Wild Singularities and Kangaroo Points for the Resolution of Singularities in Positive Characteristic

HERWIG HAUSER

Wild singularities appear when trying to apply the characteristic zero proof for the resolution of singularities to the case of characteristic $p > 0$. They are a specific type of singularities which produce under blowup at selected points of the exceptional divisor – so called kangaroo points – an increase of the characteristic zero resolution invariant. This increase destroys the induction argument in positive characteristic.

In the article, we describe the structure of wild singularities and the occurrence of kangaroo points, giving a complete characterization of both phenomena. This in turn sheds some light on the difficulties one has to overcome in order to resolve singularities in arbitrary characteristic.

Contents:

Introduction

Let X be a singular subscheme of a smooth scheme W, and let W be equipped with a normal crossings divisor E. In characteristic zero, the embedded resolution of the singularities of X with respect to W and E goes in two steps. First, construct an upper semicontinuous function $\mathrm{inv} : X \to \Gamma$ on X with values in a well ordered set Γ for which the locus Z of points where inv attains its maximal value is smooth and transversal to E. Moreover, the points of X at which X is smooth and transversal to E shall be mapped by inv to the minimal element 0 of Γ. In this way, inv defines a stratification of X in locally closed subschemes. The open dense stratum consists of the smooth points where X is transversal to E and the smallest strata correspond to the (locally) worst singularities. Take Z as the center of a blowup $\pi : W' \to W$, and let X' be the strict or weak transform of X (the precise definitions for these notions are given later on). The second step of the proof consists in showing that at each point a' of X' in $E' = \pi^{-1}(E)$, the value $\mathrm{inv}(a')$ of inv has dropped in comparison to the image point $a = \pi(a')$ in Z. If this is the case, induction shows that finitely many blowups resolve the singularities of X.

In most of the existing proofs for the embedded resolution in arbitrary dimension (at least for hypersurfaces), the function inv is constructed as a vector of non-negative rational numbers whose entries are orders of certain ideals at the points a of W. The first entry is usually the order of the ideal sheaf K defining X in W, i.e., the order of the stalk of K in the local ring $\mathcal{O}_{W,a}$ of W at a. Fix a and let $c = c(a)$ be this order. The second entry of inv is constructed via the choice of a local hypersurface of maximal contact at a and taking the order of the resulting coefficient ideal. To be more explicit, let X be a hypersurface in W with local equation $f = 0$ at a, and let $(x, y) = (x, y_m, \ldots, y_1)$ be a system of local regular parameters in W at a such that the hypersurface V of maximal contact is defined locally at a by $x = 0$.

MSC-2000: 14B05, 14E15, 12D10. Supported by projects P-18992 and P-21461 of the Austrian Science Fund FWF.

It is well known that V contains locally at a the points where K has the same order as at a. Pass to an étale neighborhood of a and expand the Weierstrass form of f as a series in x, say

$$f(x,y) = x^c + \sum_{i=0}^{c-1} a_i(y)x^i.$$

The coefficient ideal of f in V at a is the ideal J in $\mathcal{O}_V$ generated by the weighted coefficients $a_i^{k_i}$ with $k_i = \frac{c!}{c-i}$. The order $e = e(a)$ of J at a is then the natural candidate for the second entry of the invariant inv, say

$$e = \min\{\mathrm{ord}_a(a_i^{k_i}),\ 0 \leq i \leq c-1\}.$$

This gives $\mathrm{inv}_a(K) = (c, e, \ldots) = (\mathrm{ord}_a(K), \mathrm{ord}_a(J), \ldots)$ for all a in X. The remaining entries of inv are constructed in a similar fashion. It is then shown that e (and the other entries) do not depend on the choice of V and the subsequent hypersurfaces of maximal contact, and that inv is upper semicontinuous when varying a. In particular, the second entry e is upper semicontinuous along any stratum of points where the order of K is constant. Thus the first two components of inv are intrinsically defined and upper semicontinuous when considered with respect to the lexicographic ordering.

Along an actual resolution process, the definition of the invariant is slightly more involved, taking into account the decomposition of coefficient ideals into an exceptional monomial factor and a remaining, unkown factor, and also the task of how to achieve through the invariant the required transversality of the center with E.

We only address here the first issue, leaving aside the transversality problem. It is easily observed that the first entry c of inv does not increase under blowup of W in centers contained in the stratum of maximal order of K. Moreover, always in characteristic zero, the transform V' of the local hypersurface of maximal contact V at a contains all points a' of $X' \cap E'$ where c has remained constant, and V' is again a hypersurface of maximal contact for the (strict or weak) transform K' at these points. In particular, the coefficient ideal J' of K' at a' is defined in V'. Its order e', however, may have increased, since J' need not be the weak transform of J. It only equals the controlled transform. This is a transform in between the weak and total transform, defined by deleting a prescribed exceptional factor from the total transform (whereas for the weak transform the maximal possible exceptional factor is deleted).

So in order to ensure the commutativity of taking the weak transform with the descent to coefficient ideals and to ensure the non-increase of of the invariant it is necessary to factor from J and J' the entire exceptional monomial. Then take for the second entry of the invariant the order $o = o(a)$ of the remaining factor. We denote this as $J = M \cdot I$ and $J' = M' \cdot I'$, and redefine the second entry of inv as $o = \mathrm{ord}_a(I)$, $o' = \mathrm{ord}_{a'}(I')$. The factor I passes to the weak transform I' and therefore o does not increase (since the center is contained in the locus of points where (c, o) attains its maximal value, which is assumed throughout). Thus, in total, we have at any point a' of $X' \cap E'$ that the first two entries (c', o') of $\mathrm{inv}_{a'}(K')$ satisfy

$$(c', o') \leq (c, o)$$

with respect to the lexicographic ordering. This is in essence the starting point of the induction argument for the resolution of singularities in characteristic zero.

Unfortunately, the argument does not carry over to positive characteristic. Experimentation shows that the location of the points a' of $X' \cap E'$ with $c' = c$ is much more erratic. In contrast to characteristic zero, these points cannot be "caught" by the successive transforms of smooth hypersurfaces. This complicates the control of the singularities under blowup.

The present paper is concerned with the problem of extending the preceding characteristic zero definitions and constructions of coefficient ideals and their orders to the case of singularities in positive characteristic, and to investigate the behaviour of the shade under blowup. The first entry of the invariant, the order c of K at points, carries over without obstruction, and $c' \leq c$ holds again for any blowup with center contained in the maximal locus S of $\mathrm{ord}_a(K)$. But hypersurfaces of maximal contact need no longer exist. Narasimhan showed that the stratum S need not be contained locally in any smooth hypersurface [Na1, Na2], and Abhyankar's concept of Tschirnhaus transformation breaks down [Ab]. Therefore it is not clear how to choose the coefficient ideal of K and the second entry of the invariant.

The first observation is that, for any smooth local hypersurface V in W at a, the coefficient ideal $J = J_V(K)$ of K in V at a is defined, and thus one can consider $e = \mathrm{ord}_a(J)$, the order of J in V at a (we discard the factorization $J = M \cdot I$ for the moment). Clearly, e will now depend not only on a but also on V. It is therefore not a significant measure for the complexity of the singularity as long as we cannot make it intrinsic, i.e., independent of any choices.

It turns out that, back in characteristic zero, there is an alternative definition of e, and this is the one we will use in positive characteristic: the order e of J at a with respect to a hypersurface of maximal contact V equals the maximum of the orders of the coefficient ideals of K, the maximum being taken over all smooth local hypersurface U in W through a, say

$$\mathrm{ord}_a(J_V(K)) = \max \{\mathrm{ord}_a(J_U(K)),\ U \subset W \text{ smooth}\}.$$

This equality was noticed by Abhyankar (see e.g. [Ab3]) and is proven in [EH]. It suggests to use the maximum as an alternative definition of the second component of the resolution invariant in positive characteristic. Hypersurfaces V which realize the maximum will be called *hypersurfaces of weak maximal contact* [EH, FK].[1] After factoring from J the exceptional monomial (see below for the precise formula), one obtains as the order of the remaining factor a number which again is intrinsic. It will be called in this article the *shade* of K at a, and is denoted by $o = \mathrm{shade}_a(K)$. We thus get again a pair (c, o) of positive integers as the candidates for the first two entries of our resolution invariant. The definition is valid in any characteristic, and coincides in characteristic zero, as mentioned above, with the classical definition.

As a matter of fact, the pair (c, o) behaves in positive characteristic by far not as nice as in characteristic zero. For instance, in the purely inseparable case, say $f = x^c + g(y)$, Moh observed that the shade may increase under blowup along permissible centers at points where c has remained constant [Mo1, Mo2]. Even though the increase can be bounded (see Thm. 1), it prohibits to apply directly induction. Cossart has studied in his thesis extensively this type of increase ("le cas joyeux"), describing many special circumstances and further

[1] Actually, as Frühbis-Krüger observed, one would have to maximize the orders of all iterated coefficient ideals [FK] defined in flags of smooth subschemes. We do not pursue this aspect here.

intricacies [Co1]. In his recent papers with Piltant [CP1, CP2], he succeeds in a tour de force to control the pathologies for three-folds in order to establish (non-embedded) resolution in dimension three and any characteristic. Abhyankar had achieved this earlier for algebraically closed fields of characteristic $p > 5$, with a later refinement by Cossart and a substantial simplification by Cutkosky [Ab4], [Co4], [Cu2].

Another drawback of the shade is its lack of semicontinuity. This was observed by several people, among them Hironaka, Cossart, Piltant, Villamayor and Włodarczyk [Hi1, Co1, Co2, Vi1, WŁ2]. So it is not at all clear whether the shade is an appropriate measure for the complexity of a singularity along the points where the first entry, the order of the ideal, is constant.

Despite of this uncertainty about the relevance of the shade, it is instrumental to understand its behaviour under permissible blowup. Aside of the small hope that the shade can possibly be adjusted so that it does serve as the second component of the resolution invariant (as it happens e.g. for surfaces, see [HW]), the respective study may also lead to the discovery of new invariants and to a clarification of the obstructions for resolution in positive characteristic.

Understanding and explaining the occasional increase of the shade under permissible blowups is the objective of the present paper. Actually, the increase happens very rarely, and the singularity, say the ideal K, has to assume a very special shape in order to make it happen. We propose to call such singularities *wild*, since their transformation under blowup is irregular and seems hard to be controlled. In addition, the points of the exceptional divisor where the increase occurs – the *kangaroo points* – are confined to specific regions. This may raise a certain expectation to be able to take care of these rare exceptions, but this could not be confirmed yet.

Let us emphasize that the present paper focusses on one particular pathology of singularities in positive characteristic. We do not claim that this comprises all possible obstructions to resolution, nor that our study automatically paves the way towards a proof of resolution in arbitrary characteristic and dimension. But the reasons which cause the increase of the shade are intricate, subtle and interesting.

We would like to complement this introduction by mentioning some related work having appeared recently. Hironaka defines and studies the shade in the purely inseparable situation (and then calls it the *residual order*). He describes examples where the upper semicontinuity fails (taking into account non closed points) and proves results on the locus of closed points where the residual order attains its maximal value [Hi1, Hi3]. In his program towards resolution in positive characteristic he uses part of the assertions of the main result of this article [Hi1, Prop. 13.1]. In a different vein, Frühbis-Krüger investigates kangaroo points in higher codimension [FK]. She describes several types of phenomena related to the increase of the shade when passing to iterated coefficient ideals. Hauser and Wagner use the characterization of kangaroo points to give a new proof for the embedded resolution of surfaces by adjusting suitably the pair (c, o) [HW]. This adjustement exploits the local description of wild singularities and results in the definition of a *bonus* which has to be occasionally subtracted from the shade, thus making the invariant drop after each blowup. Similar invariants and adjustments appear in Abhyankar's proof for the embedded resolution of surfaces [Ab1, Cu1], and in Panazzolo's treatment of the resolution of vector fields in dimension three [Pa].

Other authors like Cossart, Benito-Bravo-Encinas-Villamayor or Kawanoue-Matsuki apply differential operators to the coefficient ideals (and variants of it) in order to extract relevant numerical information on the complexity of the singularity [Co3, BeV, BrV, EV, Vi1, Vi2, Vi3, Ka, KM]. This approach is motivated by the work of Giraud [Gi1, Gi2], but has neither been proven yet to be a decisive technique for positive characteristic. The problem with differential operators and more specifically with partial derivatives is that they do not only eliminate p-th powers of polynomials (where p is the characteristic), but also all monomials which have just in the variable of derivation a p-th power (and arbitrary powers in the remaining variables). Thus the information gets kind of distorted, and it seems difficult to recover and to control the original ideal from its derivatives.

Recent advances in resolution of positive characteristic and resolution of vector fields (which reveals in part similar phenomena) have been achieved, among others, by Cossart-Piltant, Cossart-Jannsen-Saito, Moody, McQuillan-Panazzolo, Cano, Teissier, Urabe, Zeillinger [CP1, CP2, CJS, Md1, Md2, MP, Ca, Te, Ur, Ze1]. We refer to [Ha5] for a description of the contents of some of these papers. There have been several results by Abramowich-Karu-Matsuki-Włodarczyk and Cutkosky on the monomialization, respectively toroidalization of morphisms, with partially similar behaviours of the invariants as those for varieties in positive characteristic, see e.g. [AK, Cu3].

Recent papers on resolution in characteristic zero, as well as expository articles or lecture notes, include work of Blanco, Bierstone-Milman-Temkin, Cutkosky, Faber-Hauser, González-Pérez-Teissier, Jannsen, Kollár, Nobile, Temkin, Włodarczyk, Yasuda, Zeillinger [Bl, BMT, Cu4, FH, GT, Ja, Ko, No, Tm1, Tm2, Tm3, Wł1, Ya, Ze2].

Let us now describe in some more detail the contents of the paper. For the simplicity of the exposition we restrict to hypersurface singularities $f = 0$ in W. First, we introduce the shade of a singularity at a given point a as a subordinate invariant after the order of f at a. For this we have to choose a smooth local hypersurface V in W of weak maximal contact with f at a. The shade of f is then defined with respect to an already given, not necessarily reduced normal crossings divisor D in W. All subsequent results are local at a. We consider local blowups of W at a with centers included in V and, more specifically, contained in the locus of points where the coefficient ideal J of f in V has maximal order. For such blowups, we study the behaviour of the shade at the points a' of the exceptional divisor where the order c of f has remained constant.

We prove in Theorem 1 that the shade can only increase at a' if the order c of f at a is a pure p-th power p^b. Moreover, the order of J in V at a must be an integer multiple wc of c, and the residues modulo c of the exceptional multiplicities $q_{n-1}, \dots, q_1$ of the monomial defining $D \cap V$ and factored from J must satisfy a certain arithmetic inequality. This inequality can be expressed through the comparison of lattice points in simplices of $\mathbb{R}^{n-1}$ and their q-translates, where $q = (q_{n-1}, \dots, q_1)$. The inequality implies that a' lies outside all transforms of exceptional components through a whose multiplicity was not a multiple of c.

If all these conditions hold, one may look at the weighted initial form F of f, with respect to weights $(w, 1, \dots, 1)$ on the variables $(x, y_m, \dots, y_1)$ for which V is defined in W by $x = 0$. It turns out that this form must be a purely inseparable polynomial, say of shape $F = x^c + y^q \cdot G(y)$, where G is a homogeneous polynomial of degree $wc - |q|$ which is unique up to c-th power factors and coordinate changes. In addition, the possible location of points a' where the shade may increase is determined by F. Finally, we give an explicit formula for G; for two variables, it was first given by the author, in arbitrary dimension, it is due to Hironaka [Ha4, Hi1]. Hironaka uses some of the assertions of Thm. 1 in his recent program for the resolution in positive characteristic [Hi1]. Moh showed that the shade can increase at most by c/p, and we give an alternative bound [Mo1, Mo2].

Theorem 1 is proven by translating the occurrence of wild singularities and kangaroo points into a statement about weighted homogeneous polynomials, collected in Theorem 2. This theorem represents the key technical ingredient for understanding the increase of the shade. In the last section, we briefly mention how the main theorem can be used for an alternative proof of the embedded resolution of surfaces in $\mathbb{A}^3$.

Acknowledgements. We are very much indebted to G. Müller, J. Schicho, A. Quirós, S. Encinas, O. Villamayor, D. Cutkosky, H. Hironaka, J.-J. Risler, V. Cossart, D. Panazzolo, F. Cano, A. Bravo, R. Blanco and A. Frühbis-Krüger for many discussions and very valuable suggestions on the subject. J. Włodarczyk indicated an inaccuracy in an earlier draft of the paper. It is conceivable that some aspects of the properties listed in Theorem 1 have already been observed or conjectured earlier. We apologize for any omission in this regard. It should be understood that the present article owes very much to the papers and ideas of Zariski, Abhyankar, Hironaka, Giraud, Moh, Cossart, and of many other mathematicians working in the field.

Theorem 1

Definitions. We introduce the basic concepts needed to understand the statement of the theorem. The detailed definitions are given in the section *Preliminaries.* The ground field $\underline{k}$ is of characteristic $p > 0$ and assumed to be algebraically closed; W denotes a smooth ambient scheme of finite type over $\underline{k}$. If not said differently, points of W are always closed points. When necessary, we allow to pass from Zariski-local to étale neighborhoods. For an ideal sheaf K in W, we denote by $\mathrm{ord}_a(K)$ the order of K at a, i.e, the maximal integer k such that the stalk of K at a is contained in m_a^k, for m_a the maximal ideal of the local ring $\mathcal{O}_{W,a}$ of W at a. This definition is also valid for non-closed points a. The *top locus* $\mathrm{top}(K)$ of K is the reduced closed subscheme of W of points where $\mathrm{ord}_a(K)$ attains its maximal value. Let V be a local smooth hypersurface in W at a, with a a closed point. The *coefficient ideal* of K in V at a is the ideal

$$J_V(K) = \sum_{i=0}^{c-1} \langle a_{f,i},\ f \in K\rangle^{\frac{c}{c-i}}$$

in V, where $c = \mathrm{ord}_a(K)$ is the order of K at a and any $f \in K$ is written $f(x, y) = \sum_{i\geq 0} a_{f,i}(y)x^i$ with $a_{f,i} \in \mathcal{O}_{V,a}$, for local coordinates $(x, y_m, \ldots, y_1)$ at a defining V by $x = 0$ (the dependence on the coordinates is not indicated notationally). The rational exponent in the definition has to be understood as the equivalence class of pairs of ideals and numbers, where $(J, k) \equiv (J^l, lk)$ for any $l \in \mathbb{N}$ (or replace in the definition $\frac{c}{c-i}$ by $\frac{c!}{c-i}$). We say that V has *weak maximal contact* with K at a if the order of $J_V(K)$ at a is maximal over all choices of V.

An effective normal crossings divisor D in W is *compatible* with K at a if there is a local smooth hypersurface V transversal to D and of weak maximal contact with K at a such that $J_V(K) = I_V(D \cap V) \cdot I$ for some ideal I in V, with $I_V(D \cap V)$ the ideal defining $D \cap V$ in V. The *shade* of K at a with respect to a compatible divisor D is the number $o = \mathrm{shade}_a(K) = \mathrm{ord}_a(I)$.

Let be given a blowup $\pi : W' \to W$ with smooth center Z transversal to D and contained in $\mathrm{top}(K)$ and $\mathrm{top}(I)$, locally at a. Such blowups and centers will be called *permissible* for K and D at a. Let K' be the weak transform of K in W', and define the transform D' of D as $D' = \pi^{-1}(D)+(o-c)\cdot Y'$ of D, with $Y' = \pi^{-1}(Z)$ the exceptional divisor of π. Let $a' \in Y'$ be a point above a with $\mathrm{ord}_{a'}(K') = \mathrm{ord}_a(K)$, and consider smooth local hypersurfaces V' in W' at a'. As Z is assumed to be transversal to D and contained in $\mathrm{top}(I)$ locally at a, one may choose V' transversal to D' and with weak maximal contact with K' at a' so that the coefficient ideal $J_{V'}(K')$ of K' in V' at a' factors into $J_{V'}(K') = I_{V'}(D' \cap V') \cdot I'$ for some ideal I' in V' (see the section *Preliminaries* below for details). Hence D' is compatible with K' at a' and the shade of K' at a' with respect to D' is defined.

We say that the ideal K defines a *wild singularity* at a with respect to D and a permissible blowup $\pi : W' \to W$, if there exists a point a' in W' above a such that the weak transform K' of K satisfies

$$\mathrm{ord}_{a'}(K') = \mathrm{ord}_a(K),$$

$$\mathrm{shade}_{a'}(K') > \mathrm{shade}_a(K).$$

The points a and a' are then called *antelope*, respectively *kangaroo point*, of the blowup.[2] Obviously, we may (and will) replace in the preceding definitions the ambient scheme W by a suitable neighborhood of a so that V is closed in W and Z is contained in V on whole W.

For a list r of integers and a non zero integer c, let $\phi_c(r)$ denote the number of entries r_i of r which are not divisible by c,

$$\phi_c(r) = \#\{i,\, r_i \not\equiv 0 \text{ mod } c\}.$$

Write $|\overline{r}^c|$ for the sum of the residues $0 \leq \overline{r}_i^c < c$ of the entries of r modulo c.

The following result characterizes wild singularities. For the ease of the exposition, we restrict to principal ideals. A more explicit statement will be given in Theorem 2.

Theorem 1. *Let K be a principal ideal in W of order c at a and let D be an effective normal crossings divisor in W compatible with K. Let V be a hypersurface of weak maximal contact with K at a which is transversal to D. Assume that the coefficient ideal of K in V is non-zero and factorizes into $J_V(K) = I_V(D \cap V) \cdot I$ for some ideal I in V at a. Let $\pi : W' \to W$ be the blowup of W along a smooth center Z which is transversal to D and contained in the loci of points where K and I have maximal order. Then, for K to have a wild singularity at a with respect to D, with kangaroo point a' in $Y' = \pi^{-1}(Z)$, the following conditions must be satisfied.*

[2] In the purely inseparable case, Hironaka investigates an invariant similar to the shade of an ideal, called *residual order*, and also the respective analogue of kangaroo points, called by him *metastatic points* [Hi1, Hi2].

(1) *The order of* K *at* a *is a pure* p*-th power* $c = p^b$*, for some integer* $b \geq 1$.

(2) *The order of* $J_V(K)$ *at* a *is a* c*-multiple* $e = wc$*, for some integer* $w \geq 1$.

(3) *The multiplicities* r_i *of the components of* $D \cap V$ *at* a *whose transforms do not contain* a' *satisfy*

$$(\mathbb{A}) \qquad |\overline{r}^c| \leq (\phi_c(r) - 1) \cdot c.$$

(4) *The multiplicities* s_i *of the components of* $D \cap V$ *at* a *whose transforms contain* a' *are multiples of* c.

(5) *The weighted initial form* F *of a generator of* K *at* a *with respect to the weights* $(w, 1, \ldots, 1)$ *is uniquely determined, up to the choice of local coordinates in* W *at* a *and up to* c*-th power factors, by the orders* c *and* e*, the divisor* D *and the location of* a' *on* Y'.

(5') *In suitable local coordinates* $(x, y) = (x, y_m, \ldots, y_1)$ *in* W *at* a*,* F *is given by a purely inseparable polynomial* $F(x, y) = x^c + y^q \cdot G(y)$*, with* G *a homogeneous polynomial of degree* $o = e - |q|$ *and* $y^q \cdot G(y)$ *not a* c*-th power. Here,* V *and* $D \cap V$ *are defined at* a *by* $x = 0$*, respectively* $y^q = 0$*, with* $q = (r_m, \ldots, r_{j+1}, s_j, \ldots, s_1)$ *and* j *the number of components* s_i*. The polynomial* $y^q \cdot G(y)$ *is uniquely determined, up to coordinate changes in* y *and* c*-th power factors, respectively summands, by the multiplicities* r_i *and* s_i *and the degree* o*. The coordinates* $(y_m, \ldots, y_1)$ *in* V *can be chosen so that, setting* $z = (y_{m-1}, \ldots, y_1)$ *and* $\underline{1} = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{N}^{m-1-j} \times 0^j$*, we have*

$$(\mathbb{G}) \qquad G(1, z + \underline{1}) = \lfloor \prod_{i=j+1}^{m-1} (z_i + 1)^{-r_i} \cdot N(z^c) \rfloor_o$$

for some non-zero polynomial N *of degree* $\leq o/c$.

(6) *For any pure* p*-th power* c*, any list* $q = (r_m, \ldots, r_{j+1}, s_j, \ldots, s_1)$ *as in (3) and (4) and any integer* $o \geq 0$ *so that* $|q| + o$ *is a multiple of* c*, the ideal* K *generated by* $F(x, y) = x^c + y^q \cdot G(y)$ *as in (5') has a wild singularity at* $a = 0$ *with respect to the divisor* D *defined by* $y^q = 0$.

(7) *The increase of the shade of* K *at* a' *is bounded by* $c/p = p^{b-1}$*. It is also bounded by* p^d*, where* $d \geq 0$ *is maximal so that the weighted initial form of* K *is a* p^d*-th power, and by* $|r| - |u|$*, where* $u_i \leq r_i$ *are so that* $|\overline{u}^c| > (\phi_c(u) - 1) \cdot c$.

Comments on Theorem 1. Homogeneous polynomials $P(y) = y^q \cdot G(y)$ of degree $e = wc$ as in (5') will be called *oblique* with respect to $c = p^b$ and q. For each choice of q with decomposition $q = (r_m, \ldots, r_{j+1}, s_j, \ldots, s_1)$, they are characterized and uniquely determined (up to c-th power factors and summands[3]) by the requirement $\mathrm{ord}_z^c P^+(y) > e - |q|$, where $P^+(y) = P(y_m, z + \underline{1} \cdot y_m)$ with $z = (y_{m-1}, \ldots, y_1)$ and $\underline{1} = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{N}^{m-1-j} \times 0^j$, and where $\mathrm{ord}_z^c P^+$ denotes the maximum under the addition of c-th powers of the orders of P^+ with respect to z. This will be proven in the section on oblique polynomials.

[3] We are indebted to J. Włodarczyk for pointing out an inaccuracy at this point of an earlier draft of the paper.

The exponents r_i from (3) and s_i from (4) will be called the *relevant*, respectively *silent multiplicities* of the components of $D \cap V$ at a with respect to a'.

Conditions (1) and (2) of the theorem appear, either implicitly or in the form of examples, in the work of Abhyankar, Cossart and Moh [Ab1, Co1, Mo1, Mo2]. The first bound in (7) is due to Moh, and appears for surfaces already in [Ab1] (see [Cu1] for a concise account of Abhyankar's reasoning). The remaining assertions of the theorem, especially (3), (4), (5) and the first part of (5'), are from [Ha4]. The description of the polynomial G from ($\mathbb{G}$) of (5') is given in [Ha4] only for $m = 2$, together with other characterizations. In [Hi1, Hi2], Hironaka investigates the kangaroo phenomenon in the purely inseparable case $F(x, y) = x^c + y^q \cdot G(y)$ and gives the formula for G for arbitrary m; he calls the degree of G the residual order of F and kangaroo points metastatic. The formula for G was also determined by Schicho. Notice that the multiplicity r_m does not occur in the product, and that the formula determines G. It can easily be shown that changing the polynomial N in (5') only affects $y^q \cdot G(y)$ by introducing or deleting c-th powers. Assertion (6) follows directly from (5') by computation.

The proof of the arithmetic inequality in (3) and the uniqueness of the weighted initial form in (5) is not so difficult, though it is computationally somewhat involved. It relies on an accurate analysis of the behaviour of the weighted initial form F under weighted homogeneous coordinate changes. The inequality (3) signifies in the case where $q = r$ (i.e., no multiplicities s_i occur), that the simplex $\Delta = \{\alpha \in \mathbb{N}^m,\ |\alpha| = o\}$, with $o = e - |r|$, contains more c-multiples than its translate $r + \Delta$ (cf. Lemma 5),

$$|\Delta \cap c \cdot \mathbb{N}^m| > |(r + \Delta) \cap c \cdot \mathbb{N}^m|.$$

The inequality (3) appears, but with strict inequality and a different significance, in the work of Abhyankar on good points [Ab2] (the difference being that the inequality has in our context a negative effect). The inequality will be made explicit in the example below.

By condition (3) the kangaroo point a' does not lie on at least two transforms of components of $D \cap V$ at a. This jumping-off of a', together with the jump of the shade, justifies the naming *kangaroo*. The components of $D \cap V$ at a of (4) can be eliminated by auxiliary blowups with centers of codimension 2, so that a' then only lies on the new exceptional component Y'. Similarly, all components r_i can be made smaller than c. It is reasonable to expect and confirmed by experimentation that the location of the kangaroo point a' on the exceptional divisor is completely determined by the polynomial $y^q \cdot G(y)$.

The uniqueness in (5) is an astonishing circumstance. It signifies that this obstruction to prove resolution of singularities in positive characteristic with the methods of the characteristic zero proof only occurs in very specific situations. In particular, it clearly shows why the purely inseparable case is the the most significant difficulty; it has always attracted in the literature special attention. The theorem allows us to reduce the resolution problem to this case. Even though the description of (5') gives a precise description of the weighted initial form of the defining polynomial, it is not yet clear how to overcome the obstruction of wild singularities systematically (cf. the paragraphs after the example for a couple of possible approaches).

The appearence of wild singularities is due to the failure of maximal contact in positive characteristic. This failure has been studied extensively, especially by Abhyankar, Hironaka, Giraud, Moh, Cossart and Villamayor. The theorem explains for instance, why for local uniformization of surfaces the case of non discrete rational valuations is the hardest one [Ab1, Cu1]. Indeed, to have an infinite number of increases of the shade along a valuation, there must occur infinitely many antelope points in the intersection of two exceptional components, and the successive kangaroo point has to lie outside the transforms of these components.

Example 1. For $m = 2$ and $j = 0$, say $q = (r_1, r_2)$ without silent multiplicities s_i, the inequality of (3) signifies that r_1 and r_2 are not multiples of $c = p^b$ (so that, in particular, the point a lies at the intersection of the two components of $D \cap V$), and that their residues modulo c satisfy $\overline{r_1} + \overline{r_2} \leq c$. By definition of j, the kangaroo point a' lies on the exceptional divisor Y' of the blowup outside its intersection points with the transforms of the two components of $D \cap V$ at a.

The simplest example of an oblique polynomial is $P(y, z) = yz \cdot (y^2 + z^2)$ in characteristic 2. Here, the exponents $r_1 = r_2 = 1$ satisfy inequality (3) of the theorem, $G(y, z) = y^2 + z^2$ has degree $o = 2$ and $P^+(y, z) = P(y, z + y) = y(z + y) \cdot z^2 = yz^3 + y^2z^2$ has p-order $\mathrm{ord}_z^p P^+$ with respect to z equal to $3 > o$ (since y^2z^2 is a square).

Example 2. (Communicated by O. Villamayor, see also [Hi1]) The polynomial $f = x^p + y^p z$ defines a subscheme X of three-dimensional affine space $\mathbb{A}^3$. It is isomorphic to a Cartesian product along the z-axis, since replacing z by $z + t$ with $t \neq 0$ and x by $x + \sqrt[p]{t}$ reproduces f. The coefficient ideal of f at $a = 0$ in the hypersurface V defined by $x = 0$ equals $y^p z$, and its top locus is reduced to the origin. The origin is thus the only permissible center, despite of the Cartesian product structure of X. This pathology shows that the coefficient ideal only carries reliable information at the point in question, but not in a whole neighborhood.

Impact on resolution. It could be hoped that the explicit description of wild singularities from above opens ways to prove resolution in positive characteristic and arbitrary dimension. Some caution is here in order: Statements (5) and (5') only tell us something about the weighted initial form of the defining polynomial (say, in the hypersurface case). Higher order terms are completely overlooked and discarded. But they will come into play later in the resolution process, and their control seems to be out of reach.

The theorem rather suggests that the pair formed by the order and shade of an ideal are just not the right local invariants to capture the complexity of a singularity in positive characteristic. New singularity measures have to be searched. One idea is to consider all derivatives of order $< e$ of the polynomial $y^q \cdot G(y)$ and to extract invariants from the ideal they generate.

Nevertheless, the theorem suggests some new perspectives how to approach resolution. The arithmetic condition (3) of Theorem 1 cannot occur in embedding dimension 2. Therefore, there are no wild singularities in the resolution of plane curve singularities, and hence no characteristic p obstruction [HR]. For surfaces, wild singularities may appear. The increase of the shade can be controlled since it is compensated by stronger drops before or after the blowup. The invariant formed by the pair $(\mathrm{ord}, \mathrm{shade})$ decreases in the long run. Zeillinger observed in [Ze1] that one can define at kangaroo points a correction term, the bonus, which, when subtracted from the shade, saves the induction: The modified shade interpolates monotonously the original shade and drops after each blowup (supposing, of course, that the order has remained constant). The definition of the bonus is quite systematic. It exploits the precise knowledge of wild singularities [HW]. Similar observations had already been made earlier by Abhyankar [Ab1, Cu1], and seem to play a role in Hironaka's recent approach to positive characteristic [Hi1, Hi2].

At the end of this article we give a short argument for the embedded resolution of surfaces of order $c = p^b$, based on Theorem 1. In fact, the characterization of wild singularities implies that between two subsequent kangaroo points of a sequence of local blowups with constant order there is always a point which lies in the intersection of two exceptional components. Then one shows by a simple computation that the occurrence of this point forces the shade to decrease to its half along the sequence, a drop which easily makes up the maximal increase p^{b-1} of the shade at the kangaroo point.

The description of wild singularities suggests to either try to choose a resolution process which never runs into a wild singularity, or to develop more refined measures for the complexity of a singularity. Some options are:

(A) Ensure by suitable earlier blowups that the arithmetic inequality (3) never happens.

(B) Define a semi-local resolution invariant taking into account the location of the kangaroo points and coordinates which are global along the exceptional components.

(C) Correct the occasional increase of the shade by subtracting in specific situations a bonus from it [Ab1, Cu1, HW]. The critical case for this are three-folds.

(D) The observations from the introduction suggest to consider as candidates for possible resolution invariants also the order of polynomials along curves. Indeed, for $P(y, z)$ as at the beginning, being divisible by $(y + z)^o$ can be rephrased by saying that P has order at least equal to o along the curve $y + z = 0$. It is known that in characteristic zero, this order along curves is used implicitly to bound the order at points of the strict transform at those points a' of the exceptional divisor where a translation occurs. The phenomenon of wild singularities in characteristic p turns out to be related to the failure of the upper semicontinuity of the order when poynomials are considered up to the addition of p^b-th powers, cf. [Hi1].

(E) In the resolution process of a hypersurface which is given, say, by a polynomial F of fixed degree, difficulties appear if there is an infinite sequence of wild singularities and kangaroo points along a valuation. By the theorem, each such wild singularity imposes severe restrictions on the coefficients of the original polynomial F (they must satisfy a linear system of equations). It is then tempting to try to show that these restrictions are sufficiently independent so that infinitely many of them cannot be satisfied simultaneously by any non-zero polynomial F.

(F) Try to resolve wild singularities directly by some ad-hoc method.

Comments on the proof. Assertion (1) of Theorem 1 will be seen by using Abhyankar's concept of Tschirnhaus transformation for constructing osculating hypersurfaces as in [EH]. The proof of (2) is a variant of the proof of (3) in a simple special situation and will be omitted.

The key challenge is the uniqueness assertion from (5). It implies most of the other statements of the theorem. In the case where we assume from the beginning that the weighted initial form F of K is a purely inseparable polynomial, the argument is relatively easy and given separately in the section on oblique polynomials. In the general case, the proof of (5) is more involved. It requires a detailed control of the behaviour of weighted homogeneous polynomials under coordinate changes. Even though the reasoning is quite straightforward, there appear some subtle technical complications. We recall here that (5) implies in particular via (5') that we can restrict the study and resolution of wild singularities to purely inseparable polynomials.

A first simplification is achieved by choosing local coordinates at a so that the blowup π, the hypersurfaces of weak maximal contact and the various transforms of the ideals assume a convenient form. The outcome is made explicit in Theorem 2, which, in some sense, represents the algebraic essence of Theorem 1. The increase of the shade is transcribed into an inequality relating the orders of a weighted homogeneous polynomial before and after a weighted homogeneous coordinate change. Assertion (3) of Theorem 2 gives uniqueness of the polynomial without assuming weak maximal contact. The proof uses elementary techniques of convex geometry and linear algebra.

Once the uniqueness assertion of Theorem 2 is ensured, it suffices to search for each choice of parameter values (dimension, degree, relevant and silent multiplicities, ...) a weighted homogeneous polynomial which satisfies inequality (3) of Theorem 2. If such a polynomial exists, it is the unique candidate for providing a wild singularity with these parameter values as described in Theorem 1. Now, if the inequality of (3) of Theorem 1 is not fulfilled, we show that one can take as candidate a polynomial which is c-th power $F(x,y) = (x + A(y))^c$ with A homogeneous of degree w. It will satisfy inequality (3) of Theorem 2 but not inequality (3) of Theorem 1. As the obvious coordinate change in x transforms F into x^c, the hypersurface V defined by $x = 0$ did not have weak maximal contact. This settles (3) of Theorem 1.

A similar reasoning shows that, whenever inequality (3) is satisfied, the weighted initial form must be a purely inseparable polynomial as in (5'). Indeed, it suffices to show that the polynomial from (5') defines a wild singularity as claimed in assertion (6). This is proven in the section on oblique polynomials.

Preliminaries

Setting and concepts. All schemes are of finite type over an algebraically closed field $\underline{k}$ of characteristic $p > 0$. Throughout, W denotes a regular ambient scheme of dimension $m+1$ and a a closed point of W. We write $\mathcal{O}_{W,a}$ for the local ring of W at a, with maximal ideal $m_{W,a}$ and completion $\widehat{\mathcal{O}}_{W,a}$. Regular parameter systems of $\mathcal{O}_{W,a}$ are called local coordinates of W at a; they will be denoted by $(x,y) = (x, y_m, \ldots, y_1)$.

A closed subscheme D of a smooth scheme is a normal crossings scheme if it can be defined at any of its points by a monomial ideal in local coordinates. It is an effective normal crossings divisor if this ideal is locally principal. The exponents of the defining local monomial are called the multiplicities of the components of the divisor. Two closed subschemes D and V of W meet transversally if the subscheme of W defined by the product of their ideals is a normal crossings scheme. If V is smooth, then $D \cap V$ is again a normal crossings subscheme of V. We denote by $M = I_V(D \cap V)$ the monomial ideal in V at a defining the intersection.

A local smooth hypersurface of W at a is a regular, locally closed subscheme V of W of codimension 1 passing through a. It is defined by a regular element of $\mathcal{O}_{W,a}$. Usually, coordinates (x,y) are selected so that V is defined in a neighborhood of a by $x = 0$. We identify elements of $\mathcal{O}_{W,a}$ with their representatives on small open neighborhoods of a in W, and then talk of ideals instead of stalks of ideal sheaves.

Let K be an ideal in W at a point a (not necessarily closed). We denote by $\mathrm{ord}_a(K) = \max\{k,\ K \subset m_{W,a}^k\}$ its order at a. The zero ideal has order ∞. The order of an ideal is upper semicontinuous in a. We denote by $\mathrm{top}(K)$ and call the top locus of K the reduced closed subscheme of W of points of maximal order. Similarly, one may also define the local top locus $\mathrm{top}_a(K)$ of K at a, restricting K to a sufficiently small neighborhood of a. If P is a polynomial in variables $y_m, \ldots, y_1$, we denote by $\mathrm{ord}_y(P)$ the order of P in the localization of the polynomial ring at the ideal generated by $y_m, \ldots, y_1$.

A smooth local hypersurface V at a has weak maximal contact with K at a if it maximizes the order of the coefficient ideal of K in V at a. Such hypersurfaces always exist: Either the coefficient ideal is 0 for some (possibly formal) V, and then its order is ∞, or the orders of the coefficient ideals are bounded for varying V, in which case the maximum is attained by some V. In characteristic zero, it is possible to choose V in addition so that, locally at a, it contains $\mathrm{top}(K)$, and so that V has weak maximal contact to K at all points a_1 of $\mathrm{top}(K)$ sufficiently close to a. This is no longer true in positive characteristic [Na1, Na2, Mu, Gr1, Gr2, Co1, Co2]. Given a point at $a \in \mathrm{top}(K)$, there need not exist a smooth local hypersurface V in W at a having weak maximal contact with K at a such that for a_1 a point of $\mathrm{top}(K)$ sufficiently close to a, V has weak maximal contact with K at a_1. This may not be possible even if V contains $\mathrm{top}(K)$ locally at a. Therefore, we may not be able to choose the same local hypersurface V for the points of small open subsets of $\mathrm{top}(K)$.

Let V have weak maximal contact with K at a and assume given a normal crossings divisor D in W transversal to V and compatible with K and V. The shade of K in V is the order of I at a, where $J_V(K) = M \cdot I$ is the factorization from above, say $\mathrm{shade}_a K = \mathrm{ord}_a(J_V(K)) - \mathrm{ord}_a(M)$ with $M = I_V(D \cap V)$. It is independent of the choice of V as long as D is transversal to V and compatible with K and V. The order of $D \cap V$ in V at a equals, by transversality, the order of D in W at a.

In characteristic zero, the standard resolution invariant is a vector of integers, consisting of the orders of coefficient ideals in decreasing dimensions. This vector is considered with respect to the lexicographic ordering. Its first two components are associated to an ideal K, a local hypersurface V of maximal contact with K at a in the sense of Hironaka [Hi5], and a normal crossings divisor D compatible with K and V; they are defined as the pair $(\mathrm{ord}_a(K), \mathrm{ord}_a(J_V(K)) - \mathrm{ord}_a(I_V(D \cap V)))$. It is known that hypersurfaces which have maximal contact with K also have weak maximal contact [EH], so that the second component of the invariant does not depend on the choice of V and coincides with the shade of K as defined above.

For a smooth closed subscheme Z of W, let $\pi : W' \to W$ denote the blowup of W in Z with exceptional component $Y' = \pi^{-1}(Z)$. Pull-backs of ideals and subschemes are denoted by superscripts *, strict transforms by ^{s}, weak transforms by $^\curlyvee$. Objects in W' assuming the same role as their counterparts in W will be primed $'$, and the respective type of transform will be specified. We suppose that Z is contained in $\mathrm{top}(K)$ and, locally at a, also in V, and that it is transversal to a given normal crossings divisor D. Such blowups will be called permissible. The transforms D^s and D^* are again normal crossings divisors in W' and V^s is smooth. We denote by K' the weak transform $K^\curlyvee = K^* \cdot I_{W'}(Y')^{-\mathrm{ord}_Z(K)}$ of K in W'.

Let $a' \in Y'$ be a point above a. Then $\mathrm{ord}_{a'}(K') \leq \mathrm{ord}_a(K)$ because of $Z \subset \mathrm{top}(K)$. The point a' is called equiconstant for K (or infinitely near to a) if equality holds.

The occurrence of kangaroo points depends on the configuration of the exceptional components at a and a'. The multiplicities of the components of $D \cap V$ at a whose transforms in V^s do not contain a' are the relevant multiplicities of $D \cap V$ at a with respect to a'; they are grouped in an unordered list r, which, clearly, depends on the position of a'. The remaining multiplicities of $D \cap V$ are the silent ones and grouped in the list s.

If V has weak maximal contact with K at a, the transform V^s contains the equiconstant points a' in Y' above a. Moreover, V may be chosen at a so that V^s has again weak maximal contact with K' at a'. In characteristic zero, this property can be achieved along any permissible sequence of local blowups as long as the order of K remains constant: There exists a choice of V whose iterated transforms maintain weak maximal contact with K along any sequence of equiconstant points. One may take V osculating for K at a in the sense of [EH]. In particular, the transforms of V contain all equiconstant points above a. Such hypersurfaces are said to have maximal contact with K at a. Their existence implies that the shade of K does not increase at equiconstant points,

$$(\mathrm{ord}_{a'}(K'), \mathrm{shade}_{a'}(K')) \leq_{lex} (\mathrm{ord}_a(K), \mathrm{shade}_a(K)).$$

This inequality is the starting point of the proof of resolution in characteristic zero by descent in dimension. The inequality is known to fail in characteristic $p > 0$: There exist sequences of permissible blowups for which the transforms of any local smooth hypersurface V at a lose eventually some equiconstant points of K above a. In particular, the transforms of the hypersurface cease to have weak maximal contact with the transforms of K [Na1, Na2, Mu, Mo1, Mo2, Gr1, Gr2, Co1, Co2, Ha3]. It is then necessary to replace occasionally the local hypersurface by a new one so as to ensure again weak maximal contact. It turns out that this change may produce an increase of the shade. This is the phenomenon we propose to understand and describe in the present article.

Let be given K and D in W at a as above. Choose V in W of weak maximal contact with K at a, and assume that D is compatible with K and V. Write $J_V(K) = I_V(D \cap V) \cdot I$ for some ideal I in V. Consider the blowup $\pi : W' \to W$ with center Z transversal to D and contained in $\mathrm{top}(K)$ and $\mathrm{top}(I)$, locally at a. Set $Y' = \pi^{-1}(Z)$. As we are only working locally at points $a \in Z$ and $a' \in Y'$ above a, we may shrink W so that V is closed in W and $Z \subset \mathrm{top}(I)$ holds on whole W. Hence we may assume $Z \subset V$. Set $c = \mathrm{ord}_a(K)$ and $c' = \mathrm{ord}_{a'}(K')$ so that $c' \leq c$. Assume that $a' \in V^s$ is an equiconstant point for K, $c' = c$. Set $o = \mathrm{shade}_a(K) = \mathrm{ord}_a(I)$ and define the transform D' of D in W' as $D' = D^* + (o - c) \cdot Y'$. As Z is transversal to D and $o \geq c$, this is an effective normal crossings divisor and V^s is transversal to D'.

Locally at a', the coefficient ideal $J_{V^s}(K')$ of K' at a' with respect to V^s equals the controlled transform $(J_V(K))^! = (J_V(K))^* \cdot I_{V^s}(Y' \cap V^s)^{-c}$ of $J_V(K)$ with respect to the control c. As $Z \subset \mathrm{top}(I)$ and hence $I \subset I_V(Z)^o$, we get a factorization $J_{V^s}(K') = M^\diamond \cdot I^\curlyvee$ with $M^\diamond = I_{V^s}(D' \cap V^s)$ and $I^\curlyvee$ the weak transform of I in V^s at a'. Hence D' is compatible with K' and V^s.

The strict transform V^s of V need not have weak maximal contact with K' at a'. In this case we may choose a new hypersurface V' in W' at a' which does have weak maximal contact with K' at a'. It can be obtained from V^s by a local formal automorphism of W' at a'. It is easy to see that V' can again be chosen transversal to D'. For such a V', the coefficient ideal $J_{V'}(K')$ factorizes again; but, in contrast to before, $J_{V'}(K') = M' \cdot I^\diamond$ with $M' = I_{V'}(D' \cap V')$ and $I^\diamond$ some ideal in V' at a' which is in general not (isomorphic to) the weak transform $I^\curlyvee$ of I in V^s at a'. In this situation, the shade of K' at a' with respect to D' is $o' = \mathrm{ord}_{a'}(I^\diamond)$. It does not depend on the choice of V and V'. Observe that $\mathrm{shade}_{a'}(K') > \mathrm{ord}_{a'}(I^\curlyvee)$ if V^s does not have weak maximal contact with K'. And, as indicated in the introduction, it may even happen that $\mathrm{shade}_{a'}(K') > \mathrm{shade}_a(K)$.

Coordinate choices. Assume given coordinates $(x, y_m, \ldots, y_1)$ at a so that V is defined by $x = 0$. For K of order c at a, let e be the order of the coefficient ideal $J_V(K)$ of K in V. Assume that $J_V(K)$ is non-zero and set $w = e/c$. The weighted initial forms of K with respect to the coordinates and the weight vector $(w, 1, \ldots, 1)$ are the weighted homogeneous polynomials of order c and weighted degree e which are expansions of elements of K with respect to (x, y) [AHV].

Let $\pi : W' \to W$ be a blowup with smooth center Z transversal to a normal crossings divisor D and contained in a closed smooth hypersurface V transversal to D. Assume that V has weak maximal contact with an ideal K in W at a. Let a' be an equiconstant point of $Y' = \pi^{-1}(Z)$ above a for K. We set $m = \dim(W) - 1$ and $d = \dim(Z)$.

Lemma 1. *There exist local coordinates $(x, y) = (x, y_m, \ldots, y_1)$ of W at a such that*
(1) *a has components $a = (0, \ldots, 0)$ with respect to (x, y).*
(2) *V is defined in W by $x = 0$.*
(3) *Z is defined in W by $x = y_m = \ldots = y_{d+1} = 0$.*
(4) *$D \cap V$ is defined in V locally at a by a monomial $y_m^{q_m} \cdots y_1^{q_1}$, for some $q \in \mathbb{N}^m$.*
(5) *The point a' lies in the y_m-chart of W'. With respect to the local coordinates at the origin of the y_m-chart of W' given by the chart expression of π,*

$$\pi_m : (x, y_m, \ldots, y_1) \to (xy_m, y_m, y_{m-1}y_m, \ldots, y_{d+1}y_m, y_d, \ldots, y_1),$$

it has components

$$a' = (0, 0, a'_{m-1}, \ldots, a'_{j+1}, 0, \ldots, 0)$$

for some $d \le j \le m-1$ and with $a'_{m-1}, \ldots, a'_{j+1} \neq 0$. Here, $j - d$ is the number of components of D whose transforms pass through a'.
(6) *Local coordinates in W' at a' are given by the composition of π_m with the translation $\mu_t : (x, y) \to (x, y + t)$ in W' with $t = (0, t_{m-1}, \ldots, t_{j+1}, 0, \ldots, 0)$ and $t_i = a'_i$. This composition equals the composition of the linear map $\lambda_t : (x, y) \to (x, y + ty_m)$ in W at a with π_m. The map λ_t preserves Z and V.*
(7) *If condition (4) is not imposed, the coordinates $x, y_m, \ldots, y_1$ at a' can be chosen with (1) to (3) and so that a' is the origin of the y_m-chart with local coordinates in W' at a' given by the monomial blowup π_m.*
(8) *The transform V^s of V in W' is given in the induced coordinates at a' by $x = 0$.*

(9) *The decomposition* $q = r + s$ *of the multiplicities of* $D \cap V$ *at* a *in relevant and silent multiplicities with respect to* a' *is given by* $r = (q_m, \dots, q_{j+1}, 0, \dots, 0)$ *and* $s = (0, \dots, 0, q_j, \dots, q_1)$.

Notice that if D is an effective normal crossings divisor so that $I_V(D \cap V)$ is monomial in the coordinates $x, y_m, \dots, y_1$, and if not all t_i are zero (i.e., if $j \leq m-2$), the coordinate change λ_t destroys the monomiality of $I_V(D \cap V)$ as in statement (4).

Proof. All statements are standard. It is clear that $(x, y_m, \dots, y_1)$ can be chosen satisfying (1) to (3), and (4) can be achieved because D and Z are transversal. As for (5), we know by (3) that the exceptional component Y' is covered by the charts corresponding to $x, y_m, \dots, y_{d+1}$. As V has weak maximal contact with K we know that x appears in the initial form of K. As $\mathrm{ord}_{a'}(K') = \mathrm{ord}_a(K)$ it follows that a' cannot lie in the x-chart. Hence a' lies in the other charts and satisfies there $a'_n = 0$. A permutation of $y_m, \dots, y_{d+1}$ allows to assume that a' lies in the y_m-chart. This permutation does not alter (2) and (3). As Y' is given in the y_m-chart by $y_m = 0$ and as $a' \in Y'$ we get $a'_m = 0$. From $a_d = \dots = a_1 = 0$ follows that $a'_d = \dots = a'_1 = 0$. After a permutation of $y_{m-1}, \dots, y_{d+1}$ we may assume that $a'_i \neq 0$ for $m-1 \geq i \geq j+1$ and $a'_i = 0$ for $j \geq i \geq 1$ and $i = m$ with $m-1-j$ the number of non-zero components of a'. This establishes (5). Assertions (7) to (9) follow from (5) by easy computations.

Lemma 2. *Let* K *be an ideal in* W *at* a *of order* c, *and let* V *be a smooth local hypersurface in* W *at* a. *Let be given local coordinates* $x, y_m, \dots, y_1$ *at* a *so that* V *is defined by* $x = 0$. *Assume that at least one of the homogeneous initial forms of degree* c *of elements of* K *involves the coordinate* x.
(1) *There exists a local formal automorphism* φ *in* W *at* a *of the form* $\psi(x, y) = (x + b(y), y)$ *so that* V *has weak maximal contact with* $\psi^*(K)$.
(2) *Let* D *be an effective normal crossings divisor in* W *transversal to* V *at* a *and compatible with* K *and* V. *Then* D *is also compatible with* $\psi^*(K)$ *and* V.

Condition (1) is equivalent to saying that $\psi(V)$ has weak maximal contact with K. The version given in the lemma is more convenient when working in local coordinates. If the homogeneous initial forms of degree c of elements of K involve in all coordinate systems at least two coordinates, the orders of all coefficient ideals of K are equal to c, and any local hypersurface has weak maximal contact with K. In this case, there are no wild singularities.

Proof. (1) follows from the definition of coefficient ideals and the Gauss-Bruhat decomposition of local formal automorphisms with respect to the lexicographic order [Ha1]. (2) is immediate from (1).

Wild singularities

Coefficient ideals. Wild singularities are characterized by the increase of the shade under blowup. In this section, we make this condition more explicit. Apply Lemma 2 to the weak transform K' of the ideal K and the transform V^s of a hypersurface V of weak maximal contact with K under the blowup $W' \to W$ at an equiconstant point $a' \in W'$. As V has weak maximal contact with K at a, we may arrange that it satisfies the assumption in Lemma 2 with respect to the initial forms of K. As the order of K' at a' has remained constant, these properties will then also hold for V^s with respect to K'. Moreover, it follows, as in [Hi4] and [Ha2, proof of Thm. 8.1], that the local automorphism ψ in W' at a' is induced by a local automorphism φ of W at a of the same type. Combining this with the assertions of Lemma 1, the construction of a hypersurface V' of weak maximal contact with K' at a' can be transcribed into the search of a transformation of K by a local automorphism of W at a, together with the translation λ_t as in Lemma 1. Applying these transformations to K, the passage from $\widetilde{K} = \lambda_t^*(\varphi^*(K))$ to $\psi^*(K')$ is given by a monomial substitution of the chosen coordinates. This allows to read off the shade of K' from the Newton polyhedron of $\widetilde{K}$. We specify the details.

Fix coordinates $(x, y_m, \ldots, y_1)$ at a so that V is defined by $x = 0$. Working in the polynomial ring over $\underline{k}$ in these variables, the coefficient ideal $J_x(F)$ of a polynomial F with respect to x is understood as the coefficient ideal $J_V(F)$with respect to the hyperplane V. Similarly, the order of elements of $\mathcal{O}_{W,a}$ will now be expressed as the order of their polynomial expansion with respect to the variables.

Let e be the order of $J_V(K)$ at a. Choose an element in K of order c at a with coefficient ideal in V of order e. Set $w = e/c$ and equip the variables with the weights $(w, 1, \ldots, 1)$. Let F denote the expansion of the weighted initial form of the chosen element at a as a polynomial in (x, y). We may asssume that x^c appears with non-zero coefficient in F, w.l.o.g. equal to 1. Let $x + H(y)$ be the expansion of the weighted initial form of φ so that $F(x + H(y), y)$ is the expansion of an element in the weighted initial form of $\varphi^*(K)$. Set

$$\widetilde{F}(x, y) = F(x + H(y), y + ty_m)$$

with $t = (0, t_{m-1}, \ldots, t_{j+1}, 0, \ldots, 0)$ as in Lemma 1. Observe that V has weak maximal contact with K at a if and only if $F(x + H(y), y)$ is not a c-th power, for all choices of H. The chart expression of the blowup $(W', a') \to (W, a)$ is the composition of $\lambda_t : (x, y) \to (x, y + ty_m)$ with the monomial blowup π of Z in the y_m-chart. Substituting accordingly the variables (x, y) in $\widetilde{F}$ produces a polynomial $\widetilde{F}'$ which has order c in the variables (x, y) and whose coefficient ideal with respect to x has order at least e' in the variables y. Notice that $\widetilde{F}'$ is the expansion of an element of the weighted initial form of $\psi^*(K')$.

As V' maximizes the order e' of $J_{V'}(K')$ in W' at a', the coefficient ideal of $\widetilde{F}'$ with respect to x has maximal order. This coefficient ideal is the controlled transform $(J_x(\widetilde{F}))^!$ under the monomial blowup in the y_m-chart with respect to the control $c = \mathrm{ord}_a(K)$. Therefore e' is bounded by the order of $(J_x(\widetilde{F}))^!$ with respect to y,

$$e' \leq \mathrm{ord}_y((J_x(\widetilde{F})^!).$$

As the blowup π is monomial in the y_m-chart we can interpret the preceding inequality in terms of $\widetilde{F}$. In fact, paper & pencil show that

$$\mathrm{ord}_y((J_x(\widetilde{F}))^!) \leq \mathrm{ord}_z(J_x(\widetilde{F})) + o - c,$$

with $z = (y_{m-1}, \ldots, y_1)$ and $o = \mathrm{shade}_a(K)$. By definition, we have $o' = \mathrm{shade}_{a'}(K') = e' - \mathrm{ord}_{a'}(D') = e' - (o - c) - |s|$ with s the list of silent exponents of D at a with respect to a'. This yields the inequality

$$o' \leq \mathrm{ord}_y((J_x(\widetilde{F}))^!) - \mathrm{ord}_{a'}(D') \leq \mathrm{ord}_z(J_x(\widetilde{F})) - |s|.$$

If $o < o'$ then $\mathrm{ord}_y(J_x(F)) - \mathrm{ord}_a(D) < \mathrm{ord}_z(J_x(\widetilde{F})) - |s|$ follows. Using that $\mathrm{ord}_a D = |r| + |s|$ we have proven

Lemma 3. *Assume that a is a wild singularity of K with kangaroo point a'. Let F and D be as above, let r be the list of relevant multiplicities of D at a with respect to a', and set $\widetilde{F}(x,y) = F(x + H(y), y + ty_m)$ in chosen coordinates $(x,y) = (x, y_m, \dots, y_1) = (x, y_m, z)$. Then*

$$(\mathbb{O}) \qquad \mathrm{ord}_z(J_x(\widetilde{F})) > \mathrm{ord}_y(J_x(F)) - |r|.$$

This inequality will be used to prove Theorem 1. It carries on the exponents of the polynomials F and $\widetilde{F}$, say the associated Newton polyhedra. Using these, the inequality can be made quite explicit. This will be done in the next section. In the case of purely inseparable polynomials $F = x^c + y^q \cdot G(y)$, the inequality simply reads

$$\mathrm{ord}_z^c(G(y + ty_m)) > \mathrm{ord}_y^c(G(y)) + |s|,$$

where, as before, ord_y^c and ord_z^c denote the order of polynomials up to the addition of c-th powers.

Zwickels. This section carries on the location of the exponents of the involved polynomials. Let be given $c \leq e$ in $\mathbb{N}$ and write $cw = e$ with and $w \in \mathbb{Q}$. Set $m = n - 1$. Let $L_c = \{(k, \alpha) \in \mathbb{N}^{1+m},\ k < c\}$ and consider the map

$$L_c \to \mathbb{Q}^m : \ (k, \alpha) \to \tfrac{c}{c-k} \cdot \alpha$$

projecting elements (k, α) of $\mathbb{N}^{1+m}$ to elements of $\mathbb{Q}^m$. The center of the projection is the point $(c, 0, \dots, 0)$. Let $q \in \mathbb{N}^m$ with $|q| = q_1 + \dots + q_m \leq e$ be fixed, and assume given a decomposition $q = r + s$ with $r = (q_m, \dots, q_{j+1}, 0, \dots, 0)$ and $s = (0, \dots, 0, q_j, \dots, q_1)$ for some j between $m - 1$ and $d \geq 0$. Define the upper zwickel $Z(q)$ in $\mathbb{N}^{1+m}$ as the set of points (k, α) with $0 \leq k \leq c$, $wk + |\alpha| = e$ and projection $\frac{c}{c-k} \cdot \alpha \geq_{cp} q$, denoting by $\geq_{cp}$ the componentwise order. Thus $Z(q)$ is given by

$$Z(q) : \ wk + |\alpha| = e \ \text{ and } \ \alpha \geq_{cp} \lceil \tfrac{c-k}{c} \cdot (q_m, \dots, q_1) \rceil.$$

Define the lower zwickel $Y(r,s)$ in $\mathbb{N}^{1+m}$ as the set of points (k, β) in $\mathbb{N}^{1+m}$ with $0 \leq k \leq c$, $wk + |\beta| = e$ and projection $\frac{c}{c-k} \cdot \beta \geq_{cp} (|r|, 0, \dots, 0, s)$. Thus $Y(r,s)$ is given by

$$Y(r,s) : \ wk + |\beta| = e \ \text{ and } \ \beta \geq_{cp} \lceil (\tfrac{c-k}{c} \cdot |r|, 0, \dots, 0, \tfrac{c-k}{c} \cdot q_j, \dots, \tfrac{c-k}{c} \cdot q_1) \rceil.$$

For $j = m - 1$ and hence $r = (q_m, 0, \dots, 0)$ and $s = (0, q_{m-1}, \dots, q_1)$ we have $Z(q) = Y(r,s)$. In general, the two zwickels are different. We claim that for any r and s and $0 \leq k \leq c$ the slice

$$Y(r,s)(k) = \{(k, \beta) \in Y(r,s)\} = Y(r,s) \cap (\{k\} \times \mathbb{N}^m)$$

has at least as many elements as the slice

$$Z(q)(k) = \{(k, \alpha) \in Z(q)\} = Z(q) \cap (\{k\} \times \mathbb{N}^m).$$

This holds for $k = 0$, by definition of $Z(q)$ and $Y(r,s)$. For arbitrary k, the inequality $\lceil \frac{c-k}{c} \cdot |r| \rceil \leq |\lceil \frac{c-k}{c} \cdot r \rceil|$ implies that the condition

$$wk + |\beta| = e \ \text{ and } \ \beta \geq_{cp} (|\lceil \tfrac{c-k}{c} \cdot r \rceil|, 0, \ldots, 0, \lceil \tfrac{c-k}{c} \cdot q_j \rceil, \ldots, \lceil \tfrac{c-k}{c} \cdot q_1 \rceil)$$

is equally or more restrictive than the condition

$$wk + |\beta| = e \ \text{ and } \ \beta \geq_{cp} (\lceil \tfrac{c-k}{c} \cdot |r| \rceil, 0, \ldots, 0, \lceil \tfrac{c-k}{c} \cdot q_j \rceil, \ldots, \lceil \tfrac{c-k}{c} \cdot q_1 \rceil)$$

defining $Y(r,s)(k)$. For each k, the set of pairs k, β satisfying the first condition has as many elements as $Z(q)(k)$ because $|r| + q_j + \ldots + q_1 = |q|$. The claim follows.

Lemma 4. *For given variables $(x,y) = (x, y_m, \ldots, y_1)$, let $F(x,y)$ be a weighted homogeneous polynomial of weighted degree e with respect to $(w, 1, \ldots, 1)$, $w \in \mathbb{N}$. Denote by $J_x(F)$ the coefficient ideal of F with respect to x. Let $\widetilde{F}(x,y) = F(x + H(y), y + ty_m)$ and $q = r + s$ be as in Lemma 3, with H a homogeneous polynomial of degree w and $t = (0, t_{m-1}, \ldots, t_1)$.*
(1) A monomial y^q can be factored from $J_x(F)$ if and only if F has support in $Z(q)$.
(2) If F and $\widetilde{F}$ satisfy the inequality (#) from Lemma 3, the support of $\widetilde{F} - x^c$ lies outside $Y(r,s)$.

Proof. The assertions follow immediately from the definitions.

It will be shown in Theorem 2 that in the situation of (2) of the lemma the inequality $|\overline{r}^c| > (\phi_c(r) - 1) \cdot c$ opposite to the inequality from assertion (3) of Theorem 1 implies that F was already a c-th power, say $F = (x + A(y))^c$ for some homogeneous polynomial A of degree w. If F was the weighted initial form of an element of the ideal K, this signifies that the chosen hypersurface $x = 0$ did not have weak maximal contact with K, contrary to the assumption in Theorem 1. So the presence of a wild singularity forces the inequality $|\overline{r}^c| \leq (\phi_c(r) - 1) \cdot c$. In this case the inequality (#) from Lemma 3 may really occur, as we will see in the section on oblique polynomials.

The arithmetic inequality from (3) of Theorem 1 can be interpreted as an inequality between the number of lattice points in zwickels. Let a c-ray be the segment in $\mathbb{N}^{1+m}$ between the point $(c, 0, \ldots, 0) \in \mathbb{N}^{1+m}$ and a lattice point in $\{0\} \times c \cdot \mathbb{N}^m$. Then inequality (3) is equivalent to saying that the upper zwickel $Z(q)$ contains less c-rays than the lower zwickel $Y(r,s)$. More precisely, with $Z(q)(0) = Z(q) \cap (0 \times \mathbb{N}^m)$ and $Y(r,s)(0) = Y(r,s) \cap (0 \times \mathbb{N}^m)$, we have

Lemma 5. *Let $c \in \mathbb{N}$ and $q = r + s \in \mathbb{N}^m$ with $r = (q_m, \ldots, q_{j+1}, 0, \ldots, 0)$ and $s = (0, \ldots, 0, q_j, \ldots, q_1)$ for some j between 0 and $m-1$. Then the inequality $|\overline{r}^c| \leq (\phi_c(r) - 1) \cdot c$ holds if and only if*

$$|Z(q)(0) \cap c \cdot \mathbb{N}^{1+m}| < |Y(r,s)(0) \cap c \cdot \mathbb{N}^{1+m}|.$$

Proof. To see this, let $(0, c\alpha)$ be a point of $0 \times c \cdot \mathbb{N}^m$. It belongs to $Z(q)(0) \cap c \cdot \mathbb{N}^{1+m}$ if and only if $|c\alpha| = e$ and

$$c\alpha \geq_{cp} \lceil (q_m, \ldots, q_1) \rceil = (\lceil q_m \rceil, \ldots, \lceil q_1 \rceil).$$

As the components of α are integers, the inequality is equivalent to

$$\alpha \geq_{cp} (\lceil \tfrac{q_m}{c} \rceil, \dots, \lceil \tfrac{q_1}{c} \rceil).$$

Conversely, $(0, c\beta)$ in $0 \times \mathbb{N}^m$ belongs to $Y(r,s)(0) \cap c \cdot \mathbb{N}^{1+m}$ if $|c\beta| = e$ and

$$c\beta \geq_{cp} \lceil (|r|, 0, \dots, 0, q_j, \dots, q_1) \rceil = (\lceil |r| \rceil, 0, \dots, 0, \lceil q_j \rceil, \dots, \lceil q_1 \rceil),$$

which can be written as

$$\beta \geq_{cp} (\lceil \tfrac{|r|}{c} \rceil, 0, \dots, 0, \lceil \tfrac{q_j}{c} \rceil, \dots, \lceil \tfrac{q_1}{c} \rceil).$$

The inequality $|\overline{r}^c| > (\phi_c(r) - 1) \cdot c$ is equivalent to the equality $|\lceil \frac{\overline{r}^c}{c} \rceil| = \lceil \frac{|\overline{r}^c|}{c} \rceil$, and hence also to $|\lceil \frac{r}{c} \rceil| = \lceil \frac{|r|}{c} \rceil$. This implies that the second condition on $(0, c\beta)$ can be rewritten as

$$\beta \geq_{cp} (|\lceil \tfrac{r}{c} \rceil|, 0, \dots, 0, \lceil \tfrac{q_j}{c} \rceil, \dots, \lceil \tfrac{q_1}{c} \rceil).$$

The equality

$$|\lceil \tfrac{r}{c} \rceil| = \lceil \tfrac{q_m}{c} \rceil + \ldots + \lceil \tfrac{q_{j+1}}{c} \rceil.$$

then implies that $|\overline{r}^c| > (\phi_c(r) - 1) \cdot c$ if and only if the upper zwickel $Z(q)$ contains as many c-rays as the lower zwickel $Y(r,s)$. This proves the assertion.

Theorem 2

We continue in making the statement of Theorem 1 more explicit. Recall from Lemma 3 the inequality $\mathrm{ord}_z(J_x(\widetilde{F})) > \mathrm{ord}_y(J_x(F)) - |r|$ for the orders of coefficient ideals in the presence of a wild singularity. Here, $F(x,y)$ was a weighted homogeneous polynomial of order c at 0 and weighted degree $e = c/w$ with respect to weights $(w, 1, \dots, 1)$, and not equal to x^c. We had set $\widetilde{F}(x,y) = F(x + H(y), y + ty_m) = F(x + \sum_\gamma h_\gamma y^\gamma, y + ty_m)$, where the sum ranged over $\gamma \in \mathbb{N}^m$ with $|\gamma| = w$ and where $t = (0, t_{m-1}, \dots, t_{j+1}, 0, \dots, 0)$ with non-zero t_i and j the number of silent exponents s_i. Any decomposition $q = (r_m, \dots, r_{j+1}, s_j, \dots, s_1)$ defined induced zwickels $Z(q)$ and $Y(r,s)$ in $\mathbb{N}^{1+m}$. The number of components of r not divisible by c was denoted by $\phi_c(r)$.

Theorem 2. *Let $F(x,y)$ and $\widetilde{F}(x,y) = F(x + H(y), y + ty_m)$ be polynomials as above, with support of F in $Z(q)$.*
(1) Assume that F is not a c-th power. If $w = e/c \notin \mathbb{N}$ or $|\overline{r}^c| > (\phi_c(r) - 1) \cdot c$ or $H = 0$ then

$$\mathrm{ord}_z(J_x(\widetilde{F})) \leq \mathrm{ord}_y(J_x(F)) - |r|.$$

(2) Assume that F is not a c-th power. If $w = e/c \in \mathbb{N}$ and $|\overline{r}^c| \leq (\phi_c(r) - 1) \cdot c$ then

$$\mathrm{ord}_z(J_x(\widetilde{F})) \leq \mathrm{ord}_y(J_x(F)) - |u|,$$

for any $u \in \mathbb{N}^m$ with $u_i \leq r_i$ and $|\overline{u}^c| > (\phi_c(u) - 1) \cdot c$.
(3) For each choice of c, e, r, s, H and t there is at most one F with

$$\mathrm{ord}_z(J_x(\widetilde{F})) > \mathrm{ord}_y(J_x(F)) - |r|.$$

In this case, either F is a c-th power, or the assumptions of (2) hold and F is purely inseparable of the form $x^c + y^q \cdot G(y)$ with $c = p^b$ and G a polynomial of degree $e - |q|$.

This result will imply via Lemma 3 assertions (2) to (7) of Theorem 1, see the respective section below (with the exception of Moh's bound from (7)). The list s of silent exponents can be made empty by auxiliary blowups in suitable centers (defined locally by $x = y_i = 0$ for $1 \le i \le j$), so that $q = r$ becomes the relevant case (some r_i may be equal to 0). The polynomial G in (3) will be described in the section on oblique polynomials. With $P(y) = y^q \cdot G(y)$ and $P^+(y) = P(y + ty_m)$ the inequality of (3) reads $\mathrm{ord}_z^c P^+ > \mathrm{ord}_y^c P - |r|$.

Matrices

Multinomial matrices. For the proof of Theorem 2 we will have to compute the coefficients of $\widetilde{F}$ in terms of the coefficients of F. The dependence will obviously be linear. We are thus lead to consider the respective transformation matrices. These are matrices with certain multinomial entries. We need a preparatory lemma.

Lemma 6. *Let $v \in \mathbb{N}$, $\mu \in \mathbb{N}^l$ and $U = \{\delta \in \mu + \mathbb{N}^l,\ |\delta| \le v\}$. Set $u = |U| = \binom{l+v-|\mu|}{l}$. Let $\theta \in \mathbb{N}^l$ and let $t = (t_l, \dots, t_1)$ be a vector of variables. Then*

$$\det((\tbinom{\gamma+\theta}{\delta} \cdot t^{\gamma+\theta-\delta})_{\gamma,\delta\in U}) = t^\rho$$

with $\rho = u \cdot \theta \in \mathbb{N}^l$ independent of t.

Example. Let $l = 1$, $t = t_1 = 1$ and $\mu \in \mathbb{N}$. Then $U = \{\delta \in \mathbb{N},\ \mu \le \delta \le v\}$ and the matrix has the form

$$\begin{pmatrix} \binom{\gamma+\theta}{\mu} & \binom{\gamma+\theta}{\mu+1} & \dots & \binom{\gamma+\theta}{\mu+v} \\ \binom{\gamma+\theta+1}{\mu} & \binom{\gamma+\theta+1}{\mu+1} & \dots & \binom{\gamma+\theta+1}{\mu+v} \\ \dots & \dots & \dots & \dots \\ \binom{\gamma+\theta+v}{\mu} & \binom{\gamma+\theta+v}{\mu+1} & \dots & \binom{\gamma+\theta+v}{\mu+v} \end{pmatrix}$$

with determinant 1.

Proof. Write A^θ for the $(u \times u)$-square matrix with entries $A^\theta_{\gamma\delta} = \binom{\gamma+\theta}{\delta} \cdot t^{\gamma+\theta-\delta}$. Observe that for $\theta = 0 \in \mathbb{N}^l$ we have $\det A^0 = 1$, since the matrix is upper triangular with 1's on the diagonal. From $\binom{j+1}{i} = \binom{j}{i} + \binom{j}{i-1}$ follows for any $\varepsilon \in \mathbb{N}^l$ with $|\varepsilon| = 1$ that

$$\begin{aligned} A^{\theta+\varepsilon}_{\gamma\delta} &= t^\varepsilon \cdot A^\theta_{\gamma\delta} + A^\theta_{\gamma,\delta-\varepsilon} && \text{if } \delta \in \varepsilon + \mathbb{N}^l, \\ A^{\theta+\varepsilon}_{\gamma\delta} &= t^\varepsilon \cdot A^\theta_{\gamma\delta} && \text{otherwise.} \end{aligned}$$

Therefore the matrix $A^{\theta+\varepsilon}$ is obtained from A^θ by multiplying the columns $A^\theta_{-,\delta}$ by t^ε, for all $\delta \in \varepsilon + \mathbb{N}^l$, and by then adding the column $A^\theta_{-,\delta-\varepsilon}$ to it. The other columns $A^\theta_{-,\delta}$ are only multiplied with t^ε. This implies that

$$\det(A^{\theta+\varepsilon}) = t^{u\cdot\varepsilon} \cdot \det(A^\theta),$$

and induction gives $\det(A^\theta) = t^{u\cdot\theta} \cdot \det(A^0) = t^{u\cdot\theta}$. The lemma is proven.

Transformation matrices. Let $F(x,y)$ and $\widetilde{F}(x,y) = F(x+H(y), y+ty_m)$ be weighted homogeneous polynomials of order c and weighted degree e with respect to a weight vector $(w,1,\ldots,1)$ on $(x,y)=(x,y_m,\ldots,y_1)$, with $w=e/c$. The sum $H(y)=\sum_\gamma h_\gamma y^\gamma$ ranges over $\gamma\in\mathbb{N}^m$ with $|\gamma|=w$, the coefficients h_γ and the components t_i of $t=(0,t_{m-1},\ldots,t_1)$ belong to the ground field. Write

$$F(x,y)=\sum a_{k\alpha}x^k y^\alpha \ \text{ and } \ \widetilde{F}(x,y)=\sum b_{l\beta}(t)x^l y^\beta$$

with $wk+|\alpha| = wl+|\beta| = e$. Fix a decomposition $q=r+s\in\mathbb{N}^m$ with $r=(q_m,\ldots,q_{j+1},0,\ldots,0)$ and $s=(0,\ldots,0,q_j,\ldots,q_1)$ for some index j between $m-1$ and 0. Write elements $\beta\in\mathbb{N}^m$ as (β_m,β^-) where $\beta^-=(\beta_{m-1},\ldots,\beta_1)\in\mathbb{N}^{m-1}$. Let $Y^*(r,s)$ be the subset of $Y(r,s)$ of elements $(k,\beta)\in\mathbb{N}^{1+m}$ given by

$$|\beta^-|\le e-wk-\lceil \tfrac{c-k}{c}\cdot|r|\rceil,$$

$$\beta^- \ge_{cp} \lceil \tfrac{c-k}{c}\cdot(0,\ldots,0,q_j,\ldots,q_1)\rceil.$$

By definition, for each k, the slice $Y^*(r,s)(k)=Y^*(r,s)\cap(\{k\}\times\mathbb{N}^m)$ has the same cardinality as the slice $Z(q)(k)$ of the upper zwickel $Z(q)$. For α and δ in $\mathbb{Z}^m$ set $\binom{\alpha}{\delta}=\prod_i\binom{\alpha_i}{\delta_i}$ where $\binom{\alpha_i}{\delta_i}$ is zero if $\alpha_i<\delta_i$ or $\delta_i<0$. For Γ a subset of $\mathbb{N}^m$, define for $k\in\mathbb{N}$ and $\lambda=(\lambda_\gamma)_{\gamma\in\Gamma}\in\mathbb{N}^\Gamma$ the alternative binomial coefficient

$$[\tbinom{k}{\lambda}]=\prod_{\gamma\in\Gamma}\tbinom{k-|\lambda|^\gamma}{\lambda_\gamma} \ \text{ with } \ |\lambda|^\gamma=\sum_{\varepsilon\in\Gamma,\varepsilon<_{lex}\gamma}\lambda_\varepsilon.$$

Let $\Gamma\subset\mathbb{N}^m$ be the set of $\gamma\in\mathbb{N}^m$ with $|\gamma|=w$ and write $h=(h_\gamma)_{\gamma\in\Gamma}$. Set $\lambda\cdot\Gamma=\sum_{\gamma\in\Gamma}\lambda_\gamma\cdot\gamma\in\mathbb{N}^m$ and fix $t=(0,t_{m-1},\ldots,t_{j+1},0,\ldots,0)$. In this situation we have

Lemma 7. *Let $F(x,y)$ and $\widetilde{F}(x,y)=F(x+H(y),y+ty_m)$ be weighted homogeneous polynomials as above, and assume that x^c has coefficient 1 in both. Fix $q=r+s\in\mathbb{N}^m$ with zwickels $Z(q)$ and $Y^*(r,s)\subset Y(r,s)$.*

(1) *The transformation matrix $A=(A_{k\alpha,l\beta})$ from the coefficients $a_{k\alpha}$ of F to the coefficients $b_{l\beta}$ of $\widetilde{F}$ is given by*

$$A_{k\alpha,l\beta}=\sum_{\lambda\in\mathbb{N}^\Gamma,|\lambda|=k-l}\tbinom{k}{l}[\tbinom{k-l}{\lambda}]\tbinom{\alpha}{\delta_{\alpha\beta\lambda}}\cdot h^\lambda\cdot t^{\alpha-\delta_{\alpha\beta\lambda}},$$

where $\delta_{\alpha\beta\lambda}=(\alpha_m,\beta^- -(\lambda\cdot\Gamma)^-)\in\mathbb{N}^m$ and $h^\lambda=\Pi_\gamma h_\gamma^{\lambda_\gamma}$.

(2) *The quadratic submatrix $A^\square=(A_{k\alpha,l\beta})$ of A with $(k\alpha,l\beta)$ ranging in $Z(q)\times Y^*(r,s)$ has determinant t^ρ where $\rho=\rho(r,s)$ is a vector in $\mathbb{N}^m$ independent of $h=(h_\gamma)_{\gamma\in\Gamma}$ which satisfies $\rho_m=0$ and $\rho_j=\cdots=\rho_1=0$.*

(3) *Assume that F has support in $Z(q)$. If $t_{m-1},\ldots,t_{j+1}$ are non-zero, the coefficients $b_{l\beta}$ of $\widetilde{F}$ in the lower zwickel $Y(r,s)$ determine all coefficients of F. In particular, there is, for each choice of H and t, at most one non-zero polynomial $F(x,y)$ with support in $Z(q)$ such that $\widetilde{F}(x,y)-x^c$ has support outside $Y(r,s)$.*

Proof. Multinomial expansion of $\widetilde{F}(x,y)=F(x+\sum_\gamma h_\gamma y^\gamma, y+ty_m)$ gives for each $k\alpha\in\mathbb{N}^{1+m}$

$$
\begin{aligned}
&(x+\textstyle\sum_{\gamma\in\Gamma} h_\gamma y^\gamma)^k (y+ty_m)^\alpha = \\
&= \textstyle\sum_{l\in\mathbb{N}, l\le k} \binom{k}{l} x^l (\sum_{\gamma\in\Gamma} h_\gamma y^\gamma)^{k-l} \sum_{\delta\in\mathbb{N}^m, \delta\le_{cp}\alpha} \binom{\alpha}{\delta} y^\delta t^{\alpha-\delta} y_m^{|\alpha-\delta|} = \\
&= \textstyle\sum_l \binom{k}{l} x^l \sum_{\lambda\in\mathbb{N}^\Gamma, |\lambda|=k-l} \prod_{\gamma\in\Gamma} \binom{k-l-|\lambda|^\gamma}{\lambda_\gamma} (h_\gamma y^\gamma)^{\lambda_\gamma} \cdot \sum_\delta \binom{\alpha}{\delta} y^\delta t^{\alpha-\delta} y_m^{|\alpha-\delta|} = \\
&= \textstyle\sum_l \sum_\lambda \sum_\delta \binom{k}{l} \prod_\gamma \binom{k-l-|\lambda|^\gamma}{\lambda_\gamma} \binom{\alpha}{\delta} \cdot h^\lambda \cdot t^{\alpha-\delta} \cdot x^l \cdot y^{\Sigma\gamma\lambda_\gamma} \cdot y^\delta \cdot y_m^{|\alpha-\delta|} = \\
&= \textstyle\sum_l \sum_\lambda \sum_\delta \binom{k}{l} [\binom{k-l}{\lambda}] \binom{\alpha}{\delta} \cdot h^\lambda \cdot t^{\alpha-\delta} \cdot x^l \cdot y^{\lambda\cdot\Gamma+\delta} \cdot y_m^{|\alpha-\delta|}.
\end{aligned}
$$

As $\delta_{\alpha\beta\lambda} = ((\delta_{\alpha\beta\lambda})_m, \delta^-_{\alpha\beta\lambda}) = (\alpha_m, \beta^- - (\lambda\cdot\Gamma)^-)$ we can rewrite for given k, α and l a sum $\sum e_{\lambda\delta} \cdot y^{\lambda\cdot\Gamma+\delta} \cdot y_m^{|\alpha-\delta|}$ over $\lambda\in\mathbb{N}^\Gamma$ and $\delta\in\mathbb{N}^m$ with coefficients $e_{\lambda\delta}$ as

$$
\begin{aligned}
&\textstyle\sum_{\lambda\in\mathbb{N}^m, |\lambda|=k-l} \sum_{\delta\le_{cp}\alpha} e_{\lambda\delta} \cdot y^{\lambda\cdot\Gamma+\delta} \cdot y_m^{|\alpha-\delta|} = \\
&= \textstyle\sum_{|\lambda|=k-l} \sum_{\delta\le_{cp}\alpha} e_{\lambda\delta} \cdot (y^-)^{(\lambda\cdot\Gamma)^- + \delta^-} \cdot y_m^{|\alpha-\delta| + (\lambda\cdot\Gamma+\delta)_m} = \\
&= \textstyle\sum_{\beta\in\mathbb{N}^m} \sum_{|\lambda|=k-l} e_{\lambda\delta_{\alpha\beta\lambda}} \cdot (y^-)^{\beta^-} \cdot y_m^{\beta_m} = \\
&= \textstyle\sum_{\beta\in\mathbb{N}^m} (\sum_{|\lambda|=k-l} e_{\lambda\delta_{\alpha\beta\lambda}}) \cdot y^\beta.
\end{aligned}
$$

Here the coefficients $e_{\lambda\delta_{\alpha\beta\lambda}}$ of the last two sums are set equal to zero if $\delta_{\alpha\beta\lambda} \notin \mathbb{N}^m$ or $\alpha \notin \delta_{\alpha\beta\lambda} + \mathbb{N}^m$, say if $\binom{\alpha}{\delta_{\alpha\beta\lambda}} = 0$. Thus

$$
\begin{aligned}
&\widetilde{F}(x,y) = \textstyle\sum_{k\alpha} a_{k\alpha} \cdot (x + \sum_{\gamma\in\Gamma} h_\gamma y^\gamma)^k \cdot (y+ty_m)^\alpha = \\
&= \textstyle\sum_{k\alpha} \sum_{l\beta} \sum_{|\lambda|=k-l} a_{k\alpha} \binom{k}{l} [\binom{k-l}{\lambda}] \binom{\alpha}{\delta_{\alpha\beta\lambda}} \cdot h^\lambda \cdot t^{\alpha-\delta_{\alpha\beta\lambda}} \cdot x^l \cdot y^\beta = \\
&= \textstyle\sum_{l\beta} \beta_{l\beta} \cdot x^l \cdot y^\beta.
\end{aligned}
$$

This gives assertion (1). Observe here that we have used that $a_{c0} = 1$ and $b_{c0} = 1$.

For (2), note that $A_{k\alpha,l\beta} = 0$ if $k < l$. Hence the matrix A is block triangular with blocks $A(k) = (A_{k\alpha,k\beta})_{\alpha\beta}$ on the diagonal $k = l$. By the choice of $Y^*(r,s)$, the induced blocks $A^\square(k)$ of $A^\square$ are square matrices. Therefore $A^\square$ is a square matrix. Assertion (1) yields

$$
A_{k\alpha,k\beta} = \textstyle\sum_{|\lambda|=0} \binom{\alpha}{\delta_{\alpha\beta\lambda}} \cdot h^\lambda \cdot t^{\alpha-\delta_{\alpha\beta\lambda}} = \binom{\alpha}{\delta_{\alpha\beta 0}} \cdot t^{\alpha-\delta_{\alpha\beta 0}} = \binom{\alpha^-}{\delta^-_{\alpha\beta 0}} \cdot t^{\alpha^- - \delta^-_{\alpha\beta 0}},
$$

with $\alpha = (\alpha_m, \alpha^-)$ and $\delta_{\alpha\beta 0} = ((\delta_{\alpha\beta 0})_m, \delta^-_{\alpha\beta 0}) = (\alpha_m, \beta^-)$. Recall that $k\alpha$ and $l\beta$ vary in $Z(q)$ and $Y^*(r,s) \subset Y(r,s)$, respectively, so that

$$
\begin{aligned}
&wk + |\alpha| = e \ \text{ and } \ \alpha \ge_{cp} \lceil \tfrac{c-k}{c} \cdot (q_m, \dots, q_1) \rceil, \\
&wl + |\beta| = e \ \text{ and } \ \beta \ge_{cp} \lceil \tfrac{c-l}{c} \cdot (|r|, 0, \dots, 0, q_j, \dots, q_1) \rceil.
\end{aligned}
$$

Hence, as $k = l$, we have

$$
\begin{aligned}
&|\alpha^-| = e - wk - \alpha_m \ \text{ and } \ \alpha^- \ge_{cp} \lceil \tfrac{c-k}{c} \cdot (q_{m-1}, \dots, q_1) \rceil, \\
&|\delta^-_{\alpha\beta 0}| = |\beta^-| = e - wk - \beta_m \ \text{ and } \ \delta^-_{\alpha\beta 0} \ge_{cp} \lceil \tfrac{c-k}{c} \cdot (0, \dots, 0, q_j, \dots, q_1) \rceil.
\end{aligned}
$$

The determinant of $A^\square(k)$ is given by Lemma 6, taking there $b = (c-k)\cdot w - |\lceil \frac{c-k}{c} \cdot r \rceil|$, $\mu = \lceil \frac{c-k}{c} \cdot (0, \dots, 0, q_j, \dots, q_1) \rceil$ and $\theta = \lceil \frac{c-k}{c} \cdot (q_{m-1}, \dots, q_{j+1}, 0, \dots, 0) \rceil$. Substituting the variables $t_{m-1}, \dots, t_1$ by constants with $t_{m-1}, \dots, t_{j+1} \neq 0$ the determinant is non-zero. We conclude that all $A^\square(k)$ and hence $A^\square$ are invertible. This proves (2).

The uniqueness assertion (3) follows from (2) since the transformation matrix between the $k\alpha$ in $Z(q)$ and the $l\beta$ in $Y(r,s)$ has, as $t_{m-1},\dots,t_{j+1}$ are non-zero, maximal rank equal to the cardinality of $Z(q)$. Note that F is non-zero since x^c has coefficient 1. This establishes Lemma 7.

Proof of Theorem 2

Proof. Assertion (2) holds by replacing in (1) the entries r_i by u_i. We will first show (1) using (3), and then (3).

If $w = e/c \notin \mathbb{N}$, then H is zero, and the proof is similar to the proof in case $w \in \mathbb{N}$ by setting all $h_\gamma = 0$, but without using the arithmetic condition $|\overline{r}^c| > (\phi_c(r) - 1)\cdot c$. So assume that $w \in \mathbb{N}$ and that the inequality holds. We shall construct a weighted homogeneous polynomial F with support in $Z(q)$ which is a c-th power and such that the associated polynomial $\widetilde{F} - x^c$ has support outside $Y(r,s)$. This will only be possible if $Z(q)$ contains sufficiently many c-rays as described in Lemma 5. From the uniqueness of weighted homogeneous polynomials with $\mathrm{ord}_z(J_x(\widetilde{F})) > \mathrm{ord}_y(J_x(F)) - |r|$ by assertion (3) and the description of the inequality in terms of $Y(r,s)$ from Lemma 4 we will then conclude that any F which is not a c-th power must satisfy the opposite inequality $\mathrm{ord}_z(J_x(\widetilde{F})) \leq \mathrm{ord}_y(J_x(F)) - |r|$ as claimed in (1).

Let $l = m-1$. The set T of γ's in $\mathbb{N}^m$ satisfying

$$|\gamma| = w \ \text{ and } \ \gamma \geq_{cp} (\lceil|\tfrac{r}{c}|\rceil, 0, \dots, 0, \lceil\tfrac{q_j}{c}\rceil, \dots, \lceil\tfrac{q_1}{c}\rceil)$$

forms an equilateral l-dimensional simplex in $\Gamma = \{\gamma \in \mathbb{N}^m,\ |\gamma| = w\} \subset \mathbb{N}^m$. Consider its projection T^- in $\mathbb{N}^l$ obtained by omitting the first component γ_m. It consists of elements γ^- in $\mathbb{N}^l$ subject to

$$T^- : \ |\gamma^-| \leq w - \lceil|\tfrac{r}{c}|\rceil \ \text{ and } \ \gamma^- \geq_{cp} (0, \dots, 0, \lceil\tfrac{q_j}{c}\rceil, \dots, \lceil\tfrac{q_1}{c}\rceil).$$

Thus T^- forms an equilateral l-dimensional simplex in $\mathbb{N}^l$ with side length $w - \lceil|\frac{r}{c}|\rceil - |\lceil\frac{s}{c}\rceil|$ and l-dimensional volume $\frac{1}{l!}\cdot(w - \lceil|\frac{r}{c}|\rceil - |\lceil\frac{s}{c}\rceil|)^l$. As $\gamma^- \in \mathbb{N}^l$ determines $\gamma \in \Gamma$ we may write h_{γ^-} for h_γ. Consider the system of equations

$$h_{\gamma^-} = -\sum_{\delta^- \geq_{cp} \gamma^-} \binom{\delta^-}{\gamma^-} t^{\delta^- - \gamma^-} g_{\delta^-}, \quad \gamma^- \in T^-,$$

with unknowns $g_{\delta^-} = g_\delta$ and indices δ^- ranging in the equilateral simplex S^- in $\mathbb{N}^l$,

$$S^- : \ |\delta^-| \leq w - \lceil\tfrac{q_m}{c}\rceil \ \text{ and } \ \delta^- \geq_{cp} (\lceil\tfrac{q_l}{c}\rceil, \dots, \lceil\tfrac{q_1}{c}\rceil).$$

Thus S^- has side length $w - |\lceil\frac{q}{c}\rceil|$ and hence l-dimensional volume $\frac{1}{l!}\cdot(w - |\lceil\frac{q}{c}\rceil|)^l$. The assumption $|\overline{r}^c| > (\phi_c(r) - 1)\cdot c$ is equivalent to

$$|\lceil\tfrac{r}{c}\rceil| \leq \lceil|\tfrac{r}{c}|\rceil,$$

which in turn is equivalent to

$$|\lceil\tfrac{q}{c}\rceil| \leq \lceil|\tfrac{r}{c}|\rceil + |\lceil\tfrac{s}{c}\rceil|.$$

The set T^- equals the set U of Lemma 6, taking $\mu = (0, \dots, 0, \lceil\frac{q_j}{c}\rceil, \dots, \lceil\frac{q_1}{c}\rceil)$ and $v = w - \lceil|\frac{r}{c}|\rceil$. The lemma implies together with $t_l, \dots, t_{j+1} \neq 0$ that the system

$$h_{\gamma^-} = -\sum_{\delta^- \geq_{cp} \gamma^-} \binom{\delta^-}{\gamma^-} t^{\delta^- - \gamma^-} g_{\delta^-}, \ \ \gamma^- \in T^-,$$

admits solutions g_{δ^-} with $\delta^- \in S^-$. Set $F(x,y) = (x + \sum g_\delta y^\delta)^c$ with $\delta = (\delta_m, \delta^-)$ satisfying $\delta^- \in S^-$. This polynomial is a weighted homogeneous c-th power of weighted degree e and with support in $Z(q)$, by definition of S^-. Moreover, as $t_m = 0$,

$$
\begin{aligned}
&F(x, y + ty_m) = (x + \textstyle\sum g_\delta \cdot (y + ty_m)^\delta)^c = \\
&= (x + \textstyle\sum g_\delta \cdot y_m^{\delta_m} \cdot (y^- + t^- y_m)^{\delta^-})^c = \\
&= (x + \textstyle\sum_{\delta^- \in S^-} g_\delta \cdot y_m^{w - |\delta^-|} \cdot \sum_{\gamma \in \Gamma, \gamma^- \leq_{cp} \delta^-} \binom{\delta^-}{\gamma^-} \cdot (y^-)^{\gamma^-} \cdot (t^-)^{\delta^- - \gamma^-} \cdot y_m^{|\delta^- - \gamma^-|})^c = \\
&= (x + \textstyle\sum_{\gamma \in \Gamma} \sum_{\delta^- \in S^-, \delta^- \geq_{cp} \gamma^-} g_\delta \cdot y_m^{w - |\gamma^-|} \cdot \binom{\delta^-}{\gamma^-} \cdot (y^-)^{\gamma^-} \cdot (t^-)^{\delta^- - \gamma^-})^c = \\
&= (x + \textstyle\sum_{\gamma \in \Gamma} y^\gamma \cdot \sum_{\delta^- \in S^-, \delta^- \geq_{cp} \gamma^-} g_\delta \cdot \binom{\delta^-}{\gamma^-} \cdot (t^-)^{\delta^- - \gamma^-})^c = \\
&= (x - \textstyle\sum_{\gamma \in T} y^\gamma \cdot h_\gamma + \sum_{\gamma \in \Gamma \setminus T} y^\gamma \cdot (\ldots))^c
\end{aligned}
$$

with some unspecified sum $(\ldots)$. Observe that if $h_\gamma = 0$ for all $\gamma \in T$, then all $g_\delta = 0$. The equalities imply that

$$
\begin{aligned}
&\widetilde{F}(x,y) = F(x + \textstyle\sum_{\gamma \in \Gamma} h_\gamma y^\gamma, y + ty_m) = \\
&= F(x + \textstyle\sum_{\gamma \in T} h_\gamma y^\gamma, y + ty_m) + R(x,y) = \\
&= x^c + R(x,y),
\end{aligned}
$$

where R is a polynomial with support outside $Y(r,s)$, by definition of T. Thus $\widetilde{F} - x^c$ has support outside $Y(r,s)$. This establishes assertion (1) of Theorem 2 using the uniqueness assertion from (3).

For assertion (3) we need to consider the matrix relating the coefficients of F and $\widetilde{F}$. The statement is now an immediate consequence of Lemma 7, using the characterization of the inequality (#) from Lemma 3 as given in Lemma 4. This concludes the proof of Theorem 2.

Proof of Theorem 1

Proof. Assertion (1) can be seen by using Abhyankar's concept of Tschirnhausen transformation for constructing osculating hypersurfaces as in [EH]. Assume that $c = c' \cdot p^b$ for some $c' > 1$ prime to p. Choose an element f of K of order c whose coefficient ideal in V has order e. There exist local coordinates $(x,y) = (x, y_m, \ldots, y_1)$ so that, after passing to the completion of the local ring of W at a and applying the Weierstrass Preparation Theorem, f is a polynomial in x whose coefficients are series in y. As c' is not divisible by p, we may change the coordinates so that the coefficient of $x^{(c'-1)p^b}$ becomes zero. Then V defined by $x = 0$ has weak maximal contact with K at a, and its transform V' has again weak maximal contact with the weak transform K' of K at all equiconstant points a'. As $J_{V'}(K') = I_{V'}(D' \cap V') \cdot I'$ with I' the weak transform of I and as Z is contained in the locus of points where I has maximal order, we get $\mathrm{shade}_{a'}(K') = \mathrm{ord}_{a'}(I') \leq \mathrm{ord}_a(I) = \mathrm{shade}_a(K)$. Therefore a was not a wild singularity. This proves (1).

The divisibility of the order e by c in assertion (2), the arithmetic inequality of (3) and the last bound in (7) are direct consequences of Theorem 2. Observe here that weak maximal contact implies that F is not a c-th power. Assertion (4) on the silent multiplicities is implied by inequality (1) of Theorem 2 and the description of $D' \cap V'$ in local coordinates as in Lemma 1.

The uniqueness (5) of the weighted initial form follows from the uniqueness assertion (3) in Theorem 2, which is based on the invertibility of the transformation matrix between the coefficients of F in $Z(q)$ and of $\widetilde{F}$ in $Y(r,s)$ given in Lemma 7. The pure inseparability is then a consequence of the description of the weighted initial form in (5'). Assertion (6) follows from the formula (5') by computation.

The second bound in (7) follows from Moh's bound in case $c = p$, using uniqueness and raising the weighted homogeneous polynomial of the wild singularity for $c = p$ to the d-th power. We are left with the proof of the description of oblique polynomials from assertion (5') of Theorem 1.

Oblique polynomials. Recall that a homogeneous polynomial $P(y) = y^q \cdot G(y)$ of degree $e = wc = |q| + o$ is called oblique with respect to c and q if $\mathrm{ord}_z^c P^+ > o$ for $P^+(y) = P(y_m, z + \underline{1} \cdot y_m)$, where $z = (y_{m-1}, \ldots, y_1)$ and $\underline{1} = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{N}^{m-1-j} \times 0^j$. Here, $\mathrm{ord}_z^c P^+$ denotes the order of P^+ with respect to z up to the addition of c-th powers, and the exponent $q = (r_m, \ldots, r_{j+1}, s_j, \ldots, s_1)$ has been decomposed into relevant and silent multiplicities.[4] Instead of $P^+(y) = P(y_m, z + \underline{1} \cdot y_m)$ we could have taken also $P^+(y) = P(y + ty_m)$ for $t = (0, t_{m-1}, \ldots, t_{j+1}, 0, \ldots, 0)$ with non-zero constants t_i. By a suitable homothety in the variables, all t_i can be made equal to 1, so we restrict to this case.

We assume for convenience that no silent multiplicities s_i occur, so that $r = q$. From Theorem 1 we know that c is a pure p-th power, say $c = p^b$ with $b \geq 1$, and that the degree e of $P(y) = y^r \cdot G(y)$ is a multiple wc of c. Let $o = e - |r|$ be the degree of G and set $l = m - 1$. Observe that $\mathrm{ord}_z^c P^+ > o$ implies, by the uniqueness (5) of Theorem 1 and the proof given above of the second bound in (7), that actually $\mathrm{ord}_z^c P^+ \geq o + p^{b-1}$.

Dehomogenizing P with respect to y_m by setting $y_m = 1$ preserves c-th powers. Moreover, as P is homogeneous of degree divisible by c, the dehomogenization creates no new c-th powers in $P(1, z)$. Therefore, the characterization of obliqueness can be transcribed one to one to the dehomogenized situation.

Set $Q(z) = P(1, z) = z^{r^-} \cdot G(1, z)$ with $r^- = (r_l, \ldots, r_1)$ and $G(1, z)$ a polynomial of degree $\leq o$. The dehomogenization of P^+ equals $Q^+(z) = Q(z + \underline{1}) = (z + \underline{1})^{r^-} \cdot G(1, z + \underline{1})$, where $\underline{1}$ now denotes $(1, \ldots, 1) \in \mathbb{N}^l$. The inequality $\mathrm{ord}_z^c(P^+) \geq o + c/p$ signifies that the order with respect to z of Q^+ modulo c-th powers is at least $o + c/p$. Thus P being oblique with respect to c is equivalent to $Q^+ \in \langle z \rangle^{o+c/p}$ modulo c-th powers, where $\langle z \rangle$ denotes the ideal in the polynomial ring $\underline{k}[z]$ generated by $z_l, \ldots, z_1$. Let us write this as

$$(z + \underline{1})^{r^-} \cdot G(1, z + \underline{1}) - N(z^c) \in \langle z \rangle^{o+c/p}$$

for some polynomial N of degree less or equal to $(|r^-| + o)/c$. As $(z + \underline{1})^{r^-}$ has order zero and $G(1, z)$ has order $\leq o$, the polynomial N cannot be zero. We may assume moreover that $N(z^c)$ has degree $< o + c/p$, say, that N has degree $\leq o/c$. Changing N to some N' of degree less or equal to $(|r^-| + o)/c$ modifies $(z + \underline{1})^{r^-} \cdot G(1, z + \underline{1})$ and hence also P only by the addition of c-th powers. It follows that for each choice of c, r and o there is at most one

[4] In [Ha5], p. 18, a slightly different definition of oblique was chosen, discarding the monomial factor.

oblique polynomial P, up to the addition of c-th powers. Using that $(z+\underline{1})^{r^-}$ is invertible in the formal power series ring we get $G(1,z+\underline{1}) = \lfloor (z+\underline{1})^{-r^-}\cdot N(z^c)\rfloor_o$ as required. This proves assertion (5') and completes the proof of Theorem 1.

Remarks. The reasoning above shows directly that there is precisely one oblique polynomial P, up to the addition of p^b-th powers, for each choice of $c=p^b$, r and o satisfying $|\overline{r}^c| \leq (\phi_c(r)-1)\cdot c$ and $|r|+o = wc$, without referring to the proof of the uniqueness statement (5) of Theorem 1 (which is used essentially to reduce to the purely inseparable case).

Assume that $r_i < c = p^b$ for all i and let $(z+\underline{1})^{-1}$ denote the product of all $(z_i+1)^{-1}$. Setting $M(z) = (z+\underline{1})^{-1}\cdot N(z)$ as a formal power series, we get

$$G(1,z+\underline{1}) = \lfloor (z+\underline{1})^{r'}\cdot M(z^c)\rfloor_o,$$

where r' denotes the vector of residues r'_i of $-r_i$ modulo c, say $0\leq r'_i < c$. Inverting the translation $\tau(z) = z+\underline{1}$ we get

$$\begin{aligned} z^{r^-}\cdot G(1,z) &= z^{r^-}\cdot \tau^{-1}\{\lfloor (z+\underline{1})^{-r^-}\cdot N(z^c)\rfloor_o\} \\ &= z^{r^-}\cdot \tau^{-1}\{\lfloor (z+\underline{1})^{r'}\cdot M(z^c)\rfloor_o\}. \end{aligned}$$

The homogenization of this polynomial with respect to y_m followed by the multiplication with $y_m^{r_m}$ then yields the oblique polynomial $P(y) = y^r\cdot G(y)$. It does not depend on the choice of N, up to the addition of c-th powers.

Note that for $c=p$, the monomial y^r is the maximal power which can be factored from an oblique polynomial P. Otherwise we would have some factor $y^{\tilde{r}}$ where the inequality $\tilde{r}_i \geq r_i$ is strict for at least one i. This would cause $P(y) = y^{\tilde{r}}\cdot \widetilde{G}(y)$ with $\widetilde{G}$ of degree $\tilde{o} < o$. Then $\operatorname{ord}_z^p P^+ \leq \tilde{o}+1$ would imply $\operatorname{ord}_z^p P^+ \leq o$ and contradiction to obliqueness. Thus $\Delta = \{\alpha\in\mathbb{N}^m,\ |\alpha| = o\}$ is the smallest equilateral simplex such that the support of P is contained in $r+\Delta$, for some $r\in\mathbb{N}^m$. This property could turn out to be crucial in further studies of the resolution of wild singularities.

Resolution of surfaces

Moh's bound on the increase of the shade and the description of wild singularities from Theorem 1 (actually, only of a small portion of it) allow us to adapt the characteristic zero proof of the embedded resolution of surfaces to the case of arbitary characteristic. We shall only treat hypersurfaces in a smooth three-dimensional ambient variety W and skip technical details. By a classical argument [Za1, Za2, Hi4, Ha2] one can reduce, using the embedded resolution of curves, by a sequence of blowups to the case where the locus of points of maximal order of the defining function of the hypersurface consists of isolated points and smooth curves intersecting transversally. Blowing up these curves, the shade cannot increase, by inequality (3) of Theorem 1. So we are left with point blowups. We have to prove that in any given sequence of point blowups along which the order remains constant the shade cannot increase or remain constant infinitely many times (if there is a curve blowup in between, the argument is similar). To this end we shall observe that between a kangaroo point and the antelope point of its subsequent kangaroo point, the shade drops at least to its half. As the increase between antelope point and kangaroo point is bounded by c/p, our goal follows (the case where the shade is equal to 2 requires a small extra argument).

So let us fix a sequence of point blowups in a three-dimensional ambient space along which the order of the defining equation remains constant. We call oasis point of a given antelope point a_1 the last point a_0 prior to a_1 where none of the two exceptional components through a_1 has existed yet. So the blowup of a_0 creates an exceptional component whose strict transform passes through a_1, and the second component through a_1 has appeared later (recall that a_1 must lie on two exceptional components). It turns out that the shade actually drops at least to its half already between the oasis point and its successive antelope point.

The definition of oasis point implies via Lemma 1 that there exist local coordinates (x, y, z) at a_0 for which the whole sequence of blowups between a_0 and a_1 is given by monomial substitutions of the coordinates. Moreover, the centers of blowups are always contained in the hypersurfaces $x = 0$ (we lift the coordinates at a_0 through all blowups). This allows to work in dimension 2. The claim now follows from the following easy observation, whose proof is left as an amusing exercise.

Lemma 8. *Let $\pi : V_1 \to V_0$ be a sequence of point blowups of a smooth two-dimensional variety V_0. Let $a_0 \in V_0$ be the first center of blowup, and assume that there are coordinates y, z at a_0 which make all blowups monomial. Assume moreover, that a_1 is contained in two exceptional components. Let $G(y, z)$ be a non-zero polynomial in V_0 at a_0, and denote by $G'(y, z)$ its strict transform in V_1 at a_1. Then*

$$\mathrm{ord}_{yz} G' < \tfrac{1}{2} \cdot \mathrm{ord}_{yz} G.$$

The above reasoning seems to be restricted to surfaces, since the strong decrease need not happen in higher dimension. For a more systematic treatment of the embedded resolution of surfaces see [Ab1, Cu1, Hi4, CJS, HW].

References

[Ab1] Abhyankar, S.: Nonsplitting of valuations in extensions of two dimensional regular local domains. Math. Ann. 170 (1967), 87-144.

[Ab2] Abhyankar, S.: Good points of a hypersurface. Adv. Math. 68 (1988), 87-256.

[Ab3] Abhyankar, S.: Desingularization of plane curves. In: Summer Institute on Algebraic Geometry. Arcata 1981. Proc. Symp. Pure Appl. Math. 40 Amer. Math. Soc. 1981.

[Ab4] Abhyankar, S.: Resolution of singularities of embedded algebraic surfaces. Acad. Press 1966. 2nd edition, Springer 1998.

[AHV] Aroca, J.-M., Hironaka, H., Vicente, J.-L.: The theory of the maximal contact. Memorias Mat. Inst. Jorge Juan Madrid 29 (1975).

[AK] Abramovich, D., Karu, K., Matsuki, K., Włodarczyk, J.: Torification and factorization of birational maps. J. Amer. Math. Soc. 15 (2002), 531-572.

[Be] Benito, A.: The τ-invariant and elimination. Preprint 2009.

[BeV] Benito, A., Villamayor, O.: Singularities in positive characteristic: elimination and monoidal transformations. arXiv:math/0811.4148.

[Bl] Blanco, R.: Complexity of Villamayor's algorithm in the monomial case. arXiv:0704.3416v.

[BM] Bierstone, E., Milman, P.: Functoriality in resolution of singularities. Publ. Res. Inst. Math. Sci. 44 (2008), 609-640.

[BMT] Bierstone, E., Milman, P., Temkin, M.: Q-universal desingularization. arXiv:0905.3580.

[BrV] Bravo, A., Villamayor, O.: Singularities in positive characteristic, stratification and simplification of the singular locus. Adv. Math. 224 (2010), 1349-1418.

[Ca] Cano, F.: Reduction of the singularities of codimension one singular foliations in dimension three. Ann. Math. 160 (2004), 907-1011.

[CJS] Cossart, V., Jannsen, U., Saito, S.: Canonical embedded and non-embedded resolution of singularities for excellent two-dimensional schemes. arXiv:math/0905.2191.

[Co1] Cossart, V.: Polyèdre caractéristique d'une singularité. Thèse d'État, Orsay 1987.

[Co2] Cossart, V.: Contact maximal en caractéristique positive et petite multiplicité. Duke Math. J. 63 (1991), 57-64.

[Co3] Cossart, V.: Resolution of singularities of threefolds in positive characteristic. A few obstructions to embedded resolution. Notes from lectures at RIMS, Kyoto, December 2008.

[Co4] Cossart, V.: Modèle projectif régulier et désingularisation. Math. Ann. 293 (1992), 115-122.

[CP1] Cossart, V., Piltant, O.: Resolution of singularities of threefolds in positive characteristic I. J. Algebra 320 (2008), 1051-1082.

[CP2] Cossart, V., Piltant, O.: Resolution of singularities of threefolds in positive characteristic II. J. Algebra 321 (2009), 1836-1976.

[Cu1] Cutkosky, D.: A skeleton key to Abhyankar's proof of embedded resolution of characteristic p surfaces. Preprint 2010.

[Cu2] Cutkosky, D.: Resolution of singularities for 3-folds in positive characteristic, Amer. J. Math. 131 (2009), 59—127.

[Cu3] Cutkosky, D.: Toroidalization of birational morphisms of 3-folds. Memoirs Amer. Math. Soc. 190, 890 (2007).

[Cu4] Cutkosky, D.: Resolution of Singularities. Graduate Studies in Math. 63. Amer. Math. Soc. 2004.

[EH] Encinas, S., Hauser, H.: Strong resolution of singularities in characteristic zero. Comment. Math. Helv. 77 (2002), 421-445.

[EV] Encinas, S., Villamayor, O.: Rees algebras and resolution of singularities. Rev. Mat. Iberoam. Proceedings XVI-Coloquio Latinoamericano de Álgebra 2006.

[FH] Faber, E., Hauser, H.: Today's menu: Geometry and resolution of singular algebraic surfaces. Bull. Amer. Math. Soc. 47 (2010), 373—417.

[FK] Frühbis-Krüger, A.: A short note on Hauser's Kangaroo Phenomena and weak maximal contact in higher dimensions. Preprint 2010.

[Gi1] Giraud, J.: Sur la théorie du contact maximal. Math. Zeitschrift 137 (1974), 285-310.

[Gi2] Giraud, J.: Contact maximal en caractéristique positive. Ann. Scient. Éc. Norm. Sup. Paris 8 (1975), 201-234.

[GT] González-Pérez, P., Teissier, B.: Toric geometry and the Semple-Nash modification. arXiv: 0912.0593v3.

[Ha1] Hauser, H.: Three power series techniques. Proc. London Math. Soc. 88 (2004), 1-24.

[Ha2] Hauser, H.: Excellent surfaces over a field and their taut resolution. In: Resolution of Singularities (Obergurgl 1997). Progress in Math. 181, Birkhäuser 2000.

[Ha3] Hauser, H.: Seventeen obstacles for resolution of singularities. In: The Brieskorn Anniversary volume, Progress in Math. 162, Birkhäuser 1997.

[Ha4] Hauser, H.: Why the characteristic zero proof of resolution fails in positive characteristic. Manuscript 2003.

[Ha5] Hauser, H.: On the problem of resolution of singularities in positive characteristic (Or: A proof we are still waiting for). Bull. Amer. Math. Soc. 47 (2010), 1-30.

[Hi1] Hironaka, H.: Program for resolution of singularities in characteristics $p > 0$. Notes from lectures at the Clay Mathematics Institute, September 2008.

[Hi2] Hironaka, H.: A program for resolution of singularities, in all characteristics $p > 0$ and in all dimensions. Lecture Notes ICTP Trieste, June 2006.

[Hi3] Hironaka, H.: Theory of infinitely near singular points. J. Korean Math. Soc. 40 (2003), 901-920.

[Hi4] Hironaka, H.: Desingularization of excellent surfaces. Notes by B. Bennett at the Conference on Algebraic Geometry, Bowdoin 1967. Reprinted in: Cossart, V., Giraud, J., Orbanz, U.: Resolution of surface singularities. Lecture Notes in Math. 1101, Springer 1984.

[Hi5] Hironaka, H.: Resolution of singularities of an algebraic variety over a field of characteristic zero. Ann. Math. 79 (1964), 109-326.

[HR] Hauser, H., Regensburger, G.: Explizite Auflösung von ebenen Kurvensingularitäten in beliebiger Charakteristik. L'Enseign. Math. 50 (2004), 305-353.

[HW] Hauser, H., Wagner, D.: Alternative invariants for the embedded resolution of surfaces in positive characteristic. Preprint 2009.

[Ja] Jannsen, U.: Resolution of singularities for embedded curves. Appendix to: Saito, S., Sato, K.: A finiteness theorem for zero-cycles over p-adic fields. Annals of Math., to appear.

[Ka] Kawanoue, H.: Toward resolution of singularities over a field of positive characteristic. Part I. Publ. Res. Inst. Math. Sci. 43 (2007), 819-909.

[KM] Kawanoue, H., Matsuki, K.: Toward resolution of singularities over a field of positive characteristic (The Kawanoue program). Part II. Basic invariants associated to the idealistic filtration and their properties. Preprint 2008.

[Md1] Moody, J.: On resolving singularities. J. London Math. Soc. 64 (2001), 548-564.

[Md2] Moody, J.: On resolving vector fields. J. Algebra 189 (1997), 90-100.

[Mo1] Moh, T.-T.: On a stability theorem for local uniformization in characteristic p. Publ. Res. Inst. Math. Sci. 2 (1987), 965-973.

[Mo2] Moh, T.-T.: On a Newton polygon approach to the uniformization of singularities of characteristic p. In: Algebraic Geometry and Singularities. Proc. Conference on Singularities, La Rábida. Birkhäuser 1996.

[MP] McQuillan, M., Panazzolo, D.: Almost étale resolution of foliations. Preprint 2009.

[Mu] Mulay, S.B.: Equimultiplicity and hyperplanarity. Proc. Amer. Math. Soc. 87 (1983), 407-413.

[Na1] Narasimhan, R.: Monomial equimultiple curves in positve characteristic. Proc. Amer. Math. Soc. 89 (1983), 402-413.

[Na2] Narasimhan, R.: Hyperplanarity of the equimultiple locus. Proc. Amer. Math. Soc. 87 (1983), 403-406.

[No] Nobile, A.: Algorithmic equiresolution of possibly non-reduced families of singularities. arXiv:1005.0786.

[Pa] Panazzolo, D.: Resolution of Singularities of Vector Fields in Dimension Three. Acta Math. 197 (2006), 167-289.

[Pi] Piltant, O.: On the Jung method in positive characteristic. Ann. Inst. Fourier 53 (2003), 1237-1258.

[Te] Teissier, B.: Valuations, deformations and toric geometry. In: Valuation Theory and its Applications II (eds. F.V. Kuhlmann, S. Kuhlmann, M. Marshall). Fields Institute Communications.

[Tm1] Temkin, M.: Functorial desingularization of quasi-excellent schemes in characteristic zero: the non-embedded case. arXiv: 0904.1592.

[Tm2] Temkin, M.: Functorial desingularization over Q: boundaries and the embedded case. arXiv: 0912.2570.

[Tm3] Temkin, M.: Absolute desingularization in characteristic zero. arXiv:1001.5433v1.

[Vi1] Villamayor, O.: Hypersurface singularities in positive characteristic, Adv. Math. 213 (2007), 687-733.

[Vi2] Villamayor, O.: Rees algebras on smooth schemes: integral closure and higher differential operators. Rev. Mat. Iberoam. 24 (2008), 213-242.

[Vi3] Villamayor, O.: Elimination with applications to singularities in positive characteristic. Publ. Res. Inst. Math. Sci. 44 (2008), 661-697.

[Ur] Urabe, T.: New ideas for resolution of singularities in arbitrary characteristics. arXiv: 1004.5446.

[Wł1] Włodarczyk, J.: Simple Hironaka resolution in characteristic zero. J. Amer. Math. Soc. 18 (2005), 779-822.

[Wł2] Włodarczyk, J.: Program on resolution of singularities in characteristic p. Notes from lectures at RIMS, Kyoto, December 2008.

[Ya] Yasuda, T.: Higher Nash blowups. Compositio Math. 143 (2007), 1493-1510.

[Za1] Zariski, O.: Reduction of singularities of algebraic three dimensional varieties. Ann. Math. 45 (1944), 472-542.

[Za2] Zariski, O.: A new proof of the total embedded resolution theorem for algebraic surfaces. Amer. J. Math. 100 (1978), 411-442.
[Ze1] Zeillinger, D.: Polyederspiele und Auflösen von Singularitäten. Ph.D. Thesis, University of Innsbruck, 2005.
[Ze2] Zeillinger, D.: A short solution to Hironaka's polyhedral game. Enseign. Math. 52 (2006), 143-158.

Faculty of Mathematics
University of Vienna, Austria
herwig.hauser@univie.ac.at