18.727, Topics in Algebraic Geometry (rigid analytic geometry)

Kiran S. Kedlaya

Fall 2004

Contents

Chapter 1

Introduction: the Tate curve

We are going to start the course precisely where the subject of rigid analytic geometry itself has its origin, with Tate's groundbreaking work on parametrization of elliptic curves. Of course this will be somewhat informal since we haven't done any theory yet; the idea is to give some sense both of the "how" (of doing analytic geometry over nonarchimedean fields) and the "why" (with an application due to Serre).

This is only one of the applications I hope to discuss during the semester; this one is pretty much number theory, as are many of the others I'm fond of, but there are lots of other applications that get into other areas. I'll provide a long (if not exhaustive) list sometime soon.

Note that we are starting with the very *last* section of [BGR]! That should give you some sense of the difference between their attitudes and mine.

References: Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves* (cited as [Sil]), whose development of the Tate curve in Section V.3 is liberally plagiarized here. References in the language of rigid geometry won't do you any good yet, but here they are for future use: [BGR, Section 9.7], [FvdP, 5.1]. Note that here and throughout the course, [BGR] means *Non-Archimedean Analysis*, by Bosch, Güntzer, and Remmert, while [FvdP] means *Rigid Analytic Geometry and its Applications*, by Fresnel and van der Put. (The former is unfortunately out of print. Then again, maybe that's just as well; it's a useful reference book but of no pedagogical value, as it's pretty dry, devoid of examples, devoid of exercises, and aside from the Tate curve, devoid of geometry and of applications!)

Warmup: Weierstrass parametrizations

An *elliptic curve* over a field is a genus 1 algebraic curve over that field equipped with the choice of a distinguished point on the curve (the *origin*). Over $\mathbb{C}$, this is the same as a Riemann surface of genus 1 equipped with a choice of a distinguished point.

Once upon a time, Weierstrass discovered that every elliptic curve E over $\mathbb{C}$ can be viewed in a canonical fashion as a complex torus, i.e., as a quotient of $\mathbb{C}$ by a lattice of the form $\mathbb{Z}\alpha + \mathbb{Z}\beta$. This parametrization makes a lot of facts about genus 1 curves over $\mathbb{C}$ much more transparent, e.g., the group law, the shape of the torsion subgroups, and the possibilities for the ring of endomorphisms.

Once upon a somewhat more recent time, Tate realized that one could do something similar over the field $\mathbb{Q}_p$ of p-adic numbers. (I'm not going to define $\mathbb{Q}_p$ here; if you don't know what it is, probably this is not the course for you!) This is not at all clear from the "$\mathbb{C}$ mod a lattice" description above; we have to break a little symmetry first. Namely, by rescaling the lattice, we may as well assume that it is generated by 1 and τ, where $\tau \in \mathcal{H}$ and $\mathcal{H}$ is the upper half-plane

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

The Weierstrass functions on $\mathbb{C}$, which generate the field of meromorphic functions on E, are now doubly periodic, and in particular they are invariant under translation by 1. So why not write them as Fourier series? In other words, apply the exponential map:

$$\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z}) \overset{\exp}{\to} \mathbb{C}^*/q^{\mathbb{Z}},$$

where $q = e^{2\pi i\tau}$.

In the Weierstass setting, there are special functions of the parameter τ that describe the elliptic curve, but of course we can write them in terms of q also. Namely, set

$$s_k(q) = \sum_{n=1}^{\infty} \frac{n^k q^n}{1-q^n} \qquad (k \in \mathbb{N})$$

and put

$$a_4(q) = -s_3(q), \qquad a_6(q) = -\frac{5s_3(q)+7s_5(q)}{12}.$$

Then [Sil, V.1.1] the elliptic curve $E_q \cong \mathbb{C}^*/q^{\mathbb{Z}}$ is isomorphic to

$$E_q : y^2 + xy = x^3 + a_4(q)x + a_6(q), \tag{1.1}$$

its discriminant is

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^{24}, \tag{1.2}$$

and its j-invariant is

$$j(E_q) = q^{-1} + 744 + 196884q + \cdots . \tag{1.3}$$

The parametrization $\mathbb{C}^*/q^{\mathbb{Z}} \to E_q$ is given by the functions (rewrites of the Weierstrass $\wp$ function and its derivative)

$$X(u,q) = \sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^2} - 2s_1(q) \tag{1.4}$$

$$Y(u,q) = \sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^3} + s_1(q); \tag{1.5}$$

that is, if you fix q, for each u the point $(X(u,q), Y(u,q))$ lies on E_q, and these form an analytic isomorphism between the complex torus and the elliptic curve. (The point $1 \in \mathbb{C}^*/q^{\mathbb{Z}}$ maps to the point at infinity on the Weierstrass equation (1.1).)

Tate curves

Now suppose K is a finite extension of the field $\mathbb{Q}_p$, and let $|\cdot|$ denote the absolute value on K normalized so that $|p| = p^{-1}$. Then the series $s_k(q)$ converges whenever $|q| < 1$, so one might wonder whether in that case we get an "isomorphism" between on one hand the quotient of the multiplicative group by the subgroup generated by $q^{\mathbb{Z}}$, and on the other hand the elliptic curve (1.1). To make this really make sense, we need to have some category of geometric objects in which we can take the quotient $\mathbb{G}_m/q^{\mathbb{Z}}$; certainly we can't do it using schemes. One of the purposes of this course is to construct precisely the sorts of geometric objects that one needs to form the quotient and say that it's isomorphic to the elliptic curve.

But in the meantime, one can prove that everything works correctly at the level of points.

Theorem 1.1. *For $q \in K$ with $|q| < 1$, the equation* (1.1) *defines an elliptic curve over K with discriminant* (1.2) *and j-invariant* (1.3)*; the series $X(u,q)$ and $Y(u,q)$ defined in* (1.4) *and* (1.5) *converge for all $u \in (K^{\mathrm{alg}})^* \setminus q^{\mathbb{Z}}$; and the map $\phi : (K^{\mathrm{alg}})^* \to E_q(K^{\mathrm{alg}})$ given by*

$$u \mapsto \begin{cases} (X(u,q), Y(u,q)) & u \notin q^{\mathbb{Z}} \\ O & u \in q^{\mathbb{Z}} \end{cases}$$

(where O is the point at infinity on E_q) is a Galois-equivariant surjection with kernel $q^{\mathbb{Z}}$. (In fact, the same is true with K^{alg} replaced by any finite extension of K, which says a bit more.)

Proof. This is verified in detail in [Sil, Theorem V.3.1], but it's a bit cumbersome (precisely because we don't have any of the rigid geometric machinery available yet!). So I'll only summarize for now.

I already noted that s_k converges for $|q| < 1$, so a_4 and a_6 are defined. The computation of Δ and j from a_4 and a_6 is formal, so the series expressions don't change between the complex case and this case. Oh, and

$$|\Delta(q)| = |q - 24q^2 + 252q^3 + \cdots| = |q| \neq 0$$

since all the other terms are of strictly smaller size, so $\Delta(q) \neq 0$ and E_q really is an elliptic curve.

To see that $X(u,q)$ and $Y(u,q)$ converge, it is easiest to rewrite them as in [Sil, (V.1.2)]:

$$\begin{aligned} X(u,q) &= \frac{u}{(1-u)^2} + \sum_{n=1}^{\infty} \left(\frac{q^n u}{(1-q^n u)^2} + \frac{q^{-n} u}{(1-q^{-n} u)^2} - 2\frac{q^n}{(1-q^n)^2} \right) \\ &= \frac{1}{u + u^{-1} - 2} + \sum_{n=1}^{\infty} \left(\frac{q^n u}{(1-q^n u)^2} + \frac{q^n u^{-1}}{(1-q^n u^{-1})^2} - 2\frac{q^n}{(1-q^n)^2} \right) \\ Y(u,q) &= \frac{u^2}{(1-u)^3} + \sum_{n=1}^{\infty} \left(\frac{(q^n u)^2}{(1-q^n u)^3} + \frac{(q^{-n} u)^2}{(1-q^{-n} u)^3} + \frac{q^n}{(1-q^n)^2} \right) \\ &= \frac{u^2}{(1-u)^3} + \sum_{n=1}^{\infty} \left(\frac{(q^n u)^2}{(1-q^n u)^3} + \frac{(q^n u^{-1})^2}{(1-q^n u^{-1})^3} + \frac{q^n}{(1-q^n)^2} \right). \end{aligned}$$

From this description the convergence for $u \in (K^{\mathrm{alg}})^* \setminus q^{\mathbb{Z}}$ is clear: everything is converging like a geometric series as long as none of the denominators vanish.

Aside, but an important one: the fact that I'm plugging in values from K^{alg} makes the previous conclusion much stronger than if I only took values from K, because there's a hard limit on how close together two elements of K can be (because K is discretely valued). So in general, you can't really assess convergence properties (like the radius of convergence of a power series) by testing values in a finite extension of $\mathbb{Q}_p$.

Anyway, one can check from the series above that $X(qu,q) = X(u,q) = X(u^{-1},q)$, that $Y(qu,q) = Y(u,q)$, and that $Y(u^{-1},q) = -Y(u,q) - X(u,q)$, as you would expect of points on E_q.

To check that the image of ϕ lands on E_q amounts to checking that plugging $X(u,q)$ and $Y(u,q)$ into E_q yields an identity of formal power series; this can be deduced from the complex side. Ditto for the fact that multiplication in $(K^{\mathrm{alg}})^*$ converts into the addition law on E_q.

The only bit that remains is to check surjectivity of ϕ. Unfortunately, this is the hardest part, and the proof in [Sil, V.4] is too ugly and not insightful enough to be worth reproducing here. I would recommend looking at it just to gain an appreciation for why one needs a real theory of analytic spaces over nonarchimedean fields. □

In case you know what this means: one interesting feature of the ugly calculation in [Sil, V.4] is the fact that he sorts the points on E_q into classes as follows. (Here π is a uniformizer of K.)

$$\begin{aligned} E_{q,0}(K) &= \{(x,y) \in E_q(K) : \max\{|x|,|y|\} \geq 1\} \\ U_n &= \{(x,y) \in E_q(K) : |\pi|^n = |y| > |x+y|\} \\ V_n &= \{(x,y) \in E_q(K) : |\pi|^n = |x+y| > |y|\} \\ W &= \{(x,y) \in E_q(K) : |y| = |x+y| = |q|^{1/2}\} \end{aligned}$$

The point is that these classes are "neighborhoods" of the various components of the special fibre of the Néron model of E_q. Note that $|q| < 1$ means that $|j| > 1$, so E_q necessarily has bad reduction; in fact, its reduction is multiplicative of type I_n, where n is the valuation of q in K.

The parametrization theorem

Here's the point of the construction we just made.

Theorem 1.2 (Tate)**.** *Let K be a finite extension of $\mathbb{Q}_p$, and let E be an elliptic curve with $|j(E)| > 1$.*

(a) There is a unique $q \in K^*$ with $|q| < 1$ such that $E \cong E_q$ over K^{alg}.

(b) For q as in (a), $E \cong E_q$ over K if and only if E has split multiplicative reduction.

Proof. For (a), note that one can invert the series for j given in (1.3) to solve for j in terms of q. For (b), see [Sil, Theorem V.5.3]. $\square$

In other words, Tate's construction gives a universal description of elliptic curves with bad reduction at a prime! Furthermore, it's clear from the description that the parametrization commutes with the action of $\mathrm{Gal}(K^{\mathrm{alg}}/K)$; that's one reason why number theorists like this parametrization more than the Weierstrass parametrization, which has no such property.

Application: j-invariants and complex multiplication

As an application of Tate's construction, we mention the following argument due to Serre.

Proposition 1.3. *Let K be a finite extension of $\mathbb{Q}_p$ with normalized valuation v, let E be an elliptic curve over K with $|j(E)| > 1$, and let $\ell \geq 3$ be a prime not dividing $v(j(E))$ (but $\ell = p$ is allowed). Then there exists σ in the inertia subgroup of $\mathrm{Gal}(K^{\mathrm{alg}}/K)$ and a basis P_1, P_2 of the ℓ-torsion $E[\ell]$ of E over K^{alg} such that*

$$P_1^\sigma = P_1, \qquad P_2^\sigma = P_1 + P_2.$$

Proof. There is no harm in replacing K by a finite extension of degree prime to ℓ (doing so multiplies $v(j(E))$ by some divisor of that degree, namely the degree of the residual extension). In particular, by going up a quadratic extension, we can ensure that E is congruent to its corresponding Tate curve E_q over K; and we may assume that K contains a primitive ℓ-th root of unity ζ_ℓ.

Let $Q = q^{1/\ell} \in K^{\mathrm{alg}}$ be a fixed ℓ-th root of q. Then the ℓ-torsion in $(K^{\mathrm{alg}})^*/q^{\mathbb{Z}}$ is generated by ζ_ℓ and Q. So all we have to do is notice that the Kummer extension $K(Q)/K$ is totally ramified of degree ℓ, so there exists σ in the inertia subgroup of $\mathrm{Gal}(K^{\mathrm{alg}}/K)$ such that $Q^\sigma = Q\zeta_\ell$ (and $\zeta_\ell^\sigma = \zeta_\ell$ since we put ζ_ℓ into K); the images P_1 and P_2 of ζ_ℓ and Q, respectively, do what we wanted. $\square$

Aside: strictly speaking, we didn't need the surjectivity of the Tate parametrization, since we were able to produce enough ℓ-torsion within $(K^{\mathrm{alg}})^*/q^{\mathbb{Z}}$. (That's the part where I was whining earlier because we haven't done any rigid geometry yet.) But in other applications we may not be so lucky!

Anyway, here's what Serre deduced from this. (There are other proofs too; see [Sil, V.6] for more discussion.)

Theorem 1.4. *Let $K/\mathbb{Q}$ be a number field, and let E/K be an elliptic curve whose j-invariant is not an integer in K. Then* $\mathrm{End}(E) = \mathbb{Z}$.

By contrast, an elliptic curve over a field of characteristic zero can also have endomorphism ring equal to an order in an imaginary quadratic field (the "complex multiplication" case). This theorem proves that the j-invariants of CM-curves, which one can show are algebraic numbers, are actually algebraic integers!

Proof. Suppose on the contrary that $\mathrm{End}(E)$ includes an endomorphism $\psi \notin \mathbb{Z}$. Then (see exercises) ψ must satisfy some polynomial of the form

$$a\psi^2 + b\psi + c = 0 \qquad (a, b, c \in \mathbb{Z}).$$

Moreover, the polynomial $ax^2 + bx + c$ must have distinct nonreal roots, which generate some quadratic imaginary field L.

In particular, there are lots of primes ℓ which split (and are not ramified) in ℓ, and which don't divide a. Pick one: then ψ acts on the ℓ-torsion $E[\ell]$ (which is a two-dimensional vector space over $\mathbb{F}_\ell$) via a matrix which is diagonalizable with distinct roots.

On the other hand, if we pass from K to its completion $K_{\mathfrak{p}}$ at some prime ideal $\mathfrak{p}$ (of the ring of integers of K) at which $j(E)$ is nonintegral, then by our proposition, there is an element of $\mathrm{Gal}(K_{\mathfrak{p}}^{\mathrm{alg}}/K_{\mathfrak{p}})$ which acts via a nontrivial unipotent matrix. But this element of Galois has to commute with ψ, which is clearly impossible! Contradiction. □

Exercises

1. Verify some of the formal stuff that I didn't check, like the fact that multiplication on $(K^{\mathrm{alg}})^*/q^{\mathbb{Z}}$ translates into addition on the Tate curve. Or if you prefer, just look it up in [Sil, Chapter V].

2. Check all the statements I made about the endomorphism ψ of an elliptic curve, e.g., that it satisfies a quadratic polynomial whose roots are nonreal. (Hint: use the Weierstrass parametrization.)

3. (Sil, Exercise 5.10) Prove that if the Tate curves E_q and $E_{q'}$ are isogenous, then some power of q equals some power of q'. The converse is also true, but is a bit complicated to show without doing the construction more honestly (i.e., actually using rigid geometry).

Chapter 2

A little p-adic functional analysis (part 1 of 2)

I'm going to start with a little bit of terminology and notation about nonarchimedean Banach spaces (trusting that you can fill in a few details that are similar to the real/complex case). There's a lot more where this came from, but we won't need the rest of it just yet.

Thanks to Abhinav Kumar for providing the corrections incorporated into this version (of 16 Sep 04).

References: [FvdP], Chapter 1; the stuff is also in [BGR], but you'll have to tease it out of Chapter 2 with some effort, as it's scattered among many sections. In case you come down with a craving for more p-adic functional analysis, I recommend *Nonarchimedean functional analysis*, by Schneider. This book is available online; see the notes page for a link. At worst, you can always pick up a standard functional analysis book (e.g., *Espaces vectoriels topologiques* by Bourbaki) and redo all the constructive proofs (i.e., skip anything involving Hahn-Banach) yourself in the nonarchimedean context!

Ultrametric spaces

An *ultrametric* (or *nonarchimedean metric*) on a set X is function $d : X \times X \to \mathbb{R}_{\geq 0}$ with the following properties.

(a) For $x_1, x_2 \in X$, $x_1 = x_2$ if and only if $|x| = 0$.

(b) For $x_1, x_2 \in X$, $d(x_1, x_2) = d(x_2, x_1)$.

(c) For $x_1, x_2, x_3 \in X$, $d(x_1, x_3) \leq \max\{d(x_1, x_2), d(x_2, x_3)\}$ (strong triangle inequality)

Note that if $d(x_1, x_2) \neq d(x_2, x_3)$, then in fact $d(x_1, x_3) = \max\{d(x_1, x_2), d(x_2, x_3)\}$, otherwise you get a contradiction by applying (c) to x_1, x_2, x_3 in another order.

A *Cauchy sequence* in X is a sequence $\{x_n\}_{n=1}^{\infty}$ such that for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, $d(x_m, x_n) < \epsilon$. Note that by the strong triangle inequality, this is equivalent to $d(x_n, x_{n+1}) < \epsilon$ for $n \geq N$; this is the first of many instances in which nonarchimedean analysis turns out to be easier than traditional analysis!

We say X is *complete* if every Cauchy sequence converges to a limit (necessarily unique because of (a)). We say X is *spherically complete* if every decreasing sequence of closed balls has nonempty intersection: that is, given $x_1, x_2, \dots \in X$ and $r_1, r_2, \dots \in \mathbb{R}_{\geq 0}$ such that the sets

$$D_n = \{x \in X : d(x, x_n) \leq r_n\}$$

satisfy $D_1 \supseteq D_2 \supseteq \cdots$, then $\cap_{n=1}^{\infty} D_n \neq \emptyset$. Note that if X is spherically complete, then X is also complete: given a Cauchy sequence $\{x_1, x_2, \dots\}$, we can pass to a subsequence if needed to ensure that $d(x_n, x_{n+1}) \geq d(x_{n+1}, x_{n+2})$ for all n. Then the balls

$$D_n = \{x \in X : d(x, x_n) \leq d(x_{n+1}, x_n)\}$$

have nonempty intersection, which must be a limit of the sequence. (In other words, complete means spherically complete when the balls have radii going to 0.)

Ultrametric fields

An *ultrametric field*, or *nonarchimedean valued field* (in the terminology of [FvdP]), is a field K equipped with a function $|\cdot| : K \to \mathbb{R}_{\geq 0}$ with the following properties.

(a) For $x \in K$, $x = 0$ if and only if $|x| = 0$.

(b) For $x_1, x_2 \in K$, $|x_1 x_2| = |x_1||x_2|$.

(c) For $x_1, x_2 \in K$, $|x_1 + x_2| \leq \max\{|x_1|, |x_2|\}$.

The function $d(x_1, x_2) = |x_1 - x_2|$ is then an ultrametric on K, so we know what it means for K to be complete. In this course, we will usually be working over a complete ultrametric field. Oh, and there is always a trivial absolute value given by

$$|x| = \begin{cases} 0 & x = 0 \\ 1 & x \neq 0; \end{cases}$$

I'm always going to assume (unless otherwise specified) that my absolute value function is not the trivial one.

If K is an ultrametric field, then the set of $x \in K$ with $|x| \leq 1$ is a subring of K. I'll denote it by $\mathfrak{o}_K$, or sometimes by $\mathfrak{o}$ in case K is to be understood. I might also call it the "valuation subring". The set of $x \in K$ with $|x| < 1$ is a maximal ideal of $\mathfrak{o}_K$, which I'll denote $\mathfrak{m}_K$. The field $\mathfrak{o}_K/\mathfrak{m}_K$ is called the *residue field* of K, and I'll typically call it k. (Note that k is really a field, and not a one-element ring, because the absolute value function is nontrivial!)

We call $|K^*|$ the *value group* of K. We say K is *discretely valued* if its value group is a discrete subgroup of $\mathbb{R}_{>0}$; that means it must be isomorphic to $\mathbb{Z}$. Our favorite examples of complete ultrametric fields are discretely valued, namely:

(a) the field of formal power series (or really Laurent series, but I'll call it the "field of formal power series" from here on) $k((t))$ over a field k;

(b) the field $\mathbb{Q}_p$ of p-adic numbers;

(c) any finite extension of either of these (see exercises).

Another example is the completion of the maximal unramified extension $\mathbb{Q}_p^{\mathrm{unr}}$ of $\mathbb{Q}_p$ (or of a finite extension of $\mathbb{Q}_p$).

For non-discretely valued examples, keep reading. Then again, if you want to pretend for the rest of the course that all ultrametric fields are discretely valued, you will not lose too much of the flavor of the course. (I'll try to make explicit warnings at points where it makes a difference.)

Spherically complete fields

Terminology warning: the term *maximally complete* is used interchangeably with *spherically complete* when talking about ultrametric fields. (For instance, [FvdP] uses "maximally complete" consistently.) The reason: the ultrametric field K is spherically complete if and only if it is maximal among ultrametric fields with the same value group and residue field. (I think this is due to Kaplansky; see his papers "Maximal fields with valuations I, II". See also the exercises.)

Note that any discretely valued ultrametric field is spherically complete, since the radii of balls in a decreasing sequence must stabilize. The canonical example of an ultrametric field which is complete but not spherically complete is $\mathbb{C}_p$, the completion of the algebraic closure of $\mathbb{Q}_p$.

Every ultrametric field can be embedded in a spherically complete ultrametric field, or even an algebraically closed spherically complete ultrametric field. I don't know a reference for this offhand, but see the exercises for some examples.

Nonarchimedean Banach spaces

Assume for the rest of this installment and the next (and pretty much for the rest of the course!) that K is a *complete* ultrametric field, and let $|\cdot|$ denote the norm on K.

Let V be a vector space over K. A *seminorm* on V is a function $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ satisfying the following conditions.

(a) For $a \in K$ and $v \in V$, $\|av\| = |a|\|v\|$.

(b) For $v, w \in V$, $\|v + w\| \leq \max\{\|v\|, \|w\|\}$.

If moreover $\|v\| = 0$ implies $v = 0$, we say $\|\cdot\|$ is a *norm.* If V comes equipped with a norm, we call it a *normed space (over K).*

If V is a normed space which is complete under its norm (or rather, under the induced ultrametric $d(v, w) = \|v - w\|$), we say V is a *Banach space (over K).* For instance, any finite dimensional K-vector space is a Banach space. As in the traditional setting, there are lots of simple examples of Banach spaces, e.g., the set of all null sequences $(a_0, a_1, \dots)$ over K (that is, sequences with $|a_i| \to 0$ as $i \to \infty$) with the supremum norm, or all bounded sequences with the supremum norm, or all convergent sequences with the supremum norm, or...

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same space V are *equivalent* if there exist $\alpha, \beta > 0$ such that for all $v \in V$,

$$\alpha\|v\|_1 \leq \|v\|_2 \leq \beta\|v\|_1;$$

clearly this is actually an equivalence relation on norms. Equivalent norms induce the same topology on V, but *not conversely.*

One makes subspaces and quotient spaces as follows. If $f : V \to W$ is an injective map of vector spaces over K, W is a Banach space, and $\operatorname{im}(f)$ is closed, then the restriction of the norm on W to V gives V the structure of a Banach space. If $f : V \to W$ is a surjective map of vector spaces over K, V is a Banach space, and $\ker(f)$ is closed, then the *quotient norm*

$$\|w\|_W = \inf\{\|v\| : v \in V, f(v) = w\}$$

is obviously a seminorm, but in fact it is also a norm. Namely, if $\|w\|_W = 0$, we can choose $v_1, v_2, \dots \in V$ with $f(v_i) = w$ for all i and $\|v_i\| \to 0$. Then $v_i - v_j \in \ker(f)$ for all i, j; fixing i, letting j tend to ∞ and recalling that $\ker(f)$ is closed, we see that $v_i \in \ker(f)$. Hence $w = 0$; in other words, $\|\cdot\|_W$ is a norm.

By a similar argument, any Cauchy sequence in W converges: if $\{w_i\}$ is a Cauchy sequence in W, we can choose lifts $\{v_i\}$ of the w_i to V so that $\|v_i - v_{i+1}\| \to 0$ as $i \to \infty$. Since V is complete, the v_i converge to a limit v such that $\|f(v) - w_i\|_W \to 0$ as $i \to \infty$, so $\{w_i\}$ has a limit. We conclude that W W inherits from W the structure of a Banach space.

A lot of stuff you know about Banach spaces over $\mathbb{R}$ or $\mathbb{C}$ carries over to this setting (so I'm not going to bother redoing the classical proofs in these cases; see the references). Typical examples:

- Any finite dimensional vector space over K is a Banach space, and any two norms on it are equivalent.
- A map $f : V \to W$ between Banach spaces is continuous (for the norm topologies) if and only if it is bounded (i.e., there exists $c > 0$ such that $\|f(v)\| \leq c\|v\|$ for all $v \in V$).
- Open mapping theorem: if $f : V \to W$ is a bounded surjective linear map between Banach spaces, then f is an open map (the image of an open subset is open), and the norm topology on W coincides with the quotient topology. More precisely, there exists $c > 0$ such that any $w \in W$ is the image of some $v \in V$ with $\|v\| \leq c\|w\|$. Corollary: any bijective bounded linear map between Banach spaces is an isomorphism. (If anyone wants me to go through the proof of this, let me know and I'll prepare it for next time.)
- Closed graph theorem: the linear map $f : V \to W$ between two Banach spaces is bounded if and only if its graph is closed under the product topology on $V \oplus W$. (Apply the open mapping theorem to the map between W and the quotient of $V \oplus W$ by the graph of f.)

However, the Hahn-Banach theorem extends verbatim *only* to spherically complete fields. For more general fields, one needs an extra restriction on V. We say V is *of countable type* if it contains a countable subset whose linear span is dense in V. (I think such a space in the classical setting is said to be "separable", but that word is sufficiently overburdened in algebraic geometry!)

Lemma 2.1. *Suppose V is a Banach space of countable type over K. Then for each $R > 1$, there exists an at most countable set $\{e_i\}$ such that each $v \in V$ can be written as a convergent sum $\sum c_i e_i$ with $|c_i| \cdot \|e_i\| \to 0$, and that any such sum satisfies*

$$R^{-1} \max_i\{|c_i| \cdot \|e_i\|\} \leq \| \sum c_i e^i \| \leq \max |c_i| \cdot \|e_i\|.$$

Proof. Exercise, or see [FvdP, Proposition 1.2.1]. □

Theorem 2.2 (Hahn-Banach). *Let $W \subset V$ be an inclusion of normed spaces, with V complete, and suppose $f : W \to K$ is a bounded K-linear map; that is, there exists $c > 0$ such that $|f(w)| \leq c\|w\|$ for all $w \in W$. Suppose further that* either*:*

(a) V is of countable type, or

(b) K is spherically complete.

Then for any $R > 1$, there exists a K-linear map $g : V \to K$ extending f such that $|g(v)| \leq cR\|v\|$ for all $v \in V$. Moreover, in case (b), we also have this conclusion with $R = 1$.

Proof. For (a), we may assume W is closed because any bounded linear map extends uniquely from W to its closure. In that case, apply Lemma 2.1 to V/W to produce $d_i \in V$ such that each $v \in V$ has a unique presentation as $w + \sum c_i d_i$ with $w \in W$, $c_i \in K$ and $|c_i| \to 0$. Then extend f by setting $g(w + \sum c_i d_i) = f(w)$. For (b), see Schneider's book; we won't use this part very much. (Again, let me know if you want to see this in detail.) □

Note that the Hahn-Banach theorem *always* fails in general if K is not spherically complete: if L is a spherically complete field containing K, and $D_1, D_2, \dots$ is a decreasing sequence of balls in K with empty intersection, then the identity map $K \to K$ cannot extend to a bounded map $L \to K$, because any element of L in the intersection of the D_i has nowhere to go in K! (Thanks to Damiano Testa for noticing this.)

Exercises

1. Let K be an ultrametric field and let L be a finite extension of K. Show that the absolute value of K extends uniquely to an absolute value on L, and that L is complete if K is.

2. Prove that $\mathbb{C}_p$ (the completion of the algebraic closure of $\mathbb{Q}_p$) is algebraically closed (this is basically Krasner's Lemma), but not spherically complete.

3. Prove that every ultrametric field is contained in a spherically complete field with the same value group and residue field. (Hint: take a bad descending sequence of balls, stick something in it without changing the value group or residue field, then Zornicate.) This implies the equivalence between "spherically complete" and "maximally complete".

4. Prove Lemma 2.1 (or look it up in [FvdP, Proposition 1.2.1]).

5. Let k be a field, and let $k((t^{\mathbb{Q}}))$ denote the set of formal sums $\sum_{i\in\mathbb{Q}} c_i t^i$, with each $c_i \in k$, whose support (the set of i such that $c_i \neq 0$) is well-ordered (contains no infinite decreasing subsequence). Prove that formal addition and multiplication of these are well-defined, and that they form a field under these operations. These gadgets are variously called *Hahn series* (because Hahn introduced them in 1907), *Mal'cev-Neumann series* (because Mal'cev and Neumann independently gave vast generalizations), or *generalized power series*. If you really must look this up, see Chapter 13 of Passman's book *The Algebraic Structure of Group Rings*.

6. Prove that $k((t^{\mathbb{Q}}))$ is spherically closed. Deduce that if k is algebraically closed, then so is $k((t^{\mathbb{Q}}))$.

7. (from Bjorn Poonen's undergraduate thesis) Give an explicit construction of a spherically complete field containing $\mathbb{C}_p$. (Hint: you want to do something like making $k((t^{\mathbb{Q}}))$, but starting from $\mathbb{Q}_p^{\mathrm{unr}}$ and writing down "generalized power series in p". You can make that make sense by quotienting an appropriate ring of things looking like generalized power series by the ideal $p - t$. Or see my paper "Power series and p-adic algebraic closures".)

Chapter 3

A little p-adic functional analysis (part 2 of 2)

Thanks to Abhinav Kumar for providing the corrections incorporated into this version (of 16 Sep 04).

Orthonormal bases

An *orthogonal basis* of a Banach space V is a subset $\{e_i\}_{i \in I}$ of V with the property that each $m \in M$ has a unique representation as a convergent sum $\sum_{i \in I} c_i e_i$, and if one always has $\|m\| = \max_{i \in I}\{\|c_i e_i\|\}$. (Note that "convergent" only makes sense if I is at most countable; but note also that there is no distinction between "convergent" and "absolutely convergent" in the nonarchimedean setting!) The basis is *orthonormal* if $\|e_i\| = 1$ for each i. Lemma 1 from last time says that the norm on a Banach space of countable type can be approximated by equivalent norms which admit orthogonal bases; see below for an example.

Banach algebras

A *Banach algebra (over K)* is a Banach space A over K which is also a commutative K-algebra, and which satsifies the following additional restrictions.

(a) $\|1\| = 1$.

(b) for $x, y \in A$, $\|xy\| \leq \|x\| \cdot \|y\|$.

A *Banach module over A* is a Banach space M equipped with an A-module structure, such that $\|am\| \leq \|a\| \cdot \|m\|$ for $a \in A$ and $m \in M$.

Here's an example where life turns out to be easier than in the real/complex case [FvdP, Lemma 1.2.3]. I'll start next time with an example of this.

Lemma 3.1. *Let A be a Banach algebra over K which is noetherian as a ring. Let M be a Banach module over A which is module-finite over A. Then any A-submodule of M is closed.*

Proof. Let N be a submodule of M and let $\tilde{N}$ be the closure of N. Since A is noetherian, $\tilde{N}$ is module-finite over A; choose generators $e_1, \dots, e_n$ of $\tilde{N}$ over A. Consider the A-module homomorphism $A^n \to \tilde{N}$ defined by $(a_1, \dots, a_n) \mapsto \sum a_i e_i$, where A^n is equipped with the norm $\|(a_1, \dots, a_n)\| = \max_i\{\|a_i\|\}$. By the open mapping theorem, there exists $c \in (0, 1)$ such that each $x \in \tilde{N}$ can be written as $\sum a_i e_i$ with

$c \max_i\{\|a_i\|\} \leq \|x\|$. Choose $f_1, \dots, f_n \in N$ with $\|e_i - f_i\| \leq c^2$; we show that $f_1, \dots, f_n$ also generate $\tilde{N}$, which implies that $N = \tilde{N}$.

Given $x \in \tilde{N}$, define the sequence $x_0, x_1, \dots$ as follows. Set $x_0 = x$; given x_j, write $x_j = \sum_i a_{j,i} e_i$ with $c \max_i\{\|a_{j,i}\|\} \leq \|x_j\|$, and put

$$x_{j+1} = \sum a_{j,i}(e_i - f_i),$$

so that $\|x_{j+1}\| \leq c\|x_j\|$. This means $x_j \to 0$ and so for each i, the series $\sum_j a_{j,i}$ converges to a limit a_i satisfying $x = \sum a_i f_i$. Thus $N = \tilde{N}$, as desired. □

Tensor products

The tensor product of two Banach spaces V and W is not complete, so instead we will typically work with the *completed tensor product* $V \widehat{\otimes} W$, which as the name suggests is just the completion of the ordinary tensor product as K-vector spaces.

The completed tensor product is also a Banach space, but this takes a bit of work to check. One gets a seminorm on $V \otimes W$ from the formula

$$\|x\| = \inf\{\max_i\{\|v_i\| \cdot \|w_i\|\}\},$$

the infimum taken over all presentations $x = \sum_{i=1}^m v_i \otimes w_i$, and this extends to the completion. To check that it's a norm, one needs to check that if $\{x_j\}_{j=1}^\infty$ is a sequence of elements of $V \otimes W$, with $x_j = \sum_{j=1}^{m_j} v_{ij} \otimes w_{ij}$, and $\max_i\{\|v_{ij}\| \cdot \|w_{ij}\|\} \to 0$ as $j \to \infty$, then $x_j \to 0$ in $V \otimes W$. We may check this after replacing V and W by the span of the v_{ij} and w_{ij}, respectively, dropping us into the countable type case. Then writing everything in terms of an orthogonal basis of an equivalent norm (using Lemma 1 from last time) yields the claim.

The completed tensor product is exact in the following sense. We say a sequence

$$0 \to M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \to 0$$

of maps between Banach spaces is *exact* if it is exact in the usual sense, and also f and g are isometric. (For f, this means the norm on M_1 is the restriction from M_2; for g, the norm on M_3 is the quotient norm from M_2.) Then for any Banach space N, the sequence

$$0 \to M_1 \widehat{\otimes} N \xrightarrow{f} M_2 \widehat{\otimes} N \xrightarrow{g} M_3 \widehat{\otimes} N \to 0$$

is exact.

Warning: if A is a Banach algebra, the product seminorm on a tensor product $M \otimes_A N$ of Banach modules M and N over A may not be a norm! But see exercises for an important case where this is okay.

Exercises

1. (Leftover from last time) Prove that $\cup_{n=1}^\infty k((t^{1/n}))$ is algebraically closed for $k = \mathbb{C}$, but not for any field k of positive characteristic. (Hint: look at $P(x) = x^p - x - t^{-1}$.)

2. Suppose K is a complete discretely valued (ultrametric) field. Let M be a Banach space over K such that $\|m\| \in |K|$ for each $m \in M$. Put $\mathfrak{o}_M = \{m \in M : \|m\| \leq 1\}$ and $\overline{M} = \mathfrak{o}_M \otimes_{\mathfrak{o}_K} k$. Prove that a subset of M forms an orthonormal basis if and only if its image in $\overline{M}$ is a basis of $\overline{M}$ as a k-vector space. (Hint: see [FvdP, Lemma 1.2.2].)

3. Let A be a Banach algebra which is noetherian as a ring, and let M and N be Banach modules over A which are module-finite over A. Show that the product seminorm

$$\|x\| = \inf\{\max_i\{\|m_i\| \cdot \|n_i\| : x = \sum m_i \otimes n_i\}$$

is a norm on $M \otimes N$, and that $M \otimes N$ is complete for this norm. (Hint: choose finite presentations of M and N and reduce to the case of free modules.)

Chapter 4

Tate algebras (or, Commutative algebra revisited)

We will now talk a bit about Tate algebras, which play a role like that of the polynomial rings over a field in ordinary algebraic geometry.

As usual, K is a complete ultrametric field, which you may assume is discretely valued if you prefer, and k is its residue field. Reminder: I write $\mathfrak{o}_K$ and $\mathfrak{m}_K$ for the valuation subring of K and its maximal ideal, rather than the bizarre K^o and K^{oo} used in [FvdP].

Convention: when I write $\sum_I c_I x^I$, the sum will be running over tuples (normally of nonnegative integers) $I = (i_1, \dots, i_n)$, and $x^I = x_1^{i_1} \cdots x_n^{i_n}$.

References: The main reference is [FvdP, Section 3.1], but I have lots of issues with the presentation, so I've supplemented from [BGR, 5.1]. Also, I'll cite Lang's *Algebra* (third edition; numbering may differ in the current edition) in the exercises as [L].

Tate algebras

The *Tate algebra* (or *standard affinoid algebra*)

$$T_n = T_{n,K} = K\langle x_1, \dots, x_n \rangle$$

is the subring of the ring of formal power series $K[\![x_1, \dots, x_n]\!]$ consisting of sums $\sum_I c_I x^I$ such that $|c_I| \to 0$ as $I \to \infty$; what that really means is that for any $\epsilon > 0$, there are only finitely many tuples I such that $|c_I| > \epsilon$. Such series are sometimes called *strictly convergent power series* (as in [BGR]).

Define the *Gauss norm* on T_n by the formula

$$\left\| \sum_I c_I x^I \right\| = \max_I \{|c_I|\};$$

note that the max is really a max and not a supremum, since the $|c_I|$ tend to 0. This is in fact a norm, under which T_n becomes a Banach algebra over K of countable type (since the monomials x^I have dense span); in fact, it's isomorphic as a Banach space to the space c_0 of null sequences. (Trivial but handy consequence of this definition: the image of the Gauss norm is the same as the image of K under its norm. This will let us do some "normalization" arguments.)

As usual for normed rings, I write $\mathfrak{o}_{T_n}$ to mean the subring of T_n consisting of elements of norm ≤ 1, and $\mathfrak{m}_{T_n}$ to mean the ideal of $\mathfrak{o}_{T_n}$ consisting of elements of norm < 1. Then $\mathfrak{o}_{T_n}/\mathfrak{m}_{T_n} = k[x_1, \dots, x_n]$; given

$f \in \mathfrak{o}_{T_n}$, I'll write $\overline{f}$ for its image in $k[x_1, \dots, x_n]$ and call it the *reduction* of f. (That's why I don't use the overbar for algebraic closures!)

Here are a few basic facts about Tate algebras.

Lemma 4.1. *(a) A series $\sum c_I x^I \in K[\![x_1, \dots, x_n]\!]$ belongs to T_n if and only if the sum $\sum c_I \alpha_1^{i_1} \cdots \alpha_n^{i_n}$ converges in K for any $\alpha_1, \dots, \alpha_n \in \mathfrak{o}_K$.*

(b) Suppose that the residue field $k = \mathfrak{o}_K/\mathfrak{m}_K$ is infinite. Then given a series $\sum c_I x^I \in T_n$, there exist $\alpha_1, \dots, \alpha_n \in \mathfrak{o}_K$ such that

$$\left\| \sum_I c_I x^I \right\| = \left| \sum_I c_I \alpha_1^{i_1} \cdots \alpha_n^{i_n} \right|.$$

(c) The Gauss norm is fully *multiplicative: for all $f, g \in T_n$, $\|fg\| = \|f\| \cdot \|g\|$. (Remember, a Banach algebra is only required to have $\|fg\| \leq \|f\| \cdot \|g\|$.)*

Proof. (a) This is clear; you need only check $\alpha_1 = \cdots = \alpha_n = 1$.

(b) We may as well assume $\sum_I c_I x^I \neq 0$. Let $P(x_1, \dots, x_n)$ be the sum of $c_I x^I$ over all tuples I for which $|c_I|$ is maximal. Then P is a polynomial in $x_1, \dots, x_n$, $\|P\| = \|\sum_I c_I x^I\|$, and

$$\|P(x_1, \dots, x_n) - \sum_I c_I x^I\| < \|P\|.$$

It thus suffices to prove the claim for P instead of the original series. But P can be written as the product of some $c \in K$ with a polynomial P_0 which has coefficients in $\mathfrak{o}_K$ but not all in $\mathfrak{m}_K$. Since k is infinite, the reduction $\overline{P_0}$ does not vanish everywhere. Pick $\alpha_1, \dots, \alpha_n \in \mathfrak{o}_K$ so that $P_0(\alpha_1, \dots, \alpha_n) \notin \mathfrak{m}_K$; then these are a good choice.

(c) Write $f = \sum_I c_I x^I$, $g = \sum_J d_J x^J$, and $fg = \sum_H e_H x^H$. Let I be the tuple maximizing $|c_I|$ which is first in lexicographic order. (That is, you compare tuples by first comparing their first components, then their second if the firsts are tied, then their thirds, and so on.) Likewise, let J be the tuple maximizing $|d_J|$ which is first in lexicographic order. Then e_{I+J} equals $c_I d_J$ plus some other terms of the form $c_{I'} d_{J'}$, where I' and J' are two other tuples adding up to $I + J$. But that means that either I' must precede I in lexicographic order, in which case $|c_{I'}| < |c_I|$ and $|d_{J'}| \leq |d_J|$, or J' must precede J in lexicographic order, in which case $|c_{I'}| \leq |c_I|$ and $|d_{J'}| \leq |d_J|$. In any case, we see that $|e_{I+J} - c_I d_J| < |c_I d_J|$, so $|e_{I+J}| = |c_I d_J$. It follows that $\|fg\| \geq \|f\| \cdot \|g\|$; since we already know the other inequality, we have $\|fg\| = \|f\| \cdot \|g\|$.

□

It is worth noting what the units are in T_n; since we can normalize in T_n, we just treat units of norm 1. (This is [BGR, Proposition 5.1.3/1].)

Lemma 4.2. *For $f \in T_n$ with $\|f\| = 1$, the following are equivalent.*

(a) f is a unit in $\mathfrak{o}_{T_n}$.

(b) f is a unit in T_n.

(c) $\overline{f}$ is constant (i.e., is a unit in $k[x_1, \dots, x_n]$).

(d) $|f(0)| = 1$ and $\|f - f(0)\| < 1$.

Proof. Note that (a) and (b) are equivalent because $\|\cdot\|$ is fully multiplicative. Clearly (a) implies (c), and (c) and (d) are equivalent. Finally, given (d), the series $f(0)\sum_i(1-f/f_0)^i$ converges to a reciprocal of f in $\mathfrak{o}_{T_n}$, so (a) follows. □

Also, note that there is an isomorphism

$$K\langle x_1,\ldots,x_m\rangle\widehat{\otimes}K\langle y_1,\ldots,y_n\rangle \cong K\langle x_1,\ldots,x_m,y_1,\ldots,y_n\rangle.$$

The Tate algebra is meant to be the "ring of functions on the closed unit polydisc". Its properties are a blend between properties of polynomial rings and formal power series rings. This hybrid nature makes it possible to do analytic geometry using Tate algebras using a lot of our intuition from algebraic geometry.

Weierstrass preparation

We say $f \in \mathfrak{o}_{T_n}$ is *(normalized) distinguished (in x_n) of degree d* if

$$\overline{f} = c_0 + c_1 x_n + \cdots + c_d x_n^d$$

with $c_d \in k^*$ and $c_i \in \mathfrak{o}_{T_n}$ for $i = 0,\ldots,d-1$. Note that $\|f\| = 1$, whence the "normalized"; if you allow something of this form times an element of K^*, you get elements which are *distinguished* in the terminology of [BGR]. However, for this lecture, all my distinguished elements will also be normalized, so I won't keep saying "normalized". (Terminology rant: [FvdP] use "regular" for my "normalized distinguished", but the word "regular" will come up later with a more useful meaning.)

Then one has the following results; see [FvdP, Theorem 3.1.1] and [BGR, 5.2]. (Caution: [FvdP] incorrectly applies the moniker "preparation" to (c) instead of (b).)

Theorem 4.3. *(a) (Division algorithm) Suppose $f \in \mathfrak{o}_{T_n}$ is distinguished in x_n of degree d. Then any $g \in T_n$ can be uniquely written as $qf + r$ such that $q \in T_n$ and $r \in T_{n-1}[x_n]$, where the degree of r in x_n is less than d. Moreover, $\|g\| = \max\{\|q\|,\|r\|\}$.*

(b) (Preparation) If $f \in \mathfrak{o}_{T_n}$ is distinguished in x_n of degree d, then there is a unique way to write $f = gh$ with $g \in \mathfrak{o}_{T_{n-1}}[x_n]$ monic of degree d (and hence distinguished) and $h \in \mathfrak{o}_{T_n}^*$.

(c) (Distinction) If $f_1,\ldots,f_m \in \mathfrak{o}_{T_n}$ all have norm 1, then there exists an automorphism τ of T_n (preserving Gauss norms) such that $f_1^\tau,\ldots,f_m^\tau$ are distinguished in x_n.

The last parenthetical is actually moot, as all automorphisms of T_n will preserve the Gauss norm, but this will only become obvious a bit later.

Proof. (a) We first verify uniqueness. If $qf + r = q'f + r'$ are two different decompositions of the same g, then $(q-q')f = r'-r$. By the full multiplicativity of the Gauss norm, this means $\|q-q'\| = \|r'-r\|$; pick some $c \in K$ with $|c| = \|q-q'\|^{-1}$. Then $c(q-q')f = c(r'-r)$ and so the same is true with bars everywhere; but that contradicts uniqueness in the ordinary division algorithm for polynomials.

We next verify the claim about the norm. If $g = qf + r$, then on one hand

$$\|g\| \le \max\{\|qf\|,\|r\|\} = \max\{\|q\|,\|r\|\},$$

and we know we have equality if $\|q\| \neq \|r\|$. But if $\|q\| = \|r\| > \|g\|$, we can choose $c \in K$ with $|c| = \|q\|^{-1}$, and then $cg = cqf + cr$. But the reduction of cg is zero, so by the uniqueness in the ordinary division algorithm, cq and cr would also have to have zero reductions, contradiction. So even in case $\|q\| = \|r\|$ we must have $\|g\| = \max\{\|q\|,\|r\|\}$.

Now for existence. We first check this in case $f = f_0 = c_0 + c_1 x_n + \cdots + c_d x_n^d$ with each $c_i \in \mathfrak{o}_{T_{n-1}}$; this forces $c_d \in \mathfrak{o}_K^*$. Put $g = \sum d_I x^I$, and apply the ordinary ordinary division algorithm to write $x^I = q_I f + r_I$ with $q_I, r_I \in T_{n-1}[z_n]$, with the degree of r_I in z_n being less than d. By what we already showed, we have $\max\{\|q_I\|, \|r_I\|\} = \|x^I\| = 1$; thus the series

$$q = \sum_I d_I q_I, \qquad r = \sum_I d_I r_I$$

converge in T_n and $T_{n-1}[z_n]$, respectively (the latter because I've bounded the degrees, so I really get a polynomial in z_n and not a series). By design, $g = qf + r$.

Now for the general case; write $f = f_0 - D$ where f_0 is as in the previous case and $\|D\| < 1$. Given g, we put $g_0 = g$; given g_i, apply what I just did to write

$$g_i = q_i f_0 + r_i = q_i f + r_i + q_i D$$

and put $g_{i+1} = q_i D$. Then $q = \sum_i q_i$ and $r = \sum r_i$ converge to limits satisfying $g = qf + r$, and r is again a polynomial in z_n of degree less than d.

(b) We first check existence. Apply division to obtain q', r' such that $x_n^d = q'f + r'$, and put $q = x_n^d - r'$; then $q \in \mathfrak{o}_{T_{n-1}}[x_n]$ is monic of degree d and $q'f = q$. On reductions, we have $\overline{q} = \overline{q'f}$, and $\overline{q}$ and $\overline{f}$ are polynomials of the same degree. Hence $\overline{q'}$ is a unit, and so q' is a unit by Lemma 4.2. We can thus factor $f = gh$ with $g = q$ and $h = (q')^{-1}$.

We next check uniqueness. If $f = gh$ is a presentation of the desired form, we have

$$x_n^d = h^{-1} f + (x_n^d - g),$$

and this is what you get from an application of (a). Thus specifying f uniquely determines $x_n^d - g$, and hence determines g and h.

(c) I'll let you find the easy proof for k infinite for yourself; instead, I'll give the slightly more elaborate argument that also works for k finite. Write $f_l = \sum c_{I,l} t^I$ for $l = 1, \dots, m$, and choose integers $e_1, \dots, e_{n-1} \geq 0$ such that whenever $I \neq J$ are among the finitely many tuples with $|c_{I,l}| = |c_{J,l}| = 1$ for some l, we have

$$e_1 i_1 + \cdots + e_{n-1} i_{n-1} + i_n \neq e_1 i_j + \cdots + e_{n-1} j_{n-1} + j_n.$$

(For instance, you can choose the e_i to be successive powers of a bigger than any value of $i_1, \dots, i_n$ that shows up in the tuples.) Let τ be the automorphism which substitutes $x_i + x_n^{e_i}$ in place of x_i for $i = 1, \dots, n-1$ (and fixes x_n). Then

$$\overline{f_l^\tau} = \sum \overline{c_{I,l}} (x_1 + x_n^{e_1})^{i_1} \cdots (x_{n-1} + x_n^{e_{n-1}})^{i_{n-1}} x_n^{i_n};$$

if you pick out the unique tuple $i_1, \dots, i_n$ maximizing $e = e_1 i_1 + \cdots + e_{n-1} i_{n-1} + i_n$, then in the reduction, the unique term of highest degree that you see is x_n^e. That means each f_l^τ is distinguished. □

More properties of Tate algebras

Weierstrass preparation immediately yields the analogue of the Hilbert basis theorem for Tate algebras; in the case of K discretely valued, this is a theorem of Fulton. (Yes, *that* Fulton! See: A note on weakly complete algebras, *Bull. Amer. Math. Soc.* **75** (1969), 591–593.)

Proposition 4.4 (Hilbert basis theorem)**.** *The ring T_n is noetherian.*

Proof. Induction on n. Given a nonzero ideal I of T_n, choose $f \in I$ nonzero, and apply distinction to find an automorphism τ of T_n such that f^τ is distinguished in x_n of some degree d. Using division, we see that I^τ is generated by f^τ together with $I^\tau \cap T_{n-1}[z_n]$. By the induction hypothesis, T_{n-1} is noetherian, as then is $T_{n-1}[z_n]$ by the usual Hilbert basis theorem. Thus I^τ is finitely generated, as then is I. □

Proposition 4.5. *The ring T_n is a unique factorization domain.*

Proof. Again, induct on n. Given $f \in T_n$, suppose $f = g_1 \cdots g_m = h_1 \cdots h_n$ are two factorizations of f into irreducibles; by pushing scalars around, we may reduce to the case where $\|g_i\| = \|h_j\| = 1$ for all i, j. By distinction, there exists an automorphism τ of T_n such that f, the g_i^τ, and the h_j^τ are all distinguished. By preparation, we can write each $g_i^\tau = P_i u_i$ with $P_i \in \mathfrak{o}_{T_{n-1}}[x_n]$ monic and $u_i \in \mathfrak{o}_{T_n}^*$, and likewise write $h_j^\tau = Q_j v_j$ with $Q_j \in \mathfrak{o}_{T_{n-1}}[x_n]$ monic and $v_j \in \mathfrak{o}_{T_n}^*$. Then $P_1 \cdots P_m$ equals $Q_1 \cdots Q_n$ times a unit, but both sides are monic polynomials in x_n over $\mathfrak{o}_{T_{n-1}}$, necessarily of the same degree (since that's true on the reductions). Thanks to the uniqueness of the preparation of $P_1 \cdots P_m$, we must in fact have $P_1 \cdots P_m = Q_1 \cdots Q_n$. Moreover, each P_i and Q_j is irreducible in T_n, hence also in $T_{n-1}[x_n]$; thus the factorizations agree up to units, by the unique factorization theorem for polynomials over a UFD (i.e., unique factorization over a field plus "Gauss's lemma"). That proves that the original factorizations of f^τ, and hence of f, agree up to units. □

Proposition 4.6. *The Krull dimension of T_n is n.*

Proof. The sequence of prime ideals

$$(0) \subset (x_1) \subset (x_1, x_2) \subset \cdots (x_1, x_2, \cdots, x_n)$$

shows that the Krull dimension is at least n. On the other hand, for any irreducible $f \in T_n$, by distinction plus preparation, $T_n/(f)$ is finite over T_{n-1}, and so has Krull dimension $n-1$. (Remember that making a finite ring extension of a noetherian ring cannot increase its Krull dimension.) Hence T_n has Krull dimension at most n, yielding the claim. □

Affinoid algebras and Noether normalization

An *affinoid algebra* is a K-algebra A of the form $T_n/\mathfrak{a}$ for some ideal $\mathfrak{a}$. By Fulton's theorem, A is noetherian.

Note that there are a couple of minor discrepancies between this definition and the one in [FvdP, 3.1]. For one, they use the term "Tate algebra" to mean *any* affinoid algebra; I have seen this elsewhere, but I still think it is nonstandard (e.g., [BGR] does not do this). For another, they define affinoid algebras as integral extensions of Tate algebras. This amounts to using the business end of the Noether normalization theorem (see below), and strikes me as bizarre.

Proposition 4.7 (Noether normalization)**.** *Let $\mathfrak{a}$ be an ideal of T_n, and let $A = T_n/\mathfrak{a}$ be the corresponding affinoid algebra. Then there exists a finite injective map $T_d \to A$ for some d; moreover, the Krull dimension of A is equal to d.*

Proof. We first prove the existence of the map, by induction on n. We may as well assume $\mathfrak{a}$ is a nontrivial ideal; by distinction and preparation, after applying an automorphism of T_n we may assume that $\mathfrak{a}$ contains a monic polynomial $f \in T_{n-1}[x_n]$. Then $T_n/(f)$ is finite over T_{n-1}; if we put $\mathfrak{b} = \mathfrak{a} \cap T_{n-1}$, then $T_n/\mathfrak{a}$ is finite over $T_{n-1}/\mathfrak{b}$. By the induction hypothesis, $T_{n-1}/\mathfrak{b}$ is finite over some T_d, yielding the claim.

For the statement about the Krull dimension, we need only recall that for $A \to B$ a finite injective homomorphism of noetherian rings, the rings A and B have the same Krull dimension, and that the Krull dimension of T_d is d by Proposition 4.6. □

Warning: unlike in the polynomial situation, an affinoid algebra can have an affinoid subalgebra of greater Krull dimension! See [FvdP, Exercises 3.2.2].

Note that the Nullstellensatz for Tate algebras falls out as a consequence.

Corollary 4.8. *For any maximal ideal $\mathfrak{m}$ of T_n, the field $T_n/\mathfrak{m}$ is finite over K.*

Proof. A field has Krull dimension 0, so by Noether normalization, there must exist a finite map $K = T_0 \to T_n/\mathfrak{m}$. $\square$

Affinoid algebras are Banach algebras (and canonically so!)

Recall the following lemma (Lemma 1 from "functional analysis, part 2").

Lemma 4.9. *Let A be a Banach algebra over K which is noetherian as a ring. Let M be a Banach module over A which is module-finite over A. Then any A-submodule of M is closed.*

In particular, since T_n is noetherian, any ideal of T_n is closed. Hence any affinoid algebra A inherits from a presentation $T_n \to A$ a quotient norm, under which it becomes a Banach algebra.

It turns out that the topology of an affinoid algebra is uniquely determined by its K-algebra structure, and all K-algebra homomorphisms of affinoid algebras. To see this, we need to back up and do a little more functional analysis; the following is [BGR, Proposition 3.7.5/2].

Proposition 4.10. *Let B be a noetherian Banach algebra over K, and suppose there exists a collection S of ideals of B such that:*

(a) for any $\mathfrak{b} \in S$, $\dim_K B/\mathfrak{b} < \infty$;

(b) $\cap_{\mathfrak{b} \in S} \mathfrak{b} = (0)$.

Then any K-algebra homomorphism of a noetherian Banach algebra A over K into B is continuous (hence bounded).

Proof. Let $f : A \to B$ be a K-algebra homomorphism; we will apply the closed graph theorem to the graph of f. Namely, we need to show that if $\{x_n\}$ is a null sequence in A and $f(x_n) \to y$ in B, then $y = 0$.

Pick an ideal $\mathfrak{b} \in S$; note that $\mathfrak{b}$ is closed, by Lemma 4.9. Put $\mathfrak{a} = f^{-1}(\mathfrak{b})$, so that $\mathfrak{a}$ is closed in A, each of $A/\mathfrak{a}$ and $B/\mathfrak{b}$ inherits a quotient norm, and $A/\mathfrak{a} \to B/\mathfrak{b}$ is injective. Since $B/\mathfrak{b}$ is finite dimensional over K, so is $A/\mathfrak{a}$, and any linear map between finite dimensional K-vector spaces is continuous. (Remember that there is only one equivalence class of norms on a finite dimensional K-vector space!) Thus we must have $y \in \mathfrak{b}$ for each $\mathfrak{b}$; since the $\mathfrak{b}$ have trivial intersection, we have $y = 0$. $\square$

Corollary 4.11. *Any K-algebra homomorphism between affinoid algebras is continuous. In particular, for a fixed affinoid algebra A, the quotient norms induced by different presentations $T_n \to A$ are all equivalent.*

Now I can clarify my remark from earlier about norm-preserving automorphisms of T_n.

Corollary 4.12. *Any automorphism of T_n preserves the Gauss norm.*

Proof. This follows from the previous corollary and the fact that we can recover the Gauss norm from the topology of T_n. Namely, an element $f \in T_n$ satisfies $\|f\| \leq 1$ if and only if for any null sequence $\{c_n\}$ in K, $\{c_n f\}$ is null in T_n. $\square$

In other words, the Gauss norm is a "canonical" norm for T_n. There is an analogous function on an arbitrary affinoid algebra, but it's only a seminorm in general; we'll talk more about it soon.

Exercises

These exercises are from Section 6 of my preprint "Full faithfulness for overconvergent F-isocrystals", which you can grab from the arXiv if you get stuck.

1. Using distinction, prove that the action of $\mathrm{GL}_n(T_n)$ on n-tuples which generate the unit ideal is transitive. (Hint: compare to a proof of the analogous statement for polynomials; see [L, Theorem XXI.3.4].)

2. Prove that every finitely generated module over T_n has a finite free resolution. (Hint: this time, see [L, Theorem XXI.3.6].)

3. Prove the analogue of the Quillen-Suslin theorem for T_n: every finitely generated projective module over T_n is free. (Hint: use [L, Theorem XXI.2.1] to show that a finite projective is stably free, that is, its direct sum with some free is free. Then compare [L, Theorem XXI.3.6].)

Chapter 5

Affinoid algebras and their spectra

Last time we discussed Tate algebras and defined affinoid algebras to be the quotients of Tate algebras. This time, we'll begin the process of turning those algebras into "spaces" by studying their maximal spectra. Of course, this is not the right way to go in the long run; this is analogous to constructing varieties, whereas one should be doing something more "schematic". We'll return to this point when we discuss Berkovich spaces.

Warning: I dashed these notes off in a bit of a hurry, so there are probably lots of mistakes, which we'll doubtless find in class. As usual, if you send me corrections by email, I'll change the file on the web accordingly. Sorry about that.

References: [FvdP, Sections 3.3 and 3.4], [BGR, Section 3.8] (and various other sections of BGR which you'll find via the index, sigh).

Review: Newton polygons

Probably you've all seen this before, but just in case, let me review the theory of the Newton polygon of a polynomial over an ultrametric field. (I really should have done this back in the ultrametric field section.)

Lemma 5.1. *Let L be a complete ultrametric field, and let $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ be a polynomial over L with roots $\alpha_1, \dots, \alpha_n \in L^{\mathrm{alg}}$ (repeated with appropriate multiplicities). Sort the roots so that $|\alpha_1| \geq |\alpha_2| \geq \cdots |\alpha_n|$. Then for $i = 1, \dots, n$,*

$$|a_{n-i}| \leq |\alpha_1 \cdots \alpha_i|,$$

with equality whenever $|\alpha_i| > |\alpha_{i+1}|$, or when $i = n$.

Proof. Write $P(z) = \prod_{i=1}^n (z - \alpha_i)$. Then a_{n-i} is, up to sign, the sum of the i-fold products of the α's, so the desired inequality is clear. As for the equality, note that if $|\alpha_i| > |\alpha_{i+1}|$, then $\alpha_1 \cdots \alpha_i$ is the unique i-fold product of maximum norm. □

Corollary 5.2. *The maximum of $|\alpha|$ over the roots of P is equal to $\max_j\{|a_{n-j}|^{1/j}\}$.*

If you want all of the absolute values of the roots, you get them from the lemma as follows. Consider the set of points $(n - i, -\log a_i)$ for $i = 0, \dots, n$ (where $a_n = 1$ by convention). Form the lower convex hull of this set of points: the lower boundary of this hull is called the *Newton polygon* of P. If r occurs as a slope in this polygon of a segment of width j, then there are exactly j roots of P of norm $\exp(r)$.

The maximal spectrum and the spectral seminorm

Let A be an affinoid algebra, and let $\operatorname{Max} A$ denote the set of maximal ideals of A, a/k/a the *maximal spectrum* of A. We refer to $\operatorname{Max} A$ as the *affinoid space* associated to A, although for the time being it's just a set; we'll give it a "topology" and a ringed space structure later.

By the Nullstellensatz for Tate algebras (proved last time), for each $\mathfrak{m} \in A$, $A/\mathfrak{m}$ is finite dimensional over K; in particular, it admits a unique extension of the norm on K. We will use function-theoretic notation to speak about $\operatorname{Max} A$; that is, if x is a "point" of $\operatorname{Max} A$ corresponding to the maximal ideal $\mathfrak{m}_x$ and $f \in A$, we will write $f(x)$ to mean the image of f in $A/\mathfrak{m}_x$.

Define the *spectral seminorm* of A (or $\operatorname{Max} A$) as

$$\|f\|_{\mathrm{spec}} = \sup_{x \in \operatorname{Max} A} |f(x)|.$$

(The term *supremum seminorm* is used interchangeably, as in [BGR].) It is straightforward to check that $\|\cdot\|_{\mathrm{spec}}$ is actually a seminorm, and that $\|fg\|_{\mathrm{spec}} \leq \|f\|_{\mathrm{spec}} \|g\|_{\mathrm{spec}}$. The spectral seminorm is a norm if and only if the intersection of the maximal ideals of A is the zero ideal; in that case, we also call it the *spectral norm*. It will turn out that this happens if and only if A is reduced; see below.

Note that from what we showed last time, the spectral norm on T_n is precisely the Gauss norm, and the supremum defining the spectral norm is achieved at some point. (More precisely, last time we showed that if K has infinite residue field, the supremum is actually achieved at some K-rational point. But the same argument shows that for any K, the supremum is achieved at some point defined over a finite extension of K.) That this holds in general is the content of the "maximum modulus principle" for affinoid spaces; see below.

Proposition 5.3 (Maximum modulus principle)**.** *For A an affinoid algebra and $f \in A$, there exists $x \in \operatorname{Max} A$ such that $\|f\|_{\mathrm{spec}} = |f(x)|$. Moreover, if $\|f\|_{\mathrm{spec}} = 0$, then f is nilpotent.*

Proof. There is no harm in quotienting A by its nilradical, since computing $|f(x)|$ is insensitive to nilpotents. That is, we may assume A is reduced. We may also assume A is integral, since otherwise we can check the claim on each connected component of A.

So assume that A is an integral domain. By Noether normalization, A can be written as a finite integral extension of some T_d. That means there is an irreducible polynomial $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ over T_d such that $P(f) = 0$. (Note that the coefficients of P lie in T_d and not its fraction field, because T_d is a unique factorization domain and so "Gauss's lemma" applies.) From the theory of the Newton polygon, for any $x \in \operatorname{Max} A$ lying over $y \in \operatorname{Max} T_d$, we have $|f(x)| = \max_i\{|a_{n-i}(y)|^{1/i}\}$. Running this over all x and y, we get

$$\|f\|_{\mathrm{spec}} = \max_i\{\|a_{n-i}\|_{\mathrm{spec}}^{1/i}\},$$

and the maximum on the right is achieved because we already know the maximum modulus principle for T_d. This proves the first claim. As for the second claim, note that $\|f\|_{\mathrm{spec}} = 0$ implies $\|a_{n-i}\|_{\mathrm{spec}} = 0$ for all i. Since the spectral seminorm on T_d is a norm, we have $a_{n-i} = 0$ for all i, and so f is nilpotent. $\square$

Corollary 5.4. *For A an affinoid algebra, the intersection of the maximal ideals of A equals the nilradical of A. (That is, A is a "Jacobson ring".) In particular, the spectral seminorm is a norm if and only if A is reduced.*

Let $\mathfrak{o}_A^{\mathrm{spec}}$ be the subring of A consisting of those $f \in A$ for which $|f|_{\mathrm{spec}} \leq 1$.

Lemma 5.5. *For $\phi : A \to B$ a finite injective homomorphism of affinoid algebras and $f \in A$, one has $\|f\|_{\mathrm{spec}} = \|\phi(f)\|_{\mathrm{spec}}$.*

Proof. The finiteness of ϕ means that the map $\operatorname{Max} B \to \operatorname{Max} A$ is surjective, from which the claim follows. $\square$

Lemma 5.6. *Suppose $\phi : T_d \to A$ is a finite injective K-algebra homomorphism. Then $\mathfrak{o}_A^{\mathrm{spec}}$ is integral over $\phi(\mathfrak{o}_{T_d}^{\mathrm{spec}})$.*

Note that $\phi(\mathfrak{o}_{T_d}^{\mathrm{spec}}) \subseteq \mathfrak{o}_A^{\mathrm{spec}}$ by the previous lemma, so the statement makes sense.

Proof. If A is an integral domain, the proof of the maximum modulus principle yields that any $f \in \mathfrak{o}_A^{\mathrm{spec}}$ is the root of a polynomial over $\phi(\mathfrak{o}_{T_d}^{\mathrm{spec}})$. For the reduction to this case, see [FvdP, Proposition 3.4.5]. $\square$

Lemma 5.7. *For A an affinoid algebra, $\|\cdot\|$ a norm obtained as the quotient norm from some surjection $T_n \to A$, $f \in A$, and $x \in \operatorname{Max} A$, one has $\|f\| \geq |f(x)|$.*

It will turn out that this also holds for an arbitrary Banach norm on A; see next handout.

Proof. Since $\|\cdot\|$ is a quotient norm, we can find $g \in T_n$ mapping to f in A such that $\|g\| = \|f\|$. Let $y \in \operatorname{Max} T_n$ be the point over which x lies; then $\|g\| \geq |g(y)|$ because the Gauss norm on T_n coincides with the spectral seminorm. But $g(y) = f(x)$, so the claim follows. $\square$

Lemma 5.8. *For A an affinoid algebra under some Banach norm $\|\cdot\|$ and $f \in A$, one has $\|f\|_{\mathrm{spec}} \leq 1$ if and only if the sequence $\{\|f^j\|\}_{j=1}^{\infty}$ is bounded.*

Proof. Both conditions are preserved under changing to an equivalent norm, and any two Banach norms on A are equivalent (see last handout), so we may as well take $\|\cdot\|$ to be the quotient norm from some surjection $T_n \to A$, so that we can apply the previous lemma.

If $\|f^j\|$ is bounded, then $|f^j(x)|$ is bounded for each $x \in \operatorname{Max} A$ (because $|f^j(x)| \leq \|f^j\|$ by the previous lemma), so $|f(x)| \leq 1$, and so $\|f\|_{\mathrm{spec}} \leq 1$. Conversely, suppose $\|f\|_{\mathrm{spec}} \leq 1$. Write A as a finite integral extension of some T_d. By the previous lemma, f is integral over $\mathfrak{o}_{T_d}^{\mathrm{spec}}$, i.e., it is the root of some $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ with $a_i \in T_d$ and $\|a_i\|_{\mathrm{spec}} \leq 1$. Each power of f can be written as a linear combination of $1, f, \dots, f^{n-1}$ whose coefficients are polynomials in the a_i with integer coefficients. The set of such polynomials is bounded under the Gauss norm in T_d, so by the continuity of the map $T_d \to A$, it is also bounded under the norm on A. Thus the powers of f are bounded. $\square$

The following corollary of the previous proposition might have been what you were expecting when I first said "spectral seminorm".

Corollary 5.9. *Let A be an affinoid algebra with norm $\|\cdot\|$. Then for $f \in A$,*

$$\|f\|_{\mathrm{spec}} = \lim_{n \to \infty} \|f^n\|^{1/n}.$$

More on spectral seminorms in the next handout.

Exercises

1. Suppose A is an affinoid algebra and $f \in A$. Prove that the following are equivalent:
   (a) $\inf\{|f(x)| : x \in \operatorname{Max} A\} > 0$;
   (b) $f(x) \neq 0$ for all $x \in \operatorname{Max} A$;
   (c) $f \in A^*$.

   (Note: this is exactly [FvdP, Exercise 3.3.4(1).])

2. Suppose $\rho_1, \dots, \rho_n \in (0, \infty)$ are such that some power of each ρ_i belongs to $|K^*|$. Prove that the "modified Tate algebra"

$$T_{n,\rho} = \{\sum_I c_I x^I \in K[\![x_1, \dots, x_n]\!] : |c_I| \rho_1^{i_1} \cdots \rho_n^{i_n} \to 0\}$$

is an affinoid algebra. (See [FvdP, Exercise 3.3.4(5)].)

3. Suppose $\rho \in (0, \infty)$ is such that no power of ρ belongs to $|K^*|$. Prove that $T_{1,\rho}$ is *not* an affinoid algebra. (Hint: what is the spectral norm of t_1?)

Chapter 6

More on affinoid algebras (amended 4 Oct 04)

Addenda on the spectral seminorm

A norm on a Banach algebra A is *power-multiplicative* if $\|f^n\| = \|f\|^n$ for any $f \in A$ and any positive integer n. Our proof that the Gauss norm on T_n has a topological characterization adapts to show that for any affinoid algebra A, there is at most one power-multiplicative Banach norm on A. We now know that if a power-multiplicative norm exists, it must be the spectral seminorm; hence such a norm exists if and only if A is reduced.

In fact, the spectral seminorm is "minimal" in the following sense [BGR, Corollary 3.8.2/2].

Proposition 6.1. *Let A be an affinoid algebra equipped with a Banach norm $\|\cdot\|$. Then for all $f \in A$, $\|f\|_{\text{spec}} \leq \|f\|$. In particular, $|f(x)| \leq \|f\|$ for any $x \in \operatorname{Max} A$ (we already showed this if $\|\cdot\|$ is a quotient norm from a Tate algebra).*

Proof. Apply the formula

$$\|f\|_{\text{spec}} = \lim_{n\to\infty} \|f^n\|^{1/n}$$

and note that $\|f^n\| \leq \|f\|^n$ because $\|\cdot\|$ is a Banach algebra norm. □

This yields the following characterization of nilpotent elements, in the vein of our characterization of power-bounded elements [BGR, Proposition 6.2.3/2].

Proposition 6.2. *For A an affinoid algebra and $f \in A$, the following statements are equivalent:*

(a) f is topologically nilpotent (i.e., $\{f^n\}$ is a null sequence in A);

(b) $|f(x)| < 1$ for all $x \in \operatorname{Max} A$;

(c) $\|f\|_{\text{spec}} < 1$.

Proof. The equivalence of (b) and (c) follows from the maximum modulus principle, and (a) implies (c) by the previous proposition. Given (c), choose $c \in K$ and $m \in \mathbb{N}$ such that $|c| > 1$ but $\|cf^m\|_{\text{spec}} \leq 1$. Then cf^m is power-bounded (from last time), so $c^{-1}(cf^m) = f^m$ is topologically nilpotent, as then is f. Thus (c) implies (a), and we are done. □

Spectral norms are Banach norms

We now know that the spectral seminorm on a reduced affinoid algebra is a norm. However, more than that is true: it is a Banach norm. (This proof is from [FvdP, Theorem 3.4.9]; the proof in [BGR, Theorem 6.2.4/1] is a bit more intricate.)

We say a bounded map $f : A \to B$ of Banach spaces is *strict* if the bijection between $A/\ker(f)$ with its quotient topology and $\operatorname{im}(f)$ with its subspace topology is an isomorphism of normed spaces. Since $\ker(f)$ is closed, by the open mapping theorem this is equivalent to saying that $\operatorname{im}(f)$ is closed in B.

Note that if A and B are affinoid algebras and f is finite, it is automatically strict, because $f(A)$ is an A-submodule of B, and any submodule of a finitely generated module over an affinoid algebra is closed (see lemma from "p-adic functional analysis 2"). In fact, if A is an affinoid algebra and M_1, M_2 are finitely generated modules over A, then every A-algebra homomorphism $M_1 \to M_2$ is strict.

On the other hand, here's an example of a bounded non-strict map: the homomorphism $T_1 \to T_1$ sending x_1 to cx_1, where $c \in K$ satisfies $|c| < 1$.

Lemma 6.3. *Let $A \hookrightarrow B$ be a strict inclusion of affinoid algebras. Then the spectral seminorm on B restricts to the spectral seminorm on A.*

Proof. Choose a Banach norm $\|\cdot\|$ on B; it then restricts to a Banach norm on A by strictness (and the fact that any two Banach norms on an affinoid algebra are equivalent), and applying the formula $\|f\|_{\text{spec}} = \lim_{n\to\infty} \|f^n\|^{1/n}$ gives us the claim. □

Theorem 6.4. *Let A be a reduced affinoid algebra. Then A is complete under $\|\cdot\|_{\text{spec}}$. In particular, every Banach algebra norm on A is equivalent to the spectral norm.*

Proof. The last sentence will follow from what we showed earlier: any two Banach algebra norms on an affinoid algebra are equivalent. So we focus on showing that A is complete.

We first reduce to the case where A is an integral domain. Let $\mathfrak{p}_1, \dots, \mathfrak{p}_m$ be the minimal primes of A. Choose a Banach norm $\|\cdot\|_A$ on A, put $A_i = A/\mathfrak{p}_i$ (which is an integral domain), and equip each A_i with the quotient norm induced by $\|\cdot\|_A$. Then $A_1 \oplus \cdots \oplus A_m$ becomes a finitely generated Banach module over A under the max norm

$$(a_1, \dots, a_m) = \max_i \{\|a_i\|\}.$$

Let $\iota : A \to A_1 \oplus \cdots \oplus A_m$ be the canonical injection; then $\iota(A)$ is an A-submodule of $A_1 \oplus \cdots \oplus A_m$. By the lemma from "$p$-adic functional analysis 2", $\iota(A)$ is closed in $A_1 \oplus \cdots \oplus A_m$, and ι is an isomorphism onto its image by the open mapping theorem. The map ι is isometric for the spectral norms (because $\operatorname{Max} A = \cup_i \operatorname{Max} A_i$), so proving that the spectral norm on each A_i is equivalent to the quotient norm proves that the spectral norm on A is equivalent to $\|\cdot\|_A$.

From now on, we assume A is an integral domain. If B is a reduced affinoid algebra containing A and $A \hookrightarrow B$ is strict (e.g., if $A \hookrightarrow B$ is finite), showing that B is complete under its spectral seminorm implies that A is complete under its spectral seminorm, thanks to the previous lemma. In particular, we can write A as a finite integral extension of T_d for some $d \geq 0$, and take B to be the integral closure of T_d in the normal closure of $\operatorname{Frac} A$ over $\operatorname{Frac} T_d$. This lets us break the problem into two steps.

(a) Show that if A is finite over T_d and $\operatorname{Frac} A$ is purely inseparable over $\operatorname{Frac} T_d$, then the spectral norm on A is complete.

(b) Show that if B is finite over A, $\operatorname{Frac} B$ is Galois over $\operatorname{Frac} A$, and the spectral norm on A is fully multiplicative and complete, then the spectral norm on B is complete.

So we work on these two steps separately.

(a) There is nothing to show unless the characteristic of K is $p > 0$. Note that for any affinoid algebra A over K and any field extension K' containing K and complete for a norm extending that on K, the inclusion of A into the completed tensor product $A\widehat{\otimes}_K K'$ is strict. In particular, we may reduce to the case where K is algebraically closed. In that case,

$$A \subseteq K\langle x_1^{1/p^m}, \dots, x_d^{1/p^m}\rangle$$

for some m, so it's enough to check the completeness of $K\langle x_1^{1/p^m}, \dots, x_d^{1/p^m}\rangle$ under its spectral norm. But again, this is true because that spectral norm is the Gauss norm.

(b) Put $G = \mathrm{Gal}(\operatorname{Frac} B/\operatorname{Frac} A)$, and put

$$\mathrm{Trace}(f) = \sum_{g\in G} f^g;$$

this gives a map from $\operatorname{Frac} B$ to $\operatorname{Frac} A$ such that $\|\mathrm{Trace}(f)\|_{A,\mathrm{spec}} \leq \|f\|_{B,\mathrm{spec}}$. (This trace is the same as the trace of multiplication by f as a $\operatorname{Frac} A$-linear transformation on $\operatorname{Frac} B$.) From basic algebra, we know that the $\operatorname{Frac} A$-linear pairing

$$(x, y) \mapsto \mathrm{Trace}(xy)$$

on $\operatorname{Frac} B$ is nondegenerate.

Choose $e_1, \dots, e_n \in B$ which form a basis for $\operatorname{Frac} B$ over $\operatorname{Frac} A$, and let $e_1^*, \dots, e_n^*$ be the dual basis for the trace pairing. We show that $Ae_1 + \cdots + Ae_n$ is a Banach module under the spectral norm, by showing that the spectral norm is equivalent to the maximum norm

$$\|f_1e_1 + \cdots + f_1e_n\|_{B,\mathrm{spec}} = \max_i\{\|f_i\|_{A,\mathrm{spec}}\}.$$

Choose $a_0 \in A$ such that $a_0e_j^* \in \mathfrak{o}_B^{\mathrm{spec}}$ for $j = 1, \dots, n$; now given $f_1e_1 + \cdots + f_ne_n \in Ae_1 + \cdots + Ae_n$, we have

$$\begin{aligned} a_0f_j &= \mathrm{Trace}(a_0e_j^* \sum_i f_ie_i) \\ \|\mathrm{Trace}(a_0e_j^* \sum_i f_ie_i)\|_{A,\mathrm{spec}} &\leq \|a_0e_j^* \sum_i f_ie_i\|_{B,\mathrm{spec}} \\ &\leq \|\sum_i f_ie_i\|_{B,\mathrm{spec}}. \end{aligned}$$

Therefore

$$\|\sum_i f_ie_i\|_{B,\mathrm{spec}} \geq \|a_0\|_{A,\mathrm{spec}} \max_i\{\|f_i\|_{A,\mathrm{spec}}\}.$$

Since we also have

$$\|\sum_i f_ie_i\|_{B,\mathrm{spec}} \leq \max_i\{\|f_i\|_{A,\mathrm{spec}}\} \max_i\{\|e_i\|_{B,\mathrm{spec}}\},$$

the spectral norm restricted to $Ae_1 + \cdots + Ae_n$ is equivalent to the maximum norm, and so is a Banach norm.

For some $a \in A$, $aB \subseteq Ae_1 + \cdots + Ae_n$; since the spectral norm on A is multiplicative, the spectral norm on B is thus complete.

□

Note that this theorem can also be interpreted as follows: a sequence of elements of A converges to zero under some (any) Banach algebra norm if and only if it converges uniformly to zero on $\operatorname{Max} A$.

The reduction of an affinoid algebra

In case you are wondering when the spectral seminorm is not just a norm but is actually fully multiplicative (like the Gauss norm), here is your answer. Recall that for A an affinoid algebra, we defined

$$\mathfrak{o}_A^{\mathrm{spec}} = \{f \in A : \|f\|_{\mathrm{spec}} \leq 1\}.$$

Now define

$$\mathfrak{m}_A^{\mathrm{spec}} = \{f \in A : \|f\|_{\mathrm{spec}} < 1\}$$

and $\overline{A}^{\mathrm{spec}} = \mathfrak{o}_A^{\mathrm{spec}}/\mathfrak{m}_A^{\mathrm{spec}}$; we call the latter the *reduction* of A. Then we have the following [BGR, Proposition 6.2.3/5].

Proposition 6.5. *The spectral seminorm is a fully multiplicative norm if and only if A is reduced and $\overline{A}$ is an integral domain.*

Note that A being an integral domain is not enough; see exercises.

Proof. We already know that the spectral seminorm is a norm if and only if A is reduced; also, if A is reduced and the spectral norm is fully multiplicative, then the product of elements of spectral norm 1 again has spectral norm 1, so $\overline{A}$ is an integral domain. Conversely, suppose A is reduced and $\overline{A}$ is an integral domain. Given $f, g \in A$ nonzero, there exists an integer n such that $\|f\|_{\mathrm{spec}}^n$ and $\|g\|_{\mathrm{spec}}^n$ belong to $|K^*|$, by the maximum modulus principle. (Namely, the spectral seminorm is always the norm of the evaluation of f at some point whose residue field is finite over K.) Choose $c, d \in K^*$ with $c\|f\|_{\mathrm{spec}}^n = d\|g\|_{\mathrm{spec}}^n = 1$. Then the product of the images of cf^n and dg^n in $\overline{A}$ must be nonzero because $\overline{A}$ is an integral domain; that is, $\|cf^n dg^n\|_{\mathrm{spec}} = 1$. Hence

$$1 = \|cf^n dg^n\|_{\mathrm{spec}} = c\|f\|_{\mathrm{spec}}^n \cdot d\|g\|_{\mathrm{spec}}^n$$

and (by power-multiplicativity of the spectral seminorm) it follows that $\|fg\|_{\mathrm{spec}} = \|f\|_{\mathrm{spec}} \cdot \|g\|_{\mathrm{spec}}$. □

Exercises

1. Give an explicit example of an affinoid algebra A which is an integral domain, but whose spectral seminorm is not fully multiplicative. (Hint: consider power series in x and x^{-1} which converge on a suitable annulus; your motivation should functions on a complex annulus which have their suprema on opposite boundary components. We'll look more at this geometric situation in the next few lectures.)

2. Let A be an affinoid algebra, and suppose $f \in A$ satisfies $\|f\|_{A,\mathrm{spec}} = 1$. Prove that the natural map from A to $B = A\langle x\rangle/(fx - 1)$ is strict. (Such a B is called a *strict localization* of A.)

Chapter 7

Even more on affinoid algebras

Ruochuan caught me last time using the following fact without justification, in the proof that spectral norms are Banach norms. In fact, this is [FvdP, Theorem 3.5.1], so I might as well explain this. (Actually, I only used part (b) in that proof, so the annoying part (a) below wasn't really needed. But let's explain it anyway.)

Theorem 7.1. *Let A be a reduced integral affinoid algebra. Then the integral closure of A in its fraction field is finitely generated as an A-module. (Corollary: the integral closure of A in any finite extension of its fraction field is also finite over A.)*

Proof. We can write A as a finite integral extension of some T_d by Noether normalization, so it is enough to show that the integral closure of T_d in any finite extension of its fraction field is finite over T_d. As in the previous proof, this breaks down into two steps.

(a) If L is a finite purely inseparable extension of T_d, then the integral closure of T_d in L is finite over T_d.

(b) If A is an affinoid algebra which is a normal (integrally closed in its fraction field) domain, and L is a finite Galois extension of $\operatorname{Frac} A$, then the integral closure of A in L is finite over A.

Part (b) is easy: let B be the integral closure of A in L, choose $e_1, \dots, e_n \in B$ which form a basis of L over $\operatorname{Frac} A$, and let $e_1^*, \dots, e_n^*$ be the dual basis for the trace pairing. Now simply pick $a \in A$ such that $ae_1^*, \dots, ae_n^* \in B$; then note that $af \in Ae_1 + \cdots + Ae_n$ for any $f \in B$. So B is contained in a finitely generated A-module; since A is noetherian, A is finitely generated.

Part (a) is a bit more annoying. It's clear if K is perfect, as in that case we can pass from L to the fraction field of $K\langle x_1^{1/p^n}, \dots, x_d^{1/p^n}\rangle$ for some n, and the integral closure in that field is clearly the bigger Tate algebra, which is visibly finite over T_d. You can still argue like this if $[K : K^p] < \infty$, but otherwise it gets messy. Let K' be the completed algebraic closure of K; the point is that whatever generators you get of the integral closure of $K'\langle x_1, \dots, x_d\rangle$ in $K'\langle x_1^{1/p^n}, \dots, x_d^{1/p^n}\rangle$ can be approximated by generators which are actually integral over T_d. (Compare the argument in the proof of the "closed submodule principle", i.e., [FdvP, Lemma 1.2.3], or just look this up in [FvdP, Theorem 3.5.1].)

□

Chapter 8

Subsets of the projective line, part 1

Before proceeding with more general theory, let's look at some concrete examples of affinoid spaces. We'll also look at some spaces that by all rights should be admissible in rigid analytic geometry but which are not affinoid.

Reference: [FvdP, Chapter 2].

Affinoid subsets of $\mathbb{P}^1$

As usual, $\mathbb{P}^1$ will denote the projective line over $\operatorname{Spec}\mathbb{Z}$, which you may then base-change to any base scheme. Here I'm going to work with a complete ultrametric field K. [FvdP] work with the set of K-valued points $\mathbb{P}^1(K)$, but this is not really the right way to look at this; I will instead define the set $\mathbb{P}$ on which I work to be the set of *closed points* of the scheme $\mathbb{P}^1_K$. However, when it's convenient to do so, I will represent a closed point by a point of $\mathbb{P}^1(K^{\text{alg}})$ contained in it; I hope this won't cause too much confusion.

Also, let Γ be the divisible closure in $\mathbb{R}_{>0}$ of the group $|K^*|$; that is, $\Gamma = |(K^{\text{alg}})^*|$.

A *closed disc* in $\mathbb{P}$ is a subset of one of the forms

$$\{x \in K^{\text{alg}} : |x - a| \leq r\}$$

or

$$\{x \in K^{\text{alg}} : |x - a| \geq r\} \cup \{\infty\}$$

for some $a \in K$ and $r \in \Gamma$. An *open disc* is the same thing but with strict inequalities.

A *connected affinoid subset* of $\mathbb{P}$ is the complement of a nonempty finite union of open discs. An *affinoid subset* of $\mathbb{P}$ is a finite union of connected affinoid subsets.

We want to say that the affinoid subsets are "affinoid spaces" in some natural sense, but we can't yet say that because we haven't abstractly defined an affinoid space. The best we can do for now is "naturally" identify them with the maximal spectra of certain affinoid algebras.

Example: the closed disc $|x| \leq 1$ is just $\operatorname{Max} K\langle x\rangle$. Keep this example firmly in mind as we go along!

Holomorphic functions

Let F be an affinoid subset of $\mathbb{P}$, and let $R(F) \subset K(x)$ be the set of rational functions whose poles do not lie in F. Let $\|\cdot\|_F$ denote the supremum norm on $R(F)$:

$$\|f\|_F = \sup_{a \in F}\{|f(a)|\}.$$

Let A_F be the completion of $R(F)$ with respect to $\|\cdot\|_F$.

Theorem 8.1.

(a) The ring A_F is a reduced affinoid algebra.

(b) The natural map $F \to \operatorname{Max} A_F$ is a bijection.

(c) Under the map in (b), the supremum norm corresponds to the spectral norm on A_F.

Proof. We can decompose F uniquely as a disjoint union of connected affinoid subsets (see exercises), so it suffices to do the case where F is connected. Also, by performing a fractional linear transformation, I can reduce to the case where F is contained in the closed unit disc (so in particular $\infty \notin F$).

Suppose F is defined by the equations

$$|x - a_i| \leq r_i \quad (i = 1, \dots, m), \qquad |x - b_j| \geq s_j \quad (j = 1, \dots, n).$$

Choose positive integers $e_1, \dots, e_m, f_1, \dots, f_m$ and elements $\rho_1, \dots, \rho_m, \sigma_1, \dots, \sigma_n \in K$ with $|\rho_i| = r_i^{e_i}$ and $|\sigma_j| = s_j^{f_j}$. Let I be the ideal of $K\langle x, y_1, \dots, y_m, z_1, \dots, z_n\rangle$ generated by

$$\rho_i y_i - (x - a_i)^{e_i} \quad (i = 1, \dots, m), \qquad \sigma_j - (x - b_j)^{f_i} z_j \quad (j = 1, \dots, n).$$

Then one checks (using the hypothesis on the placement of F) that A_F is the maximal reduced quotient of

$$K\langle x, y_1, \dots, y_m, z_1, \dots, z_n\rangle / I$$

and that the other properties hold. Actually the way to do this is in reverse: first observe (by the Nullstellensatz) that $\operatorname{Max} A_F = F$, so the spectral norm coincides with the supremum norm, then use the theorem from last time that A_F is complete for the supremum norm. □

Exercises

1. (from [FvdP, Exercise 2.1.1]) Prove that every affinoid subset can be written uniquely as a disjoint union of connected affinoid subsets.

2. Suppose that in the proof of Theorem 8.1, the e_i and f_j are chosen to be as small as possible. Prove that the quotient
$$K\langle x, y_1, \dots, y_m, z_1, \dots, z_n\rangle / I$$
is already reduced, so it is actually equal to A_F.

3. Suppose that K is spherically complete. Give an explicit example of an affinoid algebra A which is an integral domain, but whose spectral seminorm is not fully multiplicative. (Hint: consider power series in x and x^{-1} which converge on a suitable annulus; your motivation should functions on a complex annulus which have their suprema on opposite boundary components. We'll look more at this geometric situation in the next few lectures.)

Chapter 9

Subsets of the projective line (and G-topologies), part 2

Reference: [FvdP, Chapter 2]. Note: I'm going to go through the examples a bit quickly, because I think it's easy enough for you to fill in the details; or see [FvdP]. See also [BGR, 9.1] for G-topologies.

Affinoid subsets of $\mathbb{P}^1$ revisited

Let $F \subset \mathbb{P}^1$ be an affinoid subset of $\mathbb{P}^1$. Last time, we showed that the ring A_F, the completion for the supremum norm of the ring of rational functions with poles outside F, is an affinoid algebra whose maximal ideals are precisely the points of F.

It will sometimes be useful to cover a connected affinoid subset with "standard" pieces. Namely, let F be a connected affinoid subset of $\mathbb{P}$ containing ∞, and write F as the complement of the union of the open discs

$$D_i = \{a \in \mathbb{P} : |a - a_i| < r_i\} \qquad (a_i \in K, r_i \in \Gamma).$$

Put $F_i = \mathbb{P} \setminus D_i$. Put $A = A_F$ and $A_i = A_{F_i}$. If F doesn't contain ∞, you can do likewise but with one "everted" open disc that contains ∞.

Proposition 9.1. *Let F be a connected affinoid subset of $\mathbb{P}$, and pick some $a \in \mathbb{P}^1(K) \setminus F$. Then any element of A_F can be written uniquely as a rational function, with all zeroes inside F and all poles at a, times a unit of A_F.*

Proof. First note that if F is the closed unit disc and $a = \infty$, then we already know (by Weierstrass preparation) that every element of $A_F = K\langle x \rangle$ is equal to a polynomial in x times a unit. Moreover, we can factor that polynomial into a part whose roots have norm ≤ 1 and a part whose roots have norm > 1, and the latter is invertible in A_F. This gives existence of the desired factorization; uniqueness follows because every point in the closed unit disc really does give rise to a maximal ideal of A_F.

Now consider the general statement. We first check that each $f \in A_F$ has only finitely many zeroes. It suffices to check this on each F_i, and also it doesn't hurt to replace K by a finite extension. But note that after tensoring with a finite extension of K (whose image under $|\cdot|$ contains the radius of the disc D_i), A_{F_i} becomes isomorphic to a Tate algebra, which we already looked at above.

There exists a unique rational function g, with all zeroes inside F and all poles at a, whose zeroes are precisely those of f with the same multiplicities, and this rational function visibly belongs to f. Moreover, f is divisible by g in the localization of A_F at each maximal ideal, so $f/g \in A_F$; likewise, $g/f \in A_F$. This gives existence of the desired factorization; uniqueness again follows because each point of F really corresponds to a maximal ideal of A_F. □

Corollary 9.2. *If F is connected, then A_F is a principal ideal domain.*

Another interesting class of examples are the annuli

$$F = \{x \in \mathbb{P} : r_1 \leq |x| \leq r_2\},$$

on which analytic functions are given by Laurent series $\sum_{n=-\infty}^{\infty} c_n x^n$, with $c_n \in K$, which converge for $r_1 \leq |x| \leq r_2$. (Note that if one of r_1 or r_2 is in $\Gamma = |(K^{\text{alg}})^*|$ but not in $|K^*|$, you have to check this radius of convergence using points in $\mathbb{P}$, not just K-rational points.) If $r_1 = r_2 = 1$, you get what [FvdP] calls a *ring domain*; see [FvdP,Example 2.2.5].

G-topologies

One has a natural topology on $\mathbb{P}$ induced by the metric topology on K, but this topology is much too fine to be of any use in doing analytic geometry. In fact the same is true of $\operatorname{Max} A$ of any affinoid space. To view these objects as locally ringed spaces in a sensible fashion, we need a better topology; unfortunately, this will have to be a Grothendieck topology, but one of a particularly simple form.

Let X be a set. A *G-topology* on X consists of the following data:

(i) a family of subsets of X containing $\emptyset$ and F, and closed under finite intersections (the *admissible subsets*);

(ii) for each admissible subset, a set of (set-theoretic) coverings of U by admissible subsets (the *admissible coverings*);

subject to the following conditions.

(a) The covering $\{U\}$ of an admissible subset by itself is always admissible.

(b) If U, V are admissible subsets with $V \subset U$ and $\{U_i\}_{i \in I}$ is an admissible covering of U, then $\{U_i \cap V\}_{i \in I}$ is an admissible covering of V.

(c) If U is an admissible subset, $\{U_i\}_{i \in I}$ is an admissible covering, and we are given an admissible covering of each U_i, then the union of these coverings is an admissible covering of U.

We also call admissible subsets *admissible open subsets*, or even *admissible opens*. If we give the topology a name T, we will speak of T-admissible opens and coverings, or even just T-opens and T-coverings.

If you've seen Grothendieck topologies before, you should be on familiar territory. If not, keep in mind that the point of this definition is to isolate, out of the usual concept of a topology, the bare minimum needed to work with sheaves. Namely, a *presheaf* on a G-topology is a contravariant functor $\mathcal{F}$ from the category of admissible subsets (with morphisms being inclusions) to sets (or whatever other objects you have in mind). That is, for each inclusion $U \subseteq V$ of admissible opens, you get a restriction map $\operatorname{Res}_{V,U} : \mathcal{F}(V) \to \mathcal{F}(U)$, and these compose as you expect. (so in particular $\operatorname{Res}_{U,U}$ is the identity). The presheaf is a *sheaf* if the sheaf axiom is satisfied: whenever $\{U_i\}_{i \in I}$ is an admissible covering of an admissible open U, specifying an $f_i \in \mathcal{F}(U_i)$ such that $\operatorname{Res}_{U_i, U_i \cap U_j}(f_i) = \operatorname{Res}_{U_j, U_i \cap U_j}(f_j)$ for all i, j uniquely specifies an $f \in \mathcal{F}(U)$ such that $\operatorname{Res}_{U,U_i}(f) = f_i$. Basically everything you know about sheaves carries over to this context: there is a "sheafification" functor, one can make Čech complexes, and so on.

Since all we care about is the sheaf theory on a topology, we want to consider two topologies with the same sheaf theory to be "equivalent". To wit, we say a G-topology T' is *finer* than another G-topology T on the same set if every T-open is T'-open and every T-covering is a T'-covering. We say T' is *slightly finer* than T if T' is finer than T, and also:

(a) every T'-open has a T'-covering by T-opens;

(b) every T'-covering of a T-open can be refined to a T-covering.

The categories of (pre)sheaves on T and T' are the same, and the computation of Čech cohomology is the same; this is all easy but boring to show, so see [BGR,Chapter 9] for details. (In more precise abstract nonsense terms, these two topologies determine the same "topos".)

A map between sets equipped with G-topologies is *continuous* if every admissible open pulls back to an admissible open, and every admissible covering pulls back to an admissible covering.

Next time, we'll construct some G-topologies on $\mathbb{P}$ and play with them a bit.

Exercises

1. Let $I \subset [0, \infty)$ be an interval whose left endpoint is either 0 or lies in Γ, and whose right endpoint lies in Γ, and put
$$F = \{x \in \mathbb{P} : |x| \in I\}$$
(so that F is either a closed disc or an annulus). Let M be an invertible $n \times n$ matrix over A_F. Then there exists an invertible $n \times n$ matrix U over $K[t]$ in case $0 \in I$, or $K[t, t^{-1}]$ in case $0 \notin I$, such that $|MU - I_n|_{A_F,\mathrm{spec}} < 1$. (Hint: perform "approximate Gaussian elimination". If you get stuck, see Section 3.2 of my preprint "Semistable reduction II", on my web site.)

2. (Mittag-Leffler decompositions; [FvdP, Proposition 2.2.6]) Let F be a connected affinoid subset of $\mathbb{P}$ containing ∞, and write F as the complement of the union of the open discs
$$D_i = \{a \in \mathbb{P} : |a - a_i| < r_i\} \qquad (a_i \in K, r_i \in \Gamma).$$
Put $F_i = \mathbb{P} \setminus D_i$. Put $A = A_F$ and $A_i = A_{F_i}$, and put
$$A_+ = \{f \in A : f(\infty) = 0\}, \qquad A_{i,+} = \{f \in A_i : f(\infty) = 0\}.$$
Prove that $A_+ = \oplus_i A_{i,+}$, and that for any elements $f_i \in A_{i,+}$, one has
$$\|f_i\|_F = \|f_i\|_{F_i}, \qquad \|\sum_i f_i\|_F = \max_i\{\|f_i\|_F\}.$$

3. Given a G-topology T, prove there is a unique finest G-topology T' on the same set among those which are slightly finer than T. (Hint: see [BGR, 9.1].)

Chapter 10

More on G-topologies, part 1 (of 2)

Leftover from last time: separating discs

Here's a more precise answer to Andre's question about separating discs (which you need in order to do the reduction of the theorem identifying A_F to the connected case). I'll leave it to you to work out the (easy) reduction of the general case to this specific case.

Proposition 10.1. *Given $r_1 < r_2$ and $c > 0$, there exists a rational function $f \in K(x)$ such that*

$$\sup\{|f(x) - 1| : |x| \leq r_1\} \leq c, \qquad \sup\{|f(x)| : |x| \geq r_2\} \leq c.$$

Proof. For simplicity, I'm going to consider the special case where there exists $r \in |K^*|$ with $r_1 < r < r_2$, and leave the general case as an exercise. Choose $a \in K^*$ with $|a| = r$, and put $g(x) = a/(a - x)$. Then for $|x| \leq r_1$,

$$|g(x) - 1| = |x/(a - x)| = |x|/r_1 < 1$$

while for $|x| \geq r_2$,

$$|g(x)| = |a/(a - x)| = r/|x| < 1.$$

Now take N large enough that $(r/r_2)^N \leq c$. Then g^N has the desired bound on the outer disc; in fact, so does any polynomial in g^N with integer coefficients. On the other hand, we also have that $|g^N - 1| \leq r_1/r < 1$ on the inner disc; so we can take

$$f = (1 - g^N)^M - 1$$

for M so large that $(r_1/r)^M \leq c$. $\square$

G-topologies and Grothendieck topologies

The notion of a G-topology is a special case of the concept of a Grothendieck topology. Given a category $\mathcal{C}$ admitting finite products, a *Grothendieck topology* on $\mathcal{C}$ consists of, for each $X \in \mathrm{Obj}(\mathcal{C})$, a family $\mathrm{Cov}(X)$ of "coverings" of X, where a covering is a set of arrows $\{U_i \to X\}_{i \in I}$ in $\mathcal{C}$. (Note that I am wantonly ignoring foundational set-theoretic issues; for instance, $\mathrm{Cov}(X)$ is typically a proper class, not a set.) The coverings must then satisfy the following properties.

(a) For each $X \in \mathrm{Obj}(\mathcal{C})$, $\{X \to X\} \in \mathrm{Cov}(X)$ (trivial coverings always exist).

(b) If $Y \to X$ is an arrow in $\mathcal{C}$ and $\{U_i \to X\}_{i \in I} \in \mathrm{Cov}(X)$, then $\{U_i \times_X Y \to Y\}_{i \in I} \in \mathrm{Cov}(Y)$ (coverings pull back).

(c) If $\{U_i \to X\}_{i \in I} \in \mathrm{Cov}(X)$, and for each $i \in I$, $\{V_{ij} \to U_i\}_{j \in J_i} \in \mathrm{Cov}(U_i)$, then $\{V_{ij} \to U_i \to X\}_{i \in I, j \in J_i} \in \mathrm{Cov}(X)$ (coverings can be composed).

A G-topology is then just a Grothendieck topology on a category of subsets of a set in which morphisms are inclusions, the empty set and the whole set both appear, and in which finite products (i.e., finite intersections) exist.

Consequence: whatever you know about Grothendieck topologies will be true here, and is not really any easier to prove here than in general. For instance, the category of sheaves of abelian groups has enough injectives.

Sheaf cohomology and Čech cohomology

Say I have a presheaf $\mathcal{F}$ on a space X equipped with a G-topology, and an admissible covering $\{U_i\}_{i \in I}$ of X. Then one can make a Čech complex in the usual fashion, as follows. Let C^n be the product of $\mathcal{F}(U_{i_0} \cap \cdots U_{i_n})$ for all $(n+1)$-tuples $(i_0, \dots, i_n) \in I^{n+1}$. Then one can make a map $d^n : C^n \to C^{n+1}$ such that given an element $\xi = (\xi(i_0, \dots, i_n))$, $d(\xi)$ has its $(i_0, \dots, i_{n+1})$ component equal to

$$\sum_{j=0}^{n+1} (-1)^j \xi(i_0, \dots, \widehat{i_j}, \dots, i_{n+1}),$$

where the hat means you omit that term.

As usual, the maps d^n satisfy $d^{n+1} \circ d^n = 0$, so you get a complex $0 \to C_0 \to C_1 \to \cdots$. We write $\check{H}^n(\mathcal{F}, \{U_i\})$ for the cohomology of this complex, and call it the *Čech cohomology for the sheaf $\mathcal{F}$ and the covering $\{U_i\}$*. Note that the only 1-element covering is $\{X\}$, so $C_0 = \mathcal{F}(X)$, and you have a natural map $\mathcal{F}(X) \to \check{H}^0(\mathcal{F}, \{U_i\})$; the presheaf $\mathcal{F}$ is a sheaf if and only if this map is a bijection for any covering $\{U_i\}$.

Again as usual, if $\{V_j\}$ is a refinement of $\{U_i\}$ (in the sense that each V_j is contained in some U_i), you get natural maps $\check{H}^n(\mathcal{F}, \{U_i\}) \to \check{H}^n(\mathcal{F}, \{V_j\})$. The direct limit of these is called $\check{H}^n(\mathcal{F})$, or $\check{H}^n(X, \mathcal{F})$ if you want to remind yourself which space you are working on. Leray's theorem applies in this context: if $\mathcal{F}$ is acyclic (for the direct limit Čech cohomology) on each U_i, then the direct limit map $\check{H}^n(\mathcal{F}, \{U_i\}) \to \check{H}^n(X, \mathcal{F})$ is an isomorphism.

We will mostly talk about Čech cohomology here, because that's what we can write down. But there is an issue here about whether Čech cohomology coincides with sheaf cohomology (defined in terms of injective resolutions). [BGR] simply ignores this issue entirely and pretends that sheaf cohomology does not exist. (How barbaric.) [FvdP] comments that "for the G-topologies considered in this book, one can show that for all abelian sheaves the H^n coincide with the $\check{H}^n$", referring to van der Put's paper "Cohomology of affinoid spaces". I don't plan to spend any time on this point now, but I may touch on it when we talk about the cohomology of coherent sheaves later.

One thing you have to beware of when working with G-topologies is that you can't lean on the crutch of defining things in terms of stalks. This is more clear for Grothendieck topologies, when there may not be any "points" to speak of at all; but even for G-topologies, and in particular for the ones in rigid geometry, there are not enough points readily available to distinguish sheaves. That is, you can write down a nonzero abelian sheaf whose stalks are all zero. For instance, this shows up when you try to define exactness. A *short exact sequence* of abelian sheaves is a sequence

$$0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to 0$$

such that

(a) For any admissible U, the sequence

$$0 \to \mathcal{F}_1(U) \to \mathcal{F}_2(U) \to \mathcal{F}_3(U)$$

is exact.

(b) For any admissible U and any $s \in \mathcal{F}_3(U)$, there is an admissible covering $\{U_i\}$ of U such that the restriction of s to each U_i is the image of some element of $\mathcal{F}_2(U_i)$.

Another way to say it is that $\mathcal{F}_1$ is the kernel of $\mathcal{F}_2 \to \mathcal{F}_3$, and $\mathcal{F}_3$ is the cokernel of $\mathcal{F}_1 \to \mathcal{F}_2$. The asymmetry arises for the usual reason: given a map between sheaves, its kernel in the category of presheaves is again a sheaf, but its cokernel in the category of presheaves need not be a sheaf and so must be sheafified to get the cokernel in the category of sheaves.

More about this when we come back to considering sheaf cohomology more systematically.

A G-topology on $\mathbb{P}$

The *weak G-topology* on $\mathbb{P}$ is the G-topology in which the admissible sets are $\emptyset$, $\mathbb{P}$, and the affinoid subsets, and a covering (of an admissible sets by admissible sets) is admissible if it contains a finite subcovering.

The principal virtue of this topology is that the presheaf $\mathcal{O}$ defined by

$$\mathcal{O}(F) = A_F$$

turns out to be a sheaf.

Lemma 10.2. *Let $\mathcal{F}$ be a presheaf on $\mathbb{P}$ such that*

(a) If the affinoid U is a disjoint union of connected affinoids $U_1, \dots, U_n$, then $\mathcal{F}(U) \to \oplus_i \mathcal{F}(U_i)$ is a bijection.

(b) If U_1, U_2 are connected affinoids and $U_1 \cap U_2$ is nonempty, then

$$0 \to \mathcal{F}(U_1 \cup U_2) \to \mathcal{F}(U_1) \oplus \mathcal{F}(U_2) \xrightarrow{d^0} \mathcal{F}(U_1 \cap U_2)$$

is exact.

Then $\mathcal{F}$ is a sheaf. Moreover, if d^0 is always surjective in (b), then $\check{H}^m(\mathcal{F}, \{U_i\}) = 0$ for any $m > 0$ and any admissible covering $\{U_i\}$.

Proof. This is completely formal, so I'll leave it to the reader. Or see [FvdP, Proposition 2.4.6] (although they don't do all the details either). □

Proposition 10.3. *The presheaf $\mathcal{O}$ on $\mathbb{P}$ is a sheaf, and its higher Čech cohomologies vanish.*

Proof. Condition (a) from Lemma 10.2 follows from the structure theorem we proved for A_F. (Or rather, it follows from the reduction to the connected case, for which we need the calculation at the top of the handout.) To check condition (b) and the surjectivity of d^0, we may assume that $\infty \in U_1 \cap U_2$ (by applying a fractional linear transformation); then we can replace each sheaf with the subsheaf of functions vanishing at ∞ without affecting the exactness. Then the exactness of the new sequence follows from the Mittag-Leffler decomposition (i.e., from the partial fractions decomposition of a rational function); see exercises from previous handout. □

This is a special case of a much more general theorem of Tate, which we will return to soon.

Exercises

1. Extend Proposition 10.1 to the case where there is no element of $|K^*|$ between r_1 and r_2. (Hint: replace $x - a$ with a separable irreducible polynomial whose roots fall between the two discs, and find a good starting g using Lagrange interpolation.)

2. Write down a sheaf on $\mathbb{P}$, equipped with the weak G-topology, whose stalks are all zero.

3. Show that the presheaf $\mathcal{O}^*$ on $\mathbb{P}$ given by $\mathcal{O}^*(F) = A_F^*$ is a sheaf (see [FvdP, Definition 2.5.9], but they don't check any details either).

Chapter 11

More on G-topologies, part 2

Corrections from last time

Correction 1: I flubbed the answer to Rebecca's question about bases of an ordinary topology and a G-topology. Before fixing that, I should maybe say some things more precisely.

If T is a G-topology on a set X, the finest G-topology T' on X which is slightly finer than T (which I'll also call the *strong refinement* of T, or the *strong topology* slightly finer than T) is given as follows [BGR, Lemma 9.1.2/3]. (This was left as an exercise in earlier notes, but I said it in class; however, I think I forgot the last restriction in (a).)

(a) A T'-open is any set which can be covered set-theoretically (i.e., is the union, *not* just a subset of the union!) by T-opens, in such a way that the restriction to any T-open can be refined to a T-covering.

(b) A T'-covering of a T'-open U is any covering by T'-opens which, when restricted to a T-open $V \subset U$, becomes a covering which is refined by some T-covering. (In particular, coverings by T-opens are T'-admissible if and only if they satisfy the condition in (a).)

In particular, any T'-open admits an admissible cover by T-opens.

Various additional things you might expect to be true about T are only true about T'. (See [BGR, 9.1.4] for verifications.)

- If $\{U_i\}$ is a T'-covering of X, then $U \subseteq X$ is T'-admissible if and only if $U \cap U_i$ is T'-admissible for each i.
- Any set-theoretic covering which can be refined to a T'-covering is also a T'-covering.
- If $\{U_i\}$ is a T'-covering of X, then a set-theoretic covering of X is a T'-covering if and only if its restriction to each U_i is a T'-covering.
- Adding some T'-opens to a T'-covering yields another T'-covering (i.e., T' is "saturated"). That's because the new cover is refined by the old cover!

If you start with an ordinary topology and make a corresponding G-topology T in which admissibles are the opens in the usual topology and coverings are open coverings, then T'-opens are already T-open because the union of opens in an ordinary topology is open. And any T'-covering is already a set-theoretic covering by opens, so already occurs in T.

So yes, you *can* recover the same G-topology by starting with a basis B of the ordinary topology which is closed under intersections, taking only basis sets to be admissible, and taking set-theoretic coverings among them.

Correction 2: Abhinav points out that my proof that a nonzero function on a connected affinoid subset has only finitely many zeroes is bogus. (I tried to reduce to the disc case by writing a connected affinoid as an intersection of discs, but the function you chose need not extend to any of those larger discs.) The proof of this is actually a real headache given what we know right now; see [FvdP, Theorem 2.2.9]. It will be a bit easier once we talk about analytic subsets; see below.

Bigger correction: the G-topology on $\mathbb{P}$

There's a more serious problem coming up: the G-topology I described on $\mathbb{P}$ is not quite the right one, because I was sloppy in defining affinoid subsets. (In fact, it was probably a bad pedagogical idea to even call them "affinoid subsets". I blame [FvdP] for leading me astray.) Here's a better definition to use in general.

A *rational subset* of $\mathbb{P}$ is one defined by the inequalities $|f_i(x)| \leq 1$ for some rational functions $f_1, \dots, f_n \in K(x)$. This allows some sets I excluded before; for instance, you can take a nonrational point and its conjugates, and take the union of the discs of some radius about those points. The right "weak G-topology" on $\mathbb{P}$ is the one in which all rational subsets are admissible, and admissible covers are those containing a finite subcover.

By the way, note that this means that rational subsets are by fiat quasicompact in the weak G-topology.

Another G-topology on $\mathbb{P}$

Remember that I mentioned (see previous exercises) that there is always a finest topology slightly finer than any given G-topology. I now want to make that topology explicit on $\mathbb{P}$.

Let T be the finest topology slightly finer than the G-topology introduced last time. Then the T-opens are precisely the opens in the usual topology on $\mathbb{P}$, since every open neighborhood of a point contains a rational subset. (Here's where you need to fix the definition that I gave last time.)

A cover $\{U_i\}$ of some U is T-admissible if for any rational $F \subset U$, there exists a finite subset J of I and rational subsets $F_j \subset U_j$ for $j \in J$ such that $F \subset \cup_{j \in J} F_j$.

We now have the notion of an "analytic function" on any open subset F of $\mathbb{P}$, namely a section of $\mathcal{O}$. More explicitly, an analytic function on F can be viewed as a limit of a sequence of rational functions which is uniform (i.e., converges under the supremum norm) on each rational subset of F. (Compare the construction of analytic functions on a complex domain as limits of polynomials which are uniform on compacts.) That means we've given F the structure of a "G-ringed space", i.e., a space with a G-topology and a sheaf of rings for the topology. In fact, F is a *locally G-ringed space*, that is, the stalks of $\mathcal{O}$ at each $x \in F$ are local rings (because the stalk of $\mathcal{O}$ at x doesn't depend on the particular choice of F).

In general, when starting with an affinoid space, its strong G-topology will be generated by "affinoid subspaces", which can be very complicated to describe. On $\mathbb{P}$, however, it will turn out that the only such spaces will be the rational ones, which explains why we get a reasonable description above.

Examples

For any $r_1 < r_2$, the cover of $\mathbb{P}$ by the discs

$$\{x \in \mathbb{P} : |x| \leq r_2\} \qquad \{x \in \mathbb{P} : |x| \geq r_1\}$$

is admissible. However, the cover by the disjoint discs

$$\{x \in \mathbb{P} : |x| < 1\}, \qquad \{x \in \mathbb{P} : |x| \geq 1\}$$

is *not* admissible. To check this, pick a closed disc $D_r : |x| \leq r$ with $r > 1$; if our original cover were admissible, then the disc $|x| < 1$ and the annulus $1 \leq |x| \leq r$ form an admissible cover of D_r. But if that were the case, I'd be able to pick out a rational subspace of $|x| < 1$ which together with the annulus covers D_r, but that's clearly imposible since the disc and the annulus form a disjoint cover of D_r and the open disc is not itself rational.

An example of an admissible cover of the open unit disc $D = \{x \in \mathbb{P} : |x| < 1\}$: let $r_1 < r_2 < \cdots$ be a sequence of elements of $r \in (0,1) \cap \Gamma$ (where again Γ is the divisible closure of $|K^*|$) increasing to 1. Then the discs $D_i = \{x \in \mathbb{P} : |x| \leq r_i\}$ form an admissible cover of D. In fact, this cover has the following interesting property: the maps $\mathcal{O}(D_i) \to \mathcal{O}(D_{i+1})$ have dense image for all i (because already $K[x]$ is dense in each $\mathcal{O}(D_i)$). A space admitting an admissible cover by a sequence of affinoids with this density property is called a *quasi-Stein space*; they turn out to have cohomological properties similar to those of affinoid spaces (and those of Stein spaces in complex analysis).

Exercises

The purpose of these exercises is to work out some details for a special class of affinoid subsets of $\mathbb{P}^1$, namely open annuli. Throughout, choose $r_1 < r_2$ and put $X = \{x \in \mathbb{P} : r_1 < |x| < r_2\}$.

1. Prove that $\mathcal{O}(X)$ consists of Laurent series series $\sum_{n \in \mathbb{Z}} c_n x^n$ over K such that
$$\lim_{n \to \pm\infty} |c_n| r^n = 0 \qquad (r_1 < r < r_2).$$
(Hint: consider an admissible cover by closed subannuli, as in my last example.)

2. Prove that a function $\sum c_n x^n \in \mathcal{O}(X)$ is bounded if and only if $|c_n| r_1^n$ remains bounded as $n \to -\infty$ and $|c_n| r_2^n$ remains bounded as $n \to +\infty$.

3. Suppose $f \in \mathcal{O}(X)$ is bounded on X. Prove that f is equal to a polynomial times a unit of $\mathcal{O}(X)$, and deduce that f has finitely many zeroes on X.

4. Prove that $\mathcal{O}(X)$ is not a Noetherian ring, by exhibiting an ideal which is not finitely generated. In particular, the boundedness hypothesis in the previous exercise is absolutely necessary! (Hint: consider functions which vanish on all but finitely many of some infinite but non-accumulating sequence of points.)

5. Suppose that K is spherically complete. Prove that $\mathcal{O}(X)$ is a *Bézout ring*, a ring in which each finitely generated ideal is principal (even though $\mathcal{O}(X)$ is not typically Noetherian). In fact, this property is equivalent to the spherical completeness of K, but you don't have to check this. (Hint: try the case where K is discretely valued first. This result is a theorem of Lazard; see IHES 14 (1962) 47–75, or [FvdP, Theorem 2.7.6].)

Chapter 12

The G-topologies of an affinoid space

The goal of this lecture is to introduce some G-topologies on an affinoid space. Next time we'll prove Tate's theorem, which will establish the existence of the structure sheaf and of coherent sheaves of modules over the structure sheaf.

References: [FvdP, Section 4.1] and [BGR, 9.1.4]. Note that the *canonical topology* on $\operatorname{Max} A$ in [BGR] is not one of the G-topologies we're aiming for: it's just the topology induced by the supremum norm.

Affinoid subspaces

As usual, let K be a complete ultrametric field. Let A be an affinoid algebra over K and write $X = \operatorname{Max} A$. An *affinoid subspace* (or *affinoid subset*) of X is a subset $Y \subseteq X$ for which there exists a morphism $\phi : A \to B$ of affinoid algebras with $\phi(\operatorname{Max} B) \subseteq Y$, with the following universal property: given any morphism $\psi : A \to C$ of affinoid algebras with $\phi(\operatorname{Max} C) \subseteq Y$, there exists a unique morphism $\tau : B \to C$ with $\psi = \tau \circ \phi$. (I'll let you rewrite that in terms of the representability of an appropriate functor.) We will see shortly that ϕ is uniquely determined by Y; we call B the *coordinate ring* of Y.

The affinoid subspaces of X are analogues of the open affine subsets of an affine scheme (which obey an analogous universal property, though maybe you never noticed this before).

From [FvdP, Remarks 4.1.5], we collect the following observations.

Proposition 12.1. *Let Y be an affinoid subspace of X.*

(a) The map $\phi : A \to B$ is unique up to unique isomorphism.

(b) The induced map $\phi : \operatorname{Max} B \to Y$ is a bijection.

(c) For $y \in Y$, let $\mathfrak{m}_y$ and $\mathfrak{m}'_y$ are the maximal ideals of A and B, respectively, corresponding to y. Then the map $A/\mathfrak{m}_y^n \to B/(\mathfrak{m}'_y)^n$ is an isomorphism for each positive integer n.

(d) If Y is an affinoid subspace of X and Z is an affinoid subspace of Y, then Z is an affinoid subspace of X.

(e) If $\psi : A \to C$ is a morphism of affinoid algebras and Y is an affinoid subspace of $\operatorname{Max} A$ with coordinate ring B, $\psi(Y)$ is an affinoid subspace of $\operatorname{Max} B$ with coordinate ring $B \widehat{\otimes}_A C$. (The hat is missing in [FvdP, Remarks 4.1.5(4)].)

Proof. (a) is clear because the definition is via a universal mapping property. For (b) and (c), choose a point $y \in Y \subseteq X$ and an integer $n \geq 1$, and form the affinoid algebra $A/\mathfrak{m}_y^n$. Then the projection $A \to A/\mathfrak{m}_y^n$ factors uniquely through B, from which (b) and (c) follow. (d), (e), (f) are straightforward. $\square$

Corollary 12.2. *If $Y_1, \dots, Y_n$ are affinoid subspaces of* $\operatorname{Max} A$ *with coordinate rings $B_1, \dots, B_n$, then $Y_1 \cap \cdots \cap Y_n$ is an affinoid subspace with coordinate ring $B_1 \widehat{\otimes}_A \cdots \widehat{\otimes}_A B_n$.*

We may now define our first G-topology on $\operatorname{Max} A$. The *somewhat weak G-topology* on $\operatorname{Max} A$ is the G-topology in which admissible opens are affinoid subdomains, and admissible covers are covers containing a finite subcover. This (or rather, the one in which admissible covers are actually finite, but this is slightly finer than that) is the "weak G-topology" of [BGR], but it's a bit of a nuisance to prove anything about it. So following [FvdP], we are going to "sandwich" this topology between two others that yield the same topos.

Rational subspaces

Motivation: when proving the basic properties of schemes, one doesn't work with all affine opens. One restricts to the distinguished opens obtained by inverting elements, because those form a basis of the same topology. That's quite analogous to the way we are going to proceed here.

Let A be an affinoid algebra with $X = \operatorname{Max} A$. We say a subset Y of X is *rational* if there exist $f_0, \dots, f_n \in A$ generating the unit ideal, such that

$$Y = \{x \in X : |f_i(x)| \leq |f_0(x)| \qquad i = 1, \dots, n\}.$$

We now have the following result analogous to the structure theorem we proved on rational (affinoid) subsets of $\mathbb{P}$.

Proposition 12.3. *With notation as above, Y is an affinoid subspace of X with coordinate ring*

$$B = A\langle y_1, \dots, y_n\rangle/(f_1 - f_0 y_1, \dots, f_n - f_0 y_n).$$

Proof. Note that f_0 is a unit in B because $f_0, \dots, f_n$ generate the unit ideal. Also, note that the obvious map $\phi : A \to B$ carries $\operatorname{Max} B$ into Y.

Let's now check the universal property for ϕ. Given $\psi : A \to C$ carrying $\operatorname{Max} C$ into Y, we must have $\psi(f_0) \in C^*$, and the spectral norms of the $\psi(f_i)/\psi(f_0)$ must be bounded by 1. We thus obtain a well-defined and unique morphism $\tau^* : A\langle y_1, \dots, y_n\rangle \to C$ sending A to C via ψ and sending z_i to $\psi(f_i)/\psi(f_0)$ (see [FvdP, Proposition 3.4.7]). This map kills $f_i - f_0 y_i$ for each i, so factors uniquely through a map $\tau : B \to C$. Thus the universal property checks out. $\square$

The restriction that the f_i generate the unit ideal rules out such things as the subset of $\operatorname{Max} K\langle x, y\rangle$ on which $|x| \leq |y|$, for good reason: if you construct

$$K\langle x, y, z\rangle/(zx - y),$$

you get not the subset you want, but a blowup of it at the point $x = y = 0$.

Note that $Y = \emptyset$ if and only if $|f_0(x)| < \max_i\{|f_i(x)|\}$.

It may also be useful to note that $Y = \emptyset$ if and only if for some (or any sufficiently large) integer ℓ. For another equivalent form of this criterion, see the exercises.

I'll write $\mathcal{O}(Y)$ for the coordinate ring of Y. Define the *very weak G-topology* on X to be the one in which the admissible opens are rational subspaces, and the admissible covers are covers that include a finite subcover. Define the *weak G-topology* on X to be the one in which the admissible opens are finite unions of rational subspaces, and the admissible covers are again covers that include a finite subcover. Clearly the weak G-topology is slightly finer than the very weak G-topology.

We will show next time that every every affinoid subspace is a *finite* union of rational subspaces. ([FvdP] references the paper "Die Azyklizität der affinoiden Überdeckungen" by Gerritzen and Grauert; but it's also in [BGR, 8.2.2], which is where I'll take it from.) That means that on one hand, the somewhat weak

G-topology is slightly finer than the very weak G-topology, but on the other hand the weak G-topology is slightly finer than the somewhat weak G-topology. So from the point of view of the sheaf theory, I can prove everything (like Tate's acyclicity theorem) using the very weak G-topology, where it's much easier.

Incidentally, it is possible to write down affinoid subspaces which are not rational; see [FvdP, Exercise 4.1.6] for one example.

Exercises

1. Let A be an affinoid algebra. Prove that the rational subspace of A defined by $f_0, \ldots, f_n$ is empty if and only if for any sufficiently large positive integer ℓ, there exists an expression

$$f_0^\ell = \sum_\alpha c_\alpha f_1^{\alpha_1} \cdots f_n^{\alpha_n}$$

of f_0^ℓ as a homogeneous polynomial of degree ℓ in $f_1, \ldots, f_n$ with coefficients in $\mathfrak{m}_K$. (Hint: see [FvdP, Proposition 4.1.2(4).])

Chapter 13

Tate's acyclicity theorem (amended 25 Oct 04)

The goal of this lecture is to prove Tate's acyclicity theorem (or more properly, the Gerritzen-Grauert-Tate theorem): coherent sheaves on an affinoid space are acyclic for Čech cohomology.

References: [FvdP, Section 4.2] and [BGR, 8.2]; also, [BGR, 7.3.5] for the Gerritzen-Grauert theorem. Also see Tate's original notes (Rigid analytic spaces, *Invent. Math.* **12** (1971), 257–289); more on these below. I haven't read Gerritzen and Grauert's original paper (Die Azyklizität der affinoiden Überdeckungen, in *Global Analysis, Papers in honor of K. Kodaira*, Princeton University Press, 1969, 159–184), if for no other reason than that I don't read German, so I can't comment on it.

Comments I should have made last time

All of the G-topologies on an affinoid space from last time depend only on the reduced quotient of the affinoid algebra. Namely, it's clear that the notion of a rational subspace is insensitive to nilpotents, since it's defined by evaluating functions on $\operatorname{Max} A$. But also the notion of an affinoid subspace is insensitive: if $\phi : A \to B$ is the homomorphism corresponding to an affinoid subspace of $\operatorname{Max} A$, we just take $\phi^{\mathrm{red}} : A^{\mathrm{red}} \to B^{\mathrm{red}}$ to be the homomorphism corresponding to the same subspace of $\operatorname{Max} A^{\mathrm{red}} = \operatorname{Max} A$. That means there's no harm in working only with reduced affinoid algebras for the moment (though it won't make that much difference either way).

Also, if $\phi : A \to B$ is a homomorphism of affinoids, we say ϕ is a *closed immersion* if it is surjective. Also, we say ϕ is a *locally closed immersion* (resp. *open immersion*) if the homomorphism on stalks $\mathcal{O}_{Y,\phi(x)} \to \mathcal{O}_{X,x}$ is always surjective (resp. bijective). Any map defining an affinoid subspace is a locally closed immersion (see previous handout). See [BGR, 7.3.3] for more on immersions, including some of the results that go into the argument in the next section.

A bit about the reduction process

The general idea in this handout is to reduce the checking of acyclicity for one type of coverings to checking for a simpler type. Here is the basic statement you use.

Lemma 13.1. *Let X be a space equipped with a G-topology, let $\mathcal{F}$ be a presheaf on X, let $\{U_i\}_{i\in I}$ be an admissible covering of X, and let $\{V_j\}_{j\in J}$ be an admissible covering of X refining $\{U_i\}$. Suppose that $\mathcal{F}$ is an acyclic sheaf on the restriction of the covering $\{V_j\}$ to each intersection $U_{i_0} \cap \cdots \cap U_{i_n}$. Then $\mathcal{F}$ is an acyclic sheaf for the covering $\{U_i\}$ if and only if it is an acyclic sheaf for the covering $\{V_j\}$.*

Proof. This is really saying that a certain Leray spectral sequence degenerates; to see it more explicitly, see [BGR, Corollary 8.1.4/3]. □

Affinoid subspaces and rational subspaces

The goal of this section is to sketch a proof of the following theorem. A complete proof appears in [BGR, 7.3.5].

Theorem 13.2 (Gerritzen-Grauert). *Let A be an affinoid algebra, and let X be an affinoid subspace of* $\operatorname{Max} A$*. Then A is a finite union of rational subspaces of* $\operatorname{Max} A$*.*

Beware that the converse is not true: not every finite union of rational subspaces of $\operatorname{Max} A$ is an affinoid subspace (as mentioned last time).

Let's start the argument to see where a naïve approach gets stuck. As noted above, I may assume A is reduced. Let $\phi : A \to B$ denote the representing homomorphism for X, and choose a surjection $\psi : A\langle x_1, \dots, x_n\rangle \to B$ that sections ϕ. We want to argue by induction on n, but this is a bit tricky because the image of $A\langle x_1, \dots, x_i\rangle$ in B need not be the coordinate ring of an affinoid subspace. One already sees this in the rational case: if f_0, f_1, f_2 generate the unit ideal but f_0, f_1 do not, then $A\langle x_1, x_2\rangle/(f_1 - f_0x_1, f_2 - f_0x_2)$ is the coordinate ring of an affinoid subspace but $A\langle x_1\rangle/(f_1 - f_0x_1)$ is not.

To get around this, we use a relative form of of Weierstrass preparation. Recall that a series f in $T_n = K\langle x_1, \dots, x_n\rangle$ was said to be *distinguished (in x_n) of degree d* if, when we write $f = \sum c_i x_n^i$ with $c_i \in T_{n-1}$, we have:

- c_d is a unit in T_{n-1};
- $\|c_d\|_{T_{n-1}} = \max_i\{\|c_i\|_{T_{n-1}}\}$;
- $\|c_d\|_{T_{n-1}} > \|c_i\|_{T_{n-1}}$ for $i > d$.

(Remember we said f was *normalized distinguished* if in fact $\|c_d\| = 1$, and we habitually dropped the word "normalized". However, it'll be more convenient here not to normalize.)

For $f \in A\langle x_1, \dots, x_n\rangle$ and $x \in \operatorname{Max} A$, we say f is *distinguished of degree d at x* if the image of f in $(A/\mathfrak{m}_x)\langle x_1, \dots, x_n\rangle$ is distinguished of degree d (resp. $\leq d$). We say f is *distinguished of degree d* if f is distinguished of degree d at each $x \in \operatorname{Max} A$; if instead f is distinguished of some degree $\leq d$ at each $x \in \operatorname{Max} A$, but not necessarily always the *same* degree, we say f is *distinguished of degree $\leq d$*. Note that f is distinguished of degree 0 (or ≤ 0) if and only if it is a unit.

Lemma 13.3. *Suppose $f \in A\langle x_1, \dots, x_n\rangle$ is distinguished of degree $\leq d$. Then the set*

$$U = \{x \in \operatorname{Max} A : f \textit{ is distinguished of degree } d \textit{ at } x\}$$

is a rational subspace of A.

Proof. Write $f = \sum_i c_i x_n^i$ with $c_i \in A\langle x_1, \dots, x_{n-1}\rangle$, and let f_i denote the constant coefficient of c_i. Then it is easy to check (or see [BGR, Lemma 7.3.5/7]) that

$$U = \{x \in \operatorname{Max} A : |f_i(x)| \leq |f_d(x)| \qquad (i = 0, \dots, d-1)\}.$$

Since f is distinguished of some degree at each $x \in \operatorname{Max} A$, the functions $f_0, \dots, f_d$ have no common zero, and hence generate the unit ideal. Thus U is a rational subspace, as desired. □

Lemma 13.4 ("Relative distinction"). *Suppose the coefficients of $f \in A\langle x_1, \dots, x_n\rangle$ have no common zero on* $\operatorname{Max} A$*. Then there is an A-algebra automorphism τ of $A\langle x_1, \dots, x_n\rangle$ such that f^τ is distinguished of degree $\leq d$, for some nonnegative integer d.*

Proof. Exercise; it's pretty similar to the proof we gave when $A = K$. $\square$

Lemma 13.5. *Suppose that $f \in A\langle x_1, \dots, x_n\rangle$ is distinguished of degree d. Then the map*

$$A\langle x_1, \dots, x_{n-1}\rangle \to A\langle x_1, \dots, x_n\rangle/(f)$$

is finite.

Proof. As in the nonrelative case, the point is that $1, x_n, \dots, x_{n-1}^{d-1}$ generate the quotient over $A\langle x_1, \dots, x_{n-1}\rangle$. (Note that if you write f as a series in x_n, its coefficient of x_n^d has no zeroes in $\operatorname{Max} A\langle x_1, \dots, x_{n-1}\rangle$ and so is a unit.) $\square$

Now back to the proof. We have our map $\phi : A \to B$ defining an affinoid subdomain, and we chose a surjection $\psi : A\langle x_1, \dots, x_n\rangle \to B$ that sections ϕ. By Lemma 13.4, we can arrange for the kernel of ψ to contain a series f which is distinguished of degree $\leq d$ for some d.

Let U be the subset of $\operatorname{Max} A$ at which f is distinguished of degree d; then U is a rational subspace by Lemma 13.3. (It could of course be empty!) We now show that $X \cap U$ is a finite union of rational subspaces of U, and hence of $\operatorname{Max} A$. Let A' be the coordinate ring of U; then the map $A'\langle x_1, \dots, x_{n-1}\rangle \to A'\langle x_1, \dots, x_n\rangle/(f)$ is finite by Lemma 13.5, as then is the map

$$A'\langle x_1, \dots, x_{n-1}\rangle \to B\widehat{\otimes}_A A'.$$

However, $\psi : A' \to B\widehat{\otimes}_A A'$ represents an affinoid subspace of $U = \operatorname{Max} A'$, which means that the local rings at its points are isomorphic to the corresponding local rings at the points of $\operatorname{Max} A'$. It follows that the map $A'\langle x_1, \dots, x_{n-1}\rangle \to B\widehat{\otimes}_A A'$, being finite but not inducing any nontrivial extensions of local rings, must actually be surjective. (For more clarification, see [BGR, Proposition 7.3.3/8].) So we may apply the induction hypothesis to conclude that $X \cap U$ is a finite union of rational subspaces $U_1 \cup \dots \cup U_s$.

To finish, we need one more lemma.

Lemma 13.6. *Let X be an affinoid subspace of $\operatorname{Max} A$, and let U be the rational subspace given by*

$$U = \{x \in \operatorname{Max} A : |f_i(x)| \leq |f_0(x)| \qquad (i = 1, \dots, m)\},$$

for some $f_0, \dots, f_m$ generating the unit ideal in A. Suppose that $U \subseteq X$. Then for some $\epsilon > 1$ in the divisible closure of $|K^*|$, the rational subspace

$$U_\epsilon = \{x \in \operatorname{Max} A : |f_i(x)| \leq \epsilon |f_0(x)| \qquad (i = 1, \dots, m)\}$$

has the property that $X \cap U_\epsilon$ is also a rational subspace.

Proof. The proof is an approximation argument adapted from [BGR, Extension Lemma 7.3.4/10]. Again, let $\psi : A\langle x_1, \dots, x_n\rangle \to B$ be a surjection sectioning the map $A \to B$. Since $U \subseteq X$, we can choose $a_1, \dots, a_n \in A$ such that $|a_i - f_0^N x_i|_U < 1$ for some $N \in \mathbb{N}$. By a continuity argument (see exercises, or [BGR, Proposition 7.3.4/8]), we can choose $\epsilon > 1$ such that $|a_i - f_0^N x_i|_{U_\epsilon} < 1$ for all i, and such that f_0 has no zeroes on U_ϵ.

If we then write A_ϵ and B_ϵ for the coordinate rings of U_ϵ and $X \cap U_\epsilon$, respectively, then the map $A_\epsilon \to B_\epsilon$ factors as

$$A_\epsilon \to A_\epsilon\langle y_1, \dots, y_n\rangle/(y_1 - f_0^N x_1, \dots, a_n - f_0^N x_n) \to B_\epsilon.$$

and that the latter map is surjective. Thus we can cut out $X \cap U_\epsilon$ within U_ϵ by imposing the conditions that $|a_i| \leq 1$ and that some other functions actually vanish. $\square$

Using the previous lemma, we can grow each of $U_1, \cdots, U_s$ slightly to some $U_1', \dots, U_s'$ and still know that each $X \cap U_i'$ is a union of rational subspaces. By so doing, we can ensure that the complement $X \setminus (U_1' \cup \cdots \cup U_n')$ is contained in a union of rational subspaces V of $\operatorname{Max} A$ which is in turn contained in the subspace of $\operatorname{Max} A$ on which f is *not* distinguished of degree d, i.e., is distinguished of degree $\leq d-1$. (That is, each U_i' is a "strict neighborhood" of U_i, and this ensures that $X \setminus (U_1' \cup \cdots \cup U_n')$ is contained in a union of rational subspaces of $\operatorname{Max} A$ not meeting U. See next mini-section for more details.) By induction on d, we can cover $X \cap V$ with rational subspaces, completing the argument.

Strict neighborhoods

Here's a clarification of the last argument in the previous section. For U an admissible open in a strong G-topological space X having the property that $X \setminus U$ is also admissible, we say an admissible open V containing U is a *strict neighborhood* of U in X if the covering $\{V, X \setminus U\}$ of X is admissible.

We then have the following lemma; compare this to the examples of admissible covers we discussed for $\mathbb{P}$.

Lemma 13.7. *Let A be an affinoid algebra, and let U be the rational subspace of* $\operatorname{Max} A$ *given by*

$$U = \{x \in \operatorname{Max} A : |f_i(x)| \leq |f_0(x)| \qquad (i = 1, \dots, m)\},$$

for some $f_0, \dots, f_m$ generating the unit ideal in A. For $\epsilon > 1$ in the divisible closure of $|K^*|$, define the rational subspace

$$U_\epsilon = \{x \in \operatorname{Max} A : |f_i(x)| \leq \epsilon |f_0(x)| \qquad (i = 1, \dots, m)\}.$$

Then an admissible open subset V of $\operatorname{Max} A$ *for the strong G-topology is a strict neighborhood of U if and only if it contains some U_ϵ. Moreover, if V is a strict neighborhood, then there is a finite union of rational subspaces of* $\operatorname{Max} A$*, contained in* $\operatorname{Max} A \setminus U$*, which together with some affinoid subspace contained in U form a cover (automatically admissible) of* $\operatorname{Max} A$*.*

Proof. We first check that each U_ϵ is a strict neighborhood of U. Choose δ in the divisible closure of $|K^*|$ with $\epsilon > \delta > 1$; for $j = 1, \dots, m$, put

$$V_{\delta,j} = \{x \in \operatorname{Max} A : \delta |f_0(x)| \leq |f_j(x)|, |f_i(x)| \leq |f_j(x)| \qquad (i \neq j)\}.$$

Then $V_{\delta,j}$ is rational: if $\delta^N = |c|$ for $c \in K$, then $V_{\delta,j}$ is the rational subspace defined by the functions f_j^N, cf_0^N and the f_i^N for $i \neq j$. The union of the $V_{\delta,j}$ consists of those x for which $\max_{i>0}\{|f_i(x)|\} \geq \delta |f_0(x)|$, so this plus U_ϵ covers $\operatorname{Max} A$. This gives an admissible cover refining the cover $\{U_\epsilon, \operatorname{Max} A \setminus U\}$, so U_ϵ is a strict neighborhood.

Conversely, suppose V is a strict neighborhood of U_ϵ. Then there exist finitely many affinoid subspaces of $\operatorname{Max} A$ contained in the complement of U, whose union together with V covers $\operatorname{Max} A$. Let $W_1, \dots, W_r$ be those subspaces; then the function

$$\frac{|f_0(x)|}{\max_{i>0}\{|f_i(x)|\}}$$

takes a maximum value on each W_j, by a suitable application of the maximum modulus principle. If the maximum over all of $W_1, \dots, W_r$ is $\delta < 1$, then $U_\epsilon \subseteq V$ whenever $\epsilon < \delta$. □

Acyclicity of affinoid coverings

By the Gerritzen-Grauert theorem, we know that every finite covering of an affinoid space by affinoid subspaces can be refined to a finite covering by rational subspaces. Tate proved that finite rational coverings are acyclic for Čech cohomology, by which it follows that they are also acyclic for finite affinoid coverings. The latter is known as "Tate's acyclicity theorem", though the name "Gerritzen-Grauert-Tate theorem" would be more apt.

First, however, we need some further simplification of the types of coverings we are using. Let A be an affinoid algebra. Given $f_1, \dots, f_n \in A$ having no common zero (i.e., generating the unit ideal), put

$$U_i = \{x \in \operatorname{Max} A : |f_j(x)| \leq |f_i(x)| \qquad (j \neq i)\}.$$

Then $U_1, \dots, U_n$ form a covering of $\operatorname{Max} A$ by rational subspaces; such a covering is called a *standard rational covering* of $\operatorname{Max} A$.

Lemma 13.8. *Every finite covering of* $\operatorname{Max} A$ *by affinoid subspaces can be refined to a standard rational covering.*

Proof. By the Gerritzen-Grauert theorem, we may assume without loss of generality that we are starting with a finite covering by rational subspaces $V_1, \dots, V_m$. Suppose V_i is defined by the inequalities $|g_{ij}(x)| \leq |g_{i0}(x)|$ for $g_{ij} \in A$ generating the unit ideal, with j running from 1 to some n_i. Now take the f's to be the products of the form

$$g_{1j_1} \cdots g_{nj_n} \qquad (1 \leq j_i \leq n_i)$$

in which some j_i is equal to 0; these have no common zero because g_{i0} has no zero on V_i, and the V_i cover X. Moreover, the standard cover given by these f's refines the given cover (exercise). □

In fact, we can do better. Given $f_1, \dots, f_n \in A$, put

$$U_i^{\leq} = \{x \in \operatorname{Max} A : |f_i(x)| \leq 1\}, \qquad U_i^{\geq} = \{x \in \operatorname{Max} A : |f_i(x)| \geq 1\}.$$

Then the collection of sets of the form $U_1^* \cap \cdots \cap U_n^*$, where each $* \in \{\leq, \geq\}$, form a covering of $\operatorname{Max} A$; such a covering is called a *Laurent covering* of $\operatorname{Max} A$. (You might be surprised that such a covering is permitted! After all, in $\mathbb{P}$ the analogous thing would have been the closed unit disc and its inverse, which only meet along their boundary. But in fact that is indeed an admissible cover, since both spaces are rational! The point is that glueing two discs along their "boundary" is completely reasonable from our point of view, since that boundary is huge–and rational.)

Lemma 13.9. *Every finite covering of* $\operatorname{Max} A$ *by affinoid subspaces can be refined to a Laurent covering.*

Proof. By Lemma 13.8, it suffices to start with a standard rational covering, say the one generated by $f_1, \dots, f_n$. By the maximum modulus principle, there exists $c \in K^*$ such that

$$|c|^{-1} < \inf_{x \in X} \{\max_i \{|f_i(x)|\}\}.$$

(More explicitly, the subspace on which $|f_i(x)|$ is maximized by $i = j$ is a rational subspace, hence affinoid, and $|f_i(x)|$ achieves its maximum there by the maximum modulus principle.) The point then is that

$$\max_i \{|cf_i(x)|\} > 1 \qquad \text{for all } x \in \operatorname{Max} A.$$

Then the Laurent covering generated by $cf_1, \dots, cf_n$ may not refine the original standard covering, but it has the following convenient property: for each U in my new covering, the restriction of my old (standard rational) covering to U is a standard rational covering generated by *units* on U. (If U is the open consisting of those x where $|f_i(x)| \geq 1$ for $i \in S$ and $|f_i(x)| \leq 1$ for $i \notin S$, then the restriction of the old covering to U is the standard rational covering generated by the cf_i for $i \in S$, which are units in U.)

It thus suffices to check that a standard rational covering generated by units $g_1, \dots, g_m$ can be refined to a Laurent covering. But this is easy: just use the functions g_i/g_j for $1 \leq i < j \leq m$ as the generators (exercise). □

Theorem 13.10 (Acyclicity theorem for the structure sheaf). *Let $X = \operatorname{Max} A$ be an affinoid space. Then for any finite covering of X by affinoid subspaces and any finitely generated A-module M, the presheaf $\mathcal{M}$ on X (for the somewhat weak G-topology) associated to M, whose sections on an affinoid space U with coordinate ring B are $M \otimes_A B$, is a sheaf and its higher Čech cohomology spaces vanish.*

Proof. By Leray's theorem and Lemma 13.8, it is enough to check this for Laurent coverings. In fact, it is enough to check a Laurent covering generated by a single element. (If you want to see this reduction written out in more detail, see [BGR, Chapter 8].)

I'll first check for $M = A$, i.e., $\mathcal{M} = \mathcal{O}$. Say the Laurent covering is generated by $f \in A$. By the same argument as we used for $\mathbb{P}$, it suffices to show that

$$0 \to A \to A\langle f\rangle \oplus A\langle f^{-1}\rangle \xrightarrow{d^0} A\langle f, f^{-1}\rangle \to 0 \tag{13.1}$$

is exact, where

$$\begin{aligned} A\langle f\rangle &= A\langle x\rangle/(x-f) \\ A\langle f^{-1}\rangle &= A\langle y\rangle/(yf-1) \\ A\langle f, f^{-1}\rangle &= A\langle x,y\rangle/(x-f, yf-1) \end{aligned}$$

and d_0 is the difference between the two canonical maps. The exactness of this sequence can be verified by chasing through the diagram

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
 & & (x-f)A\langle x\rangle \oplus (yf-1)A\langle y\rangle \arrow[r] \arrow[d] & (x-f)A\langle x,y\rangle/(xy-1) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \arrow[r] \arrow[d, "\mathrm{id}"'] & A\langle x\rangle \oplus A\langle y\rangle \arrow[r] \arrow[d] & A\langle x,y\rangle/(xy-1) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \arrow[r] & A\langle f\rangle \oplus A\langle f^{-1}\rangle \arrow[r, "d_0"] \arrow[d] & A\langle f, f^{-1}\rangle \arrow[r] \arrow[d] & 0 \\
 & & 0 & 0 &
\end{tikzcd}$$

by checking exactness of the first and second rows, and of all of the columns.

In fact, the sequence (13.1) is not only exact, but split: there is a splitting induced by continuity from the map $A[x,y] \to A\langle f\rangle \oplus A\langle f^{-1}\rangle$ that sends $x^i y^j$ to x^{i-j} if $i \geq j$ and to y^{j-i} if $i < j$. (That is, extend by continuity to a map on $A\langle x,y\rangle$ and note that the ideal $(x-f, yf-1)$ is contained in the kernel.) That means the sequence remains exact upon tensoring over A by any A-module M, so we get the desired result for any M. □

Pay careful attention to where we used the fact that M is finitely generated: it's because we only tensored in the last step. If M were not finitely generated, we would also have to have completed the tensor product, and as has been pointed out before, completing tensor products over arbitrary Banach algebras is a somewhat unpredictable operation.

Any sheaf arising from a finite A-module is called a *coherent sheaf*; note that it also gives a sheaf on the strong topology by abstract G-topology properties. (Less abstract, you compute the sections on an arbitrary open by covering it admissibly with affinoid subspaces, or even rational subspaces, and glueing sections on those together.) Tate's theorem implies on one hand that coherent sheaves can be specified on a finite cover by affinoid subspaces (by providing modules and glueing isomorphisms), and on the other hand that they always have trivial higher Čech cohomology. Life is good.

Next time: we are now ready to start glueing affinoids together to form rigid analytic spaces!

Historical note: Tate's notes

Warning: this "history" is mostly secondhand (or thirdhand), so don't rely on this too heavily.

Tate's original notes, in which he proves the acyclicity theorem for coverings of an affinoid space by rational subsets (or rather, by "special affine subsets" in his terminology, but the result is equivalent), are conventionally dated to 1962, when he lectured on the subject of rigid analytic spaces at Harvard. The subject took off like a shot after that, but Tate only distributed his notes privately and steadfastly refused to publish them. At some point, however, Tate's trusted chain of custody broke down, and the notes came into the possession of the editors of *Inventiones*, who ultimately decided (to posterity's benefit) that these should appear in print. (I believe what happened is that someone stole the notes out of the drawer in his office where he kept them, but I don't have a corroboration for this handy. And even if I did, I wouldn't tell who did it!)

While Tate did introduce the concept of an affinoid subspace (or "affine subset" in his terminology), he did not even formulate the question of whether an arbitrary finite covering by affinoid subspaces is acyclic. This was resolved by Gerritzen and Grauert, who form part of the "German school" that developed rigid analytic geometry more fully in the 1960s. (Besides the aforementioned, and also the authors of [BGR], this school most notably includes Kiehl, who proved some key finiteness theorems which we will touch on a bit later.)

Exercises

1. Prove the "relative distinction lemma" (Lemma 13.4). (Hint: see [BGR, Proposition 7.3.5/9].)

2. Let A be an affinoid algebra, let $f_0, \dots, f_n \in A$ generate the unit ideal, and for ϵ in the divisible closure of $|K^*|$, put
$$U_\epsilon == \{x \in \operatorname{Max} A : |f_i(x)| \leq \epsilon |f_0(x)| \qquad (i = 1, \dots, m)\}.$$
Prove that for any $g \in A$, the function $\epsilon \mapsto |g|_{U_\epsilon}$ extends to a continuous function from $\mathbb{R}_{>0}$ to $\mathbb{R}_{\geq 0}$.

3. Complete the verifications of Lemma 13.8 and Lemma 13.9.

4. Prove that any affinoid subset of $\mathbb{P}$ is rational. (Hint: first check that our old concept of "rational subspace of $\mathbb{P}$", restricted to subsets of the closed unit disc, is consistent with our new concept of a rational subspace. Then apply Gerritzen-Grauert to reduce to checking that a finite union of rational subspaces of the closed unit disc is rational.)

Chapter 14

Rigid analytic spaces (at last!)

We are now ready to talk about rigid analytic spaces in earnest. I'll give the definition and then some examples; we may discuss some of these examples in more detail, as interest dictates.

References: [FvdP, Chapter 4] and [BGR, Chapter 9]. Additional references are given throughout the text.

Locally G-ringed spaces and rigid spaces

A *locally G-ringed space* is a set X with a G-topology and a sheaf of rings $\mathcal{O}_X$ whose stalks at each $x \in X$ are local rings. A morphism between two such gadgets $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is a continuous map $f : X \to Y$ (i.e., one pulling admissible opens/coverings back to admissible opens/coverings) and a sheaf-of-rings homomorphism $f^{-1}(\mathcal{O}_Y) \to \mathcal{O}_X$ such that the induced homomorphisms on stalks are local homomorphisms. As usual, the pushforward f_* of a $\mathcal{O}$-module is just the sheaf-theoretic direct image and the pullback is the sheaf-theoretic inverse image tensored up to $\mathcal{O}_X$ using the homomorphism $f^{-1}(\mathcal{O}_Y) \to \mathcal{O}_X$. As pointed out last time, any affinoid space is a locally G-ringed space: the stalk at a point coincides with the local ring of the affinoid algebra at the corresponding maximal ideal.

A *(very weak, weak, somewhat weak, strong) affinoid space* is a locally G-ringed space of the form $\operatorname{Max} A$, for some affinoid algebra A, equipped with the corresponding G-topology. Note that homomorphisms between affinoid algebras give rise to morphisms of the corresponding affinoid spaces for any of the four types of G-topologies.

Lemma 14.1. *The contravariant functor from affinoid algebras to affinoid spaces (viewed as a subcategory of the locally G-ringed spaces) is an equivalence for any of the four types of G-topologies, with quasi-inverse given by the global sections functor.*

Proof. The nontrivial assertion here is that there can be at most one morphism of locally G-ringed spaces with a given action on global sections. The action on points is uniquely determined because the induced homomorphisms on stalks are local; the sheaf map is uniquely determined by the universal property defining affinoid subdomains. For more details, see [BGR, Proposition 9.3.1/1]. □

A *(very weak, weak, somewhat weak, strong) rigid analytic space* (over K) is a locally G-ringed space X for which there exists an admissible covering $\{U_i\}_{i \in I}$ of X with the following properties.

(a) Each U_i is a (very weak, weak, somewhat weak, strong) affinoid space.

(b) A subset U of X is admissible if and only if $U \cap U_i$ is admissible for each $i \in I$.

Note that the concepts become more expansive as you make the topology finer; the term "rigid analytic space" without qualification usually means a strong rigid analytic space.

A *coherent* (resp. *coherent locally free*) sheaf on a rigid analytic space is one which on the elements of some admissible affinoid covering looks like the sheaf associated to a finitely generated (resp. finite free) module over the structure sheaf. Of course, on a general space, a coherent sheaf need not be generated by its global sections.

A *closed analytic subspace* of a rigid space X is a subspace which on the elements of some admissible affinoid covering looks like the zero locus of some ideal. (This is what you might think of as being analogous to a "closed subscheme".)

Example: generic fibres

A nice set of relatively simple examples of rigid spaces come from Raynaud's "generic fibre" construction.

Let $P = \operatorname{Spf} A$ be an affine formal scheme of finite type over $\mathfrak{o}_K$, that is, A is an $\mathfrak{o}_K$-algebra complete for the ideal $\mathfrak{m}_K A$ and topologically finitely generated over $\mathfrak{o}_K$. Then $A_K = A \otimes_{\mathfrak{o}_K} K$ is an affinoid algebra; we call the corresponding affinoid space $\operatorname{Max} A_K$ the *generic fibre* of P. If P is not affine, we can construct the generic fibre by glueing this construction.

The points of the generic fibre correspond to subschemes of P which are integral and finite flat over $\mathfrak{o}_K$. In particular, there is a specialization map $\operatorname{spe} : A_K \to P_k$ (where $P_k = P \otimes_{\mathfrak{o}_K} k$, and k is as always the residue field of K) taking one of these points to its special fibre. Note that the generic fibre consists of the space *plus* the specialization map; the space itself isn't enough to recover P.

If P is projective, then the generic fibre can be identified with the closed points of the usual generic fibre. In this case, one has a form of the "GAGA principle": any coherent sheaf on the analytic generic fibre is algebraic, and the analytic (Čech) and algebraic cohomologies coincide. This is one of Kiehl's theorems, on which more at a later date.

Incidentally, one sometimes wants to form the "generic fibre" of things which are not topologically finitely generated over $\mathfrak{o}_K$, like $\mathfrak{o}_K[\![t]\!]$, whose generic fibre should be the open unit disc over K. One has to be a bit careful: the ring $\mathfrak{o}_K[\![t]\!]$ is not a valuation ring, because there are series $\sum c_n t^n$ with $|c_n| < 1$ for all n but $|c_n| \to 1$ as $n \to \infty$. In fact, I don't know a good general construction to put here; any suggestions?

Example: the Tate curve, for real this time

Let X be a rigid space, and let Γ be a group acting on X. We say the action of Γ is *discontinuous* if X admits an admissible covering $\{U_i\}_{i \in I}$ by affinoids such that for each i, the set of $\gamma \in \Gamma$ such that $X_i^\gamma \cap X_i \neq \emptyset$ is finite. If this set only ever consists of the identity element of Γ, the action is *free*.

Let $\mathbb{G}_{m,K}$ be the rigid analytic multiplicative group over K; if you like, you can think of it as the result of removing 0 and ∞ from the generic fibre of $\mathbb{P}^1_{\mathfrak{o}_K}$ (which is just the space $\mathbb{P}$ we discussed before). For $q \in K$ with $|q| > 1$, the action of $\mathbb{Z}$ on $\mathbb{G}_{m,K}$ in which n acts by multiplication by q^n is free, since we can cover $\mathbb{G}_{m,K}$ admissibly with the affinoids

$$|q|^{n/2} \leq |x| \leq |q|^{(n+1)/2}$$

and no one of these meets its image under the action of a nonzero $n \in \mathbb{Z}$.

For more, see [FvdP, Section 5.1].

Example: Mumford curves

The group $\mathrm{PGL}_2(K)$ acts on $\mathbb{P}$, the generic fibre of $\mathbb{P}^1_{\mathfrak{o}_K}$. For Γ a subgroup of $\mathrm{PGL}_2(K)$, let $\mathcal{L} \subseteq \mathbb{P}$ be the set of limit points of Γ (that is, the set of $a \in \mathbb{P}$ for which there exists a point $b \in \mathbb{P}$ and a sequence $\gamma_1, \gamma_2, \ldots$

of distinct elements of Γ with $b^{\gamma_n} \to a$ as $n \to \infty$). We say Γ is *discontinuous* if $\mathcal{L} \neq \mathbb{P}$, and the topological closure of any orbit is compact.

A *Schottky group* is a finitely generated nontrivial discontinuous group Γ with no nontrivial finite subgroup. For Γ a Schottky group, let Ω be the complement of $\mathcal{L}$ in $\mathbb{P}$. Then we may form the quotient Ω/Γ (since the action is discontinuous in the sense of the previous section), and—a la peanut butter sandwiches![1]—the result is the analytification of a smooth projective curve, called a *Mumford curve*. The group Γ is necessarily free of some finite number g of generators, and g is also the genus of the curve. There is more combinatorial data hidden in this description, on which Γ is acting (including information about the reduction type of the curve), and the whole picture is tied up with the theory of Shimura varieties, and with stable reduction of curves.

Related example: the *Drinfel'd upper half-space* of dimension n over K is the subspace of the analytified projective space $\mathbb{P}^n_K$ minus the union of all K-rational hyperplanes. The group $\mathrm{PGL}_n(K)$ acts on this space, and one gets a lot of interesting spaces by forming quotients by discrete subgroups. This sort of business goes under the heading of "p-adic uniformization".

For more, see [FvdP, Section 5.4].

G-topologies and ordinary topologies

I wanted to mention a result that compares Čech cohomology and sheaf cohomology on some rigid spaces, but first I need to fix the fact that affinoid spaces do not have enough points. This will involve Before giving the statement, I need some "extra points" on an affinoid space, or more generally on any G-topological space. Given a space X equipped with a G-topology, a *G-filter* (or simply "filter") on X is a collection $\mathcal{F}$ of admissible subsets with the following properties.

(a) $X \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$.

(b) If $U_1, U_2 \in \mathcal{F}$, then $U_1 \cap U_2 \in \mathcal{F}$.

(c) If $U_1 \subseteq U_2$ and $U_1 \in \mathcal{F}$, then $U_2 \in \mathcal{F}$.

A *prime filter* is a filter $\mathcal{F}$ also satisfying

(d) if $U \in \mathcal{F}$ and $\{U_i\}_{i \in I}$ is an admissible covering of U, then $U_i \in \mathcal{F}$ for some $i \in I$.

A *maximal filter* (or *ultrafilter*) is a filter $\mathcal{F}$ which is maximal under inclusion; such a filter is clearly also prime. For each $x \in X$, the set of admissibles containing x is a maximal filter.

Let $\mathcal{P}(X)$ and $\mathcal{M}(X)$ denote the sets of prime and maximal filters, respectively, on X, and likewise for any admissible open U of X (that is, $\mathcal{P}(U)$ consists of prime filters of X in which U appears). Equip $\mathcal{P}(X)$ with the ordinary topology generated by the $\mathcal{P}(U)$; then there is a natural morphism of sites $\sigma : X \to \mathcal{P}(X)$, and it turns out that the functors σ_* and σ^* are equivalences between the categories of abelian sheaves on X and on $\mathcal{P}(X)$ [FvdP, Theorem 7.1.2].

The topological space $\mathcal{P}(X)$ is typically pretty unwieldy. For rigid analytic spaces, Berkovich theory gives a way to recover all of the information in $\mathcal{P}(X)$ by working on $\mathcal{M}(X)$, which is still a bigger space than X itself but is small enough to be more wieldy (and is more closely analogous to spaces you see in ordinary analysis). We'll introduce this perspective sometime later in the term.

[1] The "abracadabra" phrase of The Amazing Mumford, the magician Muppet on *Sesame Street*. I have been unable to confirm reports that this Muppet is actually named after Mumford the mathematician.

Čech versus sheaf cohomology

Here's a precise comparison statement between Čech and sheaf cohomology (it may not be optimal, but almost surely some sort of finiteness hypothesis is necessary). See van der Put, Cohomology of affinoid spaces, *Comp. Math.* **45** (1982), 165–198, Proposition 1.4.4.

Proposition 14.2. *Let X be a strong rigid analytic space with an at most countable admissible covering $\{U_i\}_{i\in I}$ such that each U_i is an affinoid space, and each $U_i \cap U_j$ is an affinoid subspace of U_i. Then Čech cohomology computes sheaf cohomology for any abelian sheaf on X.*

Note that I said any abelian sheaf, not necessarily a coherent sheaf.

Sketch of proof. Put $\mathcal{P}(X) = \cup_i \mathcal{P}(U_i)$. Given a sheaf $\mathcal{F}$, pick one injective abelian group $\mathcal{G}_p$ containing $\mathcal{F}_p$ for each $p \in \mathcal{P}(X)$, and define the presheaf $\mathcal{G}_p$ on $\mathcal{P}(X)$ by $\mathcal{G}(U) = \prod_{p\in U} \mathcal{G}_p$. Then $\mathcal{G}$ is a sheaf and its stalk at p is $\mathcal{G}_p$; in particular, $\mathcal{G}$ is injective. Let $\mathcal{H}$ be the presheaf cokernel of the injection $\mathcal{F} \to \mathcal{G}$; then $\mathcal{H}^+$ is a sheaf, and coincides with the sheafification of $\mathcal{H}$.

Define $\mathcal{K}$ as the presheaf cokernel of $\mathcal{H} \to \mathcal{H}^+$, so that $\mathcal{K}^+ = 0$. At this point one must verify that for X as chosen, the fact that $\mathcal{K}^+ = 0$ implies that $\check{H}^i(X, \mathcal{K}) = 0$ for all i. This, despite being more or less formal, is the heart of the matter; see Lemma 1.4.5 of the aforecited paper. (The point of the hypothesis on X is that any cover can be refined to a countable cover by affinoid subspaces.)

Given that $\check{H}^i(X, \mathcal{K}) = 0$ for all i, the Čech cohomologies of $\mathcal{H}$ and $\mathcal{H}^+$ must coincide. We thus have an exact sequence

$$0 \to \check{H}^0(X, \mathcal{F}) \to \check{H}^0(X, \mathcal{G}) \to \check{H}^0(X, \mathcal{H}^+) \to \check{H}^1(X, \mathcal{F}) \to 0$$

and isomorphisms $\check{H}^i(X, \mathcal{H}^+) \to \check{H}^{i+1}(X, \mathcal{F})$. Comparing the first sequence with the analogous sequence in sheaf cohomology, you get $\check{H}^1(X, \mathcal{F}) \cong H^1(X, \mathcal{F})$ for all $\mathcal{F}$. Now proceed by induction and "dimension shifting": given $\check{H}^i(X, \mathcal{F}) \cong H^i(X, \mathcal{F})$ for all $\mathcal{F}$, apply this isomorphism with $\mathcal{F}$ replaced by $\mathcal{H}^+$. $\square$

Exercises

1. (Rigid Hartogs' lemma) Let X be the rigid space obtained from $\operatorname{Max} K\langle x_1, \dots, x_n\rangle$, for some $n \geq 2$, by removing the point $(0, \dots, 0)$. Prove that $\mathcal{O}(X) = K\langle x_1, \dots, x_n\rangle$.

2. Suppose that K is spherically complete. Prove that every coherent locally free sheaf on the open annulus $\{x \in \mathbb{P} : r_1 < |x| < r_2\}$ is free. (The rank 1 case is more or less an exercise I gave earlier.) If you get stuck, see my preprint "Semistable reduction... II" on my web site. Warning: a general coherent sheaf on an open annulus need not be generated by global sections!

Problems

These aren't listed as exercises because I don't know how to do them!

1. (from Richard Taylor) Let X be the rigid space obtained from $\operatorname{Max} K\langle x_1, \dots, x_n\rangle$, for some $n \geq 2$, by removing *all* of the K-rational points. Is $\mathcal{O}(X) = K\langle x_1, \dots, x_n\rangle$? How about if you remove all K-rational planes of codimension at least 2?

2. Let X be an affinoid space, let U be a connected (in the G-topological sense) affinoid subspace, and choose $f \in \mathcal{O}(U)$. Let Y be the union of all connected affinoid subspaces V containing U for which there exists $g \in \mathcal{O}(U)$ with restriction f. (I'm trying to think of Y as the "domain of definition" of f.) Does there exist $g \in \mathcal{O}(Y)$ extending f? If so, is g unique? Also, what else can you say about Y (e.g., can one have $Y = U$?)

3. What is the "right" level of generality for the assertion that Čech cohomology computes sheaf cohomology on a rigid analytic space?

Chapter 15

More on coherent sheaves

References: [FvdP, Chapter 4]. Kiehl's original papers (in German) are: Der Endlichkeitssatz für eigentliche Abbildungen in der nichtarchimedische Funktionentheorie, *Inv. Math.* **2** (1967), 191–214; and Theorem A und B in der nichtarchimedische Funktionentheorie, *Inv. Math.* **2** (1967), 256–273.

Addendum: comment about generic fibres

I forgot to mention in the previous handout that the generic fibre construction loses information in a very precise sense. If P is a formal scheme of finite type over $\mathfrak{o}_K$, and P' is a blowup of $\mathfrak{o}_K$ within the special fibre, then the induced map from the generic fibre of P' to that of P is an isomorphism.

Example: if $P = \mathrm{Spf}\,\mathfrak{o}_K\langle x, y\rangle$, and P' is the blowup along the ideal (x, y, π) for some $\pi \in \mathfrak{m}_K$, then in both cases the generic fibre consists of the closed points of the affine plane over K represented by geometric points (a, b) with $|a| \leq 1$ and $|b| \leq 1$.

In fact, Raynaud proved an equivalence of categories between a certain category of rigid spaces, and a certain category of formal schemes localized at the blowups in the special fibre. (I don't know the precise formulation offhand; will go look it up later.) This cuts both ways: sometimes it's easier to work with rigid spaces than formal schemes because you would just as soon ignore the blowups in the special fibre. But sometimes you can't make a construction in the rigid setting without keeping track of some choice of an associated formal scheme (as in Laurent Fargues's seminar talk a couple of weeks ago).

Coherent sheaves on affinoid spaces

Last time I defined a coherent sheaf on a rigid space to be a sheaf which, locally on some admissible affinoid covering, is isomorphic to the sheaf generated by a finitely generated module over the corresponding affinoid algebra. The Gerritzen-Grauert-Tate theorem implies that you actually get a sheaf this way, but it doesn't imply that this sheaf has "enough" sections. This is fixed by the following theorem of Kiehl, whose proof I'll sketch here; see [FvdP, Section 4.5] for details.

Theorem 15.1 (Kiehl)**.** *Let $\mathcal{F}$ be a coherent sheaf on an affinoid space $X = \mathrm{Max}(A)$, and put $M = \mathcal{F}(X)$. Then M is finitely generated, and $\mathcal{F}$ is isomorphic to the coherent sheaf associated to M.*

Sketch of proof. As in the Gerritzen-Grauert-Tate argument, one reduces to the case of a Laurent covering given by a single f. (See [FvdP, Section 4.5] for more on this reduction.) That puts us in the following situation. We are given $f \in A$, finitely generated modules M_+ and M_- over $A\langle f\rangle$ and $A\langle f^{-1}\rangle$, respectively, and an isomorphism between $M_+ \otimes A\langle f, f^{-1}\rangle$ and $M_- \otimes A\langle f, f^{-1}\rangle$. For $M = \mathcal{F}(X)$, we must show that the maps $M \otimes A\langle f\rangle \to M_+$ and $M \otimes A\langle f^{-1}\rangle \to M_-$ are surjective.

One can easily reduce this to just checking for M_- (as in [FvdP, Lemma 4.5.5], but they seem to get it backwards; see below). For this part, one needs a form of Cartan's lemma: there exists $c > 0$ such that any invertible $n \times n$ matrix U over $A\langle f, f^{-1}\rangle$ with $\|U - I_n\| < c$ (where the matrix norm is the maximum over entries) factors as U_+U_-, with U_+ invertible over $A\langle f\rangle$ and B invertible over $A\langle f^{-1}\rangle$. See [FvdP, Lemma 4.5.3] for this calculation. (The basic idea is to split U *additively* as $I + V_+ + V_-$, where V_+ has only nonnegative powers of V, V_- has only nonnegative powers of V_-, and both V_+ and V_- have small norm. Then you replace U by $(1 - V_+)U(1 - V_-)$ and repeat; if c is small enough, this process converges to the desired factorization.)

To check surjectivity of $M \otimes A\langle f\rangle \to M_+$ now, you choose a set of n generators of M_+ and a set of n generators of M_- (for some n), and write down change-of-basis matrices U and V between the two sets of generators over $A\langle f, f^{-1}\rangle$. Since the image of A in $A\langle f\rangle$ is dense, we can approximate U closely by a matrix over A, so that $\|(U' - U)V\| < c$ with c as above. We can then factor $I_n - (U' - U)V = U_+U_-$ as above, and changing basis from M_+ via U_+ gives us a set of generators which are defined on both subspaces, yielding the surjectivity.

Given the surjectivity, we can choose a finitely generated submodule M_1 of M which generates both M_+ and M_-. Let $\mathcal{M}_1$ be the associated sheaf: then $\mathcal{M}_1 \to \mathcal{F}$ is surjective. Let $\mathcal{G}$ be the kernel of that map; then $\mathcal{G}$ is also coherent and given by finitely generated modules on the two pieces of the cover, so we can find a surjection $\mathcal{M}_2 \to \mathcal{G}$, where $\mathcal{M}_2$ is the coherent sheaf associated to the finitely generated module M_2. Hence $\mathcal{F}$ is the cokernel of the map $\mathcal{M}_2 \to \mathcal{M}_1$; by acyclicity, its global sections are precisely M_2/M_1, and the associated sheaf is precisely $\mathcal{F}$ because they match up on the two opens. □

This means that we can glue coherent sheaves on any admissible cover of X (I may have said this earlier without justification).

Example: the sheaf of differentials

An important example of a coherent sheaf is the sheaf of continuous differentials.

Theorem 15.2. *Let A be an affinoid algebra. There exists a finitely generated A-module $\Omega_{A/K}$ equipped with a K-linear derivation $d : A \to \Omega_{A/K}$ with the following universal property: given any finitely generated A-module M equipped with a K-linear derivation $D : A \to M$, there exists a unique A-module homomorphism $l : \Omega_{A/K} \to M$ with $D = l \circ d$.*

Proof. In case $A = T_n = K\langle x_1, \dots, x_n\rangle$, this is easy to check: take $\Omega_{A/K} = Adx_1 + \cdots + Adx_n$ together with the formal total differential. That is,

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i,$$

where the partial derivation is done formally on the series.

For general A, pick a presentation $A \cong T_n/(f_1, \dots, f_m)$, and put

$$\Omega_{A/K} = \Omega_{T_n/K}/(T_n df_1 + \cdots + T_n df_m),$$

with the induced derivation from T_n. □

The module $\Omega_{A/K}$ is called the *module of finite differentials*, or the *universal finite differential module.* It is obviously unique up to unique isomorphism, so one gets from it a coherent sheaf of differentials Ω_X on any rigid space X.

Note that the module of finite differentials is *not* always the same as the module of Kähler differentials, because the latter is sometimes badly behaved. In fact, it can be shown (though I don't remember a reference offhand) that the module of finite differentials is the maximal separated quotient of the completion of the ordinary module of Kähler differentials; it is thus sometimes called the *module of continuous differentials*, because the latter description makes it universal for continuous derivations into any Banach A-algebra.

Here are some fun facts about modules of differentials. (See [FvdP, Theorem 3.6.3], although they mostly just refer to Springer LNM 38 "Differentialrechnung in der analytische Geometrie"):

Proposition 15.3. *Let A be an affinoid algebra which is an integral domain, and let F be its fraction field.*

(a) The dimension of $\Omega_{A/K} \otimes_A F$ over F equals the Krull dimension d of A.

(b) Let $\mathfrak{m}$ be a maximal ideal of A. Then the following are equivalent.

   (i) A is smooth over K at $\mathfrak{m}$. (That means that locally near $\mathfrak{m}$, A can be written as the vanishing locus in some affine space $\operatorname{Max} T^n$ of some number m of functions whose $m \times n$ matrix of partial derivatives has maximal rank. If K is perfect, this is equivalent to A being regular at $\mathfrak{m}$, i.e., the localization $A_\mathfrak{m}$ being a regular local ring.)

   (ii) The localization $A_\mathfrak{m} \otimes_A \Omega^{A/K}$ is free over $A_\mathfrak{m}$ (of rank d).

   (iii) The dimension of $\Omega_{A/K}/\mathfrak{m}\Omega_{A/K}$ over $A/\mathfrak{m}$ is d.

Closed analytic subspaces

Let X be a rigid space over k, and let $\mathcal{I}$ be a coherent sheaf of ideals of $\mathcal{O}$. Then we get a subspace of X associated to $\mathcal{I}$ whose points are the support of the coherent sheaf $\mathcal{O}/\mathcal{I}$ (with topology induced from X), and whose structure sheaf is the restriction of $\mathcal{O}/\mathcal{I}$. On an affinoid space, this of course corresponds to viewing the vanishing locus of some ideal of functions as the Max of the quotient affinoid algebra. Any such subspace of X is called a *closed analytic subspace*.

Next time: GAGA

Next time, I'll talk about separatedness and properness for rigid spaces and sketch the proof of Kiehl's rigid GAGA theorem.

Exercises

1. Assume that K has characteristic zero. Let X be a *smooth* rigid space, let $\mathcal{F}$ be a coherent sheaf on X, and suppose that there exists a K-linear connection $\nabla : \mathcal{F} \to \mathcal{F} \otimes \Omega_{X/K}$. That is, given a section s of $\mathcal{F}$ and a function f, we have $\nabla(sf) = f\nabla(s) + s \otimes df$. Prove that $\mathcal{F}$ must be locally free. (Hint: replace the local ring at a maximal ideal by its completion and check there that the torsion submodule must vanish.)

Chapter 16

Kiehl's finiteness theorems

References: [FvdP, Chapter 4]. Again, Kiehl's original papers (in German) are: Der Endlichkeitssatz für eigentliche Abbildungen in der nichtarchimedische Funktionentheorie, *Inv. Math.* **2** (1967), 191–214; and Theorem A und B in der nichtarchimedische Funktionentheorie, *Inv. Math.* **2** (1967), 256–273.

Addendum: Generic fibres more generally

Jay pointed out a reference for a more general construction of generic fibres: de Jong, Crystalline Dieudonné theory via formal and rigid geometry (see Chapter 7, where he attributes the construction to Berthelot). Let $P = \mathrm{Spf}(R)$, where R is a quotient of a ring of the form $R\langle x_1, \dots, x_m\rangle[\![y_1, \dots, y_n]\!]$. (Note that the order of the two sets of brackets is important here!) You can then take the generic fibre of P to be the appropriate subspace of the product of the closed m-dimensional unit ball with the open m-dimensional unit ball (and its points will be the formal subschemes of P which are integral and finite flat over $\mathfrak{o}_K$). Another way to say the same thing is that you take, for each $0 < \epsilon < 1$ in the divisible closure of $|K^*|$, the subring of $R\langle x_1, \dots, x_m\rangle[\![y_1, \dots, y_n]\!]$ consisting of series convergent for $|x_1|, \dots, |x_m| \leq 1$ and $|y_1|, \dots, |y_n| \leq \epsilon$, form the quotient, and nest these affinoid spaces to get your rigid space.

Addendum: Analytification of an algebraic variety

Here's the question Andre asked last time, with my commentary on it. Suppose X is a quasi-projective *algebraic* variety over K. Embed X into some projective variety $\overline{X}$; we can then equip the closed points of $\overline{X}$ with a rigid analytic structure by picking a model of $\overline{X}$ over K and taking the rigid analytic generic fibre of the completion along the special fibre. (This is independent of the choice of the model because the rigid analytic generic fibre doesn't see blowups in the special fibre.) The closed points of X form a subspace of $\overline{X}$; the question is, is this subspace independent of the choice of $\overline{X}$?

It suffices to check that if $\overline{X}_1 \to \overline{X_2}$ is a proper morphism between two compactifications, then it induces an isomorphism on X as a rigid analytic space. (To compare two general compactifications, you can then compare each to the fibre product.) I think this is okay, but I didn't check it.

Separated and proper spaces

The category of rigid spaces has fibre products: these are generated by completed tensor products of affinoid algebras. We may thus say that a rigid space X is *separated* if the diagonal $\Delta : X \to X \times_K X$ is a closed immersion (i.e., is defined by a coherent sheaf of ideals). As for schemes, one has the following criterion for separatedness.

Proposition 16.1. *A rigid space X is separated if and only if it admits an admissible affinoid covering $\{U_i\}$ such that for $i \neq j$ with $U_i \cap U_j \neq \emptyset$, the intersection $U_i \cap U_j$ is affinoid and the canonical map $\mathcal{O}(U_i)\widehat{\otimes}_K\mathcal{O}(U_j) \to \mathcal{O}(U_i \cap U_j)$ is surjective.*

In particular, any affinoid space is separated.

If U is an affinoid subset of an affinoid space X, then U is said to be an *interior subspace* of X if it is contained in a subspace of the form $|f_1(x)|, \dots, |f_n(x)| \leq \epsilon$ for some $f_1, \dots, f_n \in \mathcal{O}(X)$ such that $K\langle f_1, \dots, f_n\rangle$ surjects onto $\mathcal{O}(X)$ and some $\epsilon < 1$. (Note that "the interior of X" is not a useful notion here.)

A rigid space X is *proper* if it is separated and there exist two finite admissible affinoid coverings $\{U_i\}_{i=1,\dots,n}$ and $\{U_i'\}_{i=1,\dots,n}$ such that U_i is an interior subspace of U_i' for $i = 1, \dots, n$. For instance, $\mathbb{P}^n_K$ (with homogeneous coordinates $t_0, \dots, t_n$) is proper because you can cover it by the subspaces

$$U_{i,\epsilon} = \{x \in \mathbb{P}^n_K : |t_j(x)| \leq \epsilon |t_i(x)|\} \qquad (i = 0, \dots, n)$$

for any ϵ, and $U_{i,\epsilon}$ is an interior subspace of $U_{i,\epsilon'}$ if $\epsilon < \epsilon'$. Likewise, any closed analytic subspace of a proper space is proper, so the analytic space associated to a projective variety is also projective.

The relative version of this construction is as follows. A morphism $X \to Y$ is proper if after restricting Y to each element of some admissible affinoid cover, I can find two finite admissible affinoid coverings $\{U_i\}$ and $\{U_i'\}$ of X, such that U_i is a relative interior subspace of U_i'. The latter means that U_i belongs to a subspace of U_i' of the form $|f_1(x)|, \dots, |f_n(x)| \leq \epsilon$ for some $f_1, \dots, f_n \in \mathcal{O}(X)$ such that $\mathcal{O}(Y)\langle f_1, \dots, f_n\rangle$ surjects onto $\mathcal{O}(U_i')$ and some $\epsilon < 1$. (This is admittedly a lousy definition; Kiehl rigged it up precisely to make the next theorem work. Is there a "universally closed" version of this definition, or a valuative criterion?)

A morphism $f : X \to Y$ is *finite* if for some admissible affinoid covering $\{U_i\}$ of Y, each $f^{-1}(U_i)$ is affinoid and $\mathcal{O}(U_i) \to \mathcal{O}(f^{-1}(U_i))$ is a finite morphism of affinoid algebras. One can recover f from the pushforward sheaf $f_*\mathcal{O}_X$, which is coherent; in particular, one can show that "some" may be replaced by "any" (see [FvdP, Definition 4.5.7]). Any finite morphism is proper, because $f^{-1}(U_i)$ is a relative interior subspace of itself!

Theorem 16.2 (Kiehl)**.** *(a) Let X be a proper rigid space over K. Then the (Čech) cohomology spaces of any coherent sheaf on X are finite dimensional over K.*

(b) Let $f : X \to Y$ be a proper morphism of rigid spaces, with Y separated, and let $\mathcal{F}$ be a coherent sheaf on X. Then the direct image $f_*\mathcal{F}$ and the higher direct images $R^i f_*\mathcal{F}$ are coherent sheaves on Y. In particular, the image $f(X)$ is a closed analytic subspace of Y.

Warning: the "higher direct images" here are constructed using Čech cohomology; I think you should be able to make a Čech/sheaf comparison in this relative case, but I didn't check it.

Sketch of proof. This sketch is basically the sketch from [FvdP, Theorem 4.10.3]. The idea is to take the two coverings you have and show that, on one hand, you get the same Čech cohomology groups (by acyclicity) from both coverings, and on the other hand, the map $\mathcal{F}(U) \to \mathcal{F}(U')$ is a compact operator (a uniform limit of operators of finite rank) whenever U is an interior subspace of U'. Thus on each cohomology space, the identity map is compact, which can only happen on a finite dimensional vector space. (The relative argument proceeds basically the same way.) □

I don't know a reference for "rigid GAGA", but I seem to think it was written down somewhere by Kiehl. Anyway, [FvdP] says you can just imitate Serre's proof; rather than presume you know how this goes, I'll give a brief version here. (Serre's paper is as good a reference as any, maybe even better, because it's from the dark ages when Serre's "Faisceaux algébriques coherents", which introduced sheaves into abstract algebraic geometry, was hot off the presses. It is: Géométrie algébrique et géométrie analytique, *Ann. Inst. Fourier (Grenoble)* **6** (1955–56), 1–42. His Théorèmes 1,2,3 are our (a),(b),(c).)

Theorem 16.3 (Rigid GAGA)**.** *Let X be a projective algebraic variety over K.*

(a) For any algebraic coherent sheaf $\mathcal{F}$ on X, the natural homomorphisms from algebraic (sheaf) to analytic (Čech) cohomologies are bijections.

(b) The analytification functor from coherent sheaves on X to coherent sheaves on the analytification of X is fully faithful.

(c) Every analytic coherent sheaf comes from an algebraic coherent sheaf (so the functor in (b) is an equivalence of categories).

Sketch of proof. Note that it's enough to prove everything for $X = \mathbb{P}^n$, since we can extend coherent sheaves by zero in both the algebraic and analytic categories to get from sheaves on X to sheaves on $\mathbb{P}^n$.

For (a), you first prove it for $\mathcal{O}$, then for the sheaves $\mathcal{O}(m)$ by induction on the dimension n of the projective space: if H is a hyperplane, you have an exact sequence

$$0 \to \mathcal{O}(-1) \to \mathcal{O} \to \mathcal{O}_H \to 0$$

where the first map is multiplication by a linear form vanishing on H. Then clever use of the five lemma gives you what you need. Since every coherent sheaf on $\mathbb{P}^n$ is the kernel of a map from some $\mathcal{O}(m)$ to some $\mathcal{O}(n)$ (another theorem of Serre!), you end up getting (a) for all coherent sheaves.

For (b), you apply (a) to the sheaf Hom between any two given algebraic sheaves: homomorphisms between the sheaves correspond to elements of H^0 of the sheaf hom.

For (c) (the hard part), you again induct on dimension. The crux of the argument is to show that given $\mathcal{F}$ and a point $x \in \mathbb{P}^n$, you can find an integer m such that the analytic H^0 of the twist $\mathcal{F}(m)$ generates the stalk of $\mathcal{F}(m)$; by a compactness argument, you can choose m uniformly for all x. That is, all stalks of $\mathcal{F}(m)$ are generated by the global sections; proceeding as in the proof of Kiehl's theorem on coherent sheaves on affinoid spaces, we deduce that $\mathcal{F}$ can be written as a cokernel of a map between algebraic coherent sheaves (both of the form $\mathcal{O}(-m)^d$), and by (b) is algebraic.

The crux lemma is a bit intricate; it's Lemme 8 in Serre's paper. You again pick a hyperplane H and write down the exact sequence

$$0 \to \mathcal{O}(H) \cong \mathcal{O}(-1) \to \mathcal{O} \to \mathcal{O}_H \to 0,$$

which on the right you can tensor with $\mathcal{F}$:

$$\mathcal{F}(-1) \to \mathcal{F} \to \mathcal{F}_H \to 0;$$

let $\mathcal{K}$ be the kernel on the left. Both $\mathcal{K}$ and $\mathcal{F}_H$ are supported on H, so are algebraic by the induction hypothesis; in particular, by Serre's theorem, they lose their higher cohomology upon twisting by $\mathcal{O}(m)$ for m sufficiently large. If you split the four-term exact sequence

$$0 \to \mathcal{K}(m) \to \mathcal{F}(m-1) \to \mathcal{F}(m) \to \mathcal{F}_H(m) \to 0$$

by adding $\mathcal{G}$ in the middle, you end up with surjections

$$H^1(\mathbb{P}^{n,\mathrm{an}}, \mathcal{F}(m-1)) \to H^1(\mathbb{P}^{n,\mathrm{an}}, \mathcal{G}), \qquad H^1(\mathbb{P}^{n,\mathrm{an}}, \mathcal{G}) \to H^1(\mathbb{P}^{n,\mathrm{an}}, \mathcal{F}(m));$$

the point is that $\dim H^1(\mathbb{P}^{n,\mathrm{an}}, \mathcal{F}(m))$ is nonincreasing as m grows. It must thus stabilize after some m, at which point one finds that the global sections of $\mathcal{F}(m)$ surject onto those of $\mathcal{O}_H(m)$. The latter generate the stalk of $\mathcal{F}_H(m)$ at x for m large enough (because this is true for the algebraic stalk, and the map between the algebraic and analytic local rings is faithfully flat—they have the same completion), so you win. ☐

Quasi-Stein spaces

I would be remiss in not mentioning Kiehl's other big cohomological theorem, his analogue of Cartan's "Theorem A" and "Theorem B" in complex analysis.

A rigid space X is *quasi-Stein* if it admits an admissible affinoid covering $U_1 \subseteq U_2 \subseteq \cdots$ in which the image of $\mathcal{O}(U_{i+1})$ is dense in $\mathcal{O}(U_i)$ for each i. Besides affinoid spaces themselves, examples include open balls, one-dimensional annuli, and products of other quasi-Stein spaces.

Here's Kiehl's main theorem about quasi-Stein spaces. Kiehl's argument is a bit fragmentary (see below), so I'm looking for a better reference (preferably not in German); suggestions?

Theorem 16.4 (Kiehl)**.** *Let $\mathcal{F}$ be a coherent sheaf on a quasi-Stein space X, with a covering $\{U_i\}$ as above.*

(a) The image of $\mathcal{F}(X)$ in $\mathcal{F}(U_i)$ is dense for all i.

(b) The cohomology groups $H^i(X, \mathcal{F})$ vanish for all i ("Theorem B"). Note that van der Put's theorem applies, so sheaf = Čech here.

(c) For each $x \in X$, $\mathcal{F}_x$ is generated as an $\mathcal{O}_x$-module by global sections of $\mathcal{F}$ ("Theorem A").

Proof. Note that $\mathcal{F}(U_{i+1})$ is dense in $\mathcal{F}(U_i)$, from which apparently (a) is "immediate" ("unmittelbar"), but I don't see why offhand. (Do you?) To prove (b), it's enough to prove H^1 always vanishes (by a dimension shifting argument); this is done by an explicit calculation (see p. 271 of Kiehl's "Theorem A und B" paper). To prove (c), let $\mathcal{G}$ be the ideal sheaf of x; since $H^1(X, \mathcal{F} \otimes \mathcal{G})$ vanishes, the sequence

$$\mathcal{G}(X) = H^0(X, \mathcal{F}) \to H^0(X, \mathcal{F}/(\mathcal{F} \otimes \mathcal{G})) = \mathcal{F}_x/\mathfrak{m}_x\mathcal{F}_x \to 0$$

is exact, and (c) follows by Nakayama's lemma. □

This *almost* means you can pretend quasi-Stein spaces are just like affinoid spaces from the point of view of the cohomology of coherent sheaves. However, note that (c) does not imply that $\mathcal{F}$ is generated by *finitely many* global sections, because there is no compactness argument available. And indeed, it may not be: e.g., let $P_1, P_2, \ldots$ be points in the open unit disc such that $|P_i| \to 1$ as $i \to \infty$, and take the direct sum of the ideal sheaves of the P_i.

On the other hand, it sometimes happens that all *locally free* coherent sheaves are generated by finitely many global sections even though the underlying space is not affinoid, e.g., on an open annulus (see exercise from an earlier handout).

Chapter 17

Berkovich spaces for dummies

References: For the little I'm going to say today, [FvdP, Section 7.1] is a sufficient reference. (Note I'm using their "filters" instead of Berkovich's "nets".) For more details, you have to read Berkovich, which can be a daunting task. Definitely start with his ICM 1998 talk ("p-adic analytic spaces"), then try his IHES paper "Étale cohomology for non-Archimedean analytic spaces". Beware that his earlier monograph "Spectral theory and analytic geometry over non-Archimedean fields" makes some definitions which are not consistent with the later papers (or so Johan tells me, I only looked at the later papers).

Addenda: analytification and properness

A couple of comments from Brian Conrad, clarifying some points from earlier.

Re analytification: SGA1, Exposé XII is worth a look for the complex analytic setup; basically all the formalism there goes through without incident. (For more comments specifically on the rigid case, see section 5 of Brian's paper "Irreducible components of rigid spaces", *Ann. Inst. Fourier (Grenoble)* **49** (1999), 473–541; but be forewarned that basically that will tell you what I just said.) Abstractly, the analytification of a K-scheme X locally of finite type can be characterized by the fact that it represents the functor associating to an analytic space Y the set of ringed space maps $Y \to X$. That means it's unique if it exists; existence is a series of reductions to affine space.

Unfortunately, dealing with properness seems to really be as hard as we were finding it to be in class. Brian's paper deals with why properness for an algebraic variety is equivalent to properness for its analytification; it uses comparatively recent work of Temkin (student of Berkovich); in particular, proving that the composition of proper maps is proper is a big hassle.

Re GAGA: the GAGA theorems are also true in the proper case (I forget who pointed that out), and you can see the reduction to the projective case (using Chow's lemma) in SGA1, Exposé XII, section 4.

Filters (again)

I snuck filters onto a previous handout, but since I didn't go over them in class, I'd better repeat myself.

Given a space X equipped with a G-topology, a *G-filter* (or simply "filter") on X is a collection $\mathcal{F}$ of admissible subsets with the following properties.

(a) $X \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$.

(b) If $U_1, U_2 \in \mathcal{F}$, then $U_1 \cap U_2 \in \mathcal{F}$.

(c) If $U_1 \subseteq U_2$ and $U_1 \in \mathcal{F}$, then $U_2 \in \mathcal{F}$.

A *prime filter* is a filter $\mathcal{F}$ also satisfying

(d) if $U \in \mathcal{F}$ and $\{U_i\}_{i \in I}$ is an admissible covering of U, then $U_i \in \mathcal{F}$ for some $i \in I$.

A *maximal filter* (or *ultrafilter*) is a filter $\mathcal{F}$ which is maximal under inclusion; such a filter is clearly also prime. For each $x \in X$, the set of admissibles containing x is a maximal filter.

Let $\mathcal{P}(X)$ and $\mathcal{M}(X)$ denote the sets of prime and maximal filters, respectively, on X, and likewise for any admissible open U of X (that is, $\mathcal{P}(U)$ consists of prime filters of X in which U appears). Equip $\mathcal{P}(X)$ with the ordinary topology generated by the $\mathcal{P}(U)$; then there is a natural morphism of sites $\sigma : X \to \mathcal{P}(X)$, and it turns out that the functors σ_* and σ^* are equivalences between the categories of abelian sheaves on X and on $\mathcal{P}(X)$ [FvdP, Theorem 7.1.2].

Moreover, $\mathcal{P}(X)$ has "enough points" in the topos-theoretic sense: you can check whether a sheaf is zero by checking that its stalks at points of $\mathcal{P}(X)$ are all zero. That's cold comfort if you can't get a handle on those stalks, but for rigid spaces you can!

Filters and valuations

The spaces $\mathcal{P}(X)$ and $\mathcal{M}(X)$ are pretty unwieldy in general, but for rigid analytic spaces, we can make them more explicit.

First of all, let X be a rigid space. Then X carries a structure sheaf $\mathcal{O}$. That sheaf has a subsheaf of rings $\mathfrak{o}$, consisting of functions of spectral seminorm bounded by 1 everywhere. (That is, if U is an affinoid subspace and $A = \Gamma(\mathcal{O}, U)$, then $\Gamma(\mathfrak{o}, U) = \mathfrak{o}_{A,\mathrm{spec}}$.)

Now let $p \in \mathcal{P}(X)$ be a prime filter. Let $\mathcal{O}_p$ and $\mathfrak{o}_p$ be the stalks of $\mathcal{O}$ and $\mathfrak{o}$, respectively, at p; that is, $\mathcal{O}_p = \lim_{\to} \Gamma(\mathcal{O}, U)$ for U running over affinoid subsets U of X containing p, and similarly for $\mathfrak{o}_p$. (You may of course run the limit over any cofinal set of neighborhoods, e.g., rational subsets of a particular affinoid neighborhood.) Define the seminorm $\|\cdot\|_p$ on $\mathcal{O}_p$ by

$$\|f\|_P = \inf\{\|f\|_U : U \in p, f \in \Gamma(\mathcal{O}, U)\}.$$

Let $\mathfrak{m}_p$ be the ideal of $\mathcal{O}_p$ consisting of elements of seminorm 0; of course $\mathfrak{m}_p \subseteq \mathfrak{o}_p$ also.

To speak intelligently about this stalk, we need a bit of valuation theory. Let G be a divisible totally ordered group; then I can view G as a vector space over $\mathbb{Q}$. The completion $\widehat{G}$ of G is then a vector space over $\mathbb{R}$; its dimension is the *real rank* of G. If G is a totally ordered group but not divisible, define its real rank to be the real rank of its divisible closure. If R is a valuation ring, define the real rank of R to be the real rank of its valuation group $(\mathrm{Frac}\, R)^*/R^*$ (viewed additively, contrary to our usual convention). Note that R has real rank 1 if and only if its valuation corresponds to a nonarchimedean valuation $|\cdot| : (\mathrm{Frac}\, R)^* \to \mathbb{R}_{>0}$.

We now state [FvdP, Proposition 7.1.8]. It's enough to consider affinoid spaces, since the stalk only depends on an affinoid neighborhood.

Proposition 17.1. *Let $X = \mathrm{Max}(A)$ be an affinoid space and let p be a prime filter of X.*

(a) The ring $\mathcal{O}_p$ is a henselian local ring with maximal ideal $\mathfrak{m}_p$. (We will hereafter denote its residue field by k_p.)

(b) The ring $\mathfrak{o}_{k_p} = \mathfrak{o}_p/\mathfrak{m}_p$ is a valuation ring with fraction field k_p, and its real rank is at most $\dim(X)+1$. Also, if $\pi \in K$ satisfies $0 < |\pi| < 1$, then $\cap_{n=1}^{\infty} \pi^n \mathfrak{o}_{k_p} = 0$.

(c) Let $\mathfrak{p}$ be the kernel of $A \to k_p$, and let B be the inverse image of $\mathfrak{o}_{k_p}$ in $\mathrm{Frac}\, A/\mathfrak{p}$. Then the image of $\mathfrak{o}_{A,\mathrm{spec}}$ in $A/\mathfrak{p}$ is contained in B; moreover, $\cap_{n=1}^{\infty} \pi^n B = 0$.

Proof. See [FvdP, Proposition 7.1.8]. □

We thus have a meaningful valuation associated to any prime filter; the converse is also true. More precisely, a *valuation on* X consists of a pair $(\mathfrak{p}, B)$, in which $\mathfrak{p}$ is a prime ideal of A and B is a valuation ring of $\operatorname{Frac} A/\mathfrak{p}$ with the following properties.

(a) The image of $\mathfrak{o}_{A,\mathrm{spec}}$ in $\operatorname{Frac} A/\mathfrak{p}$ lies in B.

(b) For some (any) $\pi \in K$ with $0 < |\pi| < 1$, $\cap_{n=1}^{\infty} \pi^n B = 0$.

Proposition 17.2. *(a) The construction of Proposition 17.1 yields a bijection between the prime filters on X and the valuations on X.*

(b) Under the bijection in (a), the maximal filters on X correspond precisely to the valuations of real rank 1.

Proof. See [FvdP, Theorem 7.1.10]. □

Time out: Gelfand-Naimark

It is worth being reminded of a fundamental fact from classical functional analysis, that will put what we just did in a better context and suggest how to move forward.

Let X be a compact (Hausdorff) topological space, and let $C(X)$ denote the space of $\mathbb{C}$-valued continuous functions on X. Then X is a commutative C^*-algebra under the supremum norm (i.e., it's complete, it carries a complex conjugation $*$, and you can compute the norm of f as the square root of the norm of ff^*). The points of X carry algebraic meaning in $C(X)$: they give rise to maximal ideals on $C(X)$ (which are all distinct by Hausdorffness).

The Gelfand-Naimark theorem (which algebraic geometers might think of as the analogue of the Nullstellensatz in this context) asserts that on one hand these are all the maximal ideals of X, and on the other hand *any* commutative C^*-algebra A can be realized as $C(X)$ by putting a suitable topology on $X = \operatorname{Max} A$. Namely, for $x \in A$, $A/\mathfrak{m}_x$ is isomorphic to its subring $\mathbb{C}$ by the Gelfand-Mazur theorem (every complex commutative unital division Banach algebra is $\mathbb{C}$; and no, it's a different Mazur); the desired topology is the coarsest topology under which for each $f \in A$, the function $X \to \mathbb{C}$ sending x to $f(x) \in A/\mathfrak{m}_x \cong \mathbb{C}$ is continuous.

The space of valuations

Identify the set of maximal filters $\mathcal{M}(X)$ of an affinoid space $X = \operatorname{Max} A$ with maps $|\cdot|_a : A \to \mathbb{R}_{\geq 0}$ satisfying:

(i) $|fg|_a = |f|_a |g|_a$;

(ii) $|f + g|_a \leq \max\{|f|_a, |g|_a\}$;

(iii) $|c|_a = |c|$ for $c \in K$;

(iv) $|f|_a \leq \|f\|_X$ for $f \in A$.

We now topologize $\mathcal{M}(X)$ with the coarsest topology such that for each $F \in A$, the map $\mathcal{M}(X) \to \mathbb{R}_{\geq 0}$ sending $a \in \mathcal{M}(X)$ to $|f|_a$ is continuous; in other words, this is the topology induced by the product topology on $\mathbb{R}_{\geq 0}^A$. In particular, this space is Hausdorff and compact.

Aside: the Berkovich topology coincides not with the subspace topology on $\mathcal{P}(X)$, but for the quotient topology under a certain natural retraction $r : \mathcal{P}(X) \to \mathcal{M}(X)$ (see [FvdP, Definition 7.1.4]).

Note that this definition glues, so we can talk about $\mathcal{M}(X)$ even when X is not affinoid, although I'll refrain from doing so for the moment.

The point here is that $\mathcal{M}(X)$ with Berkovich's topology is "better connected" than X with its metric topology. Here's an easy case of this; more generally, Berkovich showed that "Smooth p-adic analytic spaces are locally contractible" (*Invent. Math.* **137** (1999), 1–84, though the proof is a lot harder.

Proposition 17.3. *Let A be a reduced affinoid algebra whose reduction is an integral domain, and put $X = \operatorname{Max} A$. Then $\mathcal{M}(X)$ is contractible.*

Proof. Let $i : \mathcal{M}(X) \to \mathcal{M}(X)$ be the identity map and let $j : \mathcal{M}(X) \to \mathcal{M}(X)$ be the map carrying everything to the "generic point" $|f|_a = \|f\|_{A,\mathrm{spec}}$. (Note that I'm using here that the spectral seminorm is multiplicative; that's precisely what is guaranteed by the condition on the reduction of A.) Then there is an explicit homotopy $F(a,t) : \mathcal{M}(X) \times [0,1] \to \mathcal{M}(X)$ given by

$$F(a,t) = |\cdot|_a^t \|\cdot\|_{A,\mathrm{spec}}^{1-t}.$$

□

More interesting things happen when you consider nonaffinoid spaces; for instance, Berkovich proved that if X is the analytification of an smooth irreducible projective curve over K of genus g having semistable reduction, then $\mathcal{M}(X)$ can be contracted to a closed subspace homeomorphic to the dual graph of the special fibre on a semistable model of X [FvdP, Theorem 7.2.4 for a more detailed statement]. I have no idea what $\mathcal{M}(X)$ looks like if X has non-semistable reduction, though I suppose you could probably figure it out by going up to an extension of K over which semistable reduction is acquired.

Example: the Berkovich closed unit disc

Let $X = \operatorname{Max} K\langle x\rangle$; let's identify $\mathcal{M}(X)$ explicitly in case K is algebraically closed. (I'll let you work out the general case for yourself.) Given an analytic point $a \in \mathcal{M}(X)$ and corresponding valuation $|\cdot|_a$, define the function $F_a : \mathfrak{o}_K \to [0,1]$ by $F_a(y) = |x-y|_a$. Such a function satisfies the triangle inequalities

$$|y-z| \leq \max\{F_a(y), F_a(z)\}, \qquad F_a(y) \leq \max\{F_a(z), |y-z|\}.$$

Let $\mathcal{N}$ be the set of functions F_a satisfying these inequalities, with the product topology (viewing functions as elements of the product of a bunch of copies of $[0,1]$ indexed by $\mathfrak{o}_K$). Then the map $\mathcal{M}(X) \to \mathcal{N}$ is a homeomorphism [FvdP, Lemma 7.2.1].

The functions F_a can be classified as follows.

- If $\inf(F_a) = 0$, then a is an ordinary point of X.
- If $\inf(F_a) > 0$ and F_a achieves its infimum at some $y \in \mathfrak{o}_K$, then a is a "generic point" of the disc $|x-y| \leq \inf(F_a)$ (which is affinoid if $\inf(F_a) \in |K^*|$), and one has $F_a(z) = \max\{|z-y|, \inf(F_a)\}$ for all $z \in \mathfrak{o}_K$ (so F_a coincides with the supremum norm on that small disc).
- If $\inf(F_a) > 0$ but F_a does not achieve its infimum, then K must fail to be spherically complete, and we can view F_a as the limit of the supremum norms on a decreasing sequence of discs. (This jibes with Jay's comment earlier that he had read somewhere that the Berkovich construction is somehow analogous to spherical completion.)

So what?

So far it looks like Berkovich's construction is a convenient gadget for visualization but doesn't suggest anything you couldn't do already. Not so! This theory turns out to be (and the Tate and Raynaud theories turn out not to be) just the thing for discussing the étale cohomology of rigid analytic spaces. The main point is that to do that, you need to have enough fibre functors for the étale topos (in more precise terms, you need a "conservative family of fibre functors"), and the analytic points give you just that. (Note that we've already run across this issue on the Zariski site.) That étale cohomology theory is vital for dealing with the sort of p-adically uniformized spaces occurring in Drinfeld's work on the Langlands correspondence for function fields.

There is a more topos-theoretic alternative if you prefer, which is Huber's theory of "adic spaces". However, in case you couldn't tell, I don't really go in for that sort of thing, and so I'm not going to discuss it further.

Exercises

1. (from [FvdP, Exercise 7.2.5]) Suppose K has characteristic $p > 0$. Put $X = \operatorname{Max} K\langle x \rangle$. Prove that the map $X \to X$ induced by $x \mapsto x^p - x$ is finite and unramified at each ordinary point of X, but ramifies at one analytic point of X

Chapter 18

More corrections on Berkovich spaces

I'll get this right eventually, I promise. (Again, many thanks to Brian Conrad for clarifications; remaining errors are of course not his fault!)

The definition of a quasi-separated rigid space is not what I said in class. It is (just as for schemes): X is quasi-separated if $X \to \operatorname{Max} K$ is quasi-separated, and a morphism $X \to Y$ is quasi-separated if the diagonal $\Delta_{X/Y} : X \to X \times_Y X$ is a quasi-compact morphism (not a locally closed immersion). Oh, and a morphism is quasi-compact if the inverse image of every quasi-compact open is quasi-compact.

Re the notion of a net: the restriction of a quasi-net to an open subset means take elements of the quasi-net contained in the open, rather than intersecting all of them with the open. So my example with the unit squares is not a net; you can fix it by adding in the closed edges and the corners.

More serious functor: the rigid-to-Berkovich functor is backwards from what I've been saying! Mea culpa. It turns Hausdorff strictly analytic spaces into quasiseparated rigid spaces, and in *that* direction is fully faithful. It is true, though, that it induces an equivalence between paracompact strictly analyic spaces and quasi-separated rigid spaces with an admissible affinoid covering of finite type. ("Finite type" means that any element of the covering meets only finitely many others. E.g., the covering of the open unit disc by closed discs around the origin is not of finite type, but there is an admissible covering of finite type by closed annuli.) The point here is that quasi-separatedness is needed in order to glue meaningfully (otherwise the attaching maps on the Berkovich affinoids are not uniquely determined by their rigid counterparts, so the cocycle condition falls apart) and the finite type condition is needed in order to build a quasi-net. (Maybe with a more sensible definition of Berkovich's spaces you could get past the finite type issue?) Anyway, see Theorem 1.6.1 of Berkovich's IHES paper.

Upshot: the Berkovich category neither contains nor is contained in the category of rigid spaces, but they share the rigid spaces which are "not too pathological", which the spaces you typically encounter in practice will be.

And one fun thing you might want to try (suggested by Berkovich in his ICM talk): compute the Gelfand spectrum of $\mathbb{Z}$. (Remember, this means you allow the normal triangle inequality, and you don't impose any upper bound on the seminorms.)

Chapter 19

p-adic cohomology, part 1

References: [FvdP, Chapter 7]. Also, Berthelot's "Géométrie rigide et cohomologie des variétés algébriques de caractéristique p" is a good (if cursory) overview; more on his other papers below.

But first, I have a bunch of corrections and addenda from the discussion of Berkovich's construction.

Corrections from last time

As noted in class, the map $\sigma : X \to \mathcal{P}(X)$ for a G-topological space X is not actually a continuous morphism of sites. Nonetheless, it "gives rise to" a "direct image" $(\sigma_* \mathcal{F})(U) = F(\mathcal{P}(U))$ and "inverse image" $\sigma^* \mathcal{F}$ the sheafification of the presheaf $U \mapsto \lim_\to \mathcal{F}(V)$, where V runs over admissible sets with $U \subseteq \mathcal{P}(V)$. See [FvdP, Theorem 7.1.2]; note that it requires the G-topological space X to be quasi-compact. The point is that σ_* is an equivalence between the category of abelian sheaves on $\mathcal{P}(X)$ and on X, with quasi-inverse σ^*. [FvdP] punts on the question of whether the same is true for sheaves of *sets*, which would say that these two maps give an isomorphism of topoi between X and $\mathcal{P}(X)$. (I had thought this followed formally by applying the "free abelian group generated by" functor; a clarification would be appreciated.)

I mentioned in class there is a retraction map $r : \mathcal{P}(X) \to \mathcal{M}(X)$ when X is a rigid space. Here's one way to define this map: if p is a prime filter, $r(p)$ is the unique maximal filter containing r. The uniqueness is a theorem, but not a hard one: if p corresponds to a valuation of some rank, m corresponds to the valuation where you "ignore all terms not of highest order". For instance, if p corresponded to a valuation into $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ ordered lexicographically, then m would correspond to the projection of that valuation onto the *first* copy of $\mathbb{R}$. For another approach, see [FvdP, Definition 7.1.4].

My alleged proof of the contractibility of the Berkovich space for an affinoid algebra whose reduction is an integral domain doesn't work, and I don't see how to fix it offhand. In fact, I'm not even sure if the result as stated is true! On the other hand, it's definitely true that the Berkovich space of $K\langle x \rangle$ is true, because from the generic point of any disc you can travel to the generic point of the disc with the same center and any radius, including 1, which is the generic point of the whole space. (Compare [FvdP, Lemma 7.2.2].) I think something similar works for $K\langle x_1, \dots, x_n \rangle$, but I didn't check details.

Addendum: Globalizing Berkovich

The discussion from last time about $\mathcal{M}(X)$ for X an affinoid space hides some of the subtle distinctions between our category of rigid spaces and Berkovich's category of "analytic spaces". Indeed, Berkovich changed his category between his original monograph and his IHES paper; in any case, his category is somewhat richer than the one we have in mind. (See his IHES paper for details.)

For Berkovich, an "affinoid algebra" is anything covered by the ring of functions on the closed polydisc $|x_1| \leq r_1, \dots, |x_n| \leq r_n$ for *any* $r_1, \dots, r_n > 0$; in particular, they don't have to be in the divisible closure of $|K^*|$. Our affinoid algebras are "strictly affinoid" in his terminology. (Berkovich also allows K to carry the trivial valuation, to make certain arguments more uniform.)

A *quasi-net* on a topological space X is a collection τ of subsets of X with the property that for any $x \in X$, we can find $V_1, \dots, V_n \in \tau$ such that $x \in V_1 \cap \dots \cap V_n$ and the set $V_1 \cup \dots \cup V_n$ is a neighborhood of x. (Note that I didn't say "open neighborhood". I believe the intent is to allow any set containing an open subset containing x.) A *net* is a quasi-net whose restriction to $U \cap V$, for any U, V in the net, is again a quasi-net. Example: any basis of open sets. More amusing example: divide the plane into unit squares like an infinite chessboard, and take the closed unit squares in this decomposition. (Warning: this is Berkovich's terminology, but is inconsistent with the preexisting term "net" in ordinary topology, which refers to a generalized sequence!)

An *affinoid atlas* on a locally Hausdorff topological space X, equipped with a net τ of compact subsets, is a homeomorphism of each $U \in \tau$ with an affinoid space, such that for $U, V \in \tau$ with $U \subseteq V$, U is identified with an affinoid subspace of V. An *analytic space* is a space equipped with an affinoid atlas, modulo inverting "quasi-isomorphisms". (Or if you prefer, you can insist that the affinoid atlas be maximal, and then there's nothing to invert. Aside: I think you can give a ringed space definition instead, but I don't know any such definition in the literature.)

The natural subspaces of an analytic space X are *analytic domains*. An analytic domain is a subset Y of X such that for any $y \in Y$, you can find affinoid subspaces $U_1, \dots, U_n$ in Y with $y \in U_1 \cup \dots \cup U_n$ and $U_1 \cup \dots \cup U_n$ being a neighborhood.

Relationship with our other construction: the natural functor from our rigid spaces to Berkovich spaces is fully faithful, and it induces an equivalence between the category of paracompact strictly analytic spaces in Berkovich's sense, and the category of quasi-separated rigid spaces admitting an admissible affinoid covering of finite type. (Of course, most spaces you will meet in practice are of this form...)

More addenda

Brian Conrad reminds me that there is an unpublished German PhD thesis (by Ulrike Kopf) that works out a lot of the details of rigid GAGA. I have a copy in my office somewhere. (I think Brian's original source for this was Richard Taylor.)

Brian also points out that Čech and sheaf cohomology agree for any sheaf on any quasi-separated rigid space (reminder: that means the diagonal $\Delta_X : X \to X \times X$ is a locally closed immersion), though this isn't in the literature. He sent me a handout (from a course he taught) with the details; I may incorporate these into a handout at some point.

Clarification: Temkin's work postdates the paper by Conrad that I mentioned last time, so you have to look there directly: "On local properties on non-Archimedean analytic spaces I, II". One of Temkin's key results is that every point of a Berkovich space admits an affinoid neighborhood. (This is true by design for the Berkovichifications of rigid analytic spaces, but it's not so obvious in general.) Aside: Berkovich's definition of "proper" is purely topological (see his ICM talk); it looks formally stronger than Kiehl's definition when applied to a rigid space, and maybe it even is stronger. In any case, it satisfies Kiehl's finiteness theorem, so maybe it's really the correct definition...

Motivation: algebraic de Rham cohomology

Now, on to today's topic.

Grothendieck's letter to Atiyah (published in IHES as "On the de Rham cohomology of algebraic varieties") explains the construction of algebraic de Rham cohomology on a smooth variety X over a field K of characteristic zero; basically, it's the hypercohomology of the complex of algebraic differentials

$0 \to \Omega^0_{X/K} \to \Omega^1_{X/K} \to \cdots$. Over K with X proper, this agrees with holomorphic de Rham cohomology by GAGA, which agrees with smooth de Rham cohomology by the Dolbeaut lemma (see Griffiths-Harris), which agrees with topological cohomology with $\mathbb{C}$-coefficients by de Rham's theorem.

Hartshorne (published in IHES with the same title!) worked out a construction that works for nonsmooth varieties also. (There is also Grothendieck's related but more coordinate-free approach via the infinitesimal site, i.e., via "crystals", which I'm going to ignore for now.) The basic idea is to replace X Zariski (or étale) locally by its formal completion in some smooth ambient variety, and check that the resulting de Rham complex is "canonically independent of the choice of compactification".

The idea behind Berthelot's rigid cohomology (which was inspired by Dwork's proof of rationality of the zeta function as well as by ideas of Grothendieck and Monsky-Washnitzer) is to use a similar paradigm to construct de Rham cohomology of varieties in characteristic $p > 0$, by first passing to a p-adic lifting. Loosely speaking, the choice of the lift should drop out "at the level of homotopy".

Dagger algebras and de Rham cohomology

Additional references: besides [FvdP, 7.5], see also the papers of Elmar Grosse-Klönne (but make sure also to read the MathSciNet review of "Rigid analytic spaces with overconvergent structure sheaf" for an important erratum).

The *Monsky-Washnitzer algebra* $K\langle x_1, \dots, x_n\rangle^\dagger$ is the set of formal power series $\sum c_I x^I$ which converge on some polydisc $|x_i| \leq c$ for some $c > 1$ (but the choice of c depends on the particular series). This is a dense subring of the Tate algebra $K\langle x_1, \dots, x_n\rangle$, and you can topologize it that way (I'll call that the "affinoid topology"). On the other hand, it's sometimes better to topologize it as as the direct limit of the coordinate rings of the closed polydiscs of all radii > 1 (I'll call that the "dagger topology").

One can verify that pretty much all of the standard facts about Tate algebras apply to these guys too. For instance, they satisfy Weierstrass preparation [FvdP, Lemma 7.5.1], Noether normalization, et cetera. The one thing they do which Tate algebras don't is admit a Poincaré lemma: the partial derivative maps $\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}$ on $K\langle x_1, \dots, x_n\rangle$ are all surjective!

A *dagger algebra* (or *overconvergent affinoid algebra*) is a quotient of a $K\langle x_1, \dots, x_n\rangle^\dagger$; it turns out any map between these is compatible with both the induced affinoid and dagger topologies. The Max of such an algebra is the same as that of its (affinoid) completion. There is an analogous notion of *dagger subspace* of $\operatorname{Max} A$ for A a dagger algebra; any rational subspace is a dagger subspace, and any dagger subspace is an affinoid subspace, but I don't know whether every affinoid subspace is a dagger subspace. (Neither does [FvdP], and as far as I can tell neither does Grosse-Klönne.) But never mind that; you still get the same strong G-topology on $\operatorname{Max} A$ using just the dagger subdomains (by Gerritzen-Grauert), and so you get a subsheaf $\mathcal{O}^\dagger$ of $\mathcal{O}$, the *overconvergent structure sheaf.*

A *dagger space* is a locally G-ringed space which locally on some admissible cover looks like an affinoid space equipped with an overconvergent structure sheaf. (Note: the forgetful functor from dagger to rigid spaces is faithful but not fully faithful; that is, there are different incompatible dagger structures on a typical affinoid algebra.) Besides the structure sheaf, every dagger space also carries a sheaf of (overconvergent) continuous differentials $\Omega^\dagger_X$, whose definition I'll leave to your imagination. The *de Rham complex* on X is of course

$$\cdots \to \Omega^{i,\dagger}_X = \wedge^i_{\mathcal{O}^\dagger_X} \Omega^\dagger_X \to \cdots$$

and its hypercohomology is the *de Rham cohomology* of X. (If you don't like hypercohomology, think for now about an overconvergent affinoid space, in which case all the sheaf cohomology drops out and it's just the cohomology of the ordinary complex of global sections.)

It turns out [FvdP, Proposition 7.5.13] that the de Rham cohomology of a dagger space actually depends only on its underlying rigid space; moreover, they are functorial in the rigid spaces. (Idea: say $f : X \to Y$

is a map between rigid spaces, and that each of X and Y has been equipped with a dagger structure. By a certain approximation argument, you can closely approximate f by an actual map between the dagger spaces. You then show that there is a homotopy between the maps on de Rham complexes induced by any two choices of f.)

Monsky-Washnitzer cohomology

I promised earlier that this would have something to do with characteristic p, and indeed it does!

Let K be a complete discretely valued field of characteristic 0, whose residue field k has characteristic $p > 0$. Let $X = \operatorname{Spec} \overline{A}$ be a smooth affine scheme over k.

By a theorem of Elkik (for a more modern treatment, see the paper by Arabia: "Relèvements des algèbres lisses et des leurs morphismes"), you can lift $\overline{A}$ to a smooth $\mathfrak{o}_K$-algebra A. The *Monsky-Washnitzer cohomology* of X is the overconvergent de Rham cohomology of the affinoid space $\operatorname{Max} \widehat{A}$; by the same arguments as in the previous section, it is independent of the choice of the lift and functorial in X. It is finite dimensional over K, but this is not so easy to prove: it was first shown by Berthelot using rigid cohomology (see next handout).

Monsky and Washnitzer, inspired by Dwork's proof of rationality of the zeta function, gave a Lefschetz trace formula for Frobenius for their cohomology: if $k = \mathbb{F}_q$ and X is of pure dimension d, then

$$\#X(\mathbb{F}_{q^n}) = \sum_i (-1)^i \operatorname{Trace}((q^d F_*^{-1})^n, H^i_{MW}(X)),$$

where F_* denotes the map induced by the algebraic Frobenius $t \mapsto t^q$. See Monsky-Washnitzer's "Formal Cohomology I" and Monsky's "Formal Cohomology II, III" (or [FvdP, 7.6]). The basic idea here is to establish enough topological formalism so that you can do excision to reduce to the case where X has *no* $\mathbb{F}_{q^n}$-rational points, and then show that the right side is forced to vanish. In fact, it already vanishes *on the chain level*!

Note: making the previous paragraph makes some sense requires a bit of p-adic functional analysis. Even if you know finite dimensionality in cohomology (which Monsky and Washnitzer didn't), the de Rham complex consists of infinite dimensional vector spaces over K. So you need to work with nuclear operators; in particular, you have to choose a so-called "Dwork operator" (essentially a one-sided inverse of Frobenius), show that it is nuclear, and then check (by a neat calculation) that its trace is forced to vanish.

It cannot go unsaid here that Monsky-Washnitzer cohomology is a fabulous tool for computing zeta functions and other cohomological invariants of varieties over finite fields, particularly when they lift nicely to characteristic zero. For instance, "Kedlaya's algorithm" computes the zeta function of a hyperelliptic curve over a finite field (of not-too-large characteristic) by computing on a dagger algebra and being careful about p-adic precision; see [FvdP, 7.6] or my paper "Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology".

Next time: rigid cohomology

Next time, I'll tack on a short discussion of how to globalize the ideas of the previous section, and how to extend them to singular varieties. This will give Berthelot's rigid cohomology.

Chapter 20

p-adic cohomology, part 2

References: this stuff is not written down all in one place. The survey article by Berthelot ("Geometrie rigide et cohomologie...") I mentioned last time is a good start. For more details, see his Inventiones article "Finitude et purité cohomologique en cohomologie rigide". Further references appear within the text.

Rigid cohomology

Last time, I described a good de Rham-type cohomology theory for smooth affine varieties over a field k of characteristic $p > 0$, based on taking liftings up to characteristic zero. You can "sheafify" that theory to get a good theory for general smooth varieties. (Be careful: the de Rham complexes were not functorial, only functorial "up to homotopy".)

What about general varieties? As I mentioned last time, for algebraic de Rham cohomology, a good approach is to locally embed a general variety into a smooth variety and work on the formal completion there. That "smudges out" the singularity and gives you a smooth space with the same "homotopy type". The idea here is similar, but the formal completion gets replaced by something a bit bigger.

Let X be a variety over a field k of characteristic $p > 0$. Suppose I have an open immersion $X \hookrightarrow Y$, where Y is another k-variety, and a closed immersion $Y \hookrightarrow P_k$, where P is a smooth formal scheme of finite type over $\mathfrak{o}_K$ (and P_k is its special fibre). A good example to visualize is $P = \mathbb{P}^n$. Let P_K denote the rigid analytic generic fibre of P; remember that this has a specialization map $\mathrm{sp} : P_K \to P_k$. For $S \subseteq P_k$, define the *tube* $]S[= \mathrm{sp}^{-1}(S)$. I'll write it as $]S[_P$ if I need to specify P, which will happen a bit later.

The most precise analogue of the formal completion construction would be to consider the tube $]X[$, but that has the same sorts of problems as we saw last time with the de Rham cohomology of a closed disc. Instead, we must take something slightly bigger.

Reminder: for U an admissible subset of a G-topological space X such that $X \setminus U$ is also admissible, we say V is a *strict neighborhood of* U *in* X if V is an admissible subset containing U, and $\{V, X \setminus U\}$ form an admissible cover of X.

So what we really want to look at is not the tube $]X[$, but a strict neighborhood of $]X[$ in $]Y[$. If P happens to be affine, this can be described more concretely: if $f_1, \dots, f_n \in \Gamma(\mathcal{O}, P)$ cut out Y within P_k, we can describe $]Y[$ as the locus where $|f_1|, \cdots, |f_n| < 1$, and in particular we can make an admissible cover out of the sets U_ϵ where $|f_1|, \dots, |f_n| \leq \epsilon$ for $0 < \epsilon < 1$. (Strictly speaking, I should be taking ϵ in the divisible closure of K^*, but to keep notation simple, I'm going to ignore that technicality consistently.) If $g_1, \dots, g_m \in \Gamma(\mathcal{O}, P)$ cut out $X \setminus Y$ within X, we can characterize strict neighborhoods of $]X[$ in $]Y[$ as follows: an admissible open V in $]Y[$ containing $]X[$ is a strict neighborhood if and only if its intersection with the affinoid U_ϵ contains the set on which

$$|f_1|, \cdots, |f_n| \leq \epsilon, \qquad |g_1|, \cdots, |g_m| \geq \delta$$

for some $\delta \in (0,1)$. (If δ were 1, we would just be picking up $]X[\cap U_\epsilon$.)

We get an exact functor $j^\dagger$ on abelian sheaves E on $]Y[$ by the formula

$$j^\dagger E = \lim_{\to} j_{V*} j_V^{-1} E,$$

the limit taken over strict neighborhoods V of $]X[$ in $]Y[$. This is sometimes called the "overconvergent sections functor"; we are particularly interested in applying it to the de Rham complex.

Example: if $X = \mathbb{A}^n$, $Y = \mathbb{P}^n_k$, and $P = \mathbb{P}^n_{\mathfrak{o}_K}$, then $]X[= \operatorname{Max} K\langle x_1, \dots, x_n\rangle$, and $\Gamma(j^\dagger \mathcal{O},]X[) = K\langle x_1, \dots, x_n\rangle^\dagger$.

Suppose you can set this situation up with P proper over $\mathfrak{o}_K$. Then we want to define the *rigid cohomology* of X with coefficients in K as the hypercohomology

$$H^i_{\mathrm{rig}}(X/K) = \mathbb{H}^i(]Y[_P, j^\dagger \Omega_{]Y[}).$$

However, one has to check that this is independent of the choice of P, as well as functorial in X. The basic idea is this: if $u : P' \to P$ is proper, you want to compare the rigid cohomology of X computed within P and within P'. (If you can do that, you can compare two choices of P by comparing each to the fibre product over $\mathfrak{o}_K$.) That you do by showing that the canonical morphism

$$j^\dagger \Omega_{]Y[_P} \to \mathbb{R} u_{K*} j^\dagger \Omega_{]Y[_{P'}}$$

is an isomorphism (e.g., see Théorème 1.4 of Berthelot's Inventiones paper). This in turn follows from the "strong fibration theorem", which says that locally on P_K, a sufficiently small strict neighborhood of $]X[_{P'}$ in $]Y[_{P'}$ looks like the product of an open unit polydisc with a strict neighborhood of $]X[_P$ in $]Y[_P$. Note how crucial it is here that the Poincaré lemma holds on an *open* polydisc! (To get functoriality along $X \to X'$, embed X into P and X' into P', then embed the graph of the map into $P \times P'$, etc.)

This is enough to define rigid cohomology of quasi-projective X. For general X, you need to cover X with, say, affines and use an appropriate Čech complex to define X. Berthelot never bothers to explain this rigorously; probably the right way to say this is in the language of simplicial sets; this is the way Shiho proceeds in his two papers "Crystalline fundamental groups..., I, II".

You won't be surprised to know that there is a comparison theorem between this construction and Monsky-Washnitzer cohomology (Théorème 1.10 in Berthelot's Inventiones paper). There is also a comparison between rigid and crystalline cohomology *after* tensoring the latter up to K (Théorème 1.9). But the latter is really an integral cohomology theory (defined over the Witt ring $W(k)$), so the passage to cohomology really loses some information (e.g., about the failure of the Hodge-de Rham spectral sequence to degenerate). Rigid cohomology does seem to be a "universal" p-adic cohomology with field coefficients, or if you like, a universal p-adic Weil cohomology.

There is also a construction of rigid cohomology with supports in a closed subscheme, and of cohomology with compact supports. These are needed to talk about excision sequences and Poincaré duality, and to correctly formulate the Lefschetz trace formula for Frobenius (which works on cohomology with supports, and doesn't require properness). See Berthelot for more details.

Isocrystals

How do you put coefficients into this theory? In de Rham cohomology, you use local systems: vector bundles equipped with an integrable connection. The point is that you need the connection in order to have complex maps on the vector bundle tensored with the original de Rham complex.

In rigid cohomology, you do something similar; this is described best in an unpublished preprint of Berthelot called "Cohomologie rigide, I", available on his web site at Rennes. (He has some other useful papers there, but the others are all published somewhere.) The basic idea goes back to Grothendieck's algebraic interpretation of integrable connections: given a vector bundle E on a space X, an integrable connection should come from a "parallel transport" isomorphism $\pi_1^* E \to \pi_2^* E$ on the formal completion along the diagonal $\Delta \subseteq X \times X$ by taking first-order infinitesimals. (On higher order infinitesimals, this isomorphism looks in coordinates like it's being computed by Taylor series.)

In my notation from before, what an "overconvergent isocrystal" on X should be is a vector bundle $\mathcal{E}$ on some unspecified strict neighborhood of $]X[_P$ in $]Y[_P$ plus a connection $\nabla : \mathcal{E} \to \mathcal{E} \otimes \Omega^1$, which induces a parallel transport isomorphism on some strict neighborhood of $]X[_{P'}$ in $]Y[_{P'}$, where $P' = P \times_{\mathfrak{o}_K} P$. This turns out to be the right construction; this category is independent of the choice of P (in that if P' is another choice, the pullback functor along $P' \to P$ is an equivalence). It's called the category of *isocrystals on X (over K) overconvergent along Z*, where $Z = Y \setminus X$. If Y is proper, these are just *overconvergent isocrystals on X*. (There are also *convergent isocrystals on X*, which are defined just on $]X[$ itself, but you still have to make the nontrivial restriction that the parallel transport isomorphism is defined on all of $]X[_{P \times P}$.)

These are known to have reasonable cohomological properties under some additional hypotheses: K should be discretely valued and the isocrystals should carry "Frobenius structure", i.e., an isomorphism between the isocrystal and its pullback along some lift of the absolute Frobenius. See my preprint "Finiteness of rigid cohomology with coefficients". You can also scrape by just enough formalism mirroring that of étale cohomology to prove the Weil conjectures, by reproducing Laumon's simplified proof of Deligne's "Weil II" theorem (essentially, the Weil conjectures with coefficients); see my preprint "Fourier transforms and p-adic 'Weil II"'.

A big problem right now is to enlarge this category of coefficients; the category of isocrystals has no hope of being stable under direct images, because objects are of constant rank. (In étale cohomology, this would be like working only with the lisse (smooth) sheaves and not the constructible ones.) The fix in algebraic de Rham theory is to work with $\mathcal{D}$-modules, for $\mathcal{D}$ a suitable sheaf of differential operators. (You can imagine vector bundles with integrable connection as carrying an action of certain differential operators via the connection.) There you have a good finiteness notion called "holonomicity"; reproducing that in some sort of p-adic $\mathcal{D}$-module context is a problem Berthelot has been stuck on for a long time. (The $\mathcal{D}$-modules in question manifest already in the Weil II argument mentioned above, but in a rather simple way that doesn't cause much trouble.)

Upshot: although rigid cohomology comes a long way towards the dream of finding a way to interpret Dwork's work on zeta functions in terms of a fully functional p-adic cohomology theory for varieties over a field of positive characteristic, there are a lot of technical issues that remain unresolved.

Chapter 21

The Lubin-Tate moduli space

Note: there is no rigid geometry in this set of notes! That will come next time, when we talk about the period mapping.

References: For starters, the original paper of Lubin-Tate (which involves no rigid geometry, only formal geometry) is: J. Lubin and J. Tate, Formal moduli for one-parameter formal Lie groups, *Bulletin de la Soc. Math. de France*, **94** (1966), 49–59. (This is not their paper on local class field theory, though of course the two are closely related.) As for "Gross-Hopkins", there are two such papers. One (here [GH1]) is: M.J. Hopkins and B.H. Gross, The rigid analytic period mapping, Lubin-Tate space, and stable homotopy theory, *Bulletin of the AMS* **30** (1994), 76–86. This paper makes the link between Lubin-Tate spaces and stable homotopy theory; I won't do that here. It is quite cursory on the geometry side (I'm not capable of judging on the homotopy side), to the point of being barely legible. The second paper (here [GH2]) is: M.J. Hopkins and B.H. Gross, Equivariant vector bundles on the Lubin-Tate moduli space, in *Topology and Representation Theory*, Contemporary Mathematics **158**, AMS, 1994, 23–88. As the page count suggests, this is *much* more detailed and focuses entirely on the rigid geometry.

Formal groups

Before we do any geometry, here's a quick review of formal groups. The standard reference is Hazewinkel, *Formal Groups*, though I think his formulas have some errors in them; use with caution.

A *(commutative) formal group of dimension* n over a ring R is a (commutative) cogroup structure on $R[\![x_1, \dots, x_n]\!]$ with identity 0, i.e., a (commutative) comultiplication satisfying the usual (co)group axioms. It's of course enough to specify how $x_1, \dots, x_n$ behave under the comultiplication; their images form an n-tuple of power series

$$F(X, Y) = (F_1(X, Y), \dots, F_n(X, Y))$$

(where X is short for $x_1, \dots, x_n$ and Y for $y_1, \dots, y_n$), such that

$$F(X, 0) = X, \quad F(0, Y) = Y$$
$$F(X, F(Y, Z)) = F(F(X, Y), Z)$$
$$F(X, Y) = F(Y, X).$$

(Exercise: the existence of inverses is automatic given the other axioms.) A morphism $f : F \to F'$ of formal groups is an n-tuple of n-variate power series such that $F'(f(X), f(Y)) = F(X, Y)$. Let $\mathrm{Lie}(F)$ denote the (relative) tangent space of $R[\![x_1, \dots, x_n]\!]$, i.e., the trivial Lie algebra over R generated by $x_1, \dots, x_n$. Every endomorphism of F induces a R-linear endomorphism on $\mathrm{Lie}(F)$, which is just look at the linear terms in the power series.

Examples: take any algebraic group of dimension n and restrict the group law to the tangent space at the origin, and you get a formal group of dimension n. For $\mathbb{G}_a$, you get

$$F(x, y) = x + y.$$

For $\mathbb{G}_m$, you get

$$F(x, y) = x + y + xy,$$

or $x + y - xy$ for another choice of coordinates, or crazier things for more bizarre choices of coordinates. (Those two are isomorphic if $\mathbb{Q} \subseteq R$, even though the original groups are not, but not in general. I'll make a stronger statement below.) Any elliptic curve gives a formal group of dimension 1, as constructed in Silverman's book. Also, abelian varieties and linear groups give you other examples in higher dimension; however, I'm mostly interested here in dimension 1.

If I knew what it was, I would mention here the connection between formal groups of dimension 1 and stable homotopy theory (which as far as I can tell is due more or less entirely to Hopkins). However, I don't; maybe Mark can enlighten us a bit at some point.

Formal $\mathfrak{o}$-modules

I'm also going to work a bit (following [GH2]) with Drinfeld's more general notion of formal $\mathfrak{o}$-modules, where $\mathfrak{o}$ is a complete DVR with finite residue field $k = \mathbb{F}_q$. Fix a choice of a uniformizer π of $\mathfrak{o}$, and put $K = \operatorname{Frac} \mathfrak{o}$ as usual. Let R be a (commutative) $\mathfrak{o}$-algebra; I'll call the structure map $i : \mathfrak{o} \to R$ if I need to refer to it. A *formal* $\mathfrak{o}$*-module of dimension* n *over* R is a (commutative) formal group F of dimension n equipped with a ring homomorphism $\theta = \theta_F : \mathfrak{o} \to \operatorname{End}_R(F)$, such that

$$\theta(a)(X) \equiv i(a)X \pmod{(x_1, \dots, x_n)^2}.$$

(That is, the action of $\theta(a)$ on $\operatorname{Lie}(F)$ is by multiplication by $i(a)$.) A formal group is automatically a formal $\mathbb{Z}_p$-module as long as $\mathbb{Z}_p \subseteq R$.

Convention: I'll write a_F instead of $\theta_F(a)$, as in [GH2].

Example: the first Lubin-Tate paper (Formal complex multiplication in local fields, *Ann. Math.* **81** (1965), 380–387) show that you can uniquely specify a formal $\mathfrak{o}$-module of dimension 1 by specifying the action of π: it must be given by a series $f(x)$ with $f(x) \equiv \pi x \pmod{x^2}$ and $f(x) \equiv x^q \pmod{\pi}$.

Example: in equal characteristic p, Drinfeld stumbled across these as the analogues of the formal group associated to an algebraic group, for what we call "Drinfeld modules". Briefly, a Drinfeld module is an action of a finite extension of the ring $\mathbb{F}_q[t]$ on the additive group of a ring in characteristic p, via "additive polynomials":

$$x \mapsto c_0 x + c_1 x^p + \cdots + c_n x^{p^n}.$$

These act like abelian varieties in many ways (e.g., producing Galois representations), but have much simpler moduli and so are useful for things like proving the Langlands correspondence for GL_2 over function fields (Drinfeld's original application).

You can speak of the "invariant differentials" of F, i.e., the elements ω of the module of formal differentials (i.e., the free $R[\![X]\!]$ over $x_1, \dots, x_n$) such that $\omega(F(X, Y)) = \omega(X) + \omega(Y)$, and $\omega(a_F(X)) = i(a)\omega(X)$ for $a \in \mathfrak{o}$. These form a free R-module of rank n, called $\omega(F)$; in fact, the "quotient mod degree 2" map from invariant differentials to $R\,dx_1 \oplus \cdots \oplus R\,dx_n$ is a bijection, and all invariant differentials are closed [GH2, Proposition 2.2].

Policy: I'm now going to assume dimension 1 forever after, because Lubin-Tate theory applies only in dimension 1. Also, I may skip the $\mathfrak{o}$-module generalizations of some statements about formal groups, but those are all straightforward to extend (or see [GH2]).

Logarithms

If $f : F \to \mathbb{G}_a$ is a homomorphism of formal $\mathfrak{o}$-modules, you can take its formal derivative

$$\omega = df(x) = \frac{df}{dx} \cdot dx.$$

That gives a homomorphism $d : \mathrm{Hom}(F, \mathbb{G}_a) \to \omega(F)$. By [GH2, Proposition 3.2], if R is flat (i.e., torsion-free) over $\mathfrak{o}$, then d is injective; if R is a K-algebra, then d is bijective. In particular, in the latter case, there is a unique isomorphism $f : F \to \mathbb{G}_a$ with df equal to any prescribed generator of $\omega(F)$. We call f a *logarithm* for F.

Height

It's an easy lemma [GH2, Lemma 4.1; beware that f is used to mean two different things in the same sentence!] that if R is a field and F is a formal group with $i(\pi) = 0$, then either $\pi_F = 0$ or there is an integer h such that

$$\pi_F(x) = f(x^{q^h})$$

for some series f with $f'(0) \neq 0$. In the second case, we say F has *height* h. If R is a complete local ring whose maximal ideal I contains $i(\pi)$ (hereafter a *local* $\mathfrak{o}$*-algebra*), we say F has height h if the reduction of F has height h over R/I. (Define height for a formal $\mathfrak{o}$-module as the height of the underlying formal group. Oh, and this definition doesn't depend on dimension 1.)

Deformations

If we fix a formal $\mathfrak{o}$-module F_0 of some height h over R/I, I will refer to a formal $\mathfrak{o}$-module F over R equipped with an isomorphism of its reduction to F_0 as a *deformation of* F_0 *over* R. Then the Lubin-Tate-Drinfeld theorem explicitly describes a universal deformation of F_0. This comes from the following fact: if R is flat over $\mathfrak{o}$, then any formal $\mathfrak{o}$-module (of dimension 1, as always here) can be presented so that its formal logarithm takes the form

$$f(x) = x + \sum_{k=1}^{\infty} b_k x^{q^k}$$

for $b_k \in R \otimes K$, and this presentation is unique. A formal $\mathfrak{o}$-module presented this way is said to be $\mathfrak{o}$*-typical.* (See [GH2, §5].)

Example: for $\mathbb{G}_m$ over $\mathbb{Z}_p$, $b_k = p^{-k}$ and you get the formal logarithm of the Artin-Hasse exponential.

As the previous example shows, it's better to work with a certain change of variable here. Keeping R flat, define $v_1, v_2, \dots$ by

$$\pi b_k = v_k + b_1 v_{k-1}^q + \cdots + b_{k-1} v_1^{q^{k-1}};$$

then the v_k turn out to be integral. In fact, unwinding the construction yields a "universal formal $\mathfrak{o}$-typical module" over the infinite polynomial ring $\mathfrak{o}[v_1, v_2, \dots]$.

It turns out that you can read off heights easily here: the $\mathfrak{o}$-typical module constructed above has height h if and only if $v_1, \dots, v_{h-1}$ vanish in R/I and v_h does not.

The Lubin-Tate-Drinfeld theorem (Lubin-Tate for formal groups, Drinfeld for formal $\mathfrak{o}$-modules) now asserts that if you pull back the universal formal $\mathfrak{o}$-typical module to $\mathfrak{o}[\![u_1, \dots, u_{h-1}]\!]$ via

$$\begin{aligned} v_i &\mapsto u_i \qquad (i = 1, \dots, h-1) \\ v_h &\mapsto 1 \\ v_i &\mapsto 0 \qquad (i \geq h+1), \end{aligned}$$

and call the result F, then F is a universal deformation of its reduction modulo $(\pi, u_1, \ldots, u_{h-1})$ (which is thus defined over $\mathbb{F}_q$, and which has height h). See [GH2, Proposition 12.10] for the (easy) cohomological computation that verifies this.

Chapter 22

Periods on the Lubin-Tate moduli space

To keep things moving, I'm going to be terse now (as in [GH1]); there are lots of details filled in in [GH2].

References: Jay notes that [GH1] (the terse one) is available online; I'll put a link on the notes page.

Corrections from last time

Thanks to Jay for these.

page 2: the first Lubin-Tate paper only deals with height 1 (that is all that is needed for local class field theory), and they show uniqueness only over the completion of the maximal unramified extension of K. (Indeed, the fact that they are not isomorphic is key to being able to use them for explicit local reciprocity!)

page 3, top of page: "free $R[\![X]\!]$ over $x_1, \dots, x_n$" should be over $dx_1, \dots, dx_n$. By $\omega(F(X,Y))$, I mean you write $\omega(X) = f_1(X)dx_1 + \cdots + f_n dx_n$, then you write $F(X,Y) = (F_1(X,Y), \dots, F_n(X,Y))$, you plug in

$$\omega(F(X,Y)) = f_1(F(X,Y))dF_1(X,Y) + ... + f_n(F(X,Y))dF_n(X,Y),$$

then expand each $dF_i(X,Y)$ by the chain rule.

page 3, Height: change "formal group" in the second line to "formal $\mathfrak{o}$-module" and scratch the reference to formal $\mathfrak{o}$-modules in the parenthetical.

A group action

Last time, we built a universal deformation F over $A = \mathfrak{o}[\![v_1, \dots, v_{h-1}]\!]$ of a formal $\mathfrak{o}$-module over $\mathbb{F}_q$ of height h, which I'll denote by F_0. That means that the group $G = \operatorname{Aut}(F_0)$ acts on the deformation space $\operatorname{Spf} A$, and on the corresponding rigid analytic space X, which is the open unit polydisc in $v_1, \dots, v_{h-1}$.

It turns out that $D = \operatorname{End}(F_0)$ is a division algebra of degree n, G is the group of units in some maximal order therein, and $D \otimes K \cong M_n(K)$ is split. That means G has a natural n-dimensional linear representation V_K over K, as does D^*. In particular, G acts on the hyperplanes of V_K, i.e., on $\mathbb{P}(V_K^\vee)$; the latter carries a rigid analytic space structure, and the group action is by analytic morphisms.

The crystalline period mapping, to be defined, is a rigid analytic G-equivariant étale morphism $\Phi : X \to \mathbb{P}(V_K^\vee)$ which classifies deformations "up to isogeny" as follows. For A an affinoid algebra over K, let F_a and F_b be deformations of F_0 over $\mathfrak{o}_A$, corresponding to points $a, b \in X(A)$. Then an isogeny of F_0, viewed as an element $T \in D^*$, deforms to an isogeny $F_a \to F_b$ if and only if $T\Phi(a) = \Phi(b)$.

The universal additive extension

In order to specify Φ, I have to give you a G-equivariant line bundle $\mathcal{L}$ on X and a $K[G]$-homomorphism $V_K \to H^0(X, \mathcal{L})$ whose image is basepoint-free (i.e., the images don't all vanish at a point). For $x \in X$, we then take $\Phi(x)$ to be the hyperplane of V_K which maps to sections of $\mathcal{L}$ vanishing at x.

First, $\mathcal{L}$ is the inverse of the analytification of the sheaf ω of invariant differentials, a/k/a the Lie algebra $\mathrm{Lie}(F)$. In order to make the map, I must consider the *universal additive extension* E of F; it sits in an exact sequence

$$0 \to N \to E \to F \to 0$$

with $N = \mathbb{G}_a \otimes \mathrm{Ext}(F, \mathbb{G}_a)^\vee$, and it is universal: if $0 \to N' \to E' \to F \to 0$ is an extension of F by an additive $\mathfrak{o}$-module, then there are unique homomorphisms $i : E \to E'$, $j : N \to N'$ such that

$$\begin{tikzcd}
0 \arrow[r] & N \arrow[r] \arrow[d, "j"] & E \arrow[r] \arrow[d, "i"] & F \arrow[r] \arrow[d, "\mathrm{id}_F"] & 0 \\
0 \arrow[r] & N' \arrow[r] & E \arrow[r] & F \arrow[r] & 0
\end{tikzcd}$$

commutes. (This is straightforward, modulo some cohomology arguments which have already been exploited in constructing the universal deformation, namely, that $\mathrm{Hom}(F, \mathbb{G}_a) = 0$ and $\mathrm{Ext}(F, \mathbb{G}_a)$ is free of rank $n-1$. See [GH2, Proposition 11.3].) On the level of Lie algebras, we have

$$0 \to \mathrm{Lie}(N) \to \mathrm{Lie}(E) \to \mathrm{Lie}(F) \to 0$$

and this sequence is G-equivariant.

The bundle $\mathrm{Lie}(E)$ turns out to be the covariant Dieudonné module of F_0, so it is an "F-crystal": it carries an integrable connection $\nabla : \mathrm{Lie}(E) \to \mathrm{Lie}(E) \otimes \Omega^1_{A/\mathfrak{o}}$ plus a "Frobenius structure". The latter can be viewed as an isomorphism $\sigma^* \mathrm{Lie}(E) \to \mathrm{Lie}(E)$ for any $\sigma : A \to A$ lifting the q-power map on the special fibre. Let $\mathcal{M}$ be the analytification of $\mathrm{Lie}(E)$, as a rigid vector bundle over X; then by "Dwork's trick", $\mathcal{M}$ admits a basis of horizontal sections over X. (The idea: by formal integration, you get a basis over a small polydisc. But then you use the Frobenius pullback to "grow" this polydisc.) If you prefer, this can all be described in formulas in this case: see [GH2, Section 22].

We now have our representation $V_K = H^0(X, \mathcal{M})^\nabla$ on the horizontal sections of $\mathcal{M}$: the surjection $\mathcal{M} \to \mathcal{L} \to 0$ gives the map $V_K \to H^0(X, \mathcal{L})$ whose image is basepoint-free. The verification that Φ is étale and detects isogenies can be found in [GH2, Section 23].

What is Dwork's trick?

This is worth explaining a bit more, because it also comes up all over the place in p-adic cohomology. Say you have a vector bundle $\mathcal{M}$ over the open unit polydisc X over K with coordinates $x_1, \dots, x_n$. (Note: what was $n-1$ before is n now for notational simplicity.) It turns out that $\mathcal{M}$ is in fact generated freely by global sections (I had this on a homework earlier; I later found it in an old paper by Gruson).

Say you also have an integrable connection $\nabla : \mathcal{M} \to \mathcal{M} \otimes \Omega^1$, i.e., commuting actions of $\partial_i = \frac{\partial}{\partial x_i}$ for $i = 1, \dots, n$. You can then try to formally solve for the horizontal sections around $x_1 = \cdots = x_n = 0$. The resulting sections will converge on some polydisc, but its radius may be much smaller than 1. E.g., if $n = 1$, the rank is 1, and $\partial_i \mathbf{v} = c\mathbf{v}$, the horizontal section is $\exp(-\int c)\mathbf{v}$, which converges on some disc but possibly a small one.

What having a Frobenius structure does is give you an isomorphism between $\mathcal{M}$ and its pullback along some map σ lifting the q-power Frobenius on the reduction of $\Gamma(\mathcal{O}, X)$, e.g., one taking K into itself and taking x_i to x_i^q. That pulls back your sections on a tiny polydisc to sections on a larger polydisc (in my example, sections on the disc $|x_i| \leq \rho$ pull back to $|x_i| \leq \rho^{1/q}$); but in fact the K-vector space of horizontal sections is unique, so these sections actually extend the ones you started with. Repeat *ad infinitum.*

Chapter 23

p-adic uniformization and Shimura curves

I'm going to wrap things up by, in a sense, coming full circle; we're going to discuss p-adic uniformization, which ties in closely to the Tate curve we constructed at the very beginning of the course.

References: I'm pulling a lot of this out of: J.-F. Boutot et H. Carayol, Uniformisation p-adique des courbes de Shimura: les théorèmes de Čerednik et de Drinfeld, *Astérisque* **196-197**. I will cite this as [BC]. Thanks to Samit for the reference. Also see [FvdP, Section 5.4] for more details on how to take quotients of $\mathbb{P}^1$ to make "Mumford curves".

Quaternion algebras

Can someone suggest a good reference for this?

A few reminders about quaternion algebras: a *quaternion algebra* over a field F is a central simple algebra of degree 2 over F, i.e., a division algebra with center F of dimension 4 as an F-vector space, i.e., an element of the Brauer group of F of order 2. (Whether you allow the exceptional case $M_2(F)$, the "split quaternion algebra", is a matter of convention.)

Over a complete discretely valued field, there is a unique "ring of integers" in a quaternion algebra: it is the set of elements whose norm (element times its conjugate, a/k/a the determinant of the multiplication-by-said-element map) has absolute value ≤ 1. We call this the *maximal order* of the quaternion algebra. Over a number field F, by a maximal order of a quaternion algebra D we mean an $\mathfrak{o}_F$-subalgebra $\mathfrak{o}_D$ with $\mathfrak{o}_D \otimes_{\mathfrak{o}_F} F = D$. There exists at least one such, and any two are conjugate.

We say a quaternion algebra D over a number field *splits* at a place v (a completion, either archimedean or corresponding to a finite prime) if when you complete at v, you get the split quaternion algebra rather than an honest one. By class field theory, the number of nonsplit places is always even.

The Drinfeld upper half-plane

Throughout this lecture and the next, let K be a *discretely valued* complete nonarchimedean field with *finite* residue field $k = \mathbb{F}_q$. Let $\mathbb{C}$ be the completed algebraic closure of K. Let π be a uniformizer of $\mathfrak{o}$. See [BC, Part I] for details.

Define the *Drinfeld upper half-plane* Ω to be the analytic subspace of $\mathbb{P}^1_{\mathbb{C}}$ obtained by removing the K-rational points; those form a closed subset in the metric topology, so their complement is admissible and this really makes sense as a rigid analytic space.

Note that the group $\mathrm{PGL}_2(K)$ acts on Ω. The action is not free, because you can have fixed points over quadratic extensions of K.

One may think of $\mathbb{P}^1(\mathbb{C})$ as the set of $\mathbb{C}$-homothety classes of K-linear maps $K^2 \to \mathbb{C}$; Ω corresponds to those classes represented by injective maps. So you may think of Ω as classifying "rank two K-lattices inside $\mathbb{C}$", much as the complex upper half plane classifies rank two $\mathbb{Z}$-lattices inside $\mathbb{C}$.

There is also a natural formal scheme $\widehat{\Omega}$ with generic fibre Ω; it is most easily described in terms of the *building* associated to $\mathrm{PGL}_2(K)$, which is the following graph. Consider the set of homothety classes of $\mathfrak{o}$-lattices in K^2 as the vertices; join two classes by an edge if you can find representing lattices L_1, L_2 with $\pi L_1 \subset L_2 \subset L_1$. The edges out of each vertex can be identified with $\mathbb{P}^1_{\mathbb{F}_q}$; if you imagine the graph as a CW-complex, its points correspond to homothety classes of norms on K^2. (You should be thinking Berkovich here...) Namely, if $\mathbf{v}_1, \mathbf{v}_2$ is a basis, and I take a point on the edge joining $\mathfrak{o}\mathbf{v}_1 + \mathfrak{o}\mathbf{v}_2$ with $\mathfrak{o}\mathbf{v}_1 + \pi\mathfrak{o}\mathbf{v}_2$, its corresponding norm is

$$\|a_1\mathbf{v}_1 + a_2\mathbf{v}_2\| = \max\{|a_1|, q^t|a_2|\}$$

with $t \in [0, 1]$.

Note that one has a natural map from Ω to the building: given an element of Ω describing an injective map from K^2 to $\mathbb{C}$, compose with the norm on $\mathbb{C}$ to get a norm on K^2. Call this map λ.

Each vertex v of the building gives us a way to identify $\mathbb{P}^1_K$ with the generic fibre of a projective line over $\mathfrak{o}$; call the latter $\mathbb{P}_v$. Let $\widehat{\Omega}_v$ be the formal completion, along the special fibre, of the complement in $\mathbb{P}_v$ of the rational points on the special fibre; the corresponding rigid space is $\lambda^{-1}(v)$.

If vw is an edge corresponding to a pair of lattices L_1, L_2 with $\pi L_1 \subset L_2 \subset L_1$, let $\mathbb{P}_{vw}$ be the blowup of $\mathbb{P}_v$ at the point on its special fibre defined by L_2; we have a canonical identification $\mathbb{P}_{vw} \cong \mathbb{P}_{vw}$. Let $\widehat{\Omega}_{vw}$ be the formal completion, along the special fibre, of the complement in $\mathbb{P}_{vw}$ of the *nonsingular* rational points on the special fibre (so leave in the crossing point); the corresponding rigid space is $\lambda^{-1}(vw)$. We can now paste the $\widehat{\Omega}_v$ and $\widehat{\Omega}_{vw}$ together to get a formal scheme $\widehat{\Omega}$ with generic fibre Ω, and $\mathrm{PGL}_2(K)$ acts on it.

There are various ways to interpret $\widehat{\Omega}$ as a "moduli space" (i.e., it represents some natural functors); see [BC, I.4 and I.5].

Aside: I won't discuss it further here, but there is also a Drinfeld upper half-space of any dimension n. You get it by taking $\mathbb{P}^{n+1}(\mathbb{C})$ and removing all K-rational hyperplanes (not just K-rational points!). This carries an action of $\mathrm{PGL}_{n+1}(K)$; I think you can also analogize to other linear groups, but I don't know where anyone has done that.

Drinfeld's half-plane and formal $\mathfrak{o}_D$-modules

Let D be a quaternion algebra with center K. Fix a choice of a quadratic extension K' of K contained in D, and let $\mathfrak{o}'$ be its ring of integers.

We are interested in *special formal $\mathfrak{o}_D$-modules* over a $\mathfrak{o}$-algebra B; such a thing is a formal $\mathfrak{o}$-module X of dimension 2 (in the sense of last time) plus an action $i : \mathfrak{o}_D \to \mathrm{End}(X)$ compatible with the $\mathfrak{o}$-action, such that $\mathrm{Lie}(X)$ becomes a free $B \otimes_{\mathfrak{o}} \mathfrak{o}'$-module of rank 1. In particular, we are going to insist that these also have height 4 (I think this is as a $\mathfrak{o}$-algebra, but I'm not positive).

Anyway, Drinfeld showed that there exists a universal deformation of such a formal module over $\mathbb{F}_q$ plus some extra data (a "quasi-isogeny of height 0" on the reduction), which lives on $\widehat{\Omega}\widehat{\otimes}_{\mathfrak{o}}\widehat{\mathfrak{o}^{\mathrm{unr}}}$. This is kind of a long story, and it's in the same spirit as the discussion of Lubin-Tate formal groups from last time (even though the formal groups are all two-dimensional, the extra endomorphisms make the situation look like the Lubin-Tate case) so I'm not going to discuss it further here. See [BC, Section 2] for all the gory details.

Shimura curves and the upper half plane

Things like modular curves are obtained complex analytically by taking the complex upper half plane and quotienting by the action of a discrete group, like $\mathrm{PGL}_2(\mathbb{Z})$. If you take a group commensurable with $\mathrm{PGL}_2(\mathbb{Z})$ (a *congruence subgroup*), you get a classical modular curve; in that case, you have to compactify by adding in the orbits of $\mathbb{P}^1_{\mathbb{Q}}$ (the *cusps*). However, Shimura noticed you can also take quotients by things like the unit group of a maximal order in a quaternion algebra (the maximal order is , and get algebraic curves over $\mathbb{Q}$ (not just over $\mathbb{C}$). Moreover, like the modular curves, which are moduli spaces for elliptic curves plus some extra structure, Shimura's curves are moduli spaces for certain two-dimensional abelian varieties with extra endomorphisms by that maximal order in the quaternion algebra (so-called "false elliptic curves").

Let me make this more precise before proceeding to the rigid analogue of this description. Let Δ be an indefinite quaternion algebra with center $\mathbb{Q}$. ("Indefinite" means that $\Delta \otimes_{\mathbb{Q}} \mathbb{R}$ is split, so is congruent to $M_2(\mathbb{R})$ rather than the Hamilton quaternions.) For R a commutative $\mathbb{Q}$-algebra, let $\Delta^*(R)$ denote the group of units in the noncommutative ring $\Delta \otimes_{\mathbb{Q}} R$. Let $\mathbb{A}_f = \mathbb{Z} \otimes_{\mathbb{Q}} \widehat{\mathbb{Z}}$ denote the ring of finite adèles (where $\widehat{\mathbb{Z}} = \prod_p \mathbb{Z}_p$ is the profinite completion of $\mathbb{Z}$), and let $\mathcal{H}^{\pm} = \mathbb{P}^1(\mathbb{C}) \setminus \mathbb{P}^1(\mathbb{R})$ be the upper and lower half planes in $\mathbb{C}$. Then Shimura's observation is that for any open compact subgroup U of $\Delta^*(\mathbb{A}_f)$, the (left) quotient

$$\Delta^*(\mathbb{Q})\backslash[\mathcal{H}^{\pm} \times \Delta^*(\mathbb{A}_f)/U]$$

can be naturally identified with the $\mathbb{C}$-points of a *projective* algebraic curve S_U over $\mathbb{Q}$ (look Ma, no cusps!), and that the curve is a coarse moduli space for a moduli problem concerning two-dimensional abelian varieties with endomorphisms by $\mathfrak{o}_\Delta$ and "level U structure". (And of course if U is "sufficiently small", the moduli space is even fine.)

Shimura curves and the rigid upper half plane

Let p be a prime where Δ does not split, so that $\Delta_p = \Delta \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is an honest quaternion algebra over $\mathbb{Q}_p$. Then Δ_p^* has a unique maximal compact subgroup U_p^0, namely the units of the maximal order of Δ_p, and there is a decreasing sequence of compact subgroups U_p^n given by units congruent to 1 modulo p^n.

Suppose my level structure U looks like U_p^n times some compact open subgroup U^p of the prime-to-p part of $\Delta^*(\mathbb{A}_f)$. The theorem of Cherednik-Drinfeld (you can spell that first one "Čerednik" if you prefer; you can spell that second one "Drinfel'd" if you prefer too) says that if you form the rigid analytic quotient

$$\mathrm{GL}_2(\mathbb{Q}_p)\backslash[(\widehat{\Omega}\widehat{\otimes}\widehat{\mathbb{Z}_p^{\mathrm{unr}}}) \times Z_U],$$

you get the analytification of S_U again! Here

$$Z_U = U^p\backslash\overline{\Delta}^*(\mathbb{A}_f)/\overline{\Delta}^*(\mathbb{Q})$$

is a double coset space whose quotient by $\overline{\Delta}^*(\mathbb{Q}_p)$ is finite.

But what the heck is $\overline{\Delta}$? I'll come back to that in a moment. First, let me point out that (if I read [BC] correctly) Cherednik proved this originally by some not-so-enlightening argument; Drinfeld's contribution was to match up the moduli interpretation of the Shimura curve with the moduli interpretation of $\widehat{\Omega}$, giving a much more conceptual proof of the theorem in the bargain.

But what the heck is $\overline{\Delta}$?

Glad you asked. $\overline{\Delta}$ is the quaternion algebra obtained from Δ by switching the invariants at p and ∞; i.e., $\overline{\Delta}$ looks like Δ at all finite primes other than p, but it splits at p and is definite (nonsplit in the real place).

Why does that come up? By Tate-Honda (see [BC, Proposition III.2]):

- there is a unique isogeny class of two-dimensional abelian varieties over $\overline{\mathbb{F}_p}$ equipped with an action of $\mathfrak{o}_\Delta$;
- any such abelian variety A is isogenous to the product of two supersingular elliptic curves;
- the algebra $\mathrm{End}_{\mathfrak{o}_\Delta}(A) \otimes \mathbb{Q}$ is isomorphic to $\overline{\Delta}$.

The fact that there is this "switcheroo" between the quaternion algebra which acts on the abelian surfaces and the quaternion algebra you use in forming the rigid analytic quotient may seem like a trifle or even a nuisance, but it's actually a wonderful thing! It underlies certain "hidden symmetries" in the theory of automorphic forms, like the Jacquet-Langlands correspondence. A certain geometric realization of this correspondence is a crucial part of Ribet's proof of Serre's epsilon conjecture, and in particular that the modularity of elliptic curves implies Fermat's last theorem. (I have now told you every last thing I know about this. Go talk to David Helm for more information.)

That's all, folks!

Thanks for attending the course. Oh, and if you have corrections marked in your notes, I'd love to get some by email so I can assemble a more definitive compilation of the notes.