed geodesic $\gamma_{n_3} \subset \partial H_{n_3}$. The closed curve $\beta := \eta + \gamma_{n_3} - \eta + \gamma_{n_2}$ is not homotopic to a point or to a puncture in S, since $n_3 \ge n_2 + 2$ (recall the remark about β before). Then there is a simple closed geodesic α freely homotopic to β . Furthermore, $\alpha \cap \sigma \ne \emptyset$, since $\beta \cap \sigma \ne \emptyset$, $\gamma_{n_2}, \gamma_{n_3}$ are not homotopic to a point, and σ is a simple (infinite) geodesic. Lemma 4.2 gives that $\alpha \subset H$, which is a contradiction to $\alpha \cap \partial H \ne \emptyset$. This finishes the proof of Lemma 4.3.

In particular, we have $S \setminus H$ is the closure of its interior, and then each connected component of $S \setminus H$ is path-connected.

We have also seen that ∂H is a union of pairwise disjoint non-closed simple geodesics. We now show that each connected component J of $S \setminus H$ is a half-disk. Firstly we prove that J is simply connected.

We know that each connected component J of $S \setminus H$ is the closure of its interior and its boundary is the union of pairwise disjoint non-closed simple geodesics. Assume that there is a simple closed curve $\delta \subset J$ which is not homotopic to a point in J. If δ is homotopic to a puncture r, then $C(r, \alpha_1)$ is contained in J for some $0 < \alpha_1 < 1$. Lemma 4.2 gives that $C(r, 1) \subset H$ and then $C(r, \alpha_1) \subseteq J \cap H$, which is a contradiction. If δ is not homotopic to a puncture, then Proposition 3.1 gives that there is a simple closed geodesic γ freely homotopic to δ in J, since J is a bordered hyperbolic Riemann surface such that its border is a pairwise disjoint union of nonclosed simple geodesics. Lemma 4.2 gives that $\gamma \subset H$ and then $\gamma \subseteq J \cap H$, which is a contradiction.

Therefore, in order to see that J is a half-disk, it is enough to see that its boundary is a single non-closed simple geodesic. Assume that ∂J contains two nonclosed simple geodesics σ_1, σ_2 . Let us consider $q_1 \in \sigma_1, q_2 \in \sigma_2$, simply connected neighbourhoods $q_1 \in V_1, q_2 \in V_2$, simple closed geodesics $\gamma_{n_1} \subset \partial H_{n_1}, \gamma_{n_2} \subset \partial H_{n_2}$ with $\gamma_{n_1} \cap V_1 \neq \emptyset, \gamma_{n_2} \cap V_2 \neq \emptyset$ and $n_1 \ge n_2 + 2$, and curves $\eta_1 \subset V_1, \eta_2 \subset V_2$ joining, respectively, γ_{n_1} with q_1 and q_2 with γ_{n_2} . Since J is path-connected, we can take a curve $\eta_3 \subset J$ joining q_1 with q_2 and consider $\eta := \eta_1 + \eta_3 + \eta_2$ and $\beta := \eta + \gamma_{n_2} - \eta + \gamma_{n_1}$.

As in the proof of Lemma 4.3, we can see that there is a simple closed geodesic α freely homotopic to β in S with $\alpha \cap \sigma_1 \neq \emptyset$ and $\alpha \cap \sigma_2 \neq \emptyset$. Then Lemma 4.2 gives $\alpha \subset H$, which is a contradiction.

We choose the set G as H minus the funnels in S, that is, $G = \bigcup_n G_n$. Furthermore, we have obtained that if S is not **D** or an annulus, then we have $G = \bigcup_{t>0} G(t) \neq \emptyset$. The following particular case of Theorem 1.2 is used in [9].

COROLLARY 4.1. If a hyperbolic Riemann surface does not contain any halfdisk, then it can be exhausted by geodesic domains.

Proof. It is enough to remark that both $\mathbf{D} \setminus \{0\}$ and any funnel contain a half-disk.

One may think that in the decomposition of Theorem 1.2 we perhaps do not need half-disks. The following example shows that we need them.

EXAMPLE 4.1. Let $\{x_n\}_{n \ge 1}$ be any increasing sequence converging to 1, contained in the interval (0, 1). Let us consider $S := \mathbf{D} \setminus X$ with $X := \bigcup_{n \ge 1} \{x_n\}$ and γ_n the simple closed geodesic in S which surrounds the points x_1, \ldots, x_n , for n > 1. The curve γ_n is the boundary of a geodesic domain G_n . It is not difficult to see that $\{\gamma_n\}$ 'converges' to a non-closed simple geodesic γ in S and that γ is the boundary curve of a half-disk. If we consider other geodesic domains, then we also need a half-disk, since a non-closed simple closed geodesic cannot intersect a half-disk.

THEOREM 2.2. Every simple bordered hyperbolic Riemann surface is the union (with pairwise disjoint interiors) of funnels, half-disks and a set V which can be exhausted by the closures of geodesic domains.

Proof. Let us consider $\{\gamma_j\}$ the pairwise disjoint simple closed geodesics in the border of the simple bordered hyperbolic Riemann surface S. Observe that we can construct a hyperbolic Riemann surface S_0 (without border) by gluing to S a funnel F_j in each γ_j , with $L(\partial F_j) = L(\gamma_j)$.

Since S_0 cannot be $\mathbf{D} \setminus \{0\}$ (there are no simple closed geodesics in $\mathbf{D} \setminus \{0\}$), by Theorem 1.2 it is the union (with pairwise disjoint interiors) of funnels, half-disks and a set G which can be exhausted by geodesic domains. We obtain the desired result by deleting the funnels $\{F_j\}$ in this union. When we delete the funnels, we are also deleting the curves $\{\gamma_j\}$, the border of S; this is the reason why in this situation we consider the set $V = \overline{G}$, which is exhausted by the closures of geodesic domains.

THEOREM 1.1. Every topological orientable surface except for the sphere, the plane and the torus is the union (with pairwise disjoint interiors) of Y-pieces and cylinders.

Proof. It is well known that every topological surface S has a C^{∞} structure compatible with its topological structure. The isothermal coordinates give to S a conformal structure compatible with its C^{∞} structure; if, furthermore, S is orientable, this conformal structure is also a structure of a Riemann surface.

Then S is conformally equivalent to the sphere, the complex plane \mathbf{C} , $\mathbf{C} \setminus \{0\}$, a torus or a hyperbolic Riemann surface. We do not consider the first, second and fourth cases, since they are excluded in the statement of the theorem.

If S is conformally equivalent to $\mathbf{C} \setminus \{0\}$, it is the union of the two cylinders $\{z \in \mathbf{C} : 0 < |z| \leq 1\}$ and $\{z \in \mathbf{C} : |z| \geq 1\}$. If S is conformally equivalent to $\mathbf{D} \setminus \{0\}$, it is the union of the two cylinders $\{z \in \mathbf{C} : 0 < |z| \leq 1/2\}$ and



 $\{z \in \mathbf{C} : 1/2 \leq |z| < 1\}$. If S is conformally equivalent to another hyperbolic Riemann surface, we can apply Theorem 1.2.

Recall that a funnel is a cylinder, a Löbell Y-piece is a Y-piece, and a generalized Löbell Y-piece is the union of a Y-piece with at most three cylinders. Proposition 3.2 now gives the result if we allow half-disks in the decomposition. In order to remove the half-disks, we modify some Y-pieces as follows.

We use the construction and notations in the proof of Theorem 1.2. Consider a non-closed simple geodesic $\sigma \subseteq \partial H$. Recall that $H = \bigcup_n H_n$, and there exists a sequence of simple closed geodesics $\gamma_n \subseteq \partial H_n$ 'converging' to σ . Assume first that $S \setminus H$ is connected.

If $q \in \sigma$, define η_q as the geodesic perpendicular to σ in q, with $\eta_q(0) = q$ and $\|\eta'_q(t)\| = 1$. For fixed $p \in \sigma$ and any $m \in \mathbf{N}$, let us consider $\sigma^m = \sigma \cap B_S(p, m)$. Our arguments in the proof of Theorem 1.2 give that if $\sigma_{\varepsilon}^m := \{z \in H : z \in \eta_q(t) \text{ with } |t| < \varepsilon$ and $q \in \sigma^m\}$, then $\gamma_n \cap \sigma_{\varepsilon}^m$ converges uniformly to σ^m for any $\varepsilon > 0$. Since σ^m is relatively compact, we can choose $\varepsilon_m > 0$ such that $\sigma_{\varepsilon_m}^m$ is simply connected. If $b_m = \{z \in \partial \sigma_{\varepsilon_m}^m : d_S(z, \sigma) = \varepsilon_m\}$, choose recursively n_m as a natural number greater than n_{m-1} and such that $\gamma_{n_m} \cap \sigma_{\varepsilon_m}^m \neq \emptyset$, $\gamma_{n_m} \cap b_m = \emptyset$, and $n_0 = 0$.

Define $r_m := n_{m+1} - n_m - 1$. Assume that $r_m = 0$ for every m, that is, $n_m = m$. Define $A_m := \sigma_{\varepsilon_m}^m \cup (B_S(p,m) \cap (S \setminus H))$ and \widetilde{H}_m as the relatively compact domain $\widetilde{H}_m := H_m \cup A_m$. Observe that \widetilde{H}_m is homeomorphic to H_m , and then $\widetilde{H}_{m+1} \setminus \widetilde{H}_m$ have a similar decomposition in Y-pieces and cylinders to $H_{m+1} \setminus H_m$. It is clear that $S = \bigcup_m \widetilde{H}_m$, since $S \setminus H$ is connected. This finishes the proof in this case (see Figure 2).

If we have $r_m > 0$ for some m, we only need to choose $A_{n_m}, A_{n_m+1}, \ldots, A_{n_m+r_m}, A_{n_{m+1}}$ in a similar way with the condition $\overline{A}_j \subset A_{j+1}$ (see Figure 3).

Finally, if $S \setminus H$ is not connected, we repeat this construction in each connected component of $S \setminus H$.

THEOREM 2.1. Every simple bordered topological orientable surface except for the bordered disk and the cylinder with two boundary curves is the union (with pairwise disjoint interiors) of Y-pieces and cylinders.

Proof. Let us consider such a surface S and the topological double S_1 of S. Roughly speaking, S_1 is the union of S and S^* , the symmetric surface of S,



FIGURE 3.

identifying the symmetric points in ∂S and ∂S^* . As in the proof of Theorem 1.1 there is a structure of a Riemann surface for S_1 compatible with its topological structure. Furthermore, the holomorphic atlas of S_1 can be taken as symmetric with respect to ∂S .

Assume that S_1 is not a hyperbolic surface, that is, S_1 is the sphere, the plane, the punctured plane or a torus. S_1 cannot be the sphere or a torus since S is neither the bordered disk nor the cylinder with two boundary curves. S_1 is not the plane, since the plane is not the double of a simple bordered topological orientable surface. Then S_1 is the punctured plane, which is the union of the two cylinders S and S^* . Consequently, S is a cylinder and Theorem 2.1 holds in this case.

Assume now that S_1 is a hyperbolic surface. The simple closed curves $\{\gamma_j\}$ in ∂S are geodesics in the hyperbolic metric of S_1 , since the holomorphic atlas is symmetric with respect to ∂S . We have S is a simple bordered hyperbolic Riemann surface with the restriction of the hyperbolic metric of S_1 .

If we recall that a funnel is a cylinder, a Löbell Y-piece is a Y-piece, and a generalized Löbell Y-piece is the union of a Y-piece with at most three cylinders, then Theorem 2.2 and Proposition 3.2 give the result if we allow half-disks in the decomposition. We can remove the half-disks as in the proof of Theorem 1.1. \Box

5. Maximal half-disks

In this section we study and classify the different types of half-disk contained in a Riemann surface S, and determine whether they appear in our decomposition. First we see that half-disks are always contained in some of the basic pieces of our decomposition.

PROPOSITION 5.1. Let S be a hyperbolic Riemann surface not conformally equivalent to **D**. Suppose that S is the union (with disjoint interiors) of a set G exhausted by geodesic domains and a collection of pieces $\{P_{\alpha}\}_{\alpha \in A}$, where $P_{\alpha} \subset S$ is either a funnel or a half-disk. If $P \subset S$ is a half-disk, then there exists an index $\alpha \in A$ such that $P \subseteq P_{\alpha}$.

Proof. First, we see that P cannot intersect any simple closed geodesic. The boundary of P is a non-closed simple geodesic σ . P cannot contain a simple closed geodesic, since it is a simply connected domain. Suppose that σ intersects a simple

closed geodesic γ . If the intersection of both curves is a single point, then they are tangent, which is impossible. If they intersected in, at least, two points, then we would have a simply connected domain $D \subset P$ limited by an arc of σ and an arc of γ , which is also impossible.

If $P \cap G \neq \emptyset$, then P intersects a generalized Löbell Y-piece, and then contains a simple closed geodesic or a puncture; therefore P is not simply connected, which is a contradiction.

Therefore, there exists a set P_{α} such that $P_{\alpha} \cap P \neq \emptyset$. If P_{α} is a half-disk, then $P \subseteq P_{\alpha}$, because otherwise P would also intersect a simple closed geodesic (recall that there are simple closed geodesics arbitrarily close to any point of ∂P_{α} , since S is not conformally equivalent to **D**). If P_{α} is a funnel, then $P \subset P_{\alpha}$, because funnels are limited by a simple closed geodesic.

In any case, we conclude that $P \subseteq P_{\alpha}$, and this finishes the proof of Proposition 5.1.

Next, we prove a corollary relating the maximal half-disks of a surface and the half-disks appearing in decompositions.

DEFINITION 5.1. We say that a half-disk $P \subset S$ is canonical if there is a decomposition of the surface S in a union (with disjoint interiors) of generalized Löbell Y-pieces, funnels and half-disks, such that P is one of the half-disks in the decomposition.

DEFINITION 5.2. We say that a half-disk $P \subset S$ is maximal if $P \subseteq Q$ implies P = Q for any half-disk $Q \subset S$.

PROPOSITION 5.2. If S is a hyperbolic Riemann surface not conformally equivalent to **D** or **D** \ $\{0\}$, then a half-disk $P \subset S$ is canonical if and only if it is maximal. Furthermore, any maximal half-disk appears in every decomposition of S.

Proof. If P is canonical, each point $p \in \partial P$ is the limit of points belonging to a simple closed geodesic, so if Q is another half-disk with $P \subseteq Q$, and we assume that $P \subset Q$, then Q intersects a simple closed geodesic, which is impossible, as in the proof of Proposition 5.1. Therefore P = Q.

On the other hand, if P is a maximal half-disk, consider any decomposition of S. Then, using Proposition 5.1, P is contained in a funnel F or in a canonical half-disk Q. If it is contained in F, then P cannot be maximal, because given a half-disk in a funnel, there is always another half-disk containing it strictly. Therefore $P \subseteq Q$, and for the maximality of P, P = Q and P is canonical.

REMARK 5.1. The Poincaré disk **D** does not contain maximal half-disks, and every half-disk is canonical.

5.1. Plane domains

We finish this section by describing a situation in which half-disks arise. If $S \subset \hat{\mathbf{C}}$ is a hyperbolic plane domain, then we see that under certain circumstances, there are half-disks near a continuum in ∂S .

Let us consider a hyperbolic plane domain S and a continuous C that is a connected component of ∂S (in the topology of $\hat{\mathbf{C}}$). Without loss of generality, we can assume that $S \subseteq \mathbf{D}$ and $C = \partial \mathbf{D}$.

Consider the following subset of $\partial \mathbf{D}$, $A = \partial \mathbf{D} \setminus \overline{\mathbf{D} \setminus S}$. The set A is an open subset in the relative topology of $\partial \mathbf{D}$.

PROPOSITION 5.3. If the arc $B = (\alpha, \beta) \neq \partial \mathbf{D}$ is a connected component of A, then there is a non-closed simple geodesic γ with endpoints in α and β such that γ is the boundary of a canonical half-disk.

Proof. Given a point $x \in B$, there exists a closed neighbourhood Q_x of x in $\overline{\mathbf{D}}$ such that $P_x := \mathbf{D} \cap Q_x$ is a half-disk in the surface S. This half-disk P_x cannot be contained in a funnel because in that case A would 'limit' this funnel and so the arc B would be $\partial \mathbf{D}$. Using Proposition 5.1, we see that the set P_x does not intersect any generalized Löbell Y-piece, so there exists a canonical half-disk P with $\bigcup_{x \in B} P_x \subseteq P$, since the set $\bigcup_{x \in B} P_x$ is connected. Let $\gamma = \partial P$. By construction, the points α and β are in the closure of P in $\overline{\mathbf{D}}$. Furthermore, α and β are the limit points of γ , since otherwise we have that there are points of $\mathbf{D} \setminus S$ in P since B is a connected component of A.

REMARK 5.2. It is not always possible to find an arc $B \subset A$ that is a connected component of A. If $A = \partial \mathbf{D}$ and $\#(\mathbf{D} \setminus S) \ge 2$, then there is a funnel 'limited' by $\partial \mathbf{D}$. The other possibility is that $A = \emptyset$.

Acknowledgements. We would like to thank Professor J. L. Fernández and Professor J. Gonzalo for many useful discussions. We thank the referee of this paper for a careful reading of the manuscript.

References

- 1. L. V. AHLFORS and L. SARIO, Riemann surfaces (Princeton University Press, Princeton, 1960).
- V. ÁLVAREZ, D. PESTANA and J. M. RODRÍGUEZ, 'Isoperimetric inequalities in Riemann surfaces of infinite type', Rev. Mat. Iberoamericana 15 (1999) 353–427.
- L. BERS, 'An inequality for Riemann surfaces', Differential geometry and complex analysis. H. E. Rauch memorial volume (Springer, 1985).
- 4. P. BUSER, Geometry and spectra of compact Riemann surfaces (Birkhäuser, Boston, 1992).
- R. D. CANARY, D. B. A. EPSTEIN and P. GREEN, 'Notes on notes of Thurston', Analytical and geometrical aspects of hyperbolic spaces (Cambridge University Press, 1987) 3–92.
- 6. I. CHAVEL, Eigenvalues in Riemannian geometry (Academic Press, New York, 1984).
- J. CHEEGER, 'A lower bound for the smallest eigenvalue of the Laplacian', Problems in analysis (Princeton University Press, 1970) 195–199.
- W. FENCHEL and J. NIELSEN, 'On discontinuous groups of isometric transformations of the non-Euclidean plane', Studies and essays presented to R. Courant (Interscience, New York, 1948).
- J. L. FERNÁNDEZ and M. V. MELIÁN, 'Escaping geodesics of Riemannian surfaces', Acta Math. 187 (2001) 213–236.
- J. L. FERNÁNDEZ and J. M. RODRÍGUEZ, 'The exponent of convergence of Riemann surfaces'. Bass Riemann surfaces', Ann. Acad. Sci. Fenn. Ser. A I 15 (1990) 165–183.
- 11. I. KRA, Automorphic forms and Kleinian groups (W. A. Benjamin, 1972).
- 12. W. M. MASSEY, Algebraic topology: an introduction (Harcourt, Brace, New York, 1967).
- P. J. NICHOLLS, The ergodic theory of discrete groups, Lecture Notes Series 143 (Cambridge University Press, Cambridge, 1989).
- 14. J. G. RATCLIFFE, Foundations of hyperbolic manifolds (Springer, New York, 1994).
- H. SHIMIZU, 'On discontinuous groups operating on the product of upper half-planes', Ann. of Math. 77 (1963) 33–71.

- D. SULLIVAN, 'Related aspects of positivity in Riemannian geometry', J. Differential Geom. 25 (1987) 327–351.
- W. P. THURSTON, The geometry and topology of three-manifolds, Lecture Notes (Princeton University, Princeton, NJ, 1979).

Venancio Álvarez	José M. Rodríguez
Departamento de Análisis Matemático	Departamento de Matemáticas
Facultad de Ciencias	Universidad Carlos III de Madrid
Universidad de Málaga	Avenida de la Universidad, 30
Campus de Teatinos	28911 Leganés
29071 Málaga	Spain
Spain	jomaro@math.uc3m.es

nancho@anamat.cie.uma.es