Higher Ramification

June Terzioğlu

May 16, 2025

Contents

1 Abstract

This paper will be a sort of survey of various topics in the theory of local fields related to higher ramification: first I will cover some of the background theory required for the subject (much of which is elaboration on a treatise of Serre); then I will introduce the lower and upper numberings and how to translate between them (and why we care about them); then I will discuss an easier way to find this translation in nice cases via a lemma of Tate (which makes use of the so-called Newton copolygon); and eventually I will give a partial exposition of a major application of higher ramification groups to understanding certain extensions of local fields - namely a theorem of Deligne which states that the category of at-most-s-upper-ramified extensions of a local field K is determined (up to equivalence) by $\mathcal{O}_K/\mathfrak{m}_K^s$. In other words, even if you only have knowledge

of the field up to (modulo) a power of the maximal ideal, you can still recover knowledge about its extensions that are not too 'wildly' ramified (very roughly speaking).

2 Preliminaries

Much of this section is an elaboration on [4].

2.1 Valued fields

Definition 1. *A* ***valuation*** *on a field K is a map $v_K : K \to \mathbb{R} \cup \{\infty\}$ such that*

- $v_K(0) = \infty$
- $v_K|_{K^\times}$ *is a group homomorphism $K^\times \to (\mathbb{R}, +)$*
- $v_K(x + y) \geq \min(v_K(x), v_K(y))$ *for all $x, y \in K$ (where we say $\min(\infty, a) = a$ for all $a \in \mathbb{R} \cup \{\infty\}$)*

A field K equipped with a valuation v_K is called a ***valued field****, and we sometimes instead say this as that (K, v_K) is a valued field.*

One can define these in greater generality by having v_K instead map into a totally ordered abelian group Γ, but I will not need this. Really what I have defined above is called by some as a *rank one valuation*. Typically I will drop the subscript K when there is no ambiguity.

Here are a few easy lemmas about valuations; the last will be fundamental in our discussion later about Newton copolygons / valuation functions:

Theorem 1

Let (K, v) be a valued field.

- $v(1) = v(-1) = 0$ (in particular, $v(-x) = v(x)$).
- If $x \in K^\times$, then $v(x^{-1}) = -v(x)$.
- If $x_1, \ldots, x_n \in K$ with $v(x_j) < v(x_i)$ for all $i \neq j$, then $v\left(\sum_{i=1}^{n} x_i\right) = v(x_j)$ (i.e. even though $v\left(\sum_{i=1}^{n} x_i\right) \geq \min_{1 \leq i \leq n} \{v(x_i)\}$ in general, equality holds when the minimum is unique).

Proof: The first is because $v(1^2) = v(1) + v(1)$ and $0 = v(1) = v((-1)^2) = v(-1) + v(-1)$; the second is because $v(1) = v(xx^{-1}) = v(x) + v(x^{-1})$. For the third, let's assume $j = 1$ by relabelling. Write $r \triangleq v(x_1)$, so that

$v(x_i) > r$ and so $v(-x_i) > r$ for all $i > 1$. Then if $v\left(\sum_{i=1}^{n} x_i\right) > r$ as well, we would have

$$v(x_1) = v\left(\sum_{i=1}^{n} x_i + \sum_{i=2}^{n}(-x_i)\right) \geq \min\left(v\left(\sum_{i=1}^{n} x_i\right), \min_{2\leq i\leq n}\{v(-x_i)\}\right) > r$$

which is a contradiction.

□

Definition 2. *A* ***discrete valuation*** *is simply a valuation whose image is $r\mathbb{Z} \cup \{\infty\}$, for some $r \in \mathbb{R}_{>0}$ (which, by post-composing with a scaling, we will always assume is 1 unless otherwise stated); a field K equipped with a discrete valuation v is called a* ***discretely valued field*** *(sometimes written as a pair (K, v)).*

A common situation where we *don't* assume the image of a discrete valuation is $\mathbb{Z}$ is when we have a valuation on a field 'normalized' for a smaller subfield - this will be explained in section 2.4.

A concept closely related to valuations is absolute values:

Definition 3. *An* ***absolute value*** *on a field K is a map $|\cdot|_K : K \to \mathbb{R}_{\geq 0}$ such that*

- $|0|_K = 0$
- $|\cdot|_K|_{K^\times}$ *is a homomorphism* $K^\times \to (\mathbb{R}_{>0}, \cdot)$
- $|x + y|_K \leq |x|_K + |y|_K$ *(triangle inequality)*

If $|\cdot|_K$ satisfies the ultrametric inequality $|x + y|_K \leq \max(|x|_K, |y|_K)$ and not just the weaker triangle inequality, we call it non-archimedean.

There is a bijection between non-archimedean absolute values and valuations on a field K; simply fix some $c \in (0, 1)$ and map a valuation v to c^{-v}. For this reason we can essentially think of them as the same thing - topologically, it makes little difference which c we choose, since c^{-v} is equivalent to d^{-v} for any $c, d \in (0, 1)$ (in the sense that they give K the same topology). [1]

Definition 4. *A* ***discrete valuation ring*** *(abbreviated* ***DVR****) is a local PID that is not a field.*

Equivalently, one can show this is equivalent to being a local Dedekind domain; then in an arbitrary DVR R, nonzero proper ideals factor uniquely into a product of maximal ideals (this is actually one way to *define* Dedekind domains). So since there is only one maximal ideal $\mathfrak{m}$, often written $\mathfrak{m}_R$ (which is nonzero since DVRs are not fields), then the set of proper ideals equals $\{\langle 0\rangle, \mathfrak{m}, \mathfrak{m}^2, \mathfrak{m}^3, \ldots\}$ (where all are distinct by uniqueness of the factorization). R being a PID tells us we can even write this as $\{\langle 0\rangle, \langle\pi\rangle, \langle\pi^2\rangle, \langle\pi^3\rangle, \ldots\}$ for some π generating $\mathfrak{m}$ - this suggests the following definition:

[1] Though when K has finite residue field k, it's preferred to take $c = 1/|k|$.

Definition 5. *A* ***uniformizer*** *for a DVR R is a generator π for its maximal ideal.*

So we often write $\mathfrak{m}_R$ as πR, where π is implicitly a uniformizer. Note the uniformizers are exactly the elements of $\mathfrak{m} - \mathfrak{m}^2$, and that relative to a fixed uniformizer π, any $x \in R^\times$ can be written as $u\pi^n$ where u is a unit, $n \geq 0$, *and both are unique relative to* x.

Another few natural definitions:

Definition 6. *The* ***residue field*** *for a DVR R with maximal ideal $\mathfrak{m}$ is the field $R/\mathfrak{m}$; the* ***residue field*** *for the pair $(\mathrm{frac}(R), R)$ is defined as the same.*

If R is considered fixed and $K = \mathrm{frac}(R)$; we often write the residue field as k; similarly if $L = \mathrm{frac}(R)$, we write the residue field as l.

There is a natural bijective correspondence between discrete valuation subrings of a field K whose fraction field is K and discrete valuations on K. This is because, for a discrete valuation v on K, $\{x \in K : v(x) \geq 0\}$ is a DVR subring with fraction field K (and maximal ideal $\{x \in K : v(x) \geq 1\}$), and inversely such a DVR subring R with maximal ideal $\mathfrak{m}$ gives rise to a valuation

$$v_{\mathfrak{m}}(x) \triangleq \begin{cases} \sup\{i \geq 0 : x \in \mathfrak{m}^i\}, & x \in R \\ -\sup\{i \geq 0 : x^{-1} \in \mathfrak{m}^i\}, & x \notin R \end{cases}.$$

Note this is well-defined since, as noted previously, any element of R can be written as $u\pi^n$ for a uniformizer π, and every $x \in K = \mathrm{frac}(R)$ has either $x \in R$ or $x^{-1} \in R$. Also note that these two processes of going from DVR to discrete valuation and vice versa are inverses.

This correspondence gives a natural translation of definition 5: the uniformizers for a DVR associated with a discrete valuation v are exactly the elements x with $v(x) = 1$, and in general the generators of $\mathfrak{m}^i$ are the elements with valuation i (for $i \geq 0$, and even for $i < 0$ if you define negative powers of $\mathfrak{m}$ appropriately - but then you need to talk about 'fractional ideals' and I will not make this detour now). Another comment about uniformizers: we often say that a uniformizer for a DVR R is also one for its fraction field, and instead of saying π is a uniformizer for R or $\mathrm{frac}(R)$ we sometimes instead say that π *uniformizes* R or $\mathrm{frac}(R)$.

Definition 7. *Given a discretely valued field (K, v), the DVR subring associated with v is called the* ***ring of integers*** *of K and denoted $\mathcal{O}_K$ (or perhaps $\mathcal{O}_{K,v}$ when the valuation needs to be made clear). The maximal ideal of $\mathcal{O}_K$ is often written $\mathfrak{m}_K$.*

A discrete valuation v on a field K gives it a natural topology under which the field operations (addition, negation, multiplication and inversion) become continuous; such a field is called a *topological field*. The topology induced by the valuation can be described by simply giving a neighborhood basis of 0 and declaring that its additive translates are also open, as is doable for any topological group. So we declare that $\{x \in K : v(x) > n\}$ is open for all $n \in \mathbb{Z}$.

There is a certain condition on the topology of a discretely valued field that proves to be very useful:

Theorem 2

A discretely valued field K is locally compact (w.r.t. its valuation topology) iff it is complete and its residue field is finite.

Proof: If K is locally compact, then it has a compact subset C containing some open neighborhood of 0; since the $\mathfrak{m}^i$ (letting $\mathfrak{m} \triangleq \mathfrak{m}_K$) form a neighborhood basis for 0, this means it contains some $\mathfrak{m}^k$. But the $\mathfrak{m}^k$ are clopen in K, so $\mathfrak{m}^k$ is also closed in C and so compact. Then $\pi^{-k}\mathfrak{m}^k = \mathcal{O}_K$ (with π a uniformizer of K) is compact. In particular, since the cosets of $\mathfrak{m}$ cover $\mathcal{O}_K$, and all are homeomorphic (and so open since $\mathfrak{m}$ is), then they must admit a finite cover, so that there are only finitely many cosets of $\mathfrak{m}$ in $\mathcal{O}_K$ (meaning $k = \mathcal{O}_K/\mathfrak{m}$ is finite).

Additionally, since $\mathcal{O}_K$ is compact, then it is complete. So given a cauchy sequence $(x_i)_1^\infty$ in K, $(v_K(x_i))_1^\infty$ must stabilize since v_K is continuous with discrete codomain, so there is $N, r \in \mathbb{Z}$ with $v_K(x_i) = r$ when $i \geq N$. But then $(\pi^{-r}x_i)_1^\infty$ is cauchy and eventually in $\mathcal{O}_K$, so it converges to some $\alpha \in \mathcal{O}_K$, meaning x_i converges to $\pi^r\alpha$. So K is complete too.

Conversely, if K is complete and $\mathcal{O}_K/\mathfrak{m}$ is finite, then $\mathcal{O}_K/\mathfrak{m}^i$ is finite too for any $i \geq 1$ (since multiplication by π^i gives a isomorphism of $\mathcal{O}_K/\mathfrak{m}$ with $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ as additive groups for all $i \geq 1$, and we can multiply indexes). Then since $\mathcal{O}_{\widehat{K}} \cong \varprojlim \mathcal{O}_K/\mathfrak{m}^i$ (see section 2.3), it is profinite and so compact; but $\mathcal{O}_{\widehat{K}} \cong \mathcal{O}_K$ by completeness, so $\mathcal{O}_K$ is compact. And since $\mathcal{O}_K$ contains the open neighborhood $\mathcal{O}_K$ of 0, K is locally compact at 0, and so at every point (since we can additively translate $\mathcal{O}_K$ as needed).

□

Definition 8. *Any discretely valued field satisfying the hypotheses of theorem 2 (or equivalently the conclusion) is called a **local field**.*

Nowadays it seems like people like to use a slightly more general definition, namely that a local field should instead be a complete discretely valued field with *perfect* residue field; in fact, this is the definition we will use in the Deligne section.

Complete discretely valued fields are quite nice in general, and they also allow their valuation to be 'extended' in a nice way to finite extensions, so it would be profitable to have a way to 'complete' a discretely valued field …but in order to explore this, we need to first define some basic concepts related to ramification.

2.2 Ramification basics

If A is a Dedekind domain (remember that this means nonzero proper A-ideals factor uniquely as a product of prime ideals) and L is a finite extension of $K \triangleq \operatorname{frac}(A)$, then the integral closure $B \triangleq \overline{A}^L$ of A in L is also a Dedekind domain - let's keep this setup throughout this section. Given a nonzero prime A-ideal $\mathfrak{p}$, we can uniquely factorize $\mathfrak{p}B$ as a product of nonzero prime B-ideals. If $\mathfrak{p}B = \prod_1^r \mathfrak{P}_i^{e_i}$, where $e_i \geq 1$ for all i (and the

$\mathfrak{P}_i$ are all distinct), then the $\mathfrak{P}_i$ are exactly the prime B-ideals lying above $\mathfrak{p}$ (i.e. whose intersection with A is $\mathfrak{p}$). Then we can make the following definitions:

Definition 9. *Let $\mathfrak{p}B = \prod_1^r \mathfrak{P}_i^{e_i}$ as above.*

- *The **ramification index** e_i is defined as the e_i in the above equation.*
- *The **residue degree** f_i is defined as $[B/\mathfrak{P}_i : A/\mathfrak{p}]$* [2]

If $\mathfrak{P} = \mathfrak{P}_i$, we also write e_i as $e_{\mathfrak{P}/\mathfrak{p}}$ and f_i as $f_{\mathfrak{P}/\mathfrak{p}}$.

- *If $e_i > 1$ for some i, we say $\mathfrak{p}$ **ramifies** in L.*
- *If $e_i = 1$ and $f_i = 1$ for all i, we say $\mathfrak{p}$ **splits completely** in L.*
- *If $e_i = 1$ for all i and $r = 1$, we say $\mathfrak{p}$ **remains prime** in L.*

Theorem 3

If L/K is also galois, then e_i, f_i are independent of i (in this case, we usually omit the subscript and just write e, f instead).

Proof: This is essentially because the action $\mathrm{gal}(L/K) \curvearrowright \{\mathfrak{P}_1, \dots, \mathfrak{P}_r\}$ via element-wise images (which is a well-defined action since any $\sigma \in \mathrm{gal}(L/K)$ fixes B (it preserves polynomials over A), fixes A pointwise, and actually maps a prime ideal lying above $\mathfrak{p}$ to a prime ideal lying above $\mathfrak{p}$) is transitive, as given $j \neq k$ and $x \in \mathfrak{P}_j$, we have

$$N_K^L(x) = x \cdot \prod_{\mathbb{1}_L \neq \sigma \in \mathrm{gal}(L/K)} \sigma(x) \in A \cap \mathfrak{P}_j = \mathfrak{p} \subseteq \mathfrak{P}_k$$

so that $\sigma(x) \in \mathfrak{P}_k$ for some σ by primality of $\mathfrak{P}_k$.

□

Theorem 4

$\sum_{i=1}^r e_{\mathfrak{P}_i/\mathfrak{p}} f_{\mathfrak{P}_i/\mathfrak{p}} = [L : K]$; in particular, if L/K is galois then $efr = [L : K]$.

Proof: I will omit the details, but the idea is to show $[B/\mathfrak{p}B : A/\mathfrak{p}] = [L : K]$ and then use that $B/\mathfrak{p}B \cong \prod_1^r (B/\mathfrak{P}_i)^{e_i}$ (via the Chinese Remainder Theorem).

Definition 10 (Ramification definitions for extensions of local fields)**.** *We will see in section 2.4 that in the case K and L are complete and discretely valued (for example, if they are local fields), $\mathcal{O}_K$ and $\mathcal{O}_L$ have only one nonzero prime ideal, so $\mathfrak{m}_K \mathcal{O}_L$ factors as $\mathfrak{m}_L^{e_K^L}$ for some e_K^L. In particular, the r in definition 9 always equals* 1.

[2]Note this embedding makes sense since $\mathfrak{P}_i = \mathfrak{p} \cap A$ (so that $\mathfrak{p}$ is actually the kernel of the canonical map $A \hookrightarrow B \twoheadrightarrow B/\mathfrak{P}_i$)

- e^L_K *is called the ramification index of the* extension L/K.
- *Similarly,* $f^L_K \triangleq [\mathcal{O}_L/\mathfrak{m}_L : \mathcal{O}_K/\mathfrak{m}_K] = [l : k]$ *is called the residue degree of the* extension L/K.
- *If* $e^L_K = 1$, *the* extension L/K *is called* ***unramified***.
- *If* $f^L_K = 1$ *(or equivalently* $e^L_K = [L : K]$*, since* $e^L_K f^L_K = [L : K]$*), the* extension L/K *is called* ***totally ramified***.

I'll alternatively write these as $e_{L/K}$, $f_{L/K}$ respectively, depending on which one looks nicer in a given context. In the case it is ambiguous which valuation on our fields we are considering, I may even write them as $e_{w/v}$ and $f_{w/v}$ (where v, w are the valuations being considered on the lower / higher fields respectively).

Now we can get back to completing and extending discretely valued fields!

2.3 Completion

With the familiar construction using cauchy sequences, one can topologically complete a discretely valued field K (with respect to its valuation topology); further, one can check that the result $\widehat{K}$ has the structure of a topological field.given by applying the operations of K element-wise to sequences. Additionally, one can check that this completion topology is induced by an extension of v: define $\overline{v}((x_i)_1^\infty) \triangleq \lim_{n\to\infty} v(x_n)$ (which is well-defined because v is continuous, if we give $\mathbb{Z} \cup \{\infty\}$ the order topology); then this is a discrete valuation whose restriction to the copy of K contained in $\widehat{K}$ (via the map $x \mapsto (x)_1^\infty$) equals v, since $\widehat{v}((x)_1^\infty) = \lim_{n\to\infty} v(x) = v(x)$ for all constant sequences $(x)_1^\infty$.

The ring of integers of $\widehat{K}$ is often written $\widehat{\mathcal{O}_K}$ instead of $\mathcal{O}_{\widehat{K}}$, for reasons made clear in (1) of the following theorem:

Theorem 5

(1) $\mathcal{O}_{\widehat{K}} = \overline{\mathcal{O}_K} = \varprojlim \widehat{\mathcal{O}_K}/\pi^n\widehat{\mathcal{O}_K}$ (i.e. the DVR associated with $\widehat{v}$ is the closure in $\widehat{K}$ of the one associated with v)

(2) any uniformizer for $\mathcal{O}_K$ is also one for $\widehat{\mathcal{O}_K}$ (i.e. if $v(x) = 1$, then $\widehat{v}(x) = 1$) - this also means that if $\mathfrak{m}_K = \pi\mathcal{O}_K$, then $\mathfrak{m}_{\widehat{K}} = \pi\widehat{\mathcal{O}_K}$, and that $\pi\widehat{\mathcal{O}_K} \cap K = \pi\mathcal{O}_K$

(3) $\mathcal{O}_{\widehat{K}}/\pi\mathcal{O}_{\widehat{K}} \cong \mathcal{O}_K/\pi\mathcal{O}_K$

Proof: (1) is because any uniformizer π for $\mathcal{O}_K$ also uniformizes $\widehat{K}$, so that the $\pi^i\widehat{\mathcal{O}_K}$ form a neighborhood basis for $\widehat{\mathcal{O}_K}$ at 0, as we've seen in general; (3) is because the composition $\mathcal{O}_K \hookrightarrow \widehat{\mathcal{O}_K} \twoheadrightarrow \widehat{\mathcal{O}_K}/\pi\widehat{\mathcal{O}_K}$ has kernel $\pi\mathcal{O}_K$ (since for $x \in \mathcal{O}_K$ we have $x \in \pi\widehat{\mathcal{O}_K} \Leftrightarrow \widehat{v}(x) = 1 \Leftrightarrow v(x) = 1 \Leftrightarrow x \in \pi\mathcal{O}_K$) and is surjective: if $(x_i)_1^\infty \in \widehat{\mathcal{O}_K}$, then there is N so that $i, j \geq N \implies v(x_i - x_j) \geq 1$, meaning the terms of $(x_i)_1^\infty - (x_N)_{i=1}^\infty$ eventually have valuation at least 1 and so $(x_N)_{i=1}^\infty + \pi\widehat{\mathcal{O}_K} = (x_i)_1^\infty + \pi\widehat{\mathcal{O}_K}$ (i.e. x_N maps to $(x_i)_1^\infty + \pi\widehat{\mathcal{O}_K}$).

□

2.4 Extension

Given L/K, we say that v_L *extends* v_K iff $v_L|_K$ differs from v_K multiplicatively by a constant (which - assuming as usual that our valuations are discrete and have image $\mathbb{Z} \cup \{\infty\}$ - must necessarily be e_K^L, since if π uniformizes K then $\pi\mathcal{O}_L = \mathfrak{m}_L^{e_K^L}$ and so $v_L(\pi) = e_K^L$).

Theorem 6

If K is a complete discretely valued field, and L is a finite extension of K, then:

(1) $\mathcal{O}_L \triangleq \overline{\mathcal{O}_K}^L$ (denoting the integral closure of $\mathcal{O}_K$ in L) is a DVR and a free $\mathcal{O}_K$-module of rank $[L:K]$

(2) the valuation v_L that $\mathcal{O}_L$ induces on L makes L complete

(3) the valuation v_L is the unique one on L extending v_K

Proof: I will omit it, but the main idea is to use 'dévissage' to break into the separable / purely inseparable cases and combine appropriately.

We again get $e_{L/K} f_{L/K} = [L:K]$ in this case (we can't directly apply the discussion in section 2.2 to get this, since there we assumed our extension to be separable, but it turns out that $\mathcal{O}_L$ being a finitely generated $\mathcal{O}_K$-module suffices). We can additionally get an explicit form of the valuation on L:

Theorem 7

If K is a complete discretely valued field, and L is a finite extension of K, then $v_L(x) = v_K(N_K^L(x))/f_{L/K}$.

Proof: Let $M \triangleq \mathrm{spl}(L/K)$ be the normal closure of L/K; it is a finite extension of both L and K, so v_M extends both v_K and v_L uniquely. Since $v_M \circ \sigma$ is also a valuation on M extending v_K for any $\sigma \in \mathrm{gal}(M/K)$, it must equal v_M, and so since any conjugate of $x \in L$ can be written $\sigma(x)$ for some $\sigma \in \mathrm{gal}(M/K)$ we have $v_M(x) = v_M(\sigma(x))$.

Then

$$
\begin{aligned}
[M:L]\left(v_K(N_K^L(x))\right) &= v_K\left(N_K^L(x)^{[M:L]}\right) \\
&= v_K\left(\prod_{\sigma\in\text{gal}(M/K)} \sigma(x)\right) \\
&= e_{M/K}^{-1} v_M\left(\prod_{\sigma\in\text{gal}(M/K)} \sigma(x)\right) \\
&= e_{M/K}^{-1} \sum_{\sigma\in\text{gal}(M/K)} v_M(\sigma(x)) \\
&= e_{M/K}^{-1} \sum_{\sigma\in\text{gal}(M/K)} v_M(x) \\
&= \frac{|\operatorname{gal}(M/K)|}{e_{M/K}} v_M(x) \\
&= \frac{|\operatorname{gal}(M/K)|}{e_{M/K}} e_{M/L} v_L(x) \\
&= f_{M/K} v_M(x) \\
&= f_{M/K} e_{M/L} v_L(x) \\
&= [M:L] v_L(x)
\end{aligned}
$$

at which point the result follows.

□

Here is an important comment about normalization of valuations in extensions: given L/K finite with K complete and discretely valued, the valuations v_K and v_L are related via $v_K(x) = e_K^L v_L(x)$ for all $x \in K$; the result is that v_L may not actually equal v_K on K, though they differ multiplicatively by a constant. Sometimes we wish to circumvent this by considering the valuation $v \triangleq v_L/e_K^L$; this is a (still discrete) valuation on all of L such that $v(\pi) = 1$ when π uniformizes K. In this case we say v is the valuation of L normalized for K. In contrast, if we take our valuation instead as just $v \triangleq v_L$, this is a valuation on all of L such that $v(\pi) = 1$ when π uniformizes L; in this case we say v is the valuation of L normalized for L.

We can really just do the same thing for arbitrary subextensions E between K and L; we say v is the valuation on L normalized for E iff $v = v_L/e_E^L$ (i.e. iff v differs from v_L multiplicatively by a constant, and $v(\pi) = 1$ when π normalizes E).

One consequence of all this is that we can consider a valuation on the algebraic closure $\overline{K}$, gotten by essentially taking the union over all valuations of finite extensions of K that are normalized for K (so in particular this valuation on $\overline{K}$ is normalized for K - note it may no longer be discrete, however).

2.5 Completion and extension: all together now!

Theorem 8

Let K be a discretely valued field with valuation v, L a finite extension, and $w_1, \ldots, w_r$ all the extensions of v (note these will be the discrete valuations induced by the $\mathfrak{P}_i$ lying above $\mathfrak{m}_K$). Let $\widehat{L}_i$ be the completion of L w.r.t. w_i.

(1) $e_{\widehat{L_i}/\widehat{K}} = e_{w_i/v}$, $f_{\widehat{L_i}/\widehat{K}} = f_{w_i/v}$

(2) $[\widehat{L_i} : \widehat{K}] = [L : K]$

Proof: (1) follows from (3) of theorem 5 and the fact that if π uniformizes K, then $e_{w_i/v} = w_i(\pi) = \widehat{w_i}((\pi)_1^\infty) = e_{\widehat{L_i}/\widehat{K}}$; (2) is an immediate consequence of (1) and the fact that the ramification index and the residue degree multiply to the field extension degree.

□

One can show that if L/K is separable (resp. galois), then every $\widehat{L_i}/\widehat{K}$ is as well, but we will not need this (since we will assume all our fields are complete from the get-go).

2.6 Generators for extensions

Here I'll go on a brief detour to showcase two very useful facts relating to how a lower ring of integers 'downstairs' can generate a higher one 'upstairs' (which will be used a lot later):

Theorem 9

If L/K is a finite separable extension of discretely valued fields, with separable residue field extension, then $\mathcal{O}_L = \mathcal{O}_K[\alpha]$ for some $\alpha \in \mathcal{O}_L$.

Proof: Since L/K is separable and $\mathcal{O}_K$ is a PID, we already know (i.e. it can be shown) that $\mathcal{O}_L$ is a free $\mathcal{O}_K$-module of rank $[L : K] = ef$ (letting $e \triangleq e_K^L$ and $f \triangleq f_K^L$). Take x so that $x + \mathfrak{m}_L$ generates the residue field extension l/k (note I am using separability of l/k here!) and let π be any uniformizer of L; then to show $\{x^i\pi^j : 0 \leq i < f, 0 \leq j < e\}$ is a basis for $\mathcal{O}_L$ over $\mathcal{O}_K$, it suffices to show it spans $\mathcal{O}_L$ (since it has $e_K^L f_K^L$ elements).

By Nakayama's lemma (more specifically, by its corollary Proposition 2.8 in [3]), if the images of the $x^i\pi^j$ in $\mathcal{O}_L/\mathfrak{m}_K\mathcal{O}_L = \mathcal{O}_L/\pi^e\mathcal{O}_L$ generate it as an $\mathcal{O}_K/\mathfrak{m}_K$-vector space, then the $x^i\pi^j$ generate $\mathcal{O}_L$ as an $\mathcal{O}_K$-module. So it will suffice to show this. Since the $x^i + \pi\mathcal{O}_L$ generate l/k, we know that for any $y \in \mathcal{O}_L$, we can find $a_{i,0} \in A$ $(0 \leq i < f - 1)$ so that $\sum_{i=0}^{f-1} a_{i,0}x^i - y \in \pi\mathcal{O}_L$. This will serve as our 'base case'.

Inductively, let's say that we have $a_{i,j}$ $(0 \le i < f, 0 \le j < m)$ so that

$$\sum_{i=0}^{f-1}\sum_{j=0}^{m-1} a_{i,j}x^i\pi^j - y \in \pi^m\mathcal{O}_L.$$

Then we can write $\sum_{i=0}^{f-1}\sum_{j=0}^{m-1} a_{i,j}x^i\pi^j - y = b\pi^m$ for some $b \in \mathcal{O}_L$. By our base case, we know $b + \sum_{i=0}^{f-1} a_{i,m}x^i \in \pi\mathcal{O}_L$ for some $a_{i,m} \in A$ $(0 \le i < f-1)$, where I am flipping some signs for later convenience. This gets us $b\pi^m + \sum_{i=0}^{f-1} a_{i,m}\pi^m \in \pi^{m+1}\mathcal{O}_L$, and by substituting for $b\pi^m$ this becomes

$$\sum_{i=0}^{f-1}\sum_{j=0}^{m} a_{i,j}x^i\pi^j - y \in \pi^{m+1}\mathcal{O}_L.$$

Then in this way we can get $a_{i,j}$ $(0 \le i < f, 0 \le j < e)$ so that $\sum_{i=0}^{f-1}\sum_{j=0}^{e-1} a_{i,j}x^i\pi^j - y \in \pi^e\mathcal{O}_L$, implying that the $x^i\pi^j$ generate $\mathcal{O}_L/\pi^e\mathcal{O}_L$ over $\mathcal{O}_K/\mathfrak{m}_K$!

Now let's take $g(T)$ whose reduction $\overline{g}$ modulo $\mathfrak{m}_L$ equals $\operatorname{irr}(x + \mathfrak{m}_L, k)$; since $\overline{g}(x + \mathfrak{m}_L) = 0$, then $g(x) \in \mathfrak{m}_L$. If $v_L(g(x)) > 1$, then we can write $g(x + \varpi) = g(x) + \varpi g'(x) + \varpi^2 b$ for some $b \in \mathcal{O}_L$ (where ϖ is any uniformizer for L), just by expanding out; since l/k is separable, then $\overline{g}(x + \mathfrak{m}_L) \neq 0$, and so $g'(x)$ is a unit in $\mathcal{O}_L$. Then $v_L(\varpi g'(x)) = 1$, so since $v_L(g(x)) > 1$ and $v_L(\varpi^2 b) \ge 2$, then by theorem 1 we have $v_L(g(x + \varpi)) = 1$.

So by replacing x with $x + \varpi$, we can assume $g(x)$ uniformizes L, so that applying the earlier result we get that $\{x^i g(x)^j : 0 \le i < f, 0 \le j < e\}$ is a basis for $\mathcal{O}_L$ over $\mathcal{O}_K$. Note all of these are $\mathcal{O}_K$-linear combinations of nonnegative powers of x. By going through and subtracting out each one by a linear combination of the ones with smaller degree, we get that $\{x_i : 0 \le i < ef\}$ is a basis for $\mathcal{O}_L$ over $\mathcal{O}_K$, so that $\mathcal{O}_L = \mathcal{O}_K[x]$. So take $\alpha = x$.

□

It is sometimes useful to be able to guarantee α is a uniformizer (this will provide a nice alternative characterization of the higher ramification groups in 3.2); we can get this with slightly different hypotheses (remember that an *Eisenstein polynomial* over K is one where all coefficients but the leading one lie in $\mathfrak{m}_K$ and the constant coefficient does not lie in $\mathfrak{m}_K^2$):

Theorem 10

If L/K is a finite totally ramified extension of discretely valued fields, then $\mathcal{O}_L = \mathcal{O}_K[\pi]$ for any uniformizer $\pi \in \mathcal{O}_L$ (which also has an Eisenstein irreducible polynomial over K).

Proof: Write $f(T) \triangleq \operatorname{irr}(\pi, K)(T)$ as $\sum_0^n a_i T^i$. Let our valuation v on L be normalized for L (so that $v(\pi) = 1$ and the valuation of anything in K is necessarily a multiple of $e_K^L = [L : K] = n$ - note we are using the totally ramified hypothesis here).

Then since $\sum_0^n a_i\pi^i = 0$, which has valuation ∞, there cannot be any j with $v(a_j\pi^j) < v(a_i\pi^i)$ for all $i \neq j$ (otherwise by theorem 1 we would have $\infty = v(0) = v\left(\sum_0^n a_i\pi^i\right) = v(a_j\pi^j) = v(a_j) + j$, implying $a_j = 0$, further implying $a_i = 0$ for all i as well (since $v(a_j\pi^j)$ was supposed to be the minimum over the valuations of all the terms) and contradicting that $f(T) \neq 0$).

Then if $v(a_0) = 0$, we would have $v(a_0\pi^0) < v(a_i\pi^i) = v(a_i) + i$ for all $i > 0$, contradicting the above, so $v(a_0) > 0$ (i.e. $v(a_0) \geq n$) necessarily. Then inductively - for $j < n$ - if $v(a_j) = 0$ and $v(a_i) > 0$ (i.e. $v(a_i) \geq n$) for all $i < j$, then

$$v(a_j\pi^j) = j < k \leq v(a_k\pi^k)$$

for all $k > j$ and also

$$v(a_j\pi^j) = j < n \leq v(a_k) \leq v(a_k\pi^k)$$

for all $k < j$. So $v(a_i) \geq n$ for all $i < n$. $f(T)$ is monic by definition, so $a_n = 1$ and so $v(a_n) = 0$.

If we suppose now that we had $v(a_0) > n$ (i.e. $v(a_0) \geq 2n$), since we also have $v(a_i) \geq n$ for all $i > 0$ and so $v(a_i\pi^i) > n$ for all $i > 0$, $v(a_n\pi^n) = n$ would be the minimum among the terms of $f(T)$, again a contradiction. So $v(a_0) = n$.

Remember that for $x \in K$, $v(x) = kn$ iff $x \in \mathfrak{m}_L^k$. Then $a_0 \in \mathfrak{m}_L - \mathfrak{m}_L^2$ and $a_i \in \mathfrak{m}_L$ for all $0 < i < n$, so $f(T) = \mathrm{irr}(\pi, K)(T)$ is Eisenstein.

Then - by lemma 4 in section 1.6 of [4] - $k[T]/\langle f(T)\rangle$ has exactly one maximal ideal $\langle \mathfrak{m}_K, T\rangle$ (because $f(T)$ factors as just T^n modulo $\mathfrak{m}_K$), so it is local. It is also noetherian as a quotient of a noetherian ring. Additionally, we have $\langle \mathfrak{m}_K, T\rangle = \langle T\rangle$, because $\mathfrak{m}_K$ is generated by a_0 and $a_0 = -\sum_1^n a_iT^i$ - this means that the maximal ideal of $\mathcal{O}_K[T]/\langle f(T)\rangle$ is generated by the non-nilpotent element T (more precisely $T + \langle f(T)\rangle$, but I'll write this as just T), so by proposition 1.2 in [4] it follows that it is a DVR with uniformizer T. If we identify L with $K[T]/\langle f(T)\rangle$, then $\overline{\mathcal{O}_K}^L$ (the integral closure of $\mathcal{O}_K$ in L) equals $\mathcal{O}_K[T]/\langle f(T)\rangle$ and so is a DVR uniformized by T (which corresponds to π under our identification). So $\mathcal{O}_L = \mathcal{O}_K[\pi]$.

□

The above theorem can actually sort of be phrased as an if and only if:

Theorem 11

Let K be a discretely valued field, with v the valuation on $\overline{K}$ normalized for K; if no roots of $f(T) \in \mathcal{O}_K[T]$ have valuation 0, then $f(T)$ is an Eisenstein polynomial iff it is irreducible and the extension generated by any one of its roots is totally ramified, having said root as a uniformizer.

Proof: ($\Leftarrow$) was done in theorem 10, so I'll do ($\Rightarrow$): Let $f(T)$ be Eisenstein with roots $\alpha_1, \ldots, \alpha_n$ and consider the valuation v on the splitting field of $f(T)$ over K, normalized for K. We have that $\prod_{i=1}^n \alpha_i$ is the constant coefficient of $f(T)$ up to sign, and so

$$1 = v\left(\prod_{i=1}^{n} \alpha_i\right) = \sum_{i=1}^{n} v(\alpha_i).$$

Since $v(\alpha_i) > 0$ for all i (they have valuation at least 0 since they are integral over $\mathcal{O}_K$, and they don't have valuation 0 by hypothesis), then

$$v(\alpha_i) \geq \frac{1}{e_K^{K(\alpha_i)}} \geq \frac{1}{n}$$

for all i. For both the above equations to hold, we must have $v(\alpha_i) = 1/n$ for all i, so that $n = e_K^{K(\alpha_i)}$, implying (1) that $n = e_K^{K(\alpha_i)} \leq [K(\alpha_i) : K] \leq n$ i.e. $[K(\alpha_i) : K] = \deg(f)$ and so $f = \operatorname{irr}(\alpha_i, K)$, and (2) that $K(\alpha_i)/K$ is totally ramified, with α_i uniformizing (since it has valuation $1/e_K^{K(\alpha_i)}$).

□

---

3 Higher ramification groups

Much of this section is an elaboration on [4]. Throughout, I will assume we have a finite galois extension L/K of discretely valued fields with separable residue field extension l/k, where L has valuation v_L extending the valuation v_K of K. Remember that for $x \in K$, we have $v_L(x) = v_K(x)/e_K^L$. Additionally, using theorem 9 from the previous section, write $\mathcal{O}_L = \mathcal{O}_K[\alpha]$. Assume from here that this generator α is fixed.

Definition 11.

- *The **inertia group** $G_0(L/K)$ of L/K (written G_0 if the extension is clear) is the subgroup*

  $$\{\sigma \in \operatorname{gal}(L/K) : (\forall x \in \mathcal{O}_L)[\sigma(x) \equiv x \bmod \mathfrak{m}_L]\}$$

  of $\operatorname{gal}(L/K)$; it is sometimes denoted $\mathcal{I}_{L/K}$.

- *The **inertia field** of L/K is L_{G_0}, the fixed field of G_0 under the Galois correspondence; it is sometimes denoted $L_{\mathcal{I}}$.*

$L_{\mathcal{I}}$ is the maximal unramified subextension L of K, meaning if $K \leq E \leq L$ then $\mathfrak{m}_K$ does not ramify in E iff $E \subseteq L_{G_0}$. $L_{\mathcal{I}}/K$ is unramified, while $L/L_{\mathcal{I}}$ is totally ramified.

Definition 12. *The i**-th ramification group** of L/K (for $i \geq 0$) is defined as*

$$G_i(L/K) \triangleq \{\sigma \in \operatorname{gal}(L/K) : (\forall x \in \mathcal{O}_L)[\sigma(x) \equiv x \bmod \mathfrak{m}_L^{i+1}]\}.$$

We also define $G_{-1}(L/K) \triangleq \operatorname{gal}(L/K)$, and as before we simply write G_i if the extension is clear.

The purpose of these higher ramification groups is to essentially track how much elements of the Galois group 'move things'. In particular, if it is known that $G_i(L/K)$ is trivial, then any two K-automorphisms that act the same mod $\mathfrak{m}_L^{i+1}$ must actually be the same.

Note that the 0-th ramification group of L/K is the same as its inertia group. The ramification groups can also be thought about in a slightly different way:

Definition 13. *Given $\sigma \in \operatorname{gal}(L/K)$, let $i_L^K(\sigma) \triangleq v_L(\sigma(\alpha) - \alpha)$.*

Note that $i_K^L|_{\operatorname{gal}(L/E)} = i_E^L$ because any generator for $\mathcal{O}_L$ over $\mathcal{O}_K$ also generates $\mathcal{O}_L$ over $\mathcal{O}_E$. The upshot of this definition is that it's enough to consider just $\sigma(x) - x$ for $x = \alpha$, instead of for all $x \in \mathcal{O}_L$. Namely, the function i_K^L determines our ramification groups and does not depend on our choice of generator α for $\mathcal{O}_L$ over $\mathcal{O}_K$; this is shown by the following theorem:

Theorem 12

$G_i(L/K) = \{\sigma \in \operatorname{gal}(L/K) : i_L^K(\sigma) \geq i+1\}$; in other words,

$$i_K^L(\sigma) \geq i+1 \Leftrightarrow (\forall x \in \mathcal{O}_L)[v_L(\sigma(x) - x) \geq i+1] \Leftrightarrow v_L(\sigma(\alpha) - \alpha) \geq i+1.$$

Proof: For any $r \in \mathbb{N}$ and $a_0, \ldots, a_r \in \mathcal{O}_K$ we have

$$\begin{aligned} v_L\left(\sigma\left(\sum_0^r a_i\alpha^i\right) - \sum_0^r a_i\alpha^i\right) &= v_L\left(\sum_1^r a_i(\sigma(\alpha)^i - \alpha^i)\right) \\ &= v_L\left(\sum_{i=1}^r a_i(\sigma(\alpha) - \alpha)\left(\sum_{j=0}^{i-1} \sigma(\alpha)^j\alpha^{r-1-j}\right)\right) \\ &= v_L(\sigma(\alpha) - \alpha) + v_L\left(\sum_{i=1}^r a_i\left(\sum_{j=0}^{i-1} \sigma(\alpha)^j\alpha^{r-1-j}\right)\right) \\ &\geq v_L(\sigma(\alpha) - \alpha). \end{aligned}$$

The last step is because $\sum_{i=1}^r a_i\left(\sum_{j=0}^{i-1}\sigma(\alpha)^j\alpha^{r-1-j}\right) \in \mathcal{O}_L$.

□

Here is another essential property of the ramification groups:

Theorem 13

$G_i \trianglelefteq \operatorname{gal}(L/K)$ for all i, and there is some N such that G_i is trivial for all $i \geq N$; in summary; the G_i form a *normal series* for $\operatorname{gal}(L/K)$.

Proof: Normality is because if $\sigma \in \operatorname{gal}(L/K)$ and $\tau \in G_i$, then we have $v_L(\tau(\sigma^{-1}(\alpha)) - \sigma^{-1}(\alpha)) \geq i+1$, i.e.

$\tau(\sigma^{-1}(\alpha)) - \sigma^{-1}(\alpha) \in \mathfrak{m}_L^{i+1}$; but since $\sigma(\mathfrak{m}_L) = \mathfrak{m}_L$ (since it must map maximal ideals to maximal ideals), then $\sigma(\mathfrak{m}_L^{i+1}) = \mathfrak{m}_L^{i+1}$, so that

$$\tau(\sigma^{-1}(\alpha)) - \sigma^{-1}(\alpha) \in \mathfrak{m}_L^{i+1} \implies \sigma(\tau(\sigma^{-1}(\alpha))) - \alpha \in \mathfrak{m}_L^{i+1}$$

and so $\sigma \circ \tau \circ \sigma^{-1} \in G_i$.

The G_i stabilize because $\mathrm{gal}(L/K)$ is finite by hypothesis, so we can take N with $N \geq i_K^L(\sigma)$ for all $\sigma \in \mathrm{gal}(L/K) - \{\mathbb{1}_L\}$ (since for all nontrivial σ, $i_K^L(\sigma)$ is finite); then for each such σ we have $\sigma \notin G_N$ since $i_K^L(\sigma) \leq N < N + 1$.

□

Let's now turn our attention towards studying how the ramification groups behave in towers; i.e. let's introduce an intermediate field E with $K \leq E \leq L$ and try to describe how the ramification groups of L/E and possibly E/K (if E is galois over K) interact with the ones for L/K. For L/E, the situation is quite simple:

Theorem 14

$G_i(L/E) = G_i(L/K) \cap \mathrm{gal}(L/E)$ for all $i \geq -1$.

Proof: As noted previously, for any $\sigma \in \mathrm{gal}(L/E)$ we have $i_E^L = i_K^L|_{\mathrm{gal}(L/E)}$. Then

$$\begin{aligned} \sigma \in G_i(L/E) &\Leftrightarrow \sigma \in \mathrm{gal}(L/E) \wedge i_E^L(\sigma) \geq i + 1 \\ &\Leftrightarrow \sigma \in \mathrm{gal}(L/E) \wedge i_K^L(\sigma) \geq i + 1 \\ &\Leftrightarrow \sigma \in \mathrm{gal}(L/E) \cap G_i(L/K). \end{aligned}$$

□

In particular, if $E = L_{\mathcal{I}}$ (the inertia field of L/K from definition 11), then $\mathrm{gal}(L/E) = \mathrm{gal}(L/L_{\mathcal{I}}) = G_0(L/K)$ by definition, so that the above theorem tells us $G_i(L/L_{\mathcal{I}}) = G_i(L/K) \cap G_0(L/K) = G_i(L/K)$ when $i \geq 0$. In words, the ramification groups for L/K and $L/L_{\mathcal{I}}$ are the exact same (excluding the -1-th ramification group, which is just the whole galois group anyway). In particular, we can always reduce to the totally ramified case when studying the higher ramification groups.

For E/K, the situation is not so nice; we know that $G_{-1}(E/K) = \mathrm{gal}(E/K) \cong \mathrm{gal}(L/K)/\mathrm{gal}(L/E)$ from basic Galois theory, but the other ramification groups (i.e. $G_i(E/K)$ with $i \geq 0$) don't have so nice a description - the best we can do is the following:

Theorem 15

If $\sigma \in \mathrm{gal}(E/K)$, then
$$i_K^E(\sigma) = \frac{1}{e_K^L} \sum_{\tau \in \mathrm{gal}(L/K) : \tau|_E = \sigma} i_K^L(\tau)$$

Proof: I'll skip the proof, but it's proposition 3 of section 4.1 in [4].

There is a better way to express the relationship between the ramification groups of L/K and E/K, but it requires us to 'raise' our perspective, so to speak.

3.1 The Hasse-Herbrand transition function

Let's keep the same notation and conventions we've been using, and make a couple 'out-of-pocket' definitions:

Definition 14. *For $t \geq -1$, let $G_t(L/K) \triangleq G_{\lceil t \rceil}(L/K)$ (so that $\sigma \in G_t(L/K) \Leftrightarrow v_L(\sigma(\alpha) - \alpha) \geq t + 1$ holds in this case too).*

Definition 15. *The **Hasse-Herbrand transition function** of L/K is $\varphi_K^L : \mathbb{R}_{\geq -1} \to \mathbb{R}_{\geq -1}$ given by*
$$\varphi_K^L(t) \triangleq \int_0^t \frac{|G_x(L/K)|}{|G_0(L/K)|}\, dx.$$

The integral may seem a bit strange - I only write it because everyone else insists on doing so in the literature - but it's really just an overly cutesy way to define a piecewise linear function whose pieces are the integer intervals $[i, i+1]$ for $i \geq -1$, and whose slope on $[i, i+1]$ is simply the size of $G_{i+1}(L/K)$ normalized by the size of $G_0(L/K)$. It's literally nothing deeper than that. If we want, we can write out a very explicit formula:
$$\varphi(t) = |G_0|^{-1} \left(\sum_{i=1}^{\lfloor t \rfloor} |G_i| + (t - \lfloor t \rfloor)|G_{\lceil t \rceil}| \right).$$

In particular, this formula makes it clear that if $\varphi(t) \in \mathbb{Z}$, then $t \in \mathbb{Z}$ (because every $|G_j|$ divides every $|G_i|$ for $i \leq j$).

Note that $(\varphi_K^L)' = 1$ on $(-1, 0)$, and if $G_i(L/K)$ is trivial for $i \geq N$ then $(\varphi_K^L)' = |G_0(L/K)|^{-1}$ on (N, ∞). Also note that φ is a homeomorphism $[-1, \infty) \to [-1, \infty)$, so that it has a well-defined inverse ψ, which we will use to 'shift' our indexing for the ramification groups.

Definition 16. *Given $s \in [-1, \infty)$, let $G^s \triangleq G_{\psi(s)}$. The G^s are said to be the ramification groups of L/K in the **upper numbering**.*

Equivalently, we have $G_t = G^{\varphi(t)}$ for $t \in [-1, \infty)$. Symmetrically, the G_t are said to be in the **lower numbering**.

The key reason to introduce and distinguish between these two numberings is that the lower numbering interacts well with subgroups (as we saw in theorem 14), while the upper numbering interacts well with quotients, as theorem 16 below shows:

Theorem 16

If $K \leq E \leq L$ with E/K galois, then $G^s(E/K) \cong G^s(L/K)\operatorname{gal}(L/E)/\operatorname{gal}(L/E)$ (under the natural identification of $\operatorname{gal}(E/K)$ with $\operatorname{gal}(L/K)/\operatorname{gal}(L/E)$).

Proof: There are three ingredients, after which a little symbol manipulation suffices:

(1) The first step is to prove that, if $\sigma \in \operatorname{gal}(E/K)$, and $\tau \in \operatorname{gal}(L/K)$ extends σ and is such that $i_K^L(\tau)$ is maximal among all extensions of σ to L, then $i_K^E(\sigma) = \varphi_E^L(i_K^L(\tau) - 1) + 1$.

This is because for any τ' extending σ, we have $i_K^L(\tau'\tau) = \min(i_K^L(\tau), i_K^L(\tau'))$, which in turn is simply because the minimum among $i_K^L(\tau)$, $i_K^L(\tau')$, $i_K^L(\tau\tau')$ cannot be unique (in general i_K^L evaluated at some element of the Galois group is $\geq i+1$ iff that element is in G_i, and if two of $\tau, \tau', \tau\tau'$ are in G_i then the third must be as well by group properties). Then by theorem 15, we have

$$\begin{aligned} i_K^E(\sigma) &= \sum_{\tau' \in \operatorname{gal}(L/K) : \tau'|_E = \sigma} i_K^L(\tau') \\ &= \sum_{\tau' \in \operatorname{gal}(L/E)} i_K^L(\tau\tau') \\ &= \sum_{\tau' \in \operatorname{gal}(L/E)} \min(i_K^L(\tau), i_K^L(\tau')) \\ &= \sum_{\tau' \in \operatorname{gal}(L/E)} \min(i_K^L(\tau), i_E^L(\tau')). \end{aligned}$$

In general we have $\sum_{\tau' \in \operatorname{gal}(L/E)} \min(t, i_E^L(\tau')) = |G_0(L/E)|(\varphi_E^L(t-1)+1)$, because they are continuous piecewise linear and the derivative at $t \notin \mathbb{Z}$ of both is equal to the number of $\tau' \in \operatorname{gal}(L/E)$ that are in G_{t-1}. Then letting $t \triangleq i_K^L(\tau)$ and combining with the above calculation, we get

$$i_K^E(\sigma) = \sum_{\tau' \in \operatorname{gal}(L/E)} \min(i_K^L(\tau), i_E^L(\tau')) = \varphi_E^L(i_K^L(\tau) - 1) + 1.$$

(2) The second step is to show that $G_t(L/K)\operatorname{gal}(L/E)/\operatorname{gal}(L/E) \cong G_{\varphi_E^L(t)}(E/K)$ (under the natural identi-

fication of $\operatorname{gal}(E/K)$ with $\operatorname{gal}(L/K)/\operatorname{gal}(L/E)$); this is because

$$\begin{aligned}
\sigma \operatorname{gal}(L/E) \in G_t(L/K)\operatorname{gal}(L/E)/\operatorname{gal}(L/E) &\Leftrightarrow i_K^L(\sigma) \geq t+1 \\
&\Leftrightarrow \varphi_E^L(i_K^L(\sigma)-1) \geq \varphi_E^L(t+1) \\
&\Leftrightarrow i_K^E(\sigma|_E) \geq \varphi_E^L(t+1) && \text{(by (1))} \\
&\Leftrightarrow \sigma|_E \in G_{\varphi_E^L(t)}(E/K) \\
&\Leftrightarrow \sigma|_E \in G^t(E/K)
\end{aligned}$$

(3) The final step is to show that φ composes in towers, i.e. that $\varphi_K^E \circ \varphi_E^L = \varphi_K^L$; using the usual strategy of comparing derivatives for continuous piecewise linear functions, we get

$$\begin{aligned}
(\varphi_K^E \circ \varphi_E^L)'(t) &= (\varphi_K^E)'(\varphi_E^L(t)) \cdot (\varphi_E^L)'(t) \\
&= \frac{|G_{\varphi_E^L(t)}(E/K)|}{|G_0(E/K)|} \cdot \frac{|G_t(L/E)|}{|G_0(L/E)|} \\
&= \frac{|G_{\varphi_E^L(t)}(E/K)||G_t(L/E)|}{e_K^L} \\
&= \frac{|G_t(L/K)\operatorname{gal}(L/E)/\operatorname{gal}(L/E)||G_t(L/E)|}{e_K^L} \\
&= \frac{|G_t(L/E)||G_t(L/K)|}{e_K^L|G_t(L/K)\cap\operatorname{gal}(L/E)|} \\
&= \frac{|G_t(L/E)||G_t(L/K)|}{e_K^L|G_t(L/E)|} \\
&= \frac{|G_t(L/K)|}{|G_0(L/K)|} \\
&= (\varphi_K^L)'(t)
\end{aligned}$$

so that $\varphi_K^E \circ \varphi_E^L = \varphi_K^L$ since both equal -1 at -1.

Putting the above all together, we get

$$\begin{aligned}
G^s(L/K)\operatorname{gal}(L/E)/\operatorname{gal}(L/E) &= G_{\psi_K^L(s)}\operatorname{gal}(L/E)/\operatorname{gal}(L/E) \\
&\cong G_{\varphi_E^L(\psi_K^L(s))}(E/K) && \text{(by (2))} \\
&= G^{\varphi_K^E(\varphi_E^L(\psi_K^L(s)))}(E/K) \\
&= G^{\varphi_K^L(\psi_K^L(s))}(E/K) && \text{(by (3))} \\
&= G^s(E/K)
\end{aligned}$$

□

Serre sets up his notation so that he can write this statement very cutely as $(G/H)^s = G^s H/H$, but I write it as I did above for transparency's sake.

3.2 Factors of the ramification series

Keep the same conventions as before, but let $\mathfrak{m}$ denote $\mathfrak{m}_L$ throughout this section (because I write it way too many times in the proofs to justify putting the subscripts everywhere).

Since we've seen that the $G_i(L/K)$ form a normal series for the whole Galois group, one natural question is what the factor groups of this series look like, as this should give us information about the G_i.

The first step is to reframe some things we've done previously in terms of multiplication. Namely, we said $\sigma \in G_i(L/K) \Leftrightarrow \sigma(\alpha) \equiv \alpha \bmod \mathfrak{m}^{i+1}$ for a generator α of $\mathcal{O}_L$ over $\mathcal{O}_K$; if we instead look at $L/L_{\mathcal{I}}$, which we know is totally ramified, then we can take a uniformizing generator π of $\mathcal{O}_L$ over $\mathcal{O}_{L_{\mathcal{I}}}$ by theorem 10. In this case we still have $\sigma \in G_i(L/K) \Leftrightarrow \sigma(\pi) \equiv \pi \bmod \mathfrak{m}^{i+1}$, and - using the fact that π has valuation $1(!)$ - this latter condition is equivalent to

$$\frac{\sigma(\pi)}{\pi} \equiv 1 \bmod \mathfrak{m}^i.$$

In other words, fixing the valuation of our generator has allowed us to control how much valuation we 'lose' when multiplying / dividing by it, which is what permits us to reframe things multiplicatively here.

We saw in theorem 14 that $G_i(L/K) = G_i(L/L_{\mathcal{I}})$ for $i \geq 0$, so as noted previously restricting to the extension $L/L_{\mathcal{I}}$ makes no difference (again, $G_{-1}(L/K)$ was always just the whole galois group anyway, so we are not losing generality here).

So we should be all set to introduce the unit groups now!

Definition 17. *Let $U_L^{(0)} \triangleq \mathcal{O}_L^\times$ and $U_L^{(i)} \triangleq 1 + \mathfrak{m}_L^i$ for $i \geq 1$.*

These $U_L^{(i)}$ form a descending neighborhood basis for 1 (i.e. $U_L^{(0)} \geq U_L^{(1)} \geq U_L^{(2)} \geq \dots$), and are all complete, so $\mathcal{O}_L^\times \cong \varprojlim \mathcal{O}_L^\times / U_L^{(i)}$. Their quotients have quite a nice structure, as the following two theorems show:

Theorem 17

$\mathcal{O}_L^\times / U_L^{(1)} \cong \ell^\times$.

Proof: We can just define a map explicitly, via $x(1+\mathfrak{m}) \mapsto x + \mathfrak{m}$; this actually maps into $\ell^\times$ because $x \in \mathcal{O}_L^\times$ implies $x \notin \mathfrak{m}$, so that $x + \mathfrak{m} \neq 0 + \mathfrak{m}$. For injectivity + well-definedness, note that $x + \mathfrak{m} = y + \mathfrak{m} \Leftrightarrow xy^{-1} \in 1 + \mathfrak{m} \Leftrightarrow x(1+\mathfrak{m}) = y(1+\mathfrak{m})$ for $x, y \in \mathcal{O}_L^\times$. It's also clear it is a surjective homomorphism.

□

Theorem 18

$U_L^{(i)}/U_L^{(i+1)} \cong (\ell, +)$ for $i \geq 1$.

Proof: First we define a map $U_L^{(i)}/U_L^{(i+1)} \to \mathfrak{m}^i/\mathfrak{m}^{i+1}$, via $x(1+\mathfrak{m}^{i+1}) \mapsto (x-1)+\mathfrak{m}^{i+1}$. This is a group homomorphism because for $x, y \in 1+\mathfrak{m}^i$ we have $(xy-1)+\mathfrak{m}^{i+1} = (x-1+y-1)+\mathfrak{m}^{i+1}$ (as $xy-x-y+1 = (x-1)(y-1) \in \mathfrak{m}^{i+1}$, since $x-1, y-1 \in \mathfrak{m}^i$). It is injective since $y \in 1+\mathfrak{m}^i$ implies y is a unit in $\mathcal{O}_L$, so that

$$\begin{aligned}
(x-1)+\mathfrak{m}^{i+1} = (y-1)+\mathfrak{m}^{i+1} &\implies x-y \in \mathfrak{m}^{i+1} \\
&\implies \frac{x}{y} - 1 \in \mathfrak{m}^{i+1} \\
&\implies xy^{-1} \in 1+\mathfrak{m}^{i+1} \\
&\implies x(1+\mathfrak{m}^{i+1}) = y(1+\mathfrak{m}^{i+1}).
\end{aligned}$$

And surjectivity is clear.

Now let's show that $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ is an ℓ-vector space: this is because the homomorphism $\Theta : \mathcal{O}_L \to \mathrm{aut}(\mathfrak{m}^i/\mathfrak{m}^{i+1})$ given by $\Theta(a)(x+\mathfrak{m}^{i+1}) \triangleq ax+\mathfrak{m}^{i+1}$ is constant on every coset of $\mathfrak{m}$ (as $a \in \mathfrak{m} \implies ax \in \mathfrak{m}^{i+1} \implies ax+\mathfrak{m}^{i+1} = 0$), so it induces an action $\mathcal{O}_L/\mathfrak{m} \to \mathrm{aut}(\mathfrak{m}^i/\mathfrak{m}^{i+1})$.

Additionally, since $1+\mathfrak{m}^{i+1}$ generates $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ over ℓ (as $x+\mathfrak{m}^{i+1} = (x+\mathfrak{m})(1+\mathfrak{m}^{i+1})$), then $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ is a 1-dimensional vector space over ℓ, and so is isomorphic to $(\ell, +)$ as a group. So in summary we have $U_L^{(i)}/U_L^{(i+1)} \cong \mathfrak{m}^i/\mathfrak{m}^{i+1} \cong (\ell, +)$.

□

The upshot of understanding the structure of the $U_L^{(i)}/U_L^{(i+1)}$ is that we can embed the factors G_i/G_{i+1} in them (this embedding is induced by the map $\sigma \mapsto \sigma(\pi)/\pi : G_i(L/K) \to 1+\mathfrak{m}^i$, which is well-defined since $\sigma \in G_i(L/K) \implies \sigma(\pi) \equiv \pi \bmod \mathfrak{m}^{i+1} \implies \sigma(\pi)/\pi \equiv 1 \bmod \mathfrak{m}^i$), and so we get information about them from the above two theorems.

For example, if $\mathrm{char}(\ell) = 0$, then finite subgroups of $(\ell, +)$ are trivial, and so G_i/G_{i+1} is trivial for $i \geq 1$ - but since G_N is also trivial for large N, by multiplying orders / indexes we get that G_i is trivial for $i \geq 1$. Note this implies that G_0 is cyclic, since G_1 being trivial means $G_0 \cong G_0/G_1$ embeds into $\mathcal{O}_L^\times/U_L^{(1)} \cong \ell^\times$.

And if $\mathrm{char}(\ell) = p$, then subgroups of $(\ell, +)$ are vector spaces over $\mathbb{F}_p$, and so $G_i/G_{i+1} \leq (\mathbb{F}_{p^k}, +) \cong \mathbb{Z}_p^k$ meaning it is elementary abelian. Since G_N is trivial for large N, again by multiplying orders / indexes we get that the G_i themselves are all p-groups (for $i \geq 1$).

4 Newton polygons and copolygons

Much of this section is an elaboration on [2].

4.1 Definitions and Tate's lemma

Now, since the Hasse-Herbrand transition function allows us to translate between the lower and upper numberings on our ramification groups, it would be useful to have a nice way to calculate it. The Newton copolygon / valuation function gives us a nice way to do this in certain cases.

Definition 18. *For $f(T) \in \mathcal{O}_K[T]$ (with $f(T) = \sum_{i=0}^{n} c_i T^i$), where K is a local field with valuation v on $\overline{K}$ normalized for K.*

- *the **valuation function** or **Newton copolygon** of f is the piecewise linear function $\Psi_{v,f} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ given by $\Psi_{v,f}(t) \triangleq \min_{0 \leq i \leq n} \{it + v(c_i)\}$.*
- *the **Newton polygon** of f is the piecewise linear function going through the points (β, b), where β is the slope of a piece of the Newton copolygon of f and b is its y-intercept (if it were extended to a full line). There will be some point $(\alpha, 0)$; define the Newton polygon to equal 0 for $x > \alpha$.*

We call the non-differentiable points (i.e. the boundary points of the intervals of the linear pieces) of either the *vertices* of the function.

The basic intuition behind the first definition (which is where the name 'valuation function' probably stems from) is that $v(f(x)) \geq \min_{0 \leq i \leq n} \{v(c_i x^i)\} = \min_{0 \leq i \leq n} \{iv(x) + v(c_i)\}$, with - by theorem 1 - equality when the $iv(x) + v(c_i)$ have a unique minimum, which happens for all but finitely many values of $v(x)$ (one way to see this is that the $iv(x) + v(c_i)$ are linear functions in $v(x)$, each with different slopes). When there is such a unique minimum, this means that $v(f(x))$ depends *entirely* on $v(x)$, and $\Psi_{v,f}(v(x)) = v(f(x))$ at all but finitely many points $t_1, \ldots, t_k$. Then $\Psi_{v,f}$ is linear on each interval $[t_i, t_{i+1}]$ for $1 \leq i \leq k-1$, as well as on $[0, t_1]$ and $[t_k, \infty)$.

As the co- prefix indicates, the Newton polygon and copolygon are dual in the sense that there is a one-to-one correspondence between points of one and lines of the other. Because of this, they essentially carry the same information, so anything that can be done with one can be done with the other (though one might be preferable to use over another for a given situation).

More precisely, we have the following correspondence: to get from the Newton copolygon to the Newton polygon, a line segment with equation $y = at + b$ becomes a vertex (a, b), while a vertex (a, b) becomes a line segment with equation $y = b - at$. Conversely, to get from the Newton polygon to the Newton copolygon, a line segment with equation $y = a + bt$ becomes a vertex $(-b, a)$, and a vertex (a, b) becomes a line segment with equation $y = at + b$. A visualized example of this correspondence for a specific polynomial can be found in the appendix of [2].

Additionally, note that the Newton copolygon of a polynomial with nonzero constant term is eventually constant, while the Newton copolygon of a polynomial with zero constant term is eventually linear with positive slope.

For certain extensions, Newton copolygons turn out to be very nice to use in calculating its Hasse-Herbrand transition function, via the following theorem:

Theorem 19: Tate's Lemma (proposition 1.2 from [2])

If $K_0 \leq K \leq L$ are local fields, with L/K galois and L/K_0 finite, v the valuation on L normalized for K_0, π a uniformizer for L and $f(T) \triangleq \operatorname{irr}(\pi, L_{G_0(L/K)})(T+\pi)$, then

$$\varphi_K^L(t-1)+1 = e_{K_0}^K \Psi_{v,f}\left(\frac{t}{e_{K_0}^L}\right)$$

for $t \geq 0$.

Proof: Denote $L_{G_0(L/K)}$ as just $L_{\mathcal{I}}$. In this case, we have

$$\begin{aligned} v(f(T+\pi)) &= v\left(\prod_{\sigma \in \operatorname{gal}(L/L_{\mathcal{I}})} (T-(\sigma(\pi)-\pi))\right) \\ &= \sum_{\sigma \in \operatorname{gal}(L/L_{\mathcal{I}})} v(T-(\sigma(\pi)-\pi)) \\ &\geq \sum_{\sigma \in \operatorname{gal}(L/L_{\mathcal{I}})} \min(v(T), v(\sigma(\pi)-\pi)) \end{aligned}$$

with equality whenever $v(T)$ does not equal $v(\sigma(\pi)-\pi)$ for any $\sigma \in \operatorname{gal}(L/L_{\mathcal{I}})$. Since this only happens for finitely many values of $v(T)$, it suffices to prove the theorem when we have equality (at which point the full theorem will follow by continuity). So we can assume $\Psi_{v,f}(t) = \sum_{\sigma \in \operatorname{gal}(L/L_{\mathcal{I}})} \min(t, v(\sigma(\pi)-\pi))$; we can even assume this holds in a neighborhood of a given t.

This means that $\Psi'_{v,f}(t)$ equals the number of $\sigma \in G_0(L/K)$ with $v(\sigma(\pi)-\pi) > t$ (which equals the number of $\sigma \in G_0(L/K)$ with $v(\sigma(\pi)-\pi) \geq t$, since we were assuming none of them equal t). Since $v(\sigma(\pi)-\pi) = v_L(\sigma(\pi)-\pi)/e_{K_0}^L$, this also equals the number of $\sigma \in G_0(L/K)$ with $v_L(\sigma(\pi)-\pi) \geq e_{K_0}^L t$.

If $e_{K_0}^L t \geq 1$, this is just $|G_{e_{K_0}^L t-1}(L/K)|$; otherwise it is simply $|G_0(L/K)|$ (since every $\sigma \in G_0(L/K)$ has $v_L(\sigma(\pi)-\pi) \geq 1$).

Then for $t \geq 1$ (so that $e_{K_0}^L(t/e_{K_0}^L) \geq 1$), we have (assuming φ_K^L is differentiable at t, which we can do since it is such at all but finitely many points):

$$\begin{aligned}
\left(\Psi_{v,f} \circ \frac{-}{e^L_{K_0}}\right)'(t) &= \frac{1}{e^L_{K_0}} \Psi'_{v,f}\left(\frac{t}{e^L_{K_0}}\right) \\
&= \frac{1}{e^L_{K_0}} |G_{e^L_{K_0}(t/e^L_{K_0})-1}(L/K)| \quad (\text{since } e^L_{K_0}(t/e^L_{K_0}) \geq 1) \\
&= \frac{1}{e^K_{K_0} e^L_K} |G_{t-1}(L/K)| \\
&= \frac{1}{e^K_{K_0}} \cdot \frac{|G_{t-1}(L/K)|}{|G_0(L/K)|} \\
&= \frac{1}{e^K_{K_0}} (\varphi^L_K)'(t-1).
\end{aligned}$$

Similarly, if $t < 1$ then we get

$$\begin{aligned}
\left(\Psi_{v,f} \circ \frac{-}{e^L_{K_0}}\right)'(t) &= \frac{1}{e^L_{K_0}} \Psi'_{v,f}\left(\frac{t}{e^L_{K_0}}\right) \\
&= \frac{1}{e^L_{K_0}} |G_0(L/K)| \qquad (\text{since } e^L_{K_0}(t/e^L_{K_0}) < 1) \\
&= \frac{e^L_K}{e^L_{K_0}} \\
&= \frac{1}{e^K_{K_0}} \\
&= \frac{1}{e^K_{K_0}} (\varphi^L_K)'(t-1)
\end{aligned}$$

so that the derivatives are equal always. Then since $\Psi_{v,f}(0/e^L_{K_0}) = 0$ and $\varphi^L_K(0-1) = -1$, the result follows. $\square$

The existence of K_0 in the above theorem is just so that we can scale things if we feel like it (for instance, if it might be easier to calculate the ramification indexes of K and L over K_0 than of L over K).

Next I'll give an example of how theorem 19 can be used to explicitly calculate a Hasse-Herbrand transition function.

4.2 Tate's lemma computation example

Let K be a finite extension of $\mathbb{Q}_7$ (the 7-adic rationals) containing a primitive 7-th root of unity ζ_7, π a uniformizer of K, and denote by $\pi^{1/7}$ a fixed root of $T^7 - \pi$ in the algebraic closure $\overline{\mathbb{Q}_7}$. We want to calculate $\varphi_K^{K(\pi^{1/7})}$. Note that $K(\pi^{1/7})$ is galois over K with degree 7 because $\zeta_7 \in K$. I'll let K_0 equal $\mathbb{Q}_7$, so that our valuation (which I'll write v_7) is normalized for $\mathbb{Q}_7$ (for example, $v_7(7^k) = k$). I'll also write $n \triangleq [K : \mathbb{Q}_7]$.

Since $T^7 - \pi \in \mathcal{O}_K[T]$ is Eisenstein with none of its roots having valuation 0, then by theorem 11 it is irreducible and $K(\pi^{1/7})/K$ is totally ramified. In particular, this means the maximal unramified subextension of $K(\pi^{1/7})/K$ is just K, so our $f(T)$ in Tate's lemma will be

$$\operatorname{irr}(\pi, K)(T + \pi) = (T + \pi)^7 - \pi = \sum_{i=0}^{6} \binom{7}{i} T^{7-i} \pi^{i/7}$$

So let's calculate $\Psi_{v,f}$ for this f. Since $v_7(\binom{7}{i}) = \delta_{i,0}$ (as can be easily verified directly), we have

$$\psi_{v,f}(t) = \min_{0 \le i \le 6} \left\{ (7 - i)t + \delta_{i,0} + \frac{i}{7n} \right\}.$$

It's easiest to see what this minimum really equals by writing out the equations individually:

- $\star$ $7t$
- $\star$ $6t + \frac{1}{7n} + 1$
- $\star$ $5t + \frac{2}{7n} + 1$

$\vdots$

- $\star$ $t + \frac{6}{7n} + 1$

For small t (near 0), it's clear that $7t$ will be the minimum, with the first tie at $t = \frac{1}{7n} + \frac{1}{6}$ (between $7t$ and $t + \frac{6}{7n} + 1$). And for $t > \frac{1}{7n}$, we have

$$t + \frac{6}{7n} + 1 = t + \frac{(i-1) + (7-i)}{7n} + 1 < t + (i-1)t + \frac{7-i}{7n} + 1 = it + \frac{7-i}{7n} + 1,$$

so that $t + \frac{6}{7n} + 1$ will be the minimum for all $t > \frac{1}{7n} + 1/6$.

Then $\Psi_{v,f}(t)$ has slope 7 on $(0, \frac{1}{7n} + \frac{1}{6})$ and slope 1 on $(\frac{1}{7n} + \frac{1}{6}, \infty)$, so Tate's lemma tells us $\varphi_K^{K[\pi^{1/7}]}$ has slope $7e_{\mathbb{Q}_7}^K / e_{\mathbb{Q}_7}^{K[\pi^{1/7}]} = 1$ on $(-1, \frac{7n}{6})$ and slope $e_{\mathbb{Q}_7}^K / e_{\mathbb{Q}_7}^{K[\pi^{1/7}]} = \frac{1}{7}$ on $(\frac{7n}{6}, \infty)$. Note that from the definition, the points where the Hasse-Herbrand transition function changes slope (the 'lower breaks') must necessarily be integers. In this example there is a lower break at $\frac{7n}{6}$, but there is no contradiction here - $\frac{7n}{6}$ is an integer since we assumed K contained a primitive 7-th root of unity ζ_7 (and so has degree over $\mathbb{Q}_7$ divisible by the degree of ζ_7 over $\mathbb{Q}_7$, which is 6). This was needed to ensure that $K(\pi^{1/7})/K$ was actually galois.

Note that we can also recover information about the sizes of the higher ramification groups from the Hasse-Herbrand transition function; in particular, we have

$$G_t(K(\pi^{1/7})/K) = \begin{cases} \operatorname{gal}(K(\pi^{1/7})/K), & t \le \frac{7n}{6} \\ \{\mathbb{1}_{K(\pi^{1/7})}\}, & t > \frac{7n}{6} \end{cases}$$

4.3 Altitude of extensions

Definition 19. *The* ***altitude*** *of a finite separable totally ramified extension L/K (with $L \neq K$) is the value of φ_K^L at its rightmost 'vertex' (where the vertices are the points at which the slope changes) - in other words, if t is the infimum over the ones so that $G_t(L/K)$ is trivial, then $\mathrm{alt}_K^L \triangleq \varphi_K^L(t)$. If $L = K$, then we say $\mathrm{alt}_K^L = 0$.*

The altitude of a finite separable extension L/K is the altitude of $L/L_{\mathcal{I}}$, and L^s $(s > 0)$ is the compositum of all subfields of L/K with altitude $< s$ (which also has altitude $< s$ by theorem 20 below).

Definition 20. *Call a finite separable extension L/K 'at-most-s-upper-ramified' iff $G^s(L/K) = \mathbb{1}_L$.*

Note that the altitude is always nonnegative, that (finite separable) unramified extensions have altitude 0, and that the altitude is essentially the infimum over the s such that L/K is at-most-s-upper-ramified (i.e. L/K is at-most-s-upper-ramified iff $\mathrm{alt}_K^L < s$). Below are a few additional properties:

Theorem 20

Let L, E be finite separable over K.

(1) $\mathrm{alt}_K^{LE} \leq \max(\mathrm{alt}_K^L, \mathrm{alt}_K^E)$ (in particular, $\mathrm{alt}_K^{LE} = \mathrm{alt}_K^L$ when E/K is unramified).

(2) $\mathrm{alt}_K^L = \mathrm{alt}_K^{\mathrm{spl}(L/K)}$.

(3) $E^s = L^s \cap E$ for all $s > 0$.

Proof: I will omit it, but details can be found in [2].

---

5 Deligne's main result

This section will use a lot of the previous material to give a partial exposition up to theorem 2.8 of [1].

First we need a few preliminary things, which may seem a bit random but will come into play later. The very first is that throughout this section, a *local field* will mean a complete discretely valued field with perfect residue field (not necessarily finite). The second is that we will shift our numberings for the ramification groups; redefine $G_t(L/K) \triangleq \{\sigma \in \mathrm{gal}(L/K) : i_K^L(\sigma) \geq t\}$ (for $t \geq 0$), so that our new G_t is our old G_{t-1}. The others are a barrage of definitions:

Definition 21. *Given an R-module M, let $M^{\otimes n}$ for $n > 0$ denote the tensor product $\bigotimes_{i=1}^{n} M$; let $M^{\otimes 0}$ denote R; and let $M^{\otimes n}$ for $n < 0$ denote $\hom(M^{\otimes -n}, R)$.*

The direct sum $\bigoplus_{n\in\mathbb{Z}} M^{\otimes n}$ has a natural R-algebra structure given by tensoring / function application. The structure can be described completely in the case M is free of rank 1 over R (which it always will be throughout this section) by the fact that $\bigoplus_{n\in\mathbb{Z}} M^{\otimes n}$ is isomorphic to $R[T, \frac{1}{T}]$ via fixing a generator α of M over

R and mapping it to T. (In particular, the element $\frac{1}{T}$ will correspond to the homomorphism ϕ determined by $\phi(\alpha) = 1$.)

Definition 22. *A **truncated valuation ring** is a local principal ideal ring R whose maximal ideal $\mathfrak{m}_R$ is nilpotent. The smallest s such that $\mathfrak{m}_R^s = 0$ is the **length** of R and denoted $\lg(R)$.*

Note that because the generator of the maximal ideal is nilpotent, any (non-field) truncated valuation ring is necessarily *not* an integral domain and so not a PID, even though all ideals are generated by one element.

It can be shown that an equivalent definition of a truncated valuation ring is a quotient of a complete DVR R by a power $\mathfrak{m}^s$ of its maximal ideal (where $\lg(R/\mathfrak{m}^s)$ will equal s necessarily). Truncated valuation rings inherit a 'truncated valuation' from the ring of which they are a quotient; the truncated valuation of $R/\mathfrak{m}^s$ takes values in $[0, s-1] \cup \{\infty\}$, and is truncated in the sense that if for $x \in R$ we have $v_\mathfrak{m}(x) \in [0, s-1]$, then the truncated valuation at $x + \mathfrak{m}^s$ equals $v_\mathfrak{m}(x)$, and if $v_\mathfrak{m}(x) \geq s$ then the truncated valuation at $x + \mathfrak{m}^s$ is ∞.

Definition 23. *We will reserve the term "**triple**" for a tuple (R, M, ε) where R is a truncated valuation ring, M is a free R-module of rank 1, and $\varepsilon : M \to \mathfrak{m}_R$ is a surjective homomorphism.*

If K is a local field, for every $s \in \mathbb{Z}_+$ we can associate with it the triple $\mathrm{tr}_s(K) \triangleq (\mathcal{O}_K/\mathfrak{m}_K^s, \mathfrak{m}_K/\mathfrak{m}_K^{s+1}, \varepsilon : \mathfrak{m}_K/\mathfrak{m}_K^{s+1} \to \mathfrak{m}_K/\mathfrak{m}_K^s)$ (where ε is the natural map $x + \mathfrak{m}_K^{s+1} \mapsto x + \mathfrak{m}_K^s$)

Using the fact that any truncated valuation ring is a quotient of a complete DVR R by a power $\mathfrak{m}^s$ of its maximal ideal, it can be shown that every triple can be realized as $\mathrm{tr}_s(K)$ for some $s \in \mathbb{Z}_+$ and some local field K.

Definition 24. *Given a triple (R, M, ε), if $r < s$, then $\varepsilon_{r,s}$ is the map $M^{\otimes s} \to M^{\otimes r}$ determined by $\varepsilon_{r,s}(\alpha^{\otimes s}) \triangleq \varepsilon(\alpha)^{s-r}\alpha^{\otimes r}$, where α is a generator of M over R (so that $\alpha^{\otimes s}$ is one of $M^{\otimes s}$ over R).*

Note this map is independent of the generator we choose, since if $\beta = u\alpha$ (so that $\beta^{\otimes s} = u^s\alpha^{\otimes s}$) then $\varepsilon(\beta)^{s-r}\beta^{\otimes r} = u^s\varepsilon(\alpha)^{s-r}\alpha^{\otimes r}$.

Definition 25. *Let $\mathcal{T}$ be the category of triples, with morphisms $(R, M, \varepsilon) \to (R', M', \varepsilon')$ as tuples $(\varphi : R \to R', \eta : M \to M'^{\otimes e}, e)$ (where $e \geq 1$, φ and η are ring / R-module homomorphisms resp. (with M' and so $M'^{\otimes e}$ given the R-module structure induced by φ via 'restriction of scalars'), and η maps a generator of M to one of $M'^{\otimes e}$) so that the following diagram commutes:*

$$\begin{array}{ccc} M & \xrightarrow{\eta} & M'^{\otimes e} \\ \downarrow{\scriptstyle \varepsilon} & & \downarrow{\scriptstyle \varepsilon_{0,e}} \\ R & \xrightarrow{\varphi} & R' \end{array}$$

Composition of morphisms is given $(\varphi', \eta', e') \circ (\varphi, \eta, e) \triangleq (\varphi' \circ \varphi, \eta'^{\otimes e} \circ \eta, e'e)$. Note that a generator of $\mathfrak{m}_R$ necessarily maps to a generator of $\mathfrak{m}_{R'}^e$ under φ, so in particular $\lg(R') \leq e\lg(R)$.

Given two triples S, T and a morphism between them, we sometimes say we have an *extension* T/S. Also, η induces an isomorphism $\overline{\eta} : M \otimes_R R' \hookrightarrow\!\!\!\!\rightarrow M'^{\otimes e}$ via $m \otimes_R r' \mapsto r'\eta(m)$.

We saw previously that every triple can be realized as $\mathrm{tr}_s(K)$ for some $s \geq 1$ and some local field K; the purpose of the above definition of morphisms between triples is to properly encode extensions of their corresponding local fields. In particular, the e in the above definition will correspond to the ramification index of the corresponding field extension.

Looking at it in the reverse direction, given a finite extension L/K of local fields with ramification index e, for any $s \in \mathbb{Z}_+$ the inclusion $\iota : K \hookrightarrow L$ induces a morphism $\widetilde{\iota} \triangleq (\varphi, \eta, e)$ between the triples $\mathrm{tr}_s(K)$ and $\mathrm{tr}_{es}(L)$, where φ is the natural map $\mathcal{O}_K/\mathfrak{m}_K^s \to \mathcal{O}_L/\mathfrak{m}_L^{es}$ (note the inclusion $\mathcal{O}_K \to \mathcal{O}_L$ maps exactly $\mathfrak{m}_K^s$ to exactly $\mathfrak{m}_L^{es}$ since $v_L|_K = ev_K$), and η is induced by the natural map $\mathfrak{m}_K/\mathfrak{m}_K^{s+1} \to \mathfrak{m}_L^e/\mathfrak{m}_L^{e(s+1)}$ and the isomorphism $\mathfrak{m}_L^e/\mathfrak{m}_L^{e(s+1)} \cong (\mathfrak{m}_L/\mathfrak{m}_L^{es+1})^{\otimes e}$ (induced by $\left(\prod_1^e x_i\right) + \mathfrak{m}_L^{e(s+1)} \mapsto \bigotimes_1^e \left(x_i + \mathfrak{m}_L^{es+1}\right)$).

Further building this correspondence, we can define the following:

Definition 26. *We say a morphism (φ, η, e) between two triples (R, M, ε) and (R', M', ε') is:*

- ***separable*** *iff $\lg(R') = e\lg(R)$*
- ***finite*** *iff R' is a finitely generated R-module*
- ***unramified*** *iff it is separable + finite and $e = 1$*
- ***totally ramified*** *iff $R/\mathfrak{m}_R \cong R'/\mathfrak{m}_{R'}$*

We might also say that (R', M', ε') over (R, M, ε) is the same, if the morphism is understood.

As the names suggest, these properties holding for an extension of triples will correspond to it holding for their corresponding field extension. Analogously to the field case, it can be shown that any finite morphism $S \to S''$ of triples can be factored as $S \to S' \to S''$, where the first morphism is unramified and the second is totally ramified.

To finally make this notion of correspondence concrete, Deligne defines the following categories:

Definition 27.

- *If K is a local field, $\mathcal{E}(K)$ is the category whose objects are finite separable extensions L of K, with morphisms $L \to L'$ as K-homomorphisms $L \to L'$.*
- *If $S_0 \triangleq (R, M, \varepsilon)$ is a triple, $\mathcal{E}(S_0)$ is the category whose objects are pairs (S, f), where S is a triple and $f : S_0 \to S$ is a flat + finite morphism, and whose morphisms $(S, f) \to (S', f')$ are morphisms $g : S \to S'$ with $f' = g \circ f$.*

Given a morphism $L \to L'$ in $\mathcal{E}(K)$, we can talk about $e_L^{L'}$ by equating L with its image in L' (remember that field homomorphisms are always injective).

He then constructs, given a local field K and some fixed $s \in \mathbb{Z}_+$, a functor T_s from $\mathcal{E}(K)$ to $\mathcal{E}(\mathrm{tr}_s(K))$, mapping L to $\mathrm{tr}_{e_K^L s}(L)$ and mapping $\iota : L \hookrightarrow L'$ to the morphism $\mathrm{tr}_{e_K^L s}(L) \to \mathrm{tr}_{e_K^{L'} s}(L')$ described in the remarks after definition 25. Note that this morphism is both finite (since $\mathcal{O}_{L'}/\mathfrak{m}_{L'}^{e_K^{L'} s}$ is a finitely generated $\mathcal{O}_L/\mathfrak{m}_L^{e_K^L s}$-module) and separable (by definition, since $\lg(\mathcal{O}_L/\mathfrak{m}_L^{e_K^L s}) = e_K^L s$ and $\lg(\mathcal{O}_{L'}/\mathfrak{m}_{L'}^{e_K^{L'} s}) = e_L^{L'} e_K^L s$).

The power of this whole set-up is that the functor T_s will turn out to be almost an equivalence - namely, it's a general fact that a functor is an equivalence of categories iff it is full, faithful, and essentially surjective on objects (I remember doing this exercise in Leinster ...), and T_s is faithful and essentially surjective. The missing ingredient is fullness - and it turns out T_s is not actually such. To get around this, Deligne employs a sort of funny but also natural trick - he mods out the codomain category $\mathcal{E}(\mathrm{tr}_s(K))$ appropriately (in a way that also maintains faithfulness).

Firstly:

Theorem 21

T_s is essentially surjective on objects.

Proof: Here is a sketch: given a object $((R', M', \varepsilon'), (\varphi, \eta, e)) \in \mathcal{E}(\mathrm{tr}_s(K))$, it is finite and separable over $\mathrm{tr}_s(K)$; we can also assume it is totally ramified. We have the following (commutative) diagram:

$$\begin{array}{ccc} \mathfrak{m}_K/\mathfrak{m}_K^{s+1} & \xrightarrow{\eta} & M'^{\otimes e} \\ \downarrow{\scriptstyle \varepsilon} & & \downarrow{\scriptstyle \varepsilon'_{0,e}} \\ \mathcal{O}_K/\mathfrak{m}_K^s & \xrightarrow{\varphi} & R' \end{array}$$

If β is a generator of M', then R' is a free $\mathcal{O}_K/\mathfrak{m}_K^s$-module with basis $1, \varepsilon'(\beta), \ldots, \varepsilon'(\beta)^{r-1}$; then using our induced isomorphism $\overline{\eta} : \mathfrak{m}/\mathfrak{m}_K^{s+1} \otimes_{\mathcal{O}_K/\mathfrak{m}_K^s} R' \hookrightarrow\!\!\!\!\rightarrow M'^{\otimes e}$, we have a unique system of elements $-a_0, \ldots, -a_{r-1} \in \mathfrak{m}_K/\mathfrak{m}_K^{s+1}$ so that $x^{\otimes} = \overline{\eta}\left(\sum_0^{r-1} -a_i \otimes_{\mathcal{O}_K/\mathfrak{m}_K^s} \varepsilon'(\beta)^i\right)$ i.e. so that

$$x^{\otimes} + \overline{\eta}\left(\sum_0^{r-1} a_i \otimes_{\mathcal{O}_K/\mathfrak{m}_K^s} \varepsilon'(\beta)^i\right) = 0$$

Then R' can be reconstructed from the a_i, since $R' \cong (\mathcal{O}_K/\mathfrak{m}_K^s)[T]/\left\langle T^r + \sum_0^{r-1} \varepsilon(a_i)T^i\right\rangle$; M' and ε' can be reconstructed similarly by taking any free R'-module of rank 1 and mapping its generator to T; φ can be reconstructed as the canonical map $\mathcal{O}_K/\mathfrak{m}_K^s \to (\mathcal{O}_K/\mathfrak{m}_K^s)[T]/\left\langle T^r + \sum_0^{r-1} \varepsilon(a_i)T^i\right\rangle$; and η can be reconstructed as well. Then $((R', M', \varepsilon'), (e, \varphi, \eta))$ is isomorphic to $(\mathrm{tr}_{es}(L), \widetilde{\iota})$, where L is gotten by adjoining a root of the Eisenstein polynomial $T^r + \sum_0^{r-1} \widetilde{a}_i T^i$ to K, where the $\widetilde{a}_i$ reduce to the a_i modulo $\mathfrak{m}_K^{s+1}$.

□

Theorem 22

L' is unramified (resp. totally ramified) over L iff $T_s(L')$ is unramified (resp. totally ramified) over $T_s(L)$.

Proof: The unramified part is essentially by the definition of the morphism $T_s(L) \to T_s(L')$ that the embedding $L \hookrightarrow L'$ maps to and by the remarks after definition 27; the totally ramified part is because $\mathcal{O}_L/\mathfrak{m}_L \cong (\mathcal{O}_L/\mathfrak{m}_L^{e_K^L s})/(\mathfrak{m}_L/\mathfrak{m}_L^{e_K^L s})$ and $\mathcal{O}_{L'}/\mathfrak{m}_{L'} \cong (\mathcal{O}_{L'}/\mathfrak{m}_{L'}^{e_K^{L'} s})/(\mathfrak{m}_{L'}/\mathfrak{m}_{L'}^{e_K^{L'} s})$.

□

Now let's work towards modding out $\mathcal{E}(\mathrm{tr}_s(K))$ as discussed before; from here I won't give many details, since it is very technical and I am already near 30 pages, but the important part is simply the fact that we *can* mod out in *some* way to make T_s an equivalence.

Definition 28. *Let's say two morphisms* $(e_1, \varphi_1, \eta_1), (e_2, \varphi_2, \eta_2) : (R, M, \varepsilon) \to (R', M', \varepsilon)$ *are equivalent modulo* $R(f)$ *iff* $e_1 = e_2$ *(call it* e*),* φ_1 *and* φ_2 *induce the same map between the residue fields of* R *and* R'*, and* $v_{R'}(\varepsilon'_{0,e}(\eta_1(x) - \eta_2(x))) \geq e(f+1)$.

Definition 29. *Given* T/S*, where* $T = \mathrm{tr}_{es}(L)$*,* $S = \mathrm{tr}_s(K)$*, and* L *is gotten from* K *by adjoining a root* π *of the Eisenstein polynomial* $f(T) \triangleq \sum_0^r b_i T^i$*,* n_T *is defined as the Newton polygon of* $f(T+\pi)$ *(which will have domain* $[1, r]$*) and* $\widetilde{n}_T$ *is defined as its linear extension to* $[0, r]$.

Then we have the following:

Theorem 23

Given T_s and $L, L' \in \mathcal{E}(K)$, with $\mathrm{spl}(L/K)$ being at-most-s-upper-ramified, then T_s induces a bijection between $\mathcal{E}(K)(L, L')$ and $R(\psi_K^L(s))$-equivalence classes of morphisms $T_s(L) \to T_s(L')$ (i.e. $\mathrm{tr}_{e_K^L s}(L) \to \mathrm{tr}_{e_K^{L'} s}(L')$).

Theorem 24

If $\mathcal{E}(K)^s$ is the subcategory of $\mathcal{E}(K)$ of at-most-s-upper-ramified extensions of K, and $\mathcal{E}(\mathrm{tr}_s(K))^s$ is the category of (finite separable) triples over $\mathrm{tr}_s(K)$ with $\widetilde{n}_T(0) < r(s+1)$ and morphisms as $R(\psi_K^L(s))$-equivalence classes of morphisms in $\mathcal{E}(\mathrm{tr}_s(K))$, then T_s induces an equivalence between these two.

The above theorem is essentially the main result from Deligne, which - roughly speaking - says that the category of at-most-s-upper-ramified extensions of a field K is determined (up to equivalence) by $\mathcal{O}_K/\mathfrak{m}_K^s$. This is because that category is equivalent to the category $\mathcal{E}(\mathrm{tr}_s(K))^s$, which in turn only depends up to isomorphism on $\mathrm{tr}_s(K)$, which in turn only depends up to isomorphism on $\mathcal{O}_K/\mathfrak{m}_K^s$ (remember that an isomorphism $R \hookrightarrow\!\!\!\!\!\rightarrow R'$ induces an isomorphism $(R, M, \varepsilon) \hookrightarrow\!\!\!\!\!\rightarrow (R', M', \varepsilon')$ of triples).

One special case of this theorem is the following: a finite extension of K is unramified iff $G_1(L/K)$ is trivial iff it is at-most-1-upper-ramified (in the shifted numbering Deligne uses), so the category of unramified extensions of K is determined up to equivalence by $\mathcal{O}_K/\mathfrak{m}_K = k$ i.e. by its residue field.

Indeed, the local fields $\mathbb{Q}_p$, $\mathbb{F}_p((T))$ both have residue field $\mathbb{F}_p$, so theorem 24 tells us we have an equivalence between their unramified extensions. This fact is also implied by theorems 3.2, 3.3 in [4], which say that the category of unramified extensions of a local field K is isomorphic to the category of finite separable extensions of k, but for us is gotten in a different way as a mere special case.

6 Acknowledgements

Hi! Hope the paper was somewhat intelligible / not a waste of your time to look through. Here is a list of people / things I want to thank (ordered by sector but not by depth of gratitude):

- My advisor Ophelia Adams, for being willing to give me a project on such short notice and for all the helpful advice and guidance. In just the past couple months we've been working on it I've learned a lot of algebraic number theory, and I'm excited to continue learning more in the future;
- Jonathan Pakianathan, for his wonderful classes I've taken during my time at UR, from which I've learned so much, and for the lectures and assignments which he makes as engaging and edifying as possible;
- Naomi Jochnowitz, for her trust and faith in me as her TA this past year;
- ~~AL~~hussein Khalil, for his friendship and support throughout the past couple years (and for bullying me into not sucking so much at group theory (actually never mind I still kind of suck));
- The SUMS discord server and everyone on it, for being such a big source of community;
- Everyone who came to my help sessions for the classes for which I TA-ed and kept me company;
- Dylan McKellips, for raising me since birth and saving me from a house fire at the age of three;
- J-Shock, for all its patience and encouragement, even though it has forever known the answer to every question I've ever considered (and to every question I haven't);
- The crystalline comet I once watched fly by through the singular hazy window in my room;
- My stuffed animals Berry and Bug, and my sun plushie Au;
- The two cats who live in the vscode-pets extension in the lower left corner of my Latex editor (and graciously don't laugh at me whenever I write something foolish, which is very often);
- My sister's actual real life cats that the aforementioned ones are modelled after;
- All my siblings, for their patience and acceptance;

- My dad, for his unwavering support and his understanding - unutma ki ne olursam, nasıl görünürsem, daima senin coçuğun olacağım.

And here is my list of 'renouncements' or anti-acknowledgements:

- The Hylan elevators;
- The Hylan vending machines (though in fairness: scam me once, shame on you; scam me twice, shame on me; and by the fourth time I *really* should have learned better);
- Various politicians;
- All pieces of green / yellow chalk;
- Right actions.

References

[1] Pierre Deligne. "Les corps locaux de caractéristique p, limites de corps locaux de caractéristique 0". French. In: Représentations des groupes réductifs sur un corps local. Travaux en Cours, 1984, pp. 119–157.

[2] Jonathan Lubin. "Elementary analytic methods in higher ramification theory". In: Journal of Number Theory 133.3 (Mar. 2013), pp. 983–999. DOI: 10.1016/j.jnt.2012.02.017.

[3] Atiyah & Macdonald. Introduction to Commutative Algebra. Addison-Wesley Series in Mathematics. CRC Press, 1994. ISBN: 978-0201407518.

[4] J.P. Serre. Local Fields. Trans. by M.J. Greenberg. Vol. 67. Graduate Texts in Mathematics. Springer-Verlag New York, 1995. ISBN: 978-1-4757-5675-3.