Resolving Surface Singularities in Positive Characteristic

Herwig HAUSER, Stefan PERLEGA

Abstract

The paper contains a systematic proof for the embedded resolution of two-dimensional hypersurface singularities over a field of arbitrary characteristic. It relies on the construction of a local upper semicontinuous invariant whose top locus defines at each stage of the resolution process the center of a permissible blowup under which the invariant improves. As the invariant belongs to a well-ordered set, resolution is achieved in finitely many steps. To define the invariant we analyze in detail the obstructions which occur when transcribing the inductive characteristic 0 proof à la Hironaka. These difficulties are then overcome by introducing finer invariants than in zero characteristic: they reflect and control typical characteristic p phenomena. This may give new ideas for approaching the characteristic p resolution in higher dimensions.

1 Introduction

oo In resolution of singularities, one distinguishes between non-embedded and embedded resolution. The first aims at finding, for a given singular variety X, a regular variety $\widetilde{X}$ together with a proper birational morphism $\pi : \widetilde{X} \to X$. The second starts with a singular subvariety X embedded in some regular ambient variety W and searches for a proper birational morphism $\pi : \widetilde{W} \to W$ leaving $\widetilde{W}$ regular so that the total transform $\pi^{-1}(X)$ of X is a normal crossings divisor in $\widetilde{W}$. This is a substantially stronger assertion. Additional properties and variants of the statements can be formulated in both settings.

While in characteristic zero the embedded resolution is ensured in any dimension by Hironaka's theorem [Hir64], the situation is much more delicate in positive characteristic. For surfaces, the existence of both non-embedded

MSC 2020: 14B05, 14J17, 14J70, 32S05, 32S10, 32S25.

. Supported by the Austrian Science Foundation FWF, projects P-25652 and P-31338.

and embedded resolution is well known [Abh56, Abh66, Hir84, Lip78]. More recent approaches with often stronger results were provided by the work of [CJS09, BV10, BV12, KM16]. The case of embedded resolution nevertheless remains "work in progress" since the existing proofs are either very complicated (Abhyankar) or not sufficiently conceptual (Hironaka) so as to extend directly to higher dimension (see also [**?**]). For three-folds, non-embedded resolution has been achieved by Cossart-Piltant [CP08, CP09, **?**], and embedded resolution is still an open problem.

We will present in this paper a new and systematic proof for the embedded resolution of surface singularities inside three-dimensional ambient varieties in arbitrary characteristic. It is commonplace that the case of surfaces has still to be understood better in order to be able to pass on to three-folds and to prove embedded resolution there. We expect that our techniques present a good starting point and give new ideas for attacking this case.

Let us briefly review the situation for surfaces: The first results for surfaces in positive characteristic are due to Abhyankar [Abh56]. He establishes local uniformization (i.e., local non-embedded resolution along a valuation) making extensive use of Galois theory for local rings and of field extensions. This is extended to embedded resolution of surfaces over perfect fields and non-embedded resolution of threefolds over algebraically closed fields of characteristic $\neq 2, 3, 5$ in 1966 [Abh66], as well as to the case of arithmetic surfaces [Abh65]. For a succinct presentation of the long proof (more than 500 pages), see [Cut11] as well as [Cut09].

Hironaka proposes in [Hir84] a resolution invariant drawn from the Newton polyhedron of the singularity in order to prove local resolution for surfaces. The argument has been completed by Cossart [Cos81], see also [Hau00] and [**?**]. The invariant is rather ad-hoc. It requires to choose maximal centers inside the equimultiple locus, otherwise it may increase.

Lipman establishes non-embedded resolution of arbitrary excellent surfaces, using cohomology and duality to first reduce the singularities to rational ones [Lip78, Lip69, Art86].

Cossart, Jannsen, and Saito treat surfaces of higher embedding dimension (actually, arbitrary excellent two-dimensional schemes) building on Hironaka's invariant for two-dimensional hypersurfaces [CJS09]. Benito and Villamayor, respectively Bravo and Villamayor replace the restriction to hypersurfaces by generic projections and achieve embedded resolution for surfaces by a careful analysis of the exceptional components in their so-called "monomial case" [BV10, BV12]. Kawanoue and Matsuki rely on their concept of idealistic filtrations and prove the embedded resolution of surfaces by somewhat similar arguments as the authors mentioned before [KM16].

All these proofs proceed by a (local) descent in dimension, thus formulating a resolution problem in one variable less (say, in two variables). Two different strategies can be observed.

1. Resolve the smaller dimensional problem by a sequence of blowups in regular centers, by induction on the dimension. Apply the same blowups to the original singularity X. Then try to draw, from the resolution in lower dimension, conclusions on the transform of X itself under the induced blowups. This works perfectly well in characteristic zero by the descent via hypersurfaces of maximal contact and the introduction of coefficient ideals. In positive characteristic, substitutes for hypersurfaces of maximal contact and for the descent have been proposed in [Vil07, BV10, BV12, Kaw07, KM10, KM16]. The problem here is that the resolution in lower dimension does not directly imply an improvement of the original variety X.
2. The second strategy consists in applying the descent in dimension after *each* blowup (in characteristic zero, the descent commutes with blowup so that both strategies coincide). The descent requires to choose at each stage of the resolution process local hypersurfaces and to measure the improvement via an invariant defined by a string of ideals in decreasing dimensions. Due to the failure of maximal contact, the strict transform of the previously chosen hypersurface may not be usable anymore after blowup and a new hypersurface has to be chosen each time. Hence, the definition of a prospective invariant and its control become much more delicate in positive characteristic.

The second approach is the one we pursue in the present paper: We will present – for surfaces X embedded in a regular three-dimensional ambient variety W defined over an algebraically closed field and equipped with a boundary divisor E – a local invariant $\text{inv}_a X$ which mimics the characteristic zero resolution invariant, and which shares important properties with it:

1. it belongs to a well ordered set;
2. it is defined at all points of the strict transforms of X;
3. it is upper semicontinuous;
4. the locus of points where $\text{inv}_a X$ is maximal defines a permissible center Z in W (i.e., Z is closed, regular, and transversal to E);
5. the maximum value of $\text{inv}_a X$ on X decreases under the blowup $W' \to W$ of W along Z until embedded resolution of singularities is achieved.

The following theorem has been proved in its entirety in [Per17]. In this article, we will outline all of the main ideas and techniques which appear in its proof.

Theorem. *Let be given a singular hypersurface X in a three-dimensional regular ambient variety W, equipped with a normal crossings divisor E. There exists, for each point a of X, a local upper semicontinuous invariant $\mathrm{inv}_a X$ depending on E,*

$$\mathrm{inv}_a X = (o, c, d, n, s, k, \ell) \in (\mathbb{N}^7, <_{\mathrm{lex}}),$$

such that:

(1) The top locus $Z = \{a \in X, \mathrm{inv}_a X \text{ is maximal}\}$ is a closed regular subvariety of W, transversal to E, and contained in the singular locus $\mathrm{Sing}\, X$ of X.

(2) Assume that X is not resolved. Let $\pi : W' \to W$ be the blowup of W along Z, denote by X' the strict transform of X in W', and let a' be a point in W' above $a \in Z$. Then, for $E' = \pi^{-1}(E \cup Z)$, strict inequality holds,

$$\mathrm{inv}_{a'} X' <_{\mathrm{lex}} \mathrm{inv}_a X.$$

As the invariant takes only finitely many values on X and can only drop finitely many times under blowup, this proves embedded resolution of surfaces in a three-dimensional ambient space.

The entries o, d, s are orders of successive ideals in decreasing dimensions (the letters o, d, s stand for *order*, *residual order*, and *slope*), the difference $c - o$ counts the "old" exceptional components, n measures to what extent certain curves are not yet transversal to E, and k and ℓ are combinatorial entries only needed for two easy "terminal" cases.

The difference to the (classical) characteristic zero invariant occurs in the entries d and n, after the first descent in dimension. Due to the lack of hypersurfaces of maximal contact we will choose regular hypersurfaces *maximizing* the order of the coefficient ideal of the defining polynomial f of X at a (it is well known that, in characteristic zero, hypersurfaces of maximal contact maximize this order). The resulting coefficient ideal is factored into a monomial exceptional part and a residual part. The order of the residual part is a canonical candidate for the next entry of the invariant. As it does not behave well in positive characteristic, we replace it by a refinement, called d, of which we show that it does not increase under blowup (whenever the first two components (o, c) have remained constant). This is the clue to make the whole procedure work. The definition of d takes into account also orders of ideals along *curves*, not just at points. Before the next descent in dimension, a second transversality measure, denoted by n, is considered. It quantifies the tangency with respect to the exceptional components of the curve along which the order was taken. After this intermediate step one passes to the following coefficient ideal and takes for s its order.

A key ingredient to define the string (d, n, s) is the introduction of local flags $\mathcal{F}$. One associates to each of them a triple of integers $(d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}}) \in \mathbb{N}^3$.

The triple (d, n, s) is then defined as the maximum of $(d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$, taken over all flags which maximize an auxiliary invariant $m^{\mathcal{F}}$. As such it does not depend on any choices. It turns out that the strict transform of a maximizing flag does not have to be maximizing again after blowup. Hence, a new maximizing flag has to be chosen to realize the invariant. This makes it quite subtle to compare the invariant before and after blowup, and represents the heart of the proof (see Prop. 4).

To improve readability, we will restrict our considerations mostly to singular hypersurfaces which are, at least locally, defined by a *purely inseparable equation* of the form

$$z^{p^e} + F(x, y) = 0,$$

where e is a positive integer. These hypersurfaces are exemplary in the sense that they already exhibit the relevant pathologies which prohibit generalizing the proof for resolution of singularities over fields of characteristic zero to the setting of positive characteristic. Restricting to the case of purely inseparable equations has the advantage that it allows us to outline the main arguments of the proof [Per17] in concise form without the need of introducing all of the technical machinery which is necessary to treat general singular surfaces in arbitrary characteristic. The main arguments for the general case are then treated separately in the last two sections.

The structure of the article is as follows: After fixing the setting, we introduce in Section 3 two *terminal cases* which can be resolved easily and have to be treated apart. This will allow us to restrict our attention in the sequel to point blowups.

In Section 4 we will introduce the *residual order*, an invariant which is the natural transcription of a resolution invariant from characteristic zero to the setting of positive characteristic. The residual order in positive characteristic has several shortcomings: It is not upper semicontinuous and it may increase under blowup. We will show how both of these problems can be overcome by defining an appropriate modification. This modification makes use of the fact that the ambient space W is three-dimensional. So far, it was not possible to extend its definition to higher dimensional varieties.

We will then investigate the behaviour of the residual order under point blowups and analyze in detail the problematic cases in which it increases. We then also indicate in which way the residual order has to be refined so as to prevent this increase from happening. The definition of the resolution invariant which arises from this modification of the residual order will be given in 5. This invariant strictly decreases under point blowup (even in the cases in which the residual order increases) until a terminal case is reached.

The proofs for all the results in this article are gathered in Section 6. In Section 7 we will discuss which new problems appear in higher dimensions. In particular, we will see that the residual order shows a much worse behaviour under sequences of blowups. This makes it impossible to use the refined residual order to higher dimensions in a straightforward way. In Section 8 we will indicate how the resolution invariant of [Per17] is defined in the general case, when X is not necessarily given by a purely inseparable equation. The final Section 9 is devoted to showing that the respective invariant decreases also in this case, omitting some technical but plausible results. This will complete the proof of the theorem.

Acknowledgements. The work on the present article has profited from various sources and many discussions with other people. We are particularly grateful for conversations with Dale Cutkosky, Santiago Encinas, Heisuke Hironaka, Hiraku Kawanoue, Kenji Matsuki, Bernd Schober, and Orlando Villamayor.

2 Preliminaries

The most basic invariant for the resolution of hypersurfaces X in a regular ambient variety W is the *local order* of X at a point a. Generally, for an ideal J of a ring R and a prime ideal P, the order of J at P is defined as

$$\mathrm{ord}_P\, J = \max\{i \in \mathbb{N} : R_P J \subseteq R_P P^i\}.$$

If R is a local ring, we will just write $\mathrm{ord}\, J$ to denote the order of J at the maximal ideal of R. The local order of X at a point a is defined as $\mathrm{ord}_a X = \mathrm{ord}\, \mathcal{I}_{X,a}$ where $\mathcal{I}_{X,a}$ is the stalk of the ideal sheaf $\mathcal{I}_X$ of X at a. The order has several good properties which make it well-suited for the embedded resolution of singularities. Firstly, it is upper semicontinuous. In other words, the sets $X_{\geq i} = \{a \in X : \mathrm{ord}_a X \geq i\}$ are closed for all $i \in \mathbb{N}$. In particular, the *top locus* $\mathrm{top}(X)$ of points on X of maximal order is closed. Secondly, the order does not increase when passing to the strict transform of X under blowups in regular centers which are contained in the top locus. The local order of X constitutes the first entry of our resolution invariant $\mathrm{inv}_a(X)$ and will be denoted by o.

The locus of exceptional components which have accumulated so far will be denoted by E. By construction, E is a simple normal crossings divisor on W. As in many other proofs for the embedded resolution of singularities, we make use of a local subdivision of E into *old* and *new* components. Let $a \in X$ be a point. We say that a component of E which passes through a is *old* (at a) if it was created while the maximal value of the order of X was still strictly bigger than $\mathrm{ord}_a X$. The remaining components of E which pass through a are called *new*. They will be denoted by E_a.

We say that a regular system of parameters x, y, z of $\mathcal{O}_{W,a}$ (or $\widehat{\mathcal{O}}_{W,a}$) is *subordinate* to E_a if $E_a \subseteq V(xy)$ holds. A regular subvariety C of W has *normal crossings* with E if the union $C \cup E$ defines a normal crossings variety. A regular center Z in W is called *permissible* if it is contained in the top locus of X and has normal crossings with E.

The second entry c of the resolution invariant $\mathrm{inv}_a(X)$ is defined as the sum of the local order at a of oo the defining ideal of X and oo of the ideal which defines the old exceptional components at a, oo taking into account multiplicities. Since our centers are chosen as the locus of points where the invariant $\mathrm{inv}_a(X)$ is maximal, it can be shown that they automatically have normal crossings with all old components. The fact that they also have normal crossings with the new components is harder to verify and will be established through computations in local parameters.

A *formal flag* $\mathcal{F}$ in W at a consists of a regular surface $\mathcal{F}_2$ and a regular curve $\mathcal{F}_1 \subset \mathcal{F}_2$, both considered as subschemes of $\mathrm{Spec}(\widehat{\mathcal{O}}_{W,a})$, where $\widehat{\mathcal{O}}_{W,a}$ denotes the completion of $\mathcal{O}_{W,a}$. Regular parameters x, y, z of $\widehat{\mathcal{O}}_{W,a}$ are called *subordinate* to $\mathcal{F}$ if $\mathcal{F}_2 = V(z)$, and $\mathcal{F}_1 = V(z, y)$. We then view $\widehat{\mathcal{O}}_{\mathcal{F}_2,a} := \widehat{\mathcal{O}}_{W,a}/(z)$ as the "completed local ring" of $\mathcal{F}_2$ at a.

Throughout the article (with the exception of Sections 8, 9) we restrict to hypersurfaces X in a regular three-dimensional ambient variety W over an algebraically closed field $\mathbb{K}$ of characteristic $p > 0$ which is defined at a closed point $a \in \mathrm{top}(X)$ by an element $f \in \widehat{\mathcal{O}}_{W,a}$ of the form

$$f = z^{p^e} + F(x, y)$$

where x, y, z are a regular system of parameters for the completed local ring $\widehat{\mathcal{O}}_{W,a}$ and $F \in \mathbb{K}[[x, y]]$ is a power series of order $\mathrm{ord}\, F \geq p^e$ at 0 (allowing power series instead of polynomials has technical reasons). Then the order of X at a equals p^e. We will further always assume that there are no old exceptional components a and hence, the first two entries of the invariant $\mathrm{inv}_a(X)$ are of the form $(o, c) = (p^e, p^e)$ (this is no substantial restriction). Consequently, our goal is to find a sequence of blowups in permissible centers so that the order of the strict transform of X eventually drops at all points below p^e.

A change of parameters $z_1 = z - g(x, y)$ transforms the expansion of f into

$$f = z_1^{p^e} + g(x, y)^{p^e} + F(x, y).$$

Hence, we may assume that, after applying such a change of parameters, no p^e-th powers appear in the expansion of F. In this case, we say that the expansion of $F(x, y)$ is *clean*. This notion depends on the choice of the parameters.

The fact that X is locally defined by a purely inseparable equation is stable under blowups in the following sense: Consider the blowup $\pi : W' \to W$ along a regular center Z contained in the top locus of X. Assume in the following that the expansion of $F(x, y)$ is clean. We may assume that either $Z = V(x, z)$ or $Z = \{a\}$. Let X' be the strict transform of X. It is easy to see that at all points $a' \in \pi^{-1}(a)$ which are not contained in the strict transform of $V(z)$, the strict inequality $\operatorname{ord}_{a'} X' < \operatorname{ord}_a X$ holds. Since our goal of lowering the order is already achieved at these points, we may ignore them. Let $a' \in \pi^{-1}(a)$ be a closed point which is contained in the strict transform of $V(z)$. Then there exists an induced regular system of parameters x, y, z for $\widehat{\mathcal{O}}_{W',a'}$ so that the local ring map $\rho : \widehat{\mathcal{O}}_{W,a} \to \widehat{\mathcal{O}}_{W',a'}$ is of a particular form: If $Z = V(x, z)$, then the map is given by $\rho(x) = x$, $\rho(y) = y$, $\rho(z) = xz$. If $Z = \{a\}$, the map is, after possibly swapping x and y, given by $\rho(x) = x$, $\rho(y) = x(y + t)$, $\rho(z) = xz$ for some constant $t \in \mathbb{K}^*$. We call the transition from the local situation at a to the local situation at $a' \in \pi^{-1}(a)$ a *localized blowup*, denoted by $\pi : (W', a') \to (W, a)$. The strict transform X' is given at a' by

$$f' = z^{p^e} + F'(x, y),$$

where $F' = x^{-p^e} \cdot \rho(F)$. Clearly, $\operatorname{ord}_{a'} X' = \operatorname{ord}_a X$ holds if and only if $\operatorname{ord} F' \geq p^e$.

Notice that the expansion of $F'(x, y)$ is again clean if either $Z = V(x, z)$ or $Z = \{a\}$ and $t = 0$ holds. We say in this case that the localized blowup is *monomial* in the parameters x, y, z. If $t \neq 0$, new p^e-th powers may appear in the expansion of $F'(x, y)$. As we will see in Section 4, it is precisely the appearance of these new p^e-th powers which makes it difficult to measure the improvement of F under blowup.

Iterating such localized blowups, we can define induced parameters x, y, z at all points a' lying over a at which the order of the consecutive strict transform of X has not decreased. Notice that the locus E_{new} of new exceptional components is always locally defined by a monomial in x and/or y. On the other hand, z never defines a new exceptional component.

3 The terminal cases

To lower the order of X at all points, we have to find a sequence of blowups which drops the order of F below p^e. Following the strategy that is successfully used in characteristic zero, such a sequence would be comprised of two distinct parts: First, blowups are applied with the goal to transform F (up to multiplication with a unit) into a monomial supported by new exceptional components. It may have large degree. This intermediate goal will be called the *classical monomial case*. Once it is reached, further, combinatorially given blowups are applied to decrease the degree of the monomial until it is strictly smaller than p^e. Then the order of f will have dropped.

We will introduce a notion of *terminal cases* which is slightly more general than the classical monomial case, but offers more flexibility. These terminal cases will replace the classical monomial case.

In Section 7.3 of [Per17], the terminal cases are defined in the general setting where X is a singular hypersurface in a three-dimensional ambient variety W over a field of arbitrary characteristic. Here, we will only define them for hypersurfaces given by purely inseparable equations. We say that X is in a *terminal case* at a closed point a if there exists a regular system of parameters x, y, z for $\widehat{\mathcal{O}}_{W,a}$ which is subordinate to E_a and for which X is locally defined by an element $f \in \widehat{\mathcal{O}}_{W,a}$ of the form $f = z^{p^e} + F(x,y)$ such that one of the following two conditions hold:

- Either X is in the *monomial case* at a, say, F is of the form $F = x^{r_x} y^{r_y} \cdot u$ where $(r_x, r_y) \notin p^e \cdot \mathbb{N}^2$ and u is a unit. We do not require that the monomial $x^{r_x} y^{r_y}$ is supported on E_a, but if $E_a = \emptyset$ we require that either $r_x = 0$ or $r_y = 0$.
- Or X is in the *small residual case* at a, say, F is of the form (after possibly swapping x and y) $F = y^{kp^e} \cdot g$ where k is a positive integer and $g \in \mathbb{K}[[x,y]]$ a power series of order $0 < \operatorname{ord} g < p^e$.

Similar, but more general cases than the small residual case have been considered in the literature as final cases for the embedded resolution of singularities (cf. the notion of *good points* in [Abh88] and the notion of *exceptional and good* in [EV98]).

The terminal cases have several good properties:

1. There are only finitely many points in the top locus of X at which X is not in a terminal case (Corollary 8.2.2 [Per17]). We will sketch the proof for this fact for varieties defined by a purely inseparable equation in Proposition 1 below.
2. If X is in a terminal case at all closed points $a \in \operatorname{top}(X)$, then the pair (o, c) can be lowered by a combinatorially given resolution process. During this process, the consecutive strict transforms of X remain in terminal cases at all points of their top loci. For varieties defined by purely inseparable equations, the argument is sketched in Proposition 2.

The strategy of first reducing to terminal cases and then applying combinatorial resolution is reflected in the definition of our resolution invariant inv. If X is not in a terminal case at a, then $\mathrm{inv}_a X$ is of the form

$$\mathrm{inv}_a X = (o, c, d, n, s, 0, 0)$$

where $(d, n, s) > (0, 0, 0)$. The definition of (d, n, s) will be indicated in Section 5 for the purely inseparable case and in Section 8 for the general case.

If X is already in a terminal case at a, then

$$\mathrm{inv}_a X = (o, c, 0, 0, 0, k, \ell)$$

where (k, ℓ) is a *combinatorial tuple* (cf. Section 7.3 of [Per17]). This invariant governs the combinatorial resolution process in the sense that the locus where it attains its maximal value defines a permissible center (Proposition 8.2.1 [Per17]) and it strictly decreases when this center is blown up (Proposition 9.2.1, Proposition 9.2.2 [Per17]). Hence, the combinatorial resolution process terminates in finitely many steps, leading to a decrease of the pair (o, c).

Notice that upper semicontinuity is not an issue before reaching a terminal case since the invariant (d, n, s) only has to be considered at oo finitely many points of top(X). To show that we reach a situation where the final strict transform of X is in a terminal case at all points of its top locus, we need to show that the tuple (d, n, s) strictly decreases under localized point blowups until a terminal case is reached.

4 The residual order

Following the ideas from characteristic zero, we want to measure how far X is from the classical monomial case. To this end, factorize F into an exceptional monomial part and a residual part. Choose parameters so that E_a is defined by x, y or xy. Assume that the expansion of F is clean and consider the factorization

$$F(x, y) = M(x, y) \cdot G(x, y),$$

where M is the unique monomial of maximal degree supported on E_a that divides F. The order of G is called the *residual order* of X at a and denoted by the symbol d_{res}. Although its definition seems to be dependent on the choice of local parameters, it is an invariant of X, provided that the parameters are subordinate to E_a, and X is locally at a defined by $f = z^{p^e} + F(x, y)$. Then, after applying the coordinate change $z \mapsto z_1$ which eliminates all p^e-th powers from the expansion of $F(x, y)$, the order of $G(x, y)$ in the above factorization does not depend on the choice of x, y, z. (This fact follows from Proposition 4.1.4 and Proposition 5.1.3 in [Per17].) So we have

$$d_{\text{res}} = \operatorname{ord} G = \operatorname{ord} F - \operatorname{ord}_{E_a} F,$$

where the second term on the right hand side denotes the sum of the orders of F along the components of E_a. Clearly, X is in the classical monomial case at a if and only if $d_{\text{res}} = 0$ holds. In this sense, d_{res} measures the distance from this case.

On the other hand, the residual order is not well-suited as an invariant for the reduction to the classical monomial case. Since it is not upper semicontinuous, it cannot be used to define the center of blowup. And since it may increase under blowup in cases where the order of X remains constant, it cannot be used to measure an improvement. This increase cannot be prevented by the choice of the center. Indeed, the residual order may even increase in the case when X has an isolated singularity at a.

As pointed out earlier, there are only oo finitely many points in $\mathrm{top}(X)$ at which X is not yet in a terminal case. At these points, a local invariant (d, n, s) will be employed to substitute d_{res} and to measure how far X is from a terminal case. The entry d is a refinement of the residual order; n measures the transversality of a certain curve with the exceptional divisor; and s, the order of a certain coefficient ideal, can be seen as the slope of a segment of a Newton polygon. We show in Proposition 4 that the triple (d, n, s) strictly decreases lexicographically under point blowups until either the maximal order of X decreases or a terminal case is reached.

The definition of (d, n, s) requires a good understanding of the long term behaviour of the residual order under sequences of blowups. It will be described in detail in the remainder of this section.

It is well-known ([Moh87, Hau10, HP19a]) that the increase of the residual order can only happen in a very specific situation: Assume that X is defined at a point $a \in X$ by $f = z^{p^e} + F(x, y)$ where the expansion of $F(x, y)$ is clean and $E_a \subseteq V(xy)$. Let $\pi : W' \to W$ be the blowup with center the point a and let $a' \in \pi^{-1}(a)$ be a point at which the order of the strict transform X' of X has remained constant,

$$\mathrm{ord}_{a'} X' = \mathrm{ord}_a X.$$

We say that an exceptional component C of E_a is *preserved* under the localized point blowup if a' lies on its strict transform C'. Otherwise, we say that the component is *lost*.

Denote by d_{res} the residual order of X at a and by d'_{res} the residual order of X' at a'. It turns out that the increase $d'_{\mathrm{res}} > d_{\mathrm{res}}$ can only happen if E_a has two components and both are lost under the localized point blowup. Due to a result by Moh [Moh87], the increase is bounded by

$$d'_{\mathrm{res}} \leq d_{\mathrm{res}} + p^{e-1}.$$

In the case $e = 1$, say $f = z^p + F(x, y)$, the increase is at most 1 and can be remedied without too much efforts [HW14].

However, calculations show, aside from certain special cases of small order, that the residual order *decreases* by more than p^{e-1} in the sequence of point blowups preceding the blowup where the increase happens. This sequence comprises the blowups along which the two exceptional components are created which are then lost in the blowup with the increase. So, in total, the residual order decreases when considered for appropriately defined packages of blowups.[2] Our refinement of the residual order outweighs

[2]This good behaviour seems to fail for higher dimensions, see Section 7.

the momentaneous ups and downs and is able to capture this long-term decrease. It thus produces an invariant which *decreases* under each single blowup until a terminal case is reached. In many circumstances, the invariant coincides with the residual order, but at critical steps of the resolution process it differs so as to avoid increases. The clue here is to consider also the order of ideals along *curves*, not just at points. The motivation for doing so is given below.

We already mentioned that in certain special cases the residual order does not decrease sufficiently beforehand. We show that, whenever it then increases, already a terminal case is reached. Hence, its value is not significant anymore in this case and the increase can be ignored in these special cases.

Since the increase of the residual order is linked to the loss of exceptional components, our refinement has to involve the configuration of the components of E_a. The idea is to consider for each component D of E_a all finite sequences of localized point blowups which are of the following type:

- Under all but the last blowup in the sequence, the component D is preserved.
- Under the last blowup, both the component D and a second exceptional component are lost. Hence, the residual order may increase under this last blowup (and only this one).

For each such sequence, we will devise an upper bound for the value of the residual order of the strict transform of X at the end of the sequence. This can be paraphrased by saying that we predict the future values of the residual order, but we only look so far into the future until all components of E_a are lost. It is easy to see that these particular sequences of localized blowups can be entirely monomialized by a suitable choice of parameters in the local ring $\widehat{\mathcal{O}}_{W,a}$ of the beginning. This in turn makes it possible to control the future values of the residual order along such sequences.

Let us analyze more closely the situation of a single localized point blowup. As indicated in Section 2, we may assume that the map $\widehat{\mathcal{O}}_{W,a} \to \widehat{\mathcal{O}}_{W',a'}$ is of the form $x \mapsto x$, $y \mapsto x(y+t)$, $z \mapsto xz$ and X' is defined at a' by the element

$$f' = z^{p^e} + F'(x,y)$$

where $F'(x,y) = x^{-p^e} F(x, x(y+t))$. The induced parameters x, y, z are again subordinate to $E'_{a'}$. We have

$$E'_{a'} = \begin{cases} V(xy) & \text{if } t = 0 \text{ and } V(y) \subseteq E_a, \\ V(x) & \text{if } t \neq 0 \text{ or } V(y) \not\subseteq E_a. \end{cases}$$

If $t = 0$, the expansion of $F'(x,y)$ is again clean. It is immediate to see that $d'_{\mathrm{res}} \leq d_{\mathrm{res}}$ holds in this case. If $t \neq 0$ and $V(y) \not\subseteq E_a$, we may replace y

by the parameter $y + tx$ since the residual order does not depend on our choice of subordinate parameters. Since this change of parameters makes the localized blowup monomial, $d'_{\mathrm{res}} \leq d_{\mathrm{res}}$ also holds in this case. The same holds true if $t \neq 0$ and $V(x) \not\subseteq E_a$.

Hence, we only have to consider the case $t \neq 0$ and $E_a = V(xy)$. Let $z \mapsto z'_t$ be the change of parameters in $\widehat{\mathcal{O}}_{W',a'}$ which eliminates all p^e-th powers from the expansion of $F'(x, y)$. Then

$$f' = z_t'^{p^e} + F'_t(x, y)$$

and the expansion of $F'_t(x, y)$ is clean. The residual order d'_{res} is given by

$$d'_{\mathrm{res}} = \mathrm{ord}_{a'} F'_t - \mathrm{ord}_{V(x)} F'_t.$$

Now consider the change of parameters $y_t = y + tx$ in $\widehat{\mathcal{O}}_{W,a}$. The expansion of F with respect to the parameters x, y_t is in general not clean. Let

$$f = z_t^{p^e} + F_t(x, y_t)$$

be the cleaned expansion, for some z_t. Notice that the parameters x, y_t, z_t are *no longer* subordinate to E_a. It is straightforward to calculate that the local ring map $\widehat{\mathcal{O}}_{W,a} \to \widehat{\mathcal{O}}_{W',a'}$ is of the form $x \mapsto x$, $y_t \mapsto xy$, $z_t \mapsto xz'_t$. Hence, $F'_t(x, y) = x^{-p^e} F_t(x, xy)$ and $F_t(x, y)$ are related to each other by a monomial substitution of the parameters. Let us consider their associated Newton polygons. For a power series $g \in \mathbb{K}[[x, y]]$ of expansion $g = \sum_{i,j \geq 0} c_{i,j} x^i y^j$ with $c_{i,j} \in \mathbb{K}$, the associated Newton polygon is defined as

$$\mathrm{NP}(g) = \mathrm{conv}\Big(\bigcup_{i,j \geq 0, c_{i,j} \neq 0} (i, j) + \mathbb{R}^2_{\geq 0} \Big) \subseteq \mathbb{R}^2_{\geq 0}$$

where $\mathrm{conv}(\cdot)$ denotes the convex hull.

Let $\mathrm{in}(F_t)$ denote the initial form of F_t and let d_t be the order of $\mathrm{in}(F_t)$ along the curve $V(z, y)$. The following picture (Fig. 1) shows how the Newton polygons of F_t and F'_t as well as d_t and d'_{res} are related:

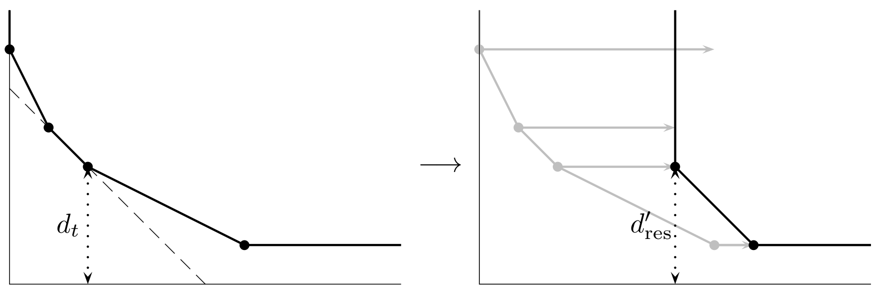

Figure 1: Newton polygons of F_t (left) and F'_t (right).

We see that the inequality $d'_{\text{res}} \leq d_t$ holds. On the other hand, the number d_t is still bounded: For all $t \in \mathbb{K}^*$, the inequality $d_t \leq d_{\text{res}} + p^{e-1}$ holds, as will be seen in Lemma 2 below. Together we get

$$d'_{\text{res}} \leq d_{\text{res}} + p^{e-1},$$

as predicted by Moh's bound.

Notice that the curve $\mathcal{C}_t = V(z_t, y + tx)$ meets both components of $V(z_t) \cap E_a = V(z_t, xy)$ transversally in a, but it does not have normal crossings with $V(z_t) \cap E_a$ since three regular curves meet in a (see Fig. 2).

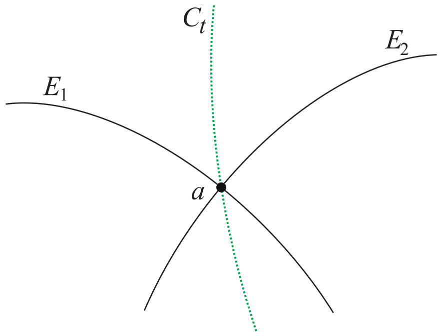

Figure 2: The curve $\mathcal{C}_t$ before blowup.

The strict transform $\mathcal{C}'_t$ of $\mathcal{C}_t$ under π contains a', and both components E_1 and E_2 are lost (see Fig. 3).

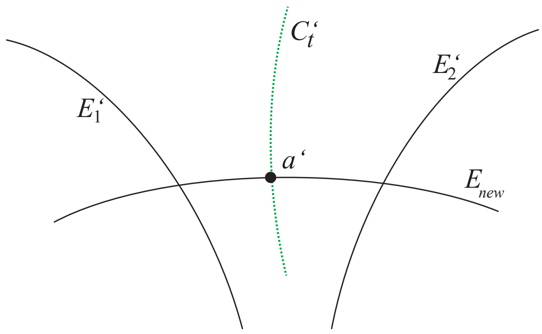

Figure 3: Transform $\mathcal{C}'_t$ of curve $\mathcal{C}_t$.

Now consider sequences of localized point blowups. Let $f \in \widehat{\mathcal{O}}_{W,a}$ be as above and assume that $V(y) \subseteq E_a$ is an exceptional component. Consider a sequence of n localized point blowups during which the component $V(y)$ is preserved in the first $n-1$ blowups and lost in the n-th blowup. Hence, the first $n-1$ blowups are given by local ring maps of the form $x \mapsto x$, $y \mapsto xy$, $z \mapsto xz$ and the last blowup is given by $x \mapsto x$, $y \mapsto x(y+t)$, $z \mapsto xz$ for a non-zero constant $t \in \mathbb{K}^*$. To estimate the residual order at the end of this sequence, consider the parameter $\widetilde{y} = y + tx^n$ in $\widehat{\mathcal{O}}_{W,a}$. This change of parameters monomializes the entire sequence of localized blowups. Let $z \mapsto \widetilde{z}$ be the change of parameters which eliminates all p^e-th powers from the expansion of $F(x, \widetilde{y})$. Hence,

$$f = \widetilde{z}^{p^e} + \widetilde{F}(x, \widetilde{y})$$

where the expansion of $\widetilde{F}(x, \widetilde{y})$ is clean. Introduce weights $\omega(x) = 1$ and $\omega(\widetilde{y}) = n$. For a non-zero power series $g \in \mathbb{K}[[x, y]]$ of expansion $g = \sum_{i,j \geq 0} c_{i,j} x^i y^j$ its *weighted order* is given as

$$\omega(g) = \min\{i + jn : c_{i,j} \neq 0\},$$

while we set $\omega(0) = \infty$. Further, the *weighted initial form* of g with respect to ω is defined as

$$\mathrm{in}_\omega(g) = \sum_{\omega(x^i y^j) = \omega(g)} c_{i,j} x^i y^j.$$

Let $\mathrm{in}_\omega(\widetilde{F})$ be the weighted initial form of $\widetilde{F}$ with respect to ω and let $\widetilde{d}$ be the order of $\mathrm{in}_\omega(\widetilde{F})$ along the curve $V(\widetilde{z}, \widetilde{y})$. It can be shown in the same way as above that $\widetilde{d}$ bounds the residual order of the final strict transform of X at the end of the sequence of localized point blowups.

Again, the curve $\widetilde{C} = V(\widetilde{z}, \widetilde{y})$ has a particular geometric configuration with respect to $V(\widetilde{z}) \cap E_a$: It meets one component of $V(\widetilde{z}) \cap E_a$ tangentially (see Fig. 4). Their intersection multiplicity equals n. This is the same as the number of point blowups which are necessary to separate the two curves.

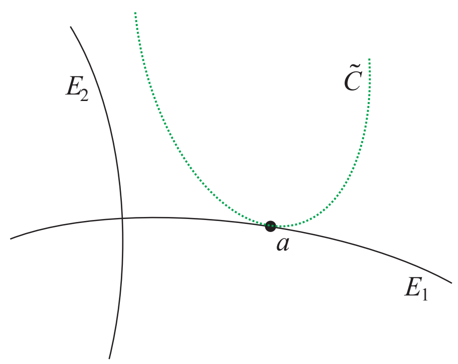

Figure 4: Tangential curve $\widetilde{C}$.

The invariant d we shall define may remain constant along a sequence of blowups as defined before (it will never increase). However, in this case the number n measuring the tangency of C with the exceptional components drops in each step. So the pair (d,n) drops lexicographically. This works well along the previously defined sequences of blowups. We then show that also for other blowups, (d,n) does not increase. This, actually, is the main body of work. It may occasionally remain constant, in which case the next entry s of the invariant comes into play: it is not hard to show that it decreases in these special situations and thus does the job. This reduces the singularity eventually to a terminal case.

5 The resolution invariant

The analysis of the behaviour of the residual order under sequences of blowup in the previous section will now enable us to define a refinement of it which does not decrease under blowup whenever the maximum of the local orders of X remains constant. It thus serves for the induction. The refined invariant replacing the residual order consists of a triple (d,n,s) of three non-negative integers which are considered with respect to the lexicographical ordering. In this section, we introduce it for hypersurfaces defined by purely inseparable equations. The general case is treated in Section 8.

We have seen in Section 3 that there are only finitely many points in $\mathrm{top}(X)$ at which X is not yet in a terminal case. It therefore suffices to define the invariant at these points.

Let $\mathcal{F}:\mathcal{F}_2\supseteq\mathcal{F}_1$ be a formal flag at a. If $\mathcal{F}_2$ has normal crossings with E_a and there exists a regular system of parameters x,y,z for $\widehat{\mathcal{O}}_{W,a}$ which is subordinate to $\mathcal{F}$ such that X is locally at a defined by an element $f\in\widehat{\mathcal{O}}_{W,a}$ which is of the form

$$f=z^{p^e}+F(x,y)$$

with $F\in\widehat{\mathcal{O}}_{\mathcal{F}_2,a}\cong\mathbb{K}[[x,y]]$ of clean expansion $F(x,y)$ with respect to x and y, we say that $\mathcal{F}$ is *adapted* to X and E. Denote by $\mathscr{F}$ the set of these flags.

To X and each flag $\mathcal{F}\in\mathscr{F}$ at a we associate a triple

$$\mathrm{inv}_a^{\mathcal{F}}(X)=(d^{\mathcal{F}},n^{\mathcal{F}},s^{\mathcal{F}}),$$

called *flag invariant* of X at a. The actual invariant (d,n,s) will then be the maximum of these triples, taken over all flags in $\mathscr{F}$. The definition of $\mathrm{inv}_a^{\mathcal{F}}(X)$ distinguishes two cases according to the position of the curve $\mathcal{F}_1$ with respect to the normal crossings divisor $\mathcal{F}_2\cap E_a$ of $\mathcal{F}_2$.

(i) (Case $n^{\mathcal{F}}=0$) First consider the case that $\mathcal{F}_1$ has normal crossings with $\mathcal{F}_2\cap E_a$. This happens either if $E_a=\emptyset$, if $\mathcal{F}_2\cap E_a$ has just one component and $\mathcal{F}_1$ meets this component transversally, or if $\mathcal{F}_1$ equals a component of $\mathcal{F}_2\cap E_a$. We encode this situation by setting $n^{\mathcal{F}}=0$.

Define

$$d^{\mathcal{F}}=d_{\mathrm{res}},$$

the residual order of X at a, as it was defined in Section 4. To define $s^{\mathcal{F}}$, we proceed as in characteristic zero. First notice that there exists a regular system of parameters x, y, z which is both subordinate to $\mathcal{F}$ and E_a. Let $f = z^{p^e} + F$ with $F \in \widehat{\mathcal{O}}_{\mathcal{F}_2,a}$ be as above. By the definition of the residual order, F has a factorization

$$F = M \cdot G,$$

where M is an exceptional monomial and $\operatorname{ord} G = d^{\mathcal{F}} = d_{\text{res}}$. Then $s^{\mathcal{F}}$ is defined as

$$s^{\mathcal{F}} = \begin{cases} \operatorname{ord} \operatorname{coeff}^{d^{\mathcal{F}}}(G) & \text{if } d^{\mathcal{F}} \geq p^e \text{ or } d^{\mathcal{F}} = 0, \\ \operatorname{ord} \operatorname{coeff}^{(p^e - d^{\mathcal{F}})d^{\mathcal{F}}}(M^{d^{\mathcal{F}}} + G^{p^e - d^{\mathcal{F}}}) & \text{if } 0 < d^{\mathcal{F}} < p^e. \end{cases}$$

Here, $\operatorname{coeff}^c(\cdot)$ denotes the *coefficient ideal* with respect to y and $c \in \mathbb{N}$; it is defined for a power series $H = \sum H_i(x) y^i \in \mathbb{K}[[x, y]]$ as

$$\operatorname{coeff}^c(H) = \left(H_i^{\frac{c!}{c-i}}, i < c \right) \subseteq \widehat{\mathcal{O}}_{\mathcal{F}_1,a} \cong \mathbb{K}[[x]].$$

The distinction of two cases in the definition of $s^{\mathcal{F}}$ has the same technical reasons as in characteristic zero, where a *companion ideal* has to be defined to perform the next descent in dimension via passage to a coefficient ideal [EH02, Wło05]. The value of $s^{\mathcal{F}}$ does not depend on the choice of x, y, z as long as these are subordinate to $\mathcal{F}$ and E_a, see Prop. 4.2.7 in [Per17]. Altogether we get in the first case

$$\operatorname{inv}_a^{\mathcal{F}}(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}}) = (d_{\text{res}}, 0, s^{\mathcal{F}}).$$

(ii) (Case $n^{\mathcal{F}} \geq 1$) Secondly, consider the case that $\mathcal{F}_1$ does not have normal crossings with $\mathcal{F}_2 \cap E_a$. This happens if either $\mathcal{F}_1$ is tangent in a to a component of $\mathcal{F}_2 \cap E_a$ (and transversal to the other component if it exists), or if $\mathcal{F}_2 \cap E_a$ has two components and both meet $\mathcal{F}_1$ in a transversally (see Figures 2 and 4). We encode this situation by setting $n^{\mathcal{F}} \geq 1$, where in, the first case, $n^{\mathcal{F}}$ denotes the intersection multiplicity of $\mathcal{F}_1$ and the exceptional component it is tangent to, and where $n^{\mathcal{F}}$ is set equal to 1 in the second case.

Let $f = z^{p^e} + F(x, y)$ be as above, and denote by $\operatorname{in}_\omega(F)$ the weighted initial form of F with respect to the weighted order function (valuation) $\omega : \mathbb{K}[[x, y]] \to \mathbb{N}_\infty$ given by $\omega(x) = 1$ and $\omega(y) = n^{\mathcal{F}}$. We denote by $\operatorname{ord}_\omega(F) = \omega(F)$ the weighted order of F. Set

$$d_{\text{curv}}^{\mathcal{F}} = \operatorname{ord}_{\mathcal{F}_1} \operatorname{in}_\omega(F),$$

the order of $\operatorname{in}_\omega(F)$ along the curve $\mathcal{F}_1$, and

$$d^{\mathcal{F}} = \begin{cases} 0 & \text{if } 0 < d_{\text{curv}}^{\mathcal{F}} < p^e \text{ and } p^e \text{ divides } \operatorname{ord}_\omega(F), \\ d_{\text{curv}}^{\mathcal{F}} & \text{otherwise.} \end{cases}$$

Both numbers $\mathrm{ord}_\omega(F)$ and $d^\mathcal{F}$ are independent of the choice of a regular system of parameters x, y, z which is subordinate to $\mathcal{F}$, see Prop. 4.1.5 in [Per17].

We set $n^\mathcal{F}$ equal to the intersection multiplicity of $\mathcal{F}_1$ with the component of E_a to which it is tangent and set $n^\mathcal{F} = 1$ in case both components of E_a are transversal to $\mathcal{F}_1$. Finally, assign to $s^\mathcal{F}$ the trivial value $s^\mathcal{F} = 0$. Altogether we get in the second case

$$\mathrm{inv}_a^\mathcal{F}(X) = (d^\mathcal{F}, n^\mathcal{F}, 0).$$

This concludes the definition of the flag invariant. Now, if X is not in a terminal case at a, we show in Proposition 3 below that there exists a *maximizing* flag $\mathcal{F} \in \mathscr{F}$, i.e., a flag $\mathcal{F}$ such that $(d^\mathcal{F}, n^\mathcal{F}, s^\mathcal{F})$ is maximal. We set

$$(d, n, s) = \max_{\mathcal{F} \in \mathscr{F}} (d^\mathcal{F}, n^\mathcal{F}, s^\mathcal{F}).$$

We will show that this invariant does the job: If $\pi : (W', a') \to (W, a)$ is a localized blowup and $\mathrm{ord}_{a'}(X') = \mathrm{ord}_a(X)$ has remained constant, then either X' is in a terminal case at a' or the triple (d, n, s) strictly decreases. Note here that the maximizing flag $\mathcal{G}$ at a' which has to be chosen to realize $\mathrm{inv}_{a'}(X')$ need not be the transform of a maximizing flag $\mathcal{F}$ at a. This non-persistence of flags under blowup is the difficulty of the whole argument in positive characteristic: We will show that even though a new maximizing flag may have to be chosen at a', the associated invariant of X' at a' can still be controlled and is lexicographically smaller than the invariant of X at a below. Therefore, after finitely many localized blowups, we must reach a terminal case. As at each stage there are at most finitely many points in $\mathrm{top}(X)$ where X is not yet terminal, the whole process terminates globally.

6 Technical results and proofs

In this section we are going to prove the various statements claimed throughout the article. As before, we will restrict our considerations to the case that X is defined by a purely inseparable equation and refer for the general case to Sections 8 and 9 below, respectively Chapters 8 and 9 of [Per17].

An important tool which we will use in the following proofs are differential operators. We refer to Chapter 1 of [Kaw07] for a very accessible account on this topic. The differential operators that we will use are of the following form: Let $\mathbb{K}$ be an algebraically closed field and let R denote either the polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$ or the power series ring $\mathbb{K}[[x_1, \ldots, x_n]]$. Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ be a multi-index. We denote by $\partial_{x^\alpha} : R \to R$ the $\mathbb{K}$-linear map which acts on the monomials $x^\beta = \prod_{i=1}^n x_i^{\beta_i}$ in R via $\partial_{x^\alpha}(x^\beta) = \binom{\beta}{\alpha} x^{\beta - \alpha}$ where $\binom{\beta}{\alpha} = \prod_{i=1}^n \binom{\beta_i}{\alpha_i}$. Notice here that binomial coefficients have a very peculiar behaviour in positive characteristic. If p is a prime number and n, k are integers with p-adic expansions $n = \sum_{i\geq 0} n_i p^i$, $k = \sum_{i\geq 0} k_i p^i$ where $0 \leq n_i, k_i < p$, then the equality $\binom{n}{k} \equiv \prod_{i\geq 0} \binom{n_i}{k_i} \pmod p$, known as Lucas' Theorem, holds.

We will mainly use the following four properties of the maps ∂_{x^α}, all of which are straightforward to verify:

(1) For all $f \in R$ and $\alpha \in \mathbb{N}^n$, $\operatorname{ord}_a \partial_{x^\alpha}(f) \geq \operatorname{ord} f - |\alpha|$ and $\operatorname{ord}_{(x_i)} \partial_{x^\alpha}(f) \geq \operatorname{ord}_{(x_i)} f - \alpha_i$. Here, a denotes a closed point in $\operatorname{Spec}(R)$.

(2) Fix an index $1 \leq i \leq n$ and let $g \in R$ be an element which does not involve the variable x_i. Define the parameters $y_1, \ldots, y_n$ via $y_i = x_i + g$ and $y_j = x_j$ for $j \neq i$. Then $\partial_{x_i^k} = \partial_{y_i^k}$ for all $k \in \mathbb{N}$.

(3) Let $\operatorname{char}(\mathbb{K}) = p > 0$, $e \in \mathbb{N}$ and $f, g \in R$. Then $\partial_{x_i^k}(f^{p^e} \cdot g) = f^{p^e} \cdot \partial_{x_i^k}(g)$ for all integers k in the range $0 < k < p^e$. In particular, $\partial_{x_i^k}(f^{p^e}) = 0$.

(4) Let $\operatorname{char}(\mathbb{K}) = p > 0$ and $n = 1$. Let $f \in R$ be an element which is not a p^e-th power for some $e \in \mathbb{N}$. Then there exists a number $0 < k < p^e$ such that $\partial_{x_1^k}(f) \neq 0$.

In the following proposition, we will sketch the proof that there are only finitely many points at which X is not in a terminal case. We restrict to the case that X is globally defined by a purely inseparable polynomial $f = z^{p^e} + F(x, y)$ and only consider points lying on the curve $C = V(y, z)$.

Proposition 1. *Let $W = \operatorname{Spec}(\mathbb{K}[x, y, z])$ where $\mathbb{K}$ is an algebraically closed field of characteristic $p > 0$ and let $X \subseteq W$ be a hypersurface which is defined by a polynomial $f \in \mathbb{K}[x, y, z]$ of the form*

$$f = z^{p^e} + y^{r_y} \cdot g(x, y)$$

where $r_y \geq p^e$ and $g \in \mathbb{K}[x, y]$. Denote by C the curve $C = V(y, z)$. Then there are only finitely many closed points $a \in C$ at which X is not in a terminal case.

Proof. By ignoring finitely many points, we may assume without loss of generality that at all points $a \in C$ the divisor E_a either is empty or of the form $E_a = V(y)$.

Further, we may assume that r_y is maximal in the sense that $\operatorname{ord}_{(y)} g = 0$. Define for each closed point $a_t \in C$ with affine coordinates $(t, 0, 0)$ the regular system of parameters $x_t = x - t, y, z$ for $\mathcal{O}_{W,a_t}$.

Consider first the case that $r_y \not\equiv 0 \pmod{p^e}$. Since $\operatorname{ord}_{(y)} g = 0$, the polynomial g is a unit in the local ring $\mathcal{O}_{W,a_t}$ for all but oo finitely many points $a_t \in C$. Hence, X is in the monomial case at these points.

Now consider the case $r_y \equiv 0 \pmod{p^e}$. Write $g(x, y) = g_0(x) + y \cdot G(x, y)$ where $g_0(x) = g(x, 0)$. If g_0 is a p^e-th power, then r_y may be increased via the change of parameters $z_1 = z - (y^{r_y} g_0(x))^{1/p^e}$. Hence, we may assume that g_0 is not a p^e-th power. By property (4) this implies that $\partial_{x^k}(g_0) \neq 0$ for some integer $0 < k < p^e$. Since $\partial_{x^k}(g) = \partial_{x^k}(g_0) + y\partial_{x^k}(G)$, this implies $\operatorname{ord}_{(y)} \partial_{x^k}(g) = 0$. Consequently, both g and $\partial_{x^k}(g)$ are units in the local ring $\mathcal{O}_{W,a_t}$ for all but oo finitely many points $a_t \in C$. Let a_t be such a point. Set $z_t = z - (y^{r_y} \cdot g(t,0))^{1/p^e}$. (Notice that $g(t,0)$ has a p^e-th root since $\mathbb{K}$ is algebraically closed.) Then f has the expansion $f = z_t^{p^e} + y^{r_y} \cdot g_t(x_t, y)$ where $g_t = g - g(t,0)$ fulfills $\operatorname{ord}_{a_t} g_t > 0$. Further, notice that

$$\partial_{x_t^k}(g_t) = \partial_{x_t^k}(g - g(t,0)) = \partial_{x_t^k}(g) = \partial_{x^k}(g)$$

by properties (2) and (3). Hence, $\partial_{x_t^k}(g_t)$ is a unit in the local ring $\mathcal{O}_{W,a_t}$. Consequently, $\operatorname{ord}_{a_t} g_t \leq k < p^e$ by property (1). Hence, X is in the small residual case at a_t. □

In the next proposition, we sketch the proof that once X is in a terminal case at all points where (o,c) is maximal, there exists a combinatorial process which lowers the pair (o,c) globally. Again, we restrict to the purely inseparable case. The main difficulty here is to show that the centers which have to be blown up to improve X are actually permissible.

Proposition 2. *Let $W = \operatorname{Spec}(\mathbb{K}[x,y,z])$ where $\mathbb{K}$ is an algebraically closed field of characteristic $p > 0$ and let $X \subseteq W$ be a hypersurface which is defined by a polynomial $f \in \mathbb{K}[x,y,z]$ of the form $f = z^{p^e} + F(x,y)$. Assume that X is in a terminal case at all closed points $a \in \operatorname{top}(X)$. Then there exists a finite sequence of blowups in permissible centers which lowers the order of the consecutive strict transform of X below p^e at all points.*

Proof. Let $a \in \operatorname{top}(X)$ be a closed point. Let $J_a \subseteq \widehat{\mathcal{O}}_{W,a}$ be the completion of the ideal which defines the top locus $\operatorname{top}(X)$ locally at a. By Lemma 1.2.3.1 [Kaw07], J_a is the radical of the ideal generated by the elements $\partial_{x^i y^j z^k}(f)$ with $i + j + k < p^e$ where x, y, z denote now any regular parameters for $\widehat{\mathcal{O}}_{W,a}$.

Consider first the case that X is in the monomial case. Hence, there exists a regular system of parameters x, y, z for $\widehat{\mathcal{O}}_{W,a}$ such that $f = z^{p^e} + x^{r_x} y^{r_y} \cdot u(x,y)$, $E_a \subseteq V(xy)$, $(r_x, r_y) \notin p^e \cdot \mathbb{N}^2$ and $u \in \mathbb{K}[[x,y]]^*$ is a unit. Assume without loss of generality that $r_x \not\equiv 0 \pmod{p^e}$. Let $0 < i < p^e$ be the residue of r_x modulo p^e. Then, by Lemma 1.2.1.2 (3) [Kaw07], we have that oo

$$\partial_{x^i}(f) = \underbrace{\binom{r_x}{i}}_{\neq 0} x^{r_x - i} y^{r_y} \cdot u + \sum_{j=1}^{i} \binom{r_x}{i-j} x^{r_x - i + j} y^{r_y} \cdot \partial_{x^j}(u) = x^{r_x - i} y^{r_y} \cdot \widetilde{u} \in J$$

for some unit $\widetilde{u} \in \mathbb{K}[[x,y]]^*$. Consequently, $z \in J_a$. By this, it is easy to see that J_a is of the form

$$J_a = \begin{cases} (xy, z) & \text{if } r_x, r_y \geq p^e, \\ (x, z) & \text{if } r_x \geq p^e,\ r_y < p^e, \\ (y, z) & \text{if } r_x < p^e,\ r_y \geq p^e, \\ (x, y, z) & \text{if } r_x, r_y < p^e. \end{cases}$$

Hence, the centers which would be blown up in the corresponding combinatorial process for the classical monomial case in characteristic zero, are also permissible in this case. If $r_x \geq p^e$ holds, the curve center $Z = V(x,z)$ is blown up. Similarly, if $r_y \geq p^e$, then $Z = V(y,z)$ is blown up. (If either of them can be blown up, it is necessary to have a *tiebreak* which decides which curve to blow up first. We will not specify the details here.) If both r_x and r_y are strictly smaller than p^e, only a point center can be blown up. With this choice of centers, it is straightforward to prove that the strict transform X' of X is again in the monomial case at all points a' lying over a at which $\operatorname{ord}_{a'} X' = \operatorname{ord}_a X$ holds and that the degree of the monomial has decreased. Hence, after finitely many blowups, the degree of the monomial has been lowered below p^e.

Now consider the case that X is in the small residual case at a. Hence, $f = z^{p^e} + y^{kp^e} \cdot g(x,y)$ where $E_a \subseteq V(xy)$, k is a positive integer and $g \in \mathbb{K}[[x,y]]$ is an element with $0 < \operatorname{ord} g < p^e$. It is easy to see that there exist numbers i, j with $0 < i + j < p^e$ such that $\partial_{x^i y^j}(g)$ is a unit. Notice that oo

$$\partial_{x^i y^j}(f) = y^{kp^e} \cdot \partial_{x^i y^j}(g)$$

by property (3). Hence, $y \in J_a$ and from this it follows that $J_a = (y,z)$. Consequently, the curve center $Z = V(y,z)$ is locally at a permissible. It is straightforward to verify that the strict transform X' is again in the small residual case at the single point a' lying over a at which the order has not decreased. Further, the order decreases at all points after repeating this process k times. □

For all of the remaining statements in this section, we will use the following local setting: Let X be a hypersurface in a three-dimensional regular ambient variety W which is given locally at a closed point $a \in \operatorname{top}(X)$ by the element $f \in \mathcal{O}_{W,a}$ of expansion

$$f = z^{p^e} + F(x,y)$$

where $\operatorname{ord} F > p^e$ and the expansion of $F(x,y)$ is clean. (In the case $\operatorname{ord} F = p^e$, it is easy to see that a is an isolated point of $\operatorname{top}(X)$ and blowing up this point lowers the order. Hence, this case can be dismissed.) Further, let the normal crossings divisor E_a be given with $E_a \subseteq V(xy)$. We will denote by

d_{res} the residual order of X at a. We always assume in the following that X is *not* in a terminal case at a. Hence, $d_{\text{res}} > 0$.

By the next lemma, it always suffices to consider the flags $\mathcal{F} \in \mathscr{F}$ which arise from (after possibly exchanging x and y) a change of parameters $y_1 = y + h(x)$ and cleaning the expansion of F with respect to x, y_1.

Lemma 1. *Let $\mathcal{F} \in \mathscr{F}$ be a flag. Then there is a flag $\mathcal{G} \in \mathscr{F}$ with $\operatorname{inv}_a^{\mathcal{F}}(X) = \operatorname{inv}_a^{\mathcal{G}}(X)$ and a regular system of parameters x_1, y_1, z_1 for $\widehat{\mathcal{O}}_{W,a}$ such that $\mathcal{G}$ is of the form $\mathcal{G}_2 = V(z_1)$, $\mathcal{G}_1 = V(z_1, y_1)$ and (after possibly swapping x and y) the parameters x_1, y_1, z_1 are of the following form: $x_1 = x$, $y_1 = y + h(x)$ for some power series $h \in \mathbb{K}[[x]]$ and $z_1 = z + g(x, y)$ with $g \in k[[x, y]]$ is the change of parameters which eliminates all p^e-th powers from the power series expansion of F with respect to x, y_1. If $n^{\mathcal{F}} > 0$ holds, then $n^{\mathcal{F}} = n^{\mathcal{G}} = \operatorname{ord} h$.*

Proof. By definition of $\mathscr{F}$, there exists a regular system of parameters x_0, y_0, z_0 for $\widehat{\mathcal{O}}_{W,a}$ such that $\mathcal{F}_2 = V(z_0)$, $\mathcal{F}_1 = V(y_0, z_0)$, $E_a \subseteq V(x_0 y_0)$ and X is locally given at a by $f_0 = z_0^{p^e} + F_0(x_0, y_0)$ with $F_0 \in \mathbb{K}[[x_0, y_0]]$. Since $f_0 = u \cdot f$ for a unit $u \in \widehat{\mathcal{O}}_{W,a}^*$ and $\operatorname{ord} F > p^e$, it is clear that z_0 is z-regular. By the Weierstrass Preparation Theorem, we may assume (after multiplying z_0 with a unit) that z_0 is of the form $z_0 = z + g(x, y)$ for some $g \in \mathbb{K}[[x, y]]$. Consequently, we may write $\mathcal{F}_1 = V(z_0, y_0)$ for some $y_0 \in \mathbb{K}[[x, y]]$ with $\operatorname{ord} y_0 = 1$. After possibly swapping x and y, we may assume that y_0 is y-regular. After applying the Weierstrass Preparation Theorem again, we may assume that $y_0 = y + h(x)$ for some $h \in \mathbb{K}[[x]]$.

Now let $z_1 = z + g_1(x, y_1)$ be the change of parameters which eliminates all p^e-th powers from the expansion of F with respect to x, y_0 and set $x_1 = x$, $y_1 = y_0 = y + h(x)$. It is straightforward to verify that the flag $\mathcal{G}$ given by $\mathcal{G}_2 = V(z_1)$, $\mathcal{G}_1 = V(z_1, y_1)$ fulfills $\operatorname{inv}_a^{\mathcal{G}}(X) = \operatorname{inv}_a^{\mathcal{F}}(X)$. Further, if $n^{\mathcal{F}} > 0$, then

$$n^{\mathcal{F}} = \dim_{\mathbb{K}}(\widehat{\mathcal{O}}_{\mathcal{F}_2,a}/(y, y_0)) = \dim_{\mathbb{K}}(\mathbb{K}[[x, y]]/(y, h(x)) = \operatorname{ord} h$$

holds as claimed. □

The following lemma is crucial to the remaining proofs, as it allows us to give an estimate of the value of $d^{\mathcal{F}}$ for all flags $\mathcal{F} \in \mathscr{F}$ which can be computed directly from the expansion of $F(x, y)$ without a change of parameters. It can be seen as a modified version of Moh's bound for the increase of the residual order under blowup in [Moh87].

Lemma 2. *Let $\mathcal{F} \in \mathscr{F}$ be a flag of the form $\mathcal{F}_2 = V(z_1)$, $\mathcal{F}_1 = V(z_1, y_1)$ where $y_1 = y + h(x)$ and $z_1 = z + g(x, y)$ is the change of parameters which eliminates all p^e-th powers from the expansion of F with respect to x, y_1. Assume that $n^{\mathcal{F}} = \operatorname{ord} h > 0$ and set $n = n^{\mathcal{F}}$. Let $\omega : \mathbb{K}[[x, y]] \to \mathbb{N}_\infty$ be the weighted order that is defined by $\omega(x) = 1$ and $\omega(y) = n$.*

Further, consider a factorization of the weighted initial form $\mathrm{in}_\omega(F)$ *of the form* $\mathrm{in}_\omega(F) = x^a y^b \cdot H(x,y)$. *Then the following inequality holds:*

$$d^{\mathcal{F}}_{\mathrm{curv}} \leq \begin{cases} \operatorname{ord} H + p^{e-1} & \textit{if } p^e \textit{ divides } \operatorname{ord}_\omega(F), \\ \operatorname{ord} H & \textit{otherwise.} \end{cases}$$

In particular, if $\operatorname{ord} H = 0$, *then* $d^{\mathcal{F}} = 0$.

Further, in the special case $\operatorname{ord} H = p^e - p^{e-1}$, *the inequality* $d^{\mathcal{F}}_{\mathrm{curv}} < p^e$ *holds.*

Proof. Write f as $f = z_1^{p^e} + F_1(x, y_1)$ where the expansion of $F_1(x, y_1)$ is clean. We may write $y_1 = y + h(x) = y + tx^n + h_1(x)$ for a non-zero constant $t \in \mathbb{K}^*$ and a power series $h_1 \in \mathbb{K}[[x]]$ with $\operatorname{ord} h_1 > n$. The associated weighted order ω_1 on $\widehat{\mathcal{O}}_{\mathcal{F}_2,a} = \mathbb{K}[[x, y_1]]$ is of the form $\omega_1(x) = 1$, $\omega_1(y_1) = n$. Hence, we may write

$$F = \mathrm{in}_\omega(F)(x, y_1 - tx^n) + K(x, y_1)$$

where $\mathrm{in}_\omega(F)(x, y_1 - tx^n)$ is weighted homogeneous with respect to ω_1 and K is some power series with $\omega_1(K) > \omega_1(F)$. Since F is clean, we know that $\mathrm{in}_\omega(F)(x, y_1 - tx^n)$ is not a p^e-th power. Since $F_1 = F + g^{p^e}$, this proves that $\operatorname{ord}_{\omega_1}(F_1) = \operatorname{ord}_\omega(F)$.

Further, $d^{\mathcal{F}}_{\mathrm{curv}} = \operatorname{ord}_{(y_1)} \mathrm{in}_{\omega_1}(F_1)$ and $\mathrm{in}_{\omega_1}(F_1) = \mathrm{in}_\omega(F) + G^{p^e}$ for some $G \in \mathbb{K}[[x,y]]$.

If $\omega(F)$ is not divisible by p^e, then $G = 0$ and $\mathrm{in}_{\omega_1}(F_1) = \mathrm{in}_\omega(F)$. Hence, it is clear that $d^{\mathcal{F}}_{\mathrm{curv}} = \operatorname{ord}_{(y_1)} \mathrm{in}_\omega(F) \leq \operatorname{ord} H$ holds in this case.

Now assume that p^e divides $\omega(F)$. Let $k < e$ be maximal with the property that $\mathrm{in}_\omega(F)$ is a p^k-th power. It is straightforward to verify that the derivative $\partial_{y^{p^k}}(\mathrm{in}_\omega(F))$ does not vanish. It has a factorization of the form $\partial_{y^{p^k}}(\mathrm{in}_\omega(F)) = x^a y^{b-p^k} \cdot H_1$ for some element $H_1 \in \mathbb{K}[[x,y]]$ with $\operatorname{ord} H_1 = \operatorname{ord} H$. Hence, using properties (1) and (2), we can compute that

$$d^{\mathcal{F}}_{\mathrm{curv}} = \operatorname{ord}_{(y_1)} \mathrm{in}_{\omega_1}(F_1) \leq \operatorname{ord}_{(y_1)} \partial_{y_1^{p^k}}(\mathrm{in}_{\omega_1}(F_1)) + p^k$$

$$= \operatorname{ord}_{(y_1)} \partial_{y^{p^k}}(\mathrm{in}_\omega(F)) + p^k \leq \operatorname{ord}_{(y_1)} H_1 + p^k \leq \operatorname{ord} H + p^{e-1}.$$

If $\operatorname{ord} H = 0$, then $d^{\mathcal{F}} = 0$ follows from the case distinction in the definition of $d^{\mathcal{F}}$.

Now consider the special case $\operatorname{ord} H = p^e - p^{e-1}$ and assume that $d^{\mathcal{F}}_{\mathrm{curv}} = p^e$. From what we have already shown, it follows that $\mathrm{in}_\omega(F)$ is a p^{e-1}-th power and

$$\partial_{y^{p^{e-1}}}(\mathrm{in}_\omega(F)) = \lambda x^a y^{b-p^{e-1}} (y - tx^n)^{p^e - p^{e-1}}$$

for some $\lambda \in \mathbb{K}^*$. Let F' be the p^{e-1}-th root of $\mathrm{in}_\omega(F)$. Set $a' = \frac{a}{p^{e-1}}$ and $b' = \frac{b}{p^{e-1}}$. Notice that $\partial_{y^{p^{e-1}}}(\mathrm{in}_\omega(F)) = (\partial_y(F'))^{p^{e-1}}$. Hence, oo

$$\partial_y(F') = x^{a'}y^{b'-1}(y - tx^n)^{p-1} = x^{a'}y^{b'-1}\sum_{i=0}^{p-1}\binom{p-1}{i}\lambda^{\frac{1}{p^{e-1}}}(-t)^{p-1-i}x^{n(p-1-i)}y^i.$$

On the other hand, since oo F' is of the form $F' = x^{a'}y^{b'}\lambda^{\frac{1}{p^{e-1}}}\sum_{i=0}^{p-1}c_ix^{n(p-1-i)}y^i$ for certain constants $c_i \in \mathbb{K}$, we also have the equality oo oo

$$\partial_y(F') = x^{a'}y^{b'-1}\lambda^{\frac{1}{p^{e-1}}}\sum_{i=0}^{p-1}(b'+i)c_ix^{n(p-1-i)}y^i.$$

Notice that there exists an index $0 \le i < p$ with $(b+i) \equiv 0 \pmod p$, but $\binom{p-1}{i} \not\equiv 0 \pmod p$. Hence, both equalities for $\partial_y(F')$ cannot be fulfilled at the same time. This proves that the strict inequality $d^{\mathcal{F}}_{\mathrm{curv}} < p^e$ holds. □

The next lemma is the essential result which will allow us to prove that $s^{\mathcal{F}}$ is bounded.

Lemma 3. *Assume that $d_{\mathrm{res}} \ge p^e$. Consider flags $\mathcal{F}, \mathcal{G} \in \mathscr{F}$ with $n^{\mathcal{F}} = n^{\mathcal{G}} = 0$ of the form $\mathcal{F}_2 = V(z)$, $\mathcal{F}_1 = V(y,z)$ and $\mathcal{G}_2 = V(z + g(x,y))$ and $\mathcal{G}_1 = V(z + g(x,y), y + h(x))$. Set $s = s^{\mathcal{F}}$ and $d = d_{\mathrm{res}}$. If $s^{\mathcal{G}} > s$ holds, then $\operatorname{ord} h = \frac{s}{d!}$.*

Proof. Assume that $s^{\mathcal{G}} > s$ holds. Hence, $h \neq 0$. Since $n^{\mathcal{G}} = 0$, we know that either $E_a = \emptyset$ or $E_a = V(x)$. Hence, we may write $F = x^{r_x} \cdot G$ where $\operatorname{ord} G = d$ and we set $r_x = 0$ if $E_a = \emptyset$.

Let f have the expansion $f = z_1^{p^e} + F_1(x, y_1)$ where $z_1 = z + g$, $y_1 = y + h$ and the expansion of F_1 is clean. Then $F_1 = F + g^{p^e}$ also has a factorization $F_1 = x^{r_x} \cdot G_1$ for some G_1 with $\operatorname{ord} G_1 = d_{\mathrm{res}}$.

Let G have the expansion $G = \sum_{i\ge 0} G_i(x)y^i$ where $G_i \in \mathbb{K}[[x]]$. Then $s = \min_{i<d}\frac{d!}{d-i}\operatorname{ord} G_i$. In particular, $\operatorname{ord} G_i \ge (d-i)\frac{s}{d!}$ holds for all indices $i \ge 0$. (Notice that this inequality holds trivially for all indices $i \ge d$.)

With respect to the parameters x and y_1, the power series G has the expansion $G = \sum_{i\ge 0}\widetilde{G}_iy_1^i$ with oo $\widetilde{G}_i = \sum_{j\ge i}\binom{j}{i}h^{j-i}G_j$. Recall that the expansion of $F = x^{r_y} \cdot G$ with respect to x, y_1 is not clean and $F_1 = F + g^{p^e}$. Hence, all terms which appear in the expansion of F with respect to x, y_1 which are *not* p^e-th powers also appear in the expansion of F_1.

Assume first that $\operatorname{ord} h > \frac{s}{d!}$ holds. Let the index $i < d$ be such that $\operatorname{ord} G_i = (d-i)\frac{s}{d!}$. Then it is clear that $\mathrm{in}(\widetilde{G}_i) = \mathrm{in}(G_i)$. By cleanness of F, no p^e-th powers appear in the expansion of $x^{r_x}\,\mathrm{in}(G_i)y^i$. Hence, $x^{r_x}\,\mathrm{in}(G_i)y_1^i$ also appears in the expansion of F_1. This proves that $s^{\mathcal{G}} \le \frac{d!}{d-i}\operatorname{ord} G_i = s$.

Now assume that $\operatorname{ord} h < \frac{s}{d!}$ holds. Hence, $s > d!$. Since this implies that $\operatorname{ord} G_i > (d-i)$ for all $i < d$ and we know that $\operatorname{ord} G = d$, it follows that

$\operatorname{in}(G) = G_d(0)y^d$ where $G_d(0) \in \mathbb{K}^*$. In other words, $\operatorname{ord} G_d = 0$. Let $k \in \mathbb{N}$ be maximal with the property that p^k divides d. Then $\binom{d}{p^k} \not\equiv 0 \pmod{p}$.

Consider first the case that $p^k < p^e$. Let $H(x)$ be the initial form of $h(x)$. It is clear that $\operatorname{in}(\widetilde{G}_{p^k}) = \binom{d}{p^k} G_d(0) H^{d-p^k}$. Since $x^{r_x}\binom{d}{p^k}G_d(0)H^{d-p^k}y_1^{p^k}$ is not a p^e-th power, we know that $s^{\mathcal{G}} \leq \frac{d!}{d-p^k} \operatorname{ord} H^{d-p^k} < s$.

Now consider the case $p^k \geq p^e$. Hence, d is divisible by p^e. Since the expansion of $F(x,y)$ is clean and $\operatorname{in}(G) = G_d(0)y^d$, this implies that $r_x \not\equiv 0 \pmod{p^e}$. As before, we know that $\operatorname{in}(\widetilde{G}_0) = G_d(0)H^d$. Notice that $x^{r_x}G_d(0)H^d$ is not a p^e-th power since H^d is a p^e-th power. Hence, we know that $s^{\mathcal{G}} \leq \frac{d!}{d} \operatorname{ord} H^d < s$. □

Proposition 3. *There exists a maximizing flag $\mathcal{F} \in \mathscr{F}$ with $\operatorname{inv}_a^{\mathcal{F}}(X) \in \mathbb{N}^3$ and $d^{\mathcal{F}} > 0$.*

Proof. By Lemma 1 and Lemma 2, the inequality $d^{\mathcal{F}} \leq d_{\text{res}} + p^{e-1}$ holds for all flags $\mathcal{F} \in \mathscr{F}$. Hence, $d^{\mathcal{F}}$ is bounded.

Now consider the weighted orders $\omega_n : \mathbb{K}[[x,y]] \to \mathbb{N}_\infty$ defined by $\omega_n(x) = 1$ and $\omega_n(y) = n$ for $n \in \mathbb{N}$. It is clear that there exists a number $N \in \mathbb{N}$ such that $\operatorname{in}_{\omega_n}(F)$ is, up to a constant, a monomial in x and y for all $n \geq N$. By Lemma 2 (and possibly swapping x and y) oo $d^{\mathcal{F}} > 0$ implies $n^{\mathcal{F}} < N$. In particular, the pair $(d^{\mathcal{F}}, n^{\mathcal{F}})$ is bounded by $(d_{\text{res}} + p^{e-1}, N)$.

Assume now that there exists no maximizing flag. By what we have already shown, this implies that there exists a sequence of flags $\mathcal{F}^i \in \mathscr{F}$ for $i \in \mathbb{N}$ with $\operatorname{inv}_a^{\mathcal{F}^i}(X) = (d_{\text{res}}, 0, s^{\mathcal{F}^i})$ and $s^{\mathcal{F}^i} \geq i$, but no flag $\mathcal{F} \in \mathscr{F}$ with $\operatorname{inv}_a^{\mathcal{F}}(X) = (d_{\text{res}}, 0, \infty)$. oo We may assume by Lemma 1 that $\mathcal{F}_2^i = V(z+g_i)$, $\mathcal{F}_1^i = V(z+g_i, y+h_i)$ for certain power series $g_i \in \mathbb{K}[[x,y]]$, $h_i \in \mathbb{K}[[x]]$. Further, we may assume that $h_0 = 0$.

Assume that $d_{\text{res}} < p^e$ holds. Since $n^{\mathcal{F}^i} = 0$ for all $i \geq 0$, we know that $E_a = V(x)$. Hence, F has a factorization oo $F = x^r \cdot G$ with $\operatorname{ord} G = d_{\text{res}}$ and for all flags $\mathcal{F} \in \mathscr{F}$ with $n^{\mathcal{F}} = 0$ the inequality

$$s^{\mathcal{F}} \leq \operatorname{ord} \operatorname{coeff}^{(p^e - d_{\text{res}})d_{\text{res}}}(x^{rd_{\text{res}}}) = rd_{\text{res}}((p^e - d_{\text{res}})d_{\text{res}} - 1)!$$

holds. This contradicts our assumption that $s^{\mathcal{F}^i}$ is unbounded.

Hence, we may assume that $d_{\text{res}} \geq p^e$. By Lemma 3 we know that $\operatorname{ord}(h_{i+1} - h_i) = \frac{s^{\mathcal{F}^i}}{d!}$. Hence, $y_\infty = y + \sum_{i\geq 0}(h_{i+1} - h_i)$ is a well-defined power series. Let z_∞ be the parameter which arises from cleaning the expansion of F with respect to x, y_∞. Let $\mathcal{F} \in \mathscr{F}$ be the flag $\mathcal{F}_2 = V(z_\infty)$, $\mathcal{F}_1 = V(z_\infty, y_\infty)$. Clearly, $n^{\mathcal{F}} = 0$. Let $k \in \mathbb{N}$ be any non-negative integer. Then $y_\infty = (y + h_k) + H_k$ where $H_k = \sum_{i\geq k}(h_{i+1} - h_i)$. We already know that oo $\operatorname{ord} H_k = \frac{s^{\mathcal{F}^k}}{d!}$. It follows from Lemma 3 that $s^{\mathcal{F}} \geq s^{\mathcal{F}^k} \geq k$. Since this holds for all $k \in \mathbb{N}$, this implies that $s^{\mathcal{F}} = \infty$.

Hence, we have shown that there exists a maximizing flag $\mathcal{F} \in \mathscr{F}$. Assume now that $\operatorname{inv}_a^{\mathcal{F}}(X) \in \mathbb{N}_\infty^3 \setminus \mathbb{N}^3$. By what we have shown, this implies that $\mathrm{inv}_a^{\mathcal{F}}(X) = (d, 0, \infty)$. Let x_1, y_1, z_1 be parameters which are subordinate to $\mathcal{F}$ and E_a and $f_1 = z_1^{p^e} + F_1(x_1, y_1)$ where the expansion of $F_1(x_1, y_1)$ is clean. Since $s^{\mathcal{F}} = \infty$, it is clear that $d_{\mathrm{res}} \geq p^e$ and F has a factorization $F_1 = M \cdot G_1$ where $\mathrm{coeff}^{d_{\mathrm{res}}}(G_1) = 0$. By the definition of the coefficient ideal, this implies that $G_1 = y_1^{d_{\mathrm{res}}} \cdot u$ where u is a unit. Since F_1 is clean oo we have $(r_x, r_y) \notin p^e \cdot \mathbb{N}^2$, which implies that X is in the monomial case at a. This contradicts our assumption that X is not in a terminal case at a.

Now assume that $d^{\mathcal{F}} = 0$ holds for the maximizing curve. This would imply that $d_{\mathrm{res}} = 0$ and X is in the monomial case at a, which contradicts our assumption. $\square$

The proof of the next result is the key ingredient in our proof of resolution of surfaces. It is here that the subtle definition of the invariant comes into play and makes the argument work.

Proposition 4. *Consider a localized point blowup and let $a' \in \pi^{-1}(a)$ be a closed point at which $\mathrm{ord}_{a'} X' = \mathrm{ord}_a X$ holds and such that X' is not in a terminal case at a'. Let $\mathcal{F}$ be a maximizing flag at a and $\mathcal{G}$ a maximizing flag at a'. Then the strict lexicographic inequality*

$$\mathrm{inv}_{a'}^{\mathcal{G}}(X') < \mathrm{inv}_a^{\mathcal{F}}(X)$$

holds. Consequently,

$$\mathrm{inv}_{a'}(X') < \mathrm{inv}_a(X).$$

Proof. Notice that it suffices to find just *any* adapted flag $\mathcal{F}$ at a which fulfills $\mathrm{inv}_{a'}^{\mathcal{G}}(X') < \mathrm{inv}_a^{\mathcal{F}}(X)$ for a chosen maximizing flag $\mathcal{G}$ at a'.

As already lined out in Section 4, we can choose a regular system of parameters x, y, z for $\widehat{\mathcal{O}}_{W,a}$ with the following properties:

- The variety X is defined at a by the element

$$f = z^{p^e} + F(x, y)$$

  where the expansion of $F(x, y)$ is clean.
- The parameters x, y, z are subordinate to E_a, hence $E_a \subseteq V(xy)$.
- The map $\rho : \widehat{\mathcal{O}}_{W,a} \to \widehat{\mathcal{O}}_{W',a'}$ is of the form $\rho(x) = x$, $\rho(y) = x(y + t)$, $\rho(z) = xz$ for some $t \in \mathbb{K}$. Here, x, y, z also denote the induced parameters for $\widehat{\mathcal{O}}_{W',a'}$.
- Either $t = 0$ (hence, one or no component of E_a are lost under the localized blowup) or $t \neq 0$ and $E_a = V(xy)$ (two components of E_a are lost).

The strict transform X' of X is then given at a' by the element

$$f' = z^{p^e} + F'(x, y)$$

where $F'(x,y) = x^{-p^e} F(x, x(y+t))$. If $t = 0$, then the expansion of $F'(x,y)$ is again clean. The divisor $E'_{a'}$ is of the form

$$E'_{a'} = \begin{cases} V(xy) & \text{if } t = 0 \text{ and } V(y) \subseteq E_a, \\ V(x) & \text{if } t \neq 0 \text{ or } V(y) \not\subseteq E_a. \end{cases}$$

Denote by d'_{res} the residual order of X' at a'.

By Lemma 1 we may assume, after possibly swapping the parameters x and y, that $\mathcal{G}$ is of the form $\mathcal{G}_2 = V(z+g(x,y))$, $\mathcal{G}_1 = V(z+g(x,y), y+h(x))$ where $z \mapsto z + g(x,y)$ is the change of parameters which eliminates all p^e-th powers from the expansion of F with respect to x, y_1. Set $y_1 = y + h(x)$ and $z_1 = z + g(x,y)$. Hence,

$$f' = z_1^{p^e} + F'_1(x, y_1)$$

where $F'_1 = F' - g^{p^e}$ and the expansion of $F'_1(x, y_1)$ is clean.

For showing the inequality $\mathrm{inv}_a^{\mathcal{F}}(X) < \mathrm{inv}_{a'}^{\mathcal{G}}(X')$, we will consider four different cases (recall that $n^{\mathcal{G}} = 0$ if $\mathcal{G}_1$ has normal crossings with $\mathcal{G}_2 \cap E'_{a'}$, and that $n^{\mathcal{G}} > 0$ equals the intersection multiplicity of $\mathcal{G}_1$ with the component of $E'_{a'}$ to which it is tangent):

(i) $n^{\mathcal{G}} = 0$ and $t = 0$.

(ii) $n^{\mathcal{G}} = 0$, $t \neq 0$ and $E_a = V(xy)$.

(iii) $n^{\mathcal{G}} \geq 1$ and $\mathcal{G}_1 = V(z + g(x,y), y + h(x))$ with $\mathrm{ord}\, h \geq 1$.

(iv) $n^{\mathcal{G}} > 1$ and $\mathcal{G}_1 = V(z + g(x,y), x + h(y))$ with $\mathrm{ord}\, h \geq 2$.

Cases (i) and (iii) are rather straightforward, whereas (ii) and (iv) are quite subtle.

(i): Since $n^{\mathcal{G}} = 0$, the curve $\mathcal{G}_1$ has normal crossings with $E'_{a'} \cap \mathcal{G}_2$. Hence, we may assume that either $\mathcal{G}_1 = V(z,x)$ or $\mathcal{G}_1 = V(z + g(x,y), y + h(x))$. Notice that $h \neq 0$ implies that $V(y) \not\subseteq E_a$. Hence, by making prior to the blowup the change of coordinates $y \mapsto y + xh(x)$ in $\widehat{\mathcal{O}}_{W,a}$, we may assume without loss of generality that $h = 0$ holds. Thus, $\mathcal{G}_2 = V(z)$ and either $\mathcal{G}_1 = V(z,y)$ or $\mathcal{G}_1 = V(z,x)$. oo Let $\mathcal{F}$ be the flag $\mathcal{F}_2 = V(z)$, $\mathcal{F}_1 = V(z,y)$. It has normal crossings with E_a (say, $n^{\mathcal{F}} = 0$).

Since either one or no component is lost under the localized blowup, we know from Section 4 that $d^{\mathcal{G}} = d'_{\mathrm{res}} \leq d_{\mathrm{res}} = d^{\mathcal{F}}$. Assume that equality holds.

Let F have the factorization $F(x,y) = M(x,y) \cdot G(x,y)$ where M is an exceptional monomial and $\mathrm{ord}\, G = d_{\mathrm{res}}$. Then $F'(x,y) = M'(x,y) \cdot G'(x,y)$ where $M'(x,y) = x^{d_{\mathrm{res}} - p^e} M(x, xy)$ and $G'(x,y) = x^{-d_{\mathrm{res}}} G(x, xy)$ is of order $\mathrm{ord}\, G' = d_{\mathrm{res}}$. Since $\mathrm{ord}\, G' = \mathrm{ord}\, G$ holds, it is easy to see that the initial forms of G and G' are of the form $\mathrm{in}(G) = c_{d_{\mathrm{res}}} y^{d_{\mathrm{res}}}$ and $\mathrm{in}(G') = \sum_{i=0}^{d_{\mathrm{res}}} c_i x^{d_{\mathrm{res}} - i} y^i$ where $c_i \in \mathbb{K}$ and $c_{d_{\mathrm{res}}} \neq 0$. If $\mathcal{G}_1 = V(z,x)$ and oo $d_{\text{res}} \geq p^e$, the numeral $s_{\mathcal{G}}$ hence has the minimal value $s_{\mathcal{G}} = d_{\text{res}}!$; oo if $0 < d_{\text{res}} < p^e$, then $s_{\mathcal{G}} = (d_{\text{res}}(p^e - d_{\text{res}}))!$. Consequently, we may assume without loss of generality that $\mathcal{G}_1 = V(z, y)$.

It is now straightforward to prove the equality of coefficient ideals

$$\mathrm{coeff}^{d^{\mathcal{G}}}(G') = x^{-d_{\text{res}}} \cdot \mathrm{coeff}^{d^{\mathcal{F}}}(G)$$

and, if $0 < d_{\text{res}} < p^e$, then

$$\mathrm{coeff}^{q^{\mathcal{G}}}(M'^{d^{\mathcal{G}}} + G'^{p^e - d^{\mathcal{G}}}) = x^{-q^{\mathcal{F}}} \cdot \mathrm{coeff}^{q^{\mathcal{F}}}(M^{d^{\mathcal{F}}} + G^{p^e - d^{\mathcal{F}}})$$

where $q^{\mathcal{G}} = (p^e - d^{\mathcal{G}})d^{\mathcal{G}}$ and $q^{\mathcal{F}} = (p^e - d^{\mathcal{F}})d^{\mathcal{F}}$. This proves that $s^{\mathcal{G}} < s^{\mathcal{F}}$.

(ii): Set $y_0 = y - tx \in \widehat{\mathcal{O}}_{W,a}$. Let $z_0 = z + g_0(x, y)$ be the change of parameters which eliminates all p^e-th powers from the expansion of F with respect to x, y_0. Let $\mathcal{F}$ be the flag at a given by $\mathcal{F}_2 = V(z_0)$, $\mathcal{F}_1 = V(z_0, y_0)$. Since $E_a = V(xy)$, we know that $n^{\mathcal{F}} = 1$. As argued in Section 4, the flag $\mathcal{F}$ fulfills the inequality $d^{\mathcal{G}} = d'_{\text{res}} \leq d^{\mathcal{F}}_{\text{curv}}$. If $d^{\mathcal{F}} = d^{\mathcal{F}}_{\text{curv}}$, this already proves that $\mathrm{inv}^{\mathcal{G}}_{a'}(X') < \mathrm{inv}^{\mathcal{F}}_a(X)$ since $n^{\mathcal{G}} < n^{\mathcal{F}}$.

Now assume that $d^{\mathcal{F}} = 0$ instead. By definition and since $n^{\mathcal{F}} \geq 1$, this holds if $0 < d^{\mathcal{F}}_{\text{curv}} < p^e$ and $\mathrm{ord}_\omega(F) = kp^e$ for some positive integer k. oo As $\mathrm{ord}(F) = \mathrm{ord}_\omega(F)$, this implies that also $\mathrm{ord}\, F = kp^e$. Consequently, $F'(x, y) = x^{(k-1)p^e} \cdot G'(x, y)$ where $\mathrm{ord}\, G' = d'_{\text{res}} \leq d^{\mathcal{F}}_{\text{curv}} < p^e$. oo We cannot have $\mathrm{ord}\, G' = 0$ since then a p^e-th power would appear in the expansion. Thus $\mathrm{ord}\, G' > 0$. But this gives that X' is in the small residual case at a'. Since we assumed that X' is not in a terminal case at a', both of these possibilities contradict our assumption.

(iii): Since $n^{\mathcal{G}} > 0$ and $\mathcal{G}_1 = V(z + g(x, y), y + h(x))$, we know that $V(y) \subseteq E'_{a'}$ holds. Hence, $t = 0$ and $V(y) \subseteq E_a$. Set $y_0 = y + xh(x)$ and let $z_0 = z + g_0(x, y)$ be the change of coordinates which eliminates all p^e-th powers from the expansion of F with respect to x, y_0. Hence, $f = z_0^{p^e} + F_0(x, y_0)$ and $F_0(x, y_0)$ is clean. It is straightforward to verify that the map $\rho : \widehat{\mathcal{O}}_{W,a} \to \widehat{\mathcal{O}}_{W',a'}$ is of the form $\rho(x) = x$, oo $\rho(y_0) = xy$, $\rho(z_0) = xz$. In particular, $F'_1(x, y) = x^{-p^e} F_0(x, xy)$.

Let $\mathcal{F}$ be the flag $\mathcal{F}_2 = V(z_0)$, $\mathcal{F}_1 = V(z_0, y_0)$ at a. Notice that $n^{\mathcal{F}} = \mathrm{ord}\, h + 1 = n^{\mathcal{G}} + 1$. Further, it is easy to verify that $\mathrm{ord}_\omega(G) = \mathrm{ord}_\omega(F) - p^e$ and $d^{\mathcal{G}}_{\text{curv}} = d^{\mathcal{F}}_{\text{curv}}$ hold. Hence, $d^{\mathcal{G}} = d^{\mathcal{F}}$. Since $n^{\mathcal{G}} < n^{\mathcal{F}}$, this proves that $\mathrm{inv}^{\mathcal{G}}_{a'}(X') < \mathrm{inv}^{\mathcal{F}}_a(X)$.

(iv): As this is the most delicate case, we will give more background information. Let $\omega : \mathbb{K}[[x, y]] \to \mathbb{N}_\infty$ be the weighted order that is defined via oo $\omega(y) = 1$ and $\omega(x) = \mathrm{ord}\, h = n^{\mathcal{G}}$. Denote by $F'_0(x, y)$ the cleaned expansion of $F'(x, y)$. Let the weighted initial form $\mathrm{in}_\omega(F'_0)$ have the factorization $\mathrm{in}_\omega(F'_0) = x^a y^b \cdot H$. By Lemma 2 we know that oo $\mathrm{ord}_\omega(F) = \mathrm{ord}_\omega(F'_0)$ and $d^{\mathcal{F}} \leq \mathrm{ord}\, H + \varepsilon$ where

$$\varepsilon = \begin{cases} p^{e-1} & \text{if } p^e \text{ divides } \mathrm{ord}_\omega(F), \\ 0 & \text{otherwise.} \end{cases}$$

Set $\mathcal{F}_2 = V(z_0)$ and $\mathcal{F}_1 = V(z_0, y_0)$ where the parameters y_0, z_0 are defined in the following way: If $t = 0$, we set $y_0 = y$ and $z_0 = 0$. Hence, $n^{\mathcal{F}} = 0$. On the other hand, if $t \neq 0$ and $E_a = V(xy)$, we set $y_0 = y - tx$ and let $z_0 = z + g_0(x, y)$ be the change of parameters which eliminates all p^e-th powers from the expansion of F with respect to $x, y - tx$. Let $f = z_0^{p^e} + F_0(x, y_0)$ where the expansion of $F_0(x, y_0)$ is clean. Then the equality $F_0'(x, y) = x^{-p^e} F_0(x, xy)$ holds. As the following picture (see Fig. 5, left) of the transformation of the associated Newton polyhedra shows, the inequality $d^{\mathcal{F}} \geq n^{\mathcal{F}} \cdot \operatorname{ord} H$ holds in both cases $n^{\mathcal{F}} = 0$ and $n^{\mathcal{F}} = 1$ (in Fig. 5, we have set $d^{\mathcal{F}} = d^{\mathcal{F}}_{(1)}$ in the case $n^{\mathcal{F}} \geq 1$ and $d^{\mathcal{F}} = d^{\mathcal{F}}_{(2)}$ in the case $n^{\mathcal{F}} = 0$, and d_1 denotes $\operatorname{ord} H$):

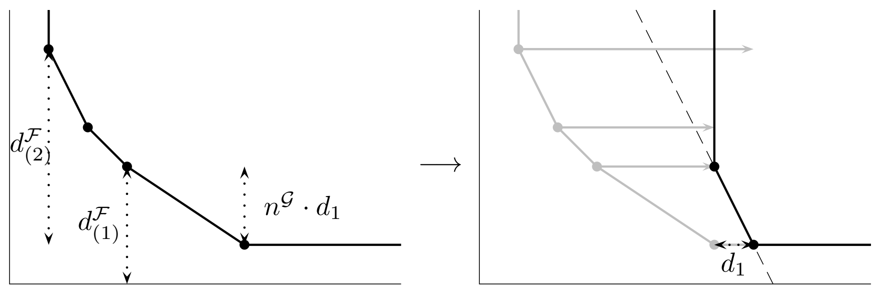

Figure 5: Newton polygons of F_0 (left) and F_0' (right).

If $\varepsilon = 0$, this already implies $d^{\mathcal{G}} \leq \operatorname{ord} H \leq \frac{d^{\mathcal{F}}}{n^{\mathcal{G}}} < d^{\mathcal{F}}$. So assume that $\varepsilon = p^{e-1}$. Hence,

$$d^{\mathcal{G}} \leq \frac{d^{\mathcal{F}}}{n^{\mathcal{G}}} + p^{e-1}.$$

Since we know that p^e divides $\operatorname{ord}_\omega(G)$ and $d^{\mathcal{G}} > 0$, the definition of $d^{\mathcal{G}}$ implies that $d^{\mathcal{G}} \geq p^e$. Now recall that $n^{\mathcal{G}} > 1$. Using this and $p \geq 2$ we can conclude that

$$d^{\mathcal{G}} \leq \operatorname{ord} H + p^{e-1} \leq \frac{d^{\mathcal{F}}}{n^{\mathcal{G}}} + \frac{p^e}{p} \leq \frac{1}{2}(d^{\mathcal{F}} + p^e).$$

But $d^{\mathcal{G}} \geq p^e$, hence $d^{\mathcal{G}} \leq d^{\mathcal{F}}$. Assume that $d^{\mathcal{G}} = d^{\mathcal{F}}$ holds. Then $d^{\mathcal{G}} = p^e$ and $\operatorname{ord} H = p^e - p^{e-1}$. But this is not possible by Lemma 2. It follows that the strict inequality $d^{\mathcal{G}} < d^{\mathcal{F}}$ holds, which implies $\operatorname{inv}^{\mathcal{G}}_{a'}(X') < \operatorname{inv}^{\mathcal{F}}_a(X)$. □

Notice that in case (iv), the image $\pi(\mathcal{G}_1) \subseteq \operatorname{Spec}(\widehat{\mathcal{O}}_{W,a})$ of the curve $\mathcal{G}_1$ is a *singular* curve. Hence, $\mathcal{F}$ necessarily has to be chosen different from the image of $\mathcal{G}$ in this case. The following picture (Fig. 6) illustrates the form of the curves $\pi(\mathcal{G}_1)$ and $\mathcal{F}_1$ in the case $t \neq 0$.

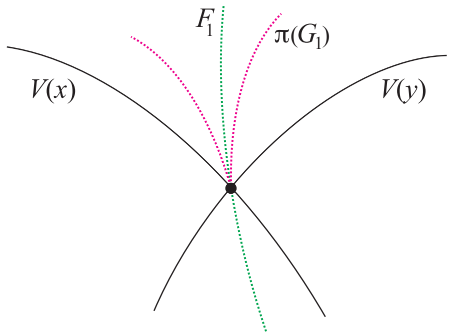

Figure 6: Image curve $\pi(\mathcal{G}_1)$ of $\mathcal{G}_1$ under the blowup π.

Remark. Combining Propositions 1, 2, and 4 establishes the embedded resolution of purely inseparable surfaces in positive characteristic as claimed in the theorem of the introduction.

7 Behaviour of the residual order in higher dimensions

The modification of the residual order which we discussed in Section 4 relies on the fact that the residual order of a purely inseparable equation $z^{p^e} + F(x, y) = 0$ decreases in the long run under sequences of point blowups, at least until a terminal case is reached. Indeed, if it were not for this decrease in the long run, it would be impossible to define a modification of the residual order which strictly decreases in each step.

In higher dimensions, that is to say for a variety defined by a purely inseparable equation $z^{p^e} + F(x_1, \dots, x_n) = 0$ with $n \geq 3$, the residual order behaves much worse than it does in the case of a three-dimensional ambient space. Moh's bound $d' \leq d + p^{e-1}$ for the increase of the residual order under a single permissible blowup (we refer to [Moh87] for the definition of *permissible* in this context) is still valid in higher dimensions. Also, an increase of the residual order can only happen if at least two exceptional components are lost under the localized blowup. But it is not true anymore that the increase of the residual order is always balanced out by the decrease which happened during the earlier blowups which were necessary to create these components. In fact, as a recent example [HP19b] shows, see below, the residual order may even increase indefinitely under sequences of point blowups.

At first sight, such an example may seem to destroy all hopes of defining a similar modification of the residual order in higher dimensions. But actually, a modified residual order in higher dimensions should be expected to strictly decrease under blowup only if the center is chosen correctly. Recall that the choice of center was trivial in the surface case since there are only finitely many points on the top locus of X at which X is not already in a terminal case. Hence, only point blowups are necessary to globally achieve terminal cases. In higher dimensions, this is certainly not true. The question whether it is possible in higher dimensions to always find a sequence of blowups which leads to a decrease of the residual order in the long run is still unsolved. To define an appropriate modification of the residual order in higher dimensions, it would be necessary to first understand which centers have to chosen in order to make the residual order decrease in the long run.

We will describe three examples which illustrate the more erratic behaviour of the residual order in higher dimensions.

Example 1. We consider a hypersurface singularity over a field of characteristic 3 which is defined by a purely inseparable equation of the form $z^9 + F(x, y, w) = 0$. All blowups are point-blowups at the origins of the respective charts. Exceptional multiplicities are only indicated by their residues modulo 9.

1. The starting equation is
   $f = z^9 + x^6 \cdot \big((y^3 + x^2)^3 w^9 + w^{14} + (\ldots)\big)$,
   where $(\ldots)$ denotes higher order terms, and where $x = 0$ defines the unique exceptional component. Therefore, in the computation of the residual order, the factor x^6 has to be neglected and we get $d_{\text{res}} = 14$ from the monomial w^{14} presenting the initial form of F. We blow up the origin.
2. At the origin of the y-chart, we get the strict transform
   $f = z^9 + x^6 y^2 \cdot \big(y(y + x^2)^3 w^9 + w^{14} + (\ldots)\big)$,
   where now both $x = 0$ and $y = 0$ define exceptional components (note that $x = 0$ defines the strict transform of the exceptional component at stage 1, and $y = 0$ defines the new exceptional component). The residual order has dropped to $d_{\text{res}} = 13$, the new initial form of F now being $y^4 w^9$. We blow up the origin.
3. At the origin of the x-chart, we get the strict transform
   $f = z^9 + x^3 y^2 \cdot \big(y(y + x)^3 w^9 + x w^{14} + (\ldots)\big)$,
   where again both $x = 0$ and $y = 0$ define exceptional components. The residual order is again $d_{\text{res}} = 13$, and the initial form of F is $y(y + x)^3 w^9$. We blow up the origin.
4. At the point of coordinates $(0, 1)$ of the x-chart, we get the strict transform
   $f = z^9 + 1 \cdot \big((y^6 - 1) w^9 + x^2 w^{14} + (\ldots)\big)$,
   which, upon replacing z by $z_1 = z - w$, gives
   $f = z_1^9 + 1 \cdot \big(y^6 w^9 + x^2 w^{14} + (\ldots)\big)$.

After this blowup, the residual order has increased to $d_{\text{res}} = 15$. Both old exceptional components are lost, the new exceptional component is defined by $x = 0$ (it does not show up in the polynomial since its multiplicity is 18, of residue 0 modulo 9).

In total, the residual order of f has increased from 14 to 15 in the entire sequence of blowups. Observe here the auxiliary role of the variable w: At the beginning, it is used to keep the residual order smaller than 15, and it appears in the critical term oo $(y^3 + x^2)^3 w^9$ with multiplicity 9. This is mandatory to have the increase of the residual order in the last blowup, see [HP19a] for details. Observe also that larger centers could be chosen, for instance the surface defined by $z = w = 0$ lies in the top locus of f (provided that this is not in conflict with the higher order terms $(\ldots)$. By such choices, the increase of the residual order would be prevented.

Example 2. We consider a hypersurface singularity over a field of characteristic 2 which is defined by a purely inseparable equation of the form $z^2 + F(x, y, w) = 0$. All blowups are point-blowups at the origins of the respective charts. Exceptional multiplicities are only indicated by their residues modulo 2.

1. The starting equation is
   $f = z^2 + \big(yw^2 + (...)\big)$.
   The residual order is $d_{\text{res}} = 3$, no exceptional components occur. We blow up the origin.
2. At the origin of the y-chart, we get the strict transform
   $f = z^2 + y \cdot \big(w^2 + (...)\big)$.
   The new exceptional component is defined by $y = 0$. The residual order has decreased to $d_{\text{res}} = 2$. We blow up the origin.
3. At the origin of the x-chart, we get the strict transform
   $f = z^2 + xy \cdot \big(w^2 + (...)\big)$,
   with the two exceptional components $x = 0$ and $y = 0$. The residual order has remained constant equal to $d_{\text{res}} = 2$. We blow up the origin.
4. At the point with coordinates $(0, 1)$ of the x-chart, we get the strict transform
   $f = z^2 + 1 \cdot \big((y + 1)w^2 + (...)\big)$,
   which, upon replacing z by $z_1 = z + w$, gives
   $f = z_1^2 + 1 \cdot \big(yw^2 + (...)\big)$
   of residual order $d_{\text{res}} = 3$. Both exceptional components were lost under this last blowup.

In total, the residual order has remained constant in the entire sequence of blowups. Note that the initial form of F has the same form as at the beginning. This implies that either the choice of the centers was too small, or that an improvement of the singularities has to be measured within the higher order terms. Again the variable w appears in the initial form with exponent 2 equal to the order of f.

Example 3. We summarize an example from [HP19b] where the residual order increases indefinitely in the long run. The ground field is of characteristic $p \geq 3$, and the singularity will be defined by an equation of the form $f = z^{p^3} + F(v, w, x, y) = 0$. Choose for d a positive integer that is divisible by $\frac{p-1}{2}p$. Set $d' = d + \frac{p-1}{2}p$, $m = \frac{2d}{p-1} + p - 1$, and $q = \frac{p+1}{2}(d + p^2 - 1)$, $q' = \frac{p+1}{2}(d' + p^2 - 1)$.

The letter A will denote an unspecified unit oo $A \in \mathbb{K}[[x_1, \ldots, x_n]]^*$, Q an unspecified power series $Q \in \mathbb{K}[[x_1, \ldots, x_n]]$, and λ an unspecified non-zero constant $\lambda \in \mathbb{K}^*$. These objects can be chosen arbitrarily at the beginning of the sequence and are from then on prescribed by the transformation rules dictated by the blowups.

1. The starting equation is
$f = z^{p^3} + x^{\frac{p-1}{2}p^2} y^{\frac{p^3+1}{2}} w^{-d}(w^d y^{\frac{p^2-1}{2}}(\lambda + v^q \cdot Q) + x^{d+\frac{p^2+1}{2}} v^q \cdot A)$.
The residual order equals $d + \frac{p^2-1}{2}$.
2. After a well chosen sequence of point blowups (see [HP19b]) one obtains the final polynomial
$f = z^{p^3} + x^{\frac{p-1}{2}p^2} y^{\frac{p^3+1}{2}} w^{-d'}(w^{d'} y^{\frac{p^2-1}{2}}(\lambda' + v^{q'} \cdot Q') + x^{d'+\frac{p^2+1}{2}} v^{q'} \cdot A')$
with new λ', Q' and A' satisfying the same properties as λ, Q, and A from the beginning. The residual order equals $d' + \frac{p^2-1}{2}$ and has hence increased by $\frac{p-1}{2}p$.

In this example, the shape of the defining polynomial at the end is the same as at the beginning, only certain exponents are shifted. Therefore, we have a cycle and the process can be repeated. Iterating the cycle produces in the long run a sequence of singularities for which the residual order tends towards infinity. Observe again that the example relies on choosing just point centers for our blowups.

8 The definition of the flag invariant in the general case

In this section we give the details on how the invariant (d, n, s) which was introduced in Section 5 for the purely inseparable case is defined in the general case.

Let X be a hypersurface in a three-dimensional ambient variety W over an algebraically closed field $\mathbb{K}$ of any characteristic. Let E be a simple normal crossings divisor on W. Let $a \in X$ be a closed point. Recall the subdivision of E into old and new components in Section 2. Denote the new components which pass through a by E_a. Let $I_3 \subseteq \widehat{\mathcal{O}}_{W,a}$ be the completion of the ideal which defines the union of X and all old components which pass through a. Set $c = \operatorname{ord} I_3$. If $c = 1$, then X is already locally resolved at a. Hence, we assume that $c > 1$ holds in the following.

A (formal) *flag* $\mathcal{F}$ at a is defined as a curve $\mathcal{F}_1$ and a hypersurface $\mathcal{F}_2$ in $\operatorname{Spec}(\widehat{\mathcal{O}}_{W,a})$ which fulfill $\mathcal{F}_1 \subseteq \mathcal{F}_2$.

A flag $\mathcal{F}$ at a is said to be *compatible* with E if $\mathcal{F}_2$ has normal crossings with E_a and $\mathcal{F}_2 \not\subseteq E_a$. The set of flags at a which are compatible with E will be denoted by $\mathscr{F}(a)$.

Let $\mathcal{F} \in \mathscr{F}(a)$. According to the geometric configuration, we define the *associated multiplicity* $n^{\mathcal{F}}$ of $\mathcal{F}$. If the union $\mathcal{F}_1 \cup (\mathcal{F}_2 \cap E_a)$ has normal crossings, we set $n^{\mathcal{F}} = 0$. Otherwise, we set

$$n^{\mathcal{F}} = \max\{\operatorname{mult}_a(\mathcal{F}_1, \mathcal{F}_2 \cap D) : D \text{ is a component of } E_a\}$$

where $\operatorname{mult}_a(.,.)$ denotes the intersection multiplicity of two curves at a. It is easy to see that $n^{\mathcal{F}} = 1$ can only hold if E_a has exactly two components. If $n^{\mathcal{F}} > 1$, there is a unique component $D_{\mathcal{F}}$ of E_a with $n^{\mathcal{F}} = \operatorname{mult}_a(\mathcal{F}_1, \mathcal{F}_2 \cap D_{\mathcal{F}})$. The component $D_{\mathcal{F}}$ is called the *associated component* of $\mathcal{F}$. Both statements are proved in Lemma 6.6.1 [Per17].

Two flags $\mathcal{F}, \mathcal{G} \in \mathscr{F}(a)$ are said to be *comparable* if they have the same associated multiplicity and (if applicable) the same associated component.

We will now define subordinate parameters and the flag invariants $m^{\mathcal{F}}, d^{\mathcal{F}}$ and $s_{\mathcal{F}}$. The definition will use a case distinction according to whether the associated multiplicity $n^{\mathcal{F}}$ is zero or positive. We will also use the definition of the coefficient ideal of an ideal. Generally, let $R = \mathbb{K}[[x_1, \ldots, x_n]]$ be the power series ring over $\mathbb{K}$ in n variables, let $J \subseteq R$ be an ideal and k a positive integer. Let each element $f \in J$ have the expansion $f = \sum_{i \geq 0} f_i x_n^i$ with $f_i \in \mathbb{K}[[x_1, \ldots, x_{n-1}]]$. Then we set

$$\operatorname{coeff}^k_{(x_1,\ldots,x_n)}(J) = \left(f_i^{\frac{k!}{k-i}} : f \in J, i < k\right) \subseteq R/(x_n) \cong \mathbb{K}[[x_1, \ldots, x_{n-1}]].$$

Let $\mathcal{F} \in \mathscr{F}(a)$ be a flag at a. A regular system of parameters $\mathbf{x} = (x, y, z)$ for $\widehat{\mathcal{O}}_{W,a}$ is said to be *subordinate* to $\mathcal{F}$ if $\mathcal{F}_2 = V(z)$, $\mathcal{F}_1 = V(y, z)$ and $E_a \subseteq V(xy)$. For parameters $\mathbf{x}$ subordinate to $\mathcal{F}$ we will work with the coefficient ideal

$$J_{2,\mathbf{x}} = \operatorname{coeff}^c_{(x,y,z)}(I_3)$$

of I_3 in $\mathbb{K}[[x, y]]$, with $c = \operatorname{ord} I_3$.

(a) Case $n^{\mathcal{F}} = 0$: Let $\mathcal{F}$ be a flag with $n^{\mathcal{F}} = 0$. The ideal $J_{2,\mathbf{x}}$ has a factorization

$$J_{2,\mathbf{x}} = M_{2,\mathbf{x}} \cdot I_{2,\mathbf{x}}$$

where $I_{2,\mathbf{x}}$ is an ideal of $\mathbb{K}[x, y]]$ and $M_{2,\mathbf{x}} = (x^{r_x} y^{r_y})$ is a principal monomial ideal where $r_x, r_y \in \mathbb{N}$ are chosen maximal with the property $V(M_{2,\mathbf{x}}) \subseteq E_a$. We set

$$m^{\mathcal{F}} = \operatorname{ord} M_{2,\mathbf{x}},$$

$$d^{\mathcal{F}} = \operatorname{ord} I_{2,\mathbf{x}}.$$

It can be shown that the numbers $m^{\mathcal{F}}$ and $d^{\mathcal{F}}$ do not depend on the choice of a parameter system $\mathbf{x}$ as long as it is subordinate to the flag $\mathcal{F}$ (Proposition 4.1.4 [Per17]).

Further, define the second coefficient ideal $J_{1,\mathbf{x}}$ as

$$J_{1,\mathbf{x}} = \begin{cases} \operatorname{coeff}^{d^{\mathcal{F}}}_{(x,y)}(I_{2,\mathbf{x}}) & \text{if } d^{\mathcal{F}} \geq c!, \\ \operatorname{coeff}^{d^{\mathcal{F}}(c!-d^{\mathcal{F}})}_{(x,y)}(M_{2,\mathbf{x}}^{d^{\mathcal{F}}} + I_{2,\mathbf{x}}^{c!-d^{\mathcal{F}}}) & \text{if } 0 < d^{\mathcal{F}} < c!, \\ 0 & \text{if } d^{\mathcal{F}} = 0. \end{cases}$$

and set

$$s^{\mathcal{F}} = \operatorname{ord} J_{1,\mathbf{x}}.$$

The number $s_{\mathcal{F}}$ is also independent of the choice of subordinate parameters (Proposition 4.2.7 [Per17]). This defines $\operatorname{inv}^{\mathcal{F}}_a(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}}) = (d^{\mathcal{F}}, 0, s^{\mathcal{F}})$ in the case $n^{\mathcal{F}} = 0$. The number $m^{\mathcal{F}}$ is not part of the invariant and only plays an auxiliary role (it imposes restrictions on how to choose the flags, see below the definition of *valid* flags).

(b) Case $n^{\mathcal{F}} > 0$: Now let $\mathcal{F} \in \mathscr{F}(a)$ be a flag with $n^{\mathcal{F}} > 0$. Take the coefficient ideal $J_{2,\mathbf{x}}$ as above. Let $\omega : \mathbb{K}[[x, y]] \to \mathbb{N}_\infty$ be the weighted order function defined by $\omega(x) = 1$ and $\omega(y) = n^{\mathcal{F}}$. We set

$$m^{\mathcal{F}}_{\mathbf{x}} = \omega(J_{2,\mathbf{x}}),$$

$$d^{\mathcal{F}}_{\mathbf{x}} = \operatorname{ord}_{(y)} \text{wk-in}_\omega(J_{2,\mathbf{x}})$$

where $\text{wk-in}_\omega(J_{2,\mathbf{x}})$ is the weighted *weak initial ideal* of $J_{2,\mathbf{x}}$ which is defined as

$$\text{wk-in}_\omega(J_{2,\mathbf{x}}) = (\text{in}_\omega(F) : F \in J_{2,\mathbf{x}},\, \omega(F) = \omega(J_{2,\mathbf{x}})).$$

Further, we define the invariants $m^{\mathcal{F}}$ and $d^{\mathcal{F}}$ via

$$m^{\mathcal{F}} = \begin{cases} m^{\mathcal{F}}_{\mathbf{x}} & \text{if } m^{\mathcal{F}}_{\mathbf{x}} \geq n^{\mathcal{F}} \cdot c!, \\ n^{\mathcal{F}} \cdot c! & \text{if } m^{\mathcal{F}}_{\mathbf{x}} < n^{\mathcal{F}} \cdot c!, \end{cases}$$

and

$$d^{\mathcal{F}} = \begin{cases} d^{\mathcal{F}}_{\mathbf{x}} & \text{if } d^{\mathcal{F}}_{\mathbf{x}} \geq c!, \\ d^{\mathcal{F}}_{\mathbf{x}} & \text{if } 0 < d^{\mathcal{F}}_{\mathbf{x}} < c! \text{ and } c! \nmid m^{\mathcal{F}}, \\ 0 & \text{if } 0 < d^{\mathcal{F}}_{\mathbf{x}} < c! \text{ and } c! \mid m^{\mathcal{F}}, \\ 0 & \text{if } d^{\mathcal{F}}_{\mathbf{x}} = 0. \end{cases}$$

The case distinctions for small values of $m^{\mathcal{F}}_{\mathbf{x}}$ and $d^{\mathcal{F}}_{\mathbf{x}}$ have technical reasons. Again, it can be shown that the numbers $m^{\mathcal{F}}$ and $d^{\mathcal{F}}$ are independent of the choice of subordinate parameters (Proposition 4.1.5 [Per17]). Assigning to $s^{\mathcal{F}}$ the trivial value $s^{\mathcal{F}} = 0$ we have defined $\mathrm{inv}^{\mathcal{F}}_a(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}}) = (d^{\mathcal{F}}, n^{\mathcal{F}}, 0)$ also in the case $n^{\mathcal{F}} > 0$. The number $m^{\mathcal{F}}$ is again not part of the invariant.

A flag $\mathcal{F} \in \mathscr{F}(a)$ is said to be *valid* if

$$m^{\mathcal{G}} \leq m^{\mathcal{F}}$$

holds for all flags $\mathcal{G} \in \mathscr{F}(a)$ which are comparable to $\mathcal{F}$: This signifies that in the case $n^{\mathcal{F}} = 0$ the order $\mathrm{ord}\, M_{2,\mathbf{x}}$ of the exceptional factor $M_{2,\mathbf{x}}$ has to be maximized, while in the case $n^{\mathcal{F}} > 0$ the weighted order $\omega(J_{2,\mathbf{x}})$ of $J_{2,\mathbf{x}}$ has to be maximized (provided that $m^{\mathcal{F}}_{\mathbf{x}} \geq n^{\mathcal{F}} \cdot c!$). Valid flags are the appropriate generalization of adapted flags which were introduced in the purely inseparable case.

A valid flag $\mathcal{F} \in \mathscr{F}(a)$ is said to be *maximizing* if

$$\mathrm{inv}^{G}_a(X) \leq \mathrm{inv}^{\mathcal{F}}_a(X)$$

holds (with respect to the lexicographic order) for all valid flags $\mathcal{G} \in \mathscr{F}(a)$. This means that $\mathrm{inv}^{\mathcal{F}}_a(X)$ is maximized only taking into account flags which maximize $m^{\mathcal{F}}$.

The existence of a maximizing flag $\mathcal{F} \in \mathscr{F}(a)$ is proved in Proposition 7.4.10 [Per17]. As in Section 5, the triple

$$(d, n, s) = \max_{\mathcal{F} \in \mathscr{F}} (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$$

is defined as the invariant of a maximizing flag.

9 The decrease of the flag invariant in the general case

We continue in the situation of the preceding section. To prove the general case of the embedded resolution of surface singularities in a three-dimensional ambient variety it remains to prove that the local invariant

$$\mathrm{inv}_a(X) = (o, c, d, n, s, k, \ell)$$

is again upper semicontinuous and that it decreases lexicographically at all points $a' \in \pi^{-1}(a)$ of the exceptional divisor when blowing up its top locus in X. Here, o is the order of X at a, $c - o$ is the number of old exceptional components E_{old} at a, and the triple (d, n, s) is defined as in the preceding section as the lexicographic maximum of the flag invariant $\mathrm{inv}_a^{\mathcal{F}}(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$ over all valid flags. Finally, k and ℓ are again combinatorial invariants used solely in the terminal case (i.e., in the monomial or small residual case, defined similiarly as in the purely inseparable situation).

Only in the terminal case the top locus of the invariant will not consist of finitely many points. As the treatment of the this case is mostly combinatorial it presents no additional difficulty in the general case (including the proof of the upper semicontinuity of the invariant). It will therefore be omitted.

In the sequel we will assume that we are not in the terminal case. oo Select for $a \in W$ one of the points where $\mathrm{inv}_a(X)$ attains its maximal value, and recall the notion of valid flags from the last section. Analogously to Proposition 4 we have

Proposition 5. *Consider a localized point blowup and let $a' \in \pi^{-1}(a)$ be a closed point at which $\mathrm{ord}_{a'} X' = \mathrm{ord}_a X$ holds and for which X' is not in a terminal case at a'. Let $\mathcal{G}$ be a valid flag at a'. Then there exists a valid flag $\mathcal{F}$ at a such that the strict lexicographic inequality*

$$\mathrm{inv}_{a'}^{\mathcal{G}}(X') < \mathrm{inv}_a^{\mathcal{F}}(X)$$

holds. Consequently,

$$\mathrm{inv}_{a'}(X') < \mathrm{inv}_a(X).$$

Proof. oo The proof is not so difficult, but somewhat complicated to follow, so we omit it here. Full details can be found in [Per17], Propositions 9.1.1 and 9.1.4. □

References

[Abh56] S. Abhyankar, *Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$*, Ann. of Math. (2) **63** (1956), 491–526.

[Abh65] ———, *Resolution of singularities of arithmetical surfaces*, Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ., 1963), Harper & Row, New York, 1965, pp. 111–152.

[Abh66] ———, *Resolution of singularities of embedded algebraic surfaces*, Pure and Applied Mathematics, Vol. 24, Academic Press, New York-London, 1966.

[Abh88] ———, *Good points of a hypersurface*, Adv. in Math. **68** (1988), no. 2, 87–256.

[Art86] M. Artin, *Lipman's proof of resolution of singularities for surfaces*, Arithmetic geometry (Storrs, Conn., 1984), Springer, New York, 1986, pp. 267–287.

[BV10] A. Bravo and O. Villamayor, *Singularities in positive characteristic, stratification and simplification of the singular locus*, Adv. Math. **224** (2010), no. 4, 1349–1418.

[BV12] A. Benito and O. Villamayor, *Techniques for the study of singularities with applications to resolution of 2-dimensional schemes*, Math. Ann. **353** (2012), no. 3, 1037–1068.

[CJS09] V. Cossart, U. Jannsen, and S. Saito, *Canonical embedded and non-embedded resolution of singularities for excellent two-dimensional schemes*, 2009, preprint, arXiv:0905.2191.

[Cos81] V. Cossart, *Desingularization of embedded excellent surfaces*, Tôhoku Math. J. (2) **33** (1981), no. 1, 25–33.

[CP08] V. Cossart and O. Piltant, *Resolution of singularities of threefolds in positive characteristic. I. Reduction to local uniformization on Artin-Schreier and purely inseparable coverings*, J. Algebra **320** (2008), no. 3, 1051–1082.

[CP09] ———, *Resolution of singularities of threefolds in positive characteristic. II*, J. Algebra **321** (2009), no. 7, 1836–1976.

[Cut09] S. D. Cutkosky, *Resolution of singularities for 3-folds in positive characteristic*, Amer. J. Math. **131** (2009), no. 1, 59–127.

[Cut11] ———, *A skeleton key to Abhyankar's proof of embedded resolution of characteristic p surfaces*, Asian J. Math. **15** (2011), no. 3, 369–416.

[EH02] S. Encinas and H. Hauser, *Strong resolution of singularities in characteristic zero*, Comment. Math. Helv. **77** (2002), no. 4, 821–845.

[EV98] S. Encinas and O. Villamayor, *Good points and constructive resolution of singularities*, Acta Math. **181** (1998), no. 1, 109–158.

[Hau00] H. Hauser, *Excellent surfaces and their taut resolution*, Resolution of singularities (Obergurgl, 1997), Progr. Math., vol. 181, Birkhäuser, Basel, 2000, pp. 341–373.

[Hau10] ———, *On the problem of resolution of singularities in positive characteristic (or: a proof we are still waiting for)*, Bull. Amer. Math. Soc. (N.S.) **47** (2010), no. 1, 1–30.

[Hir64] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II*, Ann. of Math. (2) 79 (1964), 109–203; ibid. (2) **79** (1964), 205–326.

[Hir84] ———, *Desingularization of excellent surfaces, Bowdoin 1967*, Resolution of surface singularities, V. Cossart, J. Giraud, and U. Orbanz, Lecture Notes in Mathematics, vol. 1101, Springer-Verlag, 1984.

[HP19a] H. Hauser and S. Perlega, *Characterizing the increase of the residual order under blowup in positive characteristic*, Publ. Res. Inst. Math. Sci. **55** (2019), 835–857.

[HP19b] ———, *Cycles of singularities appearing in the resolution problem in positive characteristic*, J. Alg. Geom. **28** (2019), 391–403.

[HW14] H. Hauser and D. Wagner, *Alternative invariants for the embedded resolution of purely inseparable surface singularities*, Enseign. Math. **60** (2014), no. 1-2, 177–224.

[Kaw07] H. Kawanoue, *Toward resolution of singularities over a field of positive characteristic. I. Foundation; the language of the idealistic filtration*, Publ. Res. Inst. Math. Sci. **43** (2007), no. 3, 819–909.

[KM10] H. Kawanoue and K. Matsuki, *Toward resolution of singularities over a field of positive characteristic (the idealistic filtration program) Part II. Basic invariants associated to the idealistic filtration and their properties*, Publ. Res. Inst. Math. Sci. **46** (2010), no. 2, 359–422.

[KM16] ———, *Resolution of singularities of an idealistic filtration in dimension 3 after Benito-Villamayor*, Minimal models and extremal rays (Kyoto, 2011), Adv. Stud. Pure Math., vol. 70, Math. Soc. Japan, 2016, pp. 115–214.

[Lip69] J. Lipman, *Rational singularities, with applications to algebraic surfaces and unique factorization*, Inst. Hautes Études Sci. Publ. Math. (1969), no. 36, 195–279.

[Lip78] ———, *Desingularization of two-dimensional schemes*, Ann. Math. (2) **107** (1978), no. 1, 151–207.

[Moh87] T.-T. Moh, *On a stability theorem for local uniformization in characteristic p*, Publ. Res. Inst. Math. Sci. **23** (1987), no. 6, 965–973.

[Per17] S. Perlega, *A new proof for the embedded resolution of surface singularities*, Ph.D. thesis, University of Vienna, 2017.

[Vil07] O. Villamayor, *Hypersurface singularities in positive characteristic*, Adv. Math. **213** (2007), no. 2, 687–733.

[Wło05] J. Włodarczyk, *Simple Hironaka resolution in characteristic zero*, J. Amer. Math. Soc. **18** (2005), no. 4, 779–822 (electronic).

Faculty of Mathematics,
University of Vienna, Austria
E-mail: herwig.hauser@univie.ac.at, stefan.perlega@univie.ac.at

10 Definitions and notation

- *Definitions*

  resolution invariant: strand of non-negative integers associated to X at a, considered lexicographically; requirements: is upper semicontinuous, defines the center of the next blowup, and maximal value decreases under this blowup.

  (formal) flag at a: pair $\mathcal{F} : \mathcal{F}_2 \supseteq \mathcal{F}_1$ of a regular curve $\mathcal{F}_1$ passing through a and contained in a regular surface $\mathcal{F}_2$.

  flag invariant: strand of non-negative integers associated to X at a with respect to a given flag $\mathcal{F}$.

  top locus: closed subvariety of X where a local resolution invariant attains its maximum value.

  permissible center: closed regular subvariety contained in the top locus of the invariant and transversal to exceptional normal crossings divisor.

  monomial blowup: transformation of parameters is given by monomials.

  old exceptional components: component of E through a that was created while the maximal value of the order of X was still strictly bigger than $\mathrm{ord}_a X$.

  new exceptional components: the remaining exceptional components at a, denoted by E_a.

  order: order of vanishing of a power series at a point a, or of an ideal in a local ring wrt to the maximal ideal or weighted order function (valuation).

  coefficient ideal: ideal in regular local hypersurface defined from the coefficients of a polynomial or series f when expanded with respect to a distinguished local parameter defining the hypersurface.

  clean $f = z^{p^e} + F(x, y)$: no p^e-th powers appear in expansion of F.

  exceptional multiplicity: degree of monomial factor M of F produced by earlier blowups.

  residual part: second factor G in $F = M \cdot G$.

  residual order: order of G.

  slope: order of coefficient ideal of G.

  monomial case: F is of the form $F = x^r y^s \cdot u$ where $(r, s) \notin p^e \cdot \mathbb{N}^2$ and u is a unit in $\mathbb{K}[[x, y]]$.

  small residual case: F is of the form (up to swapping x and y) $F = y^{kp^e} \cdot g$, $k > 0$, $g \in \mathbb{K}[[x, y]]$, $0 < \mathrm{ord}\, g < p^e$.

  terminal case: either monomial or small residual case.

  subordinate parameters wrt flag $\mathcal{F}$: regular system of parameters x, y, z of $\widehat{\mathcal{O}}_{W,a}$ with $\mathcal{F}_2 = V(z)$ and $\mathcal{F}_1 = V(z, y)$.

subordinate parameters wrt new exceptional components E_a: regular system of parameters x, y, z of $\widehat{\mathcal{O}}_{W,a}$ with $E_a \subseteq V(xy)$.

compatible flag $\mathcal{F}$: $\mathcal{F}_2$ has normal crossings with E_a and $\mathcal{F}_2 \not\subseteq E_a$.

associated (intersection) multiplicity: integer $n^{\mathcal{F}}$ describing the position of $\mathcal{F}_1$ with respect to E_a.

associated component: component $D_{\mathcal{F}}$ of E_a with $n^{\mathcal{F}} = \text{mult}_a(\mathcal{F}_1, \mathcal{F}_2 \cap D_{\mathcal{F}})$.

comparable flags $\mathcal{F}, \mathcal{G}$: both have the same associated multiplicity and (if applicable) the same associated component.

adapted flag $\mathcal{F}$ wrt X and E (in the purely inseparable case): $\mathcal{F}_2$ has normal crossings with E_a and there exists a regular system of parameters x, y, z for $\widehat{\mathcal{O}}_{W,a}$ subordinate to $\mathcal{F}$ such that X is locally at a defined by an element $f \in \widehat{\mathcal{O}}_{W,a}$ of the form $f = z^{p^e} + F(x, y)$ with clean $F \in \widehat{\mathcal{O}}_{\mathcal{F}_2,a} \cong \mathbb{K}[[x, y]]$.

valid flag $\mathcal{F}$ (in the general case): $\mathcal{F}$ maximizes auxiliary invariant $m^{\mathcal{F}}$, i.e., $m^{\mathcal{G}} \leq m^{\mathcal{F}}$ holds for all flags $\mathcal{G}$ which are comparable to $\mathcal{F}$.

maximizing flag: a valid flag $\mathcal{F}$ which maximizes the value of the associated flag invariant over all valid flags.

- *Notation (purely inseparable case)*

$X \subset W$ singular two-dimensional hypersurface in ambient three-dimensional regular variety W.

E normal crossings divisor in W.

E_a new exceptional components passing through $a \in W$.

E_{old} old exceptional components passing through $a \in W$.

$\mathcal{F} : \mathcal{F}_2 \supseteq \mathcal{F}_1$, flag in W; $\mathcal{F}_2$ regular surface, $\mathcal{F}_1$ regular curve.

$\mathcal{F}' : \mathcal{F}'_2 \supseteq \mathcal{F}'_1$ its transform in W'.

$\mathcal{G}$ flag in W'.

$\mathscr{F}$ collection of adapted flags $\mathcal{F}$.

$\mathscr{F}(a)$ collection of flags at a compatible with E_a.

$f = z^{p^e} + F(x, y)$ purely inseparable polynomial defining X in W.

$F = M \cdot G$ factorization into exceptional part and residual part.

$\text{coeff}^c(H) = (H_i^{\frac{c!}{c-i}}, i < c) \subseteq \mathbb{K}[[x]]$ coefficient ideal of series $H(x, y) = \sum H_i(x) y^i$ with respect to parameters (x, y) and integer c.

$\text{inv}_a X = (o, c, d, n, s, k, \ell)$ local resolution invariant.

ord order of power series or ideal in $\mathbb{K}[[x, y]]$ with respect to the maximal ideal (x, y).

$\text{ord}_{(x)}$, $\text{ord}_{(y)}$ order with respect to the ideals (x), (y).

$\text{ord}_{\mathcal{F}_1}$ order along curve $\mathcal{F}_1$.

o order of $I_{X,a}$.

c order of I_3, the local ideal of X and E_{old}.

$c - o$ order of the ideal of the old exceptional components at a.

$n^{\mathcal{F}}$ associated multiplicity of flag $\mathcal{F}$; codifies the position of the curve $\mathcal{F}_1$ with respect to the new exceptional components E_a:

$n^{\mathcal{F}} = 0$ if $\mathcal{F}_1$ has normal crossings with $\mathcal{F}_2 \cap E_a$;

$n^{\mathcal{F}} \geq 1$ the intersection multiplicity of $\mathcal{F}_1$ with the component of E_a to which it is tangent;

$n^{\mathcal{F}} = 1$ if both components of E_a are transversal to $\mathcal{F}_1$.

$\omega : \mathbb{K}[[x, y]] \to \mathbb{N}_\infty$ weighted order function given by $\omega(x) = 1$ and $\omega(y) = n^{\mathcal{F}}$.

$\operatorname{ord}_\omega(F) = \omega(F)$ weighted order of $F \in \mathbb{K}[[x, y]]$.

$d_{\text{res}} = \operatorname{ord} I_{2,\mathbf{x}} = \operatorname{ord} G$ residual order.

$d^{\mathcal{F}}_{\text{curv}} = \operatorname{ord}_{\mathcal{F}_1} \operatorname{in}_\omega(F)$ order along curve $\mathcal{F}_1$ of weighted initial form $\operatorname{in}_\omega(F)$.

$d^{\mathcal{F}}$ most refined component of invariant;

$$d^{\mathcal{F}} = \begin{cases} d_{\text{res}} & \text{if } n^{\mathcal{F}} = 0, \\ 0 & \text{if } n^{\mathcal{F}} \geq 1, \text{ and } 0 < d^{\mathcal{F}}_{\text{curv}} < p^e \text{ and } p^e \text{ divides } \operatorname{ord}_\omega(F), \\ d^{\mathcal{F}}_{\text{curv}} & \text{if } n^{\mathcal{F}} \geq 1, \text{ and } d^{\mathcal{F}}_{\text{curv}} \geq p^e \text{ or } p^e \text{ does not divide } \operatorname{ord}_\omega(F). \end{cases}$$

$s^{\mathcal{F}}$ slope;

$$s^{\mathcal{F}} = \begin{cases} \operatorname{ord} \operatorname{coeff}^{d^{\mathcal{F}}}(G) & \text{if } n^{\mathcal{F}} = 0, \text{ and } d^{\mathcal{F}} \geq p^e \text{ or } d^{\mathcal{F}} = 0, \\ \operatorname{ord} \operatorname{coeff}^{(p^e - d^{\mathcal{F}})d^{\mathcal{F}}}(M^{d^{\mathcal{F}}} + G^{p^e - d^{\mathcal{F}}}) & \text{if } n^{\mathcal{F}} = 0, \text{ and } 0 < d^{\mathcal{F}} < p^e, \\ 0 & \text{if } n^{\mathcal{F}} \geq 1. \end{cases}$$

$\operatorname{inv}^{\mathcal{F}}_a(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$ flag invariant.

$(d, n, s) = \max(d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$ maximum over all adapted flags $\mathcal{F}$.

(k, ℓ) combinatorial invariants used for terminal case.

- *Notation (general case)*

  $I_{X,a}$ the ideal defining X at a in W.

  E_{old} old exceptional components at a, created before the last decrease of the maximal order of X.

  $I_3 = I_{X,a} \cdot I_{E_{old}}$.

  $\operatorname{inv}_a X = (o, c, d, n, s, k, \ell)$ local resolution invariant.

  o order of $I_{X,a}$.

  c order of I_3.

  $c - o$ number of old exceptional components at a.

  $n^{\mathcal{F}}$ defined as in purely inseparable case.

$\mathrm{coeff}^k_{(x_1,\dots,x_n)}(J) = (f_i^{\frac{k!}{k-i}} : f \in J, i < k) \subseteq \mathbb{K}[[x_1,\dots,x_{n-1}]]$ coefficient ideal of ideal J with respect to parameters $x_1,\dots,x_n$ and integer k, where $f = \sum_{i\geq 0} f_i x_n^i$ with $f_i \in \mathbb{K}[[x_1,\dots,x_{n-1}]]$.

$J_{2,\mathbf{x}} = \mathrm{coeff}^c_{\mathbf{x}}(I_3)$ coefficient ideal of I_3 with respect to parameters $\mathbf{x} = (x,y,z)$ and integer c.

$J_{2,\mathbf{x}} = M_{2,\mathbf{x}} \cdot I_{2,\mathbf{x}}$ factorization into exceptional part and residual part.

$\mathrm{ord}\, M_{2,\mathbf{x}}$ exceptional multiplicity of $J_{2,\mathbf{x}}$.

$d_{\mathrm{res}} = \mathrm{ord}\, I_{2,\mathbf{x}}$ residual order of I_3.

$\omega : \mathbb{K}[[x,y]] \to \mathbb{N}_\infty$, $\omega(x) = 1$ and $\omega(y) = n^{\mathcal{F}}$, weighted order function.

$\mathrm{in}_\omega(F)$ weighted initial form of F.

$\text{wk-in}_\omega(J_{2,\mathbf{x}}) = (\mathrm{in}_\omega(F) : F \in J_{2,\mathbf{x}}, \omega(F) = \omega(J_{2,\mathbf{x}}))$ weighted weak initial ideal.

$\omega(J_{2,\mathbf{x}}) = m^{\mathcal{F}}_{\mathbf{x}} = \mathrm{ord}_\omega(J_{2,\mathbf{x}})$ weighted order of $J_{2,\mathbf{x}}$.

$$m^{\mathcal{F}} = \begin{cases} \mathrm{ord}\, M_{2,\mathbf{x}} & \text{if } n^{\mathcal{F}} = 0, \\ \mathrm{ord}_\omega(J_{2,\mathbf{x}}) & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } \mathrm{ord}_\omega(J_{2,\mathbf{x}}) \geq n^{\mathcal{F}} \cdot c!, \\ n^{\mathcal{F}} \cdot c! & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } \mathrm{ord}_\omega(J_{2,\mathbf{x}}) < n^{\mathcal{F}} \cdot c!. \end{cases}$$

$d^{\mathcal{F}}_{\mathbf{x}} = \mathrm{ord}_{(y)}\, \text{wk-in}_\omega(J_{2,\mathbf{x}})$ order of weighted weak initial ideal $\text{wk-in}_\omega(J_{2,\mathbf{x}})$ along the curve $y = 0$.

$d^{\mathcal{F}}$ most refined component of invariant:

$$d^{\mathcal{F}} = \begin{cases} \mathrm{ord}\, I_{2,\mathbf{x}} & \text{if } n^{\mathcal{F}} = 0, \\ d^{\mathcal{F}}_{\mathbf{x}} & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } d^{\mathcal{F}}_{\mathbf{x}} \geq c!, \\ d^{\mathcal{F}}_{\mathbf{x}} & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } 0 < d^{\mathcal{F}}_{\mathbf{x}} < c! \text{ and } c! \nmid m^{\mathcal{F}}, \\ 0 & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } 0 < d^{\mathcal{F}}_{\mathbf{x}} < c! \text{ and } c! \mid m^{\mathcal{F}}, \\ 0 & \text{if } n^{\mathcal{F}} \geq 1 \text{ and } d^{\mathcal{F}}_{\mathbf{x}} = 0. \end{cases}$$

$J_{1,\mathbf{x}}$ second coefficient ideal with respect to parameters $\mathbf{x} = (x,y)$:

$$J_{1,\mathbf{x}} = \begin{cases} \mathrm{coeff}^{d^{\mathcal{F}}}_{(x,y)}(I_{2,\mathbf{x}}) & \text{if } d^{\mathcal{F}} \geq c!, \\ \mathrm{coeff}^{d^{\mathcal{F}}(c!-d^{\mathcal{F}})}_{(x,y)}(I_{2,\mathbf{x}}^{c!-d^{\mathcal{F}}} + M_{2,\mathbf{x}}^{d^{\mathcal{F}}}) & \text{if } 0 < d^{\mathcal{F}} < c!, \\ 0 & \text{if } d^{\mathcal{F}} = 0. \end{cases}$$

$s^{\mathcal{F}}$ slope:

$$s^{\mathcal{F}} = \begin{cases} \mathrm{ord}\, J_{1,\mathbf{x}} & \text{if } n^{\mathcal{F}} = 0, \text{ and } d^{\mathcal{F}} \geq p^e \text{ or } d^{\mathcal{F}} = 0, \\ 0 & \text{if } n^{\mathcal{F}} \geq 1. \end{cases}.$$

$\mathrm{inv}^{\mathcal{F}}_a(X) = (d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$ flag invariant.

$(d,n,s) = \max(d^{\mathcal{F}}, n^{\mathcal{F}}, s^{\mathcal{F}})$ maximum over all valid flags $\mathcal{F}$.