Special 2-flags in lengths not exceeding four: a study in strong nilpotency of distributions

Piotr Mormul* and Fernand Pelletier

Abstract

In the recent years, a number of issues concerning distributions generating 1-flags (called also Goursat flags) has been analyzed. Presently similar questions are discussed as regards distributions generating *multi*-flags. (In fact, only so-called *special* multi-flags, to avoid functional moduli.) In particular and foremost, special 2-flags of small lengths are a natural ground for the search of generalizations of theorems established earlier for Goursat objects. In the present paper we locally classify, in both C^ω and C^∞ categories, special 2-flags of lengths not exceeding four. We use for that the known facts about special multi-flags along with fairly recent notions like *strong nilpotency* of distributions. In length four there are already 34 orbits, the number to be confronted with only 14 singularity classes – basic invariant sets discovered in 2003.

As a common denominator for different parts of the paper, there could serve the fact that only rarely multi-flags' germs are strongly nilpotent, whereas all of them are weakly nilpotent, or nilpotentizable (possessing a local nilpotent basis of sections).

1 Definition of special k-flags and their singularities

Special k-flags (the natural parameter $k \geq 2$ is sometimes called 'width') of lengths $r \geq 1$ can be defined in several equivalent ways, like in [KRub], [PaR], [M2]. All these approaches can be reduced to one transparent definition. (The reduction is via two early Bryant's results from [B], one lemma from [PaR], and the answer to a recent question of Zhitomirskii, cf. p. 165 in [M2].)

Namely, for a distribution D on a manifold M, the tower of consecutive Lie squares of D

$$D = D^r \subset D^{r-1} \subset D^{r-2} \subset \cdots \subset D^1 \subset D^0 = TM$$

(that is, $[D^j,\, D^j] = D^{j-1}$ for $j = r,\, r-1, \dots, 2,\, 1$) should consist of distributions of ranks, starting from the smallest object D^r: $k+1,\, 2k+1,\, \dots,\, rk+1,\, (r+1)k+1 = \dim M$ such that

* Supported by Polish KBN grant 2 P03A 010 22.

• for $j = 1, \ldots, r-1$ the Cauchy-characteristic module $L(D^j)$ of D^j sits already in the smaller object D^{j+1}: $L(D^j) \subset D^{j+1}$ **and** is regular of corank 1 in D^{j+1}, while $L(D^r) = 0$;

•• the covariant subdistribution F of D^1 (see [KRub], p. 5 for the definition extending the classical Cartan's approach from [C], p. 121) exists and is involutive. Note that, in view of Lemma 1 in [KRub], such an F is **automatically** of corank 1 in D^1; the hypotheses in that lemma are satisfied as $\text{rk}\,[D^1,\, D^1]/D^1 \;=\; k > 1$.[1]

Attention. Recently new works [Ad] and [SY] have appeared, revisiting, among other subjects, the very definition of special multi-flags. In the light of those works, the extensive condition • in the definition above is *redundant*. This condition follows from •• and the property of regular dimension growth of the flag of consecutive Lie squares of the initial distribution D, so-called the *big flag* of D. Things being so, the entire theory of special multi-flags starts to appear more compact (the more compact the better).

Note also that different properties of Cartan's original object discussed in [C] (treated nowadays in the genericness' context as a local module of vector fields) were grouped together in [MPe1].

Special multi-flags, and in particular special 2-flags, appear, from the one side, to be rich in singularities, and from the other – to possess finite-parameter families of pseudo-normal forms, with no functional moduli. It is natural, then, to search for *precise* normal forms for them, at least in small lengths. Realizing well that, from certain length onwards, some parameters may prove genuine *moduli*, as we, besides, rigorously exemplify in Section 1.5. (Our example is in length $r = 7$ and works, in fact, in all widths $k \geq 2$, not only for $k = 2$. It is likely that moduli of special multi-flags exist already in length six. Moreover, the length of the onset of moduli may decreasingly depend on flags' width.)

In the parallel framework of 1-flags (most often called Goursat or Cartan-Goursat) similar questions have led to lists of exact local models in lengths not exceeding seven and to the discovery of real moduli in lengths from eight on. A distinctive feature of Goursat flags is that for them the property • comes in automatically and that there is *plenty* of involutive corank one subdistributions of D^1 (cf. ••), although none of them is canonical, while the covariant subdistribution of D^1 is simply $L(D^1)$. The key tool for 1-flags, sufficient up to length six, has been the Jean stratification [J] describing in geometric terms (if only implicitly) the sequences of consecutive singularities showing up in 1-flags.

For special 2-flags, it is not possible to follow that way too closely, although one natural stratification, into *singularity classes*, exists ([M3, M4]). It does not, however, correspond to Jean's one, but rather to a much *coarser* stratification of Goursat objects

[1] Equivalently, using Tanaka's and Yamaguchi's approach [T, Y] (well anterior to [KRub] and not designed for special flags, although applicable for them), one stipulates in •• two things: — the distribution $\widehat{D}^1 = D^1/L(D^1)$ of rank $k+1$ on a manifold of dimension $2k+1$ is of type $\mathfrak{C}^1(1,\, k)$ of [T] and, as such, possesses its *symbol* subdistribution $\widehat{F} \subset \widehat{D}^1$ ([Y], p. 30) and — $\widehat{F}$ is involutive (cf. Prop. 1.5 in [Y]). F is then the counterimage of $\widehat{F}$ under the factoring out by $L(D^1)$. Thus, for special k-flags, $k \geq 2$, the stipulated involutive corank one subdistribution of D^1 is at the same time: the covariant subdistribution in the Cartan-Kumpera-Rubin sense *and* symbol subdistribution in the Tanaka-Yamaguchi sense.

into *Kumpera-Ruiz classes* – cf. [MonZ], p. 466. Jean-like singularities in special 2-flags (and all the more so for $k > 2$) seem to be incredibly rich and escaping any reasonable ordering. Places resembling his approach *can* be detected in the present work: we often distinguish between transverse and tangent, but it is worlds apart from the regular ternary tree of 'basic geometries' of 'car + trailers' systems.

Putting things simply, in 2-flags there is much more room for singular positions than in 1-flags. Already in length three the singularity classes evoked above fail to fully describe the orbits of the local classification; one of them splits up into three orbits. In addition, a fairly new notion of *strong nilpotency* ([M1]) appears to be useful. It allows to completely describe all orbits in lengths 3 and 4, and can be perceived as a key notion of the present work. With its use we show – this is our main result – that there are 34 orbits of the local classification (in both smooth and analytic category) of special 2-flags of length four, as contrasted with only 14 singularity classes in that length. In this way the length four appears to be 'discrete' yet, with no moduli whatsoever.

It is to be underlined at this place that the local classification of special k-flags appears to be *stable* with respect to the width $k \geq 2$ for lengths $r \leq 3$, but already not, for various (not all predictable) reasons, for $r = 4$. Compare in this respect Remark 4 and Section 7.7.

1.1 Sandwich Diagram for multi-flags.

All these requirements merge naturally into a *sandwich diagram*.[2] Note that the inclusions $L(D^{j-1}) \supset L(D^j)$ in its lower line are due to the Jacobi identity.

$$\begin{array}{ccccccccccccc} TM = D^0 & \supset & D^1 & \supset & D^2 & \supset & D^3 & \cdots & D^{r-1} & \supset & D^r & & \\ & & \cup & & \cup & & \cup & & \cup & & \cup & & \\ & & F & \supset & L(D^1) & \supset & L(D^2) & \cdots & L(D^{r-2}) & \supset & L(D^{r-1}) & \supset & L(D^r) = 0\,. \end{array}$$

As for the inclusion $F \supset L(D^1)$, it follows from [KRub] and, besides, is a part of the answer to the question mentioned in the previous paragraph. All vertical inclusions in this diagram are of codimension one, while all (drawn, we do not mean superpositions of them) horizontal inclusions are of codimension k. The squares built by these inclusions can, indeed, be perceived as certain 'sandwiches'. For instance, in the utmost left sandwich F and D^2 are as if fillings, while D^1 and $L(D^1)$ constitute the covers (of dimensions differing by $k + 1$, one has to admit). At that, the sum $k + 1$ of **co**dimensions, in D^1, of F and D^2 equals the dimension of the quotient space $D^1/L(D^1)$, so that it is natural to ask how the k-dimensional space $F/L(D^1)$ and the line $D^2/L(D^1)$ are mutually positioned in $D^1/L(D^1)$. Similar questions impose by themselves in further sandwiches 'indexed' by the upper right vertices D^3, D^4, …, D^r.

[2] so called after a similar (if not identical) diagram assembled for Goursat distributions, or 1-flags, in [MonZ]

1.2 Analogues for special multi-flags of Kumpera-Ruiz classes.

We first divide all existing germs of special k-flags of length r into 2^{r-1} pairwise disjoint *sandwich classes* in function of the geometry of the distinguished spaces in the sandwiches (at the reference point for a germ), and label those groups by words of length r over the alphabet $\{1, \underline{2}\}$ starting (on the left) with 1, having the second cipher $\underline{2}$ iff $D^2(p) \subset F(p)$, and for $3 \le j \le r$ having the j-th cipher $\underline{2}$ iff $D^j(p) \subset L(D^{j-2})(p)$.
This construction puts in relief possible non-transverse situations in the sandwiches. For instance, the second cipher is $\underline{2}$ iff the line $D^2(p)/L(D^1)(p)$ is not transverse, in the space $D^1(p)/L(D^1)$, to the codimension one subspace $F(p)/L(D^1)(p)$, and similarly in further sandwiches. This resembles the Kumpera-Ruiz classes of Goursat germs constructed in [MonZ]. In length r the number of sandwiches has then been $r-2$ (and so the # of KR classes 2^{r-2}). For multi-flags this number is $r-1$ because the covariant distribution of D^1 comes into play and gives rise to one additional sandwich.

How can one ascertain if such virtually created sandwich classes really materialize, and, if so, how to possibly sort them further? In the present paper we restrict ourselves to $k = 2$, whereas the general construction (in the framework of multi-dimensional Cartan prolongations) is given in [M2]. We will produce a huge variety of polynomial germs at $0 \in \mathbb{R}^N$, N possibly very large (odd), of rank-3 distributions. Often – this is important – certain variables x_j will appear in them in a shifted form $b + x_j$, and it will always be an issue if such shifting constants are rigid or flexible, subject to further simplifications. More precisely, for each $m \in \{1,\, 2,\, 3\}$ we are going to define an operation $\mathbf{m}$ producing new rank-3 distributions from previous ones. Technically, its outcome (indices of new incoming variables) will also depend on how many operations have been done *before* $\mathbf{m}$.

More specifically, the outcome of $\mathbf{m}$ – being performed as operation number l – on a distribution $(Z_1,\, Z_2,\, Z_3)$ defined in the vicinity of $0 \in \mathbb{R}^s(u_1, \ldots, u_s)$, is the germ at $0 \in \mathbb{R}^{s+2}(u_1, \ldots, u_s,\, x_{l+1},\, y_{l+1})$ of a new rank-3 distribution generated by

$$Z_1' = \begin{cases} Z_1 + (b_{l+1} + x_{l+1})Z_2 + (c_{l+1} + y_{l+1})Z_3\,, & \text{when } \mathbf{m} = \mathbf{1}, \\ x_{l+1}Z_1 + Z_2 + (c_{l+1} + y_{l+1})Z_3\,, & \text{when } \mathbf{m} = \mathbf{2}, \\ x_{l+1}Z_1 + y_{l+1}Z_2 + Z_3\,, & \text{when } \mathbf{m} = \mathbf{3} \end{cases}$$

and $Z_2' = \frac{\partial}{\partial x_{l+1}}$, $Z_3' = \frac{\partial}{\partial y_{l+1}}$; b and/or c are certain constants (depending on the germ under consideration). For any possible next such operation (and one is bound to perform many of them) it is important that these local generators are written precisely in this order, yielding together a new 'longer' or more involved distribution $(Z_1',\, Z_2',\, Z_3')$. Note that two operations $\mathbf{1}$ and $\mathbf{2}$, out of three typically available, bring in new numerical parameters (adding to possibly already existing previous parameters).

Extended K–R pseudo-normal forms (EKR for short), of length $r \ge 1$, denoted by $\mathbf{j}_1.\mathbf{j}_2 \ldots \mathbf{j}_r$, where $j_1, \ldots, j_r \in \{1,\, 2,\, 3\}$ and depending on numerous numerical parameters within a fixed symbol $\mathbf{j}_1.\mathbf{j}_2 \ldots \mathbf{j}_r$, are defined inductively, starting from the empty label distribution

$$\Big(\frac{\partial}{\partial t},\, \frac{\partial}{\partial x_1},\, \frac{\partial}{\partial y_1}\Big)$$

understood in the vicinity of $0 \in \mathbb{R}^3(t,\, x_1,\, y_1)$. Then, assuming the family of pseudo-normal forms $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}$ already constructed and written in coordinates that go along with the operations: first $\mathbf{j}_1$, then $\mathbf{j}_2$ and so on up to $\mathbf{j}_{r-1}$, the normal forms subsumed under the symbol $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}.\mathbf{j}_r$ are the outcome of the operation $\mathbf{j}_r$ performed as the operation number r over those coordinately written distributions $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}$.

For a moment, it is nearly directly visible that every EKR is a special 2-flag of length equal to the number of operations used to produce it. In particular, it is easy to predict what are, for any EKR of length r, the involutive subdistributions of ranks $2,\, 4, \ldots,\, 2r$; see Observation 1 below. The point is that locally the converse is also true, and one has

Theorem 1 ([M2]). *Let a rank-3 distribution D generate a special 2-flag of length $r \geq 1$ on a manifold M^{2r+3}. For every point $p \in M$, D in a neighbourhood of p is equivalent, by a local diffeomorphism that sends p to 0, to a certain* EKR $\mathbf{j}_1.\mathbf{j}_2 \ldots \mathbf{j}_r$ *in a neighbourhood of $0 \in \mathbb{R}^{2r+3}$. Moreover, that* EKR *can be taken such that $\mathbf{j}_1 = \mathbf{1}$ and the first letter* **2**, *if any, appears before the letters* **3**.

The restriction on the EKR codes mentioned in this theorem is called, after [M2], the rule of the least upward jumps: after the starting **1**, and possibly several more **1**'s, there must first appear a **2** and only later a **3**, if any. Note also that possible constants in the EKR's representing a given germ D are not, in general, defined uniquely, as shows already Example 1. For $r \leq 4$ this, in all EKRs, is duly analyzed in the present contribution, and conclusions differ sometimes from natural expectations.

Example 1. The EKR $\mathbf{1.1} \ldots \mathbf{1}$ (r ciphers **1**) subsumes a vast fan of different pseudo-normal forms – germs at $0 \in \mathbb{R}^{2r+3}$ parametrized by real parameters $b_2,\, c_2, \ldots,\, b_{r+1},\, c_{r+1}$. Under a closer inspection (Theorem 1 in [KRub]), they all are pairwise equivalent, and are equivalent to the classical *Cartan distribution* (or *jet bundle* in the terminology of [Y]) on the space $J^r(1,\, 2)$ of the r-jets of functions $\mathbb{R} \to \mathbb{R}^2$, given by the Pfaffian equations

$$dx_j - x_{j+1}dt \;=\; 0 \;=\; dy_j - y_{j+1}dt\,, \qquad j = 1,\, 2, \ldots,\, r\,.$$

All other EKRs are not equivalent to the jet bundles, as is explained in Proposition 1 below. It is to be noted that the question of a geometric characterization of Cartan distributions was addressed in many works and, in full generality (for all jet spaces $J^r(m,\, k)$), was answered in [Y].

1.3 The EKR's versus sandwich classes.

What relationship exists between the sandwich class of a given germ of a special 2-flag and its all possible EKR presentations? A key tool for answering this question is

Observation 1. *If a distribution $D = D^r$ generating a special 2-flag of length $r \geq 1$ is presented in any* EKR *form on $\mathbb{R}^{2r+3}(t,\, x_1,\, y_1, ...,\, x_{r+1},\, y_{r+1})$, then the members of the associated subflag in the sandwich diagram for D^r are canonically positioned as follows.*

- $F = \big(\partial/\partial x_2,\, \partial/\partial y_2,\, \partial/\partial x_3,\, \partial/\partial y_3,\, \ldots,\, \partial/\partial x_{r+1},\, \partial/\partial y_{r+1}\big)$,

- $L(D^j) = \big(\partial/\partial x_{j+2},\ \partial/\partial y_{j+2},\ \dots,\ \partial/\partial x_{r+1},\ \partial/\partial y_{r+1}\big)$ *for* $j \le r-1$,
- $L(D^r) = (0)$.

These extremely simplified descriptions are the analogues of similar ones for Goursat flags when viewed in Kumpera-Ruiz coordinates. Another analogue (a derivative product of Observation 1) is

Proposition 1. *Assume a germ, D, of a special* 2*-flag of length r sits in a sandwich class having the label $\mathcal{E}$. Then, for any* EKR $\mathbf{j}_1 \dots \mathbf{j}_{r-1}.\mathbf{j}_r$ *for D, $\mathbf{j}_l = \mathbf{1}$ iff the l-th cipher in $\mathcal{E}$ is* 1*.*

Therefore, the singular phenomena – pointwise inclusions in sandwiches do narrow (to **2** and **3**) the pool of operations available at the relevant steps of producing EKR visualisations for special 2-flags.

Proof. $\mathbf{j}_1$ is by default **1** and the first cipher in $\mathcal{E}$ is by definition 1. Consider now $\mathbf{j}_l$, $l \ge 2$, and recall that the operation $\mathbf{j}_l$ transforms certain EKR $(Z_1,\, Z_2,\, Z_3)$ into an EKR $(Z_1',\, Z_2',\, Z_3')$. When $\mathbf{j}_l$ is either **2** or **3**, then, by definition of these operations, $Z_1' \equiv x_{l+1}Z_1 \mod (Z_2,\, Z_3)$, where $Z_2 = \frac{\partial}{\partial x_l}$ and $Z_3 = \frac{\partial}{\partial y_l}$. (As for $Z_2' = \frac{\partial}{\partial x_{l+1}}$ and $Z_3' = \frac{\partial}{\partial y_{l+1}}$, they cause no trouble in the discussion.) Whereas for $\mathbf{j}_l = \mathbf{1}$, $Z_1' \equiv Z_1 \mod (Z_2,\, Z_3)$ and a non-zero vector $Z_1(0)$ is, by its recursive construction (in $l-1$ steps), spanned by

$$\partial/\partial t,\ \ \partial/\partial x_1,\ \ \partial/\partial y_1\,,\ \dots,\ \partial/\partial x_{l-1},\ \ \partial/\partial y_{l-1}\,.$$

Hence, in view of Observation 1, $Z_1(0)$ does not sit in: $F(0)$ when $l=2$, and $L(D^{l-2})(0)$ when $l > 2$. $\square$

Remark 1. When $k = 1$, two operations instead of three $(\mathbf{1}, \mathbf{2}, \mathbf{3})$ in the present text, lead to the well-known local Kumpera-Ruiz pseudo-normal forms for Goursat flags.

1.4 Singularity classes of special 2-flags refining the sandwich classes.

We refine further the singularities of special 2-flags and recall from [M3] how one passes from the sandwich classes to *singularity classes*. In fact, to any germ $\mathcal{F}$ of a special 2-flag associated is a word $\mathcal{W}(\mathcal{F})$ over $\{1, 2, 3\}$, called 'singularity class' of $\mathcal{F}$. It is a specification of the word 'sandwich class' for $\mathcal{F}$ (this last being over, reiterating, $\{1, \underline{2}\}$) with the letters $\underline{2}$ replaced either by 2 or 3, in function of the geometry of $\mathcal{F}$.

In the definition that follows we keep fixed the germ of a rank-3 distribution D at $p \in M$, generating on M a special 2-flag $\mathcal{F}$ of length r.

Suppose that in the sandwich class $\mathcal{C}$ of D at p there appears somewhere, for the first time when going from the left, the letter $\underline{2} = j_m$ (j_m is, as we know, not the first letter in $\mathcal{C}$) **and** that there are in $\mathcal{C}$ other letters $\underline{2} = j_s$, $m < s$, as well. We will specify each such j_s to one of the two: 2 or 3. (The specification of the first j_m will be made later and will be easy.) Let the nearest $\underline{2}$ standing to the left to j_s be $\underline{2} = j_t$, $m \le t < s$. These two 'neighbouring' letters $\underline{2}$ are separated in $\mathcal{C}$ by $l = s - t - 1 \ge 0$ letters 1.

The gist of the construction consists in taking the *small flag* of precisely original flag's member D^s,

$$D^s = V_1 \subset V_2 \subset V_3 \subset V_4 \subset V_5 \subset \cdots ,$$

$V_{i+1} = V_i + [D^s,\, V_i]$, then focusing precisely on this new flag's member V_{2l+3}. Reiterating, in the t-th sandwich, there holds the inclusion: $F(p) \supset D^2(p)$ when $t = 2$, or else $L(D^{t-2})(p) \supset D^t(p)$ when $t > 2$. This serves as a preparation to an important point.

Surprisingly perhaps, specifying j_s to 3 goes via replacing D^t by V_{2l+3} in the relevant sandwich inclusion at the reference point. That is to say, $j_s = \underline{2}$ is being specified to 3 iff $F(p) \supset V_{2l+3}(p)$ (when $t = 2$) or else $L(D^{t-2})(p) \supset V_{2l+3}(p)$ (when $t > 2$) holds.

In this way all non-first letters $\underline{2}$ in $\mathcal{C}$ are, one independently of another, specified to 2 or 3. Having that done, one simply replaces the first letter $\underline{2}$ by 2, and altogether obtains a word over $\{1,\, 2,\, 3\}$. It is the singularity class $\mathcal{W}(\mathcal{F})$ of $\mathcal{F}$ at p. Thus created $\mathcal{W}(\mathcal{F})$ clearly satisfies the least upward jumps rule.

Example 2. In length 4 there exist the following fourteen singularity classes: 1.1.1.1, 1.1.1.2; 1.1.2.1, 1.1.2.2, 1.1.2.3; 1.2.1.1, 1.2.1.2, 1.2.1.3, 1.2.2.1, 1.2.2.2, 1.2.2.3, 1.2.3.1, 1.2.3.2, 1.2.3.3.

Do singularity classes surge to surface in the mentioned local polynomial pseudo-normal forms EKR, as the sandwich classes have done? Yes, the EKR's are faithful to the underlying local flag's geometry epitomized in the singularity class, and there holds

Theorem 2 ([M3, M4]). *For every germ D of a rank-3 distribution generating a special 2-flag of length $r \ge 1$, and for every its pseudo-normal form of the type $\mathbf{j}_1.\mathbf{j}_2 \dots \mathbf{j}_r$ (subject to the least upward jumps rule), the word $j_1.\, j_2 \dots j_r$ is but $\mathcal{W}(D)$.*
In particular, the singularity class of any EKR *form $\mathbf{j}_1.\, \mathbf{j}_2 \dots \mathbf{j}_r$, regardless of its constants, is $j_1.\, j_2 \dots j_r$.*

This theorem shows additionally that all defined singularity classes are non-empty. How many singularity classes do there exist for special 2-flags, and of what codimensions are they?

On each manifold M of dimension $2r + 3$ bearing a special 2-flag of length r, the shadows of singularity classes (one says also about *materializations* of singularities) form always – and not only for 'generic' flags! – a very neat stratification by embedded submanifolds whose codimensions are directly computable. Namely,

Proposition 2. *The codimension of the materialization of any fixed singularity class $\mathcal{C}$ is equal, provided the materialization is non-empty, to*

$$\text{the number of letters 2 in } \mathcal{C} + \text{twice the number of letters 3 in } \mathcal{C}\,.$$

Once Theorem 2 shown, one proves this statement locally, using *any* fixed EKR depicting locally the flag in question.

The number of different singularity classes of special 2-flags of length $r \geq 3$ is

$$2 + 3 + 3^2 + \cdots + 3^{r-2} \;=\; \frac{1 + 3^{r-1}}{2}\,. \tag{1}$$

(One focuses attention on the position of the first letter 2 in the class' code, remembering that the codes satisfy the least upward jumps rule: no letter 2 or else that letter at the very end – account for the summand 2, that letter at the one before last position accounts for the summand 3, and so on. Then that letter at the second position accounts for the biggest summand 3^{r-2}.)

1.5 Moduli among parameters in pseudo-normal forms.

Once the singularity classes (in the present paper – only for 2-flags) and faithful to them pseudo-normal forms EKR have been recalled, one of the first imposing questions is that about the *status* of real parameters entering the EKR forms. The same question concerning parameters in normal forms for germs of 1-flags, sparked by the benchmark work [KRui], had remained without answer over a considerable period 1982–97.

With examples of moduli of 1-flags at hand, it is not long to produce an example of an EKR parameter that is a true modulus. To this end, choose the following family of EKR's **1.2.1.2.1.2.1** sitting (see Theorem 2) in the singularity class 1.2.1.2.1.2.1:

$$\begin{aligned}
dx_1 - x_2 dt &= 0 & \qquad dy_1 - y_2 dt &= 0\\
dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\
dx_3 - (1 + x_4) dx_2 &= 0 & dy_3 - y_4 dx_2 &= 0\\
dx_2 - x_5 dx_4 &= 0 & dy_4 - y_5 dx_4 &= 0\\
dx_5 - (1 + x_6) dx_4 &= 0 & dy_5 - y_6 dx_4 &= 0\\
dx_4 - x_7 dx_6 &= 0 & dy_6 - y_7 dx_6 &= 0\\
dx_7 - (c + x_8) dx_6 &= 0 & dy_7 - y_8 dx_6 &= 0\,,
\end{aligned} \tag{2}$$

where $c \in \mathbb{R}$ is an arbitrary real parameter and these objects are considered as germs at $0 \in \mathbb{R}^{17}(t,\, x_1,\, y_1, \ldots,\, x_8,\, y_8)$. (Due to the Pfaffian equations' description, it is not instantly visible that the objects sit in an EKR. Yet, by the time we prove the statement in Appendix (Section 8), it will be clear that the proposed objects belong to a concrete EKR class of normal forms). The proof is being postponed to keep the exposition balanced.

Remark 2. (a) The 1-parameter family in (2) is, as it stands, written for the width $k = 2$ (there are only two columns of Pfaffian equations). However, a similar family could be proposed for any bigger width. The reader can easily figure out the potential 3rd, ..., kth columns, all constructed on the pattern of the second column, with no additional constants (the non-zero constants, decisive for the example, always in the first column only). The proof for the analogous objects inside the EKR class **1.2.1.2.1.2.1** in the space of special k-flags, $k > 2$, would be essentially the same, only the basic vector equation would be longer and so would be equations on the levels X_5 and X_3.

(b) The germs of special k-flags equivalent to these in (2), or in analogous families for $k > 2$, are thus uni- or more-modal (their modality in Arnold's sense is at least one). We suspect that their true modality is either two or three. A lot of work is needed in this direction. Already the analysis of the class 1.2.1.2 in section 5.2 indicates possible complications.

Remark 3. We want to note that the problems of local classification of special k-flags, $k \geq 2$, (and of 1-flags, too) have important affinities with those of local classifying of unparametrized curves in $\mathbb{R}^{k+1}$. That is, with the R-L classification of germs of mappings $\mathbb{R} \to \mathbb{R}^{k+1}$, although the two sets of problems are not the same. (In the 1-flags case, which stands out by the lack of a canonical analogue of the covariant subdistribution F, we mean the local classification of unparametrized contact curves in the contact space $\mathbb{R}^3$.) The first researchers who had gradually (from 1999 onwards) discovered those remarkable affinities were Montgomery and Zhitomirskii. From 2003 there has also been an important influx of ideas by Ishikawa. Later in section 7.7 we give, with quotations from [GHo, Ar], a concrete example of a striking (if only partial) interplay between the two fields.

1.6 Simple local construction of 2-flags of length 1 and 2.

Before dealing with the special 2-flags in lengths 3 and 4, we briefly survey the lengths 1 and 2 in which the bare sandwich classes are the orbits.

Theorem 3 ([KRub]). (i) *Any special 2-flag of length 1 can be locally brought to the following particular* EKR **1**

$$dx_1 - x_2 dt = 0 \qquad\qquad dy_1 - y_2 dt = 0$$

displaying no constants.

(ii) *Any germ of a special 2-flag of length 2 sitting in the generic sandwich class* 1.1 *can be brought to the following* EKR **1.1***,*

$$\begin{aligned} dx_1 - x_2 dt &= 0 \qquad\qquad & dy_1 - y_2 dt &= 0 \\ dx_2 - x_3 dt &= 0 \qquad\qquad & dy_2 - y_3 dt &= 0\,. \end{aligned}$$

Every germ of a special 2-flag of length 2 in the sandwich class $1.\underline{2}$ *of codimension 1 can be written as the following particular* **1.2***,*

$$\begin{aligned} dx_1 - x_2 dt &= 0 \qquad\qquad & dy_1 - y_2 dt &= 0 \\ dt - x_3 dx_2 &= 0 \qquad\qquad & dy_2 - y_3 dx_2 &= 0\,. \end{aligned}$$

Proof. Lemma 2 and Theorem 1 in [KRub] imply that **1** and **1.1** are single orbits, and that **1** exhausts, up to local equivalence, all special 2-flags of length 1. That **1.2** is a single orbit and that there are only two orbits in length 2, is explicitly written (albeit without proof) on p. 10^{8-10} in [KRub]. Here is a short explanation.

It is clear from Theorem 1 that the EKR families **1.1** and **1.2** do cover all orbits in length 2. We work with the latter family and take into account the simplification coming from item (i), thus having the members of **1.2** brought to the form

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & \qquad dy_2 - (c + y_3)dx_2 &= 0\,. \end{aligned}$$

In order to get rid of c there suffices to just take two new (bar) variables $y_2 = \overline{y}_2 + cx_2$ and $y_1 = \overline{y}_1 + cx_1$. It works because plugging the new expression for y_2 in $dy_1 - y_2dt$ brings in the term x_2dt which is but dx_1 due to the first Pfaffian equation in the left column. □

2 Strong versus weak nilpotency (in length three)

It is known since certain time (Theorem 4 in [M2]) that, on top of the Goursat distributions, also all special k-flags, and all the more so special 2-flags, are locally nilpotentizable, or: weakly nilpotent in the actually prevailing terminology. In fact, local bases given in the EKR presentations for them are nilpotent, and of nilpotency orders that can be effectively computed. On the other hand, only a tiny portion of germs of special k-flags seems to be *strongly nilpotent* in the precise sense of [AGau] and [M1]; that is, equivalent to their relevant *nilpotent approximations*. (Nilpotent approximations of distributions had been investigated by numerous researchers, with outstanding contributions [ASa, AGamSa, BiSt, Be]; see also [A] for an important coordinate-free description.)

This phenomenon has been discovered recently, [M1], amongst Goursat distributions.[3] In the present work it turns out to be of key importance in handling special 2-flags in lengths exceeding those of Theorem 3. For, in view of this theorem, the neatest EKR's available in these small lengths display no constants. Whence, by the last item of Theorem 4 of [M2],

Observation 2. *All germs of special 2-flags in lengths 1 and 2 are strongly nilpotent.*

(As a matter of record, in these lengths, the same is true for special flags of any width k.)

Among 2-flags of length 3, one singularity class stands out by its complication. It is 1.2.1, visualised – see Theorem 2 – by the EKR's in the family **1.2.1**. Most of the germs in 1.2.1 appear *not* to be strongly nilpotent. In order to see this clearly, we simplify the members of the visualising family by means of item (ii) of Theorem 3. That is, write constants (b and c in the occurrence) only in the bottommost Pfaffian equations.

Proposition 3. *The germ, D, at $0 \in \mathbb{R}^9$ of an* EKR

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & \qquad dy_2 - y_3dx_2 &= 0 \\ dx_3 - (b + x_4)dx_2 &= 0 & \qquad dy_3 - (c + y_4)dx_2 &= 0\,. \end{aligned} \tag{3}$$

[3] after a question by Agrachev whether the moduli of the local classification of Goursat objects survived the passage to nilpotent approximations

is strongly nilpotent iff $b = c = 0$.

The remaining of the present chapter is devoted to a proof of this (rather unexpected) fact. The basic reference is a highly constructive algorithm from [Be] for computing nilpotent approximations; to that algorithm one can often add shortcuts pertinent to objects under consideration, as is the case for (3).

When $b = c = 0$, the germ is strongly nilpotent by Theorem 4 (last item) of [M2]. Assume now $(b,\, c) \neq (0,\, 0)$. Under this assumption, upon computing the small flag of D at 0, it becomes visible that the small growth vector[4] of D at 0 is

$$[3,\, 5,\, 7,\, 8,\, 9] \tag{4}$$

and that an imposing-by-itself collection of *linearly adapted*, for D at 0, coordinates is

$$x_4,\;\; y_4,\;\; x_2,\;\; x_3 - bx_2,\;\; y_3 - cx_2,\;\; t,\;\; y_2,\;\; x_1,\;\; y_1\,.$$

The weights (read off from the small vector) attached to these variables are $1, 1, 1, 2, 2, 3, 3, 4, 5$, respectively; compare the increments in the vector (4). Improving these coordinates to *adapted* (i. e., having non-holonomic orders not only not exceeding, but coinciding with the weights) coordinates $z_1,\, z_2, \ldots,\, z_9$,

$$z_1 = x_4,\quad z_2 = y_4,\quad z_3 = x_2;\quad z_4 = x_3 - bx_2,\quad z_5 = y_3 - cx_2;$$

$$z_6 = t - \frac{b}{2}x_2^2,\quad z_7 = y_2 - \frac{c}{2}x_2^2;\quad z_8 = x_1 - \frac{b}{3}x_2^3;\quad z_9 = y_1 - \frac{bc}{8}x_2^4\,,$$

permits to ascertain the nilpotent approximation $\widehat{D}$ of D. To this end one has to watch $D = (Z_1,\, Z_2,\, Z_3)$ in these coordinates and extract all the (nilpotent) terms of weight -1 in the Taylor expansions of the vector fields' generators. It is clear that $Z_2 = \partial/\partial x_4$ becomes now ∂_1 and $Z_3 = \partial/\partial y_4$ becomes ∂_2. After more (elementary) computations there emerges the new form of the most involved generator Z_1 in our EKR,

$$Z_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 + \big(z_4z_7 + bz_3z_7 + \frac{c}{2}z_3^2z_4\big)\partial_9\,. \tag{5}$$

The only non-nilpotent term in all three generators is $z_4z_7\partial_9$ in Z_1 of weight $2+3-5 = 0$. All the remaining terms are of weight -1 and so survive the passing to the nilpotent approximation $\widehat{D}$. Consequently, that latter distribution is spanned by $\widehat{Z}_2 = \partial_1$, $\widehat{Z}_3 = \partial_2$ and by

$$\widehat{Z}_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 + \big(bz_3z_7 + \frac{c}{2}z_3^2z_4\big)\partial_9\,.$$

At this point $\widehat{D}$ is found, but not yet well understood. In order to analyze it smoothly, we pass to other, also adapted for D at 0, variables $z_1, \ldots,\, z_5,\, \overline{z_6}$,

[4] the small growth vector of a distribution D at a point p is the sequence of linear dimensions at p of the members of the small flag of D

$\overline{z_7}$, $\overline{z_8}$, $\overline{z_9}$, where

$$\overline{z_6} = z_6 - z_3 z_4\,, \quad \overline{z_7} = z_7 - z_3 z_5\,, \quad \overline{z_8} = z_8 - \frac{1}{2} z_3^{\,2} z_4\,,$$

$$\overline{z_9} \;=\; z_9 - \frac{b}{2} z_3^{\,2} z_7 + \frac{1}{6} z_3^{\,3} (b z_5 - c z_4)\,.$$

In them, the first generator of $\widehat{D}$ becomes tractable,

$$\widehat{Z}_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 - z_1 z_3 \partial_6 - z_2 z_3 \partial_7 - \frac{1}{2} z_1 z_3^{\,2} \partial_8 + \frac{1}{6} z_3^{\,3} (b z_2 - c z_1) \partial_9\,.$$

Now observe that each product of two or more factors from among $\widehat{Z}_1$, $\widehat{Z}_2$, $\widehat{Z}_3$ has no components in ∂_1, ∂_2, ∂_3, and depends only on z_1, z_2, z_3, as $\widehat{Z}_1$, $\widehat{Z}_2$, $\widehat{Z}_3$ themselves do. Therefore, any product of two products of factors from among $\widehat{Z}_1$, $\widehat{Z}_2$, $\widehat{Z}_3$ **vanishes**. In consequence, the big flag of $\widehat{D}$ coincides with the small one.
Hence the big growth vector of $\widehat{D}$ at 0 coincides with the small growth vector of $\widehat{D}$ at 0, and the latter is but the small growth vector of D at 0 (the key property of nilpotent approximations), that is (4). In this way we know the big vector of $\widehat{D}$ at 0, and find it different from the big vector of D at 0, $[3, 5, 7, 9]$. The germs at 0 of D and $\widehat{D}$ are thus non-equivalent. □

3 Classification in length three

Suppose that there is given a special 2-flag germ of length $r \geq 2$, generated by a rank-3 distribution $D = D^r$, displaying, at the reference point, an inclusion in the second sandwich. It was explained in section 1.5 that the locus, say H, of the inclusion $F(\cdot) \supset D^2(\cdot)$ is – always and automatically – an embedded codimension-one submanifold.

In length $r = 3$, around any point p of H one can ask if D is transverse or tangent to H at p.

Example 3. At points displaying the basic geometry 1.2.2, D is always transverse to H, whereas at the 1.2.3 points it is tangent to H. The reason becomes visible in any EKR glasses: around any 1.2.2 point, the generator Z_1 has the bare component $\partial/\partial x_3$, whereas around any 1.2.3 point that generator has the component $y_4\,\partial/\partial x_3$ that vanishes at 0.
This observation offers, besides, an alternative (and very simple) way of specifying the second letter $\underline{2}$ in the sandwich class $1.\underline{2}.\underline{2}$. (And more widely, in any class with a pair of neighbouring $\underline{2}$'s in the code, concerning the refining of the second $\underline{2}$ in the pair.) This way, however, does not lend itself to full scale generalizations, while the way recapitulated in section 1. 4 is universal.

Let us ask this question at points having the geometry 1.2.1. As we know already, the proper visualisation around these points is the pseudo-normal form (3) in which b, c are certain, à priori unknown parameters.

Observation 3. *Assume that the flag of D has at p the geometry* 1.2.1 *and that $H \ni p$ is the hypersurface of the inclusion in the second sandwich. Then D is tangent to H at p if and only if $b = 0$ in any visualisation* (3) *for D around p.*

In order to prove this one recalls that then $H = \{x_3 = 0\}$, while Z_1 has the component $(b + x_4)\partial/px_3$ taking at 0 the value $b\,\partial/\partial x_3$. □

In view of this observation, the singularity class 1.2.1 splits into two invariant parts, transverse and tangent. Independently, it splits (Proposition 3) into two other invariant parts depending on the strong nilpotency holding true or not. Moreover, the latter property ($b = c = 0$ in the glasses) implies the tangency ($b = 0$ in the glasses). The class 1.2.1 thus splits into *three* invariant parts

- $1.2.1_{-\mathrm{s,\,tra}}$ — germs in 1.2.1 not strongly nilpotent and transverse,
- $1.2.1_{-\mathrm{s,\,tan}}$ — germs in 1.2.1 not strongly nilpotent and tangent,
- $1.2.1_{+\mathrm{s}}$ — the strongly nilpotent germs in 1.2.1.

We are now in a position to locally classify the special 2-flags of length three.

Theorem 4. *In length three there exist altogether* 7 *orbits of the local classification of special 2-flags. The singularity classes* 1.1.1, 1.1.2, 1.2.2, 1.2.3 *of codimensions, resp.,* 0, 1, 2, 3, *are single orbits with the normal* EKR *forms* **1.1.1**, **1.1.2**, **1.2.2**, **1.2.3** *with all constants zero as respective local models.*

The three invariant parts of the singularity class 1.2.1 *of codimension* 1 *are orbits, too. In fact, all members of* $1.2.1_{-\mathrm{s,\,tra}}$ *are equivalent to*

$$\begin{aligned} dx_1 - x_2dt &= 0 \qquad & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 \qquad & dy_2 - y_3dx_2 &= 0\\ dx_3 - (1 + x_4)dx_2 &= 0 \qquad & dy_3 - y_4dx_2 &= 0\,, \end{aligned}$$

and this orbit has codimension one.

All members of $1.2.1_{-\mathrm{s,\,tan}}$ *are equivalent to*

$$\begin{aligned} dx_1 - x_2dt &= 0 \qquad & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 \qquad & dy_2 - y_3dx_2 &= 0\\ dx_3 - x_4dx_2 &= 0 \qquad & dy_3 - (1 + y_4)dx_2 &= 0\,, \end{aligned}$$

and the orbit' codimension is two.

All members of $1.2.1_{+\mathrm{s}}$ *are equivalent to the germ at* $0 \in \mathbb{R}^9$ *of*

$$\begin{aligned} dx_1 - x_2dt &= 0 \qquad & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 \qquad & dy_2 - y_3dx_2 &= 0\\ dx_3 - x_4dx_2 &= 0 \qquad & dy_3 - y_4dx_2 &= 0\,, \end{aligned}$$

and the codimension of this orbit is three.

Corollary 1. *Strongly nilpotent germs of special 2-flags in length three are all those in: the first four and the last (seventh), orbits in the above theorem.*

(As regards the first four orbits, it is so in view of Theorem 4, last item, in [M2]. At this point, however, it should be noted that in [M2] the families like **1.1.2** or **1.2.2** were not yet ultimately simplified, cf. p. 169 there.)

Corollary 2. *It follows from the contents of* Section 2 *that the germs in different orbits* $1.2.1_{-\mathrm{s,\,tra}}$ *and* $1.2.1_{-\mathrm{s,\,tan}}$ *have at the reference points the same s. gr. v.* (4)*. Thus, for special 2-flags, the small growth vector does not discern all orbits of the local classification already in length three. (For Goursat flags the smallest such length is seven.)*

Proof of Theorem 4. Concerning 1.1.1, it is again Theorem 1 of [KRub]. Concerning 1.1.2, one can, without loss of generality, work with the following EKR's,

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0\\ dx_2 - x_3dt &= 0 & dy_2 - y_3dt &= 0\\ dt - x_4dx_3 &= 0 & dy_3 - (c + y_4)dx_3 &= 0\,. \end{aligned}$$

As in the class 1.2, it is natural to venture $y_3 = \overline{y}_3 + cx_3$. Then this expression for y_3 plugged in to $dy_2 - y_3dt$ sparks a simplification, because $x_3dt = dx_2$. Thus $y_2 = \overline{y}_2 + cx_2$ is needed. And this y_2 substituted to $dy_1 - y_2dt$ causes again a simplification due to $x_2dt = dx_1$, and $y_1 = \overline{y}_1 + cx_1$ is needed to conclude.

As regards 1.2.2, without loss of generality one can work with the following EKR's,

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0\\ dx_2 - x_4dx_3 &= 0 & dy_3 - (c + y_4)dx_3 &= 0\,, \end{aligned} \tag{6}$$

trying to reduce to 0 the constant c. The technique is similar to that employed for the previous class. One starts with $y_3 = \overline{y}_3 + cx_3$, then spots $x_3dx_2 = dt$ holding true in the system (6) and takes $y_2 = \overline{y}_2 + ct$, after which concludes with $y_1 = \overline{y}_1 + \frac{c}{2}t^2$.

As for 1.2.3, no work is needed after previous simplifications in length two, and the local model

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0\\ dx_2 - x_4dy_3 &= 0 & dx_3 - y_4dy_3 &= 0\,, \end{aligned}$$

follows. Also the part $1.2.1_{+\mathrm{s}}$ of 1.2.1, after Proposition 3, needs no work, and the same applies to the part $1.2.1_{-\mathrm{s,\,tan}}$: in the pseudo-normal form (3) there must hold $b = 0$, $c \neq 0$, and such c is easily normalizable to 1.

There remains the part $1.2.1_{-\mathrm{s,\,tra}}$ of 1.2.1 when, in (3), $b \neq 0$ and c is arbitrary. One can assume (by means of passing to the bar variables $x_4 = b\overline{x}_4$, $x_3 = b\overline{x}_3$, $t = b\overline{t}$,

$x_1 = b\overline{x}_1$, $y_1 = b\overline{y}_1$) that $b = 1$. Then starts as in previous cases with $y_3 = \overline{y}_3 + cx_3$. But $dx_3 = (1 + x_4)dx_2$ in the Pfaffian system in question. Hence

$$dy_3 - (c + y_4)dx_2 = d\overline{y}_3 + c(1 + x_4)dx_2 - (c + y_4)dx_2 = d\overline{y}_3 - (y_4 - cx_4)dx_2\,.$$

Now it imposes by itself to write $y_4 = \overline{y}_4 + cx_4$, continue with $y_2 = \overline{y}_2 + ct$, and conclude with $y_1 = \overline{y}_1 + \frac{c}{2}t^2$. Theorem 4 is proved. □

Remark 4. The local classifications of special k-flags, $k \geq 2$, of lengths not exceeding three coincide with those in width two, $k = 2$. In particular, Theorem 4 directly generalizes: there are always 7 orbits (four of them being singularity classes and the remaining three building up the class 1.2.1) having the same characterizations as in width two.
In other words, the classifications in lengths not exceeding three are *stable* with respect to width $k \geq 2$.

4 Classification in length four – simpler part

The question that stands now is about the orbits sitting inside the fourteen singularity classes listed in Example 1. We start with with an elementary

Theorem 5. *In length four, only the following* 6 *singularity classes of germs of special* 2*-flags (out of altogether* 14 *existing in that length) are single orbits of the local classification:* 1.1.1.1, 1.1.1.2, 1.1.2.2, 1.1.2.3, 1.2.2.3, 1.2.3.3. *As unique local models there can be taken, respectively, the* EKR's **1.1.1.1***,* **1.1.1.2***,* **1.1.2.2***,* **1.1.2.3***,* **1.2.2.3***,* **1.2.3.3** *with all constants appearing in them equal to* 0*. In consequence, all these classes are strongly nilpotent.*

Proofs for these classes go entirely analogously to those in length three concerning the classes 1.1.1, 1.1.2, 1.2.2, and 1.2.3; only the chains of consecutive passings from variables y to $\overline{y}$ are sometimes longer by one step.

Nextly we group together four singularity classes that split (each of them) into no more than three orbits.

Theorem 6. *In length four, the classes* 1.2.2.2 *and* 1.2.3.2 *consist of two orbits each. Whereas the classes* 1.1.2.1 *and* 1.2.1.3 *consist of three orbits each. The codimensions of orbits, and local models, are listed in the proof.*

4.1 Proof for 1.1.2.1 – the strong nilpotency at work.

The method for 1.1.2.1 is analogous to that for the class 1.2.1, and 1.1.2.1 splits into: $1.1.2.1_{+\mathrm{s}}$ – strongly nilpotent germs (an orbit of codimension three), $1.1.2.1_{-\mathrm{s,\,tra}}$ (a generic part in 1.1.2.1 and an orbit of codimension one equal to the codimension of the class) – germs not strongly nilpotent and transverse to the locus $\widetilde{H}$ of hitting the sandwich class $1.1.\underline{2}.1$, and $1.1.2.1_{-\mathrm{s,\,tan}}$ (an orbit of codimension two) – germs not strongly nilpotent and tangent to $\widetilde{H}$ at the reference point.

When searching for local models in 1.1.2.1, the unique local model for 1.1.2,

$$\begin{aligned} dx_1 - x_2 dt &= 0 \qquad\qquad & dy_1 - y_2 dt &= 0\\ dx_2 - x_3 dt &= 0 & dy_2 - y_3 dt &= 0\\ dt - x_4 dx_3 &= 0 & dy_3 - y_4 dx_3 &= 0\,, \end{aligned}$$

is to be extended by a couple of equations

$$dx_4 - (B + x_5)dx_3 = 0 \qquad\qquad dy_4 - (C + y_5)dx_3 = 0$$

with two new parameters B and C. For a representative of the strong nilpotency part $1.1.2.1_{+\mathrm{s}}$, we take $B = C = 0$. The proof that the complement in 1.1.2.1 of the germs equivalent to the particular **1.1.2.1** with $B = C = 0$, consists of not strongly nilpotent germs, now splits into two separate proofs, according to two highly different underlying geometries. (For 1.2.1, off the scope of strong nilpotency, there were also two different geometries, yet they displayed one and the same small growth vector, cf. Corollary 2, and could have been subsumed into one computation.)

Namely, the 'sandwich' locus $\widetilde{H}$ has now the equation $x_4 = 0$, and the germs with $B \neq 0$ are transversal to $\widetilde{H}$, while those with $B = 0$ are tangent to $\widetilde{H}$ at 0 (our reference point). In the normal forms for transversal ones, the constant C can be easily reduced to 0 (as in the handling of $1.2.1_{-\mathrm{s,\,tra}}$ in the proof of Thm. 4). A local model for $1.1.2.1_{-\mathrm{s,\,tra}}$ is obtained by further normalizing B to 1. In the normal forms for tangent germs, C can be reduced to 1, yielding a model for $1.1.2.1_{-\mathrm{s,\,tan}}$.

- All germs in $1.1.2.1_{-\mathrm{s,\,tra}}$ are not strongly nilpotent:

A careful computation shows that, independently of a germ in $1.1.2.1_{-\mathrm{s,\,tra}}$, the departure point – the small growth vector at the reference point – is $[3, 5, 7, 8, 9, 10, 11]$. We work with C already annihilated and $B \neq 0$ – and improve the starting EKR coordinates to linearly adapted

$$x_5,\;\; y_5,\;\; x_3,\;\; x_4 - Bx_3,\;\; y_4,\;\; t,\;\; y_3,\;\; x_2,\;\; y_2,\;\; x_1,\;\; y_1\,.$$

These coordinate functions are not yet adapted (the attached weights, read off from the small vector, are $1, 1, 1, 2, 2, 3, 3, 4, 5, 6, 7$, while the non-holonomic orders of functions are, in some cases, smaller). Improving them further, by Bellaïche adopted to the situation, yields (certain) adapted coordinates

$$z_1 = x_5, \quad z_2 = y_5, \quad z_3 = x_3; \quad z_4 = x_4 - Bx_3, \quad z_5 = y_4; \quad z_6 = t - \frac{B}{2}x_3^2,$$

$$z_7 = y_3\,; \quad z_8 = x_2 - \frac{B}{3}x_3^3\,; \quad z_9 = y_2\,; \quad z_{10} = x_1 - \frac{B^2}{15}x_3^5\,; \quad z_{11} = y_1$$

in which the nilpotent approximation can be distracted already. Namely, $Z_2 = \partial/\partial x_5$ becomes now ∂_1 and $Z_3 = \partial/\partial y_5$ becomes ∂_2, while Z_1 assumes the form

$$\begin{aligned} Z_1 \;&= \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 + (\underline{z_4z_7} + Bz_3z_7)\partial_9\\ &\quad + \Big(\underline{z_4z_8} + Bz_3z_8 + \frac{B}{3}z_3^3z_4\Big)\partial_{10} + \big(\underline{z_4z_9} + Bz_3z_9\big)\partial_{11}\,. \end{aligned}$$

The underlined terms are of degree 0, the remaining are of degree -1. Recalling, passing from a distribution to its nilpotent approximation consists in leaving out all the terms, in arbitrary adapted coordinates, of degrees exceeding -1. In the present case, thus, $\widehat{D}$ is generated by $\widehat{Z}_2 = Z_2,\ \widehat{Z}_3 = Z_3$ and by

$$\widehat{Z}_1 = \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 \\ + Bz_3z_7\partial_9 + \big(Bz_3z_8 + \frac{B}{3}z_3^3z_4\big)\partial_{10} + Bz_3z_9\partial_{11}\,. \tag{7}$$

Similarly as working earlier with 1.2.1, through (7) one does *not* see the properties of $\widehat{D}$. Hence seeks coordinates that are *more* adapted. After a careful search, $z_1, \dots, z_5$ and

$$\overline{z_6} = z_6 - z_3z_4\,, \qquad \overline{z_7} = z_7 - z_3z_5\,, \qquad \overline{z_8} = z_8 - \frac{1}{2}z_3^2z_4\,,$$

$$\overline{z_9} \;=\; z_9 - \frac{B}{2}z_3^2z_7 + \frac{B}{6}z_3^3bz_5\,, \qquad \overline{z_{10}} \;=\; z_{10} - \frac{B}{2}z_3^2z_8 + \frac{B}{24}z_3^4z_4\,,$$

$$\overline{z_{11}} \;=\; z_{11} - \frac{B}{2}z_3^2z_9 + \frac{B^2}{8}z_3^4z_7 - \frac{B^2}{40}z_3^5z_5\,,$$

are such that $\widehat{Z}_2$ and $\widehat{Z}_3$ remain unchanged, while $\widehat{Z}_1$ assumes the form

$$\widehat{Z}_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 - z_1z_3\partial_6 - z_2z_3\partial_7 - \frac{1}{2}z_1z_3^2\partial_8 \\ + \frac{B}{6}z_2z_3^3\partial_9 + \frac{B}{24}z_1z_3^4\partial_{10} - \frac{B^2}{40}z_2z_3^5\partial_{11}\,. \tag{8}$$

That is to say, the components ∂_4 through ∂_{11} in the fields $\widehat{Z}_i$, $i = 1, 2, 3$ spanning $\widehat{D}$ depend now only on z_1, z_2, z_3, while the $(\partial_1, \partial_2, \partial_3)$-components are constant. This implies, like earlier in the proof of Proposition 3, the coincidence of the small and big growth vectors of $\widehat{D}$ at the reference point 0. In consequence, the germs at 0, D and $\widehat{D}$, have different big growth vectors, hence are non-equivalent.

•• All germs in $1.1.2.1_{-\mathrm{s,tan}}$ are not strongly nilpotent:

We begin with a remark that a new proof is needed for this part because the small growth vector governing $1.1.2.1_{-\mathrm{s,tan}}$ is different from that servicing $1.1.2.1_{-\mathrm{s,tra}}$. In fact, after a delicate computation, it is $[3, 5, 7, 8, 9, 9, 10, 11]$.[5] Now $B = 0$ in the pseudo-normal forms, and we purposedly keep a general $C \neq 0$. The argument evolves, again, stepwise. Firstly one passes from the EKR coordinates to linearly adapted

$$x_5,\ y_5,\ x_3,\ x_4,\ y_4 - Cx_3,\ t,\ y_3,\ x_2,\ y_2,\ x_1,\ y_1\,,$$

whose weights are now $1, 1, 1, 2, 2, 3, 3, 4, 5, 7, 8$, respectively. It appears that, among them, only y_3 is not adapted: its non-holonomic order is 2, and weight 3; it suffices to improve it to $y_3 - \frac{C}{2}x_3^2$. In this way there emerges a set of adapted coordinates

$$z_1 = x_5\,, \quad z_2 = y_5\,, \quad z_3 = x_3\,; \quad z_4 = x_4\,, \quad z_5 = y_4 - Cx_3\,; \quad z_6 = t,$$

[5] it was not so in length three with $1.2.1_{-\mathrm{s,tra}}$ and $1.2.1_{-\mathrm{s,tan}}$! This kind of complications, when the length grows, is typical in special 2-flags.

$$z_7 = y_3 - \frac{C}{2}x_3^2\,; \quad z_8 = x_2\,; \quad z_9 = y_2\,; \quad z_{10} = x_1\,; \quad z_{11} = y_1\,.$$

In these coordinates $Z_2 = \partial_1$, $Z_3 = \partial_2$, and

$$Z_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 + \big(\underline{z_4z_7} + \frac{C}{2}z_3^2z_4\big)\partial_9 + z_4z_8\partial_{10} + z_4z_9\partial_{11}\,.$$

The nilpotent approximation $\big(\widehat{Z}_1,\,\widehat{Z}_2,\,\widehat{Z}_3\big)$ is obtained by only leaving out this one underlined term $z_4z_7\partial_9$ of degree 0 in Z_1. That is, $\widehat{Z}_2 = \partial_1$, $\widehat{Z}_3 = \partial_2$, and

$$\widehat{Z}_1 = \partial_3 + z_1\partial_4 + z_2\partial_5 + z_4\partial_6 + z_5\partial_7 + z_3z_4\partial_8 + \frac{C}{2}z_3^2z_4\partial_9 + z_4z_8\partial_{10} + z_4z_9\partial_{11}\,. \qquad (9)$$

As often in flags, nothing is visible in such Bellaïche-like vector field (9) save that it seems not possible that just leaving out the above single term results in the slowing down of the big vector, at the reference point, from $[3,5,7,9,11]$ (for D) to $[3,5,7,8,9,9,10,11]$ (for $\widehat{D}$). But this is the case! To see this, it suffices to improve the adapted coordinates to

$$\overline{z_6} = z_6 - z_3z_4\,, \qquad \overline{z_7} = z_7 - z_3z_5\,, \qquad \overline{z_8} = z_8 - \frac{1}{2}z_3^2z_4\,,$$

$$\overline{z_9} \;=\; z_9 - \frac{C}{6}z_3^3z_4\,, \qquad \overline{z_{10}} = z_{10} - z_3z_4z_8 + \frac{1}{2}z_1z_3^2z_8 + \frac{1}{3}z_3^3z_4^2 - \frac{7}{24}z_1z_3^4z_4\,,$$

$$\overline{z_{11}} \;=\; z_{11} - z_3z_4z_9 + \frac{1}{2}z_1z_3^2z_9 + \frac{C}{8}z_3^4z_4^2 - \frac{C}{10}z_1z_3^5z_4\,.$$

In these [more sophisticated] $z_1, \dots, z_5, \overline{z_6}, \dots, \overline{z_{11}}$, the involved generator (9) becomes but

$$\widehat{Z}_1 \;=\; \partial_3 + z_1\partial_4 + z_2\partial_5 - z_1z_3\partial_6 - z_2z_3\partial_7 - \frac{1}{2}z_1z_3^2\partial_8 - \frac{C}{6}z_1z_3^3\partial_9 - \frac{7}{24}z_1^2z_3^4\partial_{10} - \frac{C}{10}z_1^2z_3^5\partial_{11}\,.$$

And the reader knows already that such an expression, using only z_1, z_2, z_3 in components, guarantees that the big and small vectors of $\widehat{D}$, and hence the small of D, all coincide. Thus $\widehat{D}$ is far from being equivalent to D.

4.2 1.2.1.3 .

Concerning 1.2.1.3, the previous discussion of 1.2.1 applies to the Lie squares of members of this class, while the prolongation to length four leaves no freedom on the level of EKR pseudo-normal forms, because the last 3 in the code corresponds to the prolongation pattern **3** that brings in no new parameters. In fact, a distribution now being denoted D^4 and its [factored out] Lie square $D^3/L(D^3)$ being a distribution germ sitting in the class 1.2.1, 1.2.1.3 is split up according to the local geometry of $D^3/L(D^3)$: of the type $1.2.1_{+\mathrm{s}}$, or $1.2.1_{-\mathrm{s,\,tra}}$, or else $1.2.1_{-\mathrm{s,\,tan}}$. In either case the relevant local model for $D^3/L(D^3)$ is being extended by one precise pair of Pfaffian equations

$$dx_2 - x_5dy_4 = 0 \qquad\qquad dx_4 - y_5dy_4 = 0\,.$$

4.3 Proof for the classes 1.2.2.2 and 1.2.3.2 .

It turns out that the germs of special 2-flags sitting in 1.2.2.2 are **either** strongly nilpotent, $1.2.2.2_{+s}$, and then equivalent to the EKR

$$\begin{aligned} dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0 \\ dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0 \\ dx_2 - x_4 dx_3 &= 0 & dy_3 - y_4 dx_3 &= 0 \\ dx_3 - x_5 dx_4 &= 0 & dy_4 - y_5 dx_4 &= 0 \end{aligned} \tag{10}$$

(this orbit is of codimension four – its materialization has, for the object (10), local equations $x_3 = x_4 = x_5 = y_5 = 0$), **or else** not strongly nilpotent, and in that case equivalent to

$$\begin{aligned} dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0 \\ dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0 \\ dx_2 - x_4 dx_3 &= 0 & dy_3 - y_4 dx_3 &= 0 \\ dx_3 - x_5 dx_4 &= 0 & dy_4 - (1 + y_5) dx_4 &= 0 \end{aligned} \tag{11}$$

(this is the generic orbit of codimension three; its materialization has, for the object (11), local equations $x_3 = x_4 = x_5 = 0$).

In fact, to show that the orbit of (11) consists of not strongly nilpotent germs is rather lengthy; instead, we are going to demonstrate (what is enough for theorem) the non-equivalence to the strong nilpotency part $1.2.2.2_{+s}$.

Indeed, suppose that the object (10) is equivalent, as the germ at $0 \in \mathbb{R}^{11}$, to an EKR like (11), with a constant C in the place of 1 in the last Pfaffian equation there. That is, suppose the existence of a conjugating diffeomorphism

$$\Phi \;=\; (T,\, X_1,\, Y_1,\, X_2,\, Y_2,\, X_3,\, Y_3,\, X_4,\, Y_4,\, X_5,\, Y_5) : \; (\mathbb{R}^{11},\, 0) \hookleftarrow$$

(note its preserving of 0, for only the germs at 0 are being discussed). The aim is to show that $C = 0$. Similar situations of hypothetical conjugacies between different EKR's will frequently occur later. Because of that it is important to carefully describe restrictions such Φ (and several other conjugacies appearing later in the paper) is subject to. First of all, the EKR's that are conjugated have, by Observation 1, **the same** nicely positioned subflag of associated involutive subdistributions

$$F \supset L(D^1) \supset L(D^2) \supset L(D^3) \supset L(D^4) = 0$$

which must be preserved by Φ. It implies that

- $T,\, X_1,\, Y_1$ depend only on $t,\, x_1,\, y_1$,
- for $2 \le j \le 4$, functions $X_j,\, Y_j$ depend only on $t,\, x_1,\, y_1, \ldots,\, x_j,\, {}_j$.

In turn, it will momentarily become visible that – in the discussed situation – one knows even more about the components $X_3,\ X_4,$ and X_5.
Indeed, whenever there happens – as in our case – an inclusion in the second sandwich, $F(0) \supset D^2(0)$, of the sandwich diagram for D given by (10) or by (11), it happens not at isolated points like 0 but **in codimension one**. For, in any EKR for D in the vicinity of 0, taking again into account Observation 1, the locus of the inclusion $F(\cdot) \supset D^2(\cdot)$ has the equation $x_3 = 0$. Similar remarks apply to the inclusions in the third and fourth sandwiches, $L(D^1)(\cdot) \supset D^3(\cdot)$ and $L(D^2)(\cdot) \supset D^4(\cdot)$.
Therefore, both flags have the same singularity loci of the inclusions holding true in the indicated sandwiches, and these loci locally are but the hyperplanes $\{x_3 = 0\}$, $\{x_4 = 0\}$, and $\{x_5 = 0\}$. The mapping Φ preserves these, meaning that its relevant components are divisible, as function germs, by x_3, x_4, x_5, respectively. I. e., that there exist invertible at 0 functions K, H, G, also only depending on the variables specified above and satisfying

- $X_3(t,\, x_1, \ldots,\, y_3) = x_3 K(t,\, x_1, \ldots,\, x_3,\, y_3)\,,$
- $X_4(t,\, x_1, \ldots,\, y_4) = x_4 H(t,\, x_1, \ldots,\, x_4,\, y_4)\,,$
- $X_5(t,\, x_1, \ldots,\, y_5) = x_5 G(t,\, x_1, \ldots,\, x_5,\, y_5)\,,$

(letters are taken in this order because of the subsequent nestings $x_5G \to x_4H \to x_3K$ in (12)). Proceeding in our arguments, let us reiterate that Φ preserves the distribution $(\partial/\partial x_5,\, \partial/\partial y_5)$ (which in both cases is $L(D^3)$). In consequence there must exist an invertible at 0 function f, $f\,|_0 \neq 0$, such that

$$d\Phi(p)\begin{bmatrix} x_5\begin{pmatrix} x_4\begin{pmatrix} x_3\begin{pmatrix}1\\ x_2\\ y_2\end{pmatrix}\\ 1\\ y_3\end{pmatrix}\\ 1\\ y_4\end{pmatrix}\\ 1\\ y_5\\ 0\\ 0\end{bmatrix} = \begin{bmatrix} f(p)\begin{pmatrix} x_5G\begin{pmatrix} x_4H\begin{pmatrix} x_3K\begin{pmatrix}1\\ X_2\\ Y_2\end{pmatrix}\\ 1\\ Y_3\end{pmatrix}\\ 1\\ Y_4\end{pmatrix}\\ 1\\ C+Y_5\end{pmatrix}\\ *\\ *\end{bmatrix} \tag{12}$$

where the $*$'s are functions whose nature is irrelevant for the argumentation. In (12), for brevity, p stands for $(t,\, x_1,\, y_1, \ldots,\, x_5,\, y_5)$. The first conclusion from this rich set of conditions is

$$\frac{\partial Y_4}{\partial x_4}\,|_0 \;=\; C\,f\,|_0\,, \tag{13}$$

after which one looks for an information on Y_4. The 7-th row of (12), after dividing it sidewise by x_5, gives an expression for fGY_4 in terms of Y_3 which in turn implies

$$fG\,\frac{\partial Y_4}{\partial x_4}\,|_0 \;=\; \frac{\partial Y_3}{\partial x_2}\,|_0\,, \tag{14}$$

after which one looks for an information on Y_3. And the 5-th row of (12), after dividing it sidewise by x_5x_4, supplies an expression for $fGHY_3$ in terms of Y_2. That expression implies, among others, that

$$\frac{\partial Y_2}{\partial x_2}\,|_0 \;=\; fGH\,Y_3\,|_0 \;=\; 0 \tag{15}$$

and

$$\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0 \;=\; fGH\,\frac{\partial Y_3}{\partial x_2}\,|_0 \;. \tag{16}$$

One predicts already that, after dividing it sidewise by $x_5x_4x_3$, the 3-rd row of (12) yields an expression for $fGHKY_2$ in terms of Y_1. It is crucial that that expression is *affine* in x_2 – its second derivative wrt x_2 vanishes identically. At the same time that second derivative at 0 is equal to

$$2\frac{\partial(fGHK)}{\partial x_2}\frac{\partial Y_2}{\partial x_2} + fGHK\,\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0 \;=\; fGHK\,\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0$$

(the last equality in view of (15)). Therefore, the LHS, and hence also the RHS of (16) vanishes. Now (14) and (13) directly infer $C = 0$. So, indeed, the zero and non-zero values of C are not equivalent. On the other hand, any non-zero value can be easily rescaled to the value 1 – all such germs are equivalent to (11). The class 1.2.2.2 is settled.

As for the class 1.2.3.2, its members are either strongly nilpotent and equivalent to

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0\\ dx_2 - x_4dy_3 &= 0 & dx_3 - y_4dy_3 &= 0\\ dy_3 - x_5dx_4 &= 0 & dy_4 - y_5dx_4 &= 0\,, \end{aligned}$$

building up the orbit $1.2.3.2_{+\mathrm{s}}$ of codimension five (with local equations of its materialization being $x_3 = x_4 = y_4 = x_5 = y_5 = 0$), or else not strongly nilpotent and equivalent to

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0\\ dx_2 - x_4dy_3 &= 0 & dx_3 - y_4dy_3 &= 0\\ dy_3 - x_5dx_4 &= 0 & dy_4 - (1 + y_5)dx_4 &= 0\,, \end{aligned}$$

building up the generic orbit $1.2.3.2_{-\mathrm{s}}$ of codimension four (with local equations $x_3 = x_4 = y_4 = x_5 = 0$). The proof of the non-equivalence of these two representatives is analogous (and simpler) than that servicing the class 1.2.2.2; the lack of the strong nilpotency within the second orbit is, however, even harder to show.

5 Classification in length four – harder part

It is still more surprising that

Theorem 7. *The singularity classes* 1.2.1.2, 1.2.2.1 *and* 1.2.3.1 *split into four orbits each. The codimensions and local models are given below in the proofs.*

5.1 Proof for the class 1.2.2.1 .

As previously, the Lie square of a distribution germ, factored out by its Cauchy characteristics sits in 1.2.2 whose unique local model is known. So one can take those Pfaffian equations

$$\begin{aligned} dx_1 - x_2dt &= 0 \qquad & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 \qquad & dy_2 - y_3dx_2 &= 0\\ dx_2 - x_4dx_3 &= 0 \qquad & dy_3 - y_4dx_3 &= 0\,, \end{aligned}$$

and add to them a couple of new ones,

$$dx_4 - (B + x_5)dx_3 = 0 \qquad\qquad dy_4 - (C + y_5)dx_3 = 0$$

with unknown parameters B and C. The situations $B \neq 0$ and $B = 0$ are geometrically different, and hence non-equivalent: the vanishing of B means precisely the tangency of a distribution at the reference point (here 0) to the locus of the inclusion in the 3-rd sandwich. Moreover, in the transverse case $B \neq 0$ it is easy to normalize B to 1. Assuming this done already, now instead of B we have a discrete parameter $\epsilon \in \{0,\, 1\}$ that bears a geometric meaning: $\epsilon = 1$ is transversality, $\epsilon = 0$ – tangency. And, keeping ϵ constant, we try to conjugate, via a preserving the origin diffeomorphism $\Phi = (T,\, X_1,\, Y_1, \dots,\, X_5,\, Y_5)$ of $\mathbb{R}^{11}$ into itself, the two relevant EKR's: with $C = 0$ and $C \neq 0$. This boils down, as in the discussion in Section 4.3, to the vector equation

$$d\Phi(p)\begin{array}{c} x_4\begin{pmatrix} x_3\begin{pmatrix}1\\ x_2\\ y_2\end{pmatrix}\\ 1\\ y_3\end{pmatrix}\\ 1\\ y_4\\ \epsilon + x_5\\ y_5\\ 0\\ 0\end{array} = f(p)\begin{array}{c}\begin{pmatrix} x_4G\begin{pmatrix} x_3H\begin{pmatrix}1\\ X_2\\ Y_2\end{pmatrix}\\ 1\\ Y_3\end{pmatrix}\\ 1\\ Y_4\\ \epsilon + X_5\\ C + Y_5\end{pmatrix}\\ *\\ *\end{array} \tag{17}$$

where $f(0) \neq 0$ and now only $X_3 = x_3H$, $X_4 = x_4G$ are of such special form (inclusions holding only in 2nd and 3rd sandwich).

The 9-th row in (17), evaluated at 0, reads

$$\frac{\partial Y_4}{\partial x_3} + \epsilon \frac{\partial Y_4}{\partial x_4} \,|_0 \;=\; C\, f \,|_0 \;. \tag{18}$$

The 7-th row in (17) gives fY_4 in function of Y_3, which implies

$$0 \;=\; fY_4 \,|_0 \;=\; \frac{\partial Y_3}{\partial x_3} \,|_0 \,, \tag{19}$$

$$f \,\frac{\partial Y_4}{\partial x_3} \,|_0 \;=\; \frac{\partial^2 Y_3}{\partial x_3^2} \,|_0 \,, \tag{20}$$

and

$$f \,\frac{\partial Y_4}{\partial x_4} \,|_0 \;=\; \frac{\partial Y_3}{\partial x_2} \,|_0 \,, \tag{21}$$

In a cascade of arguments, the 5-th row of (17), after dividing it sidewise by x^4, yields an expression for fGY_3, in terms of Y_2, which is affine in x_3. Hence its second derivative wrt x_3 vanishes, and in particular

$$0 \;=\; 2\,\frac{\partial(fG)}{\partial x_3}\frac{\partial Y_3}{\partial x_3} + fG\,\frac{\partial^2 Y_3}{\partial x_3^2} \,|_0 \;=\; fG\,\frac{\partial^2 Y_3}{\partial x_3^2} \,|_0 \tag{22}$$

(in view of (19)). Now this equality (22) together with (20) show that the first summand on the LHS in (18) vanishes. Passing to the second summand, that mentioned above expression for fGY_3 implies not only (22) but also

$$0 \;=\; fGY_3 \,|_0 \;=\; \frac{\partial Y_2}{\partial x_2} \,|_0 \tag{23}$$

and

$$fG\,\frac{\partial Y_3}{\partial x_2} \,|_0 \;=\; \frac{\partial^2 Y_2}{\partial x_2^2} \,|_0 \;. \tag{24}$$

And this last equality, via (21), reduces the handling of the term $\frac{\partial Y_4}{\partial x_4} \,|_0$ in (18) to the second derivative at 0 of Y_2 with respect to x_2.
Continuing the cascade, it is the 3-rd row in (17) which, after dividing it sidewise by $x_4 x_3$, gives an affine in x_2 expression for $fGH\,Y_2$. That expression, doubly differentiated wrt x_2 to an identical zero, implies

$$0 \;=\; 2\,\frac{\partial(fGH)}{\partial x_2}\frac{\partial Y_2}{\partial x_2} + fGH\,\frac{\partial^2 Y_2}{\partial x_2^2} \,|_0 \;=\; fGH\,\frac{\partial^2 Y_2}{\partial x_2^2} \,|_0$$

(the last equality by (23)). The needed derivative turns out to be zero, and so is the LHS, hence also RHS, of (18). We have shown that $C = 0$. Thus, for either of the two values of ϵ, the zero and non-zero values of C are shown to be non-equivalent. On the other hand, a non-zero C is easily normalizable to 1. So the class 1.2.2.1 splits into four orbits having for local models the relevant EKR's with the constants

- $B = 1, \quad C = 1$ (the generic orbit of codimension two),
- $B = 1, \quad C = 0$ (an orbit of codimension three),
- $B = 0, \quad C = 1$ (an orbit of codimension three),
- $B = 0, \quad C = 0$ (the strongly nilpotent part of codimension four).

The last orbit should be denoted by $1.2.2.1_{+\mathrm{s}}$, but it is long to show that the remaining orbits contain only not strongly nilpotent distribution germs.

5.2 The discussion of 1.2.3.1 and 1.2.1.2 .

Passing to the singularity class 1.2.3.1, the orbits sitting inside it have [superficially] much similar description to those inside 1.2.2.1. We mean the equations for the class 1.2.3,

$$\begin{aligned} dx_1 - x_2 dt &= 0 & \qquad dy_1 - y_2 dt &= 0 \\ dt - x_3 dx_2 &= 0 & \qquad dy_2 - y_3 dx_2 &= 0 \\ dx_2 - x_4 dy_3 &= 0 & \qquad dx_3 - y_4 dy_3 &= 0\,, \end{aligned}$$

for the square of a distribution under consideration, extended by the pair of equations pertinent to the (last) cipher 1 in the code 1.2.3.1 ,

$$dx_4 - (B + x_5) dy_3 = 0 \qquad\qquad dy_4 - (C + y_5) dy_3 = 0$$

in which, naturally, one has to normalize the constants whenever possible. Every such EKR sits in the sandwich class $1.\underline{2}.\,\underline{2}$ and so the inclusions at the reference point 0 hold in both the 2-nd and 3-rd sandwich. The loci of them are $\{x_3 = 0\}$ and $\{x_4 = 0\}$, independently of the values of B and C. The distribution represented by a given pair of values is tangent at 0 to the latter locus if and only if $B = 0$. One can quickly inspect this tangent situation in purely geometric terms. Namely, for each of the EKR's in question the locus of the singularity class 1.2.3 (for the Lie square) is $\{x_3 = x_4 = y_4 = 0\}$. In the tangent situation $B = 0$, it is natural to ask the question whether the distribution is tangent, at the reference point 0, to this locus. And it is iff $C = 0$. Hence the germs in 1.2.3.1 equivalent to the EKR with $B = C = 0$ are simultaneously tangent to the two singularity loci: of $1.2.\underline{2}$ and 1.2.3. Whereas those equivalent to an EKR with $B = 0$, $C \neq 0$ are tangent to the locus of the inclusion $D^3 \subset L(D^1)$, but not to the locus of more fine geometry 1.2.3.

In the transvese case, it is straightforward to normalize B to 1, after which there pops up the question of the relevance of C. So we try, exactly as for 1.2.2.1, to conjugate, by means of a diffeomorphism Φ, the zero value with a non-zero C. The mentioned loci have, of course, to be preserved by $\Phi = (T, X_1, \ldots, Y_5)$, whence the components X_3 and X_4 of Φ are of special form, $X_4 = x_4 G$ and $X_3 = x_3 H$; G, H invertible at 0. Moreover, there must hold

$$
d\Phi(p)\begin{pmatrix} x_4\begin{pmatrix} x_3\begin{pmatrix}1\\ x_2\\ y_2\end{pmatrix} \\ 1\\ y_3\end{pmatrix} \\ y_4\\ 1\\ 1+x_5\\ y_5\\ 0\\ 0\end{pmatrix} \;=\; f(p)\begin{pmatrix} x_4G\begin{pmatrix} x_3H\begin{pmatrix}1\\ X_2\\ Y_2\end{pmatrix} \\ 1\\ Y_3\end{pmatrix} \\ Y_4\\ 1\\ 1+X_5\\ C+Y_5\end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ * \\ * \end{matrix} \tag{25}
$$

with an invertible at 0 factor function f. We will use this set of conditions as modestly as possible. The main relation, implied by the 9-th row in (25), reads

$$
\frac{\partial Y_4}{\partial y_3} + \frac{\partial Y_4}{\partial x_4}\,|_0 \;=\; C\,f\,|_0\;. \tag{26}
$$

It will momentarily turn out that both summands on the left disappear. Indeed, for either of the EKR's the locus of the singularity class 1.2.3 (for the Lie square) is $\{x_3 = x_4 = y_4 = 0\}$. This set has, therefore, to be preserved by Φ. Consequently,

$$
Y_4 \in (x_3,\, x_4,\, y_4)\,,
$$

the ideal of functions' germs generated by the listed generators. Thus the first summand on the LHS of (26) vanishes. Passing to the second one, let us call simply Z the vector field in (25) to which $d\Phi$ is being applied. Then the 6-th row in (25) says that

$$
fY_4 \;=\; ZX_3 \;=\; Z(x_3H) \;=\; y_4\,H + x_3\,ZH \;\in\; (x_3,\, y_4)\,.
$$

Therefore, $\frac{\partial(fY_4)}{\partial x_4}\,|_0 = 0$, implying the vanishing of the second summand on the LHS in (26). In the transverse case the non-zero values of C are not equivalent to the zero value. At the same time, the non-zero values of C are readily normalizable to 1, and so the list of local models for 1.2.3.1 reads, formally as for 1.2.2.1,

- $B = 1, \quad C = 1$ — transverse generic,
- $B = 1, \quad C = 0$ — transverse atypical,
- $B = 0, \quad C = 1$ — tangent to '1.2.$\underline{2}$', but not tangent to '1.2.3',
- $B = 0, \quad C = 0$ — tangent to both '1.2.$\underline{2}$' and '1.2.3', or: strongly nilpotent.

As regards the class 1.2.1.2, it is reasonable to split the analysis into two cases. Either

• the square of a distribution – the suspension of a 1.2.1 germ – is *tangent* at the reference point to the locus of the singularity 1.2.1, or else

•• the square of a distribution is *transverse* at the reference point to the locus of the singularity 1.2.1.

Surprisingly, it is the • case that is easy. Indeed, by our earlier Theorem 4 (in its part concerning 1.2.1), the first three pairs of equations are then simplified as follows,

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - x_4dx_2 &= 0 & dy_3 - (\epsilon + y_4)dx_2 &= 0\,, \end{aligned} \tag{27}$$

with ϵ being either 1 (when the square is not strongly nilpotent) or 0 (the square strongly nilpotent), while the last pair

$$dx_2 - x_5dx_4 = 0 \qquad dy_4 - (c + y_5)dx_4 = 0$$

is open to further simplification. We mean the standard way $y_4 = \overline{y}_4 + cx_4,\ y_3 = \overline{y}_3 + cx_3$, $y_2 = \overline{y}_2 + ct,\ y_1 = \overline{y}_1 + \frac{c}{2}t^2$. This transformation, irrespectively of the value of ϵ, annihilates the constant c, because in the Pfaffian system (27) there hold the simplifying relations $x_4dx_2 = dx_3$ and $x_3dx_2 = dt$. Therefore, the • case represents but two orbits:
$1.2.1_{-\mathrm{s,\,tan}}.2$ written down as

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - x_4dx_2 &= 0 & dy_3 - (1 + y_4)dx_2 &= 0 \\ dx_2 - x_5dx_4 &= 0 & dy_4 - y_5dx_4 &= 0\,, \end{aligned}$$

and the part, $1.2.1.2_{+\mathrm{s}}$, that is strongly nilpotent in 1.2.1.2,

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - x_4dx_2 &= 0 & dy_3 - y_4dx_2 &= 0 \\ dx_2 - x_5dx_4 &= 0 & dy_4 - y_5dx_4 &= 0\,, \end{aligned}$$

As regards the •• case, by Theorem 4 for 1.2.1 again, the first three pairs of equations can be simplified to

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - (1 + x_4)dx_2 &= 0 & dy_3 - y_4dx_2 &= 0\,, \end{aligned} \tag{28}$$

while the last pair is, for the moment, general

$$dx_2 - x_5dx_4 = 0 \qquad dy_4 - (C + y_5)dx_4 = 0\,. \tag{29}$$

We will show that the two situations $C = 0$ and $C \neq 0$ in (28) – (29) are non-equivalent. To this end, we suppose the existence of a local conjugating diffeomorphism

$$\Phi \;=\; (T,\, X_1,\, Y_1,\, X_2,\, Y_2,\, X_3,\, Y_3,\, X_4,\, Y_4,\, X_5,\, Y_5) :\; (\mathbb{R}^{11},\, 0) \hookleftarrow$$

sending the object with the zero constant to an object displaying a value C:

$$d\Phi(p) \begin{array}{c} x_5 \begin{pmatrix} x_3 \begin{pmatrix} 1 \\ x_2 \\ y_2 \end{pmatrix} \\ 1 \\ y_3 \\ 1 \\ y_4 \end{pmatrix} \\ 1 \\ y_5 \\ 0 \\ 0 \end{array} \;=\; f(p) \begin{array}{c} \begin{pmatrix} x_5 G \begin{pmatrix} x_3 H \begin{pmatrix} 1 \\ X_2 \\ Y_2 \end{pmatrix} \\ 1 \\ Y_3 \\ 1 \\ Y_4 \end{pmatrix} \\ 1 \\ C + Y_5 \end{pmatrix} \\ * \\ * \end{array}\,, \tag{30}$$

where the $*$'s are certain functions; p stands, as usual, for $(t,\, x_1,\, y_1, \ldots,\, x_5,\, y_5)$, and $f,\, G,\, H$ are *invertible* function germs. (This time Φ preserves the loci of materialization of the sandwich class $1.\underline{2}.1\underline{2}$, implying $X_5 = G x_5$ and $X_3 = H x_3$.) Two basic consequences of (30) are

$$C\, f\,|_0 \;=\; \frac{\partial Y_4}{\partial x_4}\,|_0 \tag{31}$$

and

$$fG\frac{\partial Y_4}{\partial x_4}\,|_0 \;=\; \frac{fGY_4}{\partial x_4}\,|_0 \;=\; \frac{\partial Y_3}{\partial x_3}\,|_0\,, \tag{32}$$

the latter implied by a direct expression for the function fGY_4 that is encapsulated in (30). Thus the properties of Y_3 are getting important. In this respect, the (important) normalization to 0, in both germs conjugated by Φ, of the additive constant standing next to x_4 implies

$$\frac{\partial Y_3}{\partial x_2} + \frac{\partial Y_3}{\partial x_3}\,|_0 \;=\; 0 \tag{33}$$

On the other hand, there simply holds

Lemma 1. $\frac{\partial Y_3}{\partial x_2}\,|_0 = 0\,.$

Proof. Expressing in (30) the function fGY_3 via Y_2, one gets two informations. The first is

$$\frac{\partial Y_2}{\partial x_2}\,|_0 = \;0\,, \tag{34}$$

while the second is

$$fG\frac{\partial Y_3}{\partial x_2}\,|_0 \;=\; \frac{fGY_3}{\partial x_2}\,|_0 \;=\; \frac{\partial^2 Y_2}{\partial x_2^2}\,|_0\;. \tag{35}$$

But (30) allows also to express the function $fGHY_2$ via Y_1, and that expansion is clearly affine in x_2. Hence $\frac{\partial^2(fGHY_2)}{\partial x_2^2} = 0$ identically. Evaluating this at 0,

$$0 \;=\; 2\frac{\partial(fGH)}{\partial x_2}\frac{\partial Y_2}{\partial x_2} + fGH\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0 \;=\; fGH\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0 \tag{36}$$

by (34). Hence $\frac{\partial^2 Y_2}{\partial x_2^2}\,|_0 = 0$, whence $\frac{\partial Y_3}{\partial x_2}\,|_0 = 0$ by (35). Lemma is proved.

In view of Lemma 1, $\frac{\partial Y_3}{\partial x_3}\,|_0 = 0$ by (33), and so $\frac{\partial Y_4}{\partial x_4}\,|_0 = 0$ by (32). Now $C = 0$ by (31).

On the other hand, it is elementary to normalize a non-zero value C in (29) to 1. Summarizing, in the •• case the germs are either equivalent to

$$\begin{aligned} dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0\\ dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\ dx_3 - (1 + x_4)dx_2 &= 0 & dy_3 - y_4 dx_2 &= 0\\ dx_2 - x_5 dx_4 &= 0 & dy_4 - (1 + y_5)dx_4 &= 0 \end{aligned} \tag{37}$$

or else to

$$\begin{aligned} dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0\\ dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\ dx_3 - (1 + x_4)dx_2 &= 0 & dy_3 - y_4 dx_2 &= 0\\ dx_2 - x_5 dx_4 &= 0 & dy_4 - y_5 dx_4 &= 0\,. \end{aligned} \tag{38}$$

6 The most involved class 1.2.1.1

We strive, endly, to classify the class 1.2.1.1 and start from an obvious (and rough) pseudo-normal form subsuming this entire class,

$$\begin{aligned} dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0\\ dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\ dx_3 - (b + x_4)dx_2 &= 0 & dy_3 - (c + y_4)dx_2 &= 0\\ dx_4 - (B + x_5)dx_2 &= 0 & dy_4 - (C + y_5)dx_2 &= 0\,. \end{aligned} \tag{39}$$

The first question is that concerning the strong nilpotency, and for strong nilpotency the small growth vectors are important. After not so hard computations,

Observation 4. *The small growth vector at $0 \in \mathbb{R}^{11}$ of an object (39) is*

[3, 5, 7, 9, 10, 11]	when $(b,\, c) \neq (0,\, 0)$,
[3, 5, 7, 9, 10_2, 11]	when $(b,\, c) = (0,\, 0)$ and $(B,\, C) \neq (0,\, 0)$,
[3, 5, 7, 9, 10_3, 11]	when $(b,\, c) = (B,\, C) = (0,\, 0)$.

Notation. The three disjoint parts of 1.2.1.1 emerging from this observation are denoted, respectively (for momentary need), by 10_1, 10_2, and 10_3.

Proposition 4. *The part* 10_3 *entirely consists of strongly nilpotent germs. The parts* 10_1 *and* 10_2 *contain only not strongly nilpotent germs of* 2*-flags.*

The idea of proof is the same as in Chapters 2 and 4, and we skip here all details. Instead of 10_3, one could write, then, $1.2.1.1_{+\mathrm{s}}$ — the family of all strongly nilpotent distributions in the singularity class 1.2.1.1.

On the other hand, considering the Lie squares of the germs in 1.2.1.1 (that, after factoring out by their Cauchy characteristics, sit in the class 1.2.1), one can, with some abuse of notation, partition

$$1.2.1.1 \;=\; \underbrace{1.2.1_{-\mathrm{s,\,tra}} \cup 1.2.1_{-\mathrm{s,\,tan}}}_{10_1} \cup \underbrace{1.2.1_{+\mathrm{s}}}_{10_2 \cup 10_3}\,. \tag{40}$$

Transvecting the introduced two partitions, one gets a finer partition

$$1.2.1.1 \;=\; 1.2.1_{-\mathrm{s,\,tra}} \cup 1.2.1_{-\mathrm{s,\,tan}} \;\cup\; 1.2.1_{+\mathrm{s}} \cap 10_2 \;\cup\; 1.2.1_{+\mathrm{s}} \cap 10_3\,.$$

Thus (still abusing notation for brevity) there are already *four* disjoint invariant parts

- $1.2.1_{-\mathrm{s,\,tra}} = 1.2.1_{-\mathrm{s,\,tra}} \cap 10_1\,,$
- $1.2.1_{-\mathrm{s,\,tan}} = 1.2.1_{-\mathrm{s,\,tan}} \cap 10_1\,,$
- $1.2.1_{+\mathrm{s}} \cap 10_2 = 10_2\,,$
- $1.2.1_{+\mathrm{s}} \cap 10_3 = 10_3 = 1.2.1.1_{+\mathrm{s}}\,.$

Are these just orbits of the local classification? It will eventually turn out that only the first and the last part on the list are.

To see it, we start by partitioning the second item according to the position (at the reference point) of the distribution D in question, with respect to the locus of the singularity $1.2.1_{-\mathrm{s,\,tan}}$. We denote by $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tra}}$ the germs D that are relatively (i. e., within the locus of the sandwich geometry $1.\underline{2}.1.1$) *transverse* to this locus, and by $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tan}}$ those that are *tangent* to it.

We continue by similarly partitioning the third item, even though the process is now more delicate. Namely, this time one will check the position of D with respect to the locus of an aggregated singularity

$$1.2.1_{-\mathrm{s,\,tan}} \cup 1.2.1_{+\mathrm{s}} \stackrel{\mathrm{def}}{=} 1.2.1_{\mathrm{tan}}$$

(that, in each its materialization, is still smooth, not stratified, and in any EKR coordinates for D sitting in the third item, has local equations $x_3 = x_4 = 0$). We denote by $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tra}}$ the germs that are *relatively transverse* to the locus of $1.2.1_{\mathrm{tan}}$, whereas by $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tan}}$ all those that are *tangent* to that locus.

With these (prompting by themselves) definitions taken into account, our list of invariant parts of 1.2.1.1 lengthens to *six* items:

- $1.2.1_{-\mathrm{s,\,tra}}$,
- $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tra}}$,
- $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tan}}$,
- $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tra}}$,
- $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tan}}$,
- $1.2.1.1_{+\mathrm{s}}$.

Theorem 8. *The singularity class* 1.2.1.1 *splits into six orbits of the local classification. These orbits are listed above this theorem. The codimensions and local models can be read off from the proofs.*

7 Proof of Theorem 8

We will address separately every one part on the list; the proofs for the first and third part will be quite involved.

7.1 The orbit $1.2.1_{-\mathrm{s,\,tra}}$ of codimension one.

The only generic orbit within 1.2.1.1 is the first item on the list, $1.2.1_{-\mathrm{s,\,tra}}$. (Reiterating, this symbol should be understood in the sense that checking the inclusion of a germ D in this part deals only with the 'shorter' object $[D,\,D]/L([D,\,D])$.) A proof that it is indeed an orbit is not short.

As the reader already knows (Theorem 4), the germs of special 2-flags sitting in the discussed part can be brought to the following pseudo-normal form

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - (1+x_4)dx_2 &= 0 & dy_3 - y_4dx_2 &= 0 \\ dx_4 - (B+x_5)dx_2 &= 0 & dy_4 - (C+y_5)dx_2 &= 0\,, \end{aligned} \tag{41}$$

and the issue is to reduce to zero the constants B and C. This will be done simultaneously, if starting for clarity from B. To that end, we propose to consider an artificially chosen subsystem – the left tower in (41). That is,

$$\begin{aligned} dX_1 - X_2dT &= 0 \\ dT - X_3dX_2 &= 0 \\ dX_3 - (1+X_4)dX_2 &= 0 \\ dX_4 - (B+X_5)dX_2 &= 0 \end{aligned} \tag{42}$$

(we write capital letters because are going to make a substitution in (42)). This is a Goursat system living in the space $\mathbb{R}^6(T,\,X_1,\dots,\,X_5)$. Although it has no singularities, the question of possible elimination of B in it *formally* resembles the setting in the proof of Theorem 17 in [ChM]. Therefore, we just adapt (with a shift in indices) the formulas derived there on pages 147-8:

- $T = t, \qquad X_1 = -\frac{B}{6}t^2 + x_1, \qquad X_2 = -\frac{B}{3}t + x_2$
- $X_3 = \frac{x_3}{1-\frac{B}{3}x_3}, \qquad X_4 = \frac{1+x_4}{\left(1-\frac{B}{3}x_3\right)^3} - 1,$
- $X_5 = \frac{x_5}{\left(1-\frac{B}{3}x_3\right)^4} + \frac{B(1+x_4)^2}{\left(1-\frac{B}{3}x_3\right)^5} - B\,.$

The quickest way to check these is to evaluate $d(T, X_1, \dots,\, X_5)(t,\, x_1, \dots,\, x_5)$ on the vector field $[x_3,\ x_2x_3,\ 1,\ 1+x_4,\ x_5,\ 0]^{\mathrm{T}}$ and to get

$$\left(1-\frac{B}{3}x_3\right)\left[X_3,\ X_2X_3,\ 1,\ 1+X_4,\ B+X_5,\ 0\right]^{\mathrm{T}} + (*)\partial/\partial x_5$$

with a function $(*)$ whose properties are irrelevant. Continuing the proof of Proposition, we need to find $Y_1, \dots, Y_5, Y_j$ depending on $t, x_1, y_1, \dots,\, x_j, y_j$ $(j = 1, \dots,\, 5)$ that together with the already proposed $T,\, X_1, \dots,\, X_5$ are the components of a local diffeomorphism $\Phi = (T,\, X_1,\, Y_1, \dots,\, X_5,\, Y_5)$ that should conjugate (41) to another object of the type (41) with the model values $B = C = 0$.[6]

Precisely we **require** that

(‡) $d(T, X_1, Y_1 \dots,\, X_5, Y_5)(t,\, x_1,\, y_1, \dots,\, x_5,\, y_5)$ taken on the vector field

$$[x_3,\ x_2x_3,\ y_2x_3,\ 1,\ y_3,\ 1+x_4,\ y_4,\ x_5,\ y_5,\ 0,\ 0]^{\mathrm{T}}$$

be the multiplicative coefficient $\left(1-\frac{B}{3}x_3\right)$ times the vector field

$$[X_3,\ X_2X_3,\ Y_2X_3,\ 1,\ Y_3,\ 1+X_4,\ Y_4,\ B+X_5,\ C+Y_5,\ 0,\ 0]^{\mathrm{T}}$$

modulo $(\partial/\partial x_5,\ \partial/\partial y_5)$. (The coefficient $\left(1-\frac{B}{3}x_3\right)$ is prompted by the computations in [ChM].)

The main relation implied by the conjugacy (‡) is

$$\frac{\partial Y_4}{\partial x_2} + \frac{\partial Y_4}{\partial x_3}\,|_0 = C\,. \tag{43}$$

Under (‡), Y_4 gets expressed by Y_3, and, after a short calculus, (43) boils down to

$$\frac{\partial Y_3}{\partial t} + \frac{\partial^2 Y_3}{\partial x_2^2} + 2\frac{\partial^2 Y_3}{\partial x_2 \partial x_3} + \frac{\partial^2 Y_3}{\partial x_3^2}\,|_0 = C\,. \tag{44}$$

[6] Note that X_3 is, as it should be, a multiple of x_3, meaning preservation, by the sought diffeo Φ, of the set $\{F(\cdot) \supset D^2(\cdot)\}$ that is $\{x_3 = 0\}$ for both germs.

In turn, still under (‡), Y_3 gets expressed by Y_2,

$$x_3\frac{\partial Y_2}{\partial t} + x_2x_3\frac{\partial Y_2}{\partial x_1} + x_3y_2\frac{\partial Y_2}{\partial y_1} + \frac{\partial Y_2}{\partial x_2} + y_3\frac{\partial Y_2}{\partial y_2} \;=\; \Big(1 - \frac{B}{3}x_3\Big)Y_3\,, \tag{45}$$

showing under way that

$$\frac{\partial Y_2}{\partial x_2}\,|_0 = 0 \tag{46}$$

is a must in the problem. Under (‡), also Y_2 gets expressed by Y_1,

$$\frac{\partial Y_1}{\partial t} + x_2\frac{\partial Y_1}{\partial x_1} + y_2\frac{\partial Y_1}{\partial y_1} = Y_2\,, \tag{47}$$

which in turn implies another necessary condition

$$\frac{\partial Y_1}{\partial t}\,|_0 = 0\,. \tag{48}$$

Our objective is to write (44) in a *simpler* way and so get some hints concerning terms that are important in the expansion of Y_1. (The components T, X_1, Y_1 are the most important in Φ, as they entirely determine Φ. We know T and X_1, while Y_1 remains to be proposed.) Towards that aim, note that $\frac{\partial Y_3}{\partial x_3}\,|_0 = \frac{\partial Y_2}{\partial t}\,|_0 = \frac{\partial^2 Y_1}{\partial t^2}\,|_0$, by applying, consecutively, (45) and (47). Consequently – the key moment – we stipulate that

$$\frac{\partial^2 Y_1}{\partial t^2}\,|_0 = 0\,. \tag{49}$$

This clearly implies $\frac{\partial Y_3}{\partial x_3}\,|_0 = 0$. It also implies as if for free,

$$\frac{\partial Y_3}{\partial x_2}\,|_0 = 0 \tag{50}$$

(because, under (‡), $\frac{\partial Y_3}{\partial x_2} + \frac{\partial Y_3}{\partial x_3}\,|_0 = 0$). The reader may observe at this point that (49) and (‡) together are rather powerful.

Back in the main line of arguments, the LHS of (45) is an *affine* function in x_3, hence its second derivative with respect to x_3 vanishes identically. On the RHS of (45), it implies that

$$0 \;=\; -\frac{2B}{3}\frac{\partial Y_3}{\partial x_3} + \frac{\partial^2 Y_3}{\partial x_3^2}\,|_0 = \frac{\partial^2 Y_3}{\partial x_3^2}\,|_0\;\;.$$

It is also quick to infer from (45) that $\frac{\partial^2 Y_3}{\partial x_2^2}\,|_0 = \frac{\partial^3 Y_2}{\partial x_2^3}\,|_0 = 0$ (Y_2 is affine in x_2, compare (47)). All in all, under (49), the relation (44) assumes the form

$$\frac{\partial Y_3}{\partial t} + 2\frac{\partial^2 Y_3}{\partial x_2 \partial x_3}\,|_0 = C\,.$$

Expressing it in terms of Y_2, the first summand on the LHS is, by (45), equal to $\frac{\partial^2 Y_2}{\partial t \partial x_2}\,|_0$, while the second can be got via differentiating (45) sidewise with respect to x_2 and x_3,

$$\frac{\partial^2 Y_2}{\partial t \partial x_2} + \frac{\partial Y_2}{\partial x_1}\,|_0 \;=\; -\frac{B}{3}\frac{\partial Y_3}{\partial x_2} + \frac{\partial^2 Y_3}{\partial x_2 \partial x_3}\,|_0 \;=\; \frac{\partial^2 Y_3}{\partial x_2 \partial x_3}\,|_0\,,$$

with (50) accounting for the last equality. The basic relation (43) thus becomes

$$3\frac{\partial^2 Y_2}{\partial t \partial x_2} + 2\frac{\partial Y_2}{\partial x_1}\,|_0 = C\,. \tag{51}$$

Endly, (47) directly implies that $\frac{\partial Y_2}{\partial x_1}\,|_0 = \frac{\partial^2 Y_1}{\partial t \partial x_1}\,|_0$ and $\frac{\partial^2 Y_2}{\partial t \partial x_2}\,|_0 = \frac{\partial^2 Y_1}{\partial t \partial x_1}\,|_0$, reducing (51) to

$$5\frac{\partial^2 Y_1}{\partial t \partial x_1}\,|_0 = C\,, \tag{52}$$

provided that (‡), (46), (48) and (49) simultaneously hold.

The relation (52) is a mayor step in the proof, yet the formula $\frac{C}{5}tx_1$ alone would *not* do for the component Y_1, for one strives to construct a local *diffeomorphism* around $0 \in \mathbb{R}^{11}$. But it is safe to take $Y_1 = y_1 + \frac{C}{5}tx_1$ and, following (47), $Y_2 = y_2 + \frac{C}{5}x_1 + \frac{C}{5}tx_2$. The additional requirements (46), (48) and (49) clearly hold for these proposed functions, while the whole approach is so developed as to obey (‡). For reader's convenience, here are the formulas for the two next Y components. Y_3 is computed according to (45),

$$Y_3 \;=\; \left(1 - \frac{B}{3}x_3\right)^{-1}\left(y_3 + \frac{C}{5}t + \frac{2C}{5}x_2x_3\right)\,,$$

and Y_4 is – under (‡) – a precise product derived from Y_3,

$$\begin{aligned} Y_4 \;=\; & \left(1 - \frac{B}{3}x_3\right)^{-2}\left(y_4 + \frac{3C}{5}x_3 + \frac{2C}{5}x_2\big(1 + x_4\big)\right) \\ & + \frac{B}{3}\left(1 - \frac{B}{3}x_3\right)^{-3}\big(1 + x_4\big)\left(y_3 + \frac{C}{5}t + \frac{2C}{5}x_2x_3\right). \end{aligned} \tag{53}$$

As regards the last component Y_5, there is no need to compute it: in the output EKR, the additive constant standing next to Y_5 is that given by the basic relation (43). That is, C.[7]

The diffeomorphism Φ is now produced, and B, C can indeed be reduced to zero. □

[7] One also directly sees that the function (53) substituted on the LHS of (43) produces the value C.

7.2 The orbit $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tra}}$ of codimension two.

Any D from this part can, by Theorem 4 and an elementary rescaling, be written down under the pseudo-normal form

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - x_4dx_2 &= 0 & dy_3 - (1+y_4)dx_2 &= 0 \\ dx_4 - (1+x_5)dx_2 &= 0 & dy_4 - (C+y_5)dx_2 &= 0\,, \end{aligned}$$

with certain constant C. The aim is to eliminate this constant. One starts, no wonder, from $y_4 = \overline{y}_4 + Cx_4$ and computes $dy_4 - (C+y_5)dx_2 = d\overline{y}_4 + C\underline{dx_4} - (C+y_5)dx_2 = d\overline{y}_4 + C\underline{(1+x_5)dx_2} - (C+y_5)dx_2 = d\overline{y}_4 - (y_5 - Cx_5)dx_2$, because $dx_4 = (1+x_5)dx_2$ in this pseudo-normal form. This prompts $y_5 = \overline{y}_5 + Cx_5$. Then, working still within the right tower, $dy_3 - (1+y_4)dx_2 = dy_3 - C\underline{x_4dx_2} - (1+\overline{y}_4)dx_2 = dy_3 - C\underline{dx_3} - (1+\overline{y}_4)dx_2$, because $x_4dx_2 = dx_3$ for this differential system. This prompts $y_3 = \overline{y}_3 + Cx_3$.
Similarly, upon substituting this expression for y_3 in $dy_2 - y_3dx_2$, one is led to write $y_2 = \overline{y}_2 + Ct$, then to substitute it to $dy_1 - y_2dt$, and eventually to write $y_1 = \overline{y}_1 + \frac{C}{2}t^2$. In the variables $t,\, x_1, \dots,\, x_4,\, \overline{y}_1, \dots,\, \overline{y}_4$ the constant C disappears. $\square$

7.3 The orbit $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tan}}$ of codimension three.

This time, an arbitrary D from the 'doubly tangent' family can be written under the form

$$\begin{aligned} dx_1 - x_2dt &= 0 & \qquad dy_1 - y_2dt &= 0 \\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0 \\ dx_3 - x_4dx_2 &= 0 & dy_3 - (1+y_4)dx_2 &= 0 \\ dx_4 - x_5dx_2 &= 0 & dy_4 - (C+y_5)dx_2 &= 0\,, \end{aligned}$$

with, again, a constant C that should be got rid of. We will effectively construct, giving detailed motivations first, new coordinates eating this C up. So searched is a local preserving $0 \in \mathbb{R}^{11}$ diffeo $\Phi = (T,\, X_1,\, Y_1, \dots,\, X_5,\, Y_5)$ sending the EKR with $C = 0$ to the one with any fixed value of C.

That is, we demand this time that

(††) $d(T, X_1, Y_1 \dots,\, X_5, Y_5)(t,\, x_1,\, y_1, \dots,\, x_5,\, y_5)$ taken on the vector field

$$[x_3,\; x_2x_3,\; y_2x_3,\; 1,\; y_3,\; x_4,\; 1+y_4,\; x_5,\; y_5,\; 0,\; 0]^{\mathrm{T}}$$

be a function coefficient f times the vector field

$$[X_3,\; X_2X_3,\; Y_2X_3,\; 1,\; Y_3,\; X_4,\; 1+Y_4,\; X_5,\; C+Y_5,\; 0,\; 0]^{\mathrm{T}}$$

modulo $(\partial/\partial x_5,\, \partial/\partial y_5)$, with $f\,|_0 \neq 0$. Note that f is not precised yet (in contrast to the treatment of the generic case) and will get concretized only at the end. Let us stipulate additionally that $f\,|_0 = 1$. Then the basic relation reads

$$\frac{\partial Y_4}{\partial x_2} + \frac{\partial Y_4}{\partial y_3}\,|_0 = C\,, \tag{54}$$

while (††) implies

$$x_3\frac{\partial Y_3}{\partial t} + x_2x_3\frac{\partial Y_3}{\partial x_1} + y_2x_3\frac{\partial Y_3}{\partial y_1} + \frac{\partial Y_3}{\partial x_2} + y_3\frac{\partial Y_3}{\partial y_2} + x_4\frac{\partial Y_3}{\partial x_3} + (1+y_4)\frac{\partial Y_3}{\partial y_3} \;=\; f\big(1+Y_4\big)\,. \tag{55}$$

This relation allows to reduce (54) to

$$-\frac{\partial f}{\partial x_2} - \frac{\partial f}{\partial y_3} + \frac{\partial Y_3}{\partial y_2} + \frac{\partial^2 Y_3}{\partial x_2^2} + 2\frac{\partial^2 Y_3}{\partial x_2 \partial y_3} + \frac{\partial^2 Y_3}{\partial y_3^2}\,|_0 = C\,. \tag{56}$$

But (††) implies also

$$x_3\frac{\partial Y_2}{\partial t} + x_2x_3\frac{\partial Y_2}{\partial x_1} + y_2x_3\frac{\partial Y_2}{\partial y_1} + \frac{\partial Y_2}{\partial x_2} + y_3\frac{\partial Y_2}{\partial y_2} \;=\; fY_3$$

which helps to further reduce (56). Namely, after careful computations that we skip here, that relation boils down to

$$\frac{\partial^3 Y_2}{\partial x_2^3} + 3\frac{\partial^2 Y_2}{\partial x_2 \partial y_2} - \left(\frac{\partial f}{\partial x_2} + \frac{\partial f}{\partial y_3}\right)\left(1 + 2\frac{\partial Y_2}{\partial y_2} + 2\frac{\partial^2 Y_2}{\partial x_2^2}\right)|_0 \;=\; C\,. \tag{57}$$

Naturally, the objective is to descend further to indices 1 – to have only functions X_1, Y_1 in the conditions for a conjugacy. Note that (due to the inclusion holding true in the 2nd sandwich for both germs) the component X_3 is divisibe by x_3, $X_3 = x_3G$ for certain function G, $G\,|_0 \neq 0$. Now we stipulate anew that

$$fG = 1 \quad \text{identically} \tag{58}$$

(so that, with one previous assumption, $G\,|_0 = 1$). This and (††) yield a compact expression for Y_2 in terms of Y_1,

$$\frac{\partial Y_1}{\partial t} + x_2\frac{\partial Y_1}{\partial x_1} + y_2\frac{\partial Y_1}{\partial y_1} \;=\; Y_2\,.$$

With its use, (57) gets reduced to

$$-\left(\frac{\partial f}{\partial x_2} + \frac{\partial f}{\partial y_3}\right)\left(1 + 2\frac{\partial Y_1}{\partial y_1}\right)|_0 \;=\; C \tag{59}$$

which still leaves something to be desired. But also f is expressable, under (††), by the function X_2 alone:

$$x_3\frac{\partial X_2}{\partial t} + x_2x_3\frac{\partial X_2}{\partial x_1} + y_2x_3\frac{\partial X_2}{\partial y_1} + \frac{\partial X_2}{\partial x_2} + y_3\frac{\partial X_2}{\partial y_2} \;=\; f\,.$$

On top of this, all the time under (††) and (58),

$$\frac{\partial X_1}{\partial t} + x_2\frac{\partial X_1}{\partial x_1} + y_2\frac{\partial X_1}{\partial y_1} \;=\; X_2\,.$$

These premises suffice to reduce (59) ultimately to

$$-\frac{\partial X_1}{\partial y_1}\left(1 + 2\frac{\partial Y_1}{\partial y_1}\right)\big|_0 \;=\; C\,. \tag{60}$$

This is a tremendous prompt and we are now about to finish.

Indeed, one can take, simply, $T = t$, $X_1 = x_1 - \frac{C}{3}y_1$, $Y_1 = y_1$, thus securing (60). Let us write down the remaining components, just going backwards along the presented line of arguments. Immediately we get $X_2 = x_2 - \frac{C}{3}y_2$, $Y_2 = y_2$, and X_2 determines $f = 1 - \frac{C}{3}y_3$, which in turn determines $X_3 = x_3\big(1 - \frac{C}{3}y_3\big)^{-1}$. In parallel, (††) determines $Y_3 = y_3\big(1 - \frac{C}{3}y_3\big)^{-1}$, as well as

$$X_4 \;=\; x_4\Big(1 - \frac{C}{3}y_3\Big)^{-2} + \frac{C}{3}x_3\big(1 + y_4\big)\Big(1 - \frac{C}{3}y_3\Big)^{-3}.$$

Now (55) quickly generates the key component Y_4,

$$Y_4 \;=\; \big(1 + y_4\big)\Big(1 - \frac{C}{3}\Big)^{-3} - 1$$

which clearly satisfies (54). The proof is finished; there is no need to compute explicitly X_5, Y_5. Only as a matter of record, we note that, not surprisingly within $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tan}}$, $X_4 \in (x_3,\, x_4)$ (which is visible in the formula above) and $X_5 \in (x_3,\, x_4,\, x_5)$. □

Remark 5. It is precisely in this part of the singularity class 1.2.1.1 where we have detected an unexpected loss of stability under passing from $k = 2$ to $k = 3$; see Section 7.7 for the details and interplay with the theory of singularities of curves. (Note that another, perfectly natural loss of stability is caused by the appearance of the new singularity class 1.2.3.4 for $k \geq 3$, cf. [M3].)

7.4 The orbit $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tra}}$ of codimension three.

To justify its being an orbit, there suffices just a repetition of the argument from the proof in Section 7.2. Indeed, when dealing with the preliminary normal form

$$\begin{aligned} dx_1 - x_2dt &= 0 & dy_1 - y_2dt &= 0\\ dt - x_3dx_2 &= 0 & dy_2 - y_3dx_2 &= 0\\ dx_3 - x_4dx_2 &= 0 & dy_3 - y_4dx_2 &= 0\\ dx_4 - (1 + x_5)dx_2 &= 0 & dy_4 - (C + y_5)dx_2 &= 0\,, \end{aligned}$$

and trying to eliminate the constant C, one performs the same transformations and uses virtually the same bar variables y as for the part $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tra}}$. □

7.5 The orbit $1.2.1_{+\mathrm{s}}.1_{-\mathrm{s,\,tan}}$ of codimension four.

In view of Proposition 4, it is immediate to see that all such distribution germs are equivalent to the EKR

$$
\begin{aligned}
dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0\\
dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\
dx_3 - x_4 dx_2 &= 0 & dy_3 - y_4 dx_2 &= 0\\
dx_4 - x_5 dx_2 &= 0 & dy_4 - (1 + y_5) dx_2 &= 0\,.
\end{aligned}
$$

□

7.6 The orbit $1.2.1.1_{+\mathrm{s}}$ of codimension five.

The only EKR that has remained unused until this moment, and that services all strongly nilpotent germs in 1.2.1.1 (Proposition 4 again) is

$$
\begin{aligned}
dx_1 - x_2 dt &= 0 & dy_1 - y_2 dt &= 0\\
dt - x_3 dx_2 &= 0 & dy_2 - y_3 dx_2 &= 0\\
dx_3 - x_4 dx_2 &= 0 & dy_3 - y_4 dx_2 &= 0\\
dx_4 - x_5 dx_2 &= 0 & dy_4 - y_5 dx_2 &= 0\,.
\end{aligned}
$$

(In other words, within class 1.2.1.1 there holds the converse of the last item of Theorem 4 in [M2].) □

7.7 Loss of stability when the width grows.

The general ideology underlying the work on singularities of multi-flags is as follows. For any fixed k and r, there exists a huge 'monster' manifold M of dimension $(r+1)k+1$ and a *universal* rank-$(k+1)$ distribution $\mathcal{D}$ on M generating a special k-flag which realizes *all* possible local geometries of special k-flags of length r – see Remark 3 in [M2]. In that way the points of M correspond to 'all' germs of rank-$(k+1)$ distributions generating such flags. In fact, the couple $(M,\,\mathcal{D})$ is the outcome of a series of r so-called generalized Cartan prolongations (or rank-1 prolongations in the language of [SY]) started from $(\mathbb{R}^{k+1},\,T\mathbb{R}^{k+1})$. In parallel, smooth curves in $\mathbb{R}^{k+1}$ can also be Cartan-prolonged; their r-th prolongations lie in M.

We want to give an example of prolongation of curves for $k=2$ and $r=4$. It will be in close relation with the orbit $1.2.1_{-\mathrm{s,\,tan}}.1_{-\mathrm{s,\,tan}}$ discussed in Section 7.3. Let us take the curve $\gamma(s) = (t,\,x_1,\,y_1)(s) = (s^4,\,s^5,\,s^6)$ that is excerpted from the list [GHo] of simple *space* curves. We compute its first prolongation,

$$
x_2 = \frac{dx_1}{dt} = \frac{5}{4}s\,, \qquad y_2 = \frac{dy_1}{dt} = \frac{3}{2}s^2\,,
$$

then second prolongation

$$x_3 = \frac{dt}{dx_2} = \frac{16}{5}s^3\,, \qquad y_3 = \frac{dy_2}{dx_2} = \frac{12}{5}s\,,$$

and then third

$$x_4 = \frac{dx_3}{dx_2} = \frac{192}{25}s^2\,, \qquad \frac{dy_3}{dx_2} = \frac{48}{25}\,.$$

These results show that the third prolongation of γ hits at $s = 0$ the point-germ, on the relevant three-step monster manifold, with the additive constant $\frac{48}{25}$ standing next to y_4, and that y_4 is identically zero on the prolonged curve. (The use of EKR's in this discussion is equivalent to taking a good coordinate chart in a piece of the monster.) Consequently, $y_5 = \frac{dy_4}{dx_2} = 0$ in the fourth prolongation, while $x_5 = \frac{dx_4}{dx_2} = \frac{1536}{125}s$. Indeed then, the fourth prolongation of γ hits a germ in the orbit in question. That is, the model EKR with constants 1 (next to y_4) and $C = 0$ is being hit by the fourth prolongation of the curve $\left(s^4,\, s^5,\, \frac{25}{48}s^6\right)$.

When one enlarges the underlying space from three to four dimensions, the curve $(s^4,\, s^5,\, s^6)$ gets suspended to $\widetilde{\gamma}(s) = (s^4,\, s^5,\, s^6,\, 0)$ and keeps being simple. Yet its orbit becomes adjacent to the orbit of a less singular, also simple curve $\overline{\gamma}(s) = (s^4,\, s^5,\, s^6,\, s^7)$; compare in [Ar] the lists of sporadic simple curves in dimension 4. Hence one gets two closely related, if non-equivalent, curves $\overline{\gamma}$ and $\widetilde{\gamma}$. The fourth prolongation of $\widetilde{\gamma}$ hits at $s = 0$ the EKR (61), given below, with $D = 0$. Whereas the fourth prolongation of $\overline{\gamma}$ hits the member of (61) with $D = \frac{672}{125}$.

$$\begin{array}{rrr} dx_1 - x_2dt = 0 & dy_1 - y_2dt = 0 & dz_1 - z_2dt = 0 \\ dt - x_3dx_2 = 0 & dy_2 - y_3dx_2 = 0 & dz_2 - z_3dx_2 = 0 \\ dx_3 - x_4dx_2 = 0 & dy_3 - (1 + y_4)dx_2 = 0 & dz_3 - z_4dx_2 = 0 \\ dx_4 - x_5dx_2 = 0 & dy_4 - y_5dx_2 = 0 & dz_4 - (D + z_5)dx_2 = 0 \end{array} \qquad (61)$$

From this non-equivalence of 4-dimensional curve germs one *cannot* automatically deduce that the respective EKR objects (61) are non-equivalent. Yet, surprisingly in the optics of special 2-flags, the constant $D \neq 0$ in the EKR family (61) cannot be reduced to 0, indeed. It either vanishes or can be normalized to 1. This means that a single orbit in width 2, in width 3 consists of two different orbits. In other words, it splits up into two orbits when the width grows from 2 to 3. Thus, in width three, the class 1.2.1.1 splits up into at least *seven* orbits of the local classification! Reiterating, a proof of this loss of stability phenomenon does not follow from the curves' classification in [Ar]. It exceeds the scope of the present work and will be produced in [MPe2].

Attempting right now at a (tentative) conclusion, non-equivalences in the world of curves may firmly suggest probable non-equivalences of germs – points of the monster that are hit by Cartan prolongations of curves. It was not so in the case of [MPe2]. Had we noticed, however, the pertinent sporadic curves in [Ar] earlier, we would have worked towards the non-equivalence of $D = 0$ and $D \neq 0$ in (61) in a more deterministic context.

8 Appendix

We want to show that for any two different values c and $\widetilde{c}$ the distributions (2) are non-equivalent. Suppose the existence of a diffeomorphism

$$\Phi \;=\; (T,\, X_1,\, Y_1,\, X_2,\, Y_2, \dots,\, X_8,\, Y_8):\; (\mathbb{R}^{17},\, 0) \hookleftarrow$$

conjugating these two objects. The aim is to show that $c = \widetilde{c}$. Clearly,

- $T,\, X_1,\, Y_1$ depend only on $t,\, x_1,\, y_1$,
- for $2 \le j \le 8$, functions $X_j,\, Y_j$ depend only on $t, x_1, y_1, x_2, y_2, \dots, x_j, y_j$.

In the discussed situation one knows more about the components $X_3,\; X_5$, and X_7:

- $X_3(t,\, x_1,\, y_1,\, x_2,\, y_2,\, x_3,\, y_3) = x_3 K(t,\, x_1,\, y_1,\, x_2,\, y_2,\, x_3,\, y_3)$,
- $X_5(t,\, x_1, \dots,\, y_5) = x_5 H(t,\, x_1, \dots,\, x_5,\, y_5)$,
- $X_7(t,\, x_1, \dots,\, y_7) = x_7 G(t,\, x_1, \dots,\, x_7,\, y_7)$

for certain invertible at 0 functions $G,\, H,\, K$. Moreover, the preservation of the distribution $(\partial/\partial x_8,\, \partial/\partial y_8)$ implies that there must exist an invertible at 0 function f, $f\,|_0 \neq 0$, such that

$$d\Phi(p) \begin{pmatrix} x_7 \begin{pmatrix} x_5 \begin{pmatrix} x_3 \begin{pmatrix} 1 \\ x_2 \\ y_2 \end{pmatrix} \\ 1 \\ y_3 \\ 1 + x_4 \\ y_4 \end{pmatrix} \\ 1 \\ y_5 \\ 1 + x_6 \\ y_6 \end{pmatrix} \\ 1 \\ y_7 \\ c + x_8 \\ y_8 \\ 0 \\ 0 \end{pmatrix} \;=\; f(p) \begin{pmatrix} x_7 G \begin{pmatrix} x_5 H \begin{pmatrix} x_3 K \begin{pmatrix} 1 \\ X_2 \\ Y_2 \end{pmatrix} \\ 1 \\ Y_3 \\ 1 + X_4 \\ Y_4 \end{pmatrix} \\ 1 \\ Y_5 \\ 1 + X_6 \\ Y_6 \end{pmatrix} \\ 1 \\ Y_7 \\ \widetilde{c} + X_8 \\ Y_8 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ * \\ * \end{matrix} \tag{62}$$

where $p = (t,\, x_1,\, y_1, \dots,\, x_8,\, y_8)$ and, for bigger transparence, the arguments in the functions $G,\, H,\, K,\, X_2, \dots,\, Y_8$ on the RHS are not written. This vector relation entails the set of 15 scalar equations on the consecutive components $\partial/\partial t,\, \partial/\partial x_1,\, \dots,\, \partial/\partial x_7,\, \partial/\partial y_7$; we disregard the two last components – the components in the directions of $L(D^6) \subset D^7$. In view of the first 11 components of Φ depending only, recalling, on $t,\, x_1$,...,

y_5, the upper 11 among these scalar equations can be divided sidewise by x_7. Likewise and additionally, the upper 7 among them can be divided by x_5, and the first three – additionally by x_3. Agree to call thus simplified equations 'level T', 'level X_1', 'level X_7', etc, in function of the row of $d\Phi(p)$ being involved. For instance, the level T equation is the $\partial/\partial t$–component scalar equation in (62) divided sidewise by the product $x_3x_5x_7$.

Because $\frac{\partial x_7}{\partial x_7}\,|_0 = G\,|_0$, it follows from the level X_7 that

$$cG\,|_0 \;=\; \tilde{c}f\,|_0\,, \tag{63}$$

while from the level X_6 one gets

$$f\,|_0 = \frac{\partial X_6}{\partial x_6}\,|_0\;. \tag{64}$$

In turn, the level X_5 can be written in a short form

$$(*)x_5 + \frac{\partial X_5}{\partial x_5}(1+x_6) + (*)y_6 \;=\; fG(1+X_6)\,, \tag{65}$$

and, additionally, the level X_4 is the defining equation for the factor fG on the RHS in (65). In particular that level shows that fG depends only on $t,\, x_1, \dots,\, y_5$. Hence fG, as well as X_5, do not depend on x_6, and, moreover, $\frac{\partial X_5}{\partial x_5}\,|_0 = H\,|_0$. Now it is very quick to differentiate (65) with respect to x_6 at 0:

$$H\,|_0 = fG\frac{\partial X_6}{\partial x_6}\,|_0\;. \tag{66}$$

One is already half way through because, upon evaluating (65) at 0,

$$H\,|_0 = fG\,|_0 \tag{67}$$

and this quantity is clearly non-zero. So (67), (66), (64) together imply

$$f\,|_0 = 1\,. \tag{68}$$

At this point the reader may feel already that, with one more constant 1 standing next to x_4, this line of arguments can be *repeated*, with f replaced by fG and X_6 replaced by X_4. It is indeed the case (and simultaneously a kind of explanation that, for *this* type of argumentation, needed is nothing shorter than the class 1.2.1.2.1.2.1). To conclude the justification of a modulus, we are going to just write a sequence of relations holding true, with only short indications of sources for them.

$$fG\,|_0 = \frac{\partial X_4}{\partial x_4}\,|_0 \qquad \text{from the level } X_4\,,$$

$$(*)x_3 + \frac{\partial X_3}{\partial x_3}(1+x_4) + (*)y_4 \;=\; fGH(1+X_4) \quad (\text{the level } X_3)\,,$$

fGH depends only on t, $x_1, \ldots, y_3$ (the level X_2) and $\dfrac{\partial X_3}{\partial x_3} \mid_0 = K \mid_0$,

$$K \mid_0 = fGH \frac{\partial X_4}{\partial x_4} \mid_0 \quad \text{(differentiating the level } X_3 \text{ w.r.t. } x_4\text{)},$$

$$K \mid_0 = fGH \mid_0 \neq 0 \qquad \text{evaluating the level } X_3 \text{ at } 0\,.$$

$$fG \mid_0 = 1 \qquad \text{following from all the above facts}\,.$$

This last relation together with (68) say that $f \mid_0 = G \mid_0 = 1$. Now (63) boils down to $c = \widetilde{c}$. The invariant character of the parameter c in (2) is shown.

Remark 6. Note that an analogous proof in the space of 1-flags would be false. For, in the Goursat case, there is no second sandwich, they only commence by No 3. So one could not claim (as is done above) that the function X_3 is divisible by x_3. And, besides, it is well known that in length seven the local classification of Goursat is still discrete.

References

[Ar] V. I. Arnold; *Simple singularities of curves*, Proc. Steklov Inst. Math. **226** (1999), 20 – 28.

[Ad] J. Adachi; *Global stability of special multi-flags*. To appear in Israel J. Math. **179**.

[A] A. A. Agrachev; *Canonical nilpotentization of control systems*. Preprint (1997).

[AGau] ——, J–P. Gauthier; *On subanalyticity of Carnot–Carathéodory distances*, Ann. Inst. H. Poincaré – AN **18** (2001), 359 – 382.

[AGamSa] ——, R. V. Gamkrelidze, A. V. Sarychev; *Local invariants of smooth control systems*, Acta Appl. Math. **14** (1989), 191 – 237.

[ASa] ——, A. V. Sarychev; *Filtrations of a Lie algebra of vector fields and nilpotent approximations of control systems*, Dokl. Akad. Nauk SSSR vol. 295 (1987). English transl. in Soviet Math. Dokl. **36** (1988), 104 – 108.

[Be] A. Bellaïche; *The tangent space in sub-Riemannian geometry*, in: A. Bellaïche and J–J. Risler (eds.), Sub-Riemannian Geometry, Birkhäuser, Basel 1996, 1 – 78.

[BiSt] R. M. Bianchini, G. Stefani; *Graded approximations and controllability along a trajectory*, SIAM J. Control Optim. **28** (1990), 903 – 924.

[B] R. L. Bryant; *Some aspects of the local and global theory of pfaffian systems*. PhD Thesis, University of North Carolina, Chapel Hill, 1979.

[BH] ——, L. Hsu; *Rigidity of integral curves of rank* 2 *distributions*, Invent. math. **114** (1993), 435 – 461.

[C] E. Cartan; *Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre*, Ann. Ecole Normale **27** (1910), 109 – 192.

[ChM] M. Cheaito, P. Mormul; *Rank–*2 *distributions satisfying the Goursat condition: all their local models in dimension* 7 *and* 8, ESAIM Control Optim. Calc. Var. **4** (1999), 137 – 158 (electronic).

[GHo] C. G. Gibson, C. A. Hobbs; *Simple singularities of space curves*, Math. Proc. Cambridge Phil. Soc. **113** (1993), 297 – 310.

[J] F. Jean; *The car with n trailers: characterisation of the singular configurations*, ESAIM Control Optim. Calc. Var. **1** (1996), 241 – 266 (electronic).

[KRub] A. Kumpera, J. L. Rubin; *Multi-flag systems and ordinary differential equations*, Nagoya Math. J. **166** (2002), 1 – 27.

[KRui] A. Kumpera, C. Ruiz; *Sur l'équivalence locale des systèmes de Pfaff en drapeau*, in: F. Gherardelli (ed.), Monge-Ampère Equations and Related Topics, Ist. Naz. Alta Math. F. Severi, Rome, 1982, 201 – 248.

[MonZ] R. Montgomery, M. Zhitomirskii; *Geometric approach to Goursat flags*, Ann. Inst. H. Poincaré – AN **18** (2001), 459 – 493.

[M1] P. Mormul; *Goursat distributions not strongly nilpotent in dimensions not exceeding seven*, in: D. Ovens, A. Zinober (eds.), Nonlinear and Adaptive Control, LNCIS **281**, Springer 2003, 249 – 261.

[M2] ——; *Multi-dimensional Cartan prolongation and special k-flags*, in: H. Hironaka et al (eds.), Geometric Singularity Theory, Banach Center Publications **65**, Inst. of Math., Polish Acad. Sci., Warsaw 2004, 157 – 178. Also available at http://www.mimuw.edu.pl/∼mormul/special.pdf

[M3] ——; *Geometric singularity classes for special k-flags, $k \geq 2$, of arbitrary length.* Preprint in: Singularity Theory Seminar, S. Janeczko (ed.), Warsaw Univ. of Technology, Vol. 8 (2003), pp. 87 – 100. Also available at http://www.mimuw.edu.pl/∼mormul/Mor03.ps

[M4] ——; *Singularity classes of special* 2*-flags*, SIGMA **5** (2009), 102, 22 pages (electronic).

[MPe1] ——, F. Pelletier; *Contrôlabilité complète par courbes anormales par morceaux d'une distribution de rang* 3 *générique sur des variétés connexes de dimension* 5 *et* 6, Bull. Polish Acad. Sci., Math. **45** (1997), 399 – 418.

[MPe2] ——, ——; *The local classification of special multi-flags is not stable with respect to the width*. In preparation.

[PaR] W. Pasillas-Lépine, W. Respondek; *Contact systems and corank one involutive subdistributions*, Acta Appl. Math. **69** (2001), 105 – 128.

[SY] K. Shibuya, K. Yamaguchi; *Drapeau theorem for differential systems*. Diff. Geom. Appl. **27** (2009), 793 – 808.

[T] N. Tanaka; *On differential systems, graded Lie algebras and pseudogroups*, J. Math. Kyoto Univ. **10** (1970), 1 – 82.

[Y] K. Yamaguchi; *Geometrization of Jet bundles*, Hokkaido Math. J. **12** (1983), 27 – 40.