

Lecture 9: ... continued

October 1, 2014

Today we're going to continue the discussion from last time. I want to first recap the idea of Riemannian geometry (with slightly more detail but still not complete detail), and compare Riemannian surfaces with Riemann surfaces. Then we'll turn again to the example of $\mathbb{P}^1$, and see how its automorphisms differ from its isometries.

Let X be a smooth surface. This means a $\mathcal{G}$ -structured surface, where $\mathcal{G}$ is the class of smooth bijections with smooth inverse between open subsets $U \subset \mathbb{R}^2$.

There is the notion of a *tangent space* $T_x(X)$ of X at any point $x \in X$. This is a two-dimensional vector space, whose elements are supposed to correspond to first-order movements away from x . Let me not give a definition, but just tell you its characteristic properties:

1. To every smooth map $f : X \rightarrow Y$ of smooth surfaces and every point $x \in X$, there is attached a linear map $df|_x : T_x(X) \rightarrow T_{f(x)}(Y)$, the *derivative of f at x* . This derivative is an isomorphism if and only if f is a local isomorphism near x .
2. The derivative of the composition is the composition of derivatives: $d(g \circ f)|_x = dg|_{f(x)} \circ df|_x$. ("Chain rule".)
3. If U is an open subset of $\mathbb{R}^2$, then $T_x(U) = \mathbb{R}^2$ for all $x \in U$. For smooth maps between such open subsets, the abstract derivative above identifies with the usual notion of derivative as a linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ (given by the matrix of partial derivatives).

Thus, every time you choose a chart around $x \in X$ you get to identify $T_x(X)$ with $\mathbb{R}^2$ (by a linear isomorphism). If you change the chart by a transition function, this identification changes by the linear automorphism of $\mathbb{R}^2$ given by the derivative of the transition function at the point corresponding to x . More generally, you can calculate the derivatives of a smooth map f by choosing charts around x and $f(x)$. Then, via the linear isomorphisms $T_x(X) \leftrightarrow \mathbb{R}^2$ and $T_{f(x)}(Y) \leftrightarrow \mathbb{R}^2$ provided by these charts, $df|_x$ just turns into the usual derivative of the representative of f on the charts.

Using this notion of tangent space, we can say what is a *Riemannian structure* on a smooth surface X : it is a collection of inner products on each tangent space $T_x(X)$, satisfying the condition that these inner products "vary smoothly" with x .

I won't give the precise definition of what "varies smoothly" means, but you'll see it in examples. It means that, on charts, the formula for the inner product at x should somehow be a smooth function of x .

Another name for a Riemannian structure is a *metric*. This alternate name is useful if we don't want to overdose on things named after Riemann.

Recall the idea: if you have an inner product $\langle -, - \rangle$ on a vector space, you can define all sorts of geometric measurements on vectors in that vector space. For example you can define the length

of a vector, the unoriented angle between two vectors, the area of the parallelogram spanned by two vectors, etc. These definitions recover the usual notions from Euclidean geometry in the case of $\mathbb{R}^n$ with the standard dot product, and satisfy basically the same properties as are satisfied in that standard example.

Then, by various procedures (e.g. integration), you can use the geometry on vectors in each tangent space to define actual geometric notions on the surface X . But I won't dwell on this aspect because it's not so important for us right now.

Now, let us compare the notions of Riemannian surface (surface with metric) and Riemann surface (surface specified by holomorphic charts). Both of these types of objects come with an underlying smooth surface: the first by definition, and the second because a biholomorphism is in particular a smooth bijection with smooth inverse. Thus both Riemannian surfaces and Riemann surfaces can be thought of as smooth surfaces, equipped with some extra structure.

In the case of Riemannian surfaces, this structure is specified at the level of tangent spaces. In the case of Riemann surfaces, this is not the case, at least not by definition. But nonetheless a Riemann surface structure does determine some extra structure on the tangent spaces: as we've said before, it determines a notion of oriented angle between two vectors in any tangent space.

So a metric determines an unoriented geometry, including angles but also distances, whereas a Riemann surface structure determines an oriented notion of angle. Neither covers the other's territory exactly: the metric certainly gives you more geometry, but on the other hand it doesn't give you orientation. Nonetheless one can make the following definition:

Definition 0.1. *Let X be a Riemann surface, and let $\langle -, - \rangle_x$ (for all $x \in X$) denote a Riemannian metric on X , viewed just as a smooth surface. We say that the metric is compatible with the Riemann surface structure if the notion of angle coming from the Riemann surface structure agrees with the notion of angle coming from the metric.*

One can also say what this means more explicitly: the metric and the Riemann surface structure are compatible if, when we look on any holomorphic chart, the metric $\langle -, - \rangle_x$ is a positive constant multiple of the standard metric (dot product) on the complex numbers:

$$\langle -, - \rangle_x = \lambda^2 \cdot \langle -, - \rangle_{std}.$$

(Here the constant $\lambda > 0$, which may depend on x , represents the amount distances are scaled in going from the standard metric to $\langle -, - \rangle_x$.)

The reason this is an equivalent statement is that the notion of angle on a Riemann surface is by definition inherited from the standard Euclidean geometry on $\mathbb{C}$ in the charts. On the other hand two inner products on $\mathbb{R}^2$ determine the same notion of angle if and only if they are positive scalar multiples of each other (this is a fact very similar to the first problem on your first problem set).

Let us now give an example illustrating this notion. Let X be our old friend $\mathbb{C} \setminus 0$. This is a Riemann surface defined by a single chart, the identity map. Thus we can parametrize all the metrics compatible with the Riemann surface structure by all smooth functions $\lambda : \mathbb{C} \setminus 0 \rightarrow \mathbb{R}_{>0}$: the corresponding metric is given by

$$\langle -, - \rangle_z = \lambda(z)^2 \cdot \langle -, - \rangle_{std}.$$

Among these are of course the standard metric, corresponding to $\lambda(z) = 1$, and the cylindrical metric, corresponding to $\lambda(z) = \frac{1}{|z|}$. But there are of millions of other possibilities, where "millions" means in particular an uncountable number.

Nonetheless, out of all of these possibilities, there is one metric compatible with the holomorphic structure which is certainly the best: the cylindrical metric. This can be measured in a number of

ways, but here is one: every automorphism of $\mathbb{C} \setminus 0$ as a Riemann surface is an *isometry* for the cylindrical metric.

(An isometry is a map which preserves the geometry: it is a smooth bijection $f : X \rightarrow X$ with smooth inverse such that the induced map on tangent spaces $df : V \rightarrow W$, at any point $x \in X$, is a linear isometry, meaning it preserves the inner product, i.e. it satisfies $\langle Dv_1, Dv_2 \rangle = \langle v_1, v_2 \rangle$.)

One way to argue for this is to classify the automorphisms of $\mathbb{C} \setminus 0$, for example by following a similar outline to the first problem on this week's problem set. You will find that every automorphism is either multiplication by a nonzero complex number, or it is the composition of $z \mapsto 1/z$ with multiplication by a nonzero complex number. Thus one only needs to see that $z \mapsto 1/z$ and multiplication by a nonzero complex number preserve the metric, and this is a simple calculation.

This is something that should be surprising a priori, because an automorphism of a Riemann surface, by all rights, should only preserve angles, not distances. It's somewhat of a miracle that there is this compatible Riemannian structure in which every angle-preserving automorphism also preserves distances. Nonetheless this is a completely general phenomenon: there is the following theorem, which we'll prove as a consequence of the uniformization theorem:

Theorem 0.2. *Every connected Riemann surface X has a canonical compatible Riemannian metric $\langle -, - \rangle_{\text{canon}}$.*

Okay, that's not a precise statement, because the important term "canonical" has not been precisely defined. Furthermore, the statement is not quite correct when X is a Riemann surface which is abstractly isomorphic to $\mathbb{C} \setminus \{0\}$ or $\mathbb{C}$ or $\mathbb{P}^1$, without a specific isomorphism having been chosen. In those cases one needs to choose an isomorphism to get the canonical metric. But other than in these three exceptional cases, the metric is independent of any choice of isomorphism.

Let me now try to elucidate the meaning of the word "canonical" further, by giving three properties this canonical metric satisfies:

1. If we uniformize X , writing it as a quotient $\mathbb{M}/\Gamma$ of a model Riemann surface $\mathbb{M}$ (so $\mathbb{M}$ is either $\mathbb{P}^1$, $\mathbb{C}$, or $\mathbb{D}$) by a group of automorphisms Γ acting freely, then the canonical metric is the unique metric on X for which the quotient map $\mathbb{M} \rightarrow \mathbb{M}/\Gamma$ is a local isometry, when we equip the model $\mathbb{M}$ with its natural metric (spherical, Euclidean, or hyperbolic, in the three cases).
2. The canonical metric on X is the unique compatible metric, up to a constant scalar factor, in which X is *complete* and has *constant scalar curvature*. (This is not quite true for $X = \mathbb{P}^1$.)
3. Outside the two special cases $X = \mathbb{C}$ and $X = \mathbb{P}^1$, every automorphism of X is an isometry for the canonical metric.

We already illustrated the third point in our example case $X = \mathbb{C} \setminus 0$, so let's see the other points as well. The first point is saying that the exponential map

$$\exp : \mathbb{C} \rightarrow \mathbb{C} \setminus \{0\}$$

is a local isometry when the source is equipped with the standard Euclidean metric and the target is equipped with the cylindrical metric. We already know that $\exp$ preserves angles because it's holomorphic, so this just means that the scales are preserved as well. In general, a holomorphic map f on open subsets of $\mathbb{C}$ scales the standard Euclidean metric by $|f'(z)|$ in going from z to $f(z)$. When $f = \exp$ we have $f' = f$, so scales are multiplied by $|f(z)|$. This means that scales are preserved in our cylindrical metric, which was defined exactly by dividing by this number.

(By the way, this provides a better perspective on why the automorphisms of $\mathbb{C} \setminus \{0\}$ preserve the metric: via $\exp$, they correspond to compositions of $z \mapsto -z$ and $z \mapsto z + a$ on $\mathbb{C}$, and these evidently preserve the standard Euclidean metric.)

As for the second point, let me again not give precise definitions (if you're curious you can look them up on the internet). But the completeness condition signifies that any missing points must be infinitely far away, as measured by the metric. For example $\mathbb{C}$ in the Euclidean metric is complete, because the only missing points are way out at ∞ . But $\mathbb{C} \setminus 0$ in the standard Euclidean metric is not complete because the missing point 0 is right there next to the other points. On the other hand, with the cylindrical metric it is complete: the rescaling pushed the point 0 out infinitely far away. (You can think of an infinitely long cylinder in 3-dimensional space: one end corresponds to 0, the other to ∞ .) And the constant scalar curvature condition is a kind of uniformity condition, which sort of means that there are no irregular bumps in the geometry. (That's a horrible explanation, sorry.)

In order to be able to prove this theorem, and more generally in order to grasp some of the geometric significance of the uniformization theorem, it's important to discuss the canonical metrics on the model spaces $\mathbb{P}^1$, $\mathbb{C}$, and $\mathbb{D}$, and to investigate the relationship between the Riemann surface automorphisms and the isometries of these spaces. It is in this light that we now return to our discussion of the geometry of $\mathbb{P}^1$.

In the previous lecture we saw that there is a conformal (i.e., oriented angle-preserving) isomorphism between the unit sphere S^2 and $\mathbb{P}^1$, given by stereographic projection. We can transport the standard Riemannian structure on S^2 (inherited from the standard Euclidean geometry on the ambient space $\mathbb{R}^3$) over to $\mathbb{P}^1$ via this stereographic projection; the result is then a Riemannian metric on $\mathbb{P}^1$ which is compatible with the Riemann surface structure. This is the canonical metric on $\mathbb{P}^1$, also called the *round metric*.

We can write a formula for it in coordinates: on $\mathbb{C} \subset \mathbb{P}^1$, it is given by

$$\lambda(z)^2 \cdot \langle -, - \rangle_{std},$$

where $\lambda(z) = \frac{2}{1+|z|^2}$. This follows from the formulas of the previous lecture. Indeed, in the previous lecture we saw that the scaling factor going from S^2 to $\mathbb{C}$ by stereographic projection from the north pole is given by $\frac{1}{1-h}$, where h stands for the height (i.e. third coordinate). Thus the round metric is $(1-h)^2$ times the canonical metric. Then one just inputs the formula for h in terms of z coming from the formula for the inverse of stereographic projection.

You can probe the correctness of this formula in a number of ways. For example, as $|z| \rightarrow \infty$, we're scaling the metric down a lot, which corresponds to the fact that the huge region near ∞ corresponds to just a small area on the sphere. Also, at $z = 0$ we're scaling by 2: near the south pole, the inverse to stereographic projection from the north pole looks just like multiplication by 2. And when $|z| = 1$ we're not scaling at all: that's because there the stereographic projection looks like just a 90 degree rotation. Here's another way to probe it: the formula

$$\int_{-\infty}^{\infty} \frac{2}{1+t^2} dt = 2\pi$$

corresponds to the fact that a circle of radius one has length 2π . (The left hand side calculates the length of the real axis in our spherical metric; but the real axis goes to a circle under the inverse to stereographic projection.)

Now, again, our main concern in this lecture is isometries of $\mathbb{P}^1$ in this metric, and the relation of them to arbitrary automorphisms of the Riemann surface. The main results are the following:

Theorem 0.3. *The orientation-preserving isometries of S^2 are exactly the rotations around some axis.*

Part of this claim is that the composition of two rotations is necessarily another rotation, because the composition of two isometries is an isometry. This fact is not geometrically obvious to me; in fact the proof we give will be algebraic.

Theorem 0.4. *Define subsets $K, A, N \subset \text{Aut}(\mathbb{P}^1)$ as follows. The set K is the set of rotations of S^2 . The set A is the set of automorphisms of $\mathbb{P}^1$ of the form $z \mapsto \lambda \cdot z$, where λ is a positive real number. And the set N is the set of automorphisms of the form $z \mapsto z + a$, where a is a complex number. (Everything in A and N fixes the point ∞ .) Then the map*

$$K \times A \times N \rightarrow \text{Aut}(\mathbb{P}^1)$$

given by $(k, a, n) \mapsto k \circ a \circ n$, is a bijection. (So, every automorphism of $\mathbb{P}^1$ is uniquely the composition of a translation, a positive scaling, and a rotation.)

Thus, the difference between automorphisms of $\mathbb{P}^1$ and isometries of $\mathbb{P}^1$ in its round metric is these two types of automorphisms A and N . By the way, I chose these funky letters K , A , and N because the above result is a special case of the “Iwasawa decomposition of semisimple Lie groups”, where those letters are customary. The K here is also a nice example of a “maximal compact subgroup”. Although this is one of those exceptional cases where not every automorphism is an isometry, the relationship between a Lie group and its maximal compact subgroup is still quite tight.

The plan of attack for proving these theorems is as follows. First we will show that the composition of two rotations is a rotation. Then we will prove the second theorem. Then we will deduce the first theorem.

So let us start by proving that the composition of two rotations is a rotation. We'll do this by giving an alternate algebraic description of the set of all rotations, one which makes the desired fact obvious.

This requires some set-up. Suppose that $M : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is an invertible linear map. (So, M is given by a 3x3 matrix with nonzero determinant.) We say that M is *orientation preserving* if its determinant is positive, and we say that M is an *isometry* if it preserves the standard inner product on $\mathbb{R}^3$. The collection of all M 's satisfying both of these conditions (the *oriented linear isometries of $\mathbb{R}^3$*) is denoted by

$$SO(3).$$

This is a group, because both of the defining conditions are clearly closed under passing to inverses and taking compositions.

Then the algebraic description of rotations is as follows.

Proposition 0.5. *The set of rotations of $\mathbb{R}^3$ around some axis through the origin in $\mathbb{R}^3$ is the same as the set $SO(3)$ of oriented linear isometries of $\mathbb{R}^3$.*

Proof. There are two inclusions to verify. First, suppose R is a rotation around some axis; we need to show that R lies in $SO(3)$. But such an R preserves lengths of vectors, hence preserves the inner product. It is also orientation preserving, hence has positive determinant.

Now suppose that M lies in $SO(3)$. Let p denote the characteristic polynomial of M . Then p is a degree-3 polynomial with real coefficients. It follows that p has three roots in $\mathbb{C}$ (with multiplicities), and that the set of these roots (and their multiplicities) is preserved by complex conjugation. Thus there are two possibilities: either

1. All three roots are real.
2. One of the roots is real and the other two form a pair of conjugate non-real complex numbers.

I claim that in any case, p must have a positive real root. To prove this, we use the fact that $\det(M)$ is positive. On the other hand $\det(M)$ is the product of the roots of the characteristic polynomial p , so if all the roots were non-positive we would get a contradiction in either case (recall $z \cdot \bar{z} \geq 0$ for all $z \in \mathbb{C}$).

Let λ then be a positive real root of p , i.e. a positive real eigenvalue of M . Say v_0 is a corresponding eigenvector, so

$$Mv_0 = \lambda v_0.$$

By taking the length of both sides, we deduce that in fact $\lambda = 1$, so $Mv_0 = v_0$.

Now let V_0 be the (one-dimensional) span of v_0 , and let V_1 be the orthogonal complement of V_0 . Since M sends V_0 bijectively to itself (it's the identity there) and V_1 is defined in terms of V_0 and the inner product, which is preserved by M , it follows that M sends V_1 bijectively to itself. Thus M fixes the axis V_0 and is determined by its behavior in the orthogonal direction V_1 .

Thus we've reduced the problem to one dimension lower: we need to see that an oriented isometry of a two-dimensional inner product space is necessarily a rotation by some angle. But that fact is much easier to prove, and I'll leave it as an exercise. $\square$

Note that every $M \in SO(3)$ restricts to an isometry $S^2 \rightarrow S^2$. Indeed, S^2 is defined as the set of vectors in $\mathbb{R}^3$ of norm one; this condition is described purely in terms of the inner product, so M preserves it and hence induces a map $S^2 \rightarrow S^2$. The same argument applied to M^{-1} shows that this induced map is a bijection. It is also smooth (with smooth inverse), because a linear map is differentiable. Moreover, M is its own derivative at every point; so since M preserves the standard inner product on $\mathbb{R}^3$, it restricts to a linear isometry $T_p(S^2) \rightarrow T_{M(p)}(S^2)$ for all $p \in M$. Thus every $M \in SO(3)$ induces an isometry of S^2 . Moreover, M is uniquely determined by this isometry, by the formula $M(tx) = t \cdot M(x)$ for $t \geq 0$ and x of norm 1.

Thus the proposition shows us that the set of rotations is a group of isometries of S^2 .

Now let us prove the second theorem, the KAN decomposition of $Aut(\mathbb{P}^1)$.

Proof. Recall that $Aut(\mathbb{P}^1)$ is in bijection with the set of triples of ordered distinct points of $\mathbb{P}^1$, via sending φ to $(\varphi(0), \varphi(1), \varphi(\infty))$. Thus to prove the theorem it suffices to show that for any three distinct points (p, q, r) of $\mathbb{P}^1$, there is a unique $k \in K$, $a \in A$, and $n \in N$ with

$$(k \circ a \circ n)(p, q, r) = (0, 1, \infty).$$

First let us show existence. Start by letting R be any rotation of S^2 which carries r to ∞ . Then consider $z = R(p)$ and $w = R(q)$. Since R is bijection we know these are two distinct points of $\mathbb{C}$. Thus we can find a unique $n \in N$ and $\lambda \in \mathbb{C} \setminus \{0\}$ such that

$$(mult_\lambda \circ n)(0, 1, \infty) = (z, w, \infty),$$

where $mult_\lambda$ means multiplication by λ . Namely, n must be translation by $\frac{z}{w-z}$, and λ must be $w - z$.

We can further uniquely decompose $mult_\lambda$ as $mult_c \circ a$ where c has norm one and $a \in A$, namely by taking $c = \frac{\lambda}{|\lambda|}$ and a to be scaling by $|\lambda|$.

Then, let $k = R^{-1} \circ mult_c$. Since $mult_c$ is just a rotation around the north-south axis, this k is a composition of rotations, hence itself a rotation by the proposition. Furthermore, we have arranged it so that

$$(k \circ a \circ n)(p, q, r) = (0, 1, \infty).$$

Now only uniqueness needs to be checked. The only source of non-uniqueness above was the initial choice of R . So suppose R' is a different choice, leading to the quantities z', w', n', λ', a' , and c' , following the above outline. Then $R' \circ R^{-1}$ is a rotation fixing ∞ , hence is a rotation around the

north-south axis, hence is multiplication by some complex number of norm one. Thus z', w' are obtained from z, w by multiplication by a complex number of norm one. It follows that $n' = n$ and $a' = a$. Then since $k \circ a \circ n = k' \circ a' \circ n'$ we deduce as well that $k = k'$, giving the uniqueness. $\square$

To finish we prove the first theorem, that every oriented isometry of S^2 is a rotation.

Proof. Let f be an oriented isometry of S^2 . Then f an automorphism of $\mathbb{P}^1$, so it has a unique decomposition as $f = k \circ a \circ n$ as in the previous theorem. Now, by the proposition, f will be a rotation if and only if $k^{-1} \circ f = a \circ n$ is a rotation. But it's plain to see that a composition $a \circ n$ can only be an isometry if it is the identity. (E.g., rigorously, one can calculate using the formula

$$\left(\frac{2}{1 + |z|^2} \right)^2 \cdot \langle -, - \rangle_{std}$$

for the spherical metric on on the complex plane.) $\square$