Surfaces

Sergiu Moroianu

April 29, 2021.

Contents

CHAPTER 1

Classification of topological surfaces

1. Background

We start by reviewing some notions of general topology.

Let S be a set. A *topology* on S is a subset $\mathcal{T}$ of $\mathcal{P}(S)$ (the set of subsets of S) such that

(1) $\emptyset, S \in \mathcal{T}$;
(2) if $A, B \in \mathcal{T}$ then $A \cap B \in \mathcal{T}$;
(3) if $\mathcal{T}' \subset \mathcal{T}$ then

$$\bigcup_{A \in \mathcal{T}'} A \in \mathcal{T}.$$

Elements of $\mathcal{T}$ are called *open* and their complements are called *closed* in S.

A function $f : S \to S'$ between topological spaces with topologies $\mathcal{T}, \mathcal{T}'$ is called *continuous* if for all $A' \in \mathcal{T}'$, the preimage $f^{-1}(A') := \{a \in S; f(a) \in A'\}$ belongs to $\mathcal{T}$. A *homeomorphism* is a continuous map as above which is bijective and such that the inverse map $f^{-1} : S' \to S$ is also continuous.

A *neighborhood* of a point $p \in S$ is any set which contains some open set $A \ni p$. A topological space is called *separate* if every two distinct points p, p' admit disjoint neighborhoods V, V'.

A sequence of points $(p_n)_{n \geq 1}$ in S is called *convergent* (towards $p \in S$) if for every open set $\mathcal{T} \ni A \ni p$, there exists $n(A) \in \mathbb{N}$ such that $p_n \in A$ for all $n \geq n(A)$.

A *basis* for the topology $\mathcal{T}$ is a subset $\mathcal{B} \subset \mathcal{T}$ such that for every element A of $\mathcal{T}$, there exists $\mathcal{B}' \subset \mathcal{B}$ such that

$$A = \bigcup_{B \in \mathcal{B}'} B.$$

In this book we will work with spaces with countable basis, in order to avoid talking about nets. Under this standing assumption, one can describe continuity in terms of convergent sequences (Exercise 1.1).

A topological space S is called *compact* if every open cover admits a finite sub-cover. More precisely, if $\mathcal{C} \subset \mathcal{T}$ is such that $\bigcup_{A \in \mathcal{C}} A = S$, there exists a finite subset $\mathcal{C}' \subset \mathcal{C}$ such that $\bigcup_{A \in \mathcal{C}'} A = S$.

A topological space S is called *connected* if it cannot be partitioned in two non-empty disjoint open sets. It is easy to see that the union of any family of connected subsets of S with non-empty intersection is again connected (Exercise 1.6). Thus the set of connected subsets of S with the order given by inclusion is well ordered. Any one-point subset is clearly connected. By Zorn's

Lemma, there exist therefore maximal connected subsets (called *connected components*), which are necessarily disjoint, and which cover S.

1.1. Exercises.

EXERCISE 1.1. Let S, S' be topological spaces. Show that for any continuous function $f : S \to S'$ and any convergent sequence $(x_n) \to x \in S$, the sequence $f(x_n)$ converges to $f(x)$ in S'. Conversely, if a function f has this property and the topology of S has countable basis, then f must be continuous.

EXERCISE 1.2. Let S be a topological space and S' any subset. Let

$$\mathcal{T}' := \{A \cap S'; A \in \mathcal{T}\}.$$

Show that $\mathcal{T}'$ is a topology on S' (called the *induced topology*). More generally, for every set S and any function $f : S' \to S$, define $\mathcal{T}' := f^{-1}(\mathcal{T}) \subset \mathcal{P}(S')$. Show that $\mathcal{T}'$ is a topology with respect to which f becomes continuous. If $\mathcal{T}''$ is another topology on S' such that $f : (S', \mathcal{T}'') \to (S, \mathcal{T})$ is continuous, show that $\mathcal{T}' \subset \mathcal{T}''$ (i.e., $\mathcal{T}'$ is the smallest topology on S' which makes f continuous).

EXERCISE 1.3. Construct an example of continuous map $f : [0, 1) \to S \subset \mathbb{R}^2$ which is bijective, but such that the inverse function $f^{-1} : S \to [0, 1)$ is not continuous.

EXERCISE 1.4. Show that a subspace of a compact separate topological space S is compact if and only if it is closed.

EXERCISE 1.5. Show that the image of a compact space through a continuous map is compact. Give an example of a continuous map and a compact set whose pre-image is not compact. Same questions when “compact” is replaced by “connected”.

EXERCISE 1.6. Let $\mathcal{C} \subset \mathcal{P}(S)$ be a family of connected subsets of S such that $\cap_{A \in \mathcal{C}} A \neq \emptyset$. Then $\cup_{A \in \mathcal{C}} A$ is connected.

EXERCISE 1.7. Show that the connected components of any topological space are closed, but not necessarily open.

2. Topological surfaces

If S is a topological space with topology $\mathcal{T}$ and $f : S \to S'$ a function into a set S', we can define $\mathcal{T}' \subset \mathcal{P}(S')$ by

$$\mathcal{T}' := \{A' \subset S'; f^{-1}(A') \in \mathcal{T}\}. \tag{2.1}$$

By Exercise 2.6, $\mathcal{T}'$ is a topology on S'. Thus, for any equivalence relation $\sim$ on a topological space S, the projection from S onto the set $S/\sim$ of equivalence classes gives a topology on $S/\sim$, called the *identification topology*.

DEFINITION 2.1. A topological space is called a (topological) *surface* if it is locally homeomorphic to $\mathbb{R}^2$.

More precisely, there exists an open cover $\mathcal{C}$ of S such that for every $A \in \mathcal{C}$ there exists an open set $A' \subset \mathbb{R}^2$ and a homeomorphism $\phi_A : A \to A'$. Such a homeomorphism is called a *chart*, and the set of all homeomorphisms $\{\phi_A; A \in \mathcal{C}\}$ is called an *atlas*. If ϕ_A, ϕ_B are charts, the map

$$\phi_{AB} := \phi_B \circ \phi_A^{-1} : \phi_A(A \cap B) \to \phi_B(A \cap B) \tag{2.2}$$

is called the *change of charts*. By exercise 2.7, Φ_{AB} is always a homeomorphism.

The simplest example of surface, besides open sets of $\mathbb{R}^2$, is the sphere $S^2 := \{(x, z) \in \mathbb{R} \times \mathbb{C}; x^2 + |z|^2 = 1\}$. To construct charts we use *stereographic projections* from the north and south poles, i.e., the points $e = (1, 0, 0)$ and $-e$. Every line passing through e and another point $p \in S^2$ intersects $\mathbb{C}$ in a unique point $\Phi^+(p)$. The formula is

$$\Phi^+(x, z) = \frac{z}{1 - x}, \qquad (\Phi^+)^{-1}(w) = \left(\frac{2w}{|w|^2 + 1}, \frac{|w|^2 - 1}{|w|^2 + 1}\right)$$

so Φ^+ is a homeomorphism from $S^2 \setminus \{e\}$ to $\mathbb{R}^2$. Similarly one constructs the homeomorphisms $\Phi^- : S^2 \setminus \{-e\} \to \mathbb{R}^2$.

Let S be a topological surface. An atlas with smooth (respectively holomorphic) changes of charts is called a *smooth* (respectively *holomorphic*) atlas (we identify $\mathbb{R}^2$ with $\mathbb{C}$ via $(x, y) \leftrightarrow x + iy$). Two smooth (respectively holomorphic) atlases on S are *compatible* if their union is again a smooth (respectively holomorphic) atlas. This defines an equivalence relation on the set of atlases on S.

DEFINITION 2.2. A *smooth structure* (respectively a *holomorphic structure* on a topological surface is an equivalence class of smooth (respectively holomorphic) atlases.

A smooth surface (respectively a Riemann surface) is a topological surface endowed with a smooth (respectively holomorphic) structure. We leave it to the reader to define analogously the notions of topological, differentiable and complex analytic manifolds in every dimension.

Clearly, every open subset of $\mathbb{C}$ is a complex manifold, with the atlas consisting of a single chart (the inclusion map into $\mathbb{C}$). Note that in the holomorphic setting the natural notion of dimension is the complex dimension, so a topological surface endowed with a holomorphic structure is usually called a *complex curve*. Nevertheless, in this book we keep the word "surface" for consistency.

The atlas with two charts on the sphere is smooth because the change of charts is given by

$$\Phi^+ \circ (\Phi^-)^{-1}(w) = 1/\overline{w}$$

but it is not holomorphic. However, if we replace Φ^+ by its complex conjugate $\overline{\Phi^+}$, the change of charts becomes holomorphic hence the sphere has a complex analytic structure.

EXAMPLE 2.3. Let $S := \mathbb{C} \times \{0\} \sqcup \mathbb{C} \times \{1\}$ be the disjoint union of two copies of the complex plane, with the following equivalence relation: $(x, 0) \sim (y, 1)$ if and only if $x = y \neq 0$. The quotient space is a complex plane with a "double origin". It is not separate because the images of $(0, 0)$ and $(0, 1)$ in the quotient space do not have disjoint neighborhoods.

In the sequel we will work mainly with separate surfaces. We do not include this in the definition but we will check case-by-case that our examples of surfaces are separate.

2.1. Group actions. Let Γ be a group acting on a surface S; this means that for all $\gamma \in \Gamma$ we specify a continuous map $\rho(\gamma) : S \to S$. We ask the action to be compatible with the group operation, thus $\rho(\gamma) \circ \rho(\gamma') = \rho(\gamma\gamma')$, and moreover we require $\rho(1) = 1_S$, the identity map of S. It follows easily that $\rho(\gamma)$ is a homeomorphism. When there can be no ambiguity about the action, we simply write γ for the action $\rho(\gamma)$ on S.

Our main example will be the action by homographies of $\mathrm{GL}_2(\mathbb{C})$ on the complex plane. The action of a matrix $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ on a complex number z is given by

$$\gamma z := \frac{az+b}{cz+d}. \tag{2.3}$$

EXAMPLE 2.4. Define

$$\mathbb{H} := \{z \in \mathbb{C}; \Im(z) > 0\}.$$

Then the group $\mathrm{GL}_2^+(\mathbb{R})$ of 2×2 real matrices with positive determinant acts on $\mathbb{H}$ because

$$\Im(\gamma z) = \frac{\Im(z)\det(\gamma)}{(cz+d)^2}.$$

The kernel of the action consists of multiples of the identity matrix, which is the center of $\mathrm{GL}_2^+(\mathbb{R})$. We can restrict the action to $SL_2(\mathbb{R})$ (the group of real matrices with determinant 1), in which case the kernel of the action is $\{\pm 1\}$ which is the center of $SL_2(\mathbb{R})$ (exercise 2.9). Thus the quotient group

$$\mathrm{PSL}_2(\mathbb{R}) := \mathrm{GL}_2^+(\mathbb{R})/\mathbb{R}^* \simeq SL_2(\mathbb{R})/\{\pm 1\}$$

acts *faithfully* on $\mathbb{H}$, in the sense that if $\gamma z = z$ for all $z \in \mathbb{H}$ then $\gamma = 1 \in \mathrm{PSL}_2(\mathbb{R})$.

If Γ acts on a surface S, the image of any open set $A \subset S$ in $\Gamma \backslash S$ is again open. Indeed, if we denote by ϕ the class map

$$\phi : S \to \Gamma\backslash S, \qquad p \mapsto \Gamma p := [p], \tag{2.4}$$

then $\phi^{-1}(\phi(A)) = \bigcup_{\gamma\in\Gamma} \gamma A$. Since γ is a homeomorphism and arbitrary unions of open sets are again open, we see that $\phi(A)$ is open by the definition of the quotient topology.

Assume moreover that the group action is *proper discontinuous* on the surface S in the following sense: for all $x \in S$, there exists $U \ni x$ a neighborhood such that for every $\Gamma \ni \gamma \neq 1$, $\gamma(U)$ is disjoint from U. This implies that there are no fixed points.

THEOREM 2.5. *The set of equivalence classes $\Gamma\backslash S$ of a proper discontinuous action on a surface is again a surface.*

PROOF. Let $[p]$ be the class of a point $p \in S$. Let U' be a neighborhood of p with $U \cap \gamma(U) \neq \emptyset \Rightarrow \gamma = 1$. Let U be the intersection of U' with the domain of a chart around p, thus U is homeomorphic to an open subset of $\mathbb{C}$. The image of U through the projection ϕ from (2.4) is open. Moreover $\phi : U \to \phi(U)$ is clearly bijective by the choice of U, and since it is also continuous we conclude that it is a homeomorphism. Thus $\phi(U) \ni [p]$ has a neighborhood homeomorphic to an open set in $\mathbb{C}$. $\square$

2.2. Exercises.

EXERCISE 2.6. Show that $\mathcal{T}'$ defined in (2.1) is a topology on S' such that $f : S \to S'$ becomes continuous. For any other topology $\mathcal{T}''$ on S' for which f is continuous, we have $\mathcal{T}'' \subset \mathcal{T}'$ (i.e., $\mathcal{T}'$ is the largest topology on S' which makes f continuous).

EXERCISE 2.7. Show that the function ϕ_{AB} defined in (2.2) is a homeomorphism.

EXERCISE 2.8. Show that (2.3) defines a group action on $\mathbb{C}$.

EXERCISE 2.9. Show that the center of $SL_2(R)$ (i.e., the subgroup of elements which commute with all matrices in $SL_2(R)$) is made of $\{I, -I\}$ where I is the identity matrix.

EXERCISE 2.10. Two actions of $\mathbb{Z}$ on $\mathbb{C}$ are defined by the action of the generator 1:

$$\rho_1(1)z := z + 1, \qquad \rho_2(1)z := 1 + \overline{z}$$

Show that in both cases, every open set of the form $(x - \frac{1}{2}, x + \frac{1}{2}) \times \mathbb{R}$ is mapped homeomorphically onto its image in the quotient space. The quotient surfaces are called the *cylinder*, respectively the *Möbius band*.

Try (!) to construct in both cases atlases with two charts with domains the images of $(-1/2, 1/2) \times \mathbb{R}$ and $(0, 1) \times \mathbb{R}$ such that the changes of charts are holomorphic!

EXERCISE 2.11. For every $\tau \in \mathbb{C}$ with $\Im(\tau) \neq 0$ consider the action of $\mathbb{Z}^2$ on $\mathbb{C}$ defined by the commuting actions of the generators: $\rho(1,0)z = z + 1$, $\rho(0,1)z = z + i$. Show that the resulting surfaces (tori) are homeomorphic for all τ.

EXERCISE 2.12. Let Γ be the group with two (non-commuting) generators e_1, e_2 subject to the relation $e_1 e_2 = e_2^{-1} e_1$. If we define homeomorphisms of $\mathbb{C}$ by $\rho(e_1)z := \overline{z} + 1$, $\rho(e_2)z := z + i$, show that we get a group action of Γ on $\mathbb{C}$. Show that the action is proper discontinuous. The quotient is called the *Klein bottle*. What is the abelianisation of Γ (i.e., the quotient of Γ by the commutator subgroup; here you need to make sure that in every group, the subgroup generated by elements of the form $aba^{-1}b^{-1}$ is normal)?

EXERCISE 2.13. Show that the action of the free group with two generators F_2 on $\mathbb{C}$ defined by $e_1 z = \overline{z} + 1$, $e_2 z = i - \overline{z}$, is not free (i.e., it has fixed points: there exists $1 \neq \gamma \in F_2$ and $z \in \mathbb{C}$ with $\gamma z = z$).

EXERCISE 2.14. A *covering map* $f : S' \to S$ is a continuous map with the following property: every $x \in S$ has a neighborhood $U \ni x$ such that every connected component of $f^{-1}(U)$ is homeomorphic to U by the map f. Show that each properly discontinuous group action on a surface S' gives rise to a covering map $S' \to \Gamma \backslash S'$.

3. Triangulations

A *triangle* (or a 2-simplex) in a topological space S is the image of a homeomorphism from the standard triangle

$$\Delta_2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 + x_2 + x_3 = 1, x_j \geq 0\} \tag{3.1}$$

onto a subset of S. The restrictions of this homeomorphism to the vertices and the edges of Δ_2 give well-defined vertices in S (the images of vertices of the triangle) and edges (the homeomorphic images of the three edges of Δ_2). The union of the three edges is called the boundary of the triangle. The interior of an edge is by definition the edge without its end-points. The interior of the triangle is the image of the *open* triangle Δ_2° obtained by requiring strict inequalities in (3.1). One could similarly define n-simplexes in S for any dimension $n \geq 0$.

DEFINITION 3.1. A *triangulation* of a space S is a locally finite collection of triangles $(T_\alpha)_{\alpha\in A}$ on S with the following properties:

- The subsets
$$\begin{aligned} S_0 :=& \{p \in S; p \text{ is a vertex for some triangle}\}\\ S_1 :=& \{p \in S; p \text{ belongs to the interior of some edge}\}\\ S_2 :=& \{p \in S; p \text{ belongs to the interior of some triangle}\}\end{aligned}$$
are disjoint and their union is S;
- Points in S_2 belong to a unique triangle.
- Two edges which contain a common point in S_1 must coincide.

Locally finite means that every point has a neighborhood which intersects only a finite number of triangles. In a triangulated surface, two edges belong precisely to one triangle. Moreover, each vertex is contained in a finite number of triangles and a finite number of edges $E_1, \dots, E_k$ such that E_i, E_j form a triangle if and only if $|i-j| = 1$ or $i = 1, j = k$. A different way of seeing a triangulation is via the induced topology. We take a disjoint union of a family of triangles (i.e., copies of the standard triangle) and introduce an equivalence relation on them with the following conditions:

- if $p \neq p'$ but $p \sim p'$ then p and q must live on the boundary of two different triangles;
- For every edge E, the equivalence class of $p \in E^\circ$ contains precisely 2 elements.
- If $E^\circ \ni p \sim p' \in E'^\circ$ there exists a homeomorphism $\phi : E \to E'$ with $\phi(p) = p'$ and such that for all $q \in E$ we have $q \sim \phi(q)$. In other words, the whole edge E is identified to E' via ϕ.
- If p_0 is a vertex, its equivalence class is a finite set of vertices $\{p_0, \dots, p_{n-1}\}$. We can order them so that the edges through p_j are E_j and ${E_j}'$ such that ${E_j}'$ is identified with E_{j+1} for all j modulo n.

Then one can easily check that the quotient map with its induced topology is a surface. A triangulation is a homeomorphism from such a combinatorial space to a given surface S.

We will assume without proof the following fact:

THEOREM 3.2. *Every surface can be triangulated.*

An *orientation* of a triangle T is a direction of rotation around the edges. More formally, an orientation is an equivalence class of orders on the set of vertices of T modulo even permutations (an even permutation is a product of an even number of transpositions). To specify an orientation on T it is enough to order the vertices of one edge.

DEFINITION 3.3. An orientation on a triangulated surface is a collection of orientations on each triangle which are compatible at edges in the following sense: if E is an edge which belongs to the triangles T, T', then the orders on the set of vertices of E induced from the orientations of T, T' are different.

In other words, the sense along the edge E indicated by one triangle is opposite to the sense of rotation around the second triangle. Note that an orientation need not exist. The Möbius band provides a counterexample.

Let now S be a connected triangulated surface. Notice that the vertices and the edges of the triangulation form a graph. Let Γ be a maximal tree in that graph (a tree is a sub-graph without cycles). The *dual graph* is constructed (abstractly) as follows: the set of vertices of Γ' is the set of faces of the triangulation, and two vertices F_1, F_2 are joined by an edge in Γ' if and only if the triangles F_1, F_2 intersect in an edge which does not belong to Γ.

To realize Γ' geometrically on S we introduce barycentric sub-division. Namely, we choose one point in the interior of each face (we think of it as the "barycenter") and one point in the interior of each edge. Although we do not have a notion of length, we think of this new point as the "midpoint" of the edge. Inside each triangle, we then link the barycenter with the three initial vertices, and also with the three new vertices on the boundary. Thus every triangle has been sub-divided in 6 new triangles. Inside this new triangulation, we view the dual graph as having vertices at the barycenters of the faces of the initial triangulation. The edges of Γ' are realized as concatenation of edges in the barycentric sub-division going through the barycenter of an edge which does not belong to Γ . (picture)

We denote by $S^{(n)}$ the n-th iterated barycentric subdivision of a triangulated surface S.

From now on we suppose that S is compact. Notice that the number of triangles in any triangulation of a compact surface is finite. This follows immediately from the local finiteness of the triangulation, but let us see it in a different way. Define the *open star* of a vertex V in the triangulation as the union of the interiors of all simplexes (i.e., triangles and edges) which contain V, together with V itself. This is an open subset of S. The union of all such sets covers S since every edge and triangle have some vertex. By compactness, a finite number of open stars cover S; since each open star contains the interiors of only a finite number of triangles, the total number of triangles must be finite.

PROPOSITION 3.4. *Every compact triangulated surface S can be split into two closed neighborhoods U, U' of Γ, Γ' which intersect along a circle. Moreover $U \setminus \Gamma$ is path-connected.*

A space S is called *path-connected* if for any two points $p_0, p_1 \in S$ there exists a continuous map $f : [0, 1] \to S$ (i.e., a *path*) with $f(0) = p_0$ and $f(1) = p_1$.

PROOF. Take as U, U' the union of those triangles in the second barycentric subdivision $S^{(2)}$ which touch Γ, respectively Γ'. Then U, U' intersect along their boundary. By induction on the number of edges of the tree Γ we see that the boundary of U is a circle. Moreover, clearly every point in U can be joined to the boundary without touching Γ. $\square$

As a corollary, we can prove that $S \setminus \Gamma$ (and thus Γ') is connected. Let $p, q \in S \setminus \Gamma$. There exists a path between them in S, say $c : [0,1] \to S$ with $c(0) = p, c(1) = q$. Let t_1, t_2 the infimum, resp. the supremum of all times t for which $c(t) \in U$, where U is the neighborhood of Γ constructed in Prop. 3.4. If both p and q belong to U, then we connect them to ∂U without intersecting Γ, then travel along Γ. If both p and q live in the interior of U', then $0 < t_0 \leq t_1 < 1$. Now $c(t_0), c(t_1)$ both live on the boundary of U so we can link p to q by traveling along c until $t = t_0$, then around ∂U and finally again along c from time $t = t_1$ to $t = 1$.

In conclusion, we have split out surface in two pieces, one of which is a topological disk and the other one a tubular neighborhood of a graph. It will turn out that the number of "elementary cycles" in that graph, plus some mild information about orientation around these cycles, determine completely the surface S.

3.1. Exercises.

EXERCISE 3.5. Show that a path-connected space is connected. Give counterexamples to the converse.

EXERCISE 3.6. Prove that every connected and locally path-connected space is path-connected. Convince yourselves that every connected surface is locally path-connected. Conclude that every connected surface is path-connected.

EXERCISE 3.7. Construct a triangulation of the cylinder and of the Möbius band (Exercise 2.10) using as vertices the points with coordinates $(0, j)$ and $(1, j)$ for j integer. Show that the cylinder is orientable while the Möbius band is not.

4. Euler characteristic

Let G be a connected graph with v vertices and e edges. If G is a tree, then $v - e = 1$ and this identity characterizes trees among graphs. In general, the number $v - e$ is called the *Euler characteristic* of the graph, denoted $\chi(G)$. By induction, we see that $1 - \chi(G)$ represents the number of "eyes" of the graph, in the sense that by removing $1 - \chi(G)$ edges without disconnecting the graph we obtain a tree. Clearly $\chi(G) \leq 1$.

Let now S be a compact triangulated surface with v vertices, e edges and f faces. The number

$$\chi(S) := v - e + f$$

is called the Euler characteristic of the surface S. We divide the e edges of S into the e_Γ edges in Γ and the $e_{\Gamma'}$ edges which yield edges of the dual graph. Since there are as many faces in S as vertices in Γ', we see that

$$\chi(S) = \chi(\Gamma) + \chi(\Gamma') = 2 - \#\{\text{elementary cycles in } \Gamma'\}$$

It follows that $\chi(S)$ is at most 2.

If $\chi(S) = 2$, the graph Γ' does not have cycles. The surface S is obtained from two disks (the closed sets U, U' consisting of the closed stars of Γ, Γ' in the second barycentric subdivision of S) by gluing along the boundary circle. The result is a sphere.

If $\chi(S) < 2$, there exists a cycle in Γ'. By induction, there exist $2 - \chi(S)$ segments in U' so that, after we cut U' along them, we obtain a disk. Thus any surface can be reconstructed from two disks with some identifications on the boundary.

THEOREM 4.1. *Let S be a compact, connected, orientable triangulated surface. If $\chi(S) = 2$ then S is homeomorphic to the sphere S^2. In general, $\chi(S)$ is even and S is homeomorphic to a sphere with $g = \frac{2-\chi(S)}{2}$ handles attached.*

Attaching a handle means removing two disks D_1, D_2 with boundaries C_1, C_2 from S and gluing instead a cylinder along the two circles C_1, C_2.

PROOF. We have already seen that $\chi(S) \leq 2$, with equality if and only if S is a sphere. Assume that $\chi(S) < 2$. Then the dual graph Γ' is not a tree, hence it has a simple cycle C. The simplexes in the second barycentric subdivision $S^{(2)}$ which touch C form a collar neighborhood of C, say U. If we cut U along a segment transversal to C we obtain a rectangle, thus U is obtained from a rectangle by gluing along two opposite edges, i.e. U is either a cylinder or a Möbius band. By Exercise 4.5 and the orientability hypothesis only the first possibility can occur, hence the boundary of U is made of two circles C_1, C_2. Apply now *surgery* along U, i.e. remove U and replace it by two disks with boundaries C_1, C_2. This is the opposite operation to attaching a handle. We call S_1 the surface thus obtained from S by surgery along C.

The Euler characteristic of S_1 is related to that of S as follows: suppose there were a vertices on C. Then by surgery we introduce a new vertices, a edges, two faces (the two disks), and we remove as many faces as edges when we erase the cylinder U. Thus the Euler characteristic goes up by 2.

By iterating surgery as long as we get cycles in Γ', we get the conclusion. □

There is a major problem in what we have achieved so far: our invariant $\chi(S)$ seems to depend on the choice of triangulation. Thus, although we see that every orientable surface is homeomorphic to a sphere with g handles attached, we do not know yet that surfaces of different genera are not homeomorphic, because the genus is defined in terms of a choice of triangulation. In reality it does not depend on the triangulation but in order to make things rigorous, we need to introduce a finer algebraic invariant of spaces.

4.1. Exercises.

EXERCISE 4.2. Show that in a compact triangulated surface we have the identity

$$2e = 3f.$$

EXERCISE 4.3. The faces of a polyhedron can be subdivided into triangles. Check that the Euler characteristic remains unchanged during this process.

EXERCISE 4.4. A regular polyhedron is a convex polyhedron with all faces made of regular polygons, and all solid angles congruent. Determine and draw all regular polyhedra using the fact that the Euler characteristic is 2 and an identity similar to Exercise 4.2.

EXERCISE 4.5. Show that a surface is orientable if and only if it does not contain an embedded Möbius band.

5. The fundamental group

Let S be a space and $p, q \in S$. A *path* from p to q is a continuous map c from the interval $I := [0,1]$ to S which maps 0 in p and 1 in q. Two paths are *homotopic* if there is a "path of paths" linking them. More precisely, $c_0 \simeq c_1$ if there exists a continuous map $C : I \times I \to S$ such that $C(0,\cdot) = c_0$ and $C(1,\cdot) = c_1$.

The above map C is called a homotopy between c_0 and c_1. If $C(s,0) = p$ is constant in s, the homotopy is called *relative to* $0 \in I$; of course a necessary condition for the existence of a homotopy between c_0 and c_1 relative to 0 is that $c_0(0) = c_1(0)$. Similarly, the homotopy is *relative to the end-points* if $C(s,0) = p$ and $C(s,1) = q$ are constant in s.

A *loop* (based at $p \in S$) is a closed path with endpoints p, i.e., such that $c(0) = c(1) = p$. The set of loops based at p modulo the equivalence relation given by homotopy relative to the end-points is called the *fundamental group*, or the Poincaré group, denoted $\pi_1(S,p)$.

There is an operation on loops based in p defined by concatenation:

$$(cc')(t) := \begin{cases} c'(2t) & \text{for } 0 \le t \le \frac{1}{2} \\ c(2t-1) & \text{for } \frac{1}{2} \le t \le 1 \end{cases}$$

Concatenation means that we travel along c' and then along c. It makes sense also for paths which are not necessarily closed, but such that $c(0) = c'(1)$. This operation is neither commutative nor associative.

THEOREM 5.1. *The operation of concatenation is well-defined on* $\pi_1(S,p)$ *and is associative there.*

PROOF. If C, C' are homotopies between c and c_1, respectively between c' and c_1' relative to the end-points, then $\tilde{C}(s,t) := (C_s C_s')(t)$ is a homotopy between cc' and $c_1 c_1'$ relative to the end points. Associativity is best seen in a picture. $\square$

Let c_p be constant loop $c_p(t) = p$. If c is any loop based at p, then cc_p and $c_p c$ are homotopic to c relative to the end-points. For instance,

$$C(s,t) := \begin{cases} p & \text{for } 0 \le t \le \frac{s}{2} \\ c(\frac{2t-s}{2-s}) & \text{for } \frac{s}{2} \le t \le 1 \end{cases}$$

is a homotopy from c to cc_p.

If we denote by $\overline{c}$ the walk along the loop c in the opposite order (i.e., $\overline{c}(t) = c(1-t)$) then one sees easily that $c\overline{c} \sim c_p$.

Therefore $\pi_1(S,p)$ is a group, in general non-commutative.

EXAMPLE 5.2. $\pi_1(\mathbb{R}^n, 0) = \{1\}$ for all n. Indeed, every loop c is homotopic to the constant loop 0, i.e., the unit in π_1, via homotheties:

$$C_s(t) := sc(t).$$

Similarly, for every contractible space S and any point $p \in S$, we have $\pi_1(S,p) = \{1\}$.

If we change the base point, we get another group $\pi_1(S,p')$. Assuming that S is path-connected, let d be any path from p to p'. Then the map which to $[c] \in \pi_1(S,p')$ associates $[\overline{d}cd] \in \pi_1(S,p)$ is well-defined (does not depend on the choice of the representative c), and is a group morphism that we denote Φ_d.

LEMMA 5.3. *The group morphism $\Phi_d : \pi_1(S,p') \to \pi_1(S,p)$ depends only on the homotopy class of d relative to the end points.*

If d' is a path from p' to p'', then $\Phi_d \circ \Phi_{d'} = \Phi_{d'd}$ as group morphisms from $pi_1(S,p'')$ to $\pi_1(S,p)$.

In particular, if $\Phi_{\overline{d}} : \pi_1(S,p') \to \pi_1(S,p)$ is the morphism induced by the inverse path $\overline{d}$, we infer that Φ_d is an isomorphism. Thus the fundamental group does not depend on the base point up to isomorphism.

If we choose another path e from p to p', we form a loop based in p by setting $\gamma := \overline{e}d$. We can compare the isomorphisms Φ_d and Φ_e

$$\Phi_e c = \overline{e}ce = \overline{e}d\overline{d}cd\overline{d}e = \gamma\Phi_d\overline{\gamma}$$

thus $\Phi_e = [\gamma]\Phi_d[\gamma]^{-1}$. It follows that the isomorphism Φ_d is well-defined up to conjugation. Thus, although in general we do not have a canonical isomorphism from $\pi_1(S,p')$ to $\pi_1(S,p)$, a conjugacy class of such isomorphisms does exist. If the group is Abelian, all inner automorphisms are trivial so in that case the above isomorphism Φ_d is canonical.

There exists a closely related set, called FS, made of the free homotopy classes of loops in S.

THEOREM 5.4. *Let S be a path-connected space and $p \in S$. The set FS of free homotopy classes of loops in S is canonically identified with the set of conjugacy classes inside $\pi_1(S,p)$.*

PROOF. We have a tautological function $\pi_1(S,p) \to FS$ defined by taking a loop based in p into its free homotopy class. This map is a surjection since every loop is freely homotopic to a loop based in p. If $c, c' \in \pi_1(S,p)$ are conjugate via γ, i.e., $c' = \overline{\gamma}c\gamma$, then their free homotopy class is the same, a homotopy at time s being given by $\overline{\gamma_s}c\gamma_s$, where γ_s is the portion of the path γ for time at most s. Let us now take two loops c, c' based in p which represent the same free homotopy class. We view the loops as maps from S^1 to S which map 1 to p, via Exercise 5.6. Let $C : S^1 \times I \to S$ be a (free) homotopy. Let γ be the path $s \mapsto C(1,s)$. This path is in fact a loop in p, and c is homotopic to $\overline{\gamma}c'\gamma$ relative to the end points. $\square$

Associating the fundamental group to each pair (S,p) is a *functor*: for each map of spaces $f : S \to S'$, $f(p) = p'$ we get a group morphism $f_* : \pi_1(S,p) \to \pi_1(S',p')$ by setting $f_*[c] = c \circ f$ (Exercise 5.9).

5.1. Exercises.

EXERCISE 5.5. Show that every path is homotopic to a constant path relative to 0.

EXERCISE 5.6. Show that the formula

$$\phi(e^{2\pi ix}) := f(x)$$

defines a bijection from the set of loops on S into the set of maps from S^1 to S.

EXERCISE 5.7. Prove that (relative) homotopy is an equivalence relation on the set of paths in S.

EXERCISE 5.8. Two points in S are called joinable if there exists a path linking them. Show that being joinable is an equivalence relation. The equivalence classes are called "path-connected components". If $S_p \subset S$ is the path-connected component of p, show that $\pi_1(S, p) = \pi_1(S_p, p)$.

EXERCISE 5.9. The function f_* is well-defined, and is a group morphism.

6. Computations of homotopy groups

Let $S' \to S$ be a covering map and $c : I \to S$ a path starting in $p \in S$. A path $c' : I \to S'$ is called a *lift* of c if $\pi \circ c' = c$.

LEMMA 6.1. *Choose $p' \in S'$ such that $\pi(p') = p$. Then there exists a unique lift of c with $c'(0) = p'$.*

PROOF. Using that I is compact, cover $c(I)$ with a finite number of open sets U_j so that $\pi^{-1}(U_j)$ is a disjoint union of open sets in S' homeomorphic to U_j via π. Pick points $0 = t_0 < t_1 < \ldots < t_n = 1$ so that $c(t_j) \in U_j \cap U_{j-1}$. Assume that c' has been defined up to time t_j. Set $p'_j := c'(t_j)$. Let U'_j, U'_{j-1} be the components of $\pi^{-1}(U_j)$, $\pi^{-1}(U_{j-1})$ containing p'_j. Then we can clearly continue c' up to time t_{j+1}. To prove uniqueness, let T be the supremum of all times t so that every two lifts coincide up to time t. By continuity, every two lifts coincide up to time T. Assume $T \in U_j$, let $T' := c'(T)$ and $U'_j \subset S'$ be the connected component of $\pi^{-1}(U_j)$ containing T'. Then clearly c' is uniquely defined for larger times that T, unless $T = 1$. $\square$

A more general result is the so-called *homotopy lifting lemma*:

LEMMA 6.2 (Homotopy lifting). *Let X be a locally compact space. Let $f_0, f_1 : X \to S$ be continuous maps and $F : I \times X \to S$ a homotopy between f_0 and f_1. Assume that there exists a lifting $f'_0 : X \to S'$. Then there exists a lifting $F' : I \times X \to S$ of the homotopy F.*

PROOF. By the previous result, F' is unique (if it exists), and moreover $F'(\cdot, x)$ is the lift of the curve $F(\cdot, x)$ starting at $f'_0(x)$. By Exercise 6.4 F' is continuous. $\square$

These facts allow us to compute $\pi_1(S^1, 1)$. Let $c : I \to S^1$ be a loop with $c(0) = c(1) = 1$. Lift it in the covering

$$\Phi : \mathbb{R} \to S^1, \qquad \Phi(x) = \exp(2\pi i x).$$

The point $c'(1)$ sits above $c(1) = 1$ hence $\exp(2\pi i c'(1)) = 1 \Leftrightarrow c'(1) \in \mathbb{Z}$. This defines a function $L : \pi_1(S^1, 1) \to \mathbb{Z}$. Indeed, if c_0, c_1 are homotopic relative to their end points, then by the homotopy lifting lemma we have $c'_0(1) = c'_1(1)$.

Since $\mathbb{Z}$ is commutative, the base point is not important so we write $\pi_1(S^1) = \mathbb{Z}$.

As a corollary, $\pi_1(T^2) = \mathbb{Z}^2$ where $T^2 = S^1 \times S^1$ is the torus. This follows from a general fact:

LEMMA 6.3. *Let S_1, S_2 be topological spaces and $p_1 \in X_1, p_2 \in X_2$ base points. Denote $X := X_1 \times X_2, p = (p_1, p_2) \in X$. Then $\pi_1(X, p) = \pi_1(X_1, p_1) \times \pi_1(X_2, p_2)$.*

PROOF. Every loop in (X, p) is of the form

$$c(t) = (c_1(t), c_2(t)) = (c_1 \times c_2)(t)$$

where c_1, c_2 are loops in X_1, X_2. Define a map

$$\pi_1(X, p) \to \pi_1(X_1, p_1) \times \pi_1(X_2, p_2), \qquad [c] \mapsto ([c_1], [c_2]).$$

This map is evidently well-defined, i.e., it does not depend on the choice of $c \in [c]$. By Exercise 6.6 it is an isomorphism. □

6.1. Exercises.

EXERCISE 6.4. Show that F' defined in the proof of the homotopy lifting lemma is continuous.

PROOF. Since Y is locally compact, it is enough to prove continuity when restricted to a compact subset, so we can assume that Y itself is compact. Let $J := \{t \in I; F' : [0, t] \times Y \to S' \text{ is continuous}\}$. We first prove that the set J is open. Let $t \in J$. For all $y \in Y$, there exists an open set $U_y \ni y$ and $\epsilon_y > 0$ such that $F(t - \epsilon_y, t + \epsilon_y) \times U_y)$ is contained in an open set V_y whose preimage through π is a disjoint union of open sets homeomorphic to V_y. By compactness, there exist a finite number of such sets U_j and $\epsilon > 0$ such that $F([t - \epsilon, t + \epsilon] \times U_j) \subset V_j$ where V_j is a open set adapted to the covering map π as above. It is then obvious that the unique extension of F' to time $t + \epsilon$ is continuous. To show that J is closed (and hence $J = I$ which finishes the proof) let $T := \sup(J)$. For every $y \in Y$, there exists $U \ni y$ open and $\epsilon > 0$ so that $F([T - \epsilon, T + \epsilon] \times U)$ is contained in an open set V adapted to π. Then clearly the unique extension F', which is continuous on $\{T - \epsilon\} \times U$ by the hypothesis on T, is also continuous on the whole $[T - \epsilon, T + \epsilon] \times U$. Thus F' is continuous at (T, y), which shows $T \in J$. □

EXERCISE 6.5. Show that the map $L : \pi_1(S^1, 1) \to \mathbb{Z}$ defined using liftings is a group morphism. Show that it is an isomorphism.

EXERCISE 6.6. Prove that the map $\pi_1(X, p) \to \pi_1(X_1, p_1) \times \pi_1(X_2, p_2)$ from Lemma 6.3 is an isomorphism.

7. Fundamental groups of orientable surfaces

Recall that any triangulated surface of genus g has been split into two closed sets U, U' intersecting along their common boundary circle, with U a topological disk and U' a neighborhood of the dual graph. It follows from the definition that U' can be continuously retracted onto the graph Γ'. Therefore (Exercise 7.4) they have the same fundamental groups.

LEMMA 7.1. *Every graph Γ' of Euler characteristic $1 - n$ has the same homotopy type as a bouquet of n circles. Moreover $\pi_1(\Gamma')$ is a free group with n (non-commuting) generators.*

Since $\chi(\Gamma') = \chi(S) + 1 = 1 - 2g$, it follows that Γ' is (homotopically) a bouquet of 2g circles. Thus $\pi_1(U') = F^{2g}$, the free group on $2g$ generators. We claim that the map from $\pi_1(\partial'_U)$ to $\pi_1(U')$ takes the generator of $\pi_1(S^1)$ into $a_1 b_1 a_1^{-1} b_1^{-1} \dots a_g b_g a_g^{-1} b_g^{-1}$.

Since $S = U \cap U'$ we can attempt to compute $\pi_1(S)$ in terms of the fundamental groups of U, U' and $U \cap U' = S^1$. In practice, it is more convenient to work with the open sets $V := S \setminus \Gamma'$, $V' := S \setminus \Gamma$, of which U, V are deformation retracts.

A group given by generators $\gamma_1, \dots, \gamma_n$ subject to the relations $R_1, \dots, R_m$ means the quotient of the free group F^n by the normal subgroup generated by the words $R_1, \dots, R_m$ (a *relation* is just an element in F^n, which we think of as "simplifying" in the quotient group). For example, $\mathbb{Z}^2$ is the quotient of F^2 by the relation $\gamma_1\gamma_2\gamma_1^{-1}\gamma_2^{-1}$. We write

$$\langle \gamma_1, \dots, \gamma_n; R_1, \dots, R_m \rangle$$

for the resulting group.

Let $W := V \cap V'$ and choose $p \in W$. Let

$$G := \pi_1(V, p), \qquad G' := \pi_1(V', p), \qquad G_0 := \pi_1(W, p).$$

If

$$\begin{aligned} G =& \langle \gamma_1, \dots, \gamma_n; R_1, \dots, R_m \rangle, \\ G' =& \langle \gamma'_1, \dots, \gamma'_{n'}; R'_1, \dots, R'_{m'} \rangle, \\ G_0 =& \langle \gamma^0_1, \dots, \gamma^0_n; R^0_1, \dots, R^0_{m_0} \rangle \end{aligned}$$

are presentations of G, G', G_0 with generators and relations, define a new group H with generators $\gamma_1, \dots, \gamma_n, \gamma'_1, \dots, \gamma'_{n'}$ as follows: the inclusion maps $\imath_V, \imath_{V'}$ define elements $\imath_V(\gamma^0_j) \in G$ and $\imath_{V'}(\gamma^0_j) \in G'$. Then the relations in H are $R_1, \dots, R_m, R'_1, \dots, R'_{m'}$, and $\imath_V(\gamma^0_j)\imath_{V'}(\gamma^0_j)^{-1}$. The *Van Kampen* theorem states that $\pi_1(S)$ is the group H described above. In particular, H is well-defined regardless of choices of generators and relations! The above holds true whenever V, V' are an open cover of S, and V, V' and $W = V \cap V$" are path-connected.

This result allows us to compute easily the fundamental group of a bouquet of circles, but also of a surface of genus g. Since V is contractible it does not contribute any generators or relations. We have seen that V' has $2g$ free generators. The fundamental group of the intersection has one generator, introducing an extra relation:

THEOREM 7.2. *Let S be an orientable surface of Euler characteristic $2g$. Then*

$$\pi_1(S) = \langle a_1, b_1, \dots, a_g, b_g; a_1 b_1 a_1^{-1} b_1^{-1} \dots a_g b_g a_g^{-1} b_g^{-1} \rangle.$$

As a corollary, the rank of the abelianisation is $2g$. The abelianisation of a group G is its quotient by the normal subgroup spanned by commutators, i.e., the subgroup consisting of products of elements of the form $aba^{-1}b^{-1} \in G$. Homeomorphic surfaces have thus isomorphic abelianisations of their fundamental groups. Therefore orientable surfaces of different genera are not homeomorphic, thus completing the classification theorem of orientable surfaces.

REMARK 7.3. Along the same lines, we can classify non-orientable surfaces.

7.1. Exercises.

EXERCISE 7.4. A subspace $X \subset Y$ is called a *deformation retract* of Y if the identity map of Y is homotopic relative to X to a map into X. Construct a deformation retract from U' onto Γ'.

EXERCISE 7.5. If $X \subset Y$ is a deformation retract, show that the map $\pi_1(X) \to \pi_1(Y)$ induced from the inclusion is an isomorphism.

CHAPTER 2

Riemannian geometry of surfaces

1. Smooth structures

From now on we take S to be a compact, orientable surface of genus g. Let us show it has a smooth structure. Take a triangulation, and consider the charts defined on the interiors of the faces by the inverses of the triangulation homeomorphisms, say $\Phi_T : T^\circ \to \Delta_2$. For each edge E bordering two faces T_1 and T_2 construct a chart with domain $T_1^\circ \cup T_2^\circ \cup E^\circ$ as follows: take affine transformations A_1, A_2 from Δ_2 into the upper, respectively the lower half of a diamond shape, mapping the images of the edge E onto the diagonal. Then $A_1 \circ \Phi_{T_1}$ and $A_2 \circ \Phi_{T_2}$ glue nicely, providing a chart near interior points of E. Near a vertex of multiplicity n, map each triangle afinely onto a triangle centered into the origin with angle $2\pi/n$. We leave it to the reader to convince himself that these charts form a smooth atlas.

Whenever we have a smooth atlas, it is very easy to construct many other smooth structures by composing the charts in the original atlas with an arbitrary homeomorphism. We call the induced smooth structures equivalent.

PROBLEM 1.1. *What can one say about the set of equivalence classes of smooth structures on a surface of genus g?*

How about complex structures? The smooth atlas constructed above does not have holomorphic changes of coordinates. However, it is not difficult to construct another atlas which is holomorphic. The only modification needed is near the vertices. A neighborhood of a vertex is made of a union of n equilateral triangles. There exists a homeomorphism from this union into an open neighborhood of 0 in $\mathbb{C}$, given essentially by $z \mapsto z^{6/n}$. Let us illustrate first the case $n = 6$. In this case, the 6 triangles fit precisely into $\mathbb{C}$, with vertices in 0 and two consecutive roots of order 6 of the unit. In general, the function $z^{n/6} := e^{\frac{6 \log z}{n}}$ is well-defined by starting on the first triangle with the standard cut of the $\log$ function, and continuing it analytically on subsequent triangles. It will agree on the first, resp. the last edge, since the “total angle” around the singularity is $2\pi n/6$.

If we start with a different triangulation, we may get different holomorphic structures.

PROBLEM 1.2. *What can one say about the set of equivalence classes of holomorphic structures on a oriented surface of genus g?*

The rest of this book is about answering these two questions. It turns out that smooth structures on surfaces are unique up to homeomorphisms. In every genus g, the *moduli space* of complex structures on Σ_g modulo homeomorphisms is a manifold, which has itself a complex structure.

On S^2 the complex structure is unique (again, up to homeomorphism) thus the moduli space is just a point. On the torus, the moduli space is the quotient $\mathrm{PSL}_2(\mathbb{R})\backslash\mathbb{H}$, the modular curve. In genus $g \geq 2$ we get the celebrated *Teichmüller space*, a Kähler manifold of complex dimension $3g-3$. Although topologically a ball, $\mathcal{T}_g$ has a rich geometry and is perhaps one of the most interesting objects in mathematics.

From now on a smooth structure on S is assumed fixed, at some point we will prove it is unique.

2. Review of differential geometry

2.1. Functions. Let S be a smooth surface, i.e., endowed with a fixed smooth atlas.

DEFINITION 2.1. A function $f : S \to \mathbb{R}$ is called *smooth* at $p \in S$ if there exists a chart $\phi_A : A \to A' \subset \mathbb{R}^2$ with $p \in A$ such that $f \circ \phi_A^{-1} : A' \to \mathbb{R}$ is smooth at $\phi_A(p)$. The function f is called smooth on S if it is smooth at every point of S.

Smooth functions at p form an algebra over $\mathbb{R}$, similarly smooth functions on S form an algebra denoted $C^\infty(S)$. The *support* of a function $f \in C^\infty(S)$ is the closure of the set $\{p \in S; f(p) \neq 0\}$. Let $C_c^\infty(S)$ denote the sub-algebra of functions with compact support.

Since every open subset of S is a surface, we get an algebra of smooth functions $C^\infty(U)$ for every open set $U \subset S$, with restriction maps $C^\infty(U') \to C^\infty(U)$ whenever $U \subset U'$. Conversely, extension by 0 defines a map $C_c^\infty(U) \to C_c^\infty(U')$.

LEMMA 2.2. *There exists a smooth function $\mu : \mathbb{C} \to \mathbb{R}_+ = [0,\infty)$ satisfying $\mu(z) = 1$ for $|z| \leq 1/2$, $\mu(z) > 0$ for $|z| < 1$ and $\mu(z) = 0$ whenever $|z| \geq 1$.*

PROOF. We use the following smooth functions:

$$\begin{aligned}
\mu_0(r) &:= \begin{cases} e^{-1/x} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0; \end{cases} \\
\mu_1(r) &:= \mu_0(r)\mu_0(1-r); \\
\mu_2(r) &:= \int_{-\infty}^{r} \mu_1(s)ds; \\
\mu_3(r) &:= \mu_2(r)/\mu_2(1); \\
\mu_4(r) &:= \mu_3(1-2r); \\
\mu(z) &:= \mu_4(|z|).
\end{aligned} \tag{2.1}$$

□

This lemma implies that the restriction map $C^\infty(U) \to C^\infty(S)$ is neither injective nor surjective (Exercise 3.9). At the same time, for any $K \subset U$ compact, there exists $\phi \in C_c^\infty(U)$ with $\phi_{|K} = 1$. To see this, cover each point p of $K \subset U$ by the domain of a chart whose image is the disc of radius 2, centered at p. The preimages A_j of the unit open disk also cover K, then extract a finite subcover, i.e., $K \subset \bigcup_{j=1}^m A_j \subset U$. On each disc A_j consider the function $\psi_j := \mu \circ \phi_{A_j}$; it extends to a smooth function on U with compact support. Set $\psi = \sum_{j=1}^m \psi_j$. From the properties of μ in Lemma 2.2 the function ψ is nowhere zero on the compact K, let C be its strictly positive infimum. Define

$$\phi := (1 - \mu_4) \circ (\psi / C)$$

where μ_4 was constructed in (2.1), and then check easily that it has the desired properties.

The existence of ϕ implies that

$$C_c^\infty(U) \to C^\infty(S), \qquad f \mapsto f\phi$$

defines an extension of $f_{|K}$ to S.

Another consequence of the existence of μ is the so-called partition of unity. Let $S = \bigcup_{j=1}^\infty A_j$ be a locally finite open cover (Exercise 3.10) and $S = \bigcup_{j=1}^\infty K_j$ a compact cover with $K_j \subset A_j$. Then there exist smooth non-negative functions $\phi_j \in C^\infty(S)$ with support in A_j such that

$$\sum_{j=1}^\infty \phi_j = 1.$$

Here the sum is finite at every point by the assumption on the open cover. It is enough to define

$$\phi_j := \frac{\mu_j}{\sum_{j=1}^\infty \mu_j}$$

where $\mu_j \in C_c^\infty(A_j)$ is 1 on K_j. The sum $\sum_{j=1}^\infty \mu_j$ is finite in some neighborhood of every point, hence smooth. Since the K_j's form a cover of S, this sum is greater than or equal to 1, in particular it never vanishes.

3. Vectors

A *vector* at $p \in S$ is a derivation $v_p : C^\infty(S) \to \mathbb{R}$, i.e., a linear map such that

$$v_p(fg) = v_p(f)g(p) + f(p)v_p(g). \tag{3.1}$$

The set of vectors at p is a vector space, denoted T_pS. A *vector field* is a derivation $v : C^\infty(S) \to C^\infty(S)$, i.e., a linear map such that

$$v(fg) = v(f)g + fv(g). \tag{3.2}$$

The set $\mathcal{V}(S)$ is also a real vector space. By evaluating in p, each vector field gives a vector at p for all $p \in S$ so we get a linear map from $\mathcal{V}(S)$ to T_pS.

EXAMPLE 3.1. Let $p \in S$ and $c : I \to S$ a curve starting in p. Then $f \mapsto \frac{df(c(t))}{dt}_{|t=0}$ defines a vector at p, denoted $\dot{c}$. In fact every vector at p is of this form, as we shall see below.

EXAMPLE 3.2. On $\mathbb{R}^2$, the maps $f \mapsto \frac{\partial f}{\partial x}$, $f \mapsto \frac{\partial f}{\partial y}$ define vector fields, denoted ∂_x, ∂_y.

THEOREM 3.3. *Let S be a smooth surface.*

(1) *There is a natural identification of T_pS with the set $\mathrm{Der}_p(C_c^\infty(S))$ of linear maps from $C_c^\infty(S)$ to $\mathbb{R}$ satisfying* (3.1).
(2) *Every vector field preserves the support:*

$$\mathrm{supp}(v(f)) \subset \mathrm{supp}(f).$$

(3) *Every linear map from $C_c^\infty(S)$ to $C^\infty(S)$ satisfying* (3.2) *takes values in $C_c^\infty(S)$ and preserves support.*
(4) *There is a natural identification of $\mathcal{V}(S)$ with the set of derivations on $C_c^\infty(S)$.*

PROOF. We define a linear map $R : T_p \to \mathrm{Der}_p(C_c^\infty(S))$ by restricting the action of v_p to compactly supported functions. To show that it is an isomorphism, we start with a lemma about vectors.

LEMMA 3.4. *Let f be a function which equals 1 in a neighborhood K of p. Then for every vector $v_p \in T_pS$ we have $v_p(f) = 0$. If f has compact support, then $V_p(f) = 0$ for all $V_p \in \mathrm{Der}_P(C_c^\infty(S))$.*

PROOF. Take a function μ which is 1 near p and has support in K, thus $f\mu = \mu$. From the Leibnitz identity,

$$v_p(\mu) = v_p(f\mu) = v_p(f)\mu(p) + f(p)v_p(\mu) = v_p(f) + v_p(\mu) \Rightarrow v_p(f) = 0. \tag{3.3}$$

The same proof applies to the second statement. □

This holds in particular for the constant function 1. By linearity, $v_p(1-f) = 0$. Again from the Leibnitz identity, for every $g \in C^\infty(S)$ we have $v_p(g) = v_p(fg)$, and if f, g have compact support then $V_p(g) = V_p(fg)$.

We now show that R is injective. Take v_p such that $R(v_p) = 0$, i.e., v_p applied to every compactly-supported function is 0. We want to show $v_p = 0$. Take a compactly supported f as in Lemma 3.4. Then for every $g \in C^\infty(S)$ we have $v_p(g) = v_p(fg) = 0$ since fg has compact support.

Take now $V_p \in \mathrm{Der}_p(C_c^\infty(S))$ and define $v_p(g) := V_p(fg)$ for f as above. The map associating v_p to V_p is clearly linear, we claim that it satisfies Leibnitz rule. Indeed, $V_p(fg) = V_p(f^2g)$ by Lemma 3.4 so

$$v_p(gg') = V_p(fg \cdot fg') = V_p(fg)(fg')(p) + (fg)(p)V_p(fg') = v_p(g)g'(p) + g(p)v_p(g'),$$

in other words v_p is a vector. Therefore R is surjective, which finishes the proof of the first claim of the theorem.

To prove that a vector v preserves support, let $f \in C^\infty(S)$ with support C and $p \notin C$. Since S is a separate surface, there exists a compact neighborhood K of p inside an open disk U disjoint from C. Let $\mu \in C_c^\infty(S)$ be with support in U and identically 1 on K. From Lemma 3.4, the function $v(\mu)$ is identically zero on K. From $\mu f = 0$ we deduce as in (3.3) that $\mu v(f)$ is identically zero on K, thus $v(f)$ vanishes on K. Since p was arbitrary outside the support C of f if follows that the support of $v(f)$ is contained in C.

For the third statement, note that the above proof still holds for derivations from $C_c^\infty(S)$ to $C^\infty(S)$.

For the final statement, we already know that derivations on $C^\infty(S)$ define derivations on $C_c^\infty(S)$ by restricting the action of a vector field to compactly-supported functions. To show this map is injective, assume v is a vector whose restriction to $C_c^\infty(S)$ is 0. Take $p \in S$ and μ a function with compact support which is 1 in a neighborhood K of p. For every $f \in C^\infty(S)$, $v(f)$ and $v(\mu f)$ agree on K, and since the latter has compact support it follows that $v(f)$ is zero near p; since p

was arbitrary, $v(f) = 0$ as claimed. To show surjectivity, proceed as for the first statement of the theorem. Let V be a derivation on $C_c^\infty(S)$. Take $f \in C^\infty(S)$. At every point p define $v(f)(p)$ by

$$v(f)(p) := v(\mu f)(p)$$

for some $\mu \in C_c^\infty(S)$ which is 1 near p. This defines a function $v(f)$ which is independent of the choices of μ. As in the first part of the theorem, $f \mapsto v(f)$ is a derivation. □

COROLLARY 3.5. *For $U \subset S$, the inclusion $C_c^\infty(U) \subset C_c^\infty(S)$ induces a restriction map $\mathcal{V}(S) \to \mathcal{V}(U)$. For $p \in U$, we get a map $T_pU \to T_pS$ which is an isomorphism.*

PROOF. The first statement is clear. For the second, it is also clear how the map is defined (by restriction to $C_c^\infty(U)$). To show this is isomorphism, take a function $f \in C^\infty(S)$ as in Lemma 3.4 with support in U which equals 1 near p. Then every vector $v_p \in T_pS$ is determined by its action on functions with support in U, because $v_p(g) = v_p(fg)$, proving injectivity. For surjectivity define a derivation $C^\infty(S) \to \mathbb{R}$ by $g \mapsto v_p(fg)$ for f as above and $v_p \in T_pU$. □

The real vector space $\mathcal{V}(S)$ has a natural action of the ring $C^\infty(S)$ defined by

$$(fv)(g) := f \cdot v(g).$$

This makes $\mathcal{V}(S)$ into a $C^\infty(S)$-module (Exercise 3.11).

3.1. Local structure of $C^\infty(S)$. Since every surface is locally diffeomorphic to $\mathbb{R}^2$, let us study $C^\infty(\mathbb{R}^2)$ in some detail. For every $z_0 \in \mathbb{R}^2$, the space $\mathcal{I}_{z_0}$ of functions vanishing at $z_0 = (x_0, y_0)$ is a maximal ideal in $C^\infty(S)$. We claim that $\mathcal{I}_{z_0}$ is generated by the functions $x-x_0, y-y_0$ as a $C^\infty(S)$-module. Let f be a function with $f(z_0) = 0$. For $z \in \mathbb{R}^2$ set

$$F : I \to \mathbb{R}, \qquad F(t) := f((1-t)z_0 + tz)$$

and write

$$\begin{aligned} f(z) =& F(1) - F(0) \\ =& \int_0^1 F'(t)dt \\ =& \int_0^1 [\partial_x f((1-t)z_0 + tz)(x - x_0) + \partial_y f((1-t)z_0 + tz)(y - y_0)]dt \\ =& (x - x_0)\alpha + (y - y_0)\beta \end{aligned}$$

where the last equality defines α, β in an obvious way. Note that α, β are smooth functions of z. Also note that the generators $x - x_0, y - y_0$ are clearly not free in $\mathcal{I}_{z_0}$ since $(x - x_0)(y - y_0) - (y - y_0)(x - x_0) = 0$.

PROPOSITION 3.6. *The $C^\infty(\mathbb{R}^2)$-module $\mathcal{V}(\mathbb{R}^2)$ is free, with basis $\{\partial_x, \partial_y\}$.*

PROOF. Take $v \in \mathcal{V}(\mathbb{R}^2)$ and define functions $a, b \in C^\infty(\mathbb{R}^2)$ using the coordinate functions x, y:

$$a := v(x) = v(x - x_0), \qquad b := v(y) = v(y - y_0)$$

where x_0, y_0 are any constants (recall that $v(1) = 0$ and so by linearity $v(c) = 0$ for every constant function c). For every $f \in C^\infty(\mathbb{R}^2)$ compute $v(f)(z_0)$ using Taylor series:

$$\begin{aligned} v(f)(z_0) =& v(\alpha(x - x_0) + \beta(x - x_0))(z_0) \\ =& a\alpha + b\beta \\ =& (a\partial_x + b\partial_y)(f)(z_0) \end{aligned}$$

proving that $v = a\partial_x + b\partial_y$ in the point z_0. Since z_0 was arbitrary, we see that ∂_x, ∂_y span $\mathcal{V}(\mathbb{R}^2)$. They are clearly independent (Exercise 3.12). □

In the same way, one proves that for every $p \in \mathbb{R}^2$, $\{(\partial_x)_p, (\partial_y)_p\}$ form a basis in $T_p\mathbb{R}^2$, which therefore has dimension 2.

3.2. Exercises.

EXERCISE 3.7. Prove that in the context of Definition 2.1, for any other choice of chart ϕ_B with $p \in B$, the function $f \circ \phi_B^{-1} : B' \to \mathbb{R}$ is smooth at $\phi_B(p)$.

EXERCISE 3.8. Prove that the function f constructed in Lemma 2.2 is indeed smooth.

EXERCISE 3.9. Show that for an open subset $U \subsetneq S$, the restriction map $C^\infty(S) \to C^\infty(U)$ is not surjective. If U is not dense, then the restriction map is not injective.

EXERCISE 3.10. A cover $S = \bigcup_{j=1}^\infty A_j$ is called *locally finite* if every point has a neighborhood which intersects only a finite number of A_j's. Show that it is equivalent to ask that every compact in S intersects only a finite number of A_j's.

EXERCISE 3.11. Prove that $\mathcal{V}(S)$ is a $C^\infty(S)$-module.

EXERCISE 3.12. Show that ∂_x, ∂_y are independent as generators of the $C^\infty(\mathbb{R}^2)$-module $\mathcal{V}(\mathbb{R}^2)$.

EXERCISE 3.13. Show that the commutator $[u, v]$ of two derivations on an algebra, defined by $[u, v](f) := u(v(f)) - v(u(f))$, is again a derivation. When u, v are vector fields on S, the new vector field $[u, v] \in \mathcal{V}(S)$ is called the *Lie bracket* of u, v.

EXERCISE 3.14. Let S be a compact surface. Show that every maximal ideal in $C^\infty(S)$ is of the form $\mathcal{I}_p$ for some $p \in S$.

EXERCISE 3.15. For every $p \in S$, $\mathcal{I}_p\mathcal{V}(S)$ is a submodule of $\mathcal{V}(S)$, and

$$0 \mapsto \mathcal{I}_p\mathcal{V}(S) \hookrightarrow \mathcal{V}(S) \to T_pS \to 0$$

is a short exact sequence.

EXERCISE 3.16. Prove that every vector at $p \in S$ is obtained from a curve passing through p as in Example 3.1. If (x, y) is a chart near p, compute the vector $\dot{c}$ in the basis ∂_x, ∂_y at p.

4. Forms and tensors

4.1. Change of coordinates. A map $\Phi : S \to S'$ is called *smooth* if $\Phi^* C^\infty(S') \subset C^\infty(S)$, where we define

$$(\Phi^* f')(z) := f'(\Phi(z)).$$

LEMMA 4.1. *A smooth map induces linear maps between T_pS and $T_{\Phi(p)}S'$, for all $p \in S$.*

PROOF. Simply define $\Phi_* V(f') := V(f' \circ \Phi)$. □

Let $\Phi : S \to S'$ be a diffeomorphism and V a vector field on S. Define another vector field $\Phi_* V$ on S' by $\Phi_* V(f') = V(\Phi^* f')$, or more precisely

$$\Phi_* V(f')(p') := V(f' \circ \Phi)(\Phi^{-1}(p')).$$

This is easily seen to be a derivation on $C^\infty(S')$.

EXAMPLE 4.2. Let $\Phi = \Phi_1(x, y), \Phi_2(x, y) : U' \to V'$ be a diffeomorphism between open subsets of $\mathbb{R}^2$. Then in the bases ∂_x, ∂_y, the map $\Phi_* : \mathcal{V}(U') \to \mathcal{V}(V')$ takes the form

$$\Phi_* = D\Phi = \begin{bmatrix} \partial_x \Phi_1 & \partial_y \Phi_1 \\ \partial_x \Phi_2 & \partial_y \Phi_2 \end{bmatrix}.$$

Indeed, from the definition and the chain rule,

$$(\Phi_* \partial_x)(f) = \partial_x (f \circ \Phi) = \frac{\partial \Phi_1}{\partial x} \partial_x f + \frac{\partial \Phi_2}{\partial x} \partial_y f$$

so $\Phi_*(\partial_x) = (\partial_x \Phi_1)\partial_x + (\partial_x \Phi_2)\partial_y$ and similarly $\Phi_*(\partial_y) = (\partial_y \Phi_1)\partial_x + (\partial_y \Phi_2)\partial_y$.

The *tangent bundle* of a surface S is the disjoint union

$$TS := \bigsqcup_{p \in S} T_pS.$$

It is endowed with a natural function $\pi : TS \to S$ which associates to a vector $v_p \in T_pS$ its base-point $p \in S$. If $U \subset S$ is the domain of a chart $\phi_U : U \to \mathbb{R}^2$, the maps $(\phi_U)_* : T_pS \to T_{\phi_U(p)}\mathbb{R}^2$ define a bijection

$$(\phi_U)_* : TU \to T\mathbb{R}^2.$$

We have seen that $T\mathbb{R}^2$ is canonically bijective with $\mathbb{R}^2 \times \mathbb{R}^2$ via the map

$$T_p\mathbb{R}^2 \ni v_p \mapsto (p, a, b)$$

where $v_p = a\partial_x + b\partial_y$, so $T\mathbb{R}^2$ has a canonical topology. We define on TS the following topology: we require that for all charts U, the map $(\phi_U)_* : TU \to T\mathbb{R}^2$ is a homeomorphism.

We need to check that this definition is compatible with changes of charts. Let ϕ_V be another chart, and $\Phi : U' \to V'$ the composition $\phi_V \circ \phi_U^{-1}$. Then $(\phi_V)_*(\phi_U)_*^{-1}$ maps (p, v) to $(\Phi(p), D\Phi(v))$ which is clearly continuous.

It follows that TS is a topological manifold of dimension 4. Moreover, the changes of charts on TS are clearly smooth maps, thus TS is a smooth 4-manifold. Furthermore, each fiber $\pi^{-1}(p)$ is mapped by $(\phi_U)_*$ *linearly* onto $\mathbb{R}^2$.

DEFINITION 4.3. A *real vector bundle* of rank k over a surface S is a smooth manifold E of dimension $2+k$, together with a map $\pi : E \to S$ and a smooth atlas consisting of open sets $\pi^{-1}(U)$ (where U is the domain of a chart $\phi_U : U \to U'$ on S) and homeomorphisms $\Phi_U : \pi^{-1}(U) \to U' \times \mathbb{R}^k$, such that the (smooth) changes of charts $\Phi_V \Phi_U^{-1}$ are real linear in the second variable. A *complex vector bundle* of dimension k is a real vector bundle of dimension $2k$ where we require the changes of charts to be complex linear, after identifying $\mathbb{R}^{2k}$ and $\mathbb{C}^k$.

The crucial example of a vector bundle is the trivial vector bundle

$$\underline{\mathbb{R}^k} := S \times \mathbb{R}^k$$

with the projection onto the first factor.

Any chart in the above atlas is called a *trivialization* of E over U. We will mostly be interested in real plane bundles, i.e., real vector bundles of dimension 2, and in complex vector bundles of dimension 1, also called complex line bundles.

REMARK 4.4. The fibers of a vector bundle have a well-defined structure of vector space, induced by pulling back the linear structure from $\mathbb{R}^k$ (or $\mathbb{C}^k$) via any local trivialization map. This structure (i.e., the origin, the rules for addition and multiplication by scalars) is independent of the chosen local trivialization.

A *section* in a vector bundle $E \to S$ is a smooth function $s : S \to E$ such that $\pi(s(p)) = p$ for all $p \in S$. The set of sections, denoted $C^\infty(S, E)$, is a vector space and has a natural $C^\infty(S)$-module structure.

LEMMA 4.5. $C^\infty(S, TS) = \mathcal{V}(S)$.

The algebra $C^\infty(S)$ has a canonical structure of module over itself.

LEMMA 4.6. *Let $F : C^\infty(S) \to C^\infty(S)$ be a $C^\infty(S)$-linear map. There exists a unique $f \in C^\infty(S)$ with $F(g) = fg$, for all $g \in C^\infty(S)$.*

PROOF. For $p \in S$ we set $f(p) := F(1_S)(p)$. The function $f = F(1_S)$ is smooth. For every $g \in C^\infty(S)$ we have $g = g \cdot 1_S$ so $F(g) = F(g \cdot 1_S) = gF(1_S) = fg$. □

As a corollary, a $C^\infty(S)$-linear map induces $C^\infty(U)$-linear maps from $C^\infty(U)$ to itself, for every open set U in S.

Let U, V be $\mathbb{K}$ vector spaces of finite dimension, where $\mathbb{K}$ is the field $\mathbb{C}$ or $\mathbb{R}$. The *dual* of U, denoted U^*, is the space of linear functionals on U:

$$U^* := \{\alpha : U \to \mathbb{K}; \alpha \text{ is } \mathbb{K}\text{-linear}\}.$$

The space of $\mathbb{K}$-linear maps $L(U, V)$ is also a vector space, of dimension $\dim(U)\dim(V)$. The *tensor product* of U, V (over $\mathbb{K}$) can be defined as

$$U \otimes V := L(U, V^*).$$

If $u_1, \ldots, u_n, v_1, \ldots, v_m$ are bases in U, V then $u_j \otimes v_k$ form a basis in $U \otimes V$. The tensor product is associative.

A bi-linear map $\alpha : U \times V$ to $\mathbb{K}$ induces a linear map from $U \otimes V$ to $\mathbb{K}$ by $u \otimes v \mapsto \phi(u, v)$. A bilinear map $\phi : U \times U \to \mathbb{K}$ is called *alternate* if for all $u, v \in U$ we have

$$\phi(u, v) = -\phi(v, u).$$

4.2. The morphism bundle. Take E^1, E^2 any two real vector bundles of rank k_1, k_2 over S. We can organize the disjoint union of spaces of linear maps $L(E^1_p, E^2_p)$ into a vector bundle of rank $k_1 k_2$, denoted $L(E^1, E^2)$, as follows: take $U_1, U_2 \subset S$ trivializing open sets for E^1, E^2 with trivializations ϕ_1, ϕ_2. Over $U := U_1 \cap U_2$ define the following bijection ϕ_U to be smooth:

$$\bigsqcup_{p \in U} L(E^1_p, E^2_p) \ni \Phi_p \mapsto \phi_U(\Phi_p) := \phi_2 \circ \Phi_p \circ \phi_1^{-1} \in M_{k_1, k_2}(\mathbb{R}).$$

If we choose different trivializations ϕ'_1, ϕ'_2 over U'_1, U'_2, the resulting bijection ϕ_V defined over $U' := U'_1 \cap U'_2$ differs from ϕ_U as follows:

$$\begin{aligned} \phi_{U'}(\Phi_p) =& \phi'_2 \circ \Phi_p \circ (\phi'_1)^{-1} \\ =& (\phi'_2 \phi_2^{-1}) \phi_U(\Phi_p)(\phi_1(\phi'_1)^{-1}). \end{aligned}$$

By setting $A := \phi_U(\Phi_p)$ it follows that $\phi_{U'} \circ \phi_U^{-1}(A) = (\phi'_2 \phi_2^{-1}) A (\phi_1(\phi'_1)^{-1})$ for every $k_1 \times k_2$ matrix A. The changes of trivializations in E_1, E_2

$$\begin{aligned} \psi^{E_1}_{U_1 U'_1} :=& \phi'_1(\phi_1)^{-1} : U_1 \cap V_1 \to M_{k_1}(\mathbb{R}), \\ \psi^{E_2}_{U_2 U'_2} :=& \phi'_2(\phi_2)^{-1} : U_2 \cap V_2 \to M_{k_2}(\mathbb{R}) \end{aligned}$$

are smooth matrix-valued maps defined on open sets in S, it follows that the change of trivializations in $L(E^1, E^2)$ which maps a matrix-valued function A over $U \cap U'$ into

$$\psi_{UU'} A = \psi^{E_2} A (\psi^{E_1})^{-1}$$

are smooth and fiberwise-linear, thus they form a vector bundle atlas.
By analogy with Lemma 4.6, we have:

PROPOSITION 4.7. *The space of $C^\infty(S)$-linear maps from $C^\infty(S, E^1)$ to $C^\infty(S, E^2)$ is canonically isomorphic to $C^\infty(S, L(E^1, E^2))$.*

PROOF. Given $f \in C^\infty(S, L(E^1, E^2))$, define $F : C^\infty(S, E^1) \to C^\infty(S, E^2)$ by $F(s)(p) := f(p)s(p)$.
Conversely, let $F : C^\infty(S, E^1) \to C^\infty(S, E^2)$ be $C^\infty(S)$-linear. We define a section in the endomorphism bundle as follows: for every $s_p \in E^1_p$, choose a section $s : S \to E^1$ with $s(p) = s_p$. Define $f_p(s_p) := F(s)(p)$, which is clearly linear. We must check that f_p is independent of the choice of s. For this we use Lemma 4.8. Since any two extensions s, s' satisfy $(s - s')(p) = 0$, Lemma 4.8 implies by $C^\infty(S)$-linearity $F(s - s')(p) = 0$. $\square$

LEMMA 4.8. *Let $s \in C^\infty(S, E)$ be a section in a vector bundle of rank k over a surface S, vanishing at $p \in S$ (i.e., $s(p) = 0$). Then there exist functions $f_j, j = 1, \ldots, 3$ and sections*

$s_j \in C^\infty(S, E)$ *with*

$$s = \sum_{j-1}^{3} f_j s^j.$$

PROOF. First write $s = \mu s + (1 - \mu)s$ with μ compactly supported, equal to 1 near p. We assume that the support of μ is contained in a domain of chart $\phi : U \to \mathbb{R}^2$ over which E is trivialized, such that $\phi(p) = 0$. We set $f_3 := 1 - mu$, $s^3 := s$. Over U, write $E = U \times R^k$, thus μs can be written as a function-valued k-vector, $\mu s = (u_1, \dots, u_k)$. The functions u_i vanish at p. We have seen in section 3.1 that such a function can be decomposed in terms of the coordinates x, y of the chart ϕ, so $\mu_1 = x s_i^1 + y s_i^2$ for some functions $g_i^j \in C^\infty(U)$. This can be re-written

$$\mu s = x g^1 + y g^2$$

where g^j are sections in E over U. To extend the right-hand side to S, we multiply by the square of a compactly-supported function ν on $U \simeq \mathbb{R}^2$ which is 1 on the support of μ (thus $\nu^2 \mu = \mu$) and we set $f_1 := x\nu$, $f_2 = y, nu$, $s^j := \nu g^j$. □

Over a base of dimension n, we need $n + 1$ terms in the sum from the above lemma.

Similarly, for complex vector bundles E^1, E^2 we construct the bundle of complex-linear morphisms $L_{\mathbb{C}}(E^1, E^2)$.

Using the morphism bundle construction, we define the *dual* of any real vector bundle E as

$$E^* := L(E, \underline{\mathbb{R}}),$$

respectively

$$E^*_{\mathbb{C}} := L_{\mathbb{C}}(E, \underline{\mathbb{C}})$$

when E is complex. In particular, we define the cotangent bundle T^*S as the dual bundle TS^* of TS. It is a vector bundle of rank $2 = \dim(S)$.

If ϕ_u is a local trivialization of E over U, then $(\phi_U^*)^{-1}$ is a trivialization of E^* above the same open set U. The changes of trivializations of E^* are therefore given by $(\psi_{UV}^*)^{-1}$ where ψ_{UV} is the change of trivializations in E over $U \cap V$.

DEFINITION 4.9. A *covector* at $p \in S$ is a linear functional $\alpha_p : T_pS \to \mathbb{R}$. A *differential* 1-*form* is a $C^\infty(S)$-linear map $\alpha : \mathcal{V} \to C^\infty(S)$.

From lemma 4.7, there is a canonical identification between the space of sections $C^\infty(S, T^*S)$ in the cotangent bundle, and the space of differential 1-forms.

Since vector fields act on functions by derivations, there exists a "universal derivation" map

$$d : C^\infty(S) \to C^\infty(S, T^*S)$$

defined by the pairing $(f, v) \mapsto v(f)$. More explicitly, df is the map which sends any vector field v into the function $v(f)$. This universal derivation is called the *de Rham differential*.

4.3. Exercises.

EXERCISE 4.10. Convince yourselves that the tangent bundle is an example of a real plane bundle.

EXERCISE 4.11. Prove that the trivial bundle is a vector bundle in the sense of Definition 4.3.

EXERCISE 4.12. Show that if $s : S \to E^1$, $f : S \to L(E^1, E^2)$ are smooth sections, then fs is also a smooth section in E^2.

EXERCISE 4.13. Let E be a complex vector space with complex structure J. Show that for every $\alpha : E \to \mathbb{R}$ linear, the map

$$E \ni v \mapsto \alpha(v) - i\alpha(Jv) \in \mathbb{C}$$

defines a complex-linear map, thus an element in $E^*_{\mathbb{C}}$. Conversely, for every element $\mu_{\mathbb{C}} \in E^*_{\mathbb{C}}$, the real part of μ belongs to E^*. These two applications are inverse to each other, thus E^* is canonically identified with $E^*_{\mathbb{C}}$.

EXERCISE 4.14. If f is a fixed smooth function, show that the map

$$\mathcal{V}(S) \ni v \mapsto v(f) \in C^\infty(S)$$

is $C^\infty(S)$-linear, thus it defines a 1-form.

EXERCISE 4.15. In $\mathbb{R}^2$, compute the action of dx, dy on the coordinate vectors, where $x, y \in C^\infty(\mathbb{R}^2)$ are the coordinate functions. Compute df in terms of dx, dy for arbitrary f.

5. Conformal structures and complex structures

Let E be a real vector space; define $\Lambda^2(E^*)$ to be the space of bilinear alternate maps on E, i.e.,

$$\Lambda^2(E^*) := \{\alpha : E \times E \to \mathbb{R}; \alpha(u, v) = -\alpha(v, u), \forall u, v \in E\}.$$

LEMMA 5.1. *The dimension of the space of bilinear alternate maps on a vector space of dimension* 2 *is* 1.

PROOF. Choose any basis u_1, u_2 in U. Then every bilinear alternate map α maps $u_1 \otimes u_1, u_2 \otimes u_2$ and $u_1 \otimes u_2 + u_2 \otimes u_1$ to 0. thus α is determined by its value on $u_1 \otimes u_2$. There do exists nonzero such maps, for instance set $\alpha(u_1 \otimes u_1) = \alpha(u_2 \otimes u_2) = 0$ and $\alpha(u_1 \otimes u_2) = -\alpha(u_2 \otimes u_1) = 1 \in \mathbb{K}$. □

An *orientation* on a real plane V is an equivalence class of non-zero bilinear alternate maps under the following equivalence relation: $\alpha \sim \alpha'$ if there exists $c > 0$ with $\alpha = c\alpha'$. There are clearly two orientations on every real plane. For a given orientation, an ordered pair (u, v) of non-colinear vectors is called *positively oriented* if for any representative α of the orientation, $\alpha(u, v) > 0$.

Let $(V, \langle, \rangle)$ be a real vector space of dimension 2 together with a inner product, i.e., a symmetric, bi-linear, positive definite map from $V \times V$ to $\mathbb{R}$. Two vectors X, Y are called *orthogonal* if $\langle X, Y \rangle = 0$. Two inner products on V are *conformally equivalent* if they differ by multiplication with a positive number.

Let V be a real plane. A *conformal structure* on V is a choice of a conformal class. A complex structure on V is a map $\mathbb{C} \times V \to V$ extending the map of multiplication by scalars $\mathbb{R} \times V \to V$, which is real bi-linear and associative: $z_1(z_2 v) = (z_1 z_2)v$.

LEMMA 5.2. *A complex structure on V amounts to the choice of a real endomorphism $J : V \to V$ satisfying $J^2 = -1_V$.*

PROOF. Simply define J as the action of $\mathbb{C} \ni i$ on V. □

THEOREM 5.3. *On a real plane V, complex structures are in a natural bijection to conformal structures together with an orientation.*

PROOF. Although this theorem is quite simple, understanding it will be crucial in the rest of the book. Take first a conformal structure $[g]$ and a representative g, i.e., a scalar product $g : V \times V \to \mathbb{R}$. For any vector v, define Jv as the rotation of v by $90°$ in the sense of the orientation. More precisely, Jv is defined as the unique vector in V of the same length as v, orthogonal to v and such that the basis (v, Jv) is positively oriented. We can construct J in an orthonormal oriented basis (e_1, e_2); for every vector $v = (a, b)$, Jv must be equal to $(-b, a)$ therefore it is clear that the map $v \mapsto Jv$ is linear, with matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and moreover $J^2 = -1$. If we choose another representative kg of the conformal class ($k > 0$) the geometric definition does not change, thus J is well-defined in terms of the conformal class $[g]$.

Conversely, given $J : V \to V$ with $J^2 = -1$, notice first that for every $v \neq 0$ the vectors v and Jv are linearly independent. Indeed, otherwise Jv would be colinear to v, thus v would be an eigenvector of J with eigenvalue $\lambda \in \mathbb{R}$; but then v is also an eigenvalue for J^2 of eigenvalue $\lambda^2 \geq 0$, which contradicts $J^2 = -1$. Choose any vector $v \in V$ and decree that v, Jv are an oriented orthonormal basis of V, thus constructing a scalar product g_v. If we choose another initial vector v', write $v' = av + bJv$, $Jv' = -bv + aJv$. We claim that the metric $g_{v'}$ with respect to which v', Jv' form an orthonormal basis is $(a^2 + b^2)$ times g_v, thus the conformal class of g_v is independent of the choice of v.

The above maps associating a complex structure to a conformal structure and vice-versa are clearly inverse to each-other, thus bijections. □

6. Oriented vector bundles

Let $E \to S$ be a real vector bundle of rank 2. The bundle $\Lambda^2(E^*)$ is defined as the disjoint union of $\Lambda^2(E_p^*)$. If $\phi_U : E_{|U} \to U \times \mathbb{R}^2$ is a local trivialization of E, the bijection

$$\bigsqcup_{p \in U} \Lambda^2(E_p^*) \ni \alpha_p \mapsto (\phi_U^{-1})^* \alpha_p \in \Lambda^2(\mathbb{R}^2)$$

is defined to be a homeomorphism, where

$$(\phi_U^{-1})^* \alpha_p(u, v) := \alpha_p(\phi_U^{-1}(p, u), \phi_U^{-1}(p, v)).$$

If $\phi_{U'}$ is another trivialization in E with $\phi_{U'} \circ \phi_U^{-1} = \psi_{UU'}$, the change of trivializations in $\Lambda^2(E^*)$ is given by

$$\beta \mapsto (\psi_{UU'}^{-1})^* \beta.$$

Here $*$ means pull-back, not adjoint! In this way $\Lambda^2(E^*)$ becomes a real line bundle.

DEFINITION 6.1. A vector bundle $E \to S$ of rank 2 is called *orientable* if the line bundle $\Lambda^2(E^*)$ is trivial.

An orientation in a real vector bundle of rank 2 is an equivalence class of trivializations of $\Lambda^2(E^*)$ modulo the equivalence relation given by multiplication with a strictly positive function. An orientation on E induces orientations in each fiber E_p.

EXAMPLE 6.2 (The volume form). Let S be a surface embedded in $\mathbb{R}^3$. On every tangent plane define a bi-linear form as the length of the vector product:

$$\alpha(U, V) := |U \times V|$$

where

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) := (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

This bi-linear anti-symmetric form is non-zero on every tangent plane, hence it defines a trivialization of the exterior bundle $\Lambda^2(T^*S)$. It follows that every surface embedded in $\mathbb{R}^3$ is orientable.

In general, over an oriented Riemannian surface with metric g, we can define a volume form μ_g by the requirement that $\mu(X, Y) = 1$ for every orthonormal frame with positive orientation.

6.1. Exercises.

EXERCISE 6.3. Show that a line bundle is trivial if and only if it has a non-zero section.

EXERCISE 6.4. Show that a smooth surface constructed from an oriented triangulated surface as in Section 1 is orientable.

EXERCISE 6.5. Show that the volume form μ_g is a smooth section in the bundle $\Lambda^2 T^*S$.

7. Riemannian metrics and almost complex structures

A metric on a vector bundle is a collection of inner products on the fibers, varying smoothly in the base variables.

DEFINITION 7.1. Let $E \to S$ be a real vector bundle over a smooth surface. A *metric* on E is a symmetric, positive-definite smooth section in $E^* \otimes E^*$, i.e., a symmetric bi-linear form which is positive-definite on each fiber E_p, $p \in S$.

A *Riemannian metric* on a surface S is a metric on the tangent bundle of S. For instance, if $S = \mathbb{R}^2$, the canonical metric defined by

$$\langle \partial_{x_i}, \partial_{x_j} \rangle_p := \delta_i^j$$

is an example of Riemannian metric.

If $S \subset \mathbb{R}^3$ is an embedded surface, there is a canonical metric on TS induced from the inclusion $TS \subset \underline{\mathbb{R}}^3$ of vector bundles over S.

In general, if S is compact we can construct Riemannian metrics as follows: for each $p \in S$ there exists a chart $\phi_p : U_p \to D_2$ with image the disk of radius 2. Let V_p be the preimage of the disk of radius 1, and K_p its closure. A finite number of the V_p's, say $V_1, \dots, V_n$, cover S by compactness; let μ_j be a smooth non-negative function with support in U_j which is 1 on K_j.

Let g_j be Riemannian metrics on U_j. Such metrics exist since U_j is diffeomorphic to a disk in $\mathbb{R}^2$, so we could take for instance the pull-back through ϕ_j of the canonical metric on $\mathbb{R}^2$. Then $\mu_j g_j \in C_c^\infty(U_j, T^*U_j \otimes T^*U_j)$ is a symmetric non-negative 2-tensor; moreover,

$$g := \sum \mu_j g_j \in C^\infty(S, T^*S \otimes T^*S)$$

is positive-definite. The same construction works for a *paracompact* manifold, but in this book we focus only on the compact case.

Two metrics g_1, g_2 on E are *conformally equivalent* if there exists a positive function $f \in C^\infty(S, \mathbb{R}_+^*)$ with $g_1 = f g_2$. A *conformal class* is an equivalence class of metrics under the above equivalence relation.

We now want to generalize Theorem 5.3 to bundles. A *complex structure* on a bundle E is an endomorphism J (i.e., a section in the endomorphism bundle $L(E, E)$) with $J^2 = -1_E$.

THEOREM 7.2. *Let $E \to S$ be a rank-2 real vector bundle over a compact surface. Then conformal classes of metrics on E together with an orientation are in bijection with complex structures on E.*

PROOF. Theorem 5.3 is a particular case of this result, for a vector bundle over a point. Thus the bijection is established for the fibers at every point. It remains to show that when we vary the point we obtain smooth objects.

If $[g]$ is a conformal class, we pick a representative g and define $J_{[g]}$ to be the map of rotation by $\pi/2$ with respect to g, in the sense given by the orientation. This does not depend on $g \in [g]$ and clearly $J^2 = -1_E$. To show that J is a smooth section in the endomorphism bundle, we can work locally over an open set $U \subset S$ above which E is trivial. Let $e_1, e_2 : U \to E$ be sections which form a base in E at every point of U. We assume that e_1, e_2 form a positively oriented bases at a point $p_0 \in U$ (and hence by exercise 7.8 at every point $p \in U$), otherwise just re-label them. Using the metric g, we first define $e_1' := e_1 / \|e_1\|_g$, i.e., we normalize e_1. Next we set

$$e_2' := \frac{e_2 - \langle e_2, e_1' \rangle e_1'}{\|e_2 - \langle e_2, e_1' \rangle e_1'\|_g}.$$

Using this procedure (called the Gramm-Schmidt orthonormalization procedure) we obtained a smooth orthonormal basis of sections. Let ${e_1}', e^{2'} : U \to E^*$ be the dual basis, then

$$J = e_2' \otimes e^{1'} - e_1' \otimes e^{2'}$$

is clearly smooth.

In the opposite direction, we fix a complex structure J. Let $U \subset S$ be a trivializing open set, and $e_1 : U \to E$ a nowhere zero section. We can define a metric on $E_{|U}$ by requiring $\{e, Je\}$

to be orthonormal. The conformal class of the metric is independent of the choice of e. The using a finite partition of unity, patch together the metrics to a global metric on E. It is clear that the conformal class of the resulting metric is independent of choices, and that for every $v \in E_p$, $Jv \perp v$. □

The same proof works for paracompact manifolds. If we drop the smoothness requirement, the statement is true for locally trivial vector bundles over a paracompact base.

A smooth map Φ between two surfaces endowed with Riemannian metrics (S, g) and (S', g') is called *conformal* if $\Phi^* g'$ is conformal to g. Note that the definition makes sense if we only specify conformal classes on S, S'. Assume now that S, S have holomorphic structures. These induce complex structures in TS, TS', hence conformal structures and orientations on S, S'.

LEMMA 7.3. *A smooth map between two holomorphic surfaces is holomorphic if and only if it preserves orientations and is conformal with respect to the induced conformal structures.*

PROOF. In local (holomorphic) coordinates we check easily that Φ is holomorphic if and only if $D\Phi$ commutes with J. In turn, this is equivalent to Φ being orientation-preserving and conformal. □

A rephrasing of this lemma is the following description of holomorphic surfaces as conformal surfaces. A smooth map $\phi : \mathbb{C} \to \mathbb{C}$ is called *conformal* if it is conformal for the standard metric on $\mathbb{R}^2$; clearly, by the above lemma an orientation preserving map is conformal if and only if it is holomorphic.

LEMMA 7.4. *A holomorphic structure on a surface S is the same thing as an oriented conformal atlas on S.*

7.1. Exercises.

EXERCISE 7.5. Prove that a metric on E defines an isomorphism of real vector bundles $E \simeq E^*$. Show that in this way you get a metric on E^*. Find its coefficients in a local trivialization of E^* in terms of the coefficients of the initial metric on E in the corresponding trivialization.

EXERCISE 7.6. Show that conformal equivalence is an equivalence relation.

EXERCISE 7.7. Show that two metrics g_1, g_2 are conformally equivalent if and only if there exists a function $h \in C^\infty(S, \mathbb{R})$ with $g_1 = e^{2h} g_2$.

EXERCISE 7.8. Let $e_1, e_2 : S \to E$ be non-zero (i.e., nowhere zero) sections in an oriented vector bundle E above a connected base S. Show that for points $p, p' \in S$, the basis $e_1(p), e_2(p)$ is positively oriented if and only if $e_1(p'), e_2(p')$ is positively oriented.

EXERCISE 7.9. Let E be a 3-dimensional vector space with a scalar product and $\{e_1, e_2, e_3\}$ any basis. Construct explicitly an orthonormal basis $\{e'_1, e'_2, e'_3\}$ such that e_1 and e'_1 span the same line, and $\{e_1, e_2\}$ and $\{e'_1, e'_2\}$ span the same plane.

CHAPTER 3

Uniformization of surfaces

Let S be a compact oriented smooth surface. We have seen that it admits Riemannian metrics, and that each conformal class corresponds to a unique complex structure in the tangent bundle, also called an almost complex structure on S.

We want to find in each conformal class a "nice" representative. This will be a metric with constant scalar curvature. Then the almost complex structure will turn out to be integrable, i.e., it comes from a holomorphic structure on S.

1. Connections and curvature

Let $\underline{\mathbb{R}}$ be the trivial line bundle on S. Let us denote by $\nabla_V : C^\infty(S, \underline{\mathbb{R}}) \to C^\infty(S, \underline{\mathbb{R}})$ the action by derivations of V on $C^\infty(S, \underline{\mathbb{R}}) = C^\infty(S)$. It satisfies

$$\nabla_{fV} = f\nabla_V, \qquad \nabla_V \circ f = f\nabla_V + V(f). \tag{1.1}$$

DEFINITION 1.1. Let E be a vector bundle on S. A linear action of $\mathcal{V}(S)$ on $C^\infty(S, E)$ satisfying (1.1) is called a *connection* in E.

Like vector fields, connexions can be restricted over open subsets of S. If $\nabla^E, \nabla^{E'}$ are connections in E, E', we define connections in $E \oplus E'$ and in $E \otimes E'$ by

$$\nabla_V^{E\oplus E'}(s \oplus s') := \nabla_V^E s \oplus \nabla_V^{E'} s',$$
$$\nabla_V^{E\otimes E'}(s \otimes s') := \nabla_V^E s \otimes s' + s \otimes \nabla_V^{E'} s'.$$

On the trivial bundle of rank k we have the trivial connection, obtained by applying V to each component function. Using partition of unity, we can show that there exist connections in each vector bundle.

1.1. The Levi-Civita connection.

The main example of connection is the Levi-Civita connection in the tangent bundle. Let (S, g) be a surface with a Riemannian metric. There exists a unique connection in TS satisfying for all $U, V \in \mathcal{V}$

$$\begin{aligned} \nabla_U V - \nabla_V U &= [U, V]; \\ \nabla_U g &= 0. \end{aligned} \tag{1.2}$$

The last identity means explicitly

$$V(g(X, Y)) = g(\nabla_V X, Y) + g(X, \nabla_V Y)$$

The existence of ∇ is easy: we just define $\nabla_X Y$ by the requirement that for every $Z \in \mathcal{V}$,

$$\begin{aligned} 2g(\nabla_X Y, Z) =& Xg(Y,Z) + Yg(X,Z) - Zg(X,Y) \\ &+ g([X,Y],Z) + g([Z,X],Y) + g([Z,Y],X). \end{aligned} \tag{1.3}$$

To prove uniqueness, we deduce (1.3) from (1.2) by some easy manipulations.

1.2. Curvature. For $X, Y \in \mathcal{V}(S)$ and $F \in C^\infty(S, E)$, the expression

$$R_{XY}F := \nabla_X(\nabla_Y F) - \nabla_Y(\nabla_X F) - \nabla_{[X,Y]}F \tag{1.4}$$

is again a section in S. Clearly R depends $\mathbb{R}$-linearly on each of the vector fields X, Y and on the section F, but remarkably it is also $C^\infty(S)$-linear in each of these entries (Exercise 2.9). Thus for vectors X_p, Y_p we get using Lemma 4.8 a well-defined endomorphism of E_p by setting

$$F_p \mapsto (R_{XY}F)(p)$$

where X, Y, S are arbitrary extensions of X_p, Y_p, resp. F_p, to vector fields, respectively a smooth section in S. In this way, R becomes a section in $T^*M \otimes T^*M \otimes E^* \otimes S$. Clearly $R_{XY} = -R_{YX}$ so in fact $R \in C^\infty(S, \Lambda^2 T^*S \otimes \mathrm{End}(E))$.

If E is a complex bundle with complex structure J and the commutator $[\nabla, J]$ vanishes, then R_{XY} commutes with J, thus it is a complex endomorphism of E.

If E is a line bundle (real or complex), the endomorphism bundle $\mathrm{End}(E)$ is canonically trivial of rank 1, a canonical nonzero section being the identity 1_E. Consequently (provided the connection is compatible with J in the complex case) the curvature of such a line bundle is simply a volume form, or a 2-form when the base is of higher dimension.

1.3. Gaussian curvature. When E is the tangent bundle of a surface endowed with a Riemannian metric and X, Y is a orthonormal frame, the quantity $\kappa := \langle R_{XY}Y, X\rangle$, which is a real number associated to each point of S (hence a function), does not depend on the orthonormal frame. This function is called the *Gaussian curvature* of the metric. It is half of the *scalar curvature*, a quantity associated to each Riemannian manifold of arbitrary dimension. For surfaces in $\mathbb{R}^3$ it is related to the mean curvature as follows: the Gaussian is the determinant of the Weingarten map, while the mean curvature is the trace of the same map. Therefore Gaussian curvature for a surface in $\mathbb{R}^3$ is the product of the principal curvatures (the eigenvalues of the Weingarten map).

We can easily compute the Gaussian curvature of the Euclidean plane and of the unit sphere. On $\mathbb{R}^2$ with the standard basis for the tangent bundle $X_1 = \partial_{x_1}$, $X_2 = \partial_{x_2}$, we see immediately that $\nabla_{X_i} X_j = 0$ so the curvature tensor vanishes identically so $\kappa_{\mathbb{R}^2} = 0$. For the sphere we compute the curvature in the stereographic charts, the result is $\kappa_{S^2} = 1$.

1.4. Integration of volume forms. Let $C_c^\infty(S, \Lambda^2 T^*S)$ be the space of compactly supported volume forms on an oriented surface S. There exists a natural functional on this space with real values, $\omega \mapsto \int_S \omega$, with the following properties:

(1) If $\Phi : S \to S'$ is a smooth map, then $\int_S \Phi^*\omega = \int_S \omega$.

(2) If $S \subset \mathbb{R}^2$ with the standard orientation and $\omega = a(x,y)dx \wedge dy$, then $\int_S \omega = \int_{\mathbb{R}^2} a(x,y)dxdy$.

To construct $\int$, assume first that ω has support in a domain of a chart $\phi : U \to \mathbb{R}^2$ from a fixed oriented atlas. Write $\omega = a(x,y)dx \wedge dy$ and define $\int_S \omega = \int_{\mathbb{R}^2} a(x,y)dxdy$. If we consider another chart $\phi' : U \to \mathbb{R}^2$ with coordinate functions x', y', let $\Phi := \phi \circ (\phi')^{-1}$, then $(x,y) = \Phi(x',y')$ so

$$\begin{aligned} dx =& \frac{\partial \Phi_1}{\partial'_x} dx' + \frac{\partial \Phi_1}{\partial'_y} dy' \\ dy =& \frac{\partial \Phi_2}{\partial'_x} dx' + \frac{\partial \Phi_2}{\partial'_y} dy' \end{aligned}$$

therefore

$$dx \wedge dy = \det(D\Phi) dx' \wedge dy'.$$

Thus in the chart ϕ' the form ω takes the form $\omega = a \circ \Phi \det(D\Phi)dx' \wedge dy'$. The formula of change of variable under integration

$$\int_{\mathbb{R}^2} a(x,y)dxdy = \int_{\mathbb{R}^2} a(\Phi(x',y'))|\det(D\Phi)|dx' \wedge dy'$$

together with the fact that the change of charts Φ is positively oriented (thus $\det(D\Phi) > 0$) proves independence of the definition of the integral with respect to the chart.

2. The hyperbolic plane

This will be our fundamental example of Riemannian metric. The underlying surface is the upper half-plane from Example 2.4. The metric we consider is conformal to the standard Euclidean metric on $\mathbb{R}^2$, $g_E = dx^2 + dy^2$, namely

$$g_H = \frac{g_E}{y^2} = \frac{dx^2 + dy^2}{y^2}. \tag{2.1}$$

Explicitly, this means that an orthonormal basis of the tangent space at $(x,y) \in \mathbb{H}^2$ is given by $X := y\partial_x, Y := y\partial_y$. We choose the standard orientation of $T\mathbb{H}^2$, for which the frame (X,Y) is positively oriented. From (the proof of) Theorem 7.2, the almost complex structure on $T\mathbb{H}^2$ corresponding to the conformal class of g_H is given by

$$JX = Y, \qquad\qquad JY = -X.$$

Also, the volume form of the metric g_H satisfies by definition $\mu(X,Y) = 1$. We can check easily that

$$[X,Y] = -X$$

from which we compute the Levi-Civita connection in the frame X, Y using (1.3):

$$\nabla_Y X = \nabla_Y Y = 0, \qquad \nabla_X Y = -X, \qquad \nabla_X X = Y.$$

In particular ∇ commutes with J in the sense that for every vector fields U, V we have $\nabla_U(JV) = J\nabla_U V$. The curvature is then given by $R_{XY}X = Y$, or in other words

$$R_{XY} = J : T\mathbb{H}^2 \to T\mathbb{H}^2.$$

We can re-write this as $R = \mu \otimes J \in C^\infty(\mathbb{H}^2, \Lambda^2 T^*\mathbb{H}^2 \otimes \mathrm{End}(T\mathbb{H}^2))$.
Directly from the definition, the Gaussian curvature of $\mathbb{H}^2$ is the constant function -1.

2.1. Geodesics. Let $c : [a, b] \to S$ be a smooth curve and consider the vectors $\dot{c}_{c(t)}$ tangent to c. If c is injective, we can extend $\dot{c}$ to a vector field V on S.

LEMMA 2.1. *For every $X \in \mathcal{V}(S)$, the vector field $\nabla_V X$ does not depend on the choice of V, and is thus denoted $\nabla_{\dot{c}} X$. If X' is another vector field which agrees with X along c, then $\nabla_{\dot{c}} X' = \nabla_{\dot{c}} X$.*

A vector field satisfying $\nabla_{\dot{c}} X = 0$ is called *parallel* along c. In view of this lemma, we make the following

DEFINITION 2.2. A curve c is called a *geodesic* if $\nabla_{\dot{c}} \dot{c} = 0$.

Since the Levi-Civita connexion is determined (and hence preserved) by the metric, it follows that isometries preserve geodesic curves.

LEMMA 2.3. *For every $z \in S$ and $v \in T_z S$ there exists $\epsilon > 0$ and a geodesic $c : (-\epsilon, \epsilon)$ with $c(0) = z$, $\dot{c}(0) = v$. Moreover, the geodesic $c = c(z, v)$ is unique, and depends smoothly on $(z, v) \in TS$.*

PROOF. Local coordinates + ODE. Choose local coordinates (x_1, x_2) in a neighborhood of z (i.e., a chart $\phi = (x_1, x_2) : U \to \mathbb{R}^2$). Without loss of generality we may assume that $\phi(z) = 0$. The unknown geodesic is thus determined by two unknown functions $x_1(t), x_2(t)$ with $x_1(0) = x_2(0) = 0$. The tangent vector to $c(t) = (x(t), y(t))$ is

$$\dot{c}(t) = x_1'(t)\partial_{x_1} + x_2'(t)\partial_{x_2}.$$

For every $i, j \in \{1, 2\}$, $\nabla_{\partial_{x_i}} \partial_{x_j}$ is a vector field on U; since $\{\partial_{x_1}, \partial_{x_2}\}$ for a basis for local vector fields on U we can write

$$\nabla_{\partial_{x_i}} \partial_{x_j} = \Gamma_{ij}^1 \partial_{x_1} + \Gamma_{ij}^2 \partial_{x_2}$$

for some functions $\Gamma_{ij}^k \in C^\infty(U)$ called the *Christoffel symbols*. These functions depend on the connection and on the local coordinates. With this notation we write

$$\begin{aligned}
\nabla_{\dot{c}} \dot{c} =& \nabla_{\dot{c}}(x_1' \partial_{x_1} + x_2' \partial_{x_2}) \\
=& x_1'' \partial_{x_1} + x_1' \nabla_{\dot{c}} \partial_{x_1} + x_2'' \partial_{x_2} + x_2' \nabla_{\dot{c}} \partial_{x_2} \\
=& x_1'' \partial_{x_1} + x_2'' \partial_{x_2} + \sum_{i,j,k=1}^{2} x_j' x_i' \Gamma_{ij}^k \partial_{x_k}.
\end{aligned}$$

By setting the coefficients of $\partial_{x_1}, \partial_{x_2}$ equal to 0, we deduce that c is a geodesic if and only if the following identities hold:

$$x_1'' + \sum_{i,j=1}^{2} x_j' x_i' \Gamma_{ij}^1 = 0, \qquad\qquad x_2'' + \sum_{i,j=1}^{2} x_j' x_i' \Gamma_{ij}^2 = 0$$

where Γ_{ij}^k is evaluated at the point $(x_1(t), x_2(t))$. Thus $x(t) := (x_1(t), x_2(t))$ is a solution to the second-order differential system

$$x''(t) = V(x'(t), x(t)) \tag{2.2}$$

for $V : \mathbb{R}^4 \to \mathbb{R}^2$, $V_k(X, x) := -\sum_{i,j=1}^{2} \Gamma_{ij}^k(x) X_i X_k$. For such a system, there exists $\epsilon > 0$ and a neighborhood $\mathcal{V} \subset \mathbb{R}^4$ of the initial value $(x(0), x'(0))$ (corresponding to the initial point $z = x(0)$ and direction $v = x'(0)$) such that (2.2) has a unique solution for $t \in (-\epsilon, \epsilon)$ for each initial value in $\mathcal{V}$, depending smoothly on the initial values. □

We have seen that on the hyperbolic plane, the vector field Y is self-parallel in the sense that $\nabla_Y Y = 0$. For every $x \in \mathbb{R}$, the curve

$$c_x : \mathbb{R} \to \mathbb{H}^2, \qquad\qquad c_x(t) = (x, e^t)$$

is an integral curve to Y, i.e., $\dot{c_x}(t) = Y_{c_x(t)}$ for all t. It follows that c_x is a geodesic for all x, moreover $|\dot{c_x}| = 1$. To find other geodesics, use the rich group of oriented isometries $\mathrm{PSL}_2(\mathbb{R})$ of $\mathbb{H}^2$ and Exercise 2.12. For every vector $v \in T_z\mathbb{H}^2$ we find in this way a geodesic starting in z in the direction of v. By uniqueness, every geodesic on $T\mathbb{H}^2$ is of this form.

2.2. Distance. Let $\gamma : [a, b] \to S$ be a smooth curve in a surface endowed with a Riemannian metric. At each $\gamma(t)$, we have a tangent vector $\dot{\gamma}(t) \in T_{\gamma(t)}S$ defined by $f \mapsto \frac{df(\gamma(t))}{dt}$. The *length* of γ is defined as

$$l(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt.$$

Remarkably, $l(\gamma)$ depends only on the image of γ, thus is invariant under reparametrization: If $\phi : [a', b'] \to [a, b]$ is a diffeomorphism, then $l(\gamma \circ \phi) = l(\gamma)$. Indeed, if we set $t = \phi(s)$, we have

$$\dot{\overline{\gamma \circ \phi}}(s) = \phi'(s)\dot{\gamma}(\phi(s)) \tag{2.3}$$

hence

$$\int_{a'}^{b'} \|\dot{\overline{\gamma \circ \phi}}(s)\| ds = \int_{a'}^{b'} |\phi'(s)| \|\dot{\gamma}(\phi(s))\| ds = l(\gamma)$$

using the change of variables formula and the implicit assumption that ϕ is increasing, hence its derivative has positive sign.

The distance between two points p, p' in S is defined as the infimum of the lengths of all curves linking p to p'. This defines a metric on S. For two arbitrary points in S, it may happen that no curve realizes the infimum (example: in $\mathbb{C}^*$, the distance between 1 and -1 is 2, but from the triangle identity every path from 1 to -1 has length strictly larger than 2).

PROPOSITION 2.4. *Let $p, q \in S$ and $c : I \to S$ a curve from p to q of length equal to $d(p,q)$. Then c is a geodesic.*

Here we mean that the image of c can be realized as the image of a parametrized geodesic, or in other words that after a suitable parametrization $t = t(u)$, the curve $c(t(u))$ becomes a geodesic.

PROOF. Variational computation. Let $C : I \times (-\epsilon, \epsilon) \to S$ be a smooth function with $C(t,0) = c(t)$ and $C(0,s) = p$, $C(1,s) = q$ for all $t \in I$, or in other words $c_s := C(\cdot, s)$ is a family of curves from p to q passing through c (to be thought of as a smooth path through c in the space of paths from p to q). The function $s \mapsto l(c_s)$ has a minimum at $s = 0$, in particular

$$\frac{dl(c_s)}{ds}|_{s=0} = 0.$$

For simplicity assume that $\dot{c}(t) \neq 0$ for all $t \in I$, one can always reduce to this case. Let

$$T := C_*(\partial_t), \qquad X := C_*(\partial_s)$$

be vector fields along the image of C. Since ∂_s, ∂_t commute, we deduce $[T, X] = 0$. We then compute

$$\begin{aligned}\partial_s l(c_s) &= \int_0^1 \partial_s \sqrt{\langle T, T\rangle} dt = \int_0^1 \tfrac{\langle \nabla_X T, T\rangle + \langle T, \nabla_X T\rangle}{2|T|} dt \\ &= \int_0^1 \tfrac{\langle \nabla_X T, T\rangle}{|T|} dt = \int_0^1 \tfrac{\langle \nabla_T X, T\rangle}{|T|} dt \\ &= \int_0^1 \left\langle \nabla_T X, \tfrac{T}{|T|} \right\rangle dt = \int_0^1 \partial_t \left\langle X, \tfrac{T}{|T|} \right\rangle dt - \int_0^1 \left\langle X, \nabla_T(\tfrac{T}{|T|}) \right\rangle dt.\end{aligned}$$

Since $C(0,s) = p$ for all s we deduce $\frac{\partial C}{\partial s}(0,s) = 0$ so $X(0) = 0$ and similarly $X(1) = 0$. This means that

$$\int_0^1 \partial_t \langle X, T/|T|\rangle dt = 0,$$

hence if c is a curve of minimizing length then

$$\int_0^1 \left\langle X, \nabla_T(\tfrac{T}{|T|}) \right\rangle dt = 0.$$

Now every vector field X along c satisfying $X(0) = 0, X(1) = 0$ arises from a variation of the curve c as above. By choosing $X := t(1-t)\nabla_T \frac{T}{|T|}$ we see that $\nabla_T \frac{T}{|T|} = 0$ (indeed, if we define $f(t) := \left|\nabla_T \frac{T}{|T|}\right|^2$, then $f(t) \geq 0$ and $\int_0^1 t(1-t)f(t)dt = 0$ so $f(t) = 0$). Recalling that $T = \dot{c}$, this is equivalent to c being a geodesic after a suitable re-parametrization by arc-length. $\square$

Let $c : (a,b) \to S$ be a geodesic and $s : (a', b') \to (a,b)$ a diffeomorphism. Then the re-parametrized curve $\tilde{c} : c \circ s : (a', b') \to S$ is a geodesic if and only if $s(t)$ is an affine transformation, i.e., $s(t) = \alpha t + \beta$. Indeed, set $V := c'(t)$, $\tilde{V} := \tilde{c}'(s) = c'(s(t))s'(t)$. Then

$$\nabla_{\tilde{V}} \tilde{V} = s'(t)^2 \nabla_V V + s'(t)s''(t)V = s'(t)s''(t)V$$

so $\nabla_{\tilde{V}}\tilde{V} = 0$ if and only if $s's'' = 0$, which is equivalent to $(s')^2$ constant, thus $s(t)$ affine.

2.3. The Gauss-Bonnet theorem. Few results in mathematics are more beautiful than the Gauss-Bonnet theorem. This result could well be considered one of the cornerstones of modern mathematics. Its attraction lies in its complexity – it marries in an unprecedented way Analysis, Geometry and Topology – yet its statement shines in deceptive simplicity:

THEOREM 2.5. *Let (S, h) be an orientable closed surface with a Riemannian metric, κ the Gaussian curvature function and μ_h the volume form of h. Then*

$$\int_S \kappa\mu_h = 2\pi\chi(S)$$

where $\chi(S)$ is the Euler characteristic of S, also given by $\chi(S) = 2 - 2g$ where g is the genus (or the number of handles) of S.

PROOF. For the sphere S^2 with the standard metric, we compute directly that the Gaussian curvature equals 1, thus, since the area of the sphere is 4π and the Euler characteristic is 2, we check the Gauss-Bonnet theorem in this case.

For the torus with the standard flat metric (or any flat metric), the curvature and the Euler characteristic are both 0. Together with Proposition 2.6, we have therefore proved the theorem in genus 0 and 1.

PROPOSITION 2.6. *The integral $\int_S \kappa\mu_g$ is independent of the metric g.*

PROOF. The tangent bundle is a complex line bundle, the complex structure being the rotation by angle $\pi/2$ in the positive direction. The Levi-Civita connection preserves this complex structure in the sense that $[\nabla, J] = 0$, in other words

$$\nabla_X(JV) = J\nabla_X V$$

for all vector fields X, V. It follows that ∇ is a connexion on the complex bundle TM. By definition, its curvature is the same as the Riemannian curvature:

$$R^{\mathbb{C}}_{XY}Z = R_{XY}Z$$

for all vectors X, Y, Z. From the known skew-symmetry relation $\langle R_{XY}Z, W\rangle = -\langle R_{XY}W, Z\rangle$, we deduce that R_{XY} is a real multiple of the endomorphism J. We can compute the coefficient from the definition of Gaussian curvature using an orthonormal positively oriented frame (X, Y): $R_{XY}Y = \kappa X$, therefore $R_{XY} = -\kappa J$ or equivalently the curvature is simply $R = -\kappa\mu \otimes J$. Therefore the curvature of the complex bundle TS is

$$R^{\mathbb{C}} = -i\kappa\mu.$$

We are therefore trying to prove that $\int_S R^{\mathbb{C}}$ is independent of the metric.

We prove more generally

LEMMA 2.7. *Let $E \to S$ be a complex line bundle and ∇^t a family of connections. Then the curvature forms R^t differ by exact forms.*

The curvature 2-form R is called the *first Chern form* of E, it is always closed even in dimensions higher than 2, and its cohomology class (which by the above lemma is independent of the connection) is 2π times the so-called *first Chern class* of E.

PROOF. We will prove that $\partial_t R^t$ is exact, then by integration in t we get the desired result. Let U_j be a finite open cover of S by chart domains and $\phi_j : E_{|U_j} \to \mathbb{C} \times U_j$ local trivializations of E. In each trivialization write $\nabla^t = d + \alpha_j^t$ for some complex-valued 1-forms α_j^t defined over U_j. The curvature can be computed locally by $R^t = d\alpha_j^t$ so over U_j,

$$\partial_t R^t = d(\partial_t \alpha_j^t). \tag{2.4}$$

Let $\phi_{ij} := \phi_i \circ \phi_j^{-1} : U_j \cap U_i \to \mathbb{C}^*$ be a change of trivialization. We have

$$\alpha_i^t = \alpha_j^t + \phi_{ij}^{-1} d\phi_{ij}$$

and we notice that the correction term $\phi_{ij}^{-1} d\phi_{ij}$ is independent of t, so

$$\partial_t \alpha_i^t = \partial_t \alpha_j^t.$$

This defines therefore a 1-form $\dot{\alpha}$ on S. By (2.4), $\partial_t R^t = d(\dot{\alpha})$ is exact. □

Take g_0, g_1 two metrics on S. Then the family of tensors $g_t := (1-t)g_0 + tg_1$, $t \in [0,1]$ is made of symmetric, positive definite bilinear forms, hence g_t is a Riemannian metric for all t, interpolating smoothly between g_0 and g_1. They give rise to a family of connections on TS, however as *complex* bundles, the bundle TS with the complex structure J_t changes with t. To address this problem, consider the bundle $TS \to S \times [0,1]$ with complex structure over $S \times \{t\}$ given by J_t. By homotopy invariance, this bundle is isomorphic as complex line bundles to the "constant" bundle $TM \to S \times [0,1]$ with complex structure J_0 at each time t. Let $\tilde{\nabla}^t$ denote the pull-back of ∇^t via this isomorphism. We are now in the setting of Lemma 2.7, therefore the complex curvature forms $\tilde{R}^0$ and $\tilde{R}^1$ differ by an exact form. By Stokes's theorem the integral on a closed surface S of an exact form vanishes. Thus

$$\int_S R^0 = \int_S R^1.$$

Together with the identity $R^t = -i\kappa(g_t)\mu_t \otimes J_t$ we conclude the invariance of the Gauss-Bonnet integral. □

Using this proposition, we can prove the Gauss-Bonnet formula in higher genus as follows[1]: take a torus embedded in $\mathbb{R}^3$ and squeeze it between two parallel planes, so that it acquires two flat regions U^+, U^-, one exterior curved region E and one interior region I. By removing I and replacing it with flat continuations of $U^\pm$ we get a sphere with two flat regions $V^\pm$ and the curved exterior region E. Since the curvature of $V^\pm$ vanishes, we have

$$\int_E \kappa\mu = \int_{S^2} \kappa\mu = 4\pi.$$

[1]This proof used to be taught by Professor Kostake Teleman to second year students at the University of Bucharest in the 1990's

This implies, in view of the direct treatment of the torus case, that

$$\int_I \kappa\mu = \int_{T^2} \kappa\mu - \int_E \kappa\mu = -4\pi. \tag{2.5}$$

Now take the sphere obtained from E and $V^{\pm}$ and expand it in the horizontal direction by a large constant (i.e., apply to it the linear isomorphism of $\mathbb{R}^3$ given by $(z, x_3) \mapsto (Cz, x_3)$ for a large constant C). We still have $\int_{E'} \kappa\mu = 4\pi$ by the same argument as above. In the resulting sphere, which now has two large flat regions, glue back g copies of the inner region I, where g is any genus. We get in this way a surface of genus g (together with a Riemannian metric) with one exterior region E', g inner regions isometric to E and some flat regions.

By (2.5), each of the g inner regions in S contribute -4π towards the Gauss-Bonnet integrand. The exterior region E' contributes by 4π thus in all,

$$\int_S \kappa\mu = g\int_I \kappa\mu + \int_{E'} \kappa\mu = 4\pi g - 4\pi = 2\pi(2-2g). \qquad \square$$

2.4. Isometries of $\mathbb{H}^2$. Let $\mathrm{PSL}_2(\mathbb{R})$ be the group from Example 2.4. We claim that it acts by isometries on $\mathbb{H}^2$. For this, we remark that

$$g_H = \frac{\Re(|dz|^2)}{\Im(z)^2}$$

where $|dz^2| = dz \otimes d\bar{z}$. Therefore, if we take $\mathrm{PSL}_2(\mathbb{R}) \ni \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and we set $w = \gamma z$, we have

$$\gamma^*|dz^2| = |dw|^2 = \left|\frac{(cz+d)adz - (az+b)cdz}{(cz+d)^2}\right|^2 = \left|\frac{dz}{(cz+d)^2}\right|^2 = \frac{|dz|^2}{|cz+d|^4},$$

$$\gamma^*\Im(z)^2 = \left(\frac{y}{|cz+d|^2}\right)^2$$

from which we deduce $\gamma^* g_H = g_H$.

Another space isometric to $\mathbb{H}^2$ is $\mathbb{B}^2$, the unit ball in $\mathbb{C}$, with metric

$$g_B = \frac{4\Re(|dz|^2}{(1-|z|^2)^2}. \tag{2.6}$$

As above, one proves that the map

$$\mathbb{H}^2 \ni z \mapsto w = \frac{z-i}{z+i} \in \mathbb{B}^2, \qquad z = \frac{i(1-w)}{1+w}$$

is an isometry. It is clearly bijective and bi-holomorphic. It is tautological that the group $\gamma_0 \mathrm{PSL}_2(\mathbb{R})\gamma_0^{-1} \subset \mathrm{PSL}_2\mathbb{C}$, where $\gamma_0 = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C})$, acts by holomorphic automorphisms on $\mathbb{B}^2$, which are also isometries.

THEOREM 2.8. *The group of orientation-preserving isometries of* $(\mathbb{H}^2, g_H)$ *is the same as the group of bi-holomorphic isomorphisms of* $\mathbb{H}^2 \subset \mathbb{C}$*, and is given by the group* $\mathrm{PSL}_2(\mathbb{R})$.

PROOF. From Lemma 7.3, every orientation preserving isometry is also holomorphic. We claim that every bi-holomorphic map is given by a Möbius transformation (i.e., an element of $\mathrm{PSL}_2(\mathbb{R})$) which is known to act as an isometry. To prove this, we use the ball model $\mathbb{B}^2$ of the hyperbolic plane. Take any automorphism Φ of $\mathbb{B}^2$ and set $w_0 = \Phi(0)$. There exists an automorphism $a \in G := \gamma_0 \mathrm{PSL}_2(\mathbb{R})\gamma_0^{-1}$ of $\mathbb{B}^2$ which takes 0 onto w_0, thus $f := a^{-1}\Phi$ fixes 0. Thus $f : \mathbb{B}^2 \to \mathbb{B}^2$ is a holomorphic function vanishing at 0, it follows that $g : z) := \frac{f(z)}{z}$ is also holomorphic, and from the maximum principle it takes values in $\mathbb{B}^2$. In particular, we have $|f'(0)| \geq 1$ since $f'(0) = g(0)$. Since the same reasoning applies to the inverse function f^{-1} we get $|f'(0)| = 1$ and by the equality case in the maximum principle for g we get that g is constant of absolute value 1, which means that f is a rotation, in particular an isometry. $\square$

Let Γ be a discrete subgroup in $\mathrm{PSL}_2(\mathbb{R})$. Assuming that Γ acts properly discontinuously, the quotient $S := \Gamma\backslash\mathbb{H}^2$ is a surface. Since the action is by smooth maps, the surface S is smooth. Since the action preserves the metric and the holomorphic structure, the quotient inherits a Riemannian metric of Gaussian curvature -1 (which is called a *hyperbolic* metric) as well as a structure of holomorphic surface.

2.5. Exercises.

EXERCISE 2.9. Show, using the definition of a connection, that the map $R : \mathcal{V}(S)^3 \to \mathcal{V}(S)$ defined in (1.4) is $C^\infty(S)$-linear in each of the three factors.

EXERCISE 2.10. Show that the distance function on a surface with Riemannian metric satisfies the triangle inequality.

EXERCISE 2.11. Let $\Phi : E \to E'$ be an isomorphism between two $\mathbb{K}$ vector bundles over S ($\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$), ∇' a connection on E' and ∇ its pull-back on E via Φ, i.e., $\nabla_X\psi := \Phi^{-1}(\nabla'_X\Phi(\psi))$. Show that the curvature tensors satisfy

$$R = \Phi^{-1}R'\Phi.$$

In particular, if E, E' are $\mathbb{K}$-line bundles, then

$$R = R' \in C^\infty(S, \Lambda^2(S))$$

after the canonical identifications $\mathrm{End}(E) = \mathrm{End}(E') = \mathbb{K}$.

EXERCISE 2.12. Prove that the image of vertical lines through an isometry $\gamma \in \mathrm{PSL}_2(\mathbb{R})$ is geometrically a half-circle with center on the real line.

3. Hyperbolic quotients

We will show that every complete orientable hyperbolic surface is a quotient of $\mathbb{H}^2$ by some discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$. For this, we need some additional topological and geometric notions.

3.1. The universal cover. The *universal cover* of a connected surface S is a simply-connected surface $\tilde{S}$ defined as follows: Fix a base point $p \in S$ and consider all the paths starting from p. Then $\tilde{S}$ is the set of homotopy classes of such paths relative to the end points. A base of neighborhoods of a class $[\gamma]$ is defined by considering charts $\phi_U : U \to \mathbb{R}^2$ with $\gamma(1) \in U$; we define $\mathcal{V}_U([\gamma])$ as the set of those $[\delta]$ for which $\delta(1) \in U$, and such that γ is homotopic to δ relative to 0 and U in the sense that there exists a homotopy $F : I \times I \to S$ with $F(s, 0) = p$ and $F(s, 1) \in U$, $\forall s \in I$. We will leave to the exercises the proof of the fact that $\tilde{S}$ is a simply-connected surface which covers S, and that $\tilde{S}$ is complete if S is complete. The projection

$$\tilde{S} \to S, \qquad\qquad [\gamma] \mapsto \gamma(1)$$

is a smooth map which is a local diffeomorphism.

Assume that a compact (or more generally, complete) surface S has a hyperbolic metric. Take $\tilde{S}$ to be the universal cover of S and pull back the metric through the covering map to a hyperbolic metric on the simply connected, complete surface $\tilde{X}$, which will turn out to be isometric to $\mathbb{H}^2$.

3.2. Jacobi fields and comparison theorems. We claim that up to isometry there exists precisely one simply connected, complete hyperbolic surface (which can be identified with both $\mathbb{B}^2$ and $\mathbb{H}^2$). Here completeness means that every geodesic can be extended for arbitrary positive and negative time. From Lemma 2.3, on every surface S endowed with a Riemannian metric there exists a map called *geodesic flow* from a neighborhood of $\{0\} \times T^S \subset \mathbb{R} \times TS$ into S, defined by $(t, v) \mapsto \gamma_v(t)$, where γ_v is the unique geodesic starting in the direction of v. The fact that S is complete can be restated as saying that the geodesic flow is well-defined for all times t. The *exponential map* at a point $p \in S$ is then defined by

$$T_pS \to S, \qquad\qquad \exp_p(v) := \gamma_v(1).$$

EXAMPLE 3.1. Let S be the Euclidean space $\mathbb{R}^2$, then the exponential map is the identity. If S is a torus $S = \Gamma\backslash\mathbb{R}^2$, the exponential map is the covering map $\mathbb{R}^2 \to \Gamma\backslash\mathbb{R}^2$.

LEMMA 3.2. *For every point $z \in \mathbb{H}^2$, the exponential map $\exp_z : \mathbb{R}^2 \to \mathbb{H}^2$ is a diffeomorphism.*

PROOF. Since $\exp$ clearly commutes with isometries, we can check the statement on the disk model of the hyperbolic plane, with $z = 0$. Then it is an easy computation that $\exp_0(v) = f(|v|)v$, for $f(r) = \frac{\tanh(r/2)}{r}$. The function f is even in r, thus $f(|v|)$ is smooth including at $v = 0$, and ϕ is a smooth homeomorphism, hence a diffeomorphism. □

We claim that the exponential map sends lines through $0 \in T_pS$ (parametrized with constant speed) into geodesics. Indeed, we first note that the geodesics through p are scale-invariant in the following sense:

$$\gamma_{tv}(s) = \gamma_v(ts).$$

This holds because both curves are geodesic (the right-hand side is a linear rescaling of a geodesic hence a geodesic itself) and their initial tangent vector at $s = 0$ equals tv. It follows that $\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t)$ runs along the geodesic γ_v.

Using the exponential map, we can construct near each point a distinguished chart from the Riemannian metric, using the exponential map.

LEMMA 3.3. *The differential $D(\exp_p)_p : T_pS \to T_pS$ is the identity map.*

PROOF. Immediate from the definition of the exponential and of the differential. If $c : [0,1] \to T_pS$ is a curve through p with $\dot{c}(0) = v$, then $D(\exp_p)_p(v)$ is by definition the vector tangent to the curve $\exp_p(c(t))$. Fix $v \in T_pS$ and take $c(t) := tV$, then $\exp_p(tv) = \gamma_v(t)$ where γ_v is the unique geodesic starting at p in the direction of v. Since the tangent vector at $t = 0$ to this geodesic is v, it follows that $D(\exp_p)_p$ acts as the identity. □

Therefore the exponential map $\exp_p$ remains non-degenerate in a neighborhood of p, hence it is the inverse of a chart. By taking an isometry of T_pS with $\mathbb{R}^2$, we get the so-called *normal coordinates* on S near p.

LEMMA 3.4. *In normal geodesic coordinates (x_1, x_2), the following identities hold at p:*

$$\nabla_{\partial_{x_i}} \partial_{x_j}(p) = 0, \qquad\qquad \partial_{x_i}(g_{jk})(p) = 0.$$

We stress that this vanishing occurs only at p, otherwise we would get zero curvature.

PROOF. Take $a_1, a_2 \in \mathbb{R}$ and define $v := (a_1, a_2) \in T_pS$. Set V to be the vector field tangent to the geodesic γ_v, we can write it as $V_{(ta_1,ta_2)} = a_1\partial_{x_1} + a_2\partial_{x_2}$. By expanding $\nabla_V V = 0$ we get

$$\sum_{i,j=1}^{2} a_i a_j \nabla_{\partial_{x_i}} \partial_{x_j} = 0,$$

valid along the geodesic γ_v, so in particular at p. The coefficients a_1, a_2 are arbitrary and $\nabla_{\partial_{x_1}}\partial_{x_2} = \nabla_{\partial_{x_2}}\partial_{x_1}$ (valid for coordinate vector fields) so we get the first statement. The second one is a consequence:

$$\partial_{x_i}(g_{jk}) = \partial_{x_i}\langle \partial_{x_j}, \partial_{x_k}\rangle = \langle \nabla_{\partial_{x_i}}\partial_{x_j}, \partial_{x_k}\rangle + \langle \partial_{x_j}, \nabla_{\partial_{x_i}}\partial_{x_k}\rangle$$

and the right-hand side vanishes at p. □

LEMMA 3.5. *Let X_1, X_2 be an orthonormal frame in p. Extend X_1, X_2 by parallel transport along geodesic rays starting from p. Then $[X_1, X_2]_p = 0$ and $\nabla_{X_i} X_j = 0$, for all $i, j \in \{1, 2\}$.*

PROOF. Since X_1 agrees with ∂_{x_1} at p, and they are both parallel along the x_1-geodesic γ_1, it follows that X_1 agrees with ∂_{x_1} along γ_1. By definition, the covariant derivative of X_j with respect to ∂_{x_1} vanishes along γ_1. In particular, at p we have $\nabla_{X_1} X_j = 0$. The identity $[X_1, X_2] = \nabla_{X_1} X_2 - \nabla_{X_2} X_1$ completes the proof. □

THEOREM 3.6. *Let S be a geodesically complete, simply connected hyperbolic surface. Choose a vector-space isometry ϕ between $T_0\mathbb{B}^2$ and T_pS. Then $\Phi := \exp_p \circ \phi \circ \exp_0^{-1}$ is an isometry.*

PROOF. First we prove that Φ is a local isometry, using Jacobi fields. Then for topological reasons, Φ must be injective, thus it is a diffeomorphism. □

In order to check that a surface is geodesically complete, we mention the following results:

THEOREM 3.7. *The topology on a Riemannian manifold is the same as the topology given by the distance function. A Riemannian manifold is geodesically complete if and only if it is complete as a metric space, i.e., every Cauchy sequence is convergent.*

In particular, compact surfaces are geodesically complete.

Two metrics g, g' on S are called *quasi-isomorphic* if there exists $C > 0$ such that for every $p \in S$ and $v \in T_pS$, we have

$$C^{-1}\|v\|_g \leq \|v\|_{g'} \leq C\|v\|_g.$$

COROLLARY 3.8. *Let g be a geodesically complete metric on S, and g' a quasi-isometric metric. Then g' is also geodesically complete.*

3.3. Classification of complete hyperbolic surfaces.

THEOREM 3.9. *Every complete connected hyperbolic surface is isometric to a quotient of $\mathbb{H}^2$ by a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ isomorphic to $\pi_1(S)$.*

PROOF. The pull-back through the local diffeomorphism $\pi : \tilde{S} \to S$ of the metric on S is a Riemannian metric on $\tilde{S}$. Since Gaussian curvature is a local quantity, it follows that $\tilde{S}$ is also a hyperbolic surface. From Theorem 3.6, every two simply connected hyperbolic surfaces are isometric, thus $\tilde{S}$ is isometric to $\mathbb{H}^2$. The group $\mathcal{D}(\tilde{S}, \pi)$ of *deck transformations*, i.e., those diffeomorphisms $\Phi : \tilde{S} \to \tilde{S}$ which satisfy $\pi \circ \Phi = \pi$, is isomorphic to $\pi_1(S)$. But it is evident that $S = \tilde{S}/\mathcal{D}(\tilde{S}, \pi)$. □

3.4. Exercises.

EXERCISE 3.10. Prove that the conjugate group $\gamma_0 \mathrm{PSL}_2(\mathbb{R})\gamma_0^{-1} \subset \mathrm{PSL}_2(\mathbb{C})$, where $\gamma_0 = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C})$, is precisely

$$G = \{z \mapsto \frac{az+b}{\overline{b}z+\overline{a}}; a, b \in \mathbb{C}, |a|^2 - |b|^2 = 1\}.$$

EXERCISE 3.11. Prove that $\tilde{S}$ constructed in the proof of Theorem 3.9 is a surface and the map $[\gamma] \mapsto \gamma(1)$ is a covering map. If S is smooth (or holomorphic) then $\tilde{S}$ has a canonical smooth (respectively holomorphic) structure.

4. Uniformization

So far we have described hyperbolic surfaces as metric quotients of the hyperbolic plane. In this section we will show that topologically, every surface of genus $g \geq 2$ can be realized as such a quotient. The tool for this will be the uniformization theorem 4.3. We will need smooth structures and Riemannian metrics, let therefore (S, g) be a compact orientable smooth surface together with a fixed Riemannian metric. We first compute the behaviour of Gaussian curvature under conformal change. Let us recall for this purpose the definition of the Laplacian in Riemannian setting.

The space $C_c^\infty(S,\mathbb{C})$ has a pre-Hilbert structure, i.e., a scalar product for which the induced topology is not complete. This scalar product is defined using the Riemannian metric g by the L^2 product:

$$\langle f, f'\rangle_{L^2} := \int_S f f' \mu_g,$$

where μ_g is the volume form associated to the metric g. In local coordinates, $\mu_g = \det(g_{ij})^{1/2} dx \wedge dy$, where (g_{ij}) is the 2×2 matrix of the metric g in the basis ∂_x, ∂_y.
Similarly we define an L^2 product on 1-forms. We first note that g induces an isomorphism between TS and T^*S by the correspondence $V \mapsto g(V,\cdot)$. Thus g is also, using this correspondence, a metric on the bundle T^*S. For 1-forms α, α' we set

$$\langle \alpha, \alpha'\rangle_{L^2} := \int_S g(\alpha,\alpha')\mu_g.$$

Recall the de Rham differential $d : C^\infty(S) \to C^\infty(S, T^*S)$. The operator d^* (sometimes denoted δ) is the adjoint of d with respect to the scalar products on functions and on 1-forms: the identity

$$\langle d^*\alpha, f\rangle = \langle \alpha, df\rangle$$

defines uniquely d^*. The Laplacian is defined by

$$\Delta := d^*d : C^\infty(S) \to C^\infty(S).$$

As an illustration, let us compute the Laplacian for two particular cases. First, let S be $\mathbb{R}^2$ with the flat metric. The length of the basic 1-forms dx, dy is 1 at every point. Then for $f \in C^\infty(\mathbb{R}^2)$ and $\alpha = adx + bdy \in C^\infty(\mathbb{R}^2, \Lambda^1)$ with one of them of compact support, we compute using integration by parts

$$\begin{aligned}\langle df, \alpha\rangle =& \langle \partial_x f dx + \partial_y f dy, adx + bdy\rangle \\ =& \int_{\mathbb{R}}^2 (\partial_x f a + \partial_y f b) dxdy \\ =& -\int_{\mathbb{R}}^2 (f\partial_x a + f\partial_x b) dxdy \\ =& \langle f, -\partial_x a - \partial_y b\rangle\end{aligned}$$

so we deduce $d^*\alpha = -\partial_x a - \partial_y b$. The Euclidean Laplacian is therefore given by

$$\Delta^{\mathbb{R}^2} f = -(\partial_x^2 + \partial_y^2) f.$$

Let us now compute the hyperbolic Laplacian in the $\mathbb{H}^2$ model. The volume form is $y^{-2}dxdy$, while $|dx| = |dy| = y$ (i.e., the length of the 1-forms dx, dy at (x, y) equals y). Therefore,

$$\begin{aligned}\langle df, \alpha\rangle &= \langle \partial_x f dx + \partial_y f dy, adx + bdy\rangle \\ &= \int_{\mathbb{R}}^2 (y^2\partial_x f a + y^2\partial_y f b)y^{-2}dxdy \\ &= -\int_{\mathbb{R}}^2 (f\partial_x a + f\partial_x b)dxdy \\ &= -\int_{\mathbb{R}}^2 y^2(f\partial_x a + f\partial_x b)y^{-2}dxdy \\ &= \langle f, -y^2(\partial_x a + \partial_y b)\rangle.\end{aligned}$$

The hyperbolic co-differential is thus given by $d^*\alpha = -y^2(\partial_x a + \partial_y b)$, while the Laplacian equals

$$\Delta^{\mathbb{H}^2} f = -y^2(\partial_x^2 + \partial_y^2)f. \tag{4.1}$$

There is nothing special about the function $y = \Im(z)$ in the above computation. Exactly in the same way we obtain

LEMMA 4.1. *Let $g, g' := e^{-2f}g$ be two conformal metrics on a surface S. Then*

$$\Delta_{g'} = e^{2f}\Delta_g.$$

Note however that in higher dimensions the formula is no longer valid. In local coordinates, we can compute from the definition

$$\Delta f = -\frac{1}{\sqrt{\det(g)}}\sum_{i,j=1}^n \partial_{x_i}(\sqrt{\det(g)}g^{ij}\partial_{x_j})f,$$

where $g = (g_{ij})_{i,j=1,\dots,n} \in \mathrm{GL}_n(\mathbb{R})$ is coefficient matrix for the metric in the basis ∂_{x_i}, i.e., $g_{ij} = g(\partial_{x_i}, \partial_{x_j})$, and g^{ij} are the coefficients of the inverse matrix g^{-1}. The case $n = 2$ is special because then $\sqrt{\det(g')}g'^{ij} = \sqrt{\det(g)}g^{ij}$ for any conformal change $g' = e^{-2f}g$.

As a consequence of lemma 3.4 and this formula, in geodesic normal coordinates at p we have

$$\Delta f(p) = -(\partial_{x_1}^2 + \partial_{x_2}^2)f(p).$$

We stress again that such a formula is only valid at the origin of the normal geodesic coordinates.

LEMMA 4.2. *For $f \in C^\infty(S)$, define $g' := e^{-2f}g$. The Gaussian curvature functions κ, κ' are related by*

$$\kappa' = e^{2f}(\kappa - \Delta f).$$

PROOF. Let ∇, ∇' be the Levi-Civita connections associated to g, g'. Fix arbitrary vector fields X, Y, Z on S. Using (1.3), we obtain

$$2\langle \nabla'_X Y, Z\rangle_{g'} = 2e^{-2f}\left(\langle \nabla_X Y, Z\rangle_g - X(f)\langle Y, Z\rangle_g - Y(f)\langle X, Z\rangle_g + Z(f)\langle X, Y\rangle_g\right),$$

which implies (since Z was arbitrary)

$$\nabla'_X Y = \nabla_X Y - X(f)Y - Y(f)X + \langle X, Y\rangle \nabla(f). \tag{4.2}$$

It is enough to prove the Lemma at an arbitrary point p. Choose geodesic normal coordinates around p, and let X_1, X_2 be the orthonormal frame introduced in Lemma 3.5. From (4.2),

$$\begin{aligned}\nabla'_{X_2} X_2(p) =& \nabla_{X_2} X_2 - 2X_2(f)X_2 + \nabla(f),\\ \nabla'_{X_1} X_2(p) =& \nabla_{X_1} X_2 - X_1(f)X_2 - X_2(f)X_1.\end{aligned}$$

Recall from Lemma 3.4 that at p we have $[X_i, X_j] = \nabla_{X_i} X_j = 0$. By ignoring the terms which vanish at p we compute

$$\begin{aligned}\nabla'_{X_1}\nabla'_{X_2} X_2 \equiv& \nabla_{X_1}\nabla_{X_2} X_2 - 2X_1X_2(f)X_2 + \nabla_{X_1}\nabla(f)\\ &+ 2X_1(f)X_2(f)X_2 - X_1(f)\nabla(f)\\ &+ 2X_2(f)^2X_1 - |\nabla(f)|^2 X_1 + X_1(f)\nabla(f),\\ \nabla'_{X_2}\nabla'_{X_1} X_2 \equiv& \nabla_{X_2}\nabla_{X_1} X_2 - X_2X_1(f)X_2 - X_2^2(f)X_1\\ &+ X_2(f)X_1(f)X_2 + X_2(f)^2X_1 + 2X_1(f)X_2(f)X_2 - X_1(f)\nabla(f)\end{aligned}$$

From this, we deduce that modulo terms vanishing at p

$$\langle R'_{X_1X_2}X_2, X_1\rangle_g = \langle R_{X_1X_2}X_2, X_1\rangle_g + X_1^2(f) + X_2^2(f) + \mathcal{I}_p,$$

hence since $\Delta = -(X_1^2 + X_2^2)$,

$$\langle R'_{X_1X_2}X_2, X_1\rangle_{g'} = e^{-2f}(\kappa_g - \Delta f).$$

This identity is only valid at p, which however was arbitrary. The proof is finished by noting that the orthogonal frame X_1, X_2 for g' is made of vectors of g'-length equal to e^{-f}, hence $\kappa' = e^{4f}\langle R'_{X_1X_2}X_2, X_1\rangle_{g'}$. □

We come now to one of our main theorems about surfaces - the uniformization theorem. It essentially says that conformal structures on orientable surfaces admit a distinguished representative of constant curvature, of the sign of the Euler characteristic of the surface.

THEOREM 4.3. *Let S be an orientable closed surface of genus g together with a Riemannian metric h. Then*

1. *If $g = 0$, there exists $f \in C^\infty(S)$ (not unique) such that S with the metric $e^{-2f}h$ is isometric with the standard unit sphere.*
2. *If $g = 1$, there exists $f \in C^\infty(S)$ with $e^{-2f}h$ flat, and f is unique up to an additive constant.*
3. *If $g \geq 2$, there exists a unique $f \in C^\infty(S)$ with $e^{-2f}h$ hyperbolic.*

Before proving the theorem, let us examine some of its consequences.

The third part, together with Theorem 7.2, says that almost complex structures are in bijection with hyperbolic structures for genus ≥ 2, and with flat structures of volume 1 in genus 1.

COROLLARY 4.4. *Let S be a closed orientable surface of genus $g \geq 2$ endowed with a conformal structure, or equivalently with an almost complex structure J. Then there exists a discrete subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$, unique up to conjugation in $\mathrm{PSL}_2(\mathbb{R})$, and a conformal diffeomorphism $\Gamma\backslash\mathrm{PSL}_2(\mathbb{R}) \to S$.*

PROOF. This follows immediately from the uniformization theorem and Theorem 3.9. □

In particular, J is integrable because the deck transformations group Γ acts by holomorphic automorphisms of $\mathbb{H}^2$. The second part of the uniformization theorem allows us to show below that on *every* orientable conformal surface (i.e., not necessarily compact) the induced almost complex structure is integrable. This follows as a particular case from the Newlander-Nirenberg theorem valid in higher dimensions, but here we get it essentially for free.

COROLLARY 4.5. *Let S be an orientable surface endowed with a conformal structure. Then the associated almost complex structure J is integrable, i.e., S admits a holomorphic atlas so that for each chart (x, y) in the atlas, $J\partial_x = \partial_y$.*

PROOF. We start by a trivial remark. Pick a metric h in the fixed conformal class on S.

LEMMA 4.6. *For every point $p \in S$, there exists a neighborhood $U \ni p$, a metric h' on the torus $\mathbb{T}^2$ and an isometry from (U, h) to an open subset $U' \subset \mathbb{T}^2$ endowed with the metric h'.*

PROOF. Choose $p' \in \mathbb{T}^2$, and charts around p, p', namely $\phi : V \to \mathbb{R}^2$, $\phi' : V' \to \mathbb{R}^2$. Define $\Phi := \phi'^{-1} \circ \phi : V \to V'$. Choose a relatively compact open neighborhood U' of p' in V' and a partition of unity ψ_1, ψ_2 on $\mathbb{T}^2$ relative to the open cover $\{V', \mathbb{T}^2 \setminus \overline{U'}\}$. Choose any Riemannian metric h_2 (e.g. the standard metric) on $\mathbb{T}^2 \setminus \overline{U'}$, and over V' take $h_1 := \Phi_* h$, the direct image of h through the diffeomorphism Φ. Then $\psi_1 h_1 + \psi_2 h_2 =: h'$ is a Riemannian metric on $\mathbb{T}^2$ which agrees with h_1 over U'. This means that Φ is an isometry from $(U := \Phi^{-1}(U'), h)$ onto (U', h'). □

From the uniformization theorem, h' is conformally flat, which by the above lemma implies that h was locally conformally flat. A flat metric is locally isometric to $\mathbb{R}^2$ with its standard metric. For each point p in S take a function f_p defined near p such that $e^{-2f}h$ is flat in a neighborhood $U \ni p$, and choose a positively oriented isometry from a possibly smaller neighborhood $W \ni p$ onto an open subset of $\mathbb{R}^2$, say $\phi_p : W \to W'$. Clearly when p varies, these isometries provide a topological atlas for S. Since ϕ_p is an isometry relative to the metric $e^{-2f}h$, it follows that it is conformal relative to h. Thus the atlas consists of conformal maps, it follows that the changes of charts are also conformal maps between subsets of $\mathbb{R}^2$. By Lemma 7.3, oriented conformal maps between subsets of $\mathbb{C}$ are holomorphic, thus we have constructed a holomorphic atlas. □

Note that here we have used only the genus 1 part of the uniformization theorem, which is considerably simpler to prove than the other two cases, see section 5.1.

4.1. Exercises.

EXERCISE 4.7. Show that if $\{U_i\}_{i\in I}$ is an open cover of a surface S, $\{\psi_i\}$ are a partition of unity and h_i are Riemannian metrics on U_i, then $h := \sum_{i\in I} \psi_i h_i$ defines a Riemannian metric on S.

5. Proof of the uniformization theorem

It may look surprising that the genus, a topological quantity, matters in the analytic statement of the uniformization theorem. The explanation lies in the Gauss-Bonnet formula (Theorem 2.5).

REMARK 5.1. Let (S, h) be a closed oriented surface of genus g with Riemannian metric h. If g has constant Gaussian curvature, then its sign equals the sign of $\chi(S) = 2 - 2g$.

5.1. The torus case. Let us start with the easiest case where the genus of S is 1. By Lemma 4.2, in order to uniformize the surface we would like to solve the equation

$$e^{2f}(\kappa - \Delta f) = 0,$$

or equivalently

$$\Delta f = \kappa. \tag{5.1}$$

This is the *Laplace equation*, one of the most studied partial differential equations. To solve it, consider a finite-dimensional analogy.

LEMMA 5.2. *Let $d : E \to F$ be a linear map between two finite-dimensional vector spaces with scalar product. Then the equation $d^*dx = u$ has solutions if and only if $u \perp \ker(d)$, in which case the solutions form an affine space modeled on $\ker(d)$.*

Let us therefore first determine the space of harmonic functions on S.

LEMMA 5.3. *Let (S, h) be an oriented, connected and closed Riemannian manifold. Then any $f \in C^\infty(S)$ which is harmonic (i.e., $\Delta f = 0$) must be constant.*

PROOF. Since $\Delta f = 0$ we have $\langle \Delta f, f\rangle = 0$. But

$$\langle \Delta f, f\rangle = \langle d^*df, f\rangle = \|df\|^2_{L^2}$$

so $df = 0$. Locally, $\partial_x f dx + \partial_y f dy = 0$ which implies $\partial_x f = \partial_y f = 0$ so f is locally constant, and since we implicitly assume S to be connected we see that f is constant on S. ☐

THEOREM 5.4. *Let (S, h) be an oriented, connected and closed Riemannian manifold. The Laplace equation $\Delta f = u$ has solutions if and only if the function u is orthogonal in L^2-sense to the kernel of Δ, i.e., if $\int_S u\nu_h = 0$.*

PROOF. One direction is clear, namely we see immediately that $\Delta f \perp 1$:

$$\langle \Delta u, 1\rangle_{L^2} = \langle du, d1\rangle_{L^2} = 0. \tag{5.2}$$

Notice that by definition, $\langle \Delta u, 1\rangle_{L^2} = \int_S \Delta f \nu_h$ therefore we conclude that functions of the form Δf must have zero mean. ☐

If f is a solution to (5.1) then $f + c$ is also a solution, for all constant functions c. The existence of solutions depends therefore on a “geometric” issue: is it true that on the torus $\mathbb{T}^2$ with metric h we have

$$\langle \kappa, 1\rangle_{L^2(\mathbb{T}^2)} = \int_{\mathbb{T}^2} \kappa(x) 1 \mu_h(x) = 0?$$

The reader who paid attention to the Gauss-Bonnet formula already knows the answer to be always true because the Euler characteristic of the torus is 0. This settles the second statement of Theorem 4.3.

5.2. The sphere case. Let h be any Riemannian metric on S^2. Pick a point p, by Lemma 4.6 we know that h is conformally flat in a neighborhood of p. Up to a global conformal change, we can therefore assume h to be flat in a neighborhood of p.

There exists an isometry from a neighborhood of p to a neighborhood of $0 \in \mathbb{C}$, given essentially by the geodesic exponential map (see Theorem 3.6). Let $U := S^2 \setminus \{p\}$. The curvature κ_h is of course compactly supported in U, and by the Gauss-Bonnet theorem it satisfies

$$\int_U \kappa_h \nu_h = 4\pi.$$

Let $\phi : U \to (0, \infty)$ be a smooth function which near p equals $\phi(z) = \|z\|$, where z is the standard complex variable on $\mathbb{C}$.

LEMMA 5.5. *The function $\Delta_h \log \phi$ vanishes near p, and its integral satisfies*

$$\int_U \Delta_h(\log \phi)\nu_h = 2\pi.$$

PROOF. Let ϕ, ϕ' be two strictly positive functions which equal $\|z\|$ near $z = 0$. Then $\log \phi - \log \phi'$ has compact support in U, so in particular it is smooth on S^2. For every function $u \in C^\infty(S^2)$ one has

$$\int_{S^2} \Delta u \nu_h = \langle \Delta u, 1\rangle_{L^2} = \langle du, d1\rangle_{L^2} = 0,$$

so $\int_U \Delta_h(\log \phi)\nu_h$ is independent of the choice of ϕ (i.e., of the smooth extension of $\|z\|$ to the whole S^2). We pick a particular such extension, which is 1 outside the flat region of S^2 around 0, so in particular $\phi(z)$ depends only on $\|z\|$, it is identically 1 for $\|z\| \geq \epsilon$, and $\phi(z) = \|z\|$ for $\|z\| \leq \epsilon/2$. Clearly $\Delta \log \phi = 0$ on the region where ϕ is constant, it remains to evaluate the contribution near $z = 0$. In polar coordinates, $\Delta = -r^{-1}\partial_r r \partial_r - r^{-2}\partial_\theta^2$ and $\phi = \phi(r)$ so

$$\int_{\|z\|<\epsilon} \Delta(\log \phi) r dr d\theta = -2\pi \int_0^\epsilon \partial_r r \partial_r \log \phi(r) dr = -2\pi (r\partial_r \log \phi)|_{\epsilon/2}^{\epsilon} = 2\pi. \qquad \square$$

Let ϕ be any function as in Lemma 5.5. Then $\Delta \log \phi$ has compact support in U, and

$$\int_U (\kappa_h - 2\Delta \log \phi)\nu_h = 4\pi - 2 \cdot 2\pi = 0.$$

By theorem 5.4, there exists a smooth function $f_0 \in C^\infty(S^2)$, unique up to a constant, such that $\Delta_h f_0 = \kappa_h - 2\Delta_h \log \phi$. Setting $f := f_0 + 2 \log \phi$, we have by Lemma 4.2 that the metric $e^{-2f}h$ is flat on U. Notice that $e^{-2f}h$ is not defined at p because of the singularity introduced by the conformal factor $e^{-2log\phi} = r^{-4}$ near p.

Near $z = 0$ the metric $e^{-2f}h$ is just $e^{-2f_0}\phi(r)^{-2}(dr^2 + r^2 d\theta^2)$. Note that f_0 has a finite limit at $r = 0$, which we could assume to be 0 since we may subtract from f_0 an arbitrary constant. Set $R := 1/r$, then $e^{-2f}h = e^{-2f_0}(dR^2 + R^2 d\theta^2)$ as $R \to \infty$, where $\lim_{R\to\infty} e^{-2f_0(R,\theta)} = 1$.

We observe that the flat metric $e^{-2f}h$ is complete on the simply connected set U. Indeed, we have seen above that it is quasi-isometric to the standard metric on $\mathbb{C}$, so Corollary 3.8 applies. It follows by Theorem 3.6 that $(U, e^{-2f}h)$ is isometric to $\mathbb{C}$ with its standard metric. In particular, (U, h) is bi-holomorphic to $\mathbb{C}$ by some map Φ, uniquely defined up to an isometry of $\mathbb{C}$.

Consider now the holomorphic map Φ from $U \subset S$ to $\mathbb{C} \subset \mathbb{C}P^1$ as a singular map from S to $\mathbb{C}P^1$. In holomorphic charts near p, respectively near ∞, this map is bounded, hence its singularity is removable. By the same argument, the singularity of Φ^{-1} is removable. It follows that Φ extends to a bi-holomorphism of S onto $\mathbb{C}P^1$. Since Φ is holomorphic, hence conformal from h to the standard metric g_0 on $\mathbb{C}P^1$, it follows that $\Phi^* g_0$ is a metric of curvature 1 conformal to h.

5.2.1. *The set of conformal metrics of curvature* 1. We conclude that every metric h on S^2 is conformal to a metric of curvature 1, so there exists $f \in C^\infty(S^2)$ with $\kappa_{e^{2f}h} = 1$. By Theorem 3.6, the metric $e^{2f}h$ is isometric by a map Φ_f to the standard metric on S^2. We would like to understand how many such f exist. Let f, f' be two such conformal factors and $\Phi_f, \Phi_{f'}$ such that

$$\Phi_f^* g_0 = e^{2f}h, \qquad \Phi_{f'}^* g_0 = e^{2f'}h.$$

It follows that $\Phi_f \Phi_{f'}^{-1} =: \gamma$ is a conformal transformation of the standard sphere, i.e., a Möbius transformation (sometimes also called *projective transformation*)in $\mathrm{PSL}_2\mathbb{C}$. Conversely, every $\gamma \in \mathrm{PSL}_2\mathbb{C}$ has the property that $\gamma\Phi_f$ is a conformal map from S to S^2.

Some homographies act as isometries on S^2. The isometry group of S^2 is $\mathrm{SO}(3)$, which we view therefore as a subgroup in $\mathrm{PSL}_2\mathbb{C}$. It is obvious that $\gamma\Phi_f$ induces the same conformal factor (i.e., $f = f'$) if and only if $\gamma \in \mathrm{SO}(3)$. It follows that the set of conformal metrics of curvature 1 in a given conformal class is in bijection with the homogeneous space $\mathrm{SO}(3)\backslash\mathrm{PSL}_2\mathbb{C}$.

5.3. The case of genus $g \geq 2$**.** Here we must solve the equation

$$e^{2f}(\kappa - \Delta f) = -1,$$

or equivalently

$$\Delta f - \kappa - e^{-2f} = 0. \tag{5.3}$$

Consider the functional $E : C^\infty(S) \to \mathbb{R}$ given by

$$E(f) := \int_S \left(\tfrac{1}{2}\|df\|^2 - \kappa f + \tfrac{1}{2}e^{-2f}\right)\mu_h.$$

Assume that f is a point where E attains its minimum. Then for every $\phi \in C^\infty(S)$, the function

$$E_\phi : \mathbb{R} \to \mathbb{R}, \qquad E_\phi(t) := E(f + t\phi))$$

attains its minimum at $t = 0$, therefore $E'_\phi(0) = 0$. We can compute $E'_\phi(0)$ as follows:

$$\begin{aligned} E'_\phi(0) =& \int_S (\tfrac{1}{2}(\langle d\phi, df\rangle + \langle df, d\phi\rangle) - \kappa\phi - \phi e^{-2f})\mu_h \\ =& \langle \Delta f - \kappa - e^{-2f}, \phi\rangle_{L^2(S)}. \end{aligned}$$

Since ϕ was arbitrary, it follows that a minimum point for E (or more generally, a critical point) is a solution for (5.3) because the set of smooth $\phi \in C^\infty(S)$ is dense in $L^2(S)$.

It remains to show (and this is the main part of the proof) that E does attain its minimum. Let

$$f^0 := \frac{1}{\mathrm{Vol}(S)} \int_S f\mu_h \in \mathbb{R}, \qquad\qquad f^\perp := f - f^0$$

be the orthogonal decomposition of f into its zero-mode f^0 (i.e., the constant component, or equivalently the component along the kernel of Δ) and its component in the range of Δ. Notice that by definition the average of $f^\perp$ over S is 0, i.e.,

$$\int_S f^\perp \mu_h = 0. \tag{5.4}$$

LEMMA 5.6. *The functional E is bounded from below.*

PROOF. Decompose $E(f)$ into

$$\begin{aligned} E(f) =& \int_S (\tfrac{1}{2}\|df^\perp\|^2 - \kappa f^\perp)\mu_h + \int_S (-\kappa f^0 + \tfrac{1}{2}e^{-2f})\mu_h \\ =:& E_1(f^\perp) + E_2(f) \end{aligned}$$

(notice that $df = df^\perp$). The second term $E_2(f)$ can be bounded from below using the inequality $e^x \geq 1 + x$, valid for all $x \in \mathbb{R}$ (and proved most easily on the graph):

$$\int_S e^{-2f}\mu_h = e^{-2f^0} \int_S e^{-2f^\perp}\mu_h \geq e^{-2f^0} \int_S (1 - 2f^\perp)\mu_h = e^{-2f^0}\mathrm{Vol}(S),$$

where in the last equality we used (5.4). Using Gauss-Bonnet,

$$\int_S (-\kappa f^0 + \tfrac{1}{2}e^{-2f})\mu_h = 2\pi(2g-2)f^0 + \tfrac{1}{2}\int_S e^{-2f}\mu_h \geq 2\pi(2g-2)f^0 + e^{-2f^0}\tfrac{\mathrm{Vol}(S)}{2}.$$

Both constants $2g - 2$ and $\frac{\mathrm{Vol}(S)}{2}$ are strictly positive, thus when f^0 is negative, the exponential will dominate $(2g-2)f^0$. It follows that

$$E_2(f) \geq C_1$$

for some constant C_1, as claimed.

The first term $E_1(f^\perp)$ is rather obviously bounded from below; here is a quick argument: Decompose $f^\perp$ in an orthonormal basis (for the L^2 inner product) of eigenfunctions of Δ of positive eigenvalue:

$$f^\perp = \sum_{j=1}^\infty a_j \phi_j,$$

where $\Delta\phi_j = \lambda_j\phi_j$ and $0 < \lambda_j \nearrow \infty$ as $j \to \infty$. Notice that

$$\int_S \|df^\perp\|^2\mu_f = \langle df^\perp, df^\perp\rangle_{L^2} = \langle \Delta f^\perp, f^\perp\rangle = \sum_{j=1}^\infty a_j^2\lambda_j.$$

On the other hand, by the Cauchy-Bunyakovsky-Schwartz inequality,

$$\left|\int_S -\kappa f^\perp\mu_h\right| \le \left(\int_S \kappa^2\mu_h \int_S |f^\perp|^2\mu_h\right)^{\frac{1}{2}} = C_2\left(\sum_{j=1}^\infty a_j^2\right)^{\frac{1}{2}} = C_2\|f^\perp\|_{L^2}.$$

It follows that

$$E_1(f^\perp) \ge \sum_{j=1}^\infty a_j^2\lambda_j - C_2\left(\sum_{j=1}^\infty a_j^2\right)^{\frac{1}{2}} \ge \lambda_1\|f^\perp\|^2_{L^2} - C_2\|f^\perp\|_{L^2} \ge -\frac{C_2^2}{4\lambda_1},$$

where λ_1 is the smallest non-zero eigenvalue of Δ. □

5.4. Existence of the minimum. Let now f_n be a sequence of smooth functions such that $E(f_n)$ converges towards the infimum. Up to choosing a subsequence, we can assume that the sequences $E_1(f_n^\perp)$ and $E_2(f_n)$ both converge towards some finite limits.

For integer $k \ge 0$, let us define the k-Sobolev scalar product on $C^\infty(S)$: for $k = 2m$ even,

$$\langle u, v\rangle_{H_k} = \langle \Delta^m u, \Delta^m v\rangle_{L^2} + \langle u, v\rangle_{L^2},$$

while for $k = 2m+1$ odd,

$$\langle u, v\rangle_{H_k} = \langle d\Delta^m u, d\Delta^m v\rangle_{L^2} + \langle u, v\rangle_{L^2}.$$

The k-th Sobolev space $H^k(S)$ is defined as the completion of $C^\infty(S)$ with respect to this inner product. Using distributions, one can show that $H^k(S)$ is the space of those L^2 functions on S such that $\Delta^{k/2}f$ is in L^2 (for even k), respectively such that $d\Delta^{\frac{k-1}{2}}f$ is in L^2 as a 1-form, for k odd. It is easy to see that a function $f = \sum_{j=0}^\infty a_j\phi_j$ is in H^k if and only if the sequence $(\lambda_j^{k/2}a_j)$ is in l^2.

We claim that we have the following "exponential Sobolev embedding", to be proved later:

LEMMA 5.7. *There exist constants $\alpha, C > 0$ such that for every $f \in C^\infty(S)$ with $\|f^\perp\|^2_{H^1} \le \alpha$, we have*

$$\int_S e^{f^2}\mu_g \le Ce^{(f^0)^2}. \tag{5.5}$$

where $f = f^0 + f^\perp$ with f^0 constant, and $f^0 \perp f^\perp$ in L^2 sense.

Assuming this to be true, we deduce by using the standard inequality

$$2f^\perp \le \alpha(f^\perp)^2 + \frac{1}{\alpha}$$

that for every c there exists a constant C such that for $\|\nabla f\|_{L^2} < c$ we have

$$\int_S e^{2f^\perp}\mu \le Ce^{\|\nabla f\|_{L^2}}. \tag{5.6}$$

Since $E_1(f_n^\perp)$ is convergent, it must be bounded from above by some constant c_1. This implies that the sequence $\|\nabla f_n^\perp\|_{L^2}$ is bounded: indeed, if we set $u_n := \|\nabla f_n^\perp\|_{L^2} \ge \sqrt{\lambda_1}\|f_n^\perp\|_{L^2}$, then

$$u_n^2 = E_1(f_n^\perp) + \langle \kappa, f_n^\perp\rangle_{L^2} < c_1 + \|\kappa\|\|f_n^\perp\|_{L^2} < c_1 + \frac{\|\kappa\|}{\lambda_1}u_n,$$

implying that u_n^2 is uniformly bounded by $c_1 + \frac{\|\kappa\|^2}{\lambda_1^2}$

DEFINITION 5.8. Let H be a separable Hilbert space. A sequence $\{f_n\}_{n\ge1}$ *converges weakly* to $f \in H$, denoted $f_n \rightharpoonup f$, if for every $g \in H$ we have

$$\langle f_n, g\rangle \to \langle f, g\rangle \text{ as } n \to \infty.$$

A weakly convergent sequence must be bounded.

LEMMA 5.9.
1. *Every bounded sequence in H has a weakly convergent subsequence.*
2. *Every weakly convergent sequence $f_n \rightharpoonup f$ in the Sobolev space H^2 converges to f as a sequence in L^2 in the usual sense.*

Coming back to the minimizing sequence f_n for the energy functional E, we have proved that it must be bounded in H^1 by some constant C. We deduce that there exists $f^\perp \in H^1(S)$ (the first Sobolev space on S) so that, after choosing a subsequence,

$$f_n^\perp \overset{H^1}{\rightharpoonup} f^\perp, \qquad\qquad f_n^\perp \overset{L^2}{\to} f^\perp.$$

This implies that $e^{f_n^\perp}$ converges in L^1 (and hence also in L^2) to $e^{f^\perp}$. Indeed, using the inequality

$$|e^x - 1| \le |x|e^{|x|},$$

we get

$$\int_S |e^{g_n} - e^g|\mu_h \le \int_S |g_n - g|e^{|g_n-g|}\mu_h \le \|g_n - g\|_{L^2}\|e^{|g_n-g|}\|_{L^2}.$$

The first factor converges to 0 as $n \to \infty$, while the second is uniformly bounded by (5.6), proving that $\lim E_2(f_n) = E_2(f)$.

Regarding E_1, we notice first that $\langle g_n, \kappa\rangle \to \langle g, \kappa\rangle$. Next, we have

$$0 \le \|dg_n - dg\|_{L^2}^2 = \|dg_n\|^2 - 2\langle dg_n, dg\rangle + \|dg\|^2.$$

The middle term converges to $-2\|dg\|^2$ since $f_n^\perp \overset{H^1}{\rightharpoonup} f^\perp$. Thus $\|dg\|^2 \le \liminf \|dg_n\|^2$, therefore $E_1(g) \le \liminf E_1(g_n)$.

But f_n was a minimizing sequence for $E = E_1 + E_2$, entailing that $E(f)$ realizes the infimum of E.

In turn, since $E_2(f_n)$ is convergent, we deduce that the zero modes $(f_n)_0$ also converge to some f^0.

We have thus obtained a H^1 function f which is a minimum of the functional E. It follows that f is a weak solution for (5.3).

Inductively we deduce that $f \in \cap_{k\in\mathbb{N}} H^k(S)$. Indeed, since $\Delta f \in L^2$, by elliptic regularity we have $f \in H^2$. This in turn implies $e^{2f} \in H^1$. Continuing in this way by a boot-strap argument, $f \in H^k$ implies $f \in H^{k+1}$. By the Sobolev injections, f is smooth. Thus the equation (5.3) has a solution.

5.5. Uniqueness of the solution. By the existence part, we can work with a constant (-1) curvature metric conformal to the initial metric, thus we assume $\kappa = -1$. Equation (5.3) becomes

$$\Delta f + 1 = e^{-2f}.$$

We want to show that f satisfying the above equation must be identically 0. Since S is compact, f attains its maximum and minimum, let p be a maximum point.

LEMMA 5.10. *Let $p \in S$ be a local maximum for a function $f : S \to \mathbb{R}$, where S is a surface with a Riemannian metric. Then $\Delta f(p) \geq 0$.*

PROOF. In dimension 1 we know that $-f''(p) \geq 0$ if p is a local maximum. Here we choose geodesic normal coordinates centered at p, so $\Delta f(p) = -\partial_x^2 f(p) - \partial_y^2 f(p)$. The point p is a local maximum point for the coordinate curves passing through p so both components of $\Delta f(p)$ are non-negative. $\square$

By Lemma 5.10, $\Delta f(p) \geq 0$. This implies $e^{-2f(p)} \geq 1$ so $-2f(p) \geq 0$ or in other words the maximum of f is non-positive. By the same argument, the infimum of f is non-negative, which implies that f is identically 0 as claimed.

Note that by other arguments, we show in Section 6.4 that there exists at most one complete hyperbolic metric in each conformal class on any surface (not necessaryly compact).

5.6. Exercises. All Hilbert spaces below are assumed to be separable.

EXERCISE 5.11. Prove that a weakly convergent sequence in a Hilbert space must be bounded.

EXERCISE 5.12. Conversely, show that a bounded sequence in a Hilbert space has a weakly convergent subsequence.

EXERCISE 5.13. Let $\lambda_j, j \in \mathbb{N}$ be an increasing sequence of strictly positive real numbers with limit ∞. Let l^2 be the Hilbert space of square-summable sequences $(x_j)_{j\in\mathbb{N}}$, $x_j \in \mathbb{R}$:

$$\langle x, y\rangle_{l^2} := \sum_{i=0}^{\infty} x_i y_j.$$

Define H^1 as the subspace of those sequences $x \in l^2$ such that $(\lambda_j x_j)_{j\in\mathbb{N}}$ is again in l^2. Define a scalar product on H^1 by

$$\langle x, y\rangle_{H^1} := \sum_{i=0}^{\infty} \lambda_i^2 x_i y_j.$$

Prove that the inclusion $H^1 \hookrightarrow l^2$ is *compact* in the following sense: every bounded sequence in H^1 has a subsequence which is convergent (or equivalently, Cauchy) in l^2.

6. Uniformisation revisited

Let us summarize the results derived in this notes about compact surfaces. We have shown first that every compact orientable surface S is homeomorphic to a sphere with g handles attached, g being the genus of S, linked to the Euler characteristic by the formula

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

Our proof used a triangulation for S. From any triangulation we constructed a smooth structure, and with a little more effort even a complex structure.

We have then shown that a complex structure in an orientable real plane bundle is the same as a conformal structure. Inside each conformal class over a torus we proved, by solving the Laplace equation with the help of the Gauss-Bonnet formula, that there exist flat metrics, unique up to a constant. Such metrics are locally isometric to $\mathbb{R}^2$. Since every metric on a surface is locally isometric to a metric on a torus, the above argument gives us isothermal coordinates adapted to a given conformal structure. This proves that every almost complex structure on a surface is integrable.

6.1. Conformal structures on the sphere. The first statement of the uniformization theorem, in genus 0, implies that any two smooth structures on a topological sphere differ by a diffeomorphism: indeed, pick Riemannian metrics, then we proved that there exists a conformal diffeomorphism between the two metrics. Nevertheless, in any given conformal class on the sphere there exist many metrics of curvature 1, parametrized by the symmetric space $\mathrm{SO}(3)\backslash\mathrm{PSL}_2\mathbb{C}$.

We are going to give a infinitesimal proof of this fact below, in the spirit of Teichmüller theory.

6.2. Moduli space of structures on the torus. Going back to the torus, we deduce that every conformal class on $\mathbb{T}^2$ defines uniquely a complex structure. Moreover, inside the given conforma class there exists a unique flat metric up to a homothety. By lifting the metric to the universal cover $\tilde{S}$ and choosing an isometry of $\tilde{S}$ to $\mathbb{R}^2$, the conformal structure determines therefore a rank-2 free discrete subgroup Γ in $\mathbb{R}^2$ modulo multiplication by constants in $\mathbb{C}^*$. Conversely, two flat metrics which give rise to subgroups which differ by a dilation in $\mathbb{C}^*$ are clearly conformal to each other. Thus the *moduli space* of complex structures on the torus modulo diffeomorphisms can be identified with the space of free abelian subgroups of rank 2 in $\mathbb{R}^2$ modulo complex homotheties. The algebraic classification of these subgroups is quite simple: the generator of smallest length can be made (by multiplying with a complex number in $\mathbb{C}^*$) to be 1, then the least non-real element belongs to the standard fundamental region of the modular group $\mathrm{PSL}_2(\mathbb{Z})$, so in conclusion the space of complex structures (or of conformal structures, or of flat structures up to homothethy) modulo diffeomorphisms is in bijection with the modular curve $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbb{H}^2$. If we mod out only by isotopic diffeomorphisms (i.e., homotopic inside the space of diffeomorphisms), we obtain the Teichmüller space $\mathcal{T}_1$ in genus 1, which is identified with $\mathbb{H}^2$.

6.3. Moduli space of higher genus. For surfaces of genus $g \geq 2$, we showed that every conformal class contains precisely one hyperbolic metric. The proof of this statement was more difficult, since we had to solve a non-linear Laplace equation. Such a hyperbolic metric defines a discrete co-compact subgroup Γ in $\mathrm{PSL}_2(\mathbb{R})$, unique up to conjugation, such that S is isometric to $\Gamma \backslash \mathrm{PSL}_2(\mathbb{R})$. Two hyperbolic metrics which yield conjugate co-compact subgroups in $\mathrm{PSL}_2(\mathbb{R})$ must be isometric. Thus conjugation classes of genus g co-compact subgroups in $\mathrm{PSL}_2(\mathbb{R})$ are in bijection with conformal classes of metrics on S modulo $\mathrm{Diff}(S)$, and also with the set of complex structures on S modulo $\mathrm{Diff}(S)$. This set is called the *moduli space* of complex structures on S, and it has a complex analytic structure.

6.4. Uniformisation of general surfaces. In this text we have focused on compact orientable surfaces. A famous result, which took the second half of the 19th century to be completed, asserts that a connected, simply connected surface Σ with a fixed conformal class is bi-holomorphic to one of the following surfaces:

- S^2, if Σ is compact;
- $\mathbb{H}^2$, if there exist on Σ nonconstant bounded holomorphic functions;
- $\mathbb{C}$, otherwise.

This theorem, attributed to Poincaré and Koebe (1907) implies in a rather easy way the uniformisation of compact surfaces and much more: it allows us to uniformize *every* connected Riemann surface by a *complete* metric of constant curvature.

The simply connected compact case is just case (1) of Theorem 4.3 in this text. Let S be a connected non-compact Riemann surface. Then by the Poincaré-Koebe theorem its universal cover $\tilde{S}$ is bi-holomorphic to either $\mathbb{H}^2$ or $\mathbb{C}$. Assume that we are in the first case, i.e., $\tilde{S} = \mathbb{H}^2$. We obtain a representation of the fundamental group of S in the group of bi-holomorphisms of $\mathbb{H}^2$ or $\mathbb{C}$. By Theorem 2.8, every automorphism of $\mathbb{H}^2$ is an isometry for the hyperbolic metric. Hence the deck transformation group acts *by isometries*, which means that the uniform metric on the model space $\mathbb{H}^2$ descends to S! Since $\mathbb{H}^2$ is complet so are its quotients. In conclusion, S admits a complete hyperbolic metric.

Assume now that $\tilde{S}$ is $\mathbb{C}$. The automorphism group of $\mathbb{C}$ is given by affine transformations, $z \mapsto az + b$ with $a, b \in \mathbb{C}$, $a \neq 0$. As before, we obtain a representation of $\pi_1(S)$ in this group. Moreover, the action of $\pi_1(S)$ must be free, and this means that in fact the image $\pi_1(S)$ lives in the group of translations, $z \mapsto z + b$ which are isometries for the Euclidean metric on $\mathbb{C}$. Therefore, S inherits a flat complete metric.

We can actually describe all S in this latter case. A group of translations acts discretely if and only if it is either trivial (then $S = \mathbb{C}$), or infinite cyclic (then S is a cylinder) or is free abelian with two generators, in which case S is a torus.

THEOREM 6.1. *Let S be a non-compact Riemann surface. Then exactly one of the following affirmations are true:*

- *There exists on S a complete hyperbolic metric;*
- *There exists on S a complete flat metric.*

In the first case, the metric is unique. In the second case, the metric is unique up to constant dilations, and S is $\mathbb{C}$ or a cylinder.

PROOF. The existence is contained in the discussion above, essentially based on the Poincaré-Koebe theorem. Assume that on a hyperbolic quotient $\Gamma\backslash\mathbb{H}^2$ there exists another complete hyperbolic metric $g = e^{2f} g_S$. Then by the comparison theorem 3.6, the lift of g to the universal cover $(\mathbb{H}^2, \tilde{g})$ is isometric to $(\mathbb{H}^2, g_{\mathbb{H}^2})$. Let Φ be an isometry, i.e., Φ is a diffeomorphism of $\mathbb{H}^2$ onto itself with $\Phi^* g_{\mathbb{H}^2} = \tilde{g}$. In particular, Φ is a conformal map, since $\tilde{g} = e^{2\tilde{f}} g_{\mathbb{H}^2}$ where $\tilde{f}$ is the lift of f. Theorem 2.8 implies that Φ is a hyperbolic isometry, so $\Phi^* g_{\mathbb{H}^2} = g_{\mathbb{H}^2}$. We see that $\tilde{f} = 0$ and so $f = 0$, which entails uniqueness of the complete hyperbolic metric in its conformal class.

If S is a quotient of the plane by translations, let as before $g = e^{2f} g_S$ be another complete flat metric in its conformal class. We can assume that $\tilde{S}$ is $\mathbb{C}$ with its standard metric $g_{\mathbb{C}}$. Then we get an isometry $\Phi : (\mathbb{C}, g_{\mathbb{C}}) \to (\mathbb{C}, e^{2\tilde{f}} g_{\mathbb{C}})$. Since automorphisms of $\mathbb{C}$ are affine transformations, we deduce that $\tilde{f}$ is constant and so f is also constant. $\square$

Connected Riemann surfaces can therefore be classified in terms of their natural uniform metric as follows:

- the sphere;
- the complex plane;
- cylinders, i.e., quotients of $\mathbb{C}$ by an infinite cyclic group of translations;
- flat tori;
- complete hyperbolic surfaces.

On the sphere, we have remarked that there is a family with 3 parameters of spherical metrics in each conformal class. Over flat surfaces the complete flat metric is unique up to dilations; and in the hyperbolic case, the complete hyperbolic metric is unique in each conformal class.

For non-orientable surfaces with a conformal class, the classification includes:

- the real projective plane;
- Möbius bands;
- flat Klein bottles;
- complete non-orientable hyperbolic surfaces.

The method of proof of the Poincaré-Koebe theorem is entirely diferent from the analytical methods here but it is nevertheless necessary to understand at least the statement of the above classification.

7. Teichmüller theory in genus $g \geq 2$

CHAPTER 4

The Selberg trace formula

1. Classification of homographies

A matrix $A \in \mathrm{SL}_2(\mathbb{R})$ different from the identity is called

- *hyperbolic*, if $\mathrm{tr}^2(A) > 4$;
- *parabolic*, if $\mathrm{tr}^2(A) = 4$;
- *elliptic*, if $\mathrm{tr}^2(A) < 4$.

This definition is invariant under multiplication by -1, hence every element different from the identity in $\mathrm{PSL}_2(\mathbb{R})$ is of one of these three types.

EXAMPLE 1.1. For every $a > 0$, $b \in \mathbb{R}^*$ and $\theta \in \mathbb{R}$ consider the matrices

$$d_a = \begin{bmatrix} \sqrt{a} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{bmatrix}, \qquad \tau_b = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}, \qquad \rho_\theta = \begin{bmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}). \end{bmatrix}$$

Then

- d_a is hyperbolic, and the associated homography is the dilation by a factor of a;
- τ_b is parabolic and induces the translation by b on $\mathbb{H}$;
- ρ_θ is elliptic, fixes $i \in \mathbb{H}$ and acts like the rotation of angle θ on $T_i\mathbb{H}$.

It turns out that every homography is conjugated inside $\mathrm{PSL}_2(\mathbb{R})$ to one from the above example.

PROPOSITION 1.2. *Let $\gamma \in \mathrm{PSL}_2(\mathbb{R})$, $\gamma \neq 1$.*

- *If γ is hyperbolic, there exists a unique $a > 1$ so that γ is conjugated inside $\mathrm{PSL}_2(\mathbb{R})$ to d_a.*
- *Every parabolic γ is conjugated to τ_1, the horizontal translation by 1.*
- *If γ is elliptic, it is conjugated to a unique rotation ρ_θ with $\theta \in (0, 2\pi)$.*

PROOF. Consider the action of the Möbius group $\mathrm{PSL}_2\mathbb{C}$ on the Riemann sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$:

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

Let now $\gamma \in \mathrm{PSL}_2(\mathbb{R}) \subset \mathrm{PSL}_2\mathbb{C}$ and look for fixed points for the action of γ on $\hat{\mathbb{C}}$: first, ∞ is a fixed point if and only if $c = 0$. In that case, a complex number z is a fixed point, $\gamma z = z \iff (d-a)z - b = 0$. If $b \neq a$, there is a unique solution $z_0 = b/(d-a) \in \mathbb{R}$. If $d = a$ (hence $a = b = \pm 1$), we distinguish two cases: either $b = 0$, which implies $\gamma = 1$ and every point is fixed by γ, or $b \neq 0$, in which case the only fixed point in $\hat{\mathbb{C}}$ is ∞. This last case holds when γ is a translation.

Assuming $c \neq 0$, for $z \in \mathbb{C}$,

$$\frac{az+b}{cz+d} = z \iff cz^2 + (d-a)z - b = 0.$$

The discriminant of this equation is $(d-a)^2 + 4bc = (d+a)^2 - 4 = \mathrm{tr}^2\gamma - 4$ (using $\det\gamma = 1$), hence:

- for $\mathrm{tr}^2\gamma > 4$ there are two distinct real fixed points;
- for $\mathrm{tr}^2\gamma = 4$ there is one double real solution;
- for $\mathrm{tr}^2\gamma < 4$ there are two complex conjugate solutions, hence precisely one solution in $\mathbb{H}$.

Now use the fact that $\mathrm{PSL}_2(\mathbb{R})$ acts transitively on triples of distinct points in $\hat{\mathbb{R}}$, so in particular on pairs: it follows that a transformation γ with two real fixed points can be conjugated to a transformation fixing 0 and ∞. Similarly, a transformation with only one fixed point in $\hat{\mathbb{R}}$ is conjugate to a translation. Finally, since $\mathrm{PSL}_2(\mathbb{R})$ acts transitively on $\mathbb{H}$, a transformation with a unique fixed point in $\mathbb{H}$ is conjugate to a transformation preserving i, hence a rotation. $\square$

2. Dirichlet fundamental domains for compact hyperbolic surfaces

THEOREM 2.1. *A subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ acts properly discontinuously on $\mathbb{H}$ if and only if it is discrete with respect to the topology on $\mathrm{SL}_2(\mathbb{R}) \subset \mathrm{GL}_2(\mathbb{R})$ induced from $\mathbb{R}^4$, and does not contain elliptic elements.*

PROOF. The action is continuous, so if $\gamma_n \to 1$ in $\mathrm{PSL}_2(\mathbb{R})$ then $\gamma_n z \to z$ for every $z \in \mathbb{H}$. Therefore, if Γ is not discrete, the action is clearly not properly discontinuous. Assume now Γ is discrete, and suppose by contradiction that there exists $z \in \mathbb{H}$ such that for every ball V_n of radius $1/n$ centered in z, there exists $\gamma_n \in \Gamma$ and $z_n \in V_n$ so that $\gamma_n z_n \in V_n$. Hence $z_n \to z$, $\gamma_n z_n \to z$. Since γ_n are isometries, we get $\gamma_n z \to z$. By conjugating via a fixed $\alpha \in \mathrm{PSL}_2(\mathbb{R})$ mapping i to z, we get $\alpha^{-1}z = i$ and $\alpha^{-1}\gamma_n\alpha\alpha^{-1}z \to i$, hence we may assume without loss of generality that $z = i$ and $\gamma_n i \to i$.

Write $\gamma_n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$ and isolate the real and imaginary parts of $\gamma_n i$. We get

$$\frac{a_n d_n - b_n c_n}{c_n^2 + d_n^2} \to 1, \qquad \frac{a_n c_n + b_n d_n}{c_n^2 + d_n^2} \to 0.$$

Since $\det\gamma_n = 1$, the first convergence implies that the vectors $(c_n, d_n) \in \mathbb{R}^2$ are inside the unit ball. By compactness, we may extract a convergent subsequence with limit $(c, d) \in S^1$, again indexed by n for convenience. Assume for instance $c \neq 0$ (the case $d \neq 0$ is similar). Then the second convergence above and $\det\gamma_n = 1$ imply

$$a_n + b_n\tfrac{d}{c} \to 0, \qquad a_n\tfrac{d}{c} - b_n \to \tfrac{1}{c}.$$

Together, these two limits show that $(1+d^2/c^2)a_n$ has a finite limit, and so γ_n converges to some matrix, necessarily of determinant 1. $\square$

DEFINITION 2.2. A discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ is called a *Fuchsian group*.

We will be interested in Fuchsian groups without elliptic elements, so that the quotient $\Gamma\backslash\mathbb{H}$ becomes a smooth hyperbolic surface.
Let $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ be a Fuchsian group and $p \in \mathbb{H}$.

DEFINITION 2.3. The *Dirichlet domain* for the action of Γ on $\mathbb{H}$ is the set

$$F = F_p = \{z \in \mathbb{H}; d(z,p) \leq d(z,\gamma p) \text{ for all } \gamma \in \Gamma\}.$$

Directly from the definition, this is a closed, convex subset in $\mathbb{H}$. It is a *fundamental domain* for the action of Γ in the following sense:

- the projection $\Phi : \mathbb{H} \to S$ maps F onto S surjectively;
- the restriction of the projection Φ to the interior of F is injective;
- the set
$$\{z \in S; \Phi^{-1}(z) \text{ contains at least two points in } F\}$$
is of measure 0 in S.

LEMMA 2.4. *Let Γ be a Fuchsian group and $p \in \mathbb{H}$. For every $r > 0$ let*

$$m_\Gamma(r) = \# B_r(p) \cap \Gamma p$$

(i.e., the number of points in the orbit Γp situated at distance at most r to the base point p). Then there exists $C > 0$ independent of p and r such that

$$m_\Gamma(r) \leq Ce^r.$$

PROOF. There exists $r_0 > 0$ such that the ball of radius r_0 centered at p is disjoint from $B_{\gamma p}(r_0)$ for every $\gamma \in \Gamma^*$ (because Γ acts properly discontinuously). Hence the balls $B_{\gamma p}(r_0)$ for $\gamma p \in B_p(r)$ are disjoint and contained in $B_p(r + r_0)$ by the triangle inequality. It is easy to compute the area of a hyperbolic ball (see Lemma 3.4 below):

$$\mathrm{Area}(B_r(p)) = 2\pi(\cosh r - 1).$$

It follows by comparing the areas that

$$m_\Gamma(r)(\cosh r_0 - 1) < \cosh(r + r_0) - 1.$$

$\square$

PROPOSITION 2.5. *Let $S = \Gamma\backslash\mathbb{H}$ be a compact hyperbolic surface. Then*

- *$\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ is a Fuchsian subgroup without elliptic and parabolic elements.*
- *the Dirichlet fundamental domain F defined with respect to some $p \in \mathbb{C}$ is a compact polygon.*

PROOF. For every $z \in S$, there exists a neighborhood V_z adapted to the covering map $\mathbb{H} \to S$. Extract a finite cover $V_{z_1}, \dots, V_{z_n}$ of S. Let δ be a Lebesgue number for this cover, so for every $z \in S$ the δ-ball $B_z(\delta)$ is contained in some open set from the cover. This means in particular that for every $z \in \mathbb{H}$, the ball $B_z(\delta)$ projects bijectively onto its image in S, and so it does not contain conjugate points.

However, if γ is a parabolic element, then there exist pairs of points $(z, \gamma z)$ arbitrarily close to each other: conjugate γ to the translation τ_1, then notice that $d(iy, iy+1) \to 0$ as $y \to \infty$. This proves the first statement.

The Dirichlet domain is a countable intersection of hyperbolic half-planes $\mathcal{P}_\gamma$, defined for every $\gamma \in \Gamma^*$ by

$$\mathcal{P}_\gamma = \{z \in \mathbb{H}; d(z,p) \leq d(z, \gamma p)\}.$$

Therefore F is convex.

Let $d := \sup\{d(z,p); z \in S\}$. Since S is compact, d must be finite. Every point in S lives at distance at most d to p, therefore every point in $\mathbb{H}$ lives at distance at most d to the orbit Γp. Let $A := B_{3d}(p) \cap \Gamma p$. By Lemma 2.4, this set is finite. Define

$$F_{3d} = \bigcap_{\gamma p \in B_{3d}(p)} \mathcal{P}_\gamma \subset F.$$

The intersection of this set with the annulus $\{z \in \mathbb{H}; d < d(z,p) \leq 2d\}$ is empty: indeed, for every point z in the annulus, there exists a point γp in the orbit Γp situated at distance at most d from z. By the triangle inequality, $0 < d(p, \gamma p) \leq 3d$. Thus $z \notin \mathcal{P}_\gamma$ for some $\gamma p \in B_{3d}(p)$, and so $z \notin F_{3d}$.

Hence, the convex set F_{3d} does not meet the annulus, on the other hand it contains p, so in conclusion $F_{3d} \subset B_d(p)$. Since it is bounded, the closed set F_{3d} must be compact. A compact finite intersection of half-planes is a polygon.

For every γ such that $\gamma p \notin B_{3d}(p)$, clearly $B_d(p)$ is contained in the half-plane $\mathcal{P}_\gamma$. Thus $F_{3d} \cap \mathcal{P}_\gamma = F_{3d}$, and so $F = F_{3d}$ is a polygon. □

Therefore for compact surfaces the Dirichlet domain is particularly easy to understand. Below we will use the following obvious fact: the hyperbolic plane $\mathbb{H}$ decomposes as the union over $\gamma \in \Gamma$ of all the translates $\gamma F$of F. These translates can intersect only in the interior of an edge (two by two) and in vertices (with finite incidence). In particular, the intersections have zero measure, so for every L^1 function f on $\mathbb{H}$ we have

$$\int_{\mathbb{H}} f dg_{\mathbb{H}} = \sum_{\gamma \in \Gamma} \int_{\gamma F} f dg_{\mathbb{H}}.$$

3. Eigenvalues of the Laplacian on $\mathbb{H}^2$

Let $\phi : \mathbb{R} \to \mathbb{C}$ be a compactly-supported smooth function. Define the convolution operator $\Phi : C^\infty(\mathbb{H}) \to C^\infty(\mathbb{H})$ by

$$(\Phi f)(z) = \int_{\mathbb{H}} g(z,z')f(z')dg_{\mathbb{H}}(z'), \qquad g(z,z') = \phi(2\cosh d(z,z') - 2). \tag{3.1}$$

We are interested in the behaviour under convolution of eigenfunctions of $\Delta = \Delta_H$.

EXAMPLE 3.1. Let $f_s : \mathbb{H} \to \mathbb{C}$, $f_s(z) := y^{\frac{1}{2}-is}$. By a direct computation using (4.1), this function is an eigenfunction for Δ of eigenvalue $\lambda = s^2 + \frac{1}{4}$. Note for further use that $f_s(i) = 1$.

THEOREM 3.2. *For every $s \in \mathbb{C}$ there exists $u(s) \in \mathbb{C}$ with the following property: If f is a smooth function on $\mathbb{H}$ such that*

$$\Delta f = (s^2 + \tfrac{1}{4})f,$$

then $\Phi f = u(s)f$.

Such a function u is necessarily even in s. Let us first remark that Theorem 3.2 is equivalent to a seemingly weaker statement:

THEOREM 3.3. *For every $s \in \mathbb{C}$ there exists $u(s) \in \mathbb{C}$ with the following property: If f is a smooth function on $\mathbb{H}$ such that*

$$\Delta f = (s^2 + \tfrac{1}{4})f,$$

then $(\Phi f)(i) = u(s)f(i)$.

Set $\lambda = \lambda(s) = (s^2 + \frac{1}{4})$. Clearly, Theorem 3.2 implies Theorem 3.3. Conversely, assume that $(\Phi f)(i) = u(\lambda)f(i)$ for every λ-eigenfunction f of the Laplacian. We would like to see that $(\Phi f)(z) = u(\lambda)f(z)$ for every $z \in \mathbb{H}$. Since the isometry group of $\mathbb{H}$ acts transitively, let $\gamma \in \mathrm{PSL}_2(\mathbb{R})$ be such that $\gamma(i) = z$. Let $f_\gamma = \gamma^* f$. Since γ is an isometry, the pull-back operator γ^* commutes both with $\Delta = d^*d$ and with the convolution operator Φ. Hence f_γ is a λ-eigenfunction f of the Laplacian, so by Theorem 3.3 we have

$$(\Phi f_\gamma)(i) = u(\lambda)f_\gamma(i).$$

This is the same as the conclusion of Theorem 3.2 at the arbitrarily chosen point z. It is thus enough to prove Theorem 3.3.

LEMMA 3.4. *The hyperbolic plane is isometric with $\mathbb{R}^2$ endowed with the metric written in polar coordinates*

$$g = dt^2 + \sinh^2(t)d\theta^2.$$

PROOF. This is just the pull-back of the hyperbolic metric via the exponential map in one point, written in polar coordinates. Let us give a direct proof. We know that $\mathbb{H}^2$ is isometric to $\mathbb{B}$ endowed with the metric (2.6). In polar coordonates, this metric equals

$$g = \frac{4(dr^2 + r^2d\theta^2)}{(1-r^2)^2}.$$

Use the new variable $t = t(r) = \log \frac{1+r}{1-r} \in (0, \infty)$, such that $dt = \frac{2dr}{1-r^2}$. Then one computes $\frac{2r}{1-r^2} = \sinh t$ as claimed. $\square$

PROOF OF THEOREM 3.3. We work in the model of the hyperbolic plane given by the above lemma. Let P_0 be the operator of projection onto the zero Fourier mode in the variable θ:

$$P_0 : C^\infty(\mathbb{R}^2, \mathbb{C}) \to C^\infty(\mathbb{R}^2, \mathbb{C}), \qquad (P_0 f)(z) = \int_0^1 f(e^{2\pi i\alpha} z) d\alpha.$$

Clearly $P_0^2 = P_0$, $P_0 f \in C^\infty(\mathbb{R}^2, \mathbb{C})$, and $(P_0 f)(0) = f(0)$. In polar coordinates

$$P_0 f(t, \theta) = \int_0^1 f(t, \theta + 2\pi\alpha) d\alpha$$

so $P_0 f$ is a radial function (it does not depend on θ).

LEMMA 3.5. *The projector P_0 commutes with Δ.*

PROOF. The operator P_0 is obtained by averaging f along the orbits of the group S^1 acting isometrically by rotations on $(\mathbb{R}^2, dt^2 + \sinh^2(t)d\theta^2)$. Since Δ is functorially defined in terms of the metric, the conclusion is rather clear. Let us give nevertheless a direct proof. The laplacian of the metric from Lemma 3.4 is given by

$$\Delta = -\tfrac{1}{\sinh t}(\partial_t \sinh(t)\partial_t + \tfrac{1}{\sinh t}\partial_\theta^2).$$

By integration by parts, the θ derivatives vanish, so

$$P_0 \Delta f = \int_0^1 -\tfrac{1}{\sinh t}(\partial_t \sinh(t)\partial_t f(te^{2\pi i\alpha}) d\alpha.$$

Similarly, since $P_0 f$ is θ-independent,

$$\Delta(P_0 f)(te^{i\theta}) = -\tfrac{1}{\sinh t}\partial_t \sinh(t)\partial_t \int_0^1 f(te^{2\pi i\alpha}) d\alpha.$$

The conclusion follows by reversing the order of integration in α and differentiation with respect to t. $\square$

For the convolution operator Φ we prove:

LEMMA 3.6. *The projector P_0 commutes with Φ at the origin, i.e.,*

$$(\Phi P_0 f)(0) = (P_0 \Phi f)(0).$$

PROOF. We have seen above that $(P_0 \Phi f)(0) = (\Phi f)(0)$. Now, by denoting $\psi(d) = \phi(2\cosh(d) - 2)$ and using $dg = \sinh(t)dtd\theta$, we compute

$$\begin{aligned}(\Phi P_0 f)(0) &= \int_{\mathbb{H}} g(z, 0) P_0 f(z) dg \\ &= 2\pi \int_{\mathbb{R}} \int_0^1 \psi(t) f(te^{2\pi i\alpha}) d\alpha \sinh(t) dt \\ &= (\Phi f)(0).\end{aligned}$$

□

We denote by $h_f : \mathbb{R} \to \mathbb{C}$ the restriction of $P_0 f$ to the real axis, so in particular $h_f(r) = (P_0 f)(re^{i\theta})$ for every θ. Note that h_f is a smooth even function.

If we assume that f is a λ-eigenfunction of Δ, it follows from Lemma 3.5 that $P_0 f$ is a also an eigenfunction for the same eigenvalue. Since $P_0 f$ is rotation-invariant, we have

$$\Delta P_0 f(t,\theta) = -\tfrac{1}{\sinh t}(\partial_t \sinh(t)\partial_t h_f)(t).$$

Hence h_f is an eigenfunction of Δ if and only if it is a solution of the second-order singular ordinary differential equation

$$h''(t) + \coth(t)h'(t) = -\lambda h(t). \tag{3.2}$$

LEMMA 3.7. *The equation 3.2 has at most a space of dimension* 1 *of solutions of class* C^1 *at* 0.

PROOF. The Wronskian of equation 3.2 satisfies the differential equation

$$W'(t) = -\coth(t)W(t),$$

hence $W(t) = \frac{c}{\sinh(t)}$ for some constant $c \neq 0$. Since this function is not continuous in 0, it follows that both fundamental solutions of 3.2 cannot be simultaneously C^1. □

A priori it is not clear that there is at least one solution of 3.2 smooth in 0. However we have at our disposal the function f_s from example 3.1, transported on the model $\mathbb{R}^2$ of the hyperbolic plane in such a way that $i \in \mathbb{H}$ is identified with $0 \in \mathbb{R}^2$. This function has the property $h_{f_s}(0) = 1$. From lemma 3.7, h_f must be proportional to h_{f_s}, and by comparing them at 0 we get $h_f = f(0)h_{f_s}$. Set

$$u(s) := (\Phi h_{f_s})(0).$$

We deduce be repeatedly using Lemma 3.6

$$(\Phi f)(0) = (\Phi h_f)(0) = f(0)(\Phi h_{f_s})(0) = f(0)(\Phi f_s)(0) = u(s)f(0).$$

This ends the proof of Theorem 3.3, and hence also of Theorem 3.2. □

3.1. Computation of $u(s)$. Once we know that the universal eigenvalue $u(s)$ exists, we would like to compute it. For this we use the function f_s from example 3.1.

Define the following operators on smooth functions with compact support in $[0,\infty)$:

$$A : C_c^\infty([0,\infty)) \to C_c^\infty([0,\infty)), \qquad A\phi(x) = \int_0^\infty \phi(x+y^2)dy,$$

$$B : C_c^\infty([0,\infty)) \to C_c^\infty([0,\infty)), \qquad B\psi(x) = -\tfrac{4}{\pi}\int_0^\infty \psi'(x+y^2)dy. \tag{3.3}$$

An easy computation using polar coordinates shows that

$$AB = BA = I$$

(the identity operator on $C_c^\infty([0,\infty))$).

THEOREM 3.8. *Let $\phi \in C_c^\infty(\mathbb{R})$, and Φ the convolution operator defined in* (3.1). *Define*

$$v(t) := 2(A\phi)(4\sinh^2(\tfrac{t}{2})), \qquad u = \hat{v}, \tag{3.4}$$

where $\hat{v}$ denotes the Fourier transform of v. For every $\lambda \in \mathbb{C}$ and $f \in C^\infty(\mathbb{H})$ with $\Delta f = (s^2 + \frac{1}{4})f$, we have

$$\Phi f = u(s) f.$$

PROOF. Let $f = f_s = y^{\frac{1}{2} - is}$ be the function from Example 3.1. From Theorem 3.2 we know that $u(s)$ exists. Moreover, since $\Delta f = (s^2 + \frac{1}{4})f$ and $f(0) = 1$, we have $u(s) = (\Phi f)(i)$. We compute

$$\begin{aligned}\Phi f(i) &= \int_{\mathbb{H}} \phi(2\cosh(d(i,z)) - 2) f(z) dg_{\mathbb{H}}(z)\\ &= \int_{\mathbb{R}} \int_0^\infty \phi\left(\frac{x^2 + (y-1)^2}{y}\right) y^{\frac{1}{2} - is} \frac{dx \wedge dy}{y^2}\end{aligned}$$

so by changing variables $X = xy^{-\frac{1}{2}}$ and $y = e^t$,

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \phi(X^2 + 4\sinh^2(\tfrac{t}{2})) e^{-ist} dX dt.$$

We first carry out the integral in the variable X: the integrand is even in X, so we obtain $2(A\phi)(4\sinh^2(\frac{t}{2})) = v(t)$. The integral in t is the definition of the Fourier transform of v. $\square$

4. Eigenvalues of the Laplacian on compact Riemanian manifolds

Let us recall below the properties of the Laplacian on a closed (ie., compact without boundary) Riemannian manifold.

THEOREM 4.1. *Let (M, g) be a closed Riemannian manifold of dimension n, and Δ the Laplace operator. There exists an increasing sequence $0 = \lambda_0 < \lambda_1 \leq \ldots$ converging to infinity, and a corresponding sequence $\{f_j\}_{j \in \mathbb{N}}$ of smooth real-valued functions on M, such that:*

1. $\Delta f_j = \lambda_j f_j$;
2. $\{f_j\}_{j\in\mathbb{N}}$ *form an orthonormal Hilbert basis in $L^2(M, g)$.*
3. *If $h \in C^\infty(M)$, the series $\sum_j \langle h, f_j \rangle f_j$ converges in C^∞ sense to h.*
4. *The growth of the eigenvalues is restricted: there exists $C > 1$ such that for every $\lambda \in \mathbb{R}$, the eigenvalue counting function $N(\lambda) = \max\{j \in \mathbb{N}; \lambda_j < \lambda\}$ satisfies*

$$C^{-1}\lambda^{n/2} < N(\lambda) < C\lambda^{n/2}.$$

This result is standard, see for instance [**?**].

Since the eigenvalues converge to infinity, it follows that each of them has finite multiplicity. In the corresponding eigenspace there is a choice of orthonormal basis. The family of eigenfunctions $\{f_j\}_{j\in\mathbb{N}}$ is uniquely determined up to these choices. Note that even in a 1-dimensional eigenspace, there exist two functions of L^2 norm equal to 1.

5. Eigenfunctions of the Laplacian on a compact hyperbolic surface

Let $S = \Gamma\backslash\mathbb{H}^2$ be a compact oriented surface of genus $g \geq 2$, endowed with its unique hyperbolic metric. Let $\tilde{f}_j$ be the lift of the eigenfunction f_j to $\mathbb{H}$, thus $\Delta\tilde{f}_j = \lambda_j\tilde{f}_j$. From Theorem 3.8, $\tilde{f}_j$ is also an eigenfunction of the convolution operator Φ:

$$u(s_j)\tilde{f}_j = \Phi\tilde{f}_j \tag{5.1}$$

for any solution s_j of the equation $s_j^2 + \frac{1}{4} = \lambda_j$. Write

$$g(z, z') := \phi(2\cosh d(z, z') - 2),$$

so g is a function depending only on the hyperbolic distance between z and z'. Since $\tilde{f}_j$ is Γ-invariant, we can write

$$\begin{aligned}
\Phi\tilde{f}_j(z) &= \int_{\mathbb{H}} g(z, z')\tilde{f}_j(z')dg_{\mathbb{H}}(z') \\
&= \sum_{\gamma\in\Gamma}\int_{\gamma F} g(z, z')\tilde{f}_j(z')dg_{\mathbb{H}}(z') && \text{since } \bigcup_{\gamma\in\Gamma}\gamma F = \mathbb{H} \\
&= \sum_{\gamma\in\Gamma}\int_{F} g(z, \gamma^{-1}z')\tilde{f}_j(z')dg_{\mathbb{H}}(z') && \text{since } \tilde{f}_j \text{ is } \Gamma\text{-invariant} \\
&= \int_{F} G(z, z')\tilde{f}_j(z')dg_{\mathbb{H}}(z')
\end{aligned}$$

where G is the kernel defined by

$$G(z, z') := \sum_{\gamma\in\Gamma} g(z, \gamma^{-1}z').$$

The kernel g is Γ-invariant, i.e., $g(\alpha z, \alpha z') = g(z, z')$, because $g(z, z')$ depends only on the distance between z and z'. Since G is the average of g over Γ, we see immediately that $G(\alpha z, \alpha' z') = G(z, z')$, i.e., G is $\Gamma \times \Gamma$-invariant, hence it descends to $S \times S$. Therefore, by identifying functions on S with Γ-invariant functions on $\mathbb{H}$, we get

$$u(s_j)f_j(z) = \Phi\tilde{f}_j(z) = \int_S G(z, z')f_j(z')dg_S(z').$$

In other words, if $\mathcal{G}$ denotes the smoothing operator on S

$$\mathcal{G}f(z) = \int_S G(z, z')f(z')dg_S(z'),$$

we have just proved that f_j is an $u(s_j)$-eigenfunction for $\mathcal{G}$.

The trace formula will be derived by looking at the trace of $\mathcal{G}$. First, we have a general fact:

LEMMA 5.1. *For every operator defined by a smooth kernel like $\mathcal{G}$ above, we have*

$$\operatorname{tr}(\mathcal{G}) = \int_S G(z, z)dg_S(z).$$

PROOF. The fact that $\{f_j\}$ form an orthonormal basis in $L^2(S)$ implies that for every $h \in C^\infty(S)$,

$$h = \sum_j \langle h, f_j \rangle f_j$$

where by Theorem 4.1, the sum converges in C^∞ sense:

$$h(z) = \sum_j \int_S h(z') f_j(z') dz' f_j(z).$$

For some fixed $z'' \in S$, apply the above decomposition to $h := G(z'', \cdot)$:

$$G(z'', z) = \sum_j \int_S G(z'', z') f_j(z') dz' f_j(z).$$

The lemma follows by setting $z'' = z$ and integrating on S. □

Our particular $\mathcal{G}$ is already diagonalized by the Δ-eigenfunctions f_j, thus

$$\mathrm{Tr}(\mathcal{G}) = \sum_j u(s_j).$$

Together with the lemma, we get

$$\sum_j u(s_j) = \int_S G(z, z) dg_S(z). \tag{5.2}$$

6. The trace formula for compactly-supported functions v

THEOREM 6.1 (Selberg trace formula, version 1). *Let $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ be a co-compact Fuchsian group, $S = \Gamma \backslash \mathbb{H}^2$ the associated compact surface, and $g \geq 2$ its genus. Let $\{\lambda_j\}_{j \in \mathbb{N}}$ be the increasing sequence of eigenvalues of the Laplacian on S, and choose $s_j \in \mathbb{C}$ such that $s_j^2 + \frac{1}{4} = \lambda_j$. Let $\phi \in C_c^\infty(\mathbb{R})$ and define*

$$A\phi(x) = \int_0^\infty \phi(x + y^2) dy, \qquad v(t) = 2A\phi(4 \sinh^2\left(\tfrac{t}{2}\right)), \qquad u = \hat{v}.$$

Let $\mathcal{L}_\Gamma$ be the set of lengths of oriented primitive geodesics in S, counted with their multiplicity. Then

$$\sum_{j=0}^\infty u(s_j) = (g-1) \int_{\mathbb{R}} s u(s) \tanh(\pi s) ds + \sum_{l \in \mathcal{L}_\Gamma} \sum_{n=1}^\infty \frac{l}{2 \sinh\left(\frac{nl}{2}\right)} v(nl).$$

From the definition, v is a smooth, compactly supported, even function on $\mathbb{R}$. Its Fourier transform u is therefore an even holomorphic function on $\mathbb{C}$, with rapid decay along every horizontal line. Thus, it does not matter in the right-hand side above which square root of $\lambda_j - \frac{1}{4}$ we choose for s_j. Moreover the series $\sum_{j=0}^\infty u(s_j)$ is absolutely convergent, since λ_j grows linearly in j, while u is Schwartz.

PROOF. We start from (5.2) and compute the right-hand side. We write

$$\begin{aligned}\int_S G(z,z)dg_S(z) &= \int_F G(z,z)dg_{\mathbb{H}}(z)\\ &= \int_F \sum_{\gamma\in\Gamma} g(z,\gamma z)dg_{\mathbb{H}}(z)\\ &= \sum_{\gamma\in\Gamma}\int_F g(z,\gamma z)dg_{\mathbb{H}}(z). \qquad (6.1)\end{aligned}$$

Let $\mathcal{C}$ be the set of conjugacy classes in γ. One distinguished such class consists of the identity element. We recall that conjugacy classes of hyperbolic elements are in bijection with the set of closed oriented geodesics.

LEMMA 6.2. *Let $\gamma\in\Gamma$, $\gamma\neq 1$. The commutant subgroup of γ is infinite cyclic, and there exist a unique $n\geq 1$ and a generator μ of $C(\gamma)$ such that $\gamma=\mu^n$.*

PROOF. Use the fact that Γ is discrete. All elements γ of the group (other than the identity) are hyperbolic, so γ is conjugate inside $\mathrm{PSL}_2(\mathbb{R})$ to a dilation d_a, i.e., $\alpha\gamma\alpha^{-1}=d_a$. Conjugate Γ by the same element α of $\mathrm{PSL}_2(\mathbb{R})$, thereby assuming that $\gamma=\begin{bmatrix}\sqrt{a} & 0\\ 0 & \frac{1}{\sqrt{a}}\end{bmatrix}$. Let $\nu\in C(\gamma)$, $\nu\neq 1$. Then ν preserves the set of fixed points $\{0,\infty\}$ of γ, hence it also preserves the positive imaginary axis, the unique geodesic linking 0 to ∞. If ν transposes these two points, then it must have a fixed point along the geodesic 0∞, which is impossible since the elements of Γ^* cannot be elliptic. Hence ν belongs to the stabilizer of $\{0,\infty\}$, which is the subgroup of dilations and hence it is isomorphic to the multiplicative group $\mathbb{R}^*_+$. The intersection of this subgroup with Γ is discrete, hence it is infinite cyclic and there exists a unique generator μ such that γ is a positive power of μ. ☐

An element $\gamma\in\Gamma$ for which the number n given by Lemma 6.2 equals 1 is called *primitive*.

LEMMA 6.3. *Let $\gamma\in\Gamma^*$, $\gamma=\mu^n$ for $\mu\in\Gamma$ primitive. Choose a system of representatives A_μ for the set of cosets $\Gamma/\langle\mu\rangle$. Then the conjugacy class $[\gamma]$ of γ is in bijection with A_μ:*

$$[\gamma]=\{\alpha\gamma\alpha^{-1};\alpha\in A_\mu\}.$$

PROOF. By definition

$$[\gamma]=\{\alpha\gamma\alpha^{-1};\alpha\in\Gamma\}.$$

But it is evident that $\alpha\gamma\alpha^{-1}=\beta\gamma\beta^{-1}$ if and only if $\alpha^{-1}\beta\in C(\gamma)$. By Lemma 6.2, the commutant subgroup $C(\gamma)$ is the cyclic group $\langle\mu\rangle$. Thus by restricting α to a set of representatives of the right-cosets of $\langle\mu\rangle$, we get each element in the conjugacy class precisely once. ☐

Denote by $\mathcal{C}^*_{\mathrm{prim}}$ the set of primitive conjgacy classes in Γ different from the identity. We return to the identity (6.1) and split the sum along conjugacy classes:

$$\int_S G(z,z)dg_S(z)=\int_F g(z,z)dg_{\mathbb{H}}(z)+\sum_{[\mu]\in\mathcal{C}^*_{\mathrm{prim}}}\sum_{n\geq 1}\sum_{\alpha\in A_\mu}\int_F g(z,\alpha\mu^n\alpha^{-1}z)dg_{\mathbb{H}}(z).$$

Fix $[\mu]$ and n, and compute the sum

$$\begin{aligned} I_{\mu,n} &:= \sum_{\alpha\in A_\mu}\int_F g(z,\alpha\mu^n\alpha^{-1}z)dg_{\mathbb{H}}(z)\\ &= \sum_{\alpha\in A_\mu}\int_F g(\alpha^{-1}z,\mu^n\alpha^{-1}z)dg_{\mathbb{H}}(z)\\ &= \sum_{\alpha\in A_\mu}\int_{\alpha^{-1}F} g(z,\mu^n z)dg_{\mathbb{H}}(z)\\ &= \int_{\cup_{\alpha\in A_\mu}\alpha^{-1}F} g(z,\mu^n z)dg_{\mathbb{H}}(z). \end{aligned}$$

The set $\cup_{\alpha\in A_\mu}\alpha^{-1}F$ is a fundamental domain for the quotient $\langle\mu\rangle\backslash\mathbb{H}$, hence

$$I_{\mu,n} = \int_{\langle\mu\rangle\backslash\mathbb{H}} g(z,\mu^n z)dz.$$

Let $\beta \in \mathrm{PSL}_2(\mathbb{R})$ be such that $\mu' := \beta\mu\beta^{-1}$ is the dilation by e^l, where l is the length of the geodesic determined by the conjugacy class $[\mu]$ (notice that l satisfies $|\mathrm{tr}(\mu)| = 2\cosh(l/2)$). Then $\beta : \langle\mu\rangle\backslash\mathbb{H} \to \langle\mu'\rangle\backslash\mathbb{H}$ is an isometry, and

$$\int_{\langle\mu\rangle\backslash\mathbb{H}} g(z,\mu^n z)dz = \int_{\langle\mu'\rangle\backslash\mathbb{H}} g(z,\mu'^n z)dz.$$

Since $\mu' z = e^l z$, choose as fundamental domain for this quotient the infinite band

$$\{z\in\mathbb{H}; 1\le \Im(z)\le e^l\}.$$

It follows

$$\begin{aligned} I_{\mu,n} &= \int_{\mathbb{R}}\int_1^{e^l} g(z,e^{nl}z)dg_{\mathbb{H}}(z)\\ &= \int_{\mathbb{R}}\int_1^{e^l}\phi\left(\frac{|z|^2(1-e^{nl})^2}{e^{nl}y^2}\right)\frac{dx\wedge dy}{y^2}\\ &= \int_{\mathbb{R}}\int_1^{e^l}\phi\left(\frac{x^2+y^2}{e^{nl}y^2}(1-e^{nl})^2\right)\frac{dx\wedge dy}{y^2}\\ &= \int_{\mathbb{R}}\int_0^{l}\phi((X^2+1)(4\sinh^2(nl/2))dX\,dt\\ &= \frac{l}{2\sinh(nl/2)}\int_{\mathbb{R}}\phi(x^2+4\sinh^2(nl/2))dx. \end{aligned}$$

We have successively changed variables $X = x/y$; $y = e^t$; then $x = X\cdot 2\sinh(nl/2)$. From the definition of the function v, the last formula is precisely

$$I_{\mu,n} = \frac{l}{2\sinh(nl/2)}2(A\phi)(4\sinh^2(nl/2)) = \frac{l}{2\sinh(nl/2)}v(nl).$$

It remains to compute the term corresponding to the unit conjugacy class, namely

$$\int_F g(z,z)dg_{\mathbb{H}}(z) = \mathrm{Area}(F)\phi(0).$$

But by the Gauss-Bonnet (theorem 2.5), $\mathrm{Area}(F) = \mathrm{Area}(S) = 2\pi(2g-2)$. We want to compute $\phi(0)$ in terms of v. For this, we write using (3.3):

$$\begin{aligned}
\phi(0) &= BA\phi(0)\\
&= -\frac{4}{\pi}\int_0^\infty (A\phi)'(y^2)dy\\
&= -\frac{4}{\pi}\int_0^\infty (A\phi)'(4\sinh^2\left(\tfrac{t}{2}\right))\cosh\left(\tfrac{t}{2}\right)dt\\
&= -\frac{4}{\pi}\int_0^\infty \frac{1}{4\sinh\left(\frac{t}{2}\right)\cosh\left(\frac{t}{2}\right)}\partial_t((A\phi)'(4\sinh^2\left(\tfrac{t}{2}\right)))\cosh\left(\tfrac{t}{2}\right)dt\\
&= -\frac{1}{2\pi}\int_0^\infty \frac{v'(t)}{\sinh\left(\frac{t}{2}\right)}dt.
\end{aligned}$$

The proof of the trace formula is ended by Lemma 6.4 below. □

LEMMA 6.4. *Let $u = \mathcal{F}(v)$ for $v \in C_c^\infty(\mathbb{R})$ an even function. Then*

$$-2\int_0^\infty \frac{v'(t)}{\sinh(\frac{t}{2})}dt = \int_{\mathbb{R}} su(s)\tanh(\pi s)ds.$$

PROOF. Write

$$v(t) = \tfrac{1}{2\pi}\int_{\mathbb{R}} e^{ist}u(s)ds, \qquad v'(t) = \tfrac{i}{2\pi}\int_{\mathbb{R}} e^{ist}su(s)ds.$$

Since v' is odd as a function of t, we get

$$v'(t) = \tfrac{i}{4\pi}\int_{\mathbb{R}}(e^{ist} - e^{-ist})su(s)ds.$$

It follows

$$\begin{aligned}
-2\int_0^\infty \frac{v'(t)}{\sinh(\frac{t}{2})}dt &= -\tfrac{i}{2\pi}\int_{\mathbb{R}}\int_0^\infty (e^{ist}-e^{-ist})\frac{dt}{\sinh(\frac{t}{2})}su(s)ds\\
&= -\tfrac{i}{\pi}\int_{\mathbb{R}} su(s)ds\int_0^\infty (e^{ist}-e^{-ist})e^{-\frac{t}{2}}\sum_{m=0}^\infty e^{-mt}dt\\
&= -\tfrac{i}{\pi}\int_{\mathbb{R}} su(s)ds\sum_{m=0}^\infty\left(\frac{1}{\frac{1}{2}+m-is} - \frac{1}{\frac{1}{2}+m+is}\right).
\end{aligned}$$

The result is a consequence of the well-known Lemma 6.5. □

LEMMA 6.5.

$$\frac{i}{\pi}\sum_{m=0}^{\infty}\left(\frac{1}{\frac{1}{2}+m+is}-\frac{1}{\frac{1}{2}+m-is}\right)=\tanh(\pi s). \tag{6.2}$$

PROOF. Let us give a proof using the Poisson summation formula. Assume $s>0$, the other case being similar. Define

$$f(x):=\frac{1}{\pi}\frac{s}{(\frac{1}{2}+x)^2+s^2},$$

and notice that the right-hand side of (6.2) equals $\sum_{m\in\mathbb{Z}}f(m)$. The Poisson summation formula reads $\sum_{m\in\mathbb{Z}}f(m)=\sum_{n\in\mathbb{Z}}\hat{f}(2\pi n)$. We compute $\hat{f}(y)$ for $y>0$ by changing the contour of integration from the line $\{\Im(x)=0\}$ to $\{\Im(x)<-s\}$ across the pole at $x=-is-\frac{1}{2}$:

$$\hat{f}(y)=\frac{1}{\pi}\int_{\mathbb{R}}e^{-ixy}\frac{sdx}{(\frac{1}{2}+x)^2+s^2}=e^{-y(s-\frac{i}{2})}$$

(notice a minus sign coming from the clockwise orientation of the contour). Both f and its Fourier transform $\hat{f}$ are even, hence for $y\in\mathbb{R}$, $\hat{f}(y)=e^{-|y|(s-i/2)}$. It follows

$$\begin{aligned}\sum_{n\in\mathbb{Z}}\hat{f}(2\pi n)=&\sum_{n\in\mathbb{Z}}e^{-2\pi|n|(s-\frac{i}{2})}\\=&\sum_{n\in\mathbb{Z}}e^{-2\pi|n|s}e^{i\pi|n|}\\=&-1+2\sum_{n\in\mathbb{N}}(-1)^n e^{-2\pi ns}\\=&\tanh(\pi s).\end{aligned}$$

□

7. Geodesics asymptotics

In order to extend the trace formula for more general functions v, we need some bound on the growth of the sequence $\mathcal{L}_\Gamma$ of lengths of primitive closed geodesics in $S=\Gamma\backslash\mathbb{H}$.

LEMMA 7.1. *Let $L(l)$ denote the number of oriented closed geodesics of length at most l, counted with their multiplicity. Then there exists a constant $C\in\mathbb{R}$ such that*

$$L(l)<Ce^l.$$

PROOF. Fix $z\in\mathbb{H}$ and define a set

$$M_\Gamma(r):=\{\gamma\in\Gamma; d(z,\gamma z)<r\}.$$

The cardinality of $M_\Gamma(r)$ is the number $m_\Gamma(r)$ from Lemma 2.4. Take any closed geodesic of length less that $r-2d$, where d is the diameter of S. Lift this geodesic to $\mathbb{H}$ starting from a point A in the interior of the Dirichlet domain F (possible since the Dirichlet domain surjects onto S). The end-point B of this lift belongs to the interior of γF for some uniquely determined $\gamma\in\Gamma$, and so $B=\gamma A$ and moreover the geodesic from A to B is the axis of the hyperbolic homography γ. Therefore, if we start from a different geodesic we obtain a different γ. From the triangle inequality, we obtain that $d(z,\gamma z) < r$, so in this way we have injected the set of geodesics of length less than $r-2d$ into the finite set $M_\Gamma(r) = B_r(z)\cap\Gamma z$ of points in the Γ-orbit of z at distance less than r to z. It follows that $L(r-2d) \leq m_\Gamma(r)$. $\square$

Let l_j be the j^{th} length in the sequence $\mathcal{L}_\Gamma$. From Lemma 7.1, $j = L(l_j) < Ce^{l_j}$, or equivalently

$$l_j > \log(j) - c \tag{7.1}$$

for some other constant c.

THEOREM 7.2 (Selberg trace formula, general statement). *The Selberg trace formula (Theorem 6.1) holds for functions v with the following properties:*

- *there exists $\epsilon > 0$ and $C \in \mathbb{R}$ so that $|v(x)| < Ce^{-|x|\left(\frac{1}{2}+\epsilon\right)}$.*
- *There exists $\epsilon > 0$ and $C \in \mathbb{R}$ so that $|u(s)| < C(1+s^2)^{-(1+\epsilon)}$.*

PROOF. Choose a cut-off function $\psi : \mathbb{R} \to [0,\infty)$, i.e., ψ is smooth, compactly supported in the interval $[-2,2]$, and equals 1 on the interval $[-1,1]$. Set $\psi_n(x) := \psi(x/n)$. Then $\psi_n v$ converges to v, and $|\psi_n(x)v(x)| < Ce^{-|x|(1+\epsilon)}$ for every n. $\square$

8. Weyl asymptotics

We apply the Selberg trace formula for the 1-parameter family of functions $u_t(s) = e^{-t(s^2+\frac{1}{4})}$. We have

$$v_t(x) = \tfrac{1}{2\pi}\int_{\mathbb{R}} e^{-t(s^2+\frac{1}{4})}e^{isx}ds = \frac{e^{-\frac{t}{4}}}{\sqrt{4\pi t}}e^{-\frac{x^2}{4t}}.$$

The trace formula reads

$$\sum_{j=0}^{\infty} e^{-t\lambda_j} = R_1(t) + R_2(t),$$

where

$$R_1(t) = (g-1)\int_{\mathbb{R}} s\tanh(\pi s)e^{-t(s^2+\frac{1}{4})}ds,$$
$$R_2(t) = \frac{e^{-\frac{t}{4}}}{\sqrt{4\pi t}}\sum_{l\in\mathcal{L}_\Gamma}\sum_{n=1}^{\infty}\frac{l}{2\sinh\left(\frac{nl}{2}\right)}e^{-\frac{n^2l^2}{4t}}.$$

Because of the rapidly decaying factor $e^{-\frac{n^2l^2}{4t}}$ as $t\to 0$, the term $R_2(t)$ converges to 0 as $t\to 0$ together with all its derivatives. By a change of variables $s\to\sqrt{t}s$, the term R_1 becomes

$$R_1(t) = \frac{(g-1)e^{-\frac{t}{4}}}{t}\int_{\mathbb{R}} s\tanh(\pi s/\sqrt{t})e^{-s^2}ds.$$

The integral converges to 1 as $t\to 0$ by Lebesgue's dominated convergence theorem. Moreover, all its derivatives vanish at $t=0$.

It follows that the *heat trace* function

$$h(t) = \sum_{j=0}^{\infty} e^{-t\lambda_j}$$

has the following asymptotic as $t \to 0$:

$$h(t) \sim (g-1)t^{-1}e^{-t}. \tag{8.1}$$

This implies the Weyl law for the asymptotic distribution of the eignevalues of Δ via the tauberain theorem of Karamata.

THEOREM 8.1 (Karamata). *Let $(\lambda_j)_{j\in\mathbb{N}}$ be a non-decreasing sequence of real numbers converging to ∞ such that*

$$\lim_{t\to 0} t \sum_{j=0}^{\infty} e^{-t\lambda_j} = 1.$$

Then the counting function $N(r) := \max\{j \in \mathbb{N}; \lambda_j \le r\}$ satisfies

$$\lim_{r\to\infty} \tfrac{N(r)}{r} = 1.$$

PROOF. By replacing t with nt, $n \in \mathbb{N}$ we get

$$\lim_{t\to 0} t \sum_{j=0}^{\infty} e^{-(n+1)t\lambda_j} = \tfrac{1}{n+1} = \int_0^1 x^n dx.$$

Therefore, for every polynomial $P \in \mathbb{R}[X]$, we have

$$\lim_{t\to 0} t \sum_{j=0}^{\infty} P(e^{-t\lambda_j})e^{-t\lambda_j} = \int_0^1 P(x)dx.$$

Using the Weierstrass density theorem, we can uniformly approximate every continuous function on the compact interval $[0,1]$ by polynomials. It follows that for every $f \in C^0([0,1],\mathbb{R})$ we have

$$\lim_{t\to 0} t \sum_{j=0}^{\infty} f(e^{-t\lambda_j})e^{-t\lambda_j} = \int_0^1 f(x)dx. \tag{8.2}$$

Take now f to be a positive function, $f(x) = 0$ for $0 \le x \le a$ and $f(x) = \frac{1}{x}$ for $b \le x \le 1$. Equation (8.2) implies that

$$\limsup_{t\to 0} tN\left(-\tfrac{\log b}{t}\right) \le -\log a, \qquad \liminf_{t\to 0} tN\left(-\tfrac{\log a}{t}\right) \ge -\log b$$

or equivalently

$$\limsup_{r\to\infty} \tfrac{N(r)}{r} \le \tfrac{\log a}{\log b}, \qquad \liminf_{r\to\infty} \tfrac{N(r)}{r} \ge \tfrac{\log b}{\log a}.$$

Since a can be chosen arbitrarily close to b, this finishes the proof. $\square$

COROLLARY 8.2. *Let λ_j be the j^{th} eigenvalue of Δ, in increasing order counted with multiplicity. Then*

$$\lim_{j\to\infty} \tfrac{\lambda_j}{j} = \tfrac{1}{g-1}.$$

PROOF. Follows from Karamata's theorem applied to the asymptotic law (8.1). $\square$

9. The Selberg zeta function

DEFINITION 9.1. Let $\mathcal{L}$ be the set of lengths of oriented closed *primitive* geodesics on $S = \Gamma\backslash\mathbb{H}$, counted with multiplicity. For $\Re(z) > 1$ define

$$Z(z) := \prod_{l\in\mathcal{L}} \prod_{m=0}^{\infty} (1 - e^{-l(z+m)}).$$

This function will turn out to extend holomorphically to $\mathbb{C}$. First we must however check that the product is absolutely convergent in the prescribed half-plane.

LEMMA 9.2. *The series*

$$\sum_{l\in\mathcal{L}} \sum_{m=0}^{\infty} e^{-l(z+m)}$$

is absolutely convergent for $\Re(z) > 1$.

PROOF. We show that in fact the sum over all geodesics, not only the primitive ones, is absolutely convergent.

$$\begin{aligned} \sum_{l\in\mathcal{L}_\Gamma} \sum_{m=0}^{\infty} e^{-l(z+m)} &< \sum_{j=1}^{\infty} \sum_{m=0}^{\infty} e^{-l_j(z+m)} \\ &= \sum_{j=1}^{\infty} \frac{e^{-l_j z}}{1 - e^{-l_j}} \end{aligned}$$

The denominator tends to 1. Using the inequality (7.1), the series of numerators is bounded by

$$\sum_{j=1}^{\infty} e^{-l_j z} < C \sum_{j=1}^{\infty} j^{-z}.$$

It is well-known that the Riemann zeta function is absolutely convergent for $\Re(z) > 1$. $\square$

Fix $z, z_0 \in \mathbb{C}$ with $\Re(z), \Re(z_0) > 1$. Define

$$u_z(s) = \tfrac{1}{s^2 + \left(z - \frac{1}{2}\right)}, \qquad u = u_z - u_{z_0} \in C^\infty(\mathbb{R}).$$

Notice that $|u(s)| \sim s^{-4}$ as $s \to \infty$. The Fourier transform of u_z is

$$v_z(t) = \tfrac{1}{2\pi} \int_{\mathbb{R}} e^{its} \tfrac{1}{s^2 + \left(z - \frac{1}{2}\right)} ds = \frac{e^{-t\left(z - \frac{1}{2}\right)}}{2\left(z - \frac{1}{2}\right)}.$$

Since $\Re(z) > 1$, this decays faster that $e^{-t\left(\frac{1}{2}+\epsilon\right)}$ for some $\epsilon > 0$, so we can apply Theorem 7.2. For each z, z_0 with real part greater than 1 we write the Selberg trace formula for the function $u = u_z - u_{z_0}$ as follows:

$$L(z, z_0) = R_1(z, z_0) + R_2(z) - R_2(z_0)$$

where

$$\begin{aligned} L(z, z_0) &= \sum_{j=1}^{\infty} \frac{1}{s_j^2 + \left(z - \frac{1}{2}\right)^2} - \frac{1}{s_j^2 + \left(z_0 - \frac{1}{2}\right)^2}, \\ R_1(z, z_0) &= (g-1) \int_{\mathbb{R}} s \tanh(\pi s) \left(\frac{1}{s^2 + \left(z - \frac{1}{2}\right)^2} - \frac{1}{s^2 + \left(z_0 - \frac{1}{2}\right)} \right)^2 ds, \\ R_2(z) &= \frac{1}{2\left(z - \frac{1}{2}\right)} \sum_{l \in \mathcal{L}_\Gamma} \sum_{n=1}^{\infty} \frac{l}{2 \sinh\left(\frac{nl}{2}\right)} e^{-nl\left(z - \frac{1}{2}\right)}. \end{aligned}$$

LEMMA 9.3.

$$2\left(z - \tfrac{1}{2}\right) R_2(z) = \frac{Z'}{Z}(z).$$

PROOF.

$$\begin{aligned} \sum_{l \in \mathcal{L}_\Gamma} \sum_{n=1}^{\infty} \frac{l}{2 \sinh\left(\frac{nl}{2}\right)} e^{-nl\left(z - \frac{1}{2}\right)} &= \sum_{l \in \mathcal{L}_\Gamma} l \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e^{-nl\left(z - \frac{1}{2}\right)} e^{-\frac{nl}{2}} e^{-mnl} \\ &= \sum_{l \in \mathcal{L}_\Gamma} l \sum_{m=0}^{\infty} (-1 + \frac{1}{1 - e^{-l(z+m)}}) \\ &= \sum_{l \in \mathcal{L}_\Gamma} \sum_{m=0}^{\infty} \frac{l e^{-l(z+m)}}{1 - e^{-l(z+m)}}. \end{aligned}$$

The lemma follows from the definition of the zeta function. □

We re-write the trace formula as follows:

$$\frac{Z'}{Z}(z) = \frac{z - \frac{1}{2}}{z_0 - \frac{1}{2}} \frac{Z'}{Z}(z_0) + (2z - 1) L(z, z_0) - (2z - 1) R_1(z, z_0). \tag{9.1}$$

We shall prove that the right-hand side (actually, every term in the right hand side) is meromorphic in $z \in \mathbb{C}$ with simple poles and *natural* residues. We are also interested in the behaviour of the right-hand side terms under the involution $z \to 1 - z$.

LEMMA 9.4. *The term* $\frac{z - \frac{1}{2}}{z_0 - \frac{1}{2}} \frac{Z'}{Z}(z_0)$ *is holomorphic on* $\mathbb{C}$ *and odd with respect to* $z \to 1 - z$.

PROOF. Evident. □

LEMMA 9.5. *The term $(2z-1)L(z,z_0)$ is meromorphic in the variable $z \in \mathbb{C}$, with simple poles at $z = \frac{1}{2} \pm is_j$. The residue of $(2z-1)L(z,z_0)$ at $z = \frac{1}{2} \pm is_j$ is equal to the multiplicity $\mu(\lambda_j)$ of the eigenvalue λ_j. In the special case where $\frac{1}{4}$ belongs to the spectrum of Δ, the poles at $z = \frac{1}{2} \pm 0$ are superposed, and so the residue equals $2\mu(\frac{1}{4})$. Moreover, $(2z-1)L(z,z_0)$ is odd with respect to $z \to 1-z$.*

PROOF. The term $(2z-1)L(z,z_0)$ is series of rational functions, absolutely convergent outside the poles. As such, the poles of the sum are precisely the union (with multiplicity) of the poles of each term. It is evident that $(2z-1)\frac{1}{s_j^2+\left(z-\frac{1}{2}\right)^2}$ has simple poles at $z = \frac{1}{2} \pm is_j$ with residue 1, with the exception of the case where $s_j = 0$. In that case, the pole remains simple but the residue is 2. The term L is even for the transformation $z \to 1-z$, while $(2z-1)$ is odd. □

The above poles are called *non-trivial* poles of the logarithmic derivative of the Zeta function.

LEMMA 9.6. *The term $-(2z-1)R_1(z,z_0)$ has simple poles at $z=-k$ for $k = 0, 1, \ldots$, with residue $2(g-1)(2k+1)$.*

PROOF. Move the contour of integration in the definition of R_1 from $\Im(s) = 0$ to $\Im(s) = k$ for $k \in \mathbb{N}^*$. The function $s \mapsto \tanh(\pi s)$ is periodic with period i. The remaining part of the integrand in R_1 decays cubically in $|s|$, hence as $k \to \infty$, the integral along the line $\{\Im(s) = k\}$ converges to 0. This procedure computes R_1 as a series over the residues of the integrand. These residues occur at $s = i\left(z - \frac{1}{2}\right)$; at $s = i\left(z_0 - \frac{1}{2}\right)$; and finally at the poles of $\tanh(\pi s)$ in the upper half-plane, namely at $s = i(n-\frac{1}{2})$ for $n = 1, 2, \ldots$:

$$\begin{aligned}-(2z-1)R_1(z,z_0) = & -2\pi i(g-1)(2z-1)\left[\tfrac{1}{2}\tanh(i\pi\left(z-\tfrac{1}{2}\right)) - \tfrac{1}{2}\tanh(i\pi\left(z_0-\tfrac{1}{2}\right))\right] \\ & -2\pi i(g-1)(2z-1)\sum_{n=1}^{\infty} i(n-\tfrac{1}{2})\tfrac{1}{\pi}\left[\frac{1}{\left(z-\frac{1}{2}\right)^2-(n-\frac{1}{2})^2} - \frac{1}{\left(z_0-\frac{1}{2}\right)^2-(n-\frac{1}{2})^2}\right].\end{aligned}$$

Write now $\tanh(\pi iz - \frac{i\pi}{2}) = -i\cot(\pi z)$. In the right-hand side, the first term has poles in z at $z = m$ for all integers m, with residue $-(g-1)(2m-1)$. The second term has poles at $z = n$ for $n \in \mathbb{N}^*$, with residue $(g-1)(2n-1)$, and also at $z = 1-n$ for all natural $n \geq 1$, with residue $(g-1)(2n-1)$. Hence the poles at $z = n \in \mathbb{N}^*$ cancel each other. □

From Eq. (9.1), the above three lemmas show that the right-hand side is a meromorphic function on $\mathbb{C}$, hence $\partial_z \log Z$ extends meromorphically to $\mathbb{C}$. Remarkably, all the residues are natural numbers (i.e., positive integers).

THEOREM 9.7. *The function $\partial_z \log Z(z)$ extends meromorphically from $\{\Re(z) > 1\}$ to the complex plane, with simple poles and positive integer residues:*

- $z = \frac{1}{2} \pm is_j$ *for* $s_j^2 + \frac{1}{4} = \lambda_j \in \operatorname{Spec}(\Delta)$ *and* $\lambda_j \notin \{0, \frac{1}{4}\}$, *with residue* $\mu(\lambda_j)$, *the multiplicity of the eigenvalue* λ_j.
- $z = \frac{1}{2}$ *with residue* $2\mu(\frac{1}{4})$ *(in particular if* $\frac{1}{4}$ *is not an eigenvalue,* $z = \frac{1}{2}$ *is not a pole).*
- $z = 1$ *with residue* 1.
- $z = 0$ *with residue* $2g-1$, *where* g *is the genus of* S.

TABLE 1. Poles and residues of the logarithmic derivative of the Selberg zeta function

Pole	Residue
$\frac{1}{2} \pm is_j$	$\mu(\lambda_j)$
$-k$ for $k = 0, 1, \dots$	2(g-1)(2k+1)

- *$z = -k$ for $k = 1, 2, \dots$ with residue $2(g-1)(2k+1)$.*

PROOF. The right-hand side of (9.1) is meromorphic and has the required poles and residues. □

The poles and residues coming from the term R_1 are common to every compact surface S of genus g, and for this reason they are called *trivial*. The other poles, those coming from the spectrum of Δ, are the so-called *non-trivial* poles. Note that $z = 0$ is both non-trivial (coming from the eigenvalue $\lambda_0 = 0$ with multiplicity 1) and trivial, with residue $2(g-1)$. Together the two types of poles lead to the residue $2g - 1$ at $z = 0$.

10. Holomorphic extension of the Selberg zeta function

PROPOSITION 10.1. *Let f be a meromorphic function on $\mathbb{C}$ with simple poles and residues belonging to $\mathbb{N}^*$. Then there exists a holomorphic function F on $\mathbb{C}$, unique up to a multiplicative constant, so that $F' = fF$. Moreover F vanishes exactly at the poles of f, at order equal to the residues of f.*

PROOF. Fix z_0 not a pole of f. For z not a pole, define

$$F(z) = e^{\int_{z_0}^{z} f(s)ds}.$$

Here the integral is along any piece-wise smooth path from z_0 to z. The value of the integral is path-independent as long as we deform the path in a homotopy class relative to the end-points. Choose now two different paths γ, γ' from z_0 to z, so that the concatenation $c := \bar{\gamma}.\gamma'$ is a closed loop. The poles of f are discrete, so only a finite number of them are enclosed by the loop c. The integral $\int_c f(s)ds$ equals $2\pi i$ times the sum of the residues at these poles, each of them multiplied by the winding number of c around it. By hypothesis, all residues are integers, so the exponential $\exp\left(2\pi i \int_c f(s)ds\right)$ equals 1. It follows that $F(z)$ is well-defined, independently of the path from z_0 to z. It is evident that F is holomorphic where it is defined, i.e., outside the poles of f. Near a pole $a \in \mathbb{C}$, write $f(z) = n(z-a)^{-1} + u(z)$ with u holomorphic near a and $n \in \mathbb{N}^*$. Then the primitive of f equals $n \log(z-a) + U(z)$ for some primitive of u which clearly extends to $z = a$. Although the logarithm is well-defined only up to multiples of $2\pi i$, the exponential is well-defined near a and equals $(z-a)^n e^{U(z)}$, hence it vanishes to order n at $z = a$. Clearly the exponential does not vanish outside the poles of f. □

A more intrinsic way of proving this proposition goes as follows: look at a meromorphic 1-form α on any simply-connected Riemann surface S, not necessarily $\mathbb{C}$ (in case $S = \mathbb{C}$ we can choose $\alpha = f(z)dz$). Let M denote the complement of the poles of f in S, and let $\pi : \tilde{M} \to M$ be the universal cover of M. The 1-form α lifts to $\tilde{M}$ and is ∂-exact there by simple-connectedness. Let h be a primitive of $\pi^*\alpha$ on $\tilde{M}$. Let p, p' be two points in $\tilde{M}$ in the fiber of a point $P \in M$, and γ a path from p to p'. Then $h(p) - h(p') = \int_\gamma \pi^*\alpha$, independently of γ. Let $\pi_*\gamma = \pi \circ \gamma$ be the closed path in M obtained by projecting onto M via π. By the change of variables formula,

$$\int_\gamma \pi^*\alpha = \int_{\pi_*\gamma} \alpha.$$

But the integral over the oriented loop $\pi_*\gamma$ is $\pm 2\pi i$ times the sum of the residues enclosed by $\pi_*\gamma$ (when M is the Riemann sphere $\hat{\mathbb{C}}$, the loop separates the surface in two disks, and the sum of residues of α is 0). Hence the exponential of u agrees at the points p and p', so it descends to M as a holomorphic function.

COROLLARY 10.2. *The Selberg zeta function extends analytically to $\mathbb{C}$ with zeros precisely at the points from Table 9, with orders of annulation equal to the residues of Z'/Z.*

11. The functional equation of the Selberg zeta function

We have seen in the proof of the holomorphic extension of $Z(z)$ that $Z'(z)/Z(z)$ is not only meromorphic on $\mathbb{C}$, but also that the components $(2z-1)L(z, z_0)$ and $\frac{z-\frac{1}{2}}{z_0-\frac{1}{2}}\frac{Z'}{Z}(z_0)$ are *odd* with respect to the involution $z \mapsto 1-z$. Moreover, by the proof of Lemma 9.6, the meromorphic function of z $-(2z-1)R_1(z, z_0)$ satisfies the identity

$$-(2z-1)R_1(z, z_0) - (2(1-z)-1)R_1(1-z, z_0) = -2\pi(g-1)(2z-1)\cot(\pi z).$$

It follows that

$$\frac{Z'(z)}{Z(z)} + \frac{Z'(1-z)}{Z(1-z)} = -2\pi(g-1)(2z-1)\cot(\pi z)$$

and by integrating we deduce

$$\frac{Z(z)}{Z(1-z)} = \exp\left(-2\pi(g-1)\int_{1/2}^{z}(2w-1)\cot(\pi w)dw\right).$$

The meromorphic function $-2\pi(g-1)(2w-1)\cot(\pi w)$ is odd with respect to $z \mapsto 1-z$ and its residues are *even positive integers*, hence by Proposition 10.1 the square root of the exponential in the right-hand side is well-defined. We can thus rewrite the functional equation as

$$\tilde{Z}(z) = \tilde{Z}(1-z)$$

where $\tilde{Z}(z) = Z(z)\exp\left(-\pi(g-1)\int_{1/2}^{z}(2w-1)\cot(\pi w)dw\right)$.