Lecture Notes

Algebraic Geometry III/IV

Matt Kerr

Contents

Part 1

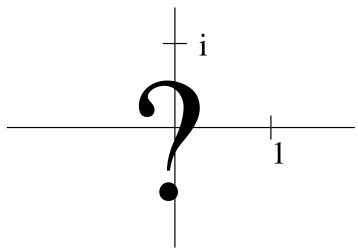
Introduction and Motivation

CHAPTER 1

Two theorems on conics in the plane

The primary subject of this course is *algebraic curves*, the simplest example of which is *the solution set of a two-variable polynomial equation in the plane*. *What* plane? $\mathbb{R}^2$, $\mathbb{C}^2$, $\mathbb{F}_p^2$, $\mathbb{Q}^2$...? For the purposes of this course, mainly $\mathbb{C}^2$, with excursions into the others.

REMARK 1.0.1. A perennial point of confusion is whether to call $\mathbb{C}$ the complex plane.



This is henceforth forbidden! It is the *complex (affine) line*, and a *real (affine) plane* via $\mathbb{C} \cong \mathbb{R}^2$. (A "complex plane" will mean something 2-dimensional over $\mathbb{C}$, so $\mathbb{C}^2$ will be the complex affine plane, $\mathbb{P}^2$ the complex projective plane, and so forth. We'll worry about affine vs. projective in the next chapter.) Note that $\mathfrak{H} := \{x + iy \,|\, x, y \in \mathbb{R},\, y > 0\} \subset \mathbb{C}$ will denote the "upper-half plane" in $\mathbb{C}$; that terminology is unavoidable.

The objects which shall concern us, then, will be 1-dimensional over $\mathbb{C}$ ("complex algebraic curves"), hence 2-dimensional over $\mathbb{R}$ ("Riemann surfaces"). Our approach will be quite intuitive and visual for the first few chapters, to get an idea of what algebraic geometry is before settling into a more measured approach. My feeling has always been that you need motivation for introducing formalism, in this case for layering lots of algebra onto geometry. In this chapter that motivation might consist of the subtle gaps that open as we try to prove some famous results on conics from (mostly) linear algebra.

EXAMPLE 1.0.2. (a) Consider the set of rational points on the Fermat quartic (degree 4) curve:

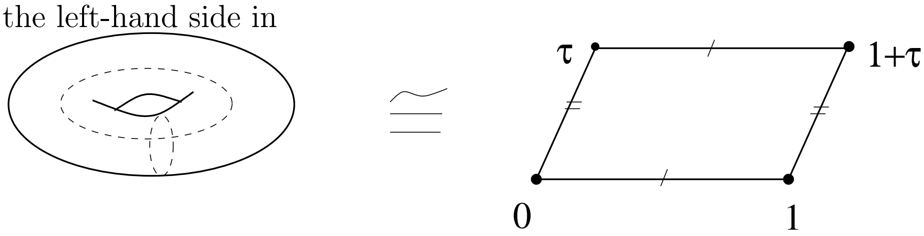
$$\{(x,y) \in \mathbb{Q}^2 \,|\, x^4 + y^4 = 1\} \subset \mathbb{Q}^2. \tag{1.0.1}$$

By a case of Fermat's last theorem, (1.0.1) is the empty set, since $\left(\frac{a}{b}\right)^4 + \left(\frac{c}{d}\right)^4 = 1$ means $(ad)^4 + (bc)^4 = (bd)^4$ with $ad, bc, bd \in \mathbb{Z}$ and $bd \neq 0$.

(b) Next we look at the Fermat cubic (degree 3) curve

$$\{(x,y) \in \mathbb{C}^2 \,|\, x^3 + y^3 = 1\} \subset \mathbb{C}^2. \tag{1.0.2}$$

We will see that (1.0.2) has the structure of a complex 1-torus, shown on the left-hand side in

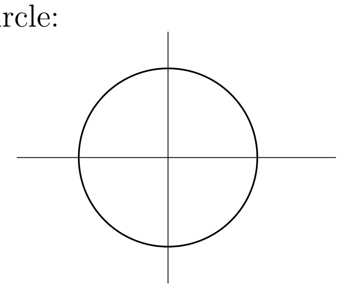


The right-hand side represents the quotient $\mathbb{C}/\Lambda$ of the complex line by the lattice $\Lambda := \mathbb{Z}\langle 1, \tau\rangle$ (for some $\tau \in \mathfrak{H}$). What is the isomorphism? It holds topologically (convince yourself of this visually) for any $\tau \in \mathfrak{H}$, but complex analytically only for the values $\tau = \frac{a\tau_0+b}{c\tau_0+d}$ where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ and $\tau_0 := \frac{1+\sqrt{-3}}{2}$. We call (1.0.2) an elliptic curve: it isn't an ellipse in any sense, but originally arose in connection with the arc-length of one.

(c) Finally, take the real degree 2 (a.k.a. "quadric" or "conic") curve

$$\{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\} \subset \mathbb{R}^2, \tag{1.0.3}$$

which is of course a circle:



First we shall pursue conics to get a preliminary feel for the interplay between algebra and geometry. These aren't too hard to visualize: you know what the real solution sets look like (ellipses, hyperbolas, pairs of lines, etc.) and the complex solutions are all spheres once we add "points at infinity".

1.1. Algebraic curves in $\mathbb{R}^2$

Let $\mathcal{P}_2^n$ denote the real polynomials of degree $\leq n$ in x and y. (If there is no possibility for confusion, I'll just write $\mathcal{P}^n$.) In an exercise below, you are asked to prove that $\mathcal{P}_2^n$ is a real vector space of dimension $\binom{n+2}{2}$.

EXAMPLE 1.1.1. (a) A basis for $\mathcal{P}^2$ is given by $\{1, x, y, xy, x^2, y^2\}$, so $\dim(\mathcal{P}_2) = 6$.

(b) For $\mathcal{P}^3$, the basis is $\{1, x, y, xy, x^2, y^2, x^3, x^2y, xy^2, y^3\}$ (and the dimension is 10).

For a configuration of distinct points

$$\mathcal{S} = \{p_1, \ldots, p_m\} \subset \mathbb{R}^2,$$

define a linear "**ev**aluation" map

$$ev_{\mathcal{S}}^n : \mathcal{P}^n \longrightarrow \mathbb{R}^m$$

via

$$f \longmapsto \begin{pmatrix} f(p_1) \\ \vdots \\ f(p_m) \end{pmatrix}.$$

Recall from linear algebra the **Rank** + **Nullity** theorem:

$$\underbrace{\dim(\mathrm{image}(ev_{\mathcal{S}}))}_{rank} + \underbrace{\dim(\ker(ev_{\mathcal{S}}))}_{nullity} = \dim(\mathcal{P}^n). \tag{1.1.1}$$

DEFINITION 1.1.2. The configuration of points $\mathcal{S}$ is called *n-general*[1] $\Longleftrightarrow$ $ev_{\mathcal{S}}^n$ is surjective (onto).

Now if $ev_{\mathcal{S}}^n$ is surjective, its rank is m; while its kernel is just the space of polynomials vanishing on $\mathcal{S}$. By (1.1.1) we have:

PROPOSITION 1.1.3. *The space of degree $\leq n$ polynomials vanishing on a general configuration of m points has dimension $\binom{n+2}{2} - m$.*

[1] or just "general" or "generic" if the context is understood

What does all this have to do with algebraic curves? Well, given a polynomial in $\mathcal{P}_n$, I can look at its solution set in $\mathbb{R}^2$. More precisely, we have the assignment[2]

$$\mathcal{P}^n \backslash \mathcal{P}^{n-1} \rightsquigarrow \{\text{degree } n \text{ real affine algebraic curves}\}$$

given by

$$f \longmapsto C_f := \{(x,y) \in \mathbb{R}^2 \,|\, f(x,y) = 0\}.$$

Notice that

$$f \in \ker(ev_{\mathcal{S}}) \iff C_f \supset \mathcal{S}. \tag{1.1.2}$$

PROPOSITION 1.1.4. *Through five (2-)general points A, B, C, D, E in the real plane, there exists a unique conic Q.*

PROOF. $6 - 5 = 1$. □

OK, OK, so what the proof means is: by Prop. 1.5, the space of degree ≤ 2 polynomials vanishing on a general configuration of 5 points has dimension $\binom{2+2}{2} - 5{=}1$. So given two degree-2 polynomials $f, g \in \mathcal{P}^2$ vanishing at all 5 points in $\mathcal{S}$, we must have $g = a \cdot f$ (for some $a \in \mathbb{R}$). But $f = 0$ and $a \cdot f = 0$ define the same curve Q; and by (1.1.2) Q contains A, B, C, D, E.

There are two issues with this. First, in order for f to define a conic, we need to know that the 1-dimensional space of solutions doesn't lie in $\mathcal{P}^1$ (inside $\mathcal{P}^2$). But if a linear (degree 1) polynomial h vanishes at all 5 points, so does h^2; and then since $h, h^2 \in \mathcal{P}^2$ are not linearly dependent, $ev_{\mathcal{S}}^2$ doesn't have maximal rank, contradicting the genericity condition on $\mathcal{S}$. (At this point, Proposition 1.6 is completely proved as stated.)

The second problem is more interesting: what exactly does it mean for A, B, C, D, E to be 2-generic? If you think about our little proof, the existence of Q has nothing to do with this genericity, but the uniqueness has everything to do with it. Moreover, the statement that "Q is unique if $\mathcal{S}$ is generic" is somewhat circular without knowing (beyond Definition 5) what the genericity condition is: namely, that *no* 4 *of the points are collinear.* It's best to wait until we're doing projective geometry to prove that, and so we will.

[2]notation: $\mathcal{A} \backslash \mathcal{B}$ denotes set-theoretic exclusion, sometimes also written $\mathcal{A} - \mathcal{B}$. "Affine" just means that the curves are in $\mathbb{R}^2$.

REMARK 1.1.5. More generally, the abstract ethos surrounding the word "general" in algebraic geometry is that you are working in the complement of finitely many algebraic conditions. The "algebraic condition" we are avoiding in this section is the vanishing of all $m \times m$ minors in a matrix representing ev^n_S.

1.2. Blaise Pascal and the mystic hexagon

Similarly, given eight points

$$A,\ B,\ C,\ D,\ E,\ P_1,\ P_2,\ P_3$$

in (3-)general position, the vector space of cubic polynomials vanishing at all of them has dimension

$$\binom{3+2}{2} - 8 = 2.$$

Call this vector space V, and write $(x_A y_A), \ldots, (x_{P_3}, y_{P_3})$ for the coordinates of given points. V lies in $\mathcal{P}^3$, which consists of elements of the form

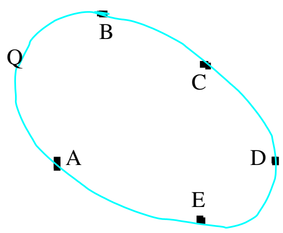
$$a_1 + a_2 x + a_3 y + \cdots + a_9 x y^2 + a_{10} y^3 = f_{\underline{a}}(x, y).$$

Asking a polynomial of this form to vanish at our 8 points yields the 8 equations

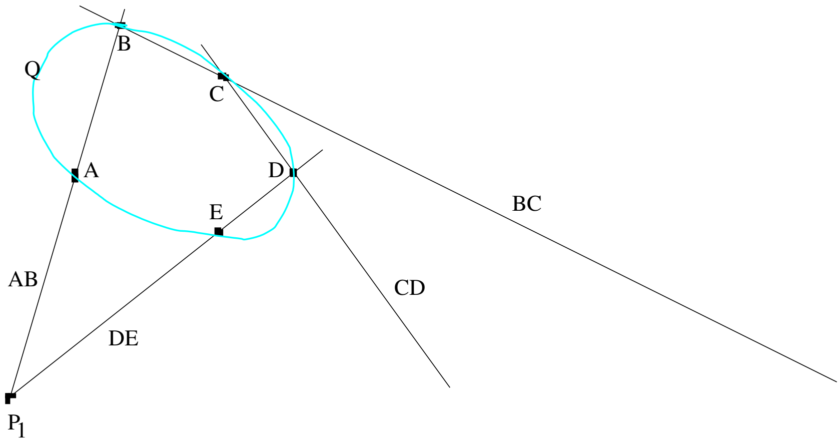
$$\begin{cases} 0 &= \quad a_1 + a_2 x_A + a_3 y_A + \cdots + a_9 x_A y_A^2 + a_{10} y_A^3 \\ \quad \vdots & \\ 0 &= \quad a_1 + a_2 x_{P_3} + a_3 y_{P_3} + \cdots + a_9 x_{P_3} y_{P_3}^2 + a_{10} y_{P_3}^3 \end{cases},$$

i.e. 8 linear constraints on the 10 variables $\{a_j\}$ expressing what it means for $f_{\underline{a}}$ to lie in V.

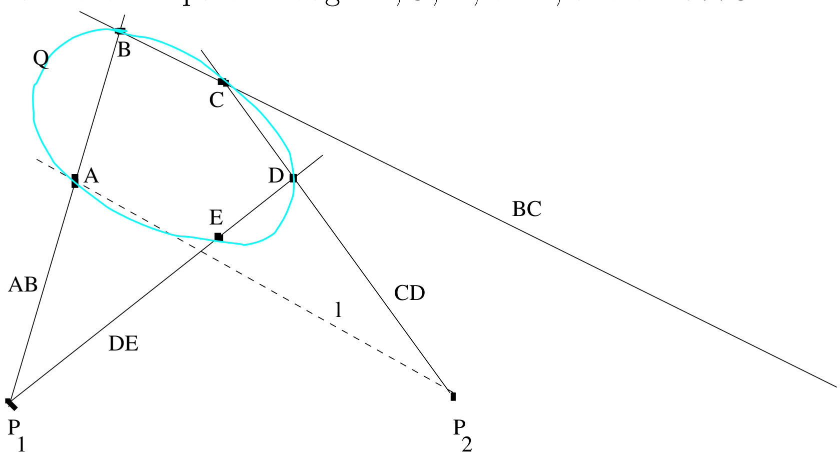
Now let A, B, C, D, E be (2-)general and Q the unique conic through them:



In fact, we shall make the stronger assumption that no three of A, B, C, D, E are collinear, so that Q is smooth (an ellipse or hyperbola, not a pair of lines). We would like to construct (arbitrarily many) points on Q using only a straightedge. Start by drawing secant lines connecting adjacent points:

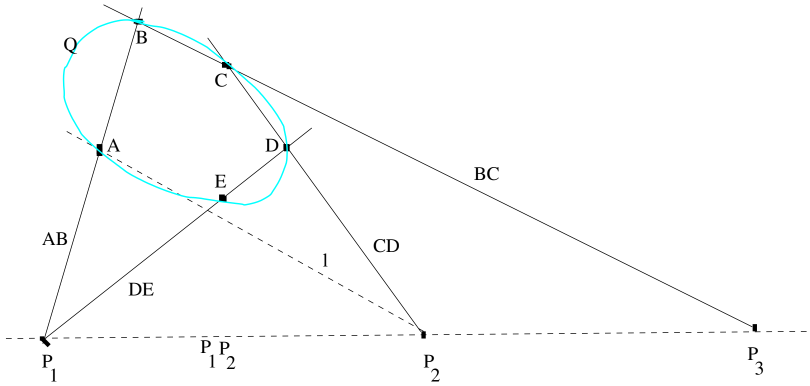


where we have labelled $AB \cap DE =: P_1$.[3] Next, draw any line ℓ through A that does not pass through B, C, D, or E, and set $\ell \cap CD =: P_2$.

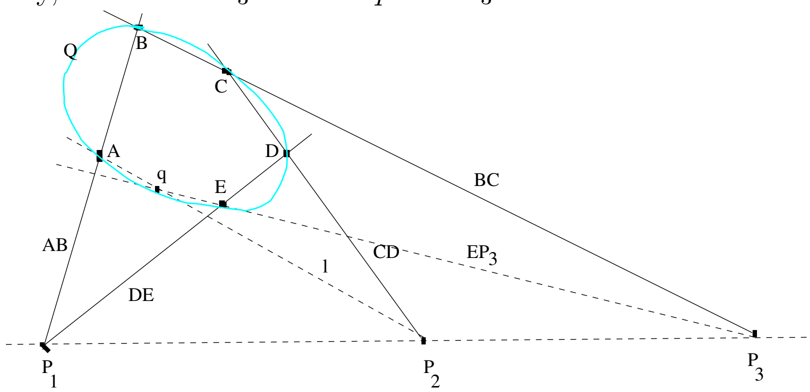


(Note that our choice of ℓ will determine the point $q \in Q$ we end up constructing.) Next, draw P_1P_2, and label $P_1P_2 \cap BC =: P_3$.

[3]Of course, the lines may be parallel. There are two ways to fix this: either make A, B, C, D, E "more" general so that none of the lines we intersect are parallel; or work projectively. Since this chapter is entirely motivational we won't worry about that level of detail..



Finally, we draw EP_3 and set $q := EP_3 \cap \ell$.



PROPOSITION 1.2.1. $q \in Q$.

To prove this, we shall require:

LEMMA 1.2.2. *$A, B, C, D, E, P_1, P_2, P_3$ is in (3-)general position, when A, B, C, D, E are general in the strong sense assumed above.*

We won't prove the lemma here. (It is a special case of something called the Cayley-Bacharach Theorem, which will be easy to prove once we know a little about residues in algebraic geometry.)

PROOF. [OF PROP. 1.8] We write $f_{AB}(x, y) = 0$ for the (linear) equation of the line AB, $f_Q(x, y) = 0$ for the conic Q, and so forth. Consider the three cubic polynomials

$$f_1(x, y) := f_Q(x, y) \cdot f_{P_1P_2}(x, y),$$

$$f_2(x, y) := f_{AB}(x, y) \cdot f_{CD}(x, y) \cdot f_{EP_3}(x, y),$$

$$f_3(x, y) := f_\ell(x, y) \cdot f_{BC}(x, y) \cdot f_{DE}(x, y).$$

Their vanishing sets are

$$C_1 := Q \cup P_1P_2,$$

$$C_2 := AB \cup CD \cup EP_3,$$

$$C_3 := \ell \cup BC \cup DE,$$

each of which contains the set

$$\mathcal{S} := \{A, B, C, D, E, P_1, P_2, P_3\}.$$

So by (1.1.2), f_1, f_2, f_3 belong to $V := \ker(ev^3_{\mathcal{S}})$. Since the dimension of V is 2, f_1, f_2, f_3 cannot be linearly independent and we have a nontrivial[4] relation

$$\alpha f_1 + \beta f_2 + \gamma f_3 = 0 \tag{1.2.1}$$

with real coefficients.

Suppose $\alpha = 0$ in (1.2.1). Then $\beta f_2 = -\gamma f_3$, so that f_2 and f_3 are proportional hence cut out the same curve:

$$AB \cup CD \cup EP_3 = \ell \cup BC \cup DE.$$

But this means $\ell = AB$, CD or EP_3, which implies ℓ contains A, B, C, D, or E — a contradiction to our choice of ℓ!

So $\alpha \neq 0$, and we may rewrite (1.2.1) as

$$f_1 = -\frac{\beta}{\alpha} f_2 - \frac{\gamma}{\alpha} f_3. \tag{1.2.2}$$

Now since ℓ and EP_3 both contain q, f_2 and f_3 both vanish at q. By (1.2.2), f_1 also is zero at q, so one of its factors has to be. Therefore q is contained in Q or P_1P_2.

Suppose $q \in P_1P_2$. Then the lines P_1P_2, EP_3, and ℓ "collapse" to the same line (look at the last diagram) and so in particular $E \in \ell$, again in contradiction to our choise of ℓ.

We conclude that $q \in Q$. □

In the construction described pictorially above, q was — in light of Proposition 1.9 — ultimately just the point where ℓ meets Q. Since our choice of ℓ was essentially free, A, B, C, D, E, and q can be thought of as 6 distinct but otherwise arbitrary points of Q. Consequently, we have proved the beautiful statement:

[4]not all of α, β, γ are zero

THEOREM 1.2.3. [B. PASCAL, 1639] *Intercepts of opposite edges of a hexagon inscribed in a conic, lie on a line.*

When we get to the notion of duality in projective geometry, Theorem 1.10 will "dualize" for free to:

COROLLARY 1.2.4. *The (three) lines joining opposite vertices of a hexagon circumscribed about a conic, meet in a single point.*

1.3. Poncelet's Porism

According to my dictionary, a *porism* is a "proposition that uncovers the possibility of finding such conditions as to make a specific problem capable of innumerable solutions".[5] The result of Poncelet I'll describe here just had a whole book devoted to it,[6] and in the late 1970's P. Griffiths (my Ph.D. advisor) and J. Harris devoted two nice articles to it. If your local pub had an ellipse-shaped pool table it would even have practical applications.

So let C and D be two conics in $\mathbb{R}^2$. They are the vanishing sets of some $f_C, f_D \in \mathcal{P}_2$. For simplicity, assume they are ellipses, with D contained in the interior of C.

Problem: *Does there exist a closed polygon (self-intersections are OK) inscribed in C and circumscribed about D?*

Solution: *Sometimes. But existence of one such polygon implies that there are infinitely many.*

This is clearly a "porism". We shall call such polygons as in the "Problem" *circuminscribed* when a specific pair C,D is understood. The precise statement is:

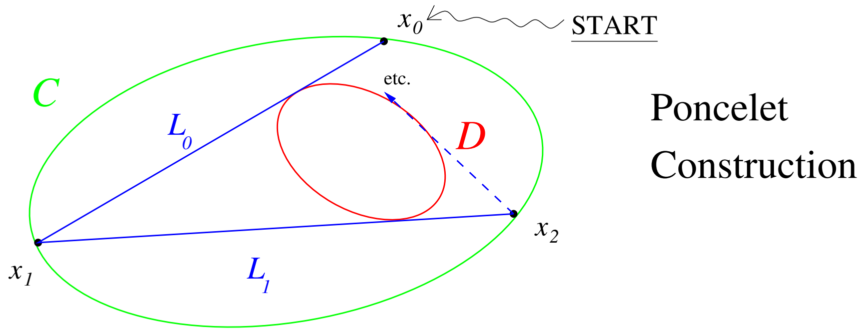
THEOREM 1.3.1. [PONCELET, 1822] *If the pair C, D has an n-sided circuminscribed polygon, then for any point on C there is a circuminscribed n-sided polygon with one vertex on C.*

Another way of putting this is that circuminscribed polygons can be rotated continuously: there are beautiful pictures of this at http://

[5]rather than a complication of the swine flu

[6]L. Flatto, "Poncelet's Theorem", AMS, 2009. It's on reserve in the library and aimed at advanced undergraduates, i.e. you. Highly recommended.

enriques.mathematik.uni-mainz.de/intgeo/poncelet.html. While this picture



is to those ones as the math building (at virtually any university) is to Durham cathedral, it allows us to characterize Theorem 1.12 in one more way:

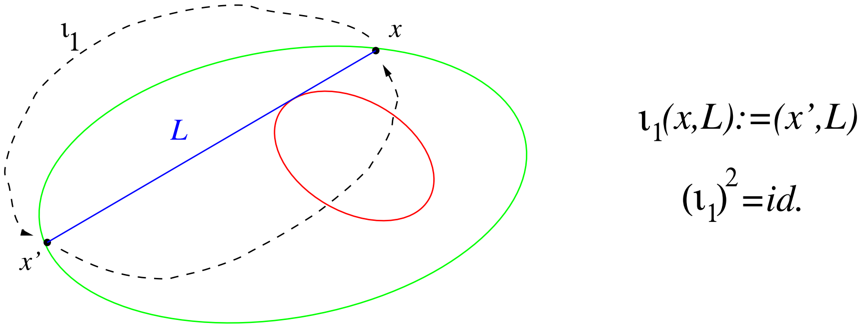
(1.3.1) If $x_n = x_0$ for *any* x_0, then $x_n = x_0$ for *every* x_0.

We will skirt around projective geometry in explaining the idea here, but can't avoid $\mathbb{C}$. Henceforth, C and D shall denote all complex solutions to $f_C = 0$ and $f_D = 0$ — that is, $C, D \subset \mathbb{C}^2$. Topologically these are "real surfaces" (in fact, spheres with one or two missing points), and are complex-analytically isomorphic to $\mathbb{C}$ or $\mathbb{C}^*$, but it won't hurt to draw them schematically as real curves on a sheet of paper. You can think of this as the real solutions standing in for the complex ones. In general, when we want to see the topology of a complex algebraic curve, we'll draw a "surface"; when we want to see how different curves intersect or how they lie in space, we'll draw a "curve".

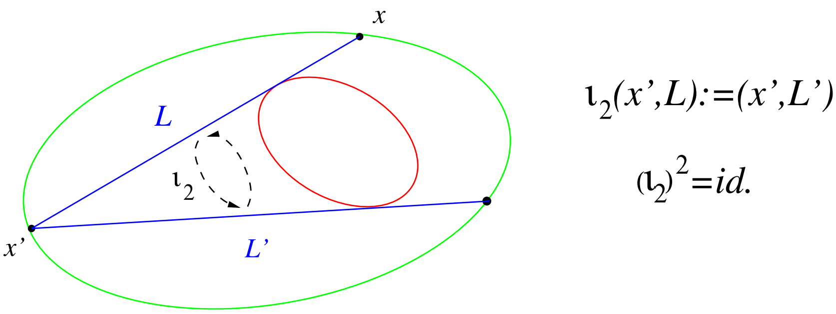
Consider the set of pairs

$$\mathcal{E} := \{(x, L) \in C \times D^* \mid x \in L\},$$

where D^* is the set of lines tangent to D. An *involution* is a map which, applied twice, gives the identity. Here are two involutions on $\mathcal{E}$:



and



The composition

$$\jmath := \iota_2 \circ \iota_1$$

takes

$$(x, L) \xmapsto{\iota_1} (x', L) \xmapsto{\iota_2} (x', L') =: (x_1, L_1)$$

(ι_1, ι_2 and $\jmath$ are all maps from $\mathcal{E}$ to $\mathcal{E}$). More generally,

$$\jmath(x_i, L_i) =: (x_{i+1}, L_{i+1})$$

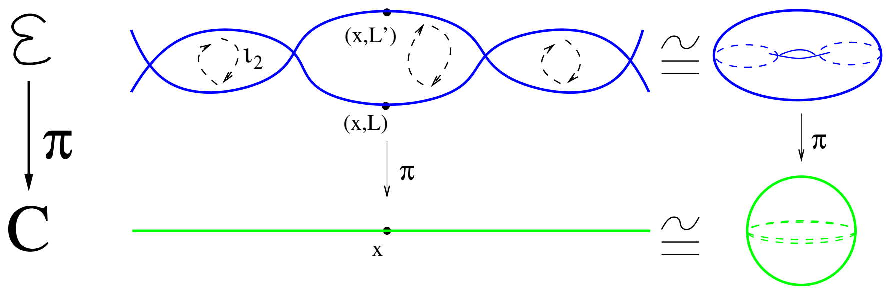
defines the i^{th} iteration of the Poncelet construction. The construction starting from some (x, L) closes if and only if

$$\jmath^n(x, L) = (x, L) \text{ for some } n. \tag{1.3.2}$$

Thinking of complex points, and assuming C and D aren't tangent anywhere and don't meet "at infinity" — that is, C and D are in general position in some sense — they meet in exactly four points. This is a first taste of Bezout's theorem, which we will prove carefully later. Now let $x \in C$: if $x \notin C \cap D$, there are exactly two lines containing x and tangent to D; if $x \in C \cap D$, there is only one. So we find that the projection

$$\begin{array}{rccc} \pi : & \mathcal{E} & \longrightarrow & C \\ & (x, L) & \longmapsto & x \end{array}$$

is a *two-sheeted branched covering* with four branch points:



This turns out to mean that $\mathcal{E}$ is an elliptic curve, hence isomorphic to $\mathbb{C}$/lattice. One deduces that $\jmath : \mathcal{E} \to \mathcal{E}$ is a translation

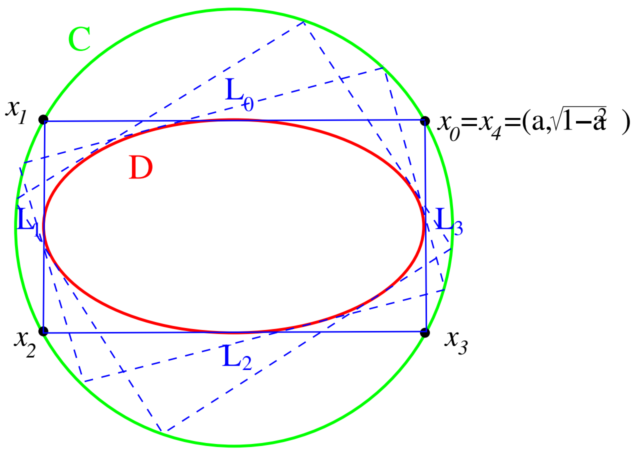
$$\mathbb{C}/\Lambda \to \mathbb{C}/\Lambda$$

$$u \longmapsto u + \xi,$$

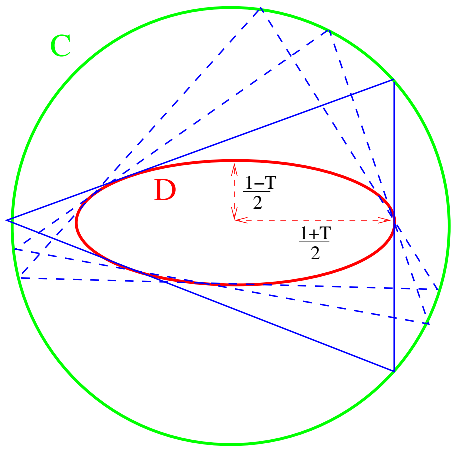
and our starting point (x, L) is some point u_0. Whether or not $\jmath^n(u_0) \equiv u_0$ has everything to do with whether $n\xi \in \Lambda$, and nothing to do with the choice of u_0, and so (1.3.1) follows.

We will see a more in-depth treatment of this after studying elliptic curves, including an algebraic criterion for deciding when $\jmath^n = \text{id.}$ for a given C, D, and an analysis of elliptical billiards. For now, here are some examples of "Poncelet in action":

Example 1.3.2. (a) $C = \{x^2 + y^2 = 1\}$, $D = \{\frac{x^2}{a^2} + \frac{y^2}{1-a^2} = 1\}$, $n = 4$.



(b) $C = \{x^2 + y^2 = 1\}$, $D = \{\frac{4x^2}{(1+T)^2} + \frac{4y^2}{(1-T)^2} = 1\}$, $n = 3$.



Exercises

(1) $\mathcal{P}_2^n$ is a vector space of dimension $\binom{n+2}{2}$. [For a more challenging problem, try it with polynomials in m variables instead of two — obviously with a different answer.]

(2) Consider the two conics $C = \{x^2 + y^2 = 1\}$ and $D = \{x^2 + y^2 = r^2\}$. The corresponding "Poncelet elliptic curve" $\mathcal{E}$ is singular (see Definition 2.9), which means that the problem "degenerates" and is solvable by hand using secondary school maths. For which r (say, between 0 and 1) does the n^{th} iterate of the Poncelet construction (starting at an arbitrary point on C) return to the starting point? (Hint: for each n, there are finitely many; just find how many, and the equation they must satisfy.)

CHAPTER 2

Riemann surfaces and algebraic curves

In this chapter we will define (complex) algebraic curves (represented by "C"),[1] complex 1-manifolds (represented by "M"), and Riemann surfaces, and start to consider under what additional hypotheses they are equivalent concepts.

2.1. Algebraic curves

DEFINITION 2.1.1. Let S_2^m denote *homogeneous* polynomials of degree m in x, y (the "2" stands for "2 variables").[2] These are polynomials of the form

$$f_m(x,y) = \sum_{\substack{j,k\geq 0\\ j+k=m}} c_{jk}x^j y^k,$$

that is, each term has total degree m. Clearly S_2^m is a subset of $\mathcal{P}_2^m$.

More generally, S_k^m is the space of degree-m homogeneous polynomials in k variables — that is, linear combinations

$$\sum_{\substack{i_1,\ldots,i_k\geq 0\\ i_1+\cdots+i_k=m}} c_{i_1,\ldots,i_k}x_1^{i_1}\cdots x_k^{i_k}$$

of monomials with total degree m. Elements $f \in S_k^m$ have the property that $f_m(\alpha x_1, \alpha x_2, \ldots, \alpha x_k) = \alpha^m f_m(x_1, x_2, \ldots, x_k)$. (See exercise 3 below.)

Given a real affine algebraic curve of degree d

$$C := \{0 = f(x,y) = f_d(x,y) + f_{d-1}(x,y) + \cdots + f_0\} \subset \mathbb{R}^2$$

[1]in the affine resp. projective plane. Later we will define algebraic curves in higher dimensional projective spaces and products thereof, but the "most intrinsic" definition of an algebraic curve as a 1-dimensional reduced scheme (some authors require this to be irreducible and over an algebraically closed field as well) is probably something to learn only once you have a first course in algebraic geometry under your belt.

[2]the field of definition, from which the c_{jk} are taken, will depend on context; S_2^m is a vector space over that field.

with f_d not identically zero, we would like to count its *intersections* with a real line given parametrically by

$$L:\, t \longmapsto (\alpha t, \beta t)$$

(where α, β are real constants). These are just the solutions of

$$0 = f_d(\alpha,\beta)t^d + f_{d-1}(\alpha,\beta)t^{d-1} + \cdots f_0. \tag{2.1.1}$$

Naively, we would like to get d points:

$$\#\{C \cap L\} = d\ ?? \tag{2.1.2}$$

Some issues arise . . .

	Problem	Solution
(a)	$\mathbb{R}$ is not algebraically closed!	pass to $\mathbb{C}^2$
(b)	solutions "at infinity"!	add a "line at infinity" to $\mathbb{C}^2$
(c)	multiple roots!	count intersections with multiplicity
(d)	C might contain L!	Uh-oh

Each "Problem" is an obstruction to (2.1.2), and the object of each "Solution" is to remove the obstruction.

In a little more depth, (a) says that in spite of the fact that the $f_i(\alpha,\beta) \in \mathbb{R}$, roots of (2.1.1) can be non-real. So we had better consider C and L as complex algebraic curves — take $x, y \in \mathbb{C}$ in the definition of C and $t \in \mathbb{C}$ in the definition of L. Their dimensions over $\mathbb{R}$ then, of course, double, and $C \cap L$ now contains the points corresponding to non-real roots of (2.1.1).

Next, if we plug $t = s^{-1}$ into (2.1.1) and multiply by s^d, then it becomes

$$0 = f_0 s^d + f_1(\alpha,\beta)s^{d-1} + \cdots + f_d(\alpha,\beta). \tag{2.1.3}$$

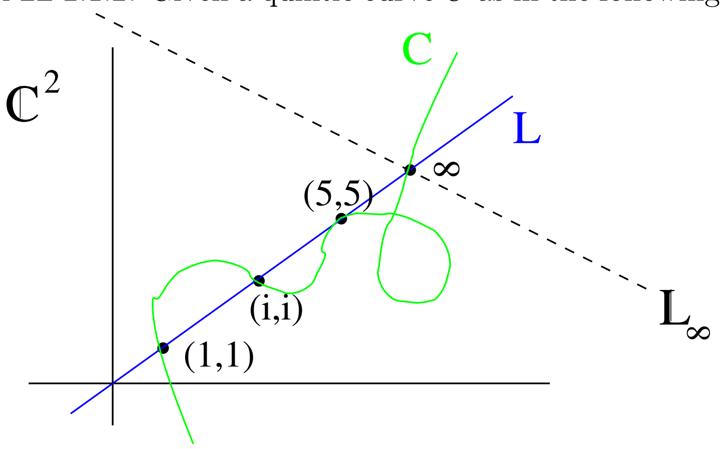
A "solution at ∞" to (2.1.1) is a solution at 0 to (2.1.3), which exists if and only if $f_d(\alpha,\beta) = 0$. For this to be counted in $C \cap L$, we must add a line $\mathbb{L}_\infty$ to $\mathbb{C}^2$ to get $\mathbb{P}^2$,[3] the complex projective plane (which we shall discuss in a moment). This adds a point to L "at ∞" corresponding to $s = 0$, yielding a $\mathbb{P}^1$ (projective line), and adds *points* at infinity to C (yielding a compact "projective" curve).

[3] called $\mathbb{CP}^2$ or $\mathbb{P}^2_{\mathbb{C}}$ in some books

We know that to get d solutions out of a degree d polynomial equation you have to count a twice repeated root as two solutions. So to get d intersection points you will certainly have to count intersections of C with L however many times the corresponding root of (2.1.1) is repeated. The curious case is an m-times repeated root at infinity: via (2.1.3), this corresponds to $f_d(\alpha, \beta) = \cdots = f_{d-m+1}(\alpha, \beta) = 0$. In that case, (2.1.1) is only in fact of degree $d - m$. One *does* have to worry about these degenerate cases, but they will look completely natural in projective coordinates.

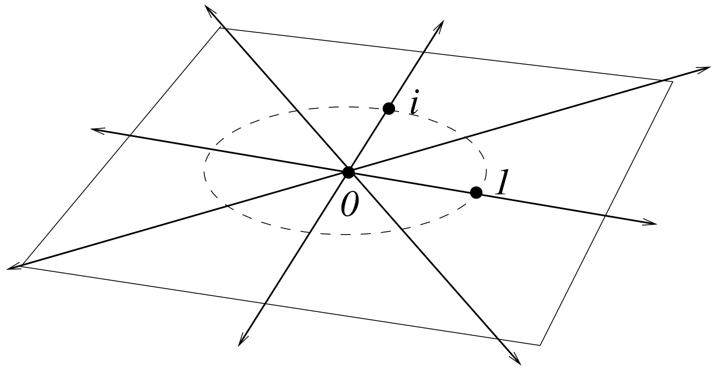
Finally, if all $f_\ell(\alpha, \beta) = 0$, then $C \supset L$ and we are in trouble — there is no way around (d). We shall have to demand that plane curves intersect "properly" (in points only), disallowing this possibility, in order to make any statement about the number of intersection points.

EXAMPLE 2.1.2. Given a quintic curve C as in the following picture

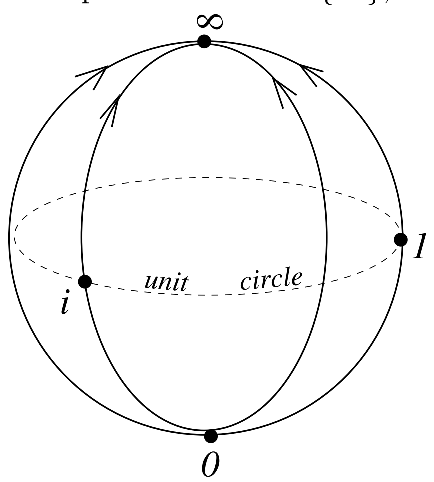


the number of intersection points in $\mathbb{R}^2$ is only 2. But the number of complex intersection points, counting multiplicities and intersections at infinity, is 5.

So . . . how does one go about adding a line (resp. point) at infinity to $\mathbb{C}^2$ (resp. L)? First, visualize L as $\mathbb{C}$



and think of all arrows as going off to the same point. Adding this point gives the "1-point" compactification $\mathbb{C} \cup \{\infty\}$, resulting in a sphere



This is an informal way of thinking of $\mathbb{P}^1$ in the following:

DEFINITION 2.1.3. Projective space $\mathbb{P}^n$ is the set of complex lines through the origin $\underline{0} \in \mathbb{C}^{n+1}$.[4] More precisely,

$$\mathbb{P}^n := \frac{(\mathbb{C}^{n+1} \backslash \{\underline{0}\})}{\left\langle \substack{(z_0, z_1, \ldots, z_n) \sim (\alpha z_0, \alpha z_1, \ldots, \alpha z_n) \\ \forall \alpha \in \mathbb{C}^*} \right\rangle}$$

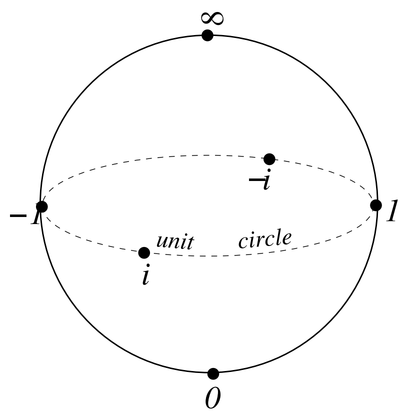
consists of nonzero vectors in $\mathbb{C}^{n+1}$, modulo the equivalence relation equating all vectors lying on a complex line. Elements are written $[z_0 : z_1 : \cdots : z_n]$.

For $n = 1$ this yields the projective line $\mathbb{P}^1$, which has the isomorphism

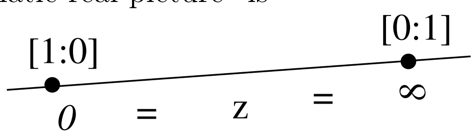
$$\begin{array}{ccc} \mathbb{P}^1 & \xrightarrow{\cong} & \mathbb{C} \cup \{\infty\} \\ [z_0 : z_1] & \longmapsto & \frac{z_1}{z_0} =: z \end{array}$$

[4]as will be proved in Chapter 5, one should really think of $\mathbb{P}^n$ as a complex manifold (but we haven't defined these yet). $\underline{0}$ denotes $(0, 0, \ldots, 0)$.

given by taking slope (of the line represented by $[z_0 : z_1]$). This map is well-defined as $[\alpha z_0 : \alpha z_1] \mapsto \frac{\alpha z_1}{\alpha z_0} = \frac{z_1}{z_0}$. The "honest topological picture" of $\mathbb{P}^1$ is



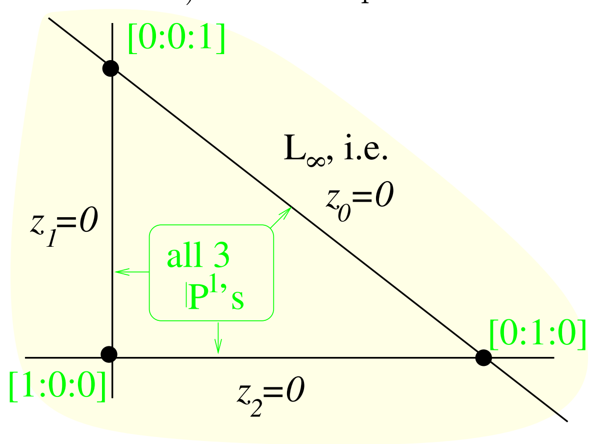
while the "schematic real picture" is



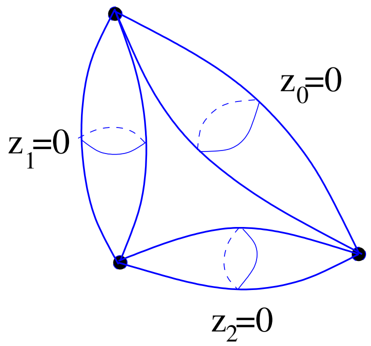
Next, setting $n = 2$ we have the projective plane $\mathbb{P}^2$, and the isomorphism

$$
\begin{array}{cccc}
 & \mathbb{P}^2 & \xrightarrow{\cong} & (\mathbb{C} \times \mathbb{C}) \cup \mathbb{P}^1 \\
(2.1.4) & [z_0 : z_1 : z_2] & \overset{\text{if } z_0 \neq 0}{\longmapsto} & \left(\frac{z_1}{z_0}, \frac{z_2}{z_0}\right) \in \mathbb{C}^2 \\
 & & \overset{\text{if } z_0 = 0}{\longmapsto} & [z_1 : z_2] \in \mathbb{P}^1
\end{array}
$$

expresses how $\mathbb{P}^2$ adds a line at infinity (the $\mathbb{P}^1$) to $\mathbb{C}^2$. For $\mathbb{P}^2$, the (rather bad, but standard) "schematic picture" is



While I'm not going to try to represent 4 real dimensions on paper, here is a mostly honest topological depiction of the 3 $\mathbb{P}^1$'s:



Remark 2.1.4. Equation (2.1.4) relates affine coordinates (on $\mathbb{C}^2$) and projective coordinates (on $\mathbb{P}^2$). Instead of the $\{z_i\}$, I will frequently use $[Z : X : Y]$ for a point in $\mathbb{P}^2$ and (asuming $Z \neq 0$) $\left(\frac{X}{Z}, \frac{Y}{Z}\right) =: (x, y)$ for the corresponding point in $\mathbb{C}^2$.

Also, a **warning** is in order: $[0 : 0 : 0]$ is not a point in $\mathbb{P}^2$. With homogeneous coordinates, some entry must be nonzero.

Returning to our degree d (and now complex) algebraic curve $C = \{f(x, y) = 0\} \in \mathbb{C}^2$, what happens to it as we compactify $\mathbb{C}^2$ to $\mathbb{P}^2$ as described above? To treat this, we first need to introduce the main object of study of this course.

Since $[Z : X : Y] = [\alpha Z : \alpha X : \alpha Y]$, in order for a polynomial equation $F(Z, X, Y) = 0$ to make sense projectively (i.e. in $\mathbb{P}^2$), we must have

$$F(Z, X, Y) = 0 \implies F(\alpha Z, \alpha X, \alpha Y) = 0 \quad (\forall \alpha \in \mathbb{C}^*). \tag{2.1.5}$$

This condition is guaranteed by homogeneity of F (cf. the property in Definition 2.1). (In fact, as we shall see later it is *equivalent* to homogeneity of F.)

Definition 2.1.5. A *projective algebraic curve* $C \subset \mathbb{P}^2$ of degree d, is the zero set of a homogeneous polynomial $F \in S_3^d$.

Here, then, is a general procedure for going between affine and projective curves:

$$f(x, y) = 0 \longmapsto Z^d f\left(\frac{X}{Z}, \frac{Y}{Z}\right) = 0 \tag{2.1.6}$$

corresponds to taking the *projective closure* $\bar{C} \subset \mathbb{P}^2$ of a given affine curve $C \subset \mathbb{C}^2$. Conversely, if the given C is already projective (defined by $F = 0$), then

$$F(Z, X, Y) = 0 \longmapsto F(1, x, y) = 0 \tag{2.1.7}$$

"restricts" C to the affine curve $C \cap \mathbb{C}^2$. Given an affine curve, taking closure then restricting gets you back to where you started.

EXAMPLE 2.1.6. Starting from the homogeneous cubic polynomial $F(Z, X, Y) = ZXY + 3Z^2Y + 4Y^3$, the affinization is $f(x, y) = F(1, x, y) = xy + 3y + 4y^3$. Conversely, if we start from $f(x, y) = x^3y - y^2 + 2x$, the projectivization is $F(Z, X, Y) = Z^4 f(\frac{X}{Z}, \frac{Y}{Z}) = X^3Y - Y^2Z^2 + 2XZ^3$.

Take F, C to be as in Definition 2.5. If $F = \prod F_i$ (so that $\deg F = \sum \deg F_i$), then writing C_i for the zero set of F_i, we have $C = \cup C_i$.

DEFINITION 2.1.7. We say that C is *irreducible* if and only if F has no proper ($\deg \geq 1$) homogeneous factors.

Now let's consider our intersection problem (2.1.2) once more, in the complex projective setting. Referring to the discussion up to Example 2.2, if C and L have an m-fold intersection at infinity, then the degree of the polynomial in (2.1.1) is $d - m$. The Fundamental Theorem of Algebra then says that (2.1.1) has $d - m$ complex roots counted with multiplicity, and we define these to be the intersection multiplicities for C and L in $\mathbb{C}^2$ as indicated in our discussion. We have proved a baby version of Bezout's theorem:

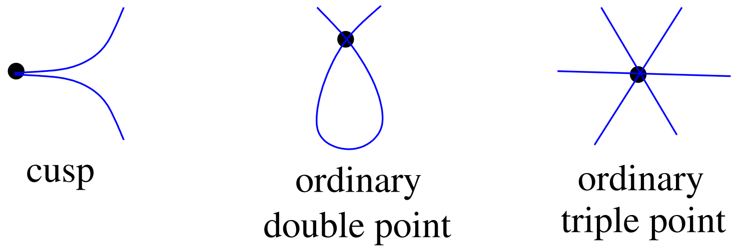
PROPOSITION 2.1.8. *Let $L \subset \mathbb{P}^2$ be a (projective) line in $\mathbb{P}^2$, i.e. an algebraic curve of degree one. A projective algebraic curve of degree d in $\mathbb{P}^2$ not containing L, meets L in d points counted with multiplicity.*

In proving this result we did a tiny bit of complex analysis on L, so were implicitly using its structure as a complex 1-manifold. In general it is quite useful to be able to do analytic computations on curves, but not all irreducible algebraic curves are complex manifolds (at least, without doing something to them called "normalization"). The obstructions are called singularities and will be explored in greater depth later. For now, we will just give a definition and a few examples.

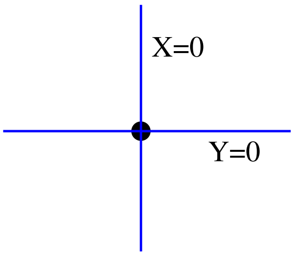
DEFINITION 2.1.9. A *singularity* or *singular point* of an affine algebraic (plane) curve $f(x, y) = 0$ is a point in $\mathbb{C}^2$ where f, $\frac{\partial f}{\partial x}$, and

$\frac{\partial f}{\partial y}$ are all zero — that is, a point on the curve where $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ vanish. A singularity of a projective algebraic curve $F(Z, X, Y) = 0$ is a point where F, $\frac{\partial F}{\partial X}$, $\frac{\partial F}{\partial Y}$, and $\frac{\partial F}{\partial Z}$ are all zero. A curve with one or more singular points is called *singular*; a curve with none is called *smooth*.

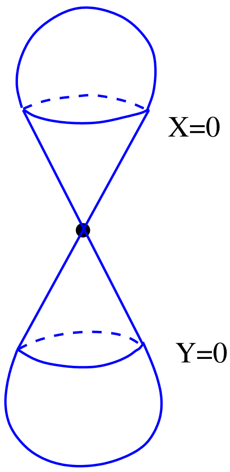
EXAMPLE 2.1.10. Here are some local real (schematic) pictures of plane curve singularities:



An example of a cusp is the point $[1 : 0 : 0]$ on $X^3 - Y^2 Z = 0$ (or $(0, 0)$ on $x^3 = y^2$); the curve $XY = 0$ has an ODP (ordinary double point, or "normal crossing") at $[1 : 0 : 0]$:



Now, this is just two $\mathbb{P}^1$'s (namely, $X = 0$ and $Y = 0$) touching at one point. A more (though not completely) "topologcially honest" picture of this is:



which makes it apparent that an ODP is actually a "bi-conical singularity".

Before we pass to the "analytic" side of our story, there are 2 more facts about homogeneous polynomials worth a quick mention. First, the map $S_3^d \to \mathcal{P}_2^d$ given by $F(Z,X,Y) \mapsto F(1,x,y)$ is an isomorphism, so $\dim(S_3^d) = \dim(\mathcal{P}_2^d)$ in particular. Second is the Euler formula

$$\sum_{i=0}^{N} Z_i \frac{\partial F}{\partial Z_i} = d \cdot F \quad \text{for } F \in S_{N+1}^d \tag{2.1.8}$$

which will be used in later chapters (cf. Chapter 6 for a proof).

2.2. Complex 1-manifolds

Recall from basic point-set topology that a *topological space* is a set X together with a collection $\{\mathfrak{U}_{\mathfrak{I}}\}_{\mathfrak{I}\in\Omega}$ of "open sets" containing X, the empty set, and all unions and all finite intersections of its members. (Here Ω is some typically huge index set. A *base* for the topology of X is a sub-collection of the $\{\mathfrak{U}_{\mathfrak{I}}\}_{\mathfrak{I}\in\Omega}$ which generates it under taking unions, and X is said to be *second countable* if it has a countable base.) X is called *Hausdorff* if points can be separated: i.e. given p and q, there exist disjoint open sets U and V containing p and q respectively.

In topology, a *homeomorphism* is a continuous, 1-to-1, open[5] map. Given a point $p \in X$, we like open sets $U \ni p$ that are homeomorphic to $\mathbb{R}^n$ (or equivalently, an open ball in $\mathbb{R}^n$) — these are called open neighborhoods of p. If these always exist, we say X is locally homeomorphic to $\mathbb{R}^n$. A second countable, Hausdorff topological space that is locally homeomorphic to $\mathbb{R}^n$, is called a *real n-manifold.*

In the case $n = 2$, we are going to layer "complex analyticity" onto this construction:

Definition 2.2.1. A *complex* 1-*manifold* consists of

(i) a connected Hausdorff topological space M;

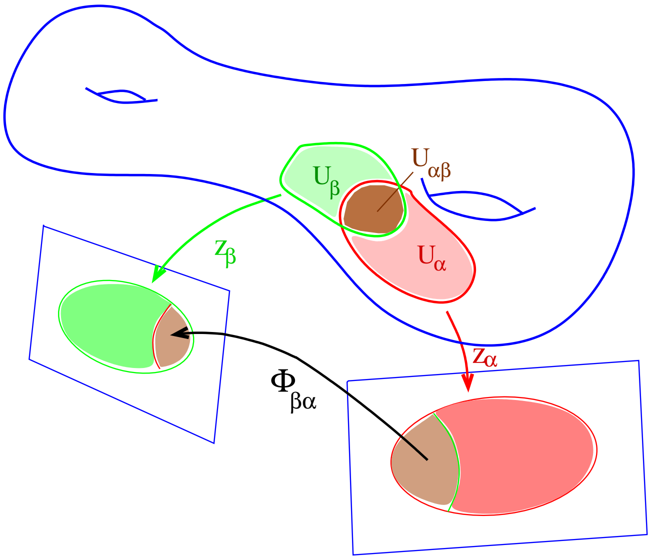
(ii) an open cover $\{U_\alpha\}$ of M (this is a finite set of open sets taken from amongst the $\{\mathfrak{U}_{\mathfrak{I}}\}$, such that $\cup_\alpha U_\alpha = M$); and

[5]a map f is open (resp. continuous) if the image (resp. preimage) of any open set is open

(iii) mappings $z_\alpha : U_\alpha \to \mathbb{C}$ that are homeomorphisms onto their image, such that the transition functions[6]

$$\Phi_{\beta\alpha} := z_\beta \circ z_\alpha^{-1} : \, z_\alpha(U_{\alpha\beta}) \to z_\beta(U_{\alpha\beta})$$

are biholomorphic (i.e., analytic isomorphisms).



The z_α are called *local coordinates*, and the $\Phi_{\beta\alpha}$ *transition (or patching) functions*; the entire collection $\{z_\alpha\}, \{\Phi_{\alpha\beta}\}$ is called an *analytic atlas*. The functions $\Phi_{\beta\alpha}$ are key: M is an complex analytic manifold because they are complex analytic. If in (iii) we replace $\mathbb{C}$ by $\mathbb{R}^n$ and require the transition functions to be smooth (i.e., have continuous partial derivatives of all orders), then M would have been a *smooth* (or "differentiable") real n-manifold instead.

If we think of the (complex analytic) transition functions in Definition 2.11 as maps from $\mathbb{R}^2$ to $\mathbb{R}^2$, then

$$\Phi(\underbrace{x}_{\substack{\text{real}\\ \text{part}}}, \underbrace{y}_{\substack{\text{imag.}\\ \text{part}}}) = (\underbrace{u(x,y)}_{\substack{\text{real}\\ \text{part}}}, \underbrace{v(x,y)}_{\substack{\text{imag.}\\ \text{part}}})$$

is smooth *and* u, v satisfy the *Cauchy-Riemann equations*. These may be expressed in terms of the Jacobian matrix

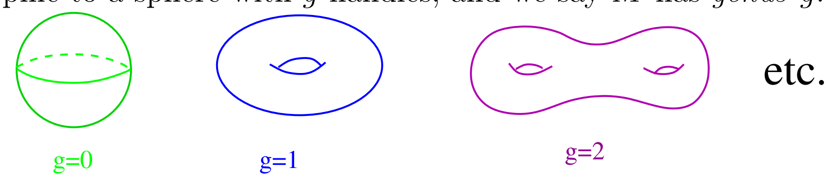
$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} u_x & -v_x \\ v_x & u_x \end{pmatrix},$$

[6]if U_α and U_β are distinct open sets in our open cover, we write $U_{\alpha\beta} := U_\alpha \cap U_\beta$. Leter we will write V_α for $z_\alpha(U_\alpha)$ and V_α^β for $z_\alpha(U_{\alpha\beta})$, so that $\Phi_{\beta\alpha}$ goes from V_α^β to V_β^α.

which consequently has positive determinant: $u_x^2 + v_x^2$ (obviously ≥ 0) cannot equal zero since Φ is *bi*holomorphic. Therefore Φ preserves orientation, so M is orientable as a real 2-manifold.[7]

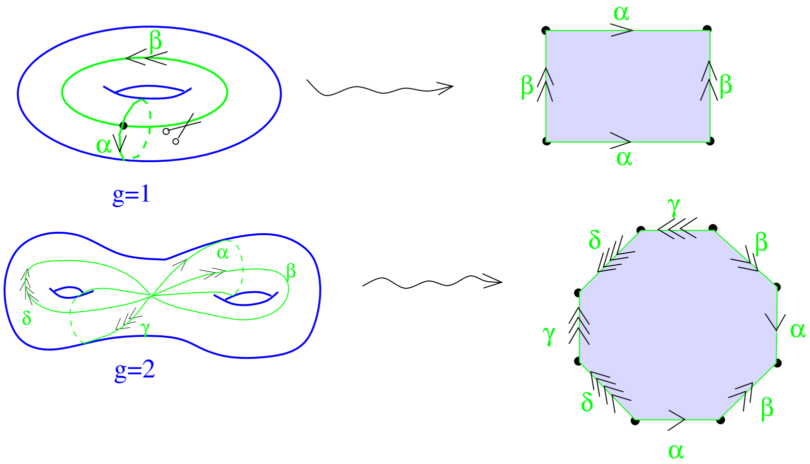
EXAMPLE 2.2.2. $\mathbb{C}$, $\mathbb{C}^*$, $\mathfrak{H}$, $\mathbb{P}^1$, and $\mathbb{C}/\Lambda$ are (the simplest) examples of complex 1-manifolds. For the first three, producing an analytic atlas is trivial (since you only need one U_α), and we will do this below for the latter two.

Now assume M is *compact*, that is, every open cover has a finite subcover. (In fact, since a complex 1-manifold always admits a metric, compactness is equivalent to every sequence of points having a convergent subsequence. Clearly $a_n = n$ has no limit in $\mathbb{C}$ but $a_n = [1 : n] = [\frac{1}{n} : 1]$ does limit to $[0 : 1]$ in $\mathbb{P}^1$, which is compact.) Then viewed over $\mathbb{R}$, M is an orientable, compact, connected, smooth 2-manifold. By a theorem in topology, this means that M is homeomorphic to a sphere with g handles, and we say M has *genus* g:



It is a fact that all $g \geq 0$ occur for complex manifolds; we'll show this in a moment for $g = 0, 1$. To do complex analysis on M, you can use the local coordinates, but for some purposes it is also convenient to cut M into a simply connected region, e.g.

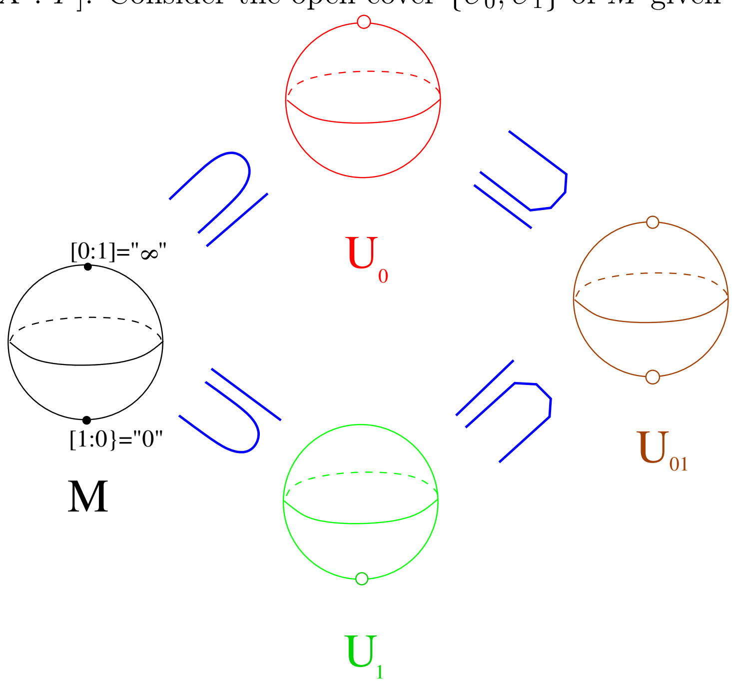
[7]in fact, a matrix of the form $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ is a rotation times a dilation, hence *preserves angles* – that is to say, under the assumption of the CR-equations, Φ is *conformal*. So a complex 1-manifold is essentially a differentiable real 2-manifold with conformal transition functions.



Extrapolating from this, one sees that if you begin with a sphere with g handles, the cut-open version is a polygon with $4g$ edges identified in pairs. (The black points on the right-hand side are *all* identified.) Using this[8] we can do a quick computation of the Euler characteristic of M:

$$\chi_M := \text{faces} - \text{edges} + \text{vertices} = 1 - 2g + 1 = 2 - 2g. \tag{2.2.1}$$

EXAMPLE 2.2.3. ($g = 0$) Let $M := \mathbb{P}^1$ with homogeneous coordinates $[X : Y]$. Consider the open cover $\{U_0, U_1\}$ of M given by:



That is, $U_0 = M \backslash \{[0 : 1]\}$ and $U_1 = M \backslash \{[1 : 0]\}$. For local coordinates,

[8]traditionally one would use the numbers of faces, edges, and vertices in a *triangulation* of M, but using a polygonal decomposition like this is also OK

we take

$$\begin{array}{rccc} z_0 : & U_0 & \to & \mathbb{C} \\ & [X:Y] & \mapsto & \frac{Y}{X} \end{array}$$

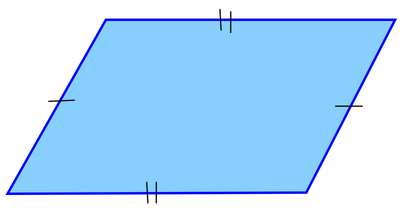
and

$$\begin{array}{rccc} z_1 : & U_1 & \to & \mathbb{C} \\ & [X:Y] & \mapsto & \frac{X}{Y} \end{array}.$$

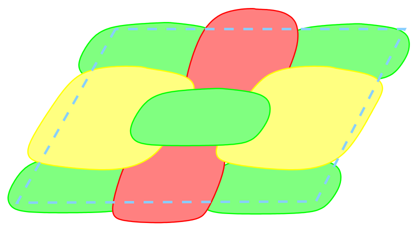
Writing $U_{01} := U_0 \cap U_1$, we have $z_0(U_{01}) = \mathbb{C}^* \subset \mathbb{C}$ and $z_1(U_{01}) = \mathbb{C}^* \subset \mathbb{C}$. The transition function (which goes from $z_0(U_{01})$ to $z_1(U_{01})$ by definition) is then

$$\begin{array}{rccc} \Phi_{10} : & \mathbb{C}^* & \to & \mathbb{C}^* \\ & u & \mapsto & \frac{1}{u} \end{array}.$$

EXAMPLE 2.2.4. ($g = 1$) Let $\lambda_1, \lambda_2 \in \mathbb{C}$ be linearly independent over $\mathbb{R}$. Then $\Lambda := \mathbb{Z}\langle \lambda_1, \lambda_2 \rangle = \mathbb{Z}\lambda_1 + \mathbb{Z}\lambda_2$ is a lattice, and we set $M := \mathbb{C}/\Lambda$. (This means that $z, z' \in \mathbb{C}$ give the same point in M if and only if $z - z' \in \Lambda$.) We endow M with local coordinates on the neighborhoods shown



M



covering by coordinate neighborhoods

basically by using the coordinate on $\mathbb{C}$ (before quotienting by Λ), and find that the transition functions Φ_{ij} are all either the idenitity or translation by some $\lambda \in \Lambda$. Topologically, M is a torus (cf. the $g = 1$ pictures above), which is evident from performing the identifications on the sides of a fundamental region for $\mathbb{C}/\Lambda$ as shown.

2.3. Riemann surfaces

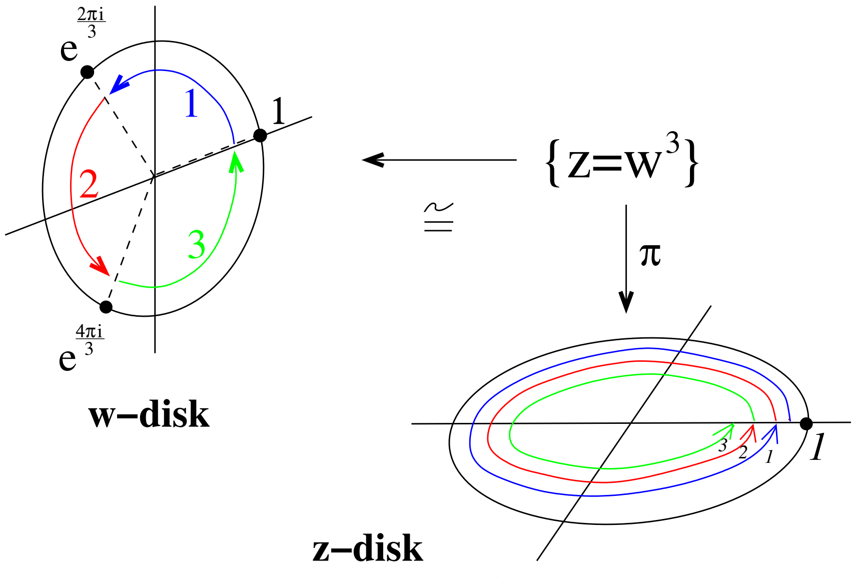
Traditionally, a *Riemann surface* M is a compact complex 1-manifold obtained as the "existence domain" of an algebraic (typically multivalued) function over $\mathbb{P}^1$. That isn't the definition I'll use here, but I do want to explain the concept.

For example, given distinct complex numbers α_i, the algebraic function

$$\mathfrak{F}(z) := \sqrt{\prod_{i=1}^{2g+2}(z-\alpha_i)}$$

on $\mathbb{P}^1$ can be made single-valued on a complex manifold M (of genus g, it turns out) constructed as follows: On two copies of $\mathbb{P}^1$, cut identical nonintersecting slits from α_{2j-1} to α_{2j} for $j = 1, \ldots, g+1$. Then glue the two copies of $\mathbb{P}^1$ together on these slits, forming a set on which $\mathfrak{F}$ becomes single-valued; finally, endow this set with an analytic atlas to get a complex manifold M. This manifold has a distinguished morphism $M \xrightarrow{\pi} \mathbb{P}^1$ presenting it as a finite branched cover of the projective line. We won't do this explicitly here — especially endowing it with an analytic atlas, since that is really a special case of normalizing an algebraic curve (cf. §3.1).[9]

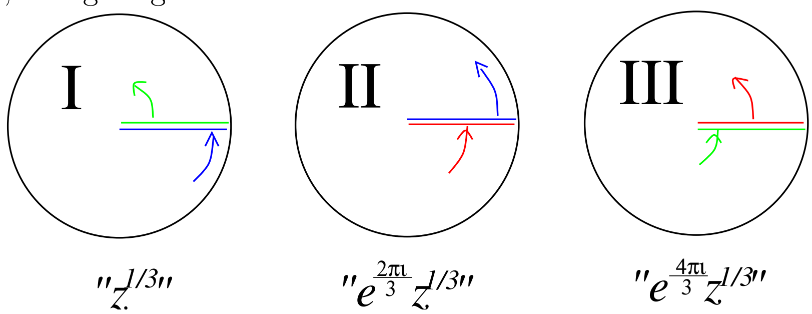
Instead, let's visualize what a couple of "existence domains" for algebraic functions look like, starting with the "Riemann surface of $(w =)z^{\frac{1}{3}}$ over the unit disk". This is some object fitting (as "$\{z = w^3\}$") into the following picture:



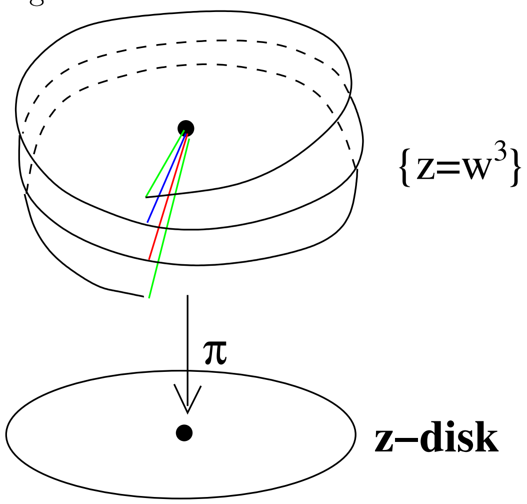
To construct it, think about following $z^{\frac{1}{3}}$ around the disk once counterclockwise: when you reach your starting point the function has become $e^{\frac{2\pi i}{3}}$ times the branch of $z^{\frac{1}{3}}$ you started with; going around once more, you get $e^{\frac{4\pi i}{3}}z^{\frac{1}{3}}$; and one more time gets you back to your original

[9]it makes a very instructive exercise though, and the next example gives a hint on how to do it

branch. So taking three unit disks, slitting them along the positive reals, and gluing them as indicated



we get the "parking lot"

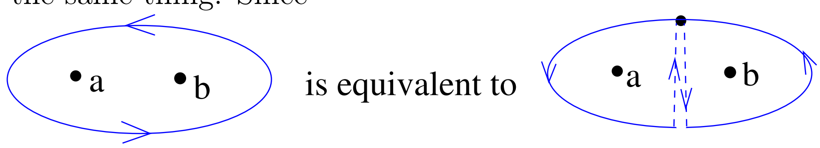


(The green segments are glued but I can't draw in 4 dimensions.) An easier way to visualize this "Riemann surface" is this: it's just the w-disk. The difficulty is in seeing the w-disk "over" the z-disk.

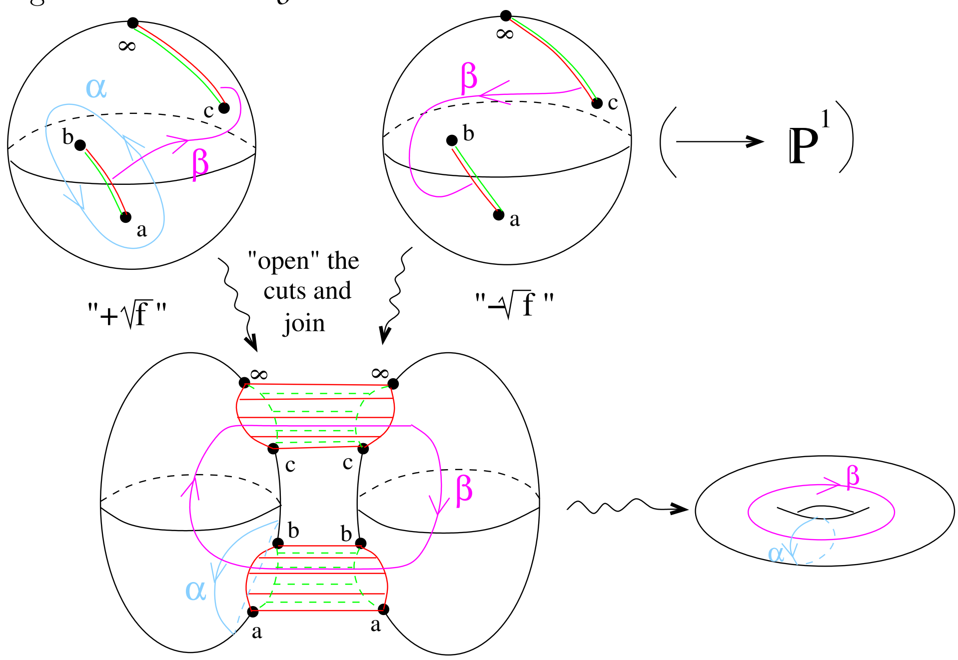
Next, let's construct an existence domain for

$$\mathfrak{F}(z) = \sqrt{(z-a)(z-b)(z-c)}$$

over $\mathbb{P}^1$. In a neighborhood of $z_0 = a, b, c$ this looks like the "Riemann surface of $(z-z_0)^{\frac{1}{2}}$ over a disk", which is the same as the construction we just did except with 2 unit disks instead of 3. Indeed, going once around a, b, or c takes $\mathfrak{F} \mapsto -\mathfrak{F}$; and furthermore, because the degree of the polynomial under the square root is odd, going once around ∞ does the same thing. Since



going around two points at once gives no change. So taking two $\mathbb{P}^1$'s and cutting and pasting them as indicated, we end up with a manifold of genus 1 on which $\mathfrak{F}$ becomes well-defined:



(In the picture, α and β are called 1-cycles; there just there to make the topology clear.) The same construction works if we replace $\mathfrak{F}(z)$ by $\sqrt{(z-a)(z-b)(z-c)(z-d)}$, with d replacing ∞.

In fact, by a deep result (on existence of nonconstant meromorphic functions on complex 1-manifolds) any compact complex 1-manifold is an "existence domain" of the sort we have just discussed: they are equivalent objects in the end. The following is motivated by this, and the desire to keep things simple:

DEFINITION 2.3.1. A *Riemann surface* is a compact complex 1-manifold.

Exercises

(1) Take projective closures of $C := \{y^2 = (x-1)(x-2)(x-3)(x-4)\}$ and $L := \{x = 0\}$ in $\mathbb{P}^2$ (find associated homogeneous equations), and determine all intersections and their multiplicities (give the projective coordinates of the points). What is the sum of multiplicities?

(2) Find the affine equation associated to $Z_0^3 + Z_1^3 + Z_2^3 = \lambda Z_0 Z_1 Z_2$. (This equation is homogeneous of degree 3 — λ is a scalar, not a coordinate).

(3) Let F be a polynomial in 3 variables. Prove that $F(z_0, z_1, z_2) = 0 \implies F(\alpha z_0, \alpha z_1, \alpha z_2) = 0$ $(\forall\, \alpha \in \mathbb{C}^*)$ forces F to be homogeneous (of some degree). [You will have to assume the following result: given two polynomials f and g (in (z_0, z_1, z_2)), with vanishing locus $\{f = 0\} \subseteq \{g = 0\}$, then f divides a power of g. This is called *Study's lemma* and will be proved later.] If this is too challenging, show the converse. If you like, do both.

CHAPTER 3

The normalization theorem

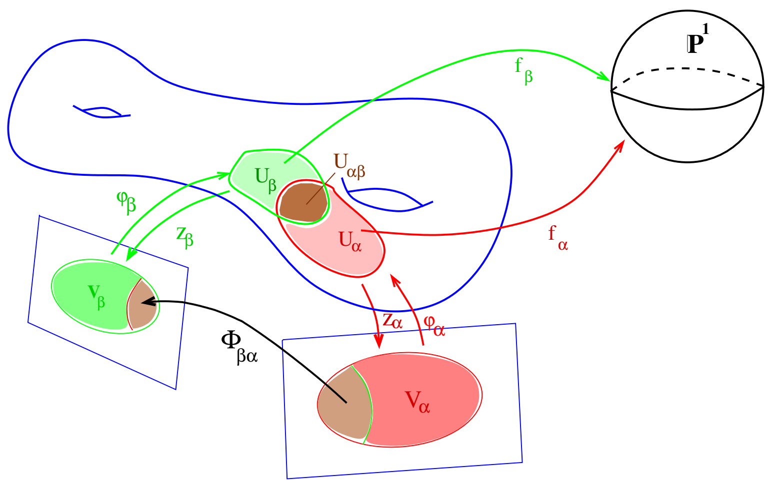
We state (but do not yet prove) the promised relationship between algebraic curves and Riemann surfaces, and explain how to work it out directly for conics. To state the general relationship, however, we need the notion of meromorphic functions on a Riemann surface, so we will first define and prove a few results about those.

3.1. Meromorphic functions on a Riemann surface

Let M be a Riemann surface (Definition 2.15) with analytic atlas $\{(U_\alpha, z_\alpha)\}$ (Definition 2.11), and write $V_\alpha := z_\alpha(U_\alpha) \subseteq \mathbb{C}$. The *local analytic chart* $\varphi_\alpha : V_\alpha \to U_\alpha(\subseteq M)$ is simply defined to be the (composition) inverse of of the local coordinate z_α. (I just don't like writing z^{-1} since in some settings this is easy to confuse with $\frac{1}{z}$.)

Definition 3.1.1. A meromorphic (resp. holomorphic) function $f \in \mathcal{K}(M)$ (resp. $\mathcal{O}(M)$) is a collection of continuous maps $f_\alpha : U_\alpha \to \mathbb{P}^1$ such that

- the $\{f_\alpha\}$ "agree" on overlaps (viz., $f_\alpha = f_\beta$ on $U_{\alpha\beta}$), and
- $f_\alpha \circ \varphi_\alpha$ is a meromorphic (resp. holomorphic) function, in the sense of complex analysis, for all α.



REMARK 3.1.2. (a) One really works with functions of the coordinate z_α, i.e. the function $f_\alpha \circ \varphi_\alpha =: g_\alpha$ (mapping $V_\alpha \to \mathbb{P}^1$), and then the compatibility condition reads

$$g_\alpha \circ \Phi_{\alpha\beta} = g_\beta. \tag{3.1.1}$$

(b) $\mathcal{K}(M)$ is a field, since you can multiply, add, and invert (additively and multiplicatively) meromorphic functions.

PROPOSITION 3.1.3. [LIOUVILLE'S THEOREM] *M compact $\implies$ $\mathcal{O}(M) \cong \mathbb{C}$ (constant functions).*

PROOF. On the one hand, $f \in \mathcal{O}(M) \implies f(M) \subset (\mathbb{P}^1 \backslash \{\infty\}) = \mathbb{C}$; while on the other, M compact and f continuous $\implies$ $f(M)$ is compact. Applying absolute value gives a compact subset $|f(M)| \subset \mathbb{R}_{\geq 0}$. This has a maximum element, whch is assumed at some point $p \in M$, and this p lies in some U_α. Hence, the absolute value of the holomorphic function $g_\alpha = f_\alpha \circ \varphi_\alpha$ attains a maximum on V_α (at $\varphi_\alpha(p)$), and by the maximum modulus principle, g_α (and thus f_α) is some constant $c \in \mathbb{C}$.

Let U_β be any open set of the atlas meeting U_α. Since $f_\beta = f_\alpha = c$ on $U_{\alpha\beta}$, and $U_{\alpha\beta}$ has accumulation points, $f_\beta = c$ on U_β. One continues this argument now for any open set meeting U_α or U_β, and so forth. By connectedness of M, this shows $f = c$ on all open sets of the atlas, hence on all of M. □

DEFINITION 3.1.4. Let $f \in \mathcal{K}(M)$ be a meromorphic function. For any $p \in M$, f is locally of the form

$$z^m h(z) \tag{3.1.2}$$

with $m \in \mathbb{Z}$, z a local coordinate vanishing at p (i.e. $z(p) = 0$), and $h(z)$ a local holomorphic function of z with $h(0) \neq 0$.[1] We say that the *order* $\nu_p(f)$ of f at p is m.

With this bit of language it is easy to compute the meromorphic function field for Riemann surfaces of genus 0 and 1.

THEOREM 3.1.5. *(a) $\mathcal{K}(\mathbb{P}^1) \cong \mathbb{C}(z)$ (z an indeterminate).*

[1]To be absolutely precise, if z is a local coordinate on $U \ni p$, with $V = z(U)$, then h is a holomorphic function on V. I'll frequently assume things like this to be "understood".

(b) Writing $\Lambda := \{m_1\lambda_1 + m_2\lambda_2 \,|\, m_i \in \mathbb{Z}\}$ *(*$\lambda_1, \lambda_2 \in \mathbb{C}$ *linearly independent over* $\mathbb{R}$*) for a lattice,* $\mathcal{K}(\mathbb{C}/\Lambda) \cong \mathbb{C}(\wp, \wp')$ *where* $\wp(u)$ *is the Weierstrass* $\wp$*-function for* Λ*.*

Proof. (a) Referring to Example 2.13, write $z = z_0$ and $w = z_1$ for the two local coordinates. I am really going to use z as a global coordinate on $\mathbb{P}^1$; the statement we want to prove is that meromorphic functions on $\mathbb{P}^1$ are precisely the rational functions of z.

In one direction, this is easy: if P, Q are polynomials in z (with $Q \neq 0$), clearly $\frac{P(z)}{Q(z)}$ is the restriction to U_0 of a meromorphic function on $\mathbb{P}^1$ (on U_1, it is $\frac{P(\frac{1}{w})}{Q(\frac{1}{w})}$).

Conversely, are all meromorphic functions rational? Given $f \in \mathcal{K}(\mathbb{P}^1)$, $\nu_p(f) < 0$ at finitely many[2] points $z_i (= p)$, and we shall for simplicity assume none of these is the point ∞. Let $\mathsf{P}_i(z) = \sum_{\nu_{z_i}(f) \leq k < 0} \beta_{ik}(z - z_i)^k$ (sum is over k) be the principal part of the Laurent expansion of f at z_i, and consider $G(z) = \sum \mathsf{P}_i(z)$. Then $f - G \in \mathcal{O}(\mathbb{P}^1)$ is constant by Liouville; and since G is rational, we're done.

(b) Next, $f \in \mathcal{K}(\mathbb{C}/\Lambda)$ if and only if f is a doubly-periodic meromorphic function on $\mathbb{C}$: that is, $f(u) = f(u + m_1\lambda_1 + m_2\lambda_2)$ for all $m_1, m_2 \in \mathbb{Z}$ (also known as an elliptic function). We will see later that these are generated (rationally) by

$$\wp(u) := \frac{1}{u^2} + \sum_{\substack{\lambda \in \Lambda \\ \lambda \neq 0}} \left(\frac{1}{(u-\lambda)^2} - \frac{1}{\lambda^2} \right)$$

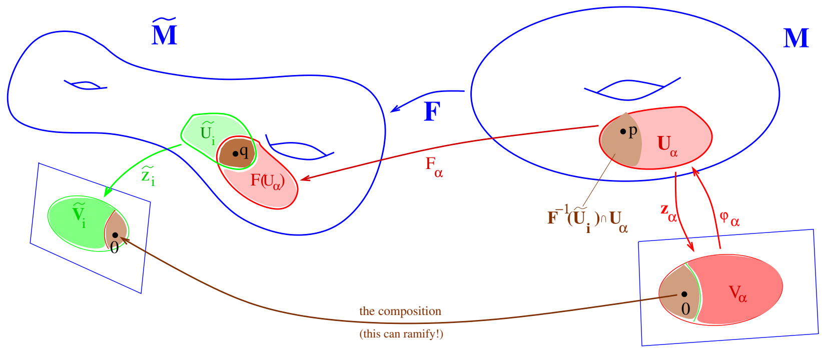
and its derivative. $\square$

Definition 3.1.6. A *morphism* (or holomorphic map) $M \xrightarrow{F} \tilde{M}$ of Riemann surfaces[3] is a collection $F_\alpha : U_\alpha \to \tilde{M}$ of continuous maps (agreeing on the $\{U_{\alpha\beta}\}$) such that the composition[4] $\tilde{z}_i \circ F_\alpha \circ \varphi_\alpha|_{z_\alpha\{F^{-1}(\tilde{U}_i) \cap U_\alpha\}}$ is holomorphic for all α, i. (Note that this definition works more generally for complex 1-manifolds — compactness is inessential.)

[2]otherwise compactness $\implies$ zeroes of $\frac{1}{f}$ have an accumutaion point $\implies$ $\frac{1}{f}$ identically 0. (Also, note that I am identifying points by the value of the coordinate z on $\mathbb{P}^1$. If M were not $\mathbb{P}^1$, I would write p_i instead of z_i.)

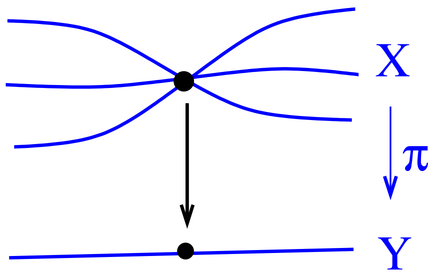
[3]write $\{U_\alpha, z_\alpha\}$ and $\{\tilde{U}_i, \tilde{z}_i\}$ for the atlases.

[4]this composition renders $\tilde{z}_i$ as a function of z_α (and is a local snapshot of F in this sense)



Now suppose we have $p \in (U_\alpha \subset)M$ and $q \in (\tilde{U}_i \subset)M$ with $F(p) = q$, $z_\alpha(p) = 0$ and $\tilde{z}_i(q) = 0$, as shown in the above figure. Assuming F is nonconstant, then after "normalizing" the local coordinates,[5] we have $\tilde{z}_i(z_\alpha) = (z_\alpha)^\mu$ for some (unique) $\mu \in \mathbb{Z}_{>0}$. One says that f has *ramification index* μ at p (over q). If this index is > 1, we say that f is *branched* over q (or ramifies at p).

REMARK 3.1.7. For $\mu = 3$, we have already seen this picture in §2.3. In general, for a holomorphic map of Riemann surfaces $\pi : X \to Y$, for all but finitely many $y \in Y$ the number $|\pi^{-1}(y)|$ is the same, and this is called the *degree* of the mapping π. (This will be explained in greater depth in a later chapter.) The branch points of π are just the remaining points of Y. Usually we will just draw a schematic picture like



and it is understood that the picture is really as in §2.3 — so that going around the point on the "base" Y moves you between *branches* of the "cover" X.

PROPOSITION 3.1.8. *Let M be a Riemann surface (or, more generally, a complex 1-manifold). The holomorphic maps $M \to \mathbb{P}^1$, excluding the constant map sending all points to $\{\infty\}$, are simply the meromorphic functions $\mathcal{K}(M)$.*

[5]see Exercise 4 below

PROOF. Again refer to Example 2.13: given a morphism $F : M \to \mathbb{P}^1$ (Definition 3.6), by definition $z_0 \circ F_\alpha \circ \varphi_\alpha$ is holomorphic on the complement of the preimage of ∞, while $z_1 \circ F_\alpha \circ \varphi_\alpha = \frac{1}{z_0 \circ F_\alpha \circ \varphi_\alpha}$ is holomorphic on the complement of the preimage of 0.[6] Hence, $F_\alpha \circ \varphi_\alpha$ is meromorphic and $\{F_\alpha\}$ defines a meromorphic function (Definition 3.1). The converse is even more tautological! □

Later we will discuss morphisms (holomorphic maps) of complex manifolds of any dimension. The following is a special case:

DEFINITION 3.1.9. Write $[Z_0 : Z_1 : \cdots : Z_n]$ for (projective) coordinates on $\mathbb{P}^n$. A map σ from a Riemann surface M to $\mathbb{P}^n$ is called holomorphic if and only if all compositions $[Z_i \circ \sigma : Z_j \circ \sigma]$ are holomorphic as maps to $\mathbb{P}^1$ on the open subsets of M where they are well-defined.

REMARK 3.1.10. If the image $\sigma(M)$ does not live in a coordinate hyperplane, this is the same as saying that *composing σ with each affine coordinate on $\mathbb{C}^n$ gives a meromorphic function.*

To see this, first write $\mathcal{M}_{ij}$ for the subsets of M where (under σ) Z_i and Z_j are not both zero; these are the open subsets in the last definition.[7] By Prop. 3.1.8, the conditions of Defn. 3.1.9 mean that $\frac{Z_j}{Z_i}$ are meromorphic functions on the $\mathcal{M}_{ij}$. We need to show that the $z_j = \frac{Z_j}{Z_0}$ extend to meromorphic functions on all of M. First, M is covered by the open sets $\{Z_i \neq 0\}$. Hence, for $p \notin \mathcal{M}_{0j}$ (i.e. Z_j and Z_0 vanish at p), we have a neighborhood U containing p where some other Z_i does not vanish, so that $U \subset \mathcal{M}_{ij}, \mathcal{M}_{i0}$. Now, on $U \cap \mathcal{M}_{0j}$ we can write $z_j = \frac{Z_j}{Z_i} \cdot \left(\frac{Z_0}{Z_i}\right)^{-1}$ as a product of functions which are meromorphic on all of U, hence showing that z_j extends as desired.

3.2. Riemann surfaces parametrize algebraic curves

Here is the Normalization Theorem. We will prove part (A) in this course.

THEOREM 3.2.1. *(A) Given an irreducible algebraic curve $C \subset \mathbb{P}^2$, there exists a Riemann surface M and a holomorphic map $\sigma : M \to \mathbb{P}^2$ with C as its image which is 1-to-1 on $\sigma^{-1}(C \backslash sing(C))$.*

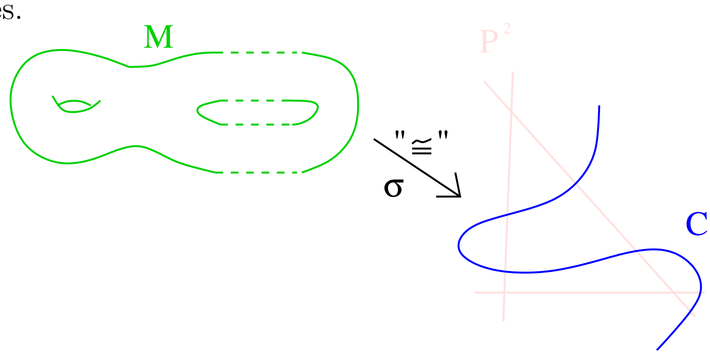
[6] The holomorphicity of $\frac{1}{z_0 \circ F_\alpha \circ \varphi_\alpha}$ guarantees, in particular, that $F_\alpha \circ \varphi_\alpha$ has only poles and not essential singularities.

[7] Some, but not all, of these $\mathcal{M}_{ij}$ may be empty if $\sigma(M)$ is contained in a coordinate hyperplane. We have excluded this "degenerate" possibility for simplicity.

(B) Given a Riemann surface M, there exists a holomorphic map $\sigma : M \to \mathbb{P}^2$ such that

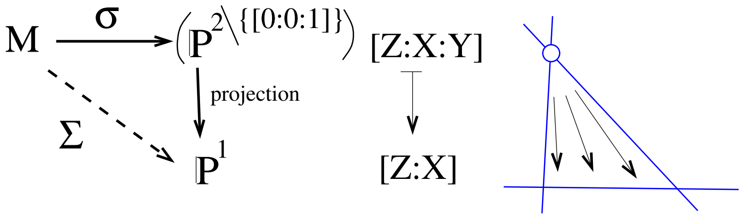
- *$\sigma(M)$ is an irreducible algebraic curve with sing$(\sigma(M))$ consisting of ordinary double points (or empty), and*
- *σ is 1-to-1 off the preimage of these ordinary double points.*

In this sense, irreducible smooth projective algebraic plane curves are equivalent to, and are isomorphically parametrized by, Riemann surfaces.



If a curve C is not smooth, then the normalization "desingularizes" it (and we shall see this quite explicitly later on). In either case, we say that M is the *normalization* of C.

Let's look briefly at the meaning of (B), which we will not prove in this course. For a given Riemann surface (i.e. compact complex 1-manifold) M, it guarantees a holomorphic map to $\mathbb{P}^2$, with image $\sigma(M) = C$ =projective closure of $\{f(x,y) = 0\}$. Changing coordinates on $\mathbb{P}^2$ if necessary, we may assume that C does not pass through $[0 : 0 : 1]$. So it makes sense to consider the composition



which exhibits M as a branched cover of $\mathbb{P}^1$ — or more precisely, as the existence domain of the algebraic function $g(x)$ obtained by solving

$$f(x, g(x)) = 0.$$

So Theorem 3.10(B) contains the statement that every complex 1-manifold is an existence domain in the sense of §2.3.

We should also note that any Riemann surface admits a holomorphic *embedding* $\sigma : M \hookrightarrow \mathbb{P}^3$, an even nicer result than part (B) above!

3.3. Stereographic projection

As a plausibility check on Theorem 3.10(A), we'd like a recipe for normalizing conics — i.e. degree-2 (conic) curves $C \subset \mathbb{P}^2$. Given a point $p \in C$, and any line line ℓ through p, by Proposition 2.8 ℓ either meets C in two points with mutliplicity 1 or in 1 point with multiplicity 2. Put differently, we have either

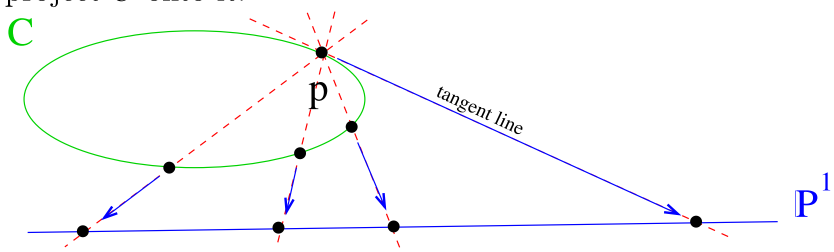
- $\ell \cap C = \{p, q\}$

or

- $\ell \cap C = 2p$, i.e. $\ell = T_pC$ is the *tangent line* to C at p.

(We will give a systematic treatment of tangent lines below.) Conversely, given p and any other point q on C, there is a unique line through them (and it doesn't meet C anywhere else).

There are two ways to think of why this gives a parametrization of C. One possibility is to take a fixed line ($\cong \mathbb{P}^1$) and use lines through p to project C onto it:



This is where the term "stereographic projection" comes from.

But this auxiliary projective line is superfluous, because the family of lines through p already gives a $\mathbb{P}^1$. (Indeed this is close to the original definition of what $\mathbb{P}^1$ is.) We can parametrize this $\mathbb{P}^1$ by the slope of the line with respect to suitable coordinates (usually $(x, y) = (\frac{Z_1}{Z_0}, \frac{Z_2}{Z_0})$). The upshot is that we get a 1-1 correspondence between lines through p and points on C, so that we are in the situation of §3.2 with $M \cong \mathbb{P}^1$.

EXAMPLE 3.3.1. Suppose we wish to find a parametrization $\mathbb{P}^1 \xrightarrow{\sigma} C$ of the conic $\{X^2 + Y^2 = Z^2\} \subset \mathbb{P}^2$, which in affine coordinates is $x^2 + y^2 = 1$. We choose a point on C, say $p = (1, 0)$, and draw lines $y = \mu(x-1)$ through p. (The slope here is μ, and this should be viewed as a choice of coordinate on $\mathbb{P}^1$.) Substituting into $x^2 + y^2 = 1$ and solving for x in terms of μ, we have

$$\begin{aligned} & x^2 + \mu^2(x-1)^2 = 1 \\ \Longrightarrow \quad & (\mu^2+1)x^2 - 2\mu^2 x + (\mu^2 - 1) = 0 \\ \Longrightarrow \quad & (x-1)\{(1+\mu^2)x + (1-\mu^2)\} = 0. \end{aligned}$$

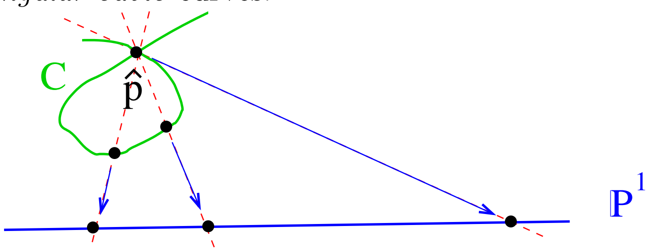
Ignoring the solution $x = 1$ (which corresponds to p), we have

$$x = \frac{\mu^2-1}{\mu^2+1}\,, \quad y = \mu\left(\frac{\mu^2-1}{\mu^2+1} - 1\right) = \frac{-2\mu}{\mu^2+1}.$$

Hence, we find

$$\sigma(\mu) = \left(\frac{\mu^2-1}{\mu^2+1}, \frac{-2\mu}{\mu^2+1}\right).$$

One can also do stereographic projection to construct normalizations of *singular cubic* curves:



The idea here is to consider lines through the singular point $\hat{p}$; since any such ℓ already meets C "twice", it will only hit C in one additional point (by Proposition 2.8). You'll work an example in the exercises below. This will not work for a smooth cubic.

Exercises

(1) Give a parametrization $m \mapsto (x(m), y(m))$ (hence an isomorphism $\mathbb{P}^1 \to C$) of the smooth conic curve C that is the projective closure of $3x^2 - y^2 = 1$. (You may work in affine coordinates.)

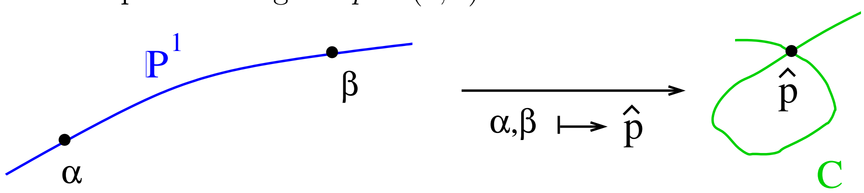
(2) Show that for any RS M and meromorphic function $(0 \neq) f \in \mathcal{K}(M)$, one has $\sum_{p \in M} \nu_p(f) = 0$. [Hint: Use the residue theorem from complex analysis. Cut open the RS as in Chapter 2, and integrate $\frac{df}{f}$ along the "boundary".]

(3) Convince yourself that the order $\nu_p(f)$ of a meromorphic function on a RS M (Definition 3.1.4) is independent of the choice of local coordinate.

(4) Prove the following, which was claimed in Definition 3.1.6: Given M, M' Riemann surfaces with a holomorphic map $f : M \to M'$

(and $f(p) = q$). Then there exist (U, z) on M and (V, w) on M' satisfying $z(p) = 0 = w(q)$, such that $w = z^\mu$ (for some $\mu \in \mathbb{N}$) is the local form taken by f near p. [Here for example "(U, z)" means an open disk $U \subset M$ with local coordinate $z : U \to \mathbb{C}$.]

(5) Find a parametrization $\mathbb{P}^1 \to C$ of the singular cubic $Y^2Z - X^2Z + X^3 = 0$ in $\mathbb{P}^2$. (C has an ordinary double point $\hat{p}$ at $[Z : X : Y] = [1 : 0 : 0]$. Check that this point is indeed a singularity of C.) To do this, convert to affine coordinates, substitute in $y = mx$, and solve for the other intersection point's coordinates as a function of m. Two points will go to $\hat{p} = (0, 0)$. Picture:



What are α and β? Change coordinates on $\mathbb{P}^1$ (fractional linear transformation) so that in your new coordinate, 0 and ∞ are sent to $\hat{p}$. Your parametrization should read now $\varphi : \mathbb{P}^1 \to C$ sending $z \mapsto (x(z), y(z))$ with $0, \infty \mapsto \hat{p}$. This will be used in a later exercise.

CHAPTER 4

Lines, conics, and duality

To complete our introduction to algebro-geometric concepts on the level of curves, in this chapter we'll study projective transformations, tangent lines, and dual curves. Our convention will be to write $\underline{Z} = \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix}$ for column vectors in $\mathbb{C}^3$ (written with respect to the "standard basis" $\underline{\mathbf{e}}$), and $[\underline{Z}] = [Z_0 : Z_1 : Z_2]$ for the corresponding point in $\mathbb{P}^2$.

4.1. The classification of complex conics

The story begins in even lower degree, with *lines* — i.e. degree 1 algebraic curves. These are subsets of $\mathbb{P}^2$ of the form

$$L_\lambda = \{ {}^t\underline{\lambda} \cdot \underline{z} = 0\}, \tag{4.1.1}$$

where $\underline{\lambda}$ is a nonzero vector in $\mathbb{C}^3$. Note that for $\alpha \in \mathbb{C}^*$, $L_{\alpha\underline{\lambda}} = L_{\underline{\lambda}}$.

By stereographic projection (cf. §3.3), lines and smooth conics are isomorphically parametrized by $\mathbb{P}^1$ in the sense of the Normalization Theorem. (For lines, the projection is done through a point not on the line; for conics, one chooses any point on the conic.) However, not all conics are smooth, and so we will need to classify conics up to projective equivalence.[1] The two key non-smooth examples to keep in mind are the *pair of lines*

$$\{XY = 0\} = \{X = 0\} \cup \{Y = 0\}$$

and the *double line*

$$\{X^2 = 0\}.$$

[1]There is a somewhat subtle point here. For smooth curves in general, projective equivalence is finer (equates fewer curves) than isomorphism as Riemann surfaces. However, you have to consider curves of degree at least 5 to see this discrepancy. As far as conics are concerned, we like projective equivalence simply because it gives a uniform and algebraic treatment of singular and smooth curves.

The first has two irreducible components (and is hence *reducible*), while the second has one component of "multiplicity two" (and is said to be *non-reduced*).[2]

To define projective equivalence, we introduce the projective general linear group

$$PGL(n,\mathbb{C}) := \frac{GL(n,\mathbb{C})}{\langle \alpha \cdot \mathrm{id.} \mid \alpha \in \mathbb{C}^* \rangle}.$$

(We have $A \equiv B \iff B = \alpha A$ for some $\alpha \in \mathbb{C}^*$.) Consider the action of $PGL(3,\mathbb{C})$ on $\mathbb{P}^2$ by

$$T\left(\begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{pmatrix}\right)[Z_0 : Z_1 : Z_2] =$$

$$\left[\begin{matrix} a_{00}Z_0 + a_{01}Z_1 \\ +a_{02}Z_2 \end{matrix} : \begin{matrix} a_{10}Z_0 + a_{11}Z_1 \\ +a_{12}Z_2 \end{matrix} : \begin{matrix} a_{20}Z_0 + a_{21}Z_1 \\ +a_{22}Z_2 \end{matrix}\right]$$

or in more compact notation

$$T(A)[\underline{Z}] = [A \cdot \underline{Z}].$$

(We are, consistently with the notation mentioned at the beginning of the chapter, letting the matrix A act on $\underline{Z}$ viewed as a column vector.[3]) This action is well-defined:

- it sends no nonzero $\underline{Z}$ to $\underline{0}$ (recall $[\underline{0}]$ is not a point in $\mathbb{P}^2$);
- if $\underline{Z} = \alpha \underline{Y}$, then $T(A)[\underline{Z}] = [A\underline{Z}] = [A \cdot \alpha \underline{Y}] = [\alpha A \underline{Y}] = [A\underline{Y}]$;
- if $A = \alpha B$, then $T(A)[\underline{Z}] = [\alpha B \underline{Z}] = [B\underline{Z}] = T(B)[\underline{Z}]$.

DEFINITION 4.1.1. The transformations $T(A) : \mathbb{P}^2 \to \mathbb{P}^2$, $A \in PGL(3,\mathbb{C})$, are the projective linear transformations (or *projectivities*) of $\mathbb{P}^2$.

---

[2]roughly speaking, "reduced" means "all of its irreducible components are of multiplcity one". So while $\{XY = 0\}$ is reduced, something like $\{XY^3 = 0\}$ is not. Obviously this is going a bit beyond the notion of an algebraic curve as a solution set, since it incorporates multiplicity. To really formalize what such an object is, we would have to work with scheme theory or algebraic cycles, which I do not want to do. So unless otherwise stated, in this course an "algebraic curve" is assumed to be reduced.

[3]the dot indicates matrix multiplication. This will often be omitted.

REMARK 4.1.2. The analogue of projectivities on $\mathbb{P}^1$ are simply the fractional linear transformations:

$$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)[Z_0 : Z_1] = [aZ_0 + bZ_1 : cZ_0 + dZ_1].$$

So writing "z" for the point $[1 : z]$,

$$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)(z) = \frac{c + dz}{a + bz}.$$

You probably know from complex analysis that such transformations preserve the cross-ratio of 4 points. Furthermore, they are the only automorphisms of $\mathbb{P}^1$ (invertible morphisms from $\mathbb{P}^1 \to \mathbb{P}^1$) as a complex 1-manifold.

How do projectivities affect algebraic curves? For a curve $C = \{F(\underline{Z}) = 0\}$ of degree d, points in $T(A)C$ are of the form $T(A)\underline{\mathfrak{z}}$ for $\underline{\mathfrak{z}} \in C$. These are precisely the solutions of the equation

$$\{F\left(T(A^{-1})(\cdot)\right) = 0\} \; (= T(A)C) \tag{4.1.2}$$

since then

$$F(T(A^{-1})T(A)\underline{\mathfrak{z}}) = F(T(A^{-1}A)\underline{\mathfrak{z}}) = F(\underline{\mathfrak{z}}) = 0.$$

Since (4.1.2) just substitutes linear forms[4] for $Z_0, \ldots, Z_n$ in the equation for C, we find:

PROPOSITION 4.1.3. *The images of (smooth resp. singular) algebraic curves of degree d under projectivities, are again (smooth resp. singular) algebraic curves of degree d.*

So lines are carried to lines, conics to conics, and so on. In general, if $T(A)C = C'$ for some A, then the curves C, C' are said to be *projectively equivalent*.

PROOF. To see why smoothness is preserved, write $\tilde{F}(\underline{\xi}) = F(T(A^{-1})\underline{\xi})$ (where we have in mind $[\underline{\xi}] = T(A)[\underline{Z}]$); and suppose (for a contradiction) that for some $\underline{\mathfrak{z}} \in C$ we have $\frac{\partial F}{\partial Z_0}(\underline{\mathfrak{z}}) \neq 0$ but $\frac{\partial \tilde{F}}{\partial \xi_i}(T(A)\underline{\mathfrak{z}}) = 0$ $(\forall i)$.

[4]i.e. homogeneous polynomials of degree 1 in $Z_0, \ldots, Z_n$

If $\tilde{F} = F \circ T(A^{-1})$, then $F = \tilde{F} \circ T(A)$, and by the (multivariable) chain rule

$$\frac{\partial F}{\partial Z_0} = \sum_i \frac{\partial T(A)_i}{\partial Z_0} \frac{\partial \tilde{F}}{\partial \xi_i} = \sum_i a_{i0} \frac{\partial \tilde{F}}{\partial \xi_i},$$

so that

$$0 \neq \frac{\partial F}{\partial Z_0}(\underline{\mathfrak{z}}) = \sum_i a_{i0} \frac{\partial \tilde{F}}{\partial \xi_i}(T(A)\underline{\mathfrak{z}}) = 0,$$

a contradiction. □

Next we want to get formulas for the effect of projectivities on lines and conics. Given $L_\lambda = \{ {}^t\underline{\lambda} \cdot \underline{Z} = 0\}$, (4.1.2) gives $0 = {}^t\underline{\lambda} A^{-1} \underline{Z} = {}^t({}^tA^{-1}\underline{\lambda})\underline{Z}$, so that

$$T(A)L_{\underline{\lambda}} = L_{({}^tA^{-1}\underline{\lambda})}. \tag{4.1.3}$$

Since $GL(3, \mathbb{C})$ acts transitively on $\mathbb{C}^3$, this implies the (relatively trivial)

PROPOSITION 4.1.4. *All lines in $\mathbb{P}^2$ are projectively equivalent.*

Let

$$Q = \{0 = aZ_0^2 + bZ_1^2 + cZ_2^2 + dZ_0Z_1 + eZ_0Z_2 + fZ_1Z_2\}$$

be an arbitrary conic. We can rewrite the equation

$$0 = \begin{pmatrix} Z_0 & Z_1 & Z_2 \end{pmatrix} \begin{pmatrix} a & \frac{d}{2} & \frac{e}{2} \\ \frac{d}{2} & b & \frac{f}{2} \\ \frac{e}{2} & \frac{f}{2} & c \end{pmatrix} \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix} =: {}^t\underline{Z}\mathcal{B}\underline{Z}$$

in terms of a (unique) *symmetric*[5] matrix $\mathcal{B}$. (The expression ${}^t\underline{Z}\mathcal{B}\underline{Z}$ is called a symetric bilinear form.) Given such a $Q_{\mathcal{B}}$, (4.1.2) substitutes in $A^{-1}\underline{Z}$ for $\underline{Z}$, yielding

$$0 = {}^t(A^{-1}\underline{Z})\mathcal{B}A^{-1}\underline{Z} = {}^t\underline{Z}({}^tA^{-1}\mathcal{B}A^{-1})\underline{Z}$$

so that

$$T(A)Q_{\mathcal{B}} = Q_{({}^tA^{-1}\mathcal{B}A^{-1})}.$$

Given an invertible complex matrix M, the transformation $\mathcal{B} \mapsto {}^tM\mathcal{B}M =: \mathcal{B}'$ is called a *cogredience*, and $\mathcal{B}$, $\mathcal{B}'$ are *cogredient over* $\mathbb{C}$. All nonzero

[5] i.e. the transpose ${}^t\mathcal{B}$ equals $\mathcal{B}$

symmetric matrices are cogredient $/\mathbb{C}$ to one of the form[6]

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ or } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We conclude:

PROPOSITION 4.1.5. *All conics in* $\mathbb{P}^2$ *are projectively equivalent to one of*

$$\{X^2 = 0\},\ \{X^2 + Y^2 = 0\},\ or\ \{X^2 + Y^2 + Z^2 = 0\}.$$

Notice that $X^2 + Y^2 = (X + \sqrt{-1}Y)(X - \sqrt{-1}Y)$ is a pair of lines, and so projectively equivalent to $XY = 0$.

COROLLARY 4.1.6. *(i) All smooth*[7] *conics are projectively equivalent.*

(ii) $Q_{\mathcal{B}}$ *is smooth* $\iff \det \mathcal{B} \neq 0$.

PROOF. (i) Since $\{X^2 + Y^2 + Z^2 = 0\}$ is the only smooth option in Prop. 4.1.5, by Prop. 4.1.3 all smooth conics must be equivalent to this hence to each other.

(ii) Cogredience $\mathcal{B} \mapsto {}^t M \mathcal{B} M$ multiplies determinant by $(\det M)^2$, which is always nonzero (as $M \in GL(3, \mathbb{C})$); so projectivities preserve non-zero-ness of $\det \mathcal{B}$. $\square$

4.2. Tangent lines

Let $C = \{F(Z_0, Z_1, Z_2) = 0\}$ be a projective algebraic curve, and suppose

$$\mathbb{R} \supset (-\epsilon, \epsilon) \to C$$
$$t \mapsto \underline{f}(t) = [Z_0(t) : Z_1(t) : Z_2(t)]$$

is a differentiable path segment in C. Then,

$$0 = (F \circ f)'(0) = \sum_{i=0}^{2} \frac{\partial F}{\partial Z_i}(f(0)) \cdot \frac{dZ_i}{dt}(0)$$

[6]This is Sylvester's theorem. For an easy proof of the *real* version, see pp. 161-162 of my linear algebra book. (To get rid of the "-1" entries, hence arrive at the simpler *complex* version, just multiply the relevant basis vectors by $\sqrt{-1}$.)

[7]or equivalently, irreducible

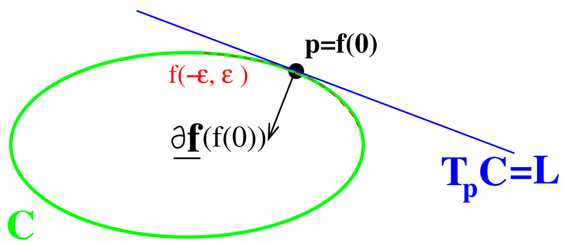
$$= \begin{pmatrix} \frac{\partial F}{\partial Z_0}(f(0)) & \frac{\partial F}{\partial Z_1}(f(0)) & \frac{\partial F}{\partial Z_2}(f(0)) \end{pmatrix} \begin{pmatrix} Z_0'(0) \\ Z_1'(0) \\ Z_2'(0) \end{pmatrix}$$

$$= {}^t\underline{\partial F}(f(0)) \cdot \underline{f}'(0),$$

and so the line $L_{\underline{\partial F}(f(0))}$ contains all tangent vectors $\underline{f}'(0)$ to all such paths in C through $f(0)$.

There is one catch: if the gradient vector $\underline{\partial F}(f(0)) = \underline{0}$, then it does not define a line at all. So we must ask C to be smooth at $f(0)$ for this computation to work.

Definition 4.2.1. The tangent line T_pC to a curve $C = \{F = 0\} \subset \mathbb{P}^2$ at a smooth point $p = [p_0 : p_1 : p_2] \in C$ is $L_{\underline{\partial F}(p)}$.



The next proposition makes the intuitively obvious statement that "projectivities respect tangent lines":

Proposition 4.2.2. *If L is the tangent line to C at p, then $T(A)L$ is the tangent line to $T(A)C$ at $T(A)p$.*

Proof. We must show

$$T(A)L_{\underline{\partial F}(p)} = L_{\underline{\partial(F \circ T(A^{-1}))}(T(A)p)}.$$

Writing $\tilde{F} = F \circ T(A)^{-1}$, this is equivalent to

$$T(A)L_{\underline{\partial(\tilde{F} \circ T(A))}(p)} = L_{\underline{\partial \tilde{F}}(T(A)p)}$$

hence to

$$L_{{}^tA^{-1}\underline{\partial(\tilde{F} \circ T(A))}(p)} = L_{\underline{\partial \tilde{F}}(T(A)p)}$$

or

$$\underline{\partial(\tilde{F} \circ T(A))}(p) \equiv {}^tA\underline{\partial \tilde{F}}(T(A)p) \tag{4.2.1}$$

where $\equiv$ means up to multiplication by $\mathbb{C}^*$. As you may wish to check by writing everything out, *equality* of both sides of (4.2.1) is just an expression of the chain rule. $\square$

Now, the tangent line to a *line* (at any point) is the line itself; for conics the story is less trivial. First we write

$$F(\underline{Z}) = {}^t\underline{Z}\mathcal{B}\underline{Z},$$

$\mathcal{B}$ symmetric, and compute the gradient: writing $\underline{e}_0, \underline{e}_1, \underline{e}_2$ for the standard basis vectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$,

$$\underline{\partial F}(\underline{Z}) = \begin{pmatrix} \partial F/\partial Z_0 \\ \partial F/\partial Z_1 \\ \partial F/\partial Z_2 \end{pmatrix} = \begin{pmatrix} {}^t\underline{Z}\mathcal{B}\underline{e}_0 + {}^t\underline{e}_0\mathcal{B}\underline{Z} \\ {}^t\underline{Z}\mathcal{B}\underline{e}_1 + {}^t\underline{e}_1\mathcal{B}\underline{Z} \\ {}^t\underline{Z}\mathcal{B}\underline{e}_2 + {}^t\underline{e}_2\mathcal{B}\underline{Z} \end{pmatrix}$$

$$= 2\begin{pmatrix} {}^t\underline{e}_0\mathcal{B}\underline{Z} \\ {}^t\underline{e}_1\mathcal{B}\underline{Z} \\ {}^t\underline{e}_2\mathcal{B}\underline{Z} \end{pmatrix} = 2\mathcal{B}\underline{Z}.$$

(Here ${}^t\underline{Z}\mathcal{B}\underline{e}_i = {}^t({}^t\underline{e}_i\,{}^t\mathcal{B}\underline{Z}) = {}^t({}^t\underline{e}_i\mathcal{B}\underline{Z}) = {}^t\underline{e}_i\mathcal{B}\underline{Z}$ uses the fact that $\mathcal{B}$ is symmetric and ${}^t\underline{e}_i\mathcal{B}\underline{Z}$ is "1×1", i.e. a scalar.)

PROPOSITION 4.2.3. *The tangent line to $Q_{\mathcal{B}}$ at $[\underline{p}] \in Q_{\mathcal{B}}$ is $L_{\mathcal{B}\underline{p}}$.*[8]

4.3. The dual projective plane

Suppose we have a vector space $V/\mathbb{C}$, a basis $\mathbf{e} = \{\underline{e}_i\}$ of V and an invertible linear transformation $\mathcal{T} : V \to V$ with matrix $[\mathcal{T}]_{\mathbf{e}} =: M$. Recall that the dual of V is the vector space

$$\check{V} := Hom(V, \mathbb{C})$$

of linear functionals ($f : V \to \mathbb{C}$); one has the tautological pairing

$$\check{V} \times V \xrightarrow{\langle\cdot,\cdot\rangle} \mathbb{C} \tag{4.3.1}$$

given by $\langle f, \underline{v}\rangle = f(\underline{v})$. We have a dual basis $\mathbf{e}^* = \{\underline{e}_i^*\}$ ($\langle \underline{e}_i^*, \underline{e}_j\rangle = \delta_{ij}$), with respect to which one has

$$\langle f, \underline{v}\rangle = \underbrace{{}^t([f]_{\mathbf{e}^*})[\underline{v}]_{\mathbf{e}}}_{\text{matrix multiplication}}.$$

Finally, there is a dual transformation $\check{\mathcal{T}} : \check{V} \to \check{V}$ defined by

$$\langle \check{\mathcal{T}} f, \mathcal{T}\underline{v}\rangle = \langle f, \underline{v}\rangle$$

[8]here we are treating $\underline{p}$ as a column vector (which is consistent with earlier notation)

with matrix

$$[\check{\mathcal{T}}]_{\mathbf{e}^*} = {}^t M^{-1}.$$

This gives a more conceptual way to look at the story of lines in $\mathbb{P}^2$ above: put $V = \mathbb{C}^3$, $[\mathcal{T}]_{\mathbf{e}} = A$, $f \in \check{V}$, $\underline{\lambda} = [f]_{\mathbf{e}^*}$ and so on. Of course $\mathbb{C}^3 \cong \check{\mathbb{C}}^3$ as vector spaces, but we want to keep them conceptually separate.

The crucial point is to projectivize V *and* $\check{V}$: writing

$$\mathbb{P}^2 = \frac{\mathbb{C}^3 \backslash \{0\}}{\mathcal{C}^*}\ , \quad \check{\mathbb{P}}^2 = \frac{\check{\mathbb{C}}^3 \backslash \{0\}}{\mathbb{C}^*},$$

we see that lines in $\mathbb{P}^2$ correspond to points $[\underline{\lambda}] \in \check{\mathbb{P}}^2$. In fact, since the notion of duality is defined by (4.3.1), it is symmetric: $\check{\check{V}} = V$, and so points in $\mathbb{P}^2$ correspond to lines in $\check{\mathbb{P}}^2$. This entire correspondence is invariant under projectivities provided one operates simultaneously on $\mathbb{P}^2$ with $T(A)$ and $\check{\mathbb{P}}^2$ with $T({}^t A^{-1})$. A bit more formally, then:

DEFINITION 4.3.1. The dual projective plane $\check{\mathbb{P}}^2$ is the space of lines in $\mathbb{P}^2$.

Now write

$$\underline{p} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix}\ , \quad \underline{\lambda} = \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{pmatrix}$$

for column vectors. Though this is maybe a little awkward, here is how I want to standardize notation:

$$p = [\underline{p}] = [p_0 : p_1 : p_2] \in \mathbb{P}^2$$

$$\lambda = [{}^t\underline{\lambda}] = [\lambda_0 : \lambda_1 : \lambda_2] \in \check{\mathbb{P}}^2,$$

in other words, points in $\mathbb{P}^2$ are thought of as *column* vectors and points in $\check{\mathbb{P}}^2$ as row vectors. As above the line $L_{\underline{\lambda}} \subset \mathbb{P}^2$ is defined by the equation

$${}^t\underline{\lambda} \cdot \underline{Z} = 0 \qquad \left(\text{solve for } \underline{Z} = \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix} \right),$$

and we say that its dual $\check{L}_{\underline{\lambda}} = \lambda$. Moreover, p defines a line $L_{\underline{p}} \subset \check{\mathbb{P}}^2$ via

$${}^t\underline{W} \cdot \underline{p} = 0 \qquad \left(\text{solve for } {}^t\underline{W} = \begin{pmatrix} W_0 & W_1 & W_2 \end{pmatrix} \right),$$

and we write $\check{p} = L_{\underline{p}}$.

What about the dual of a configuration of

(a) a point p on a line $L_{(\underline{\lambda})}$ (important in Poncelet)?
(b) a pair of lines L, $\mathcal{L}$ through a point p?
(c) a pair of points p, q on a line L?

For the first one, the equation

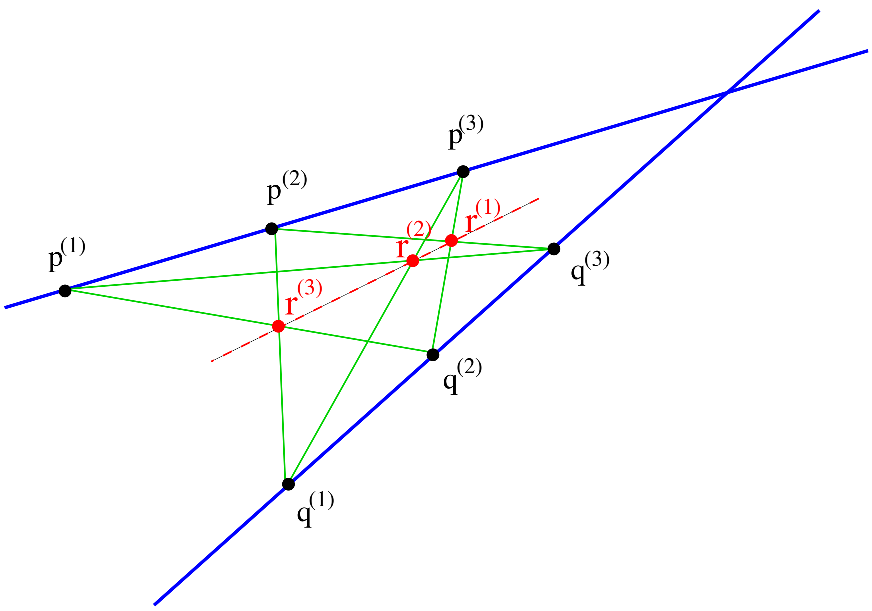
$$^t\underline{\lambda} \cdot \underline{p} = 0$$

expresses "$p \in L_{\underline{\lambda}}$"; but from the above it also expresses $\check{L}_{\underline{\lambda}} \in \check{p}$ (i.e. $\lambda \in L_{\underline{p}}$). Repeating this reasoning, we have

(ǎ) a line $\check{p}$ through a point $\check{L}$;
(b̌) a pair of points $\check{L}$, $\check{\mathcal{L}}$ on a line $\check{p}$;
(č) a pair of lines $\check{p}$, $\check{q}$ through a point $\check{L}$.

Here is something more interesting to dualize, which is left as an exercise for you.

THEOREM 4.3.2. [PAPPUS OF ALEXANDRIA, c. 300 AD] *Let $L, \mathcal{L} \subset \mathbb{P}^2$ be two distinct lines, and write $s = L \cap \mathcal{L}$. On L (resp. $\mathcal{L}$) take distinct $p^{(1)}$, $p^{(2)}$, $p^{(3)}$ (resp. $q^{(1)}$, $q^{(2)}$, $q^{(3)}$) different from s, and set (for $k = 1, 2, 3$) $r^{(k)} := \overline{p^{(i)}q^{(j)}} \cap \overline{p^{(j)}q^{(i)}}$ (where $\{i, j, k\} = \{1, 2, 3\}$). Then $r^{(1)}$, $r^{(2)}$, $r^{(3)}$ are collinear.*



PROOF. In fact, $p^{(1)}q^{(2)}p^{(3)}q^{(1)}p^{(2)}q^{(3)}$ is a hexagon "inscribed" in the conic $L \cup \mathcal{L}$, and the $\{r^{(i)}\}$ are the intercepts of its opposite edges. After changing by a projectivity (which preserves the "figure"), this conic is $XY = 0$, which is obviously the limit of the smooth conic $XY - \alpha Z^2 = 0$ as $\alpha \to 0$. Since Pascal's theorem implies collinearity of the hexagon edge intercepts for all $\alpha \neq 0$, this remains true at $\alpha = 0$. $\square$

4.4. Dual conics and polar lines

DEFINITION 4.4.1. The dual $\check{C} \subset \check{\mathbb{P}}^2$ of a smooth algebraic curve $C = \{F = 0\} \subset \mathbb{P}^2$ is the set of (dual points of) tangent lines to C. That is, $\check{C} = \{\check{T_p C} \in \check{\mathbb{P}}^2 \,|\, p \in C\}$.

This is consistent with our definition for lines. For higher degree curves, however, the dual is not one point: consider the duality map

$$\mathcal{D}_C : \mathbb{P}^2 \longrightarrow \check{\mathbb{P}}^2$$

sending[9]

$$p \longmapsto [{}^t\underline{\partial F}(p)].$$

PROPOSITION 4.4.2. *(a)* $\check{C} = \mathcal{D}_C(C)$.
(b) If $\check{C}$ *is smooth at* $\lambda = \check{T_p C}$*, then* $T_\lambda \check{C} = \check{p}$.

PROOF. (a) For $p \in C$, $T_p C = L_{\underline{\partial F}(p)} \implies \check{T_p C} = [{}^t\underline{\partial F}(p)]$. So this is practically a tautology.

(b) Here we jump into a little deep water. It suffices to show that for any path ${}^t\underline{\lambda}(\cdot) : (-\epsilon, \epsilon) \to \check{C}$ through $\check{T_p C} = [{}^t\underline{\partial F}(p)] = \mathcal{D}_C(p)$,

$$\frac{d\,{}^t\underline{\lambda}}{dt}(0) \cdot \underline{p} = 0. \tag{4.4.1}$$

Since $\check{C}$ is the image of C by $\mathcal{D}_C$, ${}^t\underline{\lambda}(t) = (\mathcal{D}_C \circ \underline{q})(t)$ for some $\underline{q}(\cdot) : (-\epsilon, \epsilon) \to C$ through p. So the left-hand side of (4.4.1) becomes

$$\frac{d}{dt}(\mathcal{D}_C \circ \underline{q})(0) \cdot \underline{p} =$$

(4.4.2)

$$\begin{pmatrix} q_0'(0) & q_1'(0) & q_2'(0) \end{pmatrix} \left.\begin{pmatrix} \frac{\partial^2 F}{\partial Z_0^2} & \frac{\partial^2 F}{\partial Z_0 \partial Z_1} & \frac{\partial^2 F}{\partial Z_0 \partial Z_2} \\ \frac{\partial^2 F}{\partial Z_1 \partial Z_0} & \frac{\partial^2 F}{\partial Z_1^2} & \frac{\partial^2 F}{\partial Z_1 \partial Z_2} \\ \frac{\partial^2 F}{\partial Z_2 \partial Z_0} & \frac{\partial^2 F}{\partial Z_2 \partial Z_1} & \frac{\partial^2 F}{\partial Z_2^2} \end{pmatrix}\right|_p \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix}.$$

[9]i.e., $[Z_0 : Z_1 : Z_2] \mapsto \left[\frac{\partial F}{\partial Z_0}(Z_0, Z_1, Z_2) : \frac{\partial F}{\partial Z_1}(Z_0, Z_1, Z_2) : \frac{\partial F}{\partial Z_2}(Z_0, Z_1, Z_2)\right]$

The matrix in the middle is the *Hessian* of F at p and has nonvanishing determinant if $\check{C}$ is nonsingular at $\mathcal{D}_C(p)$. We will return to the Hessian later in this course.

Now, since each $\frac{\partial F}{\partial Z_i}$ is homogeneous (of degree $d-1$, if $d = \deg(C)$), the Euler formula ((2.1.8), with $\underline{Z}$ set equal to $\underline{p}$) collapses (4.4.2) to

$$(d-1).\begin{pmatrix} q_0'(0) & q_1'(0) & q_2'(0) \end{pmatrix} \begin{pmatrix} \frac{\partial F}{\partial Z_0}(p) \\ \frac{\partial F}{\partial Z_1}(p) \\ \frac{\partial F}{\partial Z_2}(p) \end{pmatrix} =$$

$$(d-1).\underline{\partial F}(\underline{q}(0)) \cdot \frac{d\underline{q}}{dt}(0),$$

which is indeed zero by the beginning of §4.2. $\square$

REMARK 4.4.3. One can show that the dual of a smooth algebraic curve of degree d is an algebraic curve of degree $d(d-1)$; moreover, for $d \geq 2$ this dual is *singular*, so the duality map cannot be "reversed" as defined.

We consider the dual of the conic $Q_{\mathcal{B}}$. By the computation in §4.2 (for $F(\underline{Z}) = {}^t\underline{Z}\mathcal{B}\underline{Z}$), ${}^t\underline{\partial F}(\underline{Z}) = 2\,{}^t\underline{Z}\mathcal{B} \equiv {}^t\underline{Z}\mathcal{B}$ and so

$$\begin{aligned} \mathcal{D}_{Q_{\mathcal{B}}}(Q_{\mathcal{B}}) &= \left\{ [\,{}^t\underline{Z}\mathcal{B} \,|\, [\underline{Z}] \in C] \right\} \\ &= \{[\,{}^t\underline{Z}\mathcal{B}] \,|\, {}^t\underline{Z}\mathcal{B}\underline{Z} = 0\}. \end{aligned}$$

Making the substitition ${}^t\underline{\lambda} = {}^t\underline{Z}\mathcal{B} \longleftrightarrow \underline{Z} = {}^t\mathcal{B}^{-1}\underline{\lambda}$, this becomes

$$\begin{aligned} \check{Q}_{\mathcal{B}} &= \{[\,{}^t\underline{\lambda}] \in \check{\mathbb{P}}^2 \,|\, {}^t\underline{W}\mathcal{B}^{-1}\mathcal{B}\,{}^t\mathcal{B}^{-1}\underline{W} = 0\} \\ &= \{\lambda \in \check{\mathbb{P}}^2 \,|\, {}^t\underline{\lambda}\mathcal{B}^{-1}\underline{\lambda} = 0\} \end{aligned}$$

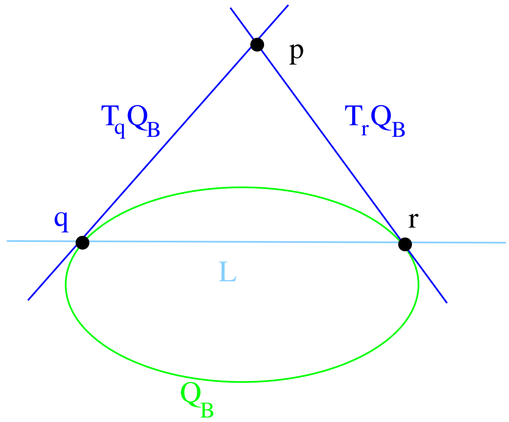
where we have used the fact that $\mathcal{B}$ is symmetric. This gives part (i) of:

PROPOSITION 4.4.4. *(i)* $\check{Q}_{\mathcal{B}} = Q_{\mathcal{B}^{-1}}$*, and* $\check{\check{Q}}_{\mathcal{B}} = Q_{\mathcal{B}}$*. (In particular, the dual of a smooth conic is a* smooth *conic, since* $\det \mathcal{B} \neq 0 \implies \det \mathcal{B}^{-1} \neq 0$*.)*

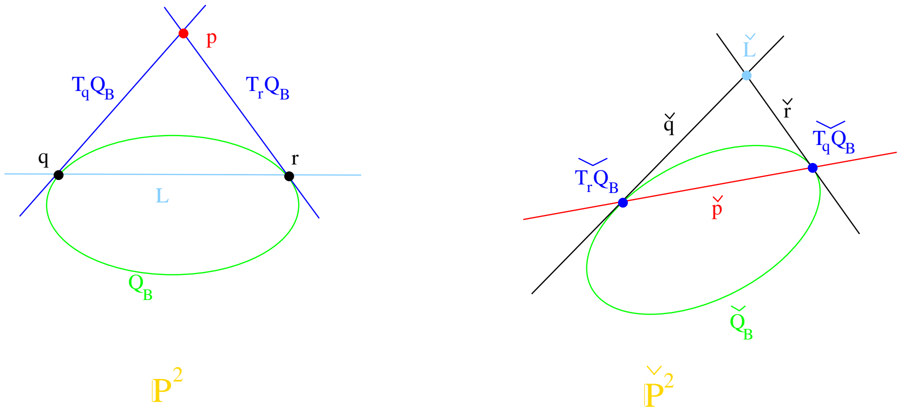
(ii) Given $p \in \mathbb{P}^2 \backslash Q_{\mathcal{B}}$*, there exist exactly two lines through* p *and tangent to* $Q_{\mathcal{B}}$*.*

PROOF. (ii) By Proposition 4.4.2(b), this is *dual* to the statement: if the line $\check{p} \subset \check{\mathbb{P}}^2$ is not tangent to $\check{Q}_{\mathcal{B}}$, then it meets $\check{Q}_{\mathcal{B}}$ in exactly 2 points. This last statement then follows from Proposition 2.1.8. $\square$

Definition 4.4.5. Let $Q_{\mathcal{B}}$ be a smooth conic and p be a point not on $Q_{\mathcal{B}}$, with $T_q Q_{\mathcal{B}}$ and $T_r Q_{\mathcal{B}}$ the two tangent lines to $Q_{\mathcal{B}}$ containing p. (Here $q, r \in Q_{\mathcal{B}}$.) Then the *polar line* $L_{(p,Q_{\mathcal{B}})} \subset \mathbb{P}^2$ of p with respect to $Q_{\mathcal{B}}$ is the line through q and r.



Proposition 4.4.6. *Let $p \in \mathbb{P}^2 \backslash Q_{\mathcal{B}}$, with polar line $L = L_{(r,Q_{\mathcal{B}})}(\subset \mathbb{P}^2)$. Then the polar line $L_{(\check{L},\check{Q}_{\mathcal{B}})} \subset \check{\mathbb{P}}^2$ (of the dual point $\check{L}$ with respect to the dual conic) is $\check{p}$ (the dual line of p). In a picture, where dual objects are the same color:*



Proof. This is an immediate consequence of the rules $(\text{a,b,c}) \longleftrightarrow (\check{\text{a}}, \check{\text{b}}, \check{\text{c}})$, the definition of the dual curve, and Proposition 4.4.2(b). $\square$

Example 4.4.7. $Q = \{-4Z_0^2 + Z_1^2 + Z_2^2 = 0\} \longrightarrow \mathcal{B} = \begin{pmatrix} -4 & & \\ & 1 & \\ & & 1 \end{pmatrix}$.

Let $p = [1:2:2]\ (\notin Q)$ so that the polar line L is given by

$$0 = \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} -4 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix} = -4Z + 2X + 2Y,$$

i.e. $L = L_{\underline{\lambda}}$ where $\lambda = [-4 : 2 : 2]$.

On the dual $\check{\mathbb{P}}^2$ side: $\check{L} = \lambda$; $\check{Q}$ has matrix $\mathcal{B}^{-1} = \begin{pmatrix} -\frac{1}{4} & & \\ & 1 & \\ & & 1 \end{pmatrix}$, hence equation $0 = -\frac{1}{4}Z_0^2 + Z_1^2 + Z_2^2$; and $\check{p}$ is the line $0 = W_0 + 2W_1 + 2W_2$. On the other hand, the polar line of λ with respect to $\check{Q}$ is

$$0 = \begin{pmatrix} -4 & 2 & 2 \end{pmatrix} \begin{pmatrix} -\frac{1}{4} & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} W_0 \\ W_1 \\ W_2 \end{pmatrix} = W_0 + 2W_1 + 2W_2,$$

agreeing with $\check{p}$.

It is instructive to think about what happens to Pascal and Poncelet under duality. While the dual of Poncelet is again just Poncelet (but in $\check{\mathbb{P}}^2$), we *do* find that if polygons inscribed in C and circumscribed about D close up after n sides, then so do the polygons (in $\check{\mathbb{P}}^2$) inscribed in $\check{D}$ and circumscribed about $\check{C}$.

The dual of Pascal, on the other hand, does give a different statement:

PROPOSITION 4.4.8. *The 3 lines through opposite vertices of a hexagon circumscribed about a conic, pass through a single point.*

Proof is basically the same as for Proposition 4.4.6; I'll let you work it out.

Exercises

(1) Given a configuration of 4 points $p, q, r, s \in \mathbb{P}^2$ in "general position", i.e. no three of them collinear, show there exists a unique projectivity sending $p \mapsto [1:0:0]$, $q \mapsto [0:1:0]$, $r \mapsto [0:0:1]$, $s \mapsto [1:1:1]$. [Hint: work with vectors $\underline{p}, \underline{q}, \underline{r}, \underline{s} \in \mathbb{C}^3$. You only have to send $\underline{p}$ (resp. $\underline{q}, \underline{r}, \underline{s}$) to a multiple of $\underline{e}_0$ (resp. $\underline{e}_1$, $\underline{e}_2$, $\underline{e}_0 + \underline{e}_1 + \underline{e}_2$).]

(2) (a) Give a direct proof of Pappus's theorem. [Hint: use the last exercise to first simplify the coordinates of several of the points.]
(b) State a dual version of Pappus's theorem, and draw a figure. [Note: it would be better to state the dual version in $\mathbb{P}^2$: think first of Pappus in $\check{\mathbb{P}}^2$ and then dualize that.]

(3) Show that the equation of the polar line of p with respect to $Q_{\mathcal{B}}$ has equation ${}^t\underline{p}\mathcal{B}\underline{Z} = 0$. Use this to give another (short) proof of Proposition 4.4.6.

(4) Prove that all automorphisms of $\mathbb{P}^1$ (as a complex manifold) are fractional linear transformations. Deduce that $Aut(\mathbb{P}^1) \cong PGL_2(\mathbb{C})$. [Use material from §3.1.]

Part 2

General definitions and results

CHAPTER 5

Complex manifolds and algebraic varieties

In Chapters 2-3 we introduced Riemann surfaces and plane algebraic curves, and stated the Normalization Theorem which produces a strong relation between them. Here we will introduce the arbitrary-dimensional generalizations of these objects. While it is true that an algebraic variety of dimension n has a desingularization[1] which is a complex n-manifold, the converse is false: already there are non-algebraic complex 2-manifolds.

However, it is true that any global analytic object (functions or differential forms, for example) on a projective algebraic variety viewed as a complex manifold, is a algebraic. This is Serre's "GAGA" (global analytic = global algebraic) principle. For example, global meromorphic functions in this context turn out to be nothing but restrictions to the algebraic variety of rational functions on the ambient projective space (elements of $\mathbb{C}(z_1, \ldots, z_n)$).

5.1. Complex n-manifolds

These are the generalization of complex 1-manifolds (or of Riemann surfaces, if we assume compactness) to higher dimension. Once again, we begin with a Hausdorff topological space X with open cover $\{U_\alpha\}$ and write $U_{\alpha\beta} = U_\alpha \cap U_\beta$. This is made into a *complex n-manifold* by the additional data of an analytic atlas on X: that is, a collection of holomorphic coordinates[2]

$$\underline{z_\alpha} : U_\alpha \xrightarrow{\simeq} V_\alpha \subseteq \mathbb{C}^n,$$

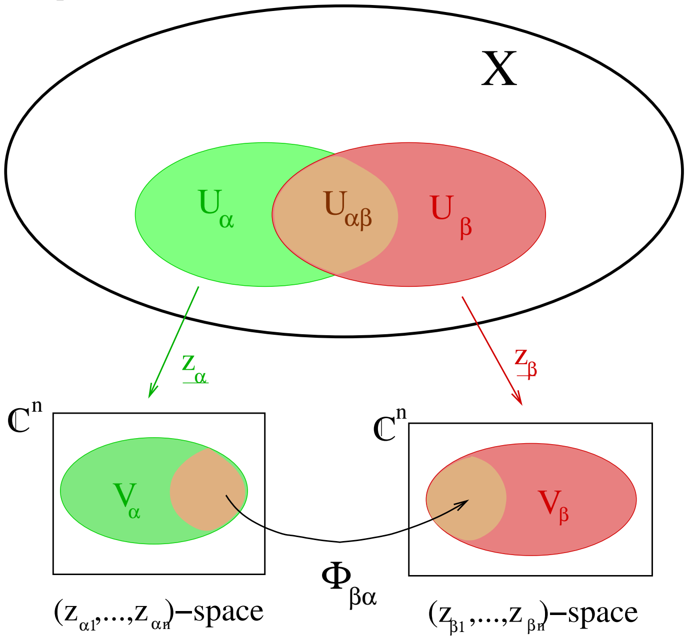
[1]"normalization" is no longer the correct term (refers to a weaker process which still produces a singular object)

[2]As usual, an underline means a vector or tuple of some kind; in this case, $\underline{z_\alpha} = (z_{\alpha 1}, \ldots, z_{\alpha n})$.

or homeomorphisms between U_α and an open set in $\mathbb{C}^n$, such that the transition functions

$$\Phi_{\beta\alpha} := \underline{z}_\beta \circ \underline{z}_\alpha^{-1} : V_\alpha^\beta \to V_\beta^\alpha$$

are biholomorphic.



Here $V_\alpha^\beta := \underline{z}_\alpha(U_{\alpha\beta})$ and $V_\beta^\alpha := \underline{z}_\beta(U_{\alpha\beta})$ are open subsets of $\mathbb{C}^n$, and we need to explain what biholomorphic means. First, a function $f : \mathbb{C}^n \to \mathbb{C}$ is holomorphic if and only if it looks locally (about each point) like $f(\underline{z}) = \sum_{I=(i_1,\ldots,i_n)} a_I \underline{z}^I$ where $\underline{z}^I$ means $z_1^{i_1} z_2^{i_2} \cdots z_n^{i_n}$ and $a_I \in \mathbb{C}$ are constants. A holomorphic map $\mathbb{C}^n \to \mathbb{C}^m$ is just m of these: $(z_1, \ldots z_n) = \underline{z} \mapsto (f_1(\underline{z}), \ldots, f_m(\underline{z}))$. (Since these definitions are local, they immediately have meaning when $\mathbb{C}^n$ etc. are replaced by open sets.) Finally, "biholomorphic" simply indicates a bijective map ($\Phi_{\beta\alpha}$ is bijective by construction) which is holomorphic in each direction.

To generalize the morphisms of Riemann surfaces introduced in §3.1, we can define a *morphism* $F : X \to Y$ *of complex manifolds.* Here X and Y need *not* be of the same dimension; for X we keep the above notation and for Y write $\underline{\mathfrak{z}}_\gamma : \mathcal{U}_\gamma \xrightarrow{\simeq} \mathcal{V}_\gamma \subseteq \mathbb{C}^m$. A morphism F is then a collection of continuous functions $F_\alpha : U_\alpha \to Y$ (agreeing on the $U_{\alpha\beta}$) such that each composition

$$\underline{\mathfrak{z}}_\gamma \circ F_\alpha \circ \underline{z}_\alpha^{-1} : \underline{z}_\alpha \left(F_\alpha^{-1}(\mathcal{U}_\gamma \cap F_\alpha(U_\alpha))\right) \to \underline{\mathfrak{z}}_\gamma \left(\mathcal{U}_\gamma \cap F_\alpha(U_\alpha)\right)$$

yields a holomorphic map (from a subset of $V_\alpha \subseteq \mathbb{C}^n$ to a subset of $\mathcal{V}_\gamma \subseteq \mathbb{C}^m$). If $n = m = 1$ then this reproduces Definition 3.1.6.[3] Moreover, compositions of morphisms are morphisms.

Basic examples of complex manifolds include (besides Riemann surfaces when $n = 1$) Cartesian products of Riemann surfaces, complex n-tori

$$\mathbb{C}^n/\mathbb{Z}\left\langle \underline{\lambda_1}, \cdots, \underline{\lambda_{2n}} \right\rangle$$

(where $\underline{\lambda_1}, \ldots, \underline{\lambda_{2n}}$ are linearly independent over $\mathbb{R}$ in $\mathbb{C}^n \cong \mathbb{R}^{2n}$), and projective n-space

$$\mathbb{P}^n := \frac{\mathbb{C}^{n+1} \backslash \{\underline{0}\}}{\langle (\xi_0, \ldots, \xi_n) \sim (\gamma\xi_0, \ldots, \gamma\xi_n) \ \ \forall \gamma \in \mathcal{C}^* \rangle}.$$

Demonstrating that $\mathbb{P}^n$ is a complex n-manifold (as we do in the next section) immediately gives meaning to a "morphism of complex manifolds from a Riemann surface to $\mathbb{P}^n$." This notion is equivalent to (but more intrinsic than) Definition 3.1.9, as we shall see.

5.2. $\mathbb{P}^n$ as a complex manifold

$\mathbb{P}^n$ is covered by the open sets $U_i := \{\xi_i \neq 0\}$, with local coordinates

$$\underline{z_i} = (z_{i1}, \ldots, z_{in}) := \left(\frac{\xi_0}{\xi_i}, \ldots, \widehat{\frac{\xi_i}{\xi_i}}, \ldots, \frac{\xi_n}{\xi_i} \right) : \ U_i \xrightarrow{\cong} \mathbb{C}^n.$$

(Here "i" replaces "α", $V_i = \mathbb{C}^n$, and $\widehat{(\cdot)}$ means to omit that term.) We need to check that the transition functions

$$\Phi_{ji} : \ V_i^j \to V_j^i$$

are holomorphic. Now, Φ_{ji} tells us how to write the $\underline{z_j} = (z_{j1}, \ldots, z_{jn})$ as functions of the $\underline{z_i} = (z_{i1}, \ldots, z_{in})$ in such a way that

$$\left(\frac{\xi_0}{\xi_i}, \ldots, \widehat{\frac{\xi_i}{\xi_i}}, \ldots, \frac{\xi_j}{\xi_i}, \ldots, \frac{\xi_n}{\xi_i} \right) \text{ is sent to } \left(\frac{\xi_0}{\xi_j}, \ldots, \frac{\xi_i}{\xi_j}, \ldots, \widehat{\frac{\xi_j}{\xi_j}}, \ldots, \frac{\xi_n}{\xi_j} \right)$$

(where for convenience we assume $j > i$). Now, $V_i^j \subset \mathbb{C}^n$ is simply the subset where $z_{ij} \neq 0$. So the correct transition function is

$$\Phi_{ji}(z_{i1}, \ldots, z_{in}) = (z_{j1}(z_{i1}, \ldots, z_{in}), \ldots, z_{jn}(z_{i1}, \ldots, z_{in}))$$

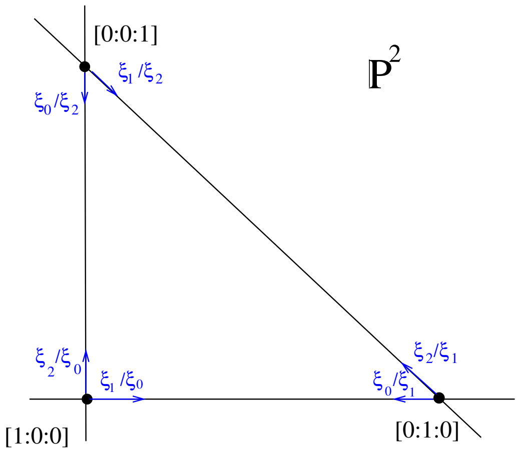
[3] it might be a good idea to glance back at the picture there (for intuition purposes)

where

$$z_{jk}(z_{i1},\ldots,z_{in}) = \begin{cases} z_{ik}/z_{ij}, & \text{for } k \leq i,\ k > j \\ z_{i,k-1}/z_{ij}, & \text{for } i+1 < k \leq j \\ 1/z_{ij}, & \text{for } k = i+1 \end{cases} . \tag{5.2.1}$$

For $\mathbb{P}^1$, $\underline{z_i}$ reduces to z_i ($i = 0,1$). More precisely, $z_0 = \frac{\xi_1}{\xi_0}$ and $z_1 = \frac{\xi_0}{\xi_1}$ are the two local coordinates, while (5.2.1) becomes $z_1(z_0) = \frac{1}{z_0}$, so that we recover Example 2.2.3. Here is a "schematic picture" of the local coordinates on $\mathbb{P}^1$:

For $\mathbb{P}^2$, we have $\underline{z_0} = (z_{01}, z_{02}) = \left(\frac{\xi_1}{\xi_0}, \frac{\xi_2}{\xi_0}\right)$, $\underline{z_1} = (z_{11}, z_{12}) = \left(\frac{\xi_0}{\xi_1}, \frac{\xi_2}{\xi_1}\right)$, and $\underline{z_2} = (z_{21}, z_{22}) = \left(\frac{\xi_0}{\xi_2}, \frac{\xi_1}{\xi_2}\right)$, with e.g. $\Phi_{20}\left(z_{01}, z_{02}\right) = \left(\frac{1}{z_{02}}, \frac{z_{01}}{z_{02}}\right)$. Again, the local coordinates can be visualized as follows:

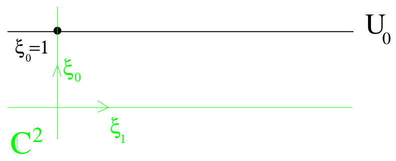


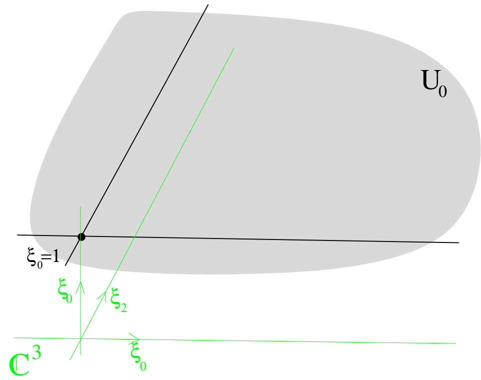
So, for instance, the coordinates $\underline{z_1} = \left(\frac{\xi_0}{\xi_1}, \frac{\xi_2}{\xi_1}\right)$ are defined on the complement U_1 of the vertical line, and both vanish at $[0:1:0]$.

Remark 5.2.1. Whenever you have a local holomorphic coordinate (system) like $\underline{z_i} : U_i \to V_i \subseteq \mathbb{C}^n$, the inverse mapping $\varphi_i = \underline{z_i}^{-1} : V_i \xrightarrow{\simeq} U_i \subset X$ (or just $V_i \hookrightarrow X$) is called a local analytic chart. In case $X = \mathbb{P}^n$, $\varphi_i : \mathbb{C}^n \hookrightarrow \mathbb{P}^n$ is given by

$$\varphi_i(z_{i1},\ldots,z_{in}) = [z_{i1} : \cdots : z_{ii} : 1 : z_{i,i+1} : \cdots : z_{in}]$$

, and one can visualize this as a map from $\mathbb{C}^n \hookrightarrow (\mathbb{C}^{n+1} \backslash \{\underline{0}\}) \twoheadrightarrow \mathbb{P}^n$. Here are pictures of the image of φ_0 for $\mathbb{P}^1$ and $\mathbb{P}^2$:





The next statement says that the notion of "holomorphic map" from a Riemann surface to projective space (Defn. 3.1.9) is just a special case of "morphism of complex manifolds". It is enough to consider (as we do) the situation where the image is not contained in a coordinate hyperplane, since these are just smaller-dimensional projective spaces (included into $\mathbb{P}^n$ by morphisms).

PROPOSITION 5.2.2. *Let M be a complex 1-manifold, and consider a continuous mapping $F : M \to \mathbb{P}^n$ with $F(M)$ not contained in any $\{Z_i = 0\}$. The following statements are then equivalent:*

(i) F is a morphism of complex manifolds;

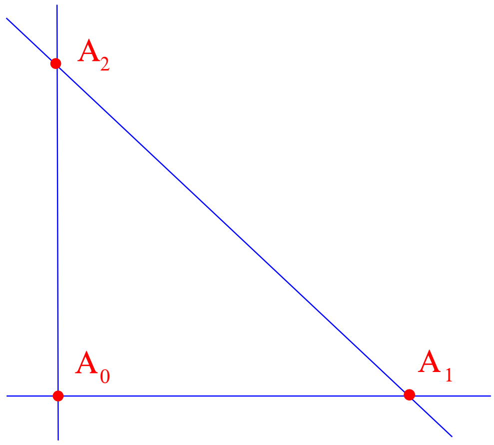
(ii) each composition $[\xi_i \circ F : \xi_j \circ F]$ is a morphism of complex manifolds to $\mathbb{P}^1$ (on the open subset of M where it is well-defined);

(iii) each $\frac{\xi_i}{\xi_j} \circ F$ gives a meromorphic function on M.

PROOF. First, write $U_i := \{Z_i \neq 0\} \subset \mathbb{P}^n$ as above. For each $\{i, j\}$ (where $i \neq j$), the projections $\pi_{ij} : U_i \cup U_j \twoheadrightarrow \mathbb{P}^1$ defined by $[\xi_0 : \cdots : \xi_n] \mapsto [\xi_i : \xi_j]$ are morphisms of complex manifolds. So if $M \xrightarrow{F} \mathbb{P}^n$ is one, then $\pi_{ij} \circ F = [\xi_i \circ F : \xi_j \circ F]$ is one too, showing $(i) \implies (ii)$. Next, $(ii) \implies (iii)$ is Remark 3.1.10. Finally, if all the $\frac{\xi_j}{\xi_i} \circ F$ give meromorphic functions on all of M, then in particular $\underline{z_i} \circ F = \left(\frac{\xi_0}{\xi_i} \circ F, \ldots, \widehat{\frac{\xi_i}{\xi_i} \circ F}, \ldots, \frac{\xi_j}{\xi_i} \circ F, \ldots, \frac{\xi_n}{\xi_i} \circ F \right)$ is holomorphic on $F^{-1}(U_j)$ (or a suitable covering of it by coordinate neighborhoods). These give the local holomorphic representations of F required for a morphism, proving $(iii) \implies (i)$. $\square$

We will refine Proposition 5.2.2 in Chapter 7 below.

Whilst we are dwelling on the subject of projective space, I would like to mention (just for $\mathbb{P}^2$) a trick for drawing the *real* solution sets of homogeneous equations on the page: *barycentric coordinates.* First draw 3 points $A^{(0)}$, $A^{(1)}$, $A^{(2)}$ on a piece of paper:

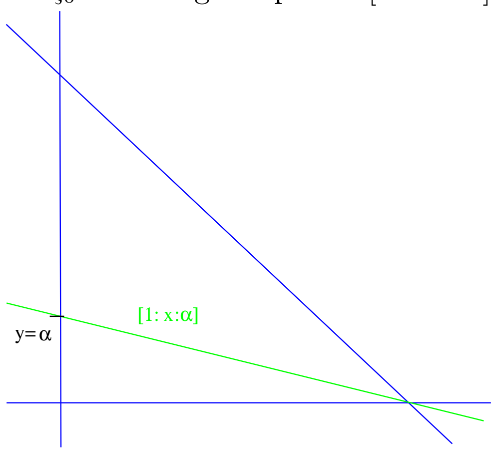


Think of these as vectors $\underline{A}^{(i)} \in \mathbb{R}^2$; it doesn't matter where the origin is. Now, plot $[\xi_0 : \xi_1 : \xi_2]$ as

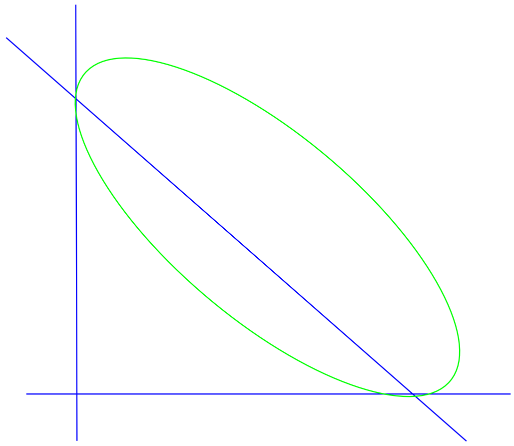
$$\sum_{i=0}^{2} \left(\frac{\xi_i}{\sum_{j=0}^{2} \xi_j} \right) \underline{A}^{(i)}. \tag{5.2.2}$$

To "draw" an algebraic curve, simply find all the solutions $[\xi_0 : \xi_1 : \xi_2]$ with $\xi_i \in \mathbb{R}$, and use (5.2.2) to plot them.

EXAMPLE 5.2.3. (i) The line $y = \alpha$ (assume $\alpha \in \mathbb{R}$) projectively completes to $\xi_2 = \alpha\xi_0$. Plotting the points $[1 : x : \alpha]$ in this way gives

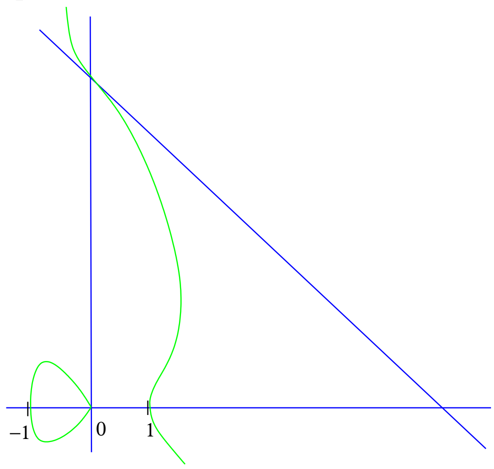


(ii) The conic $xy = 1$ completes to $\xi_1\xi_2 = (\xi_0)^2$, and we get



for its real barycentric plot.

(iii) The cubic curve $y^2 = x(x+1)(x-1)$ becomes $(\xi_2)^2\xi_0 = \xi_1(\xi_1 + \xi_0)(\xi_1 - \xi_0)$ with picture



In fact, this is the precise meaning of the "schematic" real 1-dimensional pictures of complex algebraic curves we have been drawing and will continue to draw — we are plotting the real solutions in barycentric coordinates.

5.3. Affine and projective algebraic varieties

We are going to approach this from a slightly more algebraic angle than, "take the common solution of a bunch of polynomial equations". Start with the commutative ring $S_n := \mathbb{C}[z_1, \ldots, z_n]$ of polynomials in n variables.

Let $J \subset S_n$ be an ideal. The *affine variety* associated to J is

$$V(J) := \{\underline{z} = (z_1, \ldots, z_n) \,|\, f(\underline{z}) = 0 \;\; \forall f \in J\} \subseteq \mathbb{C}^n,$$

which is the vanishing locus of all polynomials in J. By a result in algebra known as *Hilbert's basis theorem*, any ideal in S_n is finitely generated, that is, of the form $(f_1, \dots, f_k)$; consequently $V(J)$ is simply of the form $f_1(\underline{z}) = \cdots = f_k(\underline{z}) = 0$. However, working in terms of ideals does have a payoff, in the form of the famous "theorem on zeroes" or *nullstellensatz*:

THEOREM 5.3.1. [D. HILBERT, 1893] *If $g \in S_n$ vanishes identically on $V(J)$, then for some $m \in \mathbb{N}$,g^m belongs to J.*

If $J = (f)$, then this just says "if g vanishes (in $\mathbb{C}^n$) wherever f does, then f divides some power of g."

Next we consider the projective case, writing $S_{n+1} = \mathbb{C}[Z_0, \dots, Z_n]$. Its underlying additive group can be viewed as the direct sum $\oplus_d S^d_{n+1}$, where S^d_{n+1} denotes homogeneous polynomials of degree d in $n+1$ variables. Hence, any polynomial G can be written uniquely as a finite sum of homogeneous terms $\sum_d G_d$.

DEFINITION 5.3.2. An ideal $I \subset S_{n+1}$ is *homogeneous* if and only if the condition

$$G \in I \quad \Longrightarrow \quad G_d \in I \ \ (\forall d)$$

is satisfied.

The *projective variety* associated to a homogeneous ideal $I \subset S_{n+1}$is[4]

$$\bar{V}(I) := \{[\underline{Z}] = [Z_0 : \cdots : Z_n] \mid F(\underline{Z}) = 0 \ \forall F \in I\} \subseteq \mathbb{P}^n.$$

A version of the Nullstellensatz suited to this case, which is an immediate consequence of Theorem 3.1, is:

COROLLARY 5.3.3. *Given a homogeneous polynomial g vanishing on all of $\bar{V}(I)$, some power of g belongs to I.*

REMARK 5.3.4. If $F_1, \dots, F_k$ are homogeneous polynomials (of various degrees), then

(i) $I := (F_1, \dots, F_k)$ is a homogeneous ideal; and

(ii) $\bar{V}(I) = \{F_1(\underline{Z}) = \cdots = F_k(\underline{Z}) = 0\}$.

---

[4]technically, one should keep track of multiplicities of irreducible components, rather than just defining $\bar{V}(I)$ as a set. For the most part we will suppress this detail.

As in the case of curves, we want to be able to go between the affine and projective settings. To "restrict" a projective variety to the affine world, start with the ring (or algebra) homomorphism

$$S_{n+1} \to S_n$$

induced by

$$F(Z_0, Z_1, \ldots, Z_n) \mapsto F(1, z_1, \ldots, z_n).$$

If we write I° for the image of a homogeneous ideal I under this map, then

$$V(I^\circ) = \bar{V}(I) \cap \mathbb{C}^n.$$

To go the other way, recall the space $\mathcal{P}_n^d = \oplus_{j=1}^d S_n^j$ of polynomials of degree at most d. We have a homomorphism of abelian groups (or vector spaces)

$$\mathcal{P}_n^d \xrightarrow{\theta_d} S_{n+1}^d$$

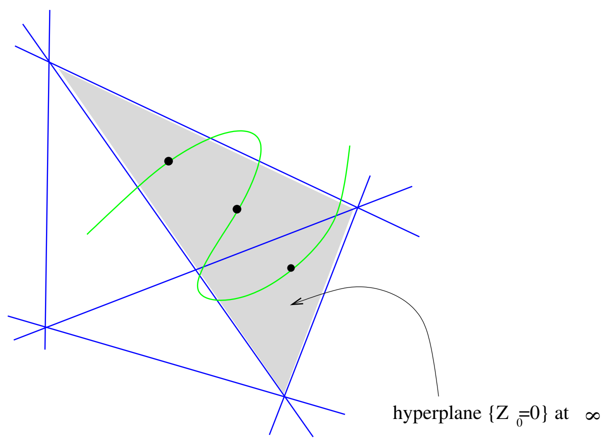
which is defined by

$$f(\underline{z}) \longmapsto (Z_0)^d f\left(\underline{Z}/Z_0\right).$$

Now, take $J \subset S_n$ an ideal. Given generators $\{f_i\}_{i=1}^k$ (f_i of degree d_i) for J, so that $J = (f_1, \ldots, f_k)$, set $\bar{J} := (\theta_{d_1}(f_1), \ldots, \theta_{d_k}(f_k))$. Then we have

$$\bar{V}(\bar{J}) \cap \mathbb{C}^n = V(J).$$

So if $n = 3$, then $\bar{V}(\bar{J})$ is adding stuff in the "$\mathbb{P}^2$ at infinity" $\{Z_0 = 0\}$ to complete your affine variety to a projective one, as suggested by the picture:



in which the black points get added in the process of completing the curve.

EXAMPLE 5.3.5. A key example of affine or projective varieties are the *hypersurfaces* cut out of $\mathbb{C}^n$ or $\mathbb{P}^n$ by a single equation. Let $F \in S^d_{n+1}$ and[5]

$$X = V(F) = \{F(\underline{Z}) = 0\} \subseteq \mathbb{P}^n$$

the corresponding projective hypersurface of degree d. (Like algebraic curves, these are called linear, quadric, cubic, quartic, quintic, etc. according as $d = 1, 2, 3, 4, 5, \ldots$.) We will define dimension rigorously below, but X is $(n-1)$-dimensional (and thus of *codimension* 1).

Now since S_n is a unique factorization domain, we can factor

$$F = \prod_{i=1}^{k} F_i^{m_i},$$

uniquely (up to order), where each F_i is prime (irreducible). We can then write unambiguously

$$X = \sum_{i=1}^{k} m_i X_i,$$

and say X is *reduced* when all $m_i = 1$, and *irreducible* when $k = 1$.

Exercises

(1) Show that each projection $\pi_{ij} : U_i \cup U_j \twoheadrightarrow \mathbb{P}^1$ described in the proof of Prop. 5.2.2 is a morphism of complex manifolds. [This is a quick one.]

(2) Sketch the real solutions of [the projective closure of] $\{y^2 = \prod_{i=1}^{2g+2}(x - a_i)\}$ in $\mathbb{P}^2$, if $a_1 < a_2 < \cdots < a_{2g+2}$ are real numbers.

[5]here we really mean $V((F))$, the variety of the ideal (F), but we shorten this to $V(F)$.

CHAPTER 6

More on projective algebraic varieties

We warm up with two examples we can get our hands on immediately: linear varieties and quadric hypersurfaces. Then we launch into what it means for an algebraic variety to be singular resp. smooth at a point, and in the latter case introduce its tangent space at that point (which is a linear variety). This leads to a careful definition of *dimension* for algebraic varieties. As a sort of "appendix" I'll give a long-overdue introduction to plane curve singularities (which were glossed over in Chapter 2).

6.1. Linear subvarieties of $\mathbb{P}^n$

We start by generalizing the "projectivities" of Chapter 4. Recall that the projective general linear group is defined as the quotient of invertible matrices by the scalar action:

$$PGL(n+1,\mathbb{C}) := \frac{GL(n+1,\mathbb{C})}{\left\langle \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix} \middle| \alpha \in \mathbb{C}^* \right\rangle}.$$

This group acts on projective space by the rule

$$PGL(n+1,\mathbb{C}) \times \mathbb{P}^n \longrightarrow \mathbb{P}^n$$

$$(M, [\underline{Z}]) \longmapsto [M.\underline{Z}] =: T(M)[\underline{Z}].$$

That is, for each $M \in PGL(n+1,\mathbb{C})$, $T(M)$ gives an automorphism of $\mathbb{P}^n$ as a complex manifold. In fact, the projectivities $T(M)$ give *all* automorphisms (generalizing the last exercise of Chapter 4), but we won't prove this here.

A system of k linear equations

$$\left\{ \begin{array}{rcl} \ell_{10}\xi_0 + \cdots + \ell_{1n}\xi_n & = & 0 \\ \vdots & & \\ \ell_{k0}\xi_0 + \cdots + \ell_{kn}\xi_n & = & 0 \end{array} \right\} \tag{6.1.1}$$

defines a linear subspace $V \subseteq \mathbb{P}^n$. Recalling that the rank of a matrix is its number of linearly independent row (or equivalently, column) vectors, the matrix

$$\begin{pmatrix} \ell_{10} & \cdots & \ell_{1n} \\ \vdots & \ddots & \vdots \\ \ell_{k0} & \cdots & \ell_{kn} \end{pmatrix} =: L$$

has $\mathrm{rank}(L) =: r \leq k$. Defining

$$\mathrm{codim}(V) := r \quad (\text{equivalently, } \dim(V) = n - r),$$

we have the

PROPOSITION 6.1.1. *(i) All projective linear subvarieties of the same (co)dimension are projectively equivalent.*

(ii) A linear subvariety of $\mathbb{P}^n$ of codimension r is isomorphic to $\mathbb{P}^{n-r}$ as a complex manifold.

PROOF. Given L and V as above, note that if the equations (6.1.1) are not independent (i.e. $k > r$), then without changing V or r, we can eliminate equations (reducing k) until they are (and $k = r$, i.e. L has maximal rank). Assume this has been done, so that reordering Z_i's if necessary,

$$\det \begin{pmatrix} \ell_{10} & \cdots & \ell_{1,k-1} \\ \vdots & \ddots & \vdots \\ \ell_{k0} & \cdots & \ell_{k,k-1} \end{pmatrix} \neq 0.$$

Let M be the $(n+1) \times (n+1)$ matrix whose first k rows are given by L (a $k \times (n+1)$ matrix) and last $n-k+1$ rows by $(0\ ,\ \mathbb{I}_{n-k+1})$, where 0 denotes a $(n-k+1) \times k$ matrix of zeroes and $\mathbb{I}_m$ always means an $m \times m$ identity matrix.

Consider the automorphism $T(M)$ of $\mathbb{P}^n$. By definition of V, $[\underline{\xi}] \in V$ if and only if (matrix multiplication by) L kills $\underline{\xi}$. So one should view $T(M)$ as taking V to the subspace $V_0 = \{\xi_0 = \cdots = \xi_{k-1} = 0\}$, which proves (i) since V was arbitrary. This also proves (ii) since V_0 is evidently a $\mathbb{P}^{n-k}$ (with homogeneous coordinates $[\xi_k : \cdots : \xi_n]$). $\square$

A linear subvariety of codimension 1 is called a *hyperplane*.

6.2. Quadric hypersurfaces

Recall that a (projective) hypersurface is a subvariety $X \subset \mathbb{P}^n$ cut out by a single homogeneous equation $F(\underline{Z}) = 0$. We are interested in the case where $F \in S^2_{n+1}$ (degree 2), so that X is a quadric. The polynomial can be written

$$F(\underline{Z}) = {}^t\underline{Z}\mathcal{B}\underline{Z} = \begin{pmatrix} Z_0 & \cdots & Z_n \end{pmatrix} \begin{pmatrix} b_{00} & \cdots & b_{0n} \\ \vdots & \ddots & \vdots \\ b_{n0} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} Z_0 \\ \vdots \\ Z_n \end{pmatrix}$$

with $\mathcal{B}$ symmetric. Under a linear change of projective coordinates

$$\begin{pmatrix} Z_0 \\ \vdots \\ Z_n \end{pmatrix} = M \begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix} \qquad (M \in GL(n+1, \mathbb{C})),$$

we find

$$F(\underline{Z}) = \begin{pmatrix} Y_0 & \cdots & Y_n \end{pmatrix} {}^tM\mathcal{B}M \begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix} =: G(\underline{Y}),$$

where (as in Chapter 4) ${}^tM\mathcal{B}M$ is said to be *cogredient* to $\mathcal{B}$.

Lemma 6.2.1. [Sylvester's Theorem /$\mathbb{C}$] Any given symmetric complex $(n+1) \times (n+1)$ matrix $\mathcal{B}$ is cogredient to exactly one of the matrices

$$M_k = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

where k is the number of 1's.

Corollary 6.2.2. *A given quadric hypersurface in $\mathbb{P}^n$ is projectively equivalent to (or transformable by a linear change of coordinates* *into) exactly one of the quadrics*

$$\mathcal{Q}_k = \left\{ \sum_{j=0}^{k-1} Y_j^2 = 0 \right\} \qquad (k = 1, \dots, n+1).$$

Note that $\mathcal{Q}_1 = \{Y_0^2 = 0\}$ is a double hyperplane, $\mathcal{Q}_2 = \{Y_0^2 + Y_1^2 = 0\}$ is a union of two hyperplanes (each $\cong \mathbb{P}^{n-1}$), and $\mathcal{Q}_k$ for $k \geq 3$ is irreducible (equation does not factor). You'll investigate these a tiny bit further in one of the exercises.

6.3. Singularities, tangent planes, and dimension

We'll need the Euler formula mentioned in §2.1, so let's prove it first:

LEMMA 6.3.1. [EULER'S FORMULA] $F \in S_{n+1}^d \implies \sum_{i=0}^n Z_i \frac{\partial F}{\partial Z_i} = d.F$

PROOF. It suffices to check this on monomials $(F =) Z_0^{d_0} \cdots Z_n^{d_n}$, $\sum d_i = d$. We have $\sum_i Z_i \frac{\partial}{\partial Z_i}(Z_0^{d_0} \cdots Z_n^{d_n}) = \sum_i Z_i \frac{d_i}{Z_i}(Z_0^{d_0} \cdots Z_n^{d_n}) = (\sum_i d_i) Z_0^{d_0} \cdots Z_n^{d_n} = d Z_0^{d_0} \cdots Z_n^{d_n}$. $\square$

Now, the definition of smoothness for hypersurfaces is similar to what we have learned for curves; the general case of varieties cut out by more than one equation is trickier. So we'll start, then, with an affine hypersurface

$$V = V(f) \subset \mathbb{C}^n,$$

and a point $p \in V$.

DEFINITION 6.3.2. (i) V is *smooth at* $p \iff \frac{\partial f}{\partial z_j}(p) \neq 0$ for some $j \in \{1, \dots, n\}$. Otherwise, p is a *singular point* (or *singularity*) of V.

(ii) If V is smooth at all of its points, V is *smooth*. Otherwise, V is *singular*.

(iii) If V is smooth at p, define the *tangent plane*

$$T_pV := \left\{ (z_1(p) + \alpha_1, \dots, z_n(p) + \alpha_n) \mid \sum_{i=1}^n \alpha_i \frac{\partial f}{\partial z_i}(p) = 0 \right\} \subset \mathbb{C}^n.$$

(Here the $\alpha_i \in \mathbb{C}$.)

So one can think of T_pV as a copy of $\mathbb{C}^{n-1}$ with origin at p and coordinates $\{\alpha_i\}$. It's also worth noting the formal correspondence between "tangent vectors" (points in T_pV) and differential operators "at p", namely $\sum_i \alpha_i \frac{\partial}{\partial z_i}$. This is not misleading at all, and in fact the intrinsic construction of tangent planes (for complex or more generally differentiable manifolds) uses local differential operators.

As for singularities, i.e. points where $f(p) = f_{z_1}(p) = \cdots = f_{z_n}(p) = 0$, we saw examples of those in Example 2.1.10 for curves. Here is one more: $y^2 = x^3 - x^2$ is singular at $(0,0)$ since both partials of $y^2 - x^3 + x^2$. On the other hand, $y^2 = x^3 - x$ is smooth because this equation together with $0 = \frac{\partial}{\partial x}(y^2 - x^3 + x) = -3x^2 + 1$ and $0 = \frac{\partial}{\partial y}(y^2 - x^3 + x) = 2y$ admit no common solution. This is easy to see: the points $(\frac{1}{\sqrt{3}}, 0)$ and $(\frac{-1}{\sqrt{3}}, 0)$ where both partials vanish, do not lie on the curve.

Next, consider a projective hypersurface

$$V = \bar{V}(F) \subset \mathbb{P}^n,$$

where F is homogeneous and $P \in V$.

DEFINITION 6.3.3. (i) V is *smooth at* $P \iff \frac{\partial F}{\partial Z_j}(P) \neq 0$ for some $j \in \{0, \ldots, n\}$. Otherwise, P is a *singular point* (or *singularity*) of V.

(ii) If V is smooth at all of its points, V is *smooth*. Otherwise, V is *singular*.

(ii) If V is smooth at P, define the *tangent plane* ($\cong \mathbb{P}^{n-1}$)

$$T_P V := \left\{ [Z_0(P) + \alpha_0 : \ldots : Z_n(P) + \alpha_n] \mid \sum_{i=0}^{n} \alpha_i \frac{\partial F}{\partial Z_i}(P) = 0 \right\} \subset \mathbb{P}^n.$$

Now *a priori*, the definition of a singular point is one at which $F(P) = F_{Z_0}(P) = \cdots = F_{Z_n}(P) = 0$; but by the Euler formula,

$$\sum_i Z_i(q) \frac{\partial F}{\partial Z_i}(P) = \deg(F) \cdot F(P) \tag{6.3.1}$$

and so it is in fact enough to check $F_{Z_0}(P) = \cdots = F_{Z_n}(P) = 0$. In fact, (6.3.1) also implies (for the projective case only!) the simplification

$$T_P V = \left\{ [\alpha_0 : \cdots : \alpha_n] \mid \sum_i \alpha_i \frac{\partial F}{\partial Z_i}(P) = 0 \right\} \tag{6.3.2}$$

Note that (6.3.2) is really just the solution set of ${}^t\underline{\partial F}(P) \cdot \underline{\alpha} = 0$, as in Chapter 4 (but now in $\mathbb{P}^n$ rather than $\mathbb{P}^2$).

As you might expect, the notions of tangent plane in affine and projective cases "agree", in the sense that – at a point on an affine hypersurface – the tangent plane of the projective completion is the completion of the tangent plane:

PROPOSITION 6.3.4. $T_pV(F(1, z_1, \ldots, z_n)) = T_{[1:p]}\bar{V}(F) \cap \mathbb{C}^n$, *where* $(P =)[1 : p]$ *means* $[1 : z_1(p) : \cdots : z_n(p)]$.

PROOF. Given $q = (z_1(q), \ldots, z_n(q)) \in \mathbb{C}^n$. Writing $f(\underline{z}) = F(1, \underline{z})$ and $Q = [1 : q]$, we want to show

$$q \in T_pV(f) \iff Q \in T_P\bar{V}(F). \tag{6.3.3}$$

The left-hand (affine) condition is, writing $z_i(q) = z_i(p) + \alpha_i$ in Definition 6.3.2(iii),

$$\sum_{i=1}^{n}(z_i(q) - z_i(p))\frac{\partial f}{\partial z_i}(p) = 0.$$

This is really

$$\sum_{i=1}^{n}(Z_i(Q) - Z_i(P))\frac{\partial F}{\partial Z_i}(P) = 0,$$

which by Euler becomes

$$\sum_{i=1}^{n} Z_i(Q) \cdot \frac{\partial F}{\partial Z_i}(P) \;-\; \deg(F) \cdot F(P) \;+\; 1 \cdot \frac{\partial F}{\partial Z_0}(P) = 0.$$

Since $F(P) = 0$, we get

$$1 \cdot \frac{\partial F}{\partial Z_0}(P) + \sum_{i=1}^{n} Z_i(Q) \cdot \frac{\partial F}{\partial Z_i}(P) = 0,$$

which is exactly the right-hand (projective) condition of (6.3.3). $\square$

Now let's have a look at singularities and smoothness in the general projective case. The definition is complicated, but after this chapter we won't use it much. Let

$$V = \bar{V}(F_1, \ldots, F_k) \subseteq \mathbb{P}^n,$$

and p be a point on V.

DEFINITION 6.3.5. (i) V is *smooth at p* if and only if there exists a neighborhood $W \subset \mathbb{P}^n$ of p and sub-index set$\{i_1, \ldots, i_c\} \subseteq \{1, \ldots, k\}$ such that[1]

(a) $V \cap W = \bar{V}(F_{i_1}, \ldots, F_{i_c}) \cap W$, and

(b) rank $\begin{pmatrix} \frac{\partial F_{i_1}}{\partial Z_0}(p) & \cdots & \frac{\partial F_{i_1}}{\partial Z_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{i_c}}{\partial Z_0}(p) & \cdots & \frac{\partial F_{i_c}}{\partial Z_n}(p) \end{pmatrix} = c.$

We say V has *codimension* c (or *dimension* $n - c$) at p.

(ii) If V is smooth at each point $p \in V$, then V is *smooth* (otherwise, V is *singular*).

(iii) If V has the same (co)dimension at each smooth point $p \in V$, then V is *equidimensional*. If moreover that codimension is c, we just say V is a variety of codimension c (dimension $n - c$).[2]

(iv) The *tangent plane* $T_pV \subset \mathbb{P}^n$ to V at a smooth point p is the solution set of $L.\underline{p} = 0$, where

$$L = \begin{pmatrix} \frac{\partial F_1}{\partial Z_0}(p) & \cdots & \frac{\partial F_1}{\partial Z_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_k}{\partial Z_0}(p) & \cdots & \frac{\partial F_k}{\partial Z_n}(p) \end{pmatrix}.$$

Condition 6.3.5(i)(a) says that locally about p, once you set $F_{i_1}(\underline{Z}) = \cdots = F_{i_c}(\underline{Z}) = 0$, the remaining equations are redundant; and roughly speaking, the condition (b) on rank says that no more (none of the F_{i_ℓ}) are redundant. In the terminology of §6.1, T_pV is a linear subvariety, and it follows from condition 6.3.5(i)(b) that its codimension is c. That is, we have really just defined the (co)dimension of a variety V at a smooth point p, to be the (co)dimension of T_pV — something we already knew how to define.

Finally, the "neighborhood" W in the definition is an analytic open set containing p (such as a "ball"), but the definition would also work if we only permitted "algebraic" open sets defined by complements of (other) subvarieties,[3] known as *Zariski open sets*. In general, if you

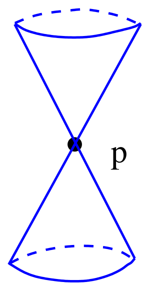
[1]the matrices in this definition assume a particular representative $(Z_0(p), \ldots, Z_n(p))$ in $\mathbb{C}^{n+1}$ of the projective coordinate $[Z_0(p) : \cdots : Z_n(p)]$. It doesn't matter which one you take, as long as you are consistent.

[2]note that the dimension of V at a *singular* point is not defined, so the definition of a m-dimensional variety *must* be "one that is of dimension m at all *smooth* points".

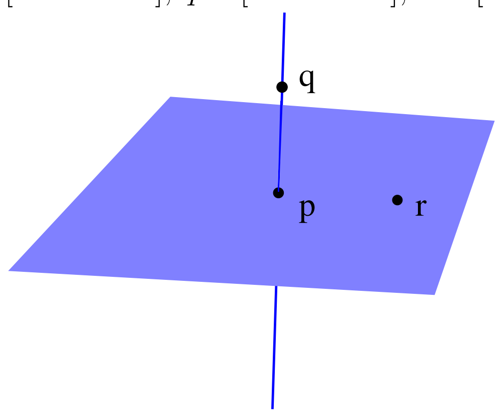
[3]this is exactly how Example 6.3.6(iii) is started below

want to view an algebraic variety as a complex analytic space (or manifold, if it is smooth), then you must use analytic open sets; on the other hand, the Zariski open sets introduce a different topology on V or $\mathbb{P}^n$, which is coarser but has the advantage of being algebraic (and is still Hausdorff). We need both. In brief, when we study varieties analytically, we use the analytic topology; when we want to make heavy use of the correspondence between varieties and ideals in commutative rings, we use the Zariski topology.

EXAMPLE 6.3.6. (i) Let V be the affine variety $\{z_1^2+z_2^2+z_3^2=0\} \subset \mathbb{C}^3$. The partial derivatives $\partial_{z_i}(\sum_j z_j^2) = 2z_i$ all vanish at $p=(0,0,0)$ and so V is singular there:



(ii) Now for a nasty one. Let $V \subset \mathbb{P}^3$ be defined by $\left\{ \begin{array}{c} Z_1Z_3 = 0 \\ Z_2Z_3 = 0 \end{array} \right\}$, and take $p=[1:0:0:0]$, $q=[1:0:0:1]$, $r=[1:1:0:0]$:



Locally about r, $Z_1 \neq 0$ and so having set $Z_1Z_3=0$ (i.e. $Z_3=0$), the second equation $Z_2Z_3=0$ is redundant. So the relevant matrix from 6.3.5(i)(b) is $\left(\begin{array}{cccc} \partial_{Z_0}(Z_1Z_3) & \partial_{Z_1}(Z_1Z_3) & \partial_{Z_2}(Z_1Z_3) & \partial_{Z_3}(Z_1Z_3) \end{array} \right)\Big|_r = \left(\begin{array}{cccc} 0 & Z_3 & 0 & Z_1 \end{array} \right)|_r = \left(\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array} \right)$, which has rank 1 proving that V has codimension 1 (dimension 2) at r. Locally about q, $Z_3 \neq 0$ and so the equations are effectively $Z_1=0$ and $Z_2=0$, neither of which is redundant. The matrix in 6.3.5(i)(b) is now $\left.\begin{pmatrix} 0 & Z_3 & 0 & Z_1 \\ 0 & 0 & Z_3 & Z_2 \end{pmatrix}\right|_q =$ $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, which does have rank 2 confirming that V has codimension 2 at q. So V is not equidimensional. Finally, at p neither equation is redundant but the matrix 6.3.5(i)(b) is $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, meaning V is singular at p.

(iii) Finally, consider the variety $C \subset \mathbb{P}^3$ defined by the three equations

$$\left\{ \begin{array}{lc} Z_0 Z_3 - Z_1 Z_2 = 0 & (\mathbf{I}) \\ Z_1^2 - Z_0 Z_2 = 0 & (\mathbf{II}) \\ Z_2^2 - Z_1 Z_3 = 0 & (\mathbf{III}) \end{array} \right\},$$

and the covering $U_i := \{Z_i \neq 0\}$ of $\mathbb{P}^3$. In U_0, we can divide by Z_0, so that $(\mathbf{II})$ becomes $Z_2 = \frac{Z_1^2}{Z_0}$. Together with $(\mathbf{I})$, this gives $Z_2^2 = Z_2 Z_2 = Z_1 Z_2 \frac{Z_1}{Z_0} = Z_0 Z_3 \frac{Z_1}{Z_0} = Z_1 Z_3$. Consquently, $(\mathbf{III})$ is redundant on U_0. Now, for any point in $U_0 \cap C$, you can check that the matrix in 6.3.5(i)(b) has rank 2, showing that C is smooth of dimension 1 at all of those points. To finish, and show that C is a 1-dimensional smooth variety, carry out a similar analysis in each of U_1, U_2, and U_3 (exercise).

Somewhat unsurprisingly, a variety of dimension 1 is called a curve, of dimension 2 a surface, and of dimension $d \geq 3$ a d-fold. So-called "Calabi-Yau threefolds" play a central role in mathematical string theory.

6.4. Singularities of plane curves

Consider a curve

$$C = \{F(\underline{Z}) = 0\} \subset \mathbb{P}^2$$

defined by a homogeneous polynomial $F \in S_3^d$ (of degree 3 in Z_0, Z_1, Z_2). A point $p \in C$ is a singularity if and only if $F_{Z_0}(p) = F_{Z_1}(p) = F_{Z_2}(p) = 0$, and (moving C by a projectivity if necessary) we may assume that $p = [1 : 0 : 0]$. To locally analyze C at p, we can pass to affine coordinates $x = \frac{Z_1}{Z_0}$, $y = \frac{Z_2}{Z_0}$ and replace F by

$$f(x, y) = F(1, x, y) = \sum_{m=k}^{d} f_m(x, y),$$

where $f_m \in S_2^m$ for each m, and $f_k \neq 0$. Now $0 = f_x(p) = f_y(p) = f(p)$ ($p = (0,0)$) translates to $0 = f_1 = f_0$, so that $k \geq 2$. We say that p is a *k-tuple point* of C, or a singularity of *order* k.

So far, we have not defined tangent planes at singular points. Indeed, this can really only be done for curves in general. To decide what the tangent lines to C at 0 should be, we think of $[x : y]$ as homogenous coordinates on the $\mathbb{P}^1$ of lines through $(0,0) = p$. The lowest-order homogeneous term f_k of f defines a 0-dimensional variety $\tau_p(C) := \{f_k(x,y) = 0\}$ in this $\mathbb{P}^1$. For each $[x_0 : y_0] \in \tau_p(C)$, one should think of $\frac{y_0}{x_0}$ as the slope of a line tangent to some "local irreducible component"[4] of the curve C at p.

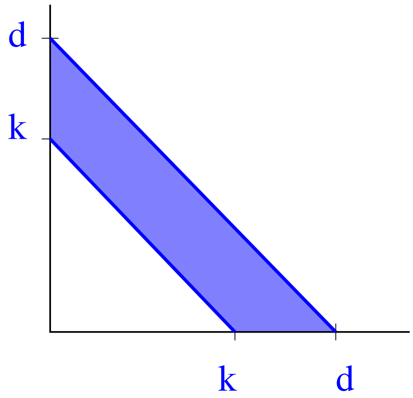
DEFINITION 6.4.1. The tangent lines to a curve C at a singularity p are the lines through p corresponding to points of $\tau_p(C)$.

Now, $f_k(x,y) = 0$ has k solutions counted with multplicity. If these are all distinct, i.e. if $\tau_p(C)$ is reduced, then we say p is an *ordinary* k-tuple point. The most geometric way to think of this is that C has k distinct tangent lines at p.

Any line through a k-tuple point p *other* than one of C's tangent lines there, meets C with multiplicity k at p: if L is given parametrically by $t \mapsto (at, bt)$ ($f_k(a,b) \neq 0$) then the intersection multiplicity is computed as in Chapter 2 by taking the order of $f(at, bt) = t^k f_k(a,b) + \cdots$ at $t = 0$.

REMARK 6.4.2. Given a polynomial $f(x,y) = \sum_{(a,b)\in\mathbb{Z}^2_{\geq 0}} \alpha_{ab} x^a y^b$ it can be useful (for various purposes) to plot the finitely many (a,b) with $\alpha_{ab} \neq 0$. If you do this when $(0,0)$ is a k-tuple point and f has degree n, then these lie in the shaded region

[4]this will be made precise when we do local normalization



This may be useful for one of the exercises below.

There is more to singularities, it turns out, than their order or even the tangent line configuration reflected by $\tau_p(C)$. A local analytic classification of so-called *simple*[5] singularities of curves has been carried out. For the purposes of this classification, if C, C' are two curves through $p = (0,0)$, their singularities at p are considered equivalent if there are small neighborhoods U, U' of $(0,0)$ in $\mathbb{C}^2$ and a biholomorphism $U \overset{\cong}{\to} U'$ carrying p to p and C to C'. The different classes of simple singularities carry "A-D-E" labels, which reflect a relation to other classifications in mathematics (simple Lie algebras/Dynkin diagrams, rational surface singularities, etc.)

The results are that double points are all equivalent to one of

$$\mathbf{A_n}: \quad x^2 + y^{n+1} = 0 \qquad (n \geq 1),$$

and (simple) triple points to one of

$$\mathbf{D_n}: \quad y(x^2 + y^{n-2}) = 0 \qquad (n \geq 4),$$

$$\mathbf{E_6}: \quad x^3 + y^4 = 0,$$

$$\mathbf{E_7}: \quad x(x^2 + y^3) = 0,$$

$$\mathbf{E_8}: \quad x^3 + y^5 = 0.$$

The ODP's (ordinary double points: two distinct tangents) are all of type A_1, as all $A_{n\geq 2}$ have only one tangent; amongst the latter, "cusps" are the singularities of type A_2.[6] OTP's (ordinary triple points) are all

[5]cf. [Barth, Hulek, Peters, and van de Ven, "Compact complex surfaces", Springer, 2004] for the definition. Simple singularities encompass all double (2-tuple) points and some triple (3-tuple) points, and nothing of higher order.

[6]refer back to Chapter 2 for a few pictures (cusps, ODP, OTP)

of type D_4; we note that the tangent lines to $y(x^2+y^2)=0$ have slopes $0, i, -i$. All $D_{n\geq 5}$ have two distinct tangents (one with "mutliplicity 2") and the $E_{6,7,8}$ each have one tangent (of "multiplicity 3").

Exercises

(1) (i) Prove that a quadric hypersurface in $\mathbb{P}^n$ defined by a symmetric bilinear form B is smooth if and only if $\det(B) \neq 0$.
(ii) Corollary 6.2.2 associates a number k to each projective quadric hypersurface in $\mathbb{P}^n$. Show that any two are projectively equivalent if and only if they have the same value of k. [This is easy.]

(2) Show that [the closure in $\mathbb{P}^2$ of] $y^2 = 4x^3 + ax + b$ is smooth unless $a^3 + 27b^2 = 0$.

(3) Find the tangent plane to the complex surface $2x^4+y^4+z^4-4xyz = 0$ (in $\mathbb{C}^3$) at the point $p = (1, 1, 1)$.

(4) Finish the proof in Example 6.3.6(iii) that C is a smooth curve.

(5) What form does a degree k projective algebraic curve (in $\mathbb{P}^2$) take if it has a singularity of order k?

(6) Analyze the singularity of $C = \{(x^2 + y^2)^2 + 3x^2y - y^3 = 0\} \subseteq \mathbb{C}^2$ at the origin. (What is its order, and type?)

(7) For which values of μ are the algebraic curves $F(X, Y, Z) = 0$ in $\mathbb{P}^2$ singular (in (a) and (b) below)? Attempt a sketch of each of the singular curves, saying where the singularities are located and what type they are.
(a) $F(X, Y, Z) = X^3 + Y^3 + Z^3 + \mu(X + Y + Z)^3$,
(b) $F(X, Y, Z) = X^3 + Y^3 + Z^3 + 3\mu XYZ$.

CHAPTER 7

Smooth varieties as complex manifolds

This Chapter starts the long slog toward a proof of part (A) of the Normalization Theorem 3.2.1. After introducing a bit of the theory of several complex variables, we'll use the holomorphic implicit function theorem to put a complex manifold structure on any smooth irreducible (affine or projective) algebraic variety:

THEOREM 7.0.1. *A smooth irreducible algebraic curve $C \subset \mathbb{P}^n$ "is" a Riemann surface. (More precisely, there exists a Riemann surface M and an injective morphism of complex manifolds $\sigma : M \hookrightarrow \mathbb{P}^2$ with C as its image.)*

This is, of course, the "smooth" case of Thm. 3.2.1(A). As for going the other way, from Riemann surfaces to algebraic curves, here is a statement which is different in character from 3.2.1(B):

THEOREM 7.0.2. *A Riemann surface M with $n+1$ linearly independent meromorphic functions $f_0, \ldots, f_n \in \mathcal{K}(M)$, yields an algebraic curve in $\mathbb{P}^n$ not contained in any proper linear subvariety.*

We won't prove this in full — just the existence of a morphism $M \to \mathbb{P}^n$ of complex manifolds which is *nondegenerate*, i.e. whose image is not contained in any $\mathbb{P}^{n-1}$. Proving that the image is described by algebraic equations (hence yields an algebraic curve) is harder.

7.1. Background from several complex variables

Let $\mathcal{O}_n$ (or $\mathbb{C}\{z_1, \ldots, z_n\}$) denote the *ring of convergent power series* $\sum_I a_I \underline{z}^I$ *in* $z_1, \ldots, z_n$, or equivalently, holomorphic functions defined on *some* neighborhood of $\underline{0} \in \mathbb{C}^n$ (cf. §5.1).

PROPOSITION 7.1.1. [W. OSGOOD, 1900] *Let f be a function on an open neighborhood of $\underline{0} \in \mathbb{C}^n$ which is holomorphic in each z_i as the other $\{z_j\}_{j \neq i}$ are held fixed; that is, $\frac{\partial f}{\partial \bar{z}_i} = 0$ $(\forall i)$. Then f is in fact a holomorphic function (and so gives an element of $\mathcal{O}_n$).*

PROOF. We will only give the proof for $n = 2$. Since f is holomorphic in z_2, we have

$$f(z_1, z_2) = \frac{1}{2\pi\sqrt{-1}} \oint \frac{f(z_1, \zeta_2)}{\zeta_2 - z_1} d\zeta_2;$$

using the holomorphicity in z_1, this

$$= \frac{1}{(2\pi\sqrt{-1})^2} \oint\oint \frac{f(\zeta_1, \zeta_2)}{(\zeta_1 - z_1)(\zeta_2 - z_2)} d\zeta_1 d\zeta_2$$

$$= \frac{1}{(2\pi\sqrt{-1})^2} \oint\oint \frac{f(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2}{\zeta_1\zeta_2 \left(1 - \frac{z_1}{\zeta_1}\right)\left(1 - \frac{z_2}{\zeta_2}\right)}.$$

Now using the power-series expansion

$$\frac{1}{1 - \frac{z_i}{\zeta_i}} = \sum_{k \geq 0} \left(\frac{z_i}{\zeta_i}\right)^k,$$

whose uniform convergence allows us to swap integration and summation, we find

$$f(z_1, z_2) = \sum_{k_1, k_2 \geq 0} \left(\frac{1}{(2\pi\sqrt{-1})^2} \oint\oint \frac{f(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2}{\zeta_1^{k_1+1} \zeta_2^{k_2+1}} \right) z_1^{k_1} z_2^{k_2}.$$

□

In order to put a complex manifold structure on a smooth variety, we will need a way to parametrize zero-loci of holomorphic functions. This is given by the *holomorphic implicit function theorem* which here I will just state and prove in the two variable case.

PROPOSITION 7.1.2. *Let* $f \in \mathcal{O}_2$ *with* $f(0, 0) = 0$, $\frac{\partial f}{\partial z_1}(0, 0) \neq 0$. *Then there exists* $w \in \mathcal{O}_1$ *such that in a neighborhood of* $(0, 0)$ *in* $\mathbb{C}^2$,

$$f(z_1, z_2) = 0 \iff z_1 = w(z_2).$$

The upshot of this is that z_2 gives a local holomorphic coordinate on $\{f(z_1, z_2) = 0\}$.

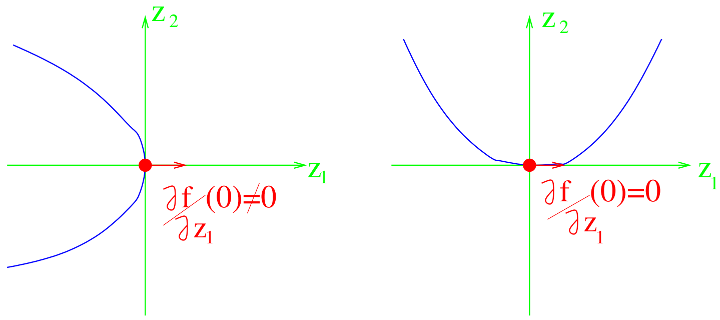
PROOF. We will assume the C^∞ implicit function theorem, and just check that the w it yields is holomorphic:

$$0 = \frac{\partial}{\partial \overline{z_2}} f(w(z_2), z_2)$$

$$= \frac{\partial f}{\partial \overline{z_2}}(w(z_2), z_2) + \frac{\partial f}{\partial z_1}(w(z_2), z_2) \cdot \frac{\partial w}{\partial \overline{z_2}} + \frac{\partial f}{\partial \overline{z_1}}(w(z_2), z_2) \cdot \frac{\partial \overline{w}}{\partial \overline{z_2}}.$$

Now since $f \in \mathcal{O}_2$, $\frac{\partial f}{\partial \bar{z}_1} = \frac{\partial f}{\partial \bar{z}_2} = 0$; moreover, by assumption $\frac{\partial f}{\partial z_1} \neq 0$ locally. So we find that $\frac{\partial w}{\partial \bar{z}_2} = 0$, so that $w \in \mathcal{O}_1$. $\square$

Here is a visual explanation of why the nonvanishing condition on $\partial f/\partial z_1$ matters:



In the left-hand picture, you can write z_1 as a function of z_2 (as desired); on the right-hand side, you cannot.

7.2. Smooth normalization

The more general statement which implies Theorem 7.0.3 is:

THEOREM 7.2.1. *Given*

- *a closed connected subset Y of a compact complex n-manifold X;*
- *a system of open neighborhoods $\{W_\alpha \subset X\}$ covering Y (with local holomorphic coordinates $\underline{z_\alpha} = (z_{\alpha 1}, \ldots, z_{\alpha n})$);*
- *holomorphic functions $f_{\alpha 1}, \ldots, f_{\alpha \ell} \in \mathcal{O}(W_\alpha)$ (for each α) such that*[1] *$Y \cap W_\alpha = V(\{f_{\alpha j}\}_{j=1,\ldots,\ell}) \cap W_\alpha$; and (also for each α)*
- *rank* $\left(\{\partial f_{\alpha j}/\partial z_{\alpha k}\}_{\substack{j=1,\ldots,\ell \\ k=1,\ldots,n}} \right) = \ell$ *[the "Jacobian condition"].*

Then Y is an $(n-\ell)$-dimensional compact complex $(n-\ell)$-manifold.

We say that Y is a codimension-ℓ complex submanifold of X. In fact, Theorem 7.2.1 immediately gives:

COROLLARY 7.2.2. *Any smooth irreducible projective algebraic variety $Y \subset \mathbb{P}^n$ of dimension d is a compact complex d-manifold.*

---

[1]this condition makes Y into an "analytic subvariety" of X; here $V(f_1, \ldots, f_\ell)$ means the vanishing locus $f_1 = \cdots = f_\ell = 0$, just as in the algebraic setting.

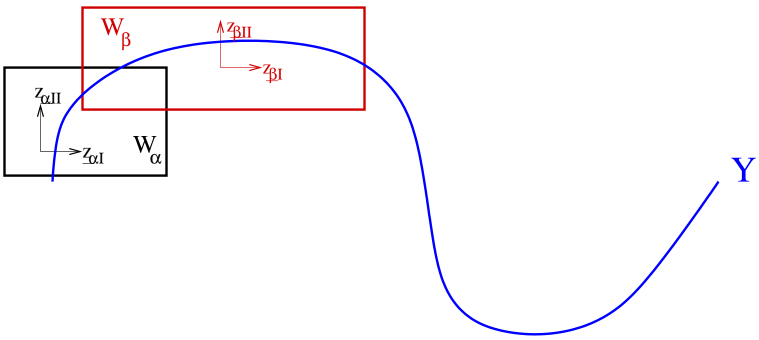
PROOF. Put $X = \mathbb{P}^n$ and $\ell = n-d$. That Y is smooth of dimension d (Defn. 6.3.5) implies the Jacobian condition required in Thm. 7.2.1. $\square$

Now we prove the Theorem.

PROOF. Refining the covering if necessary, we can arrange to have

$$\det\left(\left\{\frac{\partial f_{\alpha j}}{\partial z_{\alpha k}}\right\}_{1\leq j,k\leq \ell}\right) \neq 0. \tag{7.2.1}$$

Write "$\underline{z_\alpha}_I$" for $(z_{\alpha 1}, \ldots, z_{\alpha \ell})$ and "$\underline{z_\alpha}_{II}$" for $(z_{\alpha,\ell+1}, \ldots, z_{\alpha n})$. (So $\underline{z_\alpha} = \left(\underline{z_\alpha}_I, \underline{z_\alpha}_{II}\right)$A schematic picture:



By the condition (7.2.1), and the general holomorphic implicit function theorem, we have holomorphic functions $\{w_\alpha\}$ (mapping from open subsets of $\mathbb{C}^{n-\ell}$ to $\mathbb{C}^\ell$) such that

$$Y \cap W_\alpha = \left\{\underline{z_\alpha}_I = w_\alpha\left(\underline{z_\alpha}_{II}\right)\right\}$$

for each α. Hence, the $\left\{\underline{z_\alpha}_{II}\right\}$ give local coordinates on the $\{Y \cap W_\alpha\}$, which constitute an open cover of Y.

Consider the transition functions for X

$$\Phi_{\alpha\beta} : \underline{z_\beta}(W_{\alpha\beta}) \overset{\cong}{\to} \underline{z_\alpha}(W_{\alpha\beta})$$

$$\left(\underline{z_\beta}_I, \underline{z_\beta}_{II}\right) \mapsto \left(\phi_I\left(\underline{z_\beta}_I, \underline{z_\beta}_{II}\right), \phi_{II}\left(\underline{z_\beta}_I, \underline{z_\beta}_{II}\right)\right) =: \left(\underline{z_\alpha}_I, \underline{z_\alpha}_{II}\right)$$

corresponding to change of coordinates on $W_{\alpha\beta}$. Clearly the functions describing change of coordinates on $Y \cap W_{\alpha\beta}$ are then

$$\Phi^Y_{\alpha\beta} : \underline{z_\beta}_{II} \mapsto \phi_{II}\left(w_\beta\left(\underline{z_\beta}_{II}\right), \underline{z_\beta}_{II}\right) =: \underline{z_\alpha}_{II}.$$

This is 1-to-1 because $\Phi_{\alpha\beta}$ is, and holomorphic because ϕ_{II} and w_β are. So we have the data of an analytic atlas on Y. $\square$

7.3. Nondegenerate morphisms

The statement related to Theorem 7.0.4 which we shall prove is:

PROPOSITION 7.3.1. *Given a Riemann surface M, the following data are equivalent:*

(a) $n+1$ linearly independent meromorphic functions $f_i \in \mathcal{K}(M)$;

(b) a nondegenerate holomorphic map (morphism of complex manifolds) $\sigma : M \to \mathbb{P}^n$.

We will need the notion of a meromorphic function on a complex manifold of any dimension.

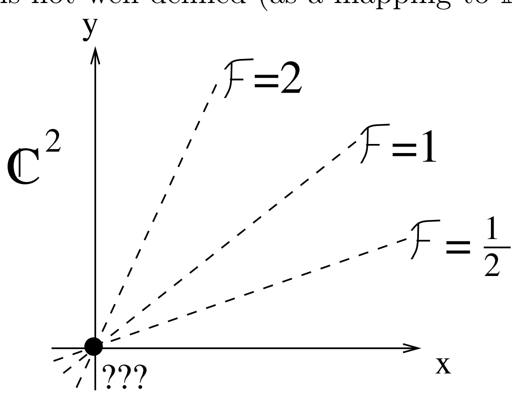
DEFINITION 7.3.2. A meromorphic function $\mathcal{F} \in \mathcal{K}(X)$ (on a complex manifold X) is a collection $\{(U_\alpha, g_\alpha, h_\alpha)\}$ such that

- $\{U_\alpha\}$ is an open cover of X;
- $g_\alpha, h_\alpha \in \mathcal{O}(U_\alpha)$ (they are holomorphic functions); and
- $g_\alpha h_\beta = g_\beta h_\alpha$ on $U_{\alpha\beta}$.

We write "$\mathcal{F} = \frac{g_\alpha}{h_\alpha}$" on U_α.[2]

REMARK 7.3.3. For $\dim(X) = 1$, this coincides with the earlier Definition 3.1.1 (via g/h); by Prop. 3.1.8 meromorphic functions on Riemann surfaces yield morphisms $X \to \mathbb{P}^1$. But this does *not* generalize: if $\dim(X) > 1$, a meromorphic function on X need not even yield a well-defined mapping $X \to \mathbb{P}^1$.

EXAMPLE 7.3.4. Consider $X = \mathbb{C}^2$ with complex coordinates x, y. Then $\mathcal{F} := x/y$ (one $U_\alpha = X$; $g = x$, $h = y$) defines a meromorphic function, which is not well-defined (as a mapping to $\mathbb{P}^1$) at $(0,0)$.



[2]the third condition says that $\frac{g_\alpha}{h_\alpha} = \frac{g_\beta}{h_\beta}$ on overlaps — at least, where the quotients are defined! (see Remark 7.3.3)

The notion of "blowing up" in algebraic geometry is motivated (in part) by the desire to remove such indeterminacies. In the last example, the idea would be to replace the origin in $X = \mathbb{C}^2$ by the $\mathbb{P}^1$of lines through the origin, yielding a new space $\tilde{X}$ (mapping down onto X) on which the meromorphic function becomes well-defined as a morphism.

EXAMPLE 7.3.5. The meromorphic functions on $\mathbb{P}^n$ and its smooth subvarieties (viewed as complex manifolds) are the *rational functions* $\mathcal{F} = \frac{P(\underline{Z})}{Q(\underline{Z})}$ for $P, Q \in S^d_{n+1}$. For instance, the affine coordinates $z_i = \frac{Z_i}{Z_0}$ are meromorphic functions (and more generally, $z_{ji} = \frac{Z_i}{Z_j}$ is one).

Here is how to see at least that "rational functions are meromorphic" in the sense of Definition 7.3.2. (That meromorphic functions are rational is more nontrivial.) In $U_j = \{Z_j \neq 0\}$, set

$$g_j(\underline{z_j}) := P(z_{j0}, \ldots, \underset{\substack{j^{\text{th}} \\ \text{entry}}}{1}, \ldots, z_{jn}) = P\left(\underline{Z}/Z_j\right) = \frac{1}{Z_j^d} P(\underline{Z})$$

$$h_j(\underline{Z_j}) := Q(z_{j0}, \ldots, \underset{\substack{j^{\text{th}} \\ \text{entry}}}{1}, \ldots, z_{jn}) = \frac{1}{Z_j^d} Q(\underline{Z});$$

then

$$g_j h_i = \frac{1}{Z_j^d} \frac{1}{Z_i^d} P(\underline{Z}) Q(\underline{Z}) = g_i h_j.$$

EXAMPLE 7.3.6. Consider a holomorphic map $f : C \to X$ from a Riemann surface to a complex manifold, and let $\mathcal{F} \in \mathcal{K}(X)$ be given by $\{(g_\alpha, h_\alpha, U_\alpha)\}$. Assume that $f(C)|_{U_\alpha} \cap \{h_\alpha = 0\}$ is a finite point set, and put $W_\alpha := f^{-1}(U_\alpha)$, $G_\alpha := g_\alpha \circ f$, $H_\alpha := h_\alpha \circ f$. Then $f^* \mathcal{F} := \{(G_\alpha, H_\alpha, W_\alpha)\}$ (or rather, G/H) belongs to $\mathcal{K}(C)$.

The last two examples will be used in the proof of Prop. 7.3.1, which we now give:

PROOF. The first issue is how we get from $n+1$ meromorphic functions to a morphism to $\mathbb{P}^n$. The set of points in M which cause a problem is

$$\Delta := \{q \in M \mid f_i(q) = 0 \text{ for all } i\} \cup \{q \in M \mid f_i(q) = \infty \text{ for some } i\}$$

Define

$$f : (M \backslash \Delta) \to \mathbb{P}^n$$

by

$$p \longmapsto [f_0(p) : \cdots : f_n(p)].$$

Near $q \in \Delta$ let z be a local holomorphic coordinate with $z(q) = 0$, then write $f_i(z) = z^{\nu_q(f_i)} h_i(z)$ (where h_i are local holomorphic functions not vanishing at q), and put $\nu := \min_{i\in\{0,\ldots,n\}}\{\nu_q(f_i)\}$. For $z \neq 0$,

$$f(z) = \left[z^{-\nu} f_0(z) : \cdots : z^{-\nu} f_n(z)\right];$$

none of the entries in this blows up locally, and at least one does not vanish at $z = 0$ (i.e. at q). Hence, f extends to all of M, and it is evident that this extension is still holomorphic as a map to $\mathbb{P}^1$.

Next, given a morphism $f : M \to \mathbb{P}^n$, we want to product an $(n+1)$-tuple of meromorphic functions. Referring to Examples 7.3.5 (for z_i) and 7.3.6 (for f^*), simply take $f_i := f^* z_i$ and you're done.

Finally, to see that f is degenerate iff the $\{f_i\}$ are linearly dependent, consider the correspondence between nonzero vectors $\underline{v} \in \mathbb{C}^{n+1}$ (up to scale) and hyperplanes in $\mathbb{P}^n$, by taking $\mathbb{P}^{n-1}_{\underline{v}}$ to be the projectification of $(\mathbb{C}^{n+1})^{\perp \underline{v}}$. Degeneracy of f occurs iff $f(M) \subset \mathbb{P}^{n-1}_{\underline{v}}$ for some $\underline{v}$, which is to say $(f_0(p), \ldots, f_n(p)) \perp \underline{v}$ for all $p \in M$. But this just reads $\sum v_i f_i(p) = 0$ $(\forall p)$, which is a nontrivial linear relation. $\square$

We give two examples of nondegenerate projective embeddings of Riemann surfaces (the first is actually a series of examples). For these cases we actually give algebraic equations for the image.

EXAMPLE 7.3.7. The so-called *rational canonical curves* are the images of the nondegenerate morphisms

$$f : \mathbb{P}^1 \hookrightarrow \mathbb{P}^n$$

given, for each $n \in \mathbb{N}$, by

$$[Z_0 : Z_1] \mapsto [Z_0^n : Z_0^{n-1} Z_1 : \cdots : Z_1^n].$$

(In affine terms, one can think of this as $z \mapsto [1 : z : \ldots : z^n]$, with $\infty \mapsto [0 : \cdots : 0 : 1]$.)

Let's see what this looks like for the first few values of n:

• for $n = 1$, f sends $[Z_0 : Z_1] \mapsto [Z_0 : Z_1]$ and so is just the identity map.

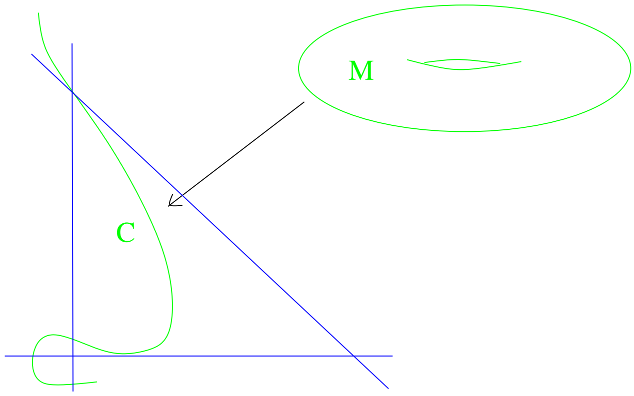
• for $n = 2$, we have $[Z_0 : Z_1] \mapsto [Z_0^2 : Z_0 Z_1 : Z_1^2]$. If we write $[Y_0 : Y_1 : Y_2]$ for the homogeneous coordinates on $\mathbb{P}^2$, then the image is the conic $\{Y_1^2 - Y_0 Y_2 = 0\} \subset \mathbb{P}^2$.

• for $n = 3$, $[Z_0 : Z_1] \mapsto [Z_0^3 : Z_0^2 Z_1 : Z_0 Z_1^2 : Z_1^3] (= [Y_0 : Y_1 : Y_2 : Y_3])$ has image

$$V := \overline{V}(Y_0 Y_3 - Y_1 Y_2, Y_1^2 - Y_0 Y_2, Y_2^2 - Y_1 Y_3) \subset \mathbb{P}^3.$$

By exercise 4 from Chapter 6 you know that V is smooth.

Example 7.3.8. Let $M = \mathbb{C}/\Lambda$ ($\Lambda \subset \mathbb{C}$ a lattice) be a complex 1-torus. We want to demonstrate that there is a (nondegenerate) morphism from M to $\mathbb{P}^2$ with a cubic curve as image. Note that this will present M as the normalization of such a cubic curve:



Some of the steps will be exercises.

First, there exists a unique meromorphic function $\wp \in \mathcal{K}(\mathbb{C})$ satisfying

• $\wp(u + \lambda) = \wp(u)$ for every $\lambda \in \Lambda$ and $u \in \mathbb{C}$

• $\wp(u) = u^{-2} + h(u)$, where $h \in \mathcal{K}(\mathbb{C})$ is holomorphic in a neighborhood of 0, has all its poles in $\Lambda \backslash \{0\}$, and $h(0) = 0$.

Existence is an exercise. Uniqueness is easy: if $\mathcal{Q}$ were another such function, $\wp - \mathcal{Q} = (\wp - u^{-2}) - (\mathcal{Q} - u^{-2})$ has no pole at 0 *and* is Λ-periodic, hence has no poles in Λ either. But the only possible zeroes were in Λ, and so $\wp - \mathcal{Q}$ is entire. By compactness of a fundamental region for Λ, any Λ-periodic entire function is bounded hence (by Liouville) constant. Since $\wp - \mathcal{Q}$ is zero at 0, this constant is zero and $\wp = \mathcal{Q}$.

In the exercises below, you will also show that $\wp$ is an even function ($\wp(u) = \wp(-u)$) and $(\wp')^2 = 4\wp^3 + a\wp + b$ for some $a, b \in \mathbb{C}$. In each case, you get equality by showing the right-hand side minus the left-hand side has no poles and is zero at some point (as in the uniqueness argument just described). The upshot is that

$$f : \mathbb{C}/\Lambda \to \mathbb{P}^2$$

defined by

$$u \mapsto [1 : \wp(u) : \wp'(u)] \quad \text{for } u \neq \bar{0}$$

and

$$\bar{0} \mapsto [0 : 0 : 1]$$

parametrizes (or normalizes) $C = \{Z_0 Z_2^2 = 4Z_1^3 + aZ_1Z_0^2 + bZ_0^3\}$, a smooth cubic with the affine equation

$$y^2 = 4x^3 + ax + b.$$

What we have said so far only gives that $f(M) \subseteq C$, but viewing the smooth curve C as a complex manifold, and f as a morphism $M \to C$, the open mapping theorem from complex analysis says the image is open; while on the other hand the image of a compact set by a continuous map is compact (hence closed in C). So $f(M)$ is open and closed in C, and thus $f(M) = C$.

Exercises

(1) Show that the rational canonical map $f : \mathbb{P}^1 \to \mathbb{P}^n$ has the following property: the image of any collection of $k\ (\leq n+1)$ distinct points $\{w_1, \ldots, w_k\} \subset \mathbb{P}^1$ is in general position (spans a $\mathbb{P}^{k-1}$ in $\mathbb{P}^n$). [Hint: Vandermonde determinant.] Then, taking $k = 2$, explain why this shows f is injective.

(2) Turning to the case $n = 3$ in Example 7.3.7, (a) actually *prove* that $V = Image(f)$ and (b) that you cannot throw out any of the three equations defining V.

(3) Show that the map $\mathbb{P}^2 \to \mathbb{P}^5$ given by $[Z_0 : Z_1 : Z_2] \mapsto [(Z_0)^2 : Z_0Z_1 : Z_0Z_2 : (Z_1)^2 : Z_1Z_2 : (Z_2)^2]$ is (a) well-defined and (b) holomorphic (i.e. a "morphism of complex manifolds"), then (c) write (polynomial) equations expressing the image as an algebraic variety. (For (c) you can just write the equations and not prove it.)

(4) Let λ_1, $\lambda_2 \in \mathbb{C}$ be two complex numbers which are $\mathbb{R}$-linearly independent, and let

$$\Lambda = \mathbb{Z}\lambda_1 + \mathbb{Z}\lambda_2 = \{n_1\lambda_1 + n_2\lambda_2 |\ n_1, n_2 \in \mathbb{Z}\}$$

be the lattice in $\mathbb{C}$ that they generate.

(a) Show that the series

$$\wp(u) = \frac{1}{u^2} + \sum_{\substack{\lambda \in \Lambda \\ \lambda \neq 0}} \left(\frac{1}{(u-\lambda)^2} - \frac{1}{\lambda^2} \right)$$

is absolutely and uniformly convergent on any compact subset of the complex u-plane which does not contain any of the points of Λ. [Hint: any compact subset is contained inside one of the following form: $|u| \leq K \ \cap \ |u - \lambda| \geq \epsilon\, (\forall \lambda)$. Break the sum into terms with $|\lambda| \leq 2K$, and $|\lambda| > 2K$, and use (essentially) the Weierstrass M-test.]

(b) Verify the pole condition in Example 7.3.8: that all poles are on Λ, and in a neighborhood U of 0, $\wp(u) = u^{-2} + h(u)$ with h holomorphic and $h(0) = 0$. [Hint: what do you know about an absolutely and uniformly convergent series of analytic functions?]

(c) Show that $\wp$ is a doubly-periodic function; that is, show that

$$\wp(u + \lambda) = \wp(u) \text{ for every } u \in \mathbb{C} \text{ and every } \lambda \in \Lambda.$$

[Hint: From (a), you can calculate the derivative $\wp'(u)$ by differentiating each term of the series defining $\wp(u)$. First prove $\wp'(u + \lambda) = \wp'(u)$, then integrate.]

(5) Now forget the explicit formula for $\wp(u)$ just given, and retain just these facts: that $\wp \in \mathcal{K}(\mathbb{C})$ is Λ-periodic with all poles $\in \Lambda$, and locally of the form $\wp(u) = u^{-2} + h(u)$ with h holomorphic (on some $U \subset \mathbb{C}$ containing a fundamental domain) and $h(0) = 0$. Prove that (a) $\wp(u) = \wp(-u)$ [$\Longrightarrow$ h even $\Longrightarrow$ h' odd] and (b) $(\wp'(u))^2 = 4(\wp(u))^3 + a\wp(u) + b$ for some $a, b \in \mathbb{C}$. [See hint given in the Example.]

CHAPTER 8

The connectedness of algebraic curves

The main theorem of this chapter will be that the smooth part[1] $C \backslash \mathrm{sing}(C)$ of an *irreducible* algebraic curve $C \subset \mathbb{P}^2$ is path-connected (and then, of course, so is C). For example, in Exercise 5 of Chapter 3, you showed that the complement of the ODP $\hat{p} = [1 : 0 : 0]$ in the singular cubic curve $\{Y^2Z - X^2Z + X^3 = 0\}$, viewed as a complex 1-manifold, is isomorphic to $\mathbb{C}^*$ — which is certainly connected.

Just so that there is no confusion, we should say what the situation is for reducible curves right away and why the result does not generalize. For plane projective algebraic curves with more than one irreducible component, say $C = \cup C_i$, the components C_i must intersect (this will be one consequence of Bezout's theorem later), making C connected. But the complement of the singularities in C will not be connected, as these will include all of the intersection points.

We begin by introducing a new, somewhat technically involved, tool for dealing with singularities, intersections, and projections of curves.

8.1. Resultants and discriminants

Let $\mathbb{D}$ be a unique factorization domain (UFD), where we recall that this is a commutative domain in which each element has a unique factorization into irreducibles, up to reordering and multiplication by units. In a UFD, amongst other things, the notion of a greatest common divisor[2] has meaning. By the Gauss lemma, $\mathbb{D}[y]$ is also a UFD. In practice we will always take $\mathbb{D}$ to be $\mathbb{C}$ or $\mathbb{C}[x]$. (Note that $\mathbb{C}[x]$ is a PID, but $\mathbb{C}[x, y]$ is not.)

Consider $f(y) = a_0y^m + a_1y^{m-1} + \cdots + a_m$, $g(y) = b_0y^n + b_1y^{n-1} + \cdots + b_n$ elements of $\mathbb{D}[y]$ with $a_0, b_0 \neq 0$.

[1] we will show that the set $\mathrm{sing}(C)$ of singular points is always finite

[2] recall that these are well-defined up to units (invertible elements); for example in $\mathbb{C}[x]$ or $\mathbb{C}[x, y]$ the units are $\mathbb{C}^*$, hence the notion of "monic gcd" (which is completely well-defined).

DEFINITION 8.1.1. The *resultant*[3] *of* f *and* g, written $\mathcal{R}(f,g)$, is the element of $\mathbb{D}$ given by the determinant of the $(n+m)\times(n+m)$ *Sylvester matrix*[4]

$$M_{(f,g)} := \left(\begin{array}{cccccccc} a_0 & a_1 & \cdots & \cdots & a_m & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & \cdots & \cdots & a_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & & \ddots & 0 \\ 0 & \cdots & 0 & a_0 & \cdots & \cdots & a_{m-1} & a_m \\ \hline b_0 & b_1 & \cdots & \cdots & b_n & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & \cdots & \cdots & b_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & & \ddots & 0 \\ 0 & \cdots & 0 & b_0 & \cdots & \cdots & b_{n-1} & b_n \end{array}\right).$$

Now writing K for the field of fractions of $\mathbb{D}$, we have the

PROPOSITION 8.1.2. $\mathcal{R}(f,g)=0 \iff gcd_{K[y]}(f,g)\neq 1.$[5]

PROOF. The gcd (say, h) is nontrivial if and only if

$$Fg = Gf \tag{8.1.1}$$

for some $F = A_0y^{m-1}+\cdots+A_{m-1}$ and $G = B_0y^{n-1}+\cdots+B_{n-1}$ in $\mathbb{D}[y]$. Indeed, if $h\neq 1$ then put $F=f/h$ and $G=g/h$. Conversely, since $\deg F<\deg f$ and $\deg G<\deg g$, and both sides of (8.1.1) factor into the same irreducibles, f and g have a common factor of degree >0.

In turn, (8.1.1) is equivalent to

$$\begin{aligned} a_0B_0 &= b_0A_0 \\ a_1B_0+a_0B_1 &= b_1A_0+b_0A_1 \\ &\vdots \\ a_mB_{n-1} &= b_nA_{m-1} \end{aligned} \tag{8.1.2}$$

being satisfied for some $\{A_i\}_{i=0}^{m-1},\{B_j\}_{j=0}^{n-1}\subset\mathbb{D}$. To get from (8.1.1) to (8.1.2), just take coefficients of y^{m+n-1}, y^{m+n-2}, ..., 1.

[3] *also called "eliminant", since y is eliminated*
[4] the line in the matrix is just an organizational device — it has no meaning
[5] two further equivalent conditions: (i) $\deg_y(\gcd_{\mathbb{D}[y]}(f,g))>0$; and, noting that $K[y]$ is a PID, so that the ideal $(f,g)_{K[y]}=(\gcd_{K[y]}(f,g))$, (ii) $(f,g)_{K[y]}\neq(1)_{K[y]}$.

Now notice that (8.1.2) can be rephrased in matrix multiplication terms: there exist $\{A_i\}, \{B_j\}$ such that

$$ {}^tM_{(f,g)} \cdot \begin{pmatrix} B_0 \\ \vdots \\ B_{n-1} \\ -A_0 \\ \vdots \\ -A_{m-1} \end{pmatrix} = \underline{0}. $$

In other words, we have shown $h \neq 1$ is the same as $\ker({}^tM_{(f,g)}) \neq \{0\}$, i.e. $\det(M_{(f,g)}) = 0$. □

DEFINITION 8.1.3. $\mathcal{D}(f) := \mathcal{R}(f, f')$ is the *discriminant* of f. Here f' denotes the formal derivative $\frac{\partial f}{\partial y}$.

EXAMPLE 8.1.4. If $f \in \mathbb{C}[y]$, then $\mathcal{D}(f) \in \mathbb{C}$ is a number, and the criterion

$$ \mathcal{D}(f) \text{ vanishes } \iff f \text{ has a multiple root} \tag{8.1.3} $$

follows immediately from Prop. 8.1.2. For the affine curve

$$ z^2 = 4y^3 + ay + b $$

to be singular, we need two of the roots of the right-hand side to coincide. That is, by (8.1.3), we need

$$ 0 = \mathcal{R}(4y^3 + ay + b, 12y^2 + a) = \begin{vmatrix} 4 & 0 & a & b & \\ & 4 & 0 & a & b \\ 12 & 0 & a & & \\ & 12 & 0 & a & \\ & & 12 & 0 & a \end{vmatrix} $$

which after a bit of row-reduction

$$ = \begin{vmatrix} 4 & 0 & a & b & 0 \\ 0 & 4 & 0 & a & b \\ 0 & 0 & -2a & -3b & 0 \\ 0 & 0 & 0 & -2a & -3b \\ 0 & 0 & 12 & 0 & a \end{vmatrix} = 16(4a^3 + 12 \cdot 9b^2) $$

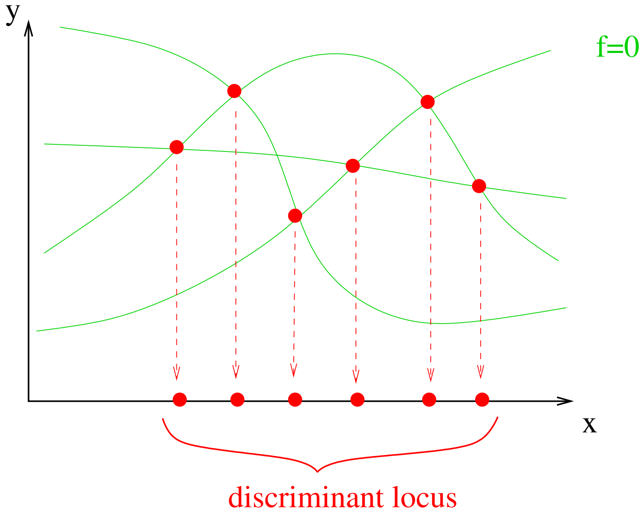
$$ = 64(a^3 + 27b^2). $$

This recovers the result from Exercise 2 of Chapter 6.

EXAMPLE 8.1.5. If $f \in \mathbb{C}[x, y]$, then $\mathcal{D}(f) \in \mathbb{C}[x]$ is a polynomial and from Prop. 8.1.2 we have:

$$\mathcal{D}(f) \text{ vanishes at } x_0 \iff f(x_0, y) \text{ has a multiple root in } y. \tag{8.1.4}$$

The collection of x_0's where this happens, that is, the set of roots of $\mathcal{D}(f)$, is called the *discriminant locus* for the projection of the affine curve $\{f(x, y) = 0\}$ onto the x-line:



PROPOSITION 8.1.6. *An irreducible (reduced) algebraic curve* $\{F = 0\} \subset \mathbb{P}^2$ *has (if any) finitely many singularities.*

PROOF. The affine polynomial $f(x, y) = F(1, x, y)$ has multiple roots in y for x in the discriminant locus $\Delta = \{(\mathcal{D}(f))(x) = 0\} \subseteq \mathbb{C}$. We may assume f has positive degree in y, since otherwise $V(f)$ is just a vertical line.

Since f is irreducible in $\mathbb{C}[x, y]$ of positive degree in y, the identical vanishing of $\mathcal{D}(f)$ would imply that $V(f)$, hence $\bar{V}(F)$, was non-reduced. So $\mathcal{D}(f)$ is a nontrivial polynomial, and Δ is finite:

$$\#\{x \in \mathbb{C} \mid \exists y \text{ such that } f(x, y) = f_y(x, y) = 0\} < \infty \tag{8.1.5}$$

It is easy to argue directly[6] that were $V(f)$ to contain a vertical line$\{x = \alpha\}$, then $(x - \alpha)$ would divide f (contradicting irreducibility). So by

[6]or you can wait for Study's lemma in the next Chapter

(8.1.5) and Prop. 2.1.8, in fact

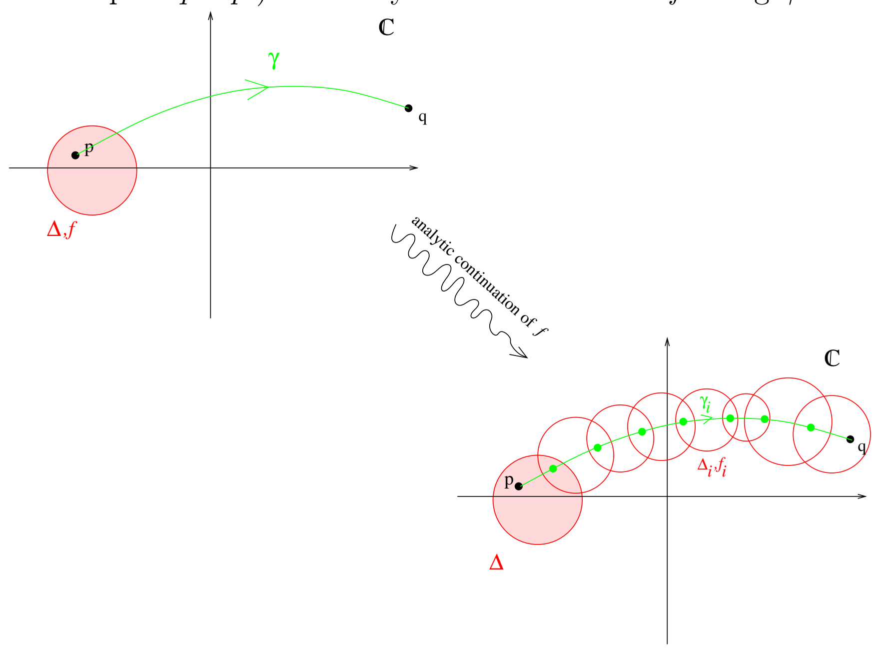
$$\#\{p \in \mathbb{C}^2 \mid f(p) = f_y(p) = 0\} < \infty.$$

The set in brackets includes all singularities of $V(f)$. The only possible *additional* singularities of $\bar{V}(F)$ are the (finitely many) points where it meets the line at ∞. $\square$

8.2. Monodromy and connectedness

Let $\Omega \subseteq \mathbb{C}$ be a *region*, that is, an open connected subset. Let $\Delta \subset \Omega$ be a small disk about a point $p \in \Omega$ on which one is given a holomorphic function, $f \in \mathcal{O}(\Delta)$. We are interested in the question of when f extends to a holomorphic function on all of Ω. To see why this doesn't always happen, take $\Omega = \mathbb{C}$ and Δ a small disk about $z = 1$: then $f = \frac{1}{z}$ only extends to a holomorphic function on $\mathbb{C}^*$. Even worse, $f = \log(z)$ becomes "multivalued" on $\mathbb{C}^*$ and so (as a holomorphic function) only extends to $\mathbb{C}\backslash\mathbb{R}_{\leq 0}$.

To give a condition which will ensure the existence of a well-defined holomorphic extension, we need the concept of *analytic continuation*. Define a *path* $\gamma \subset \Omega$ from p to q to be the image of a continuous function $\mathsf{P} : [0,1] \to \Omega$ with $\mathsf{P}(0) = p$ and $\mathsf{P}(1) = q$. (Here we are allowed to pick $q = p$.) An analytic continuation of f along γ



consists of

- a partition of γ into segments $\{\gamma_i\}_{i=0}^N$,
- a covering of γ by disks $\Delta_i \supset \gamma_i$ (with $\Delta_0 = \Delta$), and
- functions $f_i \in \mathcal{O}(\Delta_i)$ (with $f_0 = f$) satisfying $f_i \equiv f_{i+1}$ on $\Delta_i \cap \Delta_{i+1}$.

If we continue f along two different paths from p to q and compare the "results", i.e. the last function $f_N \in \mathcal{O}(\Delta_N)$ (in the neighborhood of q) in each case, these need not agree. In the above example of $f = \log(z)$ on a disk about $p = \{z = 1\}$, we can analytically continue f along any path in $\mathbb{C}^*$. However, if we take $q = p$ so that the path is closed, then we do not have $f_N(p) = f_{(0)}(p)$: they differ by $2\pi\sqrt{-1}$ times the winding number of the path about $z = 0$, hence the "multivaluedness" referred to above. This problem only occurs, however, for non-simply-connected regions:

Proposition 8.2.1. [Riemann Monodromy Principle] *Given a region $\Omega \subseteq \mathbb{C}$ which is simply connected, i.e. $\pi_1(\Omega) = \{0\}$. Let $\Delta \subset \Omega$ be a small disk, and assume that $f \in \mathcal{O}(\Delta)$ can be analytically continued along any path $\gamma \subset \Omega$ starting at $p \in \Delta$. Then there exists $\tilde{f} \in \mathcal{O}(\Omega)$ extending f.*

We will frequently use this together with the

Proposition 8.2.2. [Heredity principle] *Given $F(x, y) \in \mathcal{O}(\mathbb{C}^2)$, $f \in \mathcal{O}(\Delta)$ with*

$$F(x, f(x)) = 0. \tag{8.2.1}$$

Then the analytic continuation of f along any path γ must satisfy (8.2.1).

Proof. Since F and each f_i in the analytic continuation are holomorphic, so is each $F(x, f_i(x))$ (on Δ_i). But $F(x, f(x)) \equiv 0$ on $\Delta = \Delta_0$ by assumption, and since $f = f_0 \equiv f_1$ on $\Delta_0 \cap \Delta_1$, we have $F(x, f_1(x)) \equiv 0$ on $\Delta_0 \cap \Delta_1$ and therefore (by basic complex analysis) on all of Δ_1. Simply iterate this argument for $i = 1, \ldots, N$. $\square$

Now given an affine algebraic curve $C = \{f_0(x_0, y) = 0\}$ with f_0 of degree n, it is convenient to write C as the vanishing locus of a monic polynomial in y over $\mathbb{C}[x]$:

$$f(x, y) = y^n + a_1(x)y^{n-1} + \cdots + a_n(x) = 0. \tag{8.2.2}$$

This is acheived by performing a change of variable $x_0 = x + \lambda y$ and writing $f(x, y) := f_0(x_0, y) = f_0(x + \lambda y, y)$, which has coefficient of y^n depending polynomially on λ; choose λ so that this coefficient is 1. (The main point is that in $f_0(x_0, y)$, the y^n term may be zero, and we want to remedy that.)

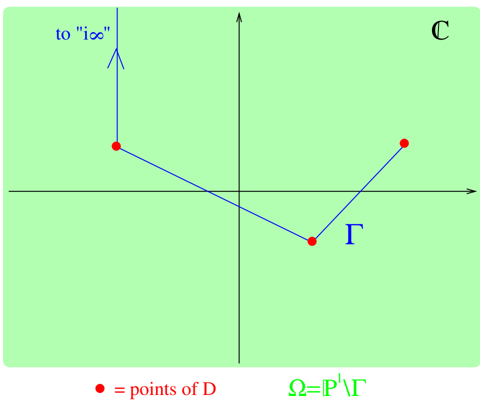
Having put the equation of C in this form, we write

$$\begin{array}{rccc} \pi: & C & \longrightarrow & \mathbb{C} \\ & (x, y) & \longmapsto & x \end{array}$$

for the projection of the curve to the x-axis. Writing $D := \{\mathcal{D}(f)(x) = 0\}$ for the *discriminant locus* of this projection, by (8.1.4) we have that for $x \in \mathbb{C} \backslash D$, the fibre $\pi^{-1}(x)$ consists of n distinct points. For some fixed disk $\Delta \subset \mathbb{C} \backslash D$, label these points $\{y_1(x), \ldots, y_n(x)\}$. Notice that $\mathcal{R}(f, \frac{\partial f}{\partial y}) = \mathcal{D}(f)(x) \neq 0$ implies that $\frac{\partial f}{\partial y} \neq 0$ on $\{f = 0\} \cap \pi^{-1}(\Delta)$, so that the holomorphic IFT (Prop. 7.1.2) gives $y_i(x) \in \mathcal{O}(\Delta)$. The point here is that the "roots" of (8.2.2) in y are algebraic — hence multivalued — functions of x, but we can take well-defined holomorphic branches of them over Δ. As we shall see, the multivaluedness will intertwine them outside Δ.

Label the points of $D = \{p_1, \ldots, p_K\}$, and let Γ be the path in $\mathbb{P}^1$ consisting of segments connecting ∞ to p_1, p_1 to p_2, and so on up to p_K. Then the region $\Omega := (\mathbb{P}^1 \backslash \Gamma) \subset \mathbb{C}$ is simply connected. By Propositions 8.2.1-2, the $\{y_i(x)\}$ extend to functions in $\mathcal{O}(\Omega)$ which still satisfy

$$f(x, y_i(x)) = 0. \tag{8.2.3}$$



Analytically continued *through* Γ in $\mathbb{C} \backslash D$, the y_i continue to satisfy (8.2.3) by the heredity principle, but may *swap*.

EXAMPLE 8.2.3. $f(x,y) = y^3 - x$, $D = \{0\}$, $\Gamma = \mathbb{R}_{\leq 0}$. Passing through Γ cyclically permutes $y_1(x) = \sqrt[3]{x}$, $y_2(x) = e^{\frac{2\pi\sqrt{-1}}{3}}\sqrt[3]{x}$, $y_3(x) = e^{\frac{4\pi\sqrt{-1}}{3}}\sqrt[3]{x}$.

This swapping (or permutation)[7] of the $y_i(x)$ gives rise to an equivalence relation "$\sim$": $y_i(x) \sim y_j(x)$ if one may be analytically continued into the other in $\mathbb{C}\backslash D$. An equivalence class is just all the $\{y_\lambda\}$ which are equivalent to a given y_i in this sense.

PROPOSITION 8.2.4. *For any equivalence class E of $\sim$, formed (reordering if necessary) by $y_1(x), \ldots, y_m(x)$,*

$$\prod_{\lambda=1}^{m}(y - y_\lambda(x)) \tag{8.2.4}$$

belongs to $\mathbb{C}[x,y]$.

Put differently: while the $\{y_\lambda(x)\}_{\lambda=1}^m$ are multivalued algebraic functions on $\mathbb{C}\backslash D$, the elementary symmetric polynomials in them are not multivalued; in fact, they are polynomials![8]

COROLLARY 8.2.5. *C irreducible $\Longrightarrow C\backslash\pi^{-1}(D)$ is connected ($\Longrightarrow$ C connected).*

PROOF. [assuming Prop. 8.2.4] If $f \in \mathbb{C}[x,y]$ doesn't factor, then by the Proposition there can be only one equivalence class: $E = \{1,\ldots,n\}$. So the complete set of "branches" $\{y_i(x)\}$ is acted on transitively by monodromy about D, and one can therefore draw a path on $C\backslash\pi^{-1}(D)$ connecting any two points. $\square$

We now prove Proposition 8.2.4, using some theorems from complex analysis. In particular, recall that *Rouché's theorem* asserts that for two holomorphic functions $f, g \in \mathcal{O}(\mathfrak{R})$ on a simply connected region[9] satisfying $|f| > |g|$ on a path $\gamma \subset \Omega$, $f+g$ and f have the same number of zeroes (counted with multiplicity) inside γ.

[7] the transformations of an algebraic structure arising from its transport around loops (in this case, loops in $\mathbb{C}$ about points of D) are what is meant by the word *monodromy* in general. So the Riemann monodromy principle is really a statement about the *absence* of monodromy.
[8] we use λ to index E (i.e. $1,\ldots,m$) and i to index $\{1,\ldots,n\}$
[9] the main point is that $\mathfrak{R}$ should contain the "interior" of γ

PROOF. The product (8.2.4) is clearly well-defined on $\mathbb{C}\backslash D$, since monodromy about D simply swaps its factors; hence it is in $\mathcal{O}(\mathbb{C}\backslash D)$. Write

$$\prod\nolimits_{\lambda=1}^{m}(y-y_\lambda(x))=\sum_{j=0}^{m}e_{m-j}(y_1(x),\dots y_m(x))y^j \tag{8.2.5}$$

where $e_{m-j}(y_1(x),\dots,y_m(x)) =: e_{m-j}(x)$ denotes the elementary symmetric polynomials in the $\{y_\lambda\}$. Again, because these are not changed under monodromy, we have $e_{m-j}(x)\in\mathcal{O}(\mathbb{C}\backslash D)$. Observe that given $\alpha\in D$ with neighborhood $\mathcal{N}_\alpha$ (a small disk about α), the polynomials $a_j(x)$ from (8.2.2) satisfy

$$x\in\mathcal{N}_\alpha \implies |a_j(x)|\le M\ (\forall j)$$

for some $M\in\mathbb{N}$. Fixing $x_0\in\mathcal{N}_\alpha$, put $a_j=a_j(x_0)$ and

$$\mathfrak{F}(y)=y^n\ ,\qquad \mathfrak{G}(y)=y^n+a_1y^{n-1}+\cdots+a_n,$$

so that the $\{y_i(x_0)\}$ are the roots of $\mathfrak{G}$. On $\gamma=\{|y|=M+1\}\subset\mathbb{C}$, we have

$$\begin{aligned}|\mathfrak{F}-\mathfrak{G}|=|a_1y^{n-1}+\cdots+a_n|&\le M\left((M+1)^n+\cdots+1\right)\\ &=(M+1)^n-1<(M+1)^n=|\mathfrak{F}|.\end{aligned}$$

By Rouché, $\mathfrak{F}$ and $\mathfrak{G}$ have the same number of zeroes inside γ; since $\mathfrak{F}=y^n$ has n zeroes (at $y=0$!), we find that

$$|y_j(x_0)|<M+1\ \text{ for all } j=1,\dots,n \text{ and } x_0\in\mathcal{N}_\alpha.$$

Consequently the $e_k(x)\in\mathcal{O}(\mathbb{C}\backslash D)$ are bounded on $\mathcal{N}_\alpha\cap(\mathbb{C}\backslash D)=\mathcal{N}_\alpha\backslash\{\alpha\}$, and so by the Riemann removable singularity theorem extend across $\{\alpha\}$. Doing this for each $\alpha\in D$, we conclude that $e_k(x)\in\mathcal{O}(\mathbb{C})$.

So the coefficients of the y^j's in (8.2.5) are entire functions of x. To prove that they are polynomials in x, we shall have to consider their behavior about $x=\infty$. If we work in the local coordinates $\tilde{x}=\frac{1}{x}$, $\tilde{y}=\frac{y}{x}$ about $[0:1:0]$ in $\mathbb{P}^2$, then the polynomial (8.2.2) defining C becomes[10]

$$\tilde{x}^n f\left(\frac{1}{\tilde{x}},\frac{\tilde{y}}{\tilde{x}}\right)=\tilde{y}^n+\left(\tilde{x}a_1\left(\frac{1}{\tilde{x}}\right)\right)\tilde{y}^{n-1}+\cdots+\tilde{x}^n a_n\left(\frac{1}{\tilde{x}}\right),$$

[10] here we are essentially taking the projective completion of C and restricting that to $U_1\subset\mathbb{P}^2$.

with roots

$$\tilde{y}_i(\tilde{x}) = \tilde{x} y_i \left(\frac{1}{\tilde{x}}\right). \tag{8.2.6}$$

Let $\mathcal{N}_\infty$ be a small neighborhood of $\tilde{x} = 0$ and $\mathcal{N}_\infty^* := \mathcal{N}_\infty \backslash \{\tilde{x} = 0\}$. By (8.2.6), the monodromy of the $\{\tilde{y}_i\}_{i=1}^n$ about $\tilde{x} = 0$ stabilizes the subset $\{\tilde{y}_\lambda\}_{\lambda=1}^m$, so that the $e_k(\tilde{y}_1(\tilde{x}), \ldots, \tilde{y}_m(\tilde{x})) = \tilde{x}^k e_k(\frac{1}{\tilde{x}})$ are well-defined holomorphic functions on $\mathcal{N}_\infty^*$. Since $\deg(a_j(x)) \leq j$, the $\tilde{x}^j a_j(\frac{1}{\tilde{x}})$ are polynomials in $\tilde{x}$ hence bounded on $\mathcal{N}_\infty$. Using Rouché as above, the $e_k(\{\tilde{y}_\lambda(\tilde{x})\}_{\lambda=1}^m)$ are also bounded on $\mathcal{N}_\infty^*$, and thus extend to holomorphic functions on $\mathcal{N}_\infty$.

In other words, $e_k(x) = e_k(\frac{1}{\tilde{x}})$ has a pole at $x = \infty$ of order at most k. Since $e_k(x)$ was also holomorphic on $\mathbb{C}$, we have $e_k \in \mathcal{K}(\mathbb{P}^1)$. Now $\mathcal{K}(\mathbb{P}^1) \cong \mathbb{C}(\mathbb{P}^1)$ and so $e_k(x) = \frac{P(x)}{Q(x)}$ where P, Q are polynomials; since its only pole is at ∞, Q is a constant. Therefore each $e_k \in \mathbb{C}[x]$, and with (8.2.5) we see that (8.2.4) is a polynomial in $\mathbb{C}[x, y]$. $\square$

Exercises

(1) Are the real points[11] of a smooth algebraic curve $\subset \mathbb{P}^2$ necessarily connected?

(2) For what values of a, b does $x^4 + ax + b$ have a multiple root?

[11] i.e. points on the curve which can be written $[X_0 : X_1 : X_2]$, with all $X_i \in \mathbb{R}$

CHAPTER 9

Hilbert's nullstellensatz

In something of an algebraic detour, we will now prove Theorem 5.3.1 for affine hypersurfaces. In the general case, we shall also state (but not prove) a more precise theorem which lays out the correspondence between affine algebraic varieties and ideals in commutative rings. (This is part of the foundation for scheme theory, which you can explore further in the books by R. Hartshorne and E. Kunz.)

9.1. Resultants (bis)

We will need some more results on resultants. As in §8.1 let $\mathbb{D}$ be a UFD with fraction field K; and for $f = a_0Y^n + a_1Y^{n-1} + \cdots + a_n$ and $g = b_0Y^m + b_1Y^{m-1} + \cdots + b_m$ polynomials in $\mathbb{D}[Y]$, define $\mathcal{R}(f,g) :=$ $\det M_{(f,g)}$. (In case $\mathbb{D}$ is itself a polynomial ring, we will sometimes write $\mathcal{R}_Y(f,g)$ to make it clear that Y is the variable being eliminated.)

Proposition 9.1.1. *$\mathcal{R}(f,g) = Gf + Fg$ for some $F, G \in \mathbb{D}[Y]$ with $\deg G < \deg g$, $\deg F < \deg f$.*

Proof. If $\mathcal{R}(f,g) = 0$, then we are done by (8.1.1). Otherwise, write

(9.1.1)

$$\begin{array}{ccccccc}
Y^{m-1}f & = & a_0Y^{n+m-1} & +a_1Y^{n+m-2} & +\cdots & +a_nY^{m-1} & & \\
Y^{m-2}f & = & & a_0Y^{n+m-2} & +\cdots & \cdots & +a_nY^{m-2} & \\
\vdots & & & & & & & \\
f & = & & & a_0Y^n & +\cdots & \cdots & +a_n \\
Y^{n-1}g & = & b_0Y^{n+m-1} & +b_1Y^{n+m-2} & +\cdots & +b_mY^{n-1} & & \\
Y^{n-2}g & = & & b_0Y^{n+m-2} & +\cdots & \cdots & +b_mY^{n-2} & \\
\vdots & & & & & & & \\
g & = & & & b_0Y^m & +\cdots & \cdots & +b_m.
\end{array}$$

Working in $K[Y]$, divide by a_0 in the first m equations, then use "elementary row operations" to kill the first m columns (to the right of

"=") apart from the Y^j's along the diagonal. Then, divide out the equations by the leading coefficients to the right of "=", normalizing those leading coefficients to 1. That is, we have essentially carried out row-reduction on $M_{(f,g)}$ in the context of the linear system, under the assumption that its determinant is nonzero.

The new system takes the form

$$\begin{array}{ccllll} ?? & = & Y^{n+m-1} & & & \\ ?? & = & & Y^{n+m-2} & & \\ \vdots & & & & \ddots & \\ ?? & = & & & & 1 \end{array} \tag{9.1.2}$$

where each "??" is a K-linear combination of the entries to the left of "=" in (9.1.1). In particular, the last row of (9.1.2) is

$$G_0 f + F_0 g = 1$$

where $G_0, F_0 \in K[Y]$ satisfy $\deg G_0 \leq m-1$, $\deg F_0 \leq n-1$. Now multiply by the product of all the elements of $\mathbb{D}$ we divided the equations by, clearing denominators of G_0, F_0 to give elements G, F of $\mathbb{D}[Y]$. But this product is in fact one definition of the determinant of $M_{(f,g)}$, namely $\mathcal{R}_{(f,g)}$. The result follows. $\square$

We should mention the formula for the resultant of two polynomials whose irreducible factors are all linear (or constant) in y, although we will probably neither use nor prove it:

PROPOSITION 9.1.2. *If f and g decompose into linear factors $f = a_0 \prod_i (Y - x_i)$, $g = b_0 \prod_j (Y - y_j)$ (for $x_i, y_j \in \mathbb{D}$), then $\mathcal{R}(f,g) = a_0^m b_0^n \prod_{i,j}(x_i - y_j)$.*

9.2. Study's lemma

We continue to assume that $\mathbb{D}$ is a UFD with $f \in \mathbb{D}[Y]$ of degree n. Given $\delta \in \mathbb{D}$, we have the ring homomorphism given by "evaluation at δ":

$$\begin{array}{ccc} \mathbb{D}[Y] & \xrightarrow{\theta_\delta} & \mathbb{D} \\ G(Y) & \longmapsto & G(\delta) \end{array}.$$

PROPOSITION 9.2.1. *(i) If $f(\delta)(= \theta_\delta(f)) = 0$, i.e. δ is a root of f, then $(Y - \delta) \mid f(Y)$.*

(ii) f has at most n roots in $\mathbb{D}$.

PROOF. (i) By the division algorithm,

$$f = q.(Y - \delta) + r \tag{9.2.1}$$

where $\deg r < \deg(Y - \delta) = 1$, i.e. $r \in \mathbb{D}$. Applying θ_δ to (9.2.1), we have

$$0 = f(\delta) = q(\delta).0 + r$$

and thus $r = 0$, so that $(Y - r)$ divides f.

(ii) Follows from (i) (and the fact that $\mathbb{D}[Y]$ is a UFD) since f can have at most $n = \deg(f)$ linear factors. $\square$

Now we will specialize to the case $\mathbb{D} = \mathbb{C}[X]$; more generally, the results of this section will hold with any algebraically closed field replacing $\mathbb{C}$ and $(X_1, \ldots, X_{n-1})$ replacing X.

Let $F \in \mathbb{D}[Y] = \mathbb{C}[X, Y] = S_2$.

PROPOSITION 9.2.2. *If $V(F) = \mathbb{C}^2$, i.e. F vanishes on all of $\mathbb{C}^2$, then $F = 0$ as an element of S_2.*

PROOF. Suppose $F \neq 0$. By Prop. 9.2.1(ii), viewed as an element of $\mathbb{D}[Y]$, F has a finite number of roots in $\mathbb{C}[X]$. Some of these may be constants in $\mathbb{C}$. Since $\mathbb{C}$ is an infinite field, there exists $\beta \in \mathbb{C}$ such that β is not one of these roots, and then $F(X, \beta)(= \theta_\beta(F)) \neq 0$ in $\mathbb{C}[X]$. Again by Prop. 9.2.1(ii), $F(X, \beta)$ itself has finitely many roots, so there exists $\alpha \in \mathbb{C}$ such that $F(\alpha, \beta) \neq 0$. Hence, F is not identically zero on $\mathbb{C}^2$. $\square$

PROPOSITION 9.2.3. [STUDY'S LEMMA] *Given $f, g \in S_2$, with f irreducible and $V(f) \subseteq V(g)$. Then f divides g.*

REMARK 9.2.4. Suppose we drop the requirement that f be irreducible, so that $f = \prod f_i^{m_i}$ (f_i irreducible in S_2). Then $V(f_i) \subset V(f)$ for each i, and by the Proposition $f_i | g$ for each i. This implies that $f | g^{\sum m_i}$, i.e. f divides a power of g.

PROOF. Since $f|0$ is trivial, we take $g \neq 0$. By Prop. 9.2.2, we have $V(g) \neq \mathbb{C}^2$, which implies $V(f) \neq \mathbb{C}^2$ hence $f \neq 0$. We may assume that $f \notin \mathbb{C}$ (since a constant divides anything), and furthermore that $\deg_Y(f) \neq 0$ (otherwise just swap X and Y). Writing

$$f = a_0(X)Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X) \ \notin \ \mathbb{C}[X]$$

($n > 0$ and $a_0(X) \neq 0$), I make the **claim**:[1] *we can assume that* $g \notin \mathbb{C}[X]$.

Assuming the **claim**, f and g are of degree > 0 in Y, so by Prop. 9.1.1 (with $\mathbb{D} = \mathbb{C}[X]$), $\mathcal{R}_Y(f,g) = Fg + Gf \in \mathbb{C}[X]$ for $\deg_Y F < \deg_Y f$, $\deg_Y G < \deg_Y g$. Given any $\alpha \in \mathbb{C} \backslash V(a_0)$, since $\mathbb{C}$ is algebraically closed there exists a root $\beta \in \mathbb{C}$ of $f(\alpha, Y)$. From $V(f) \subseteq V(g)$ we see that $(\alpha, \beta) \in V(f(\alpha, Y)) \subseteq V(f(\beta, Y)) \subseteq \mathbb{C}$, so that $f(\alpha, Y)$ and $g(\alpha, Y)$ have a common root for every $\alpha \in \mathbb{C} \backslash V(a_0)$. It follows that $a_0 \mathcal{R}_Y(f,g) \in \mathbb{C}[X]$ evaluates to zero at every $\alpha \in \mathbb{C}$, hence is zero. As $a_0 \neq 0$, we find $\mathcal{R}_Y(f,g) = 0$ in $\mathbb{C}[X]$; and then by Prop. 8.1.2, $\deg_Y(\gcd_{S_2}(f,g)) > 0$. (Alternately, $Fg = (-G)f \implies f, g$ have a divisor of nonzero degree in Y.) But f is irreducible, so divides any nonzero non-unit dividing it; we conclude that $f \mid \gcd_{S_2}(f,g) \mid g$.

To prove the **claim**, suppose $g \in \mathbb{C}[X] \backslash \{0\}$. Then there exists $\alpha \in \mathbb{C} \backslash V(g.a_0)$. Viewed as a function on $\mathbb{C}^2$, g is constant in Y, so $g(\alpha, \beta) \neq 0 \; \forall \beta \in \mathbb{C}$. But since $a_0(\alpha) \neq 0$, $\deg_Y(f(\alpha, Y)) > 0$; and then (as $\mathbb{C}$ is algebraically closed) $\exists \beta \in \mathbb{C}$ such that $f(\alpha, \beta) = 0$. By assumption, $V(f) \subseteq V(g)$ and so $g(\alpha, \beta) = 0$, a contradiction. $\square$

9.3. The nullstellensatz

The proof of Study immediately generalizes to $\mathbb{C}^n$. This yields a version of Hilbert's nullstellensatz for hypersurfaces:

COROLLARY 9.3.1. *If $V(f) = V(g)$ for $f, g \in S_n$ and ...*

(i) f, g are irreducible, then $f = \lambda g$ ($\lambda \in \mathbb{C}^*$)

(ii) f, g are not irreducible, then $\exists\, M, N \in \mathbb{N}$ such that $f | g^N$, $g | f^M$. Equivalently, f and g have the same irreducible factors.

PROOF. (i) Study $\implies f|g$ and $g|f$; (ii) is by Remark 9.2.4. $\square$

The point of this is that, modulo issues with powers, there is a *bijection* between hypersurfaces and principal ideals (i.e. polynomials up to multiplication by constants) in S_n which reverses inclusion. That is, provided f and g are "reduced" (all irreducible factors occur with multplicity 1), $(f) \supset (g) \iff f|g \iff V(f) \subset V(g)$.

[1] at this point, of course, we can't "just swap X and Y"

To get a more general perspective on this, we introduce a few new ideas. First, given a subset $\mathfrak{X} \subseteq \mathbb{C}^n$, we define the ideal of $\mathfrak{X}$ by

$$I(\mathfrak{X}) := \{f \in S_n \mid f(\underline{z}) = 0 \ \forall \underline{z} \in \mathfrak{S}\}.$$

For example, if f is "reduced", we clearly have $I(V(f)) = (f)$ by Study's Lemma. A subset $\mathfrak{X} \subseteq \mathbb{C}^n$ is *algebraic* if it is of the form $V(J)$ for some ideal $J \subset S_n$. (Indeed, this is just an affine algebraic variety.) The statement $V(I(\mathfrak{X})) = \mathfrak{X}$ is true (almost a tautology) for algebraic subsets.

Given any ideal $J \subset S_n$, we let $\mathrm{rad}(J)$ denote the *radical* of J, which is the ideal comprising all elements of S_n some power of which belongs to J. A *radical ideal* is an ideal which equals its own radical. Finally, J is *prime* $\iff$ S_n/J is a domain ($\iff$ J is irreducible in the monoid of ideals in S_n), and *maximal* $\iff$ S_n/J is a field.

THEOREM 9.3.2. *Let k be an algebraically closed field.*

(i) Every maximal ideal $\mathfrak{m} \subset S_n$ is of the form $(Z_1 - \alpha_1, \ldots, Z_n - \alpha_n) = \mathfrak{m}_{\underline{\alpha}}$ (here $\underline{\alpha} \in \mathbb{C}^n$);

(ii) Let $J \subsetneq S_n$ be an ideal: then $V(J) \neq \emptyset$;

(iii) For any ideal $J \subset S_n$, $I(V(J)) = rad(J)$.

COROLLARY 9.3.3. *The correspondence*

$$\begin{array}{ccc} \text{ideals} & & \text{subsets} \\ \{J \subset S_n\} & \overset{I}{\underset{V}{\leftrightarrows}} & \{\mathfrak{X} \subset \mathbb{C}^n\} \end{array}$$

induces bijections

$$\begin{array}{ccc} \{\text{radical ideals}\} & \longleftrightarrow & \{\text{algebraic subsets}\} \\ \cup & & \cup \\ \{\text{prime ideals}\} & \longleftrightarrow & \{\text{irred. alg. subsets}\} \end{array}.$$

This is clear from the Theorem; to see the last correspondence, notice that $V(J_1 J_2) = V(J_1) \cup V(J_2)$.

One can push the relation between commutative algebra and affine algebraic geometry much further. For example, the *ring of regular functions* on an irreducible affine variety $V = V(\mathfrak{P})$ ($\mathfrak{P}$ a prime ideal) is defined by

$$\mathbb{C}[V] := S_n/\mathfrak{P},$$

and it is easy to see that this embeds (say, for V smooth) in $\mathcal{O}(V)$. (The idea is that $\mathfrak{P}$ is the kernel of the map from S_n to $\mathcal{O}(V)$ given by restricting polynomial "functions" to , and so $S_n/\mathfrak{P}$ is its image.) $\mathbb{C}[V]$ is sometimes also called the *coordinate ring* of V. Furthermore, if V is the affine part of a smooth projective variety $\bar{V}$, the field of meromorphic functions $\mathcal{K}(\bar{V})$ is isomorphic to the fraction field $\mathbb{C}(V)$ of $\mathbb{C}[V]$. Usually $\mathbb{C}(V)$ is called the *function field* of $\bar{V}$ (or V).

There is even a way to recover varieties from their coordinate rings; this is the "Spec" operation. Very roughly speaking, the affine story is this: any finitely generated commutative domain A may be presented as $\mathbb{C}[z_1, \ldots, z_N]/I$ (where $I \subseteq \mathbb{C}[z_1, \ldots, z_N]$ is an ideal), and then you take $V(I) \subseteq \mathbb{C}^N$. This gives one realization of Spec(A); of course, there are many ways of writing A in this form (different N, different I, etc.). From the standpoint of scheme theory, Spec(A) is something intrinsic, an *affine scheme* which exists in the absence of any particular embedding in an affine space $\mathbb{C}^N$. The best resources on this are the book by E. Kunz and the classic text by R. Hartshorne.

Exercises

(1) Prove: (i) that for any algebraic subset $\mathfrak{X} \subseteq \mathbb{C}^n$, $V(I(\mathfrak{X})) = \mathfrak{X}$; (ii) that for any two ideals $J_1, J_2 \subseteq \mathbb{C}[Z_1, \ldots, Z_n]$, $V(J_1 J_2) = V(J_1) \cup V(J_2)$.

CHAPTER 10

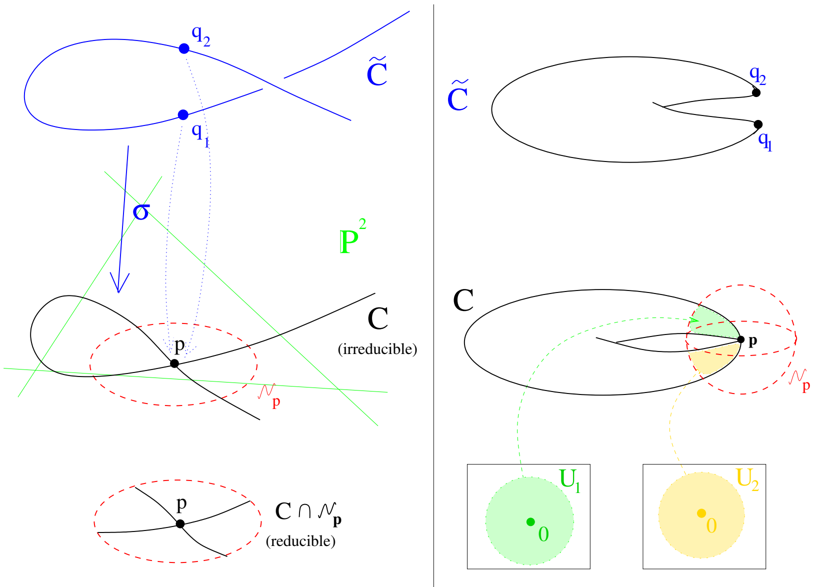
Local analytic factorization of polynomials

Recall the idea of *normalization* for an irreducible algebraic curve $C \subset \mathbb{P}^2$: there should exist a Riemann surface $\tilde{C}$ mapping holomorphically to $\mathbb{P}^2$ with C as its image. In Chapter 7 we *did* this for *nonsingular* C by using the holomorphic implicit function theorem to put a complex manifold structure on C itself. This essentially consisted, for each $p \in C$, in exhibiting a neighborhood $\mathcal{N}_p \subset \mathbb{P}^2$ of p and a (bi)holomorphic parametrization of $\mathcal{N}_p \cap C$ by some open set $U \subset \mathbb{C}$. (The holomorphicity of the transition functions was then a consequence.)

Now suppose C has an *ordinary double point* (ODP) at p — recall that this is a singularity with 2 distinct tangent lines. Denoting disjoint union by "$\amalg$", one has

$$\mathcal{N}_p \cap C \simeq \frac{U_1 \amalg U_2}{0_{U_1} \equiv 0_{U_2}};$$

that is, C locally looks like two disks $U_1, U_2\,(\subset \mathbb{C})$ glued together at one point. In order to normalize C, U_1 and U_2 must be "detached":



schematic picture **topological picture**

Our overarching goal is to produce $\tilde{C}$ and σ as in this figure. Geometrically it seems clear that the "local analytic curve" $\mathcal{N}_p \cap C$ is reducible, even though the global curve C is not. The first step, then, will be to find an appropriate formalism (in terms of 2-variable power series) for working with $\mathcal{N}_p \cap C$, which one might call "analytic localization." In this setting, the local equation can be uniquely factored. This will allow us (in the next Chapter) to carry out *local normalization* — that is, put local coordinates on the irreducible components of $\mathcal{N}_p \cap C$. Finally, we will patch these parametrizations together with those of open subsets of $C \backslash \mathrm{sing}(C)$ to obtain $\tilde{C}$.

There are algebraic approaches to "localization" of C at p. For convenience, replace C for the moment by its affinization in $\mathbb{C}^2$. From §9.3, we have the coordinate ring $R = \mathbb{C}[C]$, and to any point $p \in C$ corresponds a maximal ideal in $\mathfrak{m} \subset R$ (consisting of polynomials vanishing at p). Inverting all primes not contained in $\mathfrak{m}$, or "localizing R at $\mathfrak{m}$", replaces polynomial functions by rational functions with poles anywhere but p, which roughly corresponds to replacing C by C minus any set of points not including p. This is quite different from intersecting C with an analytic ball at p, and will not produce a local factorization of a globally irreducible C. Instead of rational functions, we need convergent power series. The closest construction in algebra is something called *completion* (or Henselian localization). If you are curious (we won't get into this), a good reference is the book by D. Eisenbud.

10.1. Analytic localization

It will suffice to think of C as an affine curve $\{f(x,y)=0\} \subset \mathbb{C}^2$ passing through $p=(0,0)$. The defining polynomial $f \in \mathbb{C}[x,y]$ is, trivially, a convergent power series; so we may consider how f factors in $\mathcal{O}_2 = \mathbb{C}\{x,y\}$ (cf. §7.1). In fact, for purposes of examining the intersection of C with a small neighborhood of the origin, we will show that f may be replaced by an element of $\mathbb{C}\{x\}[y]$ $(\subset \mathcal{O}_2)$ in a particularly nice form:

DEFINITION 10.1.1. The subset[1] $\mathfrak{W} \subset \mathbb{C}\{x\}[y]$ of *Weierstrass polynomials* comprises elements of the form

$$y^d + a_1(x)y^{d-1} + \cdots + a_{d-1}(x)y + a_d(x) \qquad (d \in \mathbb{Z}_{\geq 0})$$

where each $a_j(x) \in \mathbb{C}\{x\}$ satisfies

$$a_j(0) = 0.$$

LEMMA 10.1.2. *Let $f \in \mathcal{O}_2$ with[2] $f \not\equiv 0$ on the y-axis. Then $\exists$ $\epsilon, \rho > 0$ such that:*

(a) $f \neq 0$ on (i)$\{|x|<\rho,\ |y|=\epsilon\}$ and (ii)$\{x=0,\ 0<|y|<\epsilon\}$;

(b) the number of roots (counted with multiplicity) of $f(x,y)$ in y with $|y|<\epsilon$, is constant in x for $|x|<\rho$.

PROOF. The zeroes of $f(0,y)$ are isolated: otherwise they would have a limit point, forcing f to be identically zero. We may therefore choose ϵ so that $f(0,y) \neq 0$ for $0<|y|\leq\epsilon$. To get (a)(i) from this, just use continuity and choose ρ sufficiently small. The number of roots in

[1] technically, a submonoid – you can multiply (but not add) elements, and it has the identity element 1; the notions of "irreducible element" and "uniqueness of factorization" still have meaning. Since $\mathfrak{W}$ is inside a UFD (see proof of Thm. 10.2.2) and has 1 as its sole unit, it does indeed have unique factorization in a very strong sense. (See the discussion after the proof of Thm. 10.2.2.)

[2] in general, if $\mathcal{S}$ is some subset of the domain of definition of a function f, one should read "$f \not\equiv 0$ on $\mathcal{S}$" as "f is not identically zero on $\mathcal{S}$", and "$f \neq 0$ on $\mathcal{S}$" as "f does not vanish on $\mathcal{S}$" (i.e. f is zero at no point of $\mathcal{S}$) — two very different meanings. Henceforth the symbols will be used with no further explanation.

(b) is computed by

$$\frac{1}{2\pi\sqrt{-1}}\oint_{|y|=\epsilon}\frac{f_y(x,y)}{f(x,y)}dy \;\in\; \mathbb{Z},$$

which is continuous in x and therefore constant. □

LEMMA 10.1.3. *For f as in Lemma 10.1.2, let $\{y_\nu(x)\}_{\nu=1,\dots,d}$ be the roots described in (b).*[3] *Denote the elementary symmetric polynomials in them by $e_j(x)$ $(=\sum_{\nu_1<\cdots<\nu_j} y_{\nu_1}(x)\cdot\cdots\cdot y_{\nu_j}(x))$. Then*

$$w := y^d + e_1(x)y^{d-1} + \cdots + e_d(x)$$

is a Weierstrass polynomial.

PROOF. Note that for each ν, $y_\nu(0) = 0$ from Lemma 10.1.2(a)(i). Clearly then the $e_j(x)$ are well-defined and satisfy $e_j(0) = 0$; we must show that they are holomorphic on $\{|x| < \rho\}$. First we have

$$\frac{1}{2\pi\sqrt{-1}}\oint_{|y|=\epsilon} y^k\frac{f_y(x,y)}{f(x,y)}dy \;=\; \sum_\nu (y_\nu(x))^k \;=:\; \sigma_k(x),$$

since the residue at each $y_\nu(x)$ of the argument is $(y_\nu(x))^k\cdot Res_{y_\nu(x)}(\frac{f_y}{f}) = (y_\nu(x))^k \cdot \mathrm{ord}_{y_\nu(x)}(f(x,\cdot))$. Here the *Newton symmetric polynomials* $\sigma_k(x)$ span the same vector space over $\mathbb{C}$ as the $e_j(x)$ (a general fact from algebra). From the integral expression, the σ_k are evidently holomorphic, and therefore so are the e_j. □

Let $\mathfrak{U} := \mathcal{O}_2^* \subset \mathcal{O}_2$ denote the units, which are just the invertible convergent power series, or equivalently the convergent power series with nonzero constant term. (That is, given $g \in \mathcal{O}_2$, $g \in \mathfrak{U} \iff \frac{1}{g} \in \mathcal{O}_2$.)

LEMMA 10.1.4. *For f and w as above, there exists a unique $u \in \mathfrak{U}$ such that $uw = f$, and this holds on all of $V := \{|x| < \rho$ and $|y| \leq \epsilon\}$.*

PROOF. Write $\tilde{u} := \frac{f}{w} \in \mathcal{O}(V\backslash\{w = 0\})$. For fixed x, $w(x,y) = \prod_{\nu=1}^d (y - y_\nu(x))$, as mutliplying this out gives the $e_j(x)$ as coefficients. Consequently, for each fixed x (with $|x| < \rho$), $w(x,y)$ and $f(x,y)$ have

[3] These may well be multivalued on $\{|x| < \rho\}$ — in particular, one should expect them to be permuted as x goes about 0. So the $y_\nu(x)$ are really only well-defined on some simply-connected subset of the disk $\{|x| < \rho\}$ (e.g., deleting the positive real numbers gives a slit disk).

the same roots (in y). Therefore $\tilde{u} \neq 0$ on V, and $\tilde{u}(x,y)$ is (for each x) holomorphic in y. Now, for any given y_0 with $|y_0| < \epsilon$,

$$\tilde{u}(x, y_0) = \frac{1}{2\pi\sqrt{-1}} \oint_{|y|=\epsilon} \frac{\tilde{u}(x,y)}{y - y_0} dy\,.$$

Since $\tilde{u}(x,y)$ is holomorphic on a neighborhood of $|y| = \epsilon$, this formula shows $\tilde{u}(x, y_0)$ is holomorphic in x. By Osgood's lemma, we have $\tilde{u} \in \mathcal{O}(V)$. Since $\tilde{u} \neq 0$, it has nonzero constant term $\tilde{u}(0,0)$, and is thus a unit. Uniqueness is clear since $\tilde{u}w = f$ and $uw = f \implies (\tilde{u} - u)w = f \implies u - \tilde{u} = 0$. □

10.2. Uniqueness of local factorization

The uniqueness of u in the last Lemma was trivial. A slightly less trivial uniqueness question would be: can we write f as a product of a unit and a Weierstrass polynomial in two different ways – i.e., with a different w *and* u? We cannot:

Lemma 10.2.1. *Given $f \in \mathcal{O}_2$ (with $f \not\equiv 0$ on the y-axis), the decomposition $f = wu$ in Lemma 10.1.4 (i.e., into $w \in \mathfrak{W}$ and $u \in \mathfrak{U}$) is unique.*

Proof. Since any unit u has $u(0,0) \neq 0$, shrinking ϵ, ρ (hence V) if necessary, we have $u \neq 0$ on V. Thus if $f = wu$, the zeroes of f and w are the same. This forces $w = \prod(y - y_\nu(x)) = y^d + e_1(x)y^{d-1} + \cdots + e_d(x)$, which makes w (hence u) unique. □

Making use of the last two lemmas, we now show that $f \in \mathcal{O}_2$ factors uniquely (up to units) into irreducibles $f_i \in \mathcal{O}_2$. If f began its life as a polynomial defining an irreducible algebraic curve $C = \{f = 0\} \subset \mathbb{C}^2$, then the local piece $C \cap V$ breaks (uniquely) into irreducible components $\{f_i = 0\}$. Provided there is more than one of them, the f_i are no longer polynomials, for that would contradict (global) irreducibility of C.

Theorem 10.2.2. *$\mathcal{O}_2$ is a UFD.*

Proof. We must demonstrate that $f \in \mathcal{O}_2$ factors into irreducibles $f_1 \cdots f_\ell$ uniquely up to order and units.

First, note that $\mathcal{O}_1 = \mathbb{C}\{x\}$ is a UFD: given $g \in \mathcal{O}_1$, we have a unique decomposition $g(x) = x^{\nu_0(f)}h(x)$, where h is a unit (convergent power series with $h(0) \neq 0$) and $\nu_0(f) \in \mathbb{Z}$. The irreducibles in this case are just the factors of x.

By the Gauss lemma, it follows that $\mathbb{C}\{x\}[y]$ is a UFD.

Next, suppose that $f(x,y) = \sum_{a,b} \alpha_{ab} x^a y^b \in \mathcal{O}_2$ vanishes identically on the y-axis; that is, $0 \equiv f(0,y) = \sum_b \alpha_{0b} y^b$. It follows that all $\alpha_{0b} = 0$ for all b, so that $f = x^\nu f_0$ where $\nu > 0$ and $f_0(0,y) \not\equiv 0$. We must prove unique factorization for f_0.

Let $f \in \mathcal{O}_2$ with $f(0,y) \not\equiv 0$. Lemmas 10.1.4 and 10.2.1 give $f = uw$ uniquely. Since w belongs to the UFD $\mathbb{C}\{x\}[y]$, we have a unique decomposition $w = h_1 \cdot \cdots \cdot h_\ell$ into irreducibles $h_j \in \mathbb{C}\{x\}[y]$. Clearly also $h_j(0,y) \not\equiv 0$, and so Lemma 10.1.4 applied to each h_j gives uniquely $h_j = u_j w_j$, with each w_j a Weierstrass polynomial irreducible in $\mathbb{C}\{x\}[y]$ (since h_j is). This yields $w = (u_1 w_1) \cdot \cdots \cdot (u_\ell w_\ell) = (u_1 \cdots u_\ell) w_1 \cdots w_\ell =: \tilde{u}\tilde{w}$, and by Lemma 10.2.1 $\tilde{u}$ must be 1. So far we have $f = uw_1 \cdots w_\ell$.

We do not know yet whether w_j is irreducible in $\mathcal{O}_2$. If $w_j = v'v''$ ($v', v'' \in \mathcal{O}_2$), then $w_j(0,y) \not\equiv 0 \implies$ the same thing for v', v''. Lemma 10.1.4 applies to yield $v' = u'w'$ and $v'' = u''w''$, so that $w_j = (u'u'')(w'w'')$; applying Lemma 10.2.1 yet again gives $u'u'' = 1 \implies w_j = w'w''$. But $w', w'' \in \mathfrak{W} \subset \mathbb{C}\{x\}[y]$, contradicting irreducibility of w_j in $\mathbb{C}\{x\}[y]$.

To see uniqueness, write factorizations $f = f_1 \cdots f_\ell = g_1 \cdots g_k$ into irreducibles in $\mathcal{O}_2$; we may assume $f(0,y) \not\equiv 0$. Then Lemma 10.1.4 gives $f_j = u_j w_j$ and $g_i = \tilde{u}_i \tilde{w}_i$ with $w_j, \tilde{w}_i$ irreducible Weierstrass polynomials. We then have $(u_1 \cdots u_\ell)(w_1 \cdots w_\ell) = (\tilde{u}_1 \cdots \tilde{u}_k)(\tilde{w}_1 \cdots \tilde{w}_k)$, so that by Lemma 10.2.1 $u_1 \cdots u_\ell = \tilde{u}_1 \cdots \tilde{u}_k$ and $w_1 \cdots w_\ell = \tilde{w}_1 \cdots \tilde{w}_k$. By uniqueness of factorization in $\mathbb{C}\{x\}[y]$ (and Lemma 10.2.1), the $\{w_j\}$ and $\{\tilde{w}_i\}$ are the same (up to reordering), and $\ell = k$. $\square$

Note the key statement that comes out of this proof: given $f \in \mathcal{O}_2$ with $f(0,y) \not\equiv 0$, we have

$$f = uw_1 \cdots w_\ell, \tag{10.2.1}$$

where $u \in \mathfrak{U}$ and w_i are Weierstrass polynomials which are irreducible (as Weierstrass polynomials, as elements of $\mathbb{C}\{x\}[y]$, and as elements of $\mathcal{O}_2$). Moreover, this decomposition is completely unique, up to reordering of the w_i. Finally – this also comes out of the proof – if f

was a Weierstrass polynomial, then $u = 1$ in (10.2.1), and $\deg_y(f) = \sum_{i=1}^{\ell} \deg_y(w_i)$. This will be useful in working the following problems.

Exercises

(1) Show $f(x, y) = x^3 - x^2 + y^2$ is (a) irreducible in $\mathbb{C}[x, y]$ and (b) reducible in $\mathbb{C}\{x\}[y]$.

(2) Show $g(x, y) = x^3 - y^2$ is irreducible in $\mathbb{C}\{x\}[y]$.

CHAPTER 11

Proof of the normalization theorem

The purpose of this chapter is twofold: to find a method for explicitly parametrizing neighborhoods of singular points on algebraic curves; and, using this, to completely prove part (A) of Theorem 3.2.1. In fact, we shall prove a stronger result which contains a uniqueness statement:

THEOREM 11.0.1. *Let $C \subset \mathbb{P}^2$ be an irreducible algebraic curve, with $\mathcal{S} = sing(C)$ its set of singular points. Then there exists a Riemann surface $\tilde{C}$ and morphism (of complex manifolds) $\sigma : \tilde{C} \to \mathbb{P}^2$ such that*

(a) $\sigma(\tilde{C}) = C$
(b) $\#\{\sigma^{-1}(\mathcal{S})\} < \infty$
(c) $\sigma : (\tilde{C}\backslash\sigma^{-1}(\mathcal{S})) \to (C\backslash\mathcal{S}) =: C^*$ is injective (hence an isomorphism).

The pair $(\tilde{C}, \sigma)$ is called the normalization *of C, and is unique in the sense that if $(\tilde{C}', \sigma')$ is another, then there exists a morphism $\tau : \tilde{C} \xrightarrow{\cong} \tilde{C}'$ such that $\sigma = \sigma' \circ \tau$.*

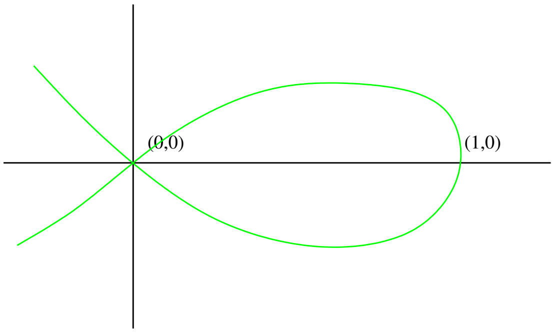
We remark that in the correspondence (cf. §9.3) between ideals $I \subset \mathbb{C}[x, y]$, varieties $V = V(I)$, and rings $\mathbb{C}[V] = \frac{\mathbb{C}[x,y]}{I}$, "normalization" means taking the *integral closure of* $\mathbb{C}[V]$ *in* $\mathbb{C}(V)$. Taking "Spec" of the result produces an affine variety $\tilde{V}$ with a morphism to V. This procedure may be carried out for projective varieties by patching affine ones together, and if this is done for curves ($V = C$), then $\tilde{V}$ is really just $\tilde{C}$ constructed algebraically. While this is beyond the scope of our course, it's instructive to look at an example.

EXAMPLE 11.0.2. If we take $V = \{x^3 - x^2 + y^2 = 0\} \subset \mathbb{C}^2$, then the coordinate ring

$$\mathbb{C}[V] = \frac{\mathbb{C}[x, y]}{(x^3 - x^2 + y^2)}$$

is not integrally closed in its fraction field $\mathbb{C}(V)$. That is, the equation $\xi^2 + (x - 1) = 0$, while irreducible in $\mathbb{C}[V][\xi]$, is "solved" by $\xi = \frac{y}{x}$, as

$\frac{y^2}{x^2} \equiv \frac{x^2-x^3}{x^2} = 1-x$ in $\mathbb{C}(V)$. A schematic picture of the irreducible cubic curve V is



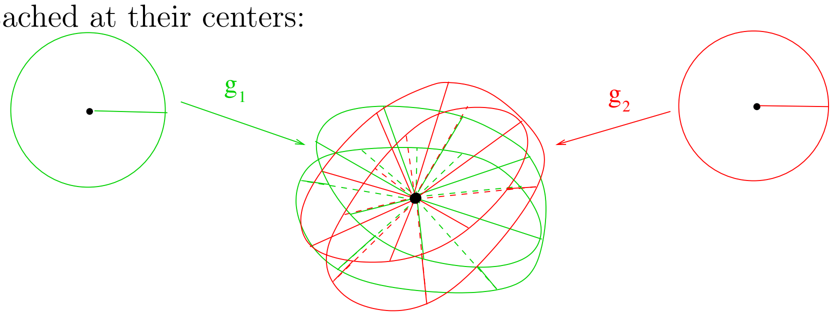
and $\frac{y}{x}$ can be viewed as "separating the branches" of V at the singular point $(0,0)$.

In an exercise above you were asked to carry out the (local) analytic approach, proving that $x^3-x^2+y^2$ is irreducible in $\mathbb{C}[x,y]$ but *reducible* in $\mathbb{C}\{x\}[y]$. Here is another such example.

EXAMPLE 11.0.3. Consider the equation $y^4+x^3-x^2(=0)$, which is irreducible in $\mathbb{C}[x,y]$ but reducible in $\mathbb{C}\{x\}[y]$, into the product of Weierstrass polynomials

$$(y^2 - x\sqrt{1-x})(y^2 + x\sqrt{1-x}).$$

Here $x\sqrt{1-x}$ is regarded as a convergent power series (in $\mathbb{C}\{x\}$) vanishing at $x=0$. The local picture (near $(0,0)$) described by this factorization is of two "parking lots" (topologically, these are just disks) attached at their centers:



We need a procedure that gives the indicated holomorphic parametrizations of these two branches.

11.1. Overview

Informally, here is the main idea of the proof of Theorem 11.0.3. Given an irreducible algebraic curve C with singular point p, we may use a $PGL(3, \mathbb{C})$-transformation (i.e. a projectivity) of $\mathbb{P}^2$ to move $p \mapsto [1:0:0]$. By another linear transformation of coordinates (cf. §8.2), we can put the affine equation in the form

$$f = y^n + a_1(x)y^{n-1} + \cdots + a_n(x)\,(=0)\ , \quad a_j(x) \in \mathbb{C}[x]. \tag{11.1.1}$$

Now we want to normalize a neighborhood of the singularity $(0, 0)$. Since f is irreducible in $\mathbb{C}[x][y]$, its discriminant $\mathcal{D}(f)(x)$ is not identically zero in $\mathbb{C}[x]$. Hence the "local factorization" $f = f_1^{m_1} \cdots f_\ell^{m_\ell}$ into irreducibles in $\mathbb{C}\{x\}[y]$ will have no repeated factors (all $m_i = 1$). Writing $\Delta = \{|x| < \rho\}$ and $W_\Delta = \Delta \times \{|y| < \epsilon\}$ for sufficiently small $\rho, \epsilon > 0$, this corresponds to the decomposition of $C \cap W_\Delta$ into $C_1^\Delta \cup \cdots \cup C_\ell^\Delta$, where each C_i^Δ is homeomorphic to a disk and the union attaches them only at their centres.

More precisely, writing

$$f = uw_1 \cdots w_\ell$$

as in §10.2, the C_i^Δ are the zero-loci $\{w_i = 0\}$ of irreducible Weierstrass polynomials. If we can write down 1-to-1 holomorphic maps $\tilde{\varphi}_i : \tilde{\Delta} \to \mathbb{C}^2$ ($\tilde{\Delta}$ is some other disk related to Δ) with image $\tilde{\varphi}_i(\tilde{\Delta}) = C_i^\Delta$, and repeat this procedure over all singular points, then the normalization $\tilde{C}$ can be constructed as follows. On $C^* = C \backslash \mathrm{sing}(C)$, we have a covering by holomorphic parametrizations $\varphi_\alpha = z_\alpha^{-1}$ (from §7.2). Composing the $\tilde{\varphi}_i$ with the z_α whenever $C_i^\Delta \cap U_\alpha$ in nonempty, yields holomorphic transition functions. Thinking of C^* as an abstract complex 1-manifold, these transition functions indicate how to attach each C_i^Δ to C^* to yield a new complex 1-manifold $\tilde{C}$. To obtain C (topologically) from this, you simply reattach the centers of the C_i^Δ.

The first step indicated in his outline, which we do not yet know how to do, was the construction of the $\{\tilde{\varphi}_i\}$. We shall now do this.

11.2. Irreducible local normalization

Let $w = y^k + b_1(x)y^{k-1} + \cdots + b_k(x)$ be a Weierstrass polynomial, irreducible in $\mathbb{C}\{x\}[y]$. Unless $k = 1$, the discriminant $(\mathcal{D}(w))(x)$ has a zero at $x = 0$. Since $\mathcal{D}(w)$ is not identically zero, this zero is isolated, and we can take ρ small enough that $x = 0$ is its only zero on $\Delta = \{|x| < \rho\}$.

Now, there is a factorization $w = \prod_{\nu=1}^{k}(y - y_\nu(x))$ which is valid in the sense of §8.2, but *not* in $\mathbb{C}\{x\}[y]$. Namely, the $\{y_\nu(x)\}$ are "multivalued" on Δ,[1] but become well-defined on Δ minus a slit. (Another, more algebraic, way to think of this factorization, if $0 < |x_0| < \rho$, is as taking place in $\mathbb{C}\{x - x_0\}[y]$.) The multivaluedness is manifested as follows: by the heredity principle, going once counterclockwise around the origin in Δ^*, permutes the roots of w by $y_\nu(x) \mapsto y_{\tau(\nu)}(x)$ where $\tau \in \mathfrak{S}_k$(=the symmetric group on k elements). This permutation must be transitive, i.e. a k-cycle: otherwise, it splits into a product of (smaller) cycles, each of which gives rise to an irreducible proper factor of w in $\mathbb{C}\{x\}[y]$, in contradiction to its irreducibility.

Here, then, is how to parametrize the set $\{w = 0\} \subset W_\Delta$:

PROPOSITION 11.2.1. *Let $w \in \mathfrak{W}$ be irreducible of degree k, and pick any $\nu \in \{1, \ldots, k\}$. Then writing $\tilde{\Delta} := \{t \in \mathbb{C} | \, |t| < \rho^{\frac{1}{k}}\}$,*

$$g : \tilde{\Delta} \to \mathbb{C}^2$$

$$t \mapsto (t^k, \tilde{y}_\nu(t^k))$$

is well-defined and injective,[2] with image the local analytic curve

$$C^\Delta := \{(x, y) \,|\, w(x, y) = 0, \ |x| < \rho, \ |y| < \epsilon\},$$

and gives a biholomorphism (of complex 1-manifolds)

$$\tilde{\Delta} \backslash \{0\} \overset{\cong}{\to} C^\Delta \backslash \{(0, 0)\}.$$

REMARK 11.2.2. Here $C^\Delta \backslash \{(0, 0)\}$ is a complex 1-manifold by the holomorphic implicit function theorem as in §7.2, and is covered by neighborhoods with local holomorphic coordinate x. One can regard the last biholomorphism as giving the transition function between (Δ_0, t) and (more generally) *any* open set in C^* with holomorphic coordinate x.

PROOF. (of Prop. 11.2.1) Recall that $y_\nu(x)$ is well-defined on the slit disk $\Delta^- := \Delta \backslash \{x \in \mathbb{R}_{\geq 0}\}$. Analytic continuation of $y_\nu(x)$ once counterclockwise around $x = 0$ yields $y_{\tau(\nu)}(x)$; going around once more gives $y_{\tau^2(\nu)}(x)$, and so on. Since τ is a k-cycle, $\tau^k(\nu) = \nu$ and *going around zero k times* returns us to $y_\nu(x)$. But t^k does *precisely this* when

[1] except at 0, since all $y_\nu(0) = 0$

[2] The meaning of "$\tilde{y}_\nu(t^k)$" will be defined in proof.

t goes around 0 once, and so $y_\nu(t^k)$ extends to a well-defined analytic function on $\tilde{\Delta}$.

A bit more carefully, we subdivide $\tilde{\Delta}^* = \cup_{j=0}^{k-1}\tilde{\Delta}^{(j)}$ into pie-slices $\tilde{\Delta}^{(j)} := \{0 < |t| < \rho^{\frac{1}{k}}$ and $\frac{j}{2\pi} \leq \arg(t) \leq \frac{j+1}{2\pi}\}$. On the interior of each slice (that is, where $\frac{j}{2\pi} < \arg(t) < \frac{j+1}{2\pi}$), we can define a holomorphic function by $y_{\tau^j(\nu)}(t^k)$, since $t \mapsto t^k$ maps this interior (isomorphically) to Δ^- where $y_{\tau^j(\nu)}(x)$ is defined. Extending these functions continuously to $\tilde{\Delta}^{(j)}$, they patch together (in fact, analytically continue into one another) to yield a single holomorphic function $\tilde{y}_\nu(t^k)$ on $\tilde{\Delta}^*$. This is bounded exactly as in §8.2, and so extends to $\mathcal{O}(\tilde{\Delta})$ by the removable singularity theorem.

Let $\zeta_k := e^{\frac{2\pi\sqrt{-1}}{k}}$. If $(t_1^k, \tilde{y}_\nu(t^k)) = (t_2^k, \tilde{y}_\nu(t_2^k))$ then

$$t_2 = (\zeta_k)^\ell t_1$$

for some $\ell \in \mathbb{Z}$, and

$$y_{\tau^\ell(\nu)}(t_1^k) = y_\nu(t_1^k).$$

Since the $\{y_\nu\}$ are all distinct away from 0, the last equation is impossible unless $k|\ell$, which implies $(\zeta_k)^\ell = 1$ so that $t_1 = t_2$. This proves that g is injective.

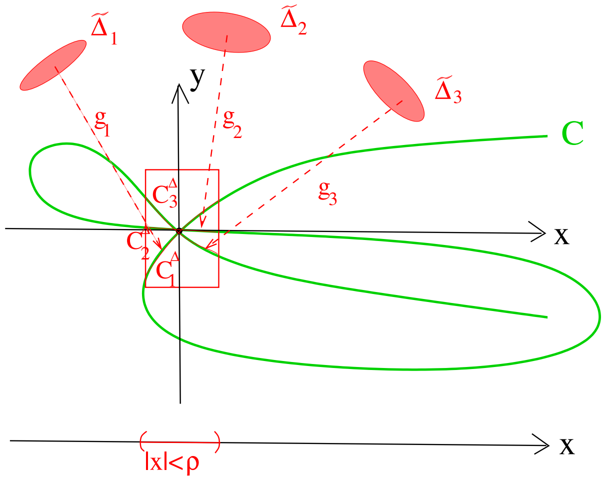
Since τ is transitive, g maps surjectively onto C^Δ. It gives, finally, a holomorphic map of Riemann surfaces on the complement of 0 since in local coordinates t (on $\tilde{\Delta}^*$) and x (for open subsets covering $C^\Delta \backslash \{(0,0)\}$) we have $x =$"$g(t)$'s x-coordinate"$= t^k$. $\square$

11.3. Finishing local normalization

Referring back to §8.1, for each of the irreducible factors w_j of f we now apply Proposition 11.2.1. This yields normalizations

$$g_j : \tilde{\Delta}_j \to C_j^\Delta$$

of the irreducible components of the local analytic curve $C \cap W_\Delta$:



Each restriction

$$g_j^\circ : \tilde{\Delta}_j^* \xrightarrow{\cong} (C_j^\Delta \backslash \{(0,0)\}) \hookrightarrow C^*$$

is biholomorphic with respect to local coordinates on $\tilde{\Delta}_j$ and an open covering of $C_j^\Delta \backslash \{(0,0)\}$. In fact, it takes the form $t \mapsto t^k (= x)$ as indicated at the end of the last proof. These may be regarded as the "glueing" maps that will attach each $\tilde{\Delta}_j$ to C^* thereby plugging the holes in C^*, which is what we do next.

Before that, we just note that one should carry out the construction of g_j's as we have done near $p = (0,0)$, at all the other singular points of C.

11.4. Global normalization (patching)

Suppose for the moment $(0,0)$ is the *only* singular point of C, so that $C^* = C \backslash \{(0,0)\}$. Then we put

$$\tilde{C} := C^* \underset{g_1^\circ}{\cup} \tilde{\Delta}_1 \underset{g_2^\circ}{\cup} \tilde{\Delta}_2 \cup \cdots \underset{g_\ell^\circ}{\cup} \Delta_\ell \,,$$

where $C^* \underset{g_1^\circ}{\cup} \tilde{\Delta}_1$ means

$$\frac{C^* \amalg \tilde{\Delta}_1}{g_1^\circ(p) \sim p \ \ (\forall p \in \tilde{\Delta}_1^*)},$$

$C^* \underset{g_1^\circ}{\cup} \tilde{\Delta}_1 \underset{g_2^\circ}{\cup} \tilde{\Delta}_2$ means

$$\frac{C^* \amalg \tilde{\Delta}_1 \amalg \tilde{\Delta}_2}{g_1^\circ(p) \sim p\,,\ g_2^\circ(q) \sim q \ \ (\forall p \in \tilde{\Delta}_1^*, q \in \tilde{\Delta}_2^*)},$$

and so forth.

If there are more singularities, then repeat this patching at each point in $\mathcal{S} = \mathrm{sing}(C)$.

To get a map $\sigma : \tilde{C} \to \mathbb{P}^2$ with image C, set

$$\sigma(c) := \begin{cases} c \ , \text{ for } c \in C^* \\ g_j(c), \text{ for } c \in \tilde{\Delta}_j \end{cases} .$$

These two prescriptions are compatible with the patching.

To see that $\tilde{C}$ is compact: given an open cover $\{U_\alpha\}$ of $\tilde{C}$, pick one $U_{\alpha(q)}$ containing each $q \in \sigma^{-1}(\mathcal{S})$. The complement $\tilde{C}'$ of these in $\tilde{C}$ is isomorphic to a closed subset of C, since σ is bijective away from $\sigma^{-1}(\mathcal{S})$. Now a closed subset of C is a closed subset in $\mathbb{P}^2$, $\mathbb{P}^2$ is compact, and a closed subset of a compact set is compact. So $\tilde{C}'$ is compact and $\{U_\alpha \cap \tilde{C}'\}$ has a finite subcover $\{U_i \cap \tilde{C}'\}$. The $\{U_i\}$ together with the $\{U_{\alpha(q)}\}$ then furnish a finite subcover of $\tilde{C}$.

We have now proved all but the uniqueness part of 11.0.3 and it is time to backtrack and get explicit.

11.5. Examples of local normalization

EXAMPLE 11.5.1. Assuming $\gcd(k, a) = 1$,

$$y^k - x^a$$

is irreducible in $\mathbb{C}\{x\}[y]$, and we shall apply the procedure of Prop. 11.2.1. The n (multivalued) roots of $y^k - x^a = 0$ in y are

$$y_1(x) = \sqrt[k]{x^a}\,,\ y_2(x) = \zeta_k \sqrt[k]{x^a}\,,\ \ldots\,,\ y_k(x) = (\zeta_k)^{k-1} \sqrt[k]{x^a};$$

they are well defined on the slit disk $\{0 < |x| < \rho,\ \arg(x) \in (0, 2\pi)\}$. If we plug t^k into $y_1(x)$ and analytically continue, we get

$$\tilde{y}_1(t^k) = t^a.$$

Hence by definition

$$g(t) = (t^k, t^a).$$

We should check that the image of g lies in $y^k - x^a = 0$: this is just the statement that $(t^k)^a = (t^a)^k$.

Example 11.5.2. Here is a more complicated example where there is more than one g_j (as in §5.3):

$$f = y^8 + y^4 - x^6 + x^3 - x^2y^4 + x^5 - x^2.$$

Viewed in $\mathbb{C}\{x\}[y]$, this is not a Weierstrass polynomial (the coefficient $1 - x^2$ of y^4 is not zero at $x = 0$), so we should expect a nontrivial unit u in (10.2.1). Indeed,

$$\begin{aligned} f &= (y^4 - x^3 + 1)(y^4 + x^3 - x^2) \\ &= \underbrace{(y^4 - x^3 + 1)}_{u}\underbrace{(y^2 - x\sqrt{1-x})}_{w_1}\underbrace{(y^2 + x\sqrt{1-x})}_{w_2}, \end{aligned}$$

where u is a unit because $u(0,0) \neq 0$.

Now w_1, w_2 are irreducible Weierstrass polynomials and so we apply 11.2.1 (with $k = 2$) to normalize their zero-sets.

Beginning with w_1, the roots are $y_{11}(x) = \sqrt{x\sqrt{1-x}}$ and $y_{12}(x) = -\sqrt{x\sqrt{1-x}}$, which are swapped as x goes around 0. So $\tilde{y}_{11}(t^2)$ is obtained by substituting t^2 for x and analytically continuing: informally, $\sqrt{t^2\sqrt{1-t^2}} = t\sqrt[4]{1-t^2}$. So

$$g_1(t) = (t^2, t\sqrt[4]{1-t^2}).$$

For w_2, the roots are $y_{21}(x) = i\sqrt{x\sqrt{1-x}}$ and $y_{22}(x) = -i\sqrt{x\sqrt{1-x}}$; and this yields

$$g_2(t) = (t^2, it\sqrt[4]{1-t^2}).$$

Let's check this parametrizes $w_2 = 0$: one need only write $(y(t))^2 + x(t)\sqrt{1-x(t)} = (it\sqrt[4]{1-t^2})^2 + t^2\sqrt{1-t^2} = 0$.

11.6. Uniqueness

Begin with two normalizations:

$$\begin{tikzcd}
\tilde{C} \arrow[r, "\sigma"] & \mathbb{P}^2 & \tilde{C}' \arrow[l, "\sigma'"'] \\
\tilde{C} \setminus \sigma^{-1}(\mathcal{S}) \arrow[u, hook] \arrow[r, "\cong"] \arrow[rr, dashed, bend right] & C \setminus \mathcal{S} \arrow[u, hook] & \tilde{C}' \setminus (\sigma')^{-1}(\mathcal{S}) \arrow[u, hook] \arrow[l, "\cong"']
\end{tikzcd}$$

with σ, σ' holomorphic maps of complex manifolds. Essentially what we have to show is that neighborhoods of the points of $\sigma^{-1}(\mathcal{S})$ (in $\tilde{C}$) and $(\sigma')^{-1}(\mathcal{S})$ (in $\tilde{C}'$) are isomorphic in a way which is compatible with

σ and σ'. Put together with the bottom dotted arrow these isomorphisms will yield the desired map $\tau : \tilde{C} \to \tilde{C}'$ of Riemann surfaces making the diagram

$$\begin{tikzcd} \tilde{C} \arrow[rr, "\tau"] \arrow[dr, "\sigma"'] & & \tilde{C}' \arrow[dl, "\sigma'"] \\ & C & \end{tikzcd}$$

To start, let $p \in \mathcal{S} \subset C$ be a singular point and $U \subset \mathbb{P}^2$ be a small open set containing it. For simplicity assume $p = [1:0:0]$ and choose coordinates so that $U \subset \{|x| < \rho,\ |y| < \epsilon\}$ and $C \cap U$ is given by the vanishing set of a Weierstrass polynomial. Write $U^* := U \backslash \{p\}$.

Now pick $q \in \sigma^{-1}(p)$; by continuity of σ, $\sigma^{-1}(U)$ is open in $\tilde{C}$. So there exists an open set V, which we may assume to be connected, with $q \in V \subset \sigma^{-1}(U)$. Since $\sigma(V \backslash \{q\}) \subset C \cap U^*$ must then be connected, and $C \cap U^*$ is homeomorphic to a disjoint union of punctured disks, σ maps $V \backslash \{q\}$ into one of these punctured disks. Consequently, V is mapped into only one (local) irreducible component[3] W of $C \cap U$. This yields the following diagram:

$$\begin{tikzcd} & V \arrow[r, "\tilde{\sigma}"] \arrow[d, "\sigma"] \arrow[dr, dashed, "\mathsf{X}"] \arrow[rr, dashed, bend left, "\mathsf{T}"] & W \arrow[d, "\pi"] & \left(\{|t| < \rho^{\frac{1}{k}}\} \subset \mathbb{C}_t\right) \arrow[l, "\cong"', "g"] \arrow[dl, "t \mapsto t^k"] \\ & \left(U \subset \mathbb{C}^2_{(x,y)}\right) \arrow[r, "pr_x"] & (\{|x| < \rho\} \subset \mathbb{C}_x) & \end{tikzcd}$$

in which pr_x and σ are morphisms of complex manifolds, so that their composition X is evidently a holomorphic (obviously bounded) function on V.

The composition T is also evidently a bounded, well-defined function on V. By the holomorphic IFT (and holomorphicity of X), it is holomorphic on $V \backslash \{q\}$; hence by the removable singularity theorem, $\mathsf{T} \in \mathcal{O}(V)$. It is also clear that $\mathsf{T}(0) = 0$. So by the open mapping theorem, T maps V onto a neighborhood $\mathcal{N}$ of 0 in $\mathbb{C}_t$ (which we may assume is a disk). Shrinking U (and thus W) if necessary, we may conclude that $\tilde{\sigma}$ – the restriction of σ to a neighborhood of q – maps V onto W. From the diagram, this $\tilde{\sigma}$ is just $g \circ \mathsf{T}$.

[3] remember that these components are homeomorphic to disks; take out p and that is where the punctured disks came from

Since σ is 1-to-1 off $\sigma^{-1}(\mathcal{S})$, no neighborhood of any other point $q_0 \in \sigma^{-1}(\mathcal{S})$ can be sent to W. Repeating the argument above by varying q, sets up a 1-to-1 correspondence between "V's" (i.e. neighborhoods of points in $\sigma^{-1}(\mathcal{S})$ in $\tilde{C}$) and "W's" (irreducible local components of C at points of $\mathcal{S}$). We can play the same game for the normalization $\tilde{C}'$, and find that for a unique $q' \in (\sigma')^{-1}(\mathcal{S})$ we have a neighborhood V' and an isomorphism $\mathsf{T}' : V' \to \mathcal{N}$ whose composition with g gives $\tilde{\sigma}' : V' \to W$.

The piece of τ carrying (V, q) to (V', q') is now defined simply by $(\mathsf{T}')^{-1} \circ \mathsf{T}$. This is automatically holomorphic, and its composition with σ' is $g \circ \mathsf{T} = \sigma$ as desired.

Exercises

(1) Locally normalize the zero-set of $f(x, y) = y^4 - (x + 1)^7$ at $(-1, 0)$.

(2) Locally normalize the zero-set of $g(x, y) = y^4 - x^6 + x^7$ at $(0, 0)$.

CHAPTER 12

Intersections of curves

Now we come to the applications of normalization, which will occupy this chapter and Chapter 14. You may recall that in Chapter 2 we studied intersections of an plane algebraic curve C with a (projective) line L. The points of $L \cap C$ were each assigned a multiplicity by restricting the equation of C under a parametrization of L, and looking at the multplicities of the roots of the resulting one-variable polynomial. With this definition, the multiplicities added up to the degree of the curve (cf. Prop. 2.1.8).

If we had tried to replace L by an arbitrary curve E at that point, we would have run into the problem of no longer knowing how to locally parametrize E near the intersection points. Now that we can do this (Prop. 11.2.1), we can pull the defining equation of C back under the parametrization and look at its order of vanishing at the intersection point. This leads to the general definition of *intersection multiplicity*, and with this in hand that we can finally state (and prove!) Bezout's theorem in general. In its proof the *intersection divisor* will make an appearance, so we begin with a short bit on divisors.

12.1. Divisors on a Riemann surface

Let M be a Riemann surface. The group of *divisors* on M is the free abelian group on points of M,

$$\mathrm{Div}(M) := \left\{ \sum_{\text{finite}} m_i[p_i] \;\middle|\; m_i \in \mathbb{Z}\,,\ p_i \in M \right\}.$$

The uncountably many symbols $[p_i]$ are the generators of this (very big) abelian group. Associated to a divisor $D = \sum m_i[p_i] \in \mathrm{Div}(M)$ is a *degree*

$$\deg(D) := \sum m_i.$$

The resulting group homomorphism

$$\mathrm{Div}(M) \xrightarrow{\deg} \mathbb{Z}. \tag{12.1.1}$$

is called the *degree map*.

The divisor of a (nontrivial) meromorphic function f is given by

$$(f) := \sum_{p \in M} \nu_p(f).[p] \ \in \ \mathrm{Div}(M),$$

where $\nu_p(f)$ is the order of f at p (Defn. 3.1.4). Note that the sum is actually finite (as required by the definition of divisor) since at all but finitely many points of M, $\nu_p(f) = 0$. Now $\mathcal{K}(M)^*$ is a multiplicative abelian group. Sending $f \to (f)$ yields a homomorphism

$$\mathcal{K}(M)^* \xrightarrow{(\cdot)} \mathrm{Div}(M) \tag{12.1.2}$$

of abelian groups, as you will show in an exercise below, which takes multplication to addition: $(fg) = (f) + (g)$, $(f^{-1}) = -(f)$.

With these definitions, the composition of (12.1.2) with (12.1.1) takes f to $\sum_{p \in M} \nu_p(f)$, which by Exercise 3.2 is zero. That is, $\deg \circ (\cdot) = 0$. Note that one can define meromorphic functions and divisors more generally on complex 1-manifolds, but it is only in the compact case (Riemann surfaces) that the divisors of meromorphic functions are always of degree 0.

EXAMPLE 12.1.1. On $\mathbb{P}^1$, the easiest meromorphic function around is $z = \frac{Z_1}{Z_0} \in \mathcal{K}(\mathbb{P}^1)^*$. Writing simply $0, \infty$ for the points $[1:0], [0:1]$, its divisor is $(z) = [0] - [\infty]$, obviously of degree 0.

12.2. Intersection multiplicities

For a polynomial in one variable $f(x)$ with $f(0) = 0$, $\deg(f)$ is the exponent of the highest degree term, while the order of vanishing $\mathrm{ord}_0(f) := \nu_0(f)$ is the exponent of the term of lowest degree. Order (unlike degree) also makes sense for power series in 1 variable.

How does all this generalize to two variables? First, a polynomial $F(x, y)$ can be written as a sum of homogeneous terms. If this is $F = F_k + F_{k+1} + \cdots + F_{d-1} + F_d$, then $\deg(F) := d$ (highest homogeneous degree) while $\mathrm{ord}_{(0,0)}(F) := k$ (lowest homogeneous degree). From §6.4, k is also the order of singularity of the curve $C = \{F = 0\}$ at $(0, 0)$, i.e. the number of tangent lines to C counted with multiplicity. When we don't want to refer to the polynomial, we will write $\mathrm{ord}_{(0,0)}C$; remember this is 1 when C is smooth at $(0,0)$, 2 when C has an ODP (normal crossing) there, and so on. Finally, $\mathrm{ord}_{(0,0)}$ also makes sense for 2-variable power series.

Now suppose $V = \{f(x,y) = 0\}$, $W = \{h(x,y) = 0\}$ are reduced affine algebraic curves that intersect properly — i.e. have no common irreducible components. Then $V \cup W$ has no repeated components, so is itself reduced. For $p \in V \cap W$,

$$(\frac{\partial}{\partial x}(fh))(p) = f_x(p)h(p) + h_x(p)f(p) = f_x(p).0 + h_x(p).0 = 0$$

and similarly $(\frac{\partial}{\partial y}(fh))(p) = 0$. Therefore $V \cap W \subset \mathrm{sing}(V \cup W)$, and Prop. 8.1.6 yields

$$\#\{V \cap W\} \leq \#\{\mathrm{sing}(V \cup W)\} < \infty. \tag{12.2.1}$$

DEFINITION 12.2.1. Assume V and W are irreducible (and distinct), and let $p \in V \cap W$. Let $U \subset \mathbb{C}^2$ be a neighborhood of p. Writing the *local* decomposition of V into irreducibles (uniquely)

$$V \cap U = V_1^{\Delta} + \cdots + V_k^{\Delta},$$

with local normalizations (again, essentially unique)

$$g_i : \Delta \to V_i^{\Delta} \, (\subset \mathbb{C}^2)$$

$$t_i \mapsto (x_i(t), y_i(t)),$$

we define the *(local) intersection multiplicity at* p

$$(V \cdot W)_p := \sum_{i=1}^{k} \mathrm{ord}_0(h(g_i(t))).$$

The *(global) intersection number* is then defined by

$$(V \cdot W) := \sum_{p \in V \cap W} (V \cdot W)_p,$$

in which the sum is finite by (12.2.1).

REMARK 12.2.2. (a) If either V or W is smooth, the intersection number is actually the degree of a divisor,

$$V \cdot W := \sum_{p \in V \cap W} (V \cdot W)_p [p].$$

This is because we can regard the smooth one (say, W) as a Riemann surface and then $V \cdot W \in \mathrm{Div}(W)$. Alternatively, you can think of $V \cdot W$ as a formal sum of points of $\mathbb{P}^2$, known as a *zero-cycle*[1] on $\mathbb{P}^2$. The degree is defined in the same way as for divisors.

(b) The composition $h \circ g_i$ appearing in Defn. 12.2.1 will frequently be written $g_i^*(h)$ – that is, we are pulling the function h back by the local normalization g_i^*.

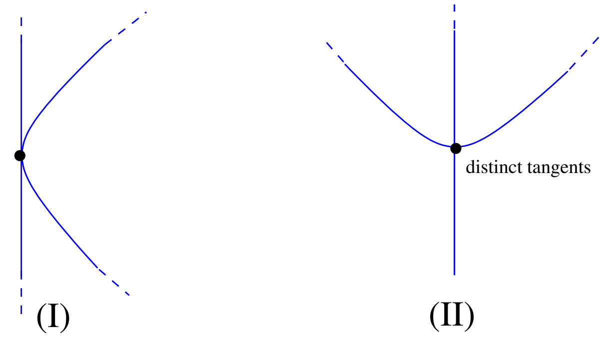
The local intersection multiplicities are well-defined essentially by the uniqueness of local normalizations. They also have some reasonable properties:

PROPOSITION 12.2.3. $(V \cdot W)_p = (W \cdot V)_p$.

PROPOSITION 12.2.4. $(V \cdot W)_p \geq ord_p(V) \cdot ord_p(W)$, *with equality precisely when none of V's tangents at p coincide with the tangents of W at p.*

We will postpone proof of these results in §12.4 − 5, since the proofs get a bit technical.

EXAMPLE 12.2.5. Here are two pictures of smooth curves meeting at a point p:



In each case, $\mathrm{ord}_p V \cdot \mathrm{ord}_p W = 1$ because the curves are smooth. But in the first case, $(V \cdot W) = 2$, while in the second (which has distinct tangents) $(V \cdot W) = 1$.

[1]"zero" refers to the fact that we are taking a formal sum of zero-dimensional subvarieties (i.e. points) in $\mathbb{P}^2$

EXAMPLE 12.2.6. Let $a, b, m, n \in \mathbb{N}$ with

$$\gcd(n, a) = \gcd(m, b) = 1.$$

Then by Prop. 12.2.4, we should have[2]

$$\left(\{y^n = x^a\} \cdot \{y^m = x^b\}\right)_{(0,0)} \geq \min(n, a) \cdot \min(m, b).$$

Let's check this by actually computing the left-hand side. The normalization of $\{y^n = x^a\}$ is just $t \overset{g}{\longmapsto} (t^n, t^a)$ by Example 11.5.1. Writing $h = y^m - x^b$, we have

$$g^*(h) = g^*(y^m - x^b) = (t^a)^m - (t^n)^b = t^{am} - t^{bn}$$

and the order of this at $(0, 0)$ is the least of am and bn:

$$\left(\{y^n = x^a\} \cdot \{y^m = x^b\}\right)_{(0,0)} = \min(am, bn).$$

This clearly satisfies the inequality, and it is easy to cook up an example where equality doesn't hold: with $n = 3$, $a = 4$, $m = 2$, $b = 9$ it becomes $8 \geq 6$.

To extend $(V \cdot W)_p$ to the more general setting where $V = \sum m_j V_j$ and $W = \sum n_k W_k$ with $\{V_j\}$ and $\{W_k\}$ irreducibles, we simply put

$$(V \cdot W)_p := \sum_{j,k} m_j n_k (V_j \cdot W_k).$$

REMARK 12.2.7. Here are two other approaches to local intersection multiplicity which give the same numbers.

(a) The commutative algebra approach makes use of localization. Recall that $\mathbb{C}(x, y)$ denotes the fraction field of $\mathbb{C}[x, y]$. Let $p = (a, b) \in \mathbb{C}^2$. The local ring at p, denoted $\mathcal{O}_p$, is the subset of $\mathbb{C}(x, y)$ consisting of rational functions $\frac{G_1}{G_2}$ (here $G_1, G_2 \in \mathbb{C}[x, y]$) with $G_2(p) \neq 0$. You can easily check that this is a ring, and it obviously contains $\mathbb{C}[x, y]$. It has a unique maximal ideal $\mathfrak{m}_p$ consisting of functions which vanish at p.

Now let $V = \{f = 0\}$, $W = \{h = 0\}$ be as above, and assume $p \in V \cap W$. Writing $(f, h)_p$ for the ideal in $\mathcal{O}_p$ generated by f and h, we define

$$(V \cdot W)_p := \dim_{\mathbb{C}} \left(\mathcal{O}_p/(f, h)_p\right)$$

[2]for instance, the polynomial $y^n - x^a$ has order given by the smallest of n and a.

by viewing the quotient $\mathcal{O}_p/(f,h)_p$ as a vector space. (Note that from this definition, invariance of $(V \cdot W)_p$ under projectivities is immediately clear.) As a simple example, we know that the intersection multiplicity at $p = (0,0)$ of $\{x = 0\}$ and $\{y^2 - x = 0\}$ should be 2. The quotient vector space, indeed, has basis $1, y$. See Chapter 4 of [L. Flatto, Poncelet's Theorem] for more on this approach.

(b) For an approach via resultants, it is convenient to work with homogeneous polynomials. Write $\bar{V} = \{F = 0\}$,$\bar{W} = \{H = 0\}$, $P = [P_0 : P_1 : P_2] \in \bar{V} \cap \bar{W}$ (in homogeneous coordinates $[Z : X : Y]$ on $\mathbb{P}^2$). Assume that $[0 : 0 : 1]$ neither belongs to (i) $C \cup D$, nor (ii) any line joining points of $C \cap D$, nor (iii) any line tangent to C or D at a point of $C \cap D$. Then we may define

$$(\bar{V} \cdot \bar{W})_P := \mathrm{ord}_{[P_0:P_1]}(\mathcal{R}_Y(F,H)).$$

Here we are thinking of F, H as elements of $\mathbb{C}[Z,X][Y]$ and $\mathcal{R}_Y(F,H)$, which eliminates Y,[3] is a polynomial in Z and X; it is in fact homogeneous and of degree $\deg(F) \cdot \deg(H)$. Its order at $[P_0 : P_1]$ is just the highest power of $(P_0X - P_1Z)$ dividing it.

Justifying this definition takes a bit of work, but it leads immediately to a proof of Bezout since the intersection multiplicities have to add up to $\deg \mathcal{R}_Y(F,H) = \deg \bar{V} \cdot \deg \bar{W}$ by construction. This is the point of view taken in [F. Kirwan, Complex Algebraic Curves].

12.3. Bezout's theorem

We first do a quick recap of Prop. 2.1.8:

Proposition 12.3.1. *Let $C = \{F(Z,X,Y) = 0\} \subset \mathbb{P}^2$ be a degree d curve, $L\,(\cong \mathbb{P}^1) \subset \mathbb{P}^2$ a line not contained in C. Then $(L \cdot C) = d$.*

Proof. By a change of coordinates, we may assume $L = \{Y = 0\}$ and $[0 : 1 : 0] \notin C$. Then by the Fundamental Theorem of Algebra,

$$F(Z,X,0) = \prod_{i=1}^{k} (X - \alpha_i Z)^{d_i},$$

where $\sum_{i=1}^{k} d_i = d$ since F has is homogeneous of degree d. Hence $C \cap L = \{[1 : \alpha_i : 0]\}_{i=1}^{k}$.

[3] as usual you can think about this resultant in terms of a projection onto the x- (or rather, $[Z : X]$-) axis.

Passing to affine coordiantes ($f = \prod(x-\alpha_i)^{d_i}$) and locally normalizing L at $(\alpha_i, 0)$ by $t \overset{g_i}{\longmapsto} \alpha_i + t$, we have

$$(L \cdot C)_{(\alpha_i,0)} := \mathrm{ord}_0(g_i^* f) = d_i.$$

We conclude that $(L \cdot C) = \sum d_i = d$. □

THEOREM 12.3.2. [E. BEZOUT, 1779] *Let $C, E \subset \mathbb{P}^2$ be properly intersecting projective algebraic curves. Then $(C \cdot E) = \deg C \cdot \deg E$.*

PROOF. Assume C is irreducible. Let $k = \deg E$, and choose lines $L_1, \ldots, L_k$ avoiding the points of $C \cap E$. Write $E = \{H(Z, X, Y) = 0\}$, $L_j = \{\Lambda_j(Z, X, Y) = 0\}$. Then by Propositions 12.2.3 and 12.3.1,

$$(C \cdot L_j) = (L_j \cdot C) = \deg C,$$

and

$$(C \cdot (\cup_{j=1}^k L_j)) = \sum_{j=1}^k (C \cdot L_j) = \deg C \cdot \deg E.$$

Now by Example 7.3.5, the quotient of two homogeneous polynomials *of the same degree* gives a meromorphic function on projective space. H is of degree k and each Λ_j is of degree 1, so we may define

$$\varphi := \frac{H}{\Lambda_1 \cdot \cdots \cdot \Lambda_k} \in \mathcal{K}(\mathbb{P}^2).$$

Writing $\sigma : \tilde{C} \to \mathbb{P}^2$ (with $\sigma(\tilde{C}) = C$) for the normalization, we have by Example 7.3.6 $\sigma^*\varphi \in \mathcal{K}(\tilde{C})$. We can compute the divisor of this meromorphic function if we notice that locally about each point of $C \cap E$ [resp. $C \cap (\cup L_j)$], φ [resp. $\frac{1}{\varphi}$] gives a defining equation for E [resp. $\cup L_j$]. So by Defn. 12.2.1,

$$(\sigma^*\varphi) = \sum_{p \in C \cap E} \nu_p(\sigma^*\varphi)[p] + \sum_{q \in C \cap (\cup L_j)} \nu_q(\sigma^*\varphi)[q]$$

$$= \sum_{p \in C \cap E} \mathrm{ord}_p(\sigma^*\varphi)[p] - \sum_{q \in C \cap (\cup L_j)} \mathrm{ord}_q(\sigma^* \frac{1}{\varphi})[q]$$

$$= \sum_{p \in C \cap E} (C \cdot E)_p[p] - \sum_{q \in C \cap (\cup L_j)} (C \cdot (\cup L_j))_q[q].$$

But as divisors of meromorphic functions on Riemann surfaces have degree 0,

$$0 = \deg((\sigma^*\varphi)) = \sum_{p \in C \cap E} (C \cdot E)_p - \sum_{q \in C \cap (\cup L_j)} (C \cdot (\cup L_j))_q$$

$$= (C \cdot E) - \deg C \cdot \deg E.$$

Finally, if C is reducible, break it into irreducible components and sum the results! □

REMARK 12.3.3. In terms of zero-cycles (cf. Remark 12.2.2(a)), Bézout is saying that $C \cdot E$ has degree $\deg C \cdot \deg E$.

12.4. Proof of Prop. 12.2.3

We now show the symmetry of intersection numbers. Write $V = \{f = 0\}$, $W = \{h = 0\}$, $p \in V \cap W$. For simplicity assume that $p = (0,0)$, V and W are irreducible, and the defining (polynomial) equations are in the form

$$f = y^m + B_1(x)y^{m-1} + \cdots + B_m(x)\ , \qquad h = y^n + b_1(x)y^{n-1} + \cdots + b_n(x).$$

We decompose these according to (10.2.1): viz.,

$$f = u_1 \cdot v_1 \cdots v_r\ , \qquad h = u_2 \cdot w_1 \cdots w_s$$

where the v_j, w_k are irreducible Weierstrass polynomials. For the roots of v_j [resp. w_k] on a slit disk $\{|x| < \rho,\ x \notin \mathbb{R}_{>0}\}$ we shall write $y_\mu^{(j)}(x)$ ($\mu = 1, \ldots, m_j$) [resp. $z_\nu^{(k)}(x)$ ($\nu = 1, \ldots, n_k$)]. On the non-slit x-disk these become multivalued, and we will assume that counterclockwise analytic continuation sends $y_\mu \mapsto y_{\mu+1}$ to keep the numbering simple. As in §11.2 the $\tilde{y}_\mu^{(j)}(t^{m_j})$ [resp. $\tilde{z}_\nu^{(k)}(t^{n_k})$] are well-defined on a small t-disk $\{|t| < \rho_0\}$, and we have[4]

$$\tilde{y}^{(j)}_{\mu+\mu_0}(t^{m_j}) = \tilde{y}^{(j)}_\mu((\zeta_{m_j}^{\mu_0} t)^{m_j})$$

for some primitive m_j^{th} root of unity ζ_{m_j}. (This changes the branch you start at when $\arg(t) = 0$.) Write $g_j(t) := (t^{m_j}, \tilde{y}_\mu^{(j)}(t^{m_j}))$ and $G_k(t) := (t^{n_k}, \tilde{z}_\nu^{(k)}(t^{n_k}))$ for the parametrizations of $\{v_j = 0\}$ and $\{w_k = 0\}$.

We then have the key identity

$$\pm \prod_{\mu_0=1}^{m_j} w_k\left(t^{n_k m_j},\ \tilde{y}^{(j)}_{\mu+\mu_0}(t^{n_k m_j})\right) \tag{12.4.1}$$

[4]Warning: you cannot write $(\zeta_{m_j}^{\mu_0} t)^{m_j} = (\zeta_{m_j})^{\mu_0 m_j} t^{m_j} = t^{m_j}$ inside the argument of $\tilde{y}_\mu^{(j)}$, since this assumes $\tilde{y}_\mu^{(j)}$ is well-defined on an entire disk (whereas only its composition with the m_j^{th}-power map is!).

$$= \prod_{\mu_0=1}^{m_j} \prod_{\nu_0=1}^{n_k} \left\{ \tilde{y}^{(j)}_{\mu+\mu_0}(t^{n_k m_j}) - \tilde{z}^{(k)}_{\nu+\nu_0}(t^{n_k m_j}) \right\}$$

$$\prod_{\nu_0=1}^{n_k} v_j \left(t^{n_k m_j},\ \tilde{z}^{(k)}_{\nu+\nu_0}(t^{n_k m_j}) \right), \tag{12.4.2}$$

which uses the factorization of each Weierstrass polynomial into a product (for each fixed x) of linear factors. Bearing in mind that rotation of a disk by $2\pi/m_j$ does not change the order of a function at 0, we compute

$$\begin{aligned} \mathrm{ord}_0((12.4.1)) &= n_k \sum_{\mu_0=1}^{m_j} \mathrm{ord}_0 \left(w_k(t^{m_j}, \tilde{y}^{(j)}_{\mu+\mu_0}(t^{m_j})) \right) \\ &= n_k m_j \mathrm{ord}_0 \left(w_k(t^{m_j}, \tilde{y}^{(j)}_{\mu}(t^{m_j})) \right) \\ &= n_k m_j \mathrm{ord}_0(g_j^* w_k). \end{aligned}$$

Dividing this by $m_j n_k$ and applying $\sum_{j=1}^{r} \sum_{k=1}^{s}$ gives

$$\sum_j \mathrm{ord}_0 \left(g_j^* \prod_k w_k \right) = \sum_j \mathrm{ord}_0(g_j^* h) = (V \cdot W)_p.$$

Similarly

$$\mathrm{ord}_0((12.4.2)) = n_k m_j \mathrm{ord}_0(G_k^* v_j),$$

and dividing out $m_j n_k$ and summing yields $(W \cdot V)_p$. Q.E.D.

12.5. Proof of prop. 12.2.4

With the same notation as in the last section, we also write out the irreducible Weierstrass polynomials

$$v_j = y^{m_j} + a^{(j)}_{m_j-1}(x) y^{m_j-1} + \cdots + a^{(j)}_0(x).$$

Note that $a^{(j)}_0(x)$ is the product of the multivalued roots $y^{(j)}_\mu(x)$. We have $\mathrm{ord}_{(0,0)} v_j \le m_j$, $\sum_j \mathrm{ord}_{(0,0)} v_j = \mathrm{ord}_{(0,0)} f$, and

$$\begin{aligned} \mathrm{ord}_{(0,0)}(v_j(x,y)) \le \mathrm{ord}_0(a^{(j)}_0(x)) &= \frac{1}{m_j} \mathrm{ord}_0(a^{(j)}_0(t^{m_j})) \\ &= \frac{1}{m_j} \mathrm{ord}_0 \left(\prod_{\mu_0=1}^{m_j} \tilde{y}^{(j)}_{\mu+\mu_0}(t^{m_j}) \right) \\ &= \mathrm{ord}_0(\tilde{y}^{(j)}_\mu(t^{m_j})). \end{aligned}$$

Therefore

$$\begin{aligned}
(V \cdot W)_p &= \sum_{j=1}^{r} \mathrm{ord}_0 \left(h(t^{m_j}, \tilde{y}_\mu^{(j)}(t^{m_j})) \right) \\
&\geq \sum_{j=1}^{r} (\mathrm{ord}_{(0,0)} h) \cdot \min \left\{ \mathrm{ord}_0(t^{m_j}), \mathrm{ord}_0(\tilde{y}_\mu^{(j)}(t^{m_j})) \right\} \\
&\geq (\mathrm{ord}_{(0,0)} h) \cdot \sum_{j=1}^{r} \min \left\{ \mathrm{ord}_0(t^{m_j}), \mathrm{ord}_{(0,0)}(v_j(x,y)) \right\} \\
&= \mathrm{ord}_{(0,0)} h \cdot \sum_{j=1}^{r} \mathrm{ord}_{(0,0)} v_j \\
&= \mathrm{ord}_{(0,0)} h \cdot \mathrm{ord}_{(0,0)} f \\
&= \mathrm{ord}_p V \cdot \mathrm{ord}_p W \ ,
\end{aligned}$$

Q.E.D.

Exercises

(1) Let M be a Riemann surface. Show that the divisor map $(\cdot) : \mathcal{K}(M)^* \to Div(M)$ is a homomorphism of (abelian) groups. [Hint: use local coordinates.]

(2) Compute the intersection multiplicity $(V \cdot W)_{(0,0)}$ for $V = \{y - \lambda x = 0\}$ and $W = \{y^2 - x^3 = 0\}$. (This will depend on $\lambda \in \mathbb{C}$.)

(3) Let $C \subset \mathbb{P}^2$ be an algebraic curve of degree $n > 1$ and L a (projective) line containing $\left\lfloor \frac{n}{2} \right\rfloor + 1$ singular points of C. (Note: $\lfloor \cdot \rfloor$ is the "greatest integer" function, which takes the greatest integer less than a given real number.) Use Bezout's theorem to prove that $C \supset L$ hence cannot be irreducible. [Hint: prove first that the intersection multiplicity of L and C at each singular point through which L passes, is at least 2.]

(4) Let $C \subset \mathbb{P}^2$ be an algebraic curve of degree 4 with 4 singular points. Using Bezout's theorem and Prop. 12.2.4, prove that C cannot be irreducible. [Hint: use the Hint from (3) together with a conic Q through the following 5 points: the 4 singularities of C plus one more point of C.]

(5) A degree d algebraic curve $C \subset \mathbb{P}^2$ can be taken to go through any $\frac{(d+1)(d+2)}{2} - 1$ distinct points. (This is just because $\dim(S_3^d) = \frac{(d+1)(d+2)}{2}$.) Prove that if all of these points are taken to lie in a single curve E of degree $e < \frac{d}{2} + 1$, then C is reducible.

(6) Compute the intersection multiplicity $(V \cdot W)_{(0,0)}$ for $V = \{y^2 - x^3 = 0\}$ and (a) $W = \{x^3 - x^2 + y^2 = 0\}$ or (b) $W = \{(y-x)^3 - 4\sqrt{2}xy = 0\}$.

CHAPTER 13

Meromorphic 1-forms on a Riemann surface

In the next chapter we will see one more application of the normalization business, via intersection numbers: the degree-genus formula. As more will be needed for its proof, presently we make a detour to define and study differential forms (with poles) on manifolds – how to patch them together via local coordinates, how to pull them back under a morphism, and so forth. Like meromorphic functions, 1-forms have an associated divisor. In contrast to the function case, the degree of this divisor is not zero: it tells you the genus of the Riemann surface, via the so-called Poincaré-Hopf theorem. This result will be key to proving the Riemann-Hurwitz and genus formulae.

13.1. Differential 1-forms

These are the expressions you integrate over paths in calculus and complex analysis. For example, on $\mathbb{R}^2$

$$\eta = F(x,y)dx + G(x,y)dy$$

is a 1-form. Given a differentiable map

$$\Phi : \mathbb{R}^2 \to \mathbb{R}^2$$

given by

$$(u,v) \mapsto (x(u,v), y(u,v)),$$

the *pullback of* η *by* Φ is

(13.1.1)

$$\begin{aligned} \Phi^*\eta \;\; := \;\; & F\left(x(u,v), y(u,v)\right) d(x(u,v)) + G(x(u,v), y(u,v)) d(y(u,v)) \\ = \;\; & \left\{ F(x(u,v), y(u,v)) \tfrac{\partial x}{\partial u}(u,v) + G(x(u,v), y(u,v)) \tfrac{\partial y}{\partial u}(u,v) \right\} du \\ & + \left\{ F(x(u,v), y(u,v)) \tfrac{\partial x}{\partial v}(u,v) + G(x(u,v), y(u,v)) \tfrac{\partial y}{\partial v}(u,v) \right\} dv. \end{aligned}$$

A "0-form" is just a function $f(x,y)$, and

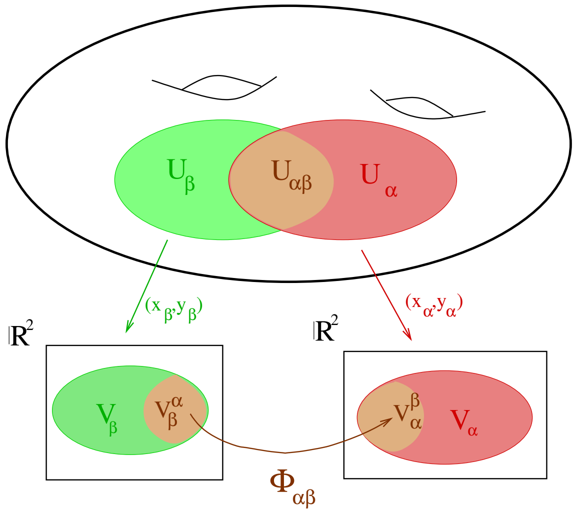
$$\Phi^* f := f \circ \Phi = f\left(x(u,v), y(u,v)\right)$$

is nothing but precomposing with Φ. (13.1.1) is simply the analogue for 1-forms of "precomposition with Φ" . This is exactly what you are doing when you change variables in an integral.

We want to generalize 1-forms from $\mathbb{R}^2$ to real 2-manifolds (and then to complex 1-manifolds), which seems to call for a bit of motivation.

Let M be a differentiable real 2-manifold, $f : M \to \mathbb{R}$ a differentiable function, and $p \in M$ a point. If $M \subset \mathbb{R}^3$, then the notion of "taking partial derivatives of f at p in directions tangent to M" makes immediate sense – you just precompose f with a (differentiable) path in M having a given tangent at p, and differentiate with respect to the variable parametrizing this path.

In abstract differential topology, one has no embedding in $\mathbb{R}^3$. Rather, the differentiability of M is arranged by requiring the transition functions $\Phi_{\alpha\beta}$ relative to local coordinates on an open cover, to be smooth:



(This was discussed at the beginning of §2.2.) One then defines the *tangent spaces*

$$T_pM := \text{vector space of linear differential operators (at } p)$$
$$\cong \mathbb{R}\left\langle \left.\frac{\partial}{\partial x_\alpha}\right|_p, \left.\frac{\partial}{\partial y_\alpha}\right|_p \right\rangle$$

and *tangent bundle*

$$TM := \cup_{p\in M} T_pM.$$

One has a projection map $\pi : TM \to M$ with $\pi^{-1}(p) = T_pM$. A global section of TM, that is, is a smooth[1] map $\sigma : M \to TM$ with $\pi \circ \sigma = \mathrm{id}_M$, is called a *vector field* on M. (Typically one writes $\vec{v}$, with the understanding that $\vec{v}(p) \in T_pM$.)

Now integration is dual to differentiation, so differentials are dual to tangent vectors. For $\frac{\partial}{\partial x_\alpha}, \frac{\partial}{\partial y_\alpha}$ a dual basis (for the dual vector space) is dx_α, dy_α: we write

$$dx\left(\frac{\partial}{\partial x}\right) = 1\ , \quad dy\left(\frac{\partial}{\partial x}\right) = 0\ ,$$
$$dx\left(\frac{\partial}{\partial y}\right) = 0\ , \quad dy\left(\frac{\partial}{\partial y}\right) = 1.$$

The *co*tangent spaces are then

$$T_p^*M \cong \mathbb{R}\left\langle dx_\alpha|_p\, , dy_\alpha|_p \right\rangle .$$

Global sections of the cotangent bundle $T^*M = \cup_{p\in M} T_p^*M$ are then the *differential* 1*-forms* on M. In local coordinates a differential 1-form η looks like:

$$\eta_\alpha = F_\alpha(x_\alpha, y_\alpha)dx_\alpha + G_\alpha(x_\alpha, y_\alpha)dy_\alpha. \tag{13.1.2}$$

Just as a function on M given locally by $\{g_\alpha : V_\alpha \to \mathbb{R}\}$ must satisfy

$$g_\beta|_{V_\beta^\alpha} = \left(g_\alpha|_{V_\alpha^\beta}\right) \circ \Phi_{\alpha\beta} \ \left(= \Phi_{\alpha\beta}^*\left(g_\alpha|_{V_\alpha^\beta}\right)\right),$$

the $\{\eta_\alpha\}$ are subject to compatibility conditions

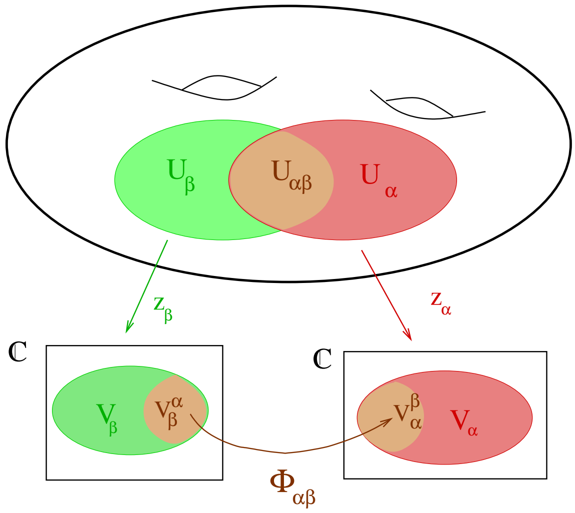
$$\eta_\beta|_{V_\beta^\alpha} = \Phi_{\alpha\beta}^*\left(\eta_\alpha|_{V_\alpha^\beta}\right).$$

Now since M (hence each $\Phi_{\alpha\beta}$) is smooth, smoothness of η_α (i.e. of F_α and G_α in (13.1.2)) is preserved under pullback, and it makes sense to define

$$\begin{aligned} A^1_{\mathbb{R}}(M) \ &:= \ \text{smooth, real-valued 1-forms on } M \\ &= \ \text{collections } \{\eta_\alpha\} \text{ with } \{F_\alpha, G_\alpha\} \text{ infinitely differentiable.} \end{aligned}$$

For a complex 1-manifold, which we recall from is a special kind of smooth real 2-manifold (the $\Phi_{\alpha\beta}$ are conformal), the labels on the diagram change:

[1] to define smoothness one has to put a manifold structure on TM, which I won't do here.



Omitting subscript α's for the moment, and writing a subscript $\mathbb{C}$ to indicate $\otimes_{\mathbb{R}}\mathbb{C}$, one has

$$T_{\mathbb{C},p}M = \mathbb{C}\left\langle \left.\frac{\partial}{\partial x}\right|_p, \left.\frac{\partial}{\partial y}\right|_p \right\rangle \cong \mathbb{C}\left\langle \left.\frac{\partial}{\partial z}\right|_p, \left.\frac{\partial}{\partial \bar{z}}\right|_p \right\rangle$$

$$T^*_{\mathbb{C},p}M \otimes \mathbb{C} = \mathbb{C}\left\langle dx|_p, dy|_p \right\rangle \cong \mathbb{C}\left\langle dz|_p, d\bar{z}|_p \right\rangle,$$

where $\frac{\partial}{\partial z} := \frac{1}{2}\left(\frac{\partial}{\partial x} - \sqrt{-1}\frac{\partial}{\partial y}\right)$, $\frac{\partial}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x} + \sqrt{-1}\frac{\partial}{\partial y}\right)$, $dz := dx + \sqrt{-1}dy$, $d\bar{z} := dx - \sqrt{-1}dy$. (This makes $dz(\frac{\partial}{\partial z}) = 1$, $dz(\frac{\partial}{\partial \bar{z}}) = 0$, $d\bar{z}(\frac{\partial}{\partial z}) = 0$, $d\bar{z}(\frac{\partial}{\partial \bar{z}}) = 1$ so that the bases are dual.) A smooth section of the complexified cotangent bundle $T^*_{\mathbb{C}}M$ thus looks locally like

$$F(x,y)dz + G(x,y)d\bar{z}$$
$$= (F+G)dx + \sqrt{-1}(F-G)dy,$$

for F and G smooth (infinitely differentiable) complex-valued functions. The 1-forms we are after are substantially more restricted:

DEFINITION 13.1.1. A holomorphic [resp. meromorphic] 1-form $\omega \in \Omega^1(M)$ [resp. $\mathcal{K}^1(M)$][2] is a collection of expressions

$$\omega_\alpha = f_\alpha(z_\alpha)dz_\alpha\,, \quad \text{with } f_\alpha : V_\alpha \to \mathbb{C} \text{ holomorphic [resp. meromorphic]},$$

satisfying

$$\omega_\beta|_{V^\alpha_\beta} = \Phi^*_{\alpha\beta}\left(\omega_\alpha|_{V^\beta_\alpha}\right) \quad \forall \alpha, \beta. \tag{13.1.3}$$

[2]recall the notation $\mathcal{K}(M)$ for meromorphic functions; this is short for $\mathcal{K}^0(M)$, as one can think of such functions as meromorphic 0-forms.

Explicitly, (13.1.3) says that

$$\begin{aligned} f_\beta(z_\beta)dz_\beta &= f_\alpha(\Phi_{\alpha\beta}(z_\beta))d(\Phi_{\alpha\beta}(z_\beta)) \\ &= f_\alpha(\Phi_{\alpha\beta}(z_\beta))\Phi'_{\alpha\beta}(z_\beta)dz_\beta, \end{aligned}$$

and is thus equivalent to

$$f_\beta(z_\beta) = f_\alpha(\Phi_{\alpha\beta}(z_\beta))\Phi'_{\alpha\beta}(z_\beta). \tag{13.1.4}$$

Given $\omega_1, \omega_2 \in \mathcal{K}^1(M)$, we can consider their quotient as a meromorphic function $\frac{\omega_1}{\omega_2} \in \mathcal{K}(M)$. This is because in local coordinates, one can "cancel the dz's" – viz., $\frac{f_\alpha(z_\alpha)dz_\alpha}{g_\alpha(z_\alpha)dz_\alpha} = \frac{f_\alpha(z_\alpha)}{g_\alpha(z_\alpha)}$ – and the compatibility condition (13.1.4) implies that such quotients do patch together (the $\Phi'_{\alpha\beta}(z_\beta)$ factors cancel). Conversely, a meromorphic function times a meromorphic 1-form gives a new meromorphic 1-form.

Example 13.1.2. On $M = \mathbb{P}^1$, let $\omega_1 = \omega$ be arbitrary and $\omega_2 = dz$. Here $z = \frac{Z_1}{Z_0}$ on $\mathbb{P}^1$ as usual, and dz looks as if it should be not just meromorphic but holomorphic. But in the "coordinate at ∞" $w = \frac{Z_0}{Z_1}$, dz becomes $d(\frac{1}{w}) = -\frac{dw}{w^2}$. So dz in fact has a pole of order 2 at $[0:1]$.

Now consider $F(z) := \frac{\omega_1}{\omega_2} = \frac{\omega}{dz} \in \mathcal{K}(\mathbb{P}^1)(\cong \mathbb{C}(z)$ by Thm. 3.1.5(a)); we have then $\omega = F(z)dz$. Therefore

$$\mathcal{K}^1(\mathbb{P}^1) = \left\{ \frac{P(z)}{Q(z)}dz \,\middle|\, P \in \mathbb{C}[z],\ Q \in \mathbb{C}[z] \setminus \{0\} \right\}.$$

Example 13.1.3. For $M = \mathbb{C}/\Lambda$ a complex 1-torus, write u for the coordinate on $\mathbb{C}$. Since each transition function $\Phi_{\alpha\beta}$ sends $u \mapsto u + \lambda$ (for some $\lambda \in \Lambda$), their derivatives $\Phi'_{\alpha\beta}$ are all identically 1. Hence, du gives a well-defined global holomorphic 1-form on M (i.e. belongs to $\Omega^1(\mathbb{C}/\Lambda)$).

So take $\omega_1 = \omega$ arbitrary, $\omega_2 = du$. The same argument as above, using Thm. 3.1.5(b), gives

$$\mathcal{K}^1(\mathbb{C}/\Lambda) \cong \{ f(u)du \mid f \ = \Lambda\text{-periodic meromorphic function on } \mathbb{C} \}.$$

Example 13.1.4. Let $f \in \mathcal{K}(M)$ be a meromorphic function. We can represent f as a collection of maps $f_\alpha : V_\alpha \to \mathbb{P}^1$. The 1-forms $df_\alpha := \frac{df_\alpha}{dz_\alpha}dz_\alpha$ are then compatible (via pullback) with the transition functions, as in (13.1.3); hence, they patch together to give a global meromorphic 1-form $df \in \mathcal{K}^1(M)$. We will refer to this as the *differential of f*.

Let $\omega \in \mathcal{K}^1(M)$ be given by a collection $\{f_\alpha(z_\alpha)dz_\alpha\}$; we would like to define its *order* $\nu_p(\omega)$ at a point $p \in U_\alpha \subset M$. We simply set

$$\nu_p(\omega) := \nu_{z_\alpha(p)}(f_\alpha);$$

if this is negative ω has a *pole* at p. As a well-definedness check, suppose $p \in U_\beta$ also. Then (using (13.1.4))

$$\nu_p(f_\beta) = \nu_p\left(f_\alpha \cdot \Phi'_{\alpha\beta}(z_\beta)\right) = \nu_p(f_\alpha)$$

since, as a biholomorphism, $\Phi_{\alpha\beta}$ must have nonvanishing derivative at every point. If ω has a pole at $p \in U_\alpha$, then its *residue* is

$$Res_p(\omega) := Res_{z_\alpha(p)}(f_\alpha) = \frac{1}{2\pi\sqrt{-1}}\oint_{C_\epsilon(p)} f_\alpha(z_\alpha)dz_\alpha$$

where $C_\epsilon(p)$ is a small circle (in V_α) about $z_\alpha(p)$. The well-definedness check boils down to change of variable in the integral.

Let $\omega = \{f_\alpha(z_\alpha)dz_\alpha\} \in \mathcal{K}^1(M)$ be a form, and $\gamma = \cup\gamma_\alpha \subset M$ be a smooth real closed curve.[3] Then we define

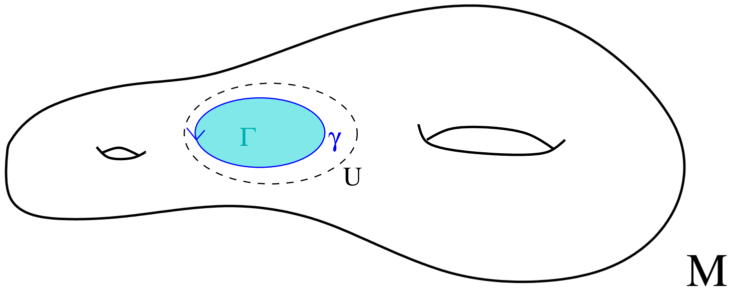
$$\int_\gamma \omega \;:=\; \sum_\alpha \int_{\gamma_\alpha} f_\alpha(z_\alpha)dz_\alpha,$$

where we observe that 1-forms have been set up so that the right-hand side is independent of choices of local coordinates and the partition of γ into local pieces. The following can be viewed as a version of either Stokes's theorem or Cauchy's theorem.

PROPOSITION 13.1.5. *Let $\Gamma \subset M$ be a closed region[4] with piecewise smooth boundary $\partial\Gamma = \gamma$.*

[3]A "real curve" means something 1-dimensional over $\mathbb{R}$ (not $\mathbb{C}$), so you should think of a closed path on the Riemann surface; and $\gamma_\alpha \subset U_\alpha$ are the segments from which the path is pieced together.

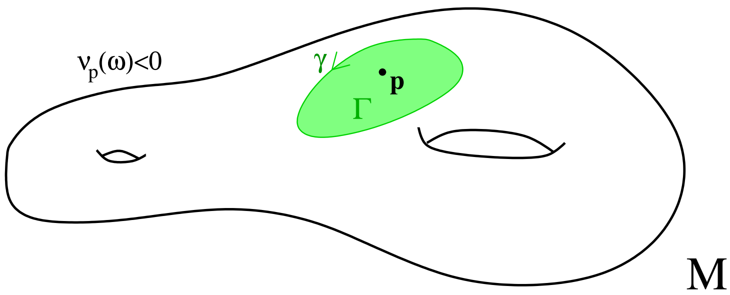
[4]the technical term here is 2-chain, though we won't get into this here



Assume that the meromorphic form ω is holomorphic on some open set U containing Γ. Then

$$\int_\gamma \omega = 0.$$

PROPOSITION 13.1.6. *Again let $\partial\Gamma = \gamma$, but assume that ω is only holomorphic on an open set containing γ (so that Γ may contain poles of ω).*



(a) Then we have the residue formula

$$\frac{1}{2\pi\sqrt{-1}}\int_\gamma \omega \quad = \sum_{\substack{p\,\in\,\Gamma \\ \nu_p(\omega)\,<\,0}} Res_p(\omega).$$

(b) In general for $\omega \in \mathcal{K}^1(M)$, $\sum_{p\in M} Res_p(\omega) = 0$.

PROOF. For the residue formula (a), take γ_0 be a sum of circular paths about those $p \in \Gamma$ where ω has poles. Let Γ_0 be the complement in M of the union of disks containing these $\{p\}$, with $\partial\Gamma_0 = \gamma_0$. Apply Prop. 13.1.5 to the pair $\Gamma - \Gamma_0$, $\gamma - \gamma_0$.

Applying the residue formula to the case $\Gamma = M$, $\gamma = \emptyset$ gives (b). $\square$

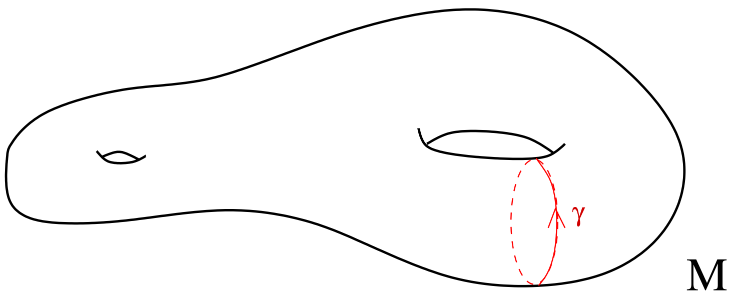
COROLLARY 13.1.7. *Consider a nonconstant meromorphic function $f \in \mathcal{K}(M)$. Then* *(a)* $\sum_{p \in M} \nu_p(f) = 0$*, i.e. the number of zeroes (counted with multplicity) equals to number of poles (counted with multplicity); and*

(b) $\#\{f^{-1}(\alpha)\}$ *(counted with multplicity) is independent of* $\alpha \in \mathbb{P}^1$*.*

Proof. (a) is Prop. 13.1.6(b) applied to $\omega = \frac{df}{f}$. Replacing f by $f - \alpha$, and noting that the number of poles doesn't change, by (a) the number of zeroes can't change either, giving (b). $\square$

Definition 13.1.8. The *degree of* f, $\deg(f)$, is defined to be the number in Cor. 13.1.7(b). Thinking of f as a covering map from $M \to \mathbb{P}^1$, $\deg(f)$ can be visualized as the number of branches (or "sheets").[5]

Remark 13.1.9. We have said nothing about $\int_\gamma \omega$ when γ is not a boundary:



Indeed, there is nothing we can say yet – this is the study of *periods*, which depend on the complex analytic structure of M. We will be able to compute some periods of holomorphic forms on algebraic curves later in the course.

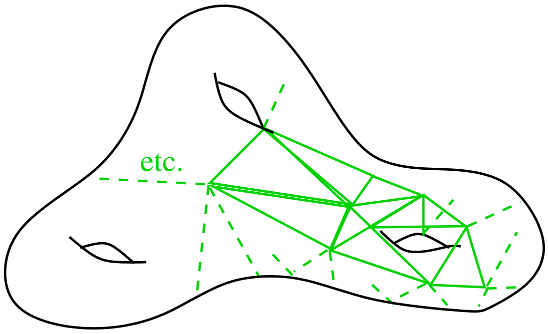
13.2. Poincaré-Hopf theorem

The usual statement of this theorem is that *the sum of indices of any*[6] *vector field* $\vec{v}$ *on a compact oriented smooth manifold* M *is equal to the Euler characteristic* χ_M; we'll only worry about the case where the real dimension of M is 2. In that case, the *index* $\mathrm{Ind}_p(\vec{v})$ of $\vec{v}$ at $p \in M$ is the number of counterclockwise rotations done by (the head of) $\vec{v}$ as one goes once counterclockwise on a small circle about p. It can only be nonzero if $\vec{v}(p) = 0$.

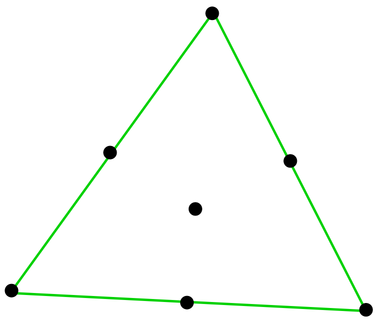
[5] you may wish to refer back to Remark 3.1.7

[6] technical point: $\vec{v}$ should have only finitely many zeroes

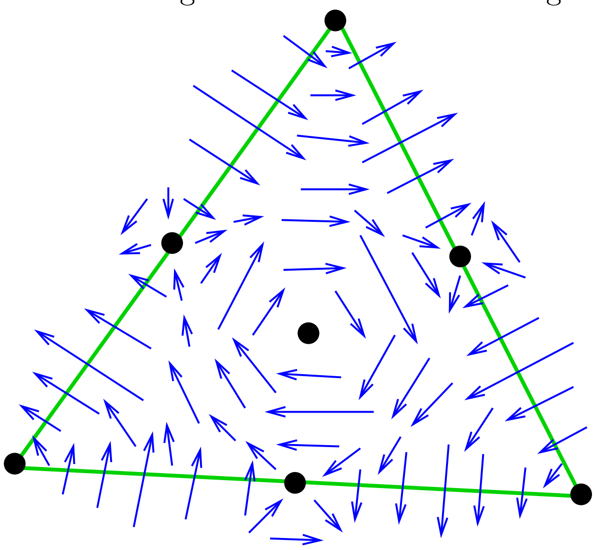
I'll give a heuristic proof of the italicized statement, which is probably more illuminating than a formal one. Subdivide a given compact smooth oriented real 2-manifold M into triangles:



Then put one marked point on each edge, vertex, and face of the triangulation:



Next draw the following vector field on each triangle:



These match up to give a global vector field on M. Evidently the index of this $\vec{v}$ is -1 at the marked points on the edges, and $+1$ at the marked points on the faces and vertices. Hence,

$$\sum_{p\in M} \mathrm{Ind}_p(\vec{v}) = \#F - \#E + \#V = \chi_M = 2-2g \tag{13.2.1}$$

where g is the genus of M. That (13.2.1) holds for *any* vector field $\vec{v}$ on M is the version of the theorem proved by Poincaré. It still holds if we allow $\vec{v}$ to have singularities at a finite set of points $\{p_1,\ldots,p_n\}$ (i.e. it is just a section over $M\backslash\{p_1,\ldots,p_n\}$), provided one adds the indices of $\vec{v}$ at the p_i to the sum.

In fact, (13.2.1) even holds if $\vec{v}$ is replaced by a smooth 1-form $\eta \in A^1_{\mathbb{R}}(M\backslash\{p_1,\ldots,p_n\})$. The idea is to use a metric on M, i.e. a nonvanishing section of $\mathrm{Sym}^2(T^*M)$, to smoothly identify TM with T^*M. The corresponding notion of index, if (in local coordinates at p) η takes the form $Fdx + Gdy$, is

$$\mathrm{Ind}_p\eta := \frac{1}{2\pi}\oint d\arctan\left(\frac{G}{F}\right), \tag{13.2.2}$$

and once again the sum in (13.2.1) must be over all zeroes of η and the $\{p_i\}$.

Now let M be a compact complex 1-manifold, and write $\omega \in \mathcal{K}^1(M)$ be locally in the form $f.dx + g.dy$ where f, g are complex-valued. To get in the above setting, we may of course view M as a smooth real 2-manifold, and take the real part of ω:

$$\eta := \Re(\omega) \overset{\text{loc}}{=} \Re(f)dx + \Re(g)dy.$$

Let p be a zero or pole of ω, and put $\nu = \nu_p(\omega)$. Of course, in a local holomorphic coordinate z about p with $z(p) = 0$, we have[7]

$$\omega \overset{\text{loc}}{\approx} z^\nu dz \;=\; r^\nu\left(\cos(\nu\theta) + \sqrt{-1}\sin(\nu\theta)\right)(dx + \sqrt{-1}dy)$$

$$=\; r^\nu\left(\cos(-\nu\theta) - \sqrt{-1}\sin(-\nu\theta)\right)dx + r^\nu\left(\sin(-\nu\theta) + \sqrt{-1}\cos(-\nu\theta)\right)dy.$$

So locally we have for the real part

$$\frac{\eta}{r^\nu} \;\approx\; \cos(-\nu\theta)dx + \sin(-\nu\theta)dy,$$

and thus by (13.2.2)

$$\mathrm{Ind}_p(\eta) = \frac{1}{2\pi}\oint d[-\nu\theta] = -\nu = -\nu_p(\omega)$$

[7]up to multiplication by a locally nonvanishing holomorphic function (which will not affect index)

$$\implies \sum_p \nu_p(\omega) = 2g - 2.$$

We have arrived at the following corollary of (13.2.1), which will henceforth be the meaning of "Poincaré-Hopf" for us:

THEOREM 13.2.1. *Let $\omega \in \mathcal{K}^1(M)^*$ be a nonvanishing meromorphic 1-form on a Riemann surface of genus g. Then*

$$\underbrace{(\#\text{ of zeroes } - \#\text{ of poles})}_{\text{counted with multiplicity}} \text{ of } \omega \ = \ 2g - 2.$$

REMARK 13.2.2. Just as for meromorphic functions we can consider the divisor

$$(\omega) := \sum_{p \in M} \nu_p(\omega)[p]$$

of a meromorphic 1-form. In this context, the Theorem says that

$$\deg((\omega)) = 2g - 2.$$

Exercises

(1) Let $E = \{y^2 - 4x^3 - 4x = 0\}$, $\omega = \frac{dx}{y}\Big|_E \in \Omega^1(E)$. (We can talk about holomorphic 1-forms on a smooth algebraic curve now, because they are Riemann surfaces by the "smooth normalization" Prop. 7.1.) Consider the complex analytic automorphism $A : E \to E$ sending $(x, y) \mapsto (-x, iy)$, and "apply" this to the 1-form: compute the pullback $A^*(\omega)$.

(2) (a) In Example 13.1.2, dz defines a meromorphic differential 1-form on $\mathbb{P}^1$. Compute its divisor (dz). Explain why $\Omega^1(\mathbb{P}^1) = \{0\}$. (b) What is the divisor of du on $\mathbb{C}/\Lambda$, from Example 13.1.3? Explain why it is the *unique* holomorphic 1-form on $\mathbb{C}/\Lambda$ up to scale.

(3) Practice with pullbacks: for the map $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ that sends $(x, y) \mapsto (u(x, y), v(x, y)) := (x^2 - 3xy,\ y^3 + x)$, compute $\Phi^*\omega$ where $\omega = udv + vdu$. Write it in the form $f(x, y)dx + g(x, y)dy$.

(4) Continuing from Exercise 5 of Chapter 3, compute the pullback of $\frac{dx}{y}$ under $\varphi : \mathbb{P}^1 \to C$. [Hint: simply plug in your final $x(z)$ and $y(z)$ from that exercise. After simplification, your answer should be very simple indeed.]

CHAPTER 14

The genus formula

We are ready to prove two formulas for the genus of a Riemann surface (RS) which are especially useful in algebraic geometry. For the first result (the Riemann-Hurwitz formula), the RS will arise as a finite branched cover of another RS whose genus is known. The proof makes essential use of Poincaré-Hopf and a *ramification divisor* which we introduce below. For the second result, which is an application of the first (and of the intersection theory from Chap. 12), the RS will arise as the normalization of an irreducible algebraic curve in $\mathbb{P}^2$ with only ordinary double point (ODP) singularities. (In fact, the 4H people will, in their reading material, learn how to deal with worse singularities later.) This is a very concrete payoff for the preceding hard work: now we can compute the genus of [the desingularization of] a projective algebraic curve!

14.1. Order and multiplicity for maps of Riemann surfaces

Consider a nonconstant morphism $f : M \to M'$ of Riemann surfaces with $f(p) = q$. In Exercise 4 of Chapter 3, the following was established: there exist

- neighborhoods $U \ni p$, $V \ni q$ with $f(U) \subset V$, and
- local holomorphic coordinates $z : U \to \mathbb{C}$ and $w : V \to \mathbb{C}$ with $z(p) = 0 = w(q)$,

such that $w \circ f = z^\nu$ for some unique $\nu \in \mathbb{N}$. More informally, in these local coordinates f "takes the form" $(w =)f(z) = z^\nu$. We write $\nu_p(f) := \nu$. This is the *ramification index*, and f *ramifies* at p precisely when it exceeds 1.

For any $q \in M'$, consider the sum

$$d(q) := \sum_{p \in M \text{ with } f(p)=q} \nu_p(f).$$

If a ramification point p of index d lies over q, then over a nearby point q_0, p is replaced by d points with ramification index 1. This is by virtue of the local form $w = z^\nu$, as is the fact that the ramification points are isolated hence finite in number (M is compact!). Evidently then, $d(q)$ is constant in q; we will call this constant $d \in \mathbb{N}$ the *degree* $\deg(f)$ of the morphism f.[1]

Here is a more gentrified way to define this. We can think of a point $q \in M'$ as a divisor $[q] \in \mathrm{Div}(M)$, and "pull it back" to a divisor on M by the formula[2]

$$f^{-1}([q]) := \sum_{f(p)=q} \nu_p(f)[p] \ \in \ \mathrm{Div}(M).$$

We then put (for any $q \in M'$, it doesn't matter)

$$\deg(f) := \deg\left(f^{-1}([q])\right) \ \left(= \sum_{f(p)=q} \nu_p(f) \right).$$

Associated to $f : M \to M'$, finally, is the *ramification divisor*

$$R_f := \sum_{p \in M} \left(\nu_p(f) - 1\right)[p] \ \in \ \mathrm{Div}(M).$$

By the above remarks, the sum is clearly finite.

14.2. Riemann-Hurwitz formula

Again take $f : M \to M'$ to be a nonconstant morphism, write $d := \deg(f)$, and put

$$r := \deg(R_f).$$

In the following g resp. g' will refer to the genus of M resp. M'.

THEOREM 14.2.1. $r = 2\{g + d - dg' - 1\}$.

REMARK 14.2.2. Some alternative ways to write this result are:
(i) $g = (g' - 1)d + \frac{r}{2} + 1$
(ii) $\chi_M = \deg(f)\chi_{M'} - \deg(R_f)$
These better represent the way you want to think of it: as a formula

---

[1] when f is a nonconstant map from M to $\mathbb{P}^1$, you can think of it as a meromorphic function and take the degree of its divisor, $\deg((f))$, which is always 0. Or, you can think of it as a morphism of Riemann surfaces and take $\deg(f)$, which is never 0. So that extra parenthesis matters!

[2] This extends linearly to define $f^{-1}(D)$ for any $D \in \mathrm{Div}(M)$.

for the genus (or Euler characteristic) of M, if you know that of M' and data about how M "sits over" M'.

PROOF. For $p \in M$ with $a = \nu_p(f)$, we choose local coordinates z, w as in §14.1 so that $z \overset{f}{\mapsto} z^a (= w)$.

We shall need to assume the existence of a nonzero meromorphic 1-form $\omega \in \mathcal{K}^1(M')$. This is obvious if M' arises as the normalization of an algebraic curve in $\mathbb{P}^2$, as you can just pull back any nonconstant meromorphic function (say, Z_1/Z_0) and take its differential. Every Riemann surface arises in this way, but to see that you need the Riemann-Roch theorem. We proceed with the proof modulo this detail.

Locally writing $\omega = g(w)dw$, we have

$$f^*\omega \overset{\text{loc}}{=} g(z^a)d(z^a) = a.g(z^a)z^{a-1}dz\,,$$

hence

$$\nu_p(f^*\omega) = a.\nu_0(g) + (a-1) = \nu_p(f).\nu_{f(p)}(\omega) + (\nu_p(f) - 1).$$

In $\text{Div}(M)$ we have therefore

$$\begin{aligned}
(f^*\omega) := \sum_{p \in M} \nu_p(f^*\omega)[p] &= \sum_p \nu_p(f).\nu_{f(p)}(\omega)[p] + \underbrace{\sum_p (\nu_p(f) - 1)\,[p]}_{R_f} \\
&= \sum_{q \in M'} \nu_q(\omega) \sum_{f(p)=q} \nu_p(f)[p] \;\;+\;\; R_f \\
&= \sum_{q \in M'} \nu_q(\omega) f^{-1}([q]) \;\;+\;\; R_f \\
&= f^{-1}\left(\sum_q \nu_q(\omega)[q]\right) \;\;+\;\; R_f \\
&= f^{-1}((\omega)) + R_f,
\end{aligned}$$

where $(\omega) \in \text{Div}(M')$.

Now $f^*\omega \in \mathcal{K}^1(M)$, so Poincaré-Hopf on M tells us that

$$2g - 2 = \deg\left((f^*\omega)\right)$$

which by the computation just done

$$= \deg\left(f^{-1}((\omega))\right) + \deg R_f$$

$$= \sum_q \nu_q(\omega) \underbrace{\sum_{f(p)=q} \nu_p(f)}_{\deg(f)} \ + \ r$$

$$= \deg(f) \underbrace{\sum \nu_q(\omega)}_{\deg((\omega))} \ + \ r.$$

Applying Poincaré-Hopf once more (but on M'), we get that this

$$= d(2g' - 2) + r.$$

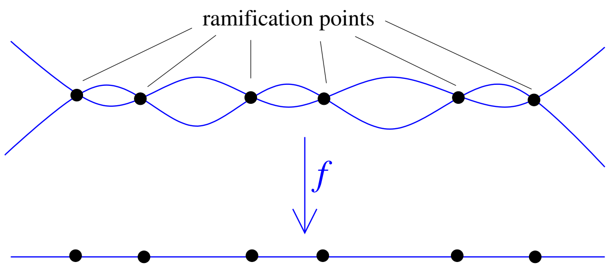
So we have shown $2 - 2g = d(2 - 2g') - r$, which is the version of R-H stated in Remark 14.2.2(ii). $\square$

We turn to some examples.

Example 14.2.3. Let $C = \{y^2 = \prod_{i=1}^{2m}(x - \alpha_i)\} \subset \mathbb{C}^2$, and let M be the normalization of its projective closure $\bar{C} \subset \mathbb{P}^2$. The original curve had a projection map to the x-axis ($(x, y) \mapsto x$), and this extends to

$$f : M \to \mathbb{P}^1 =: M',$$

as depicted below:



Clearly $g' = 0$, $d = 2$, and

$$r = \sum (\nu_p(f) - 1) = 2m$$

since $\nu_p(f) - 1 = 1$ at each of the ramification points. So by Remark 14.2.2(i)

$$g = (0 - 1).2 + \frac{2m}{2} + 1 = m - 1.$$

Example 14.2.4. Let $M = M' = \mathbb{C}/\Lambda$ be a complex 1-torus; as usual $\Lambda = \{m_1\lambda_1 + m_2\lambda_2 \mid m_1, m_2 \in \mathbb{Z}\}$, where $\lambda_1, \lambda_2 \in \mathbb{C}$ are independent over $\mathbb{R}$. Now assume $\alpha\Lambda \subseteq \Lambda$ for some $\alpha \in \mathbb{C}^*$. Then we have a "complex multiplication" map

$$M \xrightarrow{f} M'$$

$$z \longmapsto \alpha z$$

which has $R_f = 0$. You will treat this setting in an exercise below.

14.3. The genus of a projective algebraic curve

Let $C = \{F(Z, X, Y) = 0\} \subset \mathbb{P}^2$ be an irreducible algebraic curve of degree d with $\mathcal{S} = \mathrm{sing}(C)$ its set of singular points. We assume that these are all ordinary double points, and that there are exactly $|\mathcal{S}| = \delta$ of these; write $\mathcal{S} = \{p_1, \ldots, p_\delta\}$. Of course, $\delta = 0 \iff \mathcal{S} = \emptyset \iff C$ is smooth.

Denoting by $\sigma : \tilde{C} \twoheadrightarrow C$ its normalization, we shall deduce from Theorem 14.2.1 the formula:

THEOREM 14.3.1. *$\tilde{C}$ has genus*

$$g = \frac{(d-1)(d-2)}{2} - \delta.$$

To get a feel for this before launching into the proof, for C smooth we have

$$\begin{aligned} d = 1 &\implies g = 0, \\ d = 2 &\implies g = 0, \\ d = 3 &\implies g = 1, \\ d = 4 &\implies g = 3, \end{aligned}$$

and so on. For degree 3 with one ODP, we get

$$g = \frac{(3-1)(3-2)}{2} - 1 = 0,$$

as we found using stereographic projection. Indeed, we know how to parametrize all three genus 0 cases (smooth $d = 1, 2$; singular $d = 3$) by a Riemann sphere.

The rest of this section is devoted to the proof. Begin by choosing coordinates on $\mathbb{P}^2$ so that

- $L_\infty \cap C$ consists of d distinct points,
- none of the tangents to C at its ODP's are vertical (i.e. of the form $X = aZ$), and
- C does not contain $[0 : 0 : 1]$.

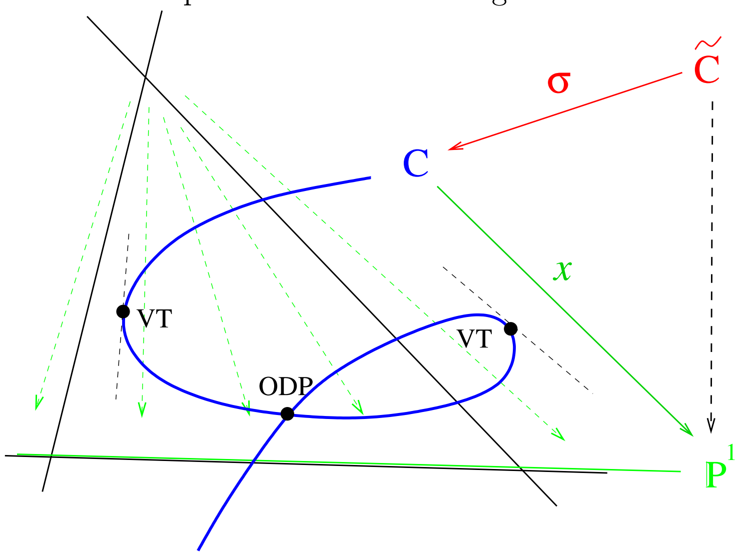
The latter requirement allows us to project from $[0 : 0 : 1]$: that is, the map

$$C \xrightarrow{\mathbf{x}} \mathbb{P}^1 =: M'$$

given by

$$[Z : X : Y] \mapsto [Z : X],$$

roughly speaking the "projection of C to the x-axis", is well-defined. Writing $M := \tilde{C}$, the main idea of the proof of to apply Riemann-Hurwitz to the composition $f = \mathbf{x} \circ \sigma : M \to M'$. In a picture, where "VT" refers to a point with vertical tangent:



Now for $M' = \mathbb{P}^1$, $g' = 0$ so that Thm. 14.2.1 gives

$$r_f = 2(\text{genus}(M) + \deg(\mathbf{x}) - 1) = 2(g + d - 1). \tag{14.3.1}$$

In particular, the degree of the map $\mathbf{x}$ is d because the projection is done along vertical lines, all but finitely many such lines meet C in d points by Bezout, and σ is 1-to-1 off finitely many such points. So we see that if we can compute the degree of the ramification divisor R_f then we are done.

To do this, let

$$E := \{F_Y = 0\}$$

where F_Y is the partial derivative. Obviously $\deg(E) = d - 1$, and so by Bézout,

$$(E \cdot C) = (d - 1)d. \tag{14.3.2}$$

Denoting by $\sum_p'$ the sum over points where C has a vertical tangent, and by $\sum_{j=1}^{\delta}$ the sum over ODP's, we have

$$(E \cdot C) = \sum_p{}'(E \cdot C)_p + \sum_{j=1}^{\delta}(E \cdot C)_{p_j}.$$

We will show

$$R_f = \sum_p{}'(E \cdot C)_p[\tilde{p}] \tag{14.3.3}$$

where $\tilde{p} = \sigma^{-1}(p) \in \tilde{C}$. (Recall that by our choice of coordinates, a point with vertical tangent cannot be a singular point, and so has a *unique* preimage point under normalization.) Taking degrees of both sides of (14.3.3) gives

$$r_f = \sum_p{}'(E \cdot C)_p = (E \cdot C) - \sum_{j=1}^{\delta}(E \cdot C)_{p_j}. \tag{14.3.4}$$

Further, we will deduce that

$$(E \cdot C)_{p_j} = 2 \quad (\forall j); \tag{14.3.5}$$

together with (14.3.2) and (14.3.3), this yields

$$r_f = d(d-1) - 2\delta.$$

Now put this together with (14.3.1) to get

$$2g + 2(d-1) = d(d-1) - 2\delta,$$

$$2g = (d-2)(d-1) - 2\delta,$$

and divide the last line by 2 to get Theorem 14.3.1. It remains only to check (14.3.3) and (14.3.5).

If C has a VT at p, then $F(p) = F_Y(p) = 0$; this implies $p \in C \cap E$. By assumption, p is a smooth point, so that[3] $F_X(p) \neq 0$. By the holomorphic implicit function theorem, we can parametrize C locally by writing $x = X/Z$ as an implicit function of $y = Y/Z$, viz.

$$0 = F(1, x(y), y).$$

[3] if $F_X(p) = 0$ then $F_Z(p) = 0$ too by the Euler formula, contradicting smoothness at p

Now, differentiating gives

$$0 = \frac{d}{dy} F(1, x(y), y) = F_X(1, x(y), y) \cdot x'(y) + F_Y(1, x(y), y).$$

For the two functions on the right-hand side to sum to zero, they must have the same order to $y(p)$:

$$\mathrm{ord}_{y(p)} F_Y(1, x(y), y) = \mathrm{ord}_{y(p)} x'(y),$$

in other words

$$\begin{aligned} (E \cdot C)_p &= \{\mathrm{ord}_{y(p)} x(y) \ - \ 1\} \\ &= \{\nu_p(\mathbf{x}) - 1\} \\ &= \{\nu_{\tilde{p}}(f) - 1\}. \end{aligned}$$

As the only ramification points of f are (σ^{-1} of) vertical tangent points,

$$R_f := \sum_{q \in \tilde{C}} (\nu_q(f) - 1)[q] = {\sum_p}' (E \cdot C)_p [\tilde{p}]$$

as claimed.

Finally, to see (14.3.5), assume for simplicity (for some j) $p_j = (0, 0)$. The local affine equation about an ODP is of the form

$$F(1, x, y) = ax^2 + 2bxy + cy^2 + \{\text{higher-order terms}\}.$$

To find the tangent lines, recall that one solves

$$0 = ax^2 + 2bxy + cy^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{B} \begin{pmatrix} x \\ y \end{pmatrix}$$

in $\mathbb{P}^1$ (for their "slopes"). That the solution Q consists of two distinct points (as p_j is an ODP) $\implies$ Q is "smooth" $\implies$ $\det B \neq 0$ $\implies$ $ac - b^2 \neq 0$. That there is no vertical tangent $\implies$ $[x : y] = [0 : 1]$ is not a solution $\implies$ $c \neq 0$. Consider the partial

$$F_Y(1, x, y) = 2bx + 2cy + \{\text{higher-order terms}\}$$

whose vanishing defines E; evidently E can be locally parametrized about p_j by

$$y = y(x) = -\frac{b}{c} x + \{\text{higher-order terms}\}.$$

To compute its intersection number against C, we pull the defining equation of C back along this parametrization and take the order at 0:

$$\begin{aligned}(E\cdot C)_{(0,0)} &= \mathrm{ord}_0(F(1,x,y(x)))\\ &= \mathrm{ord}_0\left(ax^2+2bx\cdot y(x)+c(y(x))^2+\{\text{higher-order terms}\}\right)\\ &= \mathrm{ord}_0\left(\frac{ac-b^2}{c}x^2+\{\text{higher-order terms}\}\right)\\ &= 2\,,\end{aligned}$$

Q.E.D.

14.4. Beyond stereographic projection

The genus formula is very nice, but needs to pass a smell test: if it says that a curve $C \in \mathbb{P}^2$ has genus zero normalization, then we should be able to parametrize C by the unique genus zero Riemann surface $\mathbb{P}^1$. We know that this can be done for a smooth conic and a nodal cubic (i.e. a cubic with one ODP); the first new case predicted by the formula is that of an irreducible[4] quartic curve ($d = 4$) with 3 ODP's ($\delta = 3$):

$$g = \frac{(4-1)(4-2)}{2} - 3 = 0.$$

Let's give this a try. Write $\{p_i\}_{i=0,1,2}$ for the ODP's, and suppose another curve D passes through one of these: then by 12.2.4, $(C\cdot D)_{p_i} \geq 2$. If D is a line, then it cannot pass through all 3 p_i, as then we would have

$$4 = \deg C\cdot \deg L = (C\cdot L) \geq \sum_{i=0}^{2}(C\cdot L)_{p_i} \geq 6,$$

a contradiction. So the ODP's are not collinear, and by a similar argument[5] if p_3 is any fixed smooth point of C, then no three of p_0, p_1, p_2, p_3 are collinear. We may therefore move C (and the p_i) by a projectivity of $\mathbb{P}^2$, to have $p_0 = [1:0:0]$, $p_1 = [0:1:0]$, $p_2 = [0:0:1]$, $p_3 = [1:1:1]$. (We'll do so for this abstract analysis but not for the concrete example that follows.)

The general conic in $\mathbb{P}^2$ is of the form

$$aXY + bYZ + cXZ + dX^2 + eY^2 + fZ^2 = 0.$$

[4]We have to say C is irreducible explicitly, because the union of a smooth cubic and a general line is a quartic with 3 ODP's.

[5]the 6 gets replaced by a 5 in the inequality above

By substitution, we find that the general conic through the above four points is of the form

$$Q_{[a:b]} = \{aXY + bYZ - (a+b)XZ = 0\}.$$

This is a 1-parameter family parametrized by $[a : b] \in \mathbb{P}^1$.

The zero-cycle (cf. Remark 12.2.2(a)) $Q_{[a:b]} \cdot C$ has degree 8 by Bézout, and is of the form $2[p_0] + 2[p_1] + 2[p_2] + [p_3] + \mathit{more}$. This "more" can only be one more point $q_{[a:b]}$ with multiplicity one, since what is already written has degree 7 (and by construction, one doesn't have negative intersection numbers). Naturally, q could be one of the p_i: if it is p_3, then this would say that Q is tangent to C there. Define a map

$$\sigma : \mathbb{P}^1 \to C$$

by

$$[a : b] \mapsto q_{[a:b]} \; (:= Q_{[a:b]} \cdot C - \{2[p_0] + 2[p_1] + 2[p_2] + [p_3]\}).$$

In fact, this is a morphism of complex manifolds from $\mathbb{P}^1 \to \mathbb{P}^2$ (I won't prove this carefully). Also, since C is irreducible, that σ is onto essentially follows from the open mapping theorem and compactness of $\mathbb{P}^1$.

We claim that σ is 1-to-1 off the singular points of C. Take $q \in C$ distinct from the p_i; since no three of the p_i are collinear, no four of q, p_0, p_1, p_2, p_3 are collinear, so there exists a *unique* conic Q through all five. (The uniqueness when $q = p_3$ then essentially follows from continuity of σ.)

Example 14.4.1. So what does such a normalization look like? Take the very concrete quartic curve

$$C = \{X^2Z^2 + Y^2Z^2 + 2X^2Y^2 = 0\}.$$

Irreducibility can be checked by putting the polynomial in affine form $y^2(1+2x^2)+x^2$ and showing it doesn't factor into terms of lower degree in y. I will let you check that the only singularities are $p_0 = [1 : 0 : 0]$, $p_1 = [0 : 1 : 0]$, $p_2 = [0 : 0 : 1]$; pick $p_3 := [i : 1 : 1]$ ($i = \sqrt{-1}$). The general conic through these 4 points is

$$Q_{[\alpha:\beta]} := \{\alpha XZ + \beta YZ = i(\alpha + \beta)XY\}.$$

Substituting this into α^2 times the equation of C gives

$$\begin{aligned} 0 &= (i(\alpha+\beta)XY - \beta YZ)^2 + \alpha^2 Y^2 Z^2 + 2\alpha^2 X^2 Y^2 = \cdots \\ &= (\beta^2+\alpha^2)Y^2(Z-iX)\left(Z - i\frac{2\alpha\beta+\beta^2-\alpha^2}{\beta^2+\alpha^2}X\right), \end{aligned}$$

in which the last factor gives us the $x\left(=\frac{X}{Z}\right)$-coordinate of the point $q_{[\alpha:\beta]}$. The y-coordinate is obtained by substituting into the equation of Q, and we find $\sigma([\alpha:\beta]) =$

$$\big[i(2\alpha\beta+\alpha^2-\beta^2)(2\alpha\beta+\beta^2-\alpha^2) : (\alpha^2+\beta^2)(2\alpha\beta+\alpha^2-\beta^2) \\ : (\alpha^2+\beta^2)(2\alpha\beta+\beta^2-\alpha^2)\big].$$

Or, in affine coordinates ($t = \frac{\beta}{\alpha}$ in particular),

$$\sigma(t) = \left(-i\frac{1+t^2}{t^2+2t-1}, i\frac{1+t^2}{t^2-2t-1}\right).$$

Exercises

(1) Recall the setup of Riemann-Hurwitz: C, C' compact RS with $g = genus(C)$, $g' = genus(C')$, $f : C \to C'$ nonconstant holomorphic map of degree d. Show that for any $d \geq 1$, $g \geq g'$. (The covering surface "has at least as many handles".)

(2) Let $z = \frac{Z_1}{Z_0}$ (where $[Z_0 : Z_1]$ are the homogeneous coordinates) be the "canonical coordinate" on $\mathbb{P}^1$. If a holomorphic map $f : \mathbb{P}^1 \to \mathbb{P}^1$ takes the form $f(z) = z^n + a_1 z^{n-1} + \ldots + a_n$, then
(a) What is $\deg(f)$?
(b) What can you say about the ramification divisor R_f? (at least, what is its degree?)
(c) Use Riemann-Hurwitz to check your answers.

(3) Let $C = \mathbb{C}/\Lambda$ where $\Lambda = \{m\lambda_1 + n\lambda_2 \,|\, m, n \in \mathbb{Z}\}$ is a lattice in $\mathbb{C}$. (In particular, λ_1 and λ_2 are independent over $\mathbb{R}$.) Suppose that $\alpha \in \mathbb{C}^*$ satisfies $\alpha\Lambda \subseteq \Lambda$. [Remark: if $\alpha \notin \mathbb{Z}$ this places a strong condition on Λ; we say Λ, or C, "has *complex multiplication* (or *CM*)."] The multiplication by α induces a holomorphic map $\mu_\alpha : C \to C$, i.e. an automorphism of the RS C.
(a) Show that the ramification divisor $R \in Div(C)$ for this map is zero.
(b) Prove that the degree of μ_α equals the index $[\Lambda : \alpha\Lambda]$ of the image lattice $\alpha \cdot \Lambda \subseteq \Lambda$.

(4) Find the genus of the normalization $\widetilde{C}$ of the irreducible curve C given by taking the closure of $x^2 + x^2y^2 + y^2 = 0$ in $\mathbb{P}^2$. (First convert to homogeneous coordinates and check for singularities. Then apply the genus formula. This is very similar to something above...)

(5) This problem complements (3) above, but you won't use anything from this chapter in doing it. A (holomorphic) map $f : \frac{\mathbb{C}}{\Lambda} \to \frac{\mathbb{C}}{\Lambda}$ of Riemann surfaces is nothing but an analytic map $\tilde{f} : \mathbb{C} \to \mathbb{C}$ (i.e. an entire function) such that for all $\lambda \in \Lambda$, $z \in \mathbb{C}$,

$$(*) \quad \tilde{f}(z+\lambda) - \tilde{f}(z) \ \in \ \Lambda,$$

i.e. $z_1 \equiv z_2 \bmod \Lambda \implies \tilde{f}(z_1) \equiv \tilde{f}(z_2) \bmod \Lambda$ (this is just the well-definedness condition for f). Show that such a map is necessarily affine, i.e. of the form

$$\tilde{f}(z) = \alpha z + \beta.$$

[If α other than $\alpha \in \mathbb{Z}$ works, then we are of course in the CM case described above. So a non-CM complex 1-torus, which is the "generic" case, has endomorphisms of the form $z \mapsto nz + \beta$, $n \in \mathbb{Z}$, and that's all.] Hint: consider $\tilde{f}'(z)$, and use $(*)$.

CHAPTER 15

Some applications of Bézout

We have already put Bézout's theorem to use in proving the genus formula (and in several interesting exercises at the end of Chapter 12). Now we shall use it to prove a general result on curves through configurations of points, which in particular will yield a short (and rigorous) proof of Pascal's theorem from Chapter 1. We shall also deduce some results on cubics will will come in handy in studying the group law on elliptic curves.

Throughout this Chapter we shall use the following dictionary:

algebraic curve $\subset \mathbb{P}^2$	degree	defining equation (homogeneous polynomial)
C	d	$F \in S_3^d$
D	d	$G \in S_3^d$
E	e	$H \in S_3^e$

Recall the theorem we are wanting to apply:

Bézout. *$C \cap E$ is 0-dimensional (consists of points) $\implies (C \cdot E) = de$.*

Part of the content of the (equivalent) contrapositive statement is:

tuozéB. *The number of points $|C \cap E|$ exceeds $de \implies E$ and C have a common component.*

From Chapter 9, we have:

Study. E irreducible and $E \subset C \implies H$ divides F.

Putting tuozé**B** and **S**tudy together gives:

BS. *E irreducible and $|C \cap E| > de \implies H \,|\, F$.*

We'll make use of this statement below.

15.1. Cayley-Bacharach theorem

Let $p_1, \ldots, p_n \in \mathbb{P}^2$ be distinct points, and define

$$S^d(p_1, \ldots, p_n) := \begin{array}{c}\text{homogeneous polynomials (of degree } d\text{)} \\ \text{in } [Z : X : Y] \text{ vanishing at } p_1, \ldots, p_n.\end{array}$$

LEMMA 15.1.1. *Suppose E is irreducible and $p_1, \ldots, p_a \in E$ for some $a > ed$, while $p_{a+1}, \ldots, p_n \notin E$. Then*

$$S^d(p_1, \ldots, p_n) = H \cdot S^{d-e}(p_{a+1}, \ldots, p_n).$$

PROOF. The inclusion of the RHS into the LHS is easy, since it is just saying that the product of a polynomial vanishing at the last $n-a$ points by a polynomial vanishing at the first a points, vanishes at all of them. So we turn to the reverse inclusion.

Assuming $S^d(p_1, \ldots, p_n)$ is nonzero, take a nonzero element F; this defines a degree d curve C containing $p_1, \ldots, p_n$. Clearly $p_1, \ldots, p_a \in C \cap E$, so $|C \cap E| > ed$, and by "BS", $H \mid F$. We can therefore write $F = F_0 \cdot H$ with $F_0 \in S^{d-e}$. Since $F = 0$ but $H \neq 0$ at $p_{a+1}, \ldots, p_n$, F_0 must vanish at these points. It follows that $F_0 \in S^{d-e}(p_{a+1}, \ldots, p_n)$ as desired. $\square$

THEOREM 15.1.2. *Let E be irreducible, $|C \cap D| = d^2$ with $d > e$, and assume exactly*[1] *ed of the points of $C \cap D$ lie on E. Then the remaining $d(d-e)$ points lie on a (not necessarily irreducible!) curve of degree $\leq d-e$.*

PROOF. Let $[A : B : C] \in E \backslash \{(C \cap D) \cap E\}$, and set $\lambda = F(A, B, C)$, $\mu = -G(A, B, C)$. Define $P := \lambda G + \mu F \in S^d$; this vanishes on $C \cap D$ *and* at $[A : B : C]$. Label $(C \cap D) \cap E =: \{p_1, \ldots, p_{ed}\}$, $[A : B : C] =: p_{ed+1}$, and $(C \cap D) \backslash \{(C \cap D) \cap E\} = \{p_{ed+2}, \ldots, p_{d^2+1}\}$; set $a := ed + 1$ and $n = d^2 + 1$.

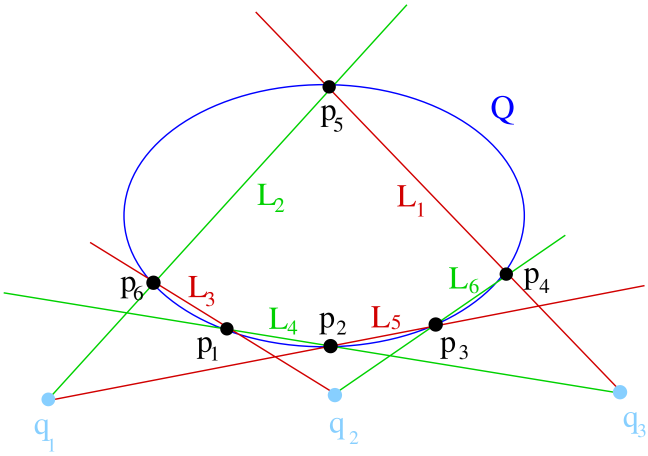
Since $a > ed$, Lemma 15.1.1 tells us that $S^d(p_1, \ldots, p_{d^2+1}) = H \cdot S^{d-e}(p_{ed+2}, \ldots, p_{d^2+1})$. But then, since $P \in S^d(p_1, \ldots, p_{d^2+1})$, we have $P = HP_0$ for some $P_0 \in S^{d-e}(p_{ed+2}, \ldots, p_{d^2+1})$. This P_0 defines the required curve. $\square$

Here is the nice application to Pascal:

[1] it is enough to check, in applying this, that ed of the points (not "exactly ed of the points") lie on E. This is because by Bézout, more than ed of these points simply *can't* lie on E.

COROLLARY 15.1.3. *The (three) intercepts of opposite sides of a hexagon inscribed in a conic are collinear.*

PROOF. Referring to the picture



we put $C := L_1 \cup L_3 \cup L_5$, $D = L_2 \cup L_4 \cup L_6$, and $E = Q$. Clearly this means $d = 3$ and $e = 2$, and we do indeed see that $de = 6$ points of $C \cap D = \{p_1, \ldots, p_6\} \cup \{q_1, q_2, q_3\}$ lie on E. So the last three points of $C \cap D$, which are the intercepts, lie on a curve of degree $d - e = 1$ by the Theorem. □

REMARK 15.1.4. If one wanted instead to plug the technical gap in the proof of Pascal suggested in Chapter 1, part of what one needs is the statement: if $p_1, \ldots, p_8 \in \mathbb{P}^2$ are distinct and in "general position" in the sense that no 4 are collinear and no 7 conconic (lying on an irreducible conic), then $\dim S^3(p_1, \ldots, p_8) = 2$. This is proved in Reid's book.

15.2. Intersections of cubics

The results of §15.1 dealt with the case where all intersections of curves have multplicity one (the "transversal" case), since we required $|C \cap D| = d^2 = (C \cdot E)$. To deal with the general case, at least assuming E is smooth and irreducible (so that we may view it as a Riemann surface), write

$$C \cdot E := \sum_{p \in E \cap C} (E \cdot C)_p [p] \ \in \ \mathrm{Div}(E).$$

If E is irreducible but singular, with an unique ODP $\hat{p}$, the same definition gives a divisor $C \cdot E \in \mathrm{Div}(\tilde{E})$ (on the normalization) provided $\hat{p} \notin E \cap C$.

THEOREM 15.2.1. *Let C, D, E be distinct cubics, with E irreducible. (If E is singular, assume moreover that $\hat{p} \notin E \cap C$, $E \cap D$.) Writing by Bézout*

$$D \cdot E \;=\; \sum_{i=1}^{9} [q_i] \;\in\; Div(E)$$

where the q_i need not be distinct, and assuming

$$C \cdot E \;=\; \sum_{i=1}^{8} [q_i] \;+\; [q] \;\in Div(E),$$

we have $q = q_9$.

In the intersection multiplicity one case, the Theorem gives immediately:

COROLLARY 15.2.2. *Let C, D, E be distinct cubics, E irreducible. If $D \cap E = \{q_1, \ldots, q_9\}$ (distinct points) and C passes through $q_1, \ldots, q_8$, then it passes through q_9.*

Actually this is true without assuming E irreducible (provided E doesn't share any components with D or C), but we won't prove that.

PROOF. *(of Theorem)* First assume E is smooth. Recall that the quotient of two homogeneous polynomials — say, F/G — yields a meromorphic function on $\mathbb{P}^2$. By Example 7.3.6, since E intersects $D = \{G = 0\}$ in points, we may pull this back to E:

$$f := \left.\frac{F}{G}\right|_E \in \mathcal{K}(E)^*.$$

Suppose (for a contradiction) that $q \neq q_9$. Since $C = \{F = 0\}$ and $D = \{G = 0\}$, the divisor of f is evidently

$$(f) = C \cdot E - D \cdot E = [q] - [q_9] \in \mathrm{Div}(E).$$

This says that f has one zero (at q) and one pole (at q_9); hence, as a holomorphic map of Riemann surfaces $E \to \mathbb{P}^1$, f has mapping degree 1. That is, f is 1-to-1; and since (using the open mapping theorem) its image must be open and closed (and $\mathbb{P}^1$ is connected), f is surjective. So f gives an isomorphism $E \cong \mathbb{P}^1$. Trouble is, this is total rubbish. Since E is a smooth cubic, its genus is 1 by the genus formula, whereas the genus of $\mathbb{P}^1$ is zero. So they can't be isomorphic for purely topological reasons. This contradication tells us that, indeed, our assumption $q \neq q_9$ was wrong, and so they are equal.

To extend this argument to the case where E is singular with ODP $\hat{p}$, first pull back $\frac{F}{G}$ along the normalization $\sigma : \tilde{E} \to \mathbb{P}^2$ (of E) to obtain $f \in \mathcal{K}(\tilde{E})$. We regard f as a map from $\tilde{E} \to \mathbb{P}^1$. As before, assuming $q \neq q_9$ leads to $\deg(f) = 1$. However, a different objection to "$\deg(f) = 1$" will be required as there is no topological obstruction: indeed, $\tilde{E} \cong \mathbb{P}^1$ by the genus formula (a nodal cubic has genus zero normalization). So argue as follows: since $\hat{p} \notin C, D$, we find that $\frac{F}{G} \in \mathcal{K}(\mathbb{P}^2)$ is well-defined at $\hat{p}$, so its pullback via σ cannot "separate" the two branches of E there. That is, at the two points of $\tilde{E}$ mapping to $\hat{p}$ (under σ), f will take the same value. But then, the mapping degree of f cannot be 1.

The other possibility is that $\hat{p}$ is a cusp.[2] We may assume $\hat{p} = [1 : 0 : 0]$ and the equation is of the form $x^3 = y^2$. Again we need to show that $f = \sigma^* \frac{F}{G}$, if nonconstant, cannot have mapping degree 1. Let $\tilde{p} = \sigma^{-1}(\hat{p})$, and write $R(x,y) := \frac{F(1,x,y)}{G(1,x,y)} - \frac{F(1,0,0)}{G(1,0,0)}$. Then $f(t) - f(0) = (\sigma^* R)(t) = R(t^2, t^3)$, and

$$\begin{aligned} \deg(f) &= \deg f^{-1}([f(0)]) \\ &\geq \mathrm{ord}_0(R(t^2,t^3)) \\ &\geq \underbrace{\mathrm{ord}_{(0,0)}(R(x,y))}_{\geq 1} \cdot \underbrace{\min\{\mathrm{ord}_0(t^2,t^3)\}}_{=2} \\ &\geq 2. \end{aligned}$$

□

Exercises

(1) Let C, D, and E be as above (defined by $F = 0$, $G = 0$, $H = 0$), of respective degrees d, d, e with $3 \leq e \leq d$. Suppose C and D intersect in d^2 distinct points, and assume that E is smooth (hence irreducible). Show that if E passes through $ed - 1$ of these, it passes through ed of them. (Imitate the argument from the proof of Theorem 15.2.1.)

---

[2] see the paragraph immediately preceding §16.1 below.

Part 3

Cubic curves

CHAPTER 16

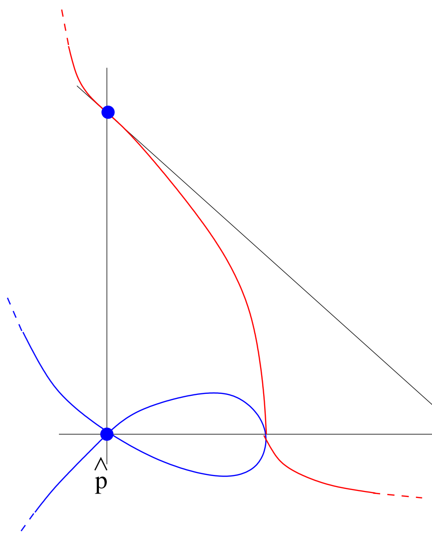
The singular cubic

Recall that a singular cubic curve[1] $D \subset \mathbb{P}^2$ is normalized via stereographic projection through its singular point $\hat{p}$; that is, we get a normalization morphism

$$\sigma : \mathbb{P}^1 \to \mathbb{P}^2$$

with image D. In particular, all singular cubics have normalization of genus zero. Moreover, they are all projectively equivalent to one of two examples.

The nodal cubic. A "node" is just an ordinary double point. Let $D = \{Y^2Z = X^2(Z - X)\}$; the affine equation is $y^2 = x^2(1 - x)$ and a schematic picture is



where I have denoted points with real coordinates in blue and points with only x-coordinate real in red. (How these sit inside the full set of complex points will be pictured below; the dotted stuff will connect up.) By Exercise 5 of Chapter 3, this is parametrized by

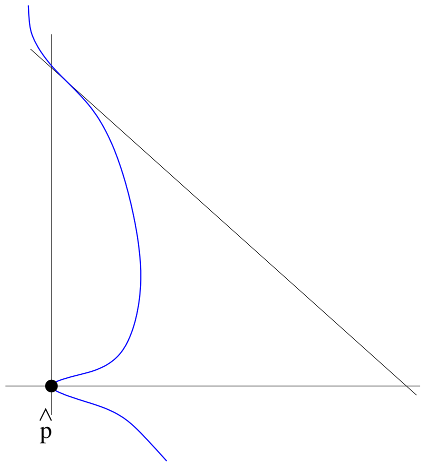
$$\varphi : \mathbb{P}^1/\{0, \infty\} \xrightarrow{\cong} D$$

[1] D is for "degenerate"

$$t \longmapsto \left(\frac{-4t}{(1-t)^2}, \frac{-t(1+t)}{(1-t)^3}\right) =: (x(t), y(t)).$$

The "$\mathbb{P}^1/\{0,\infty\}$" means the Riemann sphere with the top and bottom points identified.[2]

The cuspidal cubic. Take $D = \{Y^2Z = X^3\}$, affine equation $y^2 = x^3$, schematic picture



where I have only drawn real points. To do stereographic projection through the cusp $(0,0)$, write $y = \frac{1}{t}x$ and substitute to get $\frac{1}{t^2}x^2 = x^3 \implies x = \frac{1}{t^2}$. Hence we get a normalization

$$\varphi : \mathbb{P}^1 \to D$$

defined by

$$t \mapsto \left[1 : \frac{1}{t^2} : \frac{1}{t^3}\right].$$

One of our overarching themes in the next few chapters will be the study of algebro-geometrically defined group laws on cubics. In this chapter, we focus on the above two singular examples, as the smooth case is more difficult. For the nodal cubic, the law will turn out to be equivalent (via φ) to multiplication on $\mathbb{C}^* = \mathbb{P}^1\backslash\{0,\infty\}$; while in the cuspidal case, it identifies with addition on $\mathbb{C} = \mathbb{P}^1\backslash\{\infty\}$. In both cases, these sets are the preimages under normalization of the smooth points of D, which is where the group laws will be defined.

[2]Note that one can homogenize the formula for φ by $[T_0 : T_1] \longmapsto [(T_0 - T_1)^3 : -4T_1T_0(T_0 - T_1) : -T_1T_0(T_0 + T_1)]$, and then it is clear that $\varphi(1) = \varphi([1:1]) = [0:0:1]$.

In the course of studying such laws as well as addition theorems on these curves, we will pull back rational functions on $\mathbb{P}^2$ (quotients of homogeneous polynomials of equal degree, or equivalently elements of $\mathbb{C}(x, y)$) to get meromorphic functions on $\mathbb{P}^1$ (the normalization of our curve). So in illustrating the simplicity of the group law, hence

Principle 1: *Singularities make curves of a given degree more trivial and easier to study,*

we will be seeing a concrete example of the following

Principle 2: *Given* $C \subset \mathbb{P}^2$ *an* irreducible *algebraic curve with normalization* $\sigma : \tilde{C} \to \mathbb{P}^2$*, every* $f \in \mathcal{K}(\tilde{C})$ *is of the form* $\sigma^* F$*,* $F \in \mathbb{C}(x, y)$*.*

In other words, writing $C_0 := C \cap (\mathbb{P}^2 \backslash \{Z = 0\})$ for the affine part of C and $g_{C_0}(x, y)$ for its defining equation, if we define

$$\mathbb{C}[C_0] := \frac{\mathbb{C}[x, y]}{(g_{C_0})}, \qquad \begin{aligned} \mathbb{C}(C) :=& \quad \text{fraction field of } \mathbb{C}[C_0] \\ \cong& \quad \left\{ \sigma^* F \;\middle|\; \begin{array}{c} F \in \mathbb{C}(x, y) \\ F \not\equiv \infty \text{ on } C \end{array} \right\}, \end{aligned}$$

then Principle 2 says that

$$\begin{array}{ccc} \mathcal{K}(\tilde{C}) & \cong & \mathbb{C}(C). \\ \text{analytic} & & \text{algebraic} \end{array}$$

Since C was projective, $\tilde{C}$ is compact, and that turns out to be of fundamental importance: e.g., $\mathbb{C}[C_0]$ is only a subring of, rather than equal to, the ring of holomorphic functions on (the desingularization of) C_0.

Before continuing on, we should address one point: why should the only possible singularities of an irreducible cubic C be an ODP or cusp, and why must it have only one? First of all, if it had two singular points, then we could take a line L through those two points. Both intersection multiplicities (of C with L at these two points) would have to be ≥ 2, and so $(C \cdot L) \geq 4$ in violation of Bézout. (See what a useful theorem this is?) So C can only have one singular point, and as its equation is of degree 3 that point can only be of order 2 or 3. If it is of order 3, then by the result of Chapter 6 Exercise 5,

C is a union of 3 lines, contradicting irreducibility. Finally, the local equation about a *non*-ordinary double point of C can only be of the form $x^2 + f_3(x, y) = 0$, with f_3 homogeneous of degree 3. An explicit local analytic transformation puts this in the form $(\tilde{x})^2 + (\tilde{y})^3 = 0$. So it is a cusp. Alternately, anything which looks like $x^2 + y^n = 0$ has intersection multiplicity n with the line $x = 0$, again violating Bézout (in the context of our cubic curve) if $n > 3$.

16.1. Warm-up: Functions on a nonsingular conic

Our smooth conic is named C. Any $F \in \mathcal{K}(C)$ can be viewed as a map $C \to \mathbb{P}^1$. Composing this with the stereographically produced normalization $\sigma : \mathbb{P}^1 \overset{\cong}{\to} C$, yields

$$\begin{tikzcd} \mathbb{P}^1 \arrow[r, "\cong", "\sigma"'] \arrow[rr, bend right, "\sigma^* F"'] & C \arrow[r, "F"'] & \mathbb{P}^1, \end{tikzcd}$$

that is, a meromorphic function on $\mathbb{P}^1$. Since $\mathcal{K}(\mathbb{P}^1) = \mathbb{C}(t)$ (here $t := T_1/T_0$), $\sigma^* F$ must be of the form

$$\frac{g(t)}{h(t)} = \frac{G(T_0, T_1)}{H(T_0, T_1)}$$

where g, h, G, H are polynomials and G, H are homogeneous of the same degree. By the fundamental theorem of algebra, we can write this as

$$\gamma \cdot T_0^N \frac{\prod_i (T_1 - \alpha_i T_0)^{m_i}}{\prod_j (T_1 - \beta_j T_0)^{n_j}},$$

for some $\gamma, \alpha_i, \beta_j \in \mathbb{C}$. As $\deg G = \deg H \implies N + \sum m_i - \sum n_j = 0$, the expression simplifies to

$$\gamma \frac{\prod (t - \alpha_i)^{m_i}}{\prod (t - \beta_j)^{n_j}} \; (= (\sigma^* F)(t)).$$

Note that

$$(\sigma^* F)(\infty) \neq 0, \infty \iff \sum m_i = \sum n_j. \tag{16.1.1}$$

16.2. Functions on a nonsingular cubic (nodal case)

Let $F : D \to \mathbb{P}^1$ be

(16.2.1)
$$\begin{gathered} \text{the restriction to } D \text{ of a rational function on } \mathbb{P}^2 \\ \text{which is well-defined and } \neq 0, \infty \\ \text{at the singular point } \hat{p} \in D. \end{gathered}$$

Since the normalization $\mathbb{P}^1 \to D$ sends $0, \infty \mapsto \hat{p}$ but is otherwise 1-to-1, we get

$$\begin{tikzcd} (\mathbb{P}^1/\{0,\infty\}) \arrow[r, "\varphi", "\cong"'] \arrow[rr, bend right, "\varphi^* F"'] & D \arrow[r, "F"] & \mathbb{P}^1 \end{tikzcd}$$

with $F(0) = F(\infty) \in \mathbb{C}^*$. Henceforth we shall, by abuse of notation, refer to this composition as F.

Thinking of F as a meromorphic function on $\mathbb{P}^1$, (16.1.1) applies and we get

$$F(t) = \gamma \frac{\prod (t - \alpha_i)^{m_i}}{\prod (t - \beta_j)^{n_j}} \quad \text{with } \sum m_i = \sum n_j.$$

Furthermore,

$$\gamma = F(\infty) = F(0) = \gamma \frac{\prod \alpha_i^{m_i}}{\prod \beta_j^{n_j}}$$

so that

$$\prod \alpha_i^{m_i} = \prod \beta_j^{n_j}, \tag{16.2.2}$$

relating the z-coordinates of the zeroes and poles of F.

Now introduce the multivalued function

$$u := \int_1^* \frac{dt}{t} = \log(t)$$

on $\mathbb{P}^1$, which takes well-defined values in $\mathbb{C}/2\pi\sqrt{-1}\mathbb{Z}$. We can restate (16.2.2) in terms of u: viz.,

$$\sum_{p \in D} \nu_p(F) \cdot u(p) \equiv 0 \quad \text{mod } 2\pi\sqrt{-1}\mathbb{Z}.$$

This leads to *Abel's theorem for the singular cubic*:

PROPOSITION 16.2.1. *Given $\mathcal{P}, \mathcal{Z} \in Div(D \backslash \hat{p})$ effective*[3] *divisors of the same degree,*

$$\int_{\mathcal{P}}^{\mathcal{Z}} \frac{dz}{z} \equiv 0 \quad mod\, 2\pi\sqrt{-1}\mathbb{Z} \quad \Longleftrightarrow$$

$$\left.\begin{array}{l} \mathcal{P} = poles \\ \mathcal{Z} = zeroes \end{array}\right\} \text{ of some } F \text{ as in (16.2.1).}$$

[3]recall divisors are formal sums of points on a complex manifold with integer coefficients. A divisor is effective if none of those coefficients are negative.

Explicitly, if $\mathcal{P} = \sum n_j[\beta_j]$ and $\mathcal{Z} = \sum m_i[\alpha_i]$ are of the same degree $(d = \sum n_j = \sum m_i)$, then we may write $\mathcal{Z} - \mathcal{P} = \sum_{k=1}^{d}([z_k] - [p_k])$ and $\int_{\mathcal{P}}^{\mathcal{Z}} := \sum \int_{p_k}^{z_k}$ by some choice of paths. Also, in the statement "poles" and "zeroes" are as usual meant with multiplicity. This is a first "baby" case of a general statement for algebraic curves (Abel's theorem) connecting integrals of 1-forms to the question of when a divisor is the divisor of a meromorphic function.

16.3. Group law on the nodal cubic

Fix a normalization[4]

$$\varphi : \left(\mathbb{P}^1/\{0,\infty\}\right) \xrightarrow{\cong} D$$
$$t \longmapsto (x(t), y(t))$$
$$1 \longmapsto \varphi(1) =: \mathbf{e}.$$

Let $p, q \in D$ be arbitrary nonsingular points, and L_{pq} be the line through p and q. (If they are the same, then take L to be the tangent line T_pD.) By Bézout, $(L_{pq} \cdot D) = 3$ and so L_{pq} meets D in a third point which we call $p * q$. More precisely, everything is counted with multiplicity ($p * q$ need not be distinct from p or q) so we really mean

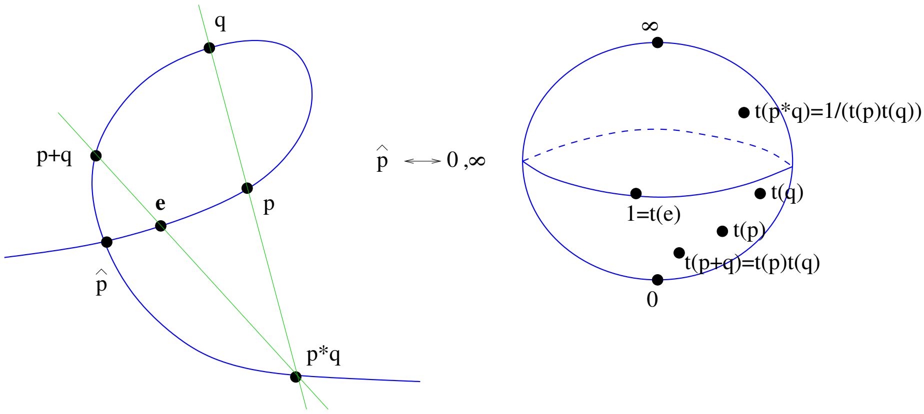
$$[p * q] := L_{pq} \cdot D - [p] - [q].$$

Now let L' be the line through $p * q$ and $\mathbf{e}$ (or $T_{\mathbf{e}}D$ if they coincide), and put

$$[p + q] := L' \cdot D - [p * q] - [\mathbf{e}].$$

That is, $p + q$ is the "extra" intersection point of this line with D guaranteed by Bézout. Here's a useful picture of the construction:

[4]when we need to use homogemeous coordinates, $\varphi(t) = [Z(t) : X(t) : Y(t)]$



Now writing f_L for the equation of a line L, observe that

$$F := \left.\frac{f_{L_{pq}}}{f_{L'}}\right|_D : D \longrightarrow \mathbb{P}^1$$

satisfies (16.2.1). In terms of the t-coordinate on $\mathbb{P}^1$, i.e. pulling F back along φ, we must have:

$$F(t) = \gamma \frac{(t-t(p))(t-t(q))(t-t(p*q))}{(t-t(p+q))(t-1)(t-t(p*q))} = \gamma \frac{(t-t(p))(t-t(q))}{(t-t(p+q))(t-1)}.$$

But since $F(0) = F(\infty)$, by (16.2.2)

$$\prod\{\text{locations of zeroes}\} = \prod\{\text{locations of poles}\}$$

$$\implies \quad t(p)\cdot t(q) = t(p+q)\cdot \underbrace{t(\mathbf{e})}_{1} = t(p+q).$$

This identifies the group law (multiplication) on $\mathbb{C}^* = \mathbb{P}^1\backslash\{0,\infty\}$ with the one just defined on $D\backslash\hat{p}$. Alternately, taking log gives

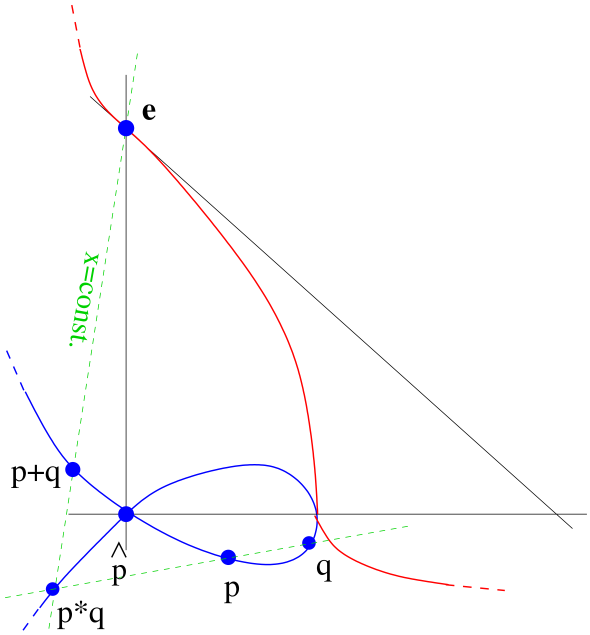
$$u(p) + u(q) \equiv u(p+q) \quad \mod 2\pi\sqrt{-1}\mathbb{Z},$$

identifying addition on $D\backslash\hat{p}$ with addition in $\mathbb{C}/2\pi\sqrt{-1}\mathbb{Z}$. This may be rewritten

$$\text{(16.3.1)} \qquad \int_1^{t(p)} \frac{dt}{t} + \int_1^{t(q)} \frac{dt}{t} \equiv \int_1^{t(p+q)} \frac{dt}{t} \quad \mod 2\pi\sqrt{-1}\mathbb{Z}.$$

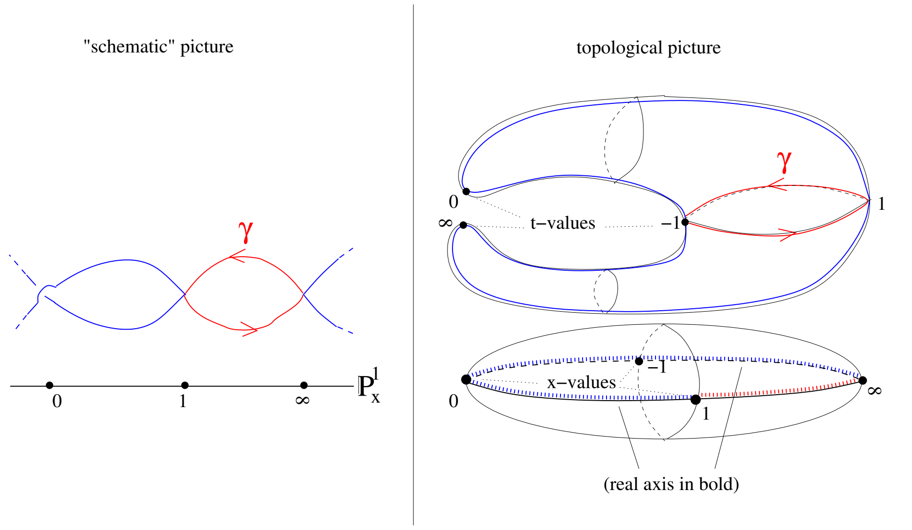
16.4. Addition theorems for the nodal cubic

Let's unwind the "equivalence of group laws" in the nodal cubic example from the beginning of the chapter. Noting that $\mathbf{e} := \varphi(1) = [0:0:1]$, here is a picture of how the group law works:



In particular, the x-coordinates of $p+q$ and $p*q$ are the same, while the y-coordinates are $\pm$ of each other.

Just to clarify the topology of the situation, here is what the projection of the normalization of D onto the x-axis looks like:



It is a 2-sheeted cover with 2 branch points, with the closed path γ indicating the "equator" (or unit circle $|t| = 1$) on the upper $\mathbb{P}^1$ (i.e. $\tilde{D}$).

Now we get to work. Start by "inverting" the equivalence $t(p) \cdot t(q) = t(p+q)$:

$$\underbrace{\varphi(t_1)}_{p} + \underbrace{\varphi(t_2)}_{q} = \underbrace{\varphi(t_1 \cdot t_2)}_{p+q}.$$

Since p, q, and $p * q$ are collinear by construction,

$$0 = \det \begin{pmatrix} Z(p) & Z(q) & Z(p*q) \\ X(p) & X(q) & X(p*q) \\ Y(p) & Y(q) & Y(p*q) \end{pmatrix}.$$

Assuming none of them is $\mathbf{e}$, $Z(p)Z(q)Z(p*q)) \neq 0$ and we get the

1st addition theorem:

$$0 = \det \begin{pmatrix} 1 & 1 & 1 \\ x(t_1) & x(t_2) & x(t_1 \cdot t_2) \\ y(t_1) & y(t_2) & -y(t_1 \cdot t_2) \end{pmatrix}.$$

This allows you to compute $x(t_1 \cdot t_2)$ from $x(t_1)$ and $x(t_2)$, using the equation of D to write $y(t) = \pm x(t)\sqrt{1 - x(t)}$.

Next, $\varphi^* \left(\frac{dx}{y}\Big|_D \right) = \frac{d(x(t))}{y(t)} = \cdots$ [use Exercise 4 from Chap. 13]$\cdots = \frac{dt}{t}$, while $\frac{dx}{y}\Big|_D = \frac{dx}{\pm x\sqrt{1-x}}$; so (16.3.1) may be expressed

$$\int_{x(\mathbf{e})(=\infty)}^{x(p)} \frac{dx}{x\sqrt{1-x}} + \int_{x(\mathbf{e})}^{x(q)} \frac{dx}{x\sqrt{1-x}} \equiv \int_{x(\mathbf{e})}^{x(p+q)} \frac{dx}{x\sqrt{1-x}}.$$

(Note that $2\pi\sqrt{-1} = \oint_{|t|=1} \frac{dt}{t} = \int_\gamma \frac{dx}{y}$. Going modulo its integer multiples, which is what "$\equiv$" means here, is necessary not to have the equation's correctness depend upon the choice of paths from ∞ to $x(p)$, to $x(q)$, and to $x(p+q)$.) Solving

$$\det \begin{pmatrix} 1 & 1 & 1 \\ x(p) & x(q) & x(p+q) \\ x(p)\sqrt{1-x(p)} & x(q)\sqrt{1-x(q)} & -x(p+q)\sqrt{1-x(p+q)} \end{pmatrix} = 0$$

for $x(p+q)$ yields

$$x(p+q) = \frac{-x(p)x(q)}{\left(\sqrt{1-x(p)} + \sqrt{1-x(q)}\right)^2}.$$

Forgetting the association with p, q, $p+q \in D$ we get the

2nd addition theorem:

$$\int_\infty^{x_1} \frac{dx}{x\sqrt{1-x}} + \int_\infty^{x_2} \frac{dx}{x\sqrt{1-x}} \equiv \int_\infty^{\frac{-x_1x_2}{(\sqrt{1-x_1}+\sqrt{1-x_2})^2}} \frac{dx}{x\sqrt{1-x}} \quad \mod 2\pi\sqrt{-1}\mathbb{Z}.$$

Note that $\int_\infty^x \frac{dx}{x\sqrt{1-x}} = \log\left(\frac{\sqrt{1-x}-1}{\sqrt{1-x}+1}\right)$ by explicit computation of the integral. (One way to view this function is $\log(t)\,(= u)$ viewed as a multivalued function of x.) So we have discovered a functional equation for $\log\left(\frac{\sqrt{1-x}-1}{\sqrt{1-x}+1}\right)$, which is ugly to check by hand.

One aspect of the game we have just played here is: start with a "natural" choice of differential 1-form on the curve (if possible, one which is smooth away from any singularities of the curve). In the above, this was $\frac{dx}{y}|_D$. You can think of this as a multivalued 1-form on the x-axis, and then D is the "existence domain of the 1-form" over $\mathbb{P}^1_x$. Then you integrate this 1-form, which gives a transcendental function which is multivalued even on D (you have to go to its universal cover to make it well-defined), and try to produce a functional equation for it (as a function of x). In the last section we'll summarize this story for a couple of other curves.

16.5. Other addition theorems (conic, cuspidal cubic)

Consider the example $C = \{y^2 + x^2 = 1\}$, parametrized by $\mathbb{P}^1$ via

$$t \overset{\varphi}{\longmapsto} \left(\frac{-2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right)$$

as in §3.3. We compute

$$\varphi^*\left(\left.\frac{dx}{y}\right|_C\right) = \frac{2dt}{t^2+1} = 2d(\arctan(t)),$$

$$\left.\frac{dx}{y}\right|_C = \frac{dx}{\sqrt{1-x^2}} = d(\arcsin(x)).$$

On the universal cover of $\mathbb{P}^1_x\backslash\{\pm 1\}$ let $\theta = \arcsin(x)$ (starting at $x = 0 \leftrightarrow t = 0 \leftrightarrow \theta = 0$).[5] Its role is similar to that of $u = \log t$ above, as the integral of our chosen differential 1-form on the curve; θ takes well-defined values in $\mathbb{C}/2\pi\mathbb{Z}$. Writing $x(\theta_1) =: x_1$, $x(\theta_2) =: x_2$, $x(\theta_1+\theta_2) =: x_{12}$, the standard trigonometry relations give

$$x_{12} = x_1\sqrt{1-x_2^2} + x_2\sqrt{1-x_1^2}.$$

[5]Note: $t \mapsto \frac{-2t}{t^2+1} = x$ is a degree-2 map (from C to the x-axis) with ramification points $t = \pm 1$ over $x = \pm 1$. On the complements of these points, we have a 2-to-1 map $\mathbb{C}^* \to \mathbb{C}^*$. The universal cover of $\mathbb{C}^*$ is $\mathbb{C}$, and so we have maps $\mathbb{C} \to \mathbb{C}^*_t \to \mathbb{C}^*_x$ sending θ to t to x. So our setup encodes the relation $\frac{-2\tan\frac{\theta}{2}}{(\tan\frac{\theta}{2})^2+1} = x(t(\theta)) = x(\theta) = \sin(\theta)$.

The "second addition formula" for the conic then reads

$$\int_0^{x_1} \frac{dx}{\sqrt{1-x^2}} + \int_0^{x_2} \frac{dx}{\sqrt{1-x^2}} \equiv \int_0^{x_{12}} \frac{dx}{\sqrt{1-x^2}} \mod 2\pi\mathbb{Z},$$

which is a functional equation for *arcsin*. More simply put, it is just the inverse of the trigonometric identity.

Next, look back to the cuspidal example from the beginning of the chapter. We have

$$\varphi^*\left(\frac{dx}{y}\bigg|_D\right) = \frac{d(x(t))}{y(t)} = \frac{d\left(\frac{1}{t^2}\right)}{\frac{1}{t^3}} = -2dt,$$

while

$$\frac{dx}{y}\bigg|_D = \frac{dx}{x^{\frac{3}{2}}}.$$

(Note that this time, the integral of $\frac{dx}{y}|_D$ is just $-2t$ and is *not* multi-valued on D.) Clearly if $t_{12} = t_1 + t_2$, then

$$\int_0^{t_1} dt + \int_0^{t_2} dt = \int_0^{t_{12}} dt$$

$$\Longrightarrow \quad \int_\infty^{\frac{1}{t_1^2}(=x(t_1))} \frac{dx}{x^{\frac{3}{2}}} + \int_\infty^{\frac{1}{t_2^2}(=x(t_2))} \frac{dx}{x^{\frac{3}{2}}} = \int_\infty^{\frac{1}{t_{12}^2}(=x(t_1+t_2))} \frac{dx}{x^{\frac{3}{2}}}.$$

So we get a functional equation for $\frac{1}{\sqrt{x}}$, which is unfortunately rather stupid: it says

$$\frac{1}{\left(\frac{1}{t_1^2}\right)^{\frac{1}{2}}} + \frac{1}{\left(\frac{1}{t_2^2}\right)^{\frac{1}{2}}} = \frac{1}{\left(\frac{1}{(t_1+t_2)^2}\right)^{\frac{1}{2}}}.$$

In an exercise below, you will show a less trivial addition theorem for the cuspidal cubic, to the effect that

$$P, Q, R \in (D \backslash \hat{p}) \text{ are collinear} \iff t(P) + t(Q) + t(R) = 0.$$

Exercises

(1) Consider the cuspidal cubic curve $D = \{Y^2 Z = X^3\} \subseteq \mathbb{P}^2$ and normalize it as above, with $\varphi : \mathbb{P}^1 \to D$ given by $t \mapsto [1 : \frac{1}{t^2} : \frac{1}{t^3}] = [Z : X : Y]$. (The singular point is $\hat{p} = [1 : 0 : 0]$.) Prove directly that the group law given by addition on $(\mathbb{P}^1 \backslash \{\infty\}) \cong \mathbb{C}$ (namely, $t_1, t_2 \mapsto t_1 + t_2$) corresponds to the following process on $(D \backslash \{\hat{p}\})$: take the line L through $\varphi(t_1)$ and $\varphi(t_2)$, then a line L' through the third intersection point $\varphi(t_1) * \varphi(t_2)$ (of L with D) and the "neutral" point $[0 : 0 : 1]$, and finally locate the third intersection

point of this L' with D to get "$\varphi(t_1)+\varphi(t_2)$" (also as above, for the nodal cubic). Do this simply *by showing that* $P, Q, R \in (D\backslash\{\hat{p}\})$ *are collinear if and only if* $t(P)+t(Q)+t(R)=0$. (Here P, Q, R are distinct.) Hint: use the determinant of the matrix

$$\begin{pmatrix} a^3 & a & 1 \\ b^3 & b & 1 \\ c^3 & c & 1 \end{pmatrix},$$

and rewrite $[1 : \frac{1}{t^2} : \frac{1}{t^3}] = [t^3 : t : 1]$.

CHAPTER 17

Putting a nonsingular cubic in standard form

An irreducible algebraic curve $E \subset \mathbb{P}^2$ is an *elliptic curve* if the genus of its normalization $\tilde{E}$ is 1 (topologically it looks like a donut). By the genus formula, all *smooth cubic curves* are elliptic. In the next two chapters we will show not only that such a curve is isomorphic to $\mathbb{C}/\Lambda$ for some lattice Λ, but will get a description of Λ which shows its dependence on E. This is important, since for two different lattices $\Lambda = \mathbb{Z}\langle\alpha, \beta\rangle$ and $\Lambda' = \mathbb{Z}\langle\alpha', \beta'\rangle$, the complex 1-tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ need not be isomorphic as Riemann surfaces. (More precisely, they are isomorphic if and only if $[\alpha : \beta]$ is carried to $[\alpha' : \beta']$ by an integral projectivity, i.e. a transformation of $\mathbb{P}^1$ induced by $A \in PSL_2(\mathbb{Z})$.)

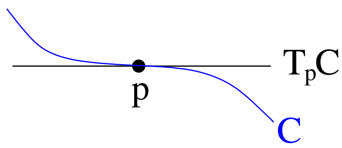
Even more significant is how we do this: by putting E in Weierstrass form, integrating a holomorphic form on it to get a map to a complex torus, and showing that the Weierstrass $\wp$-function and its derivative invert this map. To put E in this form, a choice of *flex* is required. What is that?

17.1. Flexes

Let $C = \{F(Z, X, Y) = 0\} \subset \mathbb{P}^2$ be an irreducible algebraic curve of degree $d \geq 3$. One way of thinking of the tangent line at a nonsingular point $p \in C$ is as the unique line satisfying $(C \cdot T_pC) \geq 2$.

Definition 17.1.1. A smooth point $p \in C$ is called a *flex* if the intersection multiplicity

$$(C \cdot T_pC) \geq 3.$$



Intuitively these are the inflection points of C, and can be seen to correspond to cusps of the dual curve $\check{C}$ (see §4.4). Since $\check{C}$ has finitely many singularities, this gives one proof that there are finitely many flexes; we will however take a different approach.

Denoting partial derivatives by subscript, e.g. $F_{ZX} := \frac{\partial^2 F}{\partial Z \partial X}$, the *Hessian* of F is the polynomial matrix

$$Hess_F = \begin{pmatrix} F_{ZZ} & F_{ZX} & F_{ZY} \\ F_{XZ} & F_{XX} & F_{XY} \\ F_{YZ} & F_{YX} & F_{YY} \end{pmatrix}.$$

Its determinant

$$H := \det(Hess_F)$$

is clearly a homogeneous polynomial of degree $3(d-2)$. Call $\mathcal{H}_C := \{H(Z,X,Y) = 0\} \subset \mathbb{P}^2$ the *Hessian curve* associated to C.

LEMMA 17.1.2. *Let $p \in C$ be a smooth point. Then p is a flex $\iff$ $p \in \mathcal{H}_C$.*

PROOF. Since intersection numbers are invariant under projectivities, we may assume $p = [1 : 0 : 0]$, $T_pC = \{Y = 0\}$. In affine coordinates, writing $f(x, y) := F(1, x, y)$, this means that the curve $\{f(x, y) = 0\} \subset \mathbb{C}^2$ contains $(0, 0)$ and is tangent to $\{y = 0\}$. So $f(0, 0) = 0$ and $(f_x(0, 0), f_y(0, 0)) = (\lambda, 0)$ where $\lambda \neq 0$, so that

$$f(x, y) = \lambda y + (ax^2 + 2bxy + cy^2) + \text{higher-order terms}.$$

Parametrizing T_pC by $t \mapsto (t, 0)$, we have

$$(C \cdot T_pC)_p = \text{ord}_0\,(f(t, 0)) = \text{ord}_0(at^2 + \text{h.o.t.}),$$

which is ≥ 3 (yielding a flex) if and only if $a = 0$.

Now the above form of f implies

$$F(Z, X, Y) = \lambda Y Z^{d-1} + (aX^2 + 2bXY + cY^2)Z^{d-2} + \cdots$$

so that

$$Hess_F(1, 0, 0) = \begin{pmatrix} 0 & 0 & (d-1)\lambda \\ 0 & 2a & 2b \\ (d-1)\lambda & 2b & 2c \end{pmatrix}.$$

Taking the determinant,

$$H(p) = \det(Hess_F(p)) = -2(d-1)^2\lambda^2 a.$$

This is clearly zero (i.e. $p \in \mathcal{H}_C$) if and only if $a = 0$. $\square$

Now Bezout guarantees intersections of C and $\mathcal{H}_C$. If C is singular then these might all be at singular points, so that there might be no flexes (though this isn't typical: see the exercises). On the other hand, if C is smooth then by Lemma 17.1.2 we do have flexes. Refining this observation:

PROPOSITION 17.1.3. *On a nonsingular curve C of degree $d \geq 3$, there exists at least one and at most $3d(d-2)$ flexes.*

PROOF. By Bezout,

$$\sum_{p \in C \cap \mathcal{H}_C} (C \cdot \mathcal{H}_C)_p = (C \cdot \mathcal{H}_C) = \deg(C) \cdot \deg(\mathcal{H}_C) = d \cdot 3(d-2).$$

So the number of points in $C \cap \mathcal{H}_C$ is between 1 and $3d(d-2)$, all points are smooth points, and we apply Lemma 17.1.2. $\square$

REMARK 17.1.4. Since $Hess_F$ is just the multivariable derivative (Jacobian matrix) of $\mathcal{D}_C : \mathbb{P}^2 \to \mathbb{P}^2$ (§4.4), the intersections of C and $\mathcal{H}_C$ may be viewed as degeneracies of the map $\mathcal{D}_C|_C : C \twoheadrightarrow \check{C}$. This is what gives rise to the cusps in $\check{C}$ referred to above.

DEFINITION 17.1.5. The *multiplicity* of a flex $p \in C$ is defined to be $(C \cdot \mathcal{H}_C)_p$.

Now take $C = E$ to be a smooth elliptic curve ($d = 3$). Then in the proof of Lemma 17.1.2, the precise form of the homogeneous polynomial is

(17.1.1)

$$F(Z, X, Y) = \lambda Y Z^2 + (aX^2 + 2bXY + cY^2)Z + \alpha X^3 + \beta X^2 Y + \gamma X Y^2 + \delta Y^3.$$

Assume $a = 0$ so that we have a flex at $[1:0:0]$. (Note that α must then be nonzero, in order that Y not divide F — which would make E reducible hence singular.) Then a short computation gives

$$Hess_F(1, x, y) = \begin{pmatrix} 2\lambda y & 2by & 2\lambda + 2bx \\ 2by & 6\alpha x + 2\beta y & 2\beta x + 2\gamma y + 2b \\ 2\lambda + 2bx & 2\beta x + 2\gamma y + 2b & 6\delta y + 2\gamma x + 2c \end{pmatrix}.$$

Pull this back to $T_p E = \{y = 0\}$ by making the substitution

$$Hess_F(1, t, 0) = \begin{pmatrix} 0 & 0 & 2\lambda + 2bt \\ 0 & 6\alpha t & 0 \\ 2\lambda + 2bt & 0 & 2\gamma t + 2c \end{pmatrix};$$

this has determinant

$$H(1,t,0) = -(2\lambda + 2bt)^2 6\alpha t,$$

and since $\alpha, \lambda \neq 0$

$$(T_pE \cdot \mathcal{H}_E)_p = \mathrm{ord}_0(H(1,t,0)) = 1.$$

So $\mathcal{H}_E$ is smooth at p and T_pE is *not* its tangent line. But then it intersects E transversely (since they have distinct tangent lines), so that $(E \cdot \mathcal{H}_E)_p = 1$. This computation is valid at any flex of E (after a projective change of coordinates, of course), and so proves:

PROPOSITION 17.1.6. *Any smooth cubic has* 9 *flexes, each of multiplicity one.*

PROOF. Since $\deg(\mathcal{H}_E) = 3(d-2) = 3$, Bezout gives us 9 intersection points of $\mathcal{H}_E$ and E, counted with multiplicity; and we have demonstrated that the multiplicities are all 1. $\square$

17.2. Weierstrass form

Consider an arbitrary smooth cubic curve

$$E = \{F(Z,X,Y) = 0\} \subset \mathbb{P}^2.$$

In this section we will show that there exists a choice of projective coordinates putting E uniquely into a convenient form. (Alternately, you can view this as the existence of a projectivity putting E into this form, in the same coordinates.)

We know E has a flex, and first of all we can choose coordinates so that this is at $[0:0:1] =: \mathcal{O}$ with $T_{\mathcal{O}}E = \{Z = 0\}$. To get the general equation of such a cuve: take (17.1.1), set $a = 0$ (for a flex), swap Z and Y, and (without loss of generality since $\lambda \neq 0$) normalize λ to 1; this gives

$$F(Z,X,Y) = ZY^2 + (2bXZ + cZ^2)Y + \alpha X^3 + \beta X^2 Z + \gamma XZ^2 + \delta Z^3,$$

with affine form

$$f(x,y) := F(1,x,y) = y^2 + yf_2(x) + f_3(x).$$

Now the discriminant

$$\mathcal{D}_y(f(x,y)) \;=\; \mathcal{R}_y(y^2 + yf_2(x) + f_3(x),\, 2y + f_2(x))$$

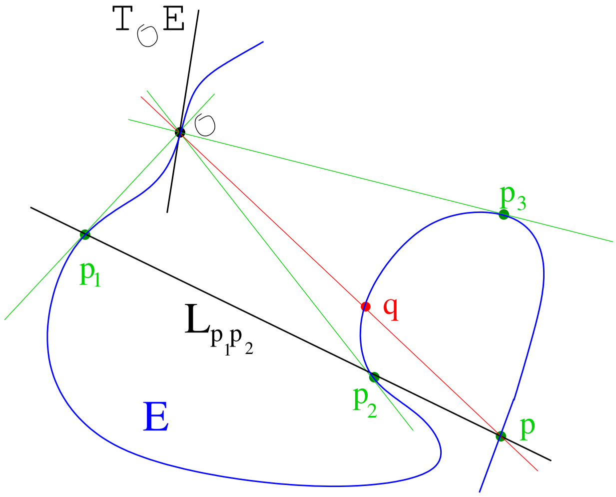
$$= \det \begin{pmatrix} 1 & f_2 & f_3 \\ 2 & f_2 & 0 \\ 0 & 2 & f_2 \end{pmatrix} = \det \begin{pmatrix} 1 & f_2 & f_3 \\ & -f_2 & -2f_3 \\ & 2 & f_2 \end{pmatrix}$$

$$= -f_2^2 + 4f_3 = -(2bx+c)^2 + 4(\alpha x^3 + \beta x^2 + \gamma x + \delta)$$

is a polynomial in x of degree 3 since $\alpha \neq 0$. Roots of $(\mathcal{D}_y(f))(x)$ correspond to vertical lines $x = x_0$ which are tangent to (the affine part of) E at some point. Bezout tells us that the intersection number there can only be 2, since $\deg(E) = 3$ and $\{X = x_0 Z\}$ already meets E at $\mathcal{O}$. Such "first order" tangencies mean the roots each have multiplicity one. Therefore E has three vertical tangents (apart from $L_\infty = \{Z = 0\}$), at p_1, p_2, p_3.

Lemma 17.2.1. *The* $\{p_i\}_{i=1}^3$ *are collinear.*

Proof. In the picture



define p to be the third intersection point of $L_{p_1p_2}$ and E, and q the third intersection point of $L_{\mathcal{O}p}$ with E. Consider the cubic curves $C_1 = E$, $C_2 = L_{\mathcal{O}p_1} + L_{\mathcal{O}p_2} + L_{\mathcal{O}p_3}$, and $C_3 = T_{\mathcal{O}}E + 2L_{p_1p_2}$. We have

$$C_1 \cdot C_2 = 3\mathcal{O} + 2p_1 + 2p_{22} + p + q$$

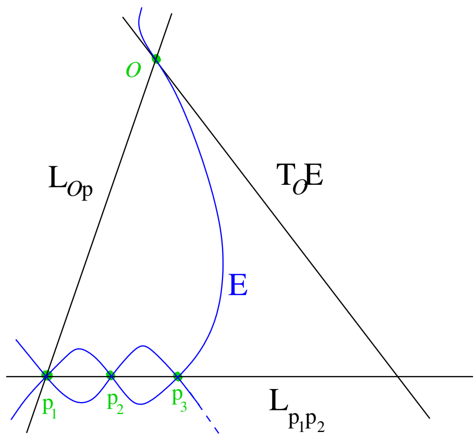
and

$$C_1 \cdot C_3 = 3\mathcal{O} + 2p_1 + 2p_2 + 2p.$$

Arguing as in §15.2, the ratio of the homogeneous polynomials defining C_2 and C_3 gives a degree 1 map $E \to \mathbb{P}^1$ (which is impossible) if $p \neq q$. So $p = q$, and $L_{\mathcal{O}p}$ is tangent to E at p. It follows that p is p_1, p_2, or

p_3. The first two are impossible since the tangent to p_1 doesn't pass through p_2 and vice versa; so $p = p_3$. Hence $p_1, p_2, p_3 \in L_{p_1p_2}$. $\square$

Now stereographic projection from $\mathcal{O}$ to $L_{p_1p_2}(\cong \mathbb{P}^1)$ presents E as a $2:1$ cover of $\mathbb{P}^1$ branched over p_1, p_2, p_3, and the image $T_{\mathcal{O}}E \cap L_{p_1p_2}$ of $\mathcal{O}$. Furthermore $L_{p_1p_2}$, $L_{\mathcal{O}p_1}$, $T_{\mathcal{O}}E$ form a triangle, and so we can choose new projective coordinates X', Y', Z' in order that $L_{p_1p_2} = \{Y' = 0\}$, $L_{\mathcal{O}p_1} = \{X' = 0\}$, and $T_{\mathcal{O}}E = \{Z' = 0\}$. For simplicity I'll drop the primes and just write X, Y, Z for this *new* coordinate system. The following picture summarizes what we know:



where (on $Y = 0$) p_1 is at $\frac{X}{Z} = 0$. Write α_1 (resp. α_2) for the value of $\frac{X}{Z}$ at p_2 (resp. p_3).

We would like an equation to correspond to this picture. Now, in the new coordinate system, the equation of E is still of the form

$$F(Z, X, Y) = ZY^2 + (2bXZ + cZ^2)Y + \alpha X^3 + \beta X^2 Z + \gamma XZ^2 + \delta Z^3,$$

because we still have a flex at $[0:0:1]$ with tangent line $Z = 0$. But now (referring to the picture) also $[1:0:0] \in E$, which implies $\delta = 0$. Moreover, $F_Y(= 2YZ + 2bXZ + cZ^2) = 0$ at $p_1 = [1:0:0]$, $p_2 = [1:\alpha_1:0]$, and $p_3 = [1:\alpha_2:0]$ since the tangents are vertical there. This yields $c = 0$, then $2b\alpha_1 = 2b\alpha_2 = 0$. As the $\{p_i\}$ are distinct (so $\alpha_i \neq 0$), we have , and

$$\begin{aligned} F(Z, X, Y) &= Y^2 Z + X(\alpha X^2 + \beta XZ + \gamma Z^2) \\ &= Y^2 Z + \alpha X(X - \alpha_1 Z)(X - \alpha_2 Z). \end{aligned}$$

Now define new coordinates by the projective transformation

$$X = \sqrt[3]{\frac{4}{\alpha}}X_0 + \frac{\alpha_1 + \alpha_2}{3}Z_0\ ,\quad Y = iY_0\ ,\quad Z = Z_0\ ,$$

which makes the equation

$$\begin{aligned}\tilde{F}(Z_0, X_0, Y_0) &= F\left(\sqrt[3]{\frac{4}{\alpha}}X_0 + \frac{\alpha_1 + \alpha_2}{3}Z_0,\ iY_0,\ Z_0\right)\\ &= -Y_0^2 Z_0 + 4X_0^3 - g_2 X_0 Z_0^2 - g_3 Z_0^3.\end{aligned}$$

Dropping the subscript 0's and taking the affine equation, we have put E in *Weierstrass form*:

PROPOSITION 17.2.2. *(a) Any smooth cubic $E \subset \mathbb{P}^2$ is projectively equivalent to a curve with affine equation of the form*[1]

$$y^2 = 4x^3 - g_2 x - g_3.$$

(b) This form is unique (given E) up to a change of the form $(g_2, g_3) \mapsto (\xi^4 g_2, \xi^6 g_3)$ where $\xi \in \mathbb{C}^*$; in particular,

$$j := \frac{g_2^3}{g_2^3 - 27g_3^2} \in \mathbb{C}$$

is an invariant of E.

PROOF. We have just seen (a). To see (b), write the projective equation $Y^2 Z = 4X^3 - g_2 XZ^2 - g_3 Z^3$. It is not difficult to see that any linear transformation preserving the form of this equation (up to rescaling) has the form $X = \varepsilon X_0$, $Y = \eta Y_0$, $Z = \frac{\varepsilon^3}{\eta^2} Z_0$. Taking $\beta := \frac{\varepsilon}{\eta}$ gives exactly the claimed effect on (g_2, g_3), and j is unchanged by this transformation. □

Exercises

(1) Show that the cubic curve $C = \{0 = X^3 + Y^3 - XY(X + Y + Z)\} \subset \mathbb{P}^2$ has one (ODP) singular point and exactly three collinear flexes. [Hint: start by computing the Hessian, then find the Hessian curve and determine its intersections with C.]

[1] note that the vanishing of the x^2 term of the right-hand side indicates that its roots sum to zero

CHAPTER 18

Canonical normalization of the Weierstrass cubic

This chapter will focus on the precise relationship between Weierstrass-form elliptic curves and complex 1-tori (or equivalently, 2-lattices in $\mathbb{C}$). We will begin by associating to a Weierstrass cubic E a "period lattice" Λ_E, and to a (full) lattice Λ a Weierstrass cubic E_Λ. These will ultimately be shown to be bijections of sets and mutual inverses. The key step is the inversion of the Weierstrass $\wp$-function and its derivative (embedding a 1-torus in $\mathbb{P}^2$) by the *Abel map* $u : E \to \mathbb{C}/\Lambda_E$. This map is closely related to the *elliptic integral*

$$\int_\infty^* \frac{dx}{\pm\sqrt{4x^3 - g_2 x - g_3}},$$

a variant of which will be studied in the exercises.

18.1. Holomorphic forms on an elliptic curve

Let E be a Weierstrass cubic, viz., the projective closure of

$$f(x,y) := y^2 - Q(x) = 0$$

in $\mathbb{P}^2$, where

$$Q(x) = (x-e_1)(x-e_2)(x-e_3)\ , \qquad e_1 + e_2 + e_3 = 0.$$

CLAIM 18.1.1. $\omega := \left.\frac{dx}{y}\right|_E \in \Omega^1(E)$ is nowhere vanishing.

REMARK 18.1.2. This statement perhaps requires clarification. You may interpret $\left.\frac{dx}{y}\right|_E$ in either of two equivalent ways:
(a) any algebraic differential form (such as $\frac{dx}{y}$) on $\mathbb{C}^2$ extends to a meromorphic form on $\mathbb{P}^2$, and you can think of $|_E$ as shorthand for pullback to E (rather than introducing $\sigma : E \hookrightarrow \mathbb{P}^2$ just to write $\sigma^* \frac{dx}{y}$);
(b) alternatively, writing $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$ exhibits x and y as meromorphic functions on $\mathbb{P}^2$ (and hence, via pullback, on E), and Example 13.1.4 tells us that $\frac{d(\,x|_E)}{y|_E}$ is a meromorphic 1-form.
Either way, we have $\omega \in \mathcal{K}^1(E)$; and part of the content of the Claim is

that ω is actually holomorphic: $\nu_p(\omega) \geq 0$ for all $p \in E$. The "nowhere vanishing" statement says that actually $\nu_p(\omega) = 0$ for all p.

PROOF. Look at the affine part $E \backslash \mathcal{O}$. Wherever $f_y \neq 0$, so that x gives a local coordinate, $\frac{dx}{y}\Big|_E$ is holomorphic and nonvanishing. We have $f = 0$ and $f_y = 0$ precisely at the three points $\{(e_i, 0)\}_{i=1,2,3}$, where $f_x = Q'(e_i) \neq 0$ so that y is a local coordinate. On E we have $0 = df = 2ydy - Q'(x)dx$ so that $\frac{dx}{y}\Big|_E =$

$$2\frac{dy}{Q'(x)}\Bigg|_E,$$

which is evidently nonvanishing and holomorphic in a neighborhood of each $(e_i, 0)$.

What about the (flex) point at infinity $\mathcal{O} = [0 : 0 : 1]$? By Poincaré-Hopf, $g = 1 \implies \sum_{p \in E} \nu_p(\omega) = 2g - 2 = 0$, so that if $\nu_p(\omega) = 0$ for all $p \in E \backslash \mathcal{O}$, there can be no contribution from $\mathcal{O}$ either. $\square$

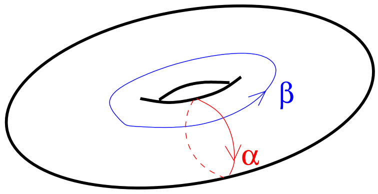
COROLLARY 18.1.3. $\Omega^1(E) \cong \mathbb{C}\langle\omega\rangle$. *That is, every holomorphic 1-form on E is a multiple of ω.*

PROOF. For any $\omega_0 \in \Omega^1(E)$, the discussion preceding Example 13.1.2 tells us $\frac{\omega_0}{\omega} \in \mathcal{K}(E)$. But since ω is nowhere vanishing, $\frac{\omega_0}{\omega}$ is actually a *holomorphic* function. Now use Liouville's theorem ($\mathcal{O}(E) \cong \mathbb{C}$). $\square$

Amongst the standard *topological* invariants of a 1-manifold M is its first homology group. An *ad hoc* definition is

$$H_1(M, \mathbb{Z}) := \frac{\left\{\begin{array}{c}\text{free abelian group generated by}\\ \text{closed piecewise-}C^\infty\text{ paths on } M\end{array}\right\}}{\left\{\begin{array}{c}\text{subgroup generated by}\\ \text{boundaries of finitely triangulable regions}\end{array}\right\}},$$

or simply "cycles modulo boundaries". From the picture



it isn't hard to convince yourself that

$$H_1(E, \mathbb{Z}) \cong \mathbb{Z}\langle \alpha, \beta \rangle .$$

That is, for any closed C^∞ path $\gamma \subset E$, there exists a closed set $\Gamma \subset E$ (with boundary $\partial\Gamma$) such that

$$\gamma = m\alpha + n\beta + \partial\Gamma.$$

The integers m, n are uniquely determined by γ. One then has

$$\begin{aligned}\int_\gamma \omega &= \int_{\partial\Gamma} \omega + m \int_\alpha \omega + n \int_\beta \omega \\ &= m \int_\alpha \omega + n \int_\beta \omega\end{aligned}$$

by Cauchy's theorem (Prop. 13.1.5). The values of the integrals $\int_\gamma \omega$ over cycles are called the *periods* of ω, and we define the *period lattice*

$$\Lambda_E := \mathbb{Z}\left\langle \int_\alpha \omega, \int_\beta \omega \right\rangle \subset \mathbb{C}.$$

This furnishes an invariant of the complex structure[1] on E which, unlike the topological invariant, actually distinguishes elliptic curves which are non-isomorphic as complex manifolds (or algebraic curves).

REMARK 18.1.4. Given a lattice of the form $\mathbb{Z}\langle \lambda_1, \lambda_2 \rangle =: \Lambda \subset \mathbb{C}$, we have a Weierstrass $\mathcal{P}$-map

$$\mathbb{C}/\Lambda \xrightarrow{\mathcal{P}} \mathbb{P}^2$$

$$u \longmapsto [1 : \wp(u) : \wp'(u)]$$

whose image (by Exercise 5 of Chapter 7) is a Weierstrass cubic! Define

$$E_\Lambda := \mathcal{P}(\mathbb{C}/\Lambda),$$

which we henceforth consider to be the range of the map $\mathcal{P}$. Obviously it is of interest to find out whether *all* Weierstrass cubics arise in this fashion (as E_Λ's).

[1]algebraic geometers call this a *transcendental* invariant

Before moving on we should note that $\mathcal{P}$ is injective. Its composition with (the x-coordinate projection) $x : E_\Lambda \to \mathbb{P}^1$ has degree 2 since $\wp$ has a unique pole on $\mathbb{C}/\Lambda$ (at 0), which is a double pole. But since mapping degrees of Riemann surfaces multiply under composition, and the degree of x itself is 2, that of $\mathcal{P} : \mathbb{C}/\Lambda \to E_\Lambda$ must be 1.

18.2. The Abel map

Let

$$E = \overline{\{y^2 = \underbrace{4x^3 - g_2 x - g_3}_{Q(x)}\}} \subset \mathbb{P}^2$$

be a Weierstrass cubic with $\omega = \left.\frac{dx}{y}\right|_E \in \Omega^1(E)$. Integrating this gives a (holomorphic) map of Riemann surfaces

$$\begin{aligned} u : E &\longrightarrow \mathbb{C}/\Lambda_E \\ p &\longmapsto \int_{\mathcal{O}}^{p} \omega \end{aligned}$$

where the integration is over any C^∞ path from 0 to p. This *Abel map* is well-defined: if γ', γ'' are two such paths, then their difference is closed and so

$$\gamma' - \gamma'' = \partial\Gamma + m\alpha + n\beta.$$

Integrating, we have

$$\int_{\gamma'} \omega - \int_{\gamma''} \omega \;=\; m \int_\alpha \omega + n \int_\beta \omega \;\in\; \Lambda_E.$$

REMARK 18.2.1. It is now easy to see that Λ_E has rank 2 (something we haven't yet addressed). Otherwise, $\mathbb{C}/\Lambda_E$ is isomorphic to $\mathbb{C}$ or[2] $\mathbb{C}^*$, both of which are noncompact. Since E is compact and u is continuous, its image must be a compact submanifold of $\mathbb{C}/\Lambda_E$. If the latter is noncompact then the image is therefore a point, meaning u is constant, clearly false since $\omega \neq 0$.

A "baby" version of Abel's theorem for elliptic curves[3] is then:

THEOREM 18.2.2. *The Abel map is injective.*

[2] $\Lambda_E \cong \mathbb{Z}$ (rank 1) $\implies$ $\mathbb{C}/\Lambda_E \cong \mathbb{C}/2\pi i\mathbb{Z} \overset{\cong}{\to} \mathbb{C}^*$ (by taking exp).

[3] The real theorem, which we will deal with later, has to be phrased in terms of divisors.

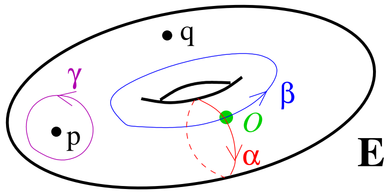
PROOF. (Sketch) Suppose $u(p) \equiv u(q) \bmod \Lambda_E$ for $p \neq q$ points of E; then

$$\int_q^p \omega = \int_{\mathcal{O}}^p \omega - \int_{\mathcal{O}}^q \omega = u(p) - u(q) \in \Lambda_E.$$

Modifying the path from q to p by $m\alpha + n\beta$ (for some $m, n \in \mathbb{Z}$), we get

$$\int_p^q \omega = 0.$$

Dirichlet's existence theorem (which we won't prove) guarantees the existence of $\eta_0 \in \mathcal{K}^1(E)$ with only simple poles, only at p and q, with $Res_p(\eta_0) = -Res_q(\eta_0) = 1$. This is true for any two (distinct) points p and q, and has nothing to do with our assumption (that $u(p) = u(q)$). Now referring to the picture



we have

$$H_1\left(E \backslash (\{p\} \cup \{q\}), \mathbb{Z}\right) \cong \mathbb{Z}\left\langle \alpha, \beta, \gamma \right\rangle \tag{18.2.1}$$

where

$$\int_\gamma \eta_0 = 2\pi i.$$

Next, "normalize" η_0, putting

$$\eta := \eta_0 - \left(\frac{\int_\alpha \eta_0}{\int_\alpha \omega} \right) \omega,$$

which has the same residues as η_0. Observe that

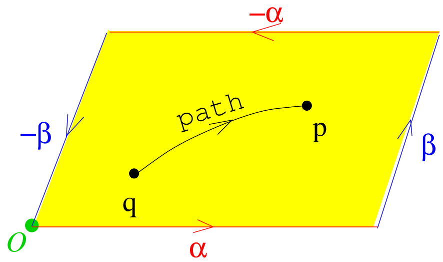
$$\int_\gamma \eta = 2\pi i,$$

while

$$\int_\alpha \eta = 0.$$

Cutting open the above figure along α and β yields the fundamental domain $\mathfrak{F}$ (the yellow region):[4]

[4]In order to accomodate the path from q to p, it may be necessary to "dilate" $\mathfrak{F}$ by an integer factor M. (You can think of this as the fundamental domain of a



On the interior of $\mathfrak{F}$, $\mathfrak{U} := \int_{\mathcal{O}}^{*} \omega$ gives a holomorphic function which is continuous on the boundary. Now

$$0 = \int_{p}^{q} \omega = \mathfrak{U}(p) - \mathfrak{U}(q)$$

which by the Residue theorem

$$= \frac{1}{2\pi i} \int_{\partial \mathfrak{F}} \mathfrak{U} \cdot \eta.$$

Noting that $\int_\alpha \omega$ (resp. $\int_\beta \omega$) is the change in $\mathfrak{U}$ from "$-\beta$" to "β" (resp. "$-\alpha$" to "α"), this

$$= \frac{1}{2\pi i} \left\{ \int_\beta \eta \int_\alpha \omega - \int_\alpha \eta \int_\beta \omega \right\}$$

$$= \frac{1}{2\pi i} \left(\int_\beta \eta \right) \left(\int_\alpha \omega \right),$$

where $\int_\alpha \omega \neq 0$. Hence,

$$\int_\beta \eta = 0.$$

By (18.2.1), any closed path on $E \backslash (\{p\} \cup \{q\})$ is, up to boundaries, of the form $n\alpha + m\beta + \ell\gamma$; and so the integral of η over such a path is $\ell \int_\gamma \omega = 2\pi i \ell$. Consequently,

$$F := \exp\left(\int_{\mathcal{O}}^{*} \eta \right)$$

is a well-defined function on E which is holomorphic off $\{p\} \cup \{q\}$. Let z (resp. w) be a local coordinate about p (resp. q) with $z(p) = 0$ (resp. $w(q) = 0$). We know that the leading term of η at p is $\frac{dz}{z}$, and at q is $-\frac{dw}{w}$. This makes F locally at p (resp. q) the product of a holomorphic function by $e^{\int \frac{dz}{z}} = e^{\log z} = z$ (resp. $e^{-\int \frac{dw}{w}} = \frac{1}{w}$), so that

finite unbranched covering of E.) This doesn't really affect the proof, except for replacing $\partial\mathfrak{F}$ by $M\alpha + M\beta - M\alpha - M\beta$.

F is meromorphic on E with divisor

$$(F) = [p] - [q].$$

Therefore $\deg(F) = 1$, making $F : E \to \mathbb{P}^1$ an isomorphism, which is impossible.

We conclude from this contradiction that p and q cannot have been distinct. $\square$

Abel's theorem is usually paired with something called "Jacobi inversion", the baby version of which is:

PROPOSITION 18.2.3. *The Abel map u is surjective (and thus an isomorphism).*

PROOF. This is trivial: u is a closed mapping (since continuous), and an open mapping (since holomorphic and not constant). The image is therefore open and closed in $\mathbb{C}/\Lambda_E$; since the latter is connected, we're done. $\square$

Essentially all of the foregoing (with the exception of Remark 18.1.4) works for any nonsingular cubic. There is a unique holomorphic 1-form up to scale; it vanishes nowhere; and integrating it from a base point gives an isomorphism from the cubic to a complex 1-torus. This follows from the last 2 sections by applying the projective transformation of Chapter 17 to put the cubic in Weierstrass form (which has just been slightly more convenient for writing down ω). For the next section, however, the Weierstrass form will be crucial.

18.3. Abel inverts Weierstrass

We now make the BIG CLAIM that

(18.3.1)

a Weierstrass cubic is always the image E_Λ of a Weierstrass $\mathcal{P}$-map

(cf. Remark 18.1.4), and we have

$$u \circ \mathcal{P} = \mathrm{id}_{\mathbb{C}/\Lambda} \tag{18.3.2}$$

and

$$\mathcal{P} \circ u = \mathrm{id}_E. \tag{18.3.3}$$

First we study the case where E is (by assumption) the image of a $\mathcal{P}$-map.

PROPOSITION 18.3.1. *Let* $\Lambda = \mathbb{Z}\langle\lambda_1, \lambda_2\rangle \subset \mathbb{C}$ *be a lattice. The composition*

$$\mathbb{C}/\Lambda \underset{\mathcal{P}}{\overset{(\cong)}{\longrightarrow}} E_\Lambda \underset{u}{\overset{(\cong)}{\longrightarrow}} \mathbb{C}/\Lambda_{E_\Lambda}$$

is the identity.

PROOF. Obviously part of the claim is that

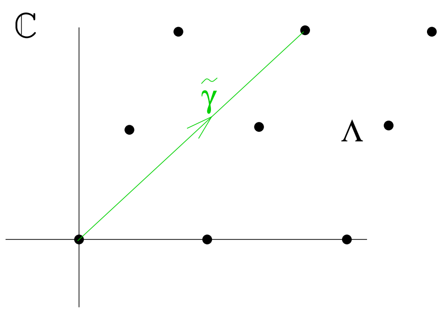
$$\Lambda_{E_\Lambda} \left(:= \left\{ \int_\gamma \omega \,\middle|\, \gamma \in H_1(E_\Lambda, \mathbb{Z}) \right\} \right) = \Lambda. \tag{18.3.4}$$

For E_Λ, not a lot is lost by working in affine (x, y) coordinates, since there is only $\mathcal{O}$ at ∞ and we know that corresponds to $u \equiv 0$ on the complex 1-tori. (Note also that "u" is used both as the Abel map and as the coordinate on $\mathbb{C}$; which one will be clear from the context.)

Since $\mathcal{P}(u) = (\wp(u), \wp'(u))$,

$$\mathcal{P}^*\omega = \mathcal{P}^*\left(\frac{dx}{y}\bigg|_{E_\Lambda} \right) = \frac{d(\wp(u))}{\wp'(u)} = \frac{\wp'(u)du}{\wp'(u)} = du.$$

Moreover, that $\mathcal{P}$ is an isomorphism means any cycle γ on E_Λ is the image of some $\tilde{\gamma} \in H_1(\mathbb{C}/\Lambda, \mathbb{Z})$



so that

$$\int_\gamma \omega = \int_{\mathcal{P}_*(\tilde{\gamma})} \omega = \int_{\tilde{\gamma}} \mathcal{P}^*\omega = \int_{\tilde{\gamma}} du$$

gives a bijection between Λ_{E_Λ} and Λ, hence (18.3.1). So then taking $u_0 \in \mathbb{C}/\Lambda$,

$$u(\mathcal{P}(u_0)) = \int_{\mathcal{O}}^{\mathcal{P}(u_0)} \frac{dx}{y} = \int_{\mathcal{P}(0)}^{\mathcal{P}(u_0)} \omega = \int_0^u \mathcal{P}^*\omega = \int_0^{u_0} du = u_0$$

proves the Proposition. □

Now let E be any Weierstrass cubic.

PROPOSITION 18.3.2. *The composition*

$$E \underset{u}{\overset{(\cong)}{\longrightarrow}} \mathbb{C}/\Lambda_E \underset{\mathcal{P}}{\overset{(\cong)}{\longrightarrow}} E_{\Lambda_E} \subset \mathbb{P}^2$$

is the identity. In fact,

$$E = E_{\Lambda_E} \tag{18.3.5}$$

exactly as subsets of $\mathbb{P}^2$.

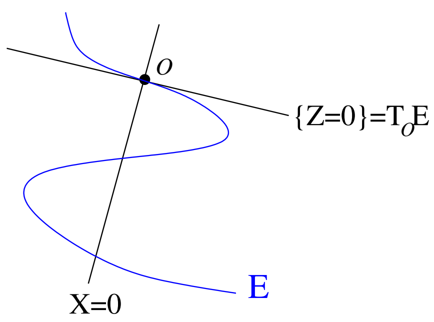
PROOF. Here is what the composition looks like, where $(x, y) \in E$:

$$(x,y) \overset{u}{\longmapsto} \int_\infty^x \frac{dx}{\pm\sqrt{Q(x)}} \overset{\mathcal{P}}{\longmapsto} \left(\wp\left(\int_\infty^x \frac{dx}{\pm\sqrt{Q(x)}}\right), \wp'\left(\int_\infty^x \frac{dx}{\pm\sqrt{Q(x)}}\right)\right),$$

where the $\pm$ is determined by y. We must show that the right-hand side recovers (x, y), or equivalently that the inverse $(x(u), y(u)) : \mathbb{C}/\Lambda_E \to E$ of the Abel map u identifies with $(\wp(u), \wp'(u))$.

Let's start with x, and compare the elliptic functions $x(u)$ and $\mathcal{P}(u)$ on $\mathbb{C}/\Lambda_E$. First I claim that both have double poles at $u = 0$: you already know that $\wp(u) = \frac{1}{u^2} +$ higher-order terms. For x, it suffices to check this on E, using

$$x = \left.\frac{X}{Z}\right|_E \in \mathcal{K}(E)^*$$



. . . which is easy:

$$\nu_{\mathcal{O}}(x) = (E \cdot \{X = 0\})_{\mathcal{O}} - (E \cdot \{Z = 0\})_{\mathcal{O}} = 1 - 3 = -2.$$

Now $x(u) = \frac{A}{u^2} + \text{h.o.t.}$, with $A =$

$$\lim_{u\to 0} x(u)\cdot u^2 = \left(\lim_{x\to\infty}\sqrt{x}\cdot u(x)\right)^2 = \left(\lim_{x\to\infty}\frac{\int_\infty^x \frac{dw}{\sqrt{Q(w)}}}{1/\sqrt{x}}\right)^2$$

$$= \left(\lim_{x\to\infty}\frac{1/\sqrt{Q(x)}}{-1/2x^{\frac{3}{2}}}\right)^2 = \lim_{x\to\infty}\frac{4x^3}{Q(x)} = 1.$$

Define an involution

$$\jmath : E \to E$$

by

$$(x, y) \mapsto (x, -y);$$

this fixes $\mathcal{O}$. For $p \in E$,

$$u(\jmath(p)) = \int_{\mathcal{O}=\jmath(\mathcal{O})}^{\jmath(p)} \frac{dx}{y} = \int_{\mathcal{O}}^{p} \jmath^* \frac{dx}{y} = -\int_{\mathcal{O}}^{p} \frac{dx}{y} = -u(p),$$

and so

$$x(-u) = x(u)\ , \quad y(-u) = -y(u).$$

All told, we now have that $x(u)$ and $\wp(u)$ are both even Λ-periodic functions locally of the form $\frac{1}{u^2} + \text{h.o.t.}$, and so their difference has no poles and must (by Liouville) be constant: $x - \wp = c$.

Next, differentiating $u = \int_\infty^x \frac{dx}{y}$ gives $\frac{du}{dx} = \frac{1}{y}$, or

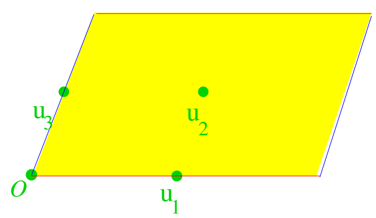
$$x'(u) = \frac{dx}{du} = y(u);$$

and then

$$0 = \frac{d}{du}(c) = x'(u) - \wp'(u) = y(u) - \wp'(u).$$

All that is left is to check that $c = 0$.

The fixed points of the involution $u \mapsto -u$ are the *2-torsion points*, i.e. those $u \in \mathbb{C}/\Lambda_E$ with $2u \equiv 0$

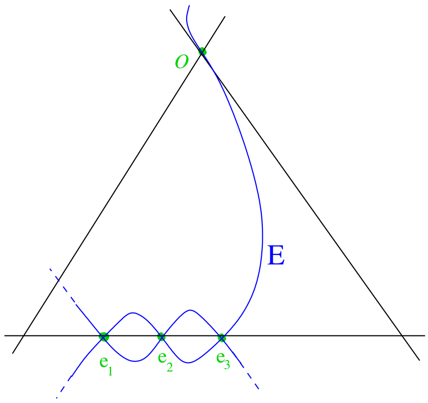


since we must have $u \equiv -u \bmod \Lambda_E$. These are, of course, the images (by u) of the fixed points of $\jmath$ in E, since $u \circ \jmath = -u$. They also must map (by $\mathcal{P}$) to the fixed points of $(x, y) \mapsto (x, -y)$ in E_{Λ_E}, since

$(\wp(-u), \wp'(-u)) = (\wp(u), -\wp'(u))$. Writing

$$y^2 = 4x^3 - g_2x - g_3 = 4(x - e_1)(x - e_2)(x - e_3)$$

for the equation of E,



$$x\,(2\text{-torsion points}) = e_1, e_2, e_3, \infty.$$

Similarly, if $E_{\Lambda_E} = \{y^2 = 4x^3 - \tilde{g}_2x - \tilde{g}_3 = 4(x - \tilde{e}_1)(x - \tilde{e}_2)(x - \tilde{e}_3)\}$ then

$$\wp(2\text{-torsion points}) = \tilde{e}_1, \tilde{e}_2, \tilde{e}_3, \infty;$$

and clearly

$$e_1 + e_2 + e_3 = \tilde{e}_1 + \tilde{e}_2 + \tilde{e}_3 = 0.$$

Since $\wp(u) = x(u) + c$,

$$\wp(u_1) + \wp(u_2) + \wp(u_3) = x(u_1) + x(u_2) + x(u_3) + 3c$$

which becomes

$$0 = 0 + 3c$$

so $c = 0$.

We conclude that $\wp(u(x, y)) = x$ and $\wp'(u(x, y)) = y$. $\square$

Exercises

(1) [Adapted from Silverman-Tate.]

Let $0 < \beta \leq \alpha$, and consider the ellipse E defined by

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1.$$

Show that the arc-length of E is given by the integral

$$4\alpha \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2 \theta} d\theta$$

for an appropriate choice of constant k depending on α and β. Then prove that this is equal to

$$4\alpha \int_0^1 \frac{1-k^2t^2}{\sqrt{(1-t^2)(1-k^2t^2)}}dt.$$

hence the problem of determining arc-length comes down to evaluating the integral

$$\int_0^1 \frac{1-k^2t^2}{u}dt$$

"on" the elliptic curve $u^2 = (1-t^2)(1-k^2t^2)$.

Remark: this curve obviously isn't cubic but is still of genus 1 because of the singularity at infinity. You can think of this as what you get if instead of writing an elliptic curve as $2:1$ over $\mathbb{P}^1$ with branching at (i.e. order 2 ramification over) 3 finite points plus infinity, you take the branching to be over 4 finite points. So it's very close to a Weierstrass cubic, even though it's quartic.

The integral in this problem is of course related to the Abel map/abelian integral above, and is meant to demonstrate why such integrals (and hence these curves) are called "elliptic".

CHAPTER 19

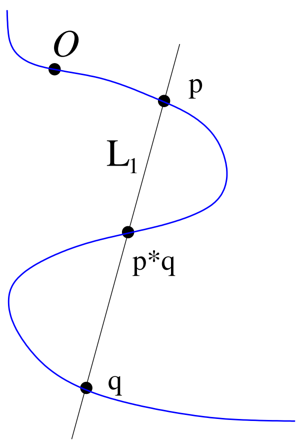
Group law on the nonsingular cubic

It is now high time for the "smooth" version of Chapter 16: group laws and addition theorems for elliptic curves. We start by introducing an algebro-geometrically defined binary operation on the points of a cubic curve, and prove it coincides with addition on $\mathbb{C}/\Lambda_E$ under the Abel isomorphism. This gives one proof that the operation defines an abelian group, and we give another more natural one as well. From the fact that it makes Abel's map into a homomorphism we will then derive functional equations for elliptic integrals.

19.1. Definition of the group law

Let $E \subset \mathbb{P}^2$ be a nonsingular cubic, which we shall not require to be in Weierstrass form, and fix a flex $\mathcal{O}$. Let p and q be points of E.

Step 1: Draw the line L_1 through p and q:

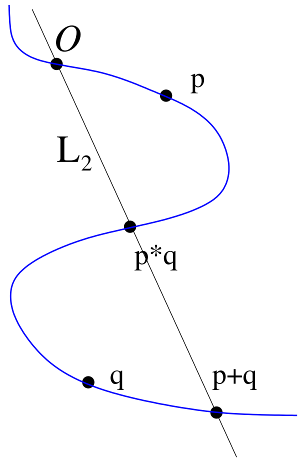


By Bezout, there is a third intersection point, which we shall denote $p * q$, so that

$$L_1 \cdot E = [p] + [q] + [p * q].$$

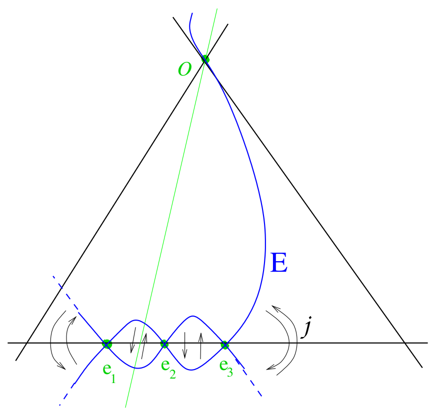
Note that the three points need not be distinct — any two, or all three, may coincide. This has the usual interpretation: double-intersection means L_1 is tangent to E at that point; triple-intersection means you have a flex.

Step 2: Draw the line L_2 through $\mathcal{O}$ and $p * q$:



The third intersection point is denoted $p + q$, with the same interpretations as above.

In the special case where E is in Weierstrass form, L_2 is a vertical line $\{X = x_0 Z\}$. So the map sending $p * q$ to $p + q$ is just the involution $\jmath : E \to E$ taking $(x, y) \mapsto (x, -y)$:



We have therefore constructed a binary operation

$$E \times E \to E$$

$$(p, q) \mapsto p + q$$

on the (set consisting of the) points of E. It is not yet clear that this defines a group. It *is* clear that it is commutative, so that *if* it defines a group, then that group is abelian.

19.2. Relation to the group structure on the 1-torus

Let $T : \mathbb{P}^2 \to \mathbb{P}^2$ be the projective transformation putting E into Weierstrass form (and taking $\mathcal{O}$ to $\mathcal{O}' := [0 : 0 : 1]$). Denote the binary operation defined on points of $E' := T(E)$ (via the method just described, with $\mathcal{O}'$ replacing $\mathcal{O}$) by $+'$. Since projectivities preserve lines, intersection multiplicities, and so forth, it is clear that $T(p) +' T(q) = T(p{+}q)$. So to show that "$+$" defines a group law for arbitrary E, it suffices to check this for Weierstrass cubics.

Hence we may assume E is in Weierstrass form (and $\mathcal{O} = [0 : 0 : 1]$). Take the Abel map

$$u : E \to \mathbb{C}/\Lambda_E$$

to be as in Chapter 18, with inverse $\mathcal{P}$. We know that "addition mod Λ_E" defines a group law on $\mathbb{C}/\Lambda_E$. The next result implies not only that our operation on E defines a group law, it says that u is an isomorphism of groups.

THEOREM 19.2.1. *The Abel map respects binary operations. That is, the equivalent formulas*

$$\mathcal{P}(u_1) + \mathcal{P}(u_2) = \mathcal{P}(u_1 + u_2) \tag{19.2.1}$$

$$u(p) + u(q) \equiv u(p + q) \ \ \textit{mod}\ \Lambda_E \tag{19.2.2}$$

hold.

PROOF. We will prove (19.2.2), in a manner reminiscent of the proof of Theorem 18.2.1. Writing F_{L_i} for the degree-1 homogeneous polynomial defining L_i, consider the meromorphic function

$$f := \left.\frac{F_{L_2}}{F_{L_1}}\right|_E \in \mathcal{K}(E)^*.$$

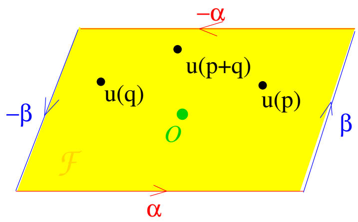
Reading its divisor off from the intersection points of the $\{L_i\}$ and E,

$$\begin{aligned}(f) \ &= \ [p+q] + [p*q] + [\mathcal{O}] - ([p] + [q] + [p*q]) \\ &= \ [p+q] - [p] - [q] + [\mathcal{O}].\end{aligned}$$

So for its pullback $\mathcal{P}^* f = f \circ \mathcal{P}$ to $\mathbb{C}/\Lambda_E$,

$$(\mathcal{P}^* f) \ = \ [u(p+q)] - [u(p)] - [u(q)] + [0].$$

Cut open $\mathbb{C}/\Lambda_E$ and put $\mathfrak{U} := \int_0^* du$ on the resulting fundamental region $\mathfrak{F}$:



Write $\mathfrak{U}(p)$ for $\mathfrak{U}(u(p))$ (and so on) for simplicity. By the residue theorem,

$$\begin{aligned} &\mathfrak{U}(p+q) - \mathfrak{U}(p) - \mathfrak{U}(q)\,(+0) \\ &= \frac{1}{2\pi i}\oint_{\partial\mathfrak{F}} \mathfrak{U}\cdot\frac{d(\mathcal{P}^*f)}{\mathcal{P}^*f} \\ &= \left\{-\frac{1}{2\pi i}\int_\alpha \mathrm{dlog}(\mathcal{P}^*f)\right\}\int_\beta du \;+\; \left\{\frac{1}{2\pi i}\int_\beta \mathrm{dlog}(\mathcal{P}^*f)\right\}\int_\alpha du. \end{aligned}$$

Since both terms in braces are integers, the whole thing belongs to

$$\mathbb{Z}\left\langle \int_\alpha du,\, \int_\beta du\right\rangle = \Lambda_E.$$

Since $\mathfrak{U}(p)$, $\mathfrak{U}(q)$, $\mathfrak{U}(p+q)$ are lifts to $\mathbb{C}$ of $u(p)$, $u(q)$, $u(p+q)$, going modulo Λ_E we see that $u(p+q) - u(p) - u(q) \equiv 0$. □

We will generalize this argument in the next chapter to get Abel's theorem for E.

REMARK 19.2.2. For an arbitrary smooth cubic C, one still has (up to scale) a unique $\omega \in \Omega^1(C)$, which gives rise to an Abel isomorphism $u : C \to \mathbb{C}/\Lambda_{(C,\omega)}$. (19.2.2) still holds in this setting by more or less the same proof; this avoids passing through Weierstrass form.

19.3. A more algebro-geometric approach

Returning to the setup of §19.1, let us suppose that the coefficients of the homogeneous cubic polynomial defining E belong to some subfield $k \subset \mathbb{C}$. We shall say *E is defined over k*. If $K \subset \mathbb{C}$ is a field extension of k (e.g. k itself, or $\mathbb{C}$), then we can consider the *K-points of E*

$$E(K) := \{[Z : X : Y] \in E \mid Z, X, Y \in K\}.$$

PROPOSITION 19.3.1. *$E(K)$ is closed under "+", and is consequently a subgroup of E.*

PROOF. If $p, q \in E(K)$ then the line $L_1 = L_{pq}$ is defined over K, and so can be parametrized $\mathbb{P}^1 \overset{\cong}{\to} L_{pq}$ over K. (This means that the formulas expressing $[Z : X : Y]$ as functions of $[T_0 : T_1] \in \mathbb{P}^1$ involve coefficients in K.) So the pullback of the homogeneous polynomial defining E is defined over K. Now this can be written $\sum_{j=0}^{3} \beta_j Z_0^{3-j} Z_1^j = \prod_{i=1}^{3}(Z_1 - \alpha_i Z_0)$, assuming without loss of generality that there are no Z_0 factors; and what we know is that the $\beta_j \in K$. This is no guarantee that the $\alpha_i \in K$. But in this case we know that two of them — say, α_1, α_2— correspond to p, q and so *must* belong to K. Consequently $\alpha_3 \in K$ as well, and its image point $p * q$ is also defined over K. Repeat the argument for L_2 and the claim follows. $\square$

REMARK 19.3.2. (a) In light of the above definition, the "correct" notation for the set of complex points of E, which we have heretofore denoted simply "E", is $E(\mathbb{C})$.

(b) As a $\mathbb{C}$-module, $E(\mathbb{C}) \cong \mathbb{C}/\Lambda_E$ has rank 1, but as an abelian group (i.e. "$\mathbb{Z}$-module"), its rank is *infinite*. (Consider a bunch of $\mathbb{Q}$-linearly independent complex numbers modulo Λ_E — there is no bound on the size of the $\mathbb{Q}$-vector space you can generate in this fashionl.) On the other hand, for the subgroup $E(\mathbb{Q})$, a famous theorem of MORDELL (1922) asserts that the rank (as an abelian group) is always finite. In problem (6) below, you will show computationally that the rank of $E(\mathbb{Q})$ in one example is at least 1.

Now the Abel map u is non-algebraic (i.e., transcendental); it should be seen as providing a link between the complex algebraic and the complex analytic. Such maps, which include multivariate abelian functions, modular and automorphic forms, are very important in arithmetic algebraic geometry. While they do not preserve the field of definition, they have nonetheless been essential to the study of things like class field theory, the proof of Fermat's last theorem, and so on. Still, in light of the subgroup structures $E(K) \subset E(\mathbb{C})$, it is a bit sloppy to prove that "+" is a group law in a manner that only works over $\mathbb{C}$.

So we will now give the "fully algebraic" approach to this proof, by checking

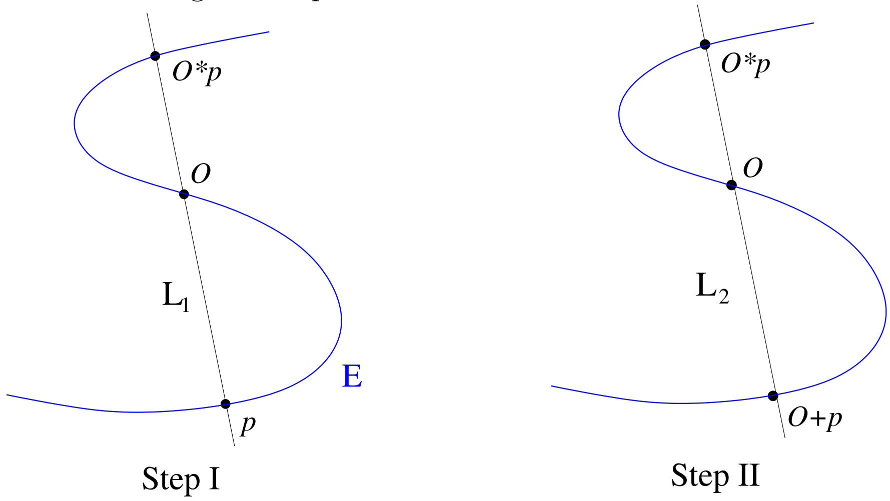
$$\mathcal{O} + p = p \quad (\forall\, p \in E), \tag{19.3.1}$$

$$(p * q) + (p + q) = \mathcal{O} \quad (\forall\, p, q \in E), \text{ and} \tag{19.3.2}$$

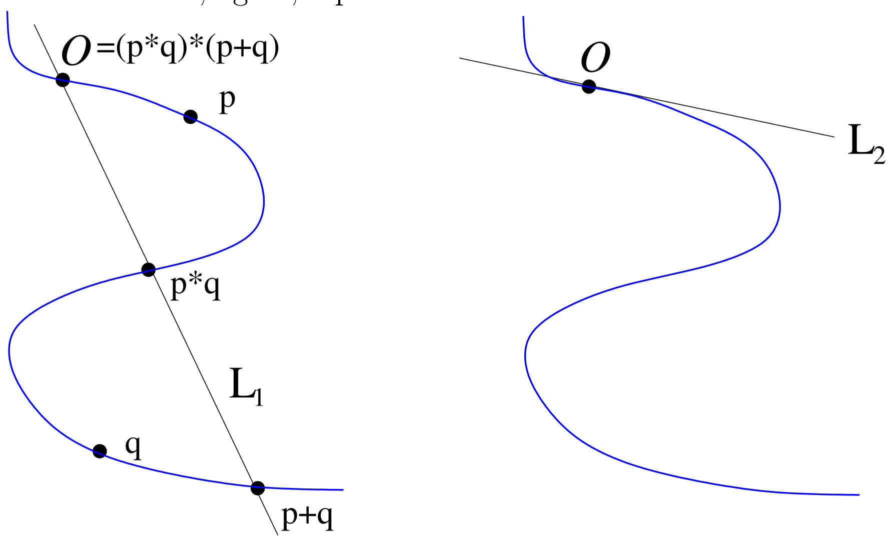
$$+ \text{ is associative.} \tag{19.3.3}$$

Notice that (19.3.2) implies $(p * \mathcal{O}) + (p + \mathcal{O}) = \mathcal{O}$, and thus that $p * \mathcal{O} = -p$, so we indeed have inverses. (19.3.1) and (19.3.3) are the other two group axioms.

To verify (19.3.1), we have the pictorial depiction of the two-step process for adding $\mathcal{O}$ and p:



in which $L_1 = L_{\mathcal{O}p}$ and $L_2 = L_{\mathcal{O},\mathcal{O}*p} = L_1$!! Consequently the third intersection point $\mathcal{O} + p$ of L_2 with E, is none other than p. (19.3.2) isn't much harder; again, a picture:



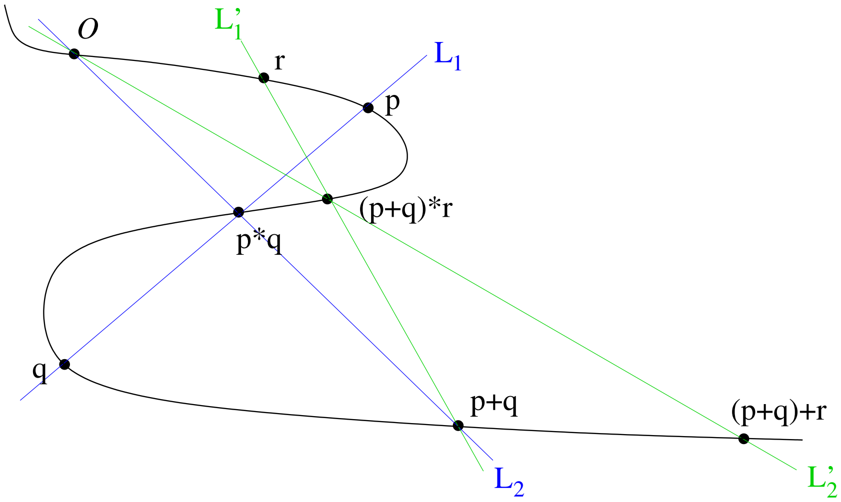
where $L_1 = L_{p*q,\,p+q}$. Since the line through $\mathcal{O}$ and $p * q$ intersects E in (and defines) $p + q$, it follows that L_1's third intersection point

$(p * q) * (p + q)$ with E is just $\mathcal{O}$. Then

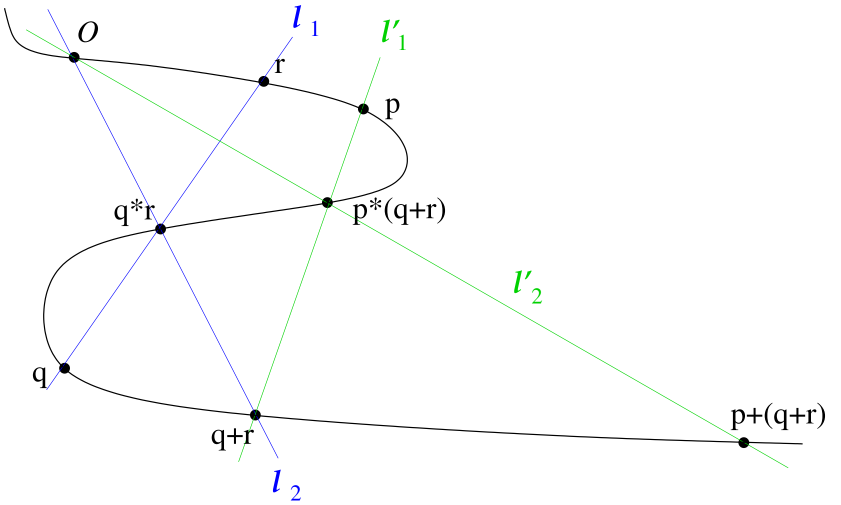
$$L_2 := L_{\mathcal{O},(p*q)*(p+q)} = L_{\mathcal{O},\mathcal{O}} = T_{\mathcal{O}}E,$$

and since $\mathcal{O}$ is a flex $((T_{\mathcal{O}}E \cdot E)_{\mathcal{O}} = 3)$, the third intersection point is again $\mathcal{O}$.

Finally we come to the associativity issue (19.3.3). Here I won't break the two steps up into two pictures. Instead, here is a depiction of $(p + q) + r$, where the blue lines compute $p + q$ and the green ones the addition of r to the result:



... and here is what $p + (q + r)$ looks like (blue lines for $q + r$, green for adding result to p):



We have to show $(p + q) + r = p + (q + r)$, or equivalently

$$(p + q) * r = p * (q + r).$$

Now look at the three cubics E, $C = L_1 \cup \ell_2 \cup L_1'$, and $D = \ell_1 \cup L_2 \cup \ell_1'$, with intersections

$$E \cdot C \;=\; [\mathcal{O}]+[p]+[q]+[r]+[p*q]+[p+q]+[q*r]+[q+r]+[(p+q)*r]$$

and

$$E \cdot D \;=\; [\mathcal{O}]+[p]+[q]+[r]+[p*q]+[p+q]+[q*r]+[q+r]+[p*(q+r)].$$

Argue à la §15.2: the ratio of the homogeneous polynomials defining C and D induces a meromorphic function on E with divisor

$$E \cdot C - E \cdot D \;=\; [(p+q)*r] - [p*(q+r)],$$

leading as usual to a contradiction *unless* these two points are the same.

19.4. Addition theorems

Now assume $E = \{y^2 = Q(x)\}$ (with $Q(x) = 4x^3 - g_2 x - g_3$) is in Weierstrass form; we would like to unwind the statements (19.2.1) and (19.2.2) that $\mathcal{P}$ and u are group homomorphisms (hence isomorphisms), to produce something more computationally explicit.

We do this first for the Weierstrass map. Writing $p = \mathcal{P}(u_1) = (\wp(u_1), \wp'(u_1))$, $q = \mathcal{P}(u_2) = (\wp(u_2), \wp'(u_2))$, we have $p * q$

$$= \jmath(p+q) = \jmath(\mathcal{P}(u_1)+\mathcal{P}(u_2)) = \jmath(\mathcal{P}(u_1+u_2)) = \jmath(\wp(u_1+u_2), \wp'(u_1+u_2))$$
$$= (\wp(u_1+u_2), -\wp'(u_1+u_2)).$$

Now p, q, and $p*q$ are collinear by construction — they lie on L_1 in the group law "process" for E. We may express this in projective coordinates by saying that

$$0 = \det \begin{pmatrix} 1 & 1 & 1 \\ \wp(u_1) & \wp(u_2) & \wp(u_1+u_2) \\ \wp'(u_1) & \wp'(u_2) & -\wp'(u_1+u_2) \end{pmatrix}.$$

This is the *first addition theorem*, and is the analogue for bi-periodic functions of the standard trigonometric angle-addition formulas. It really does express $\wp(u_1+u_2)$ in terms of $\wp(u_1)$ and $\wp(u_2)$, since $\wp'(\alpha) = \pm\sqrt{Q(\wp(\alpha))}$.

Let's actually compute the group law on E. Start by writing $y = ax+b$ for L_1 and substituting this into the equation of E to "intersect" them. This gives

$$0 = 4x^3 - g_2x - g_3 - (ax+b)^2 = 4(x - x(p))(x - x(q))(x - x(p+q))$$

since L_1 and E meet in p, q, $p * q$. (Note that $x(p+q) = x(p*q)$.)[1] From expanding these two expressions and comparing coefficients of x^2, one finds that

$$a^2 = 4(x(p) + x(q) + x(p+q));$$

and since a is the slope of L_1 it is obvious that

$$a = \frac{y(q) - y(p)}{x(q) - x(p)}.$$

Therefore we have

$$x(p+q) = \frac{1}{4}\left(\frac{y(q) - y(p)}{x(q) - x(p)}\right)^2 - x(p) - x(q).$$

Now $u(p) = \int_{\mathcal{O}}^{p} \frac{dx}{y} = \int_{\infty}^{x(p)} \frac{dx}{\sqrt{Q(x)}}$, similarly $u(q) = \int_{\infty}^{x(q)} \frac{dx}{\sqrt{Q(x)}}$ and $u(p+q) = \int_{\infty}^{x(p+q)} \frac{dx}{\sqrt{Q(x)}}$. Re-expressing

$$u(p) + u(q) \equiv u(p+q) \mod \Lambda_E$$

using all these formulas yields the *second addition theorem*:

$$\int_{\infty}^{x_1} \frac{dx}{\sqrt{Q(x)}} + \int_{\infty}^{x_2} \frac{dx}{\sqrt{Q(x)}} \underset{\Lambda_E}{\equiv} \int_{\infty}^{\left\{\frac{1}{4}\left(\frac{\sqrt{Q(x_2)} - \sqrt{Q(x_1)}}{x_2 - x_1}\right)^2 - x_2 - x_1\right\}} \frac{dx}{\sqrt{Q(x)}}$$

which is a nontrivial functional equation for the elliptic integral $\int_{\infty}^{*} \frac{dx}{\sqrt{Q(x)}}$.

* * *

The problems below (with the exception of the last one) take place on a nonsingular cubic $E := \{y^2 = 4x^3 - g_2x - g_3\} \subset \mathbb{P}^2$ *in Weierstrass form*, with base point $\mathcal{O} = [0:0:1]$, holomorphic form $\omega = \frac{dx}{y}|_E \in \Omega^1(E)$, and Abel map $u : E \to \mathbb{C}/\Lambda_E$, $u(p) = \int_{\mathcal{O}}^{p} \omega$ (recall this is an isomorphism), with inverse $\mathsf{P}(u) = [1 : \wp(u) : \wp'(u)]$. I have written

[1] the dictionary we have in mind is: $x(p) = \wp(u_1)$, $y(p) = \wp'(u_1)$; $x(q) = \wp(u_2)$, $y(q) = \wp'(u_2)$; $x(p+q) = \wp(u_1+u_2)$, $y(p+q) = \wp'(u_1+u_2)$.

everything in affine form, which you can convert to projective coordinates if needed.

Exercises

(1) Show that the 2-torsion points on $\mathbb{C}/\Lambda_E$ correspond to the x-intercepts $\{(e_i, 0)\}_{i=1}^3$ and the point $\mathcal{O}$.

(2) Assume $g_2, g_3 \in \mathbb{Q}$. Given $p, q \in E(\mathbb{Q})$ (i.e. the x, y coordinates are rational), give another proof that $p+q \in E(\mathbb{Q})$, using the addition theorems.

(3) *For this and the following three problems take* $g_2 = -4$, $g_3 = 0$. Consider the complex analytic automorphism $A : E \to E$ sending $(x, y) \mapsto (-x, iy)$. In Ch. 13 Exercise (1), you showed that $A^*\omega = i\omega$.

(a) Find A^*u (i.e. compute $u \circ A$).

(b) Prove that $i\Lambda_E = \Lambda_E$. (In fact, Λ_E is a "square" lattice — so this is a very special elliptic curve!)

(4) "Special case" of the 2nd addition theorem (or rather, what we did in §19.4 above doesn't exactly work in the case we'll do here, so you'll have to work it out from scratch): write $\wp(2u)$ in terms of $\wp(u)$, for E as in exercise (3), i.e. with equation $y^2 = 4x^3 + 4x$. [Hint: write $\wp'(u)$ and then the slope a of E at $(\wp(u), \wp'(u))$, in terms of $\wp(u)$. Write $y = ax + b$ for the line tangent to E at this point. Then factor $4x^3 + 4x - (ax+b)^2$ into linear factors (what are the roots?), multiply out both expressions, and compare coefficients of x^2.]

(5) Continuing the last problem, show that $(1, 2\sqrt{2})$ is a 4-torsion point. Use the "CM" recalled in exercise (3) to get three more 4-torsion points. Can you use the group law to find them all? (If not, why?)

(6) This problem also depends on (4). Consider a point P of E with rational x-coordinate $x_0 = \frac{p}{2^a q}$, where the fraction is written in lowest terms, a is an odd natural number and p and q are odd integers. Show that P is of infinite order (in the group). [Hint: write (x_0, y_0) for this point, and let $(x_1, y_1) := 2(x_0, y_0)$ under the group law; if $x_0 = \wp(u)$, then $x_1 = \wp(2u)$. So rewriting your

formula from (4) as a formula for x_1 in terms of x_0 and simplifying, show that x_1 is of the same form, but with larger a. Then suppose the starting point was an N-torsion point for some N and produce a contradiction via the pigeonhole principle.]

(7) This problem takes place on an arbitrary nonsingular cubic $E \subset \mathbb{P}^2$, with base point $\mathcal{O}$ = choice of flex in E, holomorphic form $\omega \in \Omega^1(E)$ with periods generating a lattice Λ_E, and Abel map $u : E \xrightarrow{\cong} \mathbb{C}/\Lambda_E$, $u(p) = \int_{\mathcal{O}}^{p} \omega$.
Prove that the 3-torsion points on $\mathbb{C}/\Lambda$ correspond (under u) to the flexes on E.

CHAPTER 20

Abel's theorem for elliptic curves

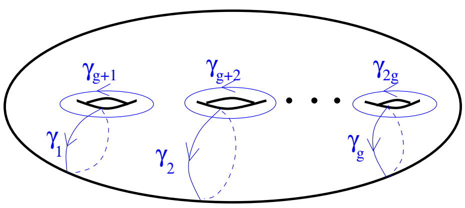
Given a divisor $D = \sum n_i[p_i]$ on an elliptic curve E, we can formally compute the sum in the group law, ending up with a single point on E. It seems of interest to ask if anything special is true if this point is the origin $\mathcal{O}$. In fact, assuming $\sum n_i = 0$, it will turn out that this is true precisely if D is the divisor of a meromorphic function on the curve. We begin by describing the statement of Abel's theorem for a curve of arbitrary genus (which does not have a group law), to place the statement for genus one in a broader context. Then we prove the genus-one case, introducing theta functions along the way.

20.1. The Jacobian of an algebraic curve

Let M be a Riemann surface of genus g. We will need to accept some facts in order to state Abel's theorem for M. (These will be returned to in later chapters, along with the proof of Abel.) It turns out that the space of holomorphic 1-forms has dimension g, whilst the abelian group of 1-cycles modulo boundaries (cf. §18.1 for definitions) has rank $2g$. In terms of bases,

$$H_1(M, \mathbb{Z}) \cong \mathbb{Z}\langle \gamma_1, \dots, \gamma_{2g} \rangle,$$
$$\Omega^1(M) \cong \mathbb{C}\langle \omega_1, \dots, \omega_g \rangle.$$

Remark 20.1.1. A visual "explanation" of the statement about homology groups may be the best one:



Chapter 3 of Griffiths's "Introduction to Algebraic Curves" gives one approach to computing the holomorphic forms of M, provided one believes that any Riemann surface is the normalization of an algebraic curve C in $\mathbb{P}^2$ with only ordinary double point (if any) singularities. (This statement relies on the existence of nonconstant meromorphic functions on M, which is nontrivial.) Since the genus g of M is $\frac{(d-1)(d-2)}{2} - \delta$ ($d = \deg(C)$, $\delta = \#$ of ODPs), it is enough to show that all meromorphic 1-forms are rational (cf. §24.1) and furthermore that *holomorphic* pullbacks of rational 1-forms from $\mathbb{P}^2$ span a space of dimension $\binom{d-1}{2} - \delta$. This is done in Griffiths, and will be discussed a little more in §24.2.

Just to get an idea of how this works, suppose $C = \{F(Z, X, Y) = 0\}$ is smooth of degree d, and recall S_3^m denotes degree-m homogeneous polynomials in 3 variables, with dimension $\binom{m+2}{2}$. If G is a homogeneous polynomial of degree n, write $g(x, y) = G(1, x, y)$ (and similarly $f(x, y) = F(1, x, y)$). Then the meromorphic 1-form on $\mathbb{P}^2$ which in affine coordinates takes the form $\frac{g \cdot dx}{f_y}$, restricts to a holomorphic 1-form on C precisely if[1] $n = d - 3$. (This is equivalent to saying $\deg(g) \leq d - 3$.) Hence,[2] $\Omega^1(C)$ has dimension $\binom{(d-3)+2}{2} = \binom{d-1}{2} = \frac{(d-1)(d-2)}{2}$.

Anyhow, let $\gamma_j \in H_1(M, \mathbb{Z})$ be a basis element; associated to it is a *period vector*

$$\pi_j := \begin{pmatrix} \int_{\gamma_j} \omega_1 \\ \vdots \\ \int_{\gamma_j} \omega_g \end{pmatrix} \in \mathbb{C}^g.$$

Together these form a $g \times 2g$ *period matrix* Π with $\mathbb{R}$-linearly independent columns. (This isn't obvious, and will be addressed in §24.2.) Hence their columns generate (over $\mathbb{Z}$) a $2g$-lattice $\Lambda_M \subset \mathbb{C}^g (\cong \mathbb{R}^{2g})$.

Recall that if V is a vector space (say, over $\mathbb{C}$) then the dual space is the space of linear functions $V^\vee := Hom(V, \mathbb{C})$.

DEFINITION 20.1.2. The *Jacobian* of M is the abelian group

$$J(M) := \frac{(\Omega^1(M))^\vee}{\text{image}\,\{H_1(M, \mathbb{Z})\}},$$

---

[1] the computation in Griffiths proving this is "ugly" but straightforward; Poincaré residues facilitate a conceptual and essentially 1-line proof (but at the cost of more sophisticated machinery).

[2] putting off to §24.2 that this formula encompasses *all* rational holomorphic forms

where the denominator means the linear functions on $\Omega^1(M)$ obtained by integrating $\omega \in \Omega^1(M)$ over 1-cycles. Evaluation of linear functions against the basis $\{\omega_1, \ldots, \omega_g\}$ induces an isomorphism

$$J(M) \xrightarrow{\cong} \frac{\mathbb{C}^g}{\Lambda_M} \,;$$

that is, the Jacobian is a complex g-torus.

LEMMA 20.1.3. *Any morphism of complex manifolds $\varphi : \mathbb{P}^1 \to \mathbb{C}^g/\Lambda_M$ is constant.*

PROOF. $\mathbb{C}^g/\Lambda_M$ has g independent holomorphic 1-forms: $du_1, \ldots, du_g$ (where $u_1, \ldots, u_g$ are just the coordinates on $\mathbb{C}^g$). Since $\varphi^*(du_i) \in \Omega^1(\mathbb{P}^1)$ and $\Omega^1(\mathbb{P}^1) = \{0\}$, we have

$$0 \;=\; \varphi^*(du_i) \underset{\text{locally}}{=} d(\varphi^* u_i)$$

which implies $\varphi^* u_i = u_i \circ \varphi$ (well-defined only locally) is constant for each $i = 1, \ldots, g$. $\square$

20.2. The Abel-Jacobi map

When is $D \in \mathrm{Div}(M)$ of the form (f), for some nontrivial meromorphic function f on M? Since $\deg((f)) = 0$ for any $f \in \mathcal{K}(M)^*$, it is clear that D must be of degree 0 — i.e. in the kernel of

$$\deg : \; \mathrm{Div}(M) \longrightarrow \mathbb{Z}$$
$$\sum n_i[p_i] \longmapsto \sum n_i.$$

So consider a divisor D in

$$\mathrm{Div}^0(M) := \ker(\deg).$$

We may write

$$D = \sum_j ([q_j] - [r_j]) = \partial \underbrace{\left(\sum_j \overrightarrow{r_j q_j} \right)}_{=:\Gamma}$$

where "∂" means topological boundary and $\overrightarrow{r_j q_j}$ is a C^∞ path from r_j to q_j.

DEFINITION 20.2.1. The *Abel-Jacobi map*

$$AJ : \; \mathrm{Div}^0(M) \to J(M)$$

sends $D\,(=\partial\Gamma)$ to

$$\int_\Gamma = \sum_j \int_{r_j}^{q_j}$$

viewed as a functional on $\Omega^1(M)$.

The first question that arises is whether this is even well-defined, which in this case means *independent of the choice of "1-chain" (sum of paths)* Γ. To check this, let $\partial\Gamma = D = \partial\Gamma'$. Then $\partial(\Gamma - \Gamma') = 0$, meaning that $\Gamma - \Gamma'$ is a 1-cycle hence represents a class in $H_1(M, \mathbb{Z})$. Consequently,

$$\int_{\Gamma-\Gamma'} = \int_\Gamma - \int_{\Gamma'}$$

"belongs to the denominator of $J(M)$". It's even easier to check that AJ is a homomorphism (of abelian groups), which is left to you.

Now suppose $D = (f)$, and consider the family of divisors

$$D_t := f^{-1}(t) \in \mathrm{Div}(M),$$

parametrized by $t \in \mathbb{P}^1$. Then $D = D_0 - D_\infty$, and the composition

$$\mathbb{P}^1 \longrightarrow \mathrm{Div}^0(M) \xrightarrow{AJ} J(M)$$

sending

$$t \longmapsto D_0 - D_t \longmapsto AJ(D_0 - D_t)$$

is constant by Lemma 20.1.3, and *zero* at $t = 0$. Thus $AJ(D) = 0$, and we observe that

$$AJ \text{ factors through } \mathrm{Pic}^0(M) := \frac{\mathrm{Div}^0(M)}{(\mathcal{K}(M)^*)}$$

in a well-defined fashion. (The denominator means "divisors of meromorphic functions", and the statement is simply that AJ kills these.) $\mathrm{Pic}^0(M)$ is called the *Picard group* of M.

THEOREM 20.2.2. [ABEL'S THEOREM] $AJ:$ *$Pic^0(M) \to J(M)$ is an isomorphism.*

Leaving aside the surjectivity part of this, the meaning of the "well-definedness + injectivity" of this map is that for $D \in \mathrm{Div}^0(M)$,

$$\underset{(\text{for some } f \in \mathcal{K}(M)^*)}{D = (f)} \iff AJ(D) \equiv 0 \mod \Lambda_M,$$

completely answering the question we asked at the outset. (Note that the forward implication [$\implies$] is just well-definedness, which is completely proved. What is nontrivial is the injectivity/backward implication, since you actually have to find some f having D as its divisor.)

EXAMPLE 20.2.3. We consider what this means in the genus-1 case, i.e. for $M = E$ (the normalization of) an elliptic curve. Let $\omega \in \Omega^1(E)$ be nonzero, and consider $D \in \mathrm{Div}^0(E)$. We can write $D = \sum n_i[p_i]$ with $\sum n_i = 0$, and

$$AJ\left(\sum n_i[p_i]\right) \;=\; AJ\left(\sum n_i([p_i]-[\mathcal{O}])\right) \;=\; \sum n_i \int_{\mathcal{O}}^{p_i} \omega \;=\; \sum n_i u(p_i)$$

where u is the Abel map. Here the right-hand sum is taking place in $\mathbb{C}/\Lambda_E$, and we see right away that

$$AJ\left(\sum n_i[p_i]\right) = 0 \iff \sum n_i u(p_i) \underset{\Lambda_E}{\equiv} 0.$$

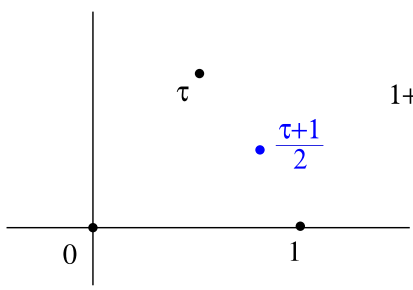
By Abel's theorem (on the left) and the fact that $u : (E,+) \to (\mathbb{C}/\Lambda_E,+)$ is a group-isomorphism (on the right), we have that

$$\text{(20.2.1)} \qquad \begin{array}{c} \sum n_i[p_i] = (f) \\ \text{for some } f \in \mathcal{K}(E)^* \end{array} \iff \begin{array}{c} \sum n_i \cdot p_i = \mathcal{O} \\ \text{in the group law on } E(\mathbb{C}). \end{array}$$

As above, the forward implication has been proved.

20.3. Direct proof of Abel for genus one

In this section we will deduce a result equivalent to the backward implication in (20.2.1), recasting it as an existence theorem for elliptic functions. For simplicity take $\Lambda = \mathbb{Z}\langle 1, \tau\rangle$, $\tau \in \mathfrak{H}$ (upper half-plane):



THEOREM 20.3.1. *Suppose* $\sum m_j = 0$ *and* $\sum m_j u_j \equiv 0$ *mod* Λ. *Then, writing* $D := \sum m_j[u_j] \in Div(\mathbb{C}/\Lambda)$, *there exists* $g \in \mathcal{K}(\mathbb{C}/\Lambda)$ *such that* $(g) = D$. *(You may think of* g *as a* Λ*-periodic meromorphic function on* $\mathbb{C}$.*)*

PROOF. Introduce the *theta function* (on $\mathbb{C}$)

$$\theta(u) := \sum_{n\in\mathbb{Z}} e^{\pi i\{n^2\tau+2nu\}}.$$

While it is *not* Λ-periodic, it has the properties
(a) $\theta(-u) = \theta(u)$
(b) $\theta(u+1) = \theta(u)$ [cf. Exercises]
(c) $\theta(u+\tau) = e^{-2\pi i\left(\frac{\tau}{2}+u\right)}\theta(u)$. To check this, write $\theta(\tau+u)$

$$= \sum_{n\in\mathbb{Z}} e^{\pi i\{n^2\tau+2nu+2n\tau\}} = \sum_{n\in\mathbb{Z}} e^{\pi i\{(n+1)^2\tau+2(n+1)u-\tau-2u\}}$$

which becomes, reindexing by $m = n+1$,

$$= e^{-\pi i\tau-2\pi iu} \sum_{m\in\mathbb{Z}} e^{\pi i(m^2\tau+2mu)}$$

as required.
(d) θ has a simple (order 1) zero at $\frac{\tau+1}{2}$ and no-where else in the fundamental domain bounded by vertices $0, 1, \tau, 1+\tau$.

Now consider

$$f(u) := \prod_j \theta\left(u - u_j + \frac{\tau+1}{2}\right)^{m_j};$$

clearly $f(u+1) = f(u)$ by property (b); but also (using property (c))

$$\begin{aligned}\frac{f(u+\tau)}{f(u)} &= \prod_j \left(\frac{\theta\left(\left\{u-u_j+\left(\frac{\tau+1}{2}\right)\right\}+\tau\right)}{\theta\left(u-u_j+\frac{\tau+1}{2}\right)}\right)^{m_j} \\ &= \prod_j \left(e^{-2\pi i\left(\tau+\frac{1}{2}+u-u_j\right)}\right)^{m_j} \\ &= e^{-2\pi i\left(\tau+\frac{1}{2}+u\right)\sum m_j}\cdot e^{2\pi i\sum m_j u_j}.\end{aligned}$$

By asssumption, $\sum m_j = 0$ and $\sum m_j u_j = M + N\tau$, so the last expression equals $e^{2\pi i N\tau}$. The function

$$g(u) := e^{-2\pi i N u} f(u)$$

will therefore satisfy $g(u+\tau) = g(u) = g(u+1)$. So it is Λ-periodic, and the definition of f together with property (d) makes it clear that $(g) = \sum m_j[u_j]$. $\square$

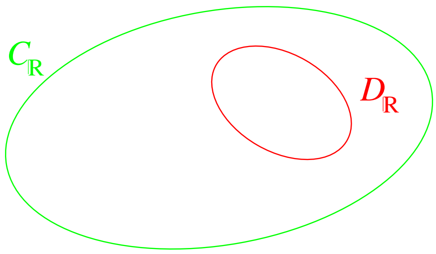
Exercises

(1) Verify property (b) for the theta function above (§20.3).

CHAPTER 21

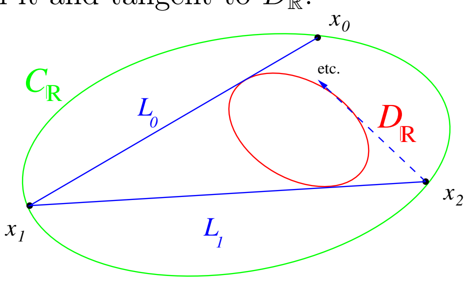
The Poncelet problem

First let's recall the most elementary statement of the "porism" from Chapter 1. One starts with two conics $C_{\mathbb{R}}$, $D_{\mathbb{R}}$ in $\mathbb{R}^2$, which for simplicity we can take to be two ellipses cut out by polynomials $f_C, f_D \in \mathcal{P}^2$ with real coefficients:



We asked in §1.3 whether there exists a closed polygon inscribed in $C_{\mathbb{R}}$ and circumscribed about $D_{\mathbb{R}}$. The result stated there, Theorem 1.3.1, said that if there is *one* then there is an *infinite family*. Our goal in this chapter is not just to flesh out the sketch of proof given there of this "porism", but to actually provide a way of deciding for *which* pairs there does exist a circuminscribed polygon.

A slight reformulation of the theorem is this: starting from some point x_0 on $C_{\mathbb{R}}$, draw a line segment tangent to $D_{\mathbb{R}}$, continue until it hits $C_{\mathbb{R}}$ again. Begin again at this new point, by drawing the other line segment through it and tangent to $D_{\mathbb{R}}$:



Iterating this construction, we may ask whether it ever closes up — i.e. returns to its starting point. (We will not care whether the path crosses itself.) What we will show is that the answer is independent of the choice of starting point x_0.

21.1. Proof of Theorem 1.3.1

This Theorem has nothing to do with C and D being ellipses, f_C and f_D being real polynomials, and so forth — it makes sense more generally for pairs of conics in the complex projective plane $\mathbb{P}^2$, and that is the context in which we view it for the proof. Namely, let $C = \{F_C(Z,X,Y) = 0\}$, $D = \{F_D(Z,X,Y) = 0\}$ be the conics cut out by homogeneous degree-2 polynomials F_C, $F_D \in S^2$. *If* the latter have coefficients in $\mathbb{R}$ (not essential for what follows), then the real points $C(\mathbb{R})$, $D(\mathbb{R})$ make sense and then $C_{\mathbb{R}}$, $D_{\mathbb{R}}$ above are just their intersections with affine space. Now, these affine real points need not meet (as in the above picture), but by Bezout C and D *must* meet in four points counted with multiplicity. We will carry out our proof *under the assumption that the multiplicities are all one*, i.e. C and D meet transversely and so $|C \cap D| = 4$.

Consider the incidence correspondence

$$\mathcal{E} := \{(x,L) \,|\, x \in L\} \subset C \times \check{D}$$

where $\check{D} \subset \check{\mathbb{P}}^2$ is the dual curve consisting of lines tangent to D (at any point). In §1.3 we defined pictorially two involutions $\iota_1 : \mathcal{E} \to \mathcal{E}$ and $\iota_2 : \mathcal{E} \to \mathcal{E}$. The idea is that each $L \in \check{D}$ meets C in two points (counted with multiplicity), and swapping those points gives ι_1; whereas each $x \in C$ is in two lines tangent to D ("counted with multiplicity"), and swapping those lines gives ι_2. Composing involutions gives $\jmath := \iota_2 \circ \iota_1$, which is no longer an involution and is the complex geometry analogue of the iteration described just above. If we pick a starting "point" $(x_0, L_0) \in \mathcal{E}$, then we are interested in whether

$$\jmath^n(x_0, L_0) = (x_0, L_0)$$

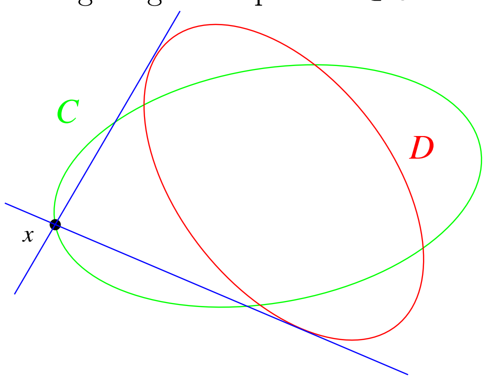
for some $n \in \mathbb{N}$.

The projection

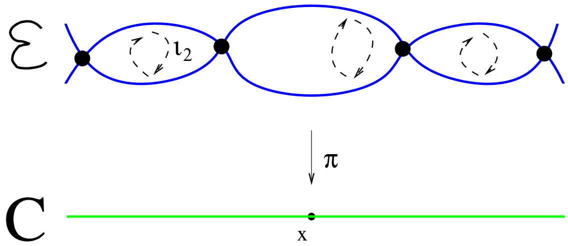
$$\begin{aligned} \pi : \mathcal{E} &\to C (\cong \mathbb{P}^1) \\ (x, L) &\mapsto x \end{aligned}$$

has

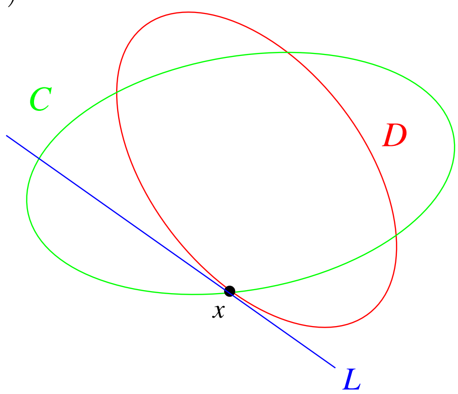
- mapping degree 2: there exist two lines L, L' tangent to D through a general point $x \in C$



- 4 ramification points (each of order two): namely, the points of $\mathcal{E}$ fixed by the involution ι_2



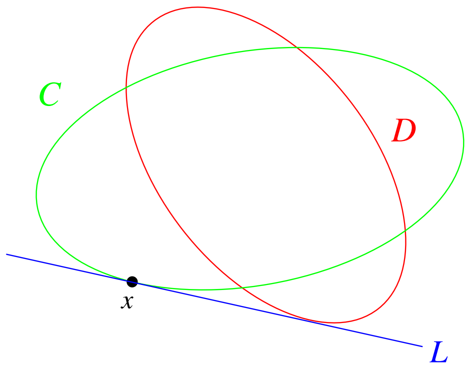
In particular, the ramification points of π identify with the points of $C \cap D$, since through each of these there is a *unique* tangent to D (rather than two):



By the Riemann-Hurwitz formula (for π), $\chi_{\mathcal{E}} = d \cdot \chi_C - r = 2 \cdot 2 - 4 = 0$. This implies $\mathcal{E}$ is elliptic, and so has an Abel map u mapping it isomorphically to a 1-torus $\mathbb{C}/\Lambda$ (where Λ depends on[1] $\mathcal{E}$ hence ultimately on C and D).

[1]to define the Abel map you also have to choose a holomorphic 1-form on $\mathcal{E}$; this affects the scaling of the lattice but not its isomorphism class

We could have carried out this same computation using $'\pi : \mathcal{E} \to \check{D}$ (sends $(x, L) \mapsto L$), whose ramification points (in $\mathcal{E}$) are the fixed points of ι_1 and hence identify with bitangents:



There are four of these since $\check{C}$ and $\check{D}$ are conics in $\check{\mathbb{P}}^2$ hence have $|\check{C} \cap \check{D}| = 4$.

Now consider an arbitrary involution I of $\mathbb{C}/\Lambda$, where the coordinate on $\mathbb{C}$ is denoted u. Any automorphism of $\mathbb{C}/\Lambda$ (in particular I) takes the form $u \mapsto au + b$ by Ch. 14 Exercise (5), and squaring this gives

$$u \longmapsto au + b \longmapsto a(au + b) + b = a^2u + b(a + 1).$$

If this is to be the identity, we must either have (i) $a = 1$ and $b \in \Lambda/2$, or (ii) $a = -1$ and $b \in \mathbb{C}$ arbitrary. Case (i) has no fixed points as it is a translation by a 2-torsion point.

By abuse of notation[2] we will think of $\iota_1, \iota_2, \jmath$ as automorphisms of $\mathbb{C}/\Lambda$. Since ι_1 and ι_2 are involutions of $\mathbb{C}/\Lambda$ with fixed points, they belong to case (ii):

$$\iota_1(u) \equiv b_1 - u\ , \quad \iota_2(u) \equiv b_2 - u \quad \pmod{\Lambda}.$$

Therefore

$$\jmath(u) = \iota_2(\iota_1(u)) \equiv b_2 - (b_1 - u) = u + \underbrace{(b_2 - b_1)}_{=:\beta},$$

i.e. $\jmath$ is a *translation* on $\mathbb{C}/\Lambda$.

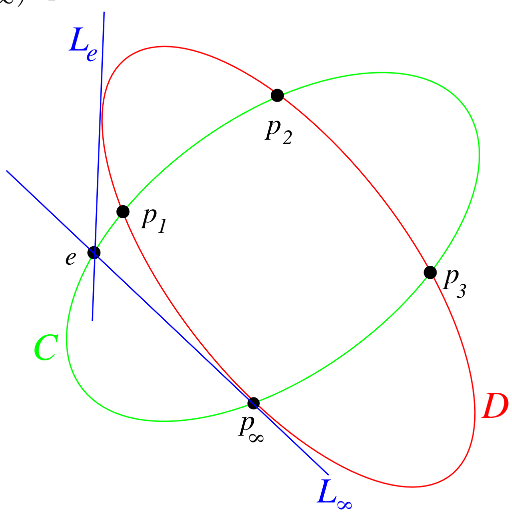
Write u_0 for the image of (x_0, L_0) under the Abel map. Clearly $\jmath^n(x_0, L_0) = (x_0, L_0)$ iff $\jmath^n(u_0) \equiv u_0 \pmod{\Lambda}$. But $\jmath^n(u_0) = u_0 + n\beta$, which $\equiv u_0$ iff $n\beta \equiv 0$, i.e. $n\beta \in \Lambda$. We conclude that the Poncelet

[2]strictly speaking one should write $u \circ \iota_1 \circ u^{-1}$ for the involution of $\mathbb{C}/\Lambda$ corresponding to ι_1 on $\mathcal{E}$

construction (starting from (x_0, L_0)) closes up at the n^{th} iteration if and only if β is n-torsion relative to the lattice. Since β depends only on $\jmath$, this has nothing to do with the choice of (x_0, L_0). Q.E.D.

21.2. Explicit solution of the Poncelet problem

The flexes are the preferred choices of origin for the group law on a cubic plane curve. On the incidence-correspondence elliptic curve $\mathcal{E}$, it turns out that the best choice for $\mathcal{O}$ is one of the fixed points of ι_2 (the four (x, L) with $x \in C \cap D$). Writing $C \cap D = \{p_1, p_2, p_3, p_\infty\}$, we set $\mathcal{O} := (p_\infty, L_\infty) \in \mathcal{E}$.



Here (e, L_e) is the first point in the "Poncelet iteration", i.e. $\jmath(\mathcal{O})$; clearly $\beta = u((e, L_e))$, with $u : \mathcal{E} \to \mathbb{C}/\Lambda$ the usual Abel isomorphism. The question of whether $\jmath^n$ is the identity can be restated in terms of the (unique) group law on $\mathcal{E}$ with origin $\mathcal{O}$:

(21.2.1) Is (e, L_e) an N-torsion point?

The approach we take to its solution in this section is work of Cayley as presented in the nice expository article [P. Griffiths and J. Harris, *On Cayley's explicit solution to Poncelet's porism*, L'Enseignement Math. 24 (1978), 31-40.].

A family of conics. Consider the collection of conics depending on $t \in \mathbb{P}^1$:

$$D_t := \left\{p \in \mathbb{P}^2 \mid tF_C(p) + F_D(p) = 0\right\},$$

where $D_\infty = C$ and $D_0 = D$. Each D_t passes through p_1, p_2, p_3, p_∞.

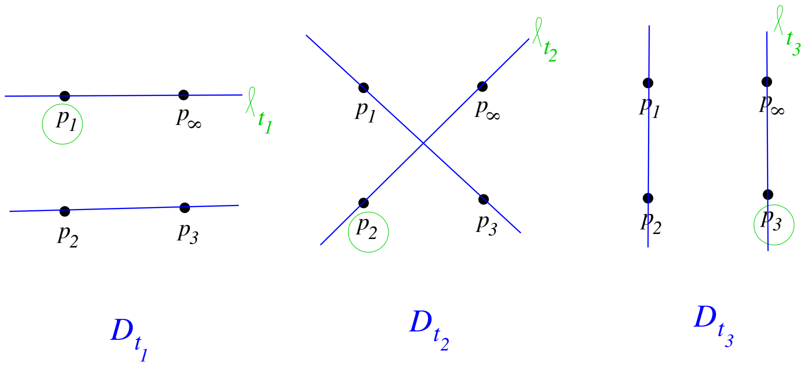
Recall that the equation of a conic may always be written

$$^t\underline{p}.M.\underline{p} = 0\ , \qquad M \text{ a symmetric 3x3 matrix;}$$

the conic is singular if and only if $\det M = 0$. Write M_C, M_D for the matrices corresponding to C, D, so that $tM_C + M_D$ corresponds to D_t. Those t for which D_t is singular, are then just the t_i in

$$\det(tM_C + M_D) = \kappa(t - t_1)(t - t_2)(t - t_3). \tag{21.2.2}$$

There are three singular conics through the $\{p_i\}_{i=1,2,3,\infty}$:



For any $t \in \mathbb{P}^1$, let

$$\ell_t := \text{tangent line (through } p_\infty) \text{ to } D_t$$

$$\ell_t \cap C \ =: \ p_\infty + q_t \quad \text{(defines } q_t).$$

From the pictures above, we see that

$$q_{t_i} = p_i \quad (i = 1, 2, 3),$$

$$q_\infty = p_\infty \quad \text{(double intersection)},$$

$$q_0 = e.$$

So stereographic projection through p_∞gives an isomorphism (normalization)

$$\mathbb{P}^1 \xrightarrow{\cong} C$$

sending

$$t \longmapsto q_t,$$

in particular

$$t_i \longmapsto p_i \ \ (i = 1, 2, 3),$$

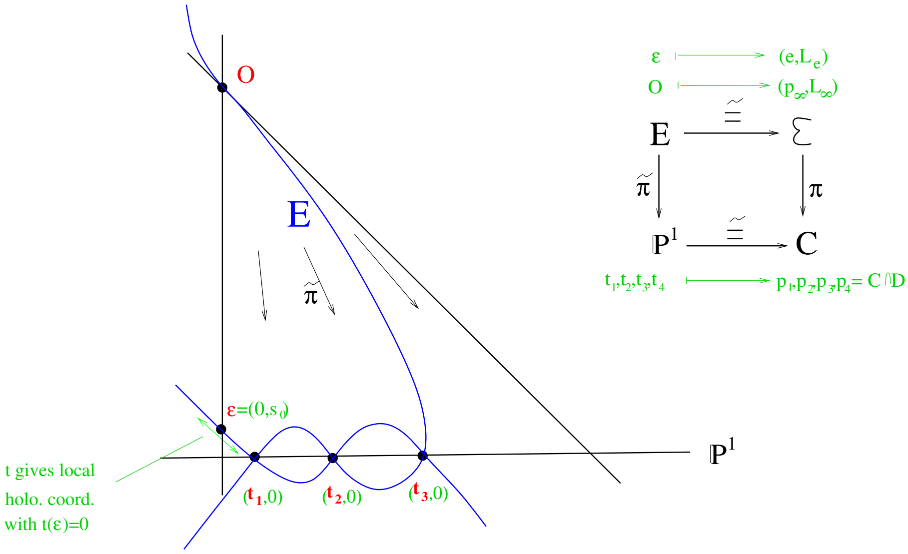
$$\infty \longmapsto p_\infty,$$

$$0 \longmapsto e.$$

This makes $\mathcal{E}$ a double-cover[3] of $\mathbb{P}^1$ branched at t_1, t_2, t_3, ∞ — i.e. the "existence domain" (cf. §2.3) of $\sqrt{(21.2.2)}$, which is to say the Riemann surface

$$\left\{s^2 = \det(tM_C + M_D)\right\} =: E. \tag{21.2.3}$$

The point (e, L_e) on $\mathcal{E}$ corresponds to a point over $t = 0$ on E; call this ε. (Moreover, $\mathcal{O} \in \mathcal{E}$ corresponds to $[0:0:1] =: \mathcal{O} \in E$, as it should.) Summarizing everything in a picture:



Our main question (21.2.1) becomes:

$$\text{Is } \varepsilon \ N\text{-torsion on } E?$$

Now $t_1 + t_2 + t_3$ may not be zero and we are lacking a factor of 4, so E is not quite in Weierstrass form. But it is easy to see that we have a normalization

$$\mathcal{P} : \mathbb{C}/\Lambda \xrightarrow{\cong} E$$

given by

$$u \longmapsto \left(\wp(u) + \frac{\sum t_i}{3}, \frac{\wp'(u)}{2\kappa^{-\frac{1}{2}}}\right).$$

Clearly this sends $0 \mapsto \mathcal{O}$; define $u_0 \in \mathbb{C}/\Lambda$ to be the point sent to ε. The question is now:

$$\text{Is } u_0 \ N\text{-torsion on } \mathbb{C}/\Lambda?$$

---

[3]remember, $\mathcal{E}$ was already (via π) a double-cover of C branched over p_1, p_2, p_3, p_∞

"Normal" elliptic curves and a "multiple addition" theorem. Put $u_j := u_0 + \Delta_j$, where $\Delta_j \in \mathbb{C}$. Abel's theorem implies

PROPOSITION 21.2.1. *There exists a Λ-periodic meromorphic function F with order-N pole at 0 and simple zeroes at $u_1, \ldots, u_N$, if and only if $u_1 + \cdots + u_N \equiv 0 \pmod{\Lambda}$.*

What we are really after here is the vector space V of meromorphic functions on E with at worst an order-N pole at $\mathcal{O}$ (and no other poles). There are $N-1$ degrees of freedom coming from pushing around the $\{u_j\}$ (while keeping $\sum u_j \equiv 0$) and one degree of freedom from multiplying the function by a constant. So $\dim V = N$; let $\{f_1, f_2, \ldots, f_N\}$ (with f_1 constant) be a basis, and define

$$\varphi_N : E \longrightarrow \mathbb{P}^{N-1}_{[w_1:\cdots:w_N]}$$

by

$$([1:t:s] =:) \, \underline{z} \longmapsto [f_1(\underline{z}) : \cdots : f_N(\underline{z})].$$

DEFINITION 21.2.2. The image of φ_N, denoted E_N, is called a *normal* elliptic curve of degree N. (Note that $E_3$is essentially E — take f_1, f_2, f_3 to be $1, t, s$.)

For $\sum u_j \equiv 0$, there exists a function $\mathfrak{F}$ on E_N with zeroes at $\varphi_N(\mathcal{P}(u_j))$, and order N pole at $\varphi_N(\mathcal{O})$. Now the "hyperplane at infinity" $\{w_1 = 0\} \subset \mathbb{P}^{N_1}$ intersects E_N only at $\varphi_N(\mathcal{O})$ (with multiplicity N). If written as the pullback to E_N of a rational function, it follows that $\mathfrak{F}$ has "denominator" w_1; the numerator must then also be a homogeneous linear form $H \in S^1_N$, i.e. $\mathfrak{F} = \left.\frac{H(\underline{w})}{w_1}\right|_{E_N}$. It follows that the $\varphi_N(\mathcal{P}(u_j))$ all lie on $\{H = 0\} \subset \mathbb{P}^{N-1}$, and so

$$0 = \det[\overbrace{\underbrace{f_i(\mathcal{P}}_{=:F_i}(u_j))}^{\substack{\text{coords. of} \\ \varphi_N(\mathcal{P}(u_j))}}]. \tag{21.2.4}$$

Conversely if this is satisfied then the $\varphi_N(\mathcal{P}(u_j))$ lie on a hyperplane $\{H = 0\}$; one writes down the function $\left.\frac{H(\underline{w})}{w_1}\right|_{E_N}$ and computes its divisor $\sum_{j=1}^N [u_j] - N[0]$, and concludes (by Abel) that $\sum u_j \equiv 0$.

We can push this computation further. Expand $F_i(u_0+\Delta_j) =$

$$F_i(u_0) + F_i'(u_0)\Delta_j + \cdots + \frac{F_i^{(N-1)}(u_0)}{(N-1)!}\Delta_j^{N-1} + \Delta_j^N(\,\cdots\,)\,,$$

then apply multilinearity of the determinant to expand the RHS of (21.2.4):[4]

$$0 \;=\; \text{const.} \times \prod_{k>\ell}(\Delta_k - \Delta_\ell) \times \det\left[F_i^{(j-1)}(u_0)\right] \;+\; \begin{pmatrix} \text{terms of higher} \\ \text{homog. degree} \\ \text{in the } \{\Delta_j\} \end{pmatrix}.$$

Dividing by $\prod_{k>\ell}(\Delta_k - \Delta_\ell)$ and taking the limit as all $\Delta_j \to 0$ (i.e. all $u_j \to u_0$), this becomes

$$0 \;=\; \det\left[F_i^{(j-1)}(u_0)\right]_{\substack{i\,=\,1,\ldots,\,N \\ j\,=\,1,\ldots,\,N}}\,. \tag{21.2.5}$$

The determinant on the RHS of (21.2.5) is called the *Wronskian* of $\varphi_N \circ \mathcal{P}$. Notice that in the limit $\sum_{j=1}^N u_j \equiv 0$ becomes $Nu_0 \equiv 0$; so this last condition is equivalent to (21.2.5)!

Example 21.2.3. Here is what the above calculation (using multilinearity of the determinant) looks like for $N = 2$, ignoring terms of degree higher than 1 in the $\{\Delta_j\}$:

$$\begin{vmatrix} F_1 + \Delta_1 F_1' & F_1 + \Delta_2 F_1' \\ F_2 + \Delta_1 F_2' & F_2 + \Delta_2 F_2' \end{vmatrix} = \begin{vmatrix} F_1 + \Delta_1 F_1' & (\Delta_2 - \Delta_1)F_1' \\ F_2 + \Delta_1 F_2' & (\Delta_2 - \Delta_1)F_2' \end{vmatrix}$$

$$= \begin{vmatrix} F_1 & (\Delta_2 - \Delta_1)F_1' \\ F_2 & (\Delta_2 - \Delta_1)F_2' \end{vmatrix} = (\Delta_2 - \Delta_1)\begin{vmatrix} F_1 & F_1' \\ F_2 & F_2' \end{vmatrix}.$$

Using the chain rule and again multilinearity of "det", one finds that the vanishing of the Wronskian is independent of the choice of local coordinate on E. So we can replace u by t (and hence F by f), which yields our "multiple addition theorem":

Theorem 21.2.4. *u_0 is N-torsion in $\mathbb{C}/\Lambda$ (and the Poncelet iteration closes up at the N^{th} step) if and only if*

$$\det\left[f_i^{(j-1)}(0)\right]_{\substack{i\,=\,1,\ldots,\,N \\ j\,=\,1,\ldots,\,N}} = 0. \tag{21.2.6}$$

[4]note: "higher homogeneous degree in the $\{\Delta_j\}$" means higher than $\prod_{k>\ell}(\Delta_k - \Delta_\ell)$

REMARK 21.2.5. The meaning of $f_i^{(j-1)}(0)$ probably requires explanation: first, we are viewing f locally as a function of t (rather than of $[1:s:t] =: \underline{z}$ on E), and the $(j-1)^{\text{st}}$ derivative is (total derivative) with respect to t. The "0" just means t is set to 0 at the end; this is because we are evaluating at ε (i.e. u_0), which has coordinates $[(s,t) =]\,(s_0, 0)$.

Application in the case N odd. Obviously we can't compute the Wronskian (21.2.6) unless we know the f_i.

Take $N = 2m+1$. Then for $f_1, \ldots, f_{m+1}; f_{m+2}, \ldots, f_{2m}$ we may choose

$$1, t, \ldots, t^m;\ s, st, \ldots, st^{m-1}.$$

These have order of pole at 0

$$0, 2, \ldots, 2m;\ 3, 5, \ldots, 2m+1.$$

The determinant in (21.2.6) is then (using that $\left.\frac{d^{j-1}t^{i-1}}{dt^{j-1}}\right|_0 = 0$ unless $j = i$)

$$\left|\begin{array}{cccccc}
1 & & 0 & 0 & \cdots & 0 \\
 & \ddots & & \vdots & \ddots & \vdots \\
0 & & m! & 0 & \cdots & 0 \\
 & \cdots & * & \left.\frac{d^{m+1}s}{dt^{m+1}}\right|_\varepsilon & \cdots & \left.\frac{d^{2m}s}{dt^{2m}}\right|_\varepsilon \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
 & \cdots & * & \left.\frac{d^{m+1}(st^{m-1})}{dt^{m+1}}\right|_\varepsilon & \cdots & \left.\frac{d^{2m}(st^{m-1})}{dt^{2m}}\right|_\varepsilon
\end{array}\right|.$$

Writing $s = s(t) = \sqrt{\det(tM_C + M_D)} = A_0 + A_1 t + A_2 t^2 + \cdots$ (here $A_0 = s_0$), this becomes a nonzero constant times

$$\left|\begin{array}{ccc}
A_{m+1} & \cdots & A_{2m} \\
\vdots & \ddots & \vdots \\
A_2 & \cdots & A_{m+1}
\end{array}\right|. \tag{21.2.7}$$

We conclude that there is a circuminscribed $(2m+1)$-gon (and hence a family of such) for the pair C, D iff (21.2.7) vanishes.

EXAMPLE 21.2.6. We work out the case $N = 3$, i.e $m = 1$. The determinant (21.2.7) is just A_2, so we can get a "Poncelet triangle"

$\iff A_2 = 0$. Writing $T_i = \frac{1}{t_i}$, calculate

$$s = \sqrt{\det(tM_C + M_D)} = \sqrt{\kappa \prod_{i=1}^{3}(t - t_i)}$$

$$= C\prod_{i=1}^{3}\sqrt{1 - \frac{t}{t_i}} = C\prod_{i=1}^{3}\left(1 - \frac{T_i}{2}t - \frac{T_i^2}{8}t^2 - \cdots\right)$$

$$\implies \quad \frac{A_2}{C} = -\frac{1}{8}\sum_{i=1}^{3} T_i^2 + \frac{1}{4}\left(T_1T_2 + T_2T_3 + T_1T_3\right).$$

If $T_1 = 1$, solving a quadratic equation we find

$$A_2 = 0 \quad \iff \quad T_2 = \frac{(1+T)^2}{4}\,,\ T_3 = \frac{(1-T)^2}{4} \ \text{ for some } T$$

$$\iff \quad \text{equation of } E \text{ reads } s^2 = \kappa(t-1)\left(t - \frac{4}{(1+T)^2}\right)\left(t - \frac{4}{(1-T)^2}\right).$$

If we take

$$M_D = \begin{pmatrix} \frac{-4}{(1+T)^2} & 0 & 0 \\ 0 & \frac{-4}{(1-T)^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

corresponding to

$$C = \left\{\frac{4x^2}{(1+T)^2} + \frac{4y^2}{(1-T)^2} = 1\right\}, \quad D = \left\{x^2 + y^2 = 1\right\},$$

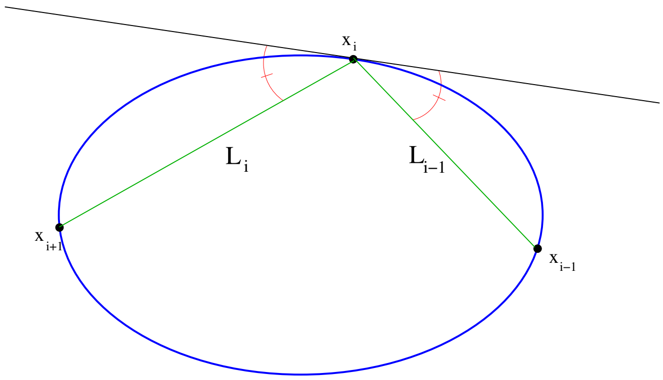
then $\kappa = -1$ and indeed

$$\det\left(tM_C + M_D\right) = \left(t - \frac{4}{(1+T)^2}\right)\left(t - \frac{4}{(1-T)^2}\right)(1-t).$$

This recovers Example 1.3.2(b) from the beginning of the course! It's easy to draw one triangle, but seems quite nontrivial that you get one independent of the starting point.

21.3. Elliptic billards

Returning to the "real" world, let $C_{\mathbb{R}} \subset \mathbb{R}^2$ be an ellipse with foci F_1 and F_2. ($C_{\mathbb{R}}$ consists of all points in $\mathbb{R}^2$, the sum of whose distances from F_1 and F_2 is a fixed constant.) We imagine that $C_{\mathbb{R}}$ is the boundary of a pool table (frictionless, of course!). A *billiard trajectory* for $C_{\mathbb{R}}$ is a sequence of pairs $(x_i, L_i)_{i\geq 0}$ with x_i, $x_{i+1} \in C \cap L_i$ and where L_{i-1}, L_i make equal angles with $T_{x_i}C_{\mathbb{R}}$



— i.e. one has "equality of angles of incidence and reflection".

If $D_{\mathbb{R}}$ is another conic (ellipse or hyperbola) then a (real) *Poncelet trajectory* for $(C_{\mathbb{R}}, D_{\mathbb{R}})$ is a sequence of pairs $(x_i, L_i)_{i\geq 0}$ with x_i, $x_{i+1} \in C \cap L_i$ and L_i tangent to $D_{\mathbb{R}}$.

THEOREM 21.3.1. [L. FLATTO, 2003] *(a) Assume $D_{\mathbb{R}}$ is confocal with*[5] *$C_{\mathbb{R}}$. Then the (real) Poncelet trajectories are billiard trajectories with respect to $C_{\mathbb{R}}$.*

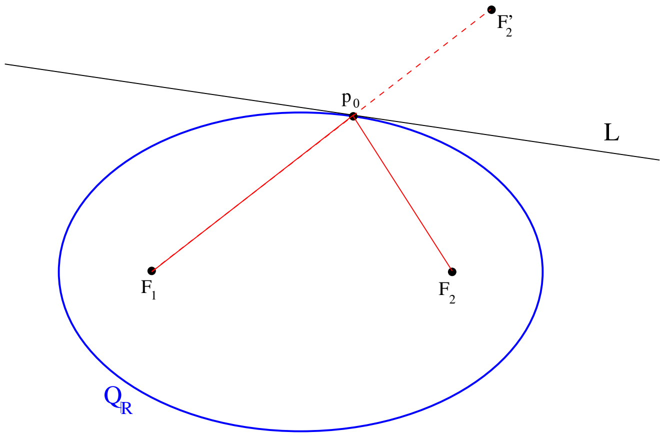
(b) Conversely, any billiard trajectory for $C_{\mathbb{R}}$ not passing through F_1 or F_2 and not along the minor axis, is a Poncelet trajectory for $C_{\mathbb{R}}$ and some $D_{\mathbb{R}}$ confocal with $C_{\mathbb{R}}$.

We will prove only (a); Flatto does (b) in Appendix E of his book [L. Flatto, "Poncelet's Theorem," AMS, 2009] (which is, by the way, written for undergraduates). At any rate, the two proofs are very similar.

REMARK 21.3.2. It's worth pointing out right away that given (x_0, L_0) (L_0 not containing F_1 or F_2 and not the minor axis), there is a unique conic $D_{\mathbb{R}}$ confocal with $C_{\mathbb{R}}$ and tangent to L_0. If L_0 passes between F_1 and F_2, $D_{\mathbb{R}}$ is a hyperbola; otherwise, it's an ellipse. One determines this $D_{\mathbb{R}}$, and then from $\sqrt{\det(tM_C + M_D)}$ obtains information (as in §21.2) on whether the Poncelet trajectory closes up. By the Theorem, this is also the billiard trajectory! You'll use this to do a computation in Exercise (2) below. But I should emphasize that if you change (x_0, L_0) (i.e. the choice of billiard trajectory), you have to change the choice of $D_{\mathbb{R}}$ accordingly.

[5]that is, F_1 and F_2 are the foci of $D_{\mathbb{R}}$. If $D_{\mathbb{R}}$ is a hyperbola, this just means that the difference of distances from its points to F_1 and F_2 must remain constant.

PROOF. (of (a)): We begin with a general principle, for a conic $Q_{\mathbb{R}}$ with foci F_1, F_2. Given $p_0 \in Q_{\mathbb{R}}$, let $L := T_{p_0} Q_{\mathbb{R}}$



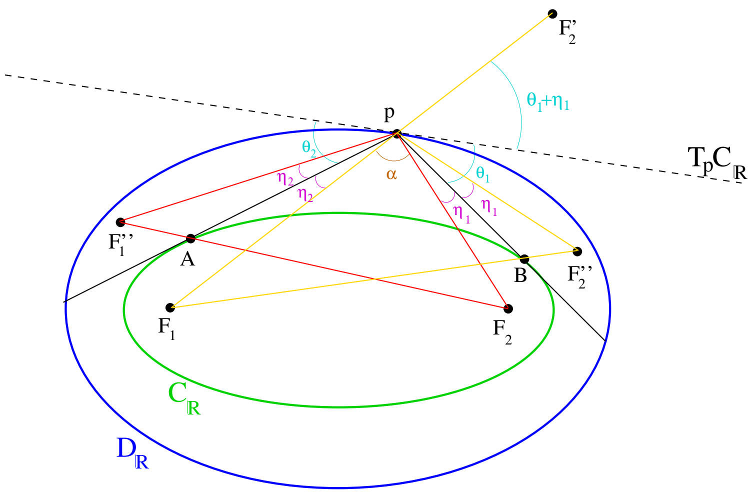
and denote by F_2' the reflection of F_2 in L. Given points p, q write pq for the segment and $|pq|$ for its length. Set $\beta := |F_1 p_0| + |p_0 F_2|$ and note that by definition of ellipse,

$$|F_1 q| + |q F_2| = \beta \qquad (\forall q \in Q_{\mathbb{R}}).$$

If $p \in L \backslash \{p_0\}$, $|F_1 p| + |p F_2'| = |F_1 p| + |p F_2|$ visibly exceeds β, meaning that taking $p = p_0$ minimizes $|F_1 p| + |p F_2'|$. It follows that

$$F_1 p_0 \cup p_0 F_2' = F_1 F_2'. \tag{21.3.1}$$

Now let $C_{\mathbb{R}}$, $D_{\mathbb{R}}$ be confocal — assume that $D_{\mathbb{R}}$ is an ellipse. Applying the principle that (21.3.1) holds for the above construction, leads to a picture



in which the solid black lines are part of a Poncelet iteration and we must show $\theta_1 = \theta_2$ (so that it is a billiard trajectory). reflection in

$T_pC_{\mathbb{R}}$ (dotted black) is denoted by one prime, reflection in solid black lines by two primes.

By definition of ellipse, $F_1A + AF_2 = F_1B + BF_2$, which implies

$$|F_1''F_2| = |F_1F_2''|.$$

From there it is clear that the triangles $F_1''pF_2$ and F_1pF_2'' are rotations of each other (through p), so that $\alpha + 2\eta_1 = \alpha + 2\eta_2$ ($\implies \eta_1 = \eta_2$). It is obvious from the picture that $\theta_1 + \eta_1 = \theta_2 + \eta_2$, and so we indeed conclude that $\theta_1 = \theta_2$. $\square$

Exercises

(1) Consider the pair of conics C, D from Exercise (2) of Chapter 1 once more — but in the following form: write

$$M_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad M_D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2 \end{pmatrix}$$

and use these to define quadratic forms by e.g.

$$Q_C(X, Y, Z) = \begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = X^2 + Y^2 - Z^2$$

So $Q_C = 0$ defines C and $Q_D = 0$ defines D as conics in $\mathbb{P}^2$. Working in homogeneous coordinates $[V : T : U]$, define an elliptic curve by

$$U^2V = \det(T \cdot M_C + V \cdot M_D).$$

In affine coordinates, this is $u^2 = \det(t \cdot M_C + M_D)$, where $t = \frac{T}{V}$, $u = \frac{U}{V}$. This is the general prescription for the elliptic curve $\mathcal{E}$ arising in the Poncelet construction, exactly as above. All you have to show is that in the present situation, with M_C and M_D as given, the elliptic curve is singular. It's practically a one-line problem, and you may use the affine setup. But now you are in a position to see "why" the curve being singular should make the Poncelet problem easier, and even to see "why" (from the abstract perspective) the solution involved trig functions.

(2) Let $a > b > 0$ and

$$C_{\mathbb{R}} = \left\{ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\};$$

it has foci $(\pm\sqrt{a^2 - b^2}\,,\ 0)$. We plan to shoot our pool ball vertically along the line $(L_0 =)\ \{x = c\}$, where $0 < c < a$ (and $c \neq \sqrt{a^2 - b^2}$). For what value of c does the resulting billiard trajectory yield a triangle? [Hint: the conics confocal with $C_{\mathbb{R}}$ are all of the form $\left\{ \frac{x^2}{a^2-\lambda} + \frac{y^2}{b^2-\lambda} = 1 \right\}$. Also: while straightforward, this is not a 1-line computation!]

CHAPTER 22

Periods of families of elliptic curves

The *periods* of an elliptic curve $E \subset \mathbb{P}^2$ are simply elements of the period lattice $\Lambda_E = \mathbb{Z}\left\langle \int_\alpha \omega, \int_\beta \omega \right\rangle$ where α, β are 1-cycles generating $H_1(E, \mathbb{Z})$ and $\omega \in \Omega^1(E)$ is some canonically chosen generator. (That is, if E is defined over $\mathbb{Q}$, then ω should be the restriction of a rational differential 1-form on $\mathbb{P}^2$ defined over $\mathbb{Q}$.) If E is taken to vary with respect to a parameter $t \in \mathbb{P}^1$, the periods give interesting multivalued transcendental functions (e.g. hypergeometric functions) which are related to modular forms.

In this Chapter we explore (via examples) two different approaches to computing "period functions" of this sort — the "Euler integral" method and the "Picard-Fuchs" method. The first of these is just a way of computing the integral using Laurent polynomials; the second derives a homogeneous linear ordinary differential equation satisfied by the periods, which yields a recurrence relation for their power-series coefficients. Actually, both methods yield power series at first but one can sometimes recognize what functions they are the power series of. This may sound like complex function theory, but in fact the power series coefficients (esp. when related to modular forms) can have arithmetic meaning, as we shall see in the next chapter; in the context of mirror symmetry (one of several interfaces between algebraic geometry and string theory), power series derived from periods are related to counting curves on threefolds.

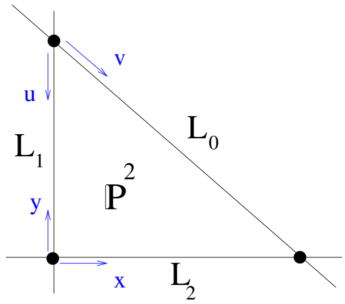
Sections §22.1 and §22.3 will have a bit of overlap with §18.1.

22.1. Holomorphic 1-forms on a smooth cubic $\subset \mathbb{P}^2$

Let $F \in S^3$ define a smooth curve $E = \{F(Z_0, Z_1, Z_2) = 0\}$; by the genus formula $g = \frac{(3-1)(3-2)}{2} = 1$, so that E is elliptic.

EXAMPLE 22.1.1. $F = Z_0 Z_1 Z_2 - t(Z_0^3 + Z_1^3 + Z_2^3)$, for any $t \in \mathbb{P}^1 \backslash \{0, \frac{1}{3}, \frac{\zeta_3}{3}, \frac{\zeta_3^2}{3}\}$ $(\zeta_3 = e^{\frac{2\pi i}{3}})$.

For the affine forms of the equation we shall use the following notation:



- on $\mathbb{P}^2\backslash L_0$, the coordinates are $x = \frac{Z_1}{Z_0}$, $y = \frac{Z_2}{Z_0}$, and equation is $f(x, y) := \frac{1}{Z_0^3}F(Z_0, Z_1, Z_2)$;
- on $\mathbb{P}^2\backslash L_2$, the coordinates are $u = \frac{Z_0}{Z_2}$, $v = \frac{Z_1}{Z_2}$, and equation is $g(u, v) := \frac{1}{Z_2^3}F(Z_0, Z_1, Z_2)$;
- the third neighborhood is left to you;
- on $\mathbb{P}^2\backslash L_0 \cup L_2$, we have $u = \frac{1}{y}$, $v = \frac{x}{y}$; $y = \frac{1}{u}$, $x = \frac{u}{v}$; and $f(x, y) = y^3 g\left(\frac{1}{y}, \frac{x}{y}\right)$.

Now define a form ω on E by

$$\left.\frac{dx}{f_y}\right|_{E\backslash(L_0\cup\mathrm{VT})} = -\left.\frac{dy}{f_x}\right|_{E\backslash(L_0\cup\mathrm{HT})} = \left.\frac{du}{g_v}\right|_{E\backslash(L_2\cup\mathrm{HT})} = -\left.\frac{dv}{g_u}\right|_{E\backslash(L_2\cup\mathrm{VT})} = \cdots$$

where

- the notation $E\backslash(L_0 \cup \mathrm{VT})$ means E minus those points where E intersects L_0 or has a vertical tangent line (similarly, HT means "horizontal tangent");
- equality of any two differentials above is meant in the sense of "where both are defined";
- for example: on E, $f = 0 \implies 0 = df = f_x dx + f_y dy \implies \frac{dx}{f_y} = -\frac{dy}{f_x}$ where ($f = 0$ and) $f_x, f_y \neq 0$;
- the "$\cdots$" means that the third neighborhood stuff is left to you.

Now consider the domains of the first two expressions: since f_x and f_y do not simultaneously vanish (E is smooth!), $\{E\backslash(L_0\cap\mathrm{VT})\}\cup\{E\backslash(L_0\cap\mathrm{HT})\}$ is all of $E\backslash L_0$. So the 6 different domains of definition glue to give $(E\backslash L_0) \cup (E\backslash L_1) \cup (E\backslash L_2)$, which is all of E. Morover, $\left.\frac{dx}{f_y}\right|_{E\backslash(L_0\cup\mathrm{VT})}$ etc. are all holomorphic where they are defined. We conclude that $\omega \in \Omega^1(E)$.

By Poincaré-Hopf, $\deg((\omega)) = 2g - 2 = 2 - 2 = 0$, and so ω's lack of poles implies it has no zeroes either. Any other $\omega' \in \Omega^1(E)$ has $\frac{\omega'}{\omega} \in \mathcal{O}(E)$, and then by Liouville ω' is a constant multiple of ω. So $\Omega^1(E)$ has dimension 1, and ω spans it.

22.2. Period of a family of cubic curves (Euler integral method)

Now consider the Hesse family E_t of elliptic curves, already given in Example 22.1.1, with affine form

$$f(x, y) = xy - t(x^3 + y^3 + 1) = 0, \qquad t \in \mathbb{C}.$$

(For the four values of t excluded in the example, E_t is singular hence not an elliptic curve. I won't write f_t because the subscript is reserved here for partial derivatives.) An alternate form of the equation, valid on $\mathbb{C}^* \times \mathbb{C}^*$, is

$$1 - t\underbrace{\left(\frac{x^3 + y^3 + 1}{xy}\right)}_{=:\varphi(x,y)} = 0,$$

where φ belongs to the ring of Laurent polynomials $\mathbb{C}[x, x^{-1}, y, y^{-1}]$.

From the last section, we have the family of holomorphic 1-forms

$$\omega_t := \left.\frac{dx}{f_y}\right|_{E_t} \in \Omega^1(E_t).$$

We can obtain a family of 1-cycles by noticing that $\{|x| = |y| = 1\} \cap E_t$ is empty for $|t| < \frac{1}{3}$, since $|\varphi(x, y)| < 3$ for x, y in the unit circle. Indeed,

$$\gamma_t := \{|x| = 1, \, |y| \le 1\} \cap E_t$$

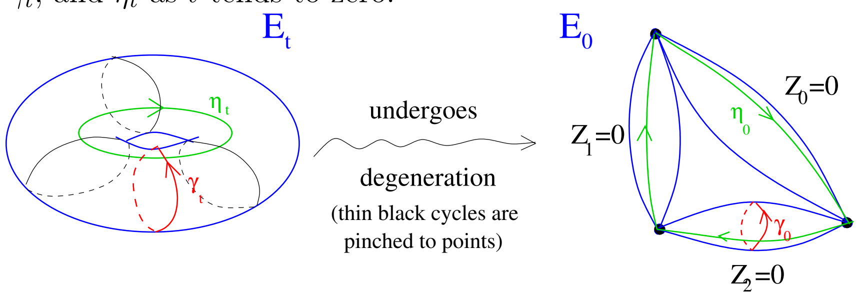
has this empty set as its boundary $\partial\gamma_t$; and so we would like to compute the period

$$P(t) := \int_{\gamma_t} \omega_t$$

as a function of t, on the open disk $|t| < \frac{1}{3}$. Now since $H_1(E_t, \mathbb{Z})$ has rank 2, there is a complementary 1-cycle η_t on E_t, and cooresponding period

$$Q(t) := \int_{\eta_t} \omega_t.$$

Noticing from the homogeneous form of the equation that $E_0 = \{Z_0Z_1Z_2 = 0\}$ is a union of 3 lines ($\cong \mathbb{P}^1$), we can easily visualize what happens to E_t, γ_t, and η_t as t tends to zero:



From the fact that ω_t tends (as $t \to 0$) to $\frac{dx}{x}$ on $Z_2 = 0$, and η_0 passes through the poles of this form while γ_0 traverses the unit circle around them, we infer that $Q(t) \to \infty$ as $t \to 0$ but $P(t) \to 2\pi i$.

We now compute $P(t)$ more precisely, by first noting that the area integral

$$\iint_{|x|=|y|=1} \frac{dx \wedge dy}{f(x,y)} = \iint \frac{dx \wedge df}{f_y \cdot f}$$

(since $df = f_x dx + f_y dy$ and $dx \wedge dx = 0$)

$$= \int_{|x|=1} \left(\int_{|y|=1} \frac{df(x,y)}{f(x,y)} \cdot \frac{1}{f_y(x,y)} \right) dx$$

(where inside the parentheses x is a fixed constant). Now thinking about the equation $f(x,y) = 0$ for $|t|$ small (and x fixed with $|x| = 1$), we have $y^3 + ay + b = 0$ where $a = \frac{x}{t}$ is big and $b = x^3 + 1$ is not. This means that two of the roots are big and one is small – in particular, there is exactly one solution $y(x)$ with modulus less than 1. Therefore, by Cauchy's residue theorem, the integral above

$$= \int_{|x|=1} \left(2\pi i \cdot \frac{1}{f_y(x, y(x))} \right) dx$$

$$= 2\pi i \int_{\substack{|x|\,=\,1 \\ y\,=\,y(x)}} \frac{dx}{f_y} = 2\pi i \int_{\gamma_t} \omega_t .$$

So

$$P(t) = \frac{1}{2\pi i} \iint_{|x|=|y|=1} \frac{dx \wedge dy}{f(x,y)} = \frac{1}{2\pi i} \iint_{|x|=|y|=1} \frac{\frac{dx}{x} \wedge \frac{dy}{y}}{1 - t\varphi(x,y)}$$

$$= \frac{1}{2\pi i} \iint_{|x|=|y|=1} \sum_{n \geq 0} t^n \varphi^n \mathrm{dlog} x \wedge \mathrm{dlog} y$$

where $\mathrm{dlog}x = \frac{dx}{x}$ and φ^n means simply the n^{th} power of $\varphi(x,y)$. Using Cauchy residue twice this

$$= 2\pi i \sum_{n\geq 0} t^n \varphi^n(0,0)$$

in which $\varphi^n(0,0) =: [\varphi^n]_0$ is the constant term of $\varphi^n = (x^2y^{-1}+x^{-1}y^2+x^{-1}y^{-1})^n$.

We can make this more explicit. Given a product

$$\underbrace{(x^2y^{-1}+x^{-1}y^2+x^{-1}y^{-1})\cdot\cdots\cdot(x^2y^{-1}+x^{-1}y^2+x^{-1}y^{-1})}_{n \text{ times}}$$

each contribution to the constant term comes from exponents summing to $(0,0)$ (i.e. multiplying to x^0y^0). But the only combinations of $(2,-1)$, $(-1,2)$, $(-1,-1)$ summming to $(0,0)$ are: $m(2,-1)+m(-1,2)+m(-1,-1)$. Hence, the only possibility for a nozero constant term is to have $n=3m$ (i.e. $3|n$), and the number of ways to choose

$$\begin{cases} x^2y^{-1} \text{ from } m \text{ factors} \\ x^{-1}y^2 \text{ from } m \text{ factors} \\ x^{-1}y^{-1} \text{ from } m \text{ factors} \end{cases}$$

is $\binom{3m}{m,m,m} := \frac{(3m)!}{m!m!m!}$. That is,

$$P(t) \;=\; 2\pi i \sum_{m\geq 0} t^{3m}\cdot \frac{(3m)!}{(m!)^3} \;=\; 2\pi i\, {}_2F_1\left(\frac{1}{3},\frac{2}{3};1;(3t)^3\right)$$

where by definition (writing $(a)_m := a(a+1)\cdots(a+m-1)$ for the "Pochhammer symbol")

$${}_2F_1(a,b;c;u) := 1 + \sum_{m\geq 1} \frac{(a)_m(b)_m}{(c)_m m!}u^m$$

is the *Gauss hypergeometric function.*

Notice that $P(t)$ is "really" a function of $(3t^3) = u$. This reflects the symmetry in the family E_t:

$$E_t \longrightarrow E_{\zeta_3 t}$$

$$(x,\, y) \longmapsto (\zeta_3 x,\, y).$$

The Gauss hypergeometric function satisfies a well-known ODE. In this case (writing $P_0(u) = P(t)$ and $Q_0(u) = Q(t)$)

$$\left\{ u(1-u)\frac{d^2}{du^2} + (1-2u)\frac{d}{du} - \frac{2}{9} \right\} P_0(u) \;=\; 0.$$

It turns out that the ODE satisfied by P_0 must be satisfied by Q_0 (the other period), which turns out to have a term of the form $\frac{\log u}{2\pi i}P_0(u)$ reflecting the fact that "following η_t around $t = 0$" yields $\eta_t + 3\gamma_t$ in $H_1(E_t, \mathbb{Z})$.

22.3. Cohomology of an elliptic curve E

Let V be a finite-dimensional vector space over $\mathbb{C}$, with basis $\{e_k\}_{k=1}^n$. The second tensor power of V, written $V \otimes V$, is the n^2-dimensional vector space consisting of finite sums $\sum_i v_i \otimes w_i$ $(v_i, w_i \in V)$ subject to bilinearity (e.g., on the left $(\alpha v + \beta w) \otimes u = \alpha v \otimes u + \beta w \otimes u$); it has basis $\{e_k \otimes e_\ell\}_{k,\ell=1}^n$. The second exterior power $\bigwedge^2 V$ consists of finite sums $\sum v_i \wedge w_i$ satisfying bilinearity and also $v \wedge w = -w \wedge v$ (so that $v \wedge v = 0$); it may be viewed as a quotient- or sub-space of $V \otimes V$, and has basis $\{e_k \wedge e_\ell\}_{1 \le k < \ell \le n}$ hence dimension $\binom{n}{2}$. In particular, if $\dim V = 2$, then $\dim(\bigwedge^2 V) = 1$; this is essentially the only case we shall use.

Dualizing the homology groups[1]

$$H_1(E, \mathbb{Z}) = \frac{\mathbb{Z}\langle \text{closed paths on } E \rangle}{\mathbb{Z}\langle \text{boundaries of regions in } E \rangle}$$

from §18.1, we define *cohomology* groups (with complex coefficients) by

$$H^1(E, \mathbb{C}) := \mathrm{Hom}\,(H_1(E, \mathbb{Z}), \mathbb{C}) \;\; (\cong \mathbb{C}^2).$$

Write $A^0(E)$ for C^∞ functions and $A^1(E)$ for C^∞ 1-forms on E, the latter locally of the form $f dx + g dy$, where $z = x + iy$; and finally we have the C^∞ *2-forms* $A^2(E)$. These are objects locally of the form

$$G dx \wedge dy \; (= -G dy \wedge dx = -\frac{i}{2} G dz \wedge d\bar{z} = \frac{i}{2} G d\bar{z} \wedge dz)$$

with f smooth, which you may think of as a field of infinitesimal area elements. In more sophisticated terms, they are C^∞ sections of the bundle $\bigwedge^2 T^* E = \cup_{p \in E} \bigwedge^2 T_p^* E$. (Refer to §13.1 for notation.)

[1]the homology class represented by a 1-cycle γ is written $[\gamma]$

The various degrees of forms are "connected" by exterior differentiation

$$d : A^0(E) \to A^1(E)$$

sending

$$\begin{aligned} F \mapsto dF &:= F_x dx + F_y dy \\ &= \frac{\partial F}{\partial z} dz + \frac{\partial F}{\partial \bar{z}} d\bar{z}, \end{aligned}$$

and

$$d : A^1(E) \to A^2(E)$$

$$\begin{aligned} f dx + g dy &\mapsto df \wedge dx + dg \wedge dy \\ &= (g_x - f_y) dx \wedge dy. \end{aligned}$$

The 1st *de Rham cohomology* group of E is then defined by

$$H^1_{dR}(E, \mathbb{C}) := \frac{\ker(d) \subset A^1(E)}{d(A^0(E))} = \frac{\text{"closed" } C^\infty \text{ forms}}{\text{"exact" } C^\infty \text{ forms}};$$

the class represented by a 1-form ω is written $[\omega]$.

LEMMA 22.3.1. *The map $\theta : H^1_{dR}(E, \mathbb{C}) \to H^1(E, \mathbb{C})$ given by*[2] $[\omega] \mapsto \{[\gamma] \mapsto \int_\gamma \omega\}$ *is well-defined, and an isomorphism.*

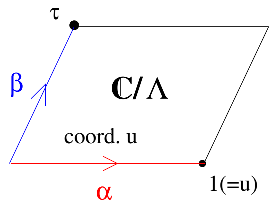
PROOF. First we check well-definedness: if γ is closed ($\partial\gamma = 0$) and ω exact ($\omega = d\eta$), then $\theta([\omega]) = 0$ since $\int_\gamma \omega = \int_\gamma d\eta = \int_{\partial\gamma} \eta = 0$. If γ is a boundary ($\gamma = \partial\Gamma$) and ω is closed ($d\omega = 0$) then $\int_\gamma \omega = \int_{\partial\Gamma} \omega = \int_\Gamma d\omega = 0$, so that $\theta([\omega])$ is defined on the level of homology classes. (The middle equality in both cases — swapping ∂ and d — is *Stokes's theorem*, a generalization of the fundamental theorem of calculus for differential forms.)

To see that θ is injective, assume $\theta([\omega]) = 0$, and let p be a point of E. Then $\mathcal{F}(q) = \int_p^q \omega$ defines a C^∞ function $\mathcal{F}$ on E. (The reason $\mathcal{F}$ isn't "multivalued" is that two paths differ by a cycle γ, and $\int_\gamma \omega = 0$ by the assumption.) Now $\omega = d\mathcal{F}$ by the fundamental theorem of calculus, and so $[\omega] = 0$.

Finally, write α, β for a basis of $H_1(E, \mathbb{Z})$. Using the identification $H^1(E, \mathbb{C}) \xrightarrow{\cong} \mathbb{C}^2$ which evaluates a functional against this basis, a nice way to think about the map θ is as sending $[\omega] \mapsto \begin{pmatrix} \int_\alpha \omega \\ \int_\beta \omega \end{pmatrix} \in \mathbb{C}^2$.

[2]"$[\gamma] \mapsto \int_\gamma \omega$" means the complex-linear functional on homology classes given by integrating ω over a representative 1-cycle

Moreover, the Abel map $E \xrightarrow{\cong} \mathbb{C}/\Lambda_E$ identifies ω with du. Rescaling ω (by a complex constant) so that $\int_\alpha \omega = 1$, the 2-vector becomes $\theta([du]) = \begin{pmatrix} \int_\alpha du \\ \int_\beta du \end{pmatrix} = \begin{pmatrix} 1 \\ \tau \end{pmatrix}$ (where we may assume $\tau \in \mathfrak{H}$), and we have the standard picture



Noting that $\theta([d\bar{u}]) = \begin{pmatrix} \int_\alpha d\bar{u} \\ \int_\beta d\bar{u} \end{pmatrix} = \begin{pmatrix} \overline{\int_\alpha du} \\ \overline{\int_\beta du} \end{pmatrix} = \begin{pmatrix} 1 \\ \bar{\tau} \end{pmatrix}$, we conclude that θ is surjective since $\begin{pmatrix} 1 \\ \tau \end{pmatrix}, \begin{pmatrix} 1 \\ \bar{\tau} \end{pmatrix}$ span $\mathbb{C}^2$. $\square$

REMARK 22.3.2. (a) Note that meromorphic 1-forms on a Riemann surface are always closed since locally (f mero.) $d\{f dz\} = df \wedge dz = \frac{\partial f}{\partial z} dz \wedge dz$ and $dz \wedge dz = -dz \wedge dz = 0$. (Here we have used $\frac{\partial f}{\partial \bar{z}} = 0$, which expresses the holomorphicity of f off its poles.) Using the same formula as above (i.e. $\omega \mapsto \{\gamma \mapsto \int_\gamma \omega\}$), we can define a map

$$\frac{\ker(Res) \subset \mathcal{K}^1(E)}{d(\mathcal{K}(E))} \xrightarrow{\tilde{\theta}} H^1(E, \mathbb{C})$$

which also turns out to be an isomorphism. (Here $\ker(Res)$ consists of forms with no residues – in particular, with no simple poles. This doesn't mean they're holomorphic though!)

(b) A nonzero holomorphic 1-form ω cannot be d of a smooth function G or meromorphic function f. (Locally the integral of ω is a holomorphic function, so in either case G or f would actually have to be holomorphic, hence by Liouville constant, making ω zero.) So we have a commuting diagram of injective homomorphisms

$$\begin{tikzcd} \Omega^1(E) \arrow[r, hook] \arrow[d, hook] \arrow[dr, hook, "\hat{\theta}"] & \frac{\ker(d) \subset A^1(E)}{\text{exact}} \arrow[d, "\theta", "\cong"'] \\ \frac{ker(Res) \subset \mathcal{K}^1(E)}{\text{exact}} \arrow[r, "\tilde{\theta}", "\cong"'] & H^1(E, \mathbb{C}) \end{tikzcd} \tag{22.3.1}$$

.In what follows $\hat{\theta}, \tilde{\theta}, \theta$ will all just be denoted θ, which you should read "take the period vector associated to this 1-form".

22.4. Differentiating cohomology classes

Given a family $\{E_t\}_{t\in\mathbb{P}^1}$ of elliptic curves (smooth but for finitely many t) with holomorphic forms $\omega_t \in \Omega^1(E_t)$, write

$$\theta(\omega_t) =: \begin{pmatrix} P(t) \\ Q(t) \end{pmatrix}.$$

This assumes a choice (this unfortunately only works locally in t) of basis α_t, β_t for $H_1(E_t, \mathbb{Z})$, so that $P(t) = \int_{\alpha_t} \omega_t$, $Q(t) = \int_{\beta_t} \omega_t$. We can differentiate this period vector to obtain

$$\begin{pmatrix} P'(t) \\ Q'(t) \end{pmatrix},$$

which for each t (considering the isomorphisms in (22.3.1)) is θ of something in $\ker(d) \subset A^1(E_t)$ or $\ker(Res) \subset \mathcal{K}^1(E)$ (but not $\Omega^1(E_t)$). We will use the latter, and we write ω_t' for a family of residue-free meromorphic 1-forms satisfying $\theta(\omega_t') = \begin{pmatrix} P'(t) \\ Q'(t) \end{pmatrix}$. The point is that by differentiating families of cohomology classes you get a new family of cohomology classes.

EXAMPLE 22.4.1. Consider the Legendre family $E_t \subset \mathbb{P}^2$ given by the projective closure of

$$y^2 = x(x-1)(x-t),$$

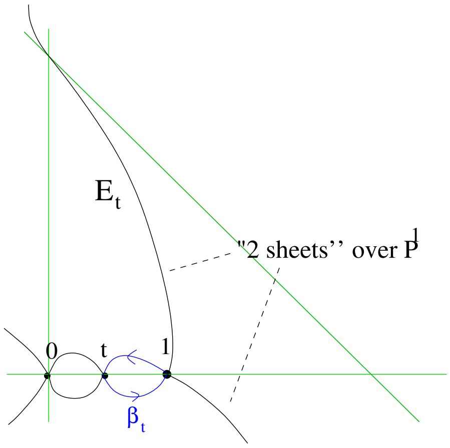
with holomorphic 1-form family

$$\omega_t = \left.\frac{dx}{y}\right|_{E_t} \in \Omega^1(E_t).$$

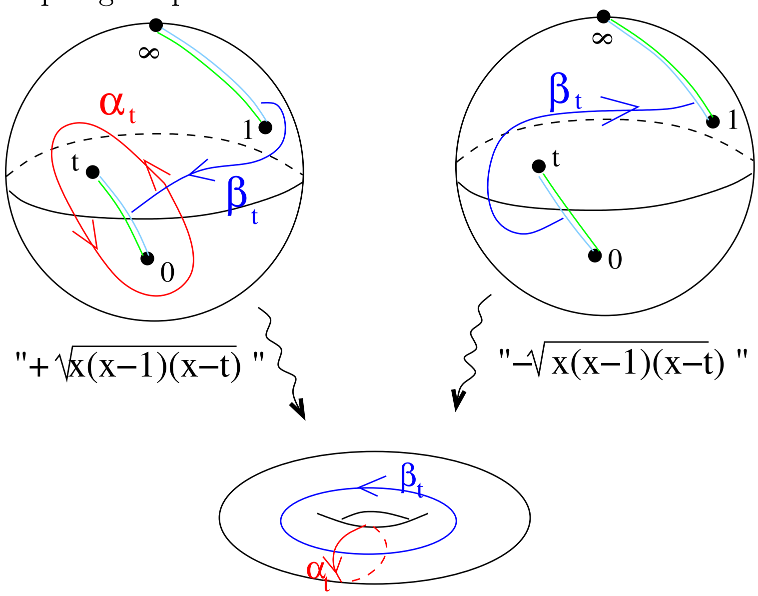
We have

$$\theta(\omega_t) = \begin{pmatrix} P(t) \\ Q(t) \end{pmatrix} = \begin{pmatrix} \int_{\alpha_t} \omega_t \\ \int_{\beta_t} \omega_t \end{pmatrix} = \begin{pmatrix} \int_{\alpha_t} \frac{dx}{\pm\sqrt{x(x-1)(x-t)}} \\ \int_{\beta_t} \frac{dx}{\pm\sqrt{x(x-1)(x-t)}} \end{pmatrix},$$

where α_t, β_t are the 1-cycles exhibited in the schmatic picture



or the topological picture



— which shows the two sheets (each is a $\mathbb{P}^1$ with slits from 0 to t and 1 to ∞) being glued together to give E_t (cf. §2.3).

From the latter picture, it is clear that for t small we may take α_t to be stationary on its sheet as t moves, and the two "pieces" of β_t on the different sheets not to change either. Therefore we may differentiate the above integrals under the integral sign (by $\frac{d}{dt}$) to obtain

$$\begin{pmatrix} P'(t) \\ Q'(t) \end{pmatrix} = \begin{pmatrix} \int_{\alpha_t} \frac{\frac{1}{2}dx}{\pm(x-t)\sqrt{x(x-1)(x-t)}} \\ \int_{\beta_t} \frac{\frac{1}{2}dx}{\pm(x-t)\sqrt{x(x-1)(x-t)}} \end{pmatrix}$$

and

$$\begin{pmatrix} P''(t) \\ Q''(t) \end{pmatrix} = \begin{pmatrix} \int_{\alpha_t} \frac{\frac{3}{4}dx}{\pm(x-t)^2\sqrt{x(x-1)(x-t)}} \\ \int_{\beta_t} \frac{\frac{3}{4}dx}{\pm(x-t)^2\sqrt{x(x-1)(x-t)}} \end{pmatrix};$$

obviously the first is $\theta\left(\frac{1/2}{x-t}\frac{dx}{y}|_{E_t}\right)$ and the second $\theta\left(\frac{3/4}{(x-t)^2}\frac{dx}{y}|_{E_t}\right)$, and so we have

$$\omega_t' = \frac{1/2}{(x-t)}\frac{dx}{y}\bigg|_{E_t}, \qquad \omega_t'' = \frac{3/4}{(x-t)^2}\frac{dx}{y}\bigg|_{E_t}.$$

These both belong to $\ker(Res) \subset \mathcal{K}^1(E_t)$, since their only poles are at $(t,0)$ (orders 2 and 4 resp.) and the sum of the residues of a meromorphic form must always be zero.

Of course, $\theta(\omega_t)$, $\theta(\omega_t')$, and $\theta(\omega_t'')$ must be linearly dependent in $\mathbb{C}^2$! Therefore, $[\omega_t]$, $[\omega_t']$, and $[\omega_t'']$ are linearly dependent in $\mathcal{K}^1(E_t)$ modulo $d(\mathcal{K}(E_t))$, i.e. as cohomology classes.

22.5. The Picard-Fuchs equation

Since what is being differentiated in the last section is really cohomology *classes* (via the identification with $\mathbb{C}^2$), it makes sense to write

$$D_t[\omega_t] = [\omega_t'], \qquad D_t^2[\omega_t] = [\omega_t''].$$

With this notation, the linear dependence observation above implies an ODE of the form

$$\underbrace{\left(A(t)D_t^2 + B(t)D_t + C(t)\right)}_{=:D_{\mathrm{PF}}}(\cdot) = 0 \tag{22.5.1}$$

satisfied by $[\omega_t]$ (as a varying cohomology class) *hence by* $P(t)$ *and* $Q(t)$*!*

However, to find A, B, and C, we have to compute. We start by differentiating a meromorphic function

$$\underbrace{d\left(\frac{2y}{(x-t)^2}\bigg|_{E_t}\right)}_{\in\mathcal{K}(E)} = \frac{-4ydx}{(x-t)^3}\bigg|_{E_t} + \frac{2dy}{(x-t)^2}\bigg|_{E_t},$$

which using $y^2 = x(x-1)(x-t) \implies dy = \frac{3x^2-2(1+t)x+t}{2y}dx$ becomes

$$= \left(\frac{-4y^2}{(x-t)^3y} + \frac{3x^2-2(1+t)x+t}{(x-t)^2y}\right)dx\bigg|_{E_t}$$

and using $y^2 = x(x-1)(x-t)$ again

$$= \left.\frac{-x^2+(2-2t)x+t}{(x-t)^2 y}dx\right|_{E_t}$$

$$= \left.\frac{-(x-t)^2-2tx+t^2+(2-2t)x+t}{(x-t)^2}\frac{dx}{y}\right|_{E_t}$$

$$= -\omega_t + \left.\frac{(2-4t)x+t^2+t}{(x-t)^2}\frac{dx}{y}\right|_{E_t}$$

$$= \cdots = -\omega_t + 4(1-2t)\omega_t' + 4t(1-t)\omega_t''.$$

So this last expression is d of a meromorphic function, hence (has both its periods 0 and) is trivial in $H^1(E_t, \mathbb{C})$. We conclude that (dividing through by 4 to simplify)

$$D_{\mathrm{PF}} = t(t-1)D_t^2 + (2t-1)D_t + \frac{1}{4}$$

kills $[\omega_t]$. From ODE theory, the associated indicial equation is

$$r(r-1) + \left(\lim_{t\to 0}\frac{B(t)}{A(t)}t\right)r + \left(\lim_{t\to 0}\frac{C(t)}{A(t)}\right) = r^2$$

which has a double root, implying one holomorphic solution (unique up to scale) and one logarithmic solution near $t = 0$.

22.6. Computation of a period (Picard-Fuchs method)

First let's compute its limit

$$\underbrace{\lim_{t\to 0}\int_{\alpha_t}\omega_t}_{P(t)} = \lim_{t\to 0}\int_{\alpha_t}\frac{dx}{\pm\sqrt{x(x-1)(x-t)}}.$$

Referring to the picture of α_t (on the slit $\mathbb{P}^1$) above, this

$$= \oint \frac{dx}{x\sqrt{x-1}} = 2\pi i \cdot Res_0\left(\frac{dx}{x\sqrt{x-1}}\right)$$

$$= 2\pi i \cdot \frac{1}{\sqrt{-1}} = 2\pi,$$

and so $P(t)$ must be "the" holomorphic solution. Write $P(t) = 2\pi\sum a_n t^n$, $a_0 = 1$, and apply D_{PF}:

$$0 = D_{\mathrm{PF}}\sum a_n t^n$$

$$= \sum_{n\geq 0}\left[t(t-1)(n+2)(n+1)a_{n+2} + (2t-1)(n+1)a_{n+1} + \frac{1}{4}a_n\right]t^n$$

where we have shifted indices after differentiating. Collecting terms with like powers of t, this

$$= \sum_{n\geq 0}\left[\frac{1}{4}a_n - (n+1)a_{n+1}\right]t^n + \sum_{n\geq 0}\left[2(n+1)a_{n+1} - (n+1)(n+2)a_{n+2}\right]t^{n+1}$$
$$+ \sum_{n\geq 0}(n+1)(n+2)a_{n+2}t^{n+2}.$$

Shifting indices once more, we have

$$= \sum_{n\geq 0}\left[\frac{1}{4}a_n - (n+1)a_{n+1} + 2na_n - n(n+1)a_{n+1} + n(n-1)a_n\right]t^n$$
$$= \sum_{n\geq 0}\left[\left(n+\frac{1}{2}\right)^2 a_n - (n+1)^2 a_{n+1}\right]t^n.$$

Since this power series is zero, we get a recurrence relation for the coefficients of $P(t)$:

$$a_{n+1} = \left(\frac{n+\frac{1}{2}}{n+1}\right)^2 a_n,$$

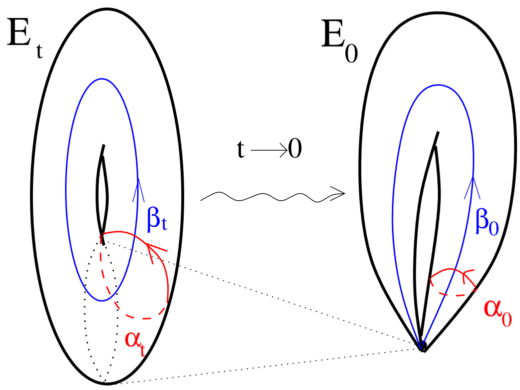
so that

$$a_n = \underbrace{a_0}_{=1}\cdot\left(\frac{1/2}{1}\frac{3/2}{2}\cdots\frac{1/2+n-1}{n}\right)^2 = \left(\frac{-1/2\cdot -3/2\cdot\cdots\cdot(-1/2-n+1)}{n!}\right)^2$$
$$= \binom{-1/2}{n}^2,$$

and

$$P(t) = 2\pi\sum_{n\geq 0}\binom{-1/2}{n}^2 t^n.$$

Again, the situation as $t \to 0$ looks like



In the next chapter the formula for $P(t)$ will be connected to counting rational points on cubics over $\mathbb{F}_p$.

Exercises

(1) Check that ${}_2F_1\left(\frac{1}{3}, \frac{2}{3}; 1; (3t)^3\right) = \sum_{m\geq 0} t^{3m} \cdot \frac{(3m)!}{(m!)^3}$ by writing out the Pochhammer symbols.

(2) Show that the curves $E_t := \{Z_0Z_1W_0W_1 - t(Z_1 - Z_0)^2(W_1 - W_0)^2 = 0\} \subset \mathbb{P}^1_{Z_0:Z_1} \times \mathbb{P}^1_{W_0:W_1}$ are in fact elliptic (except at those finitely many t — which ones? — for which they are singular). You could do this by projecting to the first $\mathbb{P}^1$ and using Riemann-Hurwitz to compute the genus. (To use R-H in this way, first find all "vertical tangents" – places on the curve where the partials with respect to W_0 and W_1 vanish.)

(3) Writing the family of curves from the last exercise in affine form, $xy - t(x-1)^2(y-1)^2 = 0$ (or $1 - t\varphi(x,y) = 0$), define a family of loops $\gamma_t \in H_1(E_t, \mathbb{Z})$ for small t, and a family of holomorphic forms $\omega_t \in \Omega^1(E_t)$, exactly as in the text. Compute the period $P(t) := \int_{\gamma_t} \omega_t$ as a power series, using the computation done above as a model.

(4) Check that $\tilde{\theta}$ in Remark 22.3.2 is well-defined and an isomorphism.

(5) Find $Q(t)$ in the Legendre example by plugging $\frac{\log t}{2\pi} P(t) + \sum_{n\geq 1} b_n t^n$ into the Picard-Fuchs equation.

CHAPTER 23

Counting $\mathbb{F}_p$-rational points on elliptic curves

In this final chapter on elliptic curves, we take a brief dip into something much more arithmetic, counting the number (mod p) of solutions in $\mathbb{P}^2(\mathbb{F}_p)$ to the equations for the Hesse and Legendre cubics from the last chapter. These cubics still depend on t, which is taken to be an integer now (rather than a complex number) so that we can reduce modulo p. In a truly bizarre twist, the number of points (over $\mathbb{F}_p$) in each case is given by nearly the same power series as the holomorphic period on the corresponding complex family of elliptic curves from Ch. 22. We briefly explain one abstract way, due to Y. MANIN, to understand this connection between arithmetic and transcendental algebraic geometry.[1]

23.1. Sum formulas

Let p be an odd prime, and $\mathbb{F}_p$ the field with p elements (i.e. $\mathbb{Z}/p\mathbb{Z}$ viewed as a ring). Equality in $\mathbb{F}_p$ will generally be denoted by "$=$", not "$\underset{(p)}{\equiv}$". (We will use the latter for counting points mod p.)

LEMMA 23.1.1. *For $k \in \mathbb{Z}$,*

$$\sum_{x\in\mathbb{F}_p\,(or\ \mathbb{F}_p^*)} x^k = \begin{cases} 0, & p-1 \nmid k \\ -1, & p-1 \mid k \end{cases}$$

in $\mathbb{F}_p$.

PROOF. Given $y \in \mathbb{F}_p^*$, the assignment $x \mapsto yx$ yields an isomorphism of additive groups $\mathbb{F}_p \to \mathbb{F}_p$. Therefore

$$y^k \sum_{x\in\mathbb{F}_p} x^k = \sum_{x\in\mathbb{F}_p} (xy)^k = \sum_{x\in\mathbb{F}_p} x^k. \tag{23.1.1}$$

[1]The interested reader is urged to study the theory of modular forms(!) and look at papers by Ahlgren, Ono, Papanikolas et al in particular. Modular forms are, on the one hand, essentially given by periods on families of (self-products of) elliptic curves; on the other hand, they are frequently constructed via point-counts over finite fields.

Now, $\mathbb{F}_p^*$ is a cyclic (multiplicative) group of order $p-1$, and so

$$(\mathbb{F}_p^*)^k = \{1\} \iff p-1 \mid k.$$

Provided $p-1 \nmid k$, then, there exists $y \in \mathbb{F}_p^*$ with $y^k \neq 1$. By (23.1.1), we have

$$0 = (y^k - 1)\sum_{x\in\mathbb{F}_p} x^k$$

which implies (dividing by $y^k - 1$)

$$0 = \sum_{x\in\mathbb{F}_p} x^k.$$

On the other hand, if $p-1 \mid k$, then $x^k = 1$ for all $x \in \mathbb{F}_p^*$ and so

$$\sum_{x\in\mathbb{F}_p} x^k \;=\; \sum_{x\in\mathbb{F}_p^*} 1 \;=\; p-1 \;=\; -1.$$

$\square$

LEMMA 23.1.2. *For $x \in \mathbb{F}_p$,*

$$\sum_{k=1}^{p-1} x^k = \begin{cases} 0, & x \neq 1 \\ -1, & x = 1 \end{cases}.$$

PROOF. If $x = 1$ then the sum is $\underbrace{1+\cdots+1}_{p-1 \text{ times}} = p-1 = -1$. If $x \neq 1$ then

$$\begin{aligned}\underbrace{(1-x)}_{\neq 0}\sum_{k=1}^{p-1} x^k \;&=\; \sum_{k=1}^{p-1} x^k - \sum_{k=2}^{p} x^k \\ &=\; x - x^p \\ &=\; x(1-x^{p-1})\end{aligned}$$

which (since $\mathbb{F}_p^*$ is cyclic of order $p-1$)

$$=\; x(1-1) \;=\; 0$$

$\square$

LEMMA 23.1.3. *Let $\xi \in \mathbb{F}_p^*$. Then $\xi^{\frac{p-1}{2}} \underset{(p)}{\equiv} \pm 1$, and*

$$\xi \in \mathbb{F}_p^2 \iff \xi^{\frac{p-1}{2}} \underset{(p)}{\equiv} 1.$$

PROOF. Since $\mathbb{F}_p^*$ is cyclic of order $p-1$, $(\xi^{\frac{p-1}{2}})^2 = \xi^{p-1} = 1$. Moreover, if a is a generator then we cannot have $a^{\frac{p-1}{2}} = 1$ ($\mathbb{F}_p^*$ would then have order $\frac{p-1}{2}$, a contradiction). Hence $x \mapsto x^{\frac{p-1}{2}}$ yields a surjective homomorphism of multiplicative groups

$$\mathbb{F}_p^* \underset{\theta}{\twoheadrightarrow} \{+1, -1\},$$

whose kernel necessarily has order $\frac{1}{2}|\mathbb{F}_p^*| = \frac{p-1}{2}$. Now if $\xi = \eta^2$ is a square, then $\xi^{\frac{p-1}{2}} = \eta^{p-1} = 1$. As $\frac{p-1}{2}$ elements of $\mathbb{F}_p^*$ are squares, these exhaust the kernel of θ and the non-square elements go to -1. $\square$

23.2. $\mathbb{F}_p$-points on the Legendre elliptic curve

Consider once again the Legendre family of cubics

$$E_t = \{Y^2 Z = X(X - Z)(X - tZ)\},$$

but this time with $t \in \mathbb{Z}$. After reducing mod p we can look at the solutions $E_t(\mathbb{F}_p) \subset E_t(\overline{\mathbb{F}_p})$, i.e. with X, Y, Z in $\mathbb{F}_p$ resp. its algebraic closure; there is a clear analogy to $E_t(\mathbb{Q}) \subset E_t(\bar{\mathbb{Q}})$.

We are going to compute the number of points $|E_t(\mathbb{F}_p)|$ modulo p, i.e. in $\mathbb{F}_p$. (Computing the number in $\mathbb{Z}$ is a much harder problem.) First we claim that

$$|E_t(\mathbb{F}_p)| \underset{(p)}{\equiv} 1 + \sum_{x \in \mathbb{F}_p} \left\{1 + [x(x-1)(x-t)]^{\frac{p-1}{2}}\right\}. \tag{23.2.1}$$

The leading "1" on the RHS counts the point $[0:0:1]$ "at ∞"; the rest of the curve is described by $y^2 = x(x-1)(x-t)$. By Lemma 23.1.3, the quantity in curly brackets yields (mod p) 2 if $x(x-1)(x-t)$ is a square, 1 if $x(x-1)(x-t) = 0$, and 0 if $x(x-1)(x-t)$ is not a square. This exactly counts pairs $(x, y) \in \mathbb{F}_p^2$ solving the affine equation, confirming (23.2.1).

Now (summing in $\mathbb{F}_p$) $\sum_{x \in \mathbb{F}_p} 1 = 0$, so the RHS of (23.2.1) is

$$1 + \sum_{x \in \mathbb{F}_p} x^{\frac{p-1}{2}} (x-1)^{\frac{p-1}{2}} (x-t)^{\frac{p-1}{2}}$$

$$= 1 + \sum_{x \in \mathbb{F}_p} x^{\frac{p-1}{2}} \left\{ \sum_{\ell=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{\ell} x^{\frac{p-1}{2} - \ell} (-t)^\ell \right\} \left\{ \sum_{k=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{k} x^k (-1)^{\frac{p-1}{2} - k} \right\}$$

$$= 1 + \sum_{x \in \mathbb{F}_p} x^{p-1} \left\{ \sum_{\ell=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{\ell} x^{-\ell} (-t)^\ell \right\} \left\{ \sum_{k=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{k} x^k (-1)^{\frac{p-1}{2} - k} \right\}.$$

The sum here can be rewritten $\sum_{x\in\mathbb{F}_p} x^{p-1}G(x)$, where $G(x) = \sum_{m=-\frac{p-1}{2}}^{\frac{p-1}{2}} a_m x^m$. By Lemma 23.1.1, this is just $-a_0$, and so the above

$$\begin{aligned}
&= 1 - \left[\left\{\sum_{\ell=0}^{\frac{p-1}{2}} \binom{p-1/2}{\ell} x^{-\ell}(-t)^{\ell}\right\}\left\{\sum_{k=0}^{\frac{p-1}{2}} \binom{p-1/2}{k} x^{k}(-1)^{\frac{p-1}{2}-k}\right\}\right]_0 \\
&= 1 - \sum_{\ell=0}^{\frac{p-1}{2}} (-t)^{\ell}(-1)^{\frac{p-1}{2}-\ell}\binom{p-1/2}{\ell}^2 \\
&= 1 - (-1)^{\frac{p+1}{2}} \sum_{\ell=0}^{\frac{p-1}{2}} \binom{p-1/2}{\ell}^2 t^{\ell} \\
&= 1 + (-1)^{\frac{p+1}{2}} \sum_{\ell\geq 0} \binom{-1/2}{\ell} t^{\ell} \\
&=: \hat{P}(t).
\end{aligned}$$

For the last step, we use the definition (in $\mathbb{F}_p$)

$$\binom{-1/2}{\ell} := \frac{-1/2 \cdot -3/2 \cdot \cdots \cdot (-1/2 - \ell + 1)}{\ell!},$$

which is evidently 0 when $p > \ell > \frac{p-1}{2}$ (since $\frac{-1}{2} - \frac{p+1}{2} + 1 = \frac{-1-p-1+2}{2} = \frac{p}{2} = 0$), and equals $\binom{\frac{p-1}{2}}{\ell}$ for $0 \leq \ell \leq \frac{p-1}{2}$ (and is *defined* to be 0 for $\ell \geq p$). We conclude:

PROPOSITION 23.2.1. *$\hat{P}(t)$ counts (mod p) the $\mathbb{F}_p$-rational points of E_t.*

Notice that

(23.2.2) $\hat{P}(t)$ is a "mod p" version of the period $P(t)$ from §22.6!

23.3. $\mathbb{F}_p$-rational points on the Hesse cubic

For this section, take p to be an odd prime with $p \equiv -1 \bmod 3$.

I can't resist doing the same exercise for the other main example from the last chapter, namely

$$E_t = \{XYZ = t(X^3 + Y^3 + Z^3)\}$$

where again we assume $t \in \mathbb{Z}$. This has affine form

$$xy = t(x^3 + y^3 + 1)$$

and toric form

$$1 = t\underbrace{(x^2y^{-1} + x^{-1}y^2 + x^{-1}y^{-1})}_{=:\varphi(x,y)},$$

where the Laurent polynomial $\varphi(x,y)$ is defined for $(x,y) \in (\mathbb{F}_p^*)^2$. Using the toric form and Lemma 23.1.2, it is easy to compute the $\mathbb{F}_p^*$-points

$$E_t^* := E_t(\mathbb{F}_p) \cap (\mathbb{F}_p^*)^2.$$

Namely, in $\mathbb{F}_p$ (i.e. mod p)

$$|E_t^*| \;=\; -\sum_{(x,y)\in(\mathbb{F}_p^*)^2} \sum_{k=1}^{p-1} t^k \left(\varphi(x,y)\right)^k,$$

the point being (besides the Lemma) that $t\varphi(x,y)$ is 1 (in $\mathbb{F}_p$) for exactly those (x,y) on E_t. Switching the order of summation this becomes

$$= \; -\sum_{k=1}^{p-1} t^k \sum_{(x,y)\in(\mathbb{F}_p^*)^2} \varphi(x,y)^k. \tag{23.3.1}$$

Now by Lemma 23.1.1

$$\sum_{(x,y)\in(\mathbb{F}_p^*)^2} x^i y^j \;=\; \left(\sum_{x\in\mathbb{F}_p^*} x^i\right)\left(\sum_{y\in\mathbb{F}_p^*} y^j\right) \;=\; \begin{cases} 1, & p-1 \mid i,j \\ 0, & \text{otherwise} \end{cases}.$$

For $k \in [1, p-2]$,

$$(\varphi(x,y))^k = [\varphi]_0 + \left\{ \begin{array}{c} \text{terms with powers of } x,y \\ \text{not both divisible by } p-1 \end{array} \right\}.$$

Our assumption on p implies that $3 \nmid p-1$, and so

$$\begin{aligned}(\varphi(x,y))^{p-1} = [\varphi^{p-1}]_0 + x^{2(p-1)}y^{-(p-1)} + x^{-(p-1)}y^{2(p-1)} + x^{-(p-1)}y^{-(p-1)} \\ + \left\{ \begin{array}{c} \text{terms with powers of } x,y \\ \text{not both divisible by } p-1 \end{array} \right\}\end{aligned}$$

(in particular, there are no $x^{-(p-1)}$, $y^{-(p-1)}$, $x^{p-1}y^{-(p-1)}$, $x^{-(p-1)}y^{p-1}$, x^{p-1} or y^{p-1} terms). So (23.3.1) becomes

$$-\sum_{k=1}^{p-1} t^k[\varphi^k]_0 \;-\; t^{p-1}\cdot 3.$$

Recall from §22.2 that $[\varphi^k]_0 = \binom{3m}{m,m,m}$ if $k = 3m$ (and 0 if $3 \nmid k$).

On the other hand, looking along the coordinate axes $X = 0$, $Y = 0$, $Z = 0$ we get (only) the points

$$[1 : -1 : 0]\,, \ \ [0 : 1 : -1]\,, \ \ [-1 : 0 : 1]$$

in $E_t(\mathbb{F}_p)$. For example, along $Z = 0$ (on E_t) we must have $X, Y \neq 0$ and so may assume $Y = 1$; then the equation is $X^3 + 1 = 0$. This has only $X = -1$ as solution: otherwise we would have an element of order 6 in $\mathbb{F}_p^*$, so $6 \mid p - 1$, contradicting our assumption on p.

We conclude that

$$|E_t(\mathbb{F}_p)| \underset{(p)}{\equiv} 3(1 - t^{p-1}) - \sum_{m=1}^{\lfloor \frac{p-1}{3} \rfloor} \binom{3m}{m, m, m} t^{3m},$$

again very reminiscent of $P(t)$ from before!

23.4. Deep reasons for (23.2.2)

The issue is this: in §23.2, why on earth does $|E_t(\mathbb{F}_p)| - 1$ (not counting the point at ∞) appear to solve the Picard-Fuchs equation $\left(t(t-1)D_t^2 + (2t-1)D_t + \frac{1}{4}\right)(\cdot) = 0$? Indeed, $\hat{P}(t) - 1 = \frac{(-1)^{\frac{p+1}{2}}}{2\pi} P(t)$, where $P(t)$ is the solution from §22.6!! The two computations were quite elementary, after all, so maybe there's an elementary explanation for their equivalence?

Nope. This is dealt with in [Clemens, "A scrapbook of complex curve theory," pp.65-69] and I'll just give a hint of the flavor here. It involves an algebro-geometric version of the Lefschetz trace formula (the formula from topology for the number of fixedmm points of a mapping), the Riemann-Roch theorem, Serre duality, and abstract sheaf theory. Not elementary, but we can give a brief summary (with no claim to total accuracy).

Consider E_t over $\overline{\mathbb{F}_p}$, $t \in \mathbb{F}_p$. Then writing FP for "number of fixed points", and $frob_p$ for the map $[Z : X : Y] \mapsto [Z^p : X^p : Y^p]$,

$$|E_t(\mathbb{F}_p)| \ = \ \mathrm{FP}\left\{frob_p :\ E_t(\overline{\mathbb{F}_p}) \to E_t(\overline{\mathbb{F}_p})\right\}.$$

This should make sense to you because as an automorphism of $\overline{\mathbb{F}_p}$, the p^{th}-power (Frobenius) map fixes exactly the elements of $\mathbb{F}_p$. By the Lefschetz-type theorem, it turns out that this

$$= \ 1 - \mathrm{trace}\left\{frob^*|_{H^1(E_t/\overline{\mathbb{F}_p}, \mathcal{O})}\right\}$$

where the H^1 is Cech cohomology computed with respect to the Zariski topology, $\mathcal{O}$ is the sheaf of regular functions, and $frob^*$ is the action by pullback (under $frob$) on cohomology classes. It is a 1-dimensional vector space, with generator represented by a certain rational function h with two simple poles, at $q = [1 : 0 : 0]$ and some other point $p \in E_t(\mathbb{F}_p)$. More precisely, $H^1(E_t/\overline{\mathbb{F}_p}, \mathcal{O})$ is isomorphic to the space of rational functions on E_t with poles allowed only at p and q modulo the subspace of rational functions with poles allowed at *either* p *or* q (not both). [That this space is 1-dimensional in the more familiar *complex* case is an exercise below.] You should also note that pulling back by $frob_p$ stabilizes the vector space we have just described, since (as p, q are taken to be in $E_t(\mathbb{F}_p)$ rather than $E_t(\overline{\mathbb{F}_p})$) p and q are fixed under $frob_p$. So the displayed expression at least makes sense.

Next, we expand h in formal power series $h = \frac{1}{y} + \sum_{\ell \geq 0} b_\ell y^\ell$ about q,[2] and also expand a generator $\omega_t \in \Omega^1(E_t/\overline{\mathbb{F}_p})$ (regular differentials) as $[1 + \sum_{k \geq 1} a_k(t) y^k]dy$. Recall also from the complex case, that residues of meromorphic functions require, and depend on, a choice of local coordinate; while residues of meromorphic 1-forms are invariant (i.e. require no such choice, as they can already be integrated around a loop without appending a "dz"). So for functions F with a pole at q, we take residue by computing $Res_q(F\omega)$; if F has no *other* pole, then (as residues sum to zero) the residue has to be zero.

Now, writing τ for the trace of $frob^*$ above, we have $frob^* h (= h \circ frob) = \frac{1}{y^p} + \sum_{\ell \geq 0} b_\ell y^{\ell p} = \tau h + f + g$. (The last equality, in which f has only a pole at q and g has only a pole at p, is by 1-dimensionality of $H^1(E_t, \mathcal{O})$ and the "explicit description" we gave of it. In that vector space, this reads $frob^*[h] = \tau[h]$.) Moreover, $Res_q(h\omega_t) = 1$ while

$$\tau = \tau Res_q(h\omega_t) + Res_q(f\omega_t) + Res_q(g\omega_t) = Res_q((frob^* h) \cdot \omega_t) = a_{p-1}(t),$$

with the last equality obtained by multiplying out the explicit expressions for $frob^* h$ and ω_t. So we end up with

$$|E_t(\mathbb{F}_p)| = 1 - a_{p-1}(t),$$

[2] note that y gives a local coordinate about $[1 : 0 : 0]$ on E_t; x does not.

and (like the periods of ω_t) $a_{p-1}(t)$ must satisfy the Picard-Fuchs equation because $[\omega_t]$ does. Again, the "regular" solution of $D_{\mathrm{PF}}(\cdot) = 0$ is unique up to scale, and from there we are essentially done.

Exercises

(1) Check that $E_t(\overline{\mathbb{F}_p})$ is closed under $frob_p$, for $t \in \mathbb{F}_p$.

(2) Let $E \subset \mathbb{P}^2$ be a smooth cubic over the complex numbers, and $p, q \in E(\mathbb{C})$ two distinct points. Let V be the vector space of meromorphic functions on E with poles only at p and q, with subspaces W_p and W_q (the meromorphic functions with poles only at p and q respectively). Using Abel's theorem, prove that the dimension of $V/(W_p + W_q)$ is one. [Hint: you will also need to use the fact that $\omega \in \Omega^1(E)$ has no zeroes, and that the "residues" of $F \in V$ given by $Res_p(F\omega)$ and $Res_q(F\omega)$ must sum to zero (cf. Prop. 13.1.6(b)).]

Part 4

Curves of higher genus

CHAPTER 24

The algebraicity of global analytic objects

To kick off the last part of this course, on curves of higher genus, this Chapter will demonstrate two approaches to the following result: *meromorphic (or holomorphic) functions and forms on normalizations of algebraic curves, all arise as pullbacks of functions and forms on projective space constructed from rational functions (quotients of homogeneous polynomials) and their differentials.* This is a special instance of Serre's GAGA principle ("global analytic is global algebraic" in the projective setting), and is proved (in §24.1) using techniques from Chapter 8 together with the primitive element theorem.[1]

For holomorphic forms, we would like a more precise result (already hinted at in Remark 20.1.1) on how to think of the holomorphic forms on a normalization "rationally". It is important at this point to recall part (B) of the Normalization Theorem 3.2.1, which says that *every Riemann surface can be obtained as the normalization of an algebraic curve in* $\mathbb{P}^2$, even one with only ordinary double point (ODP) singularities. So in the course of analyzing curves with ODP's in §24.2 we will actually have proved (cf. Prop. 24.2.1(c)) that for any Riemann surface M of genus g, $\dim(\Omega^1(M)) = g$. Featuring prominently in this section is the space of homogeneous polynomials vanishing at the set of ODP's, which will play a key role in the proof of Riemann-Roch in the next Chapter.

24.1. Chow's theorem for algebraic curves

Let $C \subset \mathbb{P}^2$ be an irreducible projective algebraic curve of degree d; applying a projective transformation if necessary we have $[0:0:1] \notin C$. Start by normalizing C; that is, express it as the image of a morphism $\sigma : \widetilde{C} \to (\mathbb{P}^2 \backslash \{[0:0:1]\})$. One may evidently produce

[1] What I haven't proved here, is that a global analytic subvariety of projective space is an algebraic subvariety; but I will use this at a couple of points in later chapters (apologies).

meromorphic functions on the Riemann surface $\widetilde{C}$, by pulling back the rational functions $\mathbb{C}(x,y)$ under σ^*. (Recall $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$ on $\mathbb{P}^2$; and the field $\mathbb{C}(x,y)$ consists of all quotients of homogeneous polynomials in X, Y, Z of the same degree.) More precisely, writing $\mathbb{C}(x,y)_C$ for the subring of rational functions whose polar set does not contain C, define the field of rational functions on $\widetilde{C}$ by

$$\mathbb{C}(\widetilde{C}) := \sigma^*\mathbb{C}(x,y)_C.$$

Next we consider the projection

$$\pi : \left(\mathbb{P}^2 \backslash \{[0:0:1]\}\right) \twoheadrightarrow \mathbb{P}^1$$
$$[Z:X:Y] \mapsto [Z:X],$$

whose composition $\tilde{\pi} := \pi \circ \sigma$ with the normalization presents $\widetilde{C}$ as a d-sheeted[2] branched cover of $\mathbb{P}^1$. Write $\mathfrak{B}\,(\subset \mathbb{P}^1)$ for the branch locus, and Γ for a path containing $\mathfrak{B}$ with $\mathbb{P}^1 \backslash \Gamma$ simply connected (cf. §8.2). We have inclusions

$$\begin{array}{ccccc} \tilde{\pi}^*\mathbb{C}(x) & \subset & \mathbb{C}(\widetilde{C}) & \subset & \mathcal{K}(\tilde{C}) \\ & & \text{rat'l} & & \text{mero.} \\ & & \text{fcns.} & & \text{fcns.} \end{array}, \tag{24.1.1}$$

where the first is obtained by noting $\pi^*\mathbb{C}(x) \subset \mathbb{C}(x,y)_C$ and $\tilde{\pi}^*\mathbb{C}(x) = \sigma^*\pi^*\mathbb{C}(x) \subset \sigma^*\mathbb{C}(x,y)_C$.

Now, one might initially speculate that the right-hand inclusion of (24.1.1) is proper when C has singularities such as ODP's, since: (a) an ODP has 2 preimage points $p, q \in \widetilde{C}$, (b) at first glance the pullback of a function would seem to have the same value at p and q, and (c) meromorphic functions on $\widetilde{C}$ ought to be able to take different values at distinct points, right? The weak link in this chain of reasoning is (b), as you can see from the following

EXAMPLE 24.1.1. $C = \{Y^2Z = X^2(X-Z)\}$ has tangent lies $Y = \pm X$ at its ODP $[1:0:0]$. The pullback of $\frac{x}{y}$ to $\widetilde{C}$ therefore takes values 1 and -1 (resp.) at the 2 points lying over the ODP.

The point is that rational functions are not well-defined at all points of $\mathbb{P}^2$, and this can be used to our advantage to get "more" functions on singular curves. So it becomes plausible that the right-hand inclusion

[2] the mapping degree $\deg(\tilde{\pi}) = d$ $(= \deg(C))$ by Bézout

of (24.1.1) is an equality, and that is exactly what we shall prove in the rest of the section.

To that end, let $\varphi \in \mathcal{K}(\widetilde{C})^*$ be a nonzero meromorphic function, and denote by P the set of poles of φ. Writing[3]

$$0 = f(x, y) = y^d + a_1(x)y^{d-1} + \cdots + a_d(x) \tag{24.1.2}$$

for the affine equation of C, we have as in §8.2 distinct solutions $\{y_j(x)\}_{j=1}^d$ to $f(x, \cdot) = 0$ over $\mathbb{P}^1\backslash\Gamma$, which are interchanged as one passes through $\Gamma\backslash\mathfrak{B}$. Moreover, by irreducibility of C (hence f), (24.1.2) is the minimal polynomial of $\sigma^* y$, proving that

$$[\mathbb{C}(\widetilde{C}) : \tilde{\pi}^*\mathbb{C}(x)] \geq d. \tag{24.1.3}$$

For each $x \in \mathbb{P}^1\backslash(\Gamma \cup \tilde{\pi}(P))$, one can think of $(x, y_j(x))$ as belonging to $\widetilde{C}$ with $\tilde{\pi}(x, y_j(x)) = x$. Consider the elementary symmetric polynomials ($i = 0, \ldots, d$, with $e_0^\varphi = 1$)

$$e_i^\varphi(x) := e_i\left(\varphi(x, y_1(x)), \ldots, \varphi(x, y_d(x))\right),$$

which are well-defined and holomorphic on $\mathbb{P}^1\backslash(\mathfrak{B} \cup \tilde{\pi}(P))$. As in §8.2, the fact that they are bounded away from $\tilde{\pi}(P)$ guarantees (by Riemann) their extension to holomorphic functions on $\mathbb{P}^1\backslash\tilde{\pi}(P)$. Further, if $x_0 \in \tilde{\pi}(P)$ has neighborhood $\Delta_{x_0} \subset \mathbb{P}^1$, then for $k \in \mathbb{N}$ sufficiently large, $\hat{\varphi} := \tilde{\pi}^*((x - x_0)^k) \cdot \varphi$ is holomorphic in $\tilde{\pi}^{-1}(\Delta_{x_0}) \subset \widetilde{C}$. By the same argument (from §8.2), $e_i^{\hat{\varphi}}(x)$ extends holomorphically across x_0; but since $e_i^{\hat{\varphi}}(x) = (x - x_0)^{ik} \cdot e_i^\varphi(x)$, $e_i^\varphi(x)$ extends *meromorphically* across x_0. Repeating this argument at all points of $\tilde{\pi}(P)$, we find that

$$e_i^\varphi \in \mathcal{K}(\mathbb{P}^1);$$

and by Theorem 3.1.5(a) $\mathcal{K}(\mathbb{P}^1) \cong \mathbb{C}(x)$.

Next observe that for any $x \in \mathbb{P}^1\backslash(\Gamma \cup \tilde{\pi}(P))$ and $j \in \{1, \ldots, d\}$,

$$0 \;=\; \prod_{i=1}^d \left(\varphi(x, y_j(x)) - \varphi(x, y_i(x))\right)$$

$$=\; \varphi(x, y_j(x))^d - e_1^\varphi(x)\varphi(x, y_j(x))^{d-1} + e_2^\varphi(x)\varphi(x, y_j(x))^{d-2} - \cdots + (-1)^d e_d^\varphi(x)\,;$$

[3] as in §8.2 we may change projective coordinates if necessary to put the equation in this form

that is, for a dense subset of points $p \in \widetilde{C}$, $\varphi(p)$ satisfies the equation

$$0 \;=\; \sum_{i=0}^{d}(-1)^i e_i^{\varphi}(\tilde{\pi}(p)) \cdot \varphi(p)^{d-i}.$$

Therefore the meromorphic function φ itself satisfies

$$0 \;=\; \sum_{i=0}^{d}(-1)^i (\tilde{\pi}^* e_i^{\varphi}) \cdot \varphi^{d-i}, \tag{24.1.4}$$

with coefficients in $\tilde{\pi}^*\mathbb{C}(x)$.

Finally, recall the *primitive element theorem*, which says that an algebraic field extension (of degree e) is generated by a single element (of degree e). (A transcendental field extension, likewise, has a transcendental element, which satisfies no algebraic equation.) Were $[\mathcal{K}(\widetilde{C}) : \tilde{\pi}^*\mathbb{C}(x)] > d$, there would thus be an element of degree $> d$; but as $\varphi \in \mathcal{K}(\widetilde{C})$ was arbitrary, (24.1.4) shows this is not so. hence

$$[\mathcal{K}(\widetilde{C}) : \tilde{\pi}^*\mathbb{C}(x)] \leq d.$$

Putting this together with (24.1.1) and (24.1.3), we see that

$$\mathcal{K}(\widetilde{C}) = \mathbb{C}(\widetilde{C}),$$

proving the

THEOREM 24.1.2. *Every meromorphic function on the normalization of an irreducible projective algebraic curve is rational, i.e. the pullback of a ratio of homogeneous polynomials.*

COROLLARY 24.1.3. *Every meromorphic 1-form on a normalization is rational (i.e. $f\,dg$ where f, g are rational).*

PROOF. Consider (say) $\sigma^*(dx) =: \omega \in \mathcal{K}^1(\widetilde{C})$, and let $\omega' \in \mathcal{K}^1(\widetilde{C})$ be any other meromorphic 1-form on $\widetilde{C}$. Then $\frac{\omega'}{\omega}$ belongs to $\mathcal{K}(\widetilde{C})$, hence is rational by Theorem 24.1.2. □

24.2. Cohomology of a Riemann surface

Let M be a Riemann surface of genus g. Recall from §20.1 that the 1st homology group $H^1(M, \mathbb{Z}) = \frac{\text{closed loops}}{\text{boundaries}}$ is an abelian group of rank $2g$, and define M's *1st cohomology group* to be the $2g$-dimensional vector space of complex-linear functionals

$$H^1(M, \mathbb{C}) := Hom(H_1(M, \mathbb{Z}), \mathbb{C}).$$

Exactly as in §22.3 (for elliptic curves) we have the *de Rham cohomology groups*

$$H^1_{dR}(M) := \frac{\ker\{A^1(M) \xrightarrow{d} A^2(M)\}}{\text{image}\{A^0(M) \xrightarrow{d} A^1(M)\}} = \frac{\text{closed } C^\infty \text{ 1-forms}}{\text{exact } C^\infty \text{ 1-forms}}.$$

To any closed 1-form ω we may assign the functional $\gamma \mapsto \int_\gamma \omega$ on loops. By the first 2 paragraphs of the proof of Lemma 22.3.1 (which work for any M), this induces a well-defined injective map

$$H^1_{dR}(M) \hookrightarrow H^1(M, \mathbb{C}). \tag{24.2.1}$$

Surjectivity also holds but will require a little more work than for elliptic curves.

Writing $\overline{\Omega^1(M)}$ for the space of "anti-holomorphic" forms (the complex conjugates of holomorphic ones), we can embed

$$\Omega^1(M) \oplus \overline{\Omega^1(M)} \hookrightarrow H^1_{dR}(M) \tag{24.2.2}$$

via

$$(\omega, \overline{\varphi}) \mapsto [\omega + \overline{\varphi}].$$

The map (24.2.1) is well-defined because $d(\Omega^1(M)) = 0 = d(\overline{\Omega^1(M)})$ (cf. Remark 22.3.2(a)). To prove injectivity, suppose $\omega + \overline{\varphi} = df$, $f \in A^0(M)$. Then

$$d(f\varphi) = f \underbrace{d\varphi}_{=0} + df \wedge \varphi = (\omega + \overline{\varphi}) \wedge \varphi = -\varphi \wedge \overline{\varphi}$$

since $\omega \wedge \varphi$ looks locally like a function times $dz \wedge dz (= 0)$. Now breaking M up into triangular regions Δ_i with local holomorphic coordinates $z_i = x_i + \sqrt{-1} y_i$,

$$\int_M \varphi \wedge \overline{\varphi} = \sum_i \int_{\Delta_i} g_i dz_i \wedge \overline{g_i}\overline{dz_i} = -2\sqrt{-1} \sum \underbrace{\int_{\Delta_i} |g_i|^2 dx_i \wedge dy_i}_{\in \mathbb{R}_{\geq 0}}.$$

Since each integral $= 0 \iff g_i \equiv 0$, we have

$$\int_M \varphi \wedge \overline{\varphi} = 0 \quad \iff \quad \varphi \equiv 0.$$

But using Stokes's theorem and $\partial M = \emptyset$,

$$\int_M \varphi \wedge \overline{\varphi} = -\int_M d(f\varphi) = -\int_{\partial M} f\varphi = 0$$

which implies $\varphi \equiv 0$. So $df = \omega \implies \frac{\partial f}{\partial \bar{z}} = 0 \implies f \in \mathcal{O}(M) \implies f$ constant (by Liouville) $\implies \omega = 0$. So (24.2.1) is injective.

By now you are quite familiar with the fact that $\dim(S_3^{d-3}) = \binom{(d-3)+2}{2} = \frac{(d-1)(d-2)}{2}$. If $\mathcal{S}$ is a set of δ points in $\mathbb{P}^2$, then the homogeneous polynomials of degree $d-3$ vanishing on each of these points are subject to δ (possibly dependent) linear conditions. Denoting the space of such polynomials by $S_3^{d-3}(-\mathcal{S})$, we therefore have

$$\dim(S_3^{d-3}(-\mathcal{S})) \geq \frac{(d-1)(d-2)}{2} - \delta. \tag{24.2.3}$$

Now assume $\sigma : M \to \mathbb{P}^2$ is injective off a finite point set, with image an algebraic curve $C = \{F(Z, X, Y) = 0\}$ of degree d ($F \in S_3^d$), having only ODP singularities (as in part (B) of the Normalization Theorem). Write $\mathcal{S}$ for the collection of these ODP's, and note $M = \widetilde{C}$. By (24.2.3) and the genus formula,

$$\dim(S_3^{d-3}(-\mathcal{S})) \geq g.$$

By Remark 20.1.1,[4] we have a map

$$S_3^{d-3}(-\mathcal{S}) \to \Omega^1(M)$$

given by

$$G \longmapsto \sigma^* \left(\frac{g\, dx}{f_y} \right)$$

where $g(x, y) = G(1, x, y)$ etc. This is necessarily injective: were $\frac{g dx}{f_y}$ to vanish on C, we would have $G \equiv 0$ on C; since F is irreducible then $F|G$ by Study, which is impossible (unless G is trivial) as $\deg(G) = d - 3 < d = \deg(F)$.

All told, we have a sequence of injective maps of complex vector spaces

$$S_3^{d-3}(-\mathcal{S}) \oplus S_3^{d-3}(-\mathcal{S}) \hookrightarrow \Omega^1(M) \oplus \overline{\Omega^1(M)} \hookrightarrow H^1_{dR}(M) \hookrightarrow H^1(M, \mathbb{C}).$$

Notice that the left-hand side has dimension $\geq 2g$ and the right-hand side has dimension exactly $2g$. All the injections are therefore isomorphisms and we conclude:

Proposition 24.2.1. *For a Riemann surface M (of genus g) normalizing an algebraic curve of degree d with ordinary double points*

[4] see Griffiths's book "Introduction to algebraic curves" for more details

$\mathcal{S} = \{p_1, \ldots, p_\delta\}$, *the holomorphic forms are all pullbacks of rational forms of the form* $\frac{g dx}{f_y}$ *(as described above). Moreover, we have:*

(a) [DE RHAM THEOREM] $H^1(M, \mathbb{C}) \cong H^1_{dR}(M)$*;*

(b) [HODGE DECOMPOSITION] $H^1_{dR}(M) \cong \Omega^1(M) \oplus \overline{\Omega^1(M)}$*; and*

(c) $\dim \Omega^1(M) = g = \frac{(d-1)(d-2)}{2} - \delta = \dim(S_3^{d-3}(-\mathcal{S}))$.

We also get an application to the period matrices Π described in §20.1. Recall that if $\gamma_1, \ldots, \gamma_{2g}$ is a basis for $H_1(M, \mathbb{Z})$ and $\omega_1, \ldots, \omega_g$ a basis for $\Omega^1(M)$, then

$$\Pi = \begin{pmatrix} \int_{\gamma_1} \omega_1 & \cdots & \cdots & \int_{\gamma_{2g}} \omega_1 \\ \vdots & \ddots & \ddots & \vdots \\ \int_{\gamma_1} \omega_g & \cdots & \cdots & \int_{\gamma_{2g}} \omega_g \end{pmatrix}.$$

PROPOSITION 24.2.2. *Viewed as vectors in* $\mathbb{R}^{2g} (\cong \mathbb{C}^g)$*, the columns* π_j *of* Π *are* $\mathbb{R}$*-linearly independent.*

PROOF. Suppose otherwise, i.e. that there exists a nonzero vector $\underline{a} \in \mathbb{R}^{2g}$ satisfying

$$\underline{0} = \Pi \underline{a}\,;$$

then we have also (by complex conjugating)

$$\underline{0} = \overline{\Pi} \underline{a}.$$

That is,

$$\begin{pmatrix} \Pi \\ \overline{\Pi} \end{pmatrix} \underline{a} = \underline{0}$$

and so the rank of $\begin{pmatrix} \Pi \\ \overline{\Pi} \end{pmatrix}$ is less than $2g$. But then there is a nonzero $\underline{b} \in \mathbb{C}^{2g}$ such that

$${}^t\underline{b} \begin{pmatrix} \Pi \\ \overline{\Pi} \end{pmatrix} = {}^t\underline{0}\,,$$

which means explicitly for each j that

$$\int_{\gamma_j} (\underbrace{\sum_{i=1}^g b_i \omega_i}_{=:\omega} + \underbrace{\sum_{i=1}^g b_{g+1} \overline{\omega_i}}_{=:\overline{\varphi}}) \;=\; 0.$$

We find then that $(\omega, \overline{\varphi}) \in \Omega^1(M) \oplus \overline{\Omega^1(M)}$ goes to zero in $Hom(H_1(M, \mathbb{Z}), \mathbb{C}) = H^1(M, \mathbb{C})$. By our sequence of injections above, $\omega = \varphi = 0$. But since $\underline{b} \neq \underline{0}$, this contradicts linear independence of $\omega_1, \ldots, \omega_g$ in $\Omega^1(M)$. $\square$

Exercises

(1) Write a basis for the holomorphic 1-forms on the (smooth) curve $C \subset \mathbb{P}^2$ with affine equation $1 + x^6 + y^6 - xy^5 = 0$. What is $\dim(\Omega^1(C))$?

CHAPTER 25

The Riemann-Roch Theorem

As you know, there are no nonconstant holomorphic functions on a Riemann surface M. What if we allow a simple pole at one point p but no poles anywhere else? Then you still get nothing, *unless* M is $\mathbb{P}^1$ (in which case there is $(z-z(p))^{-1}$). This is because for $g = \text{genus}(M) \geq 1$, there is a nonzero holomorphic form ω which doesn't vanish at p. For any meromorphic function f on M, we know that $\sum_{q \in M} Res_q(f\omega) = 0$; so if f has a simple pole at p, then $Res_p(f\omega) \neq 0$ and f must have another pole to cancel this term.

What if we are prepared to allow a double pole at p (but still no other poles)? Then the answer is more complex; if $g = 0$ or 1 there *are* nonconstant such functions (e.g. the Weierstrass $\wp$-function), while if $g \geq 2$ it can depend on the point p. In general, the vector spaces of functions $f \in \mathcal{K}(M)$ with a single pole (at p) with $\nu_p(f) \geq -k$ has dimension $\geq \max\{1, k-g+1\}$. You are guaranteed to get something nonconstant as soon as $k - g + 1 \geq 2$.

In the 1850's, Riemann proved a more general inequality which replaces p (and k) by multiple points and orders; a decade later, his student Roch turned this into an exact equality (Theorem 25.2.3 below) incorporating another term related to meromorphic 1-forms. It encompasses the equality $\dim(\Omega^1(M)) = g$ and gives a powerful tool for studying embeddings of Riemann surfaces into higher dimensional projective spaces, among other things. Its statement is in terms of spaces of functions and forms related to divisors, and we will start (§25.1) by defining these spaces precisely.

You may prefer this shorter introduction to the topic from a lecture by Lefschetz: "Well, a Riemann surface is a certain kind of Hausdorff space. You know what a Hausdorff space is, don't you? Its also compact, ok. I guess it is also a manifold. Surely you know what a manifold is. Now let me tell you one nontrivial theorem, the Riemann-Roch Theorem."[1]

25.1. Effective divisors and rational equivalence

Let M be a Riemann surface, $D = \sum_{p \in M} m_p[p]$ and $E = \sum_{p \in M} n_p[p]$ divisors on M. (Of course, only finitely many m_p and n_p are nonzero.) If for all p $m_p \geq n_p$, then we write $D \geq E$.

Definition 25.1.1. $D \in Div(M)$ is *effective* $\iff$ $D \geq 0$.

Example 25.1.2. The divisor (ω) of a holomorphic 1-form ω is effective. (Why?)

We can use this idea to put constraints on meromorphic functions and forms. For instance, suppose $D = 3[q] - 2[r]$, and $f \in \mathcal{K}(M)$ with divisor $(f) = \sum_{p \in M} \nu_p(f)[p]$. Then imposing the inequality $(f)+D \geq 0$ forces $\nu_q(f) + 3 \geq 0$ and $\nu_r(f) - 2 \geq 0$; that is, f is allowed a pole of order no worse than -3 at q, and *must* have a zero of order at least 2 at r. Likewise, if $\omega \in \mathcal{K}^1(M)$ then $(\omega) \geq D$ means ω has a zero of order at least 3 at q, and is allowed a pole of order no worse than -2 at r. The next definition formalizes this and defines the quantities which the Riemann-Roch theorem will relate.

Definition 25.1.3. For any $D \in Div(M)$,

$$\mathfrak{L}(D) := \{f \in \mathcal{K}(M)^* \mid (f) + D \geq 0\} \cup \{0\},$$

$$\mathfrak{I}(D) := \{\omega \in \mathcal{K}^1(M)^* \mid (\omega) \geq D\} \cup \{0\}.$$

(The "$\cup\{0\}$" just means that the zero-function is included, so as to produce a vector space.) Set

$$\ell(D) := \dim \mathfrak{L}(D) \ , \qquad i(D) := \dim \mathfrak{I}(D).$$

The next step is to define an equivalence relation on divisors which is ubiquitous in algebraic geometry.

Definition 25.1.4. Divisors $D, E \in Div(M)$ are rationally equivalent iff there exists[2] $f \in \mathcal{K}(M)^*$ with $(f) = D - E$; we write $D \overset{\text{rat}}{\equiv} E$.

[1] from *A Beautiful Mind* by S. Nasar

[2] by Chow's theorem all meromorphic functions are *rational*, hence the terminology "rational equivalence" (sometimes also called "linear equivalence").

PROPOSITION 25.1.5. *If* $D \overset{rat}{\equiv} E$, *then*
(i) $\deg(D) = \deg(E)$;
(ii) $\mathfrak{L}(D) \cong \mathfrak{L}(E)$;
(iii) $\mathfrak{I}(D) \cong \mathfrak{I}(E)$; *and*
(iv) $\ell(D) = \ell(E)$ *and* $i(D) = i(E)$.
Furthermore, $\overset{rat}{\equiv}$ *respects the abelian group structure of* $Div(M)$.

PROOF. By assumption $D - E = (f)$. Now Exercise 3.2 says that $\deg((f)) = 0$, which yields (i). Given $g \in \mathfrak{L}(D)$,

$$(fg) + E = (f) + (g) + E = (g) + D \geq 0\,;$$

so $g \mapsto fg$ defines a map $\mathfrak{L}(D) \to \mathfrak{L}(E)$, and $h \mapsto \frac{h}{f}$ defines an inverse map. This gives (ii), and (iii) is done in the same way. (iv) obviously follows from (ii)-(iii). The last statement about $\overset{\mathrm{rat}}{\equiv}$ is essentially just that $(D + (f)) + (E + (g)) = (D + E + (fg))$. □

REMARK 25.1.6. In fact, the Picard group $Pic(M)$ of §20.1 is just the group of equivalence classes

$$\frac{Div(M)}{\overset{\mathrm{rat}}{\equiv}}.$$

So Proposition 25.1.5(i,iv) can be thought of as saying that deg, ℓ,and i define functions from $Pic(M)$ to $\mathbb{Z}$.

DEFINITION 25.1.7. A *canonical divisor* $K \in Div(M)$ is just the divisor of any meromorphic 1-form $\omega \in \mathcal{K}^1(M)$. Since any two such are rationally equivalent (easy exercise), there is a single canonical divisor *class* $[K] \in Pic(M)$.

The next (basic) result is sometimes called "Brill-Noether reciprocity":

PROPOSITION 25.1.8. *Let* $D \in Div(M)$ *be arbitrary, and* K *a canonical divisor. Then*

$$\mathfrak{I}(D) \cong \mathfrak{L}(K - D),$$

and so $i(D) = \ell(K - D)$.

PROOF. Let $K = (\omega)$; if $(f) + K - D \geq 0$, then $(f\omega) = (f) + K \geq D - K + K = D$. So $f \mapsto f\omega$ maps $\mathfrak{L}(K - D) \to \mathfrak{I}(D)$, and $\eta \mapsto \frac{\eta}{\omega}$ gives an inverse. □

25.2. Proof and statement

Throughout this section we take C to be an irreducible degree d projective algebraic curve with ODP singularities $\mathcal{S} = \{p_1, \ldots, p_\delta\}$. Let $M := \widetilde{C} \xrightarrow{\sigma} \mathbb{P}^2$ ($\sigma(M) = C$) be its normalization, with $\sigma^{-1}(p_i) = \{q_i, r_i\}$, and define a divisor

$$\mathcal{E} := \sigma^{-1}(\mathcal{S}) = \sum_{i=1}^{\delta}[q_i] + [r_i] \ \in \ Div(M)$$

of degree 2δ. Given any line $H \subset \mathbb{P}^2$ ("H" for "hyperplane"), write

$$\mathcal{H} := \sigma^{-1}(H \cdot C) \ \in \ Div(M)$$

for the intersection divisor (of degree d).[3]

LEMMA 25.2.1. *For all sufficiently large $m \in \mathbb{N}$,*

$$\ell(m\mathcal{H} - \mathcal{E}) \geq md - 2\delta - g + 1$$

and

$$i(m\mathcal{H} - \mathcal{E}) = 0,$$

where $g = \frac{(d-1)(d-2)}{2} - \delta$ is the genus of M.

PROOF. Write $R \in S_3^1$ and $F \in S_3^d$ for the defining homogeneous polynomials of H and C (resp.). Consider the map

$$S_3^m(-\mathcal{S}) \xrightarrow{\theta} \mathfrak{L}(m\mathcal{H} - \mathcal{E})$$
$$G \longmapsto \sigma^*\left(\frac{G}{R^m}\right).$$

By Study's lemma,

$$\sigma^*\left(\frac{G}{R^m}\right) \equiv 0 \quad \Longleftrightarrow \quad G|_C \equiv 0 \quad \Longleftrightarrow \quad F|G,$$

and so $\ker\theta = F \cdot S_3^{m-d}$ $(\subset S_3^m(-\mathcal{S}))$.

Therefore, taking dimensions of

$$\mathfrak{L}(m\mathcal{H} - \mathcal{E}) \supseteq \mathrm{im}(\theta),$$

[3]If H passes through an ODP p_i, then $q_i$and r_i will both show up in $\mathcal{H}$, with multiplicities determined by the local intersection multiplicities of H with the two local analytic components of C at p_i. (See Def. 12.2.1 ff)

we find

$$\ell(m\mathcal{H}-\mathcal{E}) \geq \dim(\mathrm{im}(\theta)) = \dim\left(\frac{S_3^m(-\mathcal{S})}{F\cdot S_3^{m-d}(-\mathcal{S})}\right)$$
$$= \dim S_3^m(-\mathcal{S}) - \dim S_3^{m-d}.$$

By Prop. 24.2.1(c), this

$$= \frac{(m+1)(m+2)}{2} - \delta - \frac{(m-d+1)(m-d+2)}{2}$$
$$= dm - \delta - \frac{d(d-3)}{2}$$
$$= dm - \delta - \frac{(d-1)(d-2)}{2} + 1$$
$$= dm - 2\delta - g + 1.$$

□

LEMMA 25.2.2. *Let $D \in Div(M)$, $p \in M$. Then*

$$0 \;\leq\; \ell(D+[p]) - \ell(D) - (i(D+[p]) - i(D)) \;\leq\; 1.$$

PROOF. First note that $\mathfrak{L}(D) \subseteq \mathfrak{L}(D+[p]) \implies \ell(D+[p])-\ell(D) \geq 0$.

Next, writing $D = \sum_{q\in M} n_q[q]$, an element of $\mathfrak{L}(D+[p])\backslash\mathfrak{L}(D)$ is a function $f \in \mathcal{K}(M)^*$ satisfying

$$(f) + D + [p] \geq 0 \quad \text{and} \quad \nu_p(f) = -(n_p+1). \tag{25.2.1}$$

If f, g are two such functions, then setting $\alpha := \lim_{x\to p}\frac{f(x)}{g(x)}$, we have

$$\mathrm{ord}_p(f-\alpha g) \geq -n_p$$

so that $f-\alpha g \in \mathfrak{L}(D)$. So $\ell(D+[p]) - \ell(D) \leq 1$, and we conclude

$$0 \leq \ell(D+[p]) - \ell(D) \leq 1. \tag{25.2.2}$$

Similarly, writing K for a canonical divisor,

$$0 \leq \ell(K-D) - \ell(K-D-[p]) \leq 1$$

or equivalently

$$0 \leq i(D) - i(D+[p]) \leq 1. \tag{25.2.3}$$

Altogether,

$$0 \leq \ell(D+[p]) - \ell(D) + i(D) - i(D+[p]) \leq 2$$

and we just have to show that "2" is impossible.

Suppose (for a contradiction) that f satisfies (25.2.1), which is equivalent to "1" in (25.2.2), and $\omega \in \mathcal{K}^1(M)$ satisfies

$$(\omega) \geq D \quad \text{and} \quad \nu_p(\omega) = n_p,$$

which is equivalent to "1" in (25.2.3). Then

$$(f\omega) = (f) + (\omega) \geq -[p]$$

with

$$\nu_p(f\omega) = \nu_p(f) + \nu_p(\omega) = -1.$$

But the sum of residues of a meromorphic form must be zero (Prop. 13.1.6(b)), so $f\omega$ having a single simple pole (and no other poles) is absurd. $\square$

THEOREM 25.2.3. [RIEMANN-ROCH] *Let M be a Riemann surface of genus g, D a divisor on M. Then*

$$\ell(D) - i(D) = \deg(D) - g + 1.$$

PROOF. By part (B) of the Normalization Theorem 3.2.1, we can assume we are in the situation described in the beginning of the section, with $M = \widetilde{C}$.

By Lemma 25.2.1, there exists $m_0 \in \mathbb{Z}$ such that $m \geq m_0 \implies$

$$\ell(m\mathcal{H} - \mathcal{E}) - i(m\mathcal{H} - \mathcal{E}) \;\geq\; md - 2\delta - g + 1.$$

Now for any two lines H_1, H_2, we have $\mathcal{H}_1 \overset{\text{rat}}{\equiv} \mathcal{H}_2$; so if $H_1, \ldots, H_m$ are lines in $\mathbb{P}^2$ then by Proposition 25.1.5(iv)

$$\ell(\mathcal{H}_1 + \cdots + \mathcal{H}_m - \mathcal{E}) - i(\mathcal{H}_1 + \cdots + \mathcal{H}_m - \mathcal{E}) \;\geq\; md - 2\delta - g + 1.$$

Taking m large enough and lines through (a) all points of $\mathcal{S}$ and (b) all points in D, we can ensure that $\sum_{i=1}^m \mathcal{H}_i - \mathcal{E} - D$ is effective, so that

$$\mathcal{H}_1 + \cdots + \mathcal{H}_m - \mathcal{E} \;=\; D + [P_1] + \cdots + [P_k]$$

where $k = md - 2\delta - \deg(D)$ (and the P_j are points of M). Therefore we have

$$\ell\left(D + \sum_{j=1}^{k}[P_j]\right) - i\left(D + \sum_{j=1}^{k}[P_j]\right) \;\geq\; k + \deg(D) - g + 1.$$

Repeatedly applying the right-hand inequality of Lemma 25.2.2 gives

$$k + \ell(D) - i(D) \;\geq\; \ell\left(D + \sum_{j=1}^{k}[P_j]\right) - i\left(D + \sum_{j=1}^{k}[P_j]\right)$$

and we conclude that

$$\ell(D) - i(D) \;\geq\; \deg(D) - g + 1. \tag{25.2.4}$$

Next we show the reverse inequality. Plugging $K - D$ into (25.2.4), we have

$$\ell(K - D) - i(K - D) \;\geq\; \deg(K - D) - g + 1$$

which becomes (using Brill-Noether reciprocity)

$$i(D) - \ell(D) \;\geq\; 2g - 2 - \deg(D) - g + 1 \;=\; -(\deg(D) - g + 1),$$

so that

$$\ell(D) - i(D) \;\leq\; \deg(D) - g + 1.$$

□

Amongst the easy corollaries of this theorem are the *Riemann inequality*

$$\ell(D) \;\geq\; \deg(D) - g + 1,$$

and (by putting $D = 0$ in the theorem) the formula

$$\dim \Omega^1(M) \;=\; g.$$

Here is one more simple application:

PROPOSITION 25.2.4. *Up to isomorphism, $\mathbb{P}^1$ is the only Riemann surface of genus* 0.

PROOF. Suppose M has genus 0; then first of all the above corollary of Riemann-Roch says that $\dim \Omega^1(M) = 0$. If we take (for some $p \in M$) $D = [p]$, then $\mathfrak{I}(D) \subset \Omega^1(M) = \{0\} \implies i(D) = 0$. So by Riemann-Roch,

$$\ell(D) = \deg(D) - g + 1 = 1 - 0 + 1 = 2.$$

Now $\mathfrak{L}(D)$ consists of functions with a simple pole allowed at p (and no other poles). The constant function 1 belongs to $\mathfrak{L}(D)$; and since $\dim \mathfrak{L}(D) = 2$ there is also a nonconstant function $f \in \mathfrak{L}(D)$, which by Liouville *must* have the allowed simple pole. Therefore the mapping degree of $f : M \to \mathbb{P}^1$ is (cf. §14.1)

$$\deg(f) = \deg(f^{-1}([\infty])) = \deg([p]) = 1;$$

that is, f is an isomorphism. $\square$

Exercises

(1) Check that any two canonical divisors on a Riemann surface are rationally equivalent.

(2) Let $D \in Div(M)$, $g = \text{genus}(M)$. Prove that if $\deg D > 2g - 2$, then $i(D) = 0$. Likewise show that if $\deg D < 0$, then $\ell(D) = 0$.

(3) Let M be a genus g Riemann surface, and $p \in M$. Using Riemann-Roch, find the smallest value of k for which there *must* exist $f \in \mathcal{K}(M)^*$ having a pole at p of order no worse than k (i.e. $\nu_p(f) \geq -k$), and no other poles.

(4) Let M have genus $g \geq 2$. (a) Prove that M has a morphism to $\mathbb{P}^1$ of degree $\leq g + 1$. [Hint: use Exercise (3)] (b) Prove that M has a morphism to $\mathbb{P}^1$ of degree $\leq g$. [Hint: let $p \in M$, and look at $i((g - 2)[p])$. This is a bit harder than (a).]

(5) Assume $D > 0$. By Exercise (2), if $g \leq 1$ then $i(D) = 0$, and Riemann-Roch becomes $\ell(D) = \deg(D)$ for $g = 1$ and $\deg(D) + 1$ for $g = 0$. Prove this directly (a) for $M \cong \mathbb{P}^1$ and (b) for $M \cong \mathbb{C}/\Lambda$ (1-torus).

CHAPTER 26

Applications of Riemann-Roch, I: special Riemann surfaces

We now focus our attention on Riemann surfaces with a degree-two mapping to $\mathbb{P}^1$, starting with the case of genus 1. (The higher genus cases can be viewed as a generalization of elliptic curves, though there is no group law.) The first section begins with some general claims which will be more thoroughly investigated in the next Chapter.

26.1. Curves of genus 1

For the proof of Riemann-Roch (Theorem 25.2.3), we needed to use the (unproved) Normalization Theorem 3.2.1(B). It's actually possible to argue the other way, from Riemann-Roch to the existence of plane projective immersions (with ODP singularities) for arbitrary Riemann surfaces.

When can we do better? The degree-genus formula tells you that only curves of genera 0, 1, 3, 6, 10, 15, ... (numbers expressible as $\frac{(d-1)(d-2)}{2}$, $d \in \mathbb{N}$) can ever be embedded as *smooth* curves in $\mathbb{P}^2$. There is *no* reason to believe, from this or from the Normalization Theorem, that an *arbitrary* Riemann surface of one of these genera *can* be so embedded. In fact, *it isn't true* once you get to genera 6, 10, 15, ... That it works for genus 1 and genus 3 ("almost"; see next Chapter) is a bit of a miracle!

So: if you buy that any genus 1 Riemann surface is a complex 1-torus and any torus can be "Weierstrassed" into $\mathbb{P}^2$, the following result isn't surprising. On the other hand, it shows that Riemann-Roch is powerful and gives us a hint of how we might prove similar results in higher genus (e.g., 2 and 3) later.

THEOREM 26.1.1. *Let M be a Riemann surface of genus one. There exists an injective morphism of complex manifolds $\sigma : M \hookrightarrow \mathbb{P}^2$ with image $\sigma(M)$ a smooth algebraic curve of degree* 3.

Proof. Given $p \in M$, we know that $i(2[p]) = 0 = i([p])$ by Exercise 25.2, so that Riemann-Roch yields

$$\ell(2[p]) = \deg(2[p]) - g + 1 = 2 - 1 + 1 = 2,$$
$$\ell([p]) = \deg([p]) - g + 1 = 1.$$

In terms of the spaces of meromorphic functions, this says that

$$\mathfrak{L}(2[p]) \supsetneq \mathfrak{L}([p]) = \mathfrak{L}(0) = \mathcal{O}(M),$$

where $\dim \mathfrak{L}([p]) = 1$ means $\mathfrak{L}([p])$ consists of constant (or equivalently, holomorphic) functions. Therefore, we have an element

$$x \in \mathfrak{L}(2[p]) \backslash \mathfrak{L}([p]),$$

i.e. a meromorphic function with a double pole at p and no other poles.

Regard x as a morphism $M \to \mathbb{P}^1$. By Riemann-Hurwitz, the ramification degree

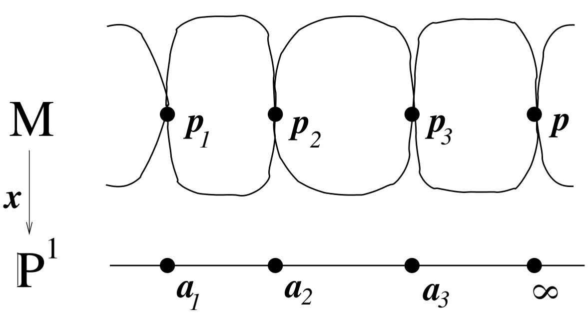
$$\deg(R_x) = 2(\deg(x) + g - 1) = 2(2 + 1 - 1) = 4,$$

whereas the ramification indices $\nu_p(x)$ for a degree two mapping are all ≤ 2. Hence, the ramification divisor is of the form (cf. §14.1 for notation)

$$R_x = [p_1] + [p_2] + [p_3] + [p_4]$$

with $p_1, p_2, p_3, p_4 \in M$ distinct. Set $a_i = x(p_i) \in \mathbb{P}^1$. The $\{a_i\}$ are still distinct points: by the form of R_x, $[p_i]$ must occur with multiplicity two in $x^{-1}([a_i])$; and since $(\deg(x) = 2 \implies) \deg(x^{-1}([a_i])) = 2$, the only possibility is $x^{-1}([a_i]) = 2[p_i]$.

Now clearly one of the p_i, say p_4, has to be p (as x has a double pole there). So also $a_4 = \infty$, and we have the picture:



In the following, $\{p_i\}$ resp. $\{a_i\}$ means $i = 1, 2, 3$.

Next, notice that $x - a_i$ is a local coordinate about a_i on $\mathbb{P}^1$. The meaning of a degree-2 ramification at p_i is simply that there is a local (holomorphic) coordinate about p_i on M such that[1] $z_i^2 = x - a_i$. Differentiating gives $dx \overset{\text{loc}}{=} 2z_i dz_i$. Again using Riemann-Roch, we have a nonvanishing holomorphic form $\omega \in \Omega^1(M)$, and by the Residue Theorem

$$0 = \sum_{q \in M} Res_q(x \cdot \omega) = Res_p(x \cdot \omega).$$

Writing $z := \int_p \omega$ for a local holomorphic coordinate at p, we have (locally) $\omega = dz$; since the residue vanishes, $x \overset{\text{loc}}{=} \frac{1}{z^2} + h(z)$ (h holomorphic) has no $\frac{1}{z}$ term.[2] Taking differentials, $dx \overset{\text{loc}}{=} \left(\frac{-2}{z^3} + h'(z)\right) dz$. Put together with the previous local computation, this tells us that dx has divisor

$$(dx) = \left(\sum_{i=1}^{3} [p_i]\right) - 3[p].$$

Set $y_0 := \frac{dx}{\omega} \in \mathcal{K}(M)^*$. In light of the fact that $(\omega) = 0$, we have that $(y_0) = (dx)$. If we put

$$g(x) := \prod_{i=1}^{3} (x - a_i),$$

then $(g(x)) = \sum_{i=1}^3 (x - a_i) = (\sum_{i=1}^3 2[p_i]) - 6[p] = 2(y_0) = (y_0^2)$. We conclude that $\frac{g(x)}{y_0^2}$ has trivial divisor and so is some constant C, and define $y := y_0\sqrt{C}$ so as to have

$$y^2 - g(x) = 0$$

on M.

Now for the embedding. Write $\sigma : M \to \mathbb{P}^2$ for the morphism defined by sending $p \mapsto [0 : 0 : 1]$ and all other points $q \mapsto [1 : x(q) : y(q)]$. The image $\sigma(M)$ is contained in the projective closure E of $\{y^2 - g(x) = 0\}$ (in $\mathbb{P}^2$), which is smooth due to distinctness of the $\{a_i\}$, and connected due to its irreducibility. By the usual arguments, $\sigma(M)$ is open and closed in E, hence equals E. At this point we have

---

[1] For an arbitrary choice of local coordinate $\mathfrak{z}_i$ it means that $x - a_i = \mathfrak{z}_i^2 h_i(\mathfrak{z}_i)$ where h doesn't vanish at 0; and then we can put $z_i := \mathfrak{z}_i \sqrt{h_i(\mathfrak{z}_i)}$.

[2] the coefficient of $\frac{1}{z^2}$ can be achieved by rescaling ω if needed

a diagram

$$
\begin{tikzcd}
C \setminus \{p\} \arrow[r, two heads, "\sigma"] \arrow[dr, "x"'] & E \setminus \{[0:0:1]\} \arrow[r, hook] \arrow[d, "\pi|_E"] & \mathbb{P}^2 \setminus \{[0:0:1]\} \arrow[dl, "\pi"] \\
& \mathbb{P}^1 &
\end{tikzcd}
\tag{26.1.1}
$$

where $\pi([Z : X : Y]) := [Z : X]$.

If σ is *not* injective, there exist distinct points $q_1, q_2 \in M \backslash \{p\}$ such that

$$\sigma(q_1) = \sigma(q_2) =: Q;$$

applying π to this gives

$$x(q_1) = x(q_2) = \pi(Q) =: \xi,$$

in which ξ is not ∞ or one of the $\{a_i\}$. Since $\deg(x) = 2$, $\deg(x^{-1}([\xi])) = 2$ and we must have $x^{-1}(\xi) = \{q_1, q_2\}$. From the equation for E it is evident that $\deg(\pi|_E) = 2$ also, with $(\pi|_E)^{-1}(\xi)$ consisting of $(\xi, \sqrt{g(\xi)})$ and $(\xi, -\sqrt{g(\xi)})$. Clearly one of these points has to be Q. From (26.1.1), it is also clear that q_1, q_2 are the only points of M that can go to these points. So whichever is not Q cannot get hit and σ fails to be surjective, a contradiction. $\square$

26.2. Hyperelliptic curves

Above we used the fact, for a genus one Riemann surface M, that $\ell(2[p]) = 2 > 1 = \ell([p])$ for $p \in M$, to construct a degree-two mapping $x : M \to \mathbb{P}^1$. Now suppose M has genus 2: how to map it to $\mathbb{P}^1$? Well, we have a basis $\{\omega_1, \omega_2\} \subset \Omega^1(M)$, and $\frac{\omega_2}{\omega_1}$ produces a meromorphic function, which does the job. By Poincaré-Hopf, $\deg((\omega_1)) = 2g - 2 = 2$, and so this map has two simple poles (or one double pole), hence has degree two.

In terms of homogeneous coordinates, we might write

$$p \mapsto [\omega_1(p) : \omega_2(p)],$$

where the meaning of the right-hand side is (expressing $\omega_i \overset{\text{loc}}{=} f_i(z)dz$ in terms of a local coordinate vanishing at p) simply $[f_1(0) : f_2(0)]$. If both f_i could simultaneously equal zero we would have a well-definedness problem (which could be gotten around by taking a limit), but this does not happen: we would have to have $i([p]) \geq 2$. By Riemann-Roch this yields $\ell([p]) = \deg([p]) - g + 1 + i([p]) \geq 2$, thereby producing an isomorphism $M \to \mathbb{P}^1$ as in the proof of 25.2.4, contradicting $g = 2$.

This discussion hopefully motivates

DEFINITION 26.2.1. A Riemann surface M is *hyperelliptic* iff there exists a (nonconstant) degree-two morphism $x : M \to \mathbb{P}^1$.

Clearly, any genus 2 Riemann surface is hyperelliptic.

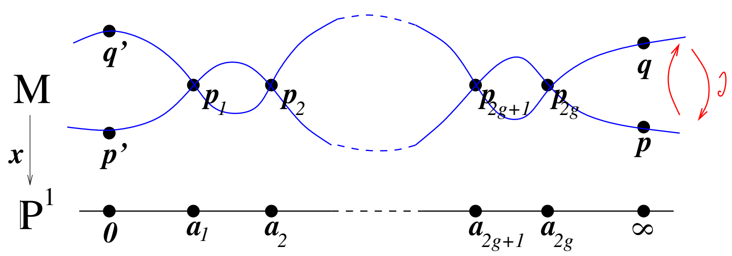
Now, let M be hyperelliptic of any genus and consider what the Riemann-Hurwitz formula has to say when applied to x:

$$\chi_M = 2\chi_{\mathbb{P}^1} - r_x$$
$$2 - 2g = 2 \cdot 2 - \sum_{p \in M} (\nu_p(x) - 1),$$

where $\deg(x) = 2 \implies \nu_p(x) \leq 2$. So the sum equals the number of ramification points, and this is just $2g + 2$:

$$R_x = [p_1] + \cdots + [p_{2g+2}].$$

By composing x with an automorphism of $\mathbb{P}^1$ if necessary, we may assume that none of the $x(p_i) =: a_i$ are 0 or ∞. Put $x^{-1}([\infty]) =: [p] + [q]$ and $x^{-1}([0]) =: [p'] + [q']$. We have the picture



in which $\jmath : M \to M$ denotes the involution exchanging the branches of M over $\mathbb{P}^1$ (cf. Exercise 1).

LEMMA 26.2.2. *Let V be a finite-dimensional vector space, $J : V \to V$ an involution. Then we have a decomposition $V = V^+ \oplus V^-$ into the $(+1)$- and (-1)- eigenspaces of J.*

PROOF. With respect to any basis for V, J is a matrix with minimal polynomial $m(t) = t^2 - 1$. This has no repeated roots, and so J is diagonalizable. Moreover, since $J^2 = \mathrm{id}_V$, any eigenvalue λ satisfies $\lambda^2 = 1$. $\square$

We apply this to the pullback map $\jmath^* : \Omega^1(M) \to \Omega^1(M)$. Notice that $\Omega^1(M)^+ = \{0\}$ since such forms would be pullbacks of holomorphic forms from $\mathbb{P}^1$ (cf. Exercise 2). Hence $\Omega^1(M) = \Omega^1(M)^-$ and so

$$\jmath^*\omega = -\omega$$

for all $\omega \in \Omega^1(M)$.

Put $D = (g+1)[p] + (g+1)[q] \in Div(M)$. We have (Exercise 25.2) $i(D) = 0$, so that Riemann-Roch gives

$$\ell(D) = 2g + 2 - g + 1 = g + 3.$$

Now apply the Lemma again, this time to $\jmath^* : \mathfrak{L}(D) \to \mathfrak{L}(D)$, noting that $\mathfrak{L}(D)^+$ contains the linearly independent set

$$\{1, x, x^2, \ldots, x^{g+1}\}.$$

In fact,[3]

$$\mathfrak{L}_M(D)^+ = x^* \underbrace{\mathfrak{L}_{\mathbb{P}^1}((g+1)[\infty])}_{\substack{\text{polynomials of} \\ \text{degree } \leq g+1}},$$

and so the above set is a basis. Therefore

$$\begin{aligned}\dim(\mathfrak{L}(D)^-) &= \ell(D) - \dim(\mathfrak{L}(D)^+) \\ &= (g+3) - (g+2) = 1,\end{aligned}$$

and there exists a nonzero $y \in \mathfrak{L}(D)$ such that $\jmath^* y = -y$.

CLAIM 26.2.3. $(y) = \sum_{i=1}^{2g+2} [p_i] - D$.

PROOF. Since the p_i are ramification points, $\jmath(p_i) = p_i$. But then

$$-y(p_i) = (\jmath^* y_i)(p_i) = y(\jmath(p_i)) = y(p_i)$$

and so $y(p_i) = 0$. That is, $y^{-1}([0]) \geq \sum_{i=1}^{2g+2} [p_i]$, which implies

$$\deg((y)) = \deg(y^{-1}([0])) \geq \deg\left(\sum [p_i]\right) = 2g + 2.$$

On the other hand, $y \in \mathfrak{L}(D) \implies y^{-1}([\infty]) \leq D \implies$

$$\deg((y)) = \deg(y^{-1}([\infty])) \leq \deg(D) = 2g + 2.$$

So $\deg((y))$ is forced to equal $2g+2$, which means also that $y^{-1}([0]) = \sum [p_i]$ and $y^{-1}([\infty]) = D$. $\square$

[3] using subscripts to denote which Riemann surface we are considering functions on (e.g. $\mathfrak{L}_M(D)$ just means $\mathfrak{L}(D)$)

Set

$$g(x) := \prod_{i=1}^{2g+2} (x - a_i) \in \mathcal{K}(M)^*,$$

and compute (in $Div(M)$)

$$(g(x)) = \sum((x-a_i)) = 2\sum[p_i] - (2g+2)x^{-1}([\infty]) = 2\sum[p_i] - 2D = (y^2).$$

But then $y^2/g(x)$ has trivial divisor, and so is a constant. Rescaling y, we have that (in $\mathcal{K}(M)$)

$$y^2 - g(x) = 0.$$

By considering the image of

$$\sigma : M \to \mathbb{P}^2$$

given by

$$m(\neq p, q) \;\mapsto\; [1 : x(m) : y(m)]$$

and

$$p, q \;\mapsto\; [0 : 0 : 1],$$

we arrive at:

THEOREM 26.2.4. *Hyperelliptic Riemann surfaces are precisely the normalizations of (plane) algebraic curves of the form*[4]

$$\left\{ Y^2 Z^{2g} = \prod_{i=1}^{2g+2} (X - a_i Z) \right\} \subset \mathbb{P}^2.$$

A basis of $\Omega^1(M)$ *is given by* $\omega_j := \frac{x^{j-1}dx}{y}$, $j = 1, \ldots, g$.

PROOF. We just need to show ω_j is holomorphic:

$$\begin{aligned}(\omega_j) &= (j-1)(x) + (dx) - (y) \\ &= \{(j-1)([p'] + [q']) - (j-1)([p] + [q])\} \\ &\quad + \left\{\sum[p_i] \;-\; 2([p] + [q])\right\} \\ &\quad - \left\{\sum[p_i] \;-\; (g+1)([p] + [q])\right\} \\ &= (j-1)\,([p'] + [q']) + (g-j)\,([p] + [q]) \;\geq\; 0.\end{aligned}$$

□

[4]note: the singular point $[0 : 0 : 1]$ is not an ODP, so the construction of g holomorphic differentials that follows shouldn't be compared with the formulas you know in that case. Also, it should be emphasized that the a_i are distinct.

A *hyperelliptic curve*, by the way, is just an irredicuble projective curve whose normalization is a hyperelliptic Riemann surface!

The first two exercises below are ones you could have done long ago, but fill in (very) small gaps in the proofs above. The same goes for the third, if you had known the definition of hyperelliptic! The fourth does make heavy use of Riemann-Roch.

Exercises

(1) Given a degree 2 holomorphic map $\varphi : M \to M'$ of compact Riemann surfaces, with corresponding involution $\jmath$ defined as follows: if $p \in M$ is a ramification point of φ, $\jmath(p) := p$; otherwise, $\varphi^{-1}(\varphi(p)) = \{p, \tilde{p}\}$ and $\jmath(p) := \tilde{p}$. Clearly $\jmath \circ \jmath = Id_M$ (i.e. $\jmath$ is an involution) and $\varphi \circ \jmath = \varphi$. Prove that $\jmath : M \to M$ is a holomorphic map of Riemann surfaces. Since $\jmath$ is injective and surjective (why?), it follows that $\jmath \in Aut(M)$.

(2) Continuing Exercise (1), let $\omega \in \Omega^1(M)$ satisfy $\jmath^*\omega = \omega$. Prove that $\omega = \varphi^*\eta$ for some $\eta \in \Omega^1(M')$.

(3) Suppose that a dth-degree irreducible algebraic curve $C \subset \mathbb{P}^2$ has a point of multiplicity $(d-2)$. Show that C is hyperelliptic.

(4) Let M be a Riemann surface of genus two. In this problem you will construct a realization of M as an algebraic curve, different to that produced above. You will need to use that M is hyperelliptic, with $x : M \to \mathbb{P}^1$ its degree-two mapping and $\jmath$ the associated involution. Take p and q (in contrast to the notation above) fixed non-ramification points on M with distinct images under x; let α and β denote arbitrary points of M.

(a) Prove that $\ell([\alpha] + [\beta]) = 1$ unless $\jmath(\alpha) = \beta$. [Hint: otherwise you get a different involution (why?). To see why this is a problem you might consider the fact that $\jmath^*\omega = -\omega$ for all holomorphic forms implies their divisors are $\jmath$-symmetric.]

(b) For any points α, β on M, show $i([\alpha] + [\beta] - [p] - [q]) = 1$ (as opposed to 2) $\iff$ $\{\alpha, \beta\} \neq \{p, q\}$. [Hint: use (a), and consider $\ell([\alpha] + [\beta] - [p] - [q])$.]

(c) Use $\mathfrak{I}(-[p] - [q])$ to construct a map $\varphi : M \to \mathbb{P}^2$. [Hint: compute $i(-[p]-[q])$.] You will need to check that φ is well-defined. [Hint: compute $i(-[p] - [q] + [\alpha])$, using Exercise 25.2.]

(d) Show that φ is injective off $\{p, q\}$, but that $\varphi(p) = \varphi(q)$. [Hint: using part (b), compute $i([\alpha] + [\beta] - [p] - [q])$.]

(e) Show that there exists a meromorphic form $\omega \in \mathfrak{I}(-[p] - [q])$ with poles at *both* p and q.

(f) Explain why the zero-divisor $(\omega)_0$ (the effective "part" of (ω)) is φ^{-1} of the intersection [divisor] of a line in $\mathbb{P}^2$ with $C := \varphi(M)$. Prove that $\deg((\omega)_0) = 4$ (easy). Assuming C is an algebraic curve (this is GAGA), conclude that $\deg(C) = 4$.

(g) Clearly $\varphi(p) = \varphi(q)$ is a singularity of C. Prove it is the only one, and a double point. [Hint: assume otherwise, and produce a genus zero normalization or similar.]

CHAPTER 27

Applications of Riemann-Roch, II: general Riemann surfaces

Next we'll use Riemann-Roch to develop two methods for mapping an arbitrary Riemann surface into a (usually higher-dimensional) projective space, with a nice application to curves of genus three. The second approach behaves differently in the hyperelliptic and nonhyperelliptic cases, so we first will want to convince ourselves that there *are* nonhyperelliptic Riemann surfaces! To see this, we will start with an heuristic argument for the "number of complex parameters" governing Riemann surfaces, and show that the hyperelliptic ones have fewer parameters. But there's much more in this chapter, which should give a glimpse of how rich the correspondence between algebraic curves and Riemann surfaces really is.

27.1. Moduli

In algebraic geometry there is the notion of *moduli spaces*, which parametrize structures of a prescribed sort modulo some equivalence relation, such as "smooth algebraic curves of degree 5 up to projective equivalence" or "Riemann surfaces of genus 4 up to isomorphism". A main point is that these spaces can be given algebraic structure themselves, i.e. turned into algebraic varieties, in many cases. Suitably refined, the structure of these varieties is one of the hotter topics of study around.[1]

We shall only be concerned with the notion of moduli as a set of local parameters (on the moduli space), and will say colloquially that some structure has a certain *number of moduli*: e.g. genus 1 (resp. 0) Riemann surfaces have one modulus (resp. zero moduli) since they can all be expressed as $\mathbb{C}/\mathbb{Z}\langle 1,\tau\rangle$ (resp. $\mathbb{P}^1$) up to isomorphism. Underlying

[1] so-called "modular curves" or "modular varieties" are a more specialized notion with an arithmetic flavor

the claim in §26.1 that not all Riemann surfaces of genus 6, 10, 15, etc. can be embedded smoothly in $\mathbb{P}^2$ is a deep calculation of Riemann:

THEOREM 27.1.1. *Riemann surfaces of genus $g \geq 2$ (up to isomorphism) have $3g-3$ moduli.*

PROOF. (Sketch) Consider a genus g Riemann surface M, and any effective divisor D of degree $2g$ on M. By Exercise 25.2, $i(D) = i(D - [p]) = 0$ for any point $p \in M$. So Riemann-Roch says

$$\ell(D) = \deg(D) - g + 1 = g + 1,$$

$$\ell(D - [p]) = g.$$

Hence there exists $f \in \mathfrak{L}(D)$ and not in any of the finitely many $\mathfrak{L}(D - [p])$ for those p appearing in D; and so $f^{-1}([\infty]) = D$, which means $\deg(f) = \deg(f^{-1}([\infty])) = 2g$. Now Riemann-Hurwitz tells us about the ramification behavior:

$$\chi_M = \deg(f) \cdot \chi_{\mathbb{P}^1} - \deg(R_f)$$

$$2 - 2g = 2g \cdot 2 - r_f$$

$$r_f = 6g - 2.$$

For "almost all" D the points in R_f will have multiplicity one (ramifications of order two) and lie over distinct points in $\mathbb{P}^1$, meaning that the branch locus $B \subset \mathbb{P}^1$ consists of $6g-2$ points. We want to use all this data to compute the number of "local deformation parameters" of M.

Look at this in a slightly more formal way, consider the set $\mathfrak{S}_1$ of 2-tuples (M, f) where M has genus g and f has degree $2g$. This maps to the set $\mathfrak{S}_2$ of 2-tuples (M, D) where $D \geq 0$ of degree $2g$ (take $D := f^{-1}([\infty])$). From there you can map to the set $\mathfrak{S}$ of Riemann surfaces of genus g, by forgetting D. It's fairly clear that (fixing M) D has $2g$ parameters, making $\dim(\mathfrak{S}_2) - \dim(\mathfrak{S}) = 2g$. Moreover, given M and D, there are $\ell(D) = g + 1$ choices of parameter for f (to have D as its poles), meaning $\dim(\mathfrak{S}_1) - \dim(\mathfrak{S}_2) = g + 1$. Our argument in the first paragraph shows that the first map is surjective (while the second obviously is), and so

$$\dim(\mathfrak{S}) = \underbrace{\{\dim(\mathfrak{S}_1) - g - 1\}}_{\dim(\mathfrak{S}_2)} - 2g = \dim(\mathfrak{S}_1) - 3g - 1. \tag{27.1.1}$$

On the other hand, you can map $\mathfrak{S}_1$ to $\mathfrak{S}_3$, the set of $(6g-2)$-tuples of (unordered) points on $\mathbb{P}^1$, by taking $f(R_f) \in Div(\mathbb{P}^1)$. This map is surjective since given a branch-point set in $\mathbb{P}^1$ you can construct an existence domain for an appropriate algebraic function,[2] and in fact the construction shows that there are only finitely many possibilities for M. Moreover, it shows that a continuous family of degree-$2g$ functions on M with the same branch-point set gives rise to a continuous family of automorphisms of M. But for $g \geq 2$ M has only finitely many automorphisms. So we see that this map is finite-to-1, and thus $\dim(\mathfrak{S}_1) = \dim(\mathfrak{S}_3) = 6g-2$. Plugging this in to (27.1.1), we get the desired result. $\square$

It's much easier to count moduli for hyperelliptic and algebraic plane curves.

PROPOSITION 27.1.2. *Hyperelliptic Riemann surfaces of genus* $g \geq 1$ *(up to isomorphism) have* $2g-1$ *moduli.*

PROOF. They are essentially just the existence domains of the algebraic functions $\sqrt{\prod_{i=1}^{2g+2}(z-\alpha_i)}$, and so are completely determined by the branch locus $\{\alpha_i\}_{i=1}^{2g+2}$. This has $2g+2$ parameters, but we have to account for change of coordinate on $\mathbb{P}^1$, which is by $PGL_2(\mathbb{C})$, by subtracting $\dim(PGL_2) = 3$. $\square$

PROPOSITION 27.1.3. *Smooth algebraic curves of degree* d *modulo projective equivalence have* $\binom{d+2}{2} - 9$ *moduli.*

PROOF. A curve is determined by a polynomial in S_3^d, which has dimension $\binom{d+2}{2}$. We have only to account for changing projective coordinates by $GL_3(\mathbb{C})$, which has dimension 9. (Here PGL_3 is not what is wanted, as we do want to quotient out the rescalings of the equation.) $\square$

Now we can compare moduli, with two very interesting results. First, consider the numbers you get for general Riemann surfaces of genus 1, 2, 3, 4, 5, 6 : the numbers of moduli are 1, 3, 6, 9, 12, 15. For hyperelliptic ones, we have instead 1, 3, 5, 7, 9, 11. So while all genus 2 Riemann surfaces are hyperelliptic, we have:

[2] like $\sqrt{(z-a)(z-b)(z-c)(z-d)}$ (for $g=1$), but more complicated (since $g \geq 2$); cf. [Griffiths and Harris], pp. 255-257

PROPOSITION 27.1.4. *A general Riemann surface of genus $g \geq 3$ is non-hyperelliptic.*

So we will need a more general method of realizing Riemann surfaces as algebraic curves than that discussed in §26.2, and that is what we will do in the remainder of this Chapter.

Finally, look at those genera which correspond to nonsingular algebraic curves in $\mathbb{P}^2$ of degree 3, 4, 5, 6, 7, ...: namely, 1, 3, 6, 10, 15, and so on. The Riemann surfaces of these genera have (by Thm. 27.1.1) moduli 1, 6, 15, 27, 42,etc. But now look at the smooth algebraic plane curves of the corresponding degrees (via Prop. 27.1.3): we get 1, 6, 12, 19, 27. The case of genus 3 will be treated in §27.3. Beyond that, we have:

PROPOSITION 27.1.5. *Smooth algebraic plane curves of degree $d \geq 5$ do* not *yield all the Riemann surfaces of genus $\frac{(d-1)(d-2)}{2}$ — only a special subset.*

27.2. Projective embeddings

Let M be a Riemann surface of genus g. For any $\mathfrak{D} \in Div(M)$ of degree $> 2g-2$, recall from Exercise 2 of Chap. 25 that $i(\mathfrak{D}) = 0$. By Riemann-Roch, we then have

$$\ell(\mathfrak{D}) = \deg(\mathfrak{D}) - g + 1.$$

This will be used repeatedly in the argument below.[3]

Now fix a divisor $D = \sum_{p \in M} n_p[p] \in Div(M)$ of degree $d \geq 2g+1$. We will define an embedding (injective morphism of complex manifolds)

$$\varphi : M \hookrightarrow \mathbb{P}^{d-g}.$$

(Since $d - g \geq g+1$, this can only give an embedding in $\mathbb{P}^2$ for $g = 1$.) The *support* of D, which is the subset of M consisting of the points aapearing in D (i.e. those p with nonzero n_p), is written $|D|$.

[3]It's very important to understand the argument in this section. Try slimming it down (I've expressed it in a somewhat bloated manner) and writing it out for a specific choice of d and g (> 1, say). (Also, if you are stuck on Chap. 26 Exercise 4, some of the steps are similar.)

First off, certainly $d > 2g-2$ and so $\ell(D) = d-g+1$. Write $\{f_0, \ldots, f_{d-g}\}$ for a basis of $\mathfrak{L}(D)$, and define for $p \notin |D|$

$$\varphi(p) := [f_0(p) : \cdots : f_{d-g}(p)]. \tag{27.2.1}$$

If $p \in |D|$, this is unsuitable since some functions may blow up (or all functions may be required to vanish). Therefore if z is a local coordinate (vanishing at p to first order), we put

$$\varphi(p) := [(z^{n_p} f_0)(p) : \cdots : (z^{n_p} f_{d-g})(p)]. \tag{27.2.2}$$

For points q in a neighborhood of p, $[(z^{n_p} f_0)(q) : \cdots : (z^{n_p} f_{d-g})(q)]$ gives the same result as $[f_0(q) : \cdots : f_{d-g}(q)]$, and so we have constructed an analytic map ... provided that (27.2.1-2) do not yield $[0 : \cdots : 0]$ at any point. That is the central well-definedness issue, and we must check it.[4]

Now for $p, q \in M$, notice that $D - [p]$, $D - 2[p]$ and $D - [p] - [q]$ each still have degree $> 2g-2$. Therefore we have

$$\ell(D - [p]) = d - g$$

and

$$\ell(D - 2[p]) = d - g - 1 = \ell(D - [p] - [q]),$$

with the immediate consequences

$$\mathfrak{L}(D - [p]) \subsetneq \mathfrak{L}(D), \tag{27.2.3}$$

$$\mathfrak{L}(D - [p] - [q]) \subsetneq \mathfrak{L}(D - [p]), \tag{27.2.4}$$

$$\mathfrak{L}(D - 2[p]) \subsetneq \mathfrak{L}(D - [p]). \tag{27.2.5}$$

To interpret these, for simplicity first assume $p, q \notin |D|$. Then (27.2.3) says that there exists $f \in \mathfrak{L}(D)$ not vanishing at p, meaning that the $\{f_i(p)\}$ are not all zero; this makes φ well-defined on $M \backslash |D|$. Next, (27.2.4) gives us $g \in \mathfrak{L}(D - [p]) \backslash \mathfrak{L}(D - [p] - [q])$, a function vanishing at p but not q, forcing φ to take different values at p and q; hence φ is injective on $M \backslash |D|$. Finally, (27.2.5) provides $h \in \mathfrak{L}(D - [p]) \backslash \mathfrak{L}(D - 2[p])$, i.e. vanishing to exactly first order at p, so that the

[4]Note that (27.2.2) isn't just a special formula for $p \in |D|$; it contains (is more general than) (27.2.1) since for $p \notin |D|$ we have $n_p = 0$.

derivative of h hence that of φ is nonzero there; together with the injectivity result, this proves that the image of $M\backslash|D|$ is smooth.

In order to extend these statements to all of M, we have to refine the argument just a bit. For a general points $p, q \in M$, (27.2.3) tells us that there exists a function $f \in \mathfrak{L}(D)$ with $\nu_p(f) = -n_p$ exactly; (27.2.4) that some $g \in \mathfrak{L}(D)$ has $\nu_p(g) > -n_p$ but $\nu_q(g) = -n_q$; and (27.2.5) that there exists an $h \in \mathfrak{L}(D)$ with $\nu_p(h) = -n_p + 1$ exactly. These give precisely the well-definedness, injectivity, and smoothness of image for the map described by (27.2.2). So the image is a compact complex analytic curve $\mathbb{P}^{d-g}$, which is algebraic by GAGA.

27.3. Canonical maps

Once again we consider a Riemann surface M, this time of genus $g \geq 2$, and let

$$\{\omega_1, \ldots, \omega_g\} \subset \Omega^1(M)$$

be a basis. Instead of choosing a divisor and go through Riemann-Roch to get a projective embedding from meromorphic functions, why not just use these? Define the *canonical map*

$$\varphi_K : M \to \mathbb{P}^{g-1}$$

by

$$p \mapsto [\omega_1(p) : \cdots : \omega_g(p)].$$

The meaning of this, as you would expect, is locally writing each ω_i as $f_i(z)dz$, and taking $[f_1(p) : \cdots : f_g(p)]$. This is well-defined, i.e. the $\{\omega_i\}$ do not all have a zero at p. Otherwise we would have $\mathfrak{I}([p]) = \mathfrak{I}(0) = g$ hence (by Riemann-Roch) $\mathfrak{L}([p]) = 2$, which we know to be false for M not isomorphic to $\mathbb{P}^1$.

Bottom line: this looks quite promising, from the standpoint of getting a convenient projective embedding. Or does it?

EXAMPLE 27.3.1. For M hyperelliptic, consider the setting of Theorem 26.2.4; we have

$$\varphi_K(p) = \left[\frac{dx}{y} : \frac{x\,dx}{y} : \cdots : \frac{x^{g-1}dx}{y}\right] = [1 : x : \cdots : x^{g-1}].$$

Notice that this looks a lot like the rational canonical map f from Example 7.3.7. In fact, it factors

$$\begin{tikzcd} C \arrow[rr, "\varphi_K"'] \arrow[dr, "x"'] & & \mathbb{P}^{g-1} \\ & \mathbb{P}^1 \arrow[ur, "f"'] & \end{tikzcd} \tag{27.3.1}$$

with $\deg(x) = 2$, and so does *not* give an embedding of M in $\mathbb{P}^{g-1}$.

All is not lost: the hyperelliptic case is very special (in a bad way), and for $g \geq 3$ we know that there are (lots of) nonhyperelliptic curves.

THEOREM 27.3.2. *Let φ_K be the canonical map for an arbitrary Riemann surface of genus $g \geq 2$. Then*

(a) φ_K is nondegenerate;[5]

(b) $\varphi_K(M)$ is smooth; and

(c) φ_K is injective $\iff$ M is nonhyperelliptic.

PROOF. (a) Were $\varphi_K(M)$ contained in a proper linear subspace of $\mathbb{P}^{g-1}$, this would produce a linear relation on the $\{\omega_i\}$. But they are linearly independent by construction, being a basis!

(b) This is clear in the hyperelliptic case, by observing that the derivative of the (injective) rational canonical map is nowhere vanishing. (We will return to this in the nonhyperelliptic case.)

(c) The implication " $\Longrightarrow$ " is already done (by contrapositive) in Example 27.3.1.

Now let z be a local coordinate vanishing to first order at a point $p \in M$, and consider the linear functionals on $\Omega^1(M)$ $(= \mathfrak{I}(0))$ given by

$$\omega \mapsto \left(\frac{\omega}{dz}\right)(p)$$
$$\omega \mapsto \left(\frac{\omega}{dz}\right)'(p)$$
$$\vdots$$
$$\omega \mapsto \left(\frac{\omega}{dz}\right)^{(k-1)}(p).$$

If the first is zero on some given ω, then $\omega \in \mathfrak{I}([p])$. If the first and second are zero, then $\omega \in \mathfrak{I}(2[p])$. If all are zero, then $\omega \in \mathfrak{I}(k[p])$. Since k linear conditions cut out a subspace of codimension at most k,

[5] cf. the beginning of Chapter 7, and also 7.3.1.

we have

$$i(k[p]) = \dim \mathfrak{I}(k[p]) \geq g - k.$$

More precisely, we have $i(k[p]) = g - k + a$ and by Riemann-Roch

$$\ell(k[p]) = k - g + 1 + i(k[p]) = 1 + a.$$

For the special case $k = 1$, there can be no redundancies in *one* linear condition and so we have $a = 0$.

Suppose $\varphi_K([p]) = \varphi_K([q])$ for $p \neq q$. Then $\omega \mapsto \left(\frac{\omega}{dz}\right)(p)$ and $\omega \mapsto \left(\frac{\omega}{dz}\right)(q)$ yield the same functional on $\Omega^1(M)$, up to a constant multiple; in particular they vanish on the same ω's. So $\mathfrak{I}([p]) = \mathfrak{I}([p]+[q])$, which yields $i([p]+[q]) = i([p]) = g - 1$ and via R-R

$$\ell([p]+[q]) = 2 - g + 1 + i([p]+[q]) = 2.$$

Therefore there exists a nonconstant meromorphic function $\mathcal{F} \in \mathfrak{L}([p]+[q])$. We know that $\mathfrak{L}([p])$ and $\mathfrak{L}([q])$ have only constant functions ($a = 0$ when $k = 1$), and so $\mathcal{F}$ has to have the allowed simple pole at both of p and q. Thus $\deg(\mathcal{F}) = 2$, and so M is hyperelliptic. This completes the proof of (c).

(b, cont'd.) Now assume M is *non*hyperelliptic. Then we must have $\ell(2[p]) = 1$, i.e. $a = 0$ for $k = 2$. Consequently

$$i(2[p]) = g - 2$$

(whilst $i([p]) = g - 1$), and we can arrange a basis of $\Omega^1(M)$ so that in local coordinates at p,

$$\omega_1 \overset{\text{loc}}{=} dz\,, \quad \omega_2 \overset{\text{loc}}{=} z h_2(z) dz\,,$$

$$\omega_j \overset{\text{loc}}{=} z^2 h_j(z) dz \quad (3 \leq j \leq g).$$

(Here the $h_i(z)$ are holomorphic, and h_2 doesn't vanish at $z = 0$.) The canonical map takes the local form

$$\varphi_K(z) = [1 : z h_2(z) : z^2 h_3(z) : \cdots : z^2 h_g(z)]$$

with derivative

$$\varphi_K(z) = [0 :\ h_2(z) + z h_2'(z) :\ * : \cdots :\ *]$$

which does not vanish at p. This gives the desired smoothness. $\square$

Consider a smooth, irreducible algebraic curve C and hyperplane $H = \{W(\underline{Z}) = 0\}$, both in $\mathbb{P}^n$. (Here W is a homogeneous polynomial in $Z_1, \ldots, Z_{n+1}$ of degree one, with affine form w.) One can define an intersection divisor $C \cdot H$ on C in a way which extends what we have done in $\mathbb{P}^2$. If C is not necessarily smooth, then the divisor lives on a normalization M of C and is denoted $\sigma^* H$; it is given simply by $\sum_{p \in M} ord_p(\sigma^* w)[p]$. We define the degree of the curve to be the degree of this divisor, called a *hyperplane section*:

$$\deg(C) := \deg(\sigma^* H).$$

Since any two hyperplane sections are rationally equivalent (why?), any two hyperplane sections have the same degree, making $\deg(C)$ well-defined.

In the case at hand, σ is φ_K and $n = g-1$. Hyperplane sections are particularly interesting because if we write $W(\underline{Z}) = \sum_{i=1}^{g} \alpha_i Z_i$, then

$$\sigma^* H \ = \ (\alpha_1 \omega_1 + \cdots + \alpha_g \omega_g)$$

is a canonical divisor on M! That's why φ_K is called the canonical map, and its image $\varphi_K(M)$ a *canonical curve.*

PROPOSITION 27.3.3. *Assume M is nonhyperelliptic of genus g. Then the degree of the canonical curve $\varphi_K(M) \subset \mathbb{P}^{g-1}$ is $2g-2$.*

PROOF. The assumption is necessary in order that M normalize $\varphi_K(M)$. (In the hyperelliptic case, it is normalized by the rational canonical map.) We then compute $\deg(\varphi_K(M)) = \deg(\varphi_K^* H) = \deg(K) = 2g-2$ by Poincaré-Hopf, and that's it. $\square$

And so, we find that "nearly all" genus 3 curves have a nice embedding into the projective plane.

COROLLARY 27.3.4. *Every nonhyperelliptic genus 3 curve is the normalization of a smooth quartic curve in $\mathbb{P}^2$.*

27.4. Weierstrass points

We began our discussion of Riemann-Roch with a naive analysis, for a fixed point p on a Riemann surface M, of what orders of pole are possible if we are after a meromorphic function with its only pole at p. To conclude, I will now briefly explain the sense in which this can depend on the choice of p and not just the genus g of M. Assume $g > 0$ for what follows.

First note that $\ell(0) = 1$ (constant functions), and $\ell([p]) = 1$ by the argument at the beginning of Chap. 25. By equation (25.2.2), we know for each k that

$$0 \leq \ell((k+1)[p]) - \ell(k[p]) \leq 1.$$

On the other hand, since the degree of $(2g-1)[p]$ exceeds $2g-2$, we have (Chap. 25 Exercise 2) that $i((2g-1)[p]) = 0$, and so (by Riemann-Roch)

$$\ell((2g-1)[p]) = (2g-1) - g + 1 = g.$$

More generally, for $k \geq 2g-1$, the fact that $i(k[p]) = 0$ yields

$$\ell(k[p]) = k - g + 1.$$

So the scenario is that $\ell(k[p])$ starts (at $k = 0$) at 1 and works its way up to g in increments of 0 or 1, as k rises to $2g-1$; thereafter it increases by 1 whenever k does.

The situation with $i(k[p])$ is "dual": it starts at g and works its way down to 0 in decrements of 0 or 1, as k rises to $2g-1$; and then it stays at 0.

Now it turns out that at all but finitely many points, $\ell(g[p]) = 1$; that is, all the increments are postponed as far as possible and the sequence $\ell(k[p])$ looks like $1, 1, 1, \ldots, 1, 2, 3, \ldots, g$, and so on. Those points where this is not the case are called the *Weierstrass points* of M. The simplest example I am aware of is, for a hyperelliptic curve, the $2g+2$ fixed points of the involution $\jmath$. For these the sequence looks like $1, 1, 2, 2, 3, 3$, etc.

Exercises

(1) Check that the definition of $\varphi_K(p)$ in §27.3 is independent of the choice of local coordinate near p.

(2) Show that any smooth quartic curve in $\mathbb{P}^2$ is a canonical curve (of genus 3), and hence also nonhyperelliptic.

(3) In this exercise you will prove a new Cayley-Bacharach type result (in $\mathbb{P}^2$): *if C_1 and C_2 (of degrees m and n respectively, with C_1 assumed smooth and irreducible) meet in mn distinct points, and*

C_3 (of degree $m+n-3$) passes through $p_1, \ldots, p_{mn-1} \in C_1 \cap C_2$, then it passes through the remaining point $p := p_{mn}$. (We shall write f_1, f_2, f_3 for the resp. defining homogeneous polynomials.) Start out by assuming C_3 does not contain p, and follow these steps:

(a) Let g denote the genus of C_1, and show that $(m-3)m = 2g-2$ and $g = \dim(S_3^{m-3})$.

(b) Let $h \in S_3^{m-3}$, set $F_h := \frac{f_2 \cdot h}{f_3}|_{C_1} \in \mathcal{K}(C_1)^*$, and write $(F_h) = [p] + (h) - D$ (this defines $D \in Div(C_1)$). Show that the map $S_3^{m-3} \to \mathfrak{L}(D)/\mathbb{C}$ (here $\mathbb{C}$ =constant functions) given by $h \mapsto F_h$ is injective, and use this to put a lower bound on $\ell(D)$.

(c) Find $\deg(D)$, $i(D)$, and obtain a contradiction.

(4) A problem on automorphisms of canonical curves:

(a) Let $\alpha : C \to C$ be an automorphism of a canonical curve of genus g. Prove that α is the restriction to C of a linear automorphism of $\mathbb{P}^{g-1}$. [Hint: consider the action of α^* on $\Omega^1(C)$.]

(b) Let M be a *non*hyperelliptic Riemann surface of genus 3 with an involution $\jmath$. How many fixed points does it have? What is the genus of the quotient Riemann surface? [Hint: consider the canonical embedding and apply (a); $\jmath$ is the restriction of what sort of linear automorphism on $\mathbb{P}^2$?]

(5) Using the embedding of §27.2, try to prove: (a) there exists, for an arbitrary Riemann surface, an embedding onto a smooth curve in $\mathbb{P}^3$ and (b) an immersion onto a curve with only ODP singularities in $\mathbb{P}^2$. [Hint: use sufficiently general projections of the complement of a linear subspace in $\mathbb{P}^n$ onto $\mathbb{P}^2$ and $\mathbb{P}^3$.] Also, (c) what degrees do these curves have?

CHAPTER 28

Abel's Theorem, part I

Recall the setup from Chapter 20: M a Riemann surface of genus $g \geq 1$, with closed paths ("1-cycles")γ_i giving a basis $\{[\gamma_i]\}_{i=1}^{2g}$ for $H_1(M,\mathbb{Z})$. We have the *Jacobian* of M, which is the complex g-torus

$$J(M) := \frac{(\Omega^1(M))^\vee}{H_1(M,\mathbb{Z})} \overset{\cong}{\to} \frac{\mathbb{C}^g}{\Lambda_M}.$$

The isomorphism is given by evaluating functionals against a basis $\{\omega_1,\ldots,\omega_g\} \subset \Omega^1(M)$, and Λ_M is called the *period lattice*. The Picard group

$$Pic^0(M) := \frac{Div^0(M)}{(\mathcal{K}(M)^*)}$$

of degree-0 divisors modulo rational equivalence is the object we want to understand. To this end, we had shown that the *Abel-Jacobi map*

$$AJ:\ Pic^0(M) \to J(M)$$
$$D \mapsto \int_{\partial^{-1}D}$$

is a well-defined homomorphism, where $\partial^{-1}D$ is just a compact way of writing "some 1-chain Γ with $\partial\Gamma = D$". The important content of this is that $AJ((f)) = 0$ for any $f \in \mathcal{K}(M)^*$.

By Abel's theorem we will henceforth mean the statement that AJ is injective, that is

$$AJ(D) = 0 \implies D = (f) \text{ for some } f \in \mathcal{K}(M)^*\,; \tag{28.0.1}$$

while the surjectivity will be known as Jacobi inversion: i.e.,

(28.0.2)

given any point in $J(M)$ [any functional on $\Omega^1(M)$, up to periods], there exists a divisor D inducing that functional via $\int_{\partial^{-1}D}(\cdot)$.

These statements will be proved in Chap. 29. Our aim here is just to explain how Abel's theorem relates to Riemann-Roch and develop a couple technical lemmas to be used in the sequel.

Before starting let's refine one aspect of the above picture just a bit. Intersecting 1-cycles on M – or more precisely, intersecting *transverse* representatives of homology classes[1] – gives a perfect, unimodular pairing

$$\text{“}\cdot\text{”} : H_1(M,\mathbb{Z}) \times H_1(M,\mathbb{Z}) \to \mathbb{Z}.$$

There is always a *symplectic basis*, that is one with the property that

$$\gamma_i \cdot \gamma_{g+j} = \delta_{ij} = -\gamma_{g+j} \cdot \gamma_i$$

$$\gamma_i \cdot \gamma_j = 0 = \gamma_{i+g} \cdot \gamma_{j+g}$$

for $1 \le i, j \le g$ (where δ_{ij} is the Kronecker delta). This is the situation *pictured* in §20.1. We will assume henceforth when writing $\{\gamma_i\}_{i=1}^{2g}$ that they form a symplectic basis for the first homology.

We should also remark on what the Picard group is really doing here. For an elliptic curve E, in $Pic^0(E)$ we have $[p+q]-[p]-[q]+[\mathcal{O}] \equiv 0$, where addition inside the brackets is the group law on E and outside the brackets means adding divisors. What this says is: while as divisors (i.e. in the free abelian group on points of E) $[p+q] \neq [p]+[q]$, working modulo rational equivalence we *do* have $[p+q]+[\mathcal{O}] \equiv [p]+[q]$. So Pic effectively recovers the group law on E. Now, curves of higher genus have no group law on points; but by "linearizing" points and working modulo divisors of functions, we get a form of generalization of the group law in genus 1. Intriguingly, a more precise form of Jacobi inversion in the next Chapter will tell us that this may "almost" be seen as a group law on *unordered g-tuples of points on M*.

28.1. From Riemann-Roch to Abel-Jacobi

Let D be a divisor on M; we have been interested in the dimensions of the vector spaces $\mathfrak{L}(D)$ and $\mathfrak{I}(D)$. In the interval

$$0 \le \deg(D) \le 2g-2$$

is where anything "of interest" lies: outside this range, either $\ell(D)$ or $i(D)$ is zero. At the extremes, Abel's theorem will tell us:

(i) $\ell(D)$ when $\deg(D) = 0$; and

(ii) $i(D)$ when $\deg(D) = 2g-2$.

[1]each intersection point contributes a ± 1 according to the "right-hand rule" and the orientation of M given by the complex structure

In case (i), if there is a meromorphic function $f \in \mathcal{K}(M)^*$ with $(f) + D \geq 0$, then

$$\begin{aligned} \deg((f) + D) &= \deg((f)) + \deg(D) = 0 + 0 = 0 \\ \Longrightarrow \quad (f) + D &= 0 \quad \Longrightarrow \quad D \overset{\text{rat}}{\equiv} 0. \end{aligned}$$

In this event, there can be only one such f (up to scale), as

$$(f) = -D = (g) \quad \Longrightarrow \quad (f/g) = 0 \quad \Longrightarrow \quad f/g \text{ constant.}$$

Together with similar reasoning in case (ii), and assuming Abel, this argument proves

Proposition 28.1.1. *(i) If* $\deg D = 0$*, then* $\ell(D) = 0$ *or* 1*; and*

$$AJ(D) = 0 \iff D \overset{rat}{\equiv} 0 \iff \ell(D) = 1.$$

(ii) If $\deg D = 2g - 2$*, then* $i(D) = 0$ *or* 1*; and*

$$AJ(K - D) = 0 \iff D \overset{rat}{\equiv} K \iff \ell(D - K) = 1 \iff i(D) = 1.$$

Another point of contact with the last few chapters comes in the context of canonical and hyperelliptic curves. First, fix $q \in M$ and look at the mapping

$$\begin{aligned} u_q :\ M &\longrightarrow J(M) \\ p \ \longmapsto\ AJ([p] - [q]) &= \begin{pmatrix} \int_q^p \omega_1 \\ \vdots \\ \int_q^p \omega_g \end{pmatrix} \text{ mod } \Lambda_M. \end{aligned}$$

Assuming Abel's theorem, we have (for genus ≥ 1)

Proposition 28.1.2. *(a)* u_q *is injective;*

(b) its differential yields the canonical map; and

(c) if M *is hyperelliptic and* q *is a fixed part of* $\jmath$*, then* $u_q(M)$ *is symmetric with respect to the involution* $\underline{u} \mapsto -\underline{u}$ *of* $J(M)$*.*

Proof. (a) Assuming $p_1 \neq p_2$ and $u_q(p_1) = u_2(p_2)$, we have

$$\begin{aligned} &AJ([p_1] - [p_2]) = 0 \\ \overset{\text{Abel}}{\Longrightarrow} \quad &\exists\, f \in \mathcal{K}(M)^* \text{ with } (f) = [p_1] - [p_2] \\ \Longrightarrow \quad &f : M \underset{(\cong)}{\longrightarrow} \mathbb{P}^1 \text{ has degree one,} \end{aligned}$$

contradicting $g \geq 1$.

(b) Given $\omega \in \Omega^1(M)$, we can consider $\omega(p) \in T_p^*M$. By the fundamental theorem of calculus, the differential

$$du_q(p): \ T_pM \ \longrightarrow \ T_{u_q(p)}J(M) \cong \mathbb{C}^g$$

is given by $(\omega_1(p), \ldots, \omega_g(p))$. (That is, if $\omega_i \overset{\text{loc}}{=} f_i(z)dz$, then $du_q(p)$ sends $\frac{\partial}{\partial z}|_p \mapsto (f_1(0), \ldots, f_g(0)) \in \mathbb{C}^g$.) This associates a line in $\mathbb{C}^g$ to each $p \in M$; projectivizing clearly recovers $\varphi_K : M \to \mathbb{P}^{g-1}$.

(c) Using $\jmath((x,y)) = (x,-y)$, we have

$$u_q(\jmath(p)) = \begin{pmatrix} \int_{q=\jmath(q)}^{\jmath(p)} \frac{dx}{y} \\ \vdots \\ \int_{\jmath(q)}^{\jmath(p)} \frac{x^{g-1}dx}{y} \end{pmatrix} = \begin{pmatrix} \int_q^p \jmath^* \frac{dx}{y} \\ \vdots \\ \int_q^p \jmath^* \frac{x^{g-1}dx}{y} \end{pmatrix}$$
$$= \begin{pmatrix} -\int_q^p \frac{dx}{y} \\ \vdots \\ -\int_q^p \frac{x^{g-1}dx}{y} \end{pmatrix} = -u_q(p).$$

$\square$

In fact, in the hyperelliptic case it is clear from (c) that the fixed points of $\jmath$ map to 2-torsion points of $J(M)$.

28.2. Differential forms of the third kind

There is a classical (and passé) terminology for meromorphic differential forms on a Riemann surface: "first kind" refers to holomorphic forms; "second kind" to meromorphic forms with trivial residues (and hence no *simple* poles); and "third kind" to everything else. In this section we'll pursue a method for constructing functions with a given divisor (if possible). The title refers to the essential use we shall make of meromorphic forms with prescribed (nonzero) residues.

Given $p, q \in M$

$$i(-[p]-[q]) \ = \ g-(-2)-1+\underbrace{\ell(-[p]-[q])}_{0} \ = \ g+1 \ (>g),$$

so there exists $\omega \in \mathfrak{I}(-[p]-[q]) \backslash \Omega^1(M)$. By the residue theorem,

$$0 \ = \ \underbrace{Res_p(\omega) + Res_q(\omega)}_{\text{both nonzero since poles simple}}$$

and so we can normalize ω so that

$$Res_p\omega = \frac{1}{2\pi\sqrt{-1}}\ , \quad Res_q\omega = \frac{-1}{2\pi\sqrt{-1}}.$$

For any meromorphic form η, write $(\eta) = (\eta)_0 - (\eta)_\infty$ where $(\eta)_0, (\eta)_\infty \geq 0$ are the zero- and polar-divisors.

LEMMA 28.2.1. *Given $D \in Div^0(M)$, there exists*

$$\eta_D \in \mathfrak{I}\left(-\sum_{p\in|D|}[p]\right)$$

such that

$$(\eta)_\infty = \sum_{p\in|D|}[p] \quad \textit{and} \quad Res_p\eta = \frac{ord_p(D)}{2\pi\sqrt{-1}}.$$

PROOF. Write $D = \sum[p_k] - \sum[q_k]$, and pick ω_k so that $Res_{p_k}\omega_k = -Res_{q_k}\omega_k = \frac{1}{2\pi\sqrt{-1}}$. Then add $\sum\omega_k =: \eta$. $\square$

Next let $D = \sum n_j[P_j]$ and η_D be as in Lemma 28.2.1, and consider a collection $\{\gamma_i\}_{i=1}^{2g}$ of closed paths with support $|\gamma_i| \subset M\backslash|D|$, such that their classes $\{[\gamma_i]\}_{i=1}^{2g} \subset H_1(M,\mathbb{Z})$ yield a basis.

LEMMA 28.2.2. *If*

$$\int_{\gamma_i}\eta_D \in \mathbb{Z} \ \ (\forall i), \tag{28.2.1}$$

then (fixing $Q \in M$)

$$f(P) := \exp\left(2\pi\sqrt{-1}\int_Q^P \eta_D\right)$$

yields a well-defined function $f \in \mathcal{K}(M)^*$ with $(f) = D$.

PROOF. We must check independence of path. Let C_j denote circular paths around the P_j. Given two paths $\overrightarrow{Q.P}$ and $\overrightarrow{Q.P}'$,

$$\overrightarrow{Q.P} - \overrightarrow{Q.P}' \ = \ \partial\Delta + \sum m_jC_j + \sum \ell_i\gamma_i$$

where Δ is a (real-2-dimensional) closed region in $M\backslash|D|$. Now

$$\int_{\partial\Delta}\eta_D = \int_\Delta d\eta_D = \int_\Delta 0 = 0,$$

$$\sum m_j\int_{C_j}\eta_D = \sum m_jn_j \in \mathbb{Z}$$

since $Res_{P_k}\eta_D = \frac{n_k}{2\pi\sqrt{-1}}$, and

$$\sum \ell_i \int_{\gamma_i} \eta_D \in \mathbb{Z}$$

by assumption (28.2.1). So for some $\mu \in \mathbb{Z}$

$$\frac{\exp\left(2\pi\sqrt{-1}\int_{\overrightarrow{Q.P}}\eta_D\right)}{\exp\left(2\pi\sqrt{-1}\int_{\overrightarrow{Q.P}'}\eta_D\right)} = \exp\left(2\pi\sqrt{-1}\mu\right) = 1,$$

and f is well-defined (and holomorphic) on $M\backslash|D|$.

For the divisor, let z be a holomorphic coordinate defined in a neighborhood of P_j (with $z(P_j) = 0$), and write

$$\eta \overset{\text{loc}}{=} \frac{n_j}{2\pi\sqrt{-1}}\frac{dz}{z} + h(z)dz$$

with h holomorphic. Without loss of generality, we can assume that Q lies in the neighborhood, with $z(Q) =: z_0$ (fixed) and $z(P) =: z$ (variable). Locally

$$\begin{aligned} f(z) &= \exp\left(2\pi\sqrt{-1}\int_Q^P \eta_D\right) \\ &= \exp\left(2\pi\sqrt{-1}\int_{z_0}^z h(w)dw\right)\cdot\exp\left(n_j\int_{z_0}^z \frac{dw}{w}\right) \\ &= H(z)\cdot\frac{\exp\left(n_j\log z\right)}{\exp\left(n_j\log z_0\right)} \end{aligned}$$

where H is holomorphic and nonvanishing in our neighborhood (being the exponential of something holomorphic). Finally, writing $H_0(z) = \frac{H(z)}{z_0^{n_j}}$, the above

$$= H_0(z)\cdot z^{n_j}.$$

This makes it clear that f is meromorphic at P_j with

$$\nu_{P_j}(f) = n_j.$$

Doing this for each j, we conclude

$$(f) = \sum n_j[P_j] = D.$$

□

In the next chapter we will take $\gamma_1, \ldots, \gamma_{2g}$ to yield a *symplectic* basis for H_1, i.e.

$$\gamma_i \cdot \gamma_{g+j} = \delta_{ij} \qquad \text{(Kronecker delta)}$$

for $1 \leq i, j \leq g$. This is the situation pictured in §20.1. It turns out that the period vectors $\pi_1, \ldots, \pi_g$ associated to $\gamma_1, \ldots, \gamma_g$ are actually linearly independent over $\mathbb{C}$,[2] and so the $g \times g$ matrix they form is invertible. Applying the inverse matrix to $\omega_1, \ldots, \omega_g$, we may replace them by $\{\omega_j\}$ satisfying

$$\int_{\gamma_i} \omega_j = \delta_{ij}.$$

Given D and η_D as in Lemma 28.2.1, then, we can modify η_D to

$$\widetilde{\eta_D} := \eta_D - \sum_{i=1}^{g} \left(\int_{\gamma_i} \eta_D \right) \omega_i$$

so that

$$\int_{\gamma_i} \widetilde{\eta_D} = 0$$

for $i = 1, \ldots, g$. We will prove that

$$\text{(28.2.2)} \qquad AJ(D) = 0 \implies \begin{array}{c} \text{there exists a further modification} \\ \widehat{\eta_D} := \widetilde{\eta_D} + \sum_{j=1}^{g} \mu_j \omega_j \\ \text{with } \int_{\gamma_i} \widehat{\eta_D} \in \mathbb{Z} \ \ (i = 1, \ldots, 2g), \end{array}$$

so as to affirm condition (28.2.1) (for $\widehat{\eta_D}$). To attack (28.2.2), we need the *Riemann bilinear relations*, our next topic.

The problems below are only loosely related to the material of his chapter. The second one is rather open ended!

Exercises

(1) This problem, in which you will prove a version of Abel's theorem for a "singular" cubic (not its normalization), is only loosely related to the chapter. Think of the cubic C as $\mathbb{P}^1$ with 0 identified to ∞ and coordinate z. We consider $\Omega^1(C)$ to be spanned by $\frac{dz}{z}$ (even though it isn't holomorphic) and $H_1(C, \mathbb{Z})$ by S^1= unit curcle. Divisors must avoid the singularity, and meromorphic functions f must have $f(0) = f(\infty) \neq 0, \infty$.

[2]to be proved in §29.1. Since the vectors are non-real, this doesn't follow from independence over $\mathbb{R}$ (which we already have).

(a) What is $J(C)$?

(b) Compute $AJ(D)$ for $D = \sum n_i[z_i]$, $\sum n_i = 0$.

(c) Show $0 = AJ((f)) \in J(C)$.

(d) Formulate and prove the injectivity statement (Abel's theorem). [Hint: the proof will use what you did in (c), even though it's a "converse", and so needn't be long.]

(2) Let M be a Riemann surface and $\Sigma \subset M$ a (nonempty) finite set of points.

(a) Define divisors on the relative variety (M, Σ) to be formal sums $\sum n_i[p_i]$ where no p_i lies in Σ; two of these are rationally equivalent if their difference is the divisor of $f \in \mathcal{K}(M)^*$ which is 1 on all points of Σ. Construct an AJ map and Jacobian for (M, Σ). [Hint: the case $M = \mathbb{P}^1$, $\Sigma = \{0, \infty\}$ should recover the results of exercise (1).]

(b) Next consider the complement $M \backslash \Sigma$. We define divisors by $Div(M)/Div(\Sigma)$, and rational equivalence by taking divisors of meromorphic functions on M (and ignoring any poles/zeroes on Σ, since that information is quaotiented out). Construct an AJ map and Jacobian for $M \backslash \Sigma$. [Hint: note that there is no such thing as degree of a divisor, since the points in Σ effectively have arbitrary coefficients. Or rather, using these points, you can make the degree of any divisor zero! This should have some bearing on your choice of path.]

CHAPTER 29

Abel's Theorem, part II

As suggested at the end of the previous chapter, on any Riemann surface M, we can produce a perfect[1] pairing on the level of homology

$$\langle\ ,\ \rangle :\ H_1(M,\mathbb{Z}) \times H_1(M,\mathbb{Z}) \to \mathbb{Z} \tag{29.0.1}$$

by intersecting 1-cycles.[2] With respect to the basis $\{\gamma_j\}_{j=1}^{2g}$ described there, this has $(2g \times 2g)$ matrix[3]

$$Q = \begin{pmatrix} 0 & \mathbb{I}_g \\ -\mathbb{I}_g & 0 \end{pmatrix}.$$

We can use this pairing to produce an isomorphism of dual spaces

(29.0.2)

$$H_1(M,\mathbb{C}) = H_1(M,\mathbb{Z}) \otimes \mathbb{C} \xrightarrow{\cong} Hom(H_1(M,\mathbb{Z}),\mathbb{C}) = H^1(M,\mathbb{C})$$

which is a special case of *Poincaré duality.*

Recalling the isomorphisms

$$\Omega^1(M) \oplus \overline{\Omega^1(M)} \xrightarrow{\cong} H^1_{dR}(M,\mathbb{C}) \xrightarrow{\cong} H^1(M,\mathbb{C}),$$

there is also a pairing (the "cup-product")

$$H^1(M,\mathbb{C}) \times H^1(M,\mathbb{C}) \to \mathbb{C}$$

[1]this means that the (bilinear) pairing is described (with respect to an integral basis of $H_1(C,\mathbb{Z})$) by an integrally invertible, i.e. unimodular, matrix.

[2]More precisely, any two given homology classes have representative 1-cycles (say, α, β) which intersect transversely. At each intersection point p there is a local holomorphic coordinate $z = x + iy$, and the tangent vectors $\underline{v}_\alpha$ and $\underline{v}_\beta$ to the 1-cycles (which have well-defined directions) can be wedged to produce an element $\underline{v}_\alpha \wedge \underline{v}_\beta = \xi_p \frac{\partial}{\partial x} \wedge \frac{\partial}{\partial y} \in \bigwedge^2 T_pM$ ($\xi_p \neq 0$). The intersection is called positively or negatively oriented depending upon the sign of ξ_p, and the intersection number $\langle [\alpha],[\beta]\rangle$ is the number of positive intersection points minus the number of negative ones.

[3]having this particular intersection form is the definition of a *symplectic basis* of $H_1(M,\mathbb{Z})$

induced on the level of 1-forms by

$$(\omega, \eta) \longmapsto \int_M \omega \wedge \eta.$$

Notice that since two holomorphic forms wedge to zero, this pairing restricts to zero on $\Omega^1(M) \times \Omega^1(M)$ (and $\overline{\Omega^1(M)} \times \overline{\Omega^1(M)}$). Yet another pairing (the "cap-product")

$$H_1(M, \mathbb{Z}) \times H^1(M, \mathbb{C}) \to \mathbb{C}$$

is induced by

$$(\gamma, \omega) \mapsto \int_\gamma \omega.$$

The restriction of this pairing to $H_1(M, \mathbb{Z}) \times \Omega^1(M)$ is caputred by the period matrix of chapter 20. An important fact is that, under (28.0.2), both of these integration-induced products are nothing but complex-linear extensions of (28.0.1).

Assuming this compatibility, we can quickly derive the *Riemann bilinear relations* as follows. If for any closed 1-form $\varphi \in \Omega^1(M) \oplus \overline{\Omega^1(M)}$, we write

$$\pi_j(\varphi) := \int_{\gamma_j} \varphi,$$

then (28.0.2) identifies

$$[\varphi] = \sum_{j=1}^{g} (\pi_j(\varphi)[\gamma_{j+g}] - \pi_{j+g}(\varphi)[\gamma_j]) \tag{29.0.3}$$

in $H^1(M, \mathbb{C})$, i.e. as functionals on homology. One has for $\omega, \varphi \in \Omega^1(M)$

$$0 = \int_M \omega \wedge \varphi = -\sum_{j=1}^{g} (\pi_j(\varphi)\pi_{j+g}(\omega) - \pi_{j+g}(\varphi)\pi_j(\omega)) \tag{29.0.4}$$

by writing $\int_M \omega \wedge \varphi = \langle [\omega], [\varphi] \rangle$ and expanding both classes as in (28.0.3). Similar reasoning together with the local computation

$$idz \wedge d\bar{z} = i(dx + idy) \wedge (dx - idy) = i(-2idx \wedge dy) = 2dx \wedge dy,$$

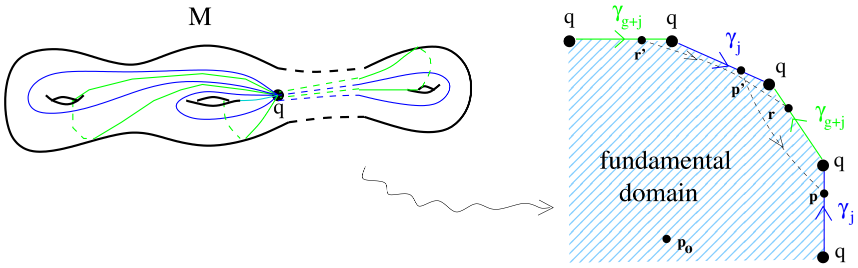
leads to

$$0 < i \int_M \omega \wedge \bar{\omega} = -i \sum_{i=1}^{g} \left(\overline{\pi_j(\omega)} \pi_{j+g}(\omega) - \overline{\pi_{j+g}(\omega)} \pi_j(\omega) \right). \tag{29.0.5}$$

This is all meant as motivation, though it can be made completely rigorous. We'll start the first section with a more concrete, classical proof of (29.0.4-5), without the compatibility assumptions on the three bilinear pairings.

29.1. Derivation of the Riemann Bilinear Relations

We start by cutting M open to get the "fundamental domain", a simply-connected closed region $\mathfrak{F}$



with boundary $\partial\mathfrak{F}$. (Only a piece of it is shown in the picture.) Let p_0 in the interior of $\mathfrak{F}$ be fixed. Given $\omega \in \Omega^1(M)$,

$$u(p) := \int_{p_0}^{p} \omega$$

then yields a well-defined (single-valued) holomorphic[4] function on $\mathfrak{F}$. If we take a second holomorphic form $\varphi \in \Omega^1(M)$, then

$$d(u\varphi) = \omega \wedge \varphi = 0.$$

That is, $u\varphi$ is a closed holomorphic form on $\mathfrak{F}$ with the consequence that

$$0 \;=\; \int_{\mathfrak{F}} d(u\varphi) \;=\; \int_{\partial\mathfrak{F}} u\varphi$$

by Stokes's theorem. Now, the picture above tells us that $\partial\mathfrak{F}$ is the composition of paths

$$\gamma_{2g}^{-1}\gamma_g^{-1}\gamma_{2g}\gamma_g \cdot \cdots \cdot \gamma_{g+2}^{-1}\gamma_2^{-1}\gamma_{g+2}\gamma_2\gamma_{g+1}^{-1}\gamma_1^{-1}\gamma_{g+1}\gamma_1,$$

[4]to be holomorphic on a closed set means that the function extends to a holomorphic function on a slightly larger open set (which, in this case, would live on the universal cover of M).

written right to left (with inverse meaning the reverse direction). So the last integral becomes

$$= \sum_{j=1}^{g}\{ \int_{\gamma_j} \underbrace{(u(p)-u(p'))}_{\int_{p'}^{q}\omega - \int_{\gamma_{g+j}}\omega + \int_{q}^{p}\omega} \varphi \; + \; \int_{\gamma_{j+g}} \underbrace{(u(r)-u(r'))}_{\int_{r'}^{q}\omega + \int_{\gamma_j}\omega + \int_{q}^{r}\omega} \varphi \; \}$$

and, noting that $\int_{p'}^{q}\omega = \int_{p}^{q}\omega = -\int_{q}^{p}\omega$, this

$$= \sum_{j=1}^{g}\left(-\int_{\gamma_{g+j}}\omega \int_{\gamma_j}\varphi + \int_{\gamma_{g+j}}\varphi \int_{\gamma_j}\omega \right).$$

Replacing φ by $i\bar{\omega}$, essentially the same computation yields

$$0 < i\int_{\mathfrak{F}} \underbrace{\omega \wedge \bar{\omega}}_{d(u\bar{\omega})} = \int_{\partial\mathfrak{F}} u(i\bar{\omega}) = i\sum_{j=1}^{g}\left(-\int_{\gamma_{j+g}}\omega \int_{\gamma_j}\bar{\omega} + \int_{\gamma_{g+j}}\bar{\omega}\int_{\gamma_j}\omega \right).$$

So we have recovered (29.0.4-5).

To reformulate this in matrix terms for any symplectic basis $\{\gamma_j\}_{j=1}^{2g}$ of $H_1(M,\mathbb{Z})$ and any basis $\{\omega_i\}_{i=1}^{g}$ of $\Omega^1(M)$, notice that the $(k,\ell)^{\text{th}}$ entry of[5]

$$\Pi \cdot Q \cdot {}^t\Pi \;=\; \begin{pmatrix} \uparrow & & \uparrow \\ \pi_1 & \cdots & \pi_{2g} \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} 0 & \mathbb{I}_g \\ -\mathbb{I}_g & 0 \end{pmatrix} \begin{pmatrix} \leftarrow & \pi_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \pi_{2g} & \rightarrow \end{pmatrix}$$

$$= \begin{pmatrix} \uparrow & & \uparrow & \uparrow & & \uparrow \\ -\pi_{g+1} & \cdots & -\pi_{2g} & \pi_1 & \cdots & \pi_g \\ \downarrow & & \downarrow & \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} \leftarrow & \pi_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \pi_{2g} & \rightarrow \end{pmatrix}$$

is

$$\sum_{j=1}^{g} \left(\pi_j(\omega_k)\pi_{g+j}(\omega_\ell) - \pi_j(\omega_\ell)\pi_{g+j}(\omega_k)\right)$$

which is zero by (29.0.4); so

$$\Pi \cdot Q \cdot {}^t\Pi = 0. \tag{29.1.1}$$

Similarly,

$$\sqrt{-1}\Pi \cdot Q \cdot {}^t\bar{P}i > 0 \tag{29.1.2}$$

in the sense that ${}^t\underline{x}(\sqrt{-1}\Pi\cdot Q\cdot{}^t\bar{\Pi})\underline{\bar{x}} \in \mathbb{R}_{>0}$ for any $\underline{x} \in \mathbb{C}^g$. In particular, the diagonal entries of (29.1.2) are positive real.

[5]recall from Chapter 20 that π_j is the complex g-vector with i^{th} entry $\pi_j(\omega_i)$

REMARK 29.1.1. Consider any two symplectic integral bases $\Gamma = \{\gamma_j\}$ and $\Gamma = \{\gamma'_j\}$ (thought of as row-vectors), so that

$$\Gamma' = \Gamma A$$

for some $A \in SL_{2g}(\mathbb{Z})$. Applying the basis $\{\omega_i\}$ (viewed as a column-vector of 1-forms) on the left yields

$$\Pi' = \Pi A.$$

Furthermore, since both bases are symplectic we have $Q = {}^t\Gamma \cdot \Gamma$ and

$$Q = {}^t\Gamma' \cdot \Gamma' = {}^tA^t\Gamma\Gamma A = {}^tAQA;$$

that is, A belongs to the symplectic group $Sp_{2g}(\mathbb{Z})$. It is for this reason that (29.1.1-2) are compatible with change of symplectic basis: e.g., assuming (29.1.1) we have

$$\Pi' Q^t\Pi' = \Pi A Q^t A^t \Pi = \Pi Q^t \Pi = 0.$$

Now thinking of the $g \times 2g$ period matrix as two $g \times g$ blocks, viz.

$$\Pi = \begin{pmatrix} \mathcal{A} & \mathcal{B} \end{pmatrix}, \tag{29.1.3}$$

we have

$$\Pi Q^t \Pi = \begin{pmatrix} \mathcal{A} & \mathcal{B} \end{pmatrix} \begin{pmatrix} & \mathbb{I}_g \\ -\mathbb{I}_g & \end{pmatrix} \begin{pmatrix} {}^t\mathcal{A} \\ {}^t\mathcal{B} \end{pmatrix} = \mathcal{A} \cdot {}^t\mathcal{B} - \mathcal{B} \cdot {}^t\mathcal{A}$$

and

$$\Pi Q^t \bar{\Pi} = \mathcal{A} \cdot {}^t\bar{\mathcal{B}} - \mathcal{B} \cdot {}^t\bar{\mathcal{A}}.$$

In these terms, (29.1.1) reads

$$\mathcal{A} \cdot {}^t\mathcal{B} = \mathcal{B} \cdot {}^t\mathcal{A} \tag{29.1.4}$$

while (29.1.2) becomes

$$\sqrt{-1}{}^t\underline{v}(\mathcal{A}^t\bar{\mathcal{B}} - \mathcal{B}^t\bar{\mathcal{A}})\bar{\underline{v}} > 0 \quad (\forall \underline{v} \in \mathbb{C}^g). \tag{29.1.5}$$

If ${}^t\mathcal{A}$ has nonzero kernel, then there exists $\underline{v} \in \mathbb{C}^g$ satisfying ${}^t\mathcal{A}\underline{v} = 0$, hence ${}^t\underline{v}\mathcal{A} = 0$ and ${}^t\bar{\mathcal{A}}\bar{\underline{v}} = 0$, contradicting (29.1.5). It follows that $\mathcal{A}$ is invertible, and so we have proved the statement on $\mathbb{C}$-linear independence asserted at the end of §28.2.

Applying $\mathcal{A}^{-1}$ to the left of Π amounts to a change of the basis $\{\omega_i\}$ for $\Omega^1(M)$, viz.[6]

$$\mathcal{A}^{-1}\Pi = \mathcal{A}^{-1}\begin{pmatrix} \omega_1 \\ \vdots \\ \omega_g \end{pmatrix}\begin{pmatrix} \gamma_1 & \cdots & \gamma_{2g} \end{pmatrix} = \begin{pmatrix} \omega'_1 \\ \vdots \\ \omega'_g \end{pmatrix}\begin{pmatrix} \gamma_1 & \cdots & \gamma_{2g} \end{pmatrix}.$$

If we apply it to (28.1.3), then we get

$$\Pi' := \mathcal{A}^{-1}\Pi = \begin{pmatrix} \mathbb{I}_g & \mathcal{A}^{-1}\mathcal{B} \end{pmatrix}.$$

We can therefore always assume that $\{\omega_i\}$ is chosen so that

$$\Pi = \begin{pmatrix} \mathbb{I}_g & Z \end{pmatrix},$$

again as claimed in §28.2. The bilinear relations (29.1.4-5) simplify to

$$(29.1.6) \qquad \begin{cases} Z = {}^tZ \\ \sqrt{-1}(\bar{Z} - Z) > 0 \end{cases},$$

which in particular tell us that the imaginary part $Im(Z)$ is a positive-definite, real symmetric matrix.

29.2. Proof of Abel's Theorem

With the holomorphic basis as normalized above, we can now quickly establish (28.2.2) and hence (28.0.1). Write $D = \sum n_i[P_i]$ (with $\sum n_i = 0$) and let $\varphi := \widetilde{\eta_D}$ be as §28.2, so that

$$(29.2.1) \qquad Res_{P_i}(\varphi) = \frac{n_i}{2\pi\sqrt{-1}} \quad (\forall i)$$

and

$$\int_{\gamma_j} \varphi = 0 \quad (j = 1, \ldots, g).$$

For each $k = 1, \ldots, g$ set

$$u_k(p) := \int_{p_0}^{p} \omega_k$$

on $\mathfrak{F}$, and let Γ be a 1-chain (sum of paths) with $\partial\Gamma = D$. Then noting $D = \sum n_i([P_i] - [p_0])$, we have

$$\int_\Gamma \omega_k = \sum_i n_i u_k(P_i) = 2\pi\sqrt{-1} \sum_{p \in |D|} Res_p(u_k \varphi)$$

[6]here the "product" of ω_i and γ_j is just the integral $\int_{\gamma_j} \omega_i$

which by the Residue Theorem

$$= \int_{\partial\mathfrak{F}} u_k\varphi \overset{\S 29.1}{=} \sum_j (\underbrace{\pi_j(\omega_k)}_{\delta_{jk}}\pi_{g+j}(\varphi) - \pi_{g+j}(\omega_k)\underbrace{\pi_j(\varphi)}_{0}) = \pi_{g+k}(\varphi).$$

If $AJ(D) = 0$ then there are integers m_j ($j = 1, \ldots, 2g$) such that for every k

$$\int_\Gamma \omega_k = \sum_{j=1}^{2g} m_j \int_{\gamma_j} \omega_k.$$

Using $\int_{\gamma_j} \omega_k = \delta_{jk}$ and $Z = {}^tZ$ (from (29.1.6)), this

$$= m_k + \sum_{j=1}^{g} m_{j+g}\pi_{j+g}(\omega_k) = m_k + \sum_{j=1}^{g} m_{j+g}\pi_{k+g}(\omega_j).$$

Now

$$\hat{\varphi} := \varphi - \sum_{j=1}^{g} m_{j+g}\omega_j$$

is still an element of $\mathfrak{I}(-\sum_{p\in|D|}[p])$ satisfying (29.2.1). Moreover, for $k \in \{1, \ldots, g\}$

$$\pi_{k+g}(\hat{\varphi}) = \pi_{k+g}(\varphi) - \sum_{j=1}^{g} m_{j+g}\pi_{k+g}(\omega_j)$$

$$= \int_\Gamma \omega_k - \sum_{j=1}^{g} m_{j+g}\pi_{k+g}(\omega_j) = m_k \in \mathbb{Z}$$

and

$$\pi_k(\hat{\varphi}) = \underbrace{\pi_k(\varphi)}_{0} - \sum_{j=1}^{g} m_{j+g}\underbrace{\pi_k(\omega_j)}_{\delta_{kj}} = -m_{k+g} \in \mathbb{Z}.$$

By Lemma 28.2.2, $\exp\left(2\pi\sqrt{-1}\int\hat{\varphi}\right)$ now gives a meromorphic function with $(f) = D$.

29.3. Proof of Jacobi Inversion

To show that AJ is surjective, we will study the image of a certain class of (degree zero) divisor on M, namely those of the form

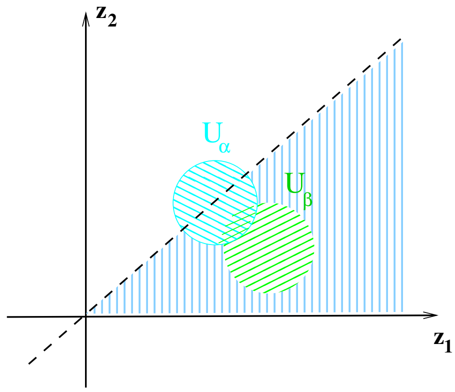
$$[p_1] + \cdots + [p_d] - d[q]$$

given some fixed point $q \in M$ and natural number d. Such divisors are obviously in 1-to-1 correspondence with unordered d-tuples of points on M, in other words with elements of the d^{th} symmetric power[7]

$$Sym^d M \;:=\; \frac{\overbrace{M \times \cdots \times M}^{d \text{ copies}}}{(p_1, \ldots, p_d) \underset{\forall\, \sigma \in \mathfrak{S}_d}{\sim} (p_{\sigma(1)}, \ldots, p_{\sigma(d)})}.$$

In order to be able to use complex analytic techniques we need to put the structure of a d-dimensional complex manifold on this.[8]

To get the idea of how this works, take $d = 2$ and consider $\mathbb{C}$ instead of a (compact) Riemann surface. The symmetric square $Sym^2\mathbb{C}$ is the quotient of $\mathbb{C} \times \mathbb{C}$ by the involution $(z_1, z_2) \mapsto (z_2, z_1)$. What causes difficulty is the locus consisting of its fixed points, i.e. the diagonal line. Take two small open sets in $Sym^2 M$, one which intersects the diagonal and one which does not:



Clearly (z_1, z_2) give local holomorphic coordinates on U_α. On U_β, they are not well-defined, but their elementary symmetric polynomials $\sigma_1(z_1, z_2) = z_1 + z_2$ and $\sigma_2(z_1, z_2) = z_1 z_2$ are. Moreover, these functions generate all polynomials in z_1, z_2 which are invariant under the involution and hence well-defined on $U_\beta \subset Sym^2\mathbb{C}$. Taking $(w_1, w_2) := (z_1 + z_2, z_1 z_2)$ as the holomorphic coordinates there,[9] the transition function $\Phi_{\beta\alpha}$ is then just (σ_1, σ_2). To see that this is invertible on $U_{\alpha\beta}$, notice that in U_α the diagonal is defined by $w_1^2 = 4w_2$ (since $z_1 = z_2 \iff (z_1 + z_2)^2 = 4z_1 z_2$). Since $U_{\alpha\beta}$ avoids this locus (and is simply connected), $\sqrt{w_1^2 - 4w_2}$ is well defined there and we can

[7] elements are written either $p_1 + \cdots + p_d$ or $\{p_1, \ldots, p_d\}$

[8] Had we started with M *itself* of dimension > 1, its symmetric powers would be *singular* complex analytic spaces, hence *not* manifolds. So what happens next is special for $\dim(M) = 1$.

[9] we could in fact take these as global coordinates, but this situation won't generalize to M

define $\Phi_{\alpha\beta}$ by

$$\begin{cases} z_1 = \frac{w_1+\sqrt{w_1^2-4w_2}}{2} \\ z_2 = \frac{w_2-\sqrt{w_1^2-4w_2}}{2} \end{cases} .$$

More generally, in a neighborhood of

$$\{\underbrace{q_1,\ldots,q_1}_{k_1 \text{ times}};\ \ldots\ ;\underbrace{q_\ell,\ldots,q_\ell}_{k_\ell \text{ times}}\} \in Sym^d M$$

(where $\sum_{j=1}^{\ell} k_j = d$), the local coordinate system is given in terms of holomorphic coordinates z_j on M near each q_j, by

$$\{\underbrace{p_1,\ldots,p_{k_1}}_{\text{all near } q_1};\ \ldots\ ;\underbrace{p_{d-k_\ell+1},\ldots,p_d}_{\text{all near } q_\ell}\} \longmapsto$$

$$(\sigma_1(z_1(p_1),\ldots,z_1(p_{k_1})),\ldots,\sigma_{k_1}(z_1(p_1),\ldots,z_1(p_{k_1}))\ ;\ \ldots\ ;$$
$$\sigma_1(z_\ell(p_{d-k_\ell+1}),\ldots,z_\ell(p_d)),\ldots,\sigma_{k_\ell}(z_\ell(p_{d-k_\ell+1}),\ldots,z_\ell(p_d)))\,.$$

Inelegant, but it gets the job done.

Now let D be any divisor of degree d on M, and consider the mapping

$$\alpha_D : \mathbb{P}(\mathfrak{L}(D)) \to Sym^d M$$

which sends (for $f \in \mathfrak{L}(D)$)

$$[f] \mapsto (f) + D.$$

(Here $(f) + D \geq 0$ by definition, and $\deg((f) + D) = \deg D = d$; so $(f) + D$ is of the form $[p_1] + \cdots + [p_d]$. The map sends to the projective equivalence class $[f]$, i.e. "f up to a constant multiple", to $\{p_1,\ldots,p_d\}$.)

Lemma 29.3.1. *α_D is (a) injective and (b) holomorphic.*

Definition 29.3.2. The *linear system*[10] $|D|$ consists of all effective divisors on M rationally equivalent to D. The Lemma evidently realizes $|D| = \text{image}(\alpha_D)$ as a subvariety of $Sym^d M$ isomorphic to $\mathbb{P}^{\ell(D)-1}$.

Proof. (of Lemma)

(a) $(f) + D = (g) + D \implies (f) = (g) \implies (f/g) = 0 \implies f/g$ constant $\implies [f] = [g]$.

(b) To show α_D holomorphic in a neighborhood of $[f_0]$, augment f_0 to a basis $\{f_0, f_1, \ldots, f_{\ell(D)}\} \subset \mathfrak{L}(D)$ and write $f_{\underline{\mu}} := f_0 + \sum_{j=1}^{\ell(D)} \mu_j f_j$

[10]the notation $|D|$ is unfortunately standard for both the linear system and the support of D, two completely different concepts!

so that $\{\mu_j\}_{j=1}^{\ell(D)}$ are the local holomorphic coordinates (on some small $U \subset \mathfrak{L}(D)$). Let $p \in |D| \cup |(f_0)|$, with open neighborhood $\mathcal{N}_p \subset M$ and local coordinate z (i.e. $\mathrm{ord}_p(z) = 1$). Set $k := \mathrm{ord}_p(f_0) + \mathrm{ord}_p(D)$, and $\mathcal{W}_{f_0,p} := Sym^k\mathcal{N}_p$ with coordinates $\sigma_1(z_1, \ldots, z_k), \ldots, \sigma_k(z_1, \ldots, z_k)$. We must show that the composition

$$\begin{array}{ccccccc} U & \longrightarrow & \mathcal{O}(\mathcal{N}_p) & \longrightarrow & \mathcal{W}_{f_0,p} & \hookrightarrow & \mathbb{C}^k \\ \underline{\mu} & \longmapsto & f_{\underline{\mu}} z^{\mathrm{ord}_p D} & \longmapsto & \begin{array}{c} \left(f_{\underline{\mu}} z^{\mathrm{ord}_p D}\right)\Big|_{\mathcal{N}_p} \\ \| \\ p_1(\underline{\mu}) + \cdots + p_k(\underline{\mu}) \end{array} & \longmapsto & \begin{pmatrix} \sigma_1\left(z(p_1(\underline{\mu})), \ldots, z(p_k(\underline{\mu}))\right) \\ \vdots \\ \sigma_k\left(z(p_1(\underline{\mu})), \ldots, z(p_k(\underline{\mu}))\right) \end{pmatrix} \end{array}$$

is holomorphic, which in turn boils down to the statement that each σ_ℓ is holomorphic in each μ_j. For $k = 1$, this is the holomorphic implicit function theorem; for $k > 1$, it is this together with Rouché and the Riemann extension theorem in a manner familiar from previous chapters. □

DEFINITION 29.3.3. An effective degree d divisor D (viewed as an element of $Sym^d M$) is called *general* $\iff D = [p_1] + \cdots + [p_d]$ with the $\{p_j\}$ distinct points of M.

Now look at the "Abel-Jacobi" mapping

$$\begin{aligned} u^d : \ Sym^d M \ &\longrightarrow \ J(C) \\ [p_1] + \cdots + [p_d] \ &\longmapsto \ AJ\left(\sum_{j=1}^{d}[p_j] \ - d[q]\right), \end{aligned}$$

where $q \in M$ is fixed. This is shown to be holomorphic by using the fundamental theorem of calculus at general D, then applying the Osgood and Riemann extension theorems. (Boundedness is clear by taking a local lifting of the image of u^d to $\mathbb{C}^g$.)

The next result does not require D to be general.

LEMMA 29.3.4. *The fiber of u^d over $u^d(D)$ is $|D|$ ($\cong \mathbb{P}^{\ell(D)-1}$).*

PROOF. (For simplicity write u for u^d.)

$\underline{u^{-1}(u(D)) \subset |D|}$: $u(E) = u(D) \implies AJ(E - D) = 0 \overset{\text{Abel}}{\implies} E - D$ is the divisor of some $f \in \mathcal{K}(M)^* \implies (f) + D = E \geq 0$ (since $E \in Sym^d M$) $\implies f \in \mathfrak{L}(D) \implies E = \alpha_D(f) \in \text{image}(\alpha_D) = |D|$.

$\underline{u^{-1}(u(D)) \supset |D|}$: Given $E \in |D|$, there exists $f \in \mathfrak{L}(D)$ such that $E = (f) + D \implies E - D = (f) \overset{\text{rat}}{\equiv} 0 \implies 0 = AJ(E-D) \implies u(E) = u(D) \implies E \in u^{-1}(u(D))$. $\square$

If $D = [p_1] + \cdots + [p_d]$ is general, then writing z_j for local coordinates about each p_j,

$$(du^d)_D : \, T_D\left(Sym^d M\right) \longrightarrow T_{u(D)}(J(M))$$

is computed by the matrix

$$\left.\begin{pmatrix} \frac{\partial}{\partial z_1} \sum_{i=1}^d \int_q^{z_i} \omega_1 & \cdots & \frac{\partial}{\partial z_1} \sum_{i=1}^d \int_q^{z_i} \omega_g \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial z_d} \sum_{i=1}^d \int_q^{z_i} \omega_1 & \cdots & \frac{\partial}{\partial z_d} \sum_{i=1}^d \int_q^{z_i} \omega_g \end{pmatrix}\right|_{\{p_1,\ldots,p_d\}}.$$

If we write locally (about each p_j) $\omega_i \overset{\text{loc}}{=} f_i(z_j)dz_j$, this

$$= \begin{pmatrix} f_1(p_1) & \cdots & f_g(p_1) \\ \vdots & \ddots & \vdots \\ f_1(p_d) & \cdots & f_g(p_d) \end{pmatrix} = \begin{pmatrix} \leftarrow & \widetilde{\varphi_K(p_1)} & \rightarrow \\ & \vdots & \\ \leftarrow & \widetilde{\varphi_K(p_d)} & \rightarrow \end{pmatrix},$$

where φ_K is the canonical map and $\widetilde{\varphi_K(p_j)} \in \mathbb{C}^g$ is a "lift" of $\varphi_K(p_j) \in \mathbb{P}^{g-1}$. (For $d = 1$ this is just Proposition 28.1.2(b).) From this we see that

$$\operatorname{rank}\left((du^d)_D\right) = \dim\left(\operatorname{span}(\varphi_K(p_1), \ldots, \varphi_K(p_d))\right) + 1,$$

where "span" means the projective linear span in $\mathbb{P}^{g-1}$. Taking $d = g$, we now have the following claim:

LEMMA 29.3.5. *rank*$((du^g)_D) = g$ *for* $D = [p_1] + \cdots + [p_g] \in Sym^g M$ *generic,*[11] *i.e. in some Zariski open subset of* $Sym^d M$.

PROOF. Choose $p_1, \ldots, p_g$ distinct with $\operatorname{span}(\varphi_K(p_1), \ldots, \varphi_K(p_g)) =$ all of $\mathbb{P}^{g-1}$. This is possible since the canonical map is always nondegenerate by Theorem 27.3.2(a). Consequently $\operatorname{rank}((du^g)_D) = g$, and this holds more generally for D in an algebraic open set. This is because its failure is equivalent to $\det(du^g) = 0$, which is an algebraic condition which will hold on some codimension-one subvariety. $\square$

[11]"general" may not be quite enough — D may have to avoid a larger number of subvarieties of $Sym^g M$ then just the ones where two or more p_j's coincide.

Theorem 29.3.6. [Jacobi inversion] *u^g is surjective and generically injective.*

Proof. By Lemma 29.3.5, du^g is generically an isomorphism of tangent spaces. So u^g takes an open ball about a general point $D \in Sym^d M$ to an open ball. But u^g is continuous and $Sym^g M$ compact, so image(u^g) is both a closed analytic subvariety of $J(M)$ and contains an open ball, and is therefore all of $J(M)$ (which is connected).

Since at a generic D, du^g is (in particular) injective, we see that any such D is an isolated point of $(u^g)^{-1}\{u^g(D)\}$. But the latter is a projective space by Lemma 29.3.4, and so the only way D is isolated is if $(u^g)^{-1}\{u^g(D)\}$ is isomorphic to $\mathbb{P}^0$, i.e. is just D itself. $\square$

Finally, to address (28.0.2) head-on, surjectivity of AJ follows from the diagram

(29.3.1)
$$\begin{tikzcd} Div^0(M) \arrow[r, "AJ"] & J(M) \\ & Sym^g M \arrow[ul, hook, "D \mapsto D - g[q]"] \arrow[u, "u^g"'] \end{tikzcd} .$$

29.4. A final remark on moduli

For any Riemann surface M (of genus ≥ 1) with given symplectic basis of $H_1(M, \mathbb{Z})$, we know that there is a unique choice of basis for $\Omega^1(M)$ making the period matrix Π of the form $\begin{pmatrix} \mathbb{I}_g & Z \end{pmatrix}$. Moreover, we know by (29.1.6) that Z is symmetric with positive definite imaginary part, i.e. belongs to the g^{th} Siegel upper half space

$$\mathfrak{H}^g := \{Z \in M_g(\mathbb{C}) \mid Z = {}^tZ,\ Im(Z) > 0\}.$$

Note that $\mathfrak{H}^1$ is just $\mathfrak{H}$, the familiar upper half plane.

The Jacobian $J(M)$ is, of course, the quotient of $\mathbb{C}^g$ by the lattice Λ_M given by integral linear combinations of the columns of Π. More generally, let Z be any $g \times g$ complex matrix such that $\begin{pmatrix} \mathbb{I}_g & Z \end{pmatrix}$ has $\mathbb{R}$-linearly independent column vectors. Writing Λ_Z for their $\mathbb{Z}$-span, we define a complex torus by $A_Z := \mathbb{C}^g/\Lambda_Z$; any complex g-torus is isomorphic to one of this form. A major result is the

THEOREM 29.4.1. [RIEMANN EMBEDDING THEOREM] *A_Z is an abelian variety (i.e., has a holomorphic embedding in projective space) if and only if $\pm Z$ belongs to $\mathfrak{H}^g$.*

(Of course, any τ $\mathbb{R}$-linearly independent from 1 is in the upper or lower half plane, so every complex 1-torus is algebraic; already for $g = 2$ this is false.) You can find (effectively) two proofs in Griffiths and Harris, one using generalized theta functions and the other using Kodaira's embedding theorem.

For us, the implications of this theorem are:

(a) Jacobians of Riemann surfaces of genus g are abelian varieties of dimension g; and

(b) abelian varieties of dimension g have $\frac{g(g+1)}{2}$ moduli.

Since Riemann surfaces of genus $g \geq 2$ have $3g-3$ moduli, we conclude:

COROLLARY 29.4.2. *For $g < 4$, all abelian g-folds are Jacobians; for $g \geq 4$, "most" of them are not.*

For $g \geq 4$, then, we have the problem of characterizing the "Jacobian locus" in the moduli space $\mathfrak{H}^g/Sp_{2g}(\mathbb{Z})$, which is the (very difficult) *Schottky problem.* There are recent results describing this locus in terms of the vanishing of theta functions.[12]

[12]for example, see the seminar talk http://www.math.princeton.edu/~sam/papers/talk3s.pdf

Appendix: genera of singular curves

Start with an irreducible projective algebraic curve

$$C \subset \mathbb{P}^2$$

of degree d, defined over $\mathbb{C}$. We know how to piece the local normalizations about singular points together with the smooth part of C to construct a Riemann surface $\tilde{C}$, together with a map

$$\sigma : \tilde{C} \to \mathbb{P}^2$$

with image$(\sigma) = C$. The genus formula of Chapter 14, derived from a generic stereographic projection and the Riemann-Hurwitz formula, said that

$$g(\tilde{C}) = \frac{(d-1)(d-2)}{2} - \delta$$

if all singularities of C (if any) are ordinary double points and there are δ such points.

Here we would like to be able to compute the genus of the normalization of an *arbitrary* irreducible curve, with singularities of any order and type. There exist formulas when the singularities are ordinary,[13] but my preference is for methods over formulas, particularly when the methods allow you to treat more general cases.[14] There are two methods: the first one computes the divisor of the pullback of a meromorphic differential 1-form on $\mathbb{P}^2$ to $\tilde{C}$ and applies Poincaré-Hopf; the second is based on projecting C to a line and applying Riemann-Hurwitz (like in the proof of the genus formula). Rather than trying to state them formally, I'll use both methods to treat an example which is "sufficiently general" that you'll be able to adapt the approaches to any other curve.

[13] cf. §6.4 for the definition of an ordinary k-tuple point. One book which proves the formula is Fulton's book on algebraic curves.
[14] and particularly when you can't remember formulas (welcome to my world)

So here is the ugly curve we will study: put

$$F(Z,X,Y) := X^3Z^3 + X^6 + Y^5Z,$$

and

$$C := \{F(Z,X,Y) = 0\} \subset \mathbb{P}^2,$$

with affine form $x^3 + x^6 + y^5 = 0$. This is a degree 6 (i.e. sextic) curve; a *smooth* curve of this degree has genus 10. That will not be the answer here.

One immediately obvious singularity is at $(0,0)$ (i.e. $[1:0:0]$ in projective coordinates $[Z:X:Y]$). The lowest-order homogeneous term (of the affine equation, in coordinates vanishing at this point) is x^3. So $[1:0:0]$ is a triple point, but very definitely *not* an *ordinary* triple point of C. Ugly enough? Well, this turns out to be the only singularity.

Method I: Poincaré-Hopf

Set $\omega = \sigma^*\left(\frac{dx}{y}\right) \in \mathcal{K}^1(\tilde{C})^*$ (it will actually turn out to be in $\Omega^1(\tilde{C})$, although this is inessential for the method). Poincaré-Hopf tells us that $\deg((\omega)) = 2g - 2$, where $g = g(\tilde{C})$. So we have to compute $(\omega) = \sum m_i[p_i] \in \mathrm{Div}(\tilde{C})$. Where might these $\{p_i\}$ lie in $\tilde{C}$? Or rather, where might the $\{\sigma(p_i)\}$ lie on C? There are four (not necessarily disjoint) possibilities:

(1) on the intersections of C with the x-axis, i.e. in $C \cap \{Y = 0\}$;
(2) at points where C has a vertical tangent, hence in $C \cap \{F_Y = 0\}$;
(3) at singularities of C, i.e. in $\mathrm{Sing}(C)$; and
(4) on the line at infinity, i.e. in $C \cap \{Z = 0\}$.

Why might one expect nontrivial $\nu_p(\omega)$ at p in one of these sets? For (1), the denominator of $\frac{dx}{y}$ is zero on the line $Y = 0$; while for (2) the pullback of dx will be zero, since at such a point the curve has no "horizontal variation" to first order. You should always be suspicious of (3) and (4). Conversely: on the smooth affine part of C, dx and y never blow up, and (1) and (2) are the only ways they can develop a zero. So (1)-(4) are actually the *only* places where ω can have a zero or pole.

Now we go through these 4 sets of points for the particular curve under consideration.

(1): we look at the affine equaton and set $y = 0$, which yields $x^3 + x^6 = 0$, hence $x = 0$, ζ_6, $\overline{\zeta_6}$, or -1. (Here, $\zeta_6 = \exp(\frac{\pi\sqrt{-1}}{3})$.) While $(0,0)$ is a singular point and will be dealt with below, it is clear that $\frac{dx}{y}|_C$ will behave in the same way near the remaining three points: $(-1,0)$, $(\zeta_6,0)$, and $(\overline{\zeta_6},0)$. We look in a neighborhood of $(-1,0)$ on C. Setting $\chi = x+1$ (or $x = \chi - 1$), the equation becomes in (χ, y):

$$0 = y^5 + (\chi-1)^3 + (\chi-1)^6 = y^5 - 3\chi + \{\text{higher-order terms in } \chi\}$$
$$= y^5 - 3\chi h(\chi),$$

where $h(0) \neq 0$.[15] The local normalization of C at $(\chi, y) = (0,0)$ is therefore $t \mapsto \left(t^5,\, t \cdot \sqrt[5]{3h(t^5)}\right)$, under which $\frac{dx}{y} = \frac{d\chi}{y}$ pulls back to $\frac{d(t^5)}{t \cdot \sqrt[5]{3h(t^5)}} = t^3 \cdot \frac{5dt}{\sqrt[5]{3h(t^5)}}$ which has a zero of order 3 at $t = 0$. So we conclude that

$$\nu_{\sigma^{-1}[1:-1:0]}(\omega) = 3,$$

and similarly that

$$\nu_{\sigma^{-1}[1:\zeta_6:0]}(\omega) = \nu_{\sigma^{-1}[1:\overline{\zeta_6}:0]}(\omega) = 3.$$

(2): for vertical tangents or singularities we will have

$$0 = F_Y = 5Y^4 Z,$$

so that these must occur along the x-axis or along the line at ∞. The intersections with the x-axis other than $[1:0:0]$ were just dealt with. Any nonsingular intersections with $\{Z = 0\}$ will be dealt with in step (4). So vertical tangents are subsumed under the other three categories.

(3): at a singular point we must have $0 = F_Y$,

$$0 = F_X = 3X^2Z^3 + 6X^5 = 3X^2(Z^3 + 2X^3),$$

and

$$0 = F_Z = 3X^3Z^2 + Y^5.$$

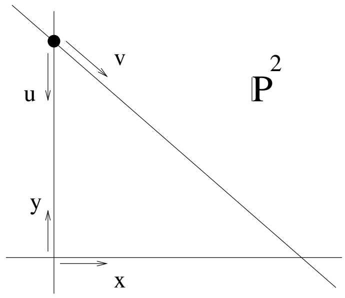
We must have $Z = 0$ or $Y = 0$. If $Z = 0$ then the last two equations imply $X = Y = 0$, a contradiction. If $Y = 0$ then the last equation gives $Z = 0$ (no!) or $X = 0$; the latter works, and so $[1:0:0]$ is the only singular point. In local coordinates about $(x,y) = (0,0)$, our curve is $0 = y^5 + x^3 + x^6 = y^5 + x^3h(x)$ (different $h(x)$ from above,

[15]Sometimes (though rarely) one may have to "remember" more about h in these types of problems, but not in this example.

again $h(0) \neq 0$), which is *locally* irreducible and has a singularity of order 3. Undeer the local normalization $t \mapsto (t^5,\, t^3 \cdot \sqrt[5]{h(t^5)})$, $\frac{dx}{y}$ pulls back to $\frac{d(t^5)}{t^3 \cdot \sqrt[5]{h(t^5)}} = t \cdot \frac{5dt}{\sqrt[5]{h(t^5)}}$, and we conclude that

$$\nu_{\sigma^{-1}[1:0:0]}(\omega) = 1.$$

(4): $C \cap \{Z = 0\}$ is the single point $[0 : 0 : 1]$. We will need to switch to affine coordinates vanishing at this point, namely $u = \frac{1}{y} = \frac{Z}{Y}$, $v = \frac{x}{y} = \frac{X}{Y}$ (or conversely $y = \frac{1}{u}$, $x = \frac{v}{u}$):



We divide $Z^3X^3 + X^6 + Y^5Z = 0$ by Y^6, obtaining

$$\left(\frac{Z}{Y}\right)^3 \left(\frac{X}{Y}\right)^3 + \left(\frac{X}{Y}\right)^6 + \frac{Z}{Y} = 0$$

$$v^6 + u^3v^3 + u = 0$$

which is a locally irreducible Weierstrass polynomial in v with (multi-valued) roots of the form

$$v_*(u) := \sqrt[3]{\frac{-u^3 \pm \sqrt{u^6 - 4u}}{2}}.$$

(Use the quadratic formula to solve for v^3, then take cube root.) Substituting in t^6 gives

$$\tilde{v}(t^6) = \sqrt[3]{\frac{-t^{18} + \sqrt{t^{36} - 4t^6}}{2}} = \sqrt[3]{\frac{-t^{18} + t^3 \cdot \sqrt{t^{30} - 4}}{2}}$$

$$= t \cdot \sqrt[3]{\frac{-t^{15} + \sqrt{t^{30} - 4}}{2}}$$

which is just t times some local analytic $H(t)$ with $H(0) \neq 0$. So the normalization is $t \mapsto (t^6,\, t \cdot H(t))$ and $\frac{dx}{y} = \frac{d\left(\frac{v}{u}\right)}{\frac{1}{u}}$ pulls back to $\frac{d\left(\frac{tH(t)}{t^6}\right)}{\frac{1}{t^6}} = t^6 d\left(\frac{H(t)}{t^5}\right) = (tH'(t) - 5H(t))\, dt$, which has neither zero nor pole at $t = 0$. Hence,

$$\nu_{\sigma^{-1}[0:0:1]}(\omega) = 0.$$

Upshot: putting everything together,

$$(\omega) = [\sigma^{-1}[1:0:0]] + 3[\sigma^{-1}[1:\zeta_6:0]] + 3[\sigma^{-1}[1:-1:0]] + 3[\sigma^{-1}[1:\overline{\zeta_6}:0]]$$

$$\implies \quad 2g-2 \;=\; \deg((\omega)) \;=\; 1+3+3+3 \;=\; 10$$

$$\implies \quad g = 6.$$

Method II: Riemann-Hurwitz

Recall that this dealt with *maps* of Riemann surfaces

$$f : M \to N,$$

and told us that

$$\chi_M = \deg(f) \cdot \chi_N - r_f.$$

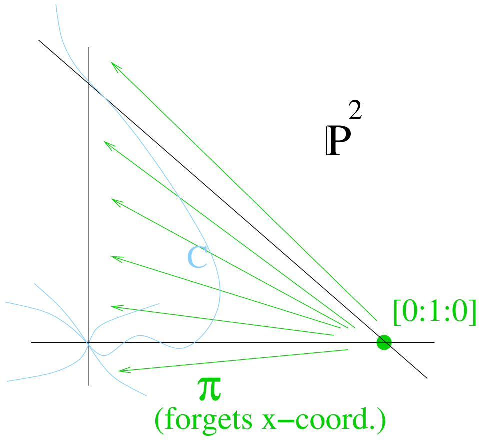
Here $\deg(f)$ is the mapping degree of f (the number of points in the preimage of a general point on N) and r_f is the degree of the ramification divisor[16] $R_f := \sum_{p \in M} (\nu_p(f) - 1)[p]$.

Now let $q \in \mathbb{P}^2 \backslash C$, $M = \tilde{C}$, $N = \mathbb{P}^1$, π =stereographic projection $(\mathbb{P}^2 \backslash \{q\}) \to \mathbb{P}^1$ through q; and take

$$f : \tilde{C} \to \mathbb{P}^1$$

to be given by $f := \pi \circ \sigma$. Usually it is easiest to take $[1:0:0]$, $[0:1:0]$, or $[0:0:1]$ as q. In our case the only one of these *not* on C is $[0:1:0]$. So our projection looks like

[16]locally about any $p \in M$ and its image $f(p) \in N$, one has local holomorphic coodinates z resp. w (with $z(p) = 0$ resp. $w(f(p)) = 0$), in which f takes the form $z \mapsto z^{\nu_p(f)} = w$.



and the mapping degree is the number of intersection points of $\{y = y_0\}$ and $\{x^3 + x^6 + y^5 = 0\}$ for general y_0 — i.e. $\deg(f) = 6$. Obviously $\chi_{\mathbb{P}^1} = 2 - 2 \cdot 0 = 2$, so we have

$$2 - 2g = \chi_{\tilde{C}} = 6 \cdot 2 - r_f = 12 - r_f$$

$$\implies \quad g = \frac{1}{2} r_f - 5.$$

So we will have to compute $R_f = \sum m_i [p_i]$ (or at least r_f) and the first issue to resolve is where the $\sigma(p_i)$ can lie on C:

(1) points having horizontal tangents (subset of $\{F = 0\} \cap \{F_X = 0\}$);
(2) singular points ($F_X = F_Y = F_Z = 0$) — i.e. $[1 : 0 : 0]$ for our example; and
(3) $L_\infty \cap C$ — i.e. $[0 : 0 : 1]$ in our example.

(1): $0 = F_X = 3X^2(Z^3 + 2X^3)$ has solutions other than $X = 0$, which corresponds to the singular point. Namely, writing $x = \frac{X}{Z}$ we get $x^3 + \frac{1}{2} = 0$ hence $x = \frac{-1}{\sqrt[3]{2}}, \frac{\zeta_6}{\sqrt[3]{2}}, \frac{\bar{\zeta}_6}{\sqrt[3]{2}}$. Plugging this into the affine equation of C yields $y^5 = \frac{1}{4}$ hence $y = \frac{1}{\sqrt[5]{4}}, \frac{\zeta_5}{\sqrt[5]{4}}, \frac{\zeta_5^2}{\sqrt[5]{4}}, \frac{\zeta_5^3}{\sqrt[5]{4}}, \frac{\zeta_5^4}{\sqrt[5]{4}}$. These are independent of the choice amongst the 3 values for x, and so we get $5 \cdot 3 = 15$ ramification points. As you may check, the intersections between $F = 0$ and $F_X = 0$ at these points are all of first order, hence correspond to ramifications of order 2 and so make a contribution of $\nu_p(f) - 1 = 2 - 1 = 1$ each to r_f.

(2): Near $(x, y) = (0, 0)$, the composition

$$t \overset{\sigma}{\longmapsto} (t^5, t^3 H(t)) \overset{\pi}{\longmapsto} t^3 H(t)$$

has $\nu_p(f) = 3$ hence contributes 2 to r_f.

(3): Near $(u, v) = (0, 0)$,

$$t \overset{\sigma}{\longmapsto} \underbrace{(t^6, tH(t))}_{u,v} \overset{\pi}{\longmapsto} t^6.$$

This is because π is supposed to take the y-coordinate, which is $\frac{1}{u}$ here; but we have to compute the image in a holomorphic coordinate vanishing at the image of $p = [0 : 0 : 1]$. So in fact u is the correct variable, and the map indeed has $\nu_p(f) = 6$ and contributes 5 to r_f.

Conclusion: $r_f = 15 \cdot 1 + 2 + 5 = 22$

$$\implies \quad g \;=\; \frac{22}{2} - 5 \;=\; 6,$$

confirming the previous computation.

Exercises

(1) Find the genus of the normalization of the curve $C \subset \mathbb{P}^2$ with affine equation $x^3 + y^2 + y^2x^3 - \frac{2}{5}y^5 = 0$. Do this in 2 different ways: (a) by using an appropriate projection, computing the degree of the ramification divisor, and applying Riemann-Hurwitz; (b) by computing the divisor of the pull-back of a meromorphic 1-form on the normalization and applying Poincare-Hopf (try it with $\frac{dy}{x}$). [*Hint: to do the local normalizations, first make sure you are dealing with an irreducible Weierstass polynomial* $f(x, y) = 0$ *– you may have to change variable, swap coordinates, divide out by a unit (which can reduce the degree of the equation!), factor into irreducibles, whatever. If you can't find a multivalued solution* $y(x)$ *by taking roots, using quadratic equation, and so on, you can always use power series. If* $f(x, y)$ *is an irreducible Weierstrass polynomial of degree* k *in* y*, then try substituting in* t^k *for* x*: write* $0 = f(t^k, y)$ *and solve for* y *as a power-series in* t*, call this* $G(t)$*. Then the local normalization is* $t \mapsto (t^k, G(t))$.]