XIII

Riemann Surfaces

Part XIII: Contents

47 Basic definitions of Riemann surfaces

Roughly speaking, the theory of Riemann surfaces is just the generalization of complex analysis using ideas from differential geometry: Just like how a 2-manifold can be viewed as a collection of patches of the real plane $\mathbb{R}^2$ smoothly welded together to form a more complicated object, we take "pieces" of the complex plane $\mathbb{C}$, *analytically* welded together .

We already know that the theory of holomorphic function is very nice — they're all analytic! The same amount of rigidity is to be expected here.

In fact, on *compact* Riemann surfaces, the theories are even nicer than the case of holomorphic functions! For example:

- For two Riemann surfaces X and Y where Y is compact, any meromorphic function $f\colon X \to Y$ must in fact be holomorphic i.e. defined everywhere.
- If X is a compact Riemann surface, then a holomorphic function $f\colon X \to \mathbb{C}$ is constant.
- In the same setting as above, furthermore we have that if $g\colon X \to \mathbb{C}$ is meromorphic, then the number of zeros of g is equal to the number of poles of g, with multiplicity.

Remark 47.0.1 (Why do we have these nice properties?) — Roughly speaking, $\mathbb{C}$ is not compact — it is isomorphic to the Riemann sphere with a hole removed. By filling in the hole, we allow meromorphic functions to be extended taking value ∞ at places that previously was a pole.

As an orientable 2-manifold, we can define the **genus** of a Riemann surface — it is a purely topological concept, yet it is crucially linked to several algebraic invariants in very surprising ways. You may have heard of the *elliptic curve* in cryptography — they also present as a Riemann surface, and a generalization, **hyperelliptic curve**, form a family of Riemann surfaces of arbitrary genus ≥ 2!

§47.1 Complex structures

Recall the definitions in the previous chapters:

- A topological n-manifold is a Hausdorff space, with a covering $\{U_i\}$, each being homeomorphic to $\mathbb{R}^n$.
- A smooth n-manifold is a topological n-manifold, where all the transition maps are smooth.

What do they have in common? Seemingly not too much. But essentially, they're all describing the same philosophy:

We take countably many patches $\{U_i\}$, and weld them together while keeping the underlying structure.

Here, a topological manifold has a *topological structure*, and a smooth manifold has a *smooth structure*. In a similar manner, a complex manifold has a *complex structure*.

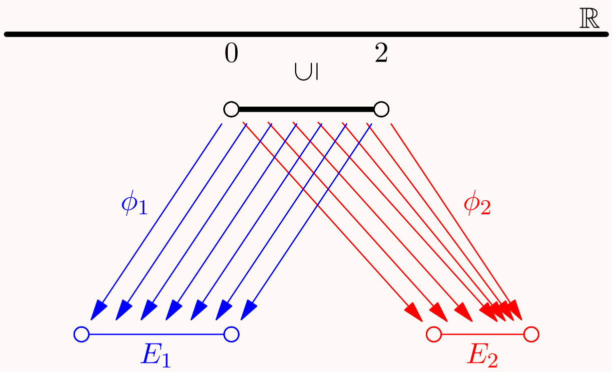
What do we mean by "structure" here?

First, a topological structure is familiar to you — it's just a topology. Formally, the topology is defined by the collection of open sets, but the actual *meaning* of a topological structure dictates:

- whether a set is considered open or closed,
- whether a sequence of points converge to a given point,
- whether a map $X \to Y$ or $Y \to X$ is continuous (given Y is another topological space),
- etc.

Given a topological n-manifold with an existing (Hausdorff) topology on it, we can tell whether a local chart *respects the topological structure*; in other words, is a homeomorphism.

Example 47.1.1 ($(0,2)$ is a topological 1-manifold)
The open interval $(0,2)$ included in $\mathbb{R}$ can be considered a topological manifold.



Two possible charts for the space, ϕ_1 and ϕ_2, are shown.

Their formulas are $\phi_1 \colon (0,2) \to (0,2), \phi_1(x) = x$ and $\phi_2 \colon (0,2) \to (0,1.3), \phi_2(x) = x + 0.35 \cdot (1 - x - |1 - x|)$.

In the example above, you may notice that, even though the chart ϕ_2 is a homeomorphism, it doesn't look *smooth*. So, you want to define a smooth 2-manifold as something like:

> A surface $S \subseteq \mathbb{R}^3$ is a smooth 2-manifold if, for each $p \in S$, there exists an open neighborhood $V \subseteq S$ that is diffeomorphic to $E \subseteq \mathbb{R}^2$.

In fact, this is the actual definition in classical differential geometry — of course, this isn't completely general, for instance, we know that the Klein bottle cannot be embedded into $\mathbb{R}^3$.

So, why didn't we define something like this in Definition 46.2.2? The problem is, the concept of a diffeomorphism isn't defined on a Hausdorff topological space — in fact it can't be defined, right in the example above, you can see a homeomorphism that is not a diffeomorphism — in other words, a topological space can be assigned different *smooth structures*.

So, the essence of what the definition Definition 46.2.2 is doing is, it implicitly defines what a *smooth structure* mean, by inducing the smooth structure from each patch $E_i \subseteq \mathbb{R}^n$ to the topological space M. The condition that the transition functions need to be smooth is, of course, to ensure that the smooth structures on M induced by different ϕ_i are the same.

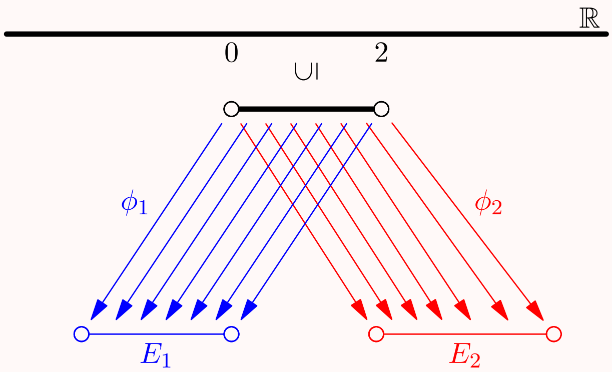
In completely the same way, we could have replaced Definition 46.1.2 by:

> A topological n-manifold M is a set with a collection of subsets $\{U_i\}$ that covers M, for each U_i there is a bijective map from it to a subset of $\mathbb{R}^n$, say
>
> $$\phi_i \colon U_i \to E_i \subseteq \mathbb{R}^n$$
>
> where each E_i is an open subset of $\mathbb{R}^n$, satisfying that all the transition maps are topological homeomorphisms.

Here, the ϕ_i are "set isomorphisms" and plays a similar role as the homeomorphisms in Definition 46.2.2, and the topological space structure is similarly induced from the patches E_i.

Example 47.1.2 ($(0,2)$ is a smooth 1-manifold)

Just as above, the open interval $(0,2)$ included in $\mathbb{R}$ can also be considered a smooth manifold.



This time around, ϕ_1 is the same as above, but $\phi_2 \colon (0,2) \to (0, 2+\frac{1}{e})$ is defined by

$$\phi_2(x) = \begin{cases} x & \text{if } x \le 1 \\ x + e^{-1/(x-1)} & \text{otherwise.} \end{cases}$$

Because all of ϕ_1, ϕ_2, and their inverses are smooth functions, the transition maps $\phi_1 \circ \phi_2^{-1}$ and $\phi_2 \circ \phi_1^{-1}$ are thus smooth, satisfying the hypothesis of Definition 46.2.2.

You should take a moment to think through this idea — because smooth functions on $\mathbb{R}^n$ are so natural, it's easy to forget that a smooth manifold carries more structure than just the topology.

Once again, as we have seen in the example above, $\mathbb{R}^n$ has more structure than just being smooth — it has an *analytic structure*. The chart ϕ_2 does not preserve this structure.

So, for Riemann surface, we will just have:

A Riemann surface is a smooth (real) 2-manifold which locally looks like $\mathbb{C}$, and carries an *complex-smooth structure*.

Of course, by the miracle of complex analysis — holomorphic functions are analytic! — this is equivalent to stating that a Riemann surface carries a complex-analytic structure.

§47.2 Riemann surface

Prototypical example for this section: The Riemann sphere, or any open subset of $\mathbb{C}$ such as $\{z \in \mathbb{C} \mid |z| < 1\}$.

From the motivation above, the definition of a Riemann surface naturally falls out:

Definition 47.2.1 (Riemann surface)**.** A **Riemann surface** X is a second countable connected Hausdorff space with an open cover $\{U_i\}$ of countably many sets homeomorphic to open subsets of $\mathbb{C}$, say by homeomorphisms

$$\phi_i \colon U_i \xrightarrow{\cong} E_i \subseteq \mathbb{C}$$

such that the **transition maps** ϕ_{ij} defined by

$$\phi_{ij} \colon E_i \cap \phi_i^{\text{img}}(U_i \cap U_j) \xrightarrow{\phi_i^{-1}} U_i \cap U_j \xrightarrow{\phi_j} E_j \cap \phi_j^{\text{img}}(U_i \cap U_j).$$

are analytic functions. Each ϕ_i is called a **complex chart**, and together they form a **complex atlas**.

We say that the complex atlas gives the Hausdorff space a **complex structure**. Thus, in other words, a Riemann surface is a (second countable, connected, Hausdorff) topological space with a complex structure.

[Mi95] has an alternative definition, by a maximal complex atlas. Both definitions are the same, but in practice, it's easier to specify finitely many complex charts than specifying infinitely many ones.

A complex chart $U_i \to E_i$ should be think of as giving a **local coordinate** on U_i. Formally:

Definition 47.2.2. For a point $p \in X$, open set $U \subseteq X$ and complex chart $\phi \colon U \to \mathbb{C}$, let $z = \phi(x)$ for each $x \in U$, we call z a **local coordinate**. We say that the local coordinate is **centered** at p if $\phi(p) = 0$.

§47.3 Complex manifold

Analogously to the definition of a real n-manifold, we can define a complex manifold. Just as above, the structure has much more rigidity than a smooth surface.

Definition 47.3.1 (Complex n-manifold)**.** A **complex n-manifold** is a Hausdorff space with an open cover $\{U_i\}$ of countably many sets homeomorphic to open subsets of $\mathbb{C}^n$, say by homeomorphisms

$$\phi_i \colon U_i \xrightarrow{\cong} E_i \subseteq \mathbb{C}^n$$

such that the **transition maps** ϕ_{ij} are analytic functions.

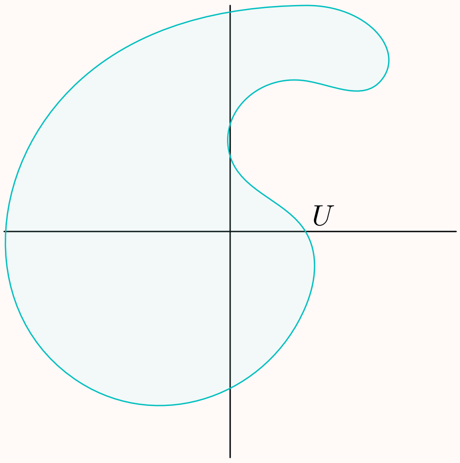
Of course, a complex n-manifold is naturally a smooth (real) $2n$-manifold.

§47.4 Examples of Riemann surfaces

In this chapter, several examples will be given.

Example 47.4.1 (Open subsets of $\mathbb{C}$)

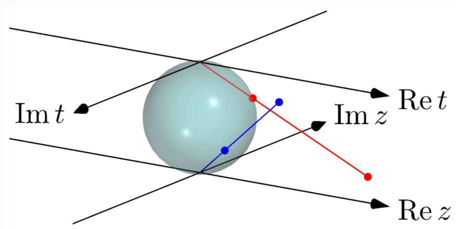
Any connected open subset $U \subseteq \mathbb{C}$ is a Riemann surface.



This is a boring example (the whole thing can be defined without any welding), but let's go on.

Example 47.4.2 (The Riemann sphere)

The Riemann sphere $\mathbb{C}_\infty$, as a smooth 2-manifold, is just a sphere.



Its complex structure is defined as follows:

Embed the sphere in $\mathbb{R}^3$ such that $N = (0, 0, 1)$ and $S = (0, 0, 0)$ are two antipodal points.

Let E_1 be the xy-plane, and let E_2 be the set of points with $z = 1$.

Then, let $\phi_1 \colon \mathbb{C}_\infty \setminus \{N\} \to E_1$ be the stereographic projection from the sphere (except the point N) to E_1 through the point N, and let $\phi_2 \colon \mathbb{C}_\infty \setminus \{S\} \to E_2$ be the stereographic projection from the sphere (except the point S) to E_2 through the point S.

We think of E_1 and E_2 as copies of the complex plane embedded in $\mathbb{R}^3$ by $z \mapsto (\operatorname{Re} z, \operatorname{Im} z, 0) \in E_1$ and $t \mapsto (\operatorname{Re} t, -\operatorname{Im} t, 1) \in E_2$. Then ϕ_1 and ϕ_2 are complex charts for $\mathbb{C}_\infty$.

The domain of ϕ_1 and ϕ_2 covers $\mathbb{C}_\infty$. To make $\mathbb{C}_\infty$ into a complex manifold, we must ensure that the complex structure induced by ϕ_1 and ϕ_2 are the same — indeed, over any open set U that contains neither N nor S, the projections are related by

$\phi_1(p) = \frac{1}{\phi_2(p)}$ for all $p \in U$.

This also explains why the minus sign is needed in $t \mapsto (\operatorname{Re} t, -\operatorname{Im} t, 1)$ — otherwise, the projections will be related by $\phi_1(p) = \frac{1}{\overline{(\phi_2(p))}}$, which is not analytic.

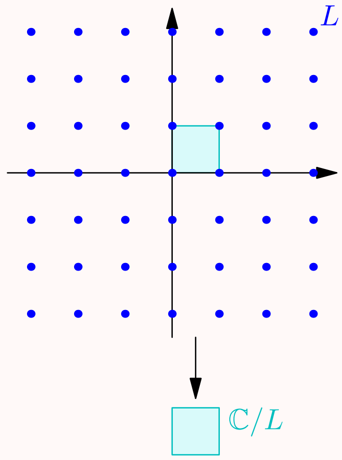
We can think of the Riemann sphere as the result of welding two copies of $\mathbb{C}$ together in order to "fill in" the missing point ∞.

In the example above, the local coordinate given by ϕ_1 is centered at S, and the local coordinate given by ϕ_2 is centered at N. The point $\phi_1^{-1}(4)$ would have local coordinate $z = 4$ under the chart ϕ_1, and local coordinate $t = \frac{1}{4}$ under the chart ϕ_2.

Example 47.4.3 (The complex torus)

Let L be the set $\mathbb{Z}[i]$ of complex numbers with both real and imaginary parts of $\mathbb{C}$. Then L forms an additive subgroup of $\mathbb{C}$.

Consider the quotient $\mathbb{C}/L$. The quotient map $\mathbb{C} \to \mathbb{C}/L$ induces a natural complex structure on $\mathbb{C}/L$.



Here we draw $\mathbb{C}/L$ as a square, but you should imagine that the top and bottom edge, as well as the left and right edges, are smoothly welded together.

For each small patch of the torus, we can isomorphically map it to $\mathbb{C}$ by taking a suitable component of the preimage of the quotient map — the different choices of the projection are related by transition functions $\phi_{ij}(x) = x + a$ for $a \in L$, this is analytic.

The complex torus is compact, thus any holomorphic function on $\mathbb{C}/L$ is constant. Meromorphic functions are more interesting, and also difficult to construct.

And some non-examples.

Example 47.4.4

The disjoint union of two Riemann spheres is not a Riemann surface, because it is not connected.

The condition that a Riemann surface must be connected is merely a technical condition such that theorems are nice — we don't lose much by requiring this, because any topological space with a complex structure can be broken down into disjoint union of Riemann surfaces, one for each connected component.

48 Morphisms between Riemann surfaces

§48.1 Definition

The definition is what we would expect — since a Riemann surface's main feature is a complex structure, a map $f: \mathbb{C} \to \mathbb{C}$ is a morphism between Riemann surfaces if and only if it is holomorphic.

Definition 48.1.1. Let X and Y be Riemann surfaces. A mapping $f: X \to Y$ is holomorphic at $p \in X$ if and only if there exists charts $\phi_1: U_1 \to E_1$ on X with $p \in U_1$ and $\phi_2: U_2 \to E_2$ on Y with $f(p) \in U_2$ such that the composition $\phi_2 \circ f \circ \phi_1^{-1}$ is holomorphic at $\phi_1(p)$. We say f is a **morphism between Riemann surfaces** if and only if it is holomorphic at all points of X.

In other words: f is holomorphic if and only if it is holomorphic as function mapping between local coordinates.

Example 48.1.2

Some examples follows.

- The function $f: \mathbb{C} \to \mathbb{C}$ by $f(x) = x^3$ is a morphism.

  Note that this function is not bijective. At each point $p \neq 0$, there is an open neighborhood on which f has an inverse, but f has no inverse at 0.

- The embedding of the complex plane into the Riemann sphere, $\mathbb{C} \hookrightarrow \mathbb{C}_\infty$, is a morphism.

§48.2 Functions to the Riemann sphere

Prototypical example for this section: The meromorphic function $\frac{1}{z}$ can be made into a holomorphic $\mathbb{C} \to \mathbb{C}_\infty$ function.

In this section, we will see that the Riemann sphere $\mathbb{C}_\infty$ can be viewed as "$\mathbb{C}$ with a point at infinity added". This interpretation allows us to interpret meromorphic functions $f: X \to \mathbb{C}$ as holomorphic maps $g: X \to \mathbb{C}_\infty$, which allows a much better handling of meromorphic functions — there's no longer any singularity, the resulting function g is holomorphic everywhere!

First, we see that $\mathbb{C}_\infty$ can be naturally interpreted as $\mathbb{C}$ with a single point added: With notation as in Example 47.4.2, identify $\mathbb{C}_\infty \setminus \{N\}$ with E_1 (and thus with $\mathbb{C}$) through the map ϕ_1, and we let ∞ be the point N.

Question 48.2.1. Convince yourself that it makes sense to call the point ∞ — for every sequence of points $\{z_i\}$ on $\mathbb{C}$ such that $|z_i| \to +\infty$, then $\phi_1^{-1}(z_i) \to \infty$ on $\mathbb{C}_\infty$ as a topological space.

So, let X be a Riemann surface, and $f\colon X \to \mathbb{C}$ be a meromorphic function on X. Naturally, g can be defined by

$$g(z) = \begin{cases} f(z) & \text{if } f(z) \neq \infty \\ \infty & \text{if } f(z) = \infty. \end{cases}$$

Then g is continuous — but furthermore, it's analytic.

Question 48.2.2. Clearly, at points $z \in X$ where $g(z) \neq \infty$, then g is analytic.

Convince yourself that g is also analytic at $z \in X$ where $g(z) = \infty$. (With notation as in Example 47.4.2, take a small open set $U \subseteq X$, and re-parametrize $g^{\text{img}}(U) \subseteq \mathbb{C}_\infty$ by $t = 1/z$.)

Therefore,

Proposition 48.2.3

There is a one-to-one correspondence between meromorphic functions $f\colon X \to \mathbb{C}$ and holomorphic maps $g\colon X \to \mathbb{C}_\infty$ such that g is not identically ∞.

Or, more informally,

Plugging in the hole at ∞ of $\mathbb{C}$ allows us to analytically extend meromorphic functions to $\mathbb{C} \cup \infty$ maps which is holomorphic everywhere.

§48.3 Some other nice properties

We have just seen in the last section that the Riemann sphere $\mathbb{C}_\infty$ allows us to remove the singularities of a meromorphic functions.

Informally speaking, this is because $\mathbb{C}_\infty$ is a "compactfication" of $\mathbb{C}$ — adding a point to make it compact — and compact Riemann surfaces enjoy many nice properties.

Proposition 48.3.1

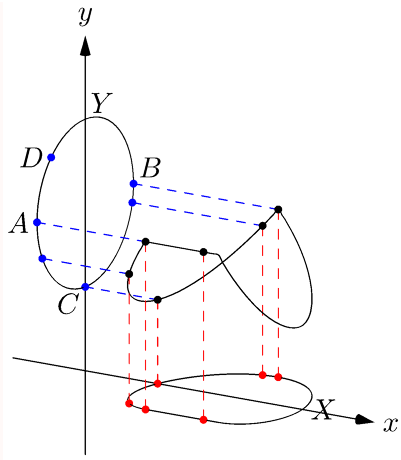
Let X and Y be compact, $f\colon X \to Y$ be holomorphic and not constant. For each point $y \in Y$, define d_y be the total multiplicity of the points in the preimage of y.

Then, d_y is well-defined and constant.

You can see why this proposition is surprising:

Example 48.3.2 (The proposition does not hold for smooth compact manifolds)

Consider the following function $f\colon X \to Y$ between compact smooth real 1-manifold, depicted as a plot with x and y-axis. (Note that a compact 1-manifold cannot be embedded into $\mathbb{R}$, because compact subsets of $\mathbb{R}$ are closed and bounded, thus necessarily have a boundary. A proper graph would live in a 4-dimensional space, which is rather difficult to visualize, so we settle with an approximate representation.)



Here, X and Y are both isomorphic to the unit circle.
We count the number of points in the fiber above each point in Y:

- Above point A, there are infinitely many points.
- Above point B, there is only one point. (You can argue that this point has "multiplicity 2" however)
- Above point C, there are two points.
- Above point D, the fiber is empty.

Definition 48.3.3. The value d_y above is called the **degree** of the map f, written $\deg(f)$.

Example 48.3.4

The map $z \mapsto z^k$, when extended to a $\mathbb{C}_\infty \to \mathbb{C}_\infty$ map, has degree k.

If $t \neq 0$, then we know that t has k distinct k-th roots. But if $t = 0$, its preimage only consist of the point 0 — in this case, we wish to say $z = 0$ is a "multiple point" — we will formalize it next section when we defines the multiplicity of a map.

If you have read Section 74.1, this is in fact the same as the concept of a degree in homology when X and Y are both Riemann spheres — it counts how many spherical bags that $\operatorname{im} f$ consists of. But, in this case, the theory is extra nice — not only that the graph is homotopy equivalent to one that covers each point d times, but each point is in fact covered *exactly* d times!

This theme will be recurrent in complex analysis and Riemann surfaces. Basically:

If the "things" are counted properly, the formula is very nice.

The proof of the proposition is not difficult — the main observation is that the theorem is true for functions of the form $f(z) = z^n$, and locally around each point $p \in X$, f is either an isomorphism or has the form above. So d_y is locally constant, and thus constant because Y is connected.

§48.4 Multiplicity of a map

Prototypical example for this section: $f(x) = (x-3)^2$ *has* $\operatorname{mult}_3(f) = 2$.

In the previous section, we informally talk about the multiplicity of a map at a point. We will rigorously define it in this section.

Example 48.4.1

Consider the map $f\colon \mathbb{C} \to \mathbb{C}$ given by $f(z) = z^5 + 1$.

Above each point $y \in \mathbb{C}$, the fiber $f^{\text{pre}}(y)$ has 5 points — except when $y = 1$, then $f^{\text{pre}}(1) = \{0\}$ has only 1 point.

This behavior is *undesirable*, and we would like to say that the function f maps 5 "identical copies" of the point 0 to the point 1. (Another way you could see it is that, for each sequences $\{y_i\}$ converging to 1, there are 5 different sequences $\{x_i\}$ converging to 0 such that $f(x_i) = y_i$ for each i.)

Inspired by this, we will define multiplicity in a way such that:

- $z \mapsto z^m$ has multiplicity m, for integer $m \ge 1$.
- If we perform an analytic reparametrization of the source or the target, then the degree does not change.

Turns out these two properties completely defines the degree! We have the following.

Proposition 48.4.2

Let $f\colon X \to Y$ be a nonconstant holomorphic map defined at $p \in X$. Then there is a unique integer $m \ge 1$ such that, for every chart $\phi_2\colon U_2 \to V_2$ on Y centered at $f(p)$ (that is, $\phi_2(f(p)) = 0$), there is a chart $\phi_1\colon U_1 \to V_1$ on X centered at p such that the induced map $\phi_2 \circ F \circ \phi_1^{-1}\colon V_1 \to V_2$ has the form $z \mapsto z^m$.

In other words, once we fix a chart of Y, there exist a chart of (an open subset of) X such that the induced map between open subsets of $\mathbb{C}$ is a power map; furthermore, the exponent is independent of the selection.

Every map looks locally like $z \mapsto z^m$.

Proof. Essentially, use the Taylor expansion to determine m, then the selection of ϕ_1 is pretty much fixed by the restrictions. □

Definition 48.4.3. The value m above is the **multiplicity** of f at point p, written $\operatorname{mult}_p(f)$.

Example 48.4.4 (More examples of multiplicity of a map at a point)

We consider some examples.

- The function $z \mapsto z^{-2}$, extended to a $\mathbb{C} \to \mathbb{C}_\infty$ map, has multiplicity 2 at point 0 — "two copies" of the point 0 is mapped to the point ∞.
- The function $f(z) = (z-1)(z-2)^5$ has $\operatorname{mult}_2(f) = 5$ — more generally, if p is a root of f, then $\operatorname{mult}_p(f)$ is the multiplicity of the root.

- The function $z \mapsto z+1$ has multiplicity 1 everywhere — in fact, the multiplicity of a nonconstant map at "most" points will be 1.

These are the official terms:

Definition 48.4.5. A point p such that $\operatorname{mult}_p(f) > 1$ is called a **ramification point**. In that case, the point $f(p)$ is called a **branch point**.

§48.5 The sum of the orders of a meromorphic function

Yet another case where we get a nice formula.

Example 48.5.1

Let us consider some meromorphic $\mathbb{C}_\infty \to \mathbb{C}_\infty$ functions (defined by extending a $\mathbb{C} \to \mathbb{C}$ function the obvious way), and list the zeros and poles of it (with multiplicity).

Function	Zeros	Poles
5	None	None
$(x+1)^2$	$-1, -1$	∞, ∞
$\frac{1}{x^2+1}$	∞, ∞	$i, -i$
$\frac{x+1}{x+2}$	-1	-2

Every time, the number of zeros equals the number of poles. This is not a coincidence!

Proposition 48.5.2

Let $f\colon X \to \mathbb{C}$ be a nonconstant meromorphic function on a compact Riemann surface X. Then

$$\sum_p \operatorname{ord}_p(f) = 0.$$

Of course, we need X to be compact — there certainly are $\mathbb{C} \to \mathbb{C}$ functions that has several zeros, but no poles.

Proof. We extend f to a $X \to \mathbb{C}_\infty$ function, then the sum of multiplicities of points in the fiber of 0 is equal to that in the fiber of ∞. ☐

§48.6 The Hurwitz formula

write this one. It's quite nice actually

§48.7 The identity theorem

The following propositions are expected — the same behavior is seen in complex analysis with holomorphic functions.

Theorem 48.7.1

Let $f, g\colon X \to Y$ be holomorphic maps between Riemann surfaces. If $f = g$ on a nonempty open subset of X, then $f = g$.

This is the analog of Problem 31C*. Note that here the assumption that X is connected is used — the disjoint union of two copies of $\mathbb{C}$ is a smooth 2-manifold, but not a Riemann manifold.

That is,

Holomorphic maps are *rigid* — the value of a function on a tiny subset determines its value everywhere.

49 Affine and projective plane curves

In this chapter, we will define affine and projective plane curves. This has two purposes:

- Many interesting curves in $\mathbb{R}^2$ can be defined as the set of roots of a polynomial. This is just a natural generalization.
- We will see that, in fact, *every* compact Riemann surfaces can be written as a projective curve! Thus, by studying the projective curves, we have in fact studied all compact Riemann surfaces.

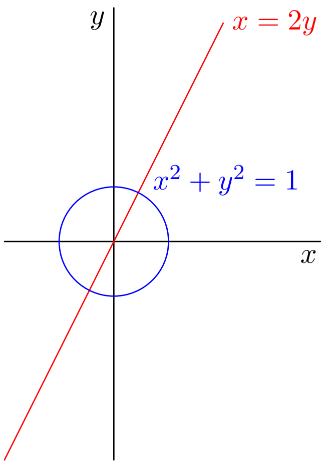
We will see what these means in the following sections.

§49.1 Affine plane curves

Consider some familiar curves on the plane.

- A line can be represented by an equation $y = ax + b$, or $x = c$.
- A circle can be written as the set of $y = \pm\sqrt{1 - x^2}$ for $-1 \le x \le 1$.

There is not much going on so far, but here is a picture.



As you can see, the definitions above are actually quite clumsy. We can do better by defining the points on the curve *implicitly*:

- A line can be represented as the set of (x, y) such that $ax + by + c = 0$.
- A circle can be represented as the set of (x, y) such that $x^2 + y^2 = 1$.

Of course, this way it is harder computationally to compute the coordinate of a point, but the definition is nicer.

The point is:

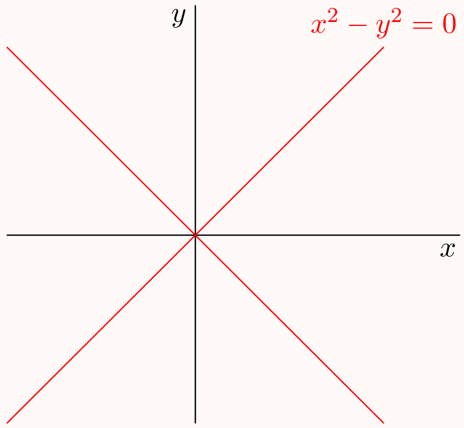
Many of the interesting curves can be written as the set of roots of a polynomial.

So we will try to do the same here — intuitively, if we start with complex dimension 2 and specify one polynomial, then the remaining part has complex dimension 1 i.e. a Riemann surface.

First, there is a technical detail we need to sort out — the set of roots of a polynomial need not be a smooth curve.

Example 49.1.1

The set of roots of $x^2 - y^2 = 0$ in $\mathbb{R}^2$ is not a curve near the origin — there are two intersecting curves.



This can be easily handled by placing a restriction on the polynomial. Let $f(x, y)$ a polynomial, and $X = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$. Then:

Theorem 49.1.2

For a point $(x, y) \in X$ such that not both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ vanishes, then X is smooth near (x, y).

If at a point $(x, y) \in X$ such that $\frac{\partial f}{\partial x} \neq 0$ or $\frac{\partial f}{\partial y} \neq 0$, we say X is **smooth** or **nonsingular** at (x, y).

In fact, we have something more. With notation as above, let $(x, y) \in X$, then:

- Suppose $\frac{\partial f}{\partial x} \neq 0$, then near the point (x, y), X can be parametrized by $x = g(y)$ for some analytic function g.
- Suppose $\frac{\partial f}{\partial y} \neq 0$, then near the point (x, y), X can be parametrized by $y = h(x)$ for some analytic function h.

All these are just the implicit function theorem.

Exercise 49.1.3. Check the statement above on the circle $x^2 + y^2 = 1$, at the points $(0, 1)$ and $(1, 0)$.

The exact same statement holds if we replace $\mathbb{R}^2$ with $\mathbb{C}^2$.

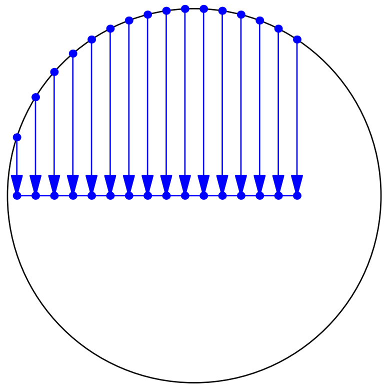
Next, we want the set of roots X to actually be a *Riemann surface*, not just a set of points in $\mathbb{C}^2$. So, we would need to find a suitable analytic structure on X.

In the circle above, what would be a suitable analytic structure? One possible thought is to unroll the circle by arc-length and map it onto $\mathbb{R}$, but for a Riemann surface this isn't even well-defined — how would you unroll, let's say a sphere onto a plane?

Another possibility is, given the statement of the implicit function above, we declare:

- On an open set $U \subseteq X$ where $\frac{\partial f}{\partial x} \neq 0$ for all points in U, suppose U is small enough such that X can be parametrized by $x = g(y)$ for some analytic function g, then the map ϕ such that $\phi(x, y) = \phi(g(y), y) = y$ is a complex chart.
- Similar for the $\frac{\partial f}{\partial y} \neq 0$ case.

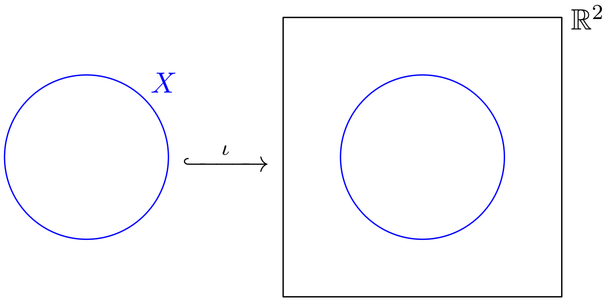
A possible complex chart is depicted below. Intuitively, the fact that $\frac{\partial f}{\partial y} = 0$ at the two points $(1, 0)$ and $(-1, 0)$ reflects that this "project-to-x" complex chart cannot be used at these points.



Actually, in the real analytic case, the two definitions above are equivalent. You can optionally do the exercise below.

Exercise 49.1.4. Show this for the circle above. (One possibility is to write down an explicit formula for the arc length and show it is analytic)

While this definition is already somewhat natural, there is something more to this. In category theory, we study properties of objects by studying the maps between them. The set X above has a natural map — the inclusion map into $\mathbb{R}^2$, and $\mathbb{R}^2$ has an obvious existing analytic structure.



The analytic structure defined above is natural in the following sense:

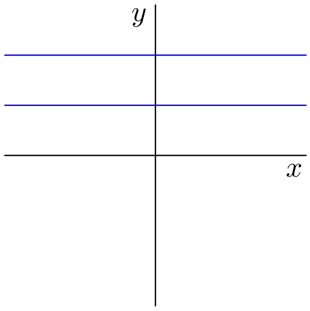
- For a function g such that $Y \xrightarrow{g} X \xrightarrow{\iota} \mathbb{R}^2$, then g is analytic if and only if $\iota \circ g$ is analytic.
- For a function $X \xrightarrow{\iota} \mathbb{R}^2 \xrightarrow{g} Y$, then g is analytic if and only if $g \circ \iota$ is analytic.
- $X \xrightarrow{\iota} \mathbb{R}^2$, then ι is analytic, and for any other complex structure $X' \xrightarrow{\iota'} \mathbb{R}^2$ such that ι' is analytic, there exists an unique analytic map $X' \to X$.

In fact, each of the bullet point uniquely determines the complex structure on X.

In some sense, this is like a universal property for our natural analytic structure.

Of course, we haven't defined what an analytic real manifold is. Brave readers may try to rigorously formalize all these concepts and prove the statement above.

There is another technical detail that needs to be sorted out. The set of zeros of $f(x, y) = (y-1)(y-2)$ is:



This is certainly smooth — but it's not connected. We required a Riemann surface to be connected.

Apart from these two issues, our final statement is:

Definition 49.1.5. Given a polynomial $f(z, w) \in \mathbb{C}[z, w]$, let $X = \{(z, w) \in \mathbb{C}^2 \mid f(z, w) = 0\}$ be the set of roots of f. Suppose that X is connected, and for all $(z, w) \in X$, f is smooth at (z, w) (that is, either $\frac{\partial f}{\partial z} \neq 0$ or $\frac{\partial f}{\partial w} \neq 0$). Then, X is a Riemann surface — we call X an **(smooth) affine plane curve**, with complex charts defined by:

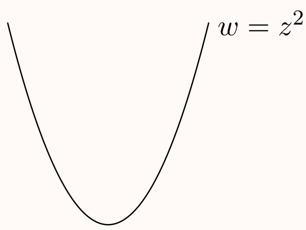
- On an open set U such that $\frac{\partial f}{\partial z} \neq 0$ everywhere on U, then $\phi\colon U \to \mathbb{C}$, $\phi(z, w) = w$ is a complex chart.
- On an open set U such that $\frac{\partial f}{\partial w} \neq 0$ everywhere on U, then $\phi\colon U \to \mathbb{C}$, $\phi(z, w) = z$ is a complex chart.

We call them affine because the plane is "flat", unlike the projective plane $\mathbb{CP}^2$ which is more "curved" in some sense.

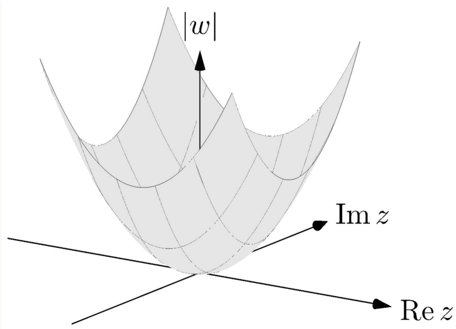
Of course, we should have some examples — with these tools, we are in a position to define an (affine) elliptic curve, and other affine curves.

Example 49.1.6 (A parabola)

Consider the Riemann surface cut out by $w = z^2$. Its real part looks like a parabola:



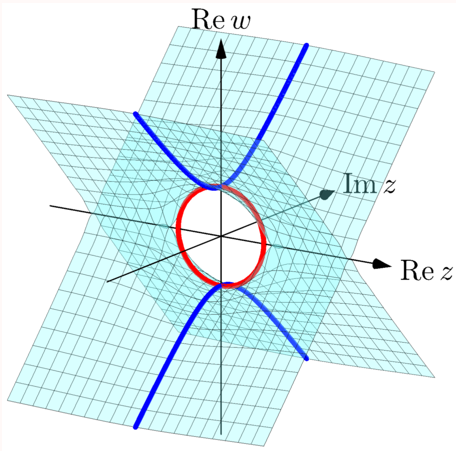
Since drawing a graph in 4 dimensions is difficult, we will project the Riemann surface onto 3 dimensions. The result is:



This Riemann surface is in fact isomorphic to the complex plane $\mathbb{C}$ by $(z, w) \mapsto z$.

Example 49.1.7 (The circle)

We all know what the real part of the circle looks like. Visualizing the whole Riemann surface is a bit more difficult, however.



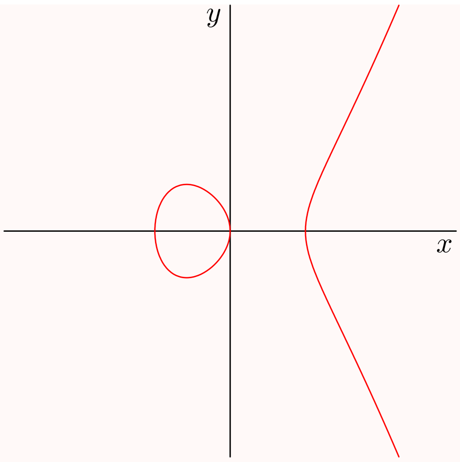
The highlighted red circle is the real part. Note that the fact that the plane is shown to be self-intersecting is merely an artifact of the projection.

Although the circle is not isomorphic to the complex plane $\mathbb{C}$ (we won't be able to prove this any time soon[a]), it is in fact isomorphic to the hyperbola $x^2 - y^2 = 1$ given by the transformation $y \mapsto y \cdot i$. With another rotation and multiplication by a constant, it is in turn isomorphic to the hyperbola $xy = 1$, which is "almost" isomorphic to the line $x = y$, missing one point $(0, 0)$.

[a] If you have read the homotopy chapter, this Riemann surface has a deformation retract to its real part — the circle, thus is homotopic to it. We know the complex plane $\mathbb{C}$ is nulhomotopic instead.

Example 49.1.8 (The elliptic curve $y^2 = x^3 - x$)

The real part looks like this. (The complex part is not drawn this time.)



While we won't be able to prove this any time soon, turns out this Riemann surface is not isomorphic to $\mathbb{C}$ — even if we allow deleting finitely many points.

§49.2 The projective line $\mathbb{CP}^1$

We will define the projective line — as it will turn out, it is isomorphic to the Riemann sphere $\mathbb{C}_\infty$ which we have already defined. So this section is only to show how our tools work.

As you might have guessed by the name: as a set of points, $\mathbb{CP}^1$ is the quotient of the set of points $\mathbb{C}^2 \setminus \{0\}$, modulo the relation $(x, y) \sim (\lambda x, \lambda y)$ for any $\lambda \in \mathbb{C} \setminus \{0\}$.

As a topological complex manifold, fortunately, it is still easy — $\mathbb{C}^2 \setminus \{0\}$ has a natural topology, and $\mathbb{CP}^1$ gets the quotient topology.

Exercise 49.2.1. Define the topology on the space $\mathbb{RP}^1$ analogously.

Exercise 49.2.2. Let $X \subseteq \mathbb{R}^2$ be a line that does not pass through the point $(0, 0)$. Show that $X \xrightarrow{f} \mathbb{R}^2 \xrightarrow{q} \mathbb{RP}^1$ is an embedding i.e. $X \xrightarrow{q \circ f} \operatorname{im}(q \circ f) \subseteq \mathbb{RP}^1$ is a homeomorphism.

As a Riemann surface, the usual textbook definition goes:

Definition 49.2.3 (Complex structure of $\mathbb{CP}^1$)**.** Cover $\mathbb{CP}^1$ by two open sets, U_1 consisting of points with nonzero x coordinate, and U_2 consisting of points with nonzero y-coordinate. Then the two complex charts $\phi_1 \colon U_1 \to \mathbb{C}$ given by $\phi_1(x, y) = y/x$ and $\phi_2 \colon U_2 \to \mathbb{C}$ given by $\phi_2(x, y) = x/y$ determines a complex structure.

And goes on to prove that the two open sets indeed cover the whole of $\mathbb{CP}^1$, the value y/x is well-defined, transition maps are holomorphic, etc.

The definition above is elementary, but uninstructive. Where does the complex charts come from?

Given what we have done in the previous chapter, it should be obvious where we should go from here. There are two things to try:

- Let X be an affine plane curve in $\mathbb{C}^2$ that does not contain the point 0. Then the map $X \hookrightarrow \mathbb{C}^2 \twoheadrightarrow \mathbb{CP}^1$ should be an isomorphism whenever some certain derivative does not vanish.

- We can also use maps: the complex structure is such that whenever we have $Y \xrightarrow{f} \mathbb{C}^2 \xrightarrow{q} \mathbb{CP}^1$ or $\mathbb{C}^2 \xrightarrow{q} \mathbb{CP}^1 \xrightarrow{g} Y$, then f is analytic if and only if $q \circ f$ is analytic; and g is analytic if and only if $g \circ q$ is analytic.

Both are equivalent to the definition above — in fact, the definition is merely a special case of the first bullet point, where X is taken to be the line $x = 1$ and $y = 1$ respectively. Coincidentally, the 2 resulting complex charts is the simplest one to write down algebraically, and they already cover the whole $\mathbb{CP}^1$, so it is often taken to be the definition. There is no reason why it must be these 2 lines however — you might as well use $x + y = 1$ and $x - y = 1$.

§49.3 Projective plane curves

Instead of using affine plane curves $X \subseteq \mathbb{C}^2$, this time around, we will define projective plane curves $X \subseteq \mathbb{CP}^2$.

Apart from "just another source of example", projective plane curves have a distinctive advantage — *they're compact*! This allows many nice properties to hold — we have seen a few in the last chapter.

We start with defining the **projective plane** $\mathbb{CP}^2$. Of course it is $\mathbb{C}^3 \setminus \{0\}$ quotient by the relation $(x, y, z) \sim (\lambda x, \lambda y, \lambda z)$. It has a natural 2-dimensional complex structure induced from $\mathbb{C}^3$ by the quotient map.

The above definition is natural, but abstract. Concretely, we can write:

Question 49.3.1. Define the three complex-manifold charts (on the open set where they're well-defined) by:

$$\begin{aligned} \phi_0(x, y, z) &= (y/x, z/x) \\ \phi_1(x, y, z) &= (x/y, z/y) \\ \phi_2(x, y, z) &= (x/z, y/z). \end{aligned}$$

Convince yourself that this complex manifold structure is the correct one.

Then, a projective plane curve X is defined to be the set of points (x, y, z) such that $f(x, y, z) = 0$ — again, satisfying certain smoothness and connectedness conditions. Unfortunately, if the polynomial were e.g. $f(x, y, z) = x - 1$, it will not be well-defined, as $f(1, 0, 0) = 0$ but $f(2, 0, 0) = 1$. So we require that f is homogeneous — that way, $f(x, y, z)$ is still not well-defined, but at least we know whether $f(x, y, z) = 0$.

The complex structure on a projective plane curve is similarly defined by the universal property.

The definition is short and natural, but abstract. A more concrete definition is given below.

Question 49.3.2. With notation as above, define U_0, U_1 and U_2 to be the domain of ϕ_0, ϕ_1 and ϕ_2 respectively. Note that $U_i \xrightarrow{\phi_i} \mathbb{C}^2$ gives an isomorphism between U_i and the affine plane $\mathbb{C}^2$.

Convince yourself that the intersection of a projective plane curve X with one of the U_i is a (possibly empty) affine plane curve, when mapped to $\mathbb{C}^2$, and all the mappings are isomorphisms.

We need some examples.

Example 49.3.3 (The Riemann sphere, again)

The Riemann sphere can alternatively be defined as the set of points where $z = 0$ in $\mathbb{CP}^2$.

There's nothing interesting about this — we already know how the Riemann surface looks like. It just serves as a trivial example.

Example 49.3.4 (An elliptic curve, again)

Let $f(x, y) = x^3 - x - y^2$. We know that the set of roots of f in the affine plane $\mathbb{C}^2$ is the elliptic curve.

Identifying $\mathbb{C}^2$ with U_2, most points in $\mathbb{CP}^2$ can be written as $(x, y, 1)$. We want to find a polynomial $g(x, y, z)$ such that its set of roots in $\mathbb{CP}^2$, restricted to U_2, equals to the elliptic curve.

Intuitively, by the identity theorem, this should suffices to uniquely determines the Riemann surface. Indeed, our target polynomial g is:

$$g(x, y, z) = x^3 - xz^2 - y^2 z.$$

This is just the laziest way to homogenize the polynomial f, multiplying the least power of z to make the result a homogeneous polynomial, and that $g(x, y, 1) = f(x, y)$.

We have that $\mathbb{CP}^2$ is compact, and the set of roots of g is closed, therefore the resulting Riemann surface is *compact*! As promised.

visualize this

As it turns out, unlike the Riemann sphere, the Riemann surface defined by the elliptic curve above has *genus 1*! We have the first example that is definitely distinct from the Riemann sphere.

Exercise 49.3.5. In the example above, what if we multiply a larger power of z? For instance

$$g(x, y, z) = x^3 z - xz^2 - y^2 z.$$

Example 49.3.6 (A hyperelliptic curve)

Let $f(x, y) = (x - x_1)(x - x_2) \cdots (x - x_{2k+1}) - y^2$, where all of $x_1, \dots, x_{2k+1}$ are distinct complex numbers.

We can homogenize f to get $g(x, y, z) = (x - x_1 z)(x - x_2 z) \cdots (x - x_{2k+1} z) - y^2 z^{2k-1}$.

As above, the set of roots of g in $\mathbb{CP}^2$ cuts out a Riemann surface — once again, this has *genus k*!

Therefore, we have seen examples of compact Riemann surfaces of all the genera simply by picking different values of k.

Saying that we have "seen" the surfaces themselves is not quite accurate — but you can try to visualize these hyperelliptic curves the same way the elliptic curve is visualized.

§49.4 Filling in the holes

Prototypical example for this section: The Riemann sphere is formed by filling in a single hole in the complex plane $\mathbb{C}$.

write this

§49.5 Nodes of a plane curve

Prototypical example for this section: The set defined by the equation $x^2 - y^2 = 0$ has a simple node.

50 Differential forms

§50.1 Differential form on $\mathbb{C}$

In this section, we will generalize the definition of what can be contour integrated.

Definition 50.1.1 (Differential 1-forms on $\mathbb{C}$)**.** A 1-form ω on an open set $U \subseteq \mathbb{C}$ is an expression of the form $f(z)d\operatorname{Re} + g(z)d\operatorname{Im}$, where $f(z)$ and $g(z)$ are smooth $U \to \mathbb{C}$ functions.

Here, smooth means being infinitely differentiable when interpreted as $\mathbb{R}^2 \to \mathbb{R}^2$ functions.

This is almost the same as the definition of a 1-form on $\mathbb{R}^2$! Here, Re and Im takes the role of $\mathbf{e}_1^\vee$ and $\mathbf{e}_2^\vee$ the obvious way.

The only difference is, as you can observe, $f(z)$ and $g(z)$ returns complex numbers instead of real numbers — but this is mostly inconsequential, by the projection principle (Theorem 43.2.1), the 1-form ω is equivalent to a pair of real-valued 1-forms $(\operatorname{Re}\omega, \operatorname{Im}\omega)$.

The reason why we want to do what we did is simply for convenience — by abuse of notation, let z be the function $z \mapsto z$, then we want dz to be a 1-form that returns the change in z.

§50.2 Visualization of differential forms

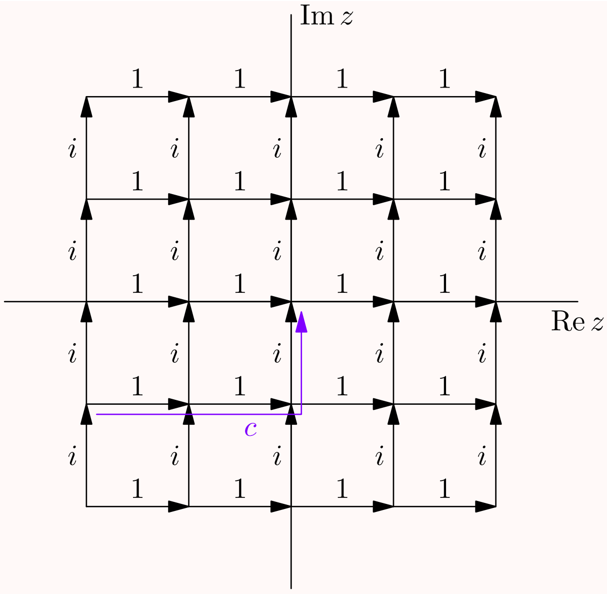
Because ω takes in a point and returns a $\mathbb{R}$-linear map from the tangent space, the obvious way to visualize it is to draw a quiver diagram — for each point, the value of $\omega_p(v)$ is plotted for vectors v, which we interpret as "if we integrate a curve c in the direction of v, with length approximately the length of v, close to the point p, then the result is approximately the labeled value.

To integrate a differential form ω over a curve c, simply add up the numbers that corresponds to the direction of movement of c that appears in the diagram.

With that visualization, here are some examples.

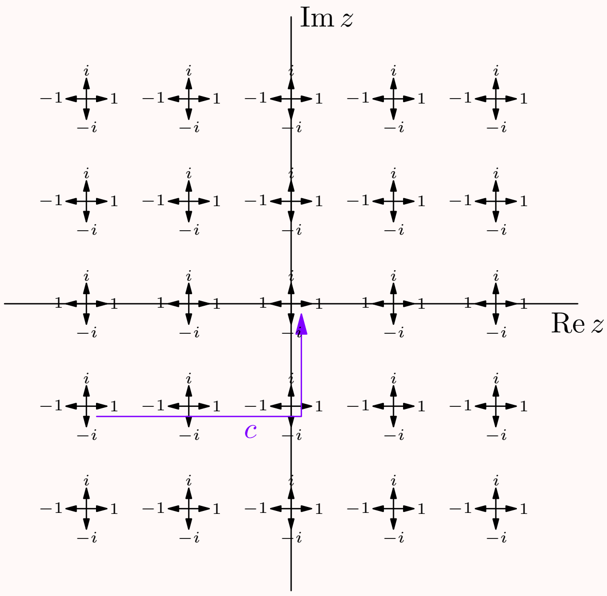
Example 50.2.1 (The 1-form dz)

We may visualize dz, which is just the change in z, as follows.



The integration $\int_c dz$ can be computed by adding up the value of the vectors together, so we get $2+i$ — this is indeed equal to the change of z as we travel along the curve, $0-(-2-i)=2+i$.

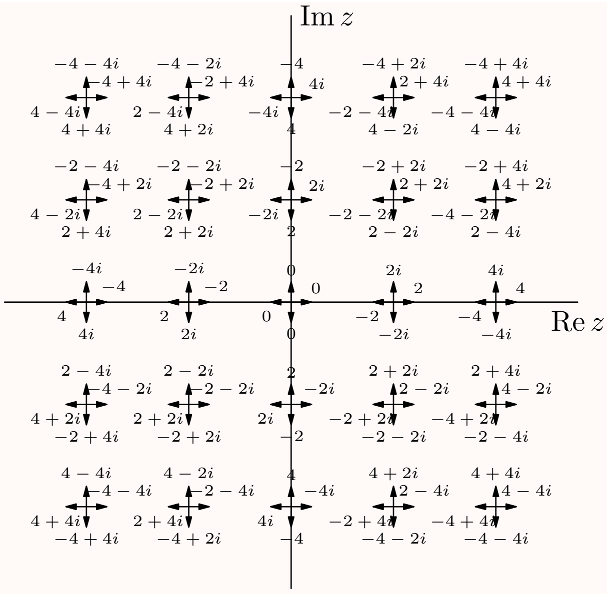
Since having the arrows extending the whole length can be confusing, we will shorten the arrow like the following.



Example 50.2.2 (Another holomorphic 1-form: $d(z^2)=2z\cdot dz$)

While we have never defined what a holomorphic 1-form is, it makes intuitive sense for the definition to satisfy that: if $f(z)$ is holomorphic, then $df(z)$ should be a holomorphic 1-form.

In any case, if you perform the same procedure as above and compute the differential change of z^2 along the tangent vectors, you will get the following.



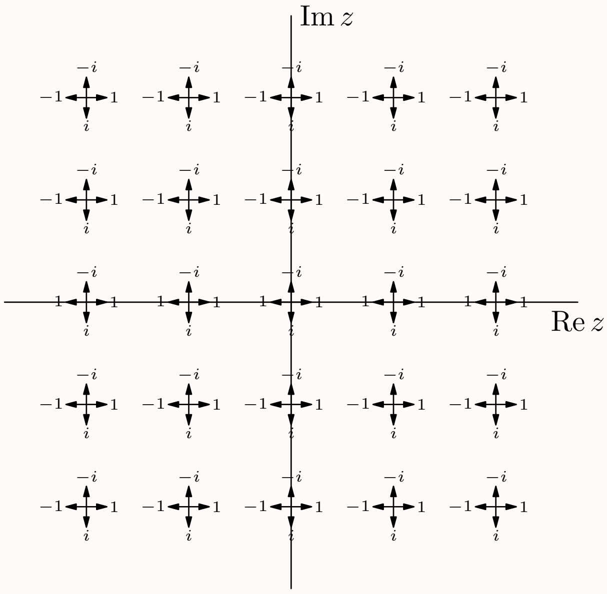
Unfortunately, it gets a bit cluttered regardless. Anyway, as you can see, at each point z and along each direction, the value of the 1-form $d(z^2)$ is $2z$ times the corresponding value of the 1-form dz, thus it makes sense for us to define multiplication such that $d(z^2) = 2z \cdot dz$.

Example 50.2.3 (A non-holomorphic form: $d\overline{z}$)

In both of the examples above, we see that, at each point z, $\omega_z(\mathbf{e}_2) = i \cdot \omega_z(\mathbf{e}_1)$, where $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ — in other words, rotating the vector counterclockwise by 90 deg multiplies the value of the differential form by i.

The natural question would be: Is this the property of all 1-forms? Turns out it isn't. (Later on, we will see that this is a common property of all holomorphic 1-forms, or more generally, all type $(1, 0)$ 1-forms.)

Consider the following example: let $\omega = d\overline{z}$ — this is just the change in value of $\overline{z}$.



Example 50.2.4 (Another non-holomorphic form: $\overline{z} \cdot dz$)

Just as how a smooth function $f(z)$ is holomorphic if and only if it is complex-differentiable, we should define a holomorphic 1-form such that a smooth 1-form ω is holomorphic if and only if we can compute its differential $d\omega$.

We certainly haven't defined a 2-form yet, nor have we defined what it means to differentiate a 1-form ω to a 2-form.

§50.2.i Holomorphic forms

With the above examples in mind, we defines:

Definition 50.2.5 (Holomorphic 1-forms on the complex plane)**.** A 1-form ω is holomorphic if and only if it can be written as $f(z) \cdot dz$ for some holomorphic function f.

And also a few other types.

Definition 50.2.6 (Type $(1,0)$ and type $(0,1)$ 1-forms)**.** A 1-form is of **type** $(1,0)$ if it is locally of the form $f(z)dz$ for smooth f. Similarly, a 1-form is of **type** $(0,1)$ if it is locally of the form $f(z)d\overline{z}$ for smooth f.

Example 50.2.7 (Some type $(1,0)$ 1-forms)

Looking at the examples above:

- The holomorphic forms, dz or $2z \cdot dz$, are of course type $(1,0)$.
- $\overline{z} \cdot dz$ is still a type $(1,0)$ form, even though it is not holomorphic.
- $d\overline{z}$, however, is not a type $(1,0)$ form.

Why do we care? Note that it is nontrivial that the definition above is well-defined — it only makes sense because a holomorphic function scales every direction the same amount!

Intuitively,

A $(1,0)$ form ω is a form such that $\omega_p(\mathbf{e}_2) = \omega_p(\mathbf{e}_1) \cdot i$.

§50.2.ii Putting the pieces together: 1-forms on a Riemann surface

write this

Unsurprisingly, now we can define a 1-form on a Riemann surface.

51 The Riemann-Roch theorem

§51.1 Motivation

Recall a basic fact in complex analysis:

> A holomorphic $\mathbb{C} \to \mathbb{C}$ function is uniquely determined by its Taylor series expansion at the origin.

Compared to the case of real smooth function, this is already very rigid — the value of the function in a small neighborhood of the origin determines the value of the function everywhere — but, in order to specify a function, you still need infinitely many coordinates!

Meanwhile, we have Liouville's theorem:

> A bounded holomorphic $\mathbb{C} \to \mathbb{C}$ function is constant.

As we have learnt earlier, this theorem, when phrased in terms of Riemann surfaces, can be more elegantly rephrased to the following:

> A holomorphic $\mathbb{C}_\infty \to \mathbb{C}$ function is constant.

In other words, in order to specify a holomorphic $\mathbb{C}_\infty \to \mathbb{C}$, you only need a *single complex number*! That is, the $\mathbb{C}$-vector space $\operatorname{Hom}(\mathbb{C}_\infty, \mathbb{C})$ has dimension 1.

Naturally, you may ask, "is there anything inbetween"? There is! And the Riemann-Roch theorem is the main ingredient to understand how these things work.

So, how are we going to define this? If you compare the two situations above, a holomorphic $\mathbb{C} \to \mathbb{C}$ function is a meromorphic $\mathbb{C}_\infty \to \mathbb{C}$ function, which is allowed to have a pole at ∞, and nowhere else.

So,

> **By smoothly interpolate between "allow pole of arbitrary order" and "must be holomorphic", we can produce many interesting spaces of functions.**

Conveniently, back in chapter 32, we have defined the **multiplicity** of a zero and the **order** of a pole of a meromorphic function. So, the natural point between these two extremes is to allow a pole of order at most d.

For notational convenience, we defines:

Definition 51.1.1 (Order of a meromorphic function)**.** Let f be meromorphic at p. We define $\operatorname{ord}_p(f)$ to be:

- d, if f has a zero of multiplicity d at p;
- $-d$, if f has a pole of order d at p;
- 0, otherwise.

Example 51.1.2 (The space of functions with pole of order at most 4 on $\mathbb{C}_\infty$)

Let $L(4 \cdot \infty)$ be the set of meromorphic $\mathbb{C}_\infty \to \mathbb{C}$ function, being holomorphic everywhere except ∞, and has a pole of order at most 4 at ∞ — in other words,

$$L(4 \cdot \infty) = \{f \text{ meromorphic on } \mathbb{C}_\infty \mid f \text{ defined on } \mathbb{C}_\infty \setminus \{\infty\}, \operatorname{ord}_\infty(f) \geq -4\}.$$

(The notation $L(-)$ will be explained later.)

Obviously, this forms a natural $\mathbb{C}$-vector space.

Consider the Taylor series of any $f \in L(4 \cdot \infty)$ at the origin:

$$f(z) = \frac{c_{-m}}{z^m} + \frac{c_{-m+1}}{z^{m-1}} + \cdots + \frac{c_{-1}}{z} + c_0 + c_1 z + \cdots$$

Obviously, because f is defined at the origin, it cannot have any nonzero coefficient c_{-m} for $m > 0$. But more importantly, it cannot have any nonzero coefficient c_m for $m > 4$ either![a]

Did you see what happened here? We start with requiring the function to be analytic and does not blow up too badly, and we end up with just the *algebraic* function — the polynomials!

In particular, $L(4 \cdot \infty)$ consists of the polynomials of degree ≤ 4, and

$$\dim L(4 \cdot \infty) = 5$$

as a $\mathbb{C}$-vector space.

[a] The reason is actually not very straightforward, but you can see for yourself why it is true: if there are only finitely many nonzero terms, then the order of the pole at ∞, $(-\operatorname{ord}_\infty(f))$, is precisely the degree of the highest nonzero coefficient.

Example 51.1.3 (More complicated $L(-)$ spaces)

There's no reason why we should restrict ourselves to considering only the functions that blow up at ∞ — as we will see, more general meromorphic functions can be considered, as long as we restrict the order of the poles.

Let $L(-1 \cdot 3 + 4 \cdot i + 5 \cdot \infty)$ be the set of meromorphic functions $f \colon \mathbb{C}_\infty \to \mathbb{C}$ that are:

- holomorphic everywhere in $\mathbb{C}_\infty$, possibly with the exception of the points 3, i, and ∞;
- at 3, it must have a root of order ≥ 1;
- at i, it cannot have a pole of order more than 4;
- at ∞, it cannot have a pole of order more than 5.

So, for example, $(z \mapsto z - 3)$, $(z \mapsto \frac{(z-3)^3}{(z-i)^2})$, or $(z \mapsto (z-3)^4)$ are functions in the set, but not $(z \mapsto (z-3)^2 + 1)$ or $(z \mapsto (z-3)^7)$.

As before, this is a $\mathbb{C}$-vector space, and furthermore, it is also finite-dimensional! What should its dimension be?

Well, note that there is a 1-1 bijection between functions $f \in L(-1 \cdot 3 + 4 \cdot i + 3 \cdot \infty)$

and functions $g \in L(-1 \cdot 3 + 7 \cdot \infty)$ by

$$g = \Phi(f) = (z \mapsto f(z) \cdot (z - i)^4),$$

where, as you could probably have guessed by now, $L(-1 \cdot 3 + 7 \cdot \infty)$ is the space of meromorphic functions that has at least a zero at 3 and at most a pole of order 7 at ∞.

Using that information, it shouldn't be too hard for you to see that the dimension should be 7.

For another piece of motivation: Later on, we will also define the concept of **divisors** and **line bundles**. If you have learned about these concepts in algebraic geometry context, you might be interested to learn what they are actually about; otherwise, it is still very surprising that these theorems can be naturally generalized to completely algebraic settings, and *your intuition from the case of analytic manifold will mostly work verbatim* — in fact, you can even define the genus of a number field, like $\mathbb{Q}[\sqrt{2}]$!

§51.2 Divisors

Prototypical example for this section: $(-3) \cdot i + (-4) \cdot \infty$ *is a divisor on* $\mathbb{C}_\infty$*.*

We start with defining a convenient notation for the above concepts.

First, observe that the condition "f must have a zero of order at most 4 at the origin" can be conveniently written as

$$z^4 \mid f.$$

In other words, z^4 must be a **divisor** of f.

This notation works if f is a polynomial, since we already know what it means for two polynomials to divide each other.

Generalizing, we could say "f cannot have a pole of order more than 3 at the point i, and f cannot have a pole of order more than 4 at the point ∞" by

$$(z - i)^{-3} \cdot (z - \infty)^{-4} \mid f.$$

Of course, at this point, the notation is purely formal — there is no interpretation as "functions" that could be assigned to the expression $(z - \infty)$, for instance.

Those objects are, appropriately enough, called **divisors**. So we come to the formal definition:

Definition 51.2.1 (Divisors)**.** Let X be a Riemann manifold, then a divisor D on X is a function $D \colon X \to \mathbb{Z}$, which is nonzero on a discrete set of points.

The formal objects $(z - i)^{-3} \cdot (z - \infty)^{-4}$ above, from now on, we will consider it as a function $D \colon \mathbb{C}_\infty \to \mathbb{Z}$ by

$$D(z) = \begin{cases} -3 & z = i \\ -4 & z = \infty \\ 0 & \text{otherwise.} \end{cases}$$

Abuse of Notation 51.2.2. For a point $p \in X$, we write p to mean the divisor that takes value 1 at p, and 0 elsewhere.

Because divisors are integer-valued functions, we can add two divisors together or multiply a divisor with an integer, the result is an integer. So,

Example 51.2.3 ($(z-i)^{-3} \cdot (z-\infty)^{-4}$ as a divisor)

The divisor D that corresponds to the formal object $(z-i)^{-3} \cdot (z-\infty)^{-4}$ above can be written as $(-3) \cdot i + (-4) \cdot \infty$.

§51.3 Degree of a divisor

Prototypical example for this section: $\deg((-3) \cdot i + (-4) \cdot \infty) = -7$.

If the surface X is compact, any discrete set of points is finite. Thus, a divisor D on X has finite support.

This allows us to define the degree of a divisor:

Definition 51.3.1 (Degree of a divisor)**.** For a divisor D on a compact surface X, its degree is $\sum_{p \in X} D(p)$.

Of course, the sum is well-defined because only finitely many terms are nonzero.

§51.4 The principal divisor of a meromorphic function

Prototypical example for this section: $\operatorname{div} \frac{(z-3)^2}{z-i} = 2 \cdot 3 + (-1) \cdot i + (-1) \cdot \infty$ *has degree* 0.

After defining a divisor, we want a convenient notation to formalize our fuzzy notation earlier of a divisor "divides" a function.

Definition 51.4.1 (Divisor of a meromorphic function)**.** Let f be meromorphic on a Riemann surface X. Then the divisor of f, $\operatorname{div}(f)$, is such that

$$\operatorname{div}(f)(p) = \operatorname{ord}_p(f).$$

We can also write it as a formal sum: $\operatorname{div}(f) = \sum_p \operatorname{ord}_p(f) \cdot p$ — by the abuse of notation above, this would be an actual sum if f has finitely many roots and poles.

If a divisor D is equal to $\operatorname{div}(f)$ for some f, we call D a **principal divisor**. (Compare this with a principal ideal, being an ideal generated by one element!)

Looking at the prototype example of this section, you might have guessed the following for the Riemann sphere. In fact, much more is true:

Proposition 51.4.2

If a divisor D on a compact surface X is principal, then $\deg D = 0$.

Let us not forget our goal of defining a convenient notation to talk about the space of functions with bounded poles. With the notation defined above, if $f = z$, $\operatorname{div} f = 1 \cdot 0 + (-1) \cdot \infty$, and we want to say f "divides" the divisor $(-1) \cdot \infty$. The natural definition would be:

Definition 51.4.3 (The partial ordering of divisors)**.** We write $D \ge 0$ if $D(p) \ge 0$ for all $p \in X$.

For two divisors D_1 and D_2, we write $D_1 \ge D_2$ if $D_1 - D_2 \ge 0$.

And finally,

Definition 51.4.4. Let D be a divisor on a Riemann surface X. Then the space of meromorphic functions with poles bounded by D is

$$L(D) = \{f \text{ meromorphic on } X \mid \operatorname{div}(f) \geq -D\}.$$

Exercise 51.4.5. This exercise is just for you to get familiar with the notation. Show the following:

- For two divisors $D_1 \leq D_2$, then $L(D_1) \subseteq L(D_2)$.
- If X is compact, then $L(0) \cong \mathbb{C}$.
- If X is compact and $\deg D < 0$, then $L(D) = \{0\}$.

§51.5 The Riemann-Roch theorem

Explain canonical divisor

Theorem 51.5.1 (The Riemann-Roch theorem)
Let D be a divisor on an algebraic curve X of genus g, and K be a canonical divisor on X. Then

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

List some applications

52 Line bundles

§52.1 Overview

You might have heard about line bundles, which is somehow "a set L with a map $\pi\colon L \to X$ where the preimage of each point is a line". And then, in the algebraic geometry section, you come across the concept of "section" which appears to be just a function.

That sounds reasonable, but you may ask, "so what? Isn't it then just another complex manifold which has one more dimension than X? Why not just study complex manifold?"

It's true, but there are more structures on a line bundle:

- You can take the product of two line bundles, which somehow "add up the twists" of both line bundles.
- A section is not just a function — you can think of a section as the graph of a function in the special case that the "graph paper" itself is flat, but if it is curved like a Möbius strip, you will see that there is no way to assign a "function value" to each point of the "graph paper" — a situation which we will call "the line bundle is not trivial".

In other words, a line bundle vastly generalizes the "space of the graph of a function".

Later on, you will see a deep hidden connection between line bundles and linearly equivalent classes of divisors, and how they are all linked by the so-called Picard group.

§52.2 Definition

Let X be a Riemann surface.

In this section, we will view X as just a curve — that is, a 1-dimensional object instead of a 2-dimensional object — because:

- It is easier to visualize things when they can be embedded in 3-dimensional space. (Try to draw the graph of ... with both real and complex part, and you will see what I mean!)
- Since all of our functions of interest are analytic, the behavior of a function elsewhere is determined by its value on the real axis.

Looking only at the real part can makes some intuition slipped however — for example, it is possible to overlook that the circle $x^2 + y^2 = 1$ and the hyperbola $x^2 = 1 + y^2$ cuts out Riemann surfaces in $\mathbb{C}^2$ of the same shape, or that the function $\frac{1}{x^2+1}$ has a pole at $x = \pm i$. So, be careful.

Definition 52.2.1. A **line bundle** L is a set, together with:

- A projection map $\pi\colon L \to X$,
- An open cover $\{U_i\}$ of X,
- For each U_i, a **line bundle chart** $\phi\colon \pi^{-1}(U) \to \mathbb{C} \times U$ that bijectively maps each point in $\pi^{-1}(p)$ to a point in $\mathbb{C} \times p$,

- For two open sets U_1 and U_2, the **transition function** $\phi_2 \circ \phi_1^{-1} \colon \mathbb{C} \times U_1 \to \mathbb{C} \times U_2$ must be a $\mathbb{C}$-*vector space isomorphism* restricted to $\mathbb{C} \times p \to \mathbb{C} \times p$ for each point $p \in U_1 \times U_2$, and the scaling factor must be an analytic function on U.

Remark 52.2.2 (Warning) — Typically, we draw a graph of the function $f(x)$ by the set of points (x, y) where $y = f(x)$.

This time, we use the notation in [Mi95] — the target of a line bundle chart is $\mathbb{C} \times U$ instead of $U \times \mathbb{C}$ — so if we consider a section the generalization of a function, the coordinate would look like (y, x) instead.

The definition is dense, but essentially:

A line bundle is a set with a line bundle structure, consisting of an analytic structure and a 1-dimensional vector space structure.

The transition maps is simply to weld the pieces of the line bundle together, just like how they welded pieces of a Riemann surface in Chapter 47.

Another definition, we will explain this one later.

Definition 52.2.3 (Sections of a line bundle)**.** Let L be a line bundle. A **section** on an open set U is a map $f \colon U \to L$ such that $\pi \circ f$ is the identity map on U.

We call a section $f \colon X \to L$ a **global section**.

The section $f \colon U \to L$ is an **analytic section** if for every $U_1 \subseteq U$ such that there is a line bundle chart $\phi \colon \pi^{-1}(U_1) \to \mathbb{C} \times U_1$, then $\phi \circ f\restriction_{U_1} \colon U_1 \to \mathbb{C} \times U_1$ is analytic.

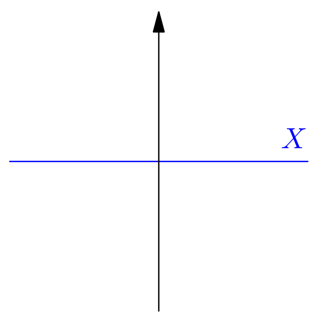
We will see this definition later on in algebraic geometry, Definition 82.2.2.

Remark 52.2.4 — In most books, they will first define what a sheaf is, then instead of "analytic section", they say "a section of the sheaf of analytic functions" (or regular functions, etc.)

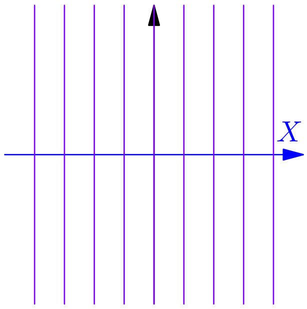
§52.3 Visualizing a line bundle

Just as how you can keep all the information of the Riemann sphere $\mathbb{C}_\infty$ in your head at once just by visualizing a sphere (with the analytic structure viewed as some "compatible grids" on the surface), you should also be able to keep all the information of a line bundle in your head at once — at least in the simplest cases.

First, we visualize $\mathbb{C} \times X$ where $X = \mathbb{C}$. Looking at only the real parts, it looks like a plane.

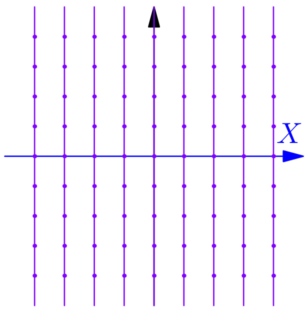


As a line bundle, the preimage of each point is a line.

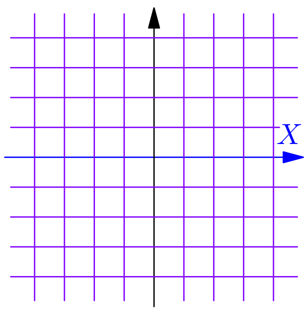


> **Question 52.3.1.** In symbols, what subset of $\mathbb{C} \times X$ does a vertical line correspond to?

They are not just disparate lines however — there are two more structures. First one is a vector space structure — of course the dimension of $\mathbb{C}$ as a $\mathbb{C}$-vector space is 1. We can visualize it by marking the points $1, 2, 3, \dots$ on the line.

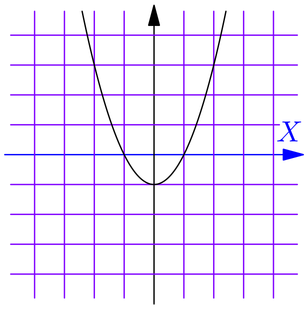


The other structure is that the lines must "smoothly varies" as p varies over X. We visualize this by drawing, well, a grid.



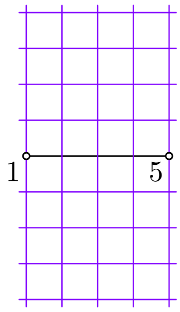
> **Question 52.3.2.** How does the picture of the grid correspond to the formal definition of a line bundle chart? (Hint: take the preimage of the vertical lines $x = c$ and the horizontal lines $y = c$ with respect to the line bundle chart $\phi \colon L \to \mathbb{C} \times U$, where $U \subseteq X$, then use the analytic structure on X to identify open subsets of U with open subsets of $\mathbb{C}$.)

So far, nothing surprising — this is just the usual grid graph, where we can draw functions on it like $y = x^2 - 1$, and a function is analytic if it is analytic with respect to the grid.

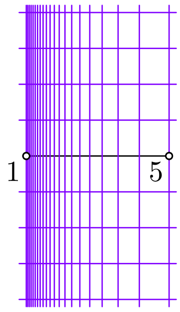


Of course, instead of a function, we call this a *section.* This particular section is in fact analytic, as you would expect.

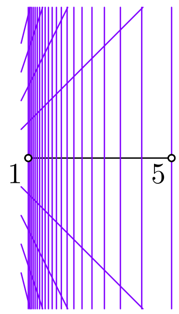
Let us take a look at which "grids" represent the same line bundle structure. For this part, we will look at $\{z \in \mathbb{C} \mid |z-3| < 2\}$, its real part being $(1, 5)$.



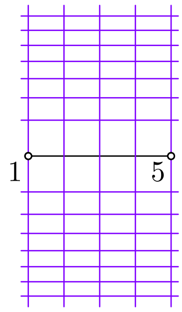
If we apply an analytic reparametrization on the segment $(1, 5)$ — for example let $t = 25/x$, then the grid becomes like the following. It still represents the same line bundle structure — in other words, the two charts are compatible.



If we rescale the vertical direction by an analytically-varying function, it still represents the same line bundle structure.



However, if we rescale the vertical direction by something that is not linear, the vector space structure will be changed. The following grid *does not* represent the same line bundle structure:



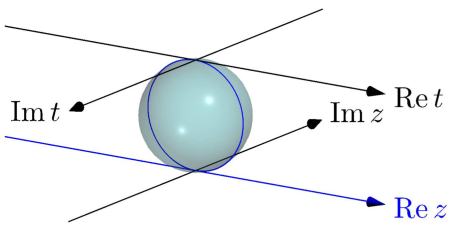
Intuitively, this makes sense — in a *vector space*, you can add two elements together and get another element — in our case, if $a, b \in L$ such that $\pi(a) = \pi(b)$, we can compute $c = a + b$ and get another element $c \in L$ with $\pi(c) = \pi(a) = \pi(b)$. If we rescale the vertical direction non-linearly, the element c will be changed.

Finally, don't forget that L still has an analytic structure — even though a section isn't necessarily a function, we are still able to say when a section is analytic.

> **Question 52.3.3.** Verify that everything explained above matches the formal definition. (This is important! Fuzzy pictures won't help you to understand the concepts; and if your intuition is incomplete or inaccurate, you will have a lot of trouble understanding the subsequent parts.)

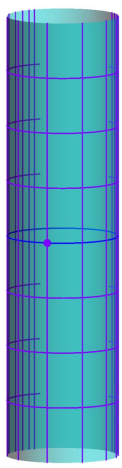
So far, everything just looks like a graph paper, on which a section looks just like a function.[1] Let us consider a more complicated space X — the Riemann sphere.

Because we are looking at the real part only, so once again, $\mathbb{C}_\infty$ looks like just a circle.



As before, we let z and t parametrize the points on the surface, with $t = \frac{1}{z}$ wherever both are defined.

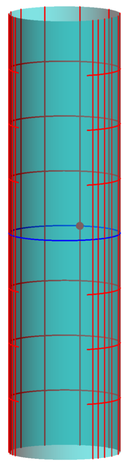
Still, we need two dimensions to embed a circle. So, the real part of $\mathbb{C} \times \mathbb{C}_\infty$ may looks something like the following:



[1] As warned above, "graph coordinate" is written (y, x).

The grid lines are drawn, and the origin $z = 0$, is marked with a dot. The vertical lines mark the position $z = 0$, $z = \frac{1}{2}$, $z = 1$, $z = \frac{3}{2}$, ...

On the opposite side, we may have something like the following. The vertical lines mark the position $t = 0$, $t = \frac{1}{2}$, $t = 1$, $t = \frac{3}{2}$, ...

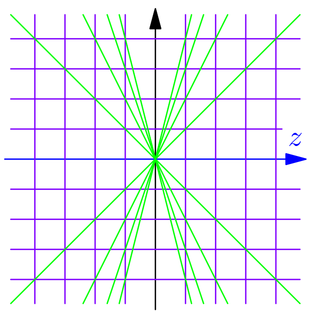


Question 52.3.4. Check that, on $\mathbb{C} \times U$ for any open set $U \subseteq \mathbb{C}_\infty$ that contains neither 0 nor ∞, the two line bundle charts above define the same line bundle structure. (What are the transition functions?)

So, we have the so-called trivial line bundle $\mathbb{C} \times \mathbb{C}_\infty$.

As promised, there are also nontrivial line bundles here.

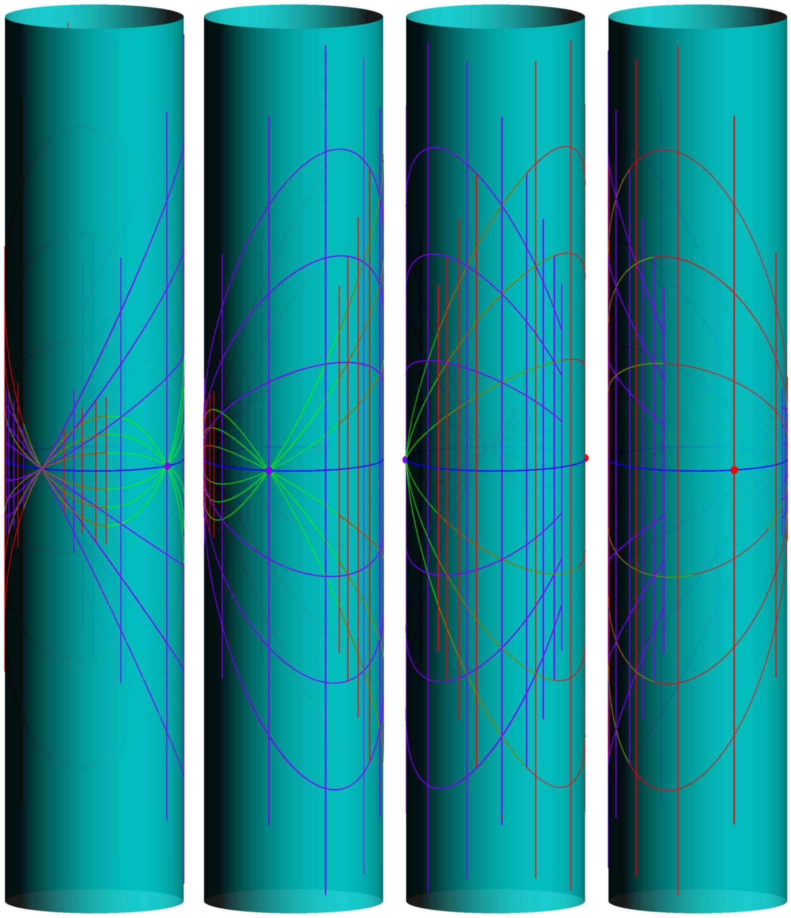
First, recall from the section above: over any open set U that contains neither 0 nor ∞, we can consider another line bundle chart that scales the vertical direction by a factor of z, this induces the same line bundle structure on U.



Question 52.3.5. Let ϕ_p and ϕ_g be the line bundle charts corresponding to the purple and green grid, respectively. Verify that if a point $q \in L$ satisfies $\phi_p(q) = (y, z)$ for $z \notin \{0, \infty\}$, then $\phi_g(q) = (\frac{y}{z}, z)$.

Now — note that the trivial line bundle $\mathbb{C} \times \mathbb{C}_\infty$ above can be seen as welding the two pieces together, such that the purple line (y, z) gets welded to the red line (y, t) for each y. There is nothing that restricts us to that specific welding method, however — this time around, we will try to weld the green line $(\frac{y}{z}, z)$ to the red line (y, t) for each y.

The thing will look like this. It looks quite complicated, so this time 4 views are shown.



The cylinder this time is only for illustrative purpose. Let us see what is going on.

- First, near the purple and the red point, the graph lines looks like our usual situation.

  Note that because $t = \frac{1}{z}$, looking from outside, the red coordinate lines will looks flipped.

- On the positive side ($z > 0$ and $t > 0$), no problem — we just need to squeeze the purple lines closer together — as depicted in the figure.

- On the negative side, however — note that the green line $(\frac{y}{z}, z)$ moves *downwards* when y increases, so we will need to "twist the graph paper" for it to go up.

  In the figure, this is depicted as a singularity where all the horizontal lines intersect, but in reality, you should think of it as we twisting the "graph paper" by 180 degrees and weld it to the other part.

This is a Möbius strip!

Thus, it appears to be obvious that this line bundle is not isomorphic to the trivial one, whatever "isomorphic" might mean.

> **Question 52.3.6.** Check that what we did above makes sense when y, z and t are not real — in particular, $\mathbb{C} \setminus \{0\}$ is a connected set, unlike $\mathbb{R} \setminus \{0\}$. You probably won't be able to visualize the "graph paper" this time (it is 4-dimensional!), so you will have to keep your intuition confined in the real part and use algebra for the rest.

§52.4 Morphisms between line bundles

In order to formally define what it means for two line bundles to be isomorphic, we need to be able to define morphisms. It is exactly what you expect — it must respect the line bundle structure (that is, the vector space structure and the analytic structure) on L_1 and L_2.

Definition 52.4.1. Let $\pi_1 \colon L_1 \to X$ and $\pi_2 \colon L_2 \to X$ be line bundles. A line bundle morphism $\alpha \colon L_1 \to L_2$ is a set morphism such that:

- $\pi_2 \circ \alpha = \pi_1$, and
- if $\phi_1 \colon \pi_1^{-1}(U_1) \to \mathbb{C} \times U_1$ and $\phi_2 \colon \pi_2^{-1}(U_2) \to \mathbb{C} \times U_2$ are line bundle charts, then the composition

$$\phi_2 \circ \alpha \circ \phi_1^{-1} \colon \mathbb{C} \times (U_1 \cap U_2) \to \mathbb{C} \times (U_1 \cap U_2)$$

  has the form $(s, p) \mapsto (f(p) \cdot s, p)$ where f is analytic on $U_1 \cap U_2$.

Exercise 52.4.2. The function $f(p)$ above must be nonzero for all $p \in U_1 \cap U_2$. Why? (Hint: invert the function by swapping the role of ϕ_1 and ϕ_2.)

Question 52.4.3. Check that the above definition is the equivalent to the following: α is a line bundle morphism if and only if

- it maps a point $q \in \pi_1^{-1}(x)$ to some point $\alpha(q) \in \pi_2^{-1}(x)$ (that is, each fiber gets mapped to the corresponding fiber), and
- for every analytic section $s \colon U \to L_1$ on open set $U \subseteq X$, then $\alpha(s)$ is an analytic section $s \colon U \to L_2$.

Example 52.4.4

Let $X = \mathbb{C}$. Then $\alpha \colon \mathbb{C} \times X \to \mathbb{C} \times X$ by $\alpha(y, x) = (y \cdot x^2, x)$ is a line bundle homomorphism.

This line bundle homomorphism is not an isomorphism, because every point $(y, 0)$ gets mapped to $(0, 0)$.

The definition of line bundle isomorphism is what you would expect.

Definition 52.4.5 (Isomorphism of line bundles)**.** Two line bundles L_1 and L_2 are **isomorphic** if there are line bundle isomorphisms $\alpha \colon L_1 \to L_2$ and $\beta \colon L_2 \to L_1$ that are inverse of each other.

§52.5 Relation to invertible sheaves

write (actually we haven't defined sheaf yet either)