GEOMETRIC SINGULARITY THEORY
BANACH CENTER PUBLICATIONS, VOLUME 65
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2004

MULTI-DIMENSIONAL CARTAN PROLONGATION AND SPECIAL k–FLAGS

PIOTR MORMUL

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

Abstract. Since the mid-nineties it has gradually become understood that the Cartan prolongation of rank 2 distributions is a key operation leading locally, when applied many times, to all so-called Goursat distributions. That is those, whose derived flag of consecutive Lie squares is a 1-flag (growing in ranks always by 1). We first observe that successive *generalized* Cartan prolongations (gCp) of rank k+1 distributions lead locally to all *special* k-flags: rank k+1 distributions D with the derived flag $\mathcal{F}$ being a k-flag possessing a corank 1 involutive subflag preserving the Lie square of $\mathcal{F}$. (Note that 1-flags are always special.)

Secondly, we show that special k-flags are effectively nilpotentizable (or: *weakly* nilpotent) in the sense that local polynomial pseudo-normal forms for such D resulting naturally from sequences of gCp's give local nilpotent bases for D. Moreover, the nilpotency orders of the generated real Lie algebras can be explicitly computed by means of simple linear algebra (for $k = 1$ this was done earlier in [M1], [M3]). For $k = 2$ we also transform our linear algebra formulas into recursive ones that resemble a bit Jean's formulas [Je] for nonholonomy degrees of Goursat germs.

Additionally it is shown that, when all parameters appearing in a local form for a special k-flag vanish, then such a distribution germ is also *strongly* nilpotent in the sense of [AGau] and [M1].

1. Generalized Cartan prolongations. Through the work [BH], followed by [Je] and [MonZ], it has become known that the Goursat distributions, that naturally generalize *Cartan distributions* on the spaces of jets of functions $\mathbb{R} \to \mathbb{R}$, and at the same time bring in a rich pattern of singularities, are locally nothing more than the outcomes of so-called Cartan prolongations [defined rigorously and very geometrically only fairly recently in [BH] and, in a less constructive manner closer to the originating paper [Ca2], in [ShSl]]

2000 *Mathematics Subject Classification*: Primary 58A15, 17, 30; Secondary 17B30, 34C20.

Research supported by Polish KBN grant 2 P03A 010 22 and by EC IHP programme HPRN-CT-00271.

The paper is in final form and no version of it will be published elsewhere.

when one departs from the tangent bundle to a 2-manifold. This helped a lot to get new insights into Goursat distributions, their geometry and symmetries. The paper is devoted to Cartan-like, but more general than in [BH], prolongations of distributions, and to their applications: building more involved (having longer small and big growth vectors, or higher nonholonomy degrees, for instance) distributions from simpler ones.

1.1. *The basic* (*if homogeneous*) *example.* We want to start from one important example illustrating the substance of the paper—the Cartan distribution (or: contact system) on the space of 1-jets of functions $\mathbb{R}(t) \longrightarrow \mathbb{R}^2(x, y)$. It is a rank 3 distribution E living on $M = \mathbb{R}^5(t, x, y, x^1, y^1)$, generated by the vector fields

$$\Big(\frac{\partial}{\partial t} + x^1\frac{\partial}{\partial x} + y^1\frac{\partial}{\partial y}\,,\ \frac{\partial}{\partial x^1}\,,\ \frac{\partial}{\partial y^1}\Big), \tag{1}$$

where x^1 (y^1) can be regarded as dx/dt (dy/dt). This E is clearly 'two step' in the sense that $[E,\, E] = TM$; later this will be expressed by saying that the flag generated by E 'has length one'. One is in the realm of [Ca1], p. 121. Any such two step E possesses (and, at that, uniquely determined!) its corank 1 subdistribution F enjoying the property

$$[F,\, F] \subset E. \tag{2}$$

In [Ca1] that F was neatly defined by the Pfaffian equations (4) there (cf. also [GeVe] and p. 5 in [KuRub]). Cartan calls such an accompanying subdistribution F *le système covariant* of [the Pfaffian system] E. For the object (1), the covariant system is but $F = \big(\partial/\partial x^1,\, \partial/\partial y^1\big)$ and is even involutive, not just satisfying (2).

To have that rank 2 distribution F involutive is very and very rare. This is situation (a) in [Ca1]; among other advantages, such F helps to describe all symmetries of E. While typically the covariant object has its 'curvature' and allows to retrieve E in that $[F,\, F] = E$; this is situation (b) in [Ca1], extremely rich geometrically and investigated much further in that seminal work. We will say that a general such E (i.e., *no* particulars about the inclusion (2)) generates a 2-flag of length 1, while an E with its covariant system F involutive will generate a *special* 2-flag of length 1. Later we will deal with flags of, no wonder, bigger lengths, but they will likewise be only special—in the sense to be defined (in Section 3.2), and not general.

Pursuing further the differences between (a) and (b), it is observed in [Ca1] (see also a much posterior to it [GeVe]) that all E generating special 2-flags of length 1 are locally equivalent to (1). This implies that special 2-flags of length 1 are homogeneous: they look identically around any point, and hence feature no singularities at all; see also Proposition 1 (i) below. (We note in parentheses that, this notwithstanding, their *global* aspects, by far not evident, only start to attract researchers, [Ad2]. For 1-flags a similar situation occurs in length 2, with locally existing only one 'classical' behaviour (or model) of Engel 1889, and globally rich geometries being possible, started to be dealt with in [Ge], [Mon], [Ad1], ...)

In the guise of a forerunner, we say also that (1) is the outcome of a single (generalized) Cartan prolongation applied to the full tangent bundle to the manifold $\mathbb{R}^3(t, x, y)$. Later in the text we will see the results of several prolongations performed in row which will often *not* be homogeneous—will feature singular behaviour(s) alongside with the generic

one. Even two prolongations, one performed after the other, will turn out sufficient to generate singularities, cf. Proposition 1 (iii).

1.2. *The definition of the prolongation.* We assume in the sequel that all discussed objects are C^∞ and propose a more general prolongation scheme than the one discussed in [BH]—*generalized* (or: multi-dimensional) *Cartan prolongation* (gCp). Its main application, spread over Sections 4–5, is the effective local (weak in the sense of [M1]) nilpotency of a quite rich family of objects related to and generalizing abstract Cartan distributions (or: contact systems in a parallel terminology) appearing in the geometric theory of ODEs (and, although not within the scope of the present paper, PDEs; see in this respect [KLV] or [BCG$^+$]). This extends considerably the range of known [weakly] nilpotent distributions.

In short retrospection, among others, the papers [HeLuSul], [LaSus], and later also [Mu], aroused interest in the search for nilpotent distributions and for possible underlying structure theorems. Goursat distributions—quite particular subbundles of the tangent bundle having the flag of consecutive Lie squares regular and growing in ranks (slowly!) always by one—are among such objects. Nilpotent bases with (what is of key importance for applications) precisely computed *nilpotency orders*([1]) of the induced real Lie algebras, were proposed in [M1].

In fact, those bases were first put forward in a much different context in [KuRui]; later they have been re-found by applying many times an important theorem, recalled in Section 2 below as Theorem 1. In [BH], in a comment to that theorem we read: 'Indeed it is the basis of the proof of Goursat's theorem.'

So what are: Cartan prolongation evoked in that theorem, and gCp needed for a new theorem appearing in the next section of the present work? In the following definition, excerpted from [BH], p. 454^{4-10}, one obtains the *definition of generalized Cartan prolongation* by simply replacing 'rank 2' by 'rank k+1', '2-dimensional' by '(k+1)-dimensional', '$\mathbb{PR}^1$' by '$\mathbb{PR}^k$', and '1-manifolds' by 'k-manifolds' (another definition, more vague if probably better suited for applications in differential equations and control theory, is given in [ShSl]). One more minor change is that later, having some further generalizations (still in progress) in mind, we will write $\pi : \mathrm{Gr}(D, 1) \longrightarrow M$ instead of $\pi : \mathbb{P}\mathcal{D} \longrightarrow M$:

'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 $\mathbb{P}\mathcal{D}_m$ is the space of 1-dimensional linear subspaces of the 2-dimensional vector space $\mathcal{D}_m$. 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 $\xi \subset \mathcal{D}_m$. The distribution $\mathcal{D}^{(1)}$ is called the (*first*) *prolongation* of $\mathcal{D}$.'

Remark 1. Certain ingredients of this definition of gCp are dispersed in the literature, like for instance in [Y2], p. 34 or [BCG$^+$], p. 148–9, or else [Gr], par. 4.1.C. Yet, to our knowledge, they are always accompanied by so-called *independence conditions* forced

([1]) i.e., minimal numbers of Lie multiplications in an algebra yielding automatically zero, cf. p. 239 in [LaSus].

by the impending applications to the geometry of PDEs, or by explicit conditions that the directions ξ be not arbitrary (not vertical in certain sense, as on p. 34^{6-8} in [Y2]) often restricting a prolongation to the domain of just one affine chart. Or—in [Gr]—allowing for different 'charts' (different submanifolds V_0' in that work) in the first starting manifold (V_0), but then considering only jets of graphs of mappings in such an arbitrarily fixed chart([2]). While the essence of the singularities that are hidden in 1-flags, and more generally in special k-flags (defined later in Section 3.2), consists in arbitrary, also vertical, positions of ξ being *allowed*. And allowed also many times either in row or intermittently, etc. To support this comment, one sees for instance in the summary of the results of an important work [Y1] (p. 110) that the canonical distributions C^r constructed there on the jet bundles of functions $\mathbb{R}^n \to \mathbb{R}$ are *regular* in the sense of Tanaka [T]—have the small growth vectors independent of points in the bundles. For 1-flags (n then should be taken 1) it is precisely—see [KuRui], [Mu]—the simplest and generic situation singling out the geometry of the abstract Cartan distributions on spaces of jets of functions $\mathbb{R} \to \mathbb{R}$. That is, keeping singularities at bay by singling out, at all prolongation steps, only non-vertical directions.

Acknowledgments. We wish to express our gratitude to Antonio Kumpera who stressed the importance of multi-flags for underdetermined differential equations, sent us in 1999 the first version of [KuRub], and with whom we discussed multi-flags on several occasions, if only by correspondence.

2. Structural theorem. A sketch of the following theorem, attributed to Élie Cartan, is given in [BH] (and, to an extent, in [Sl] where, however, independence conditions evoked in Remark 1 pervade the exposition). In its wording, and in the present text, too, D_1 means the Lie square $[D, D]$ of a distribution D, and D_2—the Lie square of D_1. A foliation $\mathcal{F}$ appearing in Theorem 1 is a classical object closely related to the hypothesis on the deficient rank of D_2 (4 instead of 5). Finally, for consistency reasons, we write $s+1$ instead of $s+2$ (and, consequently, $\mathrm{Gr}(D', 1)$ instead of $\mathbb{P}D'$).

THEOREM 1 (Cartan-Bryant-Hsu). *Let D be a rank 2 distribution on a manifold M^{s+1} and suppose that D_1 and D_2 have ranks 3 and 4, respectively. Furthermore, suppose that there is a submersion $f : M \to N^s$ 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 D' on N with the property that $D_1 = f^*(D')$ and, moreover, there exists a canonical smooth map $f^{(1)} : M \longrightarrow \mathrm{Gr}(D', 1)$ which is a local diffeomorphism, which satisfies $f = \pi \circ f^{(1)}$, and which satisfies $f^{(1)}_* D = (D')^{(1)}$.*

([2]) At one moment, p. 242^{1-4}, Gromov steps out of the framework of integral manifolds V', of [the Cartan distribution] H^r on the r-jets' prolongation manifold V^r, projecting diffeomorphically onto some submanifold $V_0' \subset V_0$. For $V' \subset V^r$ projecting badly he proposes to keep track of 'how badly' by looking at [certain] s-th jets of V' that are represented by points in $\overline{V}^{r+s}$. And fears complications from possible singularities of $\overline{V}^{r+s}$ at points of $\overline{V}^{r+s} \backslash V^{r+s}$. In our approach, that is to say, only for $\dim V_0' = 1$, the manifolds V^r do have smooth closures—just towers of projective spaces' bundles over V_0. Yet the distributions H^r, when extended to those closures (as regular constant rank objects), start to feature pertinent to them, if covered, singularities.

This theorem has wide applications, and most specifically to Goursat flags (see Remark 3 in Section 3 below). Our objective is to formulate a local structural theorem generalizing Cartan's theorem and having even wider applications. The assumptions in Theorem 1 can be *rephrased*, avoiding mentioning D_2 and placing the foliation $\mathcal{F}$ in a new context. In fact, these assumptions easily imply that there exists a (unique, besides) line subdistribution E of D preserving D_1, $[E,\, D_1] \subset D_1$. The foliation $\mathcal{F}$ is the integral of $E(^3)$. Motivated by gCp's constructed in Section 1 (hence, eventually by the seminal [BH]) and, to an extent, by the works [KuRub] and [PR], we are going to replace a line subdistribution of a rank 2 distribution by an involutive rank k subdistribution of a rank $k{+}1$ one (that is, by its corank 1 involutive subdistribution).

THEOREM 2. *Suppose D is a rank $k{+}1$ distribution on a manifold M^{s+k} such that*

a) *D_1 is a rank $2k+1$ distribution on M,*

b) *there exists a corank 1 involutive subdistribution $E \subset D$ that preserves D_1, $[E,\, D_1] \subset D_1$.*

Then locally D is 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 k-dimensional foliation in M defined by E).

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

Proof. We work all the time locally around an arbitrarily chosen point $p \in M$ and use a notation nearly identical to those appearing in [BH], Section 4.2. A far reaching affinity of proofs should facilitate the comparison of the two theorems.

The (smooth) projection mapping $f : M \longrightarrow N = M/\mathcal{F}$ serves to define $D' = f_* D_1$. By the assumptions on E, this is a well defined rank $k{+}1$ distribution on N ('D_1 descends to N' in the language of [Mon]). Then $\pi : \mathrm{Gr}(D', 1) \longrightarrow N$ is as in the definition of gCp, and a smooth mapping between manifolds of the same dimension,

$$f^{(1)} : (M, p) \longrightarrow \big(\mathrm{Gr}(D', 1),\, df(p)D(p)\big),$$

is defined by

$$f^{(1)}(q) = df(q)D(q). \tag{3}$$

So $f^{(1)}(q)$ is a line in the $(k{+}1)$-dimensional space $D'(f(q))$. From definition (3) it follows directly that

$$\pi \circ f^{(1)} = f. \tag{4}$$

The theorem will be proved upon showing that $f^{(1)}$ is a local diffeo sending D to $(D')^{(1)}$. We will first explain that $df^{(1)}(p)$ is an isomorphism. Clearly, since f is a submersion, in view of (4),

(5) $\quad \mathrm{im}\, df^{(1)}(p)$ contains an s-dimensional subspace transverse to $\ker d\pi(f^{(1)})(p)$.

A more difficult question is whether the image of $df^{(1)}(p)$ *contains* that kernel. Indeed,

(^{3}) These objects were discovered many decades ago in [Ca2], then were often rediscovered and recently given new (and surprising) interpretations in [BH] and [Z].

LEMMA 1. $\ker d\pi(f^{(1)}(p))$ *is included in* $\operatorname{im} df^{(1)}(p)$.

Proof of the lemma. This key part is technically more involved than in [BH] because the projective spaces being the fibers are now k-dimensional instead of 1-dimensional. Maybe not surprisingly, arguments are clearer in local coordinates. Also, the necessary geometric constructions are easier to describe in properly chosen coordinates.

Take a local basis of E of pairwise commuting vector fields $v_1, \dots, v_k$, and introduce local 'Frobenius' coordinates $x_1, \dots, x_k, x_{k+1}, \dots, x_{k+s}$, say centered at p, such that $v_i = \partial_i$, where (*here and in the sequel*) we use the shorthand notation $\partial_i = \partial/\partial x_i$. Visualise further by putting $N = \{x_1 = \ldots = x_k = 0\} \subset M = \mathbb{R}^{k+s}$. Put also

$$N_i = \{x_1 = \ldots = x_{i-1} = x_{i+1} = \ldots = x_k = 0\}$$

and

$$N_{i,\epsilon} = \{x_1 = \ldots = x_{i-1} = x_i - \epsilon = x_{i+1} = \ldots = x_k = 0\}.$$

Then the intersections

$$D \cap TN, \quad D \cap TN_{i,\epsilon}, \quad i = 1, \dots, k, \tag{6}$$

are fields of lines on N, $N_{i,\epsilon}$ respectively (they are at least one-dimensional as the sums of dimensions of factors are $k + s + 1$, and E, of corank 1 in D, is transverse to TN, $TN_{i,\epsilon}$). Let L be any smooth vector field, with values in D and everywhere independent of $v_1, \dots, v_k$, spanning the fields of lines (6). Clearly, such an L (and even its direction, when $k > 1$) is not defined uniquely.

By the hypotheses made on E and D_1, the vector fields $L, v_1, \dots, v_k, [L, v_1], \dots, [L, v_k]$ form a local basis of D_1. In these coordinates

$$f(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+s}) = (0, \dots, 0, x_{k+1}, \dots, x_{k+s}).$$

Define $f_i = f|_{N_i}$. For $q \in N_i$, $df_i(q)L(q)$ is a non-zero vector included in $df(q)D(q) = f^{(1)}(q)$. Hence

$$f^{(1)}(q) = \operatorname{span}\big(df_i(q)L(q)\big) \quad \text{for } q \in N_i, \quad i = 1, \dots, k.$$

Now it is visible how the line $f^{(1)}(q)$ *rotates* inside $D'(f(p))$ when q approaches p along the integral line of v_i through p. According to the fisherman derivative' definition of the Lie bracket watched—it is important—in these special coordinates,

$$\frac{d}{dt}\, df_i(p + tv_i)L(p + tv_i)\Big|_{t=0} = [L, v_i](p). \tag{7}$$

Attention. The mapping f_i is not, naturally, the time $-t$ flow of v_i, but, in (7), df_i acts on L exactly as that flow does.

The vectors on the right hand side of (7) sit in T_pN, as all f_i take values in N, and—we recall—are linearly independent. (Saying differently, the $(k+1)$-dimensional space $D'(p)$ is spanned by the vectors $L, [L, v_1], \dots, [L, v_k]$ at p.) Therefore, the evolutions of the lines $f^{(1)}(q)$ inside the space $D'(p)$, when q approaches p in the directions of $v_1, \dots, v_k$—are all differentially independent. If we see these lines as points in $\operatorname{Gr}(D', 1)(f(p))$, the *velocities* of the relevant curves in that projective space, naturally sitting in $\operatorname{im} df^{(1)}(p)$, are also linearly independent.

On the other hand, all these k curves are mapped by π to a single point $f(p)$. Thus their velocities sit in $\ker d\pi(f^{(1)}(p))$ which is exactly of dimension k. Being independent, they span that space.

The lemma is now proved. ∎

The information (5) together with Lemma 1 yield that $df^{(1)}(p)$ is a linear surjection between two spaces of the same dimension $k+s$, hence is an isomorphism. Thus (remember that the proof goes for *any* point p fixed in M)

(8) $\qquad f^{(1)}$ is a local diffeomorphism in a neighborhood of $p \in M$.

Now we will explain that

(9) $\qquad df^{(1)}(p)$ carries $D(p)$ to $D'^{(1)}(f^{(1)}(p))$.

Due to the equality of dimensions, it suffices to show that

(10) $\qquad df^{(1)}(p)D(p) \subset D'^{(1)}(f^{(1)}(p))$,

that is (see the definition of gCp in Section 1), that

(11) $\qquad df^{(1)}(p)D(p) \subset \big(d\pi(f^{(1)}(p))\big)^{-1} f^{(1)}(p)$.

The right hand side in (11) is equal to

$$\big(d\pi(f^{(1)}(p))\big)^{-1} df(p)D(p) = \big(d\pi(f^{(1)}(p))\big)^{-1} d\pi(f^{(1)}(p)) df^{(1)}(p)D(p) \supset df^{(1)}(p)D(p),$$

because $df^{(1)}(p)D(p)$ is clearly included in the domain of $d\pi(f^{(1)}(p))$—the latter being the whole tangent space to $\mathrm{Gr}(D', 1)$ at $f^{(1)}(p)$. So (9) holds. With (8) and (9) justified, Theorem 2 is proved. ∎

3. Special k-flags and extended Kumpera-Ruiz pseudo-normal forms for them. Before the precise definition we want to give certain motivation. *Exactly as* 1-flags (starting, to avoid redundancy, from a rank 2 object, cf. for instance [MonZ], p. 462) take the main property of Cartan distributions on the jet spaces of functions $\mathbb{R} \to \mathbb{R}$ as their only *defining* property—see an extensive discussion of that in [Ku], *likewise* special k-flags (starting from a rank $k+1$ object to avoid redundancy) will borrow only similar two main properties of Cartan distributions on the jet spaces of functions $\mathbb{R} \to \mathbb{R}^k$ as their defining properties. For 1-flags, we mean

a) the regular big growth vector being $[2, 3, 4, \dots]$,

and, formally, also

b) the presence of a corank 1 involutive subflag preserving the (Lie) square of the original flag([4]).

For special k-flags, we mean

i) the big growth vector being everywhere $[k+1, 2k+1, 3k+1, \dots]$ (a distribution generates a k-flag),

([4]) The Cauchy-characteristic subflag of a Goursat flag $\mathcal{F}$, accompanied by one more involutive distribution sitting in the corank 1 member of $\mathcal{F}$ (and not uniquely defined), does enjoy b). So, only for 1-flags, this b) does not add any new condition. We write it down to have later a uniform definition for all values of k, and just to notice that 1-flags are automatically special.

ii) the presence of a corank 1 involutive subflag preserving the square of the original flag.

These two conditions together lead to a rich but still tractable world of objects, with local pseudo-normal forms with plenty of real parameters. The condition i) alone is very vast, and even the short vector $[3,5]$ ($k=2$, flag's length 1) hides really wild, not well understood geometries with a functional module in the local classification (firstly rigorously evidenced in [JaPrz]). Distributions having this growth vector were extensively investigated by É. Cartan in his famous 'cinq variables' paper [Ca1]; cf. also Section 1.1. The condition ii) is, to the contrary, very restrictive, perfectly balancing i).

3.1. *Terminological note.* It is to be pointed out that 'Cartan distributions' in the prevailing terminology of [KLV], are also called 'canonical (differential) systems on jet bundles': in [Y1]—for the jets of one function of many variables, and in [Y2]—in the general case of several functions of many variables. When the jets of k functions of just one variable are concerned, Kumpera and Rubin say in [KuRub] about 'k-flags satisfying normality conditions', while those distributions are called 'regular contact systems for curves' in [PR].

As regards special k-flags generalizing the latter, in [PR] they are (somehow misleadingly) called 'contact [regular or singular] systems for curves'. In the optics of the present paper (and, ultimately, [BH]), the verticality of lines ξ, allowed in the definition in Section 1 at each prolongation step, goes well beyond the contact context, enriching enormously possible geometries of distributions—that can be so prolonged many times, sometimes intermittently with contact (i.e., non-vertical) prolongations. See also Remark 1 above.

Now we give the precise definition.

3.2. *The definition of a special k-flag.* A distribution D of rank $k+1$ on a manifold M of dimension $(r+1)k+1$ generates a special k-flag of length r, $r\geq 1$, when its consecutive Lie squares

$$D=D^{k+1}\subset D^{2k+1}\subset\ldots\subset D^{rk+1}\subset D^{(r+1)k+1}=TM \tag{12}$$

are all regular distributions of the respective ranks written as superscripts in (12) (this is i)) *and* when there exists a subflag of involutive distributions

$$E^k\subset E^{2k}\subset\ldots\subset E^{rk}$$

of ranks written also as superindices, such that $E^{jk}\subset D^{jk+1}$ for $j=1,\ldots,r$, and E^{jk} preserves $D^{(j+1)k+1}$, $[E^{jk},\,D^{(j+1)k+1}]\subset D^{(j+1)k+1}$ for $j=1,\ldots,r-1$ (this is ii)).

It is useful to write down the diagram of inclusions that are assumed to hold in this basic definition:

$$\begin{array}{ccccccccccc} D^{k+1} & \subset & D^{2k+1} & \subset\ldots\subset & D^{(r-1)k+1} & \subset & D^{rk+1} & \subset & D^{(r+1)k+1}=TM \\ \cup & & \cup & & \cup & & \cup & & \\ E^k & \subset & E^{2k} & \subset\ldots\subset & E^{(r-1)k} & \subset & E^{rk} & & \end{array}$$

Remark 2. In the work [PR] the definition of contact systems (for curves, cf. Section 3.1) is weaker, replacing ii) by only the existence of corank 1 involutive subdistri-

butions $E^{jk} \subset D^{jk+1}$, $j = 1, \dots, r$. Yet we are going to show in this section (Proposition 1 (iv) and Corollary 1) that our formally more restrictive definition yields strictly more, not less, objects than produced in [PR]. Note, therefore, that the inferential schemes in [PR], and specifically in Theorem 2.3 there, must be flawed.

On the other hand, one can quickly show that E^{rk}, the biggest among involutive distributions stipulated in our definition, is *unique*. It can be done via answering a (recent) question of Zhitomirskii: given a two-step distribution D on M, $[D, D] = TM$, possessing a corank one involutive subdistribution $E \subset D$, is it true that E is the *covariant subdistribution* of D, in the precise sense of [KuRub], p. 5 ?

The answer is yes, provided $\operatorname{cork} D > 1$, and inside E there sits then the *Cauchy characteristic* distribution $L(D)$ of D, having the rank

$$\operatorname{rk} L(D) = \operatorname{rk} E - \operatorname{cork} D. \tag{13}$$

(For D of corank 1 the covariant object $\widehat{D}$ is smaller—it coincides with $L(D)$ that still sits in E, relation (13) keeps holding, but E in the question is *not* unique.) In the setting of special k-flags, $\operatorname{cork} D^{rk+1} = k > 1$, hence $E^{rk} = \widehat{D^{rk+1}}$, the covariant subdistribution of D^{rk+1}. Cf. also Section 1.1.

3.3. *Theorem* 2 *produces extended Kumpera-Ruiz pseudo-normal forms for special k-flags.* We will produce a huge variety of polynomially written germs at $0 \in \mathbb{R}^N$, N possibly very large and always $1 \pmod k$, of rank $k{+}1$ distributions. Often certain variables x will appear in them in a 'shifted' form $X = c + x$. And always a capital letter X, typically with indices, will mean such a shifted variable, not excluding, incidentally, the value $c = 0$. For each $m \in \{1, 2, \dots, k+1\}$ we are going to define an operation $\mathbf{m}$ producing new rank $k{+}1$ distributions from older ones. The technical writing of its outcome, not operation's formal definition, will depend on how many operations were done *before* $\mathbf{m}$.

More specifically, the outcome of $\mathbf{m}$ being performed as operation number l on a distribution $(Z_1, \dots, Z_{k+1})$ defined in the vicinity of $0 \in \mathbb{R}^s(y_1, \dots, y_s)$ is a new rank $k{+}1$ distribution defined in the vicinity of $0 \in \mathbb{R}^{s+k}(y_1, \dots, y_s, x_1^l, \dots, x_k^l)$, generated by the vector field

$$Z_1' = x_1^l Z_1 + \ldots + x_{m-1}^l Z_{m-1} + Z_m + X_m^l Z_{m+1} + \ldots + X_k^l Z_{k+1}$$

and by $Z_2' = \partial/\partial x_1^l, \dots, Z_{k+1}' = \partial/\partial x_k^l$. It is important that these local generators are written precisely in this order, yielding together a new longer object $(Z_1', Z_2', \dots, Z_{k+1}')$.

For instance, when $m = 1$, the sum Z_1' starts with Z_1 and all new incoming variables in Z_1' are shifted. When $m = k+1$, the sum Z_1' ends with Z_{k+1} and all new variables are bare (with no shifting constants).

Extended Kumpera-Ruiz pseudo-normal forms (EKR for short)([5]), of length $r \geq 1$, denoted by $\mathbf{j}_1 . \mathbf{j}_2 \dots \mathbf{j}_r$, where $j_1, \dots, j_r \in \{1, 2, \dots, k+1\}$, are defined inductively, starting from the empty label distribution

$$\big(\partial_0, \partial_1, \dots, \partial_k\big) = \Big(\frac{\partial}{\partial t}\,, \frac{\partial}{\partial x_1^0}\,, \dots,\, \frac{\partial}{\partial x_k^0}\Big)$$

([5]) The name borrowed from [PR], but meaning here much more geometrically different objects than in that work, cf. Proposition 1 and Corollary 1.

in the vicinity of $0 \in \mathbb{R}^{k+1}(t, x_1^0, \dots, x_k^0)$. Then, if we assume that $\mathbf{j}_1 \dots \mathbf{j}_{r-1}$ is already known, $\mathbf{j}_1 \dots \mathbf{j}_{r-1}.\mathbf{j}_r$ is the outcome of the operation $\mathbf{j}_r$ performed as the operation number r over the distribution $\mathbf{j}_1 \dots \mathbf{j}_{r-1}$.

The reader probably feels already that the operations $\mathbf{1}$, $\mathbf{2}$, …, $\mathbf{k+1}$ are certain prolongations. In fact, they are just different gCp's viewed in possible different affine charts on the Grassmanians used in the gCp procedure. This will be explained in details in the proof of Theorem 3 below. When $k = 1$, the two operations $\mathbf{1}$ and $\mathbf{2}$ applied interweavingly lead to the well-known local Kumpera-Ruiz pseudo-normal forms, [KuRui], for Goursat flags (to be precise: in certain raw, not polished form, cf. Proposition 1 below in this respect).

For a moment, it is nearly directly visible that every EKR is a special k-flag of length equal to the number of operations used to produce it (and equal to the length of the relevant word encoding the sequence of operations). In particular, it is easy to guess what are, for any EKR of length r, the involutive subdistributions of ranks k, $2k, \dots, rk$. The gist of the matter is that locally the converse is also true, and we have

THEOREM 3. *Let a rank $k+1$ distribution D generate a special k-flag, $k \geq 1$, of length $r \geq 1$ on a manifold $M^{(r+1)k+1}$. For every point $p \in M$, D in a neighborhood of p is equivalent, by a local diffeomorphism that sends p to 0, to a certain* EKR $\mathbf{j}_1.\mathbf{j}_2 \dots \mathbf{j}_r$ *in a neighborhood of $0 \in \mathbb{R}^{(r+1)k+1}$. Moreover, that* EKR *can be taken such that $j_1 = 1$ and, for $l = 1, \dots, r-1$, if $j_{l+1} > \max(j_1, \dots, j_l)$, then $j_{l+1} = 1 + \max(j_1, \dots, j_l)$ (the rule of the least possible* new *jumps upwards in the words $\mathbf{j}_1.\mathbf{j}_2 \dots \mathbf{j}_r$).*

Note, however, that possible constants in such an EKR representing a given germ D are not, in general, defined uniquely. On the other hand, we do not yet know if a given distribution germ determines uniquely—its proper?—sequence of operations $\mathbf{j}_1, \dots, \mathbf{j}_r$ satisfying the rule of the least upward jumps([6]). It certainly does when $k = 1$, see Proposition 1 (ii).

Proof. The proof is by induction on r, starting from $r = 0$ (not mentioned in theorem's wording) where the empty word EKR $(\partial_0, \partial_1, \dots, \partial_k)$ given in the vicinity of $0 \in \mathbb{R}^{k+1}$, is, naturally, a local model for the full tangent bundle to a $(k+1)$-dimensional manifold. In the sequel we consistently use the notation introduced in Section 3.2.

Let us assume the fact under proof for a value $r-1 \geq 0$ (as remarked, in the beginning of induction for $r - 1 = 0$ we have from the outset all the necessary information), take a distribution D^{k+1} generating a special k-flag of length r together with its involutive subdistribution E^k, and apply to them Theorem 2. Then, watching the proof of *that* theorem, (D^{k+1}, p) is equivalent to the generalized Cartan prolongation

$$\left(\left(D^{2k+1}/E^k \right)^{(1)}, \left(df(p) D^{k+1}(p), f(p) \right) \right), \tag{14}$$

where f stands for the local submersion—passing to the quotient objects modulo E^k and its k-dimensional foliation $\mathcal{F}$. Since E^k preserves D^{2k+1}, it preserves also all members of the flag of D^{2k+1}. Also it preserves the involutive distributions $E^{2k}, \dots, E^{rk}$ (being

([6]) Added in proof: yes, it does; in [M4] we construct such a sequence and call it the *singularity class* of a special k-flag at a point.

their involutive subdistribution). Therefore, D^{2k+1}/E^k is a special k-flag of length $r-1$ on the quotient manifold $M/\mathcal{F}$ of dimension $rk+1$ (well defined locally around $p \in M$, considered locally around its point $f(p)$). By the inductive assumption, there exists a local diffeomorphism $\psi : (M/\mathcal{F},\, f(p)) \longrightarrow (\mathbb{R}^{rk+1},\, 0)$ such that

$$\psi_* \left(D^{2k+1}/E^k\right) \text{ is an EKR } \mathbf{j}_1 \dots \mathbf{j}_{r-1} = (Z_1, \dots, Z_{k+1}) \text{ on } (\mathbb{R}^{rk+1},\, 0) \tag{15}$$

satisfying the rule of the least upward jumps (recall that for $r = 1$, in (15) there stands an 'empty' EKR—the full tangent bundle to $\mathbb{R}^{k+1}$). This says, among others, that ψ_* sends the line directions in D^{2k+1}/E^k to directions in $(Z_1, \dots, Z_{k+1})$ inducing the *mapping* prolongation

$$\psi^\bullet : \mathrm{Gr}\big(D^{2k+1}/E^k,\, 1\big) \longrightarrow \mathrm{Gr}\big((Z_1, \dots,\, Z_{k+1}),\, 1\big)$$

of ψ, being a local diffeomorphism of these Grassmannian bundles. (Recalling, we are interested in $\psi^\bullet$ in a neighborhood of the line $df(p)D^{k+1}(p)$ at $f(p)$.) Moreover, and this is a general easily verifiable fact, this mapping prolongation sends the gCp's one to the other,

$$\psi^\bullet_* : \left(D^{2k+1}/E^k\right)^{(1)} \longrightarrow (Z_1, \dots, Z_{k+1})^{(1)}. \tag{16}$$

(The prolongation of a conjugating diffeo locally conjugates the relevant gCp's. For $k = 1$, see also Appendix in [M2] in this respect.) Returning to (15), the line $df(p)D^{k+1}(p)$ is sent by $d\psi(f(p))$ to a line

$$\mathrm{span}(Z_j + c_j Z_{j+1} + \dots + c_k Z_{k+1})(0), \tag{17}$$

and $1 \leq j \leq k+1$ is uniquely defined. This integer j decides what prolongation operation is bound to be performed over $\mathbf{j}_1 \dots \mathbf{j}_{r-1}$ (one will instantly see that it is the operation $\mathbf{j}$). What remains to be done in the proof is: *firstly*, to explicitly write down the gCp $(Z_1, \dots, Z_{k+1})^{(1)}$ in the vicinity of the line (17), and to identify it as a, longer, EKR. And *secondly*, to take care that the eventual EKR constructed satisfies the rule of the least upward jumps.

By introducing new 'angle affine' coordinates $x^r_1, \dots, x^r_{j-1}$, $X^r_j = c_j + x^r_j, \dots, X^r_k = c_k + x^r_k$, one can parametrize all the line directions in $(Z_1, \dots, Z_{k+1})$ being close to the direction (17) by k free, floating around 0, parameters $(x^r_1, \dots, x^r_k) \in \mathbb{R}^k$. To have a gCp, one has to add to the current direction the whole kernel $\ker d\pi$ of the differential of the projection $\pi : \mathrm{Gr}\big((Z_1, \dots, Z_{k+1}),\, 1\big) \longrightarrow \mathbb{R}^{rk+1}$. That is to say, add the versors of the new angle coordinates. In this way, in the vicinity of the line (17), $(Z_1, \dots, Z_{k+1})^{(1)}$ is equal to

$$\left(x^r_1 Z_1 + \dots + x^r_{j-1} Z_{j-1} + Z_j + X^r_j Z_{j+1} + \dots + X^r_k Z_{k+1}\,,\, \frac{\partial}{\partial x^r_1}, \dots,\, \frac{\partial}{\partial x^r_k}\right). \tag{18}$$

This, indeed, is an EKR $\mathbf{j}_1 \dots \mathbf{j}_{r-1}.\mathbf{j}$ understood in the vicinity of $0 \in \mathbb{R}^{rk+1+k}$. Thus D^{k+1} is locally equivalent to the distribution (14) which is equivalent—see (16)—to a distribution identified as a $\mathbf{j}_1 \dots \mathbf{j}_{r-1}.\mathbf{j}$. Now we pass to the least upward jumps rule.

When $r = 1$, it is easy to see that the obtained $\mathbf{j}$ is equivalent to certain $\mathbf{1}$: if $j > 1$ then there does to cyclicly change the first j coordinates $t \to x^0_1 \to x^0_2 \to \dots \to x^0_{j-1} \to t$

(and keep the remaining ones untouched). This change replaces the first generator in our $\mathbf{j}$,

$$x_1^1\frac{\partial}{\partial t}+x_2^1\frac{\partial}{\partial x_1^0}+\ldots+x_{j-1}^1\frac{\partial}{\partial x_{j-2}^0}+\frac{\partial}{\partial x_{j-1}^0}+X_j^1\frac{\partial}{\partial x_j^0}+\ldots,$$

by the first generator in a $\mathbf{1}$,

$$x_1^1\frac{\partial}{\partial x_1^0}+x_2^1\frac{\partial}{\partial x_2^0}+\ldots+x_{j-1}^1\frac{\partial}{\partial x_{j-1}^0}+\frac{\partial}{\partial t}+X_j^1\frac{\partial}{\partial x_j^0}+\ldots.$$

When $r>1$ and $j>1+m=1+\max(j_1,\ldots,j_{r-1})$, then one argues in a similar way that the obtained EKR $\mathbf{j}_1\ldots\mathbf{j}_{r-1}.\mathbf{j}$ is equivalent to certain $\mathbf{j}_1\ldots\mathbf{j}_{r-1}.\mathbf{m{+}1}$ (satisfying the least upward rule by the inductive assumption).

Indeed, begin with a cyclic change of coordinates $x_m^{r-1}\to x_{m+1}^{r-1}\to\ldots\to x_{j-1}^{r-1}\to x_m^{r-1}$. Then the first vector field in (18),

$$x_1^rZ_1+\ldots+x_m^rZ_m+x_{m+1}^rZ_{m+1}+\ldots+x_{j-1}^rZ_{j-1}+Z_j+X_j^rZ_{j+1}+\ldots,$$

assumes the form

$$x_1^r[Z_1]+\ldots+x_m^rZ_m+x_{m+1}^rZ_{m+2}+\ldots+x_{j-1}^rZ_j+Z_{m+1}+X_j^rZ_{j+1}+\ldots,$$

where, as previously, $Z_l=\partial/\partial x_{l-1}^{r-1}$, $l=2,\ldots,k+1$, and $[Z_1]$ is Z_1 in the actual coordinates. We are not yet done because $[Z_1]$ is not like in the EKRs construction. In fact, in view of our basic underlying relation $j_{r-1}\le m$, the middle part of that vector field reads

$$[Z_1]\;=\;\ldots+X_{m+1}^{r-1}\frac{\partial}{\partial x_m^{r-2}}+\ldots+X_{j-1}^{r-1}\frac{\partial}{\partial x_{j-2}^{r-2}}+X_m^{r-1}\frac{\partial}{\partial x_{j-1}^{r-2}}+\ldots$$

A simple remedy is to adjust the lower indices in the x_*^{r-2} variables in this highlighted part of $[Z_1]$ to these in the X_*^{r-1} variables. That is, to perform a subsequent (going backwards with respect to the superscripts) change of coordinates, also cyclic $x_m^{r-2}\to x_{m+1}^{r-2}\to\ldots\to x_{j-1}^{r-2}\to x_m^{r-2}$ after which $[Z_1]$ assumes the form

$$\ldots+X_{m+1}^{r-1}\frac{\partial}{\partial x_{m+1}^{r-2}}+\ldots+X_{j-1}^{r-1}\frac{\partial}{\partial x_{j-1}^{r-2}}+X_m^{r-1}\frac{\partial}{\partial x_m^{r-2}}+\ldots.$$

And so on backwards, always taking into account (and profiting from the fact!) that $j_l\le m$, $l=r-1,r-2,\ldots,1$. The last, necessary for arriving again at an EKR, is the cyclic change $x_m^0\to x_{m+1}^0\to\ldots\to x_{j-1}^0\to x_m^0$. This terminates our bringing the distribution germ D^{k+1} to the (better) EKR form $\mathbf{j}_1\ldots\mathbf{j}_{r-1}.\mathbf{m{+}1}$ satisfying the least upward jumps rule. Theorem 3 is now proved by induction. ∎

3.4. *Examples of* EKRs. A systematic theory of singularities of distributions generating special multi-flags remains to be developed (meaning geometric singularities assuredly underlying the vast panoply of various EKRs constructed earlier in the present section; see also footnote 6). Here we want but to state some first and basic properties of the EKRs.

To begin with, note that, for $k=2$, in the family of pseudo-normal forms $\mathbf{1}$ there sits the distribution (1). Before continuing, let us recall one of principal tools in the entire theory of geometric distributions, the *small growth vector* (s.gr.v.) of a distribution at a point, in passing mentioned already in Remark 1. For a given distribution D, it is

the sequence of linear dimensions, at a point in question, of the tower of modules (or: sheaves) of vector fields $D = V_1 \subset V_2 \subset V_3 \subset \dots$, with $V_{i+1} = V_i + [D, V_i]$, $i = 1, 2, \dots$. For D completely nonholonomic (and such are, naturally, the distributions generating special multi-flags) the s.gr.v.'s terminate, sooner or later in function of a point, by the dimension of the underlying manifold. The length of that vector is then called the nonholonomy degree of D at that point.

PROPOSITION 1.

(i) *Fix any $k \geq 1$. All germs in $\mathbf{1.1}\dots\mathbf{1}$ (fixed length r) are equivalent regardless of constants appearing in them. So appear the unique local models for Cartan distributions on the spaces of r-jets of vector k-dimensional functions of one variable* (*or*: *curves in* $\mathbb{R}^k$). *For $k = 1$, thus we obtain so-called Goursat normal forms, or 'chained systems' still actively used in geometric control theory.*

(ii) *For $k = 1$* (*the Goursat case*), *the germs in $\mathbf{1.1}$ and $\mathbf{1.2}$ are all equivalent regardless of constants* (*Engel's theorem of* 1889), *but later, at places number* 3, 4, ..., r *in the encoding words, the alternatives $\mathbf{1}$ or $\mathbf{2}$ always create geometrically different objects, representing and leading to 2^{r-2} Kumpera-Ruiz classes of Goursat germs in the terminology of* [MonZ] (p. 466)([7]).

(iii) *For $k \geq 2$, $\mathbf{1.1}$ and $\mathbf{1.2}$ are already different and disjoint* (*there is no Engel's theorem for multi-flags*).

(iv) *The first situation when the operation $\mathbf{3}$ brings in essentially new* EKRs *is $\mathbf{1.2.3}$. If we speak for simplicity for $k = 2$, these distribution germs have at $0 \in \mathbb{R}^9$ the small growth vector* [3, 5, 6, 7, 7, 8, 9] *which does not appear among the s.gr.v.'s at* 0 *of the* EKRs *of length* 3 *produced only via the operations $\mathbf{1}$ and $\mathbf{2}$.*

Proof. Two first items are classical (by subtracting a vector polynomial, the r-jet at 0 of any given mapping $\mathbb{R} \to \mathbb{R}^k$ can be assumed zero). Before proving the two last items, we want to bring in two examples of families of distributions, already in extended Kumpera-Ruiz form, generating special 2-flags of length 3. By writing the starting variables t, x_1^0, x_2^0 as t, x_1, y_1 and then x_{j+1}, y_{j+1} instead of x_1^j, x_2^j for $j = 1, 2, 3$, the germs in $\mathbf{1.1.2}$ can, with no loss of generality, be presented as

$$\left(x_4\Big(\frac{\partial}{\partial t} + x_2\,\frac{\partial}{\partial x_1} + y_2\,\frac{\partial}{\partial y_1} + x_3\,\frac{\partial}{\partial x_2} + y_3\,\frac{\partial}{\partial y_2}\Big) + \frac{\partial}{\partial x_3} + Y_4\,\frac{\partial}{\partial y_3}\,,\ \frac{\partial}{\partial x_4}\,,\ \frac{\partial}{\partial y_4}\right)$$

(with no constants standing by x_2, y_2, x_3, y_3, precisely in view of Proposition 1 (i)). And here is the family of distribution germs $\mathbf{1.2.2}$, after a similar simplification:

$$\left(x_4\Big(x_3\Big(\frac{\partial}{\partial t} + x_2\,\frac{\partial}{\partial x_1} + y_2\,\frac{\partial}{\partial y_1}\Big) + \frac{\partial}{\partial x_2} + Y_3\,\frac{\partial}{\partial y_2}\Big) + \frac{\partial}{\partial x_3} + Y_4\,\frac{\partial}{\partial y_3}\,,\ \frac{\partial}{\partial x_4},\ \frac{\partial}{\partial y_4}\right).$$

Part (iii): All germs in $\mathbf{1.1}$ have everywhere, hence also at $0 \in \mathbb{R}^{3k+1}$, the s.gr.v. $[k+1,\ 2k+1,\ 3k+1]$ coinciding with the big one, while a short calculus shows that the germs

([7]) The question of true significance of constants in the EKRs is enormously involved already in the Goursat case, to quote for instance [KuRui], [Gas], [ChMPR], [MonZ], [M2] (this list is by far not exhaustive).

in **1. 2** have at 0 the s.gr.v. $[k+1,\, 2k+1,\, 2k+2,\, 3k+1]$ which, for $k \geq 2$, is different from the previous one.

Part (iv): The small growth vector at $0 \in \mathbb{R}^9$ is:

[3, 5, 7, 9] for the germs in **1. 1. 1** ,

[3, 5, 7, 8, 9] and [3, 5, 7, 8, 8, 9] for those in **1. 2. 1** ,

[3, 5, 6, 7, 8, 9] for **1. 1. 2** ,

[3, 5, 6, 7, 8, 8, 9] for **1. 2. 2** . ∎

Corollary 1. *Regarding Theorem* 3.2 *in* [PR]*, it is visible that, up to a reindexation of coordinates, it produces only* EKRs *having codes built out of operations* **1** *and* **2**. *Hence, for $k \geq 2$, that theorem does not give pseudo-normal forms for all germs of special k-flags.* (*In Proposition* 1 *above the arguments are given only for $k = 2$, but the operation* **3** *is essential for all $k \geq 2$. In fact, for a given width k and starting from the length $k+1$, all operations* **1** *through* **k+1** *inclusively are necessary in the production of* EKRs*, see* [M4].)

Remark 3. Another consequence of Theorem 2, modulo the arguments used already in the proof of Theorem 3 above, is that, upon applying r times multi-dimensional Cartan prolongation to the tangent bundle to a $(k+1)$-dimensional manifold (for instance just to $\mathbb{R}^{k+1}$), one obtains a 'Monster Special k-Flags Manifold' (MSkFM) of dimension $(r+1)k+1$ bearing a locally universal special k-flag. Such a universal object locally models every special k-flag existing on $((r+1)k+1)$-dimensional manifolds.

Concerning 1-flags (Goursat flags), a similar conclusion leading to 'Monster Goursat Manifolds' (MGM) bearing locally universal Goursat flags can be drawn from Theorem 4.2 in [BH]; it is done independently in Section 5.4 in [MonZ].

(And, reiterating after Section 1, that same theorem virtually *produces* all local Kumpera-Ruiz pseudo-normal forms for Goursat distributions—which served as the key motivation for the present work.)

4. Nilpotent algebras in special k-flags and strong nilpotency at certain places. Let us define a linear automorphism $\overleftarrow{\mathbf{j}}$ of $\mathbb{R}^{k+1}$ associated to any fixed operation $\mathbf{j} \in \{\mathbf{1}, \ldots, \mathbf{k{+}1}\}$, by its values on a—fixed once for all—basis $e_1, \ldots, e_{k+1}$ of $\mathbb{R}^{k+1}$. Namely, let $\overleftarrow{\mathbf{j}}$ send: $e_1 \mapsto e_1 + e_2 + \ldots + e_{k+1}$, $e_2 \mapsto e_1$, …, $e_j \mapsto e_{j-1}$, $e_{j+1} \mapsto e_{j+1}$, …, $e_{k+1} \mapsto e_{k+1}$ (so that, naturally, $\overleftarrow{\mathbf{1}}$ sends $e_2 \mapsto e_2$, …, $e_{k+1} \mapsto e_{k+1}$, and $\overleftarrow{\mathbf{k{+}1}}$ sends $e_2 \mapsto e_1$, …, $e_{k+1} \mapsto e_k$). With this notation we have:

Theorem 4. *Every rank $k{+}1$ distribution D generating a special k-flag of a length $r \geq 1$ is locally nilpotentizable* (*i.e., is locally weakly nilpotent in the sense of* [M1]) *and a local nilpotent basis around a point p is $\{Z_1, \ldots, Z_{k+1}\}$, where $(Z_1, \ldots, Z_{k+1}) \in \mathbf{j}_1.\mathbf{j}_2 \ldots \mathbf{j}_r$ issuing from Theorem* 3 *is an* EKR *form for D around p satisfying the least upward jumps rule.*

The nilpotency order of the real Lie algebra $L(\mathbf{j}_1.\mathbf{j}_2 \ldots \mathbf{j}_r) \stackrel{\text{def}}{=} L_{\mathbb{R}}\big(Z_1, \ldots, Z_{k+1}\big)$ is equal to the component of e_{k+1} (i.e., the last one) in the vector $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2} \cdots \overleftarrow{\mathbf{j}_r}(e_1 + \ldots + e_{k+1})$.

Moreover, if a germ of D at a certain point admits a local EKR *form with no non-zero constants, then that germ is also strongly nilpotent in the sense of* [AGau] *and* [M1].

REMARK 4. The algebra $L(\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r)$ appearing in this theorem is well defined. Indeed, for two different EKRs $(Z_1,\dots,Z_{k+1})$ and $(\widetilde{Z}_1,\dots,\widetilde{Z}_{k+1})$ belonging to the same EKR class $\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r$, it is easy to write down a simple translation in the space $\mathbb{R}^{(r+1)k+1}$ inducing an (inner) isomorphism of $L_{\mathbb{R}}\big(Z_1,\dots,Z_{k+1}\big)$ and $L_{\mathbb{R}}\big(\widetilde{Z}_1,\dots,\widetilde{Z}_{k+1}\big)$, sending every generator Z_l to $\widetilde{Z}_l$, $l=1,\dots,k+1$. The translation vector has as components the differences between the respective constants sitting in the fields Z_1 and $\widetilde{Z}_1$ (and many other components 0). Thus $L_{\mathbb{R}}\big(Z_1,\dots,Z_{k+1}\big)$ and $L_{\mathbb{R}}\big(\widetilde{Z}_1,\dots,\widetilde{Z}_{k+1}\big)$ are two identical copies of one underlying Lie algebra [that turns out to be nilpotent by this very theorem].

In the light of (very recent) constructions [M4], all this is *geometric* and to any distribution germ D generating a special k-flag there is associated such an algebra $\mathcal{A}$. We *conjecture* here, extending the partial results and conjectures of [M3], that $\mathcal{A}$ is always minimal for D, in the sense of having the minimal nilpotency order among the Lie algebras induced over $\mathbb{R}$ by all possible nilpotent bases for D.

It is also to be noted that when $m=\max(j_1,\dots,j_r)\le k$, then not only the last component, but precisely $k+1-m$ last components in $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2}\cdots\overleftarrow{\mathbf{j}_r}(e_1+\ldots+e_{k+1})$ are mutually equal and equal to the nilpotency order of $L(\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r)$. This will be visible in the proof.

COROLLARY 2. *For distribution germs D as in the 'Moreover' part of Theorem* 4 (*for $k=1$ they are called 'tangential' in* [M1]), *one is able to effectively compute the degree of nonholonomy of D, too. It is equal to the nilpotency order of the nilpotent Lie algebra given by the nilpotent approximation of D, and in the discussed case that algebra is just $L_{\mathbb{R}}\big(Z_1,\dots,Z_{k+1}\big)$. Hence the nonholonomy degree of such D equals the last component in the vector $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2}\cdots\overleftarrow{\mathbf{j}_r}(e_1+\ldots+e_{k+1})$, too.*

Still before proving Theorem 4, let us demonstrate how it works.

EXAMPLE 1. Fix $k=2$. The nilpotency order of the algebra $L(\mathbf{1}.\mathbf{1}.\mathbf{2})$ is equal to 6, because, computing in that fixed for ever basis e_1, e_2, e_3, we obtain

$$\overleftarrow{\mathbf{1}}\,\overleftarrow{\mathbf{1}}\,\overleftarrow{\mathbf{2}}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}1&1&0\\1&0&0\\1&0&1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}2\\5\\6\end{pmatrix}.$$

Likewise, the nilpotency order of the algebra $L(\mathbf{1}.\mathbf{2}.\mathbf{2})$ is equal to 7, because

$$\overleftarrow{\mathbf{1}}\,\overleftarrow{\mathbf{2}}\,\overleftarrow{\mathbf{2}}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}1&1&0\\1&0&0\\1&0&1\end{pmatrix}\begin{pmatrix}1&1&0\\1&0&0\\1&0&1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}3\\5\\7\end{pmatrix}.$$

REMARK 5. For $k=1$, the description of nilpotency orders of the EKRs given in Theorem 4 allows to re-prove a part of Jean's results [Je] on the nonholonomy degrees of Goursat germs. A bridge to those results rests on certain simple algebraic identities satisfied (for, we repeat, $k=1$) by the operators $\overleftarrow{\mathbf{1}}$ and $\overleftarrow{\mathbf{2}}$ when applied to e_1+e_2. The details will be given elsewhere.

Proof of Theorem 4. The proof uses refinements of methods of [M1], [M3]. By Remark 4, one can choose as handy an EKR in the class $\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r$ as possible. Let us choose, then, the uniquely defined EKR $(Z_1,\dots,Z_{k+1})\in\mathbf{j}_1.\mathbf{j}_2\dots\mathbf{j}_r$ having *all constants*

zero. (With such a choice, also minimized will be the work on the 'Moreover' part in the theorem.)

We are going to associate weights $w(\cdot)$ to all coordinates t, $x_1^0, \dots, x_k^r$, and to basic vector fields (versors) $w(\partial/\partial x) = -w(x)$ in such a way that—under the classical definition of the weight of a monomial vector field that goes back to the 1970s,

$$w\Big(y_{i_1} y_{i_2} \cdots y_{i_s} \frac{\partial}{\partial x}\Big) = w(y_{i_1}) + w(y_{i_2}) + \ldots + w(y_{i_s}) + w\Big(\frac{\partial}{\partial x}\Big) \tag{19}$$

—all $Z_1, Z_2, \dots, Z_{k+1}$ will be homogeneous of degree -1, and all vector fields appearing in the stepwise construction of the involved field Z_1 will be homogeneous, of growing degrees, as we go backwards in that construction. The manner of growth will highly depend on the EKR class, and the nilpotency order will be a function of that growth.

In our situation without constants,

$$\begin{aligned} Z_1 = x_1^r Z_1^{r-1} + x_2^r \frac{\partial}{\partial x_1^{r-1}} + \ldots \\ + x_{j_r-1}^r \frac{\partial}{\partial x_{j_r-2}^{r-1}} + \frac{\partial}{\partial x_{j_r-1}^{r-1}} + x_{j_r}^r \frac{\partial}{\partial x_{j_r}^{r-1}} + \ldots + x_k^r \frac{\partial}{\partial x_k^{r-1}}\,, \end{aligned} \tag{20}$$

where the involved vector field from the one before last step is denoted by Z_1^{r-1} (and there are no constants in it, too). Let us start by declaring $w(Z_1) = w(Z_2) = \ldots = w(Z_{k+1}) = -1$, hence also $w(x_1^r) = \ldots = w(x_k^r) = 1$. By writing $w(Z_1) = -1$, we stipulate that *all* summands in (20) are of weight -1. In particular, $-1 = w(x_1^r\, Z_1^{r-1}) = w(x_1^r) + w(Z_1^{r-1}) = 1 + w(Z_1^{r-1})$, which allows to compute the weight of Z_1^{r-1}, and to proceed further. The particulars of the entire definition can be formulated inductively as follows, understanding by Z_1^r the starting vector field Z_1.

Fix an arbitrary s such that $1 \le s \le r$ and assume that Z_1^s is a homogeneous vector field of known weight $w(Z_1^s)$, and that the weights of the coordinates $x_1^s, \dots, x_k^s$ are already defined. Recall that

$$Z_1^s = x_1^s Z_1^{s-1} + x_2^s \frac{\partial}{\partial x_1^{s-1}} + \ldots + x_{j_s-1}^s \frac{\partial}{\partial x_{j_s-2}^{s-1}} + \frac{\partial}{\partial x_{j_s-1}^{s-1}} + x_{j_s}^s \frac{\partial}{\partial x_{j_s}^{s-1}} + \ldots + x_k^s \frac{\partial}{\partial x_k^{s-1}}\,.$$

By analyzing the consecutive summands on the right hand side, the homogeneity of Z_1^s implies that

$$w(Z_1^{s-1}) = w(Z_1^s) + w\Big(\frac{\partial}{\partial x_1^s}\Big) \tag{21}$$

and that

$$w\Big(\frac{\partial}{\partial x_l^{s-1}}\Big) = w(Z_1^s) + w\Big(\frac{\partial}{\partial x_{l+1}^s}\Big) \qquad \text{for } l = 1, \dots, j_s - 2, \tag{22}$$

$$w\Big(\frac{\partial}{\partial x_{j_s-1}^{s-1}}\Big) = w(Z_1^s), \tag{23}$$

$$w\Big(\frac{\partial}{\partial x_l^{s-1}}\Big) = w(Z_1^s) + w\Big(\frac{\partial}{\partial x_l^s}\Big) \qquad \text{for } l = j_s, \dots, k. \tag{24}$$

In this way we define inductively the negative weights of all $(r+1)k+1$ versors, including $\partial/\partial t = Z_1^0$, and the positive (opposite to the former) weights of all coordinates. Now

denote by d the biggest weight of a coordinate, say x (there can be, sometimes, several coordinates of the weight d).

The key property of weights defined by (19) is the *additivity* under Lie multiplication. So the Lie products of any s factors from among $\{Z_1, \dots, Z_{k+1}\}$ are homogeneous vector fields of weight $-s$. In particular, products of more than d factors automatically vanish (our vector monomials have only weights greater than or equal to $-d$). On the other hand, D, generating a special flag, is completely non-holonomic. In particular $\partial/\partial x$ at $0 \in \mathbb{R}^{(r+1)k+1}$ should be a combination of products of factors from among $\{Z_1, \dots, Z_{k+1}\}$ evaluated at 0.

In such a combination, let us reiterate, only products of at most d factors count, while products of any $e < d$ factors have the $\partial/\partial x$ components of the form $P\,\partial/\partial x$, P—a homogeneous polynomial of degree $d - e > 0$. Those P's are, naturally, polynomials in shifted variables (denoted by capital letters) of the EKR in question. But in the present EKR the variables are actually not shifted!—there are no constants shifting them. In consequence, all such P's vanish at 0. In this way only products of exactly d factors contribute in the production(s) of $\partial/\partial x$ at 0. This implies a [modest if cardinal] information that certain product(s) of *exactly* d factors from among $\{Z_1, \dots, Z_{k+1}\}$ is (are) non-zero.

At this point we know, therefore, that $L_{\mathbb{R}}\big(Z_1, \dots, Z_{k+1}\big)$ is a nilpotent Lie algebra of nilpotency order d and also, as a byproduct, that the small flag of D, $D = V_1 \subset V_2 \subset V_3 \subset \dots$ ends (locally, as a germ) on its term V_d,

$$V_d = T\,\mathbb{R}^{(r+1)k+1}. \tag{25}$$

How can we better visualise that value d?

In what follows we are going to use the matrix notation for the operators $\overleftarrow{\mathbf{j}_1}, \dots, \overleftarrow{\mathbf{j}_r}$ being written in the mentioned fixed basis $e_1, \dots, e_{k+1}$ of $\mathbb{R}^{k+1}$. Observe that the formulas (21)–(24) can be most compactly expressed as

$$\Big[w(Z_1^{s-1}),\, w\Big(\frac{\partial}{\partial x_l^{s-1}}\Big)_{l=1,\dots,k}\Big]^T \text{ is the value of } \overleftarrow{\mathbf{j}_s} \text{ at } \Big[w(Z_1^{s}),\, w\Big(\frac{\partial}{\partial x_l^{s}}\Big)_{l=1,\dots,k}\Big]^T. \tag{26}$$

Hence, by making the composition over $s = r, r-1, \dots, 1$,

$$\Big[w\Big(\frac{\partial}{\partial t}\Big),\, w\Big(\frac{\partial}{\partial x_l^{0}}\Big)_{l=1,\dots,k}\Big]^T \text{ is the value of } \overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2}\cdots\overleftarrow{\mathbf{j}_r} \text{ at } [-1, -1, \dots, -1]^T_{l=1,\dots,k}. \tag{27}$$

Each Z_1^{s-1} ($s \geq 2$) has as a component the bare versor $\partial/\partial x^{s-2}_{j_{s-1}-1}$, and Z_1^0 simply is $\partial/\partial t$. Thus $w(Z_1^{s-1})$ is a weight of a versor, and in (26) one set of versors' weights is transformed into another set of versors' weights.

Fact. The lowest negative versor weight $-d$ occurs among the components of the vector on the left hand side of (27).

Proof. Passing (for the reader's convenience) to the positive weights of variables, one starts from the vector $[1, 1, \dots, 1]^T$ and notes that each $\overleftarrow{\mathbf{j}}$ evaluated at any vector

$[a_1, \dots, a_{k+1}]^T$ with all positive components,

$$\overleftarrow{\mathbf{j}}\,[a_1, a_2, \dots, a_{k+1}]^T = [a_1 + a_2, \dots, a_1 + a_j, a_1, a_1 + a_{j+1}, \dots, a_1 + a_{k+1}]^T,$$

has the maximum of components bigger than $a_1, a_2, \dots, a_j, a_{j+1}, \dots, a_{k+1}$, hence bigger than $\max(a_1, \dots, a_{k+1})$. ∎

It remains to explain why the last component, and possibly not only it, on the left hand side of (27) is the biggest in modulus. Not surprisingly, the least upward jumps rule satisfied by our EKR is responsible for that. In fact, for $m = \max(j_1, \dots, j_r)$, the last $k + 1 - m$ components of $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2} \cdots \overleftarrow{\mathbf{j}_r}[1, 1, \dots, 1]^T$ are pairwise equal (we work in the sequel with positive weights of variables only).

If m appears for the *first* time in the sequence $j_1, \dots, j_r$ as $j_l = m$, then the m-th component of

$$\overleftarrow{\mathbf{j}_l}\,\overleftarrow{\mathbf{j}_{l+1}} \cdots \overleftarrow{\mathbf{j}_r}[1, 1, \dots, 1]^T$$

is smaller than the components with numbers $m + 1, \dots, k + 1$, and this relation keeps holding after applying each of the subsequent operators $\overleftarrow{\mathbf{j}_{l-1}}, \dots, \overleftarrow{\mathbf{j}_1}$.

Now either $m = 1$ (in which case we are already done) or, by the least upward jumps rule, the value $m - 1$ does appear in the sequence $j_1, \dots, j_r$, and so let it appear for the first time as $j_s = m - 1$, $s < l$. Then the $(m - 1)$-th component in

$$\overleftarrow{\mathbf{j}_s} \cdots \overleftarrow{\mathbf{j}_l} \cdots \overleftarrow{\mathbf{j}_r}[1, 1, \dots, 1]^T$$

is smaller than the components with numbers $m, m + 1, \dots, k + 1$, and this relation keeps holding after applying each of the subsequent operators (if there remains any) $\overleftarrow{\mathbf{j}_{s-1}}, \dots, \overleftarrow{\mathbf{j}_1}$.

Thus at this point we know that, in the eventual outcome $\overleftarrow{\mathbf{j}_1} \cdots \overleftarrow{\mathbf{j}_r}[1, 1, \dots, 1]^T$, the $(m - 1)$-th component is smaller than the m-th component which in turn is smaller than the $k + 1 - m$ last components that are pairwise equal. Then we proceed likewise downwards, constantly using the least upward jumps property of $\mathbf{j}_1.\mathbf{j}_2 \dots \mathbf{j}_r$, obtaining in the end that, in the vector $\overleftarrow{\mathbf{j}_1} \cdots \overleftarrow{\mathbf{j}_r}[1, 1, \dots, 1]^T$, the first component is smaller than the second which is smaller than ... which is smaller than the m-th which is smaller than the $k + 1 - m$ last, mutually equal, components.

The 'Moreover' part in Theorem 4. Thanks to the choice of an EKR in the beginning, we can work all the time with the same $\{Z_1, \dots, Z_{k+1}\}$. Divide the constructed weights of coordinates $t, x_1^0, \dots, x_k^r$ into groups of equal values, the highest value being d: $w_1 = w_2 = \dots = w_{n_1} = 1$, $n_1 = k + 1$ (it is clear that, on top of $1 = w(x_1^r) = \dots = w(x_k^r)$, precisely one more variable $x_{j_r-1}^{r-1}$ has weight 1), $w_{n_1+1} = \dots = w_{n_2} = 2, \dots, w_{n_{l-1}+1} = \dots = w_{n_l} = l, \dots, w_{n_{d-1}+1} = \dots = w_{n_d} = d$, where $n_d = (r + 1)k + 1$.

Attention. In this definition it may often happen that a given integer l is not the value of a weight, and then simply $n_{l-1} = n_l$.

Proposition 2. *$[n_1, n_2, \dots, n_d]$ is the small growth vector of D at $0 \in \mathbb{R}^{(r+1)k+1}$. Moreover, all members V_j of the small flag of D have at 0 the description*

$$V_j(0) = \Big(\frac{\partial}{\partial x}\,,\, x :\, w(x) \le j\Big). \tag{28}$$

That is to say, the coordinates $t, x_1^0, \dots, x_k^r$ are—for D—linearly adapted at 0.

Proof. V_j is spanned by the products of at most j factors from $\{Z_1, Z_2, \dots, Z_{k+1}\}$. Such products are, as we already know, homogeneous of weights greater than or equal to $-j$ and, evaluated at 0, are combinations of *bare* versors (because there are no constants in $(Z_1, \dots, Z_{k+1})$). Those are, therefore, the versors of coordinates x such that $w(x) \le j$. This means the inclusions $\cdot \subset \cdot$ in (28). Clearly, for $j = 1$, with $\operatorname{rk} V_1 = k + 1$, there is an equality in (28). Suppose that j is the smallest positive integer such that we have $\subsetneq$ in (28). (In view of (25), assuredly $j < d$.) That is, that there exists a combination

$$\sum_{x:\, w(x)=j} a_x \frac{\partial}{\partial x} \notin V_j(0). \tag{29}$$

Then

$$V_{j+s}(0) = \operatorname{span}\Big(V_j(0),\ \text{certain combinations of } \frac{\partial}{\partial x}\,,\ j < w(x) \le j + s\Big)$$

for $= 1, 2, \dots$. In particular, for $s = d - j$,

$$V_d(0) = \operatorname{span}\Big(V_j(0),\ \text{certain combinations of } \frac{\partial}{\partial x}\,,\ j < w(x) \le d\Big).$$

This together with (29) imply $\sum_{x:\, w(x)=j} a_x\, \partial/\partial x \notin V_d(0)$ which contradicts (25).

The description (28), now shown to hold, allows to compute the s.gr.v. of D at 0. For $j = 2, \dots, d$, $\dim V_j(0) = n_1 + (n_2 - n_1) + \dots + (n_j - n_{j-1}) = n_j$. Proposition 2 is proved. ∎

Summarizing now, we see that our weights $w(\cdot)$ are modelled on the pattern established by the s.gr.v. of D at 0. With this knowledge, Proposition 2 says precisely that the coordinates in question are linearly adapted for D at 0 (cf., for instance, [AGamS], [Bel], [M1]). In fact, they are even adapted.

This (last) part is standard, because, on the one side, 'linear adaptedness' implies that the *nonholonomic orders* of these variables $t, x_1^0, \dots, x_k^r$ do not exceed their respective weights w_i ([Bel], p. 35), that is—do not exceed our weights $w(\cdot)$. And, on the other side, any nonholonomic derivative of a variable x from this set, with respect to $\{Z_1, \dots, Z_{k+1}\}$ and of order $l < w(x)$, is—by our weights' construction—a homogeneous polynomial of positive weight $w(x) - l$ in the variables that are not shifted by constants, and vanishes as such (cf. the earlier argument in this long proof that certain products of d factors from among $\{Z_1, \dots, Z_{k+1}\}$ do not vanish). Thus the nonholonomic orders and the weights w_i coincide, and the coordinates are indeed adapted.

Once found (any) set of adapted coordinates for a distribution around a point, one can compute the nilpotent approximation of that distribution at that point. In the present case all terms in Z_1, and in $Z_2, \dots, Z_{k+1}$, too, are of degree (weight) -1, so that the nilpotent approximation of D at 0 simply coincides with D. This exactly means ([AGau]; see also [M1], Definition 5) that the germ D is strongly nilpotent. ∎

5. Algorithmic issues. The last issue is that of more algorithmic ways of computing the nilpotency orders, not via bringing in the entire matrices of operators $\overleftarrow{\mathbf{j}}$ required in Theorem 4. For $k = 1$ this task has been carried out in [M1], [M3]. In the same vein, we will announce here an answer for $k = 2$. It is less compact (if, we believe, optimal) than

in the Goursat case. For $k = 3$ the answer, not given here, is still more involved than for $k = 2$, and so on when k grows further. It is important that the computation keeps being recursive, although the recurrence patterns become (by far) more involved.

So, from now on, we set $k = 2$ and use the letter $\mathbf{C}$ for codes of different EKRs. Moreover, we denote by $f(\mathbf{C})$ the nilpotency order of the Lie algebra $L(\mathbf{C})$ issuing from Theorem 4. A string of l same ciphers, like for instance a string of l $\mathbf{1}$'s, going in row in a code will be written shortly as $\mathbf{1}_l$. Thus $\mathbf{1}_2.\mathbf{2}_3$ is the EKR $\mathbf{1.1.2.2.2}$, and so on. Here are the recipes for computing the nilpotency orders of $L(\mathbf{C})$ from Theorem 4.

THEOREM 5. *In the situation $k = 2$, for any* EKR $\mathbf{C}$ *satisfying the least upward jumps rule ($\mathbf{C} = \emptyset$ not excluded), the nilpotency order of the Lie algebra $L(\mathbf{C})$ equals $f(\mathbf{C})$, where the function $f(\cdot)$ is defined recursively below. In this definition an $*$ stands for any cipher from among $\{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$, while a $\bullet$—it is important—stands only for $\mathbf{2}$ or $\mathbf{3}$.*

(i) $f(\emptyset) = 1, \quad f(\mathbf{1}) = 2,$

(ii) $f(\mathbf{C}.*.\mathbf{1}) = 2f(\mathbf{C}.*) - f(\mathbf{C}),$

(iii) $f(\mathbf{1}_l.2) = 2l + 2, \quad l = 1, 2, 3, \dots,$

(iv) $f(\mathbf{C}.*.\bullet.\mathbf{1}_l.\mathbf{2}) = (2l+2)f(\mathbf{C}.*.\bullet) - 2lf(\mathbf{C}.*) - f(\mathbf{C})$ *for any* $l = 0, 1, 2, \dots,$

(v) $f(\mathbf{C}.*.\bullet.\mathbf{1}_l.\mathbf{3}) = (2l+1)f(\mathbf{C}.*.\bullet) - (2l-1)f(\mathbf{C}.*) + f(\mathbf{C})$ *for any* $l = 0, 1, 2, \dots.$

A proof of this theorem will appear, in a more general setting, in author's subsequent paper. It can be useful to compare right now the above recurrencies with those governing the case $k = 1$ (Theorem 1 in [M1]). In the Goursat case the nilpotency orders $g(\cdot)$ satisfy the relations

$g(\emptyset) = 1, \quad g(\mathbf{1}) = 2,$

$g(\mathbf{C}.*.\mathbf{1}) = 2g(\mathbf{C}.*) - g(\mathbf{C}),$

$g(\mathbf{C}.*.\mathbf{2}) = g(\mathbf{C}.*) + g(\mathbf{C}),$

where an $*$ is for $\mathbf{1}$ or $\mathbf{2}$, and $\mathbf{C}$ starts on the left from $\mathbf{1.1}$, cf. Proposition 1 (ii).

EXAMPLE 2. In order to illustrate the algorithm, let us compute in two ways the nilpotency order of $L(\mathbf{1}_3.\mathbf{2.3.1.2.1}_2.\mathbf{3.2.1}_4.\mathbf{3})$. Firstly, according to Theorem 4,

$$\left(\overleftarrow{\mathbf{1}}\right)^3 \overleftarrow{\mathbf{2}}\,\overleftarrow{\mathbf{3}}\,\overleftarrow{\mathbf{1}}\,\overleftarrow{\mathbf{2}} \left(\overleftarrow{\mathbf{1}}\right)^2 \overleftarrow{\mathbf{3}}\,\overleftarrow{\mathbf{2}} \left(\overleftarrow{\mathbf{1}}\right)^4 \overleftarrow{\mathbf{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 328 \\ 1128 \\ 1193 \end{pmatrix},$$

which implies that the order in question is 1193. Secondly, according to Theorem 5, one proceeds stepwise, expanding progressively the code to the right:

$f(\mathbf{1}_3.\mathbf{2}) = 2 \cdot 3 + 2 = 8$ by (iii),

$f(\mathbf{1}_3.\mathbf{2.3}) = f(\mathbf{1}_3.\mathbf{2}) + f(\mathbf{1}_3) + f(\mathbf{1}_2) = 8 + 4 + 3 = 15$ by (v), $l = 0$,

$f(\mathbf{1}_3.\mathbf{2.3.1.2}) = 4f(\mathbf{1}_3.\mathbf{2.3}) - 2f(\mathbf{1}_3.\mathbf{2}) - f(\mathbf{1}_3) = 4 \cdot 15 - 2 \cdot 8 - 4 = 40$ by (iv), $l = 1$.

Additionally, $f(\mathbf{1}_3.\mathbf{2.3.1}) = 2 \cdot 15 - 8 = 22$ by (ii), and so

$$\begin{aligned} f(\mathbf{1}_3.\mathbf{2.3.1.2.1}_2.\mathbf{3}) &= 5f(\mathbf{1}_3.\mathbf{2.3.1.2}) - 3f(\mathbf{1}_3.\mathbf{2.3.1}) + f(\mathbf{1}_3.\mathbf{2.3}) \\ &= 5 \cdot 40 - 3 \cdot 22 + 15 = 149 \end{aligned}$$

by (v), $l = 2$. Additionally, $f(\mathbf{1}_3.\mathbf{2.3.1.2.1}) = 2 \cdot 40 - 22 = 58$ and $f(\mathbf{1}_3.\mathbf{2.3.1.2.1}_2) = 2 \cdot 58 - 40 = 76$ by (ii) again, and so

$f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2.\mathbf{3}.\mathbf{2}) = 2f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2.\mathbf{3}) - f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}) = 2 \cdot 149 - 58 = 240$ by (iv), $l = 0$. Eventually, by (v), $l = 4$,

$$\begin{aligned} &f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2.\mathbf{3}.\mathbf{2}.\mathbf{1}_4.\mathbf{3}) \\ &\quad = 9f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2.\mathbf{3}.\mathbf{2}) - 7f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2.\mathbf{3}) + f(\mathbf{1}_3.\mathbf{2}.\mathbf{3}.\mathbf{1}.\mathbf{2}.\mathbf{1}_2) \\ &\quad = 9 \cdot 240 - 7 \cdot 149 + 76 = 1193, \end{aligned}$$

as in the first way.

References

[Ad1] J. Adachi, *Engel structures with trivial characteristic foliations*, Algebr. Geom. Topol. 2 (2002), 239–255 (`http://www.maths.warwick.ac.uk/agt`).

[Ad2] J. Adachi, *Global stability of distributions of higher corank of derived length one*, Int. Math. Res. Not. 2003, no. 49, 2621–2638.

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

[AGau] A. A. Agrachev, J.-P. Gauthier, *On subanalyticity of Carnot-Carathéodory distances*. Ann. Inst. H. Poincaré Anal. Non Linéaire 18 (2001), 359–382.

[Bel] A. Bellaïche, *The tangent space in sub-Riemannian geometry*, in: Sub-Riemannian Geometry, A. Bellaïche and J.-J. Risler (eds.), Progr. Math. 144, Birkhäuser, Basel, 1996, 1–78.

[BCG$^+$] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, P. A. Griffiths, *Exterior Differential Systems*, Math. Sci. Res. Inst. Publ. 18, Springer, New York, 1991.

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

[Ca1] É. Cartan, *Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre*, Ann. École Normale 27 (1910), 109–192.

[Ca2] É. Cartan, *Sur l'équivalence absolue de certains systèmes d'équations différentielles et sur certaines familles de courbes*, Bull. Soc. Math. France XLII (1914), 12–48.

[ChMPR] M. Cheaito, P. Mormul, W. Pasillas-Lépine, W. Respondek, *On local classification of Goursat structures*, C. R. Acad. Sci. Paris Sér. I Math. 327 (1998), 503–508.

[Gas] M. Gaspar, *Sobre la clasificacion de sistemas de Pfaff en bandera*, in: X jornadas Hispano Lusas de matemáticas. V. Geometria diferencial, Univ. of Murcia, Murcia 1985, 67–74 (in Spanish).

[Ge] V. Ya. Gershkovich, *Exotic Engel structures on* $\mathbb{R}^4$, Russian J. Math. Phys. 3 (1995), 207–226.

[GeVe] V. Ya. Gershkovich, A. M. Vershik, *An estimate of the functional dimension of the orbit space of germs of distributions of general position*, Mat. Zametki 44 (1988), 596–603 (in Russian); English translation: Math. Notes 44 (1988), 806–810.

[Gr] M. Gromov, *Carnot-Carathéodory spaces seen from within*, in: Sub-Riemannian Geometry, A. Bellaïche and J.-J. Risler (eds.), Progr. Math. 144, Birkhäuser, Basel, 1996, 79–323.

[HeLuSul] H. Hermes, A. Lundell, D. Sullivan, *Nilpotent bases for distributions and control systems*, J. Differential Equations 55 (1984), 385–400.

[JaPrz] B. Jakubczyk, F. Przytycki, *Singularities of* k*-tuples of Vector Fields*, Dissertationes Math. (Rozprawy Mat.) 213, Warszawa, 1984.

[Je] F. Jean, *The car with n trailers: characterisation of the singular configurations*, ESAIM Control Optim. Calc. Var. (`http://www.edpsciences.org/cocv/`) 1 (1996), 241–266.

[KLV] I. S. Krasil′shchik, V. V. Lychagin, A. M. Vinogradov, *Geometry of Jet Spaces and Nonlinear Partial Differential Equations*, Adv. Stud. Contemp. Math. 1, Gordon and Breach, New York, 1986.

[Ku] A. Kumpera, *Flag systems and ordinary differential equations*, Ann. Mat. Pura Appl. (4) 177 (1999), 315–329.

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

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

[LaSus] G. Lafferriere, H. J. Sussmann, *A differential geometric approach to motion planning*, in: Nonholonomic Motion Planning, F. Canny and Z. Li (eds.), Kluwer, Dordrecht, 1993, 235–270.

[Mon] R. Montgomery, *Engel deformations and contact structures*, in: Northern California Symplectic Geometry Seminar, Y. Eliashberg et al. (eds.), Amer. Math. Soc. Transl. Ser. 2, 196, Amer. Math. Soc., Providence, 1999, 103–117.

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

[M1] P. Mormul, *Goursat distributions not strongly nilpotent in dimensions not exceeding seven*, in: Nonlinear and Adaptive Control (Sheffield 2001), A. Zinober and D. Owens (eds.), Lecture Notes in Control and Inform. Sci. 281, Springer, Berlin, 2003, 249–261.

[M2] P. Mormul, *Discrete models of codimension-two singularities of Goursat flags of arbitrary length with one flag's member in singular position*, Tr. Mat. Inst. Steklova 236 (2002), 491–502; Proc. Steklov Math. Inst. 236 (2002), 478–489.

[M3] P. Mormul, *Minimal nilpotent local bases for Goursat distributions in dimensions not exceeding* 7, preprint in: Singularity Theory Seminar, S. Janeczko (ed.), Warsaw Univ. of Technology, 6 (2001), 44–56.

[M4] P. Mormul, *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, 8 (2003), 87–100.

[Mu] R. M. Murray, *Nilpotent bases for a class of nonintegrable distributions with applications to trajectory generation for nonholonomic systems*, Math. Control Signals Systems 7 (1994), 58–75.

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

[ShSl] W. F. Shadwick, W. M. Sluis, *On É. Cartan's absolute equivalence of differential systems*, C. R. Acad. Sci. Paris Sér. I Math. 313 (1991), 455–459.

[Sl] W. M. Sluis, *Absolute equivalence and its applications to control theory*, a thesis presented to the University of Waterloo, Ontario, Canada, 1992.

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

[Y1] K. Yamaguchi, *Contact geometry of higher order*, Japan J. Math. (N.S.) 8 (1982), 109–176.

[Y2] K. Yamaguchi, *Geometrization of jet bundles*, Hokkaido Math. J. 12 (1983), 27–40.

[Z] M. Zhitomirskii, *Rigid and abnormal line subdistributions of* 2*-distributions*, J. Dynam. Control Systems 1 (1995), 253–294.