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

October 30, 2002

Abstract

Through the works [BH] and [MonZ] it is known 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); local 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 [M3, M5]). For $k = 2$ we also transform our linear algebra formulas into recursive ones that resemble a bit Jean's formulas [J] 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 [AG] and [M3].

1 Generalized Cartan prolongations.

The paper is devoted to so-called Cartan prolongations of distributions, and to their applications: building more involved (having longer small and big growth vectors, or bigger nonholonomy degrees, for inst.) distributions from simpler

* research supported by Polish KBN Grant 2 P03A 010 22
AMS (2000) Classification: primary 58A15, 17, 30; secondary 17B30, 34C20.

ones. We assume 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 3 – 4, is the effective local (weak in the sense of [M3]) nilpotency of a quite rich family of objects related to and generalizing abstract Cartan distributions 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 [HLSul], [LSus], 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 for them were proposed in [M2] (Rem. 1 there), [PR2], and [M3]. In the latter, the *nilpotency orders* (minimal numbers of Lie multiplications in an algebra yielding automatically zero, cf. p. 239 in [LSus]) of the induced real algebras were explicitly computed. In fact, those bases were first put forward in a much different context in [KRui]; later they have been re-found by applying many times an important theorem, recalled in Sec. 2 below as Thm. 1. In [BH], in a comment to that theorem one can 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'. 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 inst. 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 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 an arbitrarily fixed chart. While the essence of the singularities that hide in 1-flags, and more generally in special k-flags (defined later in Sec. 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, as has been explained in [KRui, Mu], the simplest generic situation describing the abstract Cartan distributions (related to ODEs) on spaces of jets of functions $\mathbb{R} \to \mathbb{R}$.

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 [KRub], and with whom we discussed multi-flags on several occasions, if only by correspondence.

2 Structural theorem.

In [BH] (and in [Sl], according to the first mentioned source) there is given a sketch of the following theorem, attributed to Elie Cartan. 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 provided the latter is a regular distribution of constant rank independent of points in an underlying manifold. A foliation $\mathcal{F}$ 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 $f^{(1)}_*D = (D')^{(1)}$.*

This theorem has wide applications, and specifically to Goursat flags (see Rem. 3 in Sec. 3 below). Our objective is to formulate a local structural theorem generalizing Cartan's theorem and having even wider applications. The assumptions in Thm. 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.[1] Motivated by gCp's constructed in Sec. 1

[1] These objects were discovered many decades ago in [C], then were often rediscovered and recently given new (and surprising) interpretations in [BH] and [Z].

(hence, eventually by the seminal [BH]) and, to an extent, by the works [KRub] and [PR3], 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, and* 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 $\left(D_1/E\right)^{(1)}$ 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 notations nearly identical to those appearing in [BH], Sec. 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,

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

is defined by

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

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

$$\pi \circ f^{(1)} = f\,. \tag{2}$$

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 (2),

$$\begin{aligned} \mathrm{im}\, df^{(1)}(p) \quad &\text{contains an } s\text{-dimensional subspace} \\ &\text{transverse to } \ker d\pi(f^{(1)})(p)\,. \end{aligned} \tag{3}$$

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

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

Proof of lemma. This key part is technically more involved than in [BH] because the projective spaces being the fibres 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 $\partial_i = \frac{\partial}{\partial x_i}$. Visualise further by putting $N = \{x_1 = \cdots = x_k = 0\} \subset M = \mathbb{R}^{k+s}$. Put also

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

and

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

Then the intersections

$$D \cap TN\,, \qquad D \cap TN_{i,\,\epsilon}\,,\;\; i = 1, \dots, k\,, \tag{4}$$

are fields of lines on N, $N_{i,\,\epsilon}$ respectively (they are at least 1–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 (4). 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

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

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,

$$\frac{d}{dt}\, df_i(p + tv_i)L(p + tv_i)\bigg|_{t=0} \;=\; [L, v_i](p)\,. \tag{5}$$

Attention. The mapping f_i is not, naturally, the time $-t$ flow of v_i, but, in (5), df_i acts on L exactly as that flow does.

The vectors on the RHS of (5) 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. Seeing these

lines as points in $\mathrm{Gr}(D',\,1)(f(p))$, 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.
Lemma is now proved.

The information (3) together with Lem. 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)

$$f^{(1)} \text{ is a local diffeomorphism in a nbhd of } p \in M\,. \tag{6}$$

Now we will explain that

$$df^{(1)}(p) \text{ carries } D(p) \text{ to } {D'}^{(1)}(f^{(1)}(p))\,. \tag{7}$$

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

$$df^{(1)}(p)D(p) \subset {D'}^{(1)}(f^{(1)}(p)) \tag{8}$$

That is (see the definition of gCp in Sec. 1), that

$$df^{(1)}(p)D(p) \subset \big(d\pi(f^{(1)}(p))\big)^{-1} f^{(1)}(p)\,. \tag{9}$$

The RHS in (9) 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 (7) holds. With (6) and (7) justified, Theorem 2 is proved. $\square$

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 inst. [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 [K], **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.[2]

[2] The Cauchy–characteristic subflag of a Goursat flag does enjoy these properties (excepting the subdistribution of the corank-1 member of the flag, whose square is the full tangent bundle). So, 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.

For special k–flags, we mean $\bullet$ the big growth vector being everywhere $[k+1,\, 2k+1,\, 3k+1, \dots]$ (a distribution generates a k–flag), and $\bullet\bullet$ 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 $\bullet$ 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 E. Cartan in his famous 'cinq variables' paper of 1910. The condition $\bullet\bullet$ is, to the contrary, very restrictive, perfectly balancing $\bullet$.

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 [KRub] about 'k–flags satisfying normality conditions', while those distributions are called 'regular contact systems for curves' in [PR3].

As regards special k–flags generalizing the latter, in [PR3] 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 Sec. 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 prolongations.

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

$$D = D^{k+1} \subset D^{2k+1} \subset \cdots \subset D^{rk+1} \subset D^{(r+1)k+1} = TM \tag{10}$$

are all regular distributions of the respective ranks written as superscripts in (10) (this is $\bullet$) **and** when there exists a subflag of involutive distributions

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

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

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

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

Remark 2. In the work [PR3] the definition of contact systems (for curves, cf. Sec. 3.1) is weaker, replacing •• by only the existence of corank–1 involutive subdistributions $E^{jk} \subset D^{jk+1}$, $j = 1, \dots, r$. We are not yet sure if our defining conditions can be relaxed to these. But we are going to show in this section (in Prop. 1, (iv) and after its proof) that our formally more restrictive definition yields strictly more, not less, objects than produced in [PR3].

3.3 Theorem 2 **produces extended K–R 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, 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 contents, depends 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

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

and by $Z_2' = \frac{\partial}{\partial x_1^l}, \dots, Z_{k+1}' = \frac{\partial}{\partial x_k^l}$. 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 inst., 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 K–R pseudo-normal forms (EKR for short),[3] 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

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

[3] the name borrowed from [PR3], but meaning here much more geometrically different objects than in that work, cf. Rem. 2 and Prop. 1

in the vicinity of $0 \in \mathbb{R}^{k+1}(t,\, x_1^0, \dots,\, x_k^0)$. Then, assuming $\mathbf{j}_1 \dots \mathbf{j}_{r-1}$ 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 $\mathbf{1}, \mathbf{2}, \dots, \mathbf{k+1}$ 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 detailly explained in the proof of Thm. 3 below. When $k = 1$, the two operations **1** and **2** applied interweavingly lead to the well-known local K–R pseudo-normal forms, [KRui], for Goursat flags (to be precise: in certain raw, not polished form, cf. Prop. 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 neighbourhood 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 neighbourhood 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 know yet 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. It certainly does when $k = 1$, see Prop. 1, (ii) (compare also Obs. 1 below, answering a first non-trivial question in this respect).

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

Let us assume the fact under proof for a value $r - 1 \geq 0$ (in the beginning of induction, $r - 1 = 0$, there is nothing to prove), 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 Thm. 2. Then, watching the proof of *that* theorem, $(D^{k+1},\, p)$ is equivalent to the generalized Cartan prolongation

$$\Big(\big(D^{2k+1}/E^k \big)^{(1)}\,,\ \ \big(df(p) D^{k+1}(p),\, f(p) \big) \Big)\,, \tag{11}$$

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 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

$$(12)\quad \psi_*\left(D^{2k+1}/E^k\right) \ \text{ is an EKR } \ \mathbf{j}_1 \dots \mathbf{j}_{r-1} = (Z_1, \dots, Z_{k+1}) \ \text{ on } \ (\mathbb{R}^{rk+1}, 0)$$

satisfying the rule of the least upward jumps (recall that for $r = 1$, in (12) 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 nbhd 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,

$$\psi^\bullet_* \ : \ \big(D^{2k+1}/E^k\big)^{(1)} \longrightarrow (Z_1, \dots, Z_{k+1})^{(1)}\,. \tag{13}$$

(The prolongation of a conjugating diffeo locally conjugates the relevant gCp's. For $k = 1$, see also Appendix in [M4] in this respect.) Returning to (12), the line $df(p)D^{k+1}(p)$ is sent by $d\psi(f(p))$ to a line

$$\mathrm{span}(Z_j + c_j Z_{j+1} + \cdots + c_k Z_{k+1})(0)\,, \tag{14}$$

and $1 \le j \le k+1$ 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 (14), and to identify it as a, longer, EKR. And **secondly**, to take care that the eventual constructed EKR satisfy 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 (14) 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 (14), $(Z_1, \dots, Z_{k+1})^{(1)}$ is equal to

$$\left(x^r_1 Z_1 + \cdots + x^r_{j-1} Z_{j-1} + Z_j + X^r_j Z_{j+1} + \cdots + X^r_k Z_{k+1}\,,\ \frac{\partial}{\partial x^r_1}, \dots, \frac{\partial}{\partial x^r_k} \right) \tag{15}$$

This, indeed, is the 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 (11) which is equivalent –

see (13) – to a distribution identified as $\mathbf{j}_1 \ldots \mathbf{j}_{r-1}\mathbf{j}$. 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 the cyclic change of the first j coordinates $t \to x_1^0 \to x_2^0 \to \cdots \to x_{j-1}^0 \to t$ (and keeping the remaining ones). This change replaces the first generator in our $\mathbf{j}$,

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

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

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

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 $x_m^{r-1} \to x_{m+1}^{r-1} \to \cdots \to x_{j-1}^{r-1} \to x_m^{r-1}$. Then the first v. f. in (15),

$$x_1^r Z_1 + \cdots + x_m^r Z_m + x_{m+1}^r Z_{m+1} + \cdots + x_{j-1}^r Z_{j-1} + Z_j + X_j^r Z_{j+1} + \cdots ,$$

assumes the form

$$x_1^r [Z_1] + \cdots + x_m^r Z_m + x_{m+1}^r Z_{m+2} + \cdots + x_{j-1}^r Z_j + Z_{m+1} + X_j^r Z_{j+1} + \cdots ,$$

where, as previously, $Z_l = \frac{\partial}{\partial x_{l-1}^{r-1}}$, $l = 2, \ldots, k+1$, 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 v. f. reads

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

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 wrt the superscripts) change of coordinates, also cyclic $x_m^{r-2} \to x_{m+1}^{r-2} \to \cdots \to x_{j-1}^{r-2} \to x_m^{r-2}$ after which $[Z_1]$ assumes the form

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

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 $x_m^0 \to x_{m+1}^0 \to \cdots \to x_{j-1}^0 \to x_m^0$. 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. $\square$

3.4 Examples and a first surprise concerning EKRs for $k \geq 2$.

A systematic theory of the EKRs still remains to be done. One of basic tools, naturally, should be the *small growth vector* (s.gr.v.) of a distribution at a point. This tool, however, shows signs of being less efficient for multi-flags than it has been for 1-flags. In order to illustrate the new kind of problems, let us fix $k = 2$. 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 $j = 1,\, 2,\, 3,$ here is the family of distribution germs **1.1.2** (after the simplification of Prop. 1, (i) below):

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

These EKRs have at $0 \in \mathbb{R}^9$, regardless of the value of a constant in Y_4, the s.gr.v. [3, 5, 6, 7, 8, 9].

And here is the family of distribution germs **1.2.2** (after a similar simplification):

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

These EKRs have at $0 \in \mathbb{R}^9$ the s.gr.v. that does not depend on c_3 in Y_3, but which depends on the value of c_4 in Y_4. Namely, for $c_4 = 0$ it is [3, 5, 6, 7, 8, 8, 9], and for $c_4 \neq 0$ it is [3, 5, 6, 7, 8, 9], exactly as for **1.1.2**. This is a first sign of plausible deep differences between multi-flags and 1-flags. For 1-flags such coincidences of s.gr.v. do not happen among different classes (recalled in Prop. 1, (ii)) of Kumpera–Ruiz pseudo-normal forms defined by freezing the code in ciphers **1**, **2**. Basically, this is shown in [PR1], pp. 142–4 (if not in a flawless way): the families of s. gr. v. pertinent to different KR classes are *disjoint.*

Thus a question imposes by itself if the EKRs in **1.1.2** and **1.2.2** with $c_4 \neq 0$ really are geometrically different (no matter what precise constants appear in them). The nilpotency orders of the real Lie algebras induced by these EKRs are different: 6 and 7, as computed in Ex. 1 below. This, however, does not yet show the non-equivalence, because the EKRs' bases are by no means canonical. The answer is given only in Obs. 1 below. And here are some more general informations on EKRs.

Proposition 1. (i) *Fix any $k \geq 1$. All germs in* **1.1...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 are obtained so-called Goursat normal forms, or 'chained systems' still actively used in geometric control theory.*

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

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

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

PROOF of Part (iii): $\mathbf{1.1}$ 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 $\mathbf{1.2}$ 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:

$[3,5,7,9]$ in $\mathbf{1.1.1}$,

$[3,5,7,8,9]$ and $[3,5,7,8,8,9]$ in $\mathbf{1.2.1}$,

$[3,5,6,7,8,9]$ in $\mathbf{1.1.2}$,

$[3,5,6,7,8,9]$ and $[3,5,6,7,8,8,9]$ in $\mathbf{1.2.2}$. $\square$

It is to be noted that only the operations $\mathbf{1}$ and $\mathbf{2}$ are, modulo a reindexation of coordinates, used in producing EKRs in [PR3], Thm. 3.2 there. In consequence, the EKRs in that paper do not cover all germs of distributions generating special 2-flags (and all the more so special k-flags, $k > 2$).

Observation 1. *For $k = 2$, in the family of distribution germs* $\mathbf{1.2.2}$ *with* $c_4 \neq 0$ *there is no object equivalent to any object in the family* $\mathbf{1.1.2}$*.*

PROOF. One examines these germs more closely by computing their s.gr.v. not only at $0 \in \mathbb{R}^9(t,\, x_1,\, y_1, \dots,\, x_4,\, y_4)$, but also in its vicinity. The computations for EKRs of (very small) length 3 are straightforward, and results depend on points. The proof will be finished if the thus obtained partitions of $\mathbb{R}^9$ are not diffeomorphic as the germs at 0. And indeed, for any object in $\mathbf{1.1.2}$, understanding its defining formulas in an open set containing 0, its

$$\text{s. gr. v. at } (t,\, x_1,\, y_1, \dots,\, x_4,\, y_4) \;=\; \begin{cases} [3,\, 5,\, 7,\, 9]\,, & \text{when } x_4 \neq 0\,, \\ [3,\, 5,\, 6,\, 7,\, 8,\, 9]\,, & \text{when } x_4 = 0\,; \end{cases}$$

while for any object in $\mathbf{1.2.2}$ with $c_4 \neq 0$, its

$$\text{s. gr. v. at } (t,\, x_1,\, y_1, \dots,\, x_4,\, y_4) \;=\; \begin{cases} [3,\, 5,\, 7,\, 9]\,, & \text{when } x_4 \neq 0,\ x_3 \neq 0, \\ [3,\, 5,\, 7,\, 8,\, 9]\,, & \text{when } x_4 \neq 0,\ x_3 = 0, \\ [3,\, 5,\, 6,\, 7,\, 8,\, 9]\,, & \text{when } x_4 = 0\,. \end{cases}$$

[4] The question of true significance of constants in the EKRs is enormously involved already in the Goursat case, to quote but [KRui, Ga, ChM], [ChMPR, PR1], [M1, M2, MonZ], [M4] (this list could be longer).

These two germs at 0 of partitions (even: stratifications) of $\mathbb{R}^9$ are clearly not diffeomorphic. $\square$

Remark 3. Another consequence of Thm. 2, modulo the arguments used already in the proof of Thm. 3 above, is that, upon applying r times multi-dimensional Cartan prolongation to the tangent bundle to a $(k+1)$-dimensional manifold (for inst. 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 Thm. 4.2 in [BH]; it is done independently in Sec. 5.4 in [MonZ].

(And, reiterating after Sec. 1, that same theorem virtually *produces* all local K–R 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 $\mathbf{j} \in \{\mathbf{1}, \dots, \mathbf{k+1}\}$, by its values on a – fixed once for all – basis $e_1, \dots, e_{k+1}$ of $\mathbb{R}^{k+1}$. Namely, let $\overleftarrow{\mathbf{j}}$ send: $e_1 \mapsto e_1 + e_2 + \cdots + e_{k+1}$, $e_2 \mapsto e_1, \dots, e_j \mapsto e_{j-1}$, $e_{j+1} \mapsto e_{j+1}, \dots, e_{k+1} \mapsto e_{k+1}$ (so that, naturally, $\overleftarrow{\mathbf{1}}$ sends $e_2 \mapsto e_2, \dots, e_{k+1} \mapsto e_{k+1}$, and $\overleftarrow{\mathbf{k+1}}$ sends $e_2 \mapsto e_1, \dots, e_{k+1} \mapsto e_k$). With these notations,

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* [M3]*) and a local nilpotent basis around a point p is $\{Z_1, \dots, Z_{k+1}\}$, where $(Z_1, \dots, Z_{k+1}) \in \mathbf{j}_1 \mathbf{j}_2 \dots \mathbf{j}_r$ issuing from* Thm. 3 *is an* EKR *form for D around p satisfying the least upward jumps rule.*

The nilpotency order of the real Lie algebra $L(\mathbf{j}_1 \mathbf{j}_2 \dots \mathbf{j}_r) \overset{\text{def}}{=} L_{\mathbb{R}}(Z_1, \dots, Z_{k+1})$ is equal to the component of e_{k+1} (i. e., the last one) in the vector $\overleftarrow{\mathbf{j}_1}\, \overleftarrow{\mathbf{j}_2} \cdots \overleftarrow{\mathbf{j}_r}(e_1 + e_2 + \cdots + e_{k+1})$.

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* [AG], [M3].

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 in 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 $L_{\mathbb{R}}(Z_1, \dots, Z_{k+1})$ and $L_{\mathbb{R}}(\widetilde{Z}_1, \dots, \widetilde{Z}_{k+1})$, 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, cf. also Rem. 1 in [M5]). 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].

It is also to be noted that when $m = \max(j_1, \dots, j_r) < k$, then not only the last component, but precisely $k+1-m$ last components in $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2}\,\cdots\,\overleftarrow{\mathbf{j}_r}(e_1 + \cdots + e_{k+1})$ 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 1. *For distribution germs D as in the 'Moreover' part of* Thm. 4 *(for $k = 1$ they are called 'tangential' in* [M3]*), 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 $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2}\,\cdots\,\overleftarrow{\mathbf{j}_r}(e_1 + \cdots + e_{k+1})$, too.*

Still before proving Thm. 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,

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

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 Thm. 4 allows to re-prove a part of Jean's results [J] 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 are given in [M5].

PROOF of Theorem 4 uses refinements of methods of [M3, M5]. By Rem. 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) $w\big(\frac{\partial}{\partial x}\big) = -\,w(x)$ in such a way that – under the classical definition of the weight of a monomial vector field that goes back to the 1970s,

$$w\Big(y_{i_1}y_{i_2}\cdots y_{i_s}\frac{\partial}{\partial x}\Big) \;=\; w(y_{i_1})+w(y_{i_2})+\cdots+w(y_{y_s})+w\Big(\frac{\partial}{\partial x}\Big) \tag{16}$$

– all $Z_1,\,Z_2,\ldots,\,Z_{k+1}$ will be homogeneous of degree -1, and all v. f. 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,

$$Z_1 = x_1^r Z_1^{r-1}+x_2^r\frac{\partial}{\partial x_1^{r-1}}+\cdots+x_{j_r-1}^r\frac{\partial}{\partial x_{j_r-2}^{r-1}}+\frac{\partial}{\partial x_{j_r-1}^{r-1}}+x_{j_r}^r\frac{\partial}{\partial x_{j_r}^{r-1}}+\cdots+x_k^r\frac{\partial}{\partial x_k^{r-1}}\,, \tag{17}$$

where the involved v. f. from the one before last step is denoted by Z_1^{r-1} (and there is no constants in it, too). Let us start by declaring $w(Z_1)=w(Z_2)=\cdots=w(Z_{k+1})=-1$, hence also $w(x_1^r)=\cdots=w(x_k^r)=1$. By writing $w(Z_1)=-1$, we stipulate that *all* summands in (17) 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 v. f. Z_1.

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

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

Analyzing the consecutive summands on the RHS, the homogeneity of Z_1^s implies that

$$w(Z_1^{s-1})=w(Z_1^s)+w\Big(\frac{\partial}{\partial x_1^s}\Big) \tag{18}$$

and that

$$w\Big(\frac{\partial}{\partial x_l^{s-1}}\Big) \;=\; w(Z_1^s)+w\Big(\frac{\partial}{\partial x_{l+1}^s}\Big) \qquad \text{for } \; l=1,\ldots,\,j_s-2\,, \tag{19}$$

$$w\Big(\frac{\partial}{\partial x_{j_s-1}^{s-1}}\Big) \;=\; w(Z_1^s)\,, \tag{20}$$

$$w\Big(\frac{\partial}{\partial x_l^{s-1}}\Big) \;=\; w(Z_1^s)+w\Big(\frac{\partial}{\partial x_l^s}\Big) \qquad \text{for } \; l=j_s,\ldots,\,k\,. \tag{21}$$

In this way inductively defined are the negative weights of all $(r+1)k+1$ versors, including $\frac{\partial}{\partial t}=Z_1^0$, 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 (16) is the **additivity** under Lie multiplication. So the Lie products of any s factors from among $\{Z_1,\dots,Z_{k+1}\}$ are homogeneous v. f.'s of weight $-s$. In particular, products of more than d factors automatically vanish (our vector monomials have only weights $\geq -d$). On the other hand, D, generating a special flag, is completely non-holonomic. In particular $\frac{\partial}{\partial x}$ 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. Because there is no coordinates shifted by constants in Z_1, only products of at least d factors are relevant in such a combination. Hence – here – only products of exactly d factors. This implies 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, $D=V_1\subset V_2\subset V_3\subset\cdots$ ends (locally, as a germ) on its term V_d,

$$V_d \;=\; T\mathbb{R}^{(r+1)k+1}. \tag{22}$$

How to 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 (18) – (21) can be most compactly expressed as
(23)

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

Hence, by making the composition over $s=r,\,r-1,\dots,1$,
(24)

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

Each Z_1^{s-1} $(s\geq 2)$ has as a component the bare versor $\frac{\partial}{\partial x^{s-2}_{j_{s-1}-1}}$, and Z_1^0 simply is $\frac{\partial}{\partial t}$. Thus $w(Z_1^{s-1})$ is a weight of a versor, and in (23) 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 LHS of (24).

Pf. 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 $a_1, a_2, \ldots, a_j, a_{j+1}, \ldots, a_{k+1}$, hence bigger than $\max(a_1, \ldots, a_{k+1})$. $\square$

It remains to explain why the last component, and possibly not only it, on the LHS of (24) is the biggest in modulus. Not surprisingly, the least upward jumps rule satisfied by our EKR is responsible for that. In fact, for $m = \max(j_1, \ldots, j_r)$, the last $k+1-m$ components of $\overleftarrow{\mathbf{j}_1}\,\overleftarrow{\mathbf{j}_2} \cdots \overleftarrow{\mathbf{j}_r}[1,\, 1, \ldots,\, 1]^T$ are pairwise equal (we work in the sequel with positive weights of variables only).

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

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

is smaller than the components No $m+1, \ldots, k+1$, and this relation keeps holding after applying each of the subsequent operators $\overleftarrow{\mathbf{j}_{l-1}}, \ldots, \overleftarrow{\mathbf{j}_1}$.

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 $j_1, \ldots, j_r$, and so let it appear for the first time as $j_s = m-1$, $s < l$. Then the $(m-1)$-th component in

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

is smaller than the components No $m, m+1, \ldots, k+1$, and this relation keeps holding after applying each of the subsequent operators (if there remains any) $\overleftarrow{\mathbf{j}_{s-1}}, \ldots, \overleftarrow{\mathbf{j}_1}$.

Thus at this point we know that, in the eventual outcome $\overleftarrow{\mathbf{j}_1} \cdots \overleftarrow{\mathbf{j}_r}[1,\, 1, \ldots,\, 1]^T$, 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 $\mathbf{j}_1\,\mathbf{j}_2 \cdots \mathbf{j}_r$, obtaining in the end that, in the vector $\overleftarrow{\mathbf{j}_1} \cdots \overleftarrow{\mathbf{j}_r}[1,\, 1, \ldots,\, 1]^T$, the 1st component is smaller than the 2nd 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 Thm. 4. Thanks to the choice of an EKR in the beginning, we can work all the time with the same $\{Z_1, \ldots, Z_{k+1}\}$. Divide the constructed weights of coordinates $t, x_1^0, \ldots, x_k^r$ into groups of equal values, the highest value being d: $w_1 = w_2 = \cdots = w_{n_1} = 1$, $n_1 = k+1$ (it is clear that, on top of $1 = w(x_1^r) = \cdots = w(x_k^r)$, precisely one more variable $x_{j_r-1}^{r-1}$ has weight 1), $w_{n_1+1} = \cdots = w_{n_2} = 2, \ldots, w_{n_{l-1}+1} = \cdots = w_{n_l} = l, \ldots, w_{n_{d-1}+1} = \cdots = w_{n_d} = d$, 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 *$[n_1, n_2, \ldots, n_d]$ 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*

$$V_j(0) = \left(\frac{\partial}{\partial x}, \quad x:\ w(x) \le j \right). \tag{25}$$

That is to say, the coordinates $t,\, x_1^0, \ldots,\, x_k^r$ *are – for* D *– linearly adapted at* $0 \in \mathbb{R}^{(r+1)k+1}$.

Pf. V_j is spanned by the products of $\leq j$ factors from among $\{Z_1,\, Z_2, \ldots,\, Z_{k+1}\}$. Such products are, as we already know, homogeneous of weights $\geq -j$ and, evaluated at 0, are combinations of *bare* versors (because there is no constants in $(Z_1, \ldots,\, Z_{k+1})$). Those are, therefore, the versors of coordinates x s.t. $w(x) \leq j$. This means the inclusions . $\subset$. in (25). Clearly, for $j = 1$, with $\operatorname{rk} V_1 = k+1$, there is an equality in (25). Suppose that j is the smallest positive integer s.t. there holds $\subsetneq$ in (25). (In view of (22), assuredly $j < d$.) That is, that there exists a combination

$$\sum_{x:\ w(x)=j} a_x \frac{\partial}{\partial x} \ \notin\ V_j(0)\,. \tag{26}$$

Then

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

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

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

This together with (26) imply $\sum_{x:\ w(x)=j} a_x \frac{\partial}{\partial x} \notin V_d(0)$ which contradicts (22).

The description (25), now shown to hold, allows to compute the s.gr.v. of D at 0. For $j = 2, \ldots,\, d$, $\dim V_j(0) = n_1 + (n_2 - n_1) + \cdots + (n_j - n_{j-1}) = n_j$. Prop. 2 is proved.

Summarizing now, our weights $w(\cdot)$ are modelled on the pattern established by the s.gr.v. of D at 0. With this knowledge, Prop. 2 says precisely that the coordinates in question are linearly adapted for D at 0 (cf., for inst., [AGS, Bel, M5]). 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 $t,\, x_1^0, \ldots,\, x_k^r$ 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, wrt $\{Z_1, \ldots,\, Z_{k+1}\}$ 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. p. 257 in [M3]). 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 $Z_2, \ldots,\, Z_{k+1}$, too, are of degree (weight) -1, so that the nilpotent approximation of D at 0 simply coincides with D. This exactly means ([AG]; see also [M3], Def. 5) that the germ D is strongly nilpotent. $\square$

5 Algorithmic issues.

The last issue is that of possible more algorithmic ways of computing the nilpotency orders, not via bringing in the entire matrices of operators $\overleftarrow{\mathbf{j}}$ required in Thm. 4. For $k = 1$ this task has been satisfactorily carried out in [M3, M5] (in the latter without referencing to [J], cf. Rem. 5). 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 cumbersome than for $k = 2$, and so on. The brunt of the description keeps being recursive, but there appears plenty of side effects entirely absent for $k = 1$.

So, from now on, set $k = 2$ and use the letter $\mathbf{C}$ for codes of different EKRs. Moreover, denote by $f(\mathbf{C})$ the nilpotency order of the Lie algebra $L(\mathbf{C})$ issuing from Thm. 4. A string of l same ciphers, like for inst. a string of $\mathbf{1}$'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 $\mathbf{1}.\mathbf{1}.\mathbf{2}.\mathbf{2}.\mathbf{2}$, and so on. Despite our restricting to the EKRs satisfying the least upward jumps rule, the algorithm requires to enlarge the pool of concerned EKRs. Also pay attention that, when an EKR *not* satisfying that rule enters the function $f(\cdot)$, like for inst. in $f(\mathbf{3})$, then the respective value of f need *not* indicate the relevant nilpotency order. Here are the recipes for computing the nilpotency orders of $L(\mathbf{C})$ from Thm. 4.

Theorem 5. *In the situation $k = 2$, for any* EKR $\mathbf{C}$ *satisfying the least upward jumps rule, 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* $\mathbf{3}$.

$$f(\emptyset) = 1, \quad f(\mathbf{1}) = f(\mathbf{2}) = 2, \quad f(\mathbf{3}) = 1\,,$$
$$f(\mathbf{C}.*.\mathbf{1}) \;=\; 2f(\mathbf{C}.*) - f(\mathbf{C})\,,$$
$$f(\mathbf{C}.\mathbf{1}_l.\mathbf{2}) \;=\; 8l f(\mathbf{C}) - (2l-1)f(\mathbf{C}.\mathbf{2}) - 2l\, f(\mathbf{C}.\mathbf{3})\,,$$
$$f(\mathbf{C}.\mathbf{1}_l.\mathbf{3}) \;=\; 8l f(\mathbf{C}) - 2l\, f(\mathbf{C}.\mathbf{2}) - (2l-1)f(\mathbf{C}.\mathbf{3})\,,$$
$$f(\mathbf{C}.*.\bullet.\mathbf{2}) \;=\; 2f(\mathbf{C}.*.\bullet) - f(\mathbf{C})\,,$$
$$f(\mathbf{C}.*.\bullet.\mathbf{3}) \;=\; f(\mathbf{C}.*.\bullet) + f(\mathbf{C}.*) + f(\mathbf{C})\,.$$

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 $k = 1$ (Thm. 1 in [M3]): in the Goursat case the nilpotency orders $g(\cdot)$ satisfy the relations

$$g(\emptyset) = 1, \quad g(\mathbf{1}) = 2, \quad g(\mathbf{2}) = 1\,,$$
$$g(\mathbf{C}.*.\mathbf{1}) \;=\; 2g(\mathbf{C}.*) - g(\mathbf{C})\,,$$
$$f(\mathbf{C}.*.\mathbf{2}) \;=\; g(\mathbf{C}.*) + g(\mathbf{C})\,,$$

where an $*$ now stands for $\mathbf{1}$ or $\mathbf{2}$.

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

$$\left(\overleftarrow{\mathbf{1}}\right)^3 \overleftarrow{\mathbf{2}} \left(\overleftarrow{\mathbf{3}}\right)^3 \overleftarrow{\mathbf{2}} \left(\overleftarrow{\mathbf{1}}\right)^4 \overleftarrow{\mathbf{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 114 \\ 405 \\ 442 \end{pmatrix},$$

implying that this order is 442. Secondly, according to Thm. 5, we proceed stepwise, expanding progressively the code to the right:

$$\begin{aligned}
&f(\mathbf{1}_3.\,\mathbf{2}) = 24f(\emptyset) - 5f(\mathbf{2}) - 6f(\mathbf{3}) = 8\,,\\
&f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}) = f(\mathbf{1}_3.\,\mathbf{2}) + f(\mathbf{1}_3) + f(\mathbf{1}_2) = 8 + 4 + 3 = 15\,,\\
&f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_2) = f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}) + f(\mathbf{1}_3.\,\mathbf{2}) + f(\mathbf{1}_3) = 15 + 8 + 4 = 27\,,\\
&f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3) = f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_2) + f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}) + f(\mathbf{1}_3.\,\mathbf{2}) = 27 + 15 + 8 = 50\,,\\
&f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}) = 2f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3) - f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}) = 2\cdot 50 - 15 = 85\,,\\
&f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}.\,\mathbf{1}_4.\,\mathbf{3}) = 32f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}) - 8f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}_2) - 7f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}.\,\mathbf{3})\\
&= 32\cdot 85 - 8\cdot 143 - 7\cdot 162 = 442\,,
\end{aligned}$$

as in the first way (because $f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}_2) = 2f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}) - f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_2) = 2\cdot 85 - 27 = 143$ and $f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}.\,\mathbf{3}) = f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3.\,\mathbf{2}) + f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_3) + f(\mathbf{1}_3.\,\mathbf{2}.\,\mathbf{3}_2) = 85 + 50 + 27 = 162$).

References

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

[AG] ——, J–P. Gauthier, *On subanalyticity of Carnot–Carathéodory distances.* Preprint 25/2000/M, SISSA, Trieste.

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

[BCG+] R. Bryant, S. Chern, R. Gardner, H. Goldschmidt, P. Griffiths, *Exterior Differential Systems*, MSRI Publ. Vol. 18, Springer Verlag, Berlin 1991.

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

[C] E. 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.

[ChM] M. Cheaito, P. Mormul, *Rank–2 distributions satisfying the Goursat condition: all their local models in dimension 7 and 8*, ESAIM: Control, Optimisation and Calculus of Variations (http://www.edpsciences.org/cocv/) **4** (1999), 137 – 158.

[ChMPR] ——, ——, W. Pasillas-Lépine, W. Respondek, *On local classification of Goursat structures*, C. R. Acad. Sci. Paris **327** (1998), sér. I, 503 – 508.

[Ga] M. Gaspar, *Sobre la clasificacion de sistemas de Pfaff en bandera*, in: Proc. of the 10^{th} Spanish–Portuguese Conf. on Math., Univ. of Murcia, Murcia 1985, 67 – 74 (in Spanish).

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

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

[JaPrz] B. Jakubczyk, F. Przytycki, *Singularities of k–tuples of vector fields*, Dissertationes Math. **213**, Inst. of Math., Polish Acad. Sci, Warszawa 1984.

[J] F. Jean, *The car with n trailers: characterisation of the singular configurations*, ESAIM: Control, Optimisation and Calculus of Variations (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*, Gordon & Breach, London 1986.

[K] A. Kumpera, *Flag systems and ordinary differential equations*, Annali Mat. Pura ed Appl. **177** (1999), 315 – 329.

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

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

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

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

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

[M1] P. Mormul, *Local classification of rank–2 distributions satisfying the Goursat condition in dimension* 9, in: P. Orro and F. Pelletier (eds.), Singularités et Géométrie sous-riemannienne, Travaux en cours **62**, Hermann, Paris 2000, 89 – 119.

[M2] ——, *Goursat flags: classification of codimension–one singularities*, J. Dynam. Control Systems **6** (2000), 311 – 330.

[M3] ——, *Goursat distributions not strongly nilpotent in dimensions not exceeding seven*, in: A. Zinober and D. Owens (eds.), Nonlinear and Adaptive Control NCN4 2001, Lecture Notes in Control and Information Sciences **281**, Springer Verlag, Berlin – New York 2003, 249 – 261.

[M4] ——, *Real moduli in codimension* 2 *in local classification of Goursat flags.* Preprint in: S. Janeczko (ed.), Singularity Theory Seminar, Warsaw Univ. of Technology, Vol. 7 (2002), 58 – 92.

[M5] ——, *Minimal nilpotent bases for Goursat distributions of coranks not exceeding six.* Preprint, Warsaw 2002.

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

[PR1] W. Pasillas-Lépine, W. Respondek, *On the geometry of Goursat structures*, ESAIM: Control, Optimisation and Calculus of Variations (http://www.edpsciences.org/cocv/) **6** (2001), 119 – 181.

[PR2] ——, ——, *Nilpotentization of the kinematics of the n-trailer system at singular points and motion planning through the singular locus*, Int. J. Control **74** (2001), 628 – 637.

[PR3] ——, ——, *Contact systems and corank one involutive subdistributions*, Acta Appl. Math. **69** (2001), 105 – 128.

[Sl] W. 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. **8** (1982), 109 – 176.

[Y2] ——, *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.