Notes for Math 201A: Arithmetic of Local Fields
University of California, San Diego, Fall 2010 quarter
Kiran S. Kedlaya (kedlaya@ucsd.edu)

This cumulative notes file will be updated regularly throughout the quarter.

1 What and why: think globally, act locally

In geometry and analysis, we learn a great many things about spaces by studying them up close, in the vicinity of a single point. As algebraic number theory emerged as a coherent subject in the late 19th century, it was realized that much of this local thinking can be transferred to number theory, with fruitful consequences. The key definition is due to Hensel: for p a prime number, the *p-adic absolute value* $|\cdot|_p$ on the field $\mathbb{Q}$ of rational numbers is defined by setting $|0|_p = 0$ and

$$\left|p^m \frac{r}{s}\right|_p = p^{-m} \qquad (m, r, s \in \mathbb{Z}; \quad \gcd(r,p) = \gcd(s,p) = 1).$$

The reason this is called an *absolute value* is that it satisfies the *strong triangle inequality*: for all $x, y \in \mathbb{Q}$,

$$|x \pm y|_p \leq \max\{|x|_p, |y|_p\}.$$

This in particular implies the usual triangle inequality $|x + y|_p \leq |x|_p + |y|_p$. The strong triangle inequality is also called the *nonarchimedean triangle inequality*, because it means that the p-adic absolute value does not enjoy the usual archimedean property of positive real numbers: for any $x, y > 0$, there exists a positive integer n for which $nx > y$. By contrast, under the p-adic absolute value, the absolute values of integers are not unbounded; rather, they are all of absolute value at most 1. (Notice also that the p-adic absolute value is multiplicative: for all $x, y \in \mathbb{Q}$, $|xy|_p = |x|_p |y|_p$.)

Thanks to the triangle inequality, it makes sense to define the *completion* of $\mathbb{Q}$ with respect to $|\cdot|_p$, denoted $\mathbb{Q}_p$ and called the *field of p-adic numbers*. This can be done in much the same way as one defines $\mathbb{R}$ as the completion of $\mathbb{Q}$ for the usual absolute value, except that we can't refer to the ordering on $\mathbb{R}$. Here is an example of an explicit construction.

(a) Let R be the set of all Cauchy sequences $x_1, x_2, \ldots$ in $\mathbb{Q}$ for the absolute value $|\cdot|_p$. That is, a sequence $x_1, x_2, \ldots$ belongs to R if and only if for each $\epsilon > 0$, there exists $N > 0$ such that whenever $m, n \geq N$ we have $|x_m - x_n|_p < \epsilon$.

(b) Notice that R is closed under term-by-term addition and multiplication, and so forms a ring.

(c) Define an function $|\cdot|_p$ on R by setting $|x_1, x_2, \ldots|_p = \lim_{n\to\infty} |x_n|_p$. This limit exists by the triangle inequality (exercise) This function satisfies the strong triangle inequality and multiplicativity.

(d) The function $|\cdot|_p$ on R is not really an absolute value because it takes many elements of R to zero. In fact, the set $\mathfrak{p}$ of elements of R which map to zero is a prime ideal. Put $\mathbb{Q}_p = R/\mathfrak{p}$; this is again a ring, and $|\cdot|_p$ induces a true absolute value on R. Moreover, $\mathbb{Q}_p$ is not just a ring but a field, because any nonzero element has nonzero p-adic absolute value, so we can write down an inverse. Namely, if $x_1, x_2, \dots$ is a sequence representing a nonzero element of R, then $x_1^{-1}, x_2^{-1}, \dots$ is also Cauchy (exercise).

One can more concretely think of elements of $\mathbb{Q}_p$ as infinite sums

$$a_m p^m + a_{m+1} p^{m+1} + \cdots \qquad (a_m, a_{m+1}, \cdots \in \{0, \dots, p-1\}),$$

i.e., as numerals in base p whose expansions run infinitely far to the *left*. By contrast, real numbers may be identified with numerals in base p whose expansions run infinitely far to the *right*, provided that one disallows numerals in which after some point all of the digits equal $p-1$. No such issue arises here; in fact, the expression

$$(p-1) + (p-1)p + (p-1)p^2 + \cdots$$

defines an extremely important p-adic number, namely -1. (Roughly speaking, this is how your computer represents -1, using $p = 2$.)

If one takes the completion only of $\mathbb{Z}$ rather than $\mathbb{Q}$, one obtains a subring of $\mathbb{Q}_p$ denoted $\mathbb{Z}_p$ and called the *ring of p-adic integers*. It is not a field, but it is still a pretty simple ring: it is a principal ideal domain with only a single prime ideal (p). In fact, every ideal other than (0) has the form (p^n) for some nonnegative integer n. One often views $\mathbb{Z}_p$ as the *inverse limit* of the rings $\mathbb{Z}/p^n\mathbb{Z}$ for $n = 1, 2, \dots$; that is, an element of $\mathbb{Z}_p$ may be viewed as a sequence $(x_1, x_2, \dots)$ with $x_n \in \mathbb{Z}/p^n\mathbb{Z}$ which is coherent, i.e., x_{n+1} maps to x_n under the "reduction modulo p^n" map from $\mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$.

So why did Hensel do all of this? He realized that it could be extremely useful to be able to do analysis on $\mathbb{Q}_p$ just like one can do analysis on $\mathbb{R}$, because a lot of subtleties about $\mathbb{Q}$ disappear when you pass to the completion. For instance, you can talk about convergence of power series: for instance, in $\mathbb{Q}_p$, the "identity"

$$1 + 2 + 2^2 + 2^3 + ... = \frac{1}{1-2} = -1$$

of Euler becomes a true statement. An even more critical example for number theory is the following special case of *Hensel's lemma*: the fact that, say, 2 is a quadratic residue modulo 7 implies that 2 is also a quadratic residue modulo 7^n for all positive integers n, and hence that 2 is a perfect square in $\mathbb{Z}_p$ (and hence in $\mathbb{Q}_p$). Quick proof by induction on n: if there exists $x \in \mathbb{Z}$ with $x^2 \equiv 2 \pmod{7^n}$, then the equation $(x + 7^n y)^2 \equiv 2 \pmod{7^{n+1}}$ in y can be rewritten as $x^2 + 2 \cdot 7^n \cdot y \equiv 2 \pmod{7^{n+1}}$ and as $(x^2 - 2)/7^n + 2y \equiv 0 \pmod{7}$, and the latter obviously has a solution with $y \in \mathbb{Z}$. (We'll prove a more general statement of Hensel's lemma in the course.)

After Hensel introduced the p-adic numbers, it was discovered that they could be used to formulate some powerful *local-to-global* statements. One of the most spectacular of these is the *Hasse-Minkowski theorem*. (See Serre's *A Course in Arithmetic* for a proof.)

Theorem 1.1 (Hasse-Minkowski)**.** *Let n be a positive integer, and let A be an invertible symmetric $n \times n$ matrix over $\mathbb{Q}$. Define the quadratic form $P(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} A_{ij} x_i x_j$. Then the statement "there exists $(x_1, \ldots, x_n) \in F^n - \{0\}$ for which $P(x_1, \ldots, x_n) = 0$" holds for the field $F = \mathbb{Q}$ if and only if it holds for each of the fields $F = \mathbb{R}, \mathbb{Q}_2, \mathbb{Q}_3, \ldots$.*

Unfortunately, there are many natural instances of the *local-to-global principle* which do not hold; a famous example of Selmer is the equation $3x^3 + 4y^3 + 5z^3 = 0$ which has nonzero solutions over $\mathbb{R}, \mathbb{Q}_2, \mathbb{Q}_3, \ldots$ but not over $\mathbb{Q}$. Nonetheless, local-to-global methods are a key tool in solving Diophantine equations; for instance, the relationship between $\mathbb{Q}$ and its completions is essential to most methods of finding rational points on elliptic curves over $\mathbb{Q}$.

In this course, we will focus on the arithmetic, and largely the Galois theory, of fields like $\mathbb{Q}_p$. One reason for this is that the absolute Galois group of $\mathbb{Q}_p$ (i.e., the Galois group of an algebraic closure of $\mathbb{Q}_p$) may be viewed as a subgroup of the absolute Galois group of $\mathbb{Q}$, and much of what we understand about the latter is gathered from the former. This has to do with the fact that while it is believed that *every* finite group occurs as the Galois group of some finite extension of $\mathbb{Q}$ (the inverse Galois problem), we will prove that the Galois groups of finite extensions of $\mathbb{Q}_p$ are all *solvable* groups, and in fact have a fairly special form.

An explicit instance of the principle that Galois theory over $\mathbb{Q}$ is best understood via the $\mathbb{Q}_p$ is *class field theory*: the abelianization of the absolute Galois group of any number field K (i.e., the Galois group of the maximal abelian extension of K) can be described explicitly using a recipe cooked up from the completions of K (which are finite extensions of $\mathbb{R}$ and $\mathbb{Q}_p$, which we'll describe later). To prove class field theory, one first proves *local class field theory*, which is a description of the abelianization of the absolute Galois group of a finite extension of $\mathbb{Q}_p$. As you might imagine, this is much easier! Stating and proving local class field theory will provide an excuse to introduce many important tools in number theory, such as *Galois cohomology*.

Exercises

1. Prove that if $x_1, x_2, \ldots$ is a Cauchy sequence with respect to $|\cdot|_p$, then $\lim_{n\to\infty} |x_n|_p$ exists.

2. Prove that if $x_1, x_2, \ldots$ is a Cauchy sequence with respect to $|\cdot|_p$ representing a nonzero element of $\mathbb{Q}_p$, then $x_1^{-1}, x_2^{-1}, \ldots$ is also Cauchy.

3. I mentioned in class that the Hasse-Minkowski theorem provides a criterion for checking for the existence of a zero of a quadratic form which involves a finite amount of computation. Here is an example of this. Let A, B, C be pairwise coprime integers, and put $P(x, y, z) = Ax^2 + By^2 + Cz^2$. Prove that for any p not dividing $2ABC$, there exist $x, y, z \in \mathbb{Q}_p$ not all zero for which $P(x, y, z) = 0$. (Hint: for $p \equiv 1 \pmod 4$, find a solution with $xyz = 0$ using the special case of Hensel's lemma introduced in class. For $p \equiv 3 \pmod 4$, use the same argument to handle all cases but $P(x, y, z) = x^2 + y^2 + z^2$, for which you know solutions over $\mathbb{Q}$.)

2 Discrete valuation rings and fields

Warning: in this course, a *ring* will always be a *commutative ring with unit* unless I say otherwise (which I don't think I ever will).

A *discrete valuation ring* (DVR, not to be confused with TiVo) is a principal ideal domain with exactly one nonzero prime ideal. For example, the ring $\mathbb{Z}_{(p)}$ consisting of all rational numbers r/s with s not divisible by p is a DVR. This is not to be confused with the ring $\mathbb{Z}_p$ of p-adic integers, although that is also a DVR. Another example is the formal power series ring $K[\![t]\!]$ over a field K.

Let's take a moment to gather some properties from this definition. Let R be a DVR. Since R has a nonzero prime ideal, it can't be either the zero ring or a field. It must then have a maximal ideal, which cannot be the zero ideal, but is definitely a prime ideal. Hence our one nonzero prime ideal must be maximal; call it I. The ideal I must be principal; let π be a generator. Since I is the only maximal ideal of R, R is a local ring; that is, any element of R not in I is a unit. Therefore, if $x \in R$ is not a unit, then x is divisible by π. Similarly, if x is divisible by π, then either x/π is a unit or x is divisible by π^2 and so on. This means that every element of x is either of the form $\pi^m u$ for some nonnegative integer m and some unit u, or divisible by *every* positive power of π. But only 0 fits the latter description: otherwise, the intersection $(\pi) \cap (\pi^2) \cap (\pi^3) \cap \cdots$ would be a nonzero prime ideal (exercise).

Since R was assumed to be an integral domain, it has a fraction field K. Each nonzero element of K can be written as $\pi^m u$ for some $m \in \mathbb{Z}$ and some unit u in R. We define the *valuation function* $v : K^\times \to \mathbb{Z}$ to send $\pi^m u$ to m. (By convention, $v(0) = +\infty$.) This function behaves like the logarithm of a nonarchimedean absolute value, in that it satisfies the following properties.

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

(b) For $x, y \in K$, $v(x + y) \geq \min\{v(x), v(y)\}$.

(c) For $x, y \in K$, $v(xy) = v(x) + v(y)$.

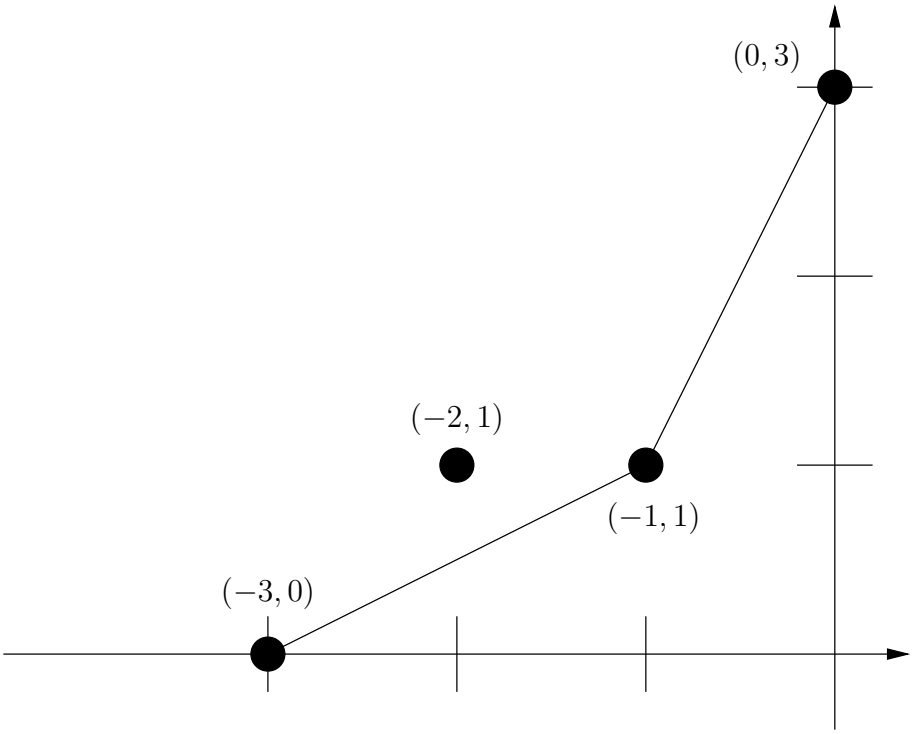
Conversely, given a surjective map $v : K \to \mathbb{Z} \cup \{+\infty\}$ satisfying these conditions, the subset $R = \{x \in K : v(x) \geq 0\}$ forms a discrete valuation ring.

For those who like commutative algebra, here is another characterization of discrete valuation rings. See Serre, §1.2 for a proof.

Proposition 2.1. *A ring R is a discrete valuation ring if and only if it is a noetherian local ring whose maximal ideal is generated by an element which is not nilpotent.*

One of the basic tools for studying valuations is the theory of *Newton polygons*. Let R, K be as above, and let $P \in K[T]$ be a polynomial. The Newton polygon of P is constructed as follows. Write $P = \sum_i P_i T^i$. For each index i, plot the point $(-i, v(P_i))$ in the xy-plane. The Newton polygon is then defined as the boundary of the *lower convex hull* of these points. In other words, it is the graph of a convex ("holds water") function that lies below all of the drawn points but otherwise is as high as possible.

An example might help make this clear. Take $K = \mathbb{Q}_2$ and $P = T^3 + 6T^2 + 10T + 8$. Here's the Newton polygon:



The relevant data from the polygon are the slopes that occur and their *multiplicities* (the horizontal widths of the corresponding segments). In the above example, the slope $1/2$ occurs with multiplicity 2 and the slope 2 occurs with multiplicity 1.

The Newton polygon has the following multiplicativity property.

Proposition 2.2. *If $P = QR$, then the Newton polygon of P is the* union *of the Newton polygons of Q and R. This means that for each slope s, the multiplicity of s as a slope of P is the sum of the multiplicities of s as a slope of Q and R.*

For example, $(T+2)^2 = T^2+4T+4$ has only the slope 1 with multiplicity 2. For another example,

$$T^3 + 6T^2 + 10T + 8 = (T^2 + 2T + 2)(T + 4)$$

and the two factors on the right account for the slopes $1/2$ and 2, respectively.

In case P splits completely into linear factors, this property tells us that the slopes of the Newton polygon compute the valuations of the roots. We will use Newton polygons a bit later to extend the valuation function from a local field to a finite extension.

To prove Proposition 2.2, we use the following definition. For $s \in \mathbb{R}$, let $v_s : K[T] \to \mathbb{R} \cup \{+\infty\}$ be the function taking $P = \sum_i P_i$ to $\min_i\{v(P_i) + si\}$. This function has the evident properties that:

(a) $v_s(P) = +\infty$ if and only if $P = 0$;

(b) $v_s(P+Q) \geq \min\{v_s(P), v_s(Q)\}$.

It also has the following geometric interpretation: it is the y-intercept of the supporting line of slope s to the Newton polygon of P. (Namely, we can find this supporting line by drawing a line of slope s through each point $(-i, v(P_i))$ and finding the one that hits the y-axis at the lowest point.)

Proposition 2.2 now follows from the following lemma.

Lemma 2.3. *The function v_s has the property that $v_s(PQ) = v_s(P) + v_s(Q)$. (That is, v_s is a valuation Moreover, if $i_{\min}, i_{\max}$ are the minimum and maximum indices i, respectively, for which $v(P_i) + si$ achieves its minimum value, and $j_{\min}, j_{\max}$ are minimum and maximum indices j, respectively, for which $v(Q_j) + sj$ achieves its minimum value, then $k_{\min} = i_{\min} + j_{\min}, k_{\max} = i_{\max} + j_{\max}$ are the minimum and maximum indices k, respectively, for which $v((PQ)_k) + sk$ achieves its minimum value.*

Proof. For any k, we have $(PQ)_k = \sum_{i+j=k} P_i Q_j$. Consequently,

$$v((PQ)_k) + sk \geq \min_{i+j=k} \{v(P_i) + v(Q_j) + s(i+j)\} \geq v_s(P) + v_s(Q).$$

If $k = k_{\min}$, then the sum $\sum_{i+j=k} P_i Q_j$ consists of the unique term with $i = i_{\min}$ and $j = j_{\min}$, plus some terms with $i < i_{\min}$, plus some terms with $j < j_{\min}$. Any term of the second type has the property that

$$v(P_i) + is > v_s(P), v(Q_j) + js \geq v_s(Q),$$

so $v(P_i Q_j) > v(P_{i_{\min}} Q_{j_{\min}})$. Likewise for the third type. This means that $(PQ)_{k_{\min}}$ consists of $P_{i_{\min}} Q_{j_{\min}}$ plus a bunch of terms of larger valuation, so $v((PQ)_{k_{\min}}) = v(P_{i_{\min}} Q_{j_{\min}})$ and so $v((PQ)_{k_{\min}}) + sk_{\min} = v_s(P) + v_s(Q)$. Similarly for $k = k_{\max}$. If $k < k_{\min}$, then every term in the sum $\sum_{i+j=k} P_i Q_j$ has either $i < i_{\min}$ or $j < j_{\min}$ or both, so $v((PQ)_k) + sk > v_s(P) + v_s(Q)$. Likewise for $k > k_{\max}$. □

Exercises

1. Let R be a noetherian domain. Prove that for any $x \in R$, the intersection $(x) \cap (x^2) \cap \cdots$ is a prime ideal of R.

2. Use Newton polygons to prove *Eisenstein's irreducibility criterion*: let $P \in \mathbb{Z}[T]$ be a monic polynomial. Suppose for some prime number p, all of the coefficients of P except the leading coefficient are divisible by p, and the constant coefficient is not divisible by p^2. Prove that P is irreducible as a polynomial over $\mathbb{Q}$ (and even over $\mathbb{Q}_p$). Polynomials for which this criterion holds, called *Eisenstein polynomials*, play a key role in the construction of local fields; see for instance the next exercise.

3. Check that the polynomial
$$\frac{(T+1)^p - 1}{T}$$
satisfies the Eisenstein criterion.

3 Hensel's lemma

Throughout this lecture, let K be the fraction field of a discrete valuation ring R with maximal ideal I. We will assume that K is a *local field*, meaning that R is *complete* for the adic topology defined by its maximal ideal. In other words, R and K are complete with respect to the absolute value $|x| = e^{-v(x)}$, for $v : K \to \mathbb{Z} \cup \{+\infty\}$ the valuation function.

We are going to investigate Hensel's fundamental observations about factoring polynomials over local fields. There are several of these, all sometimes called *Hensel's lemma*.

Proposition 3.1 (Hensel's lemma, version 1)**.** *Let $P \in R[T]$ be a polynomial. Suppose that $z \in R$ has the property that $P(z) \in I$ but $P'(z) \notin I$. Then there exists a unique $r \in R$ with $z \equiv r \pmod{I}$ and $P(r) = 0$.*

Proof. We construct r using Newton's method (also called the Newton-Raphson method) from one-variable calculus. Define a sequence $r_0, r_1, \ldots$ as follows. First put $r_0 = z$. Given r_n, we define

$$r_{n+1} = r_n - \frac{P(r_n)}{P'(r_n)}.$$

We will prove by induction on n that:

(a) $P(r_n) \equiv 0 \pmod{I^{2^n}}$ and $P'(r_n) \not\equiv 0 \pmod{I}$;

(b) $r_{n+1} \equiv r_n \pmod{I^{2^n}}$.

As the base case, we have (a) for $n = 0$ by definition. For any given n, having (a) implies (b), so to check the induction it is enough to assume (a) and (b) for a given n and then deduce the next case of (a). Given (a) and (b), it is clear that $P'(r_{n+1}) \equiv P'(r_n) \not\equiv 0 \pmod{I}$. On the other hand, by Taylor's approximation we have

$$P(r_{n+1}) \equiv P(r_n) + P'(r_n)(r_{n+1} - r_n) = 0 \pmod{(r_{n+1} - r_n)^2}.$$

Since $r_{n+1} - r_n$ belongs to I^{2^n}, this gives a congruence modulo $I^{2^{n+1}}$.

The sequence $r_0, r_1, \ldots$ converges to some $r \in R$ for which $P(r) \equiv 0 \pmod{I^{2^n}}$ for all n, and so $P(r) = 0$. To prove uniqueness, suppose $r' \in R$ also satisfies $z \equiv r' \pmod{I}$ and $P(r') = 0$. Again by Taylor's approximation, we have $P(r') \equiv P(r) + P'(r)(r' - r) \pmod{(r' - r)^2}$, which means that $P'(r)(r' - r) \equiv 0 \pmod{(r - r')^2}$. However, we have a problem: $P'(r)$ is a unit, so the congruence would force $r' - r$ to be divisible by its square. This is only possible if $r' - r$ has valuation either 0 or $+\infty$, i.e., if $r - r'$ is either zero or a unit. Since $r \equiv r' \pmod{I}$, $r - r'$ can't be a unit, so it is zero. □

The fact that we got from a congruence modulo I^{2^n} to a congruence modulo $I^{2^{n+1}}$ was not really necessary for proving the proposition; it would have been enough to get from I^n to I^{n+1}. However, just as happens over the real numbers, the exponential convergence of Newton's method is very useful when finding roots of p-adic polynomials on a computer.

When the condition $P'(z) \notin I$ fails to be satisfied, one needs a slightly better approximation in order for Newton's method to converge to a root. This amounts to a second, slightly stronger form of Hensel's lemma.

Proposition 3.2 (Hensel's lemma, version 2)**.** *Let $P \in R[T]$ be a polynomial. Suppose that $z \in R$ has the property that $P(z)/P'(z)^2 \in I$. Then there exists a unique $r \in R$ with $z \equiv r \pmod{P(z)/P'(z)}$ and $P(r) = 0$.*

Proof. Exercise. □

Here is a somewhat different form of Hensel's lemma. It can be derived from the first form, but I'll give a separate proof instead.

Proposition 3.3 (Hensel's lemma, version 3)**.** *Let $P \in R[T]$ be a monic polynomial, and suppose that $Q_1, S_1 \in R[T]$ are monic polynomials for which $P \equiv Q_1 S_1 \pmod{I}$. (That is, $P - Q_1 S_1$ has all coefficients in I.) Suppose also that Q_1 and S_1 are relatively prime as polynomials over the field R/I. Then there exist unique monic polynomials Q, S congruent to Q_1, S_1 modulo I for which $P = QS$.*

Proof. This time around, I'll be a bit less verbose (and I'll skip the uniqueness). I'll just argue that given a congruence $P \equiv Q_n S_n \pmod{I^n}$ with Q_n and S_n monic and relatively prime as polynomials over R/I, I can produce another congruence $P \equiv Q_{n+1} S_{n+1} \pmod{I^{n+1}}$ with $Q_{n+1} \equiv Q_n \pmod{I^n}$ and $S_{n+1} \equiv S_n \pmod{I^n}$.

Let π be a uniformizer of R. The plan now is to put $Q_{n+1} = Q_n + \pi^n X_n$ and $S_{n+1} = S_n + \pi^n Y_n$ for some well-chosen polynomials X_n and Y_n. To avoid changing the degree or leading coefficient, we insist that $\deg(X_n) < \deg(Q_n)$ and $\deg(Y_n) < \deg(S_n)$. The condition we'd like to achieve is that

$$P \equiv (Q_n + \pi^n X_n)(S_n + \pi^n Y_n) \pmod{\pi^{n+1}};$$

note that this can be rewritten as

$$(P - Q_n S_n) \equiv \pi^n (X_n S_n + Y_n Q_n) \pmod{\pi^{n+1}}$$

since $2n \geq n + 1$. To do this, put $Z_n = (P - Q_n S_n)/\pi^n$. The relatively prime condition modulo I means that I can produce polynomials A_n, B_n with $A_n Q_n + B_n S_n \equiv 1 \pmod{I}$. Now let X_n be the remainder upon dividing $Z_n B_n$ by Q_n, and then solve for Y_n. □

Finally, here is yet another form of Hensel's lemma which I'll need to extend valuations to finite extensions of K. By now, you should have the basic idea, so I can leave even more details as an exercise.

Proposition 3.4 (Hensel's lemma, version 4)**.** *Let $P \in K[T]$ be a polynomial whose Newton polygon contains more than one distinct slope. Then P is reducible.*

Proof. This is most easily proved using the fact (Proposition 2.2) that for any $s \in \mathbb{R}$, the function v_s taking $P = \sum_i P_i$ to $\min_i\{v(P_i) + si\}$ is a (not necessarily discrete) valuation on $K[T]$. Given this, choose an index i for which $(-i, v(P_i))$ is a vertex of the Newton polygon between the segments of slope s_1 and s_2. Choose any s in the open interval (s_1, s_2). The idea will be to construct a sequence of approximate factorizations $Q_n S_n$ of P for which

$v_s(P - Q_n S_n)$ keeps decreasing. We start by taking $Q_0 = P_i, S_0 = T^i$. Given Q_n and S_n, split $P - Q_n S_n$ as a sum $P_i X_n + T^i Y_n$ where $\deg(X_n) < i$. Then define

$$Q_{n+1} = Q_n + Y_n, S_{n+1} = S_n + X_n.$$

We claim that this converges to a good factorization; we leave this as an exercise. □

Corollary 3.5. *Any irreducible polynomial in $K[T]$ has only one slope in its Newton polygon (with some multiplicity).*

Beware that this form of Hensel's lemma does not allow you to necessary separate two *equal* slopes from each other. For instance, the polynomial

$$P(T) = T^2 + 1$$

over $\mathbb{Q}_3$ is irreducible even though its Newton polygon consists of the slope 0 with multiplicity 2, which can be written as the union of two polygons each having slope 0 with multiplicity 1. The irreducibility can be seen by noting that any factorization corresponds to a factorization over $\mathbb{Z}_3$ (Gauss's lemma; see exercises) and hence to one over $\mathbb{F}_3$, but the polynomial modulo 3 is irreducible because -1 is not a quadratic residue modulo 3.

Stylistic note: some authors prefer to derive Hensel's lemma using the *contraction mapping theorem* (as in the usual proof of the implicit function theorem). That usually gives shorter proofs, but I find the arguments using sequences of approximations a bit easier to digest the first time around.

Exercises

1. Prove that a polynomial in $R[T]$ is irreducible if and only if it is irreducible in $K[T]$. (This is sometimes called *Gauss's lemma* because Gauss made the same observation for $R = \mathbb{Z}$.)

2. Prove version 2 of Hensel's lemma. Then use it to explain why every integer congruent to 1 modulo 8 is a perfect square in $\mathbb{Q}_2$.

3. Finish the proof of version 4 of Hensel's lemma.

4 Extending valuations

Again, let R be a discrete valuation ring with fraction field K and valuation function v.

Proposition 4.1. *Suppose that K is complete. Then for any finite extension L of K, there is a unique way to extend the valuation function on K to a valuation function on L taking values not in $\mathbb{Z}$ but in $\frac{1}{N}\mathbb{Z}$ for some positive integer N.*

For example, for $K = \mathbb{Q}_p$ and $L = \mathbb{Q}_p(\sqrt{p})$, the extended valuation function will take p to $1/2$. We can always take $N = [L : K]$, the degree of L over K; that will be evident from the uniqueness proof.

Proof. Let's first check the uniqueness property using Newton polygons. Take any $x \in L$, and let $Q \in K[T]$ be the minimal polynomial of x. Since Q is irreducible, by one of the forms of Hensel's lemma, it has only one slope s in its Newton polygon. This means that the extended valuation must take x to s, or else $T - x$ could not be a factor of Q in $L[T]$. (Note that the degree of Q is the degree $[K(x) : K]$, which divides $[L : K]$.)

Now let's turn around and use the same thinking to construct the extended valuation. For each $x \in L$, define $v(x)$ to be the unique slope in the Newton polygon of the minimal polynomial of x. This definitely agrees with the original valuation for $x \in K$. To check that $v(x + y) \geq \min\{v(x), v(y)\}$, it is enough to produce a polynomial with $x + y$ as a root having all slopes in its Newton polygon at least $v(x) + v(y)$. To do this, let P and Q be the minimal polynomials of x and y, respectively. By Hensel's lemma, version 4 (Proposition 3.4), the polynomials P and Q have only one slope each, say s and t. Put $m = \deg(P)$ and $n = \deg(Q)$, and write $P = \sum_i P_i T^i$ and $Q = \sum_j Q_j T^j$. We then have the following homogeneity property:

$$v(P_{m-i}) \geq si, \qquad v(Q_{n-j}) \geq tj.$$

Over an algebraic closure of F, we may factor $P = (T - x_1) \cdots (T - x_m)$ and $Q = (T - y_1) \cdots (T - y_n)$. The polynomial

$$A(T) = \prod_{i=1}^{m} \prod_{j=1}^{n} (T - x_i - y_j)$$

then has coefficients in F and has $x+y$ as a root. The coefficient of T^{mn-k} in $A(T)$ is a symmetric polynomial in the x_i and in the y_j of total degree k, so it can be written as a polynomial in the P_i and Q_j with integer coefficients. For each term $P_{m-i_1} P_{m-i_2} \cdots Q_{n-j_1} Q_{n-j_2} \ldots$, the sum of the i's and j's is k, so the valuation is at least $\min\{s, t\}k$. But now we can turn around and say that the slopes of A are all at least $\min\{s, t\}$, because the line through the point $(-mn, 0)$ of slope $\min\{s, t\}$ lies on or below all of the points $(-k, v(A_k))$.

To check that $v(xy) = v(x) + v(y)$, we make a similar argument using the polynomial

$$M(T) = \prod_{i=1}^{m} \prod_{j=1}^{n} (T - x_i y_j).$$

This time, the coefficient of T^{mn-k} in $M(T)$ is a symmetric polynomial of degree k in the x_i and separately of degree k in the y_j, so again it can be written as a polynomial in the P_i and Q_j with integer coefficients. For each term $P_{m-i_1} P_{m-i_2} \cdots Q_{n-j_1} Q_{n-j_2} \ldots$, the sum of the i-indices is k and the sum of the j-indices is also k, so the valuation is at least $(s+t)k$. That means that the slopes of M are all at least $s+t$. But we have one extra piece of information: the final coefficient is $\pm P_0^n Q_0^m$, whose valuation is exactly $(s+t)mn$. So the slopes of M are forced to be exactly equal to $s + t$, completing the argument. □

Warning: we made heavy use of the completeness of K here. If I only assume that K is the fraction field of a discrete valuation ring, there may be multiple ways to extend the valuation to a finite extension (see exercises). This is very closely related to splitting of primes in extensions of number fields, but we'll come back to that connection a bit later.

One fact I forgot to mention earlier...

Proposition 4.2. *Let L be a finite extension of K. Then L is also complete with respect to the unique extension of the valuation on K.*

Proof. It is convenient to prove something stronger. Let V be any finite-dimensional vector space over K equipped with a function $v_V : V \to \frac{1}{N}\mathbb{Z} \cup \{+\infty\}$ for some positive integer N satisfying the following conditions.

(a) For $x \in V$, $v_V(x) = +\infty$ if and only if $x = 0$.

(b) For $x, y \in V$, $v_V(x + y) \geq \min\{v_N(x), v_V(y)\}$.

(c) For $x \in V$ and $c \in K$, $v_V(cx) = v(c) + v_V(x)$.

We will show that V is complete, i.e., any Cauchy sequence with respect to v_V has a limit. This includes the desired result, but it has the advantage that I can induct on the dimension of V (which provides more intermediate steps).

The case $\dim(V) = 1$ is clear: we may identify V with K, and then $v_V(x) = v_V(1) + v(x)$ for all $x \in K$. If $\dim(V) > 1$, choose any nonzero $x \in V$, and let W be the quotient vector space V/Kx. We can define a new valuation v_W on W by taking $v_W(y)$ to be the maximum of $v_V(z)$ over all $z \in V$ mapping to y in W. Why does the maximum exist? If it didn't, then since v_V takes values in a discrete group, we'd have some elements $z_0, z_1, z_2, \cdots \in V$ having the same image in W with $v_V(z_i) \to +\infty$ as $i \to \infty$. But $z_{i+1} - z_i = c_i x$ for some $c_i \in K$, so $v(c_i) \to \infty$ and so the sum $\sum_i c_i$ converges to a limit c. Then $z_0 + cx$ would be an element of V of infinite valuation, so the maximum exists after all (and happens to be infinite), and incidentally $y = 0$.

By the induction hypothesis, W is complete. If $x_1, x_2, \cdots$ is a Cauchy sequence in V, then the images of the x_i in W must converge to some limit y. For each i, choose $z_i \in V$ lifting $x_i - y$ with $v_V(z_i) = v_W(x_i - y)$; these converge to zero, so the sequence $x_i - z_i$ is again Cauchy. But these terms all map to y in V, so for any one lift z of y from W to V, I can write $x_i - z_i = z + c_i x$ for some $c_i \in K$, and now the c_i form a Cauchy sequence. That sequence has a limit c, so the x_i converge to the limit $z + cx$. □

Exercises

1. Put $K = \mathbb{Q}_p$, $L = \mathbb{Q}_p[T]/(T^2 - p)$. Prove that the extension of the p-adic valuation to L takes $a + bT$ to $\min\{v_p(a), \frac{1}{2} + v_p(b)\}$; in particular, it does not take values in $\mathbb{Z} \cup \{+\infty\}$. (Hint: by the uniqueness part of Proposition 4.1, it is enough to check that this satisfies the properties of a valuation.)

2. Put $K = \mathbb{Q}_3$, $L = \mathbb{Q}_3[T]/(T^2+1)$. Prove that the extension of the p-adic valuation to L takes $a + bT$ to $\min\{v_3(a), v_3(b)\}$; in particular, it again takes values in $\mathbb{Z} \cup \{+\infty\}$. (Hint: same hint as the previous exercise.)

3. Prove that the 5-adic valuation on $\mathbb{Q}$ can be extended in two different ways to $\mathbb{Q}[T]/(T^2+1)$. (Hint: construct an extension in which $2+i$ has positive valuation, then observe that this valuation is not invariant under complex conjugation because $2-i$ does not have positive valuation.)

5 Structure of complete discrete valuation rings, I: Equal characteristic case

Once again, let R be a complete discrete valuation ring with fraction field K. Let k denote the residue field of R, i.e., the quotient of R by its maximal ideal I. Choose a uniformizer π.

Lemma 5.1. *Let S be any set of coset representatives of k in R. Then every element x of R can be written uniquely as a convergent sum $\sum_{n=0}^{\infty} s_n \pi^n$ with each s_n belonging to S.*

Proof. By the definition of S, there is a unique choice of $s_0 \in S$ congruent to x modulo π. There is also a unique choice of $s_1 \in S$ congruent to $(x - s_0)/\pi$ modulo π, and so forth. □

For example, as we've already seen, every element of $\mathbb{Z}_p$ can be written as a convergent sum $s_0 + s_1 p + s_2 p^2 + \cdots$ with $s_i \in \{0, \ldots, p-1\}$. We will see an arguably better choice for the "digits" later in this section.

Before stating the next result, let me toss in a quick reminder from Galois theory. Recall that a finite extensions of fields E/F is *separable* if the trace pairing $E \times E \to F$ taking (x, y) to $\mathrm{Trace}_{E/F}(xy)$ (the trace of multiplication by xy viewed as an F-linear endomorphism of E) is nondegenerate. By the primitive element theorem, this happens if and only if we can write $E = F[T]/(P(T))$ for some polynomial $P(T)$ which is separable, i.e., for which $\gcd(P(T), P'(T)) = 1$. This condition is automatic in characteristic 0. Also, a field F is *perfect* if every finite extension of it is separable; again, this is automatic in characteristic 0. In characteristic p, this happens if and only if the p-power endomorphism of F (the *Frobenius map*) is surjective (and hence bijective, since a homomorphism of fields cannot have a nonzero kernel). For instance, any finite field is perfect.

Proposition 5.2. *Suppose that K and k are of the same characteristic, and k is perfect. Then there is an isomorphism $k[\![T]\!] \cong R$ taking T to π and inducing the identity map on residue fields. If k is of characteristic $p > 0$, then the isomorphism is unique.*

Recall that k is defined as a quotient of R, not a subring. However, if we had an isomorphism of the desired form, this would provide a copy of k sitting inside R, since k sits as a subring in $k[\![T]\!]$. Conversely, if we have a homomorphism $f : k \to R$ such that the map $k \to R \to k$ is an isomorphism, we get an isomorphism $k[\![T]\!] \cong R$ taking $s_0 + s_1 T + \cdots$ to $f(s_0) + f(s_1)\pi + \cdots$. We thus need to prove the following.

Lemma 5.3. *Suppose that K and k are of the same characteristic, and k is perfect. Then there is a homomorphism $k \to R$ for which the composition $k \to R \to k$ is the identity map. If k is of characteristic $p > 0$, then the homomorphism is unique.*

Proof in characteristic 0. We first produce a subfield F of k and a map $F \to R$ such that the composition $F \to R \to k$ coincides with the embedding of F into k. In fact, we may take $F = \mathbb{Q}$: this works because every nonzero element of $\mathbb{Z}$ has nonzero image under the map $R \to k$, so must be a unit in R.

Suppose I have such a subfield F (not necessarily equal to $\mathbb{Q}$). I claim that if $F \neq k$, then I can extend my map $F \to R$ to a map $E \to R$ for some subfield E of k strictly containing F. To see this, pick any $x \in k - F$ and put $E = F(x)$.

(a) If x is transcendental over F, then *any* choice of an element $y \in R$ lifting x defines a map $E \to R$. That is because the map $F[T] \to R$ sending T to y takes every nonzero element of $F[T]$ to an element of R of valuation zero (since the composition $F[T] \to R \to E$ is injective), so we get a map $F(y) \to R$.

(b) If x is algebraic over F, then the minimal polynomial $P(T)$ of x over F is separable because we are in characteristic 0. Use the map $F \to R$ to map $P(T)$ to a polynomial in $R[T]$. Then by Hensel's lemma, version 1, there is a unique root y of P in R reducing to x modulo π. We can then map E to R by sending x to y.

Now use your favorite equivalent of the axiom of choice (e.g., Zorn's lemma or transfinite induction) to put together an embedding of all of k into R. □

The case of positive characteristic is made tricky by the fact that in situation (b), we might hit an inseparable polynomial. This can't happen in case k is finite, since then every subfield of k is perfect. It is possible to get around this with some work, even if k itself is imperfect; this observation is due to Cohen and forms part of the *Cohen structure theorem*. However, I'll stick to the perfect case, in which case there is a uniqueness argument which locks the situation down pretty well (no axiom of choice required!).

Proof of uniqueness in characteristic p. Suppose $f_1, f_2 : k \to R$ are two homomorphisms of the desired form. Then for any $x \in k$, $f_1(x) - f_2(x)$ is divisible by π. However, we can write $x = y^p$ for some $y \in k$ since k is perfect, and then because we are in characteristic p,

$$f_1(x) - f_2(x) = f_1(y^p) - f_2(y^p) = f_1(y)^p - f_2(y)^p = (f_1(y) - f_2(y))^p$$

is divisible not just by π but by π^p. By similar reasoning, $f_1(x) - f_2(x)$ is divisible by π^{p^n} for any nonnegative integer n, and so must be zero. □

The existence is obtained by turning this around in the following way.

Proof of existence in characteristic p. Let $f_0 : k \to R$ be any map, not necessarily a homomorphism, such that $k \to R \to k$ is the identity map. For $n = 1, 2, \ldots$, define

$$f_n(x) = f_0(x^{p^{-n}})^{p^n}.$$

Since $f_0(x^{p^{-n-1}})^p \equiv f_0(x^{p^{-n}}) \mod \pi$, we have

$$f_{n+1}(x) \equiv f_n(x) \mod \pi^{p^n}.$$

Thus the sequence $f_0(x), f_1(x), \ldots$ converges to a limit $f(x)$. By a similar argument starting from the congruences

$$f_0(x^{p^{-n-1}})^p + f_0(y^{p^{n-1}})^p \equiv f_0((x+y)^{p^{-n}}) \mod \pi,$$
$$f_0(x^{p^{-n-1}})^p f_0(y^{p^{n-1}})^p \equiv f_0((xy)^{p^{-n}}) \mod \pi,$$

we find that f is a homomorphism. □

6 Structure of complete discrete valuation rings, II: Mixed characteristic case

We next consider the case where K and k are of different characteristics. This can only happen when K is of characteristic 0 and k is of characteristic $p > 0$. The typical example is of course $R = \mathbb{Z}_p$, but we can make plenty of other examples by taking extensions of $\mathbb{Q}_p$. Again, we get best results if we assume k is perfect, which include the case of most interest to us (when k is finite).

The idea here is to rerun the previous argument in the case of equal characteristic p and see what happens. The answer is quite interesting!

Lemma 6.1. *Assume that k is perfect of characteristic p (but K need not be of characteristic p). Pick any $x \in k$. For $n = 0, 1, \ldots$, choose $y_n \in R$ reducing to $x^{p^{-n}}$ modulo π. Then the sequence*

$$y_0, y_1^p, y_2^{p^2}, \ldots$$

is convergent in R, and its limit depends only on x and not on the y_n.

Proof. We start with the observation that if $a \equiv b \pmod{\pi^m}$, then $a^p \equiv b^p \pmod{\pi^{m+1}}$ because $a^p - b^p = (a-b)(a^{p-1} + \cdots + b^{p-1})$ and the second term is congruent modulo π to the sum of p identical terms, and hence to zero. Consequently, from the fact that $y_{n+1}^p \equiv y_n \pmod{\pi}$, we get that

$$y_{n+1}^{p^{n+1}} \equiv y_n^{p^n} \pmod{\pi^n}.$$

This gives the convergence. To see that the limit depends only on x, choose another element $z_n \in R$ reducing to $x^{p^{-n}}$ modulo π. Then $y_n \equiv z_n \pmod{\pi}$, so

$$y_n^{p^n} \equiv z_n^{p^n} \pmod{\pi^n}.$$

Hence the new sequence gives the same limit as the old one. □

The limit in the previous lemma is denoted $[x]$ and called the *Teichmüller lift* of x.

Lemma 6.2. *The Teichmüller map $[\cdot] : k \to R$ is multiplicative: for $x, y \in k$, $[xy] = [x][y]$.*

It would be too much to ask for the Teichmüller map to also be *additive*, though, because R is not of characteristic p: the sum $[1] + \cdots + [1]$ with p terms cannot possibly equal $[1 + \cdots + 1] = 0$.

Proof. Take $a_n, b_n \in R$ reducing to $x^{p^{-n}}, y^{p^{-n}}$, respectively. Then $a_n b_n$ reduces to $(xy)^{p^{-n}}$, so the sequences $a_n^{p^n}, b_n^{p^n}, (a_n b_n)^{p^n}$ converge to $[x], [y], [xy]$, respectively. □

For example, for $x \in \mathbb{F}_p$, the Teichmüller lift $[x] \in \mathbb{Z}_p$ satisfies $[x]^p = [x^p] = [x]$. For x nonzero, this means that $[x]$ is a $(p-1)$-st root of unity.

Lemma 6.3. *For $x \in k$, the Teichmüller lift $[x]$ is the unique lift of x to R having a p^n-th root in R for every positive integer n.*

Proof. On one hand, $[x]$ has the p^n-th root $[x^{p^{-n}}]$ for every positive integer n. On the other hand, if y is a lift of x having a p^n-th root y_n for each n, then the sequence $y_0, y_1^p, y_2^{p^2}, \ldots$ converges to both y (since it's a constant sequence) and $[x]$, so the two coincide. □

Corollary 6.4. *Suppose that p generates the maximal ideal of R. Then there is at most one automorphism of R lifting any given automorphism of k. For instance, the identity map is the only automorphism of R which acts as the identity modulo π.*

Proof. Let τ_1, τ_2 be two automorphisms of R which have the same effect modulo π. By the previous lemma, for any $x \in k$, we have $\tau_1([x]) = [\tau_1(x)] = [\tau_2(x)] = \tau_2([x])$ (since $\tau_1([x])$ and $\tau_2([x])$ both have p^n-th roots in R for any n). Also, $\tau_1(p) = \tau_2(p) = p$. Since p generates the maximal ideal of R, we can write every element x of R as a power series in p with coefficients among the Teichmüller lifts, from which it follows that $\tau_1(x) = \tau_2(x)$. □

This gives the uniqueness part of the following theorem.

Theorem 6.5 (Witt)**.** *For any perfect field k of characteristic p, there exists a unique (up to unique isomorphism) complete discrete valuation ring R with maximal ideal (p) and residue field k. Moreover, any homomorphism of fields lifts uniquely to a homomorphism of the corresponding complete DVRs.*

The existence part is less obvious; it amounts to giving an analogue of the construction of the formal power series ring over k. However, it must be intricate enough so that we can, for instance, stuff in $\mathbb{F}_p$ and get back $\mathbb{Z}_p$. We should also be able to stuff in other finite fields and get back interesting extensions of $\mathbb{Z}_p$ (more precisely, the *unramified* extensions, but more on that later).

The idea is to think about how to express the arithmetic operations of, say, $\mathbb{Z}_p$ in terms of Teichmüller lifts. If we write $x = [x_0] + [x_1]p + [x_2]p^2 + \cdots$ and $y = [y_0] + [y_1]p + [y_2]p^2 + \cdots$ with the x_i and y_j in $\mathbb{F}_p$, and then write $x + y = [z_0] + [z_1]p + [z_2]p^2 + \cdots$, then there must be some way to express z_k in terms of the x_i and y_j. In fact, z_k had better only depend on $x_0, \ldots, x_k$ and $y_0, \ldots, y_k$ and not any higher terms, since the reduction of z modulo p^{k+1} is

determined by the reductions of x and y modulo p^{k+1}. The first step is obvious: we must have $z_0 = x_0 + y_0$. But already the second step is not so clear!

Here is how to recover the expressions for the higher z_k's. Write

$$[z_1] = (x+y-[z_0])/p = ([x_0]+[x_1]p+[y_0]+[y_1]p-[z_0])/p = [x_1]+[y_1]+\frac{1}{p}([x_0]+[y_0]-[x_0+y_0]).$$

Then note that since $[x_0]+[y_0]$ reduces to x_0+y_0 modulo p, $([x_0^{1/p}]+[y_0^{1/p}])^p$ must be congruent to $[z_0]$ modulo p^2 from the definition of the Teichmüller lift). Therefore,

$$[z_1] \equiv [x_1]+[y_1]+\frac{1}{p}([x_0^{1/p}]+[y_0^{1/p}]-([x_0^{1/p}]+[y_0^{1/p}])^p) = [x_1]+[y_1]-P([x_0^{1/p}],[y_0^{1/p}]) \pmod p,$$

where $P(T,U)$ is the integer polynomial $((T+U)^p - T^p - U^p)/p$. Since a Teichmüller lift is determined by its reduction modulo p, we conclude that

$$z_1 = x_1 + y_1 - P(x_0^{1/p}, y_0^{1/p}).$$

Of course I don't really need the p-th roots in this example, because every element of $\mathbb{F}_p$ is already its own p-th root. However, leaving them in shows that a similar recipe works whenever the residue field k is perfect.

You could in principle run the same process for the higher z_k, or even do something similar for multiplication instead of addition. However, it's not so easy to guess the pattern of the resulting polynomials. The amazing discovery of Witt is that there is a much simpler recipe for generating the formulas to describe arithmetic in terms of Teichmüller lifts.

For any ring R, let $W(R)$ be the set of infinite sequences $(x_0, x_1, \dots)$ with entries in R. (We will call these sequences *Witt vectors* once we give them a bit of structure.) Define the *ghost map* w on sequences by mapping $(x_0, x_1, \dots)$ to $(w_0, w_1, \dots)$ with

$$w_0 = x_0, w_1 = x_0^p + px_1, \dots, w_n = \sum_{i=0}^{n} p^i x_i^{p^{n-i}}, \dots.$$

If p happens to be a unit in R, then the ghost map is a bijection, but otherwise not. The target is again the set of infinite sequences with entries in R, but I want to think of that as a ring using term-by-term addition and multiplication, whereas the Witt vectors will pick up rather different operations.

Theorem 6.6 (Witt)**.** *There is a unique way to equip the set $W(R)$ with a ring structure subject to the following conditions.*

(a) The construction is functorial in R. By that, I mean that whenever $R \to S$ is a ring homomorphism, so is the map $W(R) \to W(S)$ given by applying $R \to S$ term-by-term.

(b) The ghost map w is always a ring homomorphism.

The functoriality condition might sound mysterious, but it's not. It just means that the operations on $W(R)$ are defined by certain *universal polynomials* in the entries. Namely, suppose I take R to be the infinite polynomial ring $\mathbb{Z}[x_0, x_1, \dots, y_0, y_1, \dots]$. Then the universal polynomials describing Witt vector addition are none other than the components of the Witt vector sum of $(x_0, x_1, \dots)$ and $(y_0, y_1, \dots)$. (Translation into fancy language: the forgetful functor from rings to sets is represented by the ring $\mathbb{Z}[T]$.)

In fact, even before proving the theorem, one can use this description to write down candidate universal polynomials, and to see that they are uniquely determined by the condition that the ghost map is a ring homomorphism. Namely, simply apply the ghost map to $(x_0, x_1, \dots)$ and to $(y_0, y_1, \dots)$, add the results term-by-term, then apply the inverse ghost map over the ring $\mathbb{Z}[1/p, x_0, x_1, \dots, y_0, y_1, \dots]$. The entire content of Witt's theorem is that the resulting polynomials (and the corresponding ones for multiplication and negation) belong not just to $\mathbb{Z}[1/p, x_0, x_1, \dots, y_0, y_1, \dots]$ but also to $\mathbb{Z}[x_0, x_1, \dots, y_0, y_1, \dots]$.

There are a couple of ways to prove this. One is to make the following observation, attributed variously to Cartier, Dieudonné, and Dwork.

Lemma 6.7. *For $R = \mathbb{Z}[x_0, x_1, \dots, y_0, y_1, \dots]$, if I define the map $\phi : R \to R$ as the substitution x_i to x_i^p and y_i to y_i^p, then a sequence $w = (w_0, w_1, \dots)$ belongs to the image of the ghost map if and only if for $n = 1, 2, \dots$, we have $w_n \cong \phi(w_{n-1}) \pmod{p^n}$.*

The point is that this makes it clear that the image of the ghost map is a subring.

Proof. Checking whether $w = (w_0, w_1, \dots)$ is in the image of the ghost map amounts to reconstructing the sequence $(z_0, z_1, \dots)$ of elements of $R[1/p]$ giving rise to w and seeing whether those elements also belong to R. We can of course truncate the two sequences $(w_0, \dots, w_n)$ and $(z_0, \dots, z_n)$ and ask whether the integrality of the z_i is equivalent to the congruence $w_i \cong \phi(w_{i-1}) \pmod{p^i}$ for $i = 1, \dots, n$. Suppose we've checked this for some n (it's automatic for $n = 0$). We then have

$$w_{n+1} = p^{n+1} z_{n+1} + (z_0^{p^{n+1}} + \cdots + p^n z_n^p).$$

The point is that if $z_0, \dots, z_n \in R$, then the term in parentheses is congruent to $\phi(z_0^{p^n} + \cdots + p^n z_n)$ modulo p^{n+1}, by term-by-term comparison: for each i, $\phi(p^i z_i^{p^{n-i}}) = p^i \phi(z_i)^{p^{n-i}}$, and the congruence $\phi(z_i) \equiv z_i^p \pmod{p}$ implies $\phi(z_i)^{p^{n-i}} \equiv z_i^{p^{n+1-i}} \pmod{p^{n+1-i}}$. So we can divide by p^{n+1} to get z_{n+1} if and only if $w_{n+1} \equiv \phi(w_n) \pmod{p^{n+1}}$. □

There is another proof which is a bit tangential from the point of view of this course, but makes some fascinating links with symmetric functions and other combinatorial objects. I probably won't discuss this in class, but I hope to put it in the notes at some point.

So now we have succeeded in making a functor W from rings to rings. It remains to check that when k is a perfect field,

(a) the ring $W(k)$ is a complete discrete valuation ring with maximal ideal (p) and residue field k;

(b) any other such object is isomorphic to $W(k)$ (there can only be one such isomorphism because of the Teichmüller lifts).

In general, there is a surjection $W(R) \to R$ given by keeping only the zeroth component, so we have a map $W(k) \to k$. To keep going, it is useful to introduce an additional structure on Witt vectors in general. The *Verschiebung* map $V : W(R) \to W(R)$ is defined directly on Witt components: $V(x_0, x_1, \dots) = (0, x_0, x_1, \dots)$. It corresponds on the ghost side to the map $(w_0, w_1, \dots)$ to $(0, pw_0, pw_1, \dots)$, so it is additive but not multiplicative: however, it does satisfy the identity $V(x)V(y) = pV(xy)$. Also, the image of V is precisely the kernel of the map $W(R) \to R$.

Now take $R = k$ again. By calculating by hand, we find (exercise)

$$p \cdot (x_0, \dots) = (0, x_0^p, \dots) \tag{1}$$

(exercise). We can use this to deduce that the image of V is the ideal (p). On one hand, any multiple of p maps to zero in V, so it belongs to the image. On the other hand, suppose x is in the image of V. Using (1), we produce y_0 such that $x - py_0$ is in the image of V^2. If $x - py_0 = V^2(z)$, we can then produce y_1 such that $x - py_0 - pV(y_1)$ is in the image of V^3, and so on. We get a sum $y_0 + V(y_1) + \cdots$ which is convergent in the sense that each Witt component eventually stabilizes, so we get a limiting value y for which $x = py$. Hence (p) is the image of V, and hence also the kernel of $W(k) \to k$.

Also, p is not a zero divisor: for any $x \in W(R)$ which is nonzero, there must then exist a nonnegative integer n for which $x = V^n(y)$ for some y but is not in the image of V^{n+1}. From (1), py is nonzero, so $V^n(py) = px$ isn't either.

Since $p = V(x)$ for some x, we have $V(1)V(x) = pV(x)$. Since p is not a zero divisor, this forces $V(1) = p$.

For $x_1, \dots, x_n \in W(R)$, we can write

$$\begin{aligned} V(x_1) \cdots V(x_n) &= p^{n-1} V(x_1 \cdots x_n) \\ &= V(p^{n-1} x_1 \cdots x_n) \\ &= V(p^{n-2} V(*)) = V(V(p^{n-2} *)) = V(p^{n-3} V(V(*))) = \cdots . \end{aligned}$$

Consequently, $V(x_1) \cdots V(x_n)$ belongs to the image of the n-th power of V, that is, its first n components are zero. This implies that $W(k)$ is p-adically complete; since $W(k)/(p) = k$ is a field, $W(k)$ is a complete local ring. Moreover, using (1) once again, we see that any element x of the image of V^n which is not in the image of V^{n+1} is divisible by p^n but not p^{n+1}, and x/p^n is a unit in $W(k)$. Hence $W(k)$ is a principal ideal domain with one nonzero prime ideal, so it's a DVR.

That proves (a). Now we need to check that any other complete DVR R with maximal ideal p and residue field k is isomorphic to $W(k)$. Note first that the map $k \to W(k)$ given by $x \to (x, 0, 0, \dots)$ is multiplicative, and thus must coincide with the Teichmüller map (since any element of the image has p^n-th roots for all n). We can write each element of $W(k)$ as a convergent sum $[x_0] + [x_1]p + \cdots$, and likewise for R. Since there are universal polynomials to express arithmetic in terms of Teichmüller elements (never mind what they are!), we get a ring homomorphism taking Teichmüllers to Teichmüllers.

Exercises

1. Compute (using a computer if you wish) the reductions modulo 5^3 of the Teichmüller lifts in $\mathbb{Z}_5$.

2. Prove that for $p > 2$, every element of $\mathbb{Z}_p$ congruent to 1 modulo p^2 has a p-th root, by checking that the binomial series expansion for $(1 + p^2x)^{1/p}$ converges for $x \in \mathbb{Z}_p$. For $p = 2$, replace p^2 by p^3.

3. Using the previous exercise, deduce that an element $x \in \mathbb{Q}_p$ belongs to $\mathbb{Z}_p$ if and only if $1 + p^{p+1}x^{2p}$ has a p-th root in $\mathbb{Q}_p$. (Hint: the valuation of a p-th power must be a multiple of p.)

4. Using the previous exercise, deduce that the field $\mathbb{Q}_p$ has no nontrivial automorphisms. Optional: extend this to the fraction field of any complete discrete valuation ring with perfect residue field whose maximal ideal is generated by p, by proving that any automorphism of the fraction field preserves valuations (and thus acts on the DVR, and thus is determined by its action on the residue field).

5. Prove (1).

6. Prove that for any ring R, there is a ring homomorphism $F : W(R) \to W(R)$ corresponding to the left shift $(w_0, w_1, \dots) \to (w_1, w_2, \dots)$ on ghost components. Then check that when $p = 0$ in R, this map is the map induced by the Frobenius map on R via functoriality of W. (For that reason, this map is called the *Frobenius map* on $W(R)$.)

7 Local fields and number fields

I'm about to start discussing some finer structure of extensions of complete discrete valuation rings, but to motivate this I should explain the motivation from algebraic number theory. I'm hoping much of this is familiar; I'll also prove some of these facts in more detail in the next few lectures.

Let F be either a *number field*, i.e., a finite extension field of $\mathbb{Q}$, or a *function field*, i.e., a finite extension field of $\mathbb{F}_p(t)$ for some prime $p > 0$. This field comes equipped with a distinguished subring $\mathfrak{o}_F$, which is the integral closure of $\mathbb{Z}$ in the number field case, or of $\mathbb{F}_p[t]$ in the function field case. This ring is a *Dedekind domain*, a one-dimensional integrally closed noetherian integral domain. Such rings have *unique factorization of ideals*: every nonzero fractional ideal $\mathfrak{a}$ can be written as a product $\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_n^{e_n}$ for some nonnegative integer n, distinct nonzero (hence maximal) prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_n$ and some positive integers $e_1, \dots, e_n$. This representation is moreover unique up to permuting the terms.

For each prime ideal $\mathfrak{p}$ of F, we obtain a $\mathfrak{p}$-adic valuation function $v_\mathfrak{p}$ on F by taking x to the exponent of $\mathfrak{p}$ in the prime factorization of the principal fractional ideal (x). The corresponding discrete valuation ring is the localization $\mathfrak{o}_{F,\mathfrak{p}}$ of $\mathfrak{o}_F$ at $\mathfrak{p}$, i.e., the ring obtained

from $\mathfrak{o}_F$ by inverting every element of $\mathfrak{o}_F$ not belonging to $\mathfrak{p}$. Let $F_\mathfrak{p}$ denote the completion of F with respect to $v_\mathfrak{p}$.

Let E be a finite separable extension of F (the separability condition being automatic in the number field case); then $\mathfrak{o}_E$ is the integral closure of $\mathfrak{o}_F$ in E, and is again a Dedekind domain. For each prime ideal $\mathfrak{p}$ of $\mathfrak{o}_F$, we can then factor $\mathfrak{p}\mathfrak{o}_E = \mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_n^{e_n}$ as a product of powers of distinct prime ideals of $\mathfrak{o}_E$. This factorization contains the information of the *splitting* and *ramification* of $\mathfrak{p}$ in E. For instance, there are only finitely many primes $\mathfrak{p}$ of $\mathfrak{o}_F$ for which any of the exponents e_i is greater than 1; these are the *ramified primes* for the extension E/F. (More precisely, one should say that $\mathfrak{q}_i$ is *ramified* over $\mathfrak{p}$, since in general the e_i may not all be greater than 1.) The ramification determines the *discriminant* of F in a somewhat complicated way; one thing we'll do later is make that more explicit.cd

The relationship of this factorization to completion is captured by two statements. One is that

$$E \otimes_F F_\mathfrak{p} \cong E_{\mathfrak{q}_1} \oplus \cdots \oplus E_{\mathfrak{q}_n}.$$

The other is that the restriction of $v_{\mathfrak{q}_i}$ from E to F equals e_i times $v_\mathfrak{p}$. (Note that $v_\mathfrak{p}$ can only extend one way to any finite *field* extension of $F_\mathfrak{p}$, so the valuations $v_{\mathfrak{q}_i}$ must come from *different* fields!)

Suppose further that E/F is a Galois extension with group G. Then G acts transitively on the $\mathfrak{q}_i$, so the e_i are all equal. The subgroup $G_{\mathfrak{q}_i}$ of G fixing $\mathfrak{q}_i$, called the *decomposition group* of $\mathfrak{q}_i$, also acts on $E_{\mathfrak{q}_i}$; since its order equals $[E : F]/n = [E_{\mathfrak{q}_i} : F_\mathfrak{p}]$, $E_{\mathfrak{q}_i}$ is forced to be a Galois extension of $F_\mathfrak{p}$ with Galois group $G_{\mathfrak{q}_i}$. It's not too difficult to show that as $\mathfrak{p}$ varies, the groups $G_{\mathfrak{q}_i}$ cover all of G (this is less deep than the Chebotarev density theorem, on which more in a moment).

The residue field $\mathfrak{o}_F/\mathfrak{p}$ is some finite field $\mathbb{F}_q$, of which the residue field $\mathfrak{o}_E/\mathfrak{q}_i$ is a finite extension. This extension is Galois because any extension of finite fields is Galois. The group $G_{\mathfrak{q}_i}$ acts on $\mathfrak{o}_E/\mathfrak{q}_i$; the kernel of this map is called the *inertia group* $I_{\mathfrak{q}_i}$ of $\mathfrak{q}_i$. By interpreting $G_{\mathfrak{q}_i}$ also as $\mathrm{Gal}(E_{\mathfrak{q}_i}/F_\mathfrak{p})$ and using Hensel's lemma, we may see that $G_{\mathfrak{q}_i}$ actually surjects onto $\mathrm{Gal}(E_{\mathfrak{q}_i}/F_\mathfrak{p})$. (We'll give a more detailed explanation of this in a later lecture). In particular, the *Frobenius automorphism* $x \mapsto x^q$ in $\mathrm{Gal}(E_{\mathfrak{q}_i}/F_\mathfrak{p})$ can be lifted to $G_{\mathfrak{q}_i}$; the lift is unique if and only if $\mathfrak{p}$ is unramified. The Chebotarev density theorem asserts that every element of G occurs equally often as a Frobenius automorphism. The Frobenius automorphisms also appear in the Artin reciprocity law from class field theory, and in the definition of an Artin L-function (see Stark's class!).

As usual, when E/F is not Galois, you can pass to the Galois closure L of E/F, and put $G = \mathrm{Gal}(L/F)$ and $H = \mathrm{Gal}(L/E) \subseteq G$. We can't define a quotient group G/H since H is not normal, but we can split primes of F up in L and then see how they come back together in E. Using completions gives a nice interpretation of this: if we write

$$L \otimes_F F_\mathfrak{p} \cong L_{\mathfrak{q}_1} \oplus \cdots \oplus L_{\mathfrak{q}_n},$$

then the components of $E \otimes_F F_\mathfrak{p}$ correspond to H-orbits among the components. (More precisely, group each H-orbit together, then take the invariant elements of a single factor under its own stabilizer.) Since H is not normal, there is no reason why we can't have heterogeneous behavior (different residue fields, different ramification).

8 Unramified and tamely ramified extensions

In this lecture, let K be a field which is complete for a discrete valuation, normalized in the usual way (so that v_K maps surjectively to $\mathbb{Z} \cup \{+\infty\}$). Let L be a finite separable extension of K. We start by defining some key numerical invariants (which we are familiar with in the context of splitting of prime ideals in number fields).

Recall that v_K extends uniquely to a valuation on L, but with the wrong normalization. If v_L instead denotes the normalized valuation on L, then there exists a positive integer e such that $v_L(x) = ev_K(x)$ for any $x \in K$. One way to characterize e is as the L-valuation of a uniformizer in K. This number is called the *ramification index* of the extension L/K.

Next, if we write k and ℓ for the residue fields of K and L, respectively, then ℓ may be naturally viewed as an extension field of k. The degree $f = [\ell : k]$ is called the *residue degree*, or *residue field degree*, or *residue class degree* (there seems to be no naming consensus) of the extension L/K.

Proposition 8.1. *We have $ef = [L : K]$. In particular, f is finite.*

Proof. Let π be a uniformizer of L. Choose a basis of ℓ over k, lift each element to $\mathfrak{o}_L$ arbitrarily, and let B denote the result. I claim that the elements

$$\{\pi^i b : i \in \{0, \dots, e-1\}, b \in B\}$$

form a basis for L over K, which will imply the equality.

To see that these are linearly independent, we will prove that if $x = \sum_{i,b} x_{i,b}\pi^i b$ with $x_{i,b} \in K$, then

$$v_L(x) = \max_{i,b}\{v_L(x_{i,b}\pi^i b)\}. \tag{2}$$

It's clear that the left side is greater than or equal to the right side, so let's focus on the other inequality. Choose a pair (j, c) for which $v_L(x_{j,c}\pi^j c) = \max_{i,b}\{v_L(x_{i,b}\pi^i b)\}$. If we now write

$$\frac{x}{x_{j,c}\pi^j c} = \sum_{i,b} \frac{x_{i,b}\pi^i b}{x_{j,c}\pi^j c},$$

then each term on the right side belongs to $\mathfrak{o}_L$. If we map to ℓ, then all the terms with $i \neq j$ vanish because $v_L(x_{i,b}\pi^i b) \equiv i \pmod{e}$, so we can't have $v_L(x_{i,b}\pi^i b) = v_L(x_{j,c}\pi^j c)$. So we only get the sum $\sum_b \overline{x_{j,b}/x_{j,c}}\,\overline{b}$. This is a linear combination of basis elements of ℓ over k, and not all of the coefficients are zero (because the coefficient for $b = c$ equals 1), so it represents a nonzero element of ℓ. Consequently, $v_L(x) = v_L(x_{j,c}\pi^j c)$, which proves (2).

From (2), we see that we cannot have $x = 0$ but $x_{i,b}$ not all zero, so the $\pi^i b$ are linearly independent. To show that they form a basis, notice that for any nonzero $x \in L$, we can cook up a K-linear combination $y = \sum_{i,b} y_{i,b}\pi^i b$ for which $v_L(x - y) > v_L(x)$. By repeating, we can write x as a limit of linear combinations of the $\pi^i b$; the coefficients of any given $\pi^i b$ then converge to a limit, and using these limits as coefficients gives a linear combination equal to x. (Note that we have shown that the $\pi^i b$ form an *orthogonal basis* of L over K. $\square$

Side remark: if the valuation on K is not discrete, $[L : K]/(ef)$ is always an integer which is a power of the characteristic of k (meaning 1 if that characteristic is 0), but can fail to be equal to 1. This phenomenon was discovered by Ostrowski.

The case $e = 1$ is particularly interesting. We say that L/K is *unramified* if $e = 1$ *and* the residue field extension ℓ/k is *separable*. E.g., for $K = \mathbb{Q}_3$, $L = \mathbb{Q}_3[T]/(T^2 + 1)$ is unramified over K. One gets lots of examples using roots of unity; see the exercises.

Proposition 8.2. *For any finite separable extension κ of k, there exists a unique unramified extension U of K with residue field κ. Moreover, for L a finite separable extension of K with residue field ℓ, any inclusion of κ into ℓ lifts uniquely to an inclusion of U into L.*

Proof. To make U, write $\kappa = k[\overline{T}]/(\overline{P})$ for some monic irreducible polynomial $\overline{P} \in k[\overline{T}]$ (using the primitive element theorem, which holds because κ/k is separable). Then lift $\overline{P}$ to a monic polynomial $P \in \mathfrak{o}_K[T]$, which is irreducible even over K (by Gauss's lemma), and put $U = K[T]/(P(T))$. For $d = \deg(P)$, the function $v(x_0+x_1T+\cdots+x_{d-1}T^{d-1}) = \min_i\{v_K(x_i)\}$ is a valuation: it's enough to check that the product of two things with valuation 0 is again of valuation 0, but that reduces to the fact that κ is a field. This formula for the valuation means that $\mathfrak{o}_U = \mathfrak{o}_K[T]/(P(T))$, so we have a copy of κ in the residue field of U. Since $e(U/K)f(U/K) = d$ and now $f(U/K) \geq d$, we have $e = 1$ and $f = d$.

To finish, we need only show that for *this* particular choice of U, for L a finite separable extension of K with residue field ℓ, any inclusion of κ into ℓ lifts uniquely to an inclusion of U into L. But this is just Hensel's lemma for the polynomial P. □

Corollary 8.3. *If k is finite, then every unramified finite extension of K is Galois.*

Proof. This follows from the proposition and the fact that every finite extension of a finite field is a Galois extension. □

Note that an unramified extension of an unramified extension is unramified, and the compositum of unramified extensions is unramified. Moreover, if L_1, L_2 are two finite separable extensions of K, and L_1/K is unramified, then so is L_2K/L_2, but *not conversely.* (For instance, this fails if $L_1 = L_2$!)

For example, if $K = k((T))$, then for any finite separable extension ℓ of k, the unramified extension of K with residue field ℓ is precisely $\ell((T))$. For another example, if k is perfect of characteristic $p > 0$ and $K = \operatorname{Frac} W(k)$, then for any finite (hence) separable extension ℓ of k, the unramified extension of K with residue field ℓ is precisely $\operatorname{Frac} W(\ell)$. More generally, whenever k is perfect of characteristic $p > 0$, we can use Witt vectors to describe unramified extensions; see exercises.

The next simplest extensions to describe are the tamely ramified ones. We say L/K is *tame ramified* (or simply *tame*) if e is not divisible by p and ℓ/k is separable. In this case, the unramified extension U of K with residue field ℓ sits between L and K; we say that L is *totally tamely ramified* over U. For example, if n is a positive integer not divisible by p, and $a \in K$ has no d-th root in K for any $d > 1$ dividing n, then $K[T]/(T^n - a)$ is a tamely ramified extension of K (exercise).

Again, a tamely ramified extension of a tamely ramified extension is tamely ramified, and the compositum of tamely ramified extensions is tamely ramified. Moreover, if L_1, L_2 are two finite separable extensions of K, and L_1/K is tamely ramified, then so is L_2K/L_2, but not conversely.

One consequence is that for any finite separable extension L/K, the compositum T of all tamely ramified subextensions is again tamely ramified, and in fact is the maximal tamely ramified subextension of L/K. If $L \neq T$, we say L is *wildly ramified* (or simply *wild*); if $T = K$, we say L is *totally wildly ramified*.

Proposition 8.4. *Suppose L/K is totally tamely ramified of degree d. Then there exists $a \in K$ for which $L = K[T]/(T^d - a)$. In fact, a can be taken to be a uniformizer in K.*

Proof. Let π_L be a uniformizer of L; then $v_L(\pi_L^d)$ also occurs as the L-valuation of a uniformizer in K. Since $k = \ell$, we can in fact find $a \in K$ for which π_L^d/a is a unit in $\mathfrak{o}_L$ congruent to 1 modulo π_L. If we write $\pi_L^d/a = 1 + x$, then the binomial series

$$(1+x)^{1/d} = \sum_{i=0}^{\infty} \binom{d}{i} x^i$$

converges because $1/d \in \mathbb{Z}_p$, so the binomial coefficients $\binom{d}{i}$ are all in $\mathbb{Z}_p$. If we write $y = (1+x)^{1/d}$, then we can write $L = K(a^{1/d})$ as desired. □

Warning: unlike unramified extensions, tamely ramified extensions can fail to be Galois even when k is finite (exercise). The previous proposition doesn't contradict this statement because $K(a^{1/d})$ need not be Galois over K in case K doesn't contain a primitive d-th root of unity.

Exercises

1. Suppose k is perfect of characteristic $p > 0$. Prove that there is a unique ring homomorphism $W(k) \to \mathfrak{o}_K$ lifting the identity map on residue fields.

2. Suppose k is perfect of characteristic $p > 0$. Using the previous exercise, show that for any finite separable extension ℓ of k, the ring $\mathfrak{o}_K \otimes_{W(k)} W(\ell)$ is a complete discrete valuation ring whose fraction field is unramified over K with residue field ℓ.

3. Let p be a prime number, let m be any positive integer, and let ζ_m be a primitive m-th root of unity in some algebraic closure of $\mathbb{Q}_p$.

   (a) Prove that $\mathbb{Q}_p(\zeta_m)$ is unramified over $\mathbb{Q}_p$ if and only if either $p > 2$ and m is not divisible by p, or $p = 2$ and m is not divisible by 4.

   (b) Prove that $\mathbb{Q}_p(\zeta_m)$ is tamely ramified over $\mathbb{Q}_p$ if and only if either $p > 2$ and m is not divisible by p^2, or $p = 2$ and m is not divisible by 4. (In particular, for $p = 2$, $\mathbb{Q}_p(\zeta_m)$ is unramified if and only if it is tamely ramified.)

4. Let n be a positive integer not divisible by p, and suppose $a \in K$ has no d-th root in K for any $d > 1$ dividing n. Prove that $K[T]/(T^n - a)$ is a tamely ramified extension of K. (Hint: you may want to first understand the cases where $v_K(a) = 0$ and where $v_K(a)$ is coprime to n.)

5. (a) Find an example of a cubic non-Galois extension of $\mathbb{Q}_p$ for some $p > 3$.
   (b) Show that every such extension is totally tamely ramified.
   (c) Show that no such extension exists for $p \equiv 1 \pmod 3$.

6. Prove *Krasner's lemma*: let α, β be elements of an algebraic closure of K, with $K(\alpha)$ separable over K. Let $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_n$ be the conjugates of α. Prove that if $\alpha - \beta$ has greater valuation than all of $\alpha - \alpha_2, \ldots, \alpha - \alpha_n$, then $\alpha \in K(\beta)$. (Hint: use Hensel's lemma.)

9 Ramification filtrations: the lower numbering

Keep notation as in the previous lecture, but assume that L is Galois over K with Galois group G, and that ℓ is separable over k. We describe a filtration on G that picks out the maximal unramified subextension, the maximal tamely ramified extension, and some other intermediate extensions that we can use to describe the discriminant of L over K.

To begin with, recall that G acts on $\mathfrak{o}_L$ (since the valuation on K extends uniquely to L) preserving $\mathfrak{o}_K$, and thus acts on ℓ preserving k.

Lemma 9.1. *The extension ℓ/k is Galois, and the map $G \to \mathrm{Gal}(\ell/k)$ is surjective.*

Proof. Let U be the maximal unramified subextension of L/K; it has residue field ℓ because ℓ/k is separable. Any $g \in \mathrm{Gal}(L/K)$ must take U to U, so U is a Galois subextension and $\mathrm{Gal}(L/K)$ surjects onto $\mathrm{Gal}(U/K)$. But the uniqueness property of unramified extensions allows us to identify $\mathrm{Gal}(U/K)$ with $\mathrm{Gal}(\ell/k)$. □

Let $\mathfrak{p}$ denote the maximal ideal of $\mathfrak{o}_L$. For each integer $i \geq -1$, the group G acts on the quotient $\mathfrak{o}_L/\mathfrak{p}^{i+1}$; let G_i be the subgroup of G acting trivially on $\mathfrak{o}_L/\mathfrak{p}^{i+1}$. The G_i comprise the *ramification filtration on G for the lower numbering*, as to be distinguished from the *upper numbering* to come later. Obviously

$$G = G_{-1} \supset G_0 \supset \cdots,$$

and $G_0 = \mathrm{Gal}(L/U)$ from the proof of the previous lemma. We also call $G - 0$ the *inertia subgroup* of G. Also, G_i is the trivial group for i sufficiently large: e.g., choose generators $x_1, \ldots, x_n$ of $\mathfrak{o}_L$ as a $\mathfrak{o}_K$-algebra, and choose i so that $v_L(g(x_j) - x_j)) \leq i$ whenever $j \in \{1, \ldots, \}$, $g \in \mathrm{Gal}(L/K)$, and $g(x_j) \neq x_j$. (In fact, one can get away with only one generator; see exercises.)

The following fact is obvious from the definition.

Lemma 9.2. *Let K' be a subextension of L/K, and put $H = \mathrm{Gal}(L/K')$. Then $H_i = G_i \cap H$ for all $i \geq -1$.*

On the other hand, the groups G_i do not behave well with respect to taking quotients, which is arguably more natural; for instance, they don't give any useful structure on the absolute Galois group of K. For this, we will have to switch to the upper numbering, but more on that a bit later.

To get more information about the other groups, let π be a uniformizer of $\mathfrak{o}_L$.

Lemma 9.3. *For $i \geq 0$, if $g \in G_0$, then $g \in G_i$ if and only if $g(\pi)/\pi \equiv 1 \pmod{\mathfrak{p}_L^i}$.*

Proof. By the previous lemma, it is harmless to replace K by its maximal unramified subextension within L, so we may go straight to the case $G = G_0$. In this case, $\mathfrak{o}_L$ is generated by π as a $\mathfrak{o}_K$-algebra, so $g \in G_i$ if and only if $v_L(g(\pi) - \pi) \geq i + 1$. This is equivalent to $v_L(g(\pi)/pi - 1) \geq i$, proving the claim. □

Corollary 9.4. *For $i \geq 0$, the quotient G_i/G_{i+1} admits a natural (i.e., not dependent on the choice of π) injective map to U_L^i/U_L^{i+1}, for U_L^i the subgroup of the group of units of $\mathfrak{o}_L$ consisting of elements congruent to 1 modulo π^i.*

Proof. The map takes g to the image of $g(\pi)/\pi$. Injectivity, and the independence from the choice of π, are easy (or see Serre IV.2, proposition 7). □

Corollary 9.5. *The group G_0/G_1 is cyclic, of order prime to the characteristic of k.*

Proof. We have an injective map from G_0/G_1 to $\ell^\times$; since G_0/G_1 has finite order, its image must consist of a finite subgroup of the roots of unity in ℓ. But any such group is finite with order prime to the characteristic of ℓ. □

Corollary 9.6. *For $i > 0$, the group U_L^i/U_L^{i+1} is isomorphic (canonically) to the additive group $\mathfrak{p}^i/\mathfrak{p}^{i+1}$ and also (not canonically) to the additive group of ℓ. In particular, G_i/G_{i+1} is an elementary abelian p-group for p the characteristic of k, or the trivial group if k has characteristic 0.*

Corollary 9.7. *We have $G_1 = \mathrm{Gal}(L/T)$ for T the maximal tamely ramified subextension of L/K. This group is a p-group.*

Corollary 9.8. *The group G_0 is solvable. If k is a finite field, then G is solvable.*

Proof. Since G_i/G_{i+1} is abelian for $i \geq 0$ and G_i is trivial for i large, we find that G_0 is solvable, and G is solvable if and only if $G_{-1}/G_0 \cong \mathrm{Gal}(\ell/k)$ is solvable. The latter is automatic for k finite, since then $\mathrm{Gal}(\ell/k)$ is cyclic. □

This means that the analogue of the inverse Galois problem for local fields must be restricted quite far. If you just look at the inertia group, it has a unique p-Sylow subgroup, which is normal and the quotient by which is cyclic of order prime to p. This forces the group to be a semidirect product of the p-Sylow with the prime-to-p quotient (by a fact from

group theory, or see Serre IV.2 for a direct proof). Conversely, every such group occurs as the inertia group for a suitable choice of K, but I won't try to show that here.

We already know that the groups G_i/G_{i+1} are elementary abelian, so the commutator of two elements of G_i belongs to G_{i+1}. In fact, something much stronger is true. Let $\theta_i : G_i/G_{i+1} \to U_L^i/U_L^{i+1}$ be the map constructed earlier. Note first that G_0 acts on each G_i by conjugation, so G_0 has a well-defined action on G_i/G_{i+1}.

Proposition 9.9. *If $g \in G_0$ and $h \in G_i/G_{i+1}$ for some $i \geq 1$, then $\theta_i(ghg^{-1}) = \theta_0(g)^i\theta_i(h)$.*

Proof. Write $\pi' = g^{-1}(\pi)$ and $h(\pi') = \pi'(1 + \pi^i u)$ with $u \in \mathfrak{o}_L^\times$, so

$$\frac{ghg^{-1}(\pi)}{\pi} = \frac{g(\pi')g(1+\pi^i u)}{\pi} = 1 + g(\pi)^i g(u).$$

In $\mathfrak{p}_L^i/\mathfrak{p}_L^{i+1}$, $g(\pi)^i g(u)$ represents the same class as $\theta_0(g)\theta_i(h)$ because $g(u) \equiv u \pmod{\pi}$ (since $g \in G_0$). This proves the claim. □

Corollary 9.10. *If $g \in G_0$ and $h \in G_i$ for some $i \geq 1$, then $ghg^{-1}h^{-1} \in G_{i+1}$ if and only if $g^i \in G_1$ or $h \in G_{i+1}$.*

Corollary 9.11. *If G is abelian and e_0 is the order of G_0/G_1, then the integers $i \geq 1$ for which $G_i \neq G_{i+1}$ are all divisible by e_0.*

Lemma 9.12. *If $g \in G_i, h \in G_j$, and $i, j \geq 1$, then $ghg^{-1}h^{-1} \in G_{i+j}$ and $\theta_{i+j}(ghg^{-1}h^{-1}) = (j-i)\theta_i(g)\theta_j(h)$.*

Note that the expression $\theta_i(g)\theta_j(h)$ is defined by identifying U_L^i/U_L^{i+1} with $\mathfrak{p}^i/\mathfrak{p}^{i+1}$.

Proof. Write $g(\pi) = \pi(1 + \pi^i u)$ and $h(\pi) = \pi(1 + \pi^j v)$ with $u, v \in \mathfrak{o}_L$. Then

$$gh(\pi) = \pi(1 + \pi^i u)(1 + g(\pi)^j g(v)).$$

Similarly,

$$hg(\pi) = \pi(1 + \pi^j v)(1 + h(\pi)^i h(u)).$$

Since $g \in G_i$, we have $g(v) \equiv v \pmod{\mathfrak{p}_L^{i+1}}$. Also,

$$(1 + \pi^i u)^j \equiv 1 + j\pi^i u \pmod{\mathfrak{p}_L^{i+1}}.$$

Hence

$$g(\pi)^j g(v) \equiv g(\pi)v = \pi^j(1+\pi^i u)^j v \equiv \pi^j(1 + j\pi^i u)v \equiv \pi^j v + j\pi^{i+j}uv \pmod{\mathfrak{p}_L^{i+j+1}}$$

and so

$$\frac{gh(\pi)}{\pi} \equiv (1+\pi^i u)(1 + \pi^j v + j\pi^{i+j}uv) \equiv 1 + \pi^i u + \pi^j v + (j+1)\pi^{i+j}uv \pmod{\mathfrak{p}_L^{i+j+1}}.$$

Similarly,

$$\frac{hg(\pi)}{\pi} \equiv 1 + \pi^i u + \pi^j v + (i+1)\pi^{i+j} uv \pmod{\mathfrak{p}_L^{i+j+1}}.$$

Put $\pi' = hg(\pi)$. Then

$$\begin{aligned}
\frac{ghg^{-1}h^{-1}(\pi')}{\pi} &= \frac{gh(\pi)}{\pi'} \\
&= \frac{gh(\pi)}{\pi}\frac{\pi}{hg(\pi)} \\
&\equiv 1 + (\pi^i u + \pi^j v + (j+1)\pi^{i+j}uv) - (\pi^i u + \pi^j v + (i+1)\pi^{i+j}uv) \pmod{\mathfrak{p}_L^{i+j+1}} \\
&\equiv 1 + (j-i)\pi^{i+j}uv \pmod{\mathfrak{p}_L^{i+j+1}}.
\end{aligned}$$

This proves the claim. □

Lemma 9.13. *The integers $i \geq 1$ for which $G_i \neq G_{i+1}$ are all congruent modulo p.*

Proof. If $G_1 = \{1\}$ we have nothing to show, so assume $G_1 \neq \{1\}$. Take $j \geq 1$ to be the largest integer for which $G_j \neq \{1\}$, and let $i \geq 1$ be any integer for which $G_i \neq G_{i+1}$. We can choose $g \in G_i - G_{i+1}$ and $h \in G_j - G_{j+1}$ and use the previous lemma to see that $ghg^{-1}h^{-1} \in G_{i+j}$ and $\theta_{i+j}(ghg^{-1}h^{-1}) = (j-i)\theta_i(g)\theta_j(h)$. Since $i + j > j$, we have $G_{i+j} = \{1\}$ and so $ghg^{-1}h^{-1} = 1$, so $\theta_{i+j}(ghg^{-1}h^{-1}) = 0$. But $\theta_i(g), \theta_j(h)$ are nonzero, so $j - i \equiv 0 \pmod{p}$. □

Proposition 9.14. *If $g \in G_i, h \in G_j$, and $i, j \geq 1$, then $ghg^{-1}h^{-1} \in G_{i+j+1}$.*

Proof. If $g \in G_{i+1}$ or $h \in G_{j+1}$, this follows from Lemma 9.12. Otherwise, $i \equiv j \pmod{p}$ by Lemma 9.13, so $\theta_{i+j}(ghg^{-1}h^{-1}) = 0$ by Lemma 9.12 again. □

Exercises

1. Prove that for any finite separable extension L of K whose residue field is also separable over k, $\mathfrak{o}_L$ can be generated as a $\mathfrak{o}_K$-algebra by just one element. (This can fail if the residue field is not separable over k!)

2. Prove that if $K = k((T))$ and k is of characteristic 0, then the algebraic closure of $k((T))$ is the union of the $\ell((T^{1/n}))$ running over all finite extensions ℓ and all positive integers n.

3. Demonstrate that the previous exercise fails in characteristic p by showing that the ring $\cup_{\ell,n}\ell((T^{1/n}))$ does not contain a root of the polynomial $P(X) = X^p - X - T^{-1}$.

10 More ramification filtrations: the upper numbering

For various reasons, particularly motivated by local class field theory, we will need a different numbering of the ramification groups. The correct transition from the *lower numbering* to the *upper numbering* was discovered by Herbrand. Keep notation as in the previous lecture

Lemma 10.1. *There exists $x \in \mathfrak{o}_L$ such that $\mathfrak{o}_L$ is generated by x as a $\mathfrak{o}_K$-algebra. (This doesn't require L to be Galois over K, but it does require ℓ to be separable over k.)*

Proof. Recall that for π a uniformizer of $\mathfrak{o}_L$ and $b_0, \dots, b_{f-1} \in \mathfrak{o}_L$ lifting a basis of ℓ as a k-vector space, the elements $\pi^i b_j$ for $i = 0, \dots, e-1$ and $j = 0, \dots, f-1$ form an orthogonal basis of L over K, in the sense that for any choice of $c_{i,j} \in K$,

$$v_L\left(\sum_{i,j} c_{i,j}\pi^i b_j\right) = \min_{i,j}\{ev_K(c_{i,j}) + i\}.$$

In particular, the $\pi^i b_j$ form a basis for $\mathfrak{o}_L$ as a module over $\mathfrak{o}_K$.

By the primitive element theorem, we can find an element of ℓ which generates ℓ as a k-algebra. By lifting that element and then taking its powers, we may take $b_1, \dots, b_f$ to be of the form $1, x, \dots, x^{f-1}$ for some $x \in \mathfrak{o}_F$. We'll be finished if we can arrange for some polynomial $R(x)$ in x which is monic of degree f to be a uniformizer of $\mathfrak{o}_L$, in which case we take it for our value of π and then use $1, x, \dots, x^{ef-1}$ as a basis for $\mathfrak{o}_L$ as a $\mathfrak{o}_K$-algebra.

This amounts to a mutant variant of Hensel's lemma. Choose $R(T) \in \mathfrak{o}_K[T]$ to be a monic polynomial lifting the minimal polynomial of $\overline{x}$ (the image of x in ℓ) over k. We must have $v_L(R(x)) > 0$. If $v_L(R(x)) = 1$, then we're done. Otherwise, let π be any uniformizer of $\mathfrak{o}_L$, and write

$$R(x + \pi) \equiv R(x) + \pi R'(x) \pmod{\pi^2}.$$

Since ℓ is separable over k, $R'(x) \not\equiv 0 \pmod{\pi}$, so $v_L(R(x + \pi)) = 1$ and we can use $x + \pi$ instead of x. □

This gives us a handy way to compute the ramification filtration on a quotient group. Given x as above, for $g \in G$, put $i_G(g) = v_L(g(x) - x)$. Then $i_G(g) \geq i + 1$ if and only if $g \in G_i$, so this definition doesn't depend on x.

Lemma 10.2. *Let H be a normal subgroup of G with fixed field K'. Then for $\sigma \in G/H$,*

$$i_{G/H}(\sigma) = \frac{1}{e'} \sum_{g \to \sigma} i_G(g),$$

where the sum runs over $g \in G$ mapping to $\sigma \in G/H$, and $e' = e(L/K')$.

Proof. Choose $y \in \mathfrak{o}_{K'}$ generating $\mathfrak{o}_{K'}$ as an $\mathfrak{o}_K$-algebra. Pick one $g \in G$ mapping to $\sigma \in G/H$. The stated equality will follow if we show that the elements

$$a = g(y) - y, \qquad b = \prod_{h \in H} (gh(x) - x)$$

of $\mathfrak{o}_L$ have the same valuation, or equivalently, generate the same ideal. (The factor of e' occurs because $v_L(g(y) - y) = e' v_{K'}(g(y) - y) = e' i_{G/H}(\sigma)$.)

Let $f = \prod_{h \in H}(T - h(x)) \in \mathfrak{o}_{K'}[T]$ be the minimal polynomial of x over K'. Write $g(f) = \prod_{h \in H}(T - gh(x))$ for the polynomial obtained from f by applying g to each coefficient. Since $g(f) - f$ has all coefficients divisible by $g(y) - y$, $g(y) - y = a$ must divide $g(f)(x) - f(x) = \pm b$.

Conversely, we can write $y = P(x)$ for some $P(T) \in \mathfrak{o}_K[T]$. Then $P(T) - y$ is a polynomial in $\mathfrak{o}_{K'}[T]$ with $T = x$ as a root, so it is divisible by the minimal polynomial f in $K'[T]$ (and in $\mathfrak{o}_{K'}[T]$ by Gauss's lemma). Write $P(T) - y = f(T)h(T)$ with $h(T) \in \mathfrak{o}_{K'}[T]$; then apply g and put $T = x$ to get $P(x) - g(y) = g(f)(x)g(h)(x)$. The left side is $y - g(y) = -a$, which is divisible by $g(f(x)) = \pm b$. □

We can now define (following Herbrand) a new numbering on the ramification filtration, called the *upper numbering*, that is compatible with quotients rather than subgroups. It is useful to first extend the lower numbering notation to nonintegral indices: if $u \geq -1$, write G_u to mean G_i for $i = \lceil u \rceil$ the least integer greater than or equal to u. For $u \geq -1$, define

$$\phi(u) = \int_0^u \frac{dt}{[G_0 : G_t]};$$

this is the unique continuous, piecewise linear, increasing, concave function for which $\phi(0) = 0$ and $\phi'(u) = 1/[G_0 : G_u]$ for $u \geq -1$ not an integer. In particular, it has an inverse function ψ which is continuous, piecewise linear, increasing, and convex, and satisfies $\psi(0) = 0$. Also, the slopes of ψ are all integers, because they are the reciprocals of slopes of ϕ. We write $\phi_{L/K}$ and $\psi_{L/K}$ when we need to specify the extension L/K.

Lemma 10.3. *If v is an integer, then so is $u = \psi(v)$.*

Proof. Put $g_i = \#G_i$, and choose $m \in \mathbb{Z}$ for which $m \leq u \leq m + 1$. Then

$$v = \phi(u) = \frac{1}{g_0}(g_1 + \cdots + g_m + (u - m)g_{m+1}),$$

so

$$g_0 v = g_1 + \cdots + g_m + (u - m)g_{m+1}.$$

For $i \leq m$, G_{m+1} is a subgroup of G_i and so g_{m+1} is a divisor of g_i. Hence $u - m \in \mathbb{Z}$, forcing $u \in \mathbb{Z}$. □

The *upper numbering* of the ramification groups is given by

$$G^v = G_{\psi(v)}$$

or in other words

$$G^{\phi(u)} = G_u.$$

Theorem 10.4 (Herbrand)**.** *Let K' be a Galois subextension of L/K and put $H = \mathrm{Gal}(L/K')$.*

(a) We have

$$\phi_{L/K} = \phi_{K'/K} \circ \phi_{L/K'}, \qquad \psi_{L/K} = \psi_{L/K'} \circ \psi_{K'/K}.$$

(b) We have $(G/H)^v = G^v H/H$ for all $v \geq -1$.

Lemma 10.5. *We can rewrite $\phi_{L/K}(u) = \left(\frac{1}{g_0}\sum_{g\in G} \min\{i_G(g), u+1\}\right) - 1$.*

Proof. Both sides of the equation are continuous, piecewise linear functions which vanish at $u = 0$ (at which point the sum in parentheses contributes 1 for each $g \in G_0$ and 0 otherwise). Moreover, for $u \in (m, m+1)$ for any $m \in \mathbb{Z}$, the derivative of the right side is $1/g_0$ times the number of $g \in G$ for which $i_G(g) \geq m+2$, which is $g_{m+1}/g_0 = 1/[G_0 : G_{m+1}]$. □

Lemma 10.6. *For $\sigma \in G/H$, let $j(\sigma)$ be the maximum of the integers $i_G(g)$ over all $g \in G$ mapping to $\sigma \in G/H$. Then*

$$i_{G/H}(\sigma) - 1 = \phi_{L/K'}(j(\sigma) - 1).$$

Proof. Pick any $g \in G$ with image σ such that $i_G(g) = j(\sigma)$, and call this common value m. If $h \in H$ belongs to H_{m-1}, then $i_G(s) \geq m$ and so $i_G(gh) \geq m$, but this forces $i_G(gh) = m$. Otherwise, $i_G(h) < m$, so $i_G(gh) = i_G(h)$. We conclude that

$$i_G(gh) = \min\{i_G(h), m\}.$$

From Lemma 10.2,

$$i_{G/H}(\sigma) = \frac{1}{e(L/K')} \sum_{h\in H} \min\{i_G(h), m\},$$

but $e(L/K') = \#H_0$ and $i_G(h) = i_H(h)$. Applying Lemma 10.5 with the group H, we get $i_{G/H}(\sigma) = 1 + \phi_{L/K'}(m-1)$, as desired. □

Corollary 10.7. *For $v = \phi_{L/K'}(u)$, we have $G_u H/H = (G/H)_v$.*

Proof. Since ϕ is monotonic, $\sigma \in G_u H/H$ is equivalent to $j(\sigma) - 1 \geq u$ and thus to $\phi_{L/K'}(j(\sigma) - 1) \geq v$. But that just says that $i_{G/H}(\sigma) - 1 \geq v$ by Lemma 10.6, or $\sigma \in (G/H)_v$ □

Proof of Theorem 10.4. For part (a), both $\phi_{L/K}$ and $\phi_{K'/K} \circ \phi_{L/K'}$ are continuous, piecewise-linear functions vanishing at $u = 0$, so it suffices to compare their derivatives at $u \in (m, m+1)$ for any $m \in \mathbb{Z}$. By the chain rule, the derivative of $\phi_{K'/K} \circ \phi_{L/K'}$ at u is $\phi'_{K'/K}(v)\phi'_{L/K'}(u)$ for $v = \phi_{L/K'}(u)$. Recall that this equals

$$\frac{\#(G/H)_v}{e_{K'/K}} \frac{\#H_u}{e_{L/K'}}.$$

The denominators multiply to $e_{L/K}$, while the numerators multiply to $\#G_u$ by Corollary 10.7 and the fact that $\#(G_u H/H)\#H_u = \#G_u$ (because $G_u H/H \cong G_u/(G_u \cap H)$). We thus get the same thing as the derivative of $\phi_{L/K}$, as claimed. (The ψ relation follows from the ϕ relation by taking inverses.)

For (b), write $(G/H)^v = (G/H)_x$ with $x = \psi_{K'/K}(v)$. By Corollary 10.7, $(G/H)_x = G_w H/H$ with $w = \psi_{L/K'}(x)$, but this equals $\psi_{L/K}(v)$ by (a). □

I hope to get to the following fact later.

Theorem 10.8 (Hasse-Arf theorem)**.** *If G is an abelian group, then the values v at which G^v changes size are all integers.*

This is not true in general; see exercises.

Exercises

1. (from Serre) Let G be the quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$, and let $C = \{\pm 1\}$ be the center. Supose that G is the Galois group of a totally ramified extension L/K and $G_4 = \{1\}$. Show that $G^v = G$ for $v \leq 1$, $G^v = C$ for $1 < v \leq 3/2$, and $G^v = \{1\}$ for $v > 3/2$. In particular, G^v changes size at the nonintegral value $v = 3/2$, showing that the Hasse-Arf theorem cannot be extended to nonabelian groups.

2. Prove that the situation in the previous exercise can occur for K a finite extension of $\mathbb{Q}_2$. (Hint: I think this can be done for $K = \mathbb{Q}_2$, and that one can find examples in the online *Database of Local Fields*.)

11 Discriminant and different

Here's one of the key motivations for defining the ramification filtration.

Let K be a field carrying a discrete valuation v_K (completeness is not needed just yet). Let L be a finite extension of K. The *trace map* $\mathrm{Trace}_{L/K} : L \to K$ is the K-linear map obtained by viewing L as a finite-dimensional K-vector space, and defining $\mathrm{Trace}_{L/K}(x)$ to be the trace of the multiplication-by-x map on K. If L/K is Galois, this agrees with the usual definition that $\mathrm{Trace}_{L/K}(x)$ is the sum of the conjugates of x in L (because the eigenvalues of the multiplication-by-x map are precisely these conjugates, and the trace of a matrix equals the sum of its eigenvalues).

It turns out that L/K is separable if and only if the *trace pairing* $\langle x, y \rangle \to \mathrm{Trace}_{L/K}(xy)$ is nondegenerate (i.e., every nonzero element of L has nonzero pairing with something). If we let R denote the valuation subring of K and S the valuation subring of L, then the *discriminant* of L/K is defined as the ideal generated by the determinant of the matrix $(\mathrm{Trace}_{L/K}(e_i e_j))_{i,j=1}^n$ for any basis $e_1, \ldots, e_n$ of S as a module over R. (Changing the basis modifies this matrix but does not change its determinant.) For example, L/K is unramified if and only if the discriminant is the unit ideal.

An important point to note is that the discriminant makes sense even if L is not a field, but only a direct sum of fields. In particular, if $\widehat{K}$ is the completion of K, then we can talk about the discriminant of $L \otimes_K \widehat{K}$ over $\widehat{K}$. Since $L \otimes_K \widehat{K}$ splits up as a product of fields (the completions of L under the different extensions of v_K), the discriminant also splits as the product of the discriminants of these individual field extensions.

The discriminant is a somewhat crude invariant of the extension L/K. A somewhat more refined invariant is provided by computing the *codifferent*, or *inverse different*, of L/K;

this is the S-submodule of L consisting of those $x \in S$ for which $\mathrm{Trace}_{L/K}(xy) \in R$ for all $y \in S$. This turns out to be a finite S-submodule of L, so it is generated by some power of a uniformizer of S. The ideal in S generated by the inverse of that power is called the *different* of L/K. Again, it is the unit ideal if and only if L/K is unramified.

Just like the discriminant, the different splits up into contributions from different extensions of the valuation v_L. One way to say this is to again view $L \otimes_K \widehat{K}$ as a product of fields, each carrying a unique extension of v_K from $\widehat{K}$, and form the product of the valuation subrings. If we take the product of the differents corresponding to the various extensions $\widehat{L}$ of $\widehat{K}$, we get the same ideal as generated by the different of L over K. See also Proposition 10 in section III.4 of Serre.

Here's a link to the ramification groups, at least in the lower numbering.

Proposition 11.1. *Suppose K is complete (as then is L for a unique extension of v_K) and L is Galois over K with separable residue field extension. Then the L-valuation of the different of L/K equals*

$$\sum_{g \in G - \{e\}} i_G(g) = \sum_{i=0}^{\infty} (\#G_i - 1).$$

Proof. Let x be a generator of $\mathfrak{o}_L$ as a $\mathfrak{o}_K$-algebra. The quantity in question is then the valuation of $\prod_{g \neq e}(x - g(x))$, which is equal to $f'(x)$ for f the minimal polynomial of x over K. We thus need to show that the inverse different is generated by $f'(x)$; this follows from a hopefully familiar calculation: if $n = \deg(f)$, then

$$\mathrm{Trace}_{L/K}\left(\frac{x^i}{f'(x)}\right) = \begin{cases} 0 & i = 0, \dots, n-2 \\ 1 & i = n-1. \end{cases}$$

(Proof: write $1/f(T) = \sum_{i=1}^{n} 1/(f'(x_i)(T - x_i))$ for $x_1, \dots, x_n$ the roots of f, then expand in power series in T.) □

Corollary 11.2. *Let K' be the subextension of L corresponding to the (not necessarily normal) subgroup H of G. Then the K'-valuation of the different of L/K equals*

$$\frac{1}{e(L/K')} \sum_{g \notin H} i_G(h).$$

One can in principle determine the discriminant from the different, but this looks messy. The key point here is that the different lives "in L", so it is linked to the lower numbering filtration (which is defined using the valuation on L, and behaves well when fixing L and changing the field under it). The discriminant, however, lives "in K", so it is better to treat it using the upper numbering filtration, which is optimized to have good behavior with respect to K (e.g., when fixing K and changing the field over it). We'll come back to this after we discuss the Hasse-Arf theorem and Artin representations.

12 Example: cyclotomic extensions

Theorem 12.1. *Put $K = \mathbb{Q}_p$ and $L = \mathbb{Q}(\zeta_{p^n})$ for ζ_{p^n} a primitive p^n-th root of unity. Then the jumps in the upper numbering filtration on $G = \mathrm{Gal}(L/K)$ are $0, 1, \dots, n-1$ when $p \neq 2$, and $1, \dots, n-1$ when $p = 2$. In particular, they are all integers.*

There are a few glitches in the proof for $p = 2$, so we leave it as an exercise to iron those out.

Proof for $p \neq 2$. We first check that $\mathfrak{o}_L = \mathbb{Z}_p[\zeta_{p^n}]$. Note that $\zeta_{p^n} - 1$ is a root of the polynomial

$$P(T) = (T+1)^{(p-1)p^{n-1}} + (T+1)^{(p-2)p^{n-1}} + \cdots + 1.$$

Modulo p this polynomial is congruent to $((T+1)^{p^n} - 1)/((T+1)^{p^{n-1}} - 1) = T^{(p-1)p^{n-1}}$, so its nonleading coefficients are all in $p\mathbb{Z}_p$. Moreover, the constant coefficient is p. This is thus an Eisenstein polynomial, i.e., its Newton polygon is flat with slope $1/((p-1)p^{n-1})$. This forces P to be irreducible, and also forces $f(K_1/K) \geq (p-1)p^{n-1}$. Since this accounts for the whole degree, $e(K_1/K) = 1$, $f(K_1/K) = (p-1)p^{n-1}$, and $\zeta_{p^n} - 1$ is a uniformizer.

The breaks in the lower numbering filtration can now be computed as $v_L(g(\zeta_{p^n} - 1) - (\zeta_{p^n} - 1))$ for g running over the Galois group. Identify G with $(\mathbb{Z}/p^n\mathbb{Z})^\times$, so that $g(\zeta_{p^n}) = \zeta_{p^n}^g$; then $g(\zeta_{p^n} - 1) - (\zeta_{p^n} - 1) = \zeta_{p^n}(\zeta_{p^n}^{g-1} - 1)$. If $g - 1$ is exactly divisible by p^m, then $\zeta_{p^n}^{g-1}$ is a primitive p^{n-m}-th root of unity. We thus see that the lower numbering breaks occur at $p^m - 1$ for $m = 0, \dots, n-1$. It thus remains to compute the functions $\phi_{L/K}$ and $\psi_{L/K}$ to see what happens in the upper numbering; we leave this as an exercise. □

Exercises

1. Complete the proof of Theorem 12.1 by computing the functions $\psi_{L/K}$ and $\phi_{L/K}$.

2. Extend the proof of Theorem 12.1 to the case $p = 2$.

13 The norm map: unramified case

For K a field and L a finite extension of K, the *norm* $N = N_{L/K} : L \to K$ can be defined in two ways. One is that the norm of $x \in L$ is the determinant of the multiplication-by-x map viewed as a K-linear endomorphism of L. The other is that it's the product of the conjugates of x in an algebraic closure of K (provided that you count multiplicities in case L is not separable over K).

Now let K be complete for a discrete valuation, and assume that the residue field ℓ of L is separable over the residue field k of K. For $f = f(L/K)$, we have a diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U_L & \longrightarrow & L^\times & \longrightarrow & \mathbb{Z} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle N} & & \downarrow{\scriptstyle N} & & \downarrow{\scriptstyle \times f} & & \\ 0 & \longrightarrow & U_K & \longrightarrow & K^\times & \longrightarrow & \mathbb{Z} & \longrightarrow & 0 \end{array}$$

with exact rows, where U_K is the units of $\mathfrak{o}_K$ congruent to 1 modulo the maximal ideal, and similarly for L. As we saw when considering the ramification filtration, it is useful to filter the group U_K using the subgroups U_K^i of units congruent to 1 modulo the i-th power of the maximal ideal. One can then ask how the norm maps interact with these subgroups. This is easiest to see in the case of an unramified extension.

Proposition 13.1. *If L/K is Galois and unramified, then $N_{L/K}$ maps U_L^n into U_K^n for all $n \geq 1$.*

This doesn't require the Galois condition; see exercises.

Proof. Write $x \in U_L^n$ as $1 + y$ with $v_L(y) \geq n$. Put $G = \mathrm{Gal}(L/K)$ and write $N_{L/K}(x) = \prod_{g \in G}(1 + g(y))$, which equals 1 plus a sum of terms with L-valuation at least n. That sum is in K, and the unramified condition means that K-valuations and L-valuations agree for elements of K, so we have an element of U_K^n. □

If we identify $\mathfrak{o}_K^\times / U_K^1$ with the multiplicative group $k^\times$, and similarly for ℓ, then $N_{L/K}$ induces a map $\ell^\times \to k^\times$ which is just the norm map. For $n \geq 1$, we may identify U_L^n / U_L^{n+1} with $\mathfrak{p}_L^n / \mathfrak{p}_L^{n+1}$, which in turn we may canonically identify with $\mathfrak{p}_K^n / \mathfrak{p}_K^{n+1} \otimes_k \ell$. In this representation, $N_{L/K}$ induces a map $U_L^n / U_L^{n+1} \to U_K^n / U_K^{n+1}$ corresponding to the map $\mathrm{id} \otimes \mathrm{Trace}_{\ell/k} : \mathfrak{p}_K^n / \mathfrak{p}_K^{n+1} \otimes_k \ell \to \mathfrak{p}_K^n / \mathfrak{p}_K^{n+1}$. Note that this map is surjective because the trace map is surjective for the finite separable extension ℓ/k.

Proposition 13.2. *For $n \geq 1$, $N_{L/K}(U_L^n) = U_K^n$. Also $U_K / N_{L/K}(U_L) \cong k^\times / N_{L/K}(\ell^\times)$, and $K^\times / N_{L/K}(L^\times) = \mathbb{Z}/f\mathbb{Z} \times k^\times / N_{L/K}(\ell^\times)$.*

Proof. The first two statements follow from the previous discussion plus the snake lemma on the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U_L^1 & \longrightarrow & U_L & \longrightarrow & \ell^\times & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle N} & & \downarrow{\scriptstyle N} & & \downarrow{\scriptstyle \times f} & & \\ 0 & \longrightarrow & U_K^1 & \longrightarrow & U_K & \longrightarrow & k^\times & \longrightarrow & 0. \end{array}$$

The third statement follows by writing $K^\times = \mathbb{Z} \times U_K$ and similarly for K, where $1 \in \mathbb{Z}$ corresponds to your favorite uniformizer of K (which is also a uniformizer of L because L/K is unramified). □

Corollary 13.3. *If k and ℓ are finite, then $U_K = N_{L/K}(U_L)$.*

Proof. This holds because the map $N_{\ell/k} : \ell^\times \to k^\times$ is always surjective (count the elements of its kernel). □

13.1 Exercises

1. Show that Proposition 13.1 still holds without the unramified condition. (Hint: the Galois closure of an unramified extension is unramified.)

2. Do likewise for the other results in this lecture.

14 The norm map: totally ramified cases

Since we understand the behavior of the norm map on the groups U_L^n in the unramified case, it is not too serious to consider only the totally ramified case. Let's add that restriction now. The goal is to prove the following.

Theorem 14.1. *Write $\psi = \psi_{L/K}$. For $n \geq 0$ an integer, $N_{L/K}(U_L^{\psi(n)}) \subseteq U_K^n$ and $N_{L/K}(U_L^{\psi(n)+1}) \subseteq U_K^{n+1}$.*

The first thing to note is that we can induct on the order of G. Since G is solvable, if it is nontrivial then it always has a nontrivial proper normal subgroup H, corresponding to a Galois subextension K'/K of L/K. The first assertion follows easily from the induction hypothesis and Herbrand's rule $\psi_{L/K} = \psi_{L/K'} \circ \psi_{K'/K}$.

Again because G is solvable, we can always reduce to the case of a cyclic group of prime order l, which may or may not be equal to p (although the case $l \neq p$ is tame, which makes it fairly easy).

Lemma 14.2. *Assume G is cyclic of prime order l, and that the unique ramification break of G is at t (in either numbering!).*

(a) The different of L/K is $\mathfrak{p}_L^m$ for $m = (t+1)(l-1)$.

(b) For $n \geq 0$, $\mathrm{Trace}(\mathfrak{p}_L^n) = \mathfrak{p}_K^r$ for $r = \lfloor (m+n)/l \rfloor$.

Proof. Part (a) follows from Proposition 11.1. Part (b) follows from (a) by observing that from the definition of the different, $\mathrm{Trace}(\mathfrak{p}_L^n) \subseteq \mathfrak{p}_K^r$ if and only if $\mathfrak{p}_L^n$ belongs to $\mathfrak{p}_K^r$ times the inverse different. □

Here's the key calculation.

Lemma 14.3. *For $x \in \mathfrak{p}_L^n$ with $n \geq 1$, we have*

$$N_{L/K}(1+x) \equiv 1 + \mathrm{Trace}_{L/K}(x) + N_{L/K}(x) \pmod{\mathrm{Trace}(\mathfrak{p}_L^{2n})}. \tag{3}$$

Proof. If we expand $N(1+x) = \prod_{g \in G}(1+g(x))$, the initial term 1, the terms with one g, and the final term with all of the g correspond to the three terms on the right side of the desired expression. For $k = 2, \ldots, l-1$, the sum of the k-fold products of the $g(x)$ is the trace of some element of $\mathfrak{p}_L^{kn}$, giving the claim. □

To prove Theorem 14.1, we should do a bit more. Note that given Theorem 14.1 for a given n, we once again get a map $N_n : U_L^{\psi(n)}/U_L^{\psi(n)+1} \to U_K^n/U_K^{n+1}$; we need to calculate this map for a given n in order to get Theorem 14.1 for $n+1$. Here's the answer.

Proposition 14.4. *(a) For $n = 0$, $N_n : k^\times \to k^\times$ is the l-th power map. If $t \neq 0$, then $l = p$ and this map is injective. If $t = 0$, then $l \neq p$ and the map has kernel equal to the image of G in U_L/U_L^1.*

(b) *For $1 \leq n < t$, we can find $\alpha_n \in k^\times$ so that $N_n(\xi) = \alpha_n \xi^p$. This map is injective, and also surjective if k is perfect.*

(c) *For $n = t$, we can find $\alpha, \beta \in k^\times$ so that $N_n(\xi) = \alpha\xi^p + \beta\xi$. This map has kernel of order p, equal to the image of θ_t. It is surjective if k is algebraically closed.*

(d) *For $n = t$, we can find $\beta_n \in k^\times$ so that $N_n(\xi) = \beta\xi$. This map is bijective.*

Proof. Case (a) is easy. In (3), when $n < t$, the term $N(x)$ dominates; when $n > t$, the term Trace(x) dominates; and when $n = t$, the two are equal. (This follows from Lemma 14.2 and some fairly menial calculation which can be found in Serre V.3.) This gives the desired results. □

Using this reasoning, we also get the following. (See Serre V.6 for omitted details.)

Theorem 14.5. *The sequence*

$$0 \to G_{\psi(n)}/G_{\psi(n)+1} \overset{\theta_{\psi(n)}}{\to} U_L^{\psi(n)}/U_L^{\psi(n)+1} \overset{N_n}{\to} U_K^n/U_K^{n+1},$$

in which the last map is induced by the norm via the previous theorem, is exact. Consequently, the map $N_n : U_L^{\psi(n)}/U_L^{\psi(n)+1} \to U_K^n/U_K^{n+1}$ is injective if and only if $G_{\psi(n)} = G_{\psi(n)+1}$. (If this holds and also k is perfect, then N_n is also surjective.)

Proof. Again, induct on the order of G. □

Exercises

1. Put $K = \mathbb{F}_p((t))$ and $L = K[z]/(z^p - z - t^{-m})$ for m a positive integer coprime to p, so that L/K is a Galois extension with group $G = \mathbb{Z}/p\mathbb{Z}$ (an Artin-Schreier extension). Find the unique break in the ramification filtration on G, as a function of m.

15 Artin representations

The ramification filtration on the Galois group of an extension of local fields also gives useful information about linear representations of these groups. This information shows up when you try to make a careful study of Artin L-functions (without throwing out the primes of bad reduction).

Keep notation as before, but now let $\rho : G \to \mathrm{GL}(V)$ be a linear representation of G on a finite-dimensional complex vector space V. Recall that ρ is determined up to isomorphism by its character $\chi : G \to \mathbb{C}$, defined by $\chi(g) = \mathrm{Trace}(\rho(g))$. The dimension of V is also called the *degree* of χ, although this is confusing. When $\dim(V) = 1$, ρ and χ are basically the same thing.

We can think of a linear representation as giving (and being given by) a left $\mathbb{C}[G]$-module structure on V, where $\mathbb{C}[G]$ is the (not necessarily commutative!) group algebra of G over $\mathbb{C}$.

If H is a subgroup of G and W is a left $\mathbb{C}[H]$-module, then $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$ is a left $\mathbb{C}[G]$-module. The corresponding operation on representations is called *induction* of a representation (or its character) from H to G.

Here's a crucial about linear representations of finite groups. See for instance Serre's *Linear Representations of Finite Groups*.

Theorem 15.1 (Brauer)**.** *Every character of G is a $\mathbb{Z}$-linear combinations of one-dimensional characters induced from subgroups of G.*

So far we've just been talking about G as an abstract group. Now let's add the Galois theory. For $g \in G$ not equal to e, put $a_G(g) = -f(L/K) i_G(g)$. Then set $a_G(g) = f(L/K) \sum_{g \neq e} i_G(g)$, to ensure that $\sum_{g \in G} a_G(g) = 0$. Define the *conductor* of χ (or of ρ) to be the rational number

$$f(\chi) = \frac{1}{\#G} \sum_{g \in G} a_G(g^{-1}) \chi(g).$$

We can make the same definition for any class function (i.e., any conjugacy-invariant function) on G.

If we define $\chi(H)$ for H a subgroup of G to mean the *average* of χ over H, it is elementary to compute that

$$f(\chi) = \sum_{i=0}^{\infty} \frac{\#G_i}{\#G_0} (\chi(1) - \chi(G_i)).$$

Easy corollary:

$$f(\chi) = \sum_i \frac{\#G_i}{\#G_0} \operatorname{codim} V^{G_i},$$

where V^{G_i} is the subspace of V fixed by G_i.

Proposition 15.2. *Let H be a subgroup of K corresponding to the subextension K'/K. Let λ be the K-valuation of the discriminant of K' over K. Then the restriction of a_G to H equals $\lambda r_H + f(K'/K) a_H$, where r_H equals the character of the regular representation of H (which equals $\#H$ on the identity element and 0 elsewhere).*

Proof. For $g \neq e$ in H, the claim follows by writing $a_G(g) = -f_{L/K} i_G(g)$ and similarly for H, noting that $r_H(s) = 0$, and observing that $i_G(g) = i_H(g)$. For $g = e$, this reduces to the transitivity of the discriminant. □

By Frobenius reciprocity for characters, we deduce the following.

Corollary 15.3. *For χ a character of H and χ^* the induced character on H,*

$$f(\chi^*) = \lambda \chi(1) + f(K'/K) f(\chi).$$

Corollary 15.4. *Given a one-dimensional character χ with kernel H, let K' be the subextension of L/K corresponding to H. If K'/K is unramified, then $f(\chi) = 0$; otherwise, $f(\chi) - 1$ is equal to the largest ramification break of G/H for the upper numbering.*

Theorem 15.5 (Artin). *The conductor $f(\chi)$ is always an integer. Equivalently, the function a_G is itself the character of a linear representation (since $f(\chi)$ is the inner product of this character with χ).*

In the case of dimension 1, this is the Hasse-Arf theorem (for which I still owe you a proof). The general case reduces to this using Brauer's theorem and Corollary 15.3.

Corollary 15.4 has the following useful generalization: if χ is an irreducible character of dimension d with kernel H, then $f(\chi) = d(1+i)$ for i the largest ramification break of G/H for the upper numbering (taking this to be -1 if there are no breaks at all).

16 Statements of local class field theory

We now describe the formalism of *local class field theory*, the classification of abelian extensions of a local field K with finite residue field. By our earlier classification of local fields, K is either a finite extension of $\mathbb{Q}_p$, or is isomorphic to a power series field $\mathbb{F}_q((t))$. Let K^{ab} be the *maximal abelian extension* of K, i.e., the compositum of all finite abelian extensions of K within some algebraic closure of K. For example, if $K = \mathbb{Q}_p$, then K^{ab} is the union of the cyclotomic extensions $\mathbb{Q}_p(\zeta_n)$ by the local Kronecker-Weber theorem.

Theorem 16.1 (Local reciprocity law). *There is a unique map $\phi_K : K^\times \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$ satisfying the following conditions.*

(a) For any uniformizer π in K and any finite unramified extension L of K, $\phi_K(\pi)$ acts on L as the Frobenius automorphism. (That is, it acts on the residue field of L as the q-th power map.)

(b) For any finite abelian extension L of K, ϕ_K induces an isomorphism $K^\times / \mathrm{Norm}_{L/K}(L^\times) \to \mathrm{Gal}(K^{\mathrm{ab}}/K) \to \mathrm{Gal}(L/K)$.

This says that abelian subgroups of K are identified by the norm groups. One can of course take norm groups for nonabelian extensions, but this doesn't add anything.

Theorem 16.2 (Norm limitation theorem). *Let L be a finite Galois extension of K, and let M/K be the maximal abelian subextension. Then $\mathrm{Norm}_{L/K}(L^\times) = \mathrm{Norm}_{M/K}(M^\times)$.*

There is a converse which says that every possible norm group arises from some abelian extension.

Theorem 16.3 (Local existence theorem). *For every open subgroup U of $K^\times$, there exists a finite abelian extension L of K such that $U = \mathrm{Norm}_{L/K}(L^\times)$.*

The ramification filtration shows up as follows. (The Hasse-Arf theorem guarantees that we see all of the ramification breaks this way.)

Theorem 16.4 (Ramification filtration). *For L a finite abelian extension of K, $G = \mathrm{Gal}(L/K)$, and i a nonnegative integer, the inverse image of G^i under $\phi_K : K^\times \to G$ is the group U_i^K of units in $\mathfrak{o}_K$ congruent to 1 modulo π^i (for π a uniformizer of K).*

The proofs of these results involve some homological algebra involving the actions of Galois groups; we turn to this next.

17 Galois homology and cohomology

Let G be a finite group. A *left G-module* will be the same thing as a left $\mathbb{Z}[G]$-module, for $\mathbb{Z}[G]$ the (not necessarily commutative!) group algebra of G over the ring $\mathbb{Z}$. That is, a left G-module is an abelian group equipped with a left action of G.

From general results in homological algebra (see your favorite graduate algebra text), one obtains a collection of functors $M \rightsquigarrow H^i(G, M)$ from left G-modules to abelian groups, with the property that $H^0(G, M) = M^G$ (the subgroup of invariants), and every short exact sequence

$$0 \to M \to N \to P \to 0$$

gives rise to a long exact sequence

$$0 \to M^G \to N^G \to P^G \to H^1(G, M) \to H^1(G, N) \to H^1(G, P) \to H^2(G, M) \to \cdots .$$

That is, the $H^i(G, M)$ resolve the fact that taking invariants is left exact but not right exact. Also, any commuting diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M_1 & \longrightarrow & N_1 & \longrightarrow & P_1 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & M_2 & \longrightarrow & N_2 & \longrightarrow & P_2 & \longrightarrow & 0 \end{array}$$

gives rise to a commuting diagram

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & H^0(G, M_1) & \longrightarrow & H^0(G, N_1) & \longrightarrow & H^0(G, P_1) & \longrightarrow & H^1(G, M_1) & \longrightarrow & \cdots \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & H^0(G, M_1) & \longrightarrow & H^0(G, N_2) & \longrightarrow & H^0(G, P_2) & \longrightarrow & H^1(G, M_2) & \longrightarrow & \cdots \end{array}$$

There is a fairly explicit description of these groups, but we won't have much use for it except for $i = 1$ and maybe $i = 2$.

If you think as the invariants M^G of a left G-module M as comprising the largest submodule of M on which G acts trivially, it is natural to dualize and form the largest quotient of M on which G acts trivially. In other words, take the quotient by the subgroup generated by $g(x) - x$ for all $g \in G, x \in M$; this gives the group M_G of *coinvariants*. The functor of coinvariants is right exact, so more general homological algebra produces covariant functors $M \rightsquigarrow H_i(G, M)$ with the property that $H_0(G, M) = M_G$, and every short exact sequence

$$0 \to M \to N \to P \to 0$$

gives rise to a long exact sequence

$$\cdots \to H_2(G, P) \to H_1(G, M) \to H_1(G, N) \to H_1(G, P) \to M_G \to N_G \to P_G \to 0.$$

The two long exact sequences glue together as follows. Define the map $\mathrm{Norm}_G : M \to M$ by the formula

$$\mathrm{Norm}_G(x) = \sum_{g \in G} g(x).$$

(You might prefer to call this *trace*, but it's more common to think of M as being a multiplicative group.) The *Tate cohomology groups* are then defined as follows:

$$H^i_T(G, M) = \begin{cases} H^i(G, M) & i > 0 \\ M^G / \mathrm{Norm}_G(M) & i = 0 \\ \ker(\mathrm{Norm}_G)/I_G M & i = -1 \\ H_{-i-1}(G, M) & i < -1. \end{cases}$$

Here I_G denotes the *augmentation ideal* of $\mathbb{Z}[G]$, i.e., the kernel of the homomorphism $\mathbb{Z}[G] \to \mathbb{Z}$ mapping each group element to 1. (Note that $M_G = M/I_G M$ and that $I_G M \subseteq \ker(\mathrm{Norm}_G)$.) Now a short exact sequence $0 \to M \to N \to P \to 0$ turns into a doubly infinite sequence

$$\begin{aligned} &\cdots H^{-2}_T(G, P) \to H^{-1}_T(G, M) \to H^{-1}_T(G, N) \to H^{-1}_T(G, P) \to \\ &\to H^0_T(G, M) \to H^0_T(G, N) \to H^0_T(G, P) \to H^1_T(G, M) \to \cdots . \end{aligned}$$

The Tate cohomology groups might look a bit mysterious, except in indices $i = 0$ and $i = -1$ where we have explicit formulas for them. An important fact that helps us compute them is that $H^i_T(G, \mathbb{Z}[G]) = 0$ for all $i \in \mathbb{Z}$.

Lemma 17.1. *For any G-module M, if we equip $\mathbb{Z}[G] \otimes_{\mathbb{Z}} M$ with the diagonal G-action, we have $H^i_T(G, \mathbb{Z}[G] \otimes_{\mathbb{Z}} M) = 0$ for all $i \in \mathbb{Z}$.*

Proof. Suppose first that M carries the trivial action. If M is a free $\mathbb{Z}$-module, then $\mathbb{Z}[G] \otimes_{\mathbb{Z}} M$ is a free $\mathbb{Z}[G]$-module and so the claim follows. Otherwise, since $\mathbb{Z}$ is a principal ideal domain, given any surjection $F_1 \to M$ of $\mathbb{Z}$-modules with F_1 free, the kernel F_2 is also free. Now take the long exact sequence associated to

$$0 \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} F_1 \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} F_2 \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M \to 0$$

and use the previous observation.

In general, we can rewrite $\mathbb{Z}[G] \otimes_{\mathbb{Z}} M \cong \mathbb{Z}[G] \otimes_{\mathbb{Z}} N$ for N a copy of the underlying $\mathbb{Z}$-module of M, but with the trivial action: the map takes $g \otimes x$ to $g \otimes g^{-1}(x)$. We may then apply the previous paragraph. □

One important case for local class field theory: from the sequence

$$0 \to I_G \to \mathbb{Z}[G] \to \mathbb{Z} \to 0$$

and the vanishing of cohomology in the middle (Lemma 17.1), we get $H_T^{-2}(G,\mathbb{Z}) \cong H_T^{-1}(G, I_G) = I_G/I_G^2$. This turns out to be canonically isomorphic to G^{ab}. This means that for $G = \mathrm{Gal}(L/K)$, to produce the local reciprocity law it is enough to exhibit a map

$$\mathrm{Gal}(L/K)^{\mathrm{ab}} = H_T^{-2}(G,\mathbb{Z}) \to H_T^0(G, L^\times) = K^\times/\operatorname{Norm}_{L/K}(L^\times),$$

and this is exactly how we will proceed to do so.

Here's a case where the explicit formulas at $i = 0$ and $i = -1$ suffice to compute all of the cohomology groups explicitly!

Lemma 17.2 (Tate)**.** *If G is a cyclic group, then there is a isomorphism $H_T^i(G,M) \to H_T^{i+2}(G,M)$ for each $i \in \mathbb{Z}$ and each G-module M. Moreover, this isomorphism is functorial in M.*

Proof. Choose a generator g of G. The key here is the exact sequence of left $\mathbb{Z}[G]$-modules

$$0 \to \mathbb{Z} \to \mathbb{Z}[G] \to \mathbb{Z}[G] \to \mathbb{Z} \to 0$$

in which the first map is $1 \mapsto [1]$, the second is $[h] \mapsto [gh]-[h]$, and the third is $[h] \to 1$. Since everything is a free $\mathbb{Z}$-module, tensoring over $\mathbb{Z}$ with M and taking the diagonal G-actions produces another exact sequence

$$0 \to M \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M \to M \to 0$$

of left G-modules, using the diagonal action on the tensor products. Split this into two short exact sequences

$$0 \to M \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M \to P, \qquad 0 \to P \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M \to M \to 0.$$

Note that $\mathbb{Z}[G] \otimes_{\mathbb{Z}} M$ has zero Tate cohomology groups by Lemma 17.1, so by taking long exact sequences we get

$$H_T^{i+1}(G,P) \cong H_T^{i+2}(G,M), \qquad H_T^i(G,M) \cong H_T^{i+1}(G,P).$$

This gives the claim. □

In order to do induction on the size of G, we need to relate cohomology groups for different G. This depends on the fact that group cohomology is functorial in G in the following sense: if $G \to G'$ is a group homomorphism, and M is a left G'-module, then we can view M also as a left G-module, and we get maps $H^i(G',M) \to H^i(G,M)$ that are natural in the groups (i.e., composing group homomorphisms does what you want) and in M. For instance, for $i = 0$, this is the natural inclusion $M^{G'} \to M^G$. (The general case follows from this by realizing that $H^i(G,M)$ is obtained by taking an injective resolution of M, taking invariants, then taking the homology of the resulting complex.)

Lemma 17.3. *For any G-module M , $H_T^i(G, \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], M)) = 0$ for all $i \in \mathbb{Z}$. (The action of G on $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], M)$ is such that $g \in G$ acts taking $f : \mathbb{Z}[G] \to M$ to $f' : \mathbb{Z}[G] \to M$ with $f'(h) = g(f(hg))$.)*

Proof. Again for $M = \mathbb{Z}$ with the trivial action, this holds because $\mathbb{Z}[G] \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], \mathbb{Z})$ via the map taking g to the homomorphism taking h to 1 if $h = g$ and 0 otherwise. We thus deduce the claim for M free with the trivial action, and then argue for arbitrary M with the trivial action as in Lemma 17.3. Finally, for general M, rewrite $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], M) \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], N)$ for N the underlying $\mathbb{Z}$-module of M, by sending the map $f : G \to M$ to the map $f' : G \to N$ with $f'(g) = g^{-1}(f(g))$. □

Lemma 17.4 (Inflation-restriction sequence)**.** *Let G be a finite group, let H be a normal subgroup of G, and let M be a left G-module. Then the sequence*

$$0 \to H^1(G/H, M^H) \to H^1(G, M) \to H^1(H, M)$$

is exact. If moreover $H^1(H, M) = 0$, then the sequence

$$0 \to H^2(G/H, M^H) \to H^2(G, M) \to H^2(H, M)$$

and so on.

Proof. The first assertion is most easily proved by expressing $H^1(G, M)$ in terms of *crossed homomorphisms*. Namely, $H^1(G, M)$ consists of maps $f : G \to M$ satisfying

$$f(gh) = g(f(h)) + f(h),$$

modulo maps of the form $f(g) = g(x) - x$ for some $h \in M$. With this interpretation, if f is a crossed homomorphism which becomes trivial on H, then we can choose $x \in M$ with $f(h) = h(x) - x$for all $h \in H$, and then $g \mapsto f(g) - g(x) + g$ is a well-defined crossed homomorphism on G/H with values in M^H.

We reduce the second assertion to the first by shifting dimensions. Put $N = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], M)$; this module receives a map from M taking x to the map $g \to g(x)$. Form the exact sequence

$$0 \to M \to N \to P \to 0$$

and note that

$$0 \to M^H \to N^H \to P^H \to H^1(H, M) = 0$$

is exact. We now get a commuting diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & H^1(G/H, P^H) & \longrightarrow & H^1(G, P) & \longrightarrow & H^1(H, P) \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & H^2(G/H, M^H) & \longrightarrow & H^2(G, M) & \longrightarrow & H^2(H, M) \end{array}$$

in which the vertical arrows are isomorphisms because N has zero cohomology by Lemma 17.3. □

18 Tate's criterion for local class field theory

Recall that to establish the local reciprocity law, we need to produce an isomorphism

$$\mathrm{Gal}(L/K)^{\mathrm{ab}} = H_T^{-2}(G, \mathbb{Z}) \to H_T^0(G, L^\times) = K^\times / \mathrm{Norm}_{L/K}(L^\times).$$

We'll now show Tate's group-theoretic method for doing so. (This can be generalized by setting up the notion of a *class formation*, thus providing an approach to global class field theory.)

Lemma 18.1. *Let G be a finite solvable group. Let M be a left G-module. Suppose that for each subgroup H of G, $H^1(H, M) = H^2(H, M) = 0$. Then $H_T^i(G, M) = 0$ for all $i \in \mathbb{Z}$.*

Proof. We first check that $H_T^0(G, M) = 0$. For G cyclic, this follows from Lemma 17.2. To prove the general case, induct on $\#G$. If G is solvable and not cyclic, we can find a subgroup H such that G/H is cyclic and not trivial. By the induction hypothesis, $H_T^i(H, M) = 0$ for all $i \in \mathbb{Z}$. Also, since $H^1(H, M) = 0$,

$$0 \to H^2(G/H, M^H) \to H^2(G, M) \to H^2(H, M)$$

is exact by inflation-restriction (Lemma 17.4), so $H^2(G/H, M^H) = 0$. By Lemma 17.2, $H_T^0(G/H, M^H) = 0$, so any $x \in M^G$ has the form $\mathrm{Norm}_{G/H}(y)$ for some $y \in M^H$. Since $H_T^0(H, M) = 0$, $y = \mathrm{Norm}_H(z)$ for some $z \in M$, and so $x = \mathrm{Norm}_G(z)$. Hence $H_T^0(G, M) = 0$.

Now make the exact sequence

$$0 \to N \to \mathbb{Z}[G] \otimes_{\mathbb{Z}} M_{\mathbb{Z}} \to M \to 0$$

in which $M_{\mathbb{Z}}$ is the underlying $\mathbb{Z}$-module of M with the trivial G-action, and the right nontrivial arrow maps $g \otimes x$ to $g(x)$. The middle term has zero cohomology by Lemma 17.1, so we get dimension shifting isomorphisms $H_T^i(G, M) \cong H_T^{i+1}(G, N)$. In particular,

$$H^1(G, N) = H_T^1(G, N) \cong H_T^0(G, M) = 0,$$
$$H^2(G, N) = H_T^2(G, N) \cong H_T^1(G, M) = H^1(G, M) = 0,$$

so N satisfies the same hypotheses as M. So for any index i, if I can deduce that $H_T^i(G, M) = 0$ for all M satisfying the hypothesis, then I can deduce the same for $i + 1$ and $i - 1$. Since I have this for $i = 0$, I get the same for all $i \in \mathbb{Z}$ as desired. □

Lemma 18.2. *Let G be a finite solvable group, and let H be a subgroup of G. Then there are isomorphisms $H_T^i(H, \mathbb{Z}) \cong H_T^{i+1}(H, I_G)$ for all $i \in \mathbb{Z}$. In particular,*

$$H_T^1(H, I_G) \cong H_T^0(H, \mathbb{Z}) = \mathbb{Z}/(\#H)\mathbb{Z}$$
$$H_T^2(H, I_G) \cong H_T^1(H, \mathbb{Z}) = 0.$$

Proof. The isomorphisms $H^i_T(H,\mathbb{Z}) \cong H^{i+1}_T(H, I_G)$ come from the exact sequence $0 \to I_G \to \mathbb{Z}[G] \to \mathbb{Z} \to 0$ and the fact that $\mathbb{Z}[G]$ is a free $\mathbb{Z}[H]$-module (and Lemma 17.1). The computation of $H^0_T(H,\mathbb{Z})$ is immediate, since the norm map on $\mathbb{Z}$ is multiplication by $\#H$. Also, $H^1_T(H,\mathbb{Z}) = H^1(H,\mathbb{Z}) = \mathrm{Hom}(H,\mathbb{Z})$ because crossed homomorphisms to a trivial H-module are just homomorphisms, but $\mathrm{Hom}(H,\mathbb{Z}) = 0$ because H is finite. □

Theorem 18.3 (Tate)**.** *Let G be a finite solvable group. Let M be a left G-module. Suppose that for each subgroup H of G, $H^1(H,M) = 0$ and $H^2(H,M)$ is cyclic of order $\#H$. Then there are isomorphisms $H^i_T(G,\mathbb{Z}) \to H^{i+2}_T(G,M)$ which are canonical up to the choice of a generator of $H^2(G,M)$.*

Proof. Choose a generator γ of $H^2(G,M)$. For any subgroup H of G, the inclusion map $M^G \to M^H$ and the norm map $\mathrm{Norm}_{G/H} : M^H \to M^G$ extend to maps $H^i(G,M) \to H^i(H,M)$ and $H^i(H,M) \to H^i(G,M)$ (called the *restriction* and *corestriction* maps, respectively) whose composition is multiplication by $[G:H]$ on $H^i(G,M)$. It follows that γ also generates $H^2(H,M)$.

We first set up an exact sequence

$$0 \to M \to M[\phi] \to I_G \to 0$$

of left G-modules in which the map $H^2(H,M) \to H^2(H,M[\phi])$ is set up to be zero for any subgroup H of G. This is most easily expressed by describing $H^i(G,M)$ in terms of *cocycles* and *coboundaries*. Let $C^r(G,M)$ be the set of maps $G^r \to M$ (with no structural restrictions). We define a differential $d^r : C^r(G,M) \to C^{r+1}(G,M)$ by setting

$$(d^r\phi)(g_1,\ldots,g_{r+1}) = g_1(\phi(g_1,\ldots,g_r)) + \sum_{j=1}^{r}(-1)^j\phi(g_1,\ldots,g_jg_{j+1},\ldots,g_{r+1}) + (-1)^{r+1}\phi(g_1,\ldots,g_r).$$

We may then identify $H^r(G,M)$ with the r-th homology of this complex.

Choose $\phi \in C^2(G,M)$ representing γ. Let $M[\phi]$ be the direct sum of ϕ with the free abelian group on symbols x_g for $g \in G - \{e\}$, and extend the action of G by setting

$$g(x_h) = x_{gh} - x_g + \phi(g,h).$$

One checks using the cocycle condition that this gives a well-defined action. Note that by construction, ϕ is the differential of the cochain $g \mapsto x_g$. We thus have the desired sequence

$$0 \to M \to M[\phi] \to I_G \to 0$$

in which γ is killed by the map $H^2(H,M) \to H^2(H,M[\phi])$. Given this, by Lemma 18.2 we have an exact sequence

$$0 = H^1(H,M) \to H^1(H,M[\phi]) \to H^1(H,I_G) \to H^2(H,M) \to H^2(H,M[\phi]) \to H^2(H,I_G) = 0.$$

By construction, the map $H^2(H,M) \to H^2(H,M[\phi])$ is the zero map, so $H^2(H,M[\phi]) = 0$. Also, $H^1(H,I_G) \to H^2(H,M)$ is surjective and both sides have order $\#H$ by Lemma 18.2,

so the map is also injective. Hence $H^1(H, M[\phi]) \to H^1(H, I_G)$ is the zero map, forcing $H^1(H, M[\phi]) = 0$.

By Lemma 18.1, $M[\phi]$ has trivial Tate cohomology groups. We thus get dimension shifting isomorphisms $H_T^{i+1}(G, I_G) \to H_T^{i+2}(G, M)$, which when combined with the isomorphisms $H_T^i(G, \mathbb{Z}) \to H_T^{i+1}(G, I_G)$ from Lemma 18.2 yield the claim. □

19 Cohomology of local fields

We now make the calculations to apply Tate's theorem to produce the local reciprocity map.

Lemma 19.1. *For any finite Galois extension L of K, for $G = \mathrm{Gal}(L/K)$, $H^1(G, L^\times) = 0$.*

Proof. This is true for any field K by Hilbert's Theorem 90. (More exactly, Hilbert proved this in the cyclic case; the general case is due to Speiser.) □

For G cyclic and M a G-module, the *Herbrand quotient* of M is the ratio $h(M) = \#H_T^0(G, M)/\#H_T^{-1}(G, M)$ assuming that this is finite.

Lemma 19.2. *Let G be a cyclic group. Let $0 \to M \to N \to P \to 0$ be a short exact sequence of G-modules. Then $h(M)h(P) = h(N)$.*

Proof. The long exact sequence folds into a hexagon by Lemma 17.2, from which the claim is clear. □

Lemma 19.3. *Let G be a cyclic group. For any finite G-module M, $h(M) = 1$.*

Proof. Let g be a generator of G. Then the sequences

$$0 \to M^G \to M \to M \to M_G \to 0$$

$$0 \to H_T^{-1}(G, M) \to M_G \stackrel{\mathrm{Norm}_G}{\to} M^G \to H_T^0(G, M) \to 0$$

are exact if we define the map $M \to M$ to take x to $g(x) - x$. Consequently, $\#M^G = \#M_G$ and hence $\#H_T^{-1}(G, M) = \#H_T^0(G, M)$, proving the claim. □

Lemma 19.4. *The group $H^2(G, L^\times)$ has order at most $[L : K]$.*

Proof. By Lemma 19.1 plus Lemma 17.4, for M a Galois subextension with $H = \mathrm{Gal}(L/M)$, the sequence

$$0 \to H^2(G/H, M^\times) \to H^2(G, L^\times) \to H^2(H, L^\times)$$

is exact. Consequently, if we check the claim for L/M and M/K, it then follows for L/K. Since G is solvable, we may reduce to the case of G cyclic. In this case, we may replace $H^2(G, L^\times) = H_T^2(G, L^\times)$ with $H_T^0(G, L^\times)$ by Lemma 17.2. In fact, we need only compute that $h(L^\times) = [L : K]$. By Lemma 19.2, $h(L^\times) = h(\mathbb{Z})h(U_L)$. We already computed $h(\mathbb{Z}) = \#G = [L : K]$. It is easy to check that U_L admits an open subgroup of finite index V which is a free $\mathbb{Z}[G]$-module: namely, by taking logarithms, we reduce to the corresponding question for $\mathfrak{o}_L$, which we essentially did already in the section on norm calculations. Since U_L/V is finite, $h(U_L/V) = 1$ by Lemma 19.4 and so $h(U_L) = h(V) = 1$. This proves the claim. □

Lemma 19.5. *If L/K is unramified, then $H^2(G, L^\times)$ is cyclic of order $[L : K]$.*

Proof. By Lemma 17.2, we may instead compute $H^0_T(G, L^\times) = K^\times / \operatorname{Norm}_{L/K}(L^\times)$. But we already checked that the valuation defines a bijection between this group and $\mathbb{Z}/n\mathbb{Z}$ for $n = [L : K]$. □

Lemma 19.6. *The group $H^2(G, L^\times)$ is cyclic of order $[L : K]$.*

Proof. Put $n = [L : K]$. Let's write $H^2(L/K)$ as shorthand for $H^2(\mathrm{Gal}(L/K), L^\times)$. Let M be the unramified extension of K of degree n. We have a diagram

$$\begin{array}{ccccccc} & & & & H^2(M/K) & \longrightarrow & H^2(ML/L) \\ & & & & \downarrow & & \downarrow \\ 0 & \longrightarrow & H^2(L/K) & \longrightarrow & H^2(ML/K) & \longrightarrow & H^2(ML/L) \end{array}$$

in which the bottom row is exact and the vertical arrows are injective (by inflation-restriction). It's enough to show that the top horizontal arrow is zero; then we can push a generator of $H^2(M/K)$ down to $H^2(ML/K)$, where it will arise from an element of $H^2(L/K)$ of order n.

Since unramified extensions are cyclic, we may rewrite the map $H^2(M/K) \to H^2(ML/L)$ as $H^0_T(M/K) \to H^0_T(ML/L)$ by Lemma 17.2. This is the natural map $K^\times / \operatorname{Norm}_{M/K}(M^\times) \to L^\times / \operatorname{Norm}_{ML/L}(L^\times)$. The former is a cyclic group of order $n = e(L/K)f(L/K)$ generated by a uniformizer π_K of K, while the latter is a cyclic group of order $e(L/K)$ generated by a uniformizer π_L of L. Since π_K has valuation e in L, its image is zero, proving the claim. □

For L/K finite Galois, we now plug into Tate's theorem to get the map $K^\times / \operatorname{Norm}_{L/K}(L^\times) \to \mathrm{Gal}(L/K)^{\mathrm{ab}}$.