RAMIFICATION IN ALGEBRAIC NUMBER THEORY AND DYNAMICS

KENZ KALLAL

Abstract. In these notes, we introduce the theory of (nonarchimedian) valued fields and its applications to the local study of extensions of number fields. Particular attention is given to how the ramification of prime ideals in such an extension can be described locally. We will finish this discussion with a detailed proof of a famous result relating to this, namely that the different ideal of an extension of number rings is equal to the annihilator of the module of Kähler differentials corresponding to the same extension. Then, we discuss the somewhat more descriptive notion of the ramification subgroups of a Galois extension of local fields and explain the connection between the ramification subgroups and the ramification of prime ideals. The filtration of ramification groups can also be applied to the automorphism group of the local field $\mathbf{F}_p((X))$. Through the identification between the group of wild automorphisms of $\mathbf{F}_p((X))$ and the Nottingham group, we prove an original result giving a concise criterion for wild automorphisms whose iterates have ramification of a certain type. Finally, we describe the implications that such a computation has on the locations of periodic points of power series applied to nonarchimedian fields of postive characteristic.

Contents

Date: Last updated January 24, 2019.

1. Introduction

Recall the familiar notion of *ramification* in an extension of number fields L/K: A nonzero prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ (which we will refer to as a *prime* in K) is called *ramified* in L if the ideal $\mathfrak{p}\mathcal{O}_L$ in $\mathcal{O}_L$ factors into primes as

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_n^{e_n}$$

where $e_i > 1$ for some i. From this definition arises the classical question

Question 1.1. Given L/K, which primes of K are ramified in L?

Question 1.1 is answered using a reasonably natural construction, namely the *relative discriminant* of L/K, an ideal in $\mathcal{O}_K$ typically denoted $\Delta_{L/K}$. In particular, a prime $\mathfrak{p}$ in K ramifies in L if and only if $\mathfrak{p} | \Delta_{L/K}$.

More generally, for any prime $\mathfrak{q}$ in L containing a prime $\mathfrak{p}$ in K, we must have $\mathfrak{q} \supset \mathfrak{p}\mathcal{O}_L$, i.e. $\mathfrak{q} = \mathfrak{q}_i$ for some i. Conversely, all of the $\mathfrak{q}_i$'s contain $\mathfrak{p}\mathcal{O}_L$ and thus $\mathfrak{p}$. So the information above is a specific case of the *ramification index* $e(\mathfrak{q}|\mathfrak{p}) := e_i$ (namely $\mathfrak{p}$ is ramified in L if $e(\mathfrak{q}|\mathfrak{p}) > 1$ for some prime $\mathfrak{q}$ in L containing $\mathfrak{p}$).

In the other direction, for any prime $\mathfrak{q}$ in L, $\mathfrak{q}$ contains exactly one prime in K, namely $\mathfrak{q} \cap K$. So, we can ask the opposite of Question 1.1:

Question 1.2. Which primes $\mathfrak{q}$ in L are such that $e(\mathfrak{q}|\mathfrak{q} \cap \mathcal{O}_K) > 1$?

The answer to Question 1.2, along with some additional ramification information, is encoded in the *relative different* ideal $\mathfrak{D}_{L/K} \subset \mathcal{O}_L$ (just as before, this ideal is constructed so that a prime $\mathfrak{q}$ in L has $e(\mathfrak{q}|\mathfrak{q} \cap \mathcal{O}_K) > 1$ if and only if $\mathfrak{q} | \mathfrak{D}_{L/K}$), a beautiful construction which will feature prominently in this paper. It is fair to say that "the ramification data of L/K is encoded by $\Delta_{L/K}$ and $\mathfrak{D}_{L/K}$," and these constructions will be the subject of Section 2 of this paper. However, Section 2 will only present the proofs that these constructions answer their respective questions in the case of $K = \mathbf{Q}$ (and $\mathcal{O}_K = \mathbf{Z}$). The general proofs of these statements are postponed for Section 3, when the operation of localization will be introduced. The basic idea is that localizing at a prime $\mathfrak{p}$ in A will preserve ramification of primes lying over $\mathfrak{p}$, while reducing to the case where A is a PID, guaranteeing the existence of a basis for B as an A-module.

The ramification of an extension of number fields at a pair of primes $\mathfrak{q}|\mathfrak{p}$ is just as easily written in terms of the discrete valuations of the two fields corresponding to the two primes. Leveraging the properties of the completions of the two fields with respect to these valuations, this gives rise to the theory of local fields. As a result of this theory, in the setting of the completion, we may assume that $\mathcal{O}_L$ has a power basis over $\mathcal{O}_K$. In Section 3, we use this to prove the desired properties of the discriminant and different, and to prove a fascinating result which strengthens the notion that ramification is an important piece of local geometric data:

Theorem 1.3. *Let L/K be an extension of number fields. Then*

$$\mathfrak{D}_{L/K} = \mathrm{Ann}_{\mathcal{O}_L}(\Omega_{\mathcal{O}_L/\mathcal{O}_K})$$

where $\Omega_{\mathcal{O}_L/\mathcal{O}_K}$ denotes the $\mathcal{O}_L$-module of Kähler differentials of $\mathcal{O}_L$ over $\mathcal{O}_K$.

In Section 6, we consider the setting of a Galois extension L/K of completions of number fields, where we introduce an important filtration of $\mathrm{Gal}(L/K)$ called the sequence of *ramification groups*, and we show that the different ideal of a Galois extension of number fields can be determined locally by the ramification groups.

In Section 7 we will use this to make an analogous definition of ramification for "wild" automorphisms of $\mathbf{F}_p((X))$ (thus introducing the *Nottingham group* of power series tangent to the identity). It turns out that the ramification of iterates of these automorphisms is a piece of essential information about the dynamics of certain classes of power series acting on the unit disc of a nonarchimedian field.

In our case, the *ramification type* of a power series $f(X) = a_1X + a_2X^2 + \cdots \in \mathbf{F}_p[[X]]$ is the series of ramification numbers $i_n(f) := i(f^{\circ p^n})$ for $n \geq 0$, where $i(g)$ denotes the *ramification number* of g ($i(g)$ is determined by where g falls in the sequence of ramification groups). Under certain conditions on b and p, we will classify the power series f (in terms of their coefficients) for which $i_n(f) = b + bp + \cdots bp^n$ for all $n \in \mathbf{N}$. The computation in Section 8 yields an original result (in collaboration with Hudson Kirkpatrick) that achieves this classification under certain hypotheses on p and b.

2. Ramification in Number Fields: Ideal-Theoretic Methods

The material in this section is mostly drawn from Keith Conrad's online handouts [4, 3]. Let L/K be an extension of number fields of degree n. As previously mentioned, the ramification phenomena of $\mathcal{O}_L$ over $\mathcal{O}_K$ are encoded by the relative discriminant and relative different of the extension. It will be the goal of this section to define these constructions, and to prove that they indeed contain the answers to Questions 1.1 and 1.2 in the case where $K = \mathbf{Q}$. Throughout this paper, we will work in a somewhat more general setting: where L/K is an arbitrary finite extension of characteristic zero, let $A \subset K$ be a Dedekind domain with field of fractions K, and B its integral closure in L. Since any element of L can be multiplied by an element of A to get an element of B, we have in general $L = \mathrm{Frac}B$.

2.1. The Relative Discriminant Ideal.

Fortunately, defining the relative discriminant is not much more involved than the absolute discriminant of a number field.

Definition 2.1 (Relative discriminant)**.** The *relative discriminant* $\Delta_{B/A}$ is the ideal in A generated by the discriminants

$$\mathrm{disc}_{L/K}(b_1, \ldots, b_n) := \det(\mathrm{Tr}_{L/K}(b_i b_j))$$

where $(b_1, \ldots, b_n)$ runs over all K-bases of L contained in B.

One can check that in the case $A = \mathbf{Z}$ and $B = \mathbf{Q}$, the relative discriminant is the ideal in $\mathbf{Z}$ generated by the discriminant of any integral basis for $\mathcal{O}_L$. In this case, we can prove the desired result directly as in [4, Theorem 1.3]

Proposition 2.2. *Let $p \in \mathbf{Z}$ be a prime. Then (p) ramifies in L if and only if $p|\Delta_{\mathcal{O}_L/\mathbf{Z}}$.*

Proof. Recall that $\mathcal{O}_L$ has an integral basis $\omega_1, \ldots, \omega_n \in \mathcal{O}_L$. Suppose that $p\mathcal{O}_L$ factorizes into prime ideals of $\mathcal{O}_L$ as

$$p\mathcal{O}_L = \mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_m^{e_m}.$$

Recall, too, that

$$\Delta_{\mathcal{O}_L/\mathbf{Z}} = \mathrm{disc}_{L/\mathbf{Q}}(\omega_1, \ldots, \omega_n)\mathbf{Z} = \det((\mathrm{Tr}_{\mathcal{O}_L/\mathbf{Z}}(\omega_i\omega_j))_{ij})\mathbf{Z}$$

Where the trace is taken with respect to $\{\omega_i\}$. Note that $\mathcal{O}_L/p\mathcal{O}_L$ has an $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$-basis given by the $\overline{\omega_i}$'s, where $\overline{\omega_i}$ denotes the reduction of ω_i modulo $p\mathcal{O}_L$. Therefore,

$$\mathrm{Tr}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{\omega_i\omega_j}) = \mathrm{Tr}_{\mathcal{O}_L/\mathbf{Z}}(\omega_i\omega_j) \pmod p$$

since the $n \times n$ matrix given by $x \mapsto \omega_i\omega_j x$ with respect to the $\mathbf{Z}$-basis $\{\omega_i\}$ for $\mathcal{O}_L$, when reduced modulo p, is the $n \times n$ matrix given by $x \mapsto \overline{\omega_i\omega_j}x$ with respect to the $\mathbf{F}_p$-basis $\{\overline{\omega_i}\}$ for $\mathcal{O}_L/p\mathcal{O}_L$. Taking determinants, this means that

$$\mathrm{disc}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{\omega_1}, \ldots, \overline{\omega_n}) = \Delta_{\mathcal{O}_L/\mathbf{Z}} \pmod p.$$

Thus, $p|\Delta_{\mathcal{O}_L/\mathbf{Z}}$ if and only if $\mathrm{disc}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{\omega_1}, \ldots, \overline{\omega_n}) = 0$

By the Chinese Remainder Theorem, we have

$$\mathcal{O}_L/p\mathcal{O}_L = \mathcal{O}_L/\mathfrak{q}_1^{e_1} \times \cdots \times \mathcal{O}_L/\mathfrak{q}_m^{e_m}.$$

Choose bases B_i for each $\mathcal{O}_L/\mathfrak{q}_i{}^{e_i}$, and let B be the basis of their product (which we know from above is equal to $\mathcal{O}_L/p\mathcal{O}_L$) given by the concatenation of the B_i's. It is easy to see from the definition that

$$\mathrm{disc}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(B) = \prod_{i=1}^{m} \mathrm{disc}_{(\mathcal{O}_L/\mathfrak{q}_i{}^{e_i})/\mathbf{F}_p}(B_i).$$

Since the discriminants of bases are all related by unit factors, the left hand side is zero if and only if $\mathrm{disc}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{\omega_1}, \ldots, \overline{\omega_n}) = 0$, i.e. if and only if $p|\Delta_{\mathcal{O}_L/\mathbf{Z}}$. So, it suffices to determine whether $\mathrm{disc}_{(\mathcal{O}_L/\mathfrak{q}_i{}^{e_i})/\mathbf{F}_p}(B_i) = 0$ given e_i (for which we can replace B_i with any arbitrarily chosen basis).

In particular, suppose $e_i > 1$ for some i (i.e. p is ramified in L). Then we can choose an $x \in \mathfrak{q}_i \setminus \mathfrak{q}_i^{e_i}$. Letting $\overline{x}$ be the modulo-$\mathfrak{q}_i^{e_i}$ reduction of x, we are thus guaranteed that $\overline{x} \neq 0$ but that $\overline{x}^{e_i} = 0$ in $\mathcal{O}_L/\mathfrak{q}_i^{e_i}$, i.e. $\overline{x}$ is nilpotent and thus its multiples all have trace zero. Extending $\overline{x}$ to a basis of $\mathcal{O}_L/\mathfrak{q}_i^{e_i}$, this makes it immediate that the discriminant with respect to the extension of rings $(\mathcal{O}_L/\mathfrak{q}_i^{e_i})/(\mathbf{F}_p)$ of any basis of $\mathcal{O}_L/\mathfrak{q}_i^{e_i}$ is zero. Therefore, $p|\Delta_{\mathcal{O}_L/\mathbf{Z}}$ if p ramifies in L.

Conversely, suppose $e_i = 1$ for all i (i.e. p is unramified in L). The previous analysis does not apply because $\mathfrak{q}_i \setminus \mathfrak{q}_i^{e_i}$ is then empty. Instead, (since $\mathfrak{q}_i$ is a prime and is therefore maximal as opposed to its higher powers), we know that $\mathcal{O}_L/\mathfrak{q}_i$ is a (finite) field. Therefore, the extension of rings over which we wish to compute the discriminant of a basis, namely $(\mathcal{O}_L/\mathfrak{q}_i)/\mathbf{F}_p$, is actually an extension of finite fields. But it is a well-known property of discriminants over field extensions that the discriminant of a basis is always nonzero (see [23, Theorem 7]). Thus, $p \not| \Delta_{\mathcal{O}_L/\mathbf{Z}}$, as desired. □

The proof that prime $\mathfrak{p}$ in A ramifies in B if and only if $\mathfrak{p}|\Delta_{B/A}$ will be postponed until the next section, where localizing will allow us to reduce to the case where A is a PID and thus B has a finite A-basis.

2.2. **The Relative Different Ideal.** In a somewhat more complicated way than the relative discriminant, the relative different is also based on the trace form $\langle x, y\rangle := \mathrm{Tr}_{L/K}(xy)$, which is a symmetric nondegenerate bilinear form on L. To show that it is nondegenerate, it suffices to show that the trace does not vanish on

L (if $\mathrm{Tr}_{L/K}(x) \neq 0$ then $\langle y, x/y\rangle \neq 0$ for every nonzero $y \in L$ which means the map $y \mapsto \langle y, -\rangle$ is injective and thus bijective). Define

$$B^\vee := \{x \in L : \mathrm{Tr}_{L/K}(xy) \in A \text{ for all } y \in B\}.$$

Under the identification $L \cong L^\vee$ given by the trace form, $B^\vee$ is identified with a fractional ideal of B that contains B, so its inverse is an ideal in B. This is how we define the relative different:

Definition 2.3. The *relative different* $\mathfrak{D}_{B/A}$ of the ring extension B/A is the ideal $(B^\vee)^{-1} \subset B$. When the choice of A and B is obvious, i.e. $A = \mathcal{O}_K$ and $B = \mathcal{O}_L$, we sometimes denote the relative different by $\mathfrak{D}_{L/K}$.

The general proof of the desired property of the relative different, like that of the discriminant, will be postponed until local methods are available. For now, we present a proof in the case of $K = \mathbf{Q}, A = \mathbf{Z}$.

Proposition 2.4. *Let $\mathfrak{p}$ be an ideal in $\mathcal{O}_L$. Then $e(\mathfrak{p}, \mathfrak{p} \cap \mathbf{Z}) > 1$ if and only if $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$.*

Proof. Just as in the proof of Proposition 2.2, this proof will rely on the fact that $\mathcal{O}_L$ has an integral basis which is also a $\mathbf{Q}$-basis for L, which allows us to write the field trace as the trace of the multiplication map on the $\mathbf{Z}$-module $\mathcal{O}_L$ with respect to a suitable basis.

In this special case, it is easily seen that the definition of $\mathcal{O}_L^\vee$ also means that for any fractional ideal $\mathfrak{a} \subset \mathcal{O}_L$, every element of $\mathfrak{a}$ has integral trace if and only if $\mathfrak{a}$ is contained in $\mathcal{O}_L^\vee$ (in other words, "$\mathcal{O}_L^\vee$ is the largest fractional ideal of $\mathcal{O}_L$ such that all of its elements have integral trace"). So $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$ if and only if $\mathfrak{p} \supset \mathfrak{D}_{L/\mathbf{Q}}$, i.e. if and only if $\mathfrak{p}^{-1} \subset \mathcal{O}_L^\vee$, which is therefore equivalent to all of the elements of $\mathfrak{p}^{-1}$ having integral trace.

Suppose that $e(\mathfrak{p}, \mathfrak{p} \cap \mathbf{Z}) > 1$. Letting $\mathfrak{p} \cap \mathbf{Z} = p\mathbf{Z}$, this means $p\mathcal{O}_L = \mathfrak{p}\mathfrak{b}$, where $\mathfrak{b}$ is an ideal in $\mathcal{O}_L$ which is divisible by $\mathfrak{p}$. This means

$$\mathfrak{p}^{-1} = \frac{1}{p}\mathfrak{b}.$$

By the previous discussion, it follows that $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$ is equivalent to $\mathrm{Tr}_{L/\mathbf{Q}}(b) \in p\mathbf{Z}$ for all $b \in \mathfrak{b}$.

Rewriting the trace as the trace of the multiplication map on the $\mathbf{Z}$-module $\mathcal{O}_L$, we have for any $b \in \mathfrak{b}$,

$$\mathrm{Tr}_{L/\mathbf{Q}}(b) = \mathrm{Tr}_{\mathcal{O}_L/\mathbf{Z}}(b).$$

Taking the quotient of both rings by (p), we actually have

$$\mathrm{Tr}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\bar{b}) = \mathrm{Tr}_{L/\mathbf{Q}}(b) \pmod p$$

where $\bar{b}$ denotes the reduction of b modulo $p\mathcal{O}_L$. But $p\mathcal{O}_L|\mathfrak{b}^2$, which means that $b^2 \in \mathfrak{b}^2 \in p\mathcal{O}_L$, i.e. multiplication by $\bar{b}$ is a nilpotent operator on the $\mathbf{F}_p$-vector space $\mathcal{O}_L/p\mathcal{O}_L$. Hence, it has zero trace and $\mathrm{Tr}_{L/\mathbf{Q}}(b) \in p\mathbf{Z}$ for all $b \in \mathfrak{b}$, which means $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$.

Finally, suppose $e(\mathfrak{p}, \mathfrak{p} \cap \mathbf{Z}) = 1$. Then we can write $p\mathcal{O}_L = \mathfrak{p}\mathfrak{b}$, where $\mathfrak{b}$ is not divisible by $\mathfrak{p}$. As before, $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$ is still equivalent to $\mathrm{Tr}_{L/\mathbf{Q}}(b) \in p\mathbf{Z}$ for all $b \in \mathfrak{b}$. So, we still want to consider the trace of an arbitrary element $b \in \mathfrak{b}$ (and we still

know that this trace is congruent modulo p to $\mathrm{Tr}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{b})$). By the Chinese Remainder Theorem,

$$\mathcal{O}_L/p\mathcal{O}_L \cong \mathcal{O}_L/\mathfrak{p} \times \mathcal{O}_L/\mathfrak{b}.$$

So

$$\mathrm{Tr}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{b}) = \mathrm{Tr}_{(\mathcal{O}_L/\mathfrak{p})/\mathbf{F}_p}(\overline{b}) + \mathrm{Tr}_{(\mathcal{O}_L/\mathfrak{b})/\mathbf{F}_p}(\overline{b}),$$

where $\overline{b}$ denotes the reduction of b modulo the appropriate ideal. Of course, $\overline{b} = 0$ modulo $\mathfrak{b}$, so multiplication by $\mathfrak{b}$ in $\mathcal{O}_L/\mathfrak{b}$ is the zero map, which has zero trace. Hence we can ignore the last term and write

$$\mathrm{Tr}_{(\mathcal{O}_L/p\mathcal{O}_L)/\mathbf{F}_p}(\overline{b}) = \mathrm{Tr}_{(\mathcal{O}_L/\mathfrak{p})/\mathbf{F}_p}(\overline{b}).$$

But $(\mathcal{O}_L/\mathfrak{p})/\mathbf{F}_p$ is an extension of finite fields, so there must be a $b \in \mathfrak{b}$ such that $\mathrm{Tr}_{(\mathcal{O}_L/\mathfrak{p})/\mathbf{F}_p}(\overline{b}) \neq 0$. In particular, this means that $\mathrm{Tr}_{L/\mathbf{Q}}(b) \notin p\mathbf{Z}$, so $\mathfrak{p} \not| \mathfrak{D}_{L/\mathbf{Q}}$, as desired. □

In fact, the proof above can be extended to give a simple explicit description (in most cases) of $e(\mathfrak{p}|\mathfrak{p} \cap \mathbf{Z})$ based on the number of times $\mathfrak{p}|\mathfrak{D}_{L/\mathbf{Q}}$.

3. Valuation Theory

In this section, we introduce a special case the powerful theory of local fields, which is a generalization of the theory of the p-adic numbers; this will be drawn from two classic references on the topic, namely Serre's book *Local Fields* [29] and Neukirch's *Algebraic Number Theory* [25]. We begin the development of this theory with a chain of reasoning analogous to the definition of the p-adic numbers. In particular, let R be Dedekind domain with field of fractions K (a generalization of the principal ideal domain $\mathbf{Z}$ with field of fractions $\mathbf{Q}$ from which we might produce the p-adics). We can associate to each prime ideal $\mathfrak{p} \subset R$ the $\mathfrak{p}$-*adic valuation*

$$\nu_\mathfrak{p} : K \to \mathbf{R} \cap \{\infty\}$$

defined by

$$\nu_\mathfrak{p}(x) = e_\mathfrak{p}$$

for $x \in R$, where $e_\mathfrak{p}$ denotes the exponent of $\mathfrak{p}$ in the factorization of the principal ideal (x) into primes in R, i.e. so that (x) factorizes as

$$(x) = \prod_\mathfrak{p} \mathfrak{p}^{\nu_\mathfrak{p}(x)}.$$

This extends to all of K in the natural way by taking

$$\nu_\mathfrak{p}\left(\frac{a}{b}\right) = \nu_\mathfrak{p}(a) - \nu_\mathfrak{p}(a).$$

Also, we define $\nu_\mathfrak{p}(0) = \infty$. It is easy to see that $\nu_\mathfrak{p}$ satisfies some nice properties:

i) For $x \in K$, $\nu_\mathfrak{p}(x) = \infty$ if and only if $x = 0$.
ii) For $x, y \in K$, $\nu_\mathfrak{p}(xy) = \nu_\mathfrak{p}(x) + \nu_\mathfrak{p}(y)$
iii) For $x, y \in K$, $\nu_\mathfrak{p}(x + y) \geq \min(\nu_\mathfrak{p}(x), \nu_\mathfrak{p}(y))$.

A map satisfying these properties is called a *nonarchimedian valuation* on K. It is our objective to make K into a metric space (e.g. to take its completion, still in analogy to the definition of $\mathbf{Q}_p$), so we transform $\nu_\mathfrak{p}$ in the following way: Let

$$|\cdot| : K \to \mathbf{R}_{\geq 0}$$

be defined to be

$$|x| := q^{-\nu_\mathfrak{p}(x)}$$

for some fixed real number $q > 1$. The three properties that make $\nu_\mathfrak{p}$ a nonarchimedian valuation correspond to the following three properties for $|\cdot|$:

i) $|x| = 0$ if and only if $x = 0$.
ii) $|xy| = |x||y|$.
iii) $|x + y| \leq \max(|x|, |y|)$.

Since it satisfies these properties, $|\cdot|$ is a *nonarchimedian absolute value* on K. Note that the third property is a stronger version of the triangle inequality, and thus $(K, |\cdot|)$ is a metric space and we can take its completion $\hat{K}$. In fact, $\hat{K}$ has an obvious field structure, obtained from taking the ring of Cauchy sequences and modding out by the maximal ideal of Cauchy sequences which converge to zero.

The fact that $|\cdot|$ satisfies the third property is the reason why ν and $|\cdot|$ are called "nonarchimedian." If a map satisfies the triangle inequality but not (iii), it is instead called an *archimedian absolute value*, and the corresponding valuation is called an *archimedian valuation*. However, in this case, $\hat{K}$ has a very restricted theory:

Theorem 3.1 (Ostrowski)**.** *Let K be a field which is complete under some archimedian absolute value $|\cdot|$. Then $(K, |\cdot|)$ is isomorphic as a topological space to $\mathbf{R}$ or $\mathbf{C}$ with the usual topologies.*

Proof. See Neukirch [25, Ch. II, Theorem 4.2]. □

This result justifies restricting the general study of valued fields to the nonarchimedian case. In the case where K is a number field (which will obviously be the case of interest to us), Ostrowski proved another result on the general theory of valuations:

Theorem 3.2 (Ostrowski)**.** *Let K be a number field. Then all nontrivial nonarchimedian absolute values on K must be of the form $\nu_\mathfrak{p}$ for some prime $\mathfrak{p}$ in $\mathcal{O}_K$. All archimedian absolute values on K must correspond to the absolute value on an embedding of K into $\mathbf{R}$ or $\mathbf{C}$.*

Proof. See [5, Theorem 3] □

Thus, we can restrict our attention in this paper to only the valuations of the form $\nu_\mathfrak{p}$, though much of the following discussion features in the general theory of valued fields as well.

Letting $|\cdot| : K \to \mathbf{R}_{\geq 0}$ be the absolute value corresponding to $\nu_\mathfrak{p} : K \to \mathbf{R} \cup \{\infty\}$ as before, we can associate to $\mathfrak{p}$ (still in analogy to the p-adics) the ring

$$\mathcal{O}_\mathfrak{p} := \{x \in K : \nu_\mathfrak{p}(x) \geq 0\},$$

which contains all $x \in K$ such that the fractional ideal $x\mathcal{O}_K$ has prime factorization with a nonnegative exponent on $\mathfrak{p}$. Since $\nu_\mathfrak{p}$ only takes into account the behavior at the prime $\mathfrak{p}$, the other primes of $\mathcal{O}_K$ may still have negative exponents in the factorization of (x), thus $\mathcal{O}_\mathfrak{p}$ is strictly larger than $\mathcal{O}_K$ (in general it is called the *valuation ring* of K). From this intuition, we can finally observe the geometric notion of localization at a prime ideal.

Lemma 3.3. *$\mathcal{O}_\mathfrak{p}$ is equal to the localization*

$$S^{-1}\mathcal{O}_K := \left\{\frac{a}{s} \in K : a \in \mathcal{O}_K, s \in S\right\},$$

where $S = \mathcal{O}_K \setminus \mathfrak{p}$. It is a DVR, and its uniformizing parameter can be taken to be any $\pi \in \mathfrak{p}\mathcal{O}_\mathfrak{p} - \mathfrak{p}^2\mathcal{O}_\mathfrak{p}$

Proof. The main content of the lemma is actually in the second half, which almost immediately follows from the following observation: First, we show that $S^{-1}\mathcal{O}_K$ is a Dedekind domain (this actually holds for any multiplicative set $S \subset \mathcal{O}_K$ and follows from the fact that $\mathcal{O}_K$ is a Dedekind domain). Since the ideals of $S^{-1}\mathcal{O}_K$ are localizations of ideals in $\mathcal{O}_K$, which are finitely generated over $\mathcal{O}_K$, we know that the ideals of $S^{-1}\mathcal{O}_K$ are finitely generated (an ideal $S^{-1}I$ is generated over $S^{-1}\mathcal{O}_K$ by the same generating set as the ideal $I \subset \mathcal{O}_K$). So, $S^{-1}\mathcal{O}_K$ is Noetherian. The prime ideals of $S^{-1}\mathcal{O}_K$ are (by localization) in inclusion-preserving bijection with the prime ideals of $\mathcal{O}_K$ that do not intersect S. If a prime $\mathfrak{q} = S^{-1}\mathfrak{p}$ is not maximal, then it is properly contained in a maximal ideal, which must be the localization $S^{-1}\mathfrak{p}'$ of some ideal $\mathfrak{p}' \subset \mathcal{O}_K$, which must properly contain $\mathfrak{p}$ (by the above bijection). This contradicts the fact that prime ideals in $\mathcal{O}_K$ are maximal, so indeed all prime ideals in $S^{-1}\mathcal{O}_K$ are maximal. Finally, suppose that $a \in K$ is integral over $S^{-1}\mathcal{O}_K$. Then a is the root of some monic polynomial f over $S^{-1}\mathcal{O}_K$. Multiplying by $s \in S$ given by the $\deg f$ power of the product of all the denominators of the coefficients of f, we get that sa is integral over $\mathcal{O}_K$, so $sa \in \mathcal{O}_K$ (since $\mathcal{O}_K$ is integrally closed in K). Thus, $a \in S^{-1}\mathcal{O}_K$, i.e. $S^{-1}\mathcal{O}_K$ is integrally closed in its field of fractions. All of this amounts to the observation that $S^{-1}\mathcal{O}_K$ is a Dedekind domain.

Since the prime (i.e. maximal) ideals of $S^{-1}\mathcal{O}_K$ arise as localizations of prime ideals of $\mathcal{O}_K$ contained in $\mathfrak{p}$, of which there is exactly one (namely $\mathfrak{p}$), we know that $S^{-1}\mathcal{O}_K$ is a local ring with maximal ideal $\mathfrak{m} = S^{-1}\mathfrak{p}$. We can take $\pi \in \mathfrak{m} - \mathfrak{m}^2$ (these are guaranteed to be distinct ideals since $S^{-1}\mathcal{O}_K$ is a Dedekind domain), and we know that $(\pi) = \mathfrak{m}$ (it must be a power of $\mathfrak{m}$ by unique factorization of ideals, it cannot be any higher power of $\mathfrak{m}$ since π is not in any higher power of $\mathfrak{m}$). Thus, the ideals of $S^{-1}\mathcal{O}_K$ are all equal to $\mathfrak{m}^n = (\pi^n)$ for some $n \geq 0$, which means that $S^{-1}\mathcal{O}_K$ is a DVR with uniformizing parameter π.

K is the field of fractions of $S^{-1}\mathcal{O}_K \supseteq \mathcal{O}_K$, so every element of K is of the form $u \cdot \pi^n$ for $n \in \mathbf{Z}$, where u is a quotient of units in $S^{-1}\mathcal{O}_K$, i.e. $\nu_\mathfrak{p}(u) = 0$ (one can show that the units in $S^{-1}\mathcal{O}_K$ are exactly the elements of the form a/s where neither a nor s is in $\mathfrak{p}$). Then, $\mathcal{O}_\mathfrak{p}$ is just the set of all $u \cdot \pi^n$ such that $n \geq 0$, which coincides with our characterization of $S^{-1}\mathcal{O}_K$, as desired. □

This lemma and its proof are analogous to the fact that the valuation ring of $\mathbf{Q}$ under the p-adic valuation is the localization of $\mathbf{Z}$ at p. Of course, we had to be more careful to account for the fact that we cannot just write fractions in lowest form, since $\mathcal{O}_K$ is not necessarily a PID. Everything we just did still holds in the more general setting in which A is an arbitrary Dedekind domain with field of fractions K and $\mathfrak{p}$ is a prime in A. The valuation ring of K with respect to $\nu_\mathfrak{p}$ is the localization of A at $\mathfrak{p}$

The general fact that the valuation ring of a valued field is a DVR holds as long as the valuation is *discrete*, i.e. the image of the valuation on $K^\times$ has a smallest nonzero element. Any element of the valuation ring with that smallest valuation is then a uniformizing parameter.

4. Applications of Localization

4.1. **Basic Results.** Without considering more about the valuation associated with $\mathfrak{p}$, $\mathcal{O}_\mathfrak{p}$ is already a powerful gadget for studying the factors of $\mathfrak{p}$ in the decomposition of ideals in $\mathcal{O}_K$. For this reason, we devote this section to a detour explaining how to use this tool to prove the generalized versions of Propositions 2.2 and 2.4 to the relative different and discriminant.

Readers familiar with algebraic geometry will recognize the localization of $\mathcal{O}_K$ at a prime ideal as being analogous to the local ring of an affine variety at a point (see Fulton's *Algebraic Curves*, [8, Ch. 3]), and the fact that this is a DVR as some notion of non-singularity of the variety at a point. In fact, we have seen that $|\mathrm{Spec}(\mathcal{O}_\mathfrak{p})| = 1$, so the affine scheme $\mathrm{Spec}(\mathcal{O}_\mathfrak{p})$ is homeomorphic to the closed point $\{\mathfrak{p}\} \subset \mathrm{Spec}(\mathcal{O}_K)$ (recall that all primes are maximal, i.e. $\mathcal{O}_K$ has Krull dimension 1). So, we can write

$$\mathrm{Spec}(\mathcal{O}_K) = \cup_{\mathfrak{p} \in \mathrm{Spec}\mathcal{O}_K} \mathrm{Spec}(\mathcal{O}_\mathfrak{p}),$$

and we expect to be able to recover $\mathcal{O}_K$ from its localizations at all the primes. In fact, we can recover any ideal of $\mathcal{O}_K$. Note that the localizations of $\mathcal{O}_K$ are all embedded in its field of fractions K, so we can take their intersections, arriving at

Lemma 4.1. *For any ideal $I \subset \mathcal{O}_K$, we have*

$$I = \cap_\mathfrak{p} I\mathcal{O}_\mathfrak{p},$$

where the intersection is over every prime $\mathfrak{p}$ in $\mathcal{O}_K$.

Proof. $I \subset \mathcal{O}_\mathfrak{p}$ for all $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$, so the inclusion $I \subset \cap_\mathfrak{p} I\mathcal{O}_\mathfrak{p}$ is immediate.

For the opposite inclusion, let $x \notin I$ be an element of K. Then we can consider the ideal of possible denominators for x, namely

$$J = \{a \in \mathcal{O}_K : xa \in I\}.$$

If $1 \in J$, then $x \in I$, a contradiction. So J is a proper ideal, and it is contained in some maximal (i.e. prime) ideal $\mathfrak{m} \subset \mathcal{O}_K$. Then, if $x \in I\mathcal{O}_\mathfrak{m}$ we know that $x = a/s$ where $a \in I$ and $s \in \mathcal{O}_K - \mathfrak{m} \subseteq \mathcal{O}_K - J$, which means $sx \in \mathcal{O}_K$ where $s \in \mathcal{O}_K - J$, a contradiction. It follows that

$$x \notin I \implies x \notin \cap_\mathfrak{p} I\mathcal{O}_\mathfrak{p},$$

and we have the desired equality. □

This gives a useful slogan that will reduce many problems to the case of a DVR: in order to show an equality of ideals, it suffices to show the equality for their localizations at each prime. In particular, this brings us much closer to the situation in Section 2, in which the bottom ring is a PID. It is a general fact from algebra that this situation yields a basis, as was exploited in the proofs of Propositions 2.4 and 2.2.

Proposition 4.2. *Let L/K be a finite extension of characteristic zero, let $A \subset K$ be a Dedekind domain with field of fractions K, and B be its integral closure in L. Then B is a finitely-generated A-module.*

Proof. Recall from the discussion on the different ideal that we have a nondegenerate pairing

$$\langle \cdot, \cdot \rangle : L \times L \to K$$

given by

$$\langle x, y\rangle = \mathrm{Tr}_{L/K}(xy).$$

Let $x_1, \ldots, x_n$ be a basis for L over K. Clearing denominators in the minimal polynomial for x_i, if the least common denominator of the coefficients is $a_i \in A$, we can observe that $a_i x_i \in L$ is integral over A, i.e. $a_i x_i \in B$. So, we have a free A-module

$$\mathrm{Span}_A(x_1', \ldots, x_n') \subset B.$$

Taking duals, and using the identification $V \cong V^\vee$ given by the nondegenerate pairing, we get inclusions

$$\mathrm{Span}_A(x_1', \ldots, x_n') \subset B \subset B^\vee \subset \mathrm{Span}_A(x_1', \ldots, x_n')^\vee.$$

But it is easily seen that $\mathrm{Span}_A(x_1', \ldots, x_n')^\vee$ is spanned over A by the dual basis to $x_1', \ldots, x_n'$, so it is a free A-module of finite rank. Since A is Noetherian (it is Dedekind), B and $B^\vee$ are submodule of a Noetherian A-module (direct sums of Noetherian modules are Noetherian), hence B is finitely generated as an A-module. This is also the reason why $B^\vee$ is finitely generated and thus is a fractional ideal of B. □

From now on, A, B, L, K are assumed to be as in the hypotheses of Proposition 4.2.

Corollary 4.3. *If A is a PID, then B is a free A-module of finite rank*

Proof. This follows from the previous proposition and the structure theorem for modules over PIDs. □

This makes localization useful in a concrete way: it lets us use, as in the proof of Propositions 2.4 and 2.2, an A-basis for B.

4.2. **Localization and Ramification.** The machinery we have developed so far allows us to prove an important relation between the relative different and discriminant in a general setting; see [25, Ch. III, Theorem 2.9].

Theorem 4.4. *Let L/K be an extension of number fields, $A \subset K$ a Dedekind domain with field of fractions K, and B its integral closure in K. Then*

$$\Delta_{B/A} = N_{L/K}(\mathfrak{D}_{B/A}),$$

where $N_{L/K}(I)$ denotes the ideal in A generated by the norms of elements in B (a.k.a. the "relative norm ideal").

Proof. Suppose that A is a PID. Then by the corollary, B has an A-basis $\omega_1, \ldots, \omega_n$. What's more, the fractional ideal $\mathfrak{D}_{B/A}^{-1}$ in B has an A-basis $\omega_1^\vee, \ldots, \omega_n^\vee$ dual to $\omega_1, \ldots, \omega_n$ under the trace form. Since it is a Dedekind domain with finitely many prime ideals, B is a principal ideal domain, so $\mathfrak{D}_{B/A}^{-1}$ is generated by some $d \in K$ as a fractional ideal. In summary, the situation is that

$$\mathfrak{D}_{B/A}^{-1} = bB = A\omega_1 b \oplus \cdots A\omega_n b = A\omega_1^\vee \oplus \cdots \oplus A\omega_n^\vee.$$

Computing discriminants, it follows that

$$\mathrm{disc}_{L/K}(\omega_1 b, \ldots, \omega_n b) = \mathrm{disc}_{L/K}(\omega_1^\vee, \ldots, \omega_n^\vee),$$

and we obtain from the definition of the discriminant

$$\mathrm{disc}_{L/K}(\omega_1 b, \ldots, \omega_n b)A = N_{L/K}(b)^2 \mathrm{disc}_{L/K}(\omega_1, \ldots, \omega_n)A = N_{L/K}(b)^2 \Delta_{B/A}.$$

Here we used the fact that $\omega_1, \ldots, \omega_n$ is an A-basis for B to write the relative discriminant as the principal ideal generated by the discriminant of the basis. Since $\mathfrak{D}_{B/A}^{-1}$ is principally generated by b, the norm of b is just the relative norm of this fractional ideal, and we can write

$$\mathrm{disc}_{L/K}(\omega_1^\vee, \ldots, \omega_n^\vee)A = N_{L/K}(\mathfrak{D}_{B/A}^{-1})^2 \Delta_{B/A}.$$

Finally, the same analysis that allows us to write the discriminant in terms of the trace form, and the definition of the dual basis gives

$$\mathrm{disc}_{L/K}(\omega_1^\vee, \ldots, \omega_n^\vee)\mathrm{disc}_{L/K}(\omega_1, \ldots, \omega_n) = 1,$$

so we can multiply both sides of the previous equation by $\mathrm{disc}_{L/K}(\omega_1, \ldots, \omega_n) = \Delta_{B/A}$ and rearrange to get

$$N_{L/K}(\mathfrak{D}_{B/A})^2 = \Delta_{B/A}^2,$$

which implies the desired result by unique factorization.

Now, we return to the original case of $A = \mathcal{O}_K$ and $B = \mathcal{O}_L$. By the previous remarks, it suffices to show that

$$S^{-1}\Delta_{\mathcal{O}_L/\mathcal{O}_K} = S^{-1}N_{L/K}(\mathfrak{D}_{\mathcal{O}_L/\mathcal{O}_K})$$

for all S of the form $\mathcal{O}_K \setminus \mathfrak{p}$ where $\mathfrak{p}$ is a prime in $\mathcal{O}_K$. Of course,

$$S^{-1}\Delta_{\mathcal{O}_L/\mathcal{O}_K} = \Delta_{S^{-1}\mathcal{O}_L/S^{-1}\mathcal{O}_K}$$

(both inclusions follow by looking at the generators) and

$$S^{-1}\mathfrak{D}_{B/A} = \mathfrak{D}_{S^{-1}B/S^{-1}A}$$

which follows from the fact that localization is compatible with duals and inverses. Thus, it suffices to show

$$\Delta_{S^{-1}\mathcal{O}_L/S^{-1}\mathcal{O}_K} = N_{L/K}(\mathfrak{D}_{S^{-1}\mathcal{O}_L/S^{-1}\mathcal{O}_K})$$

since localization is obviously compatible with taking relative norms. By Lemma 3.3, $S^{-1}\mathcal{O}_K$ is a DVR and thus a PID. This is what we showed in the special case above, so we are done. □

Of course, we really want to be able to prove statements about divisibility of ideals by $\mathfrak{p}$ and ramification. First, we recall a basic fact about localizations of ideals [15, Ch. 1, Proposition 16]

Proposition 4.5. *Let S be a multiplicative set in A. The map $I \mapsto S^{-1}I$ is a homomorphism from the group of fractional ideals of A to the group of fractional ideals of $S^{-1}A$.*

Though we could make use of more general properties of localization, e.g. its inclusion-preserving property on ideals, or even Proposition 4.1, it ends up being easiest to apply the above principle and unique factorization of ideals in Dedekind domains.

Lemma 4.6. *Let I be an ideal in A and let $n \in \mathbf{N}$. Then I is divisible by $\mathfrak{p}^n$ if and only if $I_\mathfrak{p}$ is divisible by $\mathfrak{p}^n A_\mathfrak{p}$.*

Proof. Since A is a Dedekind domain, we know I factors as

$$I = \prod_{\mathfrak{q} \in \mathrm{Spec}(A)} \mathfrak{q}^{\nu_\mathfrak{q}},$$

and by the proposition,

$$I_{\mathfrak{p}} = \prod_{\mathfrak{q} \in \text{Spec}(A)} \mathfrak{q}^{\nu_{\mathfrak{q}}} A_{\mathfrak{p}} = \mathfrak{p}^{\nu_{\mathfrak{p}}} A_{\mathfrak{p}}$$

from which the statement of the lemma follows. □

Of course, $\mathcal{O}_p$ is a DVR and its ideals are all of the form $\mathfrak{p}^n \mathcal{O}_{\mathfrak{p}} = (\pi^n)$ for any choice of uniformizing parameter $\pi \in \mathfrak{p} - \mathfrak{p}^2$ so the lemma just means that divisibility by $\mathfrak{p}^n$ is equivalent to the localization being divisible by π^n in $\mathcal{O}_{\mathfrak{p}}$.

Thus, divisibility by powers of $\mathfrak{p}$ (i.e. ramification) and equality of ideals are both compatible with localization at $\mathfrak{p}$ in some formal sense. Abusing this machinery, we can reduce any statement about divisibility or equalities of ideals in $\mathcal{O}_K$ to the same statement in $\mathcal{O}_{\mathfrak{p}}$, which has a much simpler ideal theory. On the other hand, we want to be able to use this machinery to generalize Propositions 2.2 and 2.4 to the general case of the extension of Dedekind domains B/A with fields of fractions L/K such that B is the integral closure of A in L (whereas before we assumed $K = \mathbf{Q}$).

To do this, we would want to examine the ramification at the pairs of primes $\mathfrak{q}|\mathfrak{p}$ in $\mathcal{O}_L$ and $\mathcal{O}_K$. So, it would make sense to localize $\mathcal{O}_K$ at $\mathfrak{p}$ and to localize $\mathcal{O}_L$ at $\mathfrak{q}$. While this localization might preserve the ramification $e(\mathfrak{q}|\mathfrak{p})$ and divisibility of ideals by powers of primes, we don't expect it to be as useful because it does not necessarily preserve the fact that B is integral over A. Instead, we consider one prime of A at a time, and localize both rings at $S = A - \mathfrak{p}$.

Lemma 4.7. *Let $\mathfrak{p}$ be a prime in A, and consider the multiplicative set $S = A \setminus \mathfrak{p}$. If $\mathfrak{q}$ is a prime in B lying over $\mathfrak{p}$, then*

$$e(\mathfrak{q}|\mathfrak{p}) = e(S^{-1}\mathfrak{q}|S^{-1}\mathfrak{p}).$$

Proof. Observe that

$$(S^{-1}\mathfrak{p})B_{\mathfrak{p}} = (\mathfrak{p}A_{\mathfrak{p}})B_{\mathfrak{p}} = \mathfrak{p}B_{\mathfrak{p}} = S^{-1}(\mathfrak{p}B).$$

Since B is a Dedekind domain, $\mathfrak{p}B$ factors into primes as

$$\mathfrak{p}B = \prod_{\mathfrak{q} \in \text{Spec}B} \mathfrak{q}^{\nu_{\mathfrak{q}}} = \prod_{\substack{\mathfrak{q} \in \text{Spec}B \\ \mathfrak{q} \cap A = \mathfrak{p}}} \mathfrak{q}^{\nu_{\mathfrak{q}}}.$$

It follows from the proposition that

$$(S^{-1}\mathfrak{p})B_{\mathfrak{p}} = S^{-1}(\mathfrak{p}B) = \prod_{\substack{\mathfrak{q} \in \text{Spec}B \\ \mathfrak{q} \cap A = \mathfrak{p}}} (S^{-1}\mathfrak{q})^{\nu_{\mathfrak{q}}}.$$

Recall that if $\mathfrak{q}$ does not lie over $\mathfrak{p}$, then $S^{-1}\mathfrak{q} = B_{\mathfrak{p}}$. Otherwise, $S^{-1}\mathfrak{q}$ are distinct primes of $B_{\mathfrak{p}}$ for distinct primes $\mathfrak{q}$ in A. So the lemma follows from unique factorization of ideals in the Dedekind domain $S^{-1}B$. □

As a result, localization at $S = A - \mathfrak{p}$ preserves ramification of prime ideals in B lying over $\mathfrak{p}$, it preserves divisibility of ideals in A by $\mathfrak{p}$, and it preserves divisibility of ideals in B by primes lying over $\mathfrak{p}$. The desired results on the fundamental properties of the discriminant and different ideals soon follow:

Corollary 4.8. *Let L/K an extension of number fields, $A \subset K$ a Dedekind domain with field of fractions K and B its integral closure in L. If $\mathfrak{p}$ is a prime in A, then $\mathfrak{p}|\Delta_{B/A}$ if and only if $e(\mathfrak{p}B|\mathfrak{p}) > 1$.*

Solution. By Lemma 4.6, $\mathfrak{p}|\Delta_{B/A}$ if and only if

$$S^{-1}\mathfrak{p}|S^{-1}\Delta_{B/A} = \Delta_{S^{-1}B/S^{-1}A},$$

where $S = A - \mathfrak{p}$. Moreover, by Lemma 4.7, $e(S^{-1}(\mathfrak{p}B)|S^{-1}\mathfrak{p}) = e(\mathfrak{p}B|\mathfrak{p})$, so the conclusion of the lemma is equivalent to

$$S^{-1}\mathfrak{p}|\Delta_{S^{-1}B/S^{-1}A} \iff e(S^{-1}(\mathfrak{p}B)|S^{-1}\mathfrak{p}),$$

now a statement about prime ideals in the ring extension $S^{-1}B/S^{-1}A$. Recall that $S^{-1}A = A_{\mathfrak{p}}$ is a DVR, so in particular it is a PID. We have already shown this case of the above statement (by abusing the fact that there is a basis for B over A), so the result follows. □

Corollary 4.9. *Let L/K be an extension of number fields, $A \subset K$ a Dedekind domain with field of fractions K and B its integral closure in L. If $\mathfrak{q}$ is a prime in B, then $e(\mathfrak{q}|\mathfrak{q} \cap A) > 1$ if and only if $\mathfrak{q}|\mathfrak{D}_{B/A}$.*

Proof. By the above discussion, the result is equivalent to

$$S^{-1}\mathfrak{q}|\mathfrak{D}_{S^{-1}B/S^{-1}A} \iff e(S^{-1}\mathfrak{q}|S^{-1}\mathfrak{q} \cap S^{-1}A) > 1$$

where $S = A - A \cap \mathfrak{q}$. Of course, $S^{-1}A$ is a PID, so we showed earlier that this result holds ($S^{-1}\mathfrak{q}$ is a prime ideal in $S^{-1}B$ by the correspondence). □

5. Local Methods

Taking the valuation ring of K under $\nu_{\mathfrak{p}}$ allowed us to pass to a situation in which A is a DVR. As we will demonstrate, it will be useful to be able to pass to a complete DVR.

5.1. **Completions of Valued Fields.** To obtain a complete valued field containing K, we can take the completion of K with respect to the topology induced by $|\cdot|_{\mathfrak{p}}$. The easiest way to obtain a field structure for the topological completion of K is to let R be the ring of Cauchy sequences in K, and $\mathfrak{m}$ be the set of sequences in R which converge to zero. One can check that $\mathfrak{m}$ is a maximal ideal, so that $\widehat{K}_{\mathfrak{p}} := R/\mathfrak{m}$ is a field equipped with an embedding $K \to \widehat{K}_{\mathfrak{p}}$ taking a field element to the residue class of the corresponding constant sequence. Finally, one can check from the triangle inequality that $|\cdot|_{\mathfrak{p}}$ takes any Cauchy sequence in K to a (therefore convergent) Cauchy sequence in $\mathbf{R}$, so we can extend $|\cdot|_{\mathfrak{p}}$ to an absolute value on $\widehat{K}_{\mathfrak{p}}$ by taking

$$|\overline{\{a_n\}}| = \lim_{n\to\infty} |a_n|_{\mathfrak{p}},$$

noticing that this does not depend on the representative Cauchy sequence, verifying that it yields a valid non-archimedian absolute value on $\widehat{K}_{\mathfrak{p}}$, and that $\widehat{K}_{\mathfrak{p}}$ is complete under the topology induced by it.

We write $|\cdot|_{\hat{\mathfrak{p}}}$ for this absolute value on the completion, and $\nu_{\hat{\mathfrak{p}}}$ for the corresponding non-archimedian valuation.

Since $\nu_{\hat{\mathfrak{p}}}$ is obtained by a continuous transformation of $|\cdot|_{\hat{\mathfrak{p}}}$, we know that for all $x \in \widehat{K}_{\mathfrak{p}}$,

$$\nu_{\hat{\mathfrak{p}}}(x) = \lim_{n\to\infty} \nu_{\mathfrak{p}}(x_n)$$

where $\{x_n\}$ is a Cauchy sequence in K representing x. Since the image of $\nu_{\hat{\mathfrak{p}}}$ is $\mathbf{Z} \cup \{\infty\}$, it follows that the sequence $\{\nu_{\mathfrak{p}}(x_n)\}$ is eventually stable, and thus

$$\nu_{\mathfrak{p}}(K) = \nu_{\hat{\mathfrak{p}}}(\hat{K}_{\mathfrak{p}}) = \mathbf{Z} \cup \{\infty\}.$$

Let A be the valuation ring of K. The valuation ring of $\hat{K}$ is

$$\hat{A} := \{x \in \hat{K} : \nu_{\hat{\mathfrak{p}}}(x) \geq 0\}.$$

If $x \notin \widehat{A}$, then $1/x$ is, so $\hat{A}$ has field of fractions $\widehat{K}_{\mathfrak{p}}$ (this is a general fact about valuation rings). Every $x \in \hat{A}$ is the limit of a sequence in K whose valuation is eventually equal to $\nu_{\hat{\mathfrak{p}}}(x) \geq 0$, so it is a limit point of the valuation ring A of K. Conversely any limit point of A must have nonnegative valuation (since the elements of A do), so $\hat{A}$ is the closure of A in $\hat{K}_{\mathfrak{p}}$. In particular, $\hat{A}_{\mathfrak{p}}$ is a DVR (it is the valuation ring of the discrete valuation $\nu_{\hat{\mathfrak{p}}}$) and since it is a closed subset of the complete metric space $\widehat{K}_{\mathfrak{p}}$ it is also complete under its valuation. Since the discrete valuation on $\widehat{A}_{\mathfrak{p}}$ extends that on A, any uniformizing parameter π for A (i.e. an element of valuation equal to 1) is also a uniformizing parameter for $\widehat{A}_{\mathfrak{p}}$.

The maximal ideal of $\hat{A}_{\mathfrak{p}}$ is the set of elements of positive valuation, which by the same argument as before is the closure (i.e. completion) of $\mathfrak{p}$ in $\widehat{K}_{\mathfrak{p}}$. Note that this is equal to

$$\hat{p} := \pi \hat{A}_{\mathfrak{p}} = (\pi A)\hat{A}_{\mathfrak{p}} = \mathfrak{p}\hat{A}_{\mathfrak{p}},$$

so the closure of $\mathfrak{p}$ is equal the ideal it generates in $\hat{A}_{\mathfrak{p}}$. Since A is a DVR, its ideals are all of the form $\pi^n A$ for some $n \geq 0$. Thus, the same argument above applies: we know that the closure of $I = \pi^n A$ in $\hat{K}_{\mathfrak{p}}$ is

$$\hat{I} = \{x \in \hat{K}_{\mathfrak{p}} : \nu_{\mathfrak{p}}(x) \geq n\} = \pi^n \hat{A}_{\mathfrak{p}} = I\hat{A}_{\mathfrak{p}}.$$

Hence, the ideals of A are in natural bijection with those of $\hat{A}_{\mathfrak{p}}$ under the map $I \mapsto \hat{I}$, and $\hat{I}$ is equal to the closure of I in $\hat{A}_{\mathfrak{p}}$, as well as the ideal in $\hat{A}_{\mathfrak{p}}$ generated by I.

Our goal is typically to show a result about the ideals of $\mathcal{O}_K$, or more generally a Dedekind domain A whose field of fractions is K. In that case, the valuation ring of K is $S^{-1}A$, where $S = A - \mathfrak{p}$. For any ideal $I \subset A$, we just showed that the closure of $S^{-1}I$ is

$$(S^{-1}I)\hat{A}_{\mathfrak{p}} = I(S^{-1}A)(\hat{A}_{\mathfrak{p}}) = I\hat{A}_{\mathfrak{p}}.$$

Let $n \geq 1$ and $a/s \in S^{-1}I$. Since $s \notin \mathfrak{p}$, there exists a $b \in A$ such that $bs \equiv 1 \pmod{\mathfrak{p}^n}$. Then we have $ab \in I$ and

$$\nu_{\mathfrak{p}}\left(\frac{a}{s} - ab\right) = \nu_{\mathfrak{p}}\left(\frac{a}{s}(1 - bs)\right) \geq n.$$

It follows that I is dense in $S^{-1}I$, so the closure of I in $\hat{K}_{\mathfrak{p}}$ is just $I\hat{A}_{\mathfrak{p}}$.

In summary: passing from A to $S^{-1}A$ keeps only the ideal-theoretic information related to divisibility by powers of $\mathfrak{p}$, while passing to $\hat{A}_{\mathfrak{p}}$ loses no additional information about the ideals.

Let B/A be an extension of Dedekind domains with field of fractions L/K such that B is the integral closure of A in L.

We used in the previous section the fact that localization preserves integral closures: $S^{-1}B/S^{-1}A$ is a ring extension with the same properties as above, but $S^{-1}A$ is now guaranteed to be a DVR. Taking completions allows us to do the same thing, except both rings are DVRs (recall that the localization of B at a prime $\mathfrak{q}|\mathfrak{p}$

instead of at S is not necessarily integral over $S^{-1}A$). In particular, from [29, Ch.2, Proposition 3],

Proposition 5.1. *Let L/K be a finite extension of number fields, let A be a Dedekind domain with field of fractions K, and let B be its integral closure in L. Let $\mathfrak{q}$ be a prime of B lying over the prime $\mathfrak{p}$ of A. Then the integral closure of $\hat{A}_{\mathfrak{p}}$ in $\hat{L}_{\mathfrak{q}}$ is $\hat{B}_{\mathfrak{q}}$.*

Proof. First, observe that

$$[\hat{L}_{\mathfrak{q}} : \hat{K}_{\mathfrak{p}}] \leq [L : K] < \infty.$$

This is because any element $x \in \hat{L}_{\mathfrak{q}}$ is the limit of some Cauchy sequence of elements in L, each of which can be expanded in terms of a K-basis $\ell_1, \ldots, \ell_n$ for L. So, we can write

$$x = \overline{\{a_{i,1}\ell_1 + \cdots + a_{i,n}\ell_n\}_{i\geq 1}}$$

and verify that since this sequence is Cauchy, each sequence $\{a_{i,j}\}_{i\geq 1}$ is a Cauchy sequence that happens to live in K. It follows that

$$x = \overline{\{a_{i,1}\}_{i\geq 1}}\ell_1 + \cdots + \overline{\{a_{i,n}\}_{i\geq 1}}\ell_n$$

which means $\ell_1, \ldots, \ell_n$ span $\hat{L}_{\mathfrak{q}}$ over $\hat{K}_{\mathfrak{p}}$.

Let B' be the integral closure of $\hat{A}_{\mathfrak{p}}$ in $\hat{K}_{\mathfrak{q}}$. It is our goal to show that $B' = \hat{B}_{\mathfrak{q}}$. Since $[\hat{L}_{\mathfrak{q}} : \hat{K}_{\mathfrak{p}}] < \infty$, clearing denominators in the minimal polynomial of any $x \in \hat{L}_{\mathfrak{q}}$ tells us that $ax \in B'$ for some $a \in \hat{A}_{\mathfrak{p}}$. Therefore, Frac$B' = \hat{L}_{\mathfrak{q}}$.

Since $\hat{A}_{\mathfrak{p}}$ is a DVR, it has only one prime $\hat{\mathfrak{p}}$, and all the primes $\mathfrak{q}_i$ of B' lie over it. Each $\mathfrak{q}_i$ defines a (topologically) inequivalent valuation on B' (and hence on its field of fractions $\hat{L}_{\mathfrak{q}}$), which can be made to extend $\nu_{\hat{\mathfrak{p}}}$ by scaling by $e(\mathfrak{q}_i|\mathfrak{p})$. It follows that each $\mathfrak{q}_i$ defines an inequivalent norm on $\hat{L}_{\mathfrak{q}}$ as a finite-dimensional $\hat{K}_{\mathfrak{p}}$-vector space. But $\hat{K}_{\mathfrak{p}}$ is complete, so all norms on this finite-dimensional $\hat{K}_{\mathfrak{p}}$-vector space are equivalent. Hence, B' has only one prime ideal. Since it is a Dedekind domain, this means B' is already a DVR with maximal ideal $\mathfrak{q}'$, and the topology on $\hat{L}_{\mathfrak{q}}$ defined by $\nu_{\mathfrak{q}'}$ is the same as the one defined by $\nu_{\mathfrak{q}}$. Hence, the valuation ring of $\hat{L}_{\mathfrak{q}}$ is the same with respect to both valuations (the valuation ring can be defined purely in terms of the topology as the set of elements whose inverses do not tend to zero as they are raised to a power $n \to \infty$). By Lemma 3.3 and our original definition of $\hat{B}_{\mathfrak{q}}$, the valuation ring of $\hat{L}_{\mathfrak{q}}$ is

$$\hat{B}_{\mathfrak{q}} = (B' - \mathfrak{q}')^{-1}B'.$$

But $B' - \mathfrak{q}'$ consists of the units of B' since it is a DVR, so the result is $B' = \hat{B}_{\mathfrak{q}}$ as desired. □

Typical expositions on extensions of complete fields (cite Neukirch and Lang) exploit Hensel's lemma [cite it] to prove the above facts and to show that the unique absolute value extending $|\cdot|_{\hat{\mathfrak{p}}}$ is actually given by

$$|x| = N_{\hat{L}_{\mathfrak{q}}/\hat{K}_{\mathfrak{q}}}(x)^{1/[\hat{L}_{\mathfrak{q}}:\hat{K}_{\mathfrak{q}}]}.$$

We will not require any of these facts for our purposes, however. Finally, we observe that completion preserves ramification indices and inertial degrees: If $\mathfrak{q} \in \text{Spec}B$ lies over $\mathfrak{p} \in \text{Spec}A$, then we have

$$\nu_{\mathfrak{p}} = e(\mathfrak{q}|\mathfrak{p})\nu_{\mathfrak{q}}$$

on A. On the other hand, $\hat{\mathfrak{q}}$, the maximal ideal of $\hat{B}_{\mathfrak{q}}$, lies over $\hat{\mathfrak{p}}$, the maximal ideal of $\hat{A}_{\mathfrak{p}}$. We know that $\nu_{\hat{\mathfrak{q}}}$ agrees with $\nu_{\mathfrak{q}}$ on B, and the same is true on A for $\nu_{\hat{\mathfrak{p}}}$ and $\nu_{\mathfrak{p}}$. Thus, on A we have

$$\nu_{\mathfrak{p}} = e(\hat{\mathfrak{q}}|\hat{\mathfrak{p}})\nu_{\mathfrak{q}}$$

from which it follows that $e(\hat{\mathfrak{q}}|\hat{\mathfrak{p}}) = e(\mathfrak{q}|\mathfrak{p})$.

For the inertial degree, we use the standard fact that taking completions induces an isomorphism of residue fields $A/\mathfrak{p} \cong \hat{A}_{\mathfrak{p}}/\hat{p}$, since every Cauchy sequence in A eventually has constant image in $A/\mathfrak{p}$. The same holds for B, so we have a commutative diagram

$$\begin{array}{ccc} B/\mathfrak{q} & \xrightarrow{\sim} & \hat{B}_{\mathfrak{q}}/\hat{\mathfrak{q}} \\ \uparrow & & \uparrow \\ A/\mathfrak{p} & \xrightarrow{\sim} & \hat{A}_{\mathfrak{p}}/\hat{\mathfrak{p}} \end{array}$$

which yields

$$f(\mathfrak{q}|\mathfrak{p}) = [B/\mathfrak{q} : A/\mathfrak{p}] = [\hat{B}_{\mathfrak{q}}/\hat{\mathfrak{q}} : \hat{A}_{\mathfrak{p}}/\hat{\mathfrak{p}}] = f(\hat{q}|\hat{p}).$$

5.2. **Extensions of DVRs.** By bringing us to the case of an extension of DVRs, passing to completions allows us to exploit the following result from [29, Ch. III, Proposition 12].

Proposition 5.2. *Let A, K, L, B be such that A and B are DVRs with separable residue field extension $\kappa(B)/\kappa(A)$. Then there exists an $\alpha \in A$ such that $B = A[\alpha]$.*

Proof. Let $f = f(\mathfrak{q}|\mathfrak{p})$ and $e = e(\mathfrak{q}|\mathfrak{p})$. Since $\kappa(B)/\kappa(A)$ is separable of degree f, it admits a primitive element $\hat{x}$ (take $x \in B$ to be an arbitrary lift of the primitive element). Take π_B to be an arbitrary uniformizing parameter for B. Since $\mathfrak{q}$ is the only prime lying over $\mathfrak{p}$ we know that

$$[L : K] = ef$$

and since A is a PID, B is automatically a free A-module of rank ef. We claim that the elements $x^i\pi_B^j$ for $0 \leq i < f$ and $0 \leq j < e$ are an A-basis for B. Since there are ef of them, it suffices to show that they generate B as an A-module. By Nakayama's lemma, it suffices to show that their residues generate $B/\mathfrak{p}B$ as an $A/\mathfrak{p}$-vector space.

We know $\mathfrak{p}B = \pi_B^e B$, and the powers of x generate $B/\mathfrak{q} = B/\pi_B B$ as a $A/\mathfrak{p}$-vector space, so by multiplying by the first e powers of π_B we obtain a spanning set, as desired.

To show that the powers of x span B as an A-module, it therefore suffices to show that we can obtain a uniformizing parameter for B as a linear combination of powers of x. To do this, take $P \in \kappa(A)[X]$ to be the minimal polynomial of $\bar{x}$ over $\kappa(A)$, and take an arbitrary lift of the coefficients of P to A to obtain a polynomial $Q \in A[X]$ with the property that $Q(x) \in \mathfrak{p}$. In other words, $\nu_{\mathfrak{q}}(Q(x)) \geq 1$. If equality holds, then $Q(x)$ is a uniformizing parameter for B, and we are done. Otherwise, note that $\overline{x + \pi_B} = \bar{x}$, so $x + \pi_B$ is also a lift of a primitive element for $\kappa(B)/\kappa(A)$. In this case, however, we can compute (by Taylor expansion of polynomials)

$$Q(x + \pi_B) = Q(x) + \pi_B Q'(x) + \pi_B^2 b$$

where $b \in B$ – here we have collapsed all of the terms of valuation greater than 1. Moreover, $\kappa(B)/\kappa(A)$ is separable, so the minimal polynomial P of the primitive element $\bar{x}$ cannot share a root with P'. Hence, $P'(\bar{x}) \neq 0$, which means that $Q'(x)$ is not divisible by π_B. It follows that $\nu_{\mathfrak{q}}(Q(x+\pi_B)) = 1$, so $Q(x+\pi_B)$ is a uniformizing parameter for B.

The resulting basis (consisting either of products of powers of x and $Q(x)$ or of $x+\pi_B$ and $Q(x+\pi_B)$) is a power basis for B over A, as desired. □

5.3. **Extending Scalars by Completions.** To be able to use this property to analyze the different ideal, it still remains to show that the different is compatible with taking completions.

We showed already that L spans all of $\hat{L}_{\mathfrak{q}}$ over $\hat{K}_{\mathfrak{p}}$. In other words, the multiplication map

$$L \otimes_K \hat{K}_{\mathfrak{p}} \to \hat{L}_{\mathfrak{q}}$$

is a surjective map of $\hat{K}_{\mathfrak{p}}$-vector spaces. A careful analysis of $L \otimes_K \hat{K}_{\mathfrak{p}}$ yields a concrete way in which the data of L is related to that of its completions, from [29, Ch. II, Théorème 1]:

Proposition 5.3. *The multiplication map*

$$\varphi : L \otimes_K \hat{K}_{\mathfrak{p}} \to \prod_{\mathfrak{q}|\mathfrak{p}} \hat{L}_{\mathfrak{q}}$$

is an isomorphism of $\hat{K}_{\mathfrak{p}}$-vector spaces.

Proof. By extending scalars of a basis for L over K (i.e. tensoring each basis element with $1 \in \hat{K}_{\mathfrak{p}}$), we can see that $L \otimes_K \hat{K}_{\mathfrak{p}}$ is a K-vector space of dimension equal to $[L : K]$. The fundamental identity involving ramification indices and inertial degress (see Marcus) and the above remarks about ramification being preserved by completion tells us that that this dimension is

$$[L : K] = \sum_{\mathfrak{q}|\mathfrak{p}} e(\mathfrak{q}|\mathfrak{p}) f(\mathfrak{q}|\mathfrak{p}) = \sum_{\mathfrak{q}|\mathfrak{p}} e(\hat{\mathfrak{q}}|\hat{\mathfrak{p}}) f(\hat{\mathfrak{q}}|\hat{\mathfrak{p}})$$

which is the dimension over $\hat{K}_{\mathfrak{p}}$ of $\prod_{\mathfrak{q}|\mathfrak{p}} \hat{L}_{\mathfrak{q}}$. It therefore suffices to show that φ is surjective. Since $\hat{K}_{\mathfrak{p}}$ is complete, any (finite-dimensional) subspace of $\prod_{\mathfrak{q}|\mathfrak{p}} \hat{L}_{\mathfrak{q}}$ is complete and therefore closed. So, it suffices to show that the image of φ is dense in the codomain (since the image must be closed). Since all norms on a finite-dimensional vector space over a complete field are equivalent, we might as well take the norm on the codomain to be the sup norm. The conclusion follows from an important result from the theory of valued fields:

Lemma 5.4 (Weak Approximation)**.** *Let $|\cdot|_1, \ldots, |\cdot|_m$ be pairwise inequivalent absolute values on a field K. For any choice of $a_1, \ldots, a_m \in K$ and $\varepsilon > 0$, there exists an $x \in K$ such that $|x - a_i| < \varepsilon$ for each i.*

Proof. See [25, Ch. II, Theorem 3.4] □

Since L is dense in each $\hat{L}_{\mathfrak{q}}$, any element of $\prod_{\mathfrak{q}|\mathfrak{p}} \hat{L}_{\mathfrak{q}}$ can be approximated by elements whose coordinates are all in L, which can in turn (by weak approximation) can be approximated by elements of the form $\varphi(\ell \otimes 1)$. So the image in φ is dense in the codomain, as desired. □

The same is true for the base change of B under some hypotheses on A (note that we can achieve these hypotheses by localizing at $\mathfrak{p}$ first).

Proposition 5.5. *Suppose A is a PID. Then the multiplication map*

$$\varphi : B \otimes_A \hat{A}_{\mathfrak{p}} \to \prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}}$$

is an isomorphism of A-modules.

Proof. Since A is a PID, B is a free A-module of rank $[L : K]$. By extending the scalars of a basis for B over A, we get that $B \otimes_A \hat{A}_{\mathfrak{p}}$ is a free $\hat{A}_{\mathfrak{p}}$ Since $\hat{A}_{\mathfrak{p}}$ is a DVR (and thus a PID) for each $\mathfrak{q}|\mathfrak{p}$, we also know that $\hat{B}_{\mathfrak{q}}$ is a free $\hat{A}_{\mathfrak{p}}$-module of rank $[\hat{L}_{\mathfrak{q}} : \hat{K}_{\mathfrak{p}}]$. Thus, the codomain is a free $\hat{A}_{\mathfrak{p}}$-module of rank

$$\sum_{\mathfrak{q}|\mathfrak{p}} [\hat{L}_{\mathfrak{q}} : \hat{K}_{\mathfrak{p}}] = [L : K]$$

as well. This means it suffices to show that φ is surjective, i.e. that the image of a generating set for $B \otimes_A \hat{A}_{\mathfrak{p}}$ is a generating set for the codomain. Since $\hat{\mathfrak{p}}$ is the only maximal ideal of $\hat{A}_{\mathfrak{p}}$, it suffices to show this modulo $\hat{\mathfrak{p}}$ by Nakayama's lemma. We have a canonical identifications $A/\mathfrak{p}^n \cong \hat{A}_{\mathfrak{p}}/\hat{p}^n$ for all $n \geq 0$ (the argument is the same as for $n = 0$: all Cauchy sequences must eventually have constant image under the projection map).

In other words, it suffices to show that the map

$$(B \otimes_A \hat{A}_{\mathfrak{p}})/\hat{p} \to \left(\prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}} \right) /\hat{p}$$

is surjective. We have an isomorphism of A-modules

$$(B \otimes_A \hat{A}_{\mathfrak{p}})/\hat{\mathfrak{p}} \cong B \otimes_A (\hat{A}_{\mathfrak{p}}/\hat{p}) \cong B \otimes_A (A/\mathfrak{p}) \cong B/\mathfrak{p}B$$

and

$$\left(\prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}} \right) /\hat{p} \cong \prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}}/\hat{\mathfrak{p}}\hat{B}_{\mathfrak{q}} = \prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}}/\hat{\mathfrak{q}}^{e(\mathfrak{q}|\mathfrak{p})} \cong \prod_{\mathfrak{q}|\mathfrak{p}} B/\mathfrak{q}^{e(\mathfrak{q}|\mathfrak{p})}.$$

Following these identifications, the map we want to show is surjective is identified with the projection

$$B/\mathfrak{p}B \to \prod_{\mathfrak{q}|\mathfrak{p}} B/\mathfrak{q}^{e(\mathfrak{q}|\mathfrak{p})}$$

which is indeed surjective by the Chinese remainder theorem. □

Finally, can see that the different may be defined locally:

Proposition 5.6. *Suppose $\mathfrak{q}$ lies over $\mathfrak{p}$. Then*

$$\mathfrak{D}_{B/A}\hat{B}_{\mathfrak{q}} = \mathfrak{D}_{\hat{B}_{\mathfrak{q}}/\hat{A}_{\mathfrak{p}}}.$$

Proof. First, suppose that A is a DVR (we will reduce to this case by localizing). The previous proposition tells us that

$$B \otimes_A \hat{A}_{\mathfrak{p}} \cong \prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_{\mathfrak{q}}$$

via the multiplication map. Finitely many primes lie over $\mathfrak{p}$ so this product is a direct sum of $\hat{A}_\mathfrak{p}$-modules embedded in $L \otimes_K \hat{K}_\mathfrak{p} \cong \prod_{\mathfrak{q}|\mathfrak{p}} \hat{L}_\mathfrak{p}$. Moreover, the trace form on L extends to a $\hat{K}_\mathfrak{p}$-bilinear form on $L \otimes_K \hat{K}_\mathfrak{p}$ given by

$$\langle x, y \rangle = \sum_{\mathfrak{q}|\mathfrak{p}} \mathrm{Tr}_{\hat{L}_\mathfrak{q}/\hat{K}_\mathfrak{p}}(xy)$$

so that the dual module of $B \otimes_A \hat{A}_\mathfrak{p}$ in $L \otimes_K \hat{K}_\mathfrak{p}$ is identified with $\prod_{\mathfrak{q}|\mathfrak{p}} (\hat{B}_\mathfrak{q})^\vee$. Since A is a DVR, there is a basis $e_1, \ldots, e_n$ for B and a dual basis $e_1^\vee, \ldots, e_n^\vee$ for $B^\vee$ the dual of B in L. By computing traces, we can observe that

$$e_1^\vee \otimes 1, \ldots, e_n^\vee \otimes 1$$

is a basis for the dual module of $B \otimes_A \hat{A}_\mathfrak{p}$. This yields another identification of $\hat{A}_\mathfrak{p}$-modules

$$\prod_{\mathfrak{q}|\mathfrak{p}} (\hat{B}_\mathfrak{q})^\vee \cong (B \otimes_A \hat{A}_\mathfrak{p})^\vee \cong B^\vee \otimes_A \hat{A}_\mathfrak{p}.$$

It follows that $B^\vee$ generates the fractional ideal $(\hat{B}_\mathfrak{q})^\vee$ in $\hat{B}_\mathfrak{q}$. Taking inverses, $\mathfrak{D}_{B/A}$ generates $\mathfrak{D}_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}}$ in $\hat{B}_\mathfrak{q}$.

In the general case, let $S = A - \mathfrak{p}$. Then we have

$$\mathfrak{D}_{B/A} S^{-1} B = \mathfrak{D}_{S^{-1}B/S^{-1}A}.$$

Applying our previous analysis to $S^{-1}B/S^{-1}A$, we get

$$\mathfrak{D}_{S^{-1}B/S^{-1}A} \hat{B}_\mathfrak{q} = \mathfrak{D}_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}}.$$

Plugging in the previous identification yields the desired result. □

5.4. **Kähler Differentials and the Different.** The main theorem of this section characterizes the different in terms of another structure from algebraic geometry:

Definition 5.7. Let A be a ring and B an A-algebra. The *module of Kähler differentials* of B/A, denoted by $\Omega_{B/A}$, is obtained by taking the free B-module generated by the symbols dx for $x \in B$ and quotienting by the relations

$$\begin{aligned} d(x+y) &= dx + dy \\ d(xy) &= x\,dy + y\,dx \\ da &= 0 \end{aligned}$$

for all $x, y \in B$ and $a \in A$.

We can view the elements of $\Omega_{B/A}$ as images under the map $d : B \to \Omega_{B/A}$ with the three properties above. Such a map is called an A-derivation, and in fact $\Omega_{B/A}$ is characterized by a universal property using A-derivations:

Proposition 5.8. *For any B-module M and A-derivation $d' : B \to M$, there exists a unique B-linear map $\alpha : \Omega_{B/A} \to M$ such that $d' = \alpha \circ d$.*

Using the universal property, one can show a few basic properties of the module of differentials:

Proposition 5.9. *Let A' be an A-algebra. Then*

$$\Omega_{(B \otimes_A A')/A'} \cong \Omega_{B/A} \otimes_B (B \otimes A')$$

as B-modules.

Proof. See [6, Proposition 16.4]. □

Proposition 5.10. *Let S be a multiplicative subset of B. Then*

$$\Omega_{S^{-1}B/A} = S^{-1}\Omega_{B/A} = \Omega_{B/A} \otimes_B S^{-1}B.$$

Proof. See [6, Proposition 16.9]. □

Proposition 5.11. *Let $\{B_\alpha\}_{\alpha \in I}$ be a collection of A-algebras. Then*

$$\Omega_{\prod_{\alpha \in I} B_\alpha / A} \cong \prod_{\alpha \in I} \Omega_{B_\alpha / A}$$

Proof. See [6, Proposition 16.10]. □

Proposition 5.12. *The module of differentials is compatible with taking completions and localization: If $\mathfrak{q}|\mathfrak{p}$ and $S = A - \mathfrak{p}$ then*

$$S^{-1}\Omega_{B/A} = \Omega_{B/A} \otimes_B S^{-1}B = \Omega_{S^{-1}B/S^{-1}A}$$

and

$$\Omega_{B/A} \otimes_B \hat{B}_\mathfrak{q} = \Omega_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}}.$$

Proof. The first part of the proposition follows from Proposition 5.9, since $S^{-1}B = B \otimes_A S^{-1}A$. For the second part, assume first that A is a DVR, so that Proposition 5.5 applies. In this case, by Proposition 5.11 we have

$$\Omega_{B \otimes_A \hat{A}_\mathfrak{p}/\hat{A}_\mathfrak{p}} \cong \Omega_{\prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}} \cong \prod_{\mathfrak{q}|\mathfrak{p}} \Omega_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}}$$

and Proposition 5.9 along with the fact that finite direct products are the same as direct sums gives

$$\Omega_{B \otimes_A \hat{A}_\mathfrak{p}/\hat{A}_\mathfrak{p}} \cong \Omega_{B/A} \otimes_B \prod_{\mathfrak{q}|\mathfrak{p}} \hat{B}_\mathfrak{q} \cong \prod_{\mathfrak{q}|\mathfrak{p}} \Omega_{B/A} \otimes_B \hat{B}_\mathfrak{q}.$$

Putting this all together, we have an isomorphism

$$\prod_{\mathfrak{q}|\mathfrak{p}} \Omega_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}} = \prod_{\mathfrak{q}|\mathfrak{p}} \Omega_{B/A} \otimes_B \hat{B}_\mathfrak{q}$$

which clearly maps componentwise, giving the desired isomorphism

$$\Omega_{B/A} \otimes_B \hat{B}_\mathfrak{q} \cong \Omega_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}}.$$

Finally, in the general case we apply the above reasoning to $S^{1-}B/S^{-1}A$ to get

$$\Omega_{\hat{B}_\mathfrak{q}/\hat{A}_\mathfrak{p}} = \Omega_{S^{-1B}/S^{-1}A} \otimes_{S^{-1}B} \hat{B}_\mathfrak{q} = \Omega_{B/A} \otimes_B S^{-1}B \otimes_{S^{-1}B} \hat{B}_\mathfrak{q} = \Omega_{B/A} \otimes_B \hat{B}_\mathfrak{q}$$

by the first part of the proposition. □

Theorem 5.13. *Let L/K be an extension of number fields. Then*

$$\mathfrak{D}_{\mathcal{O}_L/\mathcal{O}_K} = \mathrm{Ann}_{\mathcal{O}_L}(\Omega_{\mathcal{O}_L/\mathcal{O}_L}).$$

Proof. By Lemma 4.1, it suffices to show that the desired equality holds when both sides are localized at all primes of $\mathfrak{q}$ of B. So we fix a prime $\mathfrak{q}$ lying over a prime $\mathfrak{p}$ of A. Let $S = B - \mathfrak{q}$. It suffices to show that

$$S^{-1}\mathfrak{D}_{B/A} = S^{-1}\mathrm{Ann}_B(\Omega_{B/A}).$$

Since the ideals of $S^{-1}B$ are in bijection with the ideals of $\hat{B}_{\mathfrak{q}}$ (see the beginning of this section), it suffices to show that

$$\mathfrak{D}_{B/A}\hat{B}_{\mathfrak{q}} = \mathrm{Ann}_B(\Omega_{B/A})\hat{B}_{\mathfrak{q}}.$$

Moreover, $\hat{B}_{\mathfrak{q}}$ is a flat B-module (see Eisenbud 7.2) so $\mathrm{Ann}_B(\Omega_{B/A})\hat{B}_{\mathfrak{q}} = \mathrm{Ann}_{\hat{B}_{\mathfrak{q}}}(\Omega_{B/A}\otimes_B \hat{B}_{\mathfrak{q}})$ which means the by Proposition 5.12 it suffices to show

$$\mathfrak{D}_{\hat{B}_{\mathfrak{q}}/\hat{A}_{\mathfrak{p}}} = \mathrm{Ann}_{\hat{B}_{\mathfrak{q}}}(\Omega_{\hat{B}_{\mathfrak{q}}/\hat{A}_{\mathfrak{p}}}),$$

i.e. we may assume that A and B are DVRs, so by Proposition 5.2 $B = A[\beta]$ for some $\beta \in B$ with minimal polynomial $f(x) \in A[x]$. By computing a dual basis (see [25, Ch. III, Proposition 2.4]) one sees that both sides of the desired equation are the ideal generated by $f'(\beta)$. □

6. Ramification Groups

Let L/K be a finite Galois extension of complete valued fields of characteristic zero (actually separable is enough but we never proved that the trace form is nondegenerate in separable extensions of positive characteristic), let B/A be their valuation rings with maximal ideals $\mathfrak{q}|\mathfrak{p}$, and suppose they have separable residue field extension. By Proposition 5.2, there is some $\beta \in B$ such that $B = A[\beta]$. The group $G = \mathrm{Gal}(L/K)$ is equipped with a decreasing series of subgroups:

Definition 6.1. Let $i \geq -1$. The i-th *ramification subgroup* of G is defined by

$$G_i := \{\sigma \in G : \nu_{\mathfrak{q}}(b - \sigma(b)) \geq i + 1 \text{ for all } b \in B\}.$$

Since B is the integral closure of A in L, note that G acts on B. Note that for $\sigma \in G$, the following are equivalent:

- $\nu_{\mathfrak{q}}(b - \sigma(b)) \geq i + 1$ for all $b \in B$
- $\nu_{\mathfrak{q}}(\beta - \sigma(\beta)) \geq i + 1$.
- σ acts trivially on $B/\mathfrak{q}^{i+1}$.

For $\sigma \in G$, take $i_G(\sigma) := \nu_{\mathfrak{q}}(\beta - \sigma(\beta))$. It is clear from the second characterization that

$$i_G(\sigma) - 1 = \sup\{i \geq -1 : \sigma \in G_i\}.$$

If $\sigma \neq \mathrm{id}$, $\beta - \sigma(\beta) \neq 0$, so $i_G(\sigma) < \infty$. Since G is finite, we can take the maximum of i_G over all non-identity elements to get that all $\sigma \neq \mathrm{id}$ are eventually not in G_i, i.e. the G_i's are eventually trivial.

One connection between the ramification groups and ramification comes from the different ideal. Since B is a DVR, $\mathfrak{D}_{B/A}$ is determined by the $\mathfrak{q}$-adic valuation of its generating element. Abusing notation, we denote this by $\nu_{\mathfrak{q}}(\mathfrak{D}_{B/A})$.

Theorem 6.2. *The different is given by*

$$\nu_{\mathfrak{q}}(\mathfrak{D}_{B/A}) = \sum_{\substack{\sigma \in G \\ \sigma \neq \mathrm{id}}} i_G(\sigma) = \sum_{i \geq 0} (|G_i| - 1).$$

Proof. Recall that since $B = A(\beta)$, the different is generated $f'(\beta)$ where $f \in A[X]$ is the minimal polynomial of β. Thus, it suffices to compute $\nu_{\mathfrak{q}}(f'(\beta))$.

By Galois theory,

$$f(X) = \prod_{\sigma \in G} (X - \sigma(\beta))$$

and we can compute

$$f'(\beta) = \prod_{\substack{\sigma \in G \\ \sigma \neq \mathrm{id}}} (\beta - \sigma(\beta)).$$

This yields by definition

$$\nu_{\mathfrak{q}}(\mathfrak{D}_B/A) = \nu_{\mathfrak{q}}(f'(\beta)) = \sum_{\substack{\sigma \in G \\ \sigma \neq \mathrm{id}}} \nu_{\mathfrak{q}}(\beta - \sigma(\beta)) = \sum_{\substack{\sigma \in G \\ \sigma \neq \mathrm{id}}} i_G(\sigma).$$

Finally, $i_G(\sigma) = \sup\{i \geq -1 : \sigma \in G_i\} + 1 = \#\{i \geq 0 : \sigma \in G_i\}$ since $G_{-1} = G$, so this sum just counts the total size of all the G_i's with $i \geq 0$ not counting id, i.e.

$$\sum_{\substack{\sigma \in G \\ \sigma \neq \mathrm{id}}} \nu_{\mathfrak{q}}(\beta - \sigma(\beta)) = \sum_{i=0}^{\infty} (|G_i| - 1)$$

as desired. □

7. Ramification in the Nottingham Group

Much of the setup in this section and the next comes from a previous paper on a special case of the main result by me and Hudson Kirkpatrick [12].

The construction of ramification groups from the previous section can be applied to any automorphism group of a valued field, though it loses any immediate connection with the ramification of prime ideals when it is not applied to the Galois group of an extension of valued fields.

For example, one can consider the global field $\mathbf{F}_p(X)$ equipped with the X-adic valuation ν_X. Its completion can be written as the field of formal Laurent series $\mathbf{F}_p((X))$, which has valuation ring $\mathbf{F}_p[[X]]$. In the same way as before, we can consider the ramification groups of $\mathrm{Aut}(\mathbf{F}_p((X)))$ by defining the *ramification numbers*

$$i(\sigma) = \nu_X(X - \sigma(X)) - 1,$$

shifting the index by 1 for the purpose of convenience in computations. Of particular interest is the group of *wild* automorphisms of $\mathbf{F}_p((X))$, namely those with $i(\sigma) \geq 1$. It turns out that these automorphisms can be identified with the following group:

Definition 7.1 (Nottingham group)**.** The *Nottingham group* $\mathcal{N}(\mathbf{F}_p)$ is the group under composition of power series of the form

$$f(X) = X + \sum_{i=2}^{\infty} a_i X^i \in \mathbf{F}_p[[X]]$$

with coefficients $a_i \in \mathbf{F}_p$.

It is easy to see that $\mathcal{N}(\mathbf{F}_p)$ is a group by explicitly computing compositional inverses from a given sequence of coefficients [10], and it is isomorphic to the group of wild automorphisms of $\mathbf{F}_p((X))$ via composition on the right [24].

The Nottingham group is of independent interest to group theorists, having first been studied by Jennings [10], Johnson [11] and York [31, 30]. One of the most important recent group-theoretic results on the Nottingham group is due to Leedham-Green–Weiss and Camina [2]; it states that every finite p-group can be embedded in $\mathcal{N}(\mathbf{F}_p)$, and every countably-based pro-p group can be embedded as a closed subgroup of $\mathcal{N}(\mathbf{F}_p)$.

We will only consider power series $f \in \mathcal{N}(\mathbf{F}_p)$ such that $f^{\circ m}(X) \neq X$ for all $m \geq 1$ (that is, elements of $\mathcal{N}(\mathbf{F}_p)$ of infinite order). This is because of a celebrated theorem of Klopsch [14] which classified all elements of $\mathcal{N}(\mathbf{F}_p)$ of order p up to conjugacy, and was generalized by Jean [9] and Lubin [22] to order p^n (see also [1]).

To make things clear in our computations, we make explicit what the X-adic valuation is on $\mathbf{F}_p((X))$ and recall the conventional ramification numbering.

Definition 7.2 (Valuation)**.** If $f(X) \in \mathbf{F}_p((X))$, then the *$X$-adic valuation* of f, $\mathrm{val}_X(f)$, is the integer i such that $f(X) \in X^i\mathbf{F}_p[[X]]$ and $f(X) \notin X^{i+1}\mathbf{F}_p[[X]]$. If f is identically zero, then we write that $\mathrm{val}_X(f) = +\infty$.

Definition 7.3 (Ramification)**.** The ramification number of $f \in \mathbf{F}_p((X))$ is

$$i(f) := \mathrm{val}_X(f(X) - X) - 1.$$

We will frequently mention the "n^{th} ramification number" of f, that is, $i_n(f) := i(f^{\circ p^n})$. It is immediate that conjugation in $\mathcal{N}(\mathbf{F}_p)$ preserves ramification. That is, for all $f, g \in \mathcal{N}(\mathbf{F}_p)$,

$$i(g \circ f \circ g^{-1}) = i(f).$$

It is therefore natural to consider the power series with certain sequences of ramification numbers and to classify them up to conjugacy.

Question 7.4. Let $b \in \mathbf{N}$ such that $p \not| b$. Which conjugacy classes of power series f of infinite order in $\mathcal{N}(\mathbf{F}_p)$ have ramification of the form $i_n(f) = b(1+p+\cdots+p^n)$ for all $n \geq 0$.

We will justify the choice of ramification of this form (which we will write as "b-ramified") with Theorem 7.11, a powerful result of Laubie and Saïne which restricts it to that form under certain conditions. Laubie, Movahhedi and Salinier [17] made progress on Question 7.4 in the case $b = 1$ by using advanced tools in number theory to prove that all such elements of $\mathcal{N}(\mathbf{F}_p)$ must be conjugate to powers of each other over some extension of $\mathbf{F}_p$. Motivated by trying to understand power series with ramification of this form for $b \geq 1$, we consider the following question:

Question 7.5. Let $b \in \mathbf{N}$ with $p \nmid b$. Given a sequence $\{a_i\}_{i\geq 1}$ of coefficients in $\mathbf{F}_p$, does

$$f(X) = X + \sum_{i=1}^{\infty} a_i X^{i+b} \in \mathcal{N}(\mathbf{F}_p)$$

have ramification $i_n(f) = b(1 + \cdots + p^n)$ for all $n \geq 0$?

We will write $a_1, \ldots, a_n$ as defined in Question 7.5 as the "first n nontrivial coefficients of f." Nordqvist [26] solved Question 7.5 for $b = 2$ by giving a finite polynomial condition on the a_i's. This itself is a generalization of the characterization of minimally ramified power series, which has been used to prove many important theorems in nonarchimedean dynamical systems [17, 21, 18, 20, 27]. Using an extension of Nordqvist's computations, Lindahl–Nordqvist [19] proved a nontrivial bound on the location of period points in nonarchimedian dynamical systems:

Theorem 7.6 (Lindahl–Nordqvist 2017)**.** *Let $p \geq 5$ be prime and let k be a nonarchimedian valued field of characteristic p. Let $f \in k[[X]]$ be of the form*

$$f(X) = X\left(1 + \sum_{j=2}^{\infty} a_j X^j\right).$$

Suppose x_0 is in the open unit disc of k and is a periodic point under the action of f with period p^n. Then

$$|x_0| \geq |a_2^{p-3}(3/2a_2^3 + a_3^2 - a_2a_4)|^{1/p}.$$

With the goal of finding a similar bound in the general case for b-ramified power series, we will answer Question 7.5 under the hypothesis that $p > b^2$.

First, we introduce some useful tools for determining the ramification numbers of a given $f \in \mathcal{N}(\mathbf{F}_p)$.

The most important historical result with regard to this is due to Sen [28, Theorem 1]:

Theorem 7.7 (Sen, 1969)**.** *Let $f \in \mathcal{N}(\mathbf{F}_p)$. Then for all $n \in \mathbf{N}$,*

$$i_n(f) \equiv i_{n-1}(f) \pmod{p^n}.$$

The proof of Sen's Theorem can be done entirely with computations on power series. The following corollary immediately follows from power series computations and Sen's Theorem.

Corollary 7.8. *Let $f \in \mathcal{N}(\mathbf{F}_p)$. Then for all $n \geq 0$,*

$$i_n(f) \geq 1 + p + \cdots + p^n. \tag{7.9}$$

Power series $f \in \mathcal{N}(\mathbf{F}_p)$ for which inequality holds in (7.9) are called *minimally ramified*. Much more difficult to prove is Keating's Theorem [13, Theorem 7], which gives a general form for $i_n(f)$ given the first two ramification numbers under certain conditions.

Theorem 7.10 (Keating, 1992)**.** *Suppose $f \in \mathcal{N}(\mathbf{F}_p)$ has infinite order. If $i_0(f) = 1$ and $i_1(f) = 1 + bp$ with $1 \leq b \leq p-2$, then $i_n(f) = 1 + bp + \cdots + bp^n$.*

We will use a powerful generalization of Theorem 7.10 from [16, Theorem 2].

Theorem 7.11 (Laubie and Saïne 1997)**.** *Let $f \in \mathcal{N}(\mathbf{F}_p)$ be of infinite order. Then the following hold:*

(1) If $p|i(f)$, then $i_n(f) = p^n i(f)$ for all $n \geq 0$.
(2) Otherwise, if $i_1(f) < (p^2 - p + 1)i(f)$, then

$$i_n(f) = i(f) + (1 + \cdots + p^{n-1})(i_1(f) - i(f))$$

It follows from Theorem 7.11 that if $f \in \mathcal{N}(\mathbf{F}_p)$ with $i(f) = b$, $i_1(f) = b + bp$ and $p \nmid b$, then $i_n(f) = b(1 + \cdots + p^n)$. We write that a power series with ramification of this form is b-ramified, and in answering Question 7.5 we seek to classify the b-ramified power series for certain values of b. The special case $b = 2$ was solved in [26, Theorem 1]:

Theorem 7.12 (Nordqvist 2016)**.** *A given power series of infinite order given by $f(X) = X + \sum_{i=1}^{\infty} a_i X^{i+2} \in \mathcal{N}(\mathbf{F}_p)$ is 2-ramified if and only if $3a_3^3 + 2a_4^2 - 2a_2a_5 \neq 0$ and $a_1 \neq 0$.*

In the next section, we will prove a similar result for the b-ramified power series under the hypothesis that $b^2 < p$.

8. Classifying Power Series by Ramification

Let

$$f(X) = X + a_1X^{b+1} + a_2X^{b+2} + \cdots \in \mathcal{N}(\mathbf{F}_p)$$

be of infinite order in $\mathcal{N}(\mathbf{F}_p)$, so that $i_0(f) = b$. By Theorem 7.11, f is b-ramified if and only if $i_1(f) = b + bp$. Therefore, it suffices to compute the coefficients on the terms of degree at most $b + bp + 1$ in $f^{\circ p}$ in terms of the a_i's. To do this, we rely on a technique established by Lindahl–Rivera-Letelier in [20, Lemma 3.6]. In particular, define the power series $\{\Delta_m\}_{m\geq 1}$ recursively by $\Delta_1 = f(X) - X$ and

$$\Delta_{m+1} = \Delta_m \circ f - \Delta_m.$$

It is straightforward to check the following lemma by induction:

Lemma 8.1. *Suppose $m \in \mathbf{N}$. Then,*

$$\Delta_m(X) = \sum_{i=0}^{m} (-1)^i \binom{m}{i} f^{\circ(m-i)}(X).$$

Since $\mathbf{F}_p$ has characteristic p, this corollary is immediate:

Corollary 8.2. $\Delta_p(X) = f^{\circ p}(X) - X$.

Therefore, to compute the coefficients of $f^{\circ p}$, it suffices to compute those of Δ_p. In particular, we wish to compute Δ_p modulo X^{b+bp+2}, since this will provide us with all the coefficients of f on the terms of degree at most $b + bp + 1$, as desired.

To do this, we prove a simple lemma that allows us to set up the necessary computational infrastructure.

Lemma 8.3. *$\mathrm{val}_X(\Delta_m) \geq bm + 1$ for all $m \in \mathbf{N}$.*

Proof. We proceed by induction. For the base case, we have already that

$$\Delta_1 = f(X) - X = a_1X^{b+1} + \cdots,$$

so $\mathrm{val}_X(\Delta_1) = b + 1$ as desired. Now, assume the inductive hypothesis that $\mathrm{val}_X(\Delta_m) \geq bm + 1$ for some $m \in \mathbf{N}$. This means that we can write

$$\Delta_m = A_1(m)X^{bm+1} + A_2(m)X^{bm+2} + \cdots$$

for some $A_1(m), \ldots \in \mathbf{F}_p$. From here it is clear (from the definition of f) that the term of least degree in

$$\Delta_{m+1}(X) = (A_1(m)f(X)^{bm+1} + A_2(m)f(X)^{bm+2} + \cdots) - (A_1(m)X^{bm+1} + A_2(m)X^{bm+2} + \cdots)$$

has degree at least $b(m + 1) + 1$ [in particular the term of least degree that is not guaranteed to be cancelled by subtracting Δ_m is produced by the term of $A_1(m)f(X)^{bm+1}$ produced by choosing bm copies of X and one copy of X^{b+1}]. The desired result follows by induction. □

Corollary 8.4. $i_1(f) \geq bp$.

Corollary 8.5. $i_1(f) \geq bp + b$

Proof. By Sen's Theorem, $i_1(f) \equiv i_0(f) = b \pmod{p}$, which means that $i_1(f) = b + kp$ for some integer k. Of course, if $k < b$, then by the assumption that $p > b$ (recall we actually assumed a slightly stronger bound), we have

$$i_1(f) = b + kp \leq b + (b - 1)p = bp + (b - p) < bp,$$

which contradicts the previous corollary. □

The above lemma means that there are no terms of degree at most bp in Δ_p. Thus, we only need to consider the terms of degree d for $bp+1 \leq d \leq b+bp+1$. In the language of the proof of the lemma, we need only to compute the coefficients $A_1(p), \ldots, A_{b+1}(p)$ in terms of the a_i's. Of course, the above corollary gives us for free that $A_1(p), \ldots, A_b(p)$ are identically zero. It remains to compute $A_{b+1}(p)$ in terms of the a_i's, which will in turn give us a criterion for b-ramification: f is b-ramified if and only if $A_{b+1}(p) \neq 0$. To do this, we compute using the recursive definition of the $A_i(m)$'s. In particular, we can compute modulo $X^{b+b(m+1)+2}$ that (since $f(X)^k$ for $k > b + bm + 1$ can contribute no term that isn't cancelled by subtracting $\Delta_m(X)$)

$$\begin{aligned}\Delta_{m+1}(X) &\equiv A_1(m)f(X)^{bm+1} + A_2(m)f(X)^{bm+2} + \cdots + A_{b+1}f(X)^{b+bm+1} - \Delta_m(X) \\ &\equiv (A_1(m)a_1(bm+1))X^{b(m+1)+1} + (A_1(m)a_2(bm+1) + A_2(m)a_1(bm+2))X^{b(m+1)+2} \\ &+ \cdots + (A_1(m)a_b(bm+1) + \cdots + A_b(m)a_1(bm+b))X^{b(m+1)+b} \\ &+ \left(A_1(m)a_1^2\binom{bm+1}{2} + A_1(m)a_{b+1}(bm+1) + \cdots + A_{b+1}(m)a_1(bm+b+1)\right)X^{b(m+1)+b+1}.\end{aligned}$$

Recalling from the definition of Δ_1 that $A_i(1) = a_i$, this is equivalent to defining the $A_i(m)$'s by the recurrence

$$\begin{pmatrix} A_1(m+1) \\ A_2(m+1) \\ \vdots \\ A_b(m+1) \\ A_{b+1}(m+1) \end{pmatrix} = \begin{pmatrix} a_1(bm+1) & & & & \\ a_2(bm+1) & a_1(bm+2) & & & \\ \vdots & \vdots & \ddots & & \\ a_b(bm+1) & \cdots & a_1(bm+b) & & \\ a_1^2\binom{bm+1}{2} + a_{b+1}(bm+1) & a_b(bm+2) & \cdots & a_1(bm+b+1) & \end{pmatrix} \begin{pmatrix} A_1(m) \\ A_2(m) \\ \vdots \\ A_b(m) \\ A_{b+1}(m) \end{pmatrix}$$

with initial conditions

$$\begin{pmatrix} A_1(1) \\ A_2(1) \\ \vdots \\ A_b(1) \\ A_{b+1}(1) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_b \\ a_{b+1} \end{pmatrix}.$$

Recursively solving these difference equations using [7, Section 1.2], one arrives at a "closed" form for these coefficients:

Proposition 8.6. *$A_i(m)$ is given by the following. For $k < b$,*

$$\begin{aligned}A_k(m) = a_k a_1^{m-1}(bm-b+k)!^{(b)} + \sum_{k>\alpha_1>\cdots>\alpha_n>0} \left(a_1^{m-n-1}a_{k-\alpha_1+1}a_{\alpha_n}\prod_{i=1}^{n-1} a_{\alpha_i-\alpha_{i-1}+1}\right) \cdot \\ \sum_{i_1=1}^{m-1}\sum_{i_2=1}^{i_1-1}\cdots\sum_{i_n=1}^{i_{n-1}-1} \frac{(bm-b+k)!^{(b)}(bi_1+\alpha_1)!^{(b)}\cdots(bi_n+\alpha_n)!^{(b)}}{(bi_1+k)!^{(b)}(bi_2+\alpha_1)!^{(b)}\cdots(bi_n+\alpha_{n-1})!^{(b)}}.\end{aligned}$$

For $k = b$,

$$A_b(m) = a_b a_1^{m-1} \frac{(bm)!^{(b)}}{b} + \sum_{k>\alpha_1>\cdots>\alpha_n>0} \left(a_1^{m-n-1} a_{b-\alpha_1+1} a_{\alpha_n} \prod_{i=1}^{n-1} a_{\alpha_i-\alpha_{i-1}+1} \right) \cdot \sum_{i_1=1}^{m-1} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_n=1}^{i_{n-1}-1} \frac{(bm)!^{(b)}(bi_1+\alpha_1)!^{(b)} \cdots (bi_n+\alpha_n)!^{(b)}}{(bi_1+b)!^{(b)}(bi_2+\alpha_1)!^{(b)} \cdots (bi_n+\alpha_{n-1})!^{(b)}}.$$

And for $k = b+1$,

$$\begin{aligned} A_{b+1}(m) = {} & a_{b+1} a_1^{m-1} \frac{(bm+1)!^{(b)}}{bm+1} + a_1^{m+1} \frac{b}{2} \sum_{r=1}^{m-1} \frac{(bm+1)!^{(b)}(br+1)!^{(b)}}{(br+b+1)!^{(b)}} r \\ & + a_b a_2 a_1^{m-2} \frac{1}{b} \sum_{r=1}^{m-1} \frac{(bm+1)!^{(b)}(br+b)!^{(b)}}{(br+b+1)!^{(b)}} \\ & + \sum_{\substack{b \geq \alpha_1 > \cdots > \alpha_n > 0 \\ \{\alpha_i\} \neq \{b\}}} \left(a_1^{m-n-1} a_{b-\alpha_1+2} a_{\alpha_n} \prod_{i=1}^{n-1} a_{\alpha_i-\alpha_{i-1}+1} \right) \cdot \\ & \sum_{i_1=1}^{m-1} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_n=1}^{i_{n-1}-1} \frac{(bm+1)!^{(b)}(bi_1+\alpha_1)!^{(b)} \cdots (bi_n+\alpha_n)!^{(b)}}{(bi_1+b+1)!^{(b)}(bi_2+\alpha_1)!^{(b)} \cdots (bi_n+\alpha_{n-1})!^{(b)}}. \end{aligned}$$

The first two of these sums we already know are identically zero by Sen's Theorem (see above remarks). Thus, it suffices to compute A_{b+1} modulo p.

The first step is to reduce the computation of the b-tuple factorials to single factorials modulo p via the following lemma:

Lemma 8.7. *Let $y \in \{0, 1, \ldots, b-1\}$ and $x \in \mathbb{Z}$. Then in $\mathbb{F}_p$,*

$$(bx+y)!^{(b)} = b^x \frac{(x+yt)!}{(yt)!}$$

where t is the smallest lift to $\mathbf{Z}$ of the inverse of b in $\mathbf{F}_p$.

It follows from this discussion that

$$\begin{aligned} A_{b+1}(p) = {} & a_{b+1} a_1^{p-1} b^p \frac{(p+t)!}{t!(b+1)} + a_1^{p+1} \frac{b}{2} \sum_{r=1}^{p-1} \frac{(p+t)!(r+t)!}{(r+1+t)!t!} b^{p-1} r + a_b a_2 a_1^{p-2} \sum_{r=1}^{p-1} b^p \frac{(p+t)!(r+1)!}{(r+1+t)!} \\ & + b^{p-1} \sum_{b>\alpha_1>\cdots>\alpha_n>0} \left(a_1^{p-n-1} a_{b-\alpha_1+2} a_{\alpha_n} \prod_{i=1}^{n-1} a_{\alpha_i-\alpha_{i+1}+1} \right) \cdot \\ & \sum_{i_1=1}^{p-1} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_n=1}^{i_{n-1}-1} \frac{(p+t)!(i_1+\alpha_1 t)! \cdots (i_n+\alpha_n t)!}{(i_1+1+t)!(i_2+\alpha_1 t)! \cdots (i_n+\alpha_{n-1} t)!(\alpha_n t)!} \\ & + b^{p-1} \sum_{b=\alpha_1>\cdots>\alpha_n>0} \left(a_1^{p-n-1} a_2 a_{\alpha_n} \prod_{i=1}^{n-1} a_{\alpha_i-\alpha_{i+1}+1} \right) \cdot \\ & \sum_{i_1=1}^{p-1} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_n=1}^{i_{n-1}-1} \frac{(p+t)!(i_1+1)!(i_2+\alpha_2 t)! \cdots (i_n+\alpha_n t)!}{(i_1+1+t)!(i_2+1)!(i_3+\alpha_2 t)! \cdots (i_n+\alpha_{n-1} t)!(\alpha_n t)!}. \end{aligned}$$

Let the inner n-tuple sums in the fourth and fifth terms be called $\varphi_1(\alpha_1, \dots, \alpha_n)$ and $\varphi_2(\alpha_1, \dots, \alpha_n)$, respectively. We compute these terms one at a time:

Proposition 8.8.

$$b^p \frac{(p+t)!}{t!(b+1)!} = b^p \sum_{r=1}^{p-1} \frac{(p+t)!(r+1)!}{(r+1+t)!} = 0, \qquad \frac{b}{2} \sum_{r=1}^{p-1} \frac{(p+t)!(r+t)!}{(r+1+t)!t!} b^{p-1} r = \frac{b+1}{2}.$$

Proof. The fact that the first term vanishes is obvious from the assumption that $p > b+1$. Moreover, since $(r+1)! = (-1)^r/(p-r-2)!$ for positive $r < p-1$ (and the $r = p-1, r < p-t-1$ terms are obviously zero) we have

$$\begin{aligned}
\sum_{r=1}^{p-1} \frac{(p+t)!(r+1)!)}{(r+1+t)!} &= \sum_{r=p-t-1}^{p-2} (-1)^r \frac{(p+t)!}{(r+1+t)!(p-r-2)!} \\
&= (p+t) \sum_{r=p-t-1}^{p-2} (-1)^r \binom{p+t-1}{r+t+1} \\
&= t \sum_{i=0}^{t-1} (-1)^{i+p-t-1} \binom{p+t-1}{p+i} \\
&= (-1)^t t \sum_{i=0}^{t-1} (-1)^i \binom{t-1}{i} \\
&= 0
\end{aligned}$$

by the binomial theorem, as desired. Finally, the only nonzero term in the third sum $\sum_{r=1}^{p-1} \frac{(p+t)!(r+t)!}{(r+1+t)!t!} b^{p-1} r$ occurs when $r+1+t = p$, i.e. $r = p-t-1$. Therefore,

$$\begin{aligned}
\frac{b}{2} \sum_{r=1}^{p-1} \frac{(p+t)!(r+t)!}{(r+1+t)!t!} b^{p-1} r &= \frac{b}{2} \cdot \frac{(p+t)!(p-1)!}{(p)!t!} b^{p-1}(p-t-1) \\
&= -\frac{b}{2} \cdot \frac{p!t!}{p!t!} b^{p-1}(p-t-1) \\
&= \frac{b}{2} \cdot (t+1) \\
&= \frac{b+1}{2}
\end{aligned}$$

which proves the proposition. □

For the remaining two terms, we proceed by induction. Let $n \in \mathbf{N}$, $\beta < b$ and $b > \alpha_1 > \cdots > \alpha_n > 0$ be integers. Then for any $i_{n-1} < p$ we have

$$\sum_{i_n=1}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!} = \sum_{i_n=1}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p)! \lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor !}{(i_n + \alpha_{n-1} t)!((\alpha_{n-1} - \alpha_n)p)!(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor)}$$

$$= \sum_{i_n=1}^{i_{n-1}-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_n + \alpha_{n-1} t} \frac{\lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor !((\alpha_{n-1} - \alpha_n)(p-t) + \beta)!}{((\alpha_{n-1} - \alpha_n)p)!(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor)}.$$

Of course, $\lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor$ can only take on the values $\lfloor \frac{\alpha_n t + \beta}{p} \rfloor$ and $\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1$ depending on whether i_n is sufficiently large (since we know a priori that $i_n < p$). In particular, the former value is taken on if and only if $i_n < \lceil \alpha_n t + \beta \rceil_p - (\alpha_n t + \beta)$

(here $\lceil x \rceil_p$ denotes the least multiple of p which is greater than x, NB this means $\lceil kp \rceil_p = (k+1)p$). If $i_{n-1} < \lceil \alpha_n t + \beta \rceil_p - (\alpha_n t + \beta)$, then $i_n \leq i_{n-1} - 1 < \lceil \alpha_n t + \beta \rceil_p - (\alpha_n t + \beta) - 1$, so we have $\lfloor \frac{i_n + \alpha_n t + \beta}{p} \rfloor = \lfloor \frac{\alpha_n t + \beta}{p} \rfloor$ (in particular it doesn't depend on any of the indices in the sum). Therefore, in this case we have

$$\sum_{i_n=1}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!} = \frac{\lfloor \frac{\alpha_n t + \beta}{p} \rfloor! ((\alpha_{n-1} - \alpha_n)(p-t) + \beta)!}{((\alpha_{n-1} - \alpha_n)p)!(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{\alpha_n t + \beta}{p} \rfloor)} \cdot \sum_{i_n=1}^{i_{n-1}-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_n + \alpha_{n-1} t}.$$

The sum on the right hand side is

$$\sum_{i_n=1}^{i_{n-1}-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_n + \alpha_{n-1} t} = \binom{i_{n-1} + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_{n-1} + \alpha_{n-1} t - 1} - \binom{1 + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{\alpha_{n-1} t}$$

$$= \frac{(i_{n-1} + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p)!}{(i_{n-1} + \alpha_{n-1} t - 1)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!} - \frac{(1 + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p)!}{(\alpha_{n-1} t)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}$$

$$= \frac{(i_{n-1} + \alpha_n t + \beta)!(\alpha_{n-1} - \alpha_n)p)! \cdot \frac{(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{i_{n-1} + \alpha_n t + \beta}{p} \rfloor)}{\lfloor \frac{i_{n-1} + \alpha_n t + \beta}{p} \rfloor!}}{(i_{n-1} + \alpha_n t - 1)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}$$

$$- \frac{(1 + \alpha_n t + \beta)!(\alpha_{n-1} - \alpha_n)p)! \frac{(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{1 + \alpha_n t + \beta}{p} \rfloor)}{\lfloor \frac{1 + \alpha_n t + \beta}{p} \rfloor!}}{(\alpha_{n-1} t)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}.$$

Recall the assumption that $i_{n-1} < \lceil \alpha_n t + \beta \rceil_p - (\alpha_n t + \beta)$, from which it follows that $\lfloor \frac{i_{n-1} + \alpha_n t + \beta}{p} \rfloor = \lfloor \frac{\alpha_n t + \beta}{p} \rfloor$ and thus

$$\sum_{i_n=1}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!} = \frac{(i_{n-1} + \alpha_n t + \beta)!}{(i_{n-1} + \alpha_{n-1} t - 1)!(\beta + 1 - (\alpha_{n-1} - \alpha_n)t)} - \frac{(\alpha_n t + \beta + 1)! \ell(\alpha_{n-1}, \alpha_n)}{(\alpha_{n-1} t)!(\beta + 1 - (\alpha_{n-1} - \alpha_n)t)}$$

where $\ell(\alpha_{n-1}, \alpha_n)$ is defined to be 1 if $\lfloor \frac{1 + \alpha_n t + \beta}{p} \rfloor = \lfloor \frac{\alpha_n t + \beta}{p} \rfloor$ (i.e. if $\alpha_n t + \beta \neq -1$ modulo p) and otherwise

$$\ell(\alpha_{n-1}, \alpha_n) := \frac{\alpha_{n-1} - \alpha_n + \lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1}{\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1}.$$

In the remaining case, $i_{n-1} \geq \lceil \alpha_n t + \beta \rceil_p - (\alpha_n t + \beta)$ and so we have

$$\sum_{i_n=1}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!} = \sum_{i_n=1}^{\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!} + \sum_{i_n=\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)}^{i_{n-1}-1} \frac{(i_n + \alpha_n t + \beta)!}{(i_n + \alpha_{n-1} t)!}$$

$$= \frac{((\alpha_{n-1} - \alpha_n)(p-t) + \beta)! \lfloor \frac{\alpha_n t + \beta}{p} \rfloor !}{((\alpha_{n-1} - \alpha_n)p)!(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{\alpha_n t + \beta}{p} \rfloor)} \cdot$$

$$\left(\sum_{i_n=1}^{\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)t}{i_n + \alpha_{n-1} t} \right.$$

$$\left. + \frac{\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1}{\alpha_{n-1} - \alpha_n + \lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1} \sum_{i_n=\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)}^{i_{n-1}-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)t}{i_n + \alpha_{n-1} t} \right).$$

We also have

$$\sum_{i_n=1}^{\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_n + \alpha_{n-1} t}$$

$$= \binom{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)p}{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)t - \beta - 1} - \binom{\alpha_n t + \beta + 1 + (\alpha_{n-1} - \alpha_n)p}{\alpha_{n-1} t}$$

while

$$\sum_{i_n=\lceil \alpha_n t+\beta \rceil_p - (\alpha_n t+\beta)}^{i_{n-1}-1} \binom{i_n + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_n + \alpha_{n-1} t}$$

$$= \binom{i_{n-1} + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_{n-1} + \alpha_{n-1} t - 1} - \binom{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)p}{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)t - \beta - 1}$$

As before, we can write

$$\binom{i_{n-1} + \alpha_n t + \beta + (\alpha_{n-1} - \alpha_n)p}{i_{n-1} + \alpha_{n-1} t - 1} = \frac{(i_{n-1} + \alpha_n t + \beta)!((\alpha_{n-1} - \alpha_n)p)! \frac{(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{i_{n-1} + \alpha_n t + \beta}{p} \rfloor)}{\lfloor \frac{i_{n-1} + \alpha_n t + \beta}{p} \rfloor !}}{(i_{n-1} + \alpha_{n-1} t - 1)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}$$

and under the assumption that $\beta + 1 \leq (\alpha_{n-1} - \alpha_n)t$,

$$\binom{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)p}{\lceil \alpha_n t + \beta \rceil_p + (\alpha_{n-1} - \alpha_n)t - \beta - 1}$$

is

$$\frac{((\alpha_{n-1} - \alpha_n)p)!(\lceil \alpha_n t + \beta \rceil_p)! \frac{(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1)}{(\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1)!}}{(\lceil \alpha_n t + \beta \rceil_p)!((\alpha_{n-1} - \alpha_n)t - \beta - 1)! \frac{(\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 2) \cdots (\lfloor \frac{\alpha_n t + \beta}{p} \rfloor + 1 + \lfloor \frac{(\alpha_{n-1} - \alpha_n)t - \beta - 1}{p} \rfloor)}{\lfloor \frac{(\alpha_{n-1} - \alpha_n)t - \beta - 1}{p} \rfloor !} ((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}$$

and

$$\binom{\alpha_n t + \beta + 1 + (\alpha_{n-1} - \alpha_n)p}{\alpha_{n-1} t} = \frac{(\alpha_n t + \beta + 1)!((\alpha_{n-1} - \alpha_n)p)! \frac{(\alpha_{n-1} - \alpha_n + 1) \cdots (\alpha_{n-1} - \alpha_n + \lfloor \frac{1 + \alpha_n t}{p} \rfloor)}{\lfloor \frac{1 + \alpha_n t}{p} \rfloor !}}{(\alpha_{n-1} t)!((\alpha_{n-1} - \alpha_n)(p-t) + \beta + 1)!}$$

so we can write (from the above expansion)

$$\sum_{i_n=1}^{i_{n-1}-1}\frac{(i_n+\alpha_n t+\beta)!}{(i_n+\alpha_{n-1}t)!}=\frac{((\alpha_{n-1}-\alpha_n)(p-t)+\beta)!\lfloor\frac{\alpha_n t+\beta}{p}\rfloor!}{((\alpha_{n-1}-\alpha_n)p)!(\alpha_{n-1}-\alpha_n+1)\cdots(\alpha_{n-1}-\alpha_n+\lfloor\frac{\alpha_n t+\beta}{p}\rfloor)}\cdot$$
$$\left(\binom{\lceil\alpha_n t+\beta\rceil_p+(\alpha_{n-1}-\alpha_n)p}{\lceil\alpha_n t+\beta\rceil_p+(\alpha_{n-1}-\alpha_n)t-\beta-1}-\binom{\alpha_n t+\beta+1+(\alpha_{n-1}-\alpha_n)p}{\alpha_{n-1}t}\right.$$
$$\left.+\frac{\lfloor\frac{\alpha_n t+\beta}{p}\rfloor+1}{\alpha_{n-1}-\alpha_n+\lfloor\frac{\alpha_n t+\beta}{p}\rfloor+1}\left(\binom{i_{n-1}+\alpha_n t+\beta+(\alpha_{n-1}-\alpha_n)p}{i_{n-1}+\alpha_{n-1}t-1}-\binom{\lceil\alpha_n t+\beta\rceil_p+(\alpha_{n-1}-\alpha_n)p)}{\lceil\alpha_n t+\beta\rceil_p+(\alpha_{n-1}-\alpha_n)t-\beta-1}\right)\right)$$
$$=\frac{(\alpha_1-\alpha_2)\lfloor\frac{(\alpha_{n-1}-\alpha_n)t-\beta-1}{p}\rfloor!}{((\alpha_{n-1}-\alpha_n)(p-t)+\beta+1)(\lfloor\frac{\alpha_n t+\beta}{p}\rfloor+1)((\alpha_{n-1}-\alpha_n)t-\beta-1)!}$$
$$\cdot\frac{1}{(\lfloor\frac{\alpha_n t+\beta}{p}\rfloor+2)\cdots(\lfloor\frac{\alpha_n t+\beta}{p}\rfloor+1+\lfloor\frac{(\alpha_{n-1}-\alpha_n)t-\beta-1}{p}\rfloor)}$$
$$-\frac{\ell(\alpha_{n-1},\alpha_n)}{((\alpha_{n-1}-\alpha_n)(p-t)+\beta+1)}\frac{(\alpha_n t+\beta+1)!}{(\alpha_{n-1}t)!}+\frac{(i_{n-1}+\alpha_n t+\beta)!}{(i_{n-1}+\alpha_{n-1}t-1)!((\alpha_{n-1}-\alpha_n)(p-t)+\beta+1)}.$$

Let $\varphi_\beta(\alpha_1,\ldots,\alpha_n)$ denote

$$\varphi_\beta(\alpha_1,\ldots,\alpha_n):=\sum_{i_1=1}^{p-1}\cdots\sum_{i_n=1}^{i_{n-1}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_n+\alpha_n t+\beta)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_n+\alpha_{n-1}t)!(\alpha_n)!}$$

(NB for $\beta=0$ this is the fourth term in the sum we want to compute). Since the last two terms in our expression for the last inner sum are the same as the entire expression in the case $i_{n-1}<\lceil\alpha_n t+\beta\rceil_p-(\alpha_n t+\beta)$, we can conclude that $\varphi_\beta(\alpha_1,\ldots,\alpha_n)$ is

$$c_1\sum_{i_1=1}^{p-1}\cdots\sum_{i_{n-1}=\lceil\alpha_n t+\beta\rceil_p-(\alpha_n+\beta)}^{i_{n-2}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_n t)!((\alpha_{n-1}-\alpha_n)t-\beta-1)!}$$
$$+c_2\sum_{i_1=1}^{p-1}\cdots\sum_{i_{n-1}=1}^{i_{n-2}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!(\alpha_n t+\beta+1)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_n t)!(\alpha_{n-1}t)!}$$
$$+\frac{1}{1+\beta-(\alpha_{n-1}-\alpha_n)t}\sum_{i_1=1}^{p-1}\cdots\sum_{i_{n-1}=1}^{i_{n-2}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!(i_{n-1}+\alpha_n t+\beta)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_n t)!(i_{n-1}+\alpha_{n-1}t-1)!}.$$

where $c_1,c_2\in\mathbf{F}_p^*$ (i.e. they have no factors of p in numerator or denominator) and they depend only on the α_i's. Note that the terms of the first sum are of the form

$$\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_n t)!((\alpha_{n-1}-\alpha_n)t-\beta-1)!}=c\frac{(i_{n-1}+\alpha_{n-1}t)!}{(\alpha_n t)!((\alpha_{n-1}-\alpha_n)t-\beta-1)!}$$

where c has a nonnegative number of factors of p. In fact, we have

$$\frac{(i_{n-1}+\alpha_{n-1}t)!}{(\alpha_n t)!((\alpha_{n-1}-\alpha_n)t-\beta-1)!}=\frac{(i_{n-1}+\alpha_n t+\beta+(\alpha_{n-1}-\alpha_n)t-\beta)!}{(\alpha_n t)!((\alpha_{n-1}-\alpha_n)t-\beta-1)!}$$

and from the fact that $i_{n-1}+\alpha_n t+\beta\geq\lceil\alpha_n t+\beta\rceil_p$ it is immediate that this is zero modulo p. Therefore, each term of the first sum is zero, and we can ignore it

altogether (this is even more obvious in the case $\beta+1>(\alpha_{n-1}-\alpha_n)t$, though we don't include it here). The second sum is equal to

$$c_2\frac{(\alpha_n t+\beta+1)!}{(\alpha_n t)!}\sum_{i_1=1}^{p-1}\cdots\sum_{i_{n-1}=1}^{i_{n-2}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_{n-1}t)!},$$

which is equal to

$$c_2\frac{(\alpha_n t+\beta+1)!}{(\alpha_n t)!}\varphi_0(\alpha_1,\ldots,\alpha_{n-1})$$

which is zero by induction (since $\alpha_{n-1}>\alpha_n\geq 1$). Therefore,

$$\begin{aligned}\varphi_\beta(\alpha_1,\ldots,\alpha_n)&=\frac{1}{1-(\alpha_{n-1}-\alpha_n)t}\cdot\\&\quad\sum_{i_1=1}^{p-1}\cdots\sum_{i_{n-1}=1}^{i_{n-2}-1}\frac{(p+t)!(i_1+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-1}t)!(i_{n-1}+\alpha_n t+\beta)!}{(i_1+1+t)!(i_2+\alpha_1 t)!\cdots(i_{n-1}+\alpha_{n-2}t)!(\alpha_n t)!(i_{n-1}+\alpha_{n-1}t-1)!}\\&=\frac{1}{1-(\alpha_{n-1}-\alpha_n)t}\varphi_{\beta+1}(\alpha_1,\ldots,\alpha_{n-2},\alpha_n)-\varphi_\beta(\alpha_1,\ldots,\alpha_{n-2},\alpha_n)\\&=-\varphi_\beta(\alpha_1,\ldots,\alpha_{n-2},\alpha_n)\end{aligned}$$

by induction since $\beta+1>0$.

The base case ($n=1$) is easy: if $\alpha>1$ or if $\beta>0$, then $i_1+\alpha t+\beta\geq i_1+1+t$ and we can write

$$\begin{aligned}\varphi_\beta(\alpha)&=\sum_{i_1=1}^{p-1}\frac{(p+t)!(i_1+\alpha t+\beta)!}{(i_1+1+t)!(\alpha t)!}=\frac{(p+t)!((\alpha-1)t+\beta-1)!}{(\alpha t)!}\sum_{i_1=1}^{p-1}\binom{i_1+\alpha t+\beta}{i_1+1+t}\\&=\frac{(p+t)!((\alpha-1)t+\beta-1)!}{(\alpha t)!}\left(\binom{p+\alpha t+\beta}{p+t}-\binom{1+\alpha t+\beta}{1+t}\right)\\&=0.\end{aligned}$$

On the other hand, if $\alpha=1$ and $\beta=0$, we find that

$$\varphi_\beta(\alpha)=\sum_{i_1=1}^{p-1}\frac{(p+t)!(i_1+t)!}{(i_1+1+t)!(t)!},$$

and the only nonzero term in this sum occurs when $i_1+1+t=p$. Thus,

$$\varphi_\beta(\alpha)=\frac{(p+t)!(p-1)!}{p!t!}=-1.$$

From all of this, it follows that

$$\varphi_\beta(\alpha_1,\ldots,\alpha_n)=(-1)^n\delta(\alpha_n,1)\delta(\beta,0).$$

Therefore,

$$A_{b+1}(p)=\frac{b+1}{2}a_1^{p+1}+\sum_{n=1}^{b}(-1)^n a_1^{p-n}\sum_{b\geq\alpha_1>\cdots>\alpha_n=1}a_{b-\alpha_1+2}\prod_{i=1}^{n-1}a_{\alpha_i-\alpha_{i+1}+1}.$$

Since $a_1\neq 0$ and $\mathbf{F}_p$ has characteristic $p>2$, we can multiply by $2a_1^{b-1}$ and apply Fermat's little theorem to get that f is b-ramified if and only if

$$(b+1)a_1^{b+1}+2\sum_{n=1}^{b}(-1)^n a_1^{b-n}\sum_{b\geq\alpha_1>\cdots>\alpha_n=1}a_{b-\alpha_1+2}\prod_{i=1}^{n-1}a_{\alpha_i-\alpha_{i+1}+1}\neq 0$$

Fixing n, the term $a_2^{e_2}\cdots a_{b+1}^{e_{b+1}}$ shows up exactly $\binom{e_2+\cdots+e_{b+1}}{e_2,\ldots,e_{b+1}}$ times. Since $n = e_2+\cdots+e_{b+1}$ (in particular it is determined by the e_i's), any monomial can only appear for at most one value of n. In fact, $a_2^{e_2}\cdots a_{b+1}^{e_{b+1}}$ shows up if and only if $e_2+2e_3+\cdots+be_{b+1}=b$. In this case,

$$e_1 = b-n = b-(e_2+\cdots+e_{b+1})$$

so this condition is equivalent to $a_1^{e_1}\cdots a_{b+1}^{e_{b+1}}$ appearing in the entire sum if and only if

$$e_1+2e_2+\cdots+(b+1)e_{b+1}=2b$$

and $e_1+e_2+\cdots+e_{b+1}=b$. So, we have finally proved the most concise version of our final result:

Theorem 8.9. *Suppose $p>b^2$. Then f is b-ramified if and only if*

$$(b+1)a_1^{b+1}+\sum_{\substack{e_1,\ldots,e_{b+1}\geq 0\\ e_1+\cdots+e_{b+1}=b\\ e_1+2e_2+\cdots+(b+1)e_{b+1}=2b}} 2(-1)^{e_2+\cdots+e_{b+1}}\binom{e_2+\cdots+e_{b+1}}{e_2,\ldots,e_{b+1}}a_1^{e_1}\cdots a_{b+1}^{e_{b+1}}=0.$$

This is a vast generalization of [26, Theorem 1], and points to possible progress on a generalization of [19, Theorem A].

Acknowledgments

I would like to thank my mentor Drew Moore for his helpful conversations regarding the new material and his encouragement, as well as J.P. May for organizing the University of Chicago REU. The last two sections of this paper are the result of a continuation of a collaboration with Hudson Kirkpatrick on a project proposed by Laurent Berger at PROMYS 2016. So, I also thank Hudson Kirkpatrick for his insights that led to our results, Laurent Berger for his mentoring on the project, and Glenn Stevens, Krishanu Shankar, David Fried, the PROMYS Foundation, and the Clay Mathematics Institute for supporting that project.

References

[1] F. M. Bleher, T. Chinburg, B. Poonen, and P. Symonds. Automorphisms of Harbater-Katz-Gabber curves. *ArXiv e-prints*, September 2015. `arXiv:1509.02139`.

[2] Rachel Camina. Subgroups of the Nottingham group. *Journal of Algebra*, 196(1):101–113, 1997.

[3] Keith Conrad. The different ideal. URL: `https://kconrad.math.uconn.edu/blurbs/gradnumthy/diff.pdf`.

[4] Keith Conrad. Discriminants and ramified primes. URL: `https://kconrad.math.uconn.edu/blurbs/gradnumthy/disc.pdf`.

[5] Keith Conrad. Ostrowski for number fields. URL: `https://kconrad.math.uconn.edu/blurbs/gradnumthy/ostrowskinumbfield.pdf`.

[6] David Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[7] Saber Elaydi. *An Introduction to Difference Equations*. Springer Science & Business Media, 2005.

[8] William Fulton. *Algebraic Curves*. Université de Versailles, 1969.

[9] Sandrine Jean. Conjugacy classes of series in positive characteristic and witt vectors. *Journal de Théorie des Nombres de Bordeaux*, 21(2):263–284, 2009.

[10] S.A. Jennings. Substitution groups of formal power series under substitution. *Canad. J. Math*, 6:325 – 340, 1954.

[11] D. L. Johnson. The group of formal power series under substitution. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 45(3):296–302, 04 2009.

[12] Kenz Kallal and Hudson Kirkpatrick. Ramification of wild automorphisms of laurent series fields. *arXiv preprints*, 2016. `arXiv:1611.01077`.

[13] Kevin Keating. Automorphisms and extensions of $k((t))$. *Journal of Number Theory*, 41(3):314 – 321, 1992.

[14] Benjamin Klopsch. Automorphisms of the Nottingham group. *Journal of Algebra*, 223(1):37 – 56, 2000.

[15] Serge Lang. *Algebraic Number Theory*, volume 110. Springer Science & Business Media, 2013.

[16] F Laubie and M Saïne. Ramification of automorphisms of $k((t))$. *Journal of Number Theory*, 63(1):143 – 145, 1997.

[17] François Laubie, Abbas Movahhedi, and Alain Salinier. Systèmes dynamiques non archimédiens et corps des norms (non-archimedean dynamic systems and fields of norms). *Compositio Mathematica*, 132(1):57–98, 2002.

[18] Karl-Olof Lindahl. The size of quadratic p-adic linearization disks. *Advances in Mathematics*, 248:872 – 894, 2013.

[19] Karl-Olof Lindahl and Jonas Nordqvist. Geometric location of periodic points of 2-ramified power series. *Journal of Mathematical Analysis and Applications*, 2018.

[20] Karl-Olof Lindahl and Juan Rivera-Letelier. Optimal cycles in ultrametric dynamics and minimally ramified power series. *Compositio Mathematica*, 152(1):187–222, 009 2015.

[21] Jonathan Lubin. Nonarchimedean dynamical systems. *Compositio Mathematica*, 94(3):321–346, 1994.

[22] Jonathan Lubin. Torsion in the Nottingham group. *Bulletin of the London Mathematical Society*, 43(3), 2011.

[23] Daniel Marcus. *Number Fields*, volume 8. Springer, 1977.

[24] Benjamin Muckenhoupt. Automorphisms of formal power series under substitution. *Transactions of the American Mathematical Society*, 99(3):373–383, 1961.

[25] Jürgen Neukirch. Algebraic number theory, volume 322 of grundlehren der mathematischen wissenschaften [fundamental principles of mathematical sciences], 1999.

[26] J. Nordqvist. Complete Classification of 2-ramified Power Series. *ArXiv e-prints*, January 2016. `arXiv:1601.03622`.

[27] J. Riviera-Letelier. *Dynamique des fonctions rationelles sur des corps locaux*. PhD thesis, Universite de Paris XI, 2000.

[28] Shankar Sen. On automorphisms of local fields. *Annals of Mathematics*, 90(1):33–46, 1969.

[29] J.P. Serre. *Corps locaux*. Hermann, Paris, 1968.

[30] I.O. York. The exponent of certain p-groups. *Proc. Edinburgh Math. Soc.*, 33:483–490, 1990.

[31] I.O. York. *The Group of Formal Power Series under Substitution*. PhD thesis, Nottingham University, 1990.