Ramification Theory for Local Fields with Imperfect Residue Fields

Liang Xiao

This note was originally written for my talk in STAGE. After that I put in some more content from [14, 15] in order to give a summary of the two papers. The argument here will be sloppy with no details but with more intuitive explanation. Thanks to Ivan Fesenko for useful comments.

1 Classical ramification theory

The ramification of a complete discretely valued field has been studied since the era of Hilbert, who introduced the notion of higher ramification groups. Later, Artin, Hasse, Tate, and other people made extraordinary contribution to generalize Hilbert's idea. A good reference for the classical ramification theory is Serre's fantastic book [13].

Notation 1.1. Let l/k be a finite Galois extension of complete discretely valued fields. Let $\mathcal{O}_k$, $\mathcal{O}_l$, π_k, π_l, $\bar{k}$, and $\bar{l}$ be rings of integers, uniformizers, and residue fields, respectively. For an element $a \in \mathcal{O}_l$, we use $\bar{a}$ to denote its reduction in $\bar{l}$.

Let $G = G_{l/k}$ be the Galois group. Use $v_l(\cdot)$ to denote the valuation on l so that $v_l(\pi_l) = 1$. We call $e = v_l(\pi_k)$ the *naïve ramification degree*; it is the index of the valuation group of k in that of l.

Hypothesis 1.2. In this section, we assume that the residue field k is perfect.

Definition 1.3. The most natural way to define higher ramification subgroups of the Galois group G is due to Hilbert:

$$g \in G_a \text{ if and only if } v_l(gx - x) \geq a + 1 \text{ for all } x \in \mathcal{O}_l.$$

Actually, we need only to test for the generator of $\mathcal{O}_l$ over $\mathcal{O}_k$; the extension of rings of integers is generated by one element [13, § III.6 Proposition 12]). In particular, $G_{-1} = G$, $G_0 = I$ is the inertia subgroup, and $G_1 = W$ is the wild inertia subgroup.

However, there is a disadvantage of this. Namely, it does not respect taking quotients and hence it does not give a filtration on the absolute Galois group G_k. Hasse and Herbrand independently defined a function ϕ and gave G an upper numbering filtration, which does extend to G_k. (We will give a working definition later.) Their ideas came from an attempt to describe the behavior of the norm map on the filtration on units. For more details, one may consult [6, Chap.3].

Definition 1.4. We set $G_u = G_{\lceil u \rceil}$ for $u \in [-1, \infty)$. Define

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

We set $G^{\phi(u)} = G_u$. The inverse of ϕ, denoted by ψ, is called the Hasse-Herbrand function.

Proposition 1.5. *It turns out, if H is a normal subgroup of G, then $(G/H)^v = G^v H/H$ for all v. Hence the upper numbering filtration patches to give a filtration $\mathrm{Fil}^v G_k$ on G_k.*

Definition 1.6. Let $\rho : G_k \to \mathrm{GL}(V)$ be a representation of finite local monodromy (the image of inertia is finite), where V is a finite dimensional vector space over a (topological) field of characteristic zero. Define the *Artin conductor* to be

$$\mathrm{Art}(\rho) \stackrel{\text{def}}{=} \sum_{a \in \mathbb{Q}_{\geq 0}} a \cdot \dim \left(V^{\mathrm{Fil}^{a+} G_k} / V^{\mathrm{Fil}^{a} G_k} \right), \tag{1.6.1}$$

where $\mathrm{Fil}^{a+} G_k = \overline{\cup_{b>a} \mathrm{Fil}^b G_k}$.

One can also define the *Swan conductor* to be

$$\mathrm{Swan}(\rho) \stackrel{\text{def}}{=} \sum_{a \in \mathbb{Q}_{\geq 1}} (a-1) \cdot \dim \left(V^{\mathrm{Fil}^{a+} G_k} / V^{\mathrm{Fil}^{a} G_k} \right), \tag{1.6.2}$$

which measures the wild ramification of ρ.

Theorem 1.7 (Hasse-Arf Theorem)**.** *The conductors $\mathrm{Art}(\rho)$ and $\mathrm{Swan}(\rho)$ are non-negative integers.*

In practise, we will consider irreducible representation ρ which exactly factors through a finite Galois extension l/k. In this case, let $b(l/k) = \max \left\{ b | G^b_{l/k} \neq \{1\} \right\}$ and then we have

$$\mathrm{Art}(\rho) = b(l/k) \cdot \dim \rho. \tag{1.7.1}$$

This $b(l/k)$ is called the *highest ramification break* of l/k.

$$b(l/k) \begin{cases} = 0 & \text{if } l/k \text{ is unramified,} \\ = 1 & \text{if } l/k \text{ is tamely ramified,} \\ > 1 & \text{if } l/k \text{ is wildly ramified.} \end{cases} \tag{1.7.2}$$

Proposition 1.8. *If $\mathcal{O}_l/\mathcal{O}_k$ is generated by one element x, we have an explicit formula for $b_{l/k}$ as follows.*

$$b_{l/k} = \frac{1}{e} \Big(\sum_{1 \neq g \in G} v_l(gx - x) + \max_{1 \neq g \in G} v_l(gx - x) \Big).$$

Alternative interpretation via rigid spaces

One of the main idea of Abbes-Saito's construction in imperfect residue field case comes from a reinterpretation of the number $b(l/k)$ via geometric connected components of certain rigid space.

Hypothesis 1.2 implies that $\mathcal{O}_l/\mathcal{O}_k$ is generated by one element x in $\mathcal{O}_l$ [13, § III.6 Proposition 12]). Let $P(x) = 0$ be its minimal polynomial.

Notation 1.9. We set $|\pi_k| = \theta$ as this number will be often mentioned later.

Proposition 1.10. *The rigid space* $X = \{u||u| \leq 1, |P(u)| < \theta^a\}$ *has* $[l : k]$ *geometric connected components if and only if* $a \geq b(l/k)$*.*

Proof. A rigorous proof can be found in [5, Lemme 2.4] or [1, Lemma 6.6]. We will give a rough idea of why this is true.

The picture here is that if a is very large, we confine u in very small neighborhoods of the roots of $P(u) = 0$, the conjugates of x. The rigid space X should be geometrically disjoint union of very small discs centered at each of conjugates of x. When a becomes smaller, the discs grow larger and, at some point, some of them crash into one disc, which decreases the number of geometric connected components.

The cut-off condition is obviously $|u - x| < \max_{1 \neq g \in G} |gx - x|$. Note that $P(u) = \prod_{g \in G}(u - gx)$. Hence, one has $|u - gx| = |gx - x|$. Thus,

$$|P(u)| = \prod_{g \in G} |u - gx| = |u - x| \prod_{1 \neq g \in G} |x - gx| < \theta^{b(l/k)}.$$

If one stares at this explanation for a moment, he may turn it into a complete proof. □

2 Abbes-Saito ramification filtrations

From now on, we drop Hypothesis 1.2 and consider the case when the residue field is imperfect. In this case, basic properties of the functions ϕ and ψ fail to hold and classical ramification theory is no longer applicable.

In the imperfect residue field case, the ramification theory is initiated by Kato in [7] for abelian representations and by Abbes and Saito in [1] for general case. Kedlaya [10] gave a different approach for equal characteristic case following the ideas of Christol, Matsuda, Mebkhout, and Tsuzuki.

When the residue field is imperfect, there are two kinds of wild ramifications. The first kind is typically given by an Eisenstein extension $y^e + \alpha_1 y^{e-1} + \cdots \alpha_e = 0$ with $p|e$. The second kind involves imperfect residue field extension but does not change the group of valuations, for example, $x^p = \alpha$ for $\bar{\alpha} \in \bar{k} \backslash \bar{k}^p$. Unfortunately, one can not separate these two kinds of ramifications like what we did for unramified extensions or tamely ramified extensions. After base change, the order of such sub-extensions may shuffle.

Another difficulty is that $\mathcal{O}_l/\mathcal{O}_k$ is no longer monogenic, for example, in the case when the residue field extension is $\mathbb{F}_p(x^{1/p}, y^{1/p})/\mathbb{F}_p(x, y)$. Thus, the naïve generalization of Proposition 1.8 is not possible. However, the rigid geometric interpretation can be generalized to this case. This is carried out by Abbes and Saito in [1] and [2].

Definition 2.1. Take $Z = (z_j)_{j\in J} \subset \mathcal{O}_l$ to be a finite set of elements generating $\mathcal{O}_l$ over $\mathcal{O}_k$, i.e., $\mathcal{O}_k[(u_j)_{j\in J}]/\mathcal{I} \simeq \mathcal{O}_l$ mapping u_j to z_j. Let $(f_i)_{i=1,\dots,n}$ be a finite set of generators of $\mathcal{I}$. Define the *Abbes-Saito space* to be

$$as^a_{l/k,Z} = \left\{ (u_J) \,\middle|\, \begin{array}{cc} |u_j| \le 1, & j \in J \\ \left|f_i\big((u_j)_{j\in J}\big)\right| \le \theta^a, & 1 \le i \le n \end{array} \right\}. \tag{2.1.1}$$

We denote the *geometric* connected components of $as^a_{l/k,Z}$ by $\pi_0^{\mathrm{geom}}(as^a_{l/k,Z})$. The *highest ramification break* $b(l/k)$ of the extension l/k is defined to be the minimal b such that $\forall a > b$, $\#\pi_0^{\mathrm{geom}}(as^a_{l/k,Z}) = [l : k]$.

The intuition here is that when $a \gg 0$, $as^a_{l/k,Z}$ is geometrically just small polydiscs around the solutions of $f_i((u_j)_{j\in J}) = 0$, and when $a \to 0^+$, $as^a_{l/k,Z}$ is approaching to the unit polydisc and, in particular, it is very likely to be geometrically connected. Thus, in the process of a starting from a very big number and getting smaller, there is an a when some components of $as^a_{l/k,Z}$ merge together. That is exactly the ramification break $b(l/k)$.

They also define a version of logarithmic ramification number, which will give rise to Swan conductors. We will not introduce the general definition, rather we will give a working definition when l/k is totally ramified.

Definition 2.2. Let $J = \{1, \dots, m\}$. We say that $\bar{k}$ has finite p-basis $(\bar{b}_1, \dots, \bar{b}_m) \in \bar{k}$ if $\{\bar{b}_1^{e_1} \cdots \bar{b}_m^{e_m} | e_j = 0, \dots, p-1 \text{ for } j \in J\}$ form a basis of $\bar{k}$ as a $\bar{k}^p$ vector space. Let b_j be a lift in $\mathcal{O}_k$ of $\bar{b}_j$ for each j. We also refer to $(b_1, \dots, b_m)$ as a p-basis of $\mathcal{O}_k$.

For the rest of this note, we always assume that $\bar{k}$ has a finite p-basis. We may always reduce to this case by certain limit argument.

Lemma 2.3. [14, Construction 3.3.5] or [15, Construction 2.1.6] *We can choose $c_1, \dots, c_m \in \mathcal{O}_l^\times$ and $b_1, \dots, b_m \in \mathcal{O}_k^\times$ such that $\mathbf{k}_j = \bar{k}(\bar{c}_1, \dots, \bar{c}_j)$ has p-basis $\bar{c}_1, \dots, \bar{c}_j, \bar{b}_{j+1}, \dots, \bar{b}_m$ and $\bar{c}_j^{p^{r_j}} = p^{r_j}$ for $p^{r_j} = [\mathbf{k}_j : \mathbf{k}_{j-1}]$ and all $j = 1, \dots, m$. We also choose the uniformizer π_l so that $\pi_k \equiv \pi_l^e \bmod \pi_l^{e+1}$. Thus, $c_1, \dots, c_m, \pi_l$ generate $\mathcal{O}_l$ over $\mathcal{O}_k$. More precisely,*

$$\left\{ c_1^{e_1} \cdots c_m^{e_m} \pi_l^i \,\middle|\, e_j \in \{0, \dots, p^{r_j} - 1\} \text{ for all } j, \text{ and } i \in \{0, \dots, e-1\} \right\} \tag{2.3.1}$$

form a basis of $\mathcal{O}_l$ over $\mathcal{O}_k$.

Let $\mathcal{O}_k\langle u_0, \dots, u_m\rangle/\mathcal{I} \xrightarrow{\sim} \mathcal{O}_l$ be the map sending u_j to c_j for $j = 1, \dots, m$ and u_0 to π_l, where the angle brackets mean to take the completion with respect to the Gauss norm. We will choose a set of generators $p_0, \dots, p_m$ of $\mathcal{I}$ as follow: for every $c_j^{p^{r_j}}$ or π_l^e, one can write it in terms of the basis listed in (2.3.1). This will give us an element p_j or p_0 in $\mathcal{I}$. Obviously,

$p_0, \ldots, p_m$ generate $\mathcal{I}$. Moreover,

$$\begin{array}{rcl} p_j & \in & u_j^{p^{r_j}} - \mathfrak{b}_j + (u_0, \pi_k) \cdot k[\![u_0, \ldots, u_m]\!], \\ p_0 & \in & u_0^e - \pi_k + (u_0\pi_k, \pi_k^2) \cdot k[\![u_0, \ldots, u_m]\!], \end{array}$$

where $\mathfrak{b}_j$ is a polynomial in $u_1, \ldots, u_{j-1}$ with coefficients in $\mathcal{O}_k$ for all $j = 1, \ldots, m$.

Definition 2.4. Keep the notation from Lemma 2.3. Define the standard *Abbes-Saito space* and *logarithmic Abbes-Saito space* to be

$$\begin{array}{rcl} as^a_{l/k} & = & \left\{ (u_0, \ldots, u_m) \middle| \begin{array}{c} |u_0| \leq 1, \ldots, |u_m| \leq 1, \\ |p_0(\underline{u})| \leq \theta^a, \ldots, |p_m(\underline{u})| \leq \theta^a \end{array} \right\} \\ \text{and } as^a_{l/k,\log} & = & \left\{ (u_0, \ldots, u_m) \middle| \begin{array}{c} |u_0| \leq 1, \ldots, |u_m| \leq 1, \\ |p_0(\underline{u})| \leq \theta^{a+1}, |p_1(\underline{u})| \leq \theta^a, \ldots, |p_m(\underline{u})| \leq \theta^a \end{array} \right\}, \end{array}$$

where $\underline{u} = (u_0, \ldots, u_m)$.

Similarly, the *highest logarithmic ramification break* $b_{\log}(l/k)$ of the extension l/k is defined to be the minimal b such that $\forall a > b$, $\#\pi_0^{\mathrm{geom}}(as^a_{l/k,\log}) = [l : k]$.

Theorem 2.5. *The Abbes-Saito spaces have the following properties.*

(1) *The ramification breaks $b(l/k)$ and $b_{\log}(l/k)$ depend only on l and k* [1, Section 3]. *In particular, we can use the standard Abbes-Saito spaces to compute the ramification breaks.*

(2) *The ramification break (resp. logarithmic ramification break) gives rise to a filtration of normal subgroups $\mathrm{Fil}^a G_k$ (resp., $\mathrm{Fil}^a_{\log} G_k$) on the Galois group G_k* [1, Theorem 3.3, 3.11]. *Moreover, for l/k a finite Galois extension, both highest ramification breaks are rational numbers* [1, Theorem 3.8, 3.16].

(3) *Let l/k be a finite separable Galois extension. If l/k is unramified, then $\mathrm{Fil}^a G_l = \mathrm{Fil}^a G_k$* [1, Proposition 3.7]. *If l/k is tamely ramified with ramification index m, then $\mathrm{Fil}^{ma}_{\log} G_l = \mathrm{Fil}^a_{\log} G_k$* [1, Proposition 3.15].

(4) *Define $\mathrm{Fil}^{a+} G_k = \overline{\cup_{b>a} \mathrm{Fil}^b G_k}$ (resp. $\mathrm{Fil}^{a+}_{\log} G_k = \overline{\cup_{b>a} \mathrm{Fil}^b_{\log} G_k}$). Then the subquotients $\mathrm{Fil}^a G_k / \mathrm{Fil}^{a+} G_k$ (resp. $\mathrm{Fil}^a_{\log} G_k / \mathrm{Fil}^{a+}_{\log} G_k$) are p-abelian groups for any $a \in \mathbb{Q}_{\geq 1}$ (resp. $a \in \mathbb{Q}_{\geq 0}$) and are 0 if $a \notin \mathbb{Q}$, except possibly false in absolutely unramified and non-logarithmic case* ([1, Theorem 3.8, 3.16], [2, Theorem 1]).

(5) *The inertia subgroup is $\mathrm{Fil}^a G_k$ if $a \in (0, 1]$ and the wild inertia subgroup is $\mathrm{Fil}^{1+} G_k = \mathrm{Fil}^{0+}_{\log} G_k$* [1, Theorem 3.7, 3.15].

(6) *When the residue field $\bar{k}$ is perfect, arithmetic ramification filtrations agree with the classical upper numbered filtration in the following way: $\mathrm{Fil}^a G_k = \mathrm{Fil}^{a-1}_{\log} G_k = G^a_k$* [1, Section 6.1].

Proof. I will sketch some of the proofs.

(1) It is straightforward to check the independence on the generators of $\mathcal{I}$. To see that does not depend on generators of $\mathcal{O}_l/\mathcal{O}_k$, one can show that if we add a new (dummy)

generator, the new Abbes-Saito space admits a fibration over the original Abbes-Saito space whose fibers are closed discs of radius θ^a. The subtle difference between the polynomial rings $\mathcal{O}_k[u_0, \dots, u_m]$ and its completion in the definition does not affect the ramification breaks.

(2) The first statement is just abstract nonsense. The second one is essentially because one defines the Abbes-Saito space over k and the geometric connect components and be seen over the algebraic closure k^{alg}, which has valued group $|k^\times|^{\mathbb{Q}}$. However, a rigorous proof needs the theory of stable formal models of rigid spaces.

(3) For the unramified case, it is essentially because, for any finite Galois extension k'/k linearly independent from l/k, $\mathcal{O}_{lk'} \simeq \mathcal{O}_l \otimes_{\mathcal{O}_k} \mathcal{O}_{k'}$. Thus, one can match up two Abbes-Saito spaces in a natural way. In the tamely ramified and logarithmic case, one can also naturally identify two Abbes-Saito spaces [1, Proposition 9.8].

(4) The proof used formal models and their reductions of Abbes-Saito spaces, which is the main theorem in [2]. Saito proved a stronger version of this in [12, Theorem 1.3.3] stating that for equal characteristic field k, the graded piece of logarithmic filtrations are actually killed by p.

(5) is an easy fact.

(6) follows from the explicit calculation in Proposition 1.8 and 1.10. □

One can define Artin conductors and Swan conductors as in the classical case, using the same Formulas (1.6.1), (1.6.2). Essentially, for an irreducible representation ρ which exactly factors through a Galois extension l/k, $\mathrm{Art}(\rho) = b(l/k) \cdot \dim \rho$ and $\mathrm{Swan}(\rho) = b_{\log}(l/k) \cdot \dim \rho$.

Who cares about imperfect residue field, anyway?

The following theorem states that the Swan conductors measure the discrepancy in the Euler characteristic formula.

Theorem 2.6 (Grothendieck-Ogg-Shavarevich Formula)**.** *Let k_0 be a perfect field of characteristic p and let l be a prime number different from p. Let X be a smooth proper curve over k_0 and let U be a dense open subset. Assume for simplicity, $X \backslash U = \{x_1, \dots, x_n\}$ are all rational points over k_0. Given a lisse $\mathbb{F}_l$-sheaf $\mathcal{F}$ on U, one can read off the Swan conductor* $\mathrm{Swan}_{x_j}(\mathcal{F})$ *at each of the missing point x_j. Then,*

$$\chi_c(U, \mathcal{F}) = \sum_{i=0}^{2} (-1)^i \dim_{\mathbb{F}_l} H^i_c(U, \mathcal{F}) = \chi_c(U, \mathbb{F}_l) \cdot \mathrm{rank}\,_{\mathbb{F}_l}\mathcal{F} - \sum_{j=0}^{n} \mathrm{Swan}_{x_j}\mathcal{F}. \tag{2.6.1}$$

There is an analogous result for overconvergent F-isocrystals.

Very vaguely speaking, one can view this as an analogue of Riemann-Roch theorem. So, it is natural to ask for a higher dimensional Theorem 2.6.

Let X be a smooth proper variety over k_0 and D a divisor with simple normal crossings. Let $\mathcal{F}$ be a lisse $\mathbb{F}_l$-sheaf on $X \backslash D$. The first difficulty is defining the Swan conductor $\mathrm{Swan}_{D_i}(\mathcal{F})$ along an irreducible components D_i of D. An obvious definition is to pass to the completion $R = \mathcal{O}^\wedge_{X,\eta_i}$ at the generic point η_i of D_i. One immediately finds out that R is a complete discrete valuation ring with residue field equal to the function field of D_i, which is typically imperfect if $\dim D_i \geq 1$.

In some sense, the above definition gives a possible way to overcome this difficulty. Indeed, Saito, in [12], proved a higher dimensional Grothendieck-Ogg-Shavarevich formula for lisse $\bar{\mathbb{Q}}_l$-sheaves under some technical conditions. We will not discuss it here. It is worth to point out that we do not expect any analogue of (2.6.1) holds for general lisse $\bar{\mathbb{Q}}_l$-sheaves in higher dimensional case. Instead, it should only hold for "clean" sheaves whose ramifications along missing divisors are controlled by the generic points of the divisors. For various definitions of cleanness, one may consult [8, 9, 12].

3 Hasse-Arf theorem

We will reproduce the main results in [14, 15]. The goal of the two papers is to prove the following.

Conjecture 3.1 (Hasse-Arf Theorem)**.** *Let k be a complete discretely valued field. Then the conductors* $\mathrm{Art}(\rho)$ *and* $\mathrm{Swan}(\rho)$ *are non-negative integers. Moreover, the subquotients* $\mathrm{Fil}^a G_k/\mathrm{Fil}^{a+} G_k$ *and* $\mathrm{Fil}^a_{\log} G_k/\mathrm{Fil}^{a+}_{\log} G_k$ *are trivial for irrational a and are abelian groups killed by p for rational $a > 1$ and 0, respectively.*

This conjecture is first raised in [2], in which Abbes and Saito proved that the subquotients of the filtrations are abelian groups expect possibly in the mixed characteristic absolutely unramified and non-logarithmic case. Later Saito [12] proved that in the equal characteristic case, $\mathrm{Fil}^a_{\log} G_k/\mathrm{Fil}^{a+}_{\log} G_k$ is an abelian group killed by p if $a \in \mathbb{Q}_{>0}$ (and is trivial if a is irrational).

Our approach is very different. It originated as follow. In the case k is of equal characteristic $p > 0$, there is another definition of Artin/Swan conductor using p-adic differential modules, which was first introduced by Christol, Matsuda, Mebkhout, and Tsuzuki in the perfect residue case and was carried out in general case by Kedlaya [10] later.

Kedlaya proved Hasse-Arf theorem for the differential conductors. Matsuda then asked if one can prove a comparison theorem between the Abbes-Saito's definition and Kedlaya's definition. Therefore, one can obtain a Hasse-Arf theorem for the ramification filtrations defined by Abbes and Saito. The first step was carried out by Bruno Chiarellotto and Andrea Pulita [4] in rank 1 case, using a different method to compare to a definition by Kato. Later, I worked out the general case [14] using the p-adic differential modules.

In the mixed characteristic case, very little was known except the result of Abbes and Saito mentioned above.

The main results in [14, 15] are the following.

Theorem 3.2. *The Conjecture 3.1 holds* except *in the following cases.*

(1) *When k is of mixed characteristic and is absolutely unramified (i.e. p is a uniformizer of k), we do not know if it is true for non-logarithmic filtration;*

(2) *When $p = 2$, instead of* $\mathrm{Swan}(\rho) \in \mathbb{Z}$, *we can only prove* $\mathrm{Swan}(\rho) \in \frac{1}{2}\mathbb{Z}$.

Remark 3.3. The first restriction also occurs in [2]. It reflects the failure of deforming the uniformizer p (not even "slightly"). Explicitly, we have a dichotomy

$$\Omega^1_{\mathcal{O}_k/\mathbb{Z}_p} \otimes \bar{k} = \begin{cases} \bigoplus_{j=1}^m \bar{k} \cdot db_j & k \text{ is absolutely unramified,} \\ \bar{k} \cdot d\pi_k \oplus \bigoplus_{j=1}^m \bar{k} \cdot db_j & \text{otherwise.} \end{cases}$$

The second restriction is purely technical. At some point, we can show that there exists $\alpha \in \mathbb{Q}$ such that $n \cdot \mathrm{Swan}(\rho) \in \alpha + \mathbb{Z}$ for all and $n \gg 0$ coprime to p. When $p = 2$ we cannot completely recover the integrality of $\mathrm{Swan}(\rho)$. (For the proof, see the last paragraph of the proof of Theorem 3.2.) In some cases, I think we can force $\alpha \in \mathbb{Z}$ using a slight variant of the method but I am not sure how to show it in general (see [15, Remark 3.5.12]), nor did I know any counterexample for the integrality of Swan conductor.

Now we sketch a proof of the theorem. Since introducing Kedlaya's definition requires carrying out the whole theory of p-adic differential modules, I will just vaguely talk about the proof, hiding Kedlaya's definition in it.

The here given proof is not optimized in the equal characteristic case, in which case one can get the integrality from slopes of Newton polygons very easily. We would like to make the proof as a process of reduction to the perfect residue field case. The equal characteristic case and the mixed characteristic case share the same strategy, but with different technical difficulties.

Proof. Let k be a complete discretely valued field. We may always reduce to the case when the residue field $\bar{k}$ has a finite p-basis, or equivalently, $\dim_{\bar{k}^p}(\bar{k}) < \infty$. We will only study the essential part when l/k is a finite Galois totally ramified extension whose wild ramification part is non-trivial. As in Lemma 2.3, the extension of ring of integers $\mathcal{O}_l/\mathcal{O}_k$ is generated by $c_1, \dots, c_m, \pi_l$, or vaguely speaking, some "good" elements which allow us to write down explicit equations by which they generate $\mathcal{O}_l$.

We use $A^m_k[0, \theta^a]$ to denote an m-dimensional polydisc over k with radius θ^a. Similarly, for a nonarchimedean field K and $\eta_0 \in (0, 1)$, we use $A^1_K[\eta_0, 1)$ to denote a half-open annulus with inner radius η_0 and outer radius 1.

We first give the outline of the proof, with serious gaps in it. Then we talk about how to remedy the problems.

Step I: $AS = TS$ **theorem.** (Make the Abbes-Saito space more functorial.)

There is a natural k-morphism $\pi' : as^a_{l/k} \to A^{m+1}_k[0, \theta^a]$ sending $(u_0, \dots, u_m)$ to $(p_0(\underline{u}), \dots, p_m(\underline{u}))$ (see Definition 2.4). The Abbes-Saito space is easy to define, but it is not easy to work with. In particular, it is not preserved under base change, in other words, for k'/k a finite extension, it is hard to link $as^a_{l/k} \times_{\pi', A^{m+1}_k[0,\theta^a]} A^{m+1}_{k'}[0, \theta^a]$ to $as^a_{l'/k'}$, where $l' = k'l$. In order to solve this problem, we need introduce the thickening space $ts^a_{l/k}$, together with $\pi : ts^a_{l/k} \to A^{m+1}_k[0, \theta^a]$.

Pretend for a moment that we have a continuous homomorphism $\psi : \mathcal{O}_k \to \mathcal{O}_k[\![\delta_0, \dots, \delta_m]\!]$ such that $\psi(\pi_k) = \pi_k + \delta_0$ and $\psi(b_j) = b_j + \delta_j$ for $j \in J$. (This homomorphism always exists in the equal characteristic case but *never* exists in the mixed characteristic case. I will come back to this point in the remedy part.) We define $ts^a_{l/k} = A^{m+1}_k[0, \theta^a] \times_{k,\psi} l$. Using standard approximation argument, we can prove that $ts^a_{l/k} \simeq as^a_{l/k}$. An alternative way to understand this construction is that we change the morphism $\pi' : as^a_{l/k} \to A_k^{m+1}[0, \theta^a]$ to make it "more functorial" on k.

Example 3.4. It is good to see the difference of two definitions in an example. Consider the extension of $\mathbb{F}_p((x))$ given by $y^p - x^{p-1}y = x$. The Abbes-Saito space is given by

$$\{(u, \delta) || u| \leq 1, |\delta| < \theta^a, u^p - x^{p-1}u = x + \delta\},$$

whereas the thickening space is given by

$$\{(u, \delta) || u| \leq 1, |\delta| < \theta^a, u^p - (x + \delta)^{p-1}u = x + \delta\}.$$

In other words, the Abbes-Saito space $as^a_{l/k}$ consists of the points which are close to the solutions to those equations; in contrast, the thickening space $ts^a_{l/k}$ consists of points which are solutions to some equations whose coefficients are close to the original equations.

Step II: generic p-th roots. (A procedure to reduce to the perfect residue case.)

It is natural to make the following observation. Let a be slightly bigger than $b(l/k)$ then $ts^a_{l/k}$ is geometrically the disjoint union of $[l : k]$ discs. What often happens is that if you only increase the radius in certain direction, $\pi_0^{\mathrm{geom}}(ts_{l/k})$ stays the same even when the radius goes across the cut-off point $\theta^{b(l/k)}$. In contrast, if you increase radius along some other direction, $\pi_0^{\mathrm{geom}}(ts_{l/k})$ will change as soon as the radius reaches $\theta^{b(l/k)}$. In the latter case, we say that direction dominates. We remark that if we change the lift of $\bar{b}_j$ from b_j to $b_j + \pi_k$, then whether the "uniformizer direction" is dominant may be changed as well.

A natural question arises along this line: when is the direction corresponding to δ_0 (uniformizer) dominant? If so, can we "forget" about other directions, in other words, can we make the residue field perfect by simply adding in p-th roots of b_j for all j? Moreover, can we make the "uniformizer direction" always dominant, say by changing a dominant b_j to $b_j + \pi_k$ if it is not dominant?

Following these questions and inspired by the work of Borger in [3], Kedlaya introduced the notation of generic rotation. Let $x_1, \dots, x_m$ be transcendental over k, let k' be the completion of $k(x_1, \dots, x_m)$ with respect to the $(1, \dots, 1)$-Gauss norm and let $l' = k'l$. It easy to see that $b(l'/k') = b(l/k)$. The upshot is that if we set the p-basis of k' to be $\{b_1 + x_1\pi_k, \dots, b_m + x_m\pi_k, x_1, \dots, x_m\}$, then the uniformizer direction is always dominant. Therefore, if the answers to the above questions are positive, we can just go ahead to add in all the p-th power roots of the prescribed p-basis and reduce to the perfect residue field case.

To realize the strategy above, we have to find a tool to detect the dominance and to study how the ramification numbers vary. This is where p-adic differential modules come into the picture.

Step III: étaleness. (Where we get the differential module from.)

Vaguely speaking, we hope that $\pi : ts^a_{l/k} \to A_k^{m+1}[0, \theta^a]$ is étale, so that we can push-forward the ring of functions on $ts^a_{l/k}$ to $A_k^{m+1}[0, \theta^a]$ to obtain a differential module $\mathcal{E}$, i.e., a locally free module with integrable connection $\nabla : \mathcal{E} \to \mathcal{E} \otimes \Omega^1_{A_k^{m+1}[0,\theta^a]}$. (In the equal characteristic p case, there is no good theory of differential modules available in the sense of Dwork. Just bear me for a moment; remedy will be explained later.)

The étaleness of π is not true in general. Of course, it is étale (in fact disjoint union of copies) when $a > b(l/k)$. Fortunately, Abbes and Saito proved in [1, 2] that π is étale for $a > b(l/k) - \epsilon$ for some $\epsilon > 0$. This extra ϵ is crucial for our approach, using which we can read off the precise $b(l/k)$ by the following step.

Step IV: relation between radii of convergence of differential modules with ramification breaks. (How to use differential modules to "bind" the space together.)

The sheaf $\mathcal{E} = \pi_* \mathcal{O}_{ts^a_{l/k}}$ is a differential module when $a > b(l/k) - \epsilon$ (not quite in the equal characteristic case, see the Remedy part). We consider its "naïve" base change of $\mathcal{E}$ to $A_l^{m+1}[0, \theta^a]$. When restricting to the fiber at $\delta_0 = \cdots = \delta_m = 0$, $\mathcal{E}$ splits as $l \otimes_k l = \prod_{g \in \mathrm{Gal}(l/k)} l_{(g)}$, where $l_{(g)}$ are just copies of l, indexed by $g \in G$. The Taylor series lift the idempotent elements to the locus where they converge. Thus, the radius θ^a where $ts^a_{l/k}$, or equivalently $as^a_{l/k}$, becomes geometrically disjoint of copies of $A_k^{m+1}[0, \theta^a]$ is the same as the radius of convergence of the differential module $\mathcal{E}$. We hence turn the question of rigid geometry to a question of p-adic differential modules.

A good thing of radius of convergence is that one can calculate it easily under base change. In particular, if we want to forget about some element in the p-basis, we just ignore the radius for the corresponding derivative. Using this, we can finish the project laid out in Step II and hence prove (modulo the lies I said) the Hasse-Arf theorem.

Step V: Logarithmic filtration. (A trick to deal with logarithmic filtration.)

In the logarithmic case, we do not expect that we can always make the uniformizer direction dominant. Instead, we expect a dichotomy:

- if the uniformizer direction is dominant, we are good anyway;
- if the uniformizer direction is not dominant, we expect that, after a large tame base change to $k_n = k(\pi_k^{1/n})$ and then a generic rotation for k_n as in Step II, $b_{\log}(l'_n/k'_n) = n b_{\log}(l/k) - 1$ and the uniformizer direction is dominant. (In the mixed characteristic case, we have a technical issue here, see the remedy later.)

Thus, we can always deduce that $n \cdot \mathrm{Swan}(\rho) \in \mathbb{Z}$ for $n \gg 0$ and $p \nmid n$. Taking two coprime numbers n_1 and n_2 will imply that $\mathrm{Swan}(\rho)$ itself is an integer.

To get the result on the subquotients, one uses a tricky wildly ramified base change due to Kedlaya, which we will not discuss here.

Remedy in the equal characteristic p case: We lift the thickening space $ts^a_{l/k}$ over k to a rigid space $TS^a_{l/k}$ over an annulus over a field of characteristic zero.

The only compliant in this case is that we do not have a good theory of differential modules over a characteristic p field. The trick is to lift the rigid space $ts^a_{l/k}$ to a rigid space $TS^a_{l/k}$ over $A_K^1[\eta_0, 1)$ for some $\eta_0 \in (0, 1)$, where K is the fraction field of a Cohen ring $\mathcal{O}_K$ of $\bar{k}$ and $A_K^1[\eta_0, 1) = \{x | \eta_0 \leq |x| < 1\}$. Roughly speaking, the lifting process is just writing down equations that define $ts^a_{l/k}$ as an affinoid subspace of $A^{m+1}_k[0,1]$ and lifting the coefficients to $\mathcal{O}_K[\![T]\!]$; then these equations will give the space $TS^a_{l/k}$ over the annulus $A^1_K[\eta_0, 1)$. Vaguely speaking, one can show that geometric connected components of $ts^a_{l/k}$ are in one-to-one correspondence with "geometric" connected components of $TS^a_{l/k}$ when $\eta_0 \to 1$. Here, "geometric" means up to a base extension from $A^1_K[\eta_0, 1)$ to $A^1_{K'}[\eta_0^{1/e_{k'/k}}, 1)$, where k' is a finite separable extension of k of naive ramification degree $e_{k'/k}$ and K' is the fraction field of a Cohen ring of the residue field of k'.

In fact, the Steps III-V of the proof are carried out for $TS^a_{l/k}$ instead of the thickening space itself.

Remedy in the mixed characteristic case: We do *not* require $\psi : \mathcal{O}_k \to \mathcal{O}_k[\![\delta_0, \dots, \delta_m]\!]$ to be a homomorphism, but just a function.

Basically, the main obstruction of constructing a homomorphism ψ is that $\psi(p)$ is forced to be p but not something like $(\pi_k + \delta_0)^{\beta_k}$ if $\pi_k^{\beta_k} = p$. Thus, we can pathetically claim that ψ is a homomorphism modulo p. When $\beta_k = v_k(p)$ is big, this turns out to be a very strong approximation, which allows us to carry out all the steps with some modification. In contrast, when $\beta_K = 1$, we can not even distinguish $\psi(p) = p$ with $\psi(p) = p + \delta_0$. This approximation is too weak to obtain the main theorem.

In Step I, we define the standard thickening space $ts^a_{l/k,\psi}$ as the rigid space associated to

$$k\langle u_0, \dots, u_m, \pi_k^{-a}\delta_0, \dots, \pi_k^{-a}\delta_m\rangle / \big(\psi(p_0(\underline{u})), \dots, \psi(p_m(\underline{u}))\big),$$

where we apply ψ termwise to $p_j(\underline{u})$. It turns out that the $AS = TS$ theorem holds, even with the existence of some reasonable error terms, namely oscillating $\psi(p_j(\underline{u}))$ by some "small" element $\mathfrak{R}_j \in \mathcal{O}_k[\![u_0, \dots, u_m, \delta_0, \dots, \delta_m]\!]$. (The $\psi(p_j(\underline{u}))$ themselves are not very well-defined anyway.)

Step II is more delicate. Since ψ is not actually a homomorphism, under base change, $ts^a_{l/k}$ will give some error terms, which is originally of the scale of $|p|$. Thus, we need to make some mild base change so that the error terms are still in the range to invoke the $AS = TS$ theorem. In particular, we cannot put in all the p-th power roots of some b_j at the same time. Instead, we base change to $k' = k(x)^{\wedge,\mathrm{unr},\wedge}(b_j + x\pi_k)^{1/p}$, where the hat is the completion with respect to the 1-Gauss norm. This operation is only valid when $\beta_k = v_k(p) \geq 2$, otherwise the error terms from ψ will exceed the restriction posed in the $AS = TS$ theorem (compare Remark 3.3). It is an easy exercise that, after finitely many such base change, we can reduce to the non-fiercely ramified case (i.e., the residue field extension is separable), which has no difference from the perfect residue field case.

It is worth to mention that when dealing with the logarithmic filtrations, we can first make a large tame base change to avoid this absolutely unramified ($\beta_k = 1$) issue.

Step III and Step IV do not need any further changes. In Step V, we omitted a key point when first introducing it. The dichotomy only appears if we have a continuous homomorphism $\phi : \mathcal{O}_k \to \mathcal{O}_k[\![\delta_1, \dots, \delta_m]\!]$ sending b_j to $b_j + \delta_j$, in which case, the differential module can be obtained from a one dimensional space "$A^1_k[0, \theta^a] \times_{k,\psi'} l \to A^1_k[0, \theta^a]$" where the *function* $\psi' : \mathcal{O}_k \to \mathcal{O}_k[\![\delta_0]\!]$ together with the above homomorphism gives ψ. If this is the case, we are essentially studying one dimensional variation of radii of convergence, which gives rise to the result in Step V.

When such a homomorphism ϕ does not exist, we are forced to work with higher dimensional variation of radii of convergence using [11]. Consequently, we have to modify the second case of the dichotomy to be $b_{\log}(l'_n/k'_n) = nb_{\log}(l/k) - \alpha_{l/k}$ for some $\alpha_{l/k} \in \mathbb{R}$ *not depending* on n. Here is my favorite argument $\heartsuit$:

By the non-logarithmic Hasse-Arf theorem, $n\mathrm{Swan}(\rho) \in \mathbb{Z} + \beta$ for $\beta = \alpha_{l/k} \dim \rho$. Take $n_1, n_2 \gg 0$ coprime to p as closer to each other as possible, in other words, take $n_1 = n_2 + 1$ when $p \neq 2$ or take $n_1 = n_2 + 2$ when $p \neq 2$. Thus, we obtain $\alpha \in \mathbb{Z}$ when $p = 2$ and $\alpha \in \frac{1}{2}\mathbb{Z}$ when $p = 2$. Then, pick two coprime $n_3, n_4 \gg 0$ both coprime to p. When $p \neq 2$, $n_3\mathrm{Swan}(\rho), n_4\mathrm{Swan}(\rho) \in \mathbb{Z}$ will imply that $\mathrm{Swan}(\rho) \in \mathbb{Z}$. When $p = 2$, same argument can only prove that $\mathrm{Swan}(\rho) \in \frac{1}{2}\mathbb{Z}$. Evil even prime number $p = 2$ only has two congruence classes....

(This argument is slightly different from the one I used in [15, Section 3.5] because in the paper I was trying to avoid the notational complication when n is not congruent to 1 modulo $p[l : k]$.) $\square$

Remarks on the relations with Abbes and Saito's approach

The essential difference between our approaches and their methods reflects the two different points of view of rigid analytic geometry: one is working with affinoid algebras and using p-adic analysis, and the other one is using formal models of affinoid algebras and studying their (stable) reductions.

In the study of rigid geometry, it often happens that these two approaches can reach the same goal using completely irrelevant arguments. However, the deep relation between these two methods is reflected in many (unexpected) places. The failure of proving Conjecture 3.1 in the mixed characteristic and absolutely unramified case sets a perfectly good example. On one hand, Abbes and Saito failed because of the reason in Remark 3.3; on the other hand, I failed because I can not deform p as in "Remedy in the mixed characteristic case". They seem to come from the same origin but appear in a complete different form.

The two approaches have their own advantages. The Abbes and Saito's approach grasps tightly the integral structure, which gives them easy access to a global theory in the vein of Grothendieck-Ogg-Shavarevich formula (see [12]). In contrast, the approach through p-adic analysis is less rigid and can "deform" better. One can obtain variational properties of Swan conductors easily (see [11] and [9]).

References

[1] Ahmed Abbes and Takeshi Saito. Ramification of local fields with imperfect residue fields. *Amer. J. Math.*, 124(5):879–920, 2002.

[2] Ahmed Abbes and Takeshi Saito. Ramification of local fields with imperfect residue fields, II. *Doc. Math.*, (Extra Vol.):5–72 (electronic), 2003. Kazuya Kato's fiftieth birthday.

[3] James M. Borger. Conductors and the moduli of residual perfection. *Math. Ann.*, 329(1):1–30, 2004.

[4] Bruno Chiarellotto and Andrea Pulita. Arithmetic and Differential Swan Conductors of rank one representations with finite local monodromy. arXiv: `math.NT/0711.0701v1`.

[5] Pierre Colmez. Conducteur d'Artin d'une représentation de de Rham. preprint.

[6] I. B. Fesenko and S. V. Vostokov. *Local fields and their extensions*, volume 121 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, second edition, 2002.

[7] Kazuya Kato. Swan conductors for characters of degree one in the imperfect residue field case. In *Algebraic K-theory and algebraic number theory (Honolulu, HI, 1987)*, volume 83 of *Contemp. Math.*, pages 101–131. Amer. Math. Soc., Providence, RI, 1989.

[8] Kazuya Kato. Class field theory, $\mathcal{D}$-modules, and ramification on higher-dimensional schemes. I. *Amer. J. Math.*, 116(4):757–784, 1994.

[9] Kiran S. Kedlaya. Swan conductors for p-adic differential modules, II: Global Variation. arXiv: `math.NT/0705.0031v2`.

[10] Kiran S. Kedlaya. Swan conductors for p-adic differential modules. I. A local construction. *Algebra Number Theory*, 1(3):269–300, 2007.

[11] Kiran S. Kedlaya and Liang Xiao. Differential modules on p-adic polyannuli. arXiv: `0804.1495v2`.

[12] Takeshi Saito. Wild ramification and the characteristic cycle of an l-adic sheaf. arXiv: `arXiv:0705.2799v2`.

[13] Jean-Pierre Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979.

[14] Liang Xiao. On ramification filtrations and p-adic differential equations, I: equal characteristic case. arXiv: `0801.4962v2`.

[15] Liang Xiao. On ramification filtrations and p-adic differential equations, II: mixed characteristic case. available at `http://www-math.mit.edu/~lxiao/`.