Singularity Classes of Special 2-Flags*

Piotr MORMUL

Institute of Mathematics, Warsaw University, 2 Banach Str., 02-097 Warsaw, Poland
E-mail: *mormul@mimuw.edu.pl*

Received April 16, 2009, in final form October 30, 2009; Published online November 13, 2009
doi:10.3842/SIGMA.2009.102

Abstract. In the paper we discuss certain classes of vector distributions in the tangent bundles to manifolds, obtained by series of applications of the so-called generalized Cartan prolongations (gCp). The classical Cartan prolongations deal with rank-2 distributions and are responsible for the appearance of the Goursat distributions. Similarly, the so-called special multi-flags are generated in the result of successive applications of gCp's. Singularities of such distributions turn out to be very rich, although without functional moduli of the local classification. The paper focuses on special 2-flags, obtained by sequences of gCp's applied to rank-3 distributions. A stratification of germs of special 2-flags of all lengths into *singularity classes* is constructed. This stratification provides invariant geometric significance to the vast family of local polynomial pseudo-normal forms for special 2-flags introduced earlier in [Mormul P., *Banach Center Publ.*, Vol. 65, Polish Acad. Sci., Warsaw, 2004, 157–178]. This is the main contribution of the present paper. The singularity classes endow those multi-parameter normal forms, which were obtained just as a by-product of sequences of gCp's, with a geometrical meaning.

Key words: generalized Cartan prolongation; special multi-flag; special 2-flag; singularity class

2000 Mathematics Subject Classification: 58A15; 58A17; 58A30

1 Introduction and main theorem

The aim of the current paper is to present a new and rather rich stratification of singularities of (special) 2-flags which naturally generalize 1-flags. Before doing that, it will be useful to briefly recall 1-flags and their singularities. These are, in the contemporary terminology, rank-2 and corank ≥ 2 subbundles $D \subset TM$ in the tangent bundle to a smooth manifold M, together with the tower of consecutive *Lie squares* $D \subset [D,D] \subset [[D,D],[D,D]] \subset \cdots$ satisfying the property that the linear dimensions of tower's members are $2, 3, 4, \dots$ at *every* point in M. (In $(\dim M - 2)$ steps the tower reaches the full tangent bundle TM.) These objects had emerged in the papers [7, 21, 6] and were later popularized in a book by Goursat in the 1920s. In the result, such distributions D are now called the *Goursat distributions*, or sometimes the *Cartan–Goursat distributions*. The respective flags are called the *Goursat flags*. Although this definition is quite restrictive, still such flags exist in all lengths. Indeed, for every $s \geq 2$, the canonical contact system $\mathcal{C}^s$ (the jet bundle or the Cartan distribution in the terminology of [9]) on the jet space $J^s(1,1)$ is a Goursat distribution of corank s; its flag has length s. However, each distribution $\mathcal{C}^s$ is homogeneous because its germs at every two points are equivalent by a local diffeomorphism of $J^s(1,1)$. Therefore, these contact systems have no singularities. It should be noted that nowadays the contact systems on $J^s(1,1)$ are *also* known under the name 'Goursat normal forms' and are characterized as such in [3] (Theorem 5.3 in Chapter II).

*This paper is a contribution to the Special Issue "Élie Cartan and Differential Geometry". The full collection is available at http://www.emis.de/journals/SIGMA/Cartan.html

For a very long time it had not been known whether Goursat flags locally featured any other geometry than that of the systems $\mathcal{C}^s$. An affirmative answer was given only in 1978 by Giaro, Kumpera and Ruiz in dimension 5, and slightly later, in [11], in all dimensions ≥ 5. Much later a geometric systematization of those findings appeared in [12]. Namely, Montgomery and Zhitomirskii defined *Kumpera–Ruiz classes* of germs of Goursat flags (KR classes for short) in every fixed length $s \geq 2$. The number of them in length s is 2^{s-2}. In fact, it is natural to encode those classes by words of length s over the alphabet $\{1, 2\}$. The words start with two 1's. In the i-th place, where $3 \leq i \leq s$, one writes 1 or 2 depending on whether the condition (GEN) from p. 466 in [12] holds true for that i or not. This specification of the way in which one puts the numbers is purely geometrical and means that either certain two (invariantly defined) lines in a plane, related with the corank-i member of the flag at the reference point, are different or merge into one line. Nearly immediately those classes appeared to perfectly match the 2^{s-2} branches in the tree of Kumpera–Ruiz [pseudo]normal forms for germs of Goursat distributions of corank s constructed in [11]. (Those were couples of *polynomial* vector fields with only finite number of *real parameters*. The construction of those fields had much in common with a KR class to which the relevant germ belonged.) It was a departure point for an entirely new, full-scale theory of Goursat flags developed in recent years.

Let us emphasize the key fact which has motivated the present article. The following two seemingly distant aspects of the theory are closely related:

- the local realizations, or KR normal forms constructed in [11], and
- the genuine KR classes of singularities defined in [12].

(The former preceded the latter by 18 years!)

Our objective is to establish an analogous, but going further, relationship for very natural generalizations of 1-flags, the so-called *special* 2-*flags.*

So, to begin with, what are special multi-flags? In the definition we will use the notion of the *Cauchy-characteristic module* (or, strictly speaking, sheaf of modules) of a distribution D, written $L(D)$ (the Japanese school adheres to the symbol $\mathrm{Ch}(D)$). It consists of all vector fields v (in the considered category of smoothness) taking values in D and preserving D: $[v, D] \subset D$.

Definition 1 (special m-flags). We fix a natural number $m \geq 2$ (called 'width'). A rank-$(m+1)$ distribution D on a manifold M generates a special m-flag of length $r \geq 1$ on M when

$\star$ the tower of consecutive Lie squares of D

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

$[D^j, D^j] = D^{j-1}$ for $j = r, r-1, \ldots, 2, 1$, consists of distributions of ranks, starting from the smallest object D^r: $m+1, 2m+1, \ldots, rm+1, (r+1)m+1 = \dim M$,

$\star\star$ for $j = 1, 2, \ldots, r-1$ 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} (i.e., such a module of vector fields has its linear dimension $\mathrm{rk}\, D^{j+1} - 1$ at every point), while $L(D^r) = 0$,

$\star\star\star$ the biggest flag's member D^1 possesses a corank-1 involutive (i.e., completely integrable) subdistribution, which we call F.

This definition is slightly more specific than the original definition from [14]. It is, however, equivalent, singling out precisely the same objects. It emphasizes the Cauchy-characteristic subflag; compare also the definition of 'generalized contact systems for curves' in [18]. On the other hand, Definition 1 is, as it stands, redundant, for condition $\star\star$ is implied by $\star$ and $\star\star\star$, see Proposition 1.3 in [1], or Corollary 6.3 in [19]. Thus the meaning of 'special' resides in the existence of an involutive corank-1 subdistribution $F \subset D^1$.[1] The involutiveness of F is critical; see Remark 2(a) in this respect.

Additional comment to Definition 1: condition $\star$ alone defines *general* m-flags, whose possible geometries are extremely rich, including functional moduli in the local classification, see for instance [5, 2, 24]. It is neatly outbalanced by conditions $\star\star$ and $\star\star\star$ (in fact, reiterating, the latter eventually covers the former).

As for $m = 1$, that time condition $\star\star$ is implied by just condition $\star$ and so 1-flags appear to be automatically special! (This is outside the scope of special multi-flags.)

There exist effective realization techniques producing distributions which generate special multi-flags of arbitrary width and length. For general $m \geq 2$ they have been constructed in Section 3.3 of [14] with an essential use of the so-called generalized Cartan prolongations; see in this respect Theorem 4 later on. (Specifically for $m = 2$, for reader's convenience, these operations are re-defined in Sections 3.2 and 3.3 of the present paper.) In the outcome one gets a vast family of polynomial [pseudo-]normal forms with many numerical parameters, the so-called Extended Kumpera–Ruiz normal forms (EKR for short; see Section 3.3 for the explanation of the origin of this name). Within a given EKR the realizations differ only by the values of the numeric parameters that enter that EKR. The classes EKR are encoded (or: labelled) by words $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}.\mathbf{j}_r$ over $\{\mathbf{1}, \mathbf{2}, \ldots, \mathbf{m}, \mathbf{m}+\mathbf{1}\}$ subject to an important limitation called *the least upward jumps rule*. Namely, admissible words should start with $\mathbf{1}$ and always a *new* but not yet used number should only *minimally* exceed the maximum of previously used numbers: for $l = 1, 2, \ldots, r-1$, if $\mathbf{j}_{l+1} > \max(\mathbf{j}_1, \ldots, \mathbf{j}_l)$ then $\mathbf{j}_{l+1} = 1 + \max(\mathbf{j}_1, \ldots, \mathbf{j}_l)$.

For $m = 2$ this rule says that, after starting from $\mathbf{1}$, the first use of $\mathbf{3}$ (if any) should occur after the first use of $\mathbf{2}$ (if any). That is to say, the number $\mathbf{3}$ does not appear without number $\mathbf{2}$ before it. It is straightforward to see (Proposition 3) that the number of EKR classes is equal to $\frac{1}{2}\left(1 + 3^{r-1}\right)$ in every length $r \geq 1$.

Instead of the Kumpera–Ruiz normal forms for Goursat flags, we now have EKR's to efficiently handle special multi-flags. Indeed, neat polynomial local realizations have been proposed in both settings $m = 1$ and $m \geq 2$. In the former case it is known (and already mentioned above) that the KR normal forms faithfully correspond to the KR classes of singularities put forward in [12]. Do, therefore, the EKR's in the latter case correspond to some partition or stratification of the space of all germs of special m-flags[2]? Or, in the least, what could be said specifically in width 2?

Our objective in the present paper is to answer the main question above affirmatively in width $m = 2$ for the rank-3 distributions generating special 2-flags of arbitrary length.

Firstly in Section 3.2 we construct an analogue of the KR classes of Goursat flags for special 2-flags, adapting the method of [12]. We call the obtained intermediate aggregates of germs of special 2-flags 'sandwich classes', because they directly emanate from the sandwich diagram for multi-flags (see Section 3.1)[3]. The sandwich classes are encoded by such words over $\{1, \underline{2}\}$ which start with 1 and are of length equal to flag's length. If the length is r, then the number of sandwich classes is 2^{r-1} (note the difference in exponent with the Goursat case, which is due to the presence of the distribution F in the sandwich diagram for multi-flags).

Secondly, we present the key part of the paper in Section 3.5. Namely, only in width 2, we refine the notion of the sandwich class to a 'singularity class'. In fact, a germ D sitting in a given sandwich class $\mathcal{S}$ of length r is being analyzed both geometrically and Lie-algebraically.

---

[1]Such F, when exists, is unique and has more than one geometrical interpretation (see Corollary 2 and Remarks 1 and 2(b) later on, and also [14, p. 165]). It is worth noting that when $m = 1$ such subdistributions F also exist but are *not* unique and have no geometrical meaning whatsoever (this concerns Goursat flags which are not considered in the present paper).

[2]This was B. Kruglikov's question asked in 2002.

[3]In this step the specification $m = 2$ is not important.

The purpose is to specify all but the first $\underline{2}$ (from the left) in the label of $\mathcal{S}$ to 2 or 3, each one independently of the others. It is the local geometry of the flag of D that decides that choice. (As for the first $\underline{2}$, it is invariably specified to 2.) In the outcome a word $j_1.j_2 \ldots j_r$ over $\{1,2,3\}$, denoted by $\mathcal{W}(D)$, is being associated to D. Given that the 'sandwich' words start with 1 and the first $\underline{2}$ in them is later specified to 2, it is clear that $\mathcal{W}(D)$ also satisfies the least upward jumps rule, exactly as the labels of the EKR classes have done. We mean that $j_1 = 1$ and if $j_{l+1} > \max(j_1,\ldots,j_l)$ then $j_{l+1} = 1 + \max(j_1,\ldots,j_l)$ for $l = 1,2,\ldots,r-1$.

Now the germs having one and the same word $\mathcal{W}(\cdot)$ build up a given *singularity class*. It follows that the partition of all germs into singularity classes is a refinement of the partition into sandwich classes, and that the cardinality of that new finer partition is the same as the cardinality of the EKRs in that length. That is, $\frac{1}{2}\left(1+3^{r-1}\right)$ in every length r (Proposition 3).

Does one know that all singularity classes are nonempty? More generally, is there a relationship among the singularity classes and the classes EKR of concrete realizations of special 2-flags in any given length r? Are singularity classes visible on the level of local polynomial pseudo-normal forms EKR? It turns out that the answer is 'yes' and the EKRs do not forget about the underlying local flag's geometry concretized by (or: discretized in) the singularity class. Namely, there holds

Theorem 1 (main theorem). *Let D be any germ of a rank-3 distribution generating a special 2-flag of length $r \geq 1$, belonging to a fixed singularity class $j_1.j_2\ldots j_r$ $(= \mathcal{W}(D))$. Then the EKR pseudo-normal forms of D are uniquely of the type $\mathbf{j}_1.\mathbf{j}_2\ldots\mathbf{j}_r$.*

Although elusive on the definition level, flag's local invariant – singularity class – acquires a concrete illustration in this theorem.

Corollary 1. *The singularity class of a germ of a special 2-flag which is already given in an EKR form $\mathbf{j}_1.\mathbf{j}_2\ldots\mathbf{j}_r$ is $j_1.j_2\ldots j_r$. That is, slightly abusing notation, $\mathcal{W}(\mathbf{j}_1.\mathbf{j}_2\ldots\mathbf{j}_r) = j_1.j_2\ldots j_r$.*

Therefore, Theorem 1 *additionally* shows that all singularity classes are non-empty. Whenever one finds an EKR for a germ of special 2-flag, one inevitably stumbles upon its singularity class. An illustrative example of retrieving the singularity classes from EKRs is given later in Appendix B. Theorem 1 is proved in Section 4.

2 Generalized Cartan prolongations produce special multi-flags

In differential geometry there exists an important operation, defined in the papers of É. Cartan and used by him in various situations. It takes rank-2 vector distributions in the tangent bundles to manifolds, processes them and yields more complicated rank-2 distributions, living on bigger manifolds, in the outcome. Nowadays it is called *Cartan prolongation* and can be applied to an arbitrary rank-2 distribution. In the modern language of [4, p. 454] its definition goes as follows.

'If $\mathcal{D}$ is a rank-2 distribution on a manifold M, then, regarding $\mathcal{D}$ as a vector bundle, we can certainly define its projectivization $\pi : \mathbb{P}\mathcal{D} \longrightarrow M$, which is a bundle over M whose typical fiber $\mathbb{P}\mathcal{D}_p$ is the space of 1-dimensional linear subspaces of the 2-dimensional vector space $\mathcal{D}_p$. Thus, the fibers of $\mathbb{P}\mathcal{D}$ are isomorphic to $\mathbb{PR}^1$ as projective 1-manifolds. There is a canonical rank-2 distribution $\mathcal{D}^{(1)}$ on $\mathbb{P}\mathcal{D}$ defined by setting $\mathcal{D}^{(1)}_\xi = (\pi')^{-1}(\xi)$ for each linear subspace $\xi \subset \mathcal{D}_p$. The distribution $\mathcal{D}^{(1)}$ is called the (*first*) *prolongation* of $\mathcal{D}$.'

Its importance stems from a key local *structural* theorem, presented (or, in authors' view, only recalled) in [4]. This theorem deals with rank-2 distributions that have mild properties of growing (in a Lie algebra sense which is explained below) neither too slowly nor too quickly. In fact, certain rank-2 distributions of corank, say, s locally turn out to be, up to the equivalence of the base manifolds, nothing but the Cartan prolongations of rank-2 distributions of corank $s-1$.

This shows that such distributions are constructed simpler than one could expect, and that they have some handy structure.

Below, $\mathcal{D}_1$ means the Lie square $[\mathcal{D}, \mathcal{D}]$ of a distribution $\mathcal{D}$, and $\mathcal{D}_2$ – the Lie square of $\mathcal{D}_1$. The foliation $\mathcal{F}$ is a classical object closely related to the hypothesis on the deficient rank of $\mathcal{D}_2$ (everywhere 4 instead of 5). In fact, under the hypotheses in Theorem 2, the Cauchy-characteristic module $L(\mathcal{D}_1)$ (for the definition of $L(\cdot)$, see the paragraph before Definition 1 earlier on) is a rank-1 subdistribution of $\mathcal{D}_1$ – a field of lines, and $\mathcal{F}$ is the 1-dimensional foliation tangent to $L(\mathcal{D}_1)$.

Theorem 2 (Cartan–Bryant–Hsu). *Let $\mathcal{D}$ be a rank-2 distribution on a manifold M^{s+2} and suppose that $\mathcal{D}_1$ and $\mathcal{D}_2$ have ranks 3 and 4 respectively. Furthermore, suppose that there is a submersion $f : M \to N^{s+1}$ with the property that the fibers of f are the leaves of the canonical foliation $\mathcal{F}$. Then there exists a unique rank-2 distribution $\mathcal{D}'$ on N with the property that $\mathcal{D}_1 = f^*(\mathcal{D}')$ and, moreover, there exists a canonical smooth map $f^{(1)} : M \longrightarrow \mathbb{P}\mathcal{D}'$ which is a local diffeomorphism, which satisfies $f = \pi \circ f^{(1)}$, and which satisfies $f^{(1)}_* \mathcal{D} = (\mathcal{D}')^{(1)}$.*

The ultimate consequence of this impressive theorem is a clear local construction of Goursat distributions. As simply as it can only be, Cartan prolongations applied in longer and longer successions produce (locally) all longer and longer Goursat flags! Montgomery and Zhitomirskii summarize the resulting situation in [12, p. 479] as follows: 'Every corank s Goursat germ can be found, up to a diffeomorphism, within the s-fold prolongation of the tangent bundle to a surface. We have called this s-fold prolongation the "monster manifold". It is a very tame monster in many respects.'

A recent big contribution [13] of the same authors demonstrates how eventually fruitful this Cartan-inspired visualisation of Goursat distributions is.

Returning to [special] multi-flags, an instance of vagueness shrouding them 10 years ago is the following. The first of the authors of [10] wrote, in a personal communication, in spring of 1999:

. . . multi-flags, they appear essentially as the usual flags. The usual flags translate, at least in the transitive case, the Cartan distribution on the jet space of a function of one variable. Multi-flags translate, in the transitive case, the same situation in the jet space of several functions of one variable. . . .

Therefore, a kind of multi-dimensional prolongation of distributions was badly needed. Various discussions around the results of [12] (existing then in a preprint form) and (drafts of) [10] remained inconclusive until the formulation of a general prolongation scheme.

One obtains the definition of *generalized Cartan prolongation* (gCp for short, p. 159 in [14]) by replacing in the definition from [4]: 'rank-2' by 'rank-$(m + 1)$', '2-dimensional' by '$(m + 1)$-dimensional', and '$\mathbb{PR}^1$' by '$\mathbb{PR}^m$'. While $\mathbb{P}D$ stands, as there, for the projectivization of the bundle $D \longrightarrow M$.

If D is a rank-$(m + 1)$ distribution on a manifold M, then, regarding D as a vector bundle, its projectivization $\pi : \mathbb{P}D \longrightarrow M$ is a bundle over M whose typical fiber $(\mathbb{P}D)_p$ is the space of 1-dimensional linear subspaces of the $(m + 1)$-dimensional vector space D_p. Thus, the fibers of $\mathbb{P}D$ are isomorphic to $\mathbb{PR}^m$ as projective m-spaces. There is a canonical rank-$(m + 1)$ distribution $D^{(1)}$ on $\mathbb{P}D$ defined by setting $D^{(1)}_\xi = (\pi')^{-1}(\xi)$ for each linear subspace $\xi \subset D_p$. This distribution $D^{(1)}$ is called the (generalized) Cartan prolongation of D.

Let us repeat that the prolonged distribution $D^{(1)}$ has the same rank $m + 1$ as the initial distribution D, but it lives on a much bigger manifold, having m dimensions more than the initial manifold M. Similarly as for the classical Cartan prolongation, immersed D-curves have canonical lifts 'upstairs' tangent to $D^{(1)}$. So it is clear what the local generators of $D^{(1)}$ are. For instance, one takes an immersed D-curve realizing any given horizontal direction ξ 'downstairs', then takes the direction of its canonical lift, and adds the m-dimensional kernel of the differential, taken at that point-direction ξ, of the projection π. Strictly speaking, a curve realizing the direction ξ is *not* necessary. It suffices to take the horizontal vectors alone and lift them upstairs, although only relatively. That is, *modulo* the kernel of π'. Having local generators of D – like in Section 3.3 of [14] – one is thus able to 'microlocally' write generators of $D^{(1)}$. (At this moment one already touches upon polynomial visualisations of the gCp's put forward in [14] and reiterated, for $m = 2$, in Section 3.3 of the present paper.)

We note that certain ingredients (but ingredients only) of the above definition of the generalized Cartan prolongation were dispersed in the literature, cf. Remark 1 in [14] for more on that.

We intend now to recall a local structural theorem generalizing Cartan's theorem from Section 1 which has geometrical applications, mainly to special multi-flags. Namely, the assumptions in Theorem 2 could be rephrased by avoiding mentioning D_2 and placing the foliation $\mathcal{F}$ in a new context. In fact, those assumptions easily implied that there existed a (unique) line subdistribution E of D preserving D_1, $[E, D_1] \subset D_1$. The foliation $\mathcal{F}$ was the integral of E. Driven by the definition of gCp's, we were going to replace a line subdistribution of a rank-2 distribution in Theorem 2 by an involutive rank-m subdistribution of a rank-$(m + 1)$ one (that is, by its corank-1 involutive subdistribution).

Theorem 3 ([14]). *Suppose D is a rank-$(m + 1)$ distribution on a manifold M^{s+m} such that* a) *D_1 is a rank-$(2m + 1)$ distribution on M, and* b) *there exists a corank-1 involutive subdistribution $E \subset D$ that preserves D_1, $[E, D_1] \subset D_1$. Then D is locally equivalent to the generalized Cartan prolongation $\big(D_1/E\big)^{(1)}$ of D_1 reduced modulo E (that lives on the quotient manifold $M/\mathcal{F}$ of dimension s, where $\mathcal{F}$ is the local m-dimensional foliation in M defined by E).*

Attention. $M/\mathcal{F}$ is to be understood only locally, to avoid topological complications. Note that $\dim M = 2m + 1$, i.e., $s = m + 1$ is not excluded in this theorem.

It appears that distributions emerging as the outputs of several applications of this theorem are precisely the jet bundles for maps $\mathbb{R} \to \mathbb{R}^m$ *together* with the neighbouring distributions prefigured by Kumpera. This should come as no surprise, for the gCp's were tailored for Theorem 3, which in turn was tailored for the objects Kumpera and Rubin wrote about – especially in the first version of [10] which was 60 + pages long. In short, it is Theorem 3 which underlies the theory of special multi-flags. In particular it 'makes possible' for Theorem 4 in Section 3 to hold true.

Aiming at completing now the discussion of the definition of special multi-flags, we note

Proposition 1. *Suppose that there is a distribution $D \subset TN$ of corank m bigger than* 1*, possessing an involutive corank-*1 *subdistribution E, and such that $[D, D] = TN$. Then, at each point $p \in N$, the value of E is described by all* (*local*) 1*-forms α on N such that $\big(\alpha \wedge d\omega\big)\big|_D = 0$ for all* (*local*) 1*-forms ω annihilating D.*

Moreover, the Cauchy-characteristic module $L(D)$ of D sits then inside E and is an involutive corank-m subdistribution of E.

This proposition is crucial for special multi-flags and, hence, also for the subject of the paper. It is proved in detail in Appendix C.

Remark 1. (a) Whenever a family $\widetilde{E}$ of subspaces of D, over points in N and of dimensions à priori possibly depending on those points, is being pointwisely described by the 1-forms α as in Proposition 1 and $\widetilde{E}$ happens to be of constant dimension, then, in [10, p. 5] it is called the *covariant subdistribution* $\widehat{D}$ of D. So then $\widehat{D} = \widetilde{E}$.

(b) Technically, the authors of [10] arrive at the covariant object not directly, but via the so-called polar spaces of $D^{\perp}$, included in $T^*N/D^{\perp}$. They continue only when the polar spaces are of constant dimension, independently of a point. In the situation in Proposition 1 that constant dimensionality turns out to be automatic (see Appendix C).

Corollary 2. *The involutive subdistribution $F \subset D^1$ from Definition 1 is unique and is nothing but the covariant subdistribution $\widehat{D^1}$. It automatically contains $L(D^1)$ as its corank-m subdistribution.*

Remark 2. (a) Alternatively, one could assume in $\star\star\star$ in Definition 1 that the covariant subdistribution of D^1 exists and is involutive. For, in view of Lemma 1 in [10], such a subdistribution is automatically of corank 1 in D^1; the hypotheses in that lemma are satisfied as $\mathrm{rk}\,[D^1, D^1]/D^1 = m > 1$.

(b) Equivalently, using Tanaka's and Yamaguchi's terminology [20, 22, 23] (well anterior to [10]), one could stipulate in $\star\star\star$ that the *symbol subdistribution* of $D^1/L(D^1)$, which is automatically of corank 1 here, be involutive. See also the detailed discussion of the symbol subdistribution on pages 28–30 in [23].

2.1 Monsters for special multi-flags

As it has been explicitly stated in [14] in Remark 3, every germ of a distribution generating a special m-flag of length r can be found within the r-fold generalized Cartan prolongation of the tangent bundle to $\mathbb{R}^{m+1}$. This follows directly from Theorem 3 coupled with the original version of the definition of special multi-flags given in section 3 of [14] (equivalent to the present Definition 1, in which the Cauchy-characteristic subdistributions are not explicitly used). Just like Goursat monster's coming into being was a direct consequence of Theorem 2. In the light of Theorem 3, locally universal objects in the theory of special multi-flags are very natural. In [14] they were abbreviated by MSkFM (from Monster Special k-Flags Manifold), and they should now be written as MSmFM, m, not k, standing now for the width.

In the recent paper [19] the gCp is named 'Rank 1 Prolongation'. The result of r consecutive gCp's applied to the tangent bundle to a manifold M of dimension $m+1$, is called there an *m-flag of length r* and is denoted by $(P^r(M), C^r)$. (Strictly speaking, the distribution C^r *generates* such a flag.) In order not to multiply symbols, we will adopt the notation $P^r(M)$ in the present paper, with a modest manifold $M = \mathbb{R}^3$.

3 Singularities of special 2-flags

It follows from the classical work [5] that special 2-flags of length 1 are homogeneous: they are identical around any point and hence feature no singularities at all. Here are two examples of rank-3 distributions generating special 2-flags of length 2. One of them is still homogeneous and the other one has a singular locus of codimension 1. The first example is the jet bundle on $J^2(1, 2)$,

$$\left(\frac{\partial}{\partial t} + x_1\frac{\partial}{\partial x} + y_1\frac{\partial}{\partial y} + x_2\frac{\partial}{\partial x_1} + y_2\frac{\partial}{\partial y_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial y_2}\right) \tag{1}$$

(it is generalized to bigger lengths in Example 1 below). The second one is the following non-homogeneous object

$$\left(x_2\Big(\frac{\partial}{\partial t} + x_1\frac{\partial}{\partial x} + y_1\frac{\partial}{\partial y}\Big) + \frac{\partial}{\partial x_1} + y_2\frac{\partial}{\partial y_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial y_2}\right), \tag{2}$$

which is singular on the hypersurface $\{x_2 = 0\}$. The fact that these two distributions are non-equivalent as germs at $0 \in \mathbb{R}^7$ will be clear in the next section. In fact, this is the starting point for the theory proposed in the paper. In width and length both equal to two, the object (1) is the local model for the [generic and the only one open] singularity class 1.1, while the object (2) is the model for the codimension-one singularity class 1.2 (see Sections 3.3 and 3.5 for precise definitions). These two classes build up the stratification of germs of special 2-flags when the length $r = 2$. The objective of the paper is to do the same in any length. Regarding further examples, a finer instance of a couple of nonequivalent 2-flags (of length 4) is given in Appendix B, with the aim of illustrating the main constructions of the paper.

3.1 Sandwich diagram for special 2-flags

Special multi-flags, and in particular special 2-flags, appear, from one side, to be rich in singularities, and from the other, to possess finite-parameter families of local pseudo-normal forms, with no functional moduli, constructed in [14]. The respective tree of normal forms is very natural and emerges in a transparent way from the sequences of gCp's being at work. Multi-parameter normal forms, in the case of 2-flags dealt with in the present paper, are indexed by certain words over the alphabet $\{\mathbf{1},\mathbf{2},\mathbf{3}\}$ of length equal to flag's length.

On the other hand, a basic partitioning in the world of special multi-flags is a stratification into singularity classes proposed in the preprint [15] and reproduced, for 2-flags, below. In their turn, the singularity classes for special 2-flags are encoded by certain words over the alphabet $\{1, 2, 3\}$ of length equal to flag's length.

Both partitions exist in their own rights, with no apparent relation to each other. A first (modest) step towards throwing bridges is the concept of sandwich classes (Section 3.2), followed by Corollary 3 which makes use of that concept.

While the eventual aim of the paper, earlier undertaken in [17] and interrupted, is to identify these two vocabularies: to show that words over $\{\mathbf{1},\mathbf{2},\mathbf{3}\}$ and words over $\{1, 2, 3\}$ label precisely the same aggregates of germs of special 2-flags – see Theorem 1. A similar issue for multi-flags of widths bigger than 2 will be addressed in a future work.

Our initial requirements $\star\star$ and $\star\star\star$ are visualised best in a *sandwich diagram*[4]

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

The inclusions in its lower line are due to the Jacobi identity ($L(D^{j-1}) \supset L(D^j)$) and to Corollary 2 ($F \supset L(D^1)$). All vertical inclusions in this diagram are of codimension one while all drawn horizontal inclusions are of codimension 2. The squares formed by these inclusions can 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 3). At that, the sum ($=3$) of codimensions, 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 2-dimensional *plane* $F/L(D^1)$ and the line $D^2/L(D^1)$ are mutually positioned in $D^1/L(D^1)$. Similar questions also arise in further sandwiches 'indexed' by the upper right 'vertices' $D^3, D^4, \ldots, D^r$.

3.2 Analogues for special 2-flags of Kumpera–Ruiz classes

We thus first divide all existing germs of special 2-flags of length r into 2^{r-1} pairwise disjoint *sandwich classes* depending on the geometry of the distinguished spaces in the sandwiches (at the reference point for a germ), and label those aggregates by words of length r over the alphabet $\{1,\underline{2}\}$ starting (on the left) with 1, having the second letter $\underline{2}$ iff $D^2(p) \subset F(p)$, and for $3 \le j \le r$ having the j-th letter $\underline{2}$ iff $D^j(p) \subset L(D^{j-2})(p)$.

It follows immediately from this definition that the sandwich classes are pairwise disjoint. On the other hand, it is not yet clear if each of them is actually nonempty; this follows only from Corollary 3 below.

[4] So-called after a similar diagram assembled for Goursat distributions, or 1-flags, in [12].

The construction of sandwich classes points to possible non-transverse situations in the sandwiches. For instance, the second letter in a sandwich label is $\underline{2}$ iff the line $D^2(p)/L(D^1)(p)$ is included in the plane $F(p)/L(D^1)(p)$, both the line and plane sitting in the 3-space $D^1(p)/L(D^1)(p)$. And it is similarly in further sandwiches. This resembles the Kumpera–Ruiz classes of Goursat germs constructed in [12]. The number of sandwiches in length r then was $r-2$ (and so the # of KR classes 2^{r-2}) due to the degenerate form of the covariant distribution of D^1: $\widehat{D^1} = L(D^1)$. Now, for 2-flags this number is $r-1$, because the covariant distribution of D^1 differs from $L(D^1)$, and gives rise to one additional sandwich.

How can one establish if such virtually created sandwich classes really materialize? And, if so, is it possible to sort them further?

We shall produce a huge variety of polynomial germs at $0 \in \mathbb{R}^N$, of rank-3 distributions, where N will be odd and possibly be very large. It is important that 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 with respect to the local classification or subject to further simplifications. More precisely, for each $\mathbf{k} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$ we are going to define an *operation* $\mathbf{k}$ producing new rank-3 distributions from previous ones. Technically, its outcome (and especially the indices of new incoming variables) will also depend on how many operations were done *before* $\mathbf{k}$.

The result of $\mathbf{k}$, being performed as an l-th operation in a succession of operations, 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 a new rank-3 distribution – the germ at $0 \in \mathbb{R}^{s+2}(u_1, \ldots, u_s, x_l, y_l)$, generated by the vector fields

$$Z_1' = \begin{cases} Z_1 + (b_l + x_l)Z_2 + (c_l + y_l)Z_3, & \text{when } \mathbf{k} = \mathbf{1}, \\ x_l Z_1 + Z_2 + (c_l + y_l)Z_3, & \text{when } \mathbf{k} = \mathbf{2}, \\ x_l Z_1 + y_l Z_2 + Z_3, & \text{when } \mathbf{k} = \mathbf{3} \end{cases}$$

and $Z_2' = \frac{\partial}{\partial x_l}$, $Z_3' = \frac{\partial}{\partial y_l}$. Here b_l and/or c_l are real parameters whose values are specified later, when one applies these operations to concrete objects. For any subsequent such operation (one will need 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'). Later (in Section 4) we will write more compactly $X_l = b_l + x_l$, $Y_l = c_l + y_l$.

3.3 Definition of EKR's

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

$$\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_0}, \frac{\partial}{\partial y_0}\right) \tag{3}$$

understood in the vicinity of $0 \in \mathbb{R}^3(t, x_0, y_0)$; this full tangent bundle to a 3-space is encoded by an empty label. (The name 'EKR' was coined in the work [18], although the very method of producing local visualisations of special multi-flags was not correct there. Namely, the authors of [18] arrived only at the operations $\mathbf{1}$ and $\mathbf{2}$. In fact, their relevant operations are just $\mathbf{1}$ and $\mathbf{2}$ modulo reindexations in the m-tuples of their variables $x_1^j, x_2^j, \ldots, x_m^j$ $(j = 0, 1, \ldots, n)$ and similar reindexations in the m-tuples of their vector fields $\kappa_1^j, \kappa_2^j, \ldots, \kappa_m^j$ $(j = 1, 2, \ldots, n)$, cf. [18, pp. 112–113]. While the operation $\mathbf{3}$ is necessary already for $m = 2$, as shows Proposition 1(iv) in [14] and the entire message of the present article. Likewise, operation $\mathbf{4}$ would turn out necessary from width 3 and length 4 onwards, operation $\mathbf{5}$ from width 4 and length 5 on, etc.) Assume that the family of pseudo-normal forms $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}$ is 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 distribution (3) when $r-1=0$). Then the normal forms subsumed under the symbol $\mathbf{j}_1 \dots \mathbf{j}_{r-1}.\mathbf{j}_r$ are the outcome of the operation $\mathbf{j}_r$ performed as the operation number r over the distributions $\mathbf{j}_1 \cdots \mathbf{j}_{r-1}$.

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 the involutive subdistributions of ranks $2, 4, \dots, 2r$ are; see also Proposition 2 below. The point is that locally the converse is also true.

Theorem 4. *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 $p \in M$, the distribution D is equivalent in a neighbourhood of p to a certain EKR $\mathbf{j}_1.\mathbf{j}_2 \dots \mathbf{j}_r$ in a neighbourhood of $0 \in \mathbb{R}^{2r+3}$, by a local diffeomorphism that sends p to* 0*. Moreover, that EKR can be taken such that $\mathbf{j}_1 = \mathbf{1}$ and the first letter* **2***, if any, appears before the first letter* **3** *(if any).*

This theorem is just the specification of Theorem 3 in [14] to the special 2-flags. In particular, the restriction on EKR's codes in it is the specification to the width $m = 2$ of the general rule of the least upward jumps put forward in [14] and already briefly explained in Section 1.

This rule looks modest in width 2. It becomes more and more restrictive in larger widths $3, 4, \dots$. Despite this, the idea standing behind it is simple. At a new stage, one Cartan-prolongs in the vicinity of a direction ξ. What operation could one use for a local description of that Cartan prolongation? Basically, any operation whose pivot is not perpendicular to ξ. Now suppose additionally that *all* such operations have their numbers (or: indices) *higher* than the indices of operations used before that stage. The rule under discussion says that one should choose the operation which has the *lowest* index among the not-yet-used indices. Technically, it boils down to a *reindexation* of the 'new' coordinates having those higher indices. Then such a reindexation can safely be extended onto the 'old' coordinates bound to operations at earlier stages, not affecting the numbering of those earlier operations. Thus, inductively, one is able to obey the rule of the least upward jumps. Details can be traced down in [14, pp. 167–168].

We stress that possible constants in the EKRs representing a given germ D (i.e., the constants in the EKRs in Theorem 4) are not, in general, defined uniquely.

Example 1. The EKR $\mathbf{1}.\mathbf{1} \dots \mathbf{1}$ (r letters $\mathbf{1}$) subsumes a vast family of different pseudo-normal forms – germs at $0 \in \mathbb{R}^{2r+3}$ parametrized by real parameters $b_1, c_1, \dots, b_r, c_r$. Under a closer inspection (Theorem 1 in [10]), they all are pairwise equivalent, and are equivalent to the classical jet bundle – Cartan distribution – on the space $J^r(1,2)$ of the r-jets of functions $\mathbb{R} \to \mathbb{R}^2$, given by the Pfaffian equations

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

All distribution germs in all other EKRs are not equivalent to the jet bundles; this follows from Corollary 3 below.

Let us note that the question of a geometric characterization of Cartan distributions as such was addressed in many papers. In full generality (for all jet spaces $J^r(n,m)$) that question was answered only in 1983 in [23].

3.4 The EKR's versus sandwich classes

What kind of a relationship does there exist between the sandwich class of a given germ of a special 2-flag and its all possible EKR presentations? In order to answer, we note

Proposition 2. *If a distribution $D = D^r$ generating a special* 2*-flag of length $r \geq 1$ is presented in any EKR form on $\mathbb{R}^{2r+3}(t, x_0, y_0, \dots, x_r, y_r)$, then the members of the associated subflag in the sandwich diagram for D^r are canonically positioned as follows*

- $F = \big(\partial/\partial x_1, \partial/\partial y_1, \partial/\partial x_2, \partial/\partial y_2, \ldots, \partial/\partial x_r, \partial/\partial y_r\big)$,
- $L(D^j) = \big(\partial/\partial x_{j+1}, \partial/\partial y_{j+1}, \ldots, \partial/\partial x_r, \partial/\partial y_r\big)$ *for* $1 \le j \le r-1$,
- $L(D^r) = (0)$.

Proof is almost immediate, because the inclusions $\supset$ (not yet equalities) are clear in view of the construction of the EKR's, while the dimensions of spaces on both sides of these inclusions always coincide by the definition of special 2-flags.

(These extremely simplified descriptions of the members of the associated subflag are the analogues of similar descriptions holding true for Goursat flags viewed in Kumpera–Ruiz coordinates.)

Proposition 2 has an important corollary. Namely,

Corollary 3. *Each given sandwich class in length r having label $\mathcal{E}$ is the aggregate of all germs admitting EKR visualisations of the forms $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}.\mathbf{j}_r$ such that $\mathbf{j}_l = \mathbf{1} \iff$ the l-th letter in $\mathcal{E}$ is* 1, *for $l = 1, 2, \ldots, r$.*

Therefore, the basic singular phenomena of the 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. The nonemptiness of sandwich classes follows. Moreover, they are embedded submanifolds in the monster manifolds $P^r(\mathbb{R}^3)$ of codimensions equal to the number of letters $\underline{2}$ in their codes. (We do not dwell on this any longer because by far more important are smaller bricks, or singularity classes, building up sandwich classes.)

Proof. $\mathbf{j}_1$ is by default **1** and the first letter 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) of length $l-1$ into an EKR (Z_1', Z_2', Z_3') of length l. When $\mathbf{j}_l$ is either **2** or **3** then, by definition of these operations,

$$Z_1' \equiv x_l Z_1 \bmod (Z_2, Z_3), \tag{4}$$

where $Z_2 = \frac{\partial}{\partial x_{l-1}}$ and $Z_3 = \frac{\partial}{\partial y_{l-1}}$. (As for $Z_2' = \frac{\partial}{\partial x_l}$ and $Z_3' = \frac{\partial}{\partial y_l}$, they cause no trouble in the discussion.) Whereas for $\mathbf{j}_l = \mathbf{1}$ we have $Z_1' \equiv Z_1 \bmod (Z_2, Z_3)$ and the non-zero vector $Z_1(0)$ is, by its recursive construction (in $l-1$ steps), spanned by

$$\partial/\partial t,\ \partial/\partial x_0,\ \partial/\partial y_0,\ \ldots,\ \partial/\partial x_{l-2},\ \partial/\partial y_{l-2}. \tag{5}$$

Hence, in view of Proposition 2, $Z_1(0)$ does not lie in $F(0)$ when $l = 2$, and in $L(D^{l-2})(0)$, when $l > 2$. ■

Remark 3. When $m = 1$ two operations, instead of three (**1**,**2**,**3**) in the present text, lead to the local Kumpera–Ruiz pseudo-normal forms for Goursat flags, evoked already in Section 1.

3.5 Singularity classes of special 2-flags refining the sandwich classes

We recall from [15] how one passes from the sandwich classes to *singularity classes*. In fact, to any germ $\mathcal{F}$ of a special 2-flag we associate a word $\mathcal{W}(\mathcal{F})$ over {1,2,3}, called a 'singularity class' of $\mathcal{F}$. It is a specification of the word 'sandwich class' for $\mathcal{F}$ (a word, recalling, over $\{1,\underline{2}\}$) with the letters $\underline{2}$ replaced either by 2 or 3, depending on the geometry of $\mathcal{F}$. It will be momentarily clear from the definition that $\mathcal{W}(\cdot)$ is an invariant of the local classification of flags with respect to diffeomorphisms in the base manifold.

Alternatively, if one restricts oneself to the locally universal flags of distributions C^r living on $P^r(\mathbb{R}^3)$, then $\mathcal{W}$ becomes essentially a function of a point in $P^r(\mathbb{R}^3)$, and it will turn out to be an invariant of the (local) symmetries of C^r. That is, an invariant of the local diffeomorphisms of $\mathbb{R}^3$, inducing after r prolongations the symmetries of C^r on $P^r(\mathbb{R}^3)$.

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

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

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

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

$V_{i+1} = V_i + [D^s, V_i]$, and then focusing on this new flag's member V_{2l+3}. Recall that, in the ν-th sandwich, there holds the inclusion: $F(p) \supset D^2(p)$, when $\nu = 2$, or else $L(D^{\nu-2})(p) \supset D^\nu(p)$, when $\nu > 2$. This is a preparation to an important, turning point decision.

Namely, writing $V_{2l+3}(p)$ instead of $D^\nu(p)$ in the relevant inclusion, and always controlling whether $\nu = 2$ or $\nu > 2$, means specifying j_s to 3. That is to say, $j_s = \underline{2}$ is being specified to 3 if and only if $F(p) \supset V_{2l+3}(p)$ (when $\nu = 2$) or else when $L(D^{\nu-2})(p) \supset V_{2l+3}(p)$ (when $\nu > 2$) holds.

In this way all non-first letters $\underline{2}$ in $\mathcal{C}$ are, one independently of another, specified to 2 or 3. Having done that, 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. The word created by such a procedure clearly satisfies the least upward jumps rule.

This is the singularity class of a given 2-flag at a point. So what is an abstract singularity class in length r, what subset of the monster manifold $P^r(\mathbb{R}^3)$ does it form? It is the union of all points in $P^r(\mathbb{R}^3)$ at which the *universal* flag has a fixed singularity class – a fixed word of length r over $\{1, 2, 3\}$ obeying the rule of least upward jumps. Hence there emerges a partition of $P^r(\mathbb{R}^3)$ into abstract, pairwise disjoint singularity classes.

Example 2. In length 4 there exist (or: $P^4(\mathbb{R}^3)$ is partitioned into) 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.[5]

(Reiterating, the emptiness of certain singularity classes has not been à priori excluded. Only Theorem 1 shows that all singularity classes are not empty – see the paragraph after Corollary 1.)

How many singularity classes do there exist for special 2-flags of fixed length?

Proposition 3. *The number of different singularity classes of special* 2*-flags of length* $r \ge 3$ *is* $2 + 3 + 3^2 + \cdots + 3^{r-2} = \frac{1}{2}(1 + 3^{r-1})$.

Proof. Let us recall that the class' code $j_1.j_2 \ldots j_r$ is subject to the least upward jumps rule. Either it is $1.1 \ldots 1$, or else it has the first from left letter $j_f = 2$ at the f-th position, $2 \le f \le r$. For $f = r$ one gets just 1 class. For $f = r - 1$ the number of classes' codes is 3^1, for $f = r - 2$ that number is 3^2, and so on downwards to $f = 2$, with the respective number of such classes 3^{r-2}. ∎

Remark 4. (a) The singularity classes discussed in the present paper are just the visible part of an iceberg. Their counterparts in the Goursat world, the KR classes, are nothing but vague approximations to the orbits of the local classification. Much finer are then *geometric classes* emanating from Jean's benchmark contribution [8] and otherwise prefigured in [12]. They are described in detail in [16][6]. Although they, too, are encoded by certain words over a three letters'

[5] In widths ≥ 3 the class 1.2.3.4 will show up as well, cf. Remark 4(b).

[6] In a different language using extensively classical Cartan prolongations and projections in the Goursat monster tower, the geometric classes, under the name of 'RVT classes', have been recently very originally treated in [13].

alphabet, one should *by no means* confuse them with singularity classes for special 2-flags. The question of further partitioning of singularity classes for special 2-flags (and/or generally for special multi-flags) is under investigation, if still open for the most part.

(b) Reiterating after Section 3.1, singularity classes for *all* widths m have been defined in [15]. To give an idea of their numbers, let us, for example, fix the length $r = 7$. Then the numbers of different singularity classes of special m-flags, for $m \in \{1, 2, \dots, 6\}$, are as follows (for $m = 1$ counted are the KR classes):

m	1	2	3	4	5	6
#	32	365	715	855	876	877

The value 365 is the value for $r = 7$ of the expression given in Proposition 3.

Remark 5. Theorem 1 naturally generalizes to *wider* special flags. For $m > 2$ the first refinement of a sandwich class – a word over {1,2,$\underline{3}$} (see [15] for details) – is not yet a singularity class. But it is a purely geometric notion, imposing natural restrictions on the EKRs representing germs that have a fixed word $j_1.j_2 \dots j_r$ over {1,2,$\underline{3}$} (satisfying the least upward jumps rule). If $\mathbf{k}_1.\mathbf{k}_2 \dots \mathbf{k}_r$ is any such an EKR, then $j_l = \min(k_l, \underline{3})$ for $l = 1, 2, \dots, r$. That is, $j_l = k_l$ for those $\mathbf{k}_l$'s that are equal to **1** or **2**, and $j_l = \underline{3}$ for all the remaining $\mathbf{k}_l$'s.

A proof of this generalization of Theorem 1 is only technically more complex, but not more difficult than the one presented in the following chapter.

Last but not least, there arises a question concerning the materializations of singularity classes for *concrete* special 2-flags. In fact, on each manifold M of dimension $2r + 3$, $r \geq 1$, bearing a special 2-flag of length r, the shadows of universal singularity classes in $P^r(\mathbb{R}^3)$ always form – and not only for 'generic' flags – a very neat stratification by embedded submanifolds whose codimensions are directly computable. Namely, we have the following

Proposition 4. *The codimension of an embedded in M submanifold of the realization of any fixed singularity class $\mathcal{C}$, if only nonempty, is equal to*

$$\text{the number of letters 2 in } \mathcal{C} \;+\; \text{twice the number of letters 3 in } \mathcal{C}. \qquad (*)$$

In particular, the same formula $(*)$ holds for each singularity class $\mathcal{C} \subset P^r(\mathbb{R}^3)$. In this case $\mathcal{C}$ is automatically nonempty because of the universality property of $P^r(\mathbb{R}^3)$: $\mathcal{C}$ is mapped by the relevant EKR coordinates into certain $\mathbb{R}^{(r+1)2+1}$ bearing the EKR forms with the label identical to the label of $\mathcal{C}$. Speaking differently, the monster manifold $P^r(\mathbb{R}^3)$ carries a universal (in length r) stratification into nonempty singularity classes.

A sketched proof of this is postponed until after the proof of Theorem 1 (Appendix A). In turn, once the codimensions are made explicit, another natural question is that about the adjacencies existing among these classes. We do not have a full answer to this question yet. We only know that, in any fixed length $r \geq 1$,

Proposition 5. *The generic class* 1.1…1 *is not adjacent to any other singularity class. An adjacency $j_1.j_2 \dots j_l \dots j_r \to j_1.j_2 \dots (j_l - 1) \dots j_r$, $2 \leq l \leq r$, holds whenever $j_l = 3$ or $j_l = 2$, provided, in the latter case, there is no letter 3 past j_l (i.e., among $j_{l+1}, \dots, j_r$).*

For instance, $1.2.3 \to 1.2.2 \to 1.1.2 \to 1.1.1$, or else $1.2.3.2 \to 1.2.3.1 \to 1.2.2.1 \to \cdots$. To completely answer the question, a deep analysis of EKRs (i.e., the effective realizations, or visualisations, of special flags) is needed.

4 Proof of Theorem 1

We assume that the reader remembers the way the sandwich classes were refined to singularity classes in Section 3.5. In the proof of Theorem 1 we stay within that same framework (and notation) and assume that:

– the ν-th letter j_ν in $\mathcal{C}$ is *not* 1,

– there follow $l \geq 0$ letters 1 past j_ν,

– the following letter j_s is *not* 1, where $s = \nu + l + 1$.

Having $D = D^r$ in a not-yet-specified EKR form $\mathbf{k}_1.\mathbf{k}_2 \ldots \mathbf{k}_r$ we know by Corollary 3 that $\mathbf{k}_\nu \neq \mathbf{1}$, $\mathbf{k}_{\nu+1} = \cdots = \mathbf{k}_{s-1} = \mathbf{1}$, $\mathbf{k}_s \neq \mathbf{1}$. And we aim to show that

$$\mathbf{k}_s = \mathbf{3} \quad \text{if and only if} \quad j_s = 3. \tag{$**$}$$

Only this is an issue. For, the first from the left letter $\mathbf{k}_f \neq \mathbf{1}$ (if any) is $\mathbf{2}$ by the least upward jumps rule satisfied by the labels of EKR classes, and the corresponding letter j_f in $\mathcal{C}$ is – by the same Corollary 3 – the first from the left not 1 letter in $\mathcal{C}$. Hence it is 2 by the very definition of singularity classes.

As for $(**)$, in Section 4.1 we will show that $\mathbf{k}_s = \mathbf{2}$ implies $j_s = 2$, and in section 4.2 that $\mathbf{k}_s = \mathbf{3}$ implies $j_s = 3$. That will do, because $\mathbf{k}_s \in \{\mathbf{2}, \mathbf{3}\} \Longleftrightarrow j_s \in \{2, 3\}$ by Corollary 3.

Prior to concrete computations, note that, automatically, the rank-3 distribution $D^s/L(D^s)$, generating a special 2-flag of length s, is in an EKR form $\mathbf{k}_1.\mathbf{k}_2 \ldots \mathbf{k}_s$. In the (rather long) computations that follow we *skip* writing down this factoring out by the Cauchy characteristics $L(D^s)$. That is, we simply leave out the variables with indices from $s+1$ onwards, upon which D^s does not depend (Proposition 2). Also, for space reasons, from now on we shall just write ∂_x and ∂_{x_k} instead of $\partial/\partial x$ and $\partial/\partial x_k$, respectively.

4.1 Easier part: $\mathbf{k}_s = \mathbf{2}$

We first deal with the case $\mathbf{k}_s = \mathbf{2}$ and aim at showing that then $j_s = 2$ (meaning the *non*-inclusion of $V_{2l+3}(0)$ in the relevant member of the Cauchy-characteristic subflag).

Proof. Let us expand the first member of the small flag of D^s

$$\begin{aligned} D^s = V_1 = \Bigg(& x_s\Big(\underline{x_\nu Z + *\big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}}\big)} + \sum_{k=\nu}^{s-2}\big(X_{k+1}\partial_{x_k} + Y_{k+1}\partial_{y_k}\big)\Big) \\ & + \partial_{x_{s-1}} + Y_s\partial_{y_{s-1}}, \partial_{x_s}, \partial_{y_s}\Bigg), \end{aligned} \tag{6}$$

where the underlined summand is the leading generator of flag's member D^ν. That is,

$$D^\nu/L(D^\nu) = \Big(x_\nu Z + *\big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}}\big), \partial_{x_\nu}, \partial_{y_\nu}\Big)$$

and the functions $*$ depend on whether $\mathbf{k}_\nu$ is $\mathbf{2}$ or $\mathbf{3}$. The capital letters X and Y stand, as in the end of Section 3.2, for variables shifted by constants: $X_{k+1} = b_{k+1} + x_{k+1}$, $Y_{k+1} = c_{k+1} + y_{k+1}$. By means of a straightforward step by step computation, stopping at each *odd* member of the small flag, one shows that

$$\begin{aligned} &\partial_{x_{s-2}} + Y_s\partial_{y_{s-2}} \in V_3, \\ &\partial_{x_{s-3}} + Y_s\partial_{y_{s-3}} \in V_5, \end{aligned}$$

$$\begin{aligned}
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\partial_{x_\nu} + Y_s\partial_{y_\nu} \in V_{2l+1}, \\
&Z + Y_s\big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}}\big) \in V_{2l+3},
\end{aligned} \tag{7}$$

where $(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}})$ stands for certain combination of the versors $\partial_{x_{\nu-1}}$ and $\partial_{y_{\nu-1}}$. In view of Proposition 2, these versors lie in $L(D^{\nu-2})$, or in F when $\nu = 2$. Whereas $Z(0)$ is, exactly as in Section 3.4, a nonzero combination of versors (5) for $l = \nu$ and as such sticks out of $L(D^{\nu-2})(0)$, or of $F(0)$ when $\nu = 2$. Therefore (7) alone implies that $V_{2l+3}(0)$ is not included in $L(D^{\nu-2})(0)$, or in $F(0)$ when $\nu = 2$. That is, $j_s = 2$. ■

4.2 Harder part: $\mathbf{k}_s = \mathbf{3}$

One should justify that now $j_s = 3$. That is, that there holds the inclusion

$$V_{2l+3}(0) \subset \big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}},\ \partial_{x_\nu}, \partial_{y_\nu}, \ldots, \partial_{x_s}, \partial_{y_s}\big)\,. \tag{8}$$

Proof. The initial object $D^s = V_1$ is now different from (6). Namely,

$$V_1 = \left(x_s\Big(x_\nu Z + *\big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}}\big) + \sum_{k=\nu}^{s-2}\big(X_{k+1}\partial_{x_k} + Y_{k+1}\partial_{y_k}\big)\Big) + \underline{y_s\partial_{x_{s-1}} + \partial_{y_{s-1}}}, \partial_{x_s}, \partial_{y_s}\right),$$

where $*$ stands for certain functions depending on the value of $\mathbf{k}_\nu$. Note the only difference, in the underlined part, with the leading generator in (6). This slight difference will turn out to be decisive in the output V_{2l+3}. Let us compute carefully some first members of the small flag of D^s:

$$\begin{aligned}
V_2 &= \left(x_s\Big(x_\nu Z + *\big(\partial_{x_{\nu-1}}, \partial_{y_{\nu-1}}\big) + \sum_{k=\nu}^{s-2}\big(X_{k+1}\partial_{x_k} + Y_{k+1}\partial_{y_k}\big)\Big), \partial_{x_{s-1}}, \partial_{y_{s-1}}, \partial_{x_s}, \partial_{y_s}\right), \\
V_3 &= \big(V_2, x_s\partial_{x_{s-2}}, x_s\partial_{y_{s-2}}, y_s\partial_{x_{s-2}} + \partial_{y_{s-2}}\big)\,, \\
V_4 &= \Big(V_3, \partial_{x_{s-2}}, \partial_{y_{s-2}}, x_s^2\partial_{x_{s-3}}, x_s^2\partial_{y_{s-3}}, x_s\big(y_s\partial_{x_{s-2}} + \partial_{y_{s-2}}\big)\Big).
\end{aligned} \tag{9}$$

Acting likewise, one keeps expressing V_{n+1} by the previous module V_n and a set of simple vector field's generators, of the cardinality growing linearly with n, for $n \le l+2$. The modules V_{l+2} and V_{l+3} are the most important in this process of computing. The reader will see that in V_{l+3} for the first time there appears the field Z standing alone, only with a monomial factor of high degree. That field requires a particular care; in the situation $\mathbf{k}_s = \mathbf{2}$ it has been responsible for the failure of the inclusion. Strictly speaking, the modules V_{l+2} and V_{l+3} look differently depending on the parity of l:

- $\star$ $l = 2k-1$, $k \ge 1$, or else
- $\star\star$ $l = 2k$, $k \ge 0$.

However, these differences are not fundamental and one common technique works in both situations. But a choice *is* necessary when it comes to details. So for the presentation in the text we choose $\star$.

For l odd, V_{l+2} is the module generated by V_{l+1} and by the following set of generators:

- $x_s\partial_{x_{\nu+k-1}},\ x_s\partial_{y_{\nu+k-1}},\ y_s\partial_{x_{\nu+k-1}} + \partial_{y_{\nu+k-1}}$;
- $x_s^3\partial_{x_{\nu+k-2}},\ x_s^3\partial_{y_{\nu+k-2}},\ x_s^2\,\big(y_s\partial_{x_{\nu+k-2}} + \partial_{y_{\nu+k-2}}\big)$;
- $\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$

- $x_s^{2k-1}\partial_{x_\nu},\ x_s^{2k-1}\partial_{y_\nu},\ x_s^{2k-2}\left(y_s\partial_{x_\nu}+\partial_{y_\nu}\right)$.

It has been straightforward to see that $V_{l+1}(0)$ is included in the RHS of (8). Hence so is $V_{l+2}(0)$. In turn, V_{l+3} is the previous module V_{l+2} extended by the generators

- $\partial_{x_{\nu+k-1}},\ \partial_{y_{\nu+k-1}}$;
- $x_s^2\partial_{x_{\nu+k-2}},\ x_s^2\partial_{y_{\nu+k-2}},\ x_s\left(y_s\partial_{x_{\nu+k-2}}+\partial_{y_{\nu+k-2}}\right)$;
- ....
- $x_s^{2k-2}\partial_{x_\nu},\ x_s^{2k-2}\partial_{y_\nu},\ x_s^{2k-3}\left(y_s\partial_{x_\nu}+\partial_{y_\nu}\right)$;
- $x_s^{l+1}Z,\ x_s^{l+1}\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big),\ x_s^l y_s Z + x_s^l\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big)$

(remember that $2k=l+1$). This gives that $V_{l+3}(0)$ is included in the RHS of (8). Having V_{l+3} thus described is a turning point in the proof. Indeed, there remains exactly l steps from V_{l+3} to V_{2l+3}. Because of that, in the bottom line of the new generators in V_{l+3}, all terms with degree $l+1$ monomials are irrelevant as they give rise only to terms vanishing at 0 under l Lie multiplications.

The remaining terms $x_s^l\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big)$ in that bottom line could contribute at 0 only by means of differentiating that monomial l times during l Lie bracketings still to be performed, because of a degree l monomial in them. Consequently they yield only the output sitting (at 0) in $\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big)(0)$.

All in all, in l steps, the bottommost line of generators of V_{l+3} will give rise uniquely to vector fields having at 0 values sitting in the RHS of (8). Therefore only the remaining lines of generators are relevant for the answer to the question whether (8) holds. Thus, for answering this question, V_{l+3} can be *replaced* till the end of computations by the module $\overline{V}_{l+3}$ generated by V_{l+2} and the smaller set of vector fields

- $\partial_{x_{\nu+k-1}},\ \partial_{y_{\nu+k-1}}$;
- $x_s^2\partial_{x_{\nu+k-2}},\ x_s^2\partial_{y_{\nu+k-2}},\ x_s\left(y_s\partial_{x_{\nu+k-2}}+\partial_{y_{\nu+k-2}}\right)$;
- ....
- $x_s^{2k-2}\partial_{x_\nu},\ x_s^{2k-2}\partial_{y_\nu},\ x_s^{2k-3}\left(y_s\partial_{x_\nu}+\partial_{y_\nu}\right)$.

So $\overline{V}_{l+3}$ is to be Lie bracketed l times with V_1. Then the result taken at 0, $\overline{V}_{2l+3}(0)$, is to be checked for its inclusion in the RHS of (8), and that *would* finish the proof. Yet, as it stands, it is not transparent at all, and a series of further simplifications is needed. Computing now the next module $\overline{V}_{l+4}=\overline{V}_{l+3}+\big[V_1,\overline{V}_{l+3}\big]$, one sees that it is the module generated by $\overline{V}_{l+3}$ and the following collection of vector fields

- $x_s\partial_{x_{\nu+k-2}},\ x_s\partial_{y_{\nu+k-2}},\ y_s\partial_{x_{\nu+k-2}}+\partial_{y_{\nu+k-2}}$;
- $x_s^3\partial_{x_{\nu+k-3}},\ x_s^3\partial_{y_{\nu+k-3}},\ x_s^2\left(y_s\partial_{x_{\nu+k-3}}+\partial_{y_{\nu+k-3}}\right)$;
- ....
- $x_s^{2k-3}\partial_{x_\nu},\ x_s^{2k-3}\partial_{y_\nu},\ x_s^{2k-4}\left(y_s\partial_{x_\nu}+\partial_{y_\nu}\right)$;
- $x_s^lZ,\ x_s^l\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big),\ x_s^{l-1}y_sZ+x_s^{l-1}\big(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\big)$

(remember that $2k-1=l$). Hence $\overline{V}_{l+4}(0)$ sits in the RHS of (8).

From now on the arguments start to repeat themselves. Only $l-1$ Lie bracketings with V_1 remain to be done, hence the last line of new generators for $\overline{V}_{l+4}$ is irrelevant for the sought inclusion of $\overline{V}_{2l+3}(0)$ in the RHS of (8). Consequently, $\overline{V}_{l+4}$ can be replaced until the end of computations by the module $\overline{\overline{V}}_{l+4}$ generated by $\overline{V}_{l+3}$ and the smaller set of vector fields

- $x_s\partial_{x_{\nu+k-2}},\ x_s\partial_{y_{\nu+k-2}},\ y_s\partial_{x_{\nu+k-2}}+\partial_{y_{\nu+k-2}}$;

- $x_s^3\partial_{x_{\nu+k-3}},\ x_s^3\partial_{y_{\nu+k-3}},\quad x_s^2\left(y_s\partial_{x_{\nu+k-3}}+\partial_{y_{\nu+k-3}}\right);$
- $\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$
- $x_s^{2k-3}\partial_{x_\nu},\ x_s^{2k-3}\partial_{y_\nu},\quad x_s^{2k-4}\left(y_s\partial_{x_\nu}+\partial_{y_\nu}\right).$

In its turn, the module $\overline{\overline{\overline{V}}}_{l+4}$ will lead in $l-1$ steps to a module $\overline{\overline{\overline{V}}}_{2l+3}$ which will suffice for the verification of the inclusion as well. In the first of these steps one is to:

– compute the module $\overline{\overline{\overline{V}}}_{l+5}=\overline{\overline{\overline{V}}}_{l+4}+\left[V_1,\overline{\overline{\overline{V}}}_{l+4}\right]$,

– check its inclusion at 0 in the RHS of (8), and then

– leave out its bottommost line of new generators, irrelevant for the verification of the inclusion of $\overline{\overline{\overline{V}}}_{2l+3}(0)$ in the RHS of (8).

And then to proceed similarly in the remaining steps.

Summarizing, the critical part of the procedure applied in the proof of Theorem 1 consists of the following bipartite steps, having numbers $l+3+\tau$, where $\tau\in\{1,\dots,l-2\}$.

Firstly in checking the inclusion at 0, of the due reduction of $V_{l+3+\tau}$, in the RHS of (8). *Secondly* in deleting the most involved, and not important for the eventual inclusion of $V_{2l+3}(0)$, among the *new* vector field's generators showing up in the Lie product of V_1 with the mentioned reduction of $V_{l+3+\tau}$.

After reducing the problem in $l-2$ steps in the described way, the situation is as follows (we use the $\widetilde{\ }$ symbols instead of writing many bars):

One knows already from the step number $l+3+l-2$ that a) the reduced module $\widetilde{V_{2l+1}}$ is such that $\widetilde{V_{2l+1}}(0)$ is included in the RHS of (8), and b) the newly emerging module $\widetilde{\widetilde{V_{2l+2}}}$, again sufficient for the verification of the inclusion in question because only irrelevant (for that verification) new generators have *just* been deleted, is generated by $\widetilde{V_{2l+1}}$ and by a tiny set of generators

- $\partial_{x_\nu},\ \partial_{y_\nu}$.

Hence, at 0, it is included in the RHS of (8) as well. Moreover, the module

$$\widetilde{\widetilde{\widetilde{V_{2l+3}}}}=\widetilde{\widetilde{\widetilde{V_{2l+2}}}}+\left[V_1,\widetilde{\widetilde{\widetilde{V_{2l+2}}}}\right],$$

when evaluated at 0, is included in the RHS of (8), too. Indeed, ∂_{y_ν} bracketed with V_1 gives the fields in $\left(\partial_{x_{\nu-1}},\partial_{y_{\nu-1}}\right)$, and ∂_{x_ν} bracketed with the generators of V_1 gives just x_sZ which vanishes at 0.

But, by virtue of the adopted procedure of reductions, the inclusion of $\widetilde{\widetilde{\widetilde{V_{2l+3}}}}(0)$ in the RHS of (8) is equivalent to the similar inclusion of $V_{2l+3}(0)$. Therefore (8) holds. That is, by Proposition 2, $V_{2l+3}(0)\subset L(D^{\nu-2})(0)$ when $\nu>2$, and $V_{2l+3}(0)\subset F(0)$ when $\nu=2$. That is to say, $j_s=3$.

The reasoning in the even situation $\star\star$ is analogous, with indices and exponents interweaving with those of the odd case $\star$. Theorem 1 is now fully proved. ■

A Proof of Proposition 4

For a given distribution D on a manifold M we take any singularity class $\mathcal{C}=j_1.j_2\dots j_r$ hit by (or: nonempty for) D. There can be many classes hit by D, if not necessarily all $\frac{1}{2}\left(1+3^{r-1}\right)$ existing in the length r under discussion. Then take any point p of the relevant singularity locus S. Here is an argument that S around p is an embedded submanifold of codimension stated in the proposition.

Proof. We take any fixed polynomial presentation for D around p (Theorem 4), being necessarily of the type $\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r$ (Theorem 1). Then, using the coordinate functions of this chosen EKR, the local equations of S around p (which now becomes 0) will be:

- $x_k = 0$ for all k such that $j_k = 2$,
- $x_s = y_s = 0$ for all s such that $j_s = 3$.

Indeed, Proposition 2 holds at all points. Hence, on analyzing the key congruence (4), it is the vanishing of x_l that is decisive for the inclusion to hold true in the l-th sandwich. This explains all the x-equations, coming from both types of letters: the $j_k = 2$ and the $j_s = 3$ in the code of $\mathcal{C}$. In other words, the x-equations are the equations of the locus of the sandwich class, say $\mathcal{E}$, which encompasses $\mathcal{C}$. Now some auxiliary equations, excising from $\mathcal{E}$ the singularity class $\mathcal{C}$, should be added to them.

Continuing then, any letter $j_s = 3$ in $\mathcal{C}$, upon analyzing the construction of the small flag of D^s, brings in the additional equation $y_s = 0$. Such is the *eventual* conclusion drawn from the arguments used in the proof of Theorem 1. We mean by this that not only the germ at 0 (previously our point p), but precisely the germs at all points q having the coordinates x_ν, x_s, y_s vanishing ($s = \nu + l + 1$ in the notation from Section 4), satisfy the inclusion $V_{2l+3}(q) \subset L(D^{\nu-2})(q)$, or $\subset F(q)$ when $\nu = 2$. That is, have a letter 3 at the s-th position in *their* codes, preceded by l letters 1, preceded in turn by a letter not 1 at the ν-th position. (Note that, naturally, each such variable x_ν is in the union of *all* variables x_k and x_s appearing in the present discussion.) This explains the y-equations joining the previous x-equations in our local description of S. ■

B Concrete example of discernment inside the sandwich class $1.\underline{2}.1.\underline{2}$

On the manifold $\mathbb{R}^{(4+1)2+1}(t, x_0, y_0, x_1, y_1, \dots, x_4, y_4)$ we will propose two non-equivalent families of EKRs, both sitting in the sandwich class $1.\underline{2}.1.\underline{2}$. One is called D and has (all members of the family) the singularity class $\mathcal{W}(D) = 1.2.1.2$, and the other is called E and has the class $\mathcal{W}(E) = 1.2.1.3$. Moreover, we will see the geometrical distinction between D and E at work. This time, for bigger transparency, we use the vertical writing of the most involved vector field's generators.

The first family of germs at $0 \in \mathbb{R}^{11}$ is generated by the following vector fields:

$$D = \left(\begin{matrix} x_4\begin{pmatrix} x_2\begin{pmatrix} 1\\ x_1\\ y_1 \end{pmatrix}\\ 1\\ y_2\\ X_3\\ Y_3 \end{pmatrix}\\ 1\\ Y_4\\ 0\\ 0 \end{matrix}, \partial_{x_4}, \partial_{y_4} \right).$$

Here, a variable x to the left of a bracket means (also in the sequel) multiplying by x all entries subsumed by that bracket. Thus, this is a 3-parameter family of rank-3 distributions (some of them might be pairwise equivalent because EKRs are only *pseudo*-normal forms). And the second 2-parameter family of germs (some of them might be pairwise equivalent as well) reads as follows:

$$E = \left(x_4\left(\begin{array}{c} x_2\left(\begin{array}{c}1\\ x_1\\ y_1\end{array}\right] \\ \left.\begin{array}{c}1\\ y_2\end{array}\right] \\ \left.\begin{array}{c}X_3\\ Y_3\end{array}\right] \\ \left.\begin{array}{c}y_4\\ 1\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \end{array}\right. , \partial_{x_4}, \partial_{y_4} \right).$$

Attention. These objects come directly from Theorem 1: on the level of local normal forms, precisely the EKR families **1.2.1.2** and **1.2.1.3** represent the sandwich class $1.\underline{2}.1.\underline{2}$ which is the union of the singularity classes 1.2.1.2 and 1.2.1.3.

It is straightforward (and based only on sandwich-like inclusions, cf. Corollary 3) that, in the process of constructing $\mathcal{W}(D)$ and $\mathcal{W}(E)$, both germs happen to belong to $1.\underline{2}.1.\underline{2}$. Then, in the passing from sandwich to singularity class(es), the first $\underline{2}$ from the left causes no trouble (see Section 3.5) while the specification of the second $\underline{2}$ is subtler.

For that second $\underline{2}$ in $1.\underline{2}.1.\underline{2}$, the values of the integers ν and l are $\nu = 2$ and $l = 1$. Because ν takes the smallest possible value, the covariant subdistributions of D^1 and E^1 (after the standard indexation of the big flags of D and E) enter into play. And, because we work with EKRs, these covariant objects are both equal to $F = \big(dt, dx_0, dy_0\big)^{\perp}$ (by Proposition 2, F is spanned by all the versors save ∂_t, ∂_{x_0}, ∂_{y_0}).

Let, for the sake of brevity, $D = D_1 \subset D_2 \subset D_3 \subset \cdots$ and $E = E_1 \subset E_2 \subset E_3 \subset \cdots$ be the respective small flags. Since $2l+3 = 5$, the algorithm of finding the singularity class requires to analyze the positions at 0 of D_5, E_5, and F. In order to do that it is helpful to watch carefully the early members D_2 and E_2:

$$D_2 = \left(x_4\left(\begin{array}{c} x_2\left(\begin{array}{c}1\\ x_1\\ y_1\end{array}\right] \\ \left.\begin{array}{c}1\\ y_2\end{array}\right] \\ \left.\begin{array}{c}X_3\\ Y_3\end{array}\right] \\ \left.\begin{array}{c}1\\ Y_4\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \end{array}\right. ,\ \begin{array}{c} x_2\left(\begin{array}{c}1\\ x_1\\ y_1\end{array}\right] \\ \left.\begin{array}{c}1\\ y_2\end{array}\right] \\ \left.\begin{array}{c}X_3\\ Y_3\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \end{array}, \partial_{x_3}, \partial_{y_3}, \partial_{x_4}, \partial_{y_4} \right), \quad E_2 = \left(x_4\left(\begin{array}{c} x_2\left(\begin{array}{c}1\\ x_1\\ y_1\end{array}\right] \\ \left.\begin{array}{c}1\\ y_2\end{array}\right] \\ \left.\begin{array}{c}X_3\\ Y_3\end{array}\right] \\ \left.\begin{array}{c}y_4\\ 1\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \end{array}\right. ,\ \begin{array}{c} x_2\left(\begin{array}{c}1\\ x_1\\ y_1\end{array}\right] \\ \left.\begin{array}{c}1\\ y_2\end{array}\right] \\ \left.\begin{array}{c}X_3\\ Y_3\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \\ \left.\begin{array}{c}0\\ 0\end{array}\right] \end{array}, \partial_{x_3}, \partial_{y_3}, \partial_{x_4}, \partial_{y_4} \right).$$

Note that the first generators on the left in these descriptions are superfluous for D_2 (E_2) as such. Yet we are to compute parts of the small flags of the departure objects (D and E), and the presence of these generators of D and E makes the necessary computations easier.

The main observation is that by multiplying the first and second generators in D_2 (E_2) we get $\partial_{x_2} + Y_4\partial_{y_2} \in D_3$ $\big(y_4\partial_{x_2} + \partial_{y_2} \in E_3\big)$, compare to (9). Because of the special role of the variable x_2 in D, this leads in two more steps to

$$\partial_t + x_1\partial_{x_0} + y_1\partial_{y_0} + Y_4\partial_{y_1} \in D_5.$$

This is just an instance of the inclusion (7) with $Z = \partial_t + x_1\partial_{x_0} + y_1\partial_{y_0}$, $\nu = 2$, $s = 4$, $V_5 = D_5$. Thus $D_5(0) \not\subset F(0)$. To distinguish between the two cases, the same operations performed for E lead to a benign inclusion $y_4 Z + \partial_{y_1} \in E_5$. Also, clearly, $x_4\partial_{x_2}, x_4\partial_{y_2} \in E_3$, which leads to another benign inclusion $x_4 Z \in E_5$. In consequence, the inclusion $E_5(0) \subset F(0)$ holds. (These statements are a simple instance of the main line of computations in the proof of Theorem 1.) Therefore, by the definition of singularity classes, the second $\underline{2}$ in the sandwich word $1.\underline{2}.1.\underline{2}$ is being specified: to 2 for D, and to 3 for E. And the first $\underline{2}$ is univocally replaced by 2. That is to say, the 3-parameter family of germs D is included in the singularity class 1.2.1.2, whereas the 2-parameter family E is included in the class 1.2.1.3.

C Proof of Proposition 1

Let $\operatorname{rk} D = n + 1$ and $\operatorname{cork} D = m > 1$.

Proof. The proof is local around any fixed point $p \in N$. In view of the Frobenius theorem there clearly exist local coordinates $x_0, x_1, \dots, x_m, y_1, y_2, \dots, y_n$ vanishing at p and such that

$$E = \big(dx_0, dx_1, \dots, dx_m\big)^{\perp} = \big(\partial_{y_1}, \partial_{y_2}, \dots, \partial_{y_n}\big) \tag{10}$$

and

$$D = \big(\partial_{x_0} + f_1\partial_{x_1} + \cdots + f_m\partial_{x_m}, E\big)$$

for certain functions f_i which may even be assumed vanishing at 0: $f_1(0) = \cdots = f_m(0) = 0$. Now the 'two-step' assumption $[D, D] = TN$ implies that $m \le n$ and

$$\operatorname{rk}\left(\frac{\partial f_i}{\partial y_j}\right)(0) = m,$$

with the indices' ranges $i = 1, \dots, m$, $j = 1, \dots, n$. Without loss of generality one can have the second range reduced to $j = 1, \dots, m$. After such a simplification the functions

$$\big(x_0, x_1, \dots, x_m, f_1, \dots, f_m, y_{m+1}, \dots, y_n\big)$$

are independent at 0 and we take them as new variables (for simplicty, we keep writing the letters x and y for the new variables). The purpose is twofold. Firstly, this coordinate change mapping, say ϕ, is clearly of the form $\phi(x_0, x_1, \dots, x_m, \dots) = (x_0, x_1, \dots, x_m, \dots)$, implying that the description (10) of E holds in both old and new variables. Secondly, the distribution D now gets an extremely simple description

$$D = \big(\partial_{x_0} + y_1\partial_{x_1} + \cdots + y_m\partial_{x_m}, \partial_{y_1}, \partial_{y_2}, \dots, \partial_{y_n}\big). \tag{11}$$

This in dual terms says

$$D^{\perp} = \big(dx_1 - y_1 dx_0, dx_2 - y_2 dx_0, \dots, dx_m - y_m dx_0\big)$$

and allows one to easily search for the covariant object. In fact, at each point close to $0 \in \mathbb{R}^{m+n+1}$ one is looking for all 1-forms α such that

$$\big(\alpha \wedge dx_0 \wedge dy_i\big)\big|_D = 0, \qquad i = 1, 2, \dots, m. \tag{12}$$

In the new coordinates the answer does not depend on a point. Indeed, upon writing

$$\alpha = a_0 dx_0 + a_1 dx_1 + \cdots + a_m dx_m + b_1 dy_1 + \cdots + b_m dy_m + \sum_{j=m+1}^{n} b_j dy_j,$$

one instantly observes that the coefficients $a_0, a_1, \dots, a_m$ are free, subject to no restrictions (for $dx_i|_D$, $i = 1, \dots, m$, are multiples of $dx_0|_D$). Concerning the coefficients b_j with $j > m$, they vanish identically, because the differentials $dx_0, dy_1, \dots, dy_n$ are free also after their restricting to D (cf. (11)). As it could be expected, the key coefficients are $b_1, \dots, b_m$. Because m is greater than 1, the conditions (12) imply that $0 = b_1 = \cdots = b_m$ identically[7]. In fact, taking $i = 1$ in (12) implies $b_2 = b_3 = \cdots = b_m = 0$ at the point under consideration. Taking $i = 2$ implies $b_1 = b_3 = \cdots = b_m = 0$ and already at this moment all coefficients $b_1, \dots, b_m$ are zero.

That is, $\alpha = a_0 dx_0 + a_1 dx_1 + \cdots + a_m dx_m$ and $a_0, a_1, \dots, a_m$ are free and this holds at every point. Therefore at every point the α's describe $\big(dx_0, dx_1, \dots, dx_m\big)^{\perp} = E$, and this is the covariant subdistribution $\widehat{D}$ of D.

Concerning the Cauchy-characteristic module $L(D)$, it is a classical fact going back to [21] (see Theorem 1 on p. 211 there) that for D under form (11) the Cauchy module is regular and

$$L(D) = \big(dx_0, dx_1, \dots, dx_m, dy_1, \dots, dy_m\big)^{\perp} = \big(\partial_{y_{m+1}}, \dots, \partial_{y_n}\big).$$

Indeed, then, in view of (10), $L(D)$ is of corank m inside E. Proposition 1 is now fully proved. ■

Acknowledgment

The author was supported by Polish Grant MNSzW N N 201 397 937.

References

[1] Adachi J., Global stability of special multi-flags, *Israel J. Math.*, to appear.

[2] Agrachev A.A., Feedback-invariant optimal control theory and differential geometry. II. Jacobi curves for singular extremals, *J. Dyn. Control Syst.* **4** (1998), 583–604.

[3] Bryant R.L., Chern S.S., Gardner R.B., Goldschmidt H.L., Griffiths P.A., Exterior differential systems, *Mathematical Sciences Research Institute Publications*, Vol. 18, Springer-Verlag, New York, 1991.

[4] Bryant R.L., Hsu L., Rigidity of integral curves of rank 2 distributions, *Invent. Math.* **114** (1993), 435–461.

[5] Cartan É., Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre, *Ann. Sci. École Norm. Sup. (3)* **27** (1910), 109–192.

[6] Cartan É., Sur l'équivalence absolue de certains systèmes d'équations différentielles et sur certaines familles de courbes, *Bull. Soc. Math. France* **42** (1914), 12–48.

[7] Engel F., Zur Invariantentheorie der Systeme von Pfaff'schen Gleichungen, *Berichte Ges. Leipzig Math.-Phys. Classe* **41** (1889), 157–176.

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

[9] Krasil'shchik I.S., Lychagin V.V., Vinogradov A.M., Geometry of jet spaces and nonlinear partial differential equations, *Advanced Studies in Contemporary Mathematics*, Vol. 1, Gordon and Breach Science Publishers, New York, 1986.

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

[11] Kumpera A., Ruiz C., Sur l'équivalence locale des systèmes de Pfaff en drapeau, in Monge–Ampère Equations and Related Topics, Editor F. Gherardelli, Ist. Naz. Alta Math. F. Severi, Rome, 1982, 201–248.

[12] Montgomery R., Zhitomirskii M., Geometric approach to Goursat flags, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **18** (2001), 459–493.

[13] Montgomery R., Zhitomirskii M., Points and curves in the monster tower, *Mem. Amer. Math. Soc.*, to appear, available at `http://www.tx.technion.ac.il/~mzhi/papers/`.

[7] It would be otherwise for $m = 1$, cf. [14, p. 165]. The dichotomy in the classical Lie–Bäcklund theorem has its roots precisely at this place.

[14] Mormul P., Multi-dimensional Cartan prolongation and special k-flags, in Geometric Singularity Theory, Editors H. Hironaka, S. Janeczko and S. Lojasiewicz, *Banach Center Publ.*, Vol. 65, Polish Acad. Sci., Warsaw, 2004, 157–178, also available at `http://www.mimuw.edu.pl/~mormul/special.pdf`.

[15] Mormul P., Geometric singularity classes for special k-flags, $k \geq 2$, of arbitrary length, in Singularity Theory Seminar, Editor S. Janeczko, Warsaw Univ. of Technology, Vol. 8, 2003, 87–100.

[16] Mormul P., Geometric classes of Goursat flags and the arithmetics of their encoding by small growth vectors, *Cent. Eur. J. Math.* **2** (2004), 859–883.

[17] Mormul P., Special 2-flags, singularity classes, and polynomial normal forms for them, *Sovrem. Mat. Prilozh.* (2005), no. 33, 131–145 (in Russian).

[18] Pasillas-Lépine W., Respondek W., Contact systems and corank one involutive subdistributions, *Acta Appl. Math.* **69** (2001), 105–128, math.DG/0004124.

[19] Shibuya K., Yamaguchi K., Drapeau theorem for differential systems, *Diff. Geom. Appl.*, to appear.

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

[21] von Weber E., Zur Invariantentheorie der Systeme Pfaff'scher Gleichungen, *Berichte Ges. Leipzig Math.-Phys. Classe* **50** (1898), 207–229.

[22] Yamaguchi K., Contact geometry of higher order, *Japan. J. Math. (N.S.)* **8** (1982), 109–176.

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

[24] Zelenko I., Fundamental form and the Cartan tensor of $(2, 5)$ distributions coincide, *J. Dynam. Control Syst.* **12** (2006), 247–276.