p-adic analysis, p-adic arithmetic*

Lecture 1

Norms

Definition 0.1. *Let K be a field. A* norm *on K is a function $|\cdot| : K \to \mathbb{R}_{\geq 0}$ satisfying*

1. $\forall x, y \in K, |x + y| \leq |x| + |y|$.
2. $\forall x, y \in K, |xy| = |x||y|$.
3. $\forall x \in K, |x| = 0 \Leftrightarrow x = 0$.

We will refer to the pair $(K, |\cdot|)$ as a normed field. Sometimes we will refer to K itself as a normed field, leaving $|\cdot|$ implicit.

Standard examples are the fields $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ with the usual (Euclidean) norm. Alternatively, any field has the trivial norm that takes any non-zero element to 1. We will suppose in the following that all our norms are in fact non-trivial.

Definition 0.2. *Fix a prime p. Let x be a non-zero rational number. We can write $x = p^n \frac{a}{b}$, where a, b are coprime to p. We then define*

$$\mathrm{ord}_p(x) = n, |x|_p = p^{-\mathrm{ord}_p(x)},$$

to be the p-adic valuation and norm of x respectively.

Definition 0.3. *We say that two norms $|\cdot|_1, |\cdot|_2$ on K are* equivalent *if there exists $\alpha > 0$ such that*

$$|\cdot|_1^\alpha = |\cdot|_2.$$

The following theorem will serve as a warm-up for using field norms.

Theorem 0.4. *Any norm $|\cdot|$ on $\mathbb{Q}$ is equivalent to either the usual norm or a p-adic norm $|\cdot|_p$.*

Recall that we are assuming all of our norms are non-trivial; otherwise the trivial norm would give a third possibility in the statement of the theorem above.

Proof. Let $a, b > 1$ be integers, and write b^n in base a as follows:

$$b^n = c_m a^m + c_{m-1} a^{m-1} + \cdots + c_0,$$

where for each i we have $0 \leq c_i \leq a - 1$.

Let $M = \sup(|1|, \ldots, |a - 1|)$. It follows that

$$|b^n| \leq |c_m||a^m| + \cdots + |c_0| \leq (m + 1)M \max(|a^m|, \ldots, 1)$$

$$\leq (n \log_a(b) + 1)M \sup(1, |a|^{n \log_a(b)}).$$

*These are the notes for a tutorial given by Jack A. Thorne at Harvard in the summer of 2010.

Taking the n^{th} root and letting $n \to \infty$ gives

$$|b| \leq \sup(1, |a|^{\log_a(b)}).$$

At this point we divide into two cases. Suppose first that there exists an integer b such that $|b| > 1$. It follows that $|a| > 1$ for any integer $a > 1$. Reversing the roles of a and b in the inequality above gives

$$|b| \leq |a|^{\log_a(b)}, |a| \leq^{\log_b(a)},$$

and hence

$$|b|^{1/\log(b)} \leq |a|^{1/\log(a)} \leq |b|^{1/\log(b)}.$$

Thus equality holds and $|a| = a^{\mu}$ for any integer $a > 1$. Thus $|\cdot|$ is equivalent to the standard Euclidean norm.

Now suppose instead that for all integers b, we have $|b| \leq 1$, hence $|b| < 1$ for some b (as otherwise $|\cdot|$ would be the trivial norm). It follows that there exists a prime p with $|p| < 1$ (take a prime factorization of b and use multiplicativity). We'll be done if we can show that for all other primes q, we have $|q| = 1$.

Suppose there exists q with $|q| < 1$. Then we can find n, m such that

$$|p^m| < \frac{1}{2}, |q^n| < \frac{1}{2}.$$

Since p, q are distinct primes, we can find integers x, y with $xp^n + yq^m = 1$. (For example, one could apply Euclid's algorithm). Then the triangle inequality gives

$$1 = |1| \leq |x||p^n| + |y||q^m| < 1.$$

This contradiction concludes the proof. □

The dichotomy in the proof above is important enough that we give it a name.

Definition 0.5. *Let K be a normed field. If $|\cdot|$ satisfies the following stronger* ultra-metric triangle inequality*:*

$$\forall x, y \in K, |x + y| \leq \sup(|x|, |y|)$$

then we say that K is non-Archimedean*. Otherwise, we say that K is* Archimedean*.*

Thus the usual norm on $\mathbb{Q}$ is Archimedean, whilst the p-adic norms are all non-Archimedean. More generally, one has the following.

Theorem 0.6. *Let K be an Archimedean normed field. Then there exists an embedding $\iota : K \to \mathbb{C}$ of normed fields (i.e. $|\cdot|_K$ is equivalent to the pullback of the usual absolute value from $\mathbb{C}$).*

The ultra-metric triangle inequality underlies many of the interesting differences between real and p-adic analysis.

Completeness

The following definition is hopefully familiar.

Definition 0.7. *Let K be a normed field. A* Cauchy sequence *in K is a sequence (x_n) such that for all $\epsilon > 0$, there exists $N > 0$ such that for all $n, m \geq N$, we have*

$$|x_n - x_m| < \epsilon.$$

We say that K is complete *if every Cauchy sequence has a limit in K.*

The rational numbers $\mathbb{Q}$ are not complete with respect to the Euclidean absolute value. One reason for passing to the completion (a.k.a. the real numbers $\mathbb{R}$) is so that the intermediate value theorem becomes true.

The following example shows that $\mathbb{Q}$ is not complete with respect to its p-adic absolute value and suggests, analogously, why a number theorist might be interested in passing to the p-adic completion of $\mathbb{Q}$.

Example (Teichmüller digits). Let $1 \leq a \leq p-1$ be an integer, and consider the sequence $(x_n) = (a^{p^n})$. This is a Cauchy sequence: for example, we have

$$x_{n+1} - x_n = a^{p^{n+1}} - a^{p^n} = a^{p^n}(a^{p^n(p-1)} - 1),$$

and the term in brackets is divisible by p^n, by the Fermat-Euler theorem. Hence

$$|x_{n+1} - x_n|_p < p^{-n}.$$

If the sequence (x_n) had a limit x in $\mathbb{Q}$, then we would have $x^p = x$ and $|x - a|_p < 1$. It follows that if $a \neq 1$ or $p-1$ then a is a non-trivial $(p-1)^{\text{st}}$ root of unity in $\mathbb{Q}$, a contradiction.

When we referred to 'passing to the completion' above, we were implicitly invoking the following theorem.

Theorem 0.8. *Let $(K, |\cdot|)$ be a normed field. Then there exists a field $\widehat{K}$ equipped with a norm, also denoted $|\cdot|$, and a map $\iota : K \to \widehat{K}$ such that*

1. *$\widehat{K}$ is complete.*
2. *ι is an isometry with dense image.*
3. *Any isometry $\phi : K \to E$ to a complete field factors uniquely through $\widehat{K}$.*

Moreover, any isometry of normed fields $\psi : K \to L$ extends uniquely to an isometry $\widehat{\psi} : \widehat{K} \to \widehat{L}$.

We leave the proof for next time. With this we can finally define the p-adic rational numbers.

Definition 0.9. *Let p be a prime. The* field of p-adic rationals*, denoted $\mathbb{Q}_p$, is by definition the completion of $\mathbb{Q}$ with respect to the p-adic norm.*

p-adic expansions

We now indicate one way to get a handle on p-adic numbers for the purpose of computation, inspired by the theory of decimal expansions for elements of $\mathbb{R}$.

Proposition 0.10. *Let $a \in \mathbb{Q}_p$. Then there exists a unique sequence of integers $0 \leq a_i \leq p-1$, $a_i = 0$ for i sufficiently negative such that*

$$a = \sum_{i>>-\infty}^{\infty} a_i p^i.$$

(To be more precise, the partial sums form a Cauchy sequence and a is the limit of this sequence). Thus for example if $|a|_p \leq 1$ then one can write

$$a = a_0 + a_1 p + a_2 p^2 + \dots$$

Proof. After multiplying by p^M, we can suppose that $|a| \leq 1$. We prove by induction that there are unique integers $0 \leq a_i \leq p-1$ for $i = 0, \dots, N$ such that

$$|a - \sum_{i=0}^{N} a_i p^i|_p < p^{-N}.$$

In fact, it's enough to show that for any $a \in \mathbb{Q}_p$, $|a|_p \leq 1$, there exists a unique integer $0 \leq b \leq p-1$ such that

$$|a - b|_p < 1.$$

To see this, we can suppose that $|a|_p = 1$. Choose a rational number c such that $|a - c|_p < 1$, and write $c = d/e$ in lowest terms. One sees that e is prime to p, hence there exist integers x, y such that

$$xe + yp = 1.$$

Hence

$$|a - xd|_p = |a - (1 - yp)c|_p \leq \sup(|a - c|_p, |ypc|_p) < 1.$$

We then subtract multiples of p from xd so that it lies in the range $0, \ldots, p-1$ and this gives the desired integer b. □

We refer to the output of the proposition as the 'p-adic expansion' of a.

Example. Extracting $\sqrt{6}$ by hand, $1/(1-p)$.

Exercises to lecture 1

1. Let (x_n) be a sequence in a non-Archimedean normed field K. Show that it is a Cauchy sequence if and only if $|x_{n+1} - x_n| \to 0$ as $n \to \infty$. (This has the useful corollary that a sum converges if and only if the individual terms tend to zero. In particular, there are no problems with conditional convergence).

2. Let K be a normed field. Show that it is non-Archimedean if and only if $|\mathbb{Z}|$ is bounded.

3. Find the p-adic expansion of $1/p!$ in $\mathbb{Q}_p$ for $p = 3, 5$.

4. In $\mathbb{Q}_p$, let $a = a_0 + a_1 p + \ldots$ be the p-adic expansion of a. What is the p-adic expansion of $-a$?

5. Show that an element of $\mathbb{Q}_p$ lies in $\mathbb{Q}$ if and only if its p-adic expansion is eventually periodic.

6. Let K be a field equipped with norms $|\cdot|_1, |\cdot|_2$. Show that these norms are equivalent if and only if they define the same topology on K, if and only if $|x|_1 < 1 \Rightarrow |x|_2 < 1$.

7. Let $|\cdot|_\infty$ denote the Euclidean norm on $\mathbb{Q}$. Show that for any $x \in \mathbb{Q}$, we have

$$\prod_{p \leq \infty} |x|_p = 1.$$

   (The validity of this formula is one reason for choosing the normalization of $|\cdot|_p$ that we did).

8. Let K be a field with norms $|\cdot|_1, \ldots, |\cdot|_n$ which are mutually inequivalent. Suppose $x_1, \ldots, x_n \in K$. Show that for every $\epsilon > 0$, there exists $x \in K$ such that $|x - x_i|_i < \epsilon$ for each $i = 1, \ldots, n$. (If you get stuck, look in the references for the Artin-Whaples approximation theorem).

   How is this statement related to the Chinese Remainder Theorem when $K = \mathbb{Q}$?

9. (For those who know the Baire category theorem): In the lecture we saw that $\mathbb{Q}$ is not p-adically complete, i.e. $\mathbb{Q} \neq \mathbb{Q}_p$. Show that in fact no countable normed field is complete (disregarding the case of the trivial norm).

Lecture 2

Completions

Last time we stated the following theorem.

Theorem 1.1. *Let $(K, |\cdot|)$ be a normed field. Then there exists a field $\widehat{K}$ equipped with a norm, also denoted $|\cdot|$, and a map $\iota : K \to \widehat{K}$ such that*

1. *$\widehat{K}$ is complete.*
2. *ι is an isometry with dense image.*
3. *Any isometry $\phi : K \to E$ to a complete field factors uniquely through $\widehat{K}$.*

Moreover, any isometry of normed fields $\psi : K \to L$ extends uniquely to an isometry $\widehat{\psi} : \widehat{K} \to \widehat{L}$.

Proof. We just show existence, the uniqueness properties being easy. Let $\mathcal{C}$ be the set of all Cauchy sequences in K. This is ring under component-wise addition and multiplication. Let $\mathcal{I} \subset \mathcal{C}$ be the set of sequences which tend to zero. $\mathcal{I}$ is in fact an ideal.

Moreover, the quotient ring $\mathcal{C}/\mathcal{I}$ is a field (as sequences not in $\mathcal{I}$ are eventually bounded away from 0, hence invertible in $\mathcal{C}/\mathcal{I}$). We define a norm on $\mathcal{C}$ by

$$|(x_n)| = \lim_{n\to\infty} |x_n|.$$

This descends to $\mathcal{C}/\mathcal{I}$ and the natural map $K \to \mathcal{C}/\mathcal{I}$ is an isometry with dense image. We take $\widehat{K} = \mathcal{C}/\mathcal{I}$. □

p-adic integers

Definition 1.2. *We write*

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p \text{ with } |x| \leq 1\} = \{x \in \mathbb{Q}_p \text{ with } \operatorname{ord}_p x \geq 0\}.$$

This is called the ring of p-adic integers. *(We are writing* ord_p *here for the natural extension of the function* ord_p *defined on* $\mathbb{Q}^\times$ *last lecture to* $\mathbb{Q}_p$*. This extension is given by* $-\log_p |\cdot|_p$*. We define formally* $\operatorname{ord}_p 0 = \infty$*).*

We note that this is indeed a ring. It is the closure of $\mathbb{Z}$ in $\mathbb{Q}_p$ (i.e. the set of elements of $\mathbb{Q}_p$ which are limits of Cauchy sequences contained in $\mathbb{Z}$). In terms of p-adic extensions $\sum_{i=-\infty}^{\infty} a_i p^i$, $\mathbb{Z}_p$ consists of those elements with $a_i = 0$ when $i < 0$. We now investigate the structure of $\mathbb{Z}_p$, beginning by introducing some basic p-adic functions.

Consider the following power series, with coefficients viewed as lying in $\mathbb{Q}_p$:

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$$

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots.$$

Suppose for the rest of this section that p is odd.

Lemma 1.3. $\log(1+x)$ *and* $\exp(x)$ *both converge in the disk* $|x| < 1$.

Proof. Only the claim about the exponential needs proof. We need to show that when $\operatorname{ord}_p x > 0$, $\operatorname{ord}_p x^n/n! \to \infty$ as $n \to \infty$. But we have

$$\operatorname{ord}_p n! = \sum_{i=0}^{\infty} \lfloor n/p^i \rfloor \leq \sum_{i=0}^{\infty} n/p^i = n/(p-1),$$

and hence

$$\operatorname{ord}_p x^n/n! \geq n(\operatorname{ord}_p x - \frac{1}{p-1}),$$

giving the result. □

Theorem 1.4. *We have a topological group isomorphism*

$$\mathbb{Z}_p^\times \cong \mathbb{Z}_p \times \mathbb{F}_p^\times.$$

Recall that $\mathbb{F}_p$ is the field with p elements.

Proof. We recall that in real analysis, the identities

$$\exp(\log(1+x)) = \log(1+x) \text{ and } \log(\exp(x)) = x$$

hold whenever these expressions make sense. It follows that the same identities hold for the formal power series written above, hence for the p-adic functions they define. (If you are unconvinced by this, see[exercise 21 to IV.1][Kob77].) Similarly, we have

$$\exp(x+y) = \exp(x)\exp(y), \log((1+x)(1+y)) = \log(1+x) + \log(1+y).$$

It follows that these functions induce isomorphisms

$$1 + p\mathbb{Z}_p \underset{\exp}{\overset{\log}{\rightleftarrows}} p\mathbb{Z}_p,$$

where the left hand side is a group under multiplication and the right hand side under addition.

Note that $p\mathbb{Z}_p = (p)$ is an ideal of $\mathbb{Z}_p$, and we have a ring homomorphism

$$\mathbb{F}_p = \mathbb{Z}/(p) \to \mathbb{Z}_p/(p).$$

In fact this is an isomorphism, and $1 + p\mathbb{Z}_p$ is just the kernel of the 'reduction modulo p' map

$$r : \mathbb{Z}_p^\times \to \mathbb{F}_p^\times.$$

The theorem will be proved if we can find a splitting of this homomorphism, i.e. a map

$$i : \mathbb{F}_p^\times \to \mathbb{Z}_p^\times$$

such that $r \circ i = 1$. This is given by the Teichmüller digits defined in the first lecture. Given an integer $a > 1$, let

$$[a] = \lim_{n\to\infty} a^{p^n}.$$

Exercise: this depends only on the residue class of a modulo p, so defines a homomorphism $\mathbb{F}_p^\times \to \mathbb{Z}_p^\times$ with the desired properties. □

Corollary 1.5. *We have* $\mathbb{Q}_p^\times \cong \mathbb{Z} \times \mathbb{Z}_p \times \mathbb{F}_p^\times$.

Proof. We have $\mathbb{Z}_p = \{|x| \leq 1\}$, hence $\mathbb{Z}_p^\times = \{|x| = 1\}$. Thus every element of $\mathbb{Q}_p^\times$ has a unique expression as a power of p times an element of $\mathbb{Z}_p^\times$. □

We now understand the algebraic structure of $\mathbb{Z}_p$ quite well. The corollary shows that the only non-trivial ideals of $\mathbb{Z}_p$ are the (p^n), n a positive integer. In particular, $\mathbb{Z}_p$ is a UFD with a unique prime element. Such rings are called discrete valuation rings - more on this soon.

The most important thing to come out of the proof is the reduction map $\mathbb{Z}_p \to \mathbb{F}_p$. A general principle in number theory is that if you can do something modulo p, then there should exist a 'p-adic interpolation' over $\mathbb{Z}_p$, that reduces back to the original construction when you apply the reduction map. One example for those who know about such things: mod p congruences between the coefficients of modular forms can often be interpolated to p-adic analytic families of modular forms.

Corollary 1.6. $\mathbb{Q}_p$ *has exactly 3 non-isomorphic quadratic extensions.*

Proof. By elementary Galois theory, the non-trivial quadratic extensions of a field K of characteristic zero correspond to non-trivial elements of $K^\times/(K^\times)^2$. Now apply the above corollary. □

Extensions of $\mathbb{Q}_p$

We take this as our cue to investigate some of the structures associated to field extensions of $\mathbb{Q}_p$.

Theorem 1.7. *Let $(K, |\cdot|)$ be a complete non-Archimedean normed field, and let L be a finite extension of degree d. Then there exists exactly one norm $|\cdot|_L$ on L extending $|\cdot|$, and L is complete with respect to this norm. For any $x \in L$, we have*

$$|x|_L = |\mathbb{N}_{L/K}(x)|^{1/d}.$$

We'll prove this when $K = \mathbb{Q}_p$.

Definition 1.8. *Let K be a non-Archimedean normed field, and let V be a K-vector space. A* norm *on V is a function $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ satisfying*

1. $\forall x, y \in V, \|x + y\| \leq \|x\| + \|y\|$.
2. $\forall \lambda \in K, x \in V, \|\lambda x\| = |\lambda|\|x\|$.
3. $\forall x \in V, \|x\| = 0 \Leftrightarrow x = 0$.

We say that two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are equivalent *if there exist $c, C > 0$ such that for all $x \in V$ we have*

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1.$$

Proposition 1.9. *Let K be a complete non-Archimedean normed field, V a finite-dimensional K-vector space. Then all norms on V are equivalent.*

Proof when $K = \mathbb{Q}_p$. Choose a basis $v_1, \ldots, v_d$ of V. We define the sup norm on V by

$$\|\lambda_1 v_1 + \cdots + \lambda_d v_d\|_{sup} = \sup_i |\lambda|.$$

We show that any other norm $\|\cdot\|$ on V is equivalent to $\|\cdot\|_{sup}$. We have

$$\|\lambda_1 v_1 + \cdots + \lambda_d v_d\| \leq (\sup_i \lambda_i)(\sup_i \|v_i\|),$$

so we can take $C = \sup_i \|v_i\|$ above. Now we just need to find c such that for all $x \in V$,

$$c\|x\|_{sup} \leq \|x\|.$$

Let $B = \{x \in V | \|x\|_{sup} = 1\}$. Then there exists $\epsilon > 0$ such that for every $x \in B$, we have $\|x\| \geq \epsilon$. Suppose for contradiction that there existed a sequence (x_n) in B with $\|x_n\| \to 0$ as $n \to \infty$.

B is sequentially compact with respect to $\|\cdot\|_{sup}$ (i.e. it satisfies the conclusion of the Bolzano-Weierstrass theorem), so after passing to a subsequence we may suppose that $x_n \to x \in B$ for some x. But then for every n we have

$$\|x\| \leq \sup(\|x - x_n\|, \|x_n\|) \leq \sup(C\|x - x_n\|_{sup}, \|x_n\|).$$

By hypothesis both terms on the right hand side tend to zero as $n \to \infty$, hence $x = 0$. This is a contradiction as $x \in B$.

The proposition follows on taking $c = 1/\epsilon$. □

We can now prove the uniqueness part of the theorem. Take $K = \mathbb{Q}_p$, and let $|\cdot|_1, |\cdot|_2$ be norms on L extending $|\cdot|_p$. Viewing L as a K-vector space and applying the proposition, we have $c, C > 0$ such that for all $x \in L, n \geq 0$,

$$c|x^n|_1 \leq |x^n|_2 \leq C|x^n|_2.$$

Taking the n^{th} root and letting $n \to \infty$ gives $|x|_1 = |x|_2$.

To see the form that this extension must take, we may suppose that L is Galois over $\mathbb{Q}_p$. Write $|\cdot|$ for the norm on L extending $|\cdot|_p$ on $\mathbb{Q}_p$. For any $\sigma \in \mathrm{Gal}(L/\mathbb{Q}_p)$, $|\cdot| \circ \sigma$ is another such norm, so they must be equal. It follows that for any $x \in L$, we have

$$|\mathbb{N}_{L/\mathbb{Q}_p}(x)| = \prod_{\sigma \in \mathrm{Gal}(L/\mathbb{Q}_p)} |\sigma(x)| = |x|^d.$$

We leave the existence part for the exercises. □

Exercises to lecture 2

1. Let $L/\mathbb{Q}_p$ be a Galois extension. In this situation we will always endow L with the norm defined above, making it a complete non-Archimedean normed field.

   Show that any $\sigma \in \mathrm{Gal}(L/\mathbb{Q}_p)$ is an isometry.

2. Find a Galois extension K of $\mathbb{Q}$ and a prime p such that $|\cdot|_p$ does not extend uniquely to K. (Hint: a quadratic extension will do).

3. Read the proof of the existence of norms on finite extensions of $\mathbb{Q}_p$ ([Kob77, §III.2, Theorem 11]).

4. In the lecture we used the estimate $\mathrm{ord}_p\, n! \leq n/(p-1)$. Show that in fact $\mathrm{ord}_p\, n! = (n - S_n)/(p-1)$, where S_n is the sum of the digits in the p-adic expansion of n.

5. Find the correct statements for the results of this lecture when $p = 2$. How many quadratic extensions does $\mathbb{Q}_2$ have? (If you get stuck then e.g. Serre's Course in Arithmetic has the relevant statements. But you should make sure you understand how to prove them using the methods given in the lecture).

6. Find generators for the quadratic extensions K of $\mathbb{Q}_5$. Compute $|K^\times|$ in each case. What do you notice?

7. Show that the set B defined in the proof of the proposition is indeed sequentially compact.

8. Consider the formal power series

   $$f(x) = \sum_{n=0}^{\infty} \binom{1/2}{n} x^n,$$

   where

   $$\binom{a}{n} = \frac{a(a-1)\dots(a-n+1)}{n!}.$$

   Show that it converges in the disk $D(0, 1^-) \subset \mathbb{Z}_p$. Compute $f(7/9)$ in $\mathbb{Q}_7$. (Hint: what power series identity, known to hold over $\mathbb{R}$, might be useful)?

9. Write $K = \mathbb{C}(x)$ for the field of rational functions in the variable x (thus every element is of the form f/g, where f, g are polynomials). If $f \in \mathbb{C}(x)$, define $\mathrm{ord}_0 f$ to be the order of vanishing of f at $x = 0$, and $|f| = 2^{-\mathrm{ord}_0 f}$. Show that this defines a norm on K. What is the completion of K with respect to this norm?

10. (A more algebraic construction of $\mathbb{Z}_p$) Show that $\mathbb{Z}_p$ is the inverse limit of the inverse system

    $$\dots \longrightarrow \mathbb{Z}/p^n\mathbb{Z} \longrightarrow \mathbb{Z}/p^{n-1}\mathbb{Z} \longrightarrow \dots \longrightarrow \mathbb{Z}/p\mathbb{Z}\,,$$

    the maps being the natural reduction maps. (This has a technical meaning in category theory, but in practical terms means that $\mathbb{Z}_p$ is isomorphic to the ring

    $$\{(x_n) \in \prod_n \mathbb{Z}/p^n\mathbb{Z} \text{ such that for all } m, x_m \bmod p^{m-1} = x_{m-1}\}.)$$

    Does the induced (product) topology agree with the norm topology?

Lecture 3

Valuation rings and ramification

Now that we have some more normed fields to play with, we define some more of the associated structures.

Definition 2.1. *Let K be a field. A* valuation *on K is a function $v : K^\times \to \mathbb{R}$ such that*

1. $\forall x, y \in K, v(x+y) \geq \min(v(x), v(y))$.
2. $\forall x, y \in K, v(xy) = v(x) + v(y)$.

We say that v is a discrete valuation *if $v(K^\times)$ is a discrete subgroup of $\mathbb{R}$.*

Thus giving a valuation is essentially equivalent to giving a non-Archimedean norm, under the rule $v(\cdot) = -\log|\cdot|$. When K is a finite extension of $\mathbb{Q}_p$, we will make the following convention: $v = v_K$ is a discrete valuation, so we may normalize it so that $v_K(K^\times) = \mathbb{Z}$. On the other hand, we will always assume that $|\cdot| = |\cdot|_K$ extends the usual norm $|\cdot|_p$ on $\mathbb{Q}_p$.

Definition 2.2. *Let K be a non-Archimedean normed field. The* valuation ring *of K is defined to be*

$$\mathcal{O}_K = \{x \in K | |x| \leq 1\}.$$

Its maximal ideal is

$$\mathfrak{m}_K = \{x \in K | |x| < 1\}.$$

Its residue field is

$$k_K = \mathcal{O}_K/\mathfrak{m}_K.$$

When v_K is discrete, $\mathfrak{m}_K$ is a principal ideal. Any element ϖ_K generating $\mathfrak{m}_K$ is called a uniformizer.

Definition 2.3. *Let L/K be a finite extension of degree d, where K is a finite extension of $\mathbb{Q}_P$. The index*

$$[v_L(L^\times) : v_L(K^\times)] = e_{L/K} = e$$

is called the ramification index *of the extension L/K. The degree*

$$[k_L : k_K] = f_{L/K} = f$$

is called the inertial degree *of the extension L/K.*

In order words, we have $v_K = ev_L$ on restriction to K.

Example. Quadratic extensions of $\mathbb{Q}_p$.

Proposition 2.4. *Let L/K be as above. Then $ef = d$.*

Proof. Since these quantities are all multiplicative in towers, we can assume that $K = \mathbb{Q}_p$. Then $\mathcal{O}_L$ is a finite free $\mathbb{Z}_p$-module of rank d. Let ϖ_L be a uniformizer of L. Then we have $(\varpi_L^e) = (p)$ as ideals of $\mathcal{O}_L$.

Now each quotient $(\varpi_L^i)/(\varpi_L^{i+1})$, $i = 0, \ldots, e-1$ is isomorphic to k_L, hence of cardinality $\#k_L = p^f$. On the other hand we have as $\mathbb{Z}_p$-modules

$$\mathcal{O}_L/(\varpi_L^e) = \mathcal{O}_L/p\mathcal{O}_L \cong \mathbb{Z}_p^d/p\mathbb{Z}_p^d,$$

which has cardinality p^d. It follows that $p^{ef} = p^d$ and hence $ef = d$. □

If $e = 1$ then we say that L/K is unramified, whereas if $f = 1$ then we say that L/K is totally ramified. The above proposition shows that these two possibilities are mutually exclusive unless $L = K$.

Finally we have the following, generalizing the p-adic expansions introduced earlier.

Proposition 2.5. *Let K be a finite extension of $\mathbb{Q}_p$, and let $\mathcal{S}$ be a set of representatives for k_K in $\mathcal{O}_K$ containing 0. Let ϖ_K be a uniformizer of K. Then every element $x \in K$ has a unique expression in the form*

$$\varpi^{v_K(x)} \sum_{i=0}^{\infty} a_i \varpi^i$$

with $a_i \in \mathcal{S}$.

Hensel's lemma

The following lemma is vital in much that follows. Let K be a complete non-Archimedean normed field.

Lemma 2.6. *Let $f(x) \in \mathcal{O}_K[x]$ be a monic polynomial, and suppose $x \in \mathcal{O}_K$ satisfies*

1. $|f(x)| < 1$;
2. $|f'(x)| = 1$.

Then there exists a unique $y \in \mathcal{O}_K$ such that $f(y) = 0$ and $|y - x| \leq |f(x)|$.

This has the following prototype.

Lemma 2.7. *Let $f(x) \in \mathbb{Z}_p[x]$ be a monic polynomial, and suppose $x \in \mathbb{Z}_p$ satisfies*

1. $f(x) \equiv 0 \mod p$;
2. $f'(x) \not\equiv 0 \mod p$.

Then there exists a unique $y \in \mathbb{Z}_p$ such that $f(y) = 0$ and $y \equiv x \mod p$.

Proof. Let us prove the second version. We construct the p-adic expansion of y, by showing inductively that for each n there is a unique set of integers $a_i, 0 \leq a_i \leq p-1$ such that if $y_n = a_0 + pa_1 + \dots$ then $f(y_n) \equiv 0 \mod p^{n+1}$.

For $n = 0$, we just take the first p-adic digit in the expansion of x.

For the induction step, suppose $a_0, \dots, a_n$ are given. We have

$$f(y_n + cp^{n+1}) = f(y_n) + f'(y_n)cp^{n+1} + (\text{terms divisible by } p^{n+2}).$$

We want this to vanish modulo p^{n+2}, and this happens if and only if

$$c \equiv \frac{f(y_n)}{p^{n+1}f'(y_n)} \mod p.$$

(Note that the right hand side is a p-adic integer by the induction hypothesis). We therefore take a_{n+1} to be the first p-adic digit of the right hand side above. □

We state without proof the following more general formulation.

Lemma 2.8. *Let $f(x) \in \mathcal{O}_K[x]$ be a monic polynomial, and suppose that $\overline{f} = f \mod \mathfrak{m}_K$ factors as $\overline{f}(x) = g_0(x)h_0(x)$, where g_0, h_0 are monic and relatively prime polynomials in $k_K[x]$. Then there exist unique monic polynomials $g, h \in \mathcal{O}_K[x]$ such that $f(x) = g(x)h(x)$ and $\overline{g} = g_0, \overline{h} = h_0$.*

Unramified extensions

Fix an algebraic closure $\overline{\mathbb{Q}}_p$. We will consider all extensions of $\mathbb{Q}_p$ as being contained in this algebraic closure.

Proposition 2.9. *Let K be a finite extension of $\mathbb{Q}_p$. Then for every integer $f \geq 1$, K has a unique unramified extension of degree f. Moreover, the composite of any two unramified extensions of K is still unramified.*

Proof. To construct such an extension, proceed as follows. Let l/k_K be an extension of k_K of degree f, and let $\overline{\alpha}$ be a primitive element. Let $\overline{g}$ be the minimal polynomial of α over k_K, and let g be any lift of f to $\mathcal{O}_K$.

Then Gauss' lemma implies that g is irreducible, hence defines an extension E of K of degree f. All of the roots of g in E in fact lie in $\mathcal{O}_E$, hence g has roots in k_E. Hence $f_{E/K} = f = [E : K]$, and E is an unramified extension.

Let us next show that the composite of two unramified extensions L, L' is unramified over K. This will follow if $L.L'$ is unramified over L, i.e. if $[L.L' : L] = [k_{L.L'} : k_L]$. Let $\overline{\alpha}$ be a primitive element for $k_{L'}/k_K$ and choose a lift $\mathcal{O}_{L'}$. Let h the minimal polynomial of α over L.

Now $\overline{h}$ is irreducible in $k_L[x]$, since any factorization would lift by Hensel's lemma to a factorization in $\mathcal{O}_L[x]$. The result now follows.

Finally, we show the uniqueness of the unramified extension of degree f. Suppose L, L' are two such. Then $L.L'$ is unramified over K, and has the same residue field as L, since a finite field has a unique extension of any given degree. Hence

$$[L.L' : K] = [L.L' : L][L : K]$$

so we see $L' \subset L$ and vice versa. This concludes the proof. □

Corollary 2.10. *The unramified extension of $\mathbb{Q}_p$ of degree f is the one obtained by adjoining the $p^f - 1$ roots of unity. It has Galois group $\mathbb{Z}/f\mathbb{Z}$ over $\mathbb{Q}_p$.*

Corollary 2.11. *Let L/K be a finite extension, K finite over $\mathbb{Q}_p$. Then L has a subfield L^{ur}, the maximal subextension of L unramified over K. Moreover, L/L^{ur} is totally ramified.*

Totally ramified extensions

In this section, K is a finite extension of $\mathbb{Q}_p$.

Definition 2.12. *Let $f(x) \in \mathcal{O}_K[x]$ be a monic polynomial:*

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n.$$

If for each $i = 1, \ldots, n$ we have $v(a_i) \geq 1$, and $v(a_n) = 1$, then we say that f is Eisenstein.

Lemma 2.13. *If f is Eisenstein then it is irreducible.*

Proof. Suppose we can factor $f = gh$ in $\mathcal{O}_K[x]$, with g, h monic polynomials. Then $\overline{f} = \overline{g}\overline{h}$, hence we have $\overline{g}(x) = x^a$, $\overline{h}(x) = x^b$ for some $a, b > 0$.

On the other hand, since $v(a_n) = 1$, one of g, h must have constant term a unit. This is a contradiction. □

Proposition 2.14. *Suppose that L/K is a totally ramified extension of degree e, and let $\varpi = \varpi_L$ be a uniformizer of L. Then the minimal polynomial of ϖ is Eisenstein.*

Conversely, if $f \in \mathcal{O}_K[x]$ is an Eisenstein polynomial of degree e, then for any root α of f, $K(\alpha)/K$ is totally ramified of degree e, and α is a uniformizer of $K(\alpha)$.

Proof. Let $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$ be the minimal polynomial of ϖ. The coefficients of f are symmetric polynomials in the conjugates of ϖ, hence have positive valuation. Also, $a_n = \mathbb{N}_{L/K}\varpi$ and hence $|a_n| = |\varpi|^e$, or alternatively $v_K(a_n) = v_L(\varpi) = 1$. It follows that f is Eisenstein.

Suppose conversely that $f \in \mathcal{O}_K[x]$ is Eisenstein, and let α be a root. Let $L = K(\alpha)$. Then as above we have $v_L(\varpi) = v_K(a_n) = 1$, hence $e_{L/K} = e$. The result follows. □

Exercises to lecture 3

1. Let $f(x) = X^p - X - 1 \in \mathbb{Z}_p[x]$. Show that f is irreducible. Let K be the splitting field of f. What is $[K : \mathbb{Q}_p]$? Is this extension ramified?

2. Let $K = \mathbb{Q}_p(\zeta_p)$, where ζ_p is a primitive p^{th} root of unity. Show that $[K : \mathbb{Q}_p] = p - 1$, and that this extension is totally ramified. Give a uniformizer of K. What about $\mathbb{Q}_p(\zeta_{p^n})$?

3. Earlier we constructed a canonical splitting of the homomorphism $\mathbb{Z}_p^\times \to \mathbb{F}_p^\times$. Construct a similar splitting for K, a finite extension of $\mathbb{Q}_p$. (Hint: Hensel's lemma).

4. Continuing with this theme, we saw $\mathbb{Z}_p^\times \cong \mathbb{Z}_p \times \mathbb{F}_p^\times$ when p is odd. Let K be a finite extension of $\mathbb{Q}_p$, and let $e = e_{K/\mathbb{Q}_p}$. Give a sufficient condition on e to have $\mathcal{O}_K^\times \cong \mathcal{O}_K \times k_K^\times$, and show that it is sharp.

5. Which quadratic extensions of $\mathbb{Q}_2$ are ramified?

6. Let L/K be finite extensions of K, and let $\alpha \in \mathcal{O}_L^\times$ be a primitive element. Let f be its minimal polynomial, g the minimal polynomial of $\overline{\alpha} \in k_L$, and $e = e_{L/K}$. Show that $\overline{f} = g^e$.

7. Let L/K be a totally ramified extension, and let ϖ be a uniformizer of L. Show that $\mathcal{O}_L = \mathcal{O}_K[\varpi]$.

Lecture 4

Worked examples

Example. Let $f(x) = x^n - p \in \mathbb{Z}_p[x]$, and suppose that $p \nmid n$. Let L be the splitting field of f over $\mathbb{Q}_p$.

Then L contains the subfield $K = \mathbb{Q}_p(\zeta_n)$ (we use ζ_r to denote a primitive r^{th} root of unity). As we've seen, K is unramified over $\mathbb{Q}_p$. Moreover, $f(x) \in \mathcal{O}_K[x]$ is an Eisenstein polynomial. Thus we have $e_{L/\mathbb{Q}_p} = n$ and $f_{L/\mathbb{Q}_p} = f_{K/\mathbb{Q}_p}$.

Example (p-cyclotomic extensions). Let $f(x) = (x^p - 1)/(x - 1) = 1 + x + \cdots + x^{p-1}$. The roots of f are the primitive p^{th} roots of unity, and we write $K = \mathbb{Q}_p(\zeta_p)$ for its splitting field.

Let $x = y + 1$. Then we have

$$f(x) = f(y+1) = \frac{(y+1)^p - 1}{y} = p + \binom{p}{2} y + \cdots + \binom{p}{p-2} y^{p-2} + y^{p-1}.$$

Thus $f(y+1)$ is an Eisenstein polynomial in the variable y. It follows from the results of the previous lecture that K is a totally ramified, Galois extension of $\mathbb{Q}_p$ of degree $p-1$, and a uniformiser is $1 - \zeta_p$.

What about the field $L = \mathbb{Q}_p(\zeta_{p^r})$, when $r > 1$? Then we take $h(x) = (x^{p^r} - 1)/(x^{p^{r-1}} - 1)$. Thus the roots of h are precisely the primitive $p^{r^{\text{th}}}$ roots of unity, and L is the splitting field of h.

In fact $h(y+1)$ is also Eisenstein. It's easy to see that h has constant term equal to p, and we have

$$h(x) \equiv \frac{(x-1)^{p^r}}{(x-1)^{p^{r-1}}} \equiv (x-1)^{(p-1)p^{r-1}} \bmod p,$$

and hence $h(y+1) \equiv y^{(p-1)p^{r-1}} \bmod p$, so that all of the non-leading terms of $h(y)$ are divisible by p. Thus L is also totally ramified and Galois over $\mathbb{Q}_p$, of degree $(p-1)p^{r-1}$. A uniformiser is $\lambda = 1 - \zeta_{p^r}$. Its Galois group H is isomorphic to $(\mathbb{Z}/p^r\mathbb{Z})^\times$. We note that this group has a natural decreasing filtration by the subgroups

$$H_n = \{x \in H \text{ such that } x \equiv 1 \mod p^n\}, n = 0, \ldots, r.$$

In fact, H_n is equal to the set of Galois automorphisms which induce the trivial action on $\mathcal{O}_L/(\lambda^{n+1})$.

This is a particular example of the so-called ramification filtration: given any Galois extension L/K of p-adic fields, there is a canonical filtration of $G = \mathrm{Gal}(L/K)$ by normal subgroups G_r, defined as above. One can show that each G_r/G_{r+1} is abelian. In particular, G is always soluble. One consequence of this is that the roots of any polynomial $f \in \mathbb{Q}_p[x]$ are expressible in radicals!

The p-adic complex numbers

We recall the following.

Proposition 3.1. *Let $\overline{\mathbb{Q}}_p$ be an algebraic closure of $\mathbb{Q}_p$. Then $|\cdot|_p$ extends uniquely to a norm $|\cdot|$ on $\overline{\mathbb{Q}}_p$.*

Theorem 3.2. *$\overline{\mathbb{Q}}_p$ is not complete.*

Definition 3.3. *We define*

$$\Omega = \widehat{\overline{\mathbb{Q}}_p},$$

and write $\mathcal{O}$ for the valuation ring of Ω.

Proof of theorem. Let $b_1 = 1$. Choose a sequence $b_n \in \overline{\mathbb{Q}}_p, n = 2, 3, \dots$ of roots of unity of order prime to p such that $b_{n-1} \in \mathbb{Q}_p(b_n)$ and

$$[\mathbb{Q}_p(b_n) : \mathbb{Q}_p(b_{n-1})] > n.$$

Put

$$c = \sum_n b_n p^n \in \Omega.$$

Suppose for contradiction that $c \in \overline{\mathbb{Q}}_p$, and let $t = [\mathbb{Q}_p(c) : \mathbb{Q}_p]$. Let

$$c_t = \sum_{n=0}^{t} b_n p^n.$$

Thus c_t is the t^{th} partial sum of the sequence defining c, and we have $|c - c_t| \leq p^{-t-1}$, by the ultrametric inequality. Let M be a Galois extension of $\mathbb{Q}_p$ containing c, c_t and b_t. For any $\sigma \in \mathrm{Gal}(M/\mathbb{Q}_p)$, we have

$$|\sigma c_t - \sigma c| \leq p^{-t-1}.$$

By construction, we can find $\sigma_1, \dots, \sigma_{t+1} \in \mathrm{Gal}(M/\mathbb{Q}_p(b_{t-1}))$ such that the $\sigma_j(b_t)$ are distinct.

We now have the inequality

$$|\sigma_i c_t - \sigma_j c_t| = |(\sigma_i b_t - \sigma_j b_t)p^t| \geq p^{-t} \text{ for } i \neq j.$$

(Since the extension generated by b_t is unramified over $\mathbb{Q}_p$, the $\sigma_j b_t$ have pairwise distinct images in the residue field, hence $|\sigma_i b_t - \sigma_j b_t| = 1$.) These two inequalities show that the $\sigma_j c$ are distinct, contradicting $t = [\mathbb{Q}_p(c) : \mathbb{Q}_p]$. This concludes the proof. □

Lemma 3.4 (Krasner's lemma)**.** *Let K be a complete non-Archimedean normed field, and let $f \in K[x]$ be a monic polynomial with roots $\alpha_1, \dots, \alpha_d$ in $\overline{K}$. Suppose $\beta \in \overline{K}$ satisfies*

$$|\beta - \alpha_1| < |\alpha_1 - \alpha_i|, i = 2, \dots, d.$$

Then $K(\alpha) \subset K(\beta)$.

Proof. Let $L = K(\beta), M = L(\alpha_1, \dots, \alpha_d)$. M/L is Galois, and for any $\sigma \in \mathrm{Gal}(M/L)$ one has

$$|\beta - \alpha_1| = |\sigma(\beta - \alpha_1)| = |\beta - \sigma(\alpha_1)|.$$

(Recall that the norm is always Galois invariant). Then

$$|\alpha_1 - \sigma(\alpha_1)| \leq \sup(|\alpha_1 - \beta|, |\beta - \sigma(\alpha_1)| = |\alpha_1 - \beta| < |\alpha_1 - \alpha_i|, i = 2, \dots, d.$$

We must therefore have $\sigma(\alpha_1) = \alpha_1$. σ was arbitrary, so it follows that $\alpha \in L$, as required. □

Proposition 3.5. *Ω is algebraically closed.*

Proof. Let $\alpha \in \overline{\Omega}$, and let $f \in \Omega[x]$ be its minimal polynomial. After scaling α we can suppose that $f \in \mathcal{O}[x]$. Write

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_n,$$

and let $C = \min_i(|\alpha - \alpha_i|)$, where $\alpha = \alpha_1, \dots, \alpha_n$ are the roots of f. Choose a polynomial $g = x^n + b_1 x^{n-1} + \dots + b_n \in \mathcal{O}_{\overline{\mathbb{Q}}_p}[x]$ such that for each i, we have $|a_i - b_i| < C^n$. Then if $\beta_1, \dots, \beta_n$ are the roots of of g, we have

$$\prod_i |\alpha - \beta_i| = |g(\alpha)| = |g(\alpha) - f(\alpha)| \leq \sup_i |a_i - b_i| < C^n.$$

In particular, we must have $|\alpha - \beta_i| < C$ for some i. It follows by Krasner's lemma that $\alpha \in \Omega(\beta_i) = \Omega$. □

It follows that Ω is a suitable domain in which to do p-adic analysis. For this reason we sometimes refer to it as the field of *p-adic complex numbers*.

Here is another interesting result that uses Krasner's lemma.

Proposition 3.6. *Let $d \geq 1$ be an integer. Then there are only finitely many extensions of $\mathbb{Q}_p$ of degree bounded by d.*

Functions defined by power series

We begin with some notation.

Definition 3.7. *Given $a \in \Omega$, $r \in \mathbb{R}$, set*

$$D(a, r) = \{x \in \Omega | |x - a| \leq r\}$$

and

$$D(a, r^-) = \{x \in \Omega | |x - a| < r\}.$$

We call these respectively the closed *and* open disks *around a of radius r.*

If $a = 0$ we will abbreviate these as $D(r)$ and $D(r^-)$.

We will be considering functions defined by formal power series

$$f(X) = \sum_{n=0}^{\infty} a_n X^n, a_n \in \Omega.$$

Lemma 3.8. *Let*

$$r = r(f) = \frac{1}{\limsup_n |a_n|^{1/n}}.$$

Then $f(x)$ converges when $|x| < r$ and diverges when $|x| > r$.

Proof. Suppose that $|x| < r$ Thus we can write $|x| = (1 - \epsilon)r$ for some positive ϵ, and hence

$$|a_n x^n| = (r|a_n|^{1/n})^n (1 - \epsilon)^n.$$

Since $\limsup_n |a_n|^{1/n} = 1/r$, we see that $|a_n|^{1/n} \leq 1/(r - \epsilon r/2)$ for all sufficiently large n, and hence

$$\lim_{n\to\infty} |a_n x^n| \leq \lim_{n\to\infty} \left(\frac{(1 - \epsilon)r}{(1 - \epsilon/2)r}\right)^n = 0.$$

A similar argument in reverse shows that if $|x| > r$ then $a_n x^n \not\to 0$ as $n \to \infty$. □

Definition 3.9. *We call r the* radius of convergence *of the power series f.*

We have the following important lemma, whose proof is left as an exercise.

Lemma 3.10. *Suppose that $f(X)$ converges in the disk $D(r)$. Then it defines a continuous function $f : D(r) \to \Omega$.*

Proposition 3.11. *The radius of convergence of* $\exp(X)$ *is $p^{-1/(p-1)}$. The radius of convergence of* $\log(1 + X)$ *is 1.*

Proof. An earlier exercise was to show that

$$\mathrm{ord}_p(n!) = \frac{n - S_n}{p - 1},$$

where S_n is the sum of the digits in the p-adic expansion of n. We compute

$$\lim_{n\to\infty} \frac{\mathrm{ord}_p\, n!}{n} = \lim_{n\to\infty} \frac{n - S_n}{n(p - 1)} = 1/(p - 1).$$

The result for log is immediate. □

Exercises to lecture 4

1. In the first example above, give a formula for $f_{K/\mathbb{Q}_p}$.
2. Let K be a finite extension of $\mathbb{Q}_p$. Show that there exists a subfield $F \subset K$ such that $[F : \mathbb{Q}] = [K : \mathbb{Q}_p]$ and F is dense in K.
3. Show that the ring of integers in $\mathbb{Q}_p(\zeta_{p^r})$ is equal to $\mathbb{Z}_p[\zeta_{p^r}]$.
4. Let $\lambda = 1-\zeta_p \in K = \mathbb{Q}_p(\zeta_p)$. Show that if $\sigma \in \mathrm{Gal}(K/\mathbb{Q}_p)$ satisfies $\sigma(\zeta_p) = \zeta_p^a$, then $\sigma(\lambda) \equiv a\lambda \bmod \lambda^2$.
5. What is $|\Omega|$?
6. Give another proof of Proposition 3.5 using the following more general formulation of Hensel's lemma:

   Lemma 3.12. *Let K be a complete non-Archimedean normed field, and let $f \in \mathcal{O}_K[x]$ be a monic polynomial. Suppose there exists $\alpha \in \mathcal{O}_K$ such that $|f(\alpha)| < |f'(\alpha)|^2$. Then there exists a unique $y \in \mathcal{O}_K$ such that $f(y) = 0$ and $|y - \alpha| \leq |f(\alpha)|/|f'(\alpha)|$.*

7. Give a proof of Proposition 3.6.
8. Decide whether $\exp(X)$ and $\log(1 + X)$ converge on the boundaries of their respective disks of convergence.
9. Compute the radii of convergence of the following power series:
   (a) $\sum_n n!X^n$
   (b) $\sum_n (\zeta_p - 1)X^n/n!$
   (c) $\sum_n p^n X^{p^n}$.

Lecture 5

Last time we introduced the field Ω of p-adic complex numbers. In this lecture we can finally start introducing some analytic tools that are completely special to the p-adic situation. Since we'll be working in Ω mostly from now on, we introduce the following valuation: $\mathrm{ord}_p : \Omega^\times \to \mathbb{R}$ is the valuation extending the usual one on $\mathbb{Q}_p$. It is given by the formula $\mathrm{ord}_p = -\log_p |\cdot|$. (This brings us into line with the notation used in the course text).

Newton polygons

Let K be a subfield of Ω, and let $f(x) = 1 + a_1 x + \cdots + a_n x^n \in K[x]$.

Definition 4.1. *The* Newton polygon $N(f)$ *of f is the lower convex hull of the points*

$$(0, 0), (1, \mathrm{ord}_p\, a_1), \ldots, (n, \mathrm{ord}_p\, a_n).$$

That is, it's the highest polygonal line such that all the points $(i, \mathrm{ord}_p\, a_i)$ lie on or above it.

See the course text for some pictures. Intuitively, it can be constructed as follows: in the (x, y)-plane, imagine the points $(i, \mathrm{ord}_p\, a_i)$ as nails sticking outwards, and the negative y-axis as a piece of string with one end fixed at the origin, and the other end free.

We rotate the string counter-clockwise until it meets one of the nails. As we continue rotating, the segment of the string between this the origin and this point will be fixed. Continuing in this manner the string forms a polygon which is necessarily convex.

Definition 4.2. *Let f be as above. We call the slopes of the segments appearing in $N(f)$ the* slopes of f. *If λ is a slope, we call the length of the projection of the corresponding segment to the x-axis the* length of λ.

Lemma 4.3. *Factor*

$$f(x) = \prod_{i=1}^{n} \left(1 - \frac{x}{\alpha_i}\right).$$

Let $\lambda_i = \operatorname{ord}_p 1/\alpha_i$. Then if λ is a slope of f of length l, then exactly l of the λ_i are equal to λ. Let

$$f_\lambda(x) = \prod_{\operatorname{ord}_p \alpha_i = -\lambda} \left(1 - \frac{x}{\alpha_i}\right).$$

Then $f_\lambda(x) \in K[x]$ and $f = \prod_\lambda f_\lambda$.

Example. Let $f(x) = 1 - x/2 - x^2/2 - x^3 \in \mathbb{Q}_2[x]$. One can check that $f(x)$ is irreducible in $\mathbb{Q}[x]$. However, its 2-adic Newton polygon has 3 distinct slopes, so it factors into 3 linear factors over $\mathbb{Q}_2$. In particular, all of its roots are in $\mathbb{Q}_2$.

Proof. The proof will be an exercise in the application of the following fact: suppose that $x_1, \ldots, x_n \in \Omega$, and $\operatorname{ord}_p x_1 < \operatorname{ord}_p x_i, i = 2, \ldots, n$. Then $\operatorname{ord}_p \sum_i x_i = \operatorname{ord}_p x_1$.

Order the α_i so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Suppose that $\lambda_1 = \cdots = \lambda_r < \lambda_{r+1}$. We want to show that the first segment of $N(f)$ is the line joining $(0,0)$ and $(r, \operatorname{ord}_p a_r)$. Now, a_i is a symmetric polynomial in the $1/\alpha_j$. Hence we have $\operatorname{ord}_p a_i \geq i\lambda_1$, so the point $(i, \operatorname{ord}_p a_i)$ lies above this line.

On the other hand, a_r is a sum of products of r of the $1/\alpha_j$, and $1/(\alpha_1 \ldots \alpha_r)$ is the unique term with minimal valuation. Thus $\operatorname{ord}_p a_r = r\lambda_1$. Moreover, we have $\operatorname{ord}_p a_{r+1} > (r+1)\lambda_1$, so there is a corner in the polygon at $(r, \operatorname{ord}_p a_r)$, and the first segment joins $(0,0)$ and $(r, \operatorname{ord}_p a_r)$.

We next suppose that $\lambda_{r+1} = \cdots = \lambda_s < \lambda_{s+1}$. The same argument works to show that the next segment of the polygon is the line joining $(r, \operatorname{ord}_p a_r)$ and $(s, \operatorname{ord}_p a_s)$. Continuing in this manner, we verify that $N(f)$ has the claimed form.

The last statement follows immediately from the Galois invariance of the norm. □

Remark 1. Using the lemma, you can give another proof of Eisenstein's criterion.

Newton polygons for power series

In this section we will work with a power series $f(X) = 1 + \sum_{i=1}^{\infty} a_i X^i \in 1 + X\Omega[[X]]$.

Definition 4.4. *For each $n \geq 1$, let $f_n(X) = 1 + \sum_{i=0}^{n} a_i X^i$. Then we define the* Newton polygon of f *to be the limit of the $N(f_n)$ as $n \to \infty$. To be more precise, we distinguish three cases.*

1. *$N(f)$ has infinitely many segments of finite length.*
2. *$N(f)$ has finitely many segments, and the last segment, which is infinitely long, has infinitely many of the points $(i, \operatorname{ord}_p a_i)$ on it.*
3. *$N(f)$ has finitely many segments, and the last segment, which is infinitely long, has finitely many of the points $(i, \operatorname{ord}_p a_i)$ on it.*

Example. Let $f(X) = -\log(1 - X)/X$.

Lemma 4.5. *Let $b = \sup_i \lambda_i$, the supremum being over the slopes of f. Then the radius of convergence of f is p^b.*

Moreover, f converges on the boundary of the disk $D(p^b)$ if and only if we are in case (3) above and the distance between the final slope and $(i, \operatorname{ord}_p a_i)$ goes to infinity with i.

Proof. Suppose that $\operatorname{ord}_p x > -b$. We can write $\operatorname{ord}_p x = -c$, with $c < b$. It follows that

$$\operatorname{ord}_p a_i x^i = \operatorname{ord}_p a_i - ic.$$

Now, $\operatorname{ord}_p a_i - ic$ is eventually monotonic and strictly increasing, so we're done. The converse direction is similar.

The final claim is clear from the proof. (Here the picture on p. 101 of the course text is helpful). □

Lemma 4.6. *Suppose λ_1 is the first slope of $N(f), c \in \Omega, \operatorname{ord}_p c = \lambda \leq \lambda_1$. Suppose that f converges on the closed disk $D(p^\lambda)$, and let*

$$g(X) = (1 - cX)f(X).$$

Then $N(g)$ is obtained by translating $N(f)$ by $(1, \lambda)$ and joining $(0, 0)$ and $(1, \lambda)$.

Suppose finally that f has last slope λ_f. Then $f(X)$ converges on the closed disk $D(p^{\lambda_f})$ if and only if $g(X)$ does.

Proof. We may suppose without loss of generality that $c = 1$ and so $\lambda = 0$. We let $g(X) = 1 + \sum_{i=0}^{\infty} b_i X^i$. Thus $b_i = a_i - a_{i-1}$ for each i. We must show that $N(g)$ is $N(f)$ translated one unit to the right, with an extra segment between the $(0, 0)$ and the endpoint.

Since $\operatorname{ord}_p b_i \geq \operatorname{ord}_p a_{i-1}$, the points $(i, \operatorname{ord}_p b_i)$ all lie above the Newton polygon of $N(f)$ translated one to the right. Suppose that $(i - 1, \operatorname{ord}_p a_{i-1})$ is a corner of $N(f)$. Then $\operatorname{ord}_p a_i > \operatorname{ord}_p a_{i-1}$, hence $\operatorname{ord}_p b_i = \operatorname{ord}_p a_{i-1}$, and $(i, \operatorname{ord}_p b_i)$ is a corner of $N(g)$.

It follows that $N(g)$ is as claimed in the lemma, except for possibly the last segment. Suppose that $N(g)$ had a slope $\lambda_g > \lambda_f$. Then for some i, $(i + 1, \operatorname{ord}_p a_i)$ lies below this segment, hence $\operatorname{ord}_p b_j > \operatorname{ord}_p a_i$ for all $j \geq i+1$. Then $b_j = a_j - a_{j-1} \Rightarrow \operatorname{ord}_p a_j = \operatorname{ord}_p a_{i+1}$ for all $j \geq i+1$. This contradicts the assumption that $f(X)$ converges in the closed disk $D(1)$.

The remaining claims follow similarly. □

Exercises to lecture 5

1. Find the Newton polygons of the following polynomials:
   (a) $1 - x + px^2$.
   (b) $1 - x^3/p^2$.
   (c) $1 + x^2 + px^4 + p^3x^6$.
   (d) $\prod_{i=1}^{p^2}(1 - ix)$.
2. Justify the first sentence in the proof of Lemma 4.6.
3. Let $f(x) = 1 + a_1x + \cdots + a_nx^n$, with $a_n \neq 0$. Suppose that $N(f)$ consists of a single line from $(0, 0)$ to (n, m), with n, m coprime integers. Show that f is irreducible.
4. Does the Newton polygon of every irreducible f have the form given in the previous exercise? Give a proof of counter-example.
5. Suppose that that $f = 1 + a_1x + \cdots + a_{2n}x^{2n}$ and that whenever α is a reciprocal root of f (i.e. $f(1/\alpha) = 0$), so also is p/α, with the same multiplicity. What does this imply about the shape of $N(f)$? Draw all possible shapes for n in the range $n = 1, 2, 3, 4$.
6. Find a power series $f(X)$ such that $N(f)$ has a segment with irrational slope.

Lecture 6

Newton polygons of power series, continued

We continue to develop the properties of Newton polygons of power series. We work with a power series $f(X) = 1 + \sum_{i=0}^{\infty} a_i X^i \in \Omega[[X]]$, assumed to have a strictly positive radius of convergence.

Lemma 5.1. *Suppose that $f(X)$ has first slope λ_1, and that f has at least two distinct slopes. Then there exists $y \in \Omega$ with $\operatorname{ord}_p 1/y = \lambda_1$ and $f(y) = 0$.*

Proof. We may suppose that $\lambda_1 = 0$. It follows that $\operatorname{ord}_p a_i \geq 0$ for all i, and that $\operatorname{ord}_p a_i \to 0$ as $i \to \infty$. Let $f_n(X) = 1 + \sum_{i=0}^{n} a_i X^i$. Let N be the largest integer such that $\operatorname{ord}_p a_N = 0$.

Then for $n \geq N$, $f_n(X)$ has exactly N zeroes $x_{n,1}, \ldots, x_{n,N}$ of valuation 0 (by our result for Newton polygons of polynomials). We define a sequence y_j as follows: let $y_N = y_{N,1}$, and for each $n \geq N$, let y_{n+1} be one of the elements of $x_{n,1}, \ldots, x_{n,N}$ minimizing $|x_{n,i} - y_n|$. We claim that $(y_n)_{n \geq N}$ is a Cauchy sequence, and that the limit y has the desired properties.

For each $n \geq N$, let S_n be the set of roots of $f_n(X)$ counted with multiplicities. For $n \geq N$, we have

$$|f_{n+1}(y_n) - f_n(y_n)| = |f_{n+1}(y_n)| = \prod_{x \in S_{n+1}} \left|1 - \frac{y_n}{x}\right|$$

$$= \prod_{i=1}^{N} \left|1 - \frac{y_n}{x_{n+1,i}}\right| = \prod_{i=1}^{N} |x_{n+1,i} - y_n| \geq |y_{n+1} - y_n|^N.$$

Rewriting this, we have

$$|y_{n+1} - y_n|^N \leq |a_{n+1} y_n^{n+1}| = |a_{n+1}|.$$

By hypothesis, $|a_n| \to 0$ as $n \to \infty$. This shows that the sequence (y_n) is in fact a Cauchy sequence. Let $y \in \Omega$ be its limit.

We have $f(y) = \lim_n f_n(y)$. On the other hand, we have

$$|f_n(y)| = |f_n(y) - f_n(y_n)| = |y - y_n| \left| \sum_{i=1}^{n} a_i \frac{y^i - y_n^i}{y - y_n} \right|.$$

Now, $(y^i - y_n^i)/(y - y_n) = (y^{i-1} + y^{i-2} y_n + \cdots + y_n^{i-1})$ has norm ≤ 1, by the ultrametric inequality. Using that $|a_i| \leq 1$ for each i gives that $|f_n(y)| \leq |y - y_n|$, which tends to 0 as n tends to infinity. Thus $f(y) = 0$. This concludes the proof. □

Lemma 5.2. *Suppose that $f(\alpha) = 0$, and that*

$$g(X) = (1 - X/\alpha)^{-1} f(X) = (1 + X/\alpha + (X/\alpha)^2 + \ldots) f(X).$$

Then $g(X)$ converges on $D(|\alpha|)$.

Proof. Let $f_n(X) = 1 + \sum_{i=1}^{n} a_i X^i$, as above. We have

$$b_i = 1/\alpha^i + a_1/\alpha^{i-1} + \cdots + a_{i-1}/\alpha + a_i,$$

hence $b_i \alpha^i = f_i(\alpha)$. But $|f_i(\alpha)| \to 0$ as $i \to \infty$, and thus $g(\alpha)$ converges. □

Weierstrass factorization

Lemma 5.3. *There exists a power series $g(X) \in 1 + X\Omega[[X]]$ such that $f(X)g(X) = 1$. Suppose that the slopes of the Newton polygon of f are strictly greater than λ. Then the same is true of g. In particular, f and g both converge in the closed disk $D(p^\lambda)$ and are non-zero there.*

Proof. We write $g(X) = 1 + \sum_{i=0}^{\infty} b_i X^i$. In order to have $f(X)g(X) = 1$ as formal power series, we must have for each i

$$b_i = -(b_{i-1} a_1 + \cdots + b_1 a_{i-1} + a_i).$$

This allows us to solve for the b_i inductively, so we find that $g(X)$ exists as a formal power series.

For the second part of the lemma, we may suppose after scaling that $\lambda = 0$. The hypotheses of the theorem imply that for each i, $\operatorname{ord}_p a_i > 0$, and $\operatorname{ord}_p a_i \to \infty$ as $i \to \infty$. We must show that the same is true of the b_i. The fact that $\operatorname{ord}_p b_i > 0$ for each i follows by induction from the above relation.

We now show that $\operatorname{ord}_p b_i \to \infty$ as $i \to \infty$. Fix $M > 0$, and choose m such that for all $i > m$, $\operatorname{ord}_p a_i > M$. Let $\epsilon = \min(\operatorname{ord}_p a_1, \ldots, \operatorname{ord}_p a_m)$. I claim that for all $i > nm$, we have $\operatorname{ord}_p b_i > \min(M, n\epsilon)$. In particular, when $i > mM/\epsilon$, $\operatorname{ord}_p b_i > M$. Thus the claim implies the lemma.

We prove the claim by induction on n, the case $n = 0$ being trivial. Suppose we've proven the claim for $n - 1$. We write

$$b_i = -(\underbrace{b_{i-1}a_1 + \cdots + b_{i-m}a_m}_{\operatorname{ord}_p > \min(M,(n-1)\epsilon)+\epsilon} + \underbrace{b_{i-m-1}a_{m+1} + \cdots + a_i}_{\operatorname{ord}_p > M}).$$

The result follows. □

Theorem 5.4. *Let $f(X) = 1 + \sum_{i=0}^{\infty} a_i X^i$, and suppose that f converges on the closed disk $D(p^\lambda)$. Let N be the total length of all segments of $N(f)$ of slope $\leq \lambda$. (We suppose in particular that N is finite).*

Then there exists a unique polynomial $h(X) \in 1 + X\Omega[X]$ of degree N and $g(X) \in 1 + X\Omega[[X]]$ convergent and non-zero in $D(p^\lambda)$ such that $h(X) = f(X)g(X)$.

Moreover, $N(h)$ is equal to the part of $N(f)$ between 0 and $(N, \operatorname{ord}_p a_N)$.

Proof. We induct on N. The case $N = 0$ is the lemma above. Suppose that the theorem has been proven in the case $N - 1$, and let $\lambda_1 \leq \lambda$ be the first slope of f. By lemma 5.1, there exists $\alpha \in \Omega$ with $\operatorname{ord}_p 1/\alpha = \lambda_1$ and $f(\alpha) = 0$. Let $f_1(X) = (1 - X/\alpha)^{-1} f(X)$. Thus $f_1(X)$ converges on the disk $D(|\alpha|) = D(p^{\lambda_1})$.

Write $c = 1/\alpha$, so that $f(X) = (1 - cX)f_1(X)$. Let $f_1'(X)$ have first slope λ_1'. If $\lambda_1' < \lambda_1$, then $f_1(X)$ has a root of slope λ_1', and hence so does f. But the previous lemma shows that $f(X)$ has no zeroes in the disk $D(\lambda_1^-)$, so this is impossible. Thus $\lambda_1' \geq \lambda_1$.

We can now apply the lemma from last time, relating $N(f)$ and $N(f_1)$. This says that $N(f)$ is obtained from $N(f_1)$ by translating by $(1, \lambda_1)$ and joining the points $(0, 0)$ and $(1, \lambda_1)$. By induction, there exists a polynomial $h_1(X)$ of degree $N - 1$ and a power series $g(X)$, convergent and non-vanishing in $D(p^\lambda)$, such that

$$h_1(X) = f_1(X)g(X).$$

Setting $h(X) = (1 - cX)h_1(X)$, we have

$$h(X) = f(X)g(X).$$

This h satisfies the hypothesis of the theorem. □

Corollary 5.5. *1. Suppose that the slope λ of f has finite length N. Then there are exactly N values of x such that $f(x) = 0$ and $\operatorname{ord}_p 1/x = \lambda$ (counting multiplicities).*

2. Suppose that $f(X)$ has infinite radius of convergence (i.e. $f(x)$ converges for any $x \in \Omega$. In this case we sometimes say that f is entire*). Then for any C, there exist only finitely many values of x with $|x| \leq C$ and $f(x) = 0$. Moreover, ordering the zeroes x_n of f to have increasing norm, we have*

$$f(X) = \prod_{i=1}^{\infty} \left(1 - \frac{X}{x_n}\right).$$

Remark 2. Compare the above corollary with the Weierstrass factorization theorem, valid over $\mathbb{C}$.

Paper topics

The aim is to give a 5-10 page exposition and a 30-45 minute lecture on one of the following topics.

1. The connections with global arithmetic. Let F be a finite extension of $\mathbb{Q}$ (a *number field*). Then F has a canonical 'ring of integers' $\mathcal{O}_F \subset F$, which plays an analogous role in the study of F to that of $\mathbb{Z}$ in the study of $\mathbb{Q}$. In particular, prime ideals of $\mathcal{O}_F$ give rise to norms on F, and completing with respect to one of these gives extensions of $\mathbb{Q}_p$ (the so-called local fields associated to the 'global field' F). There is a connection between the local and global Galois groups.

2. The Hasse principle. Consider a quadratic form $f(x_1, \ldots, x_r) = a_1x_1^2 + \cdots + a_rx_r^2$, with coefficients in the field $\mathbb{Q}$. Does there exist a rational solution to the equation $f(x_1, \ldots, x_r) = 0$? If this is true then there will certainly exist solutions in $\mathbb{Q}_p$ for every prime p (and in $\mathbb{R}$). The Hasse principle states that the converse holds: if $f = 0$ has a solution in every 'local' field, then it has a solution 'globally' (i.e. in $\mathbb{Q}$).

3. The p-adic interpolation of the Riemann zeta function. The Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

is an analytic function of a complex variable, defined a priori only where the above series converges (namely, when the real part of s is > 1). However, it admits an analytic continuation to the whole of $\mathbb{C}$, and the properties of this continuation are intimately tied up with the distribution of the primes.

It is a remarkable fact that the values of ζ at negative integers enjoy p-adic continuity properties, and in fact can be interpolated to give a p-adic analytic avatar of the Riemann zeta function.

4. The ramification filtration. Let L/K be a Galois extension of finite extensions of $\mathbb{Q}_p$. Then there is a canonical filtration of $\mathrm{Gal}(L/K)$ by normal subgroups, corresponding to the ramification behaviour of the field. Each quotient is abelian, so this shows that $\mathrm{Gal}(L/K)$ is always a soluble group.

Particularly interesting is that this filtration admits a modification, making it stable under passage to quotient; in particular, it induces a filtration of the *absolute Galois group* of K. There is also a connection with local class field theory.

5. The local Kronecker-Weber theorem. The Kronecker-Weber theorem states that any abelian extension of $\mathbb{Q}$ (i.e. a Galois extension with abelian Galois group) is contained inside a cyclotomic extension.

One way to prove this is by passing from the local Kronecker-Weber theorem, which states the same thing for $\mathbb{Q}_p$. The proof is a detailed study of the ramification properties of p-adic fields.

6. Truncated exponential polynomials. The following theorem is due to Schur:

Theorem 5.6. *Let $f_n(x) = 1 + x + x^2/2! + \cdots + x^n/n!$. Let L be the splitting field of f over $\mathbb{Q}$. Then the Galois group of $L/\mathbb{Q}$ is A_n if $4 \mid n$ and S_n otherwise.*

Coleman has given a nice proof using Newton polygons and a simple result relating local and global Galois groups, along with some results on the distributions of the primes.

Exercises to lecture 6

1. Choose a topic to write a paper on.

Lecture 7

Systems of equations and algebraic varieties

Let F a field and $\overline{F}$ its algebraic closure. We are interested in the sets of solutions to polynomial equations over F.

Definition 6.1. Affine n-space *over F is the set*

$$\mathbb{A}^n = \{(y_1, \ldots, y_n) \in \overline{F}^n\}.$$

Let $f(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n]$. The affine hypersurface *defined by f is*

$$H_f = \{(y_1, \ldots, y_n) \in \mathbb{A}^n \text{ such that } f(y_1, \ldots, y_n) = 0\}.$$

If $f_1, \ldots, f_r \in F[x_1, \ldots, x_n]$ then the affine algebraic variety *defined by these polynomials is*

$$H_{f_1,\ldots,f_r} = H_{f_1} \cap \cdots \cap H_{f_r}.$$

If H is an affine algebraic variety defined by $f_1, \ldots, f_r$ and F' is an algebraic extension of F, we write

$$H(F') = H \cap (F')^n = \{(y_1, \ldots, y_n) \in (F')^n \text{ such that } f(y_1, \ldots, y_n) = 0.\}.$$

We will speak of H_f and $H_{f_1,\ldots,f_r}$ as being defined 'over F'.

Example. Let $f(x_1, x_2) = x_1^2 + x_2^2 - 1$, and let $X = H_f$. Then $X(\mathbb{R})$ is the unit circle, while $X(\mathbb{C})$ can be put into bijection with the complex plane.

Now let $f(x, y) = y^2 - (x^3 - x)$, and take $F = \mathbb{F}_p$. Then $\#H_f(\mathbb{F}_p)$ is related to the number of elements z of $\mathbb{F}_p$ such that $z^3 - z$ is a square.

From now on we will be considering the following question: Let $F = \mathbb{F}_q$, for some q, and let X be an affine algebraic variety over $\mathbb{F}_q$. Let $N_s = \#X(\mathbb{F}_{q^s})$. How does N_s vary with s? A natural way to study sequences of integers is via generating functions, which motivates the following definition.

Definition 6.2. *With notation as above, consider the formal power series in the variable T*

$$Z(X, T) = \exp\left(\sum_{s=1}^{\infty} N_s T^s / s\right).$$

This is called the zeta function *of the algebraic variety X.*

In general, there are several different possibilities for the generating function of an integer sequence. We will see in a moment one reason why this is the right choice.

Example.

- Take $X = \mathbb{A}^n$, affine n-space. Then $N_s = \mathbb{A}^n(\mathbb{F}_{q^s}) = q^{ns}$, and hence

$$Z(X, T) = \exp\left(\sum_{s=1}^{\infty} (q^n T)^s / s\right) = \exp(-\log(1 - q^n T)) = \frac{1}{1 - q^n T}.$$

- Let $f(x, y) = y^2 - (x^3 - x)$, and let $X = H_f$. Suppose that p is odd. Then we have

$$N_s = 3 + 2\#\{x \in \mathbb{F}_{q^s}^{\times} \text{ such that } x^3 - x \text{ is a square}\}.$$

  One can show by quite indirect means that

$$Z(X, T) = \frac{1 - aT + qT^2}{1 - qT} = 1 + (q - a)T + q(1 - a + q)T^2 + \ldots$$

  for some $a \in \mathbb{Z}$. In particular, $N_1 = q - a$, and a and hence $Z(X, T)$ are determined by N_1. Thus all the numbers $N_s, s \geq 1$ are determined by N_1.

- Let $f(x) = x^p + y^p - 1$, and $X = H_f$ (the so-called Fermat hypersurface). Then the coefficients of $Z(X, T)$ are related to Gauss sums, although we won't say how.

Remark 3. One game you can play is the following: let

$$f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n].$$

Then, by reducing the coefficients of f modulo p, we obtain affine hypersurfaces H_p over $\mathbb{F}_p$ for every prime p, and hence zeta functions $Z(H_p, T)$. One can then form a 'global' zeta function

$$Z(f, s) = \prod_p Z(H_p, p^{-s}),$$

where s is a complex variable. The most trivial example is when we consider 0-dimensional affine space and take f to be the zero polynomial, so all the hypersurfaces H_p consist of a single point. Specializing the first example above to $n = 0$ gives $Z(H_p, T) = 1/(1-T)$, and hence

$$Z(f, s) = \prod_p \frac{1}{1 - p^{-s}}.$$

In other words, we recover the Riemann zeta function. Taking more interesting polynomials f gives rise to even more interesting complex analytic functions. The fact that this works is one of the reasons for defining $Z(X, T)$ in the way that we did.

More on zeta functions

One elementary property is the following.

Proposition 6.3. *Let X be an algebraic variety over $\mathbb{F}_q$. Then the coefficients of $Z(X, T)$ are positive integers.*

Proof. Note that $G_s = \mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_{q^s})$ acts on X, with $X^{G_s} = X(\mathbb{F}_{q^s})$. Let $Y = X/G_1$. Given $[x] \in Y$, we write $\deg[x] = \#[x]$ for the size of the Galois orbit of x. Note that if $x \in X(\mathbb{F}_{q^s}) - X(\mathbb{F}_{q^t})$ for every $t \mid s$ then $\deg[x] = s$.

We single out the contribution of $[x]$ to $Z(X, T)$. Let $\deg[x] = t$. Then the contribution of $[x]$ to N_s is t if $t \mid s$ and 0 otherwise, so the contribution to the zeta function is

$$\exp\left(\sum_{j=1}^{\infty} tT^{jt}/jt\right) = \exp(-\log(1 - T^t)) = \frac{1}{1 - T^t}.$$

Thus we can write

$$Z(X, T) = \prod_{[x] \in Y} \frac{1}{1 - T^{\deg[x]}} = \prod_{[x] \in Y} \left(1 + T^{\deg[x]} + T^{2\deg[x]} + \dots\right).$$

The result follows. (Note that the product on the right makes sense since, for any give N, there are only finitely many terms in the product with T^j for $j \leq N$). □

Proposition 6.4. *Suppose that X is an algebraic variety over $\mathbb{F}_q$, contained in $\mathbb{A}^n$. Then the coefficient of T^j in $Z(X, T)$ is at most q^{nj}.*

Proof. We saw above that

$$Z(\mathbb{A}^n, T) = \frac{1}{1 - q^n T} = \sum_{j=1}^{\infty} q^{nj} T^j.$$

The fact that $N_s \leq \#\mathbb{A}^n(\mathbb{F}_{q^s})$ for every s implies the corresponding result for the coefficients of $Z(X, T)$ and $Z(\mathbb{A}^n, T)$, so the result follows. □

Rationality of the zeta function

The most important basic property of the zeta function is the following.

Theorem 6.5 (Dwork)**.** *Let X be an affine algebraic variety over $\mathbb{F}_q$. Then $Z(X, T)$ is a rational function (i.e. it is the power series expansion of a rational function of T), with coefficients in $\mathbb{Q}$.*

To give you an idea of the importance of this for the numbers N_s, we give the following equivalent formulation.

Theorem 6.6. *Let X be an affine algebraic variety over $\mathbb{F}_q$. Then there exist sets of algebraic numbers $\{\alpha_1, \ldots, \alpha_t\}$, $\{\beta_1, \ldots, \beta_u\}$ which are invariant under Galois conjugation and such that*

$$N_s = \sum_{i=1}^{t} \alpha_i^s - \sum_{j=1}^{u} \beta_j^s, \text{ for every } s \geq 1.$$

Corollary 6.7. *Suppose that the degree of the numerator and denominator of $Z(X,T)$ are bounded by N. Then the numbers $N_s, s \geq 1$, are completely determined by the N_s for $s = 1, \ldots, 2N$.*

We are going to dedicate the rest of our time to proving this theorem. Here is a sketch of the proof:

1. Reduction to the case where X is an affine hypersurface. One can show that if $Y_1, Y_2 \subset X$ are algebraic varieties with $Y_1 \cup Y_2 = X$ and $Y_1 \cap Y_2 = W$, say, then

$$Z(X,T) = \frac{Z(Y_1,T) \cdot Z(Y_2,T)}{Z(W,T)}.$$

   An easy induction argument then allows us to reduce to the case where X is an affine hypersurface, defined by a single equation in $\mathbb{A}^n$.

2. A theorem of Borel and Dwork. We will prove the following result of quite independent interest:

   Theorem 6.8. *Let $F(T) = 1 + \sum_{j=1}^{\infty} A_j T^j$ be a power series with integer coefficients, and let p be a prime. Suppose that there exist $r, R > 0$ such that $rR > 1$ and*

   - *F is meromorphic in the disk $|z| < R$ in $\mathbb{C}$.*
   - *F is meromorphic in the disk $|z| < r$ in Ω.*

   Then F is the power series expansion of a rational function of T.

   (We recall that a holomorphic function in a disk $|z| < A$ is one defined by a convergent power series in that disk. A meromorphic function in a disk is one that can we written as a quotient of two holomorphic functions in the disk).

3. As we saw above, the coefficients of $Z(X,T)$ are majorized by those of $Z(\mathbb{A}^n, T) = 1/(1-q^nT)$, a function which is certainly meromorphic in $\mathbb{C}$. Thus it will suffice to exhibit $Z(X,T)$ as a meromorphic p-adic function, after which we will apply the above theorem.

A rationality criterion

The following is the first stepping stone to part 2 of the plan above.

Theorem 6.9. *Let K be a field, and let $F(T) = \sum_{i=0}^{\infty} a_i T^i$. For $m, s \geq 0$, let $A_{s,m}$ be the $m+1 \times m+1$ matrix*

$$A_{s,m} = \begin{pmatrix} a_s & a_{s+1} & a_{s+2} & \cdots & a_{s+m} \\ a_{s+1} & a_{s+2} & a_{s+3} & \cdots & a_{s+m+1} \\ a_{s+2} & a_{s+3} & a_{s+4} & \cdots & a_{s+m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{s+m} & a_{s+m+1} & a_{s+m+2} & \cdots & a_{s+2m} \end{pmatrix}.$$

Let $N_{s,m} = \det A_{s,m}$. Then $F(T)$ is a rational function if and only if there exist $M, S \geq 0$ such that $N_{s,M} = 0$ whenever $s \geq S$.

Proof. First we prove sufficiency. Write $F(T) = P(T)/Q(T)$, where $P(T) = b_0 + b_1T + \cdots + b_NT^N$ and $Q(T) = c_0 + c_1T + \cdots + c_MT^M$. Then $Q(T) \cdot F(T) = P(T)$, and comparing coefficients of the degree i term gives

$$\sum_{j=0}^{M} a_{i-M+j}c_{M-j} = 0$$

for i sufficiently large. In other words, we have

$$\begin{gathered} a_sc_M + a_{s+1}c_{M-1} + \cdots + a_{s+M}c_0 = 0 \\ a_{s+1}c_M + a_{s+2}c_{M-1} + \cdots + a_{s+M+1}c_0 = 0 \\ \vdots \\ a_{s+M}c_M + a_{s+M+1}c_{M-1} + \cdots + a_{s+2M}c_0 = 0, \end{gathered}$$

for s sufficiently large. This shows that $N_{s,M} = 0$ for s sufficiently large.

Suppose conversely that there exist M, S as in the statement of the theorem. Suppose that M is chosen to be minimal with respect to the property that $N_{s,M} = 0$ for all sufficiently large s. We claim that $N_{s,M-1} \neq 0$ for all $s \geq S$.

Suppose for contradiction that this was not the case, so that $N_{s,M-1} = N_{s,M} = 0$. Let $r_0, \ldots, r_M$ be the rows of $A_{s,M}$. Some linear combination of $r_0, \ldots, r_{M-1}$ vanishes, except for possibly the last column, say

$$\alpha_k r_k + \alpha_{k+1}r_{k+1} + \cdots + \alpha_{M-1}r_{M-1},$$

with $\alpha_k \neq 0$. Performing some row operations on $A_{s,M}$, we can replace r_k by

$$r_k + 1/\alpha_k(\alpha_{k+1}r_{k+1} + \cdots + \alpha_{M-1}r_{M-1}).$$

Suppose that $k > 0$. Then the matrix $A_{s,M}$ looks like

$$\begin{pmatrix} a_s & a_{s+1} & \ldots & a_{s+M} \\ a_{s+1} & a_{s+2} & \ldots & a_{s+m+1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \beta \\ \vdots & \vdots & & \vdots \\ a_{s+M} & a_{s+M+1} & \ldots & a_{s+2M} \end{pmatrix}.$$

The lower left $m \times m$ matrix has determinant $N_{s+1,M-1} = 0$.

Similarly, when $k = 0$ we can check that $N_{s+1,M-1} = 0$. By induction, we have $N_{s,M-1} = 0$ for all $s \geq S$, contradicting the minimality of M.

Thus for all $s \geq S$, we have $N_{s,M} = 0$ and $N_{s,M-1} \neq 0$. In particular, there exists a linear combination of the rows of $A_{s,M}$ which is zero, and in which the coefficient of the last row is non-zero, hence the last row of $A_{s,M}$ is a linear combination of the preceding M rows.

It follows that any solution $(u_0, \ldots, u_M)$ to the equations

$$\begin{gathered} a_Su_M + a_{S+1}u_{M-1} + \cdots + a_{S+M}u_0 = 0 \\ \vdots \\ a_{S+M-1}u_M + a_{S+M}u_{M-1} + \cdots + a_{S+2M-1}u_0 = 0 \end{gathered}$$

is also a solution to

$$a_su_M + a_{s+1}u_{M-1} + \cdots + a_{s+M}u_0 = 0$$

for every $s \geq S$. This says that

$$\left(\sum_{i=0}^{M} u_iT^i\right)\left(\sum_{j=1}^{\infty} a_jT^j\right)$$

is a polynomial, so we're done.

□

Exercises to lecture 7

1. Read Section V.1 of the course text.
2. Prove the equivalence between Theorems 6.5 and 6.6.
3. Deduce Corollary 6.7 from Theorem 6.5.
4. Prove the formula above relating the zeta functions of X, Y_1, Y_2 and W, and use it to show that Dwork's theorem for affine hypersurfaces implies it for all affine algebraic varieties.
5. Let $f(x, y) = y^2 - (x^3 - x)$, and compute $Z(H_f, T)$ over $\mathbb{F}_p$ for $p = 3, 5, 7$.
6. (Not to be taken too seriously) Read the following blog post for another perspective on the information contained in a statement of the form 'the generating function of this sequence is rational/algebraic/...'.
7. Try exercises V.1.14 and V.1.15 of the course text.

Lecture 8

Rationality of p-adic power series

Last time we stated the following theorem:

Theorem 7.1. *Let K be a field, and let $F(T) = \sum_{i=0}^{\infty} a_i T^i$. For $m, s \geq 0$, let $A_{s,m}$ be the $m+1 \times m+1$ matrix*

$$A_{s,m} = \begin{pmatrix} a_s & a_{s+1} & a_{s+2} & \dots & a_{s+m} \\ a_{s+1} & a_{s+2} & a_{s+3} & \dots & a_{s+m+1} \\ a_{s+2} & a_{s+3} & a_{s+4} & \dots & a_{s+m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{s+m} & a_{s+m+1} & a_{s+m+2} & \dots & a_{s+2m} \end{pmatrix}.$$

Let $N_{s,m} = \det A_{s,m}$. Then $F(T)$ is a rational function if and only if there exist $M, S \geq 0$ such that $N_{s,M} = 0$ whenever $s \geq S$.

We would like to turn this into an effective way of recognising when a power series is a rational function in disguise. (A more general version of this result was stated last time).

Theorem 7.2. *Let $F(T) = 1 + \sum_{i=1}^{\infty} a_i T^i$ be a power series, and suppose that the a_i are* integers. *Suppose that F defines a holomorphic function in the disk $|z| < R$ in $\mathbb{C}$ and an entire meromorphic function in Ω (i.e. F is the quotient of two entire holomorphic functions in Ω). Then F is the power series expansion of a rational function.*

Proof. We can suppose $R < 1$. We apply the Weierstrass preparation theorem. For any $r > 0$, we can write $B(T) = A(T) \cdot F(T)$, where A is a polynomial and B is a power series such that both $B(T)$ and $1/B(T)$ are holomorphic and non-vanishing in the disk $|x| \leq r$. Write $B(T) = 1 + \sum_{i=1}^{\infty} B_i T^i$, $A(T) = 1 + A_1 T + \dots + A_e T^e$.

Let $A_{s,m}$ and $N_{s,m}$ be as defined above. Note that $N_{s,m}$ is an integer. Equating coefficients in $B = AF$, we have

$$B_{s+e} = a_{s+e} + A_1 a_{s+e-1} + \dots + A_e a_s.$$

Choose $m > 2e$. Then the above equation shows that after doing some column operations, we can replace all but the first e columns of the matrix $A_{s,m}$ by the columns of the matrix $(B_{s+i+j})_{i,j}$. Hence

$$|N_{s,m}|_p \leq \left(\max_{j \geq s+e} |B_j|_p \right)^{m+1-e} < r^{-s(m+1-e)},$$

for s sufficiently large. Taking $r = 1/\sqrt{R}$, this gives $|N_{s,m}|_p < R^{s(m+2)}$.

But we have also

$$|N_{s,m}|_\infty < (m+1)!R^{-1/2(m+1)(m+2s)},$$

for s sufficiently large, and so

$$|N_{s,m}|_\infty |N_{s,m}|_p < (m+1)!R^{s(m+2)-1/2(m+1)(m+2s)} = (m+1)!R^{s-\frac{1}{2}m(m+1)},$$

and this expression tends to 0 as s tends to infinity. But the only integer n with $|n|_\infty |n|_p < 1$ is 0, so $N_{s,m} = 0$ for s sufficiently large. Applying the previous theorem now gives the result. □

p-adic interpolation

The above result reduces Dwork's theorem to showing that the zeta function of an affine hypersurface $X_f \subset \mathbb{A}^n$ over $\mathbb{F}_q$ is actually p-adic meromorphic. Our approach to this is based on the following observation: for each s, $\mathbb{A}^n(\mathbb{F}_q^s)$ is a group, so it makes sense to take the (discrete) Fourier transform of the indicator function of $X_f(\mathbb{F}_q^s)$, which can be viewed as taking values in Ω. We hope that this setup will give rise to a p-adic analytic function. This motivates the constructions of this lecture.

Let $\epsilon \in \Omega$ be a p^{th} root of unity. If $q = p^s$, we write $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ for the natural trace map, and $a \mapsto \epsilon^{\mathrm{tr}\, a}$ for the induced character

$$\mathbb{F}_q \to \mathbb{F}_p \to \Omega^\times.$$

We would like to find a p-adic power series $\Theta(T)$ whose value at $[a]$, $a \in \mathbb{F}_p$, is ϵ^a. More generally, for $a \in \mathbb{F}_q$ we'd like

$$\epsilon^{\mathrm{tr}\, a} = \Theta([a])\Theta([a^p]) \dots \Theta([a]^{p^{s-1}}).$$

(Recall that $[\cdot]$ denotes the Teichmuller digits, the natural section $\mathbb{F}_q \to \Omega$ of the reduction map, and that these respect multiplication).

Henceforth let $a \in \mathbb{F}_q$ and $t = [a] \in \Omega$. Let $\lambda = \epsilon - 1$. A natural first guess for Θ is the series

$$g(T) = (1+\lambda)^T = \sum_{i=0}^{\infty} \frac{T(T-1)\dots(T-i+1)}{i!}\lambda^i.$$

However, for $t \notin \mathbb{Z}_p$, we have

$$\mathrm{ord}_p \frac{t(t-1)\dots(t-i+1)}{i!}\lambda^i = i\,\mathrm{ord}_p\,\lambda - \frac{i-S_i}{p-1} = \frac{S_i}{p-1},$$

and this does not tend to infinity, so the series does not converge!

To get around this we introduce the following formal power series in two variables

$$F(X,Y) = (1+Y)^X \times (1+Y^p)^{(X^p-X)/p} \times \dots \times (1+Y^{p^n})^{(X^{p^n}-X^{p^{n-1}})/p^n} \times \dots$$

(Here $(1+Y)^X = \sum_{i=0}^\infty \frac{Y(Y-1)\dots(Y-i+1)}{i!}X^i$.) This series is contrived to satisfy the hypotheses of the two-variable version of the following.

Lemma 7.3. *Let $F(X) = 1 + \sum_{i=1}^\infty a_i X^i$, with $a_i \in \mathbb{Q}_p$ for each i. Then $a_i \in \mathbb{Z}_p$ for each i if and only if $F(X^p)/F(X)^p \in 1 + pX\mathbb{Z}_p[[X]]$.*

Proof. Suppose that $F(X) \in 1 + X\mathbb{Z}_p[[X]]$. Then we have $F(X)^p = F(X^p) + pG(X)$, for some $G(X) \in X\mathbb{Z}_p[[X]]$, hence $F(X^p)/F(X)^p = 1 - pG(X)/F(X)^p \in 1 + pX\mathbb{Z}_p[[X]]$.

Suppose conversely that $F(X^p) = F(X)^p G(X)$, with $G(X) \in 1 + pX\mathbb{Z}_p[[X]]$. Let $G(X) = 1 + \sum_{i=1}^\infty b_i X^i$. We prove by induction that $a_i \in \mathbb{Z}_p$.

Suppose $a_i \in \mathbb{Z}_p$ for $i < n$. Then equating the coefficients of X^n in $F(X^p) = F(X)^p G(X)$ gives

$$\left.\begin{array}{ll} a_{n/p} & \text{if p divides n} \\ 0 & \text{otherwise} \end{array}\right\} = \text{coefficient of } X^n \text{ in } \left(1 + \sum_{i=1}^n a_i X^i\right)^p \left(1 + \sum_{i=1}^n b_i X^i\right).$$

Re-arranging gives pa_n as a sum of terms in $p\mathbb{Z}_p$, hence $a_n \in \mathbb{Z}_p$. □

Let $F(X,Y) = \sum a_{m,n} X^n Y^m$. We calculate:

$$\frac{F(X^p,Y^p)}{F(X,Y)^p} = \frac{(1+Y^p)^{X^p}(1+Y^{p^2})^{(X^{p^2}-X^p)/p}(1+Y^{p^3})^{(X^{p^3}-X^{p^2})/p^2}\dots}{(1+Y)^{pX}(1+Y^p)^{X^p-X}(1+Y^{p^2})^{(X^{p^2}-X^p)/p}}$$

$$= \frac{(1+Y^p)^X}{(1+Y)^{pX}}.$$

We want to show that $(1+Y^p)^X/(1+Y)^{pX}$ lies in $1+pX\mathbb{Z}_p[[X,Y]]+pY\mathbb{Z}_p[[X,Y]]$. Let $(1+Y^p)/(1+Y)^p = 1+pYG(Y)$, $G(Y) \in \mathbb{Z}_p[[Y]]$. Then we have

$$\frac{(1+Y^p)^X}{(1+Y)^{pX}} = (1+pYG(Y))^X = \sum_{i=0}^{\infty} \frac{X(X-1)\dots(X-i+1)}{i!}(pYG(Y))^i.$$

Thus the hypotheses of the lemma are satisfied and we conclude $F(X,Y) \in \mathbb{Z}_p[[X,Y]]$.

Now let us view $F(X,Y) = \sum_{n=0}^{\infty}(X^n \sum_{m=n}^{\infty} a_{m,n}Y^m)$ as a series in X, with Y fixed. We take

$$\Theta(T) = F(T,\lambda) = \sum_{n=0}^{\infty} a_n T^n.$$

Note that $a_n = \sum_{m=n}^{\infty} a_{m,n}\lambda^m$ and hence $\operatorname{ord}_p a_n \geq n/(p-1)$. In particular, $\Theta(T)$ converges in the open disk $D(p^{1/(p-1),-})$.

Now let us consider once more the case where $a \in \mathbb{F}_q$ and $t = [a] \in \Omega$ is the corresponding Teichmuller digit. For fixed s we consider the formal power series

$$(1+Y)^{t+t^p+\dots+t^{p^{s-1}}} = F(t,Y)F(t^p,Y)\dots F(t^{p^{s-1}},Y).$$

(To see this identity, note that the right hand side is equal to

$$(1+Y)^{t+t^p+\dots+t^{p^{s-1}}}(1+Y^p)^{(t^{p^s}-t)/p}(1+Y^{p^2})^{(t^{p^{s+1}}-t^p)/p^2}\dots$$

Substituting $t^{p^s} = t$, we see that all terms except the first disappear). Then

$$\Theta(t)\Theta(t^p)\dots\Theta(t^{p^{s-1}}) = F(t,\lambda)F(t^p,\lambda)\dots F(t^{p^{s-1}},\lambda)$$

$$= (1+\lambda)^{t+t^p+\dots+t^{p^{s-1}}} = \epsilon^{\operatorname{tr} a},$$

which is what we wanted. We record this as a proposition.

Proposition 7.4. *There exists a series $\Theta(T) = \sum_{n=0}^{\infty} a_n T^n$ which converges in the open disk $D(p^{1/(p-1),-})$ and such that for all $a \in \mathbb{F}_{p^s}$, we have*

$$\Theta([a])\Theta([a]^p)\dots\Theta([a]^{p^{s-1}}) = \epsilon^{\operatorname{tr} a}.$$

Next time we will see how to glue these expressions into a single p-adic analytic function, which will eventually lead to a description of the zeta function as an entire meromorphic function.

Exercises to lecture 8

1. Read section IV.2 of the course text.
2. In this section is defined the Artin-Hasse exponential function $E_p(X)$. Compare with the power series $F(X,Y)$ defined above. Use Dwork's miracle lemma to give another proof, different to the one in the book, that $E_p(X)$ has coefficients in $\mathbb{Z}_p$.

3. Give another proof, different to the one above, that $F(X,Y)$ has coefficients in $\mathbb{Z}_p$, by expressing it in terms of the Artin-Hasse exponential function.

4. Compute the first p coefficients of the series $E_p(X)$. What fact in elementary number theory corresponds to the fact that the X^p coefficient lies in $\mathbb{Z}_p$?

5. In this lecture we made heavy use of the series

$$(1+X)^a = \sum_{i=0}^{\infty} \frac{a(a-1)\dots(a-i+1)}{i!} X^i$$

for $a \in \Omega$. Show that when $a \in \mathbb{Z}_p$, the coefficients of this series lie in $\mathbb{Z}_p$ (and so in particular, it converges in the disk $D(1^-)$).

Lecture 9

The space of overconvergent power series

Definition 8.1. *Let $R = \Omega[[X_1, \dots, X_n]]$ be the space of formal power series in n variables. We write*

$$U = \{(u_1, \dots, u_n) \mid \forall i, u_i \in \mathbb{Z}_{\geq 0}\}.$$

If $u = (u_1, \dots, u_n)$ then we write $|u| = \sum_{i=1}^n u_i$. Given $G(X_1, \dots, X_n) \in R$, we can write

$$G(X_1, \dots, X_n) = G(X) = \sum_{u \in U} g_u X^u.$$

We write $R_0 \subset R$ for the space of overconvergent power series

$$R_0 = \{G(X) = \sum_{u \in U} g_u X^u \in R \mid \exists M > 0, \operatorname{ord}_p g_u \geq M|u|, \forall u \in U\}.$$

The point is that the power series in R_0 converge in some disk strictly containing the closed disk of radius 1. We note that R_0 is closed under multiplication, so is a subring of R. The ring R_0 is important in several places in p-adic analysis; for example, in Monsky-Washnitzer's overconvergent cohomology. One of the features that is important in the background here is that it is actually a p-adic Fréchet space with a natural Schauder basis, given by the monomials X^u; in particular this means that the constructions with trace and determinant given below make sense.

We define several important operators on the space R_0. First, let q be a positive integer. We define T_q by

$$T_q\left(\sum_{u \in U} g_u X^u\right) = \sum_{u \in U} g_{qu} X^u.$$

If $G(X)$ is an element of R_0, then we also write G for the map $R_0 \to R_0$ given by multiplication by $G(X)$. We write $G_q(X) = G(X^q)$. (Note that in this case $G_q(X)$ also lies in R_0).

Finally, we write $\Psi_{q,G} = T_q \circ G : R_0 \to R_0$.

Lemma 8.2. *Let $G(X) = \sum_{u \in U} g_u X^u$. Then:*

1. $\Psi_{q,G}(X^u) = \sum_{v \in U} g_{qv-u} X^v$.
2. $G \circ T_q = T_q \circ G_q = \Psi_{q,G_q}$.

Proof. For the first part, we have

$$\Psi_{q,G}(X^u) = T_q \sum_{v \in U} g_{v-u} X^v = \sum_{v \in U} g_{qv-u} X^v.$$

For the second, we have

$$G \circ T_q(X^u) = \begin{cases} 0 & \text{if } q \nmid u \\ G \cdot X^{u/q} & \text{if } q \mid u \end{cases},$$

and in the latter case $G \cdot X^{u/q} = \sum_{v \in U} g_v X^{v+u/q} = \sum_{v \in U} g_{v-u/q} X^v$, while

$$\Psi_{q,G_q}(X^u) = \sum_{v \in U} g_{v-u/q} X^v.$$

(Note we are using the convention here that for $v \in \mathbb{Q}^n \supset U$, g_v has its usual meaning if $v \in U$ and is zero otherwise). □

Trace and determinant

Definition 8.3. *Let $A : R_0 \to R_0$ be a linear operator. Suppose that $A(X^u) = \sum_{v \in U} a_{v,u} X^v$. Then we write*

$$\operatorname{tr} A = \sum_{u \in U} a_{u,u}$$

for the trace *of A, when this sum exists.*

We do not worry about the question of whether this is basis independent here (although cf. the remarks above).

Lemma 8.4. *Let $G(X) \in R_0$, and let $\Psi = \Psi_{q,G}$. Then $\operatorname{tr} \Psi^s$ converges for each integer $s \geq 1$ and we have*

$$(q^s - 1)^n \operatorname{tr} \Psi^s = \sum_{x \in \Omega^n, x^{q^s - 1} = 1} G(x) \cdot G(x^q) \cdot \cdots \cdot G(x^{q^{s-1}}),$$

where we write $x = (x_1, \ldots, x_n), 1 = (1, \ldots, 1)$ and $x^{q^i} = (x_1^{q^i}, \ldots, x_n^{q^i})$.

Proof. We first treat the case $s = 1$. Then the identity we want to show is

$$(q-1)^n \operatorname{tr} \Psi = \sum_{x \in \Omega^n, x^{q-1} = 1} G(x).$$

By definition, we have

$$\operatorname{tr} \Psi = \sum_{u \in U} g_{(q-1)u},$$

and this sum exists since $G \in R_0$. Recall that

$$\sum_{x_i \in \Omega, x_i^{q-1} = 1} x_i^{w_i} = \begin{cases} q-1 & \text{if } q-1 \mid w_i \\ 0 & \text{otherwise.} \end{cases}$$

(This can be viewed as following from, for example, the orthogonality of characters for the cyclic group of $(q-1)^{\text{st}}$ roots of unity in Ω). Hence for $w \in U$, we have

$$\sum_{x \in \Omega^n, x^{q-1} = 1} x^w = \prod_{i=1}^n \left(\sum_{x_i^{q-1} = 1} x_i^{w_i} \right) = \begin{cases} (q-1)^n & \text{if } q-1 \mid w \\ 0 & \text{otherwise.} \end{cases}$$

(We write x^w here for the evaluation of X^w at the point x). Therefore

$$\sum_{x^{q-1}=1} G(x) = \sum_{w \in U} g_w \sum_{x^{q-1}=1} x^w = (q-1)^n \sum_{u \in U} g_{(q-1)u} = (q-1)^n \operatorname{tr} \Psi.$$

This proves the lemma in the case $s = 1$. For general s, we have by the previous lemma

$$\Psi^s = T_q \circ G \circ T_q \circ G \circ \Psi^{s-2} = T_{q^2} \circ G \circ G_q \circ \Psi^{s-2} = \cdots = \Psi_{q^s, G \cdot G_q \cdot \cdots G_{q^{s-1}}}.$$

Applying the above computation to this operator gives the result. □

Suppose that $A = (a_{i,j})$ is a linear map $\Omega^n \to \Omega^n$, and let T be an independent variable. Then $(1 - AT)$ is a matrix with entries in $\Omega[T]$, and we can form

$$\det(1 - AT) = \sum_{m=0}^{n} b_m T^m,$$

where

$$b_m = (-1)^m \sum_{\substack{1 \le u_1, \ldots, u_m \le n \\ \sigma \in S_m}} \left(\epsilon(\sigma) \prod_{i=1}^{m} a_{u_i, u_{\sigma(i)}} \right).$$

In this case we also have the identity of formal power series

$$\det(1 - AT) = \exp\left(- \sum_{s=1}^{\infty} \operatorname{tr} A^s T^s / s \right).$$

This can be deduced as follows. We can suppose that A is upper-triangular, with diagonal entries $a_1, \ldots, a_n$. Then the left hand side here is $\prod_{i=1}^{n} (1 - a_i T)$, while the right hand side is

$$\exp\left(- \sum_{s=1}^{\infty} \sum_{i=1}^{n} a_i^s T^s / s \right) = \exp\left(\sum_{i=1}^{n} \log(1 - a_i T) \right) = \prod_{i=1}^{n} (1 - a_i T).$$

Definition 8.5. *Let $A : R_0 \to R_0$ be a linear operator. Suppose that $A(X^u) = \sum_{v \in U} a_{v,u} X^v$. Then we write* $\det(1 - AT)$ *for the formal series*

$$\sum_{m=0}^{\infty} b_m T^m,$$

where

$$b_m = (-1)^m \sum_{\substack{u_1, \ldots, u_m \in U \\ \sigma \in S_m}} \left(\epsilon(\sigma) \prod_{i=1}^{m} a_{u_i, u_{\sigma(i)}} \right),$$

assuming that this sum exists.

The same remarks as above apply to the basis independence of this definition.

Lemma 8.6. *Suppose that $\Psi = \Psi_{q,G}$ for some $G \in R_0$. Then the formal series* $\det(1 - \Psi T)$ *exists, and has an infinite radius of convergence. Moreover, we have the identity of formal power series*

$$\det(1 - \Psi T) = \exp\left(- \sum_{s=1}^{\infty} \operatorname{tr} \Psi^s T^s / s \right).$$

Proof. To show that $\det(1-\Psi T)$ exists, we must show that the sums defining the coefficients b_m converge. Since $G = \sum_{u\in U} g_u X^u \in R_0$, we can choose $M > 0$ with $\operatorname{ord}_p g_u \geq M|u|$ for all $u \in U$. Hence for any $u_1, \ldots, u_m \in U$, and any permutation σ of $1, \ldots, m$, we have

$$\begin{aligned}
&\operatorname{ord}_p \left(g_{qu_{\sigma(1)}-u_1} \cdot g_{qu_{\sigma(2)}-u_2} \cdot \cdots \cdot g_{qu_{\sigma(m)}-u_m} \right) \\
&\geq M \left(|qu_{\sigma(1)} - u_1| + |qu_{\sigma(2)} - u_2| + \cdots + |qu_{\sigma(m)} - u_m| \right) \\
&\geq M \left(\sum_{i=1}^{m} q|u_{\sigma(i)}| - \sum_{i=1}^{m} |u_i| \right) = M(q-1) \sum_{i=1}^{m} |u_i|,
\end{aligned}$$

using that $|\cdot|$ is σ-invariant. Since there are only finitely many $u \in U$ with $|u|$ less than a given amount, this shows that the b_m and hence the series $\det(1-\Psi T)$ exists.

This also shows that $\operatorname{ord}_p b_m \to \infty$ as $m \to \infty$. In fact, we have $1/m \operatorname{ord}_p b_m \to \infty$ as $m \to \infty$, and this shows that $\det(1-\Psi T)$ has infinite radius of convergence (as if $\operatorname{ord}_p x = \lambda$ then

$$\operatorname{ord}_p b_m \lambda^m = m(1/m \operatorname{ord}_p b_m + \operatorname{ord}_p \lambda),$$

and this tends to infinity with m). The final identity follows by passing to the limit with respect to the version for finite matrices proved above. □

A particular example

Recall the following from last time:

Proposition 8.7. *There exists a series $\Theta(T) = \sum_{n=0}^{\infty} a_n T^n$ which converges in the open disk $D(p^{1/(p-1),-})$ and such that for all $a \in \mathbb{F}_{p^s}$, we have*

$$\Theta([a])\Theta([a]^p) \ldots \Theta([a]^{p^{s-1}}) = \epsilon^{\operatorname{tr} a}.$$

The following lemma will be essential.

Lemma 8.8. *Let $a \in D(1)$, and let $w \in U$. Then $\Theta(aX^w) \in R_0$.*

Proof. Recall that $\Theta(T) = F(T, \lambda) = \sum_j a_j T^j$, where $\operatorname{ord}_p a_j \geq j/(p-1)$. We must find $M > 0$ such that if we write

$$G(X) = \Theta(aX^w) = \sum_j a_j a^j X_1^{jw_1} \ldots X_n^{jw_n} = \sum_{u\in U} g_u X^u,$$

then $\operatorname{ord}_p g_u \geq M|u|$ for all $u \in U$. Taking $M = 1/(|w|(p-1))$ shows that $G \in R_0$. □

Next time we will apply the results of this and the previous lecture to find an expression for the zeta function as an entire p-adic meromorphic function.

Exercises to lecture 9

1. Prepare your talk for next week.

Lecture 10

This lecture we will finally give the proof of the following.

Theorem 9.1. *Let $X = H_f$ be an affine hypersurface over the finite field $\mathbb{F}_q$. Then the zeta function $Z(X,T)$ is a rational function.*

p-adic meromorphy

We begin by showing that $Z(X,T)$ defines an entire meromorphic function in Ω. We can suppose that $X \subset \mathbb{A}^n$. We begin by inducting on n as follows: write

$$Z(X,T) = \exp\left(\sum_{s=1}^{\infty} N_s T^s/s\right),$$

where $N_s = \#X(\mathbb{F}_{q^s})$. Define also

$$\begin{aligned} N'_s &= \#\{(x_1,\ldots,x_n) \in X(\mathbb{F}_{q^s}) \mid \forall i, x_i \neq 0\} \\ &= \#\{(x_1,\ldots,x_n) \in X \mid \forall i, x_i^{q^s-1} = 1\}, \end{aligned}$$

and

$$Z'(T) = \exp\left(\sum_{s=1}^{\infty} N'_s T^s/s\right).$$

Then we have $Z(X,T) = Z'(T) \cdot \exp\left(\sum_{s=1}^{\infty}(N_s - N'_s)T^s/s\right)$. Let $X_i = \{(x_1,\ldots,x_n) \in X \mid x_i = 0\}$. This can be viewed as an affine hypersurface in $\mathbb{A}^{n-1}$ (where we view $\mathbb{A}^{n-1} \subset \mathbb{A}^n$ as the co-ordinate hyperplane defined by $x_i = 0$). We have

$$N'_s = \#\left(X(\mathbb{F}_{q^s}) - \cup_{i=1}^n X_i(\mathbb{F}_{q^s})\right).$$

By the inclusion-exclusion principle,

$$\begin{aligned} N_s - N'_s = \sum_i \#X_i(\mathbb{F}_{q^s}) - \sum_{i<j} \#\left(X_i \cap X_j\right)(\mathbb{F}_{q^s}) \\ + \sum_{i<j<k} \#\left(X_i \cap X_j \cap X_k\right)(\mathbb{F}_{q^s}) - \ldots, \end{aligned}$$

and hence

$$\exp\left(\sum_{s=1}^{\infty}(N_s - N'_s)T^s/s\right) = \frac{\left(\prod_i Z(X_i,T)\right) \times \ldots}{\left(\prod_{i<j} Z(X_i \cap X_j, T)\right) \times \ldots}$$

(Note that $X_i \cap X_j$ can be viewed as a hypersurface in $\mathbb{A}^{n-2} \subset \mathbb{A}^n$, and so on). So by induction on n, $Z(X,T)$ will define an entire meromorphic function if the same is true for $Z'(T)$.

Now let us write $q = p^r$, and let ϵ be a primitive p^{th} root of unity. Recall that we have defined a power series $\Theta(T)$, convergent in the closed unit disk, such that for any $a \in \mathbb{F}_{q^s}, t = [a]$, we have

$$\epsilon^{\operatorname{tr} a} = \Theta(t) \cdot \Theta(t^p) \cdot \cdots \cdot \Theta(t^{p^{rs-1}}).$$

Now here comes the clever bit: for $u \in \mathbb{F}_{q^s}$, we have the relation

$$\sum_{x_0 \in \mathbb{F}_{q^s}} \epsilon^{\operatorname{tr} x_0 u} = \begin{cases} 0 & \text{if } u \neq 0 \\ q^s & \text{otherwise.} \end{cases}$$

This follows from orthogonality of characters for the abelian group $\mathbb{F}_{q^s}$. Subtracting away the term $x_0 = 0$, we have

$$\sum_{x_0 \in \mathbb{F}_{q^s}^{\times}} \epsilon^{\operatorname{tr} x_0 u} = \begin{cases} -1 & \text{if } u \neq 0 \\ q^s - 1 & \text{otherwise.} \end{cases}$$

Now apply this to $u = f(x_1,\ldots,x_n)$ and sum over $\mathbb{F}_{q^s}^{\times}$ to get the following relation:

$$\sum_{x_0,\ldots,x_n \in \mathbb{F}_{q^s}^{\times}} \epsilon^{\operatorname{tr} x_0 f(x_1,\ldots,x_n)} = q^s N'_s - (q^s - 1)^n.$$

Let $F(X_0,\dots,X_n)\in\Omega[X_0,\dots,X_n]$ be the polynomial whose coefficients are the Teichmuller lifts of those of $X_0 f(X_1,\dots,X_n)$. Write $F(X)=\sum_{i=1}^N a_i X^{w_i}$, for $w_i\in U$, with $a_i=[b_i]$, say. We then have

$$q^s N_s' = (q^s-1)^n + \sum_{x_0,\dots,x_n\in\mathbb{F}_{q^s}^\times}\prod_{i=1}^N \epsilon^{\operatorname{tr} b_i x^{w_i}}$$

$$= (q^s-1)^n + \sum_{\substack{x_0,\dots,x_n\in\Omega\\ \forall i, x_i^{q^s-1}=1}}\prod_{i=1}^N \left(\Theta(a_i x^{w_i})\cdot\Theta(a_i^p x^{pw_i})\cdot\dots\cdot\Theta(a_i^{p^{rs-1}}x^{p^{rs-1}w_i})\right).$$

Now we define

$$G(X)=G(X_0,\dots,X_n)=\prod_{i=1}^N \Theta(a_i X^{w_i})\cdot\Theta(a_i^p X^{pw_i})\cdot\dots\cdot\Theta(a_i^{p^{r-1}}X^{p^{r-1}w_i}).$$

Recall from last time that for any $a\in D(1)\subset\Omega$, $w\in U$, $\Theta(aX^w)$ is contained in the ring R_0 of overconvergent power series. Thus the same is true for G. In particular, the operator $\Psi=\Psi_{q,G}$ defined last time makes sense for this choice of G. This gives

$$q^s N_s' = (q^s-1)^n + \sum_{\substack{x_0,\dots,x_n\in\Omega\\ \forall i, x_i^{q^s-1}=1}}\prod_{i=1}^N \left(G(x)\cdot G(x^q)\cdot\dots\cdot G(x^{q^{s-1}})\right).$$

$$= (q^s-1)^n + (q^s-1)^{n+1}\operatorname{tr}\Psi^s$$

and hence

$$N_s' = \sum_{i=0}^n (-1)^n\binom{n}{i}q^{s(n-1-i)} + \sum_{i=0}^{n+1}(-1)^i\binom{n+1}{i}q^{s(n-i)}\operatorname{tr}\Psi^s.$$

Finally, we define

$$\Delta(T)=\det(1-\Psi T)=\exp\left(-\sum_{s=1}^\infty \operatorname{tr}\Psi^s T^s/s\right).$$

We saw last time that since $G\in R_0$, $\Delta(T)$ has in infinite radius of convergence, so defines an entire holomorphic function. It follows that

$$\begin{aligned} Z'(T) &= \exp\left(\sum_{s=1}^\infty N_s' T^s/s\right)\\ &= \prod_{i=0}^n\left\{\exp\left(\sum_{s=1}^\infty q^{s(n-1-i)}T^s/s\right)\right\}^{(-1)^i\binom{n}{i}} \times \prod_{i=0}^{n+1}\left\{\exp\left(\sum_{s=1}^\infty q^{s(n-i)}\operatorname{tr}\Psi^s T^s/s\right)\right\}^{(-1)^i\binom{n+1}{i}}\\ &= \prod_{i=0}^n (1-q^{n-1-i}T)^{(-1)^i\binom{n}{i}} \times \prod_{i=0}^{n+1}\Delta(q^{n-i}T)^{(-1)^{i+1}\binom{n+1}{i}}. \end{aligned}$$

The conclusion of the proof

We recall the rationality criterion proved two lectures ago.

Theorem 9.2. *Let $F(T)=1+\sum_{i=1}^\infty a_i T^i$ be a power series, and suppose that the a_i are* integers. *Suppose that F defines a holomorphic function in the disk $|z|<R$ in $\mathbb{C}$ and an entire meromorphic function in Ω (i.e. F is the quotient of two entire holomorphic functions in Ω). Then F is the power series expansion of a rational function.*

We've seen that the series $Z(X,T)$ is a formal power series with integer coefficients. The above shows that it defines an entire p-adic meromorphic function. Finally, we recall that if we write

$$Z(X,T) = 1 + \sum_{i=1}^{\infty} a_i T^i,$$

then $0 \leq a_i \leq q^{ni}$ (since X is contained in $\mathbb{A}^n$). It follows that $Z(X,T)$ defines a holomorphic function in the disk $|z| < 1/q^n$ in $\mathbb{C}$, and the above theorem applies. We conclude that $Z(X,T)$ is indeed a rational function.

The Weil conjectures

We conclude by briefly putting the function $Z(X,T)$ in a broader context. In fact, its rationality is only the first in an array of amazing properties. Suppose now that X is a projective variety of dimension n defined over $\mathbb{F}_q$, and that it is non-singular. (Projective varieties are discussed in Chapter V of the course text; they play the role of compact objects in this setting. Non-singularity is a mild non-degeneracy condition. Dimension is a measure of the number of independent parameters; thus a hypersurface in $\mathbb{A}^n$ will have dimension $n-1$. If you are not happy with these concepts, just think of affine varieties and the results we state here will not be too far from being true).

In 1949, Weil conjectured the following properties of $Z(X,T)$:

1. $Z(T)$ is a rational function.
2. Z satisfies a *functional equation*
$$Z(1/q^n T) = \pm q^{n\chi/2} T^{\chi} Z(T),$$
where χ is the *Euler characteristic* of the variety X. (This plays a very similar role to that of the Euler characteristic of a topological space in algebraic topology).
3. We can write
$$Z(T) = \frac{P_1(T)P_3(T)\dots P_{2n-1}(T)}{P_0(T)P_2(T)\dots P_{2n}(T)},$$
where $P_0(T) = 1 - T, P_{2n}(T) = 1 - q^{2n}(T)$ and
$$P_h(T) = \prod_{i=1}^{\beta_h} (1 - \alpha_{h,i} T),$$
where the $\alpha_{h,i}$ are algebraic integers of absolute value $q^{h/2}$. The numbers β_h are called the *Betti numbers* of the variety X. They should satisfy the relation $\sum_{i=0}^{2n} (-1)^i \beta_i = \chi(X)$.

Weil was led to these conjectures by extensive calculations with Fermat hypersurfaces and other varieties of interest related to Gauss sums. He seems to have been motivated by a connection with algebraic topology: indeed, one can view the above properties as corresponding to the existence of the Lefschetz trace formula and Poincaré duality!

Most amazingly of all, suppose that e.g. X is a hypersurface in $\mathbb{P}^n$ cut out by the reduction modulo p of a polynomial $f(X_0, \dots, X_n) \in \mathbb{Z}[X_0, \dots, X_n]$. Then as well as the finite sets

$$X(\mathbb{F}_{p^s}) = \{(x_0, \dots, x_n) \in \mathbb{F}_{p^s} \mid f(x_0, \dots, x_n) = 0\} / \sim$$

we have a complex manifold

$$X(\mathbb{C}) = \{(x_0, \dots, x_n) \in \mathbb{C} \mid f(x_0, \dots, x_n) = 0\} / \sim .$$

Then Weil conjectured that the Betti numbers of this complex manifold are precisely the numbers β_i defined above! Thus the topology of the complex manifold $X(\mathbb{C})$ determines the behaviour of the point counts $\#X(\mathbb{F}_{p^s})$.

The first progress on these conjectures was made by Dwork, who proved the rationality of $Z(X,T)$ in 1959 (and it is his proof we have given here). The rest of the conjectures had to wait until the construction by Grothendieck and Artin in the 1960's of *étale cohomology*, a cohomology theory that encapsulates both varieties over finite fields and complex manifolds, as well as many things in between.

Exercises to lecture 10

1. For those interested in the history of mathematics: have a look at Weil's paper [Wei49], and Dwork's paper [Dwo60].

2. Prepare for your tomorrow or Wednesday!

References

[Dwo60] Bernard Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. J. Math.*, 82:631–648, 1960.

[Kob77] Neal Koblitz. *p-adic numbers, p-adic analysis, and zeta-functions.* Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, Vol. 58.

[Wei49] André Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497–508, 1949.