UNIVERSITATIS IAGELLONICAE ACTA MATHEMATICA, FASCICULUS XLII
2004

MINIMAL NILPOTENT BASES FOR GOURSAT DISTRIBUTIONS OF CORANKS NOT EXCEEDING SIX

BY PIOTR MORMUL

Abstract. After bringing in some basic notions of the theory of nilpotent approximations, we first give a new proof, this time entirely independent of the work [**12**], of a result proved in [**19**] (and earlier still in [**18**]) that Goursat distributions of arbitrary corank locally possess nilpotent bases of *explicitly computable* orders of nilpotency of the generated Lie algebras (KR algebras). Recalling that the germs of such distributions are stratified into *geometric classes* of Jean, Montgomery and Zhitomirskii, in certain geometric classes termed *tangential*, the computed nilpotency orders of KR algebras turn out to coincide with the nonholonomy degrees, computed by Jean, at the reference points for germs. In the tangential classes, then, the nilpotency orders of KR algebras are minimal among all possible nilpotent bases. As regards the remaining (non-tangential) classes, it is a vast standing Question whether the KR algebras realize the minimum of possible nilpotency orders; the difference between their nilpotency orders and nonholonomy degrees is *unbounded* as corank/length tends to infinity. In small lengths – through 5 inclusively, and also for 8 non-tangential classes in length 6 (out of altogether 18 non-tangential existing in that length 6) – we show that the answer is *yes*: the nilpotency orders of KR algebras cannot be lowered. This extended experience makes plausible a general answer to Question in the affirmative.

1. Nilpotent approximation of a distribution at a point. We want to start from some notions related to the nilpotent approximations of geometric distributions. For any distribution D of rank d on an n–dimensional, smooth or real analytic, manifold M (i. e., a rank–d subbundle in the tangent bundle TM) its *small flag* is the nested sequence

$$V_1 \subset V_2 \subset V_3 \subset V_4 \subset \cdots$$

Research supported by Polish KBN Grant 2 P03A 010 22.

of modules (or: presheaves of modules) of vector fields (of the same category as M) tangent to M: $V_1 = D,\ V_{j+1} = V_j + [D,\, V_j]$ for $j = 1,\, 2, \ldots$ The *small growth vector* at $p \in M$ is the sequence (n_j) of linear dimensions at p of the modules V_j: $n_j = \dim V_j(p)$. Naturally, $n_1 = d$ independently of p.

D is *completely nonholonomic* when at every point of M its small growth vector attains (sooner or later) the highest value $n = \dim M$. Once this value is attained, we truncate the vector after the first appearance of n in it. The length d_{NH} of thus truncated vector is called the *nonholonomy degree* of D at p.

In the theory that we only sketch here (cf. [**9**], [**1**], [**4**], [**5**], [**10**], [**3**]; this list of references is not complete) important are the *weights* w_i related to the small flag at a point: $w_1 = \cdots = w_d = 1, \quad w_{d+1} = \cdots = w_{n_2} = 2$ (no value 2 among them when $n_2 = n_1$ $(= d)$, and generally

$$w_{n_j+1} = \cdots = w_{n_{j+1}} = j + 1 \tag{1}$$

(no value $j + 1$ among the w's when $n_j = n_{j+1}$) for $j = 1,\, 2, \ldots$

Definition 1. For a completely nonholonomic distribution D on M, coordinates $z_1,\, z_2, \ldots,\, z_n$ around $p \in M$ are *linearly adapted* at p when $D(p) = (\partial_1, \ldots, \partial_d)$, $V_2(p) = (\partial_1, \ldots,\, \partial_d, \ldots,\, \partial_{n_2})$, and so on until $V_{d_{\mathrm{NH}}}(p) = (\partial_1, \ldots,\, \partial_n) = T_pM$. Here and in the sequel we skip writing 'span' before a set of v. f. generators.
For such linearly adapted coordinates we define their *weights* $w(z_i) = w_i$, $i = 1, \ldots,\, n$.

On the other hand, having a completely nonholonomic D, every smooth function f on M near p has its *nonholonomic order* wrt D at p ($+\infty$ is not excluded). It is the minimal number of differentiations of f along the local generators of D that give at p a nonzero result.

It follows directly from the above definitions that, for linearly adapted coordinates, their nonholonomic orders **do not exceed** their weights. Linearly adapted coordinates $z_1, \ldots,\, z_n$ are *adapted* (or: *privileged*) when the nonholonomic order of z_i **equals** $w(z_i)$, $i = 1, \ldots,\, n$. (In particular, adapted coordinates must vanish at p; one says that they are centered at p.)

It is not immediate to show, but adapted coordinates always exist, and can even be algorithmically constructed from any à priori given coordinates, even in a polynomial way, as is done, e. g., in [**1**] or [**3**]. They are not unique, there remains plenty of liberty behind the requirement that the nonholonomic orders (of linearly adapted coordinates) be maximal possible. In adapted coordinates it is purposeful to attach quasihomogeneous weights also to monomial vector fields (this definition goes back to the 1970s, to the theory of differential operators; for geometric distributions, see in this respect [**1**], p. 215),

$$w(z_{i_1}\cdots z_{i_k}\partial_j) \;=\; w(z_{i_1})+\cdots+w(z_{i_k})-w(z_j)\,. \tag{2}$$

Proposition 1. *Every smooth vector field X with values in D has in its Taylor expansion in arbitrary coordinates adapted for D terms only of weights not smaller than -1 that can be grouped in homogeneous summands $X = X^{(-1)} + X^{(0)} + X^{(1)} + \cdots$*

(superscripts mean the weights defined by (2)). We denote by $\widehat{X}$ the lowest ('nilpotent') summand $X^{(-1)}$. That is, $\widehat{X} = X^{(-1)}$.

When a distribution D has local generators (vector fields) $X_1, \dots, X_d$ around p, then

Definition 2. The distribution $\widehat{D} = \big(\widehat{X_1}, \dots, \widehat{X_d}\big)$ is called the *nilpotent approximation* of D at p.

This object $\widehat{D}$ is invariantly defined, independently of the used adapted coordinates, see Prop. 5.20 in [**3**]. Its basic properties are included in the following proposition.

Proposition 2. *The nilpotent approximation $\widehat{D}$ of D has at p the same small growth vector as D (and hence the same nonholonomy degree d_{NH}). Moreover, the real Lie algebra generated by $\widehat{X_1}, \dots, \widehat{X_d}$ is a nilpotent Lie algebra of nilpotency order d_{NH}.*

(The *nilpotency order* of a Lie algebra is the minimal number of multiplications in that algebra yielding always zero, cf. p. 239 in [**14**].)

Definition 3 ([**2**]). Distribution D is *strongly nilpotent* at a point p when its germ at p is equivalent to its nilpotent approximation $\widehat{D}$ at p.

Definition 4 ([**11**], [**19**]). Distribution D is *weakly nilpotent* at a point p when D possesses, locally around p, a basis generating over $\mathbb{R}$ a nilpotent (real) Lie algebra of vector fields.

Strong nilpotency clearly implies the weak one (cf. Prop. 2). In the founding work [**11**], and posterior control theory literature, weakly nilpotent distributions are called *nilpotentizable*. Nilpotent approximations, and in particular strongly nilpotent distributions play a growing role in sub-Riemannian geometry ([**3**], [**2**], also recent works by Bonnard, Chyba, Trelat, Sachkov, to name but few).

2. Kumpera–Ruiz algebras of Goursat distributions. In the sequel we deal with Goursat distributions – a rather restricted class of objects for which preliminary (local) polynomial normal forms of [**13**] exist with real parameters only, and no functional moduli. Their definition requires that the sequence of consecutive Lie squares of the original rank–2 subbundle of TM consist of regular distributions of ranks $3, 4, \dots$ until $n = \dim M$. In [**20**] we already recalled a basic partition of Goursat germs into disjoint *geometric classes* encoded with words of length $n-2$ over the alphabet G, S, T, with the first two letters always G and such that never a T goes directly after a G. Their construction was done by Montgomery and Zhitomirskii; it is reproduced in Sec. 1.3 of [**17**]. (Implicitly these classes are already present in a pioneering work [**12**], in which the author uses a trigonometric presentation of Goursat objects.)

In dimension 4, there is but one class GG, in dimension 5 – GGG and GGS only, in dimension 6 – GGGG, GGSG, GGST, GGSS, GGGS.

The union of all geometric classes ('quarks') of fixed length with letters S in *fixed* positions in the codes is called, after [**15**], a *Kumpera–Ruiz class* (a 'particle') of Goursat germs of that corank. For instance, in dimension 6, the two geometric classes GGSG and GGST build one KR class $**\mathrm{S}*$. In dimension 7 the geometric classes GGSGG, GGSTG, and GGSTT build $**\mathrm{S}**$.

What are the mentioned polynomial (local) presentations of Goursat objects? The essence of the contribution [**13**], given in the notation of vector fields and taking into account posterior works, is as follows. Let us construct first a (not unique, depending on a number of real parameters) rank–2 distribution on $(\mathbb{R}^n(x^1, \dots, x^n), 0)$ departing from the code of a geometric class $\mathcal{C}$.

When the code starts with precisely s letters G, one puts $\overset{1}{Y} = \partial_1$, $\overset{2}{Y} = \overset{1}{Y} + x^3\partial_2\,, \dots, \overset{s+1}{Y} = \overset{s}{Y} + x^{s+2}\partial_{s+1}$. When $s < n-2$, then the $(s+1)$-th letter in $\mathcal{C}$ is S. More generally, if the mth letter in $\mathcal{C}$ is S, and $\overset{m}{Y}$ is already defined, then

$$\overset{m+1}{Y} = x^{m+2}\overset{m}{Y} + \partial_{m+1}\,.$$

But there can also be T's or G's after an S. If the m-th letter in $\mathcal{C}$ is not S, and $\overset{m}{Y}$ is already defined, then

$$\overset{m+1}{Y} = \overset{m}{Y} + \big(c^{m+2} + x^{m+2}\big)\partial_{m+1}\,,$$

where a real constant c^{m+2} is not absolutely free but

- equal to 0 when the m-th letter in $\mathcal{C}$ is T,
- not equal to 0 when the m-th letter is G going directly after a string ST…T (or after a short string S).

Now, on putting $\mathbf{X} = \partial_n$ and $\mathbf{Y} = \overset{n-1}{Y}$, and understanding $(\mathbf{X}, \mathbf{Y})$ as the germ at $0 \in \mathbb{R}^n$, we may recall the following theorem.

Theorem 1 ([**13**]). *Any Goursat germ D on a manifold of dimension n, sitting in a geometric class $\mathcal{C}$, can be put (in certain local coordinates) in a form $D = (\mathbf{X}, \mathbf{Y})$, with certain constants in the field $\mathbf{Y}$ corresponding to* G*'s past the first* S *in $\mathcal{C}$.*

Definition 5. The real Lie algebra $L_{\mathbb{R}}(\mathbf{X}, \mathbf{Y})$ generated by $\mathbf{X}$ and $\mathbf{Y}$ is called the *KR algebra* of the germ D. This algebra does not depend on the choice of coordinates in Thm. 1, although the constants in $\mathbf{Y}$ do.

Remark 1. A short analysis shows that this algebra depends only on the Kumpera–Ruiz class $\mathcal{C}$ of D; henceforth we will write it as $L_{\mathbb{R}}(\mathcal{C})$.
Indeed, take two KR pseudo-normal forms $(\mathbf{X}, \mathbf{Y})$ and $(\mathbf{X}, \widetilde{\mathbf{Y}})$ of two germs D and $\widetilde{D}$, resp., sitting in $\mathcal{C}$; these germs may well belong to different geometric classes. Only the alternatives $(\bullet)$, $(\bullet\bullet)$ occur one and the same at a time, simultaneously in the stepwise construction of both $\mathbf{Y}$ and $\widetilde{\mathbf{Y}}$ – the block structures of these fields are the same, but they differ in constants. We will construct an 'inner' isomorphism of $L_{\mathbb{R}}(\mathbf{X}, \mathbf{Y})$ and $L_{\mathbb{R}}(\mathbf{X}, \widetilde{\mathbf{Y}})$ induced by a diffeomorphism ϕ of the underlying space $\mathbb{R}^n$ into itself sending $\mathbf{X}$ to itself, and $\mathbf{Y}$ to $\widetilde{\mathbf{Y}}$. To this end, denote the constants in $\mathbf{Y}$ (resp., $\widetilde{\mathbf{Y}}$) by c^j (resp., $\tilde{c}^j$; cf. the recursive definition of $\mathbf{Y} = \overset{n-1}{Y}$), $j \in \mathcal{J}$, zero values not excluded. Then a simple $\phi = T_v$ does, where T_v means the translation in $\mathbb{R}^n$ by the vector $v = \sum_{j \in \mathcal{J}} (c^j - \tilde{c}^j)\partial_j$. (This translation replaces the constants in $\mathbf{Y}$ by those in $\widetilde{\mathbf{Y}}$.)

In 2000 we proved, answering a 1998 question of Sussmann, the following.

Theorem 2 ([**19**]). *The KR algebra of any Kumpera–Ruiz class of length r encoded with the word $\mathcal{C}$, $L_{\mathbb{R}}(\mathcal{C})$, is nilpotent of nilpotency order $O_{\mathcal{C}} = d_r$, where d_r is the last term in the sequence $d_1, d_2, \ldots, d_r$ defined as follows: $d_1 = 2$, $d_2 = 3$, $d_{j+2} = d_j + d_{j+1}$ when the $(j+2)$-th letter in $\mathcal{C}$ is* S*, and $d_{j+2} = 2d_{j+1} - d_j$ when the $(j+2)$-th letter in $\mathcal{C}$ is $*$.*

Note that, in [**22**], a weaker result stating just the nilpotency of the algebras $L_{\mathbb{R}}(\mathbf{X}, \mathbf{Y})$ – the *existence* of finite nilpotency orders for them – is given (with the last, fifth item of the relevant proof on p. 631 being, besides, incorrect – implying by far unrealistically small, sublinear in function of the length r, nilpotency orders), and no discussion of more universal objects – algebras depending only on KR classes – is undertaken.

In the present paper we propose, in Sec. 4, a new proof of this theorem. This proof, in contrast to the original one in [**18**] and [**19**], is more algebraic and entirely independent of the results of [**12**]. Cf. also Rem. 2 in this respect.

DEFINITION 6. We call *tangential* geometric classes whose codes possess letters G only at the beginning, before the first S (if any) in the code. Tangential Goursat germs are those sitting in the tangential classes.

EXAMPLE 1. Up to dimension 5, all geometric classes are tangential. The first and unique non-tangential class in dimension 6 is GGSG. In dimension 7, there are eight tangential classes and five non-tangential: GGSGG, GGSTG, GGSGS, GGGSG, GGSSG. In dimension 8, there are sixteen tangential and eighteen non-tangential geometric classes (cf. Thm. 4 below). In a general dimension n, there are as many tangential classes as KR classes, that is to say, 2^{n-4} (one replaces all $*$'s past the first S in the code of a KR class with letters T).

Tangential classes become clearly visible when one uses the polynomial presentation of Kumpera–Ruiz.

OBSERVATION 1. *A Goursat germ D is tangential $\Longleftrightarrow$ in any KR presentation* $(\mathbf{X},\, \mathbf{Y})$ *of* D *there is no non-zero constant*

(so tangential germs are easily given local models with no parameters).

3. Minimality of KR algebras in Goursat germs of small coranks. How to compute the degree of nonholonomy of Goursat germs? The answer (given by Jean originally only for the car systems) is, after all, included in [**12**], because the car systems are universal models for *germs* of Goursat distributions (i. e., for the Goursat objects understood locally). In fact, it is proved in [**15**] that the car model with $r-1$ trailers is the r times Cartan prolongation of the tangent bundle to a plane, and hence is modelling all corank–r Goursat germs. A key underlying structural theorem in this respect is Thm. 4.1 in [**6**], see [**21**] for (much) more comments on that. The essence of Jean's results in [**12**] can be conveyed – after taking into account the above remarks – as follows.

THEOREM 3 ([**12**]). *The nonholonomy degree $d_{\mathcal{G}}$ of any corank–r Goursat germ in the geometric class $\mathcal{G}$ equals the last term b_r in the sequence $b_1, b_2, \ldots, b_r$ defined only in terms of $\mathcal{G}$: $b_1 = 2,\ \ b_2 = 3$,*

$b_{j+2} = b_j + b_{j+1}$ *when the $b(j+2)$-th letter in $\mathcal{G}$ is* S,

$b_{j+2} = 2b_{j+1} - b_j$ *when the $(j+2)$-th letter in $\mathcal{G}$ is* T,

$b_{j+2} = 1 + b_{j+1}$ *when the $(j+2)$-th letter in $\mathcal{G}$ is* G.

The formulas in Thm. 2 have much in common with the ones in this theorem, so that it imposes by itself to ask about the relationship between the

nonholonomy degrees and nilpotency orders. It is a general geometric fact that, whenever they both exist, the former do not exceed the latter. And this is clearly visible on the level of Goursat distributions, for the sequences b in Thm. 3 grow not faster (and often slower) than the sequences d in Thm. 2. This is due to the third alternative $b_{j+2} = 1 + b_{j+1}$ being absent in the definition of the sequences d.
In fact, within any fixed KR class $\mathcal{C}$, the nonholonomy degrees $d_{\mathcal{G}}$ associated to geometric classes $\mathcal{G}$ sitting in $\mathcal{C}$ clearly do not exceed $O_{\mathcal{C}}$, and are equal to $O_{\mathcal{C}}$ **only** when the mentioned third alternative is not applied past the first (if any) letter S in $\mathcal{G}$. That is, only when $\mathcal{G}$ is the tangential class in $\mathcal{C}$.

So within tangential classes, the KR algebras cannot be improved in the sense of lowering nilpotency orders. (Also, the germs in tangential classes are all strongly nilpotent, cf. [**19**], and thus the KR algebras have realizations already on the level of nilpotent approximations – of *certain* Goursat germs.)

QUESTION. How is it in non-tangential classes? Do there exist there Goursat germs with better nilpotent bases – with lower nilpotency orders?

This Question makes sense from dimension 6 onwards (see Ex. 1). We announce below the full answer in dimensions 6 and 7 (addressing all non-tangential classes in these dimensions; it is an improvement over [**18**] where the dimension 7 was not yet completely settled), and a partial answer in dimension 8, addressing eight out of eighteen non-tangential classes of corank–6 Goursat distributions.

THEOREM 4.
A. *In dimension* 6*, for the germs in the class* GGSG*, the nilpotency order* $O = 7$ *is minimal among all possible local nilpotent bases for them, despite the fact that the nonholonomy degree* $d = 6$ *for these germs.*

B. *In dimension* 7*, for the germs in the geometric classes in the left column:*

Geometric class $\mathcal{G}$	$d_{\mathcal{G}}$	Encompassing KR class $\mathcal{C}$	$O_{\mathcal{C}}$
GGSGG	7	$**\text{S}**$	9
GGSTG	8	$**\text{S}**$	9
GGGSG	8	$***\text{S}*$	10
GGSSG	9	$**\text{SS}*$	11
GGSGS	11	$**\text{S}*\text{S}$	12

their respective KR algebras are of minimal possible nilpotency orders.

C. *In dimension* 8*, for the germs in the non-tangential classes in the left column:*

Geometric class $\mathcal{G}$	$d_{\mathcal{G}}$	Encompassing KR class $\mathcal{C}$	$O_{\mathcal{C}}$
GGSGGG	8	$* * \mathrm{S} * * *$	11
GGSTGG	9	$* * \mathrm{S} * * *$	11
GGSTTG	10	$* * \mathrm{S} * * *$	11
GGSGSG	12	$* * \mathrm{S} * \mathrm{S} *$	17
GGSTSG	13	$* * \mathrm{S} * \mathrm{S} *$	17
GGSGST	16	$* * \mathrm{S} * \mathrm{S} *$	17
GGGSGS	15	$* * * \mathrm{S} * \mathrm{S}$	17
GGSSGS	17	$* * \mathrm{SS} * \mathrm{S}$	19

their respective KR algebras are of minimal possible nilpotency orders, too.

These first results show that belonging to a specific geometric class, as far as nilpotency orders are concerned, is not additionally advantageous. There seems to be one common nilpotent algebra covering all 'quarks' – geometric classes building one 'particle' – a KR class of Goursat germs. We do not yet know the answer for the remaining ten non-tangential classes in dimension 8. In Sec. 5 below, a proof of the part **A** is given. Proofs of parts **B** and **C**, using the local models found in [**13**], [**8**], will be presented elsewhere. They are ideologically similar, but much longer in some fragments.

COROLLARY 1. *In dimension* 6 *and* 7*, for all Goursat germs in the non-tangential geometric classes, their nilpotent approximations are not equivalent to the departure germs. The same concerns the germs in dimension* 8 *sitting in the eight non-tangential classes listed in* Thm. 4, C. *Thus those germs are not strongly nilpotent.*

(It is so because of the second property of nilpotent approximations recalled in Prop. 2.) The weak form of nilpotency – possession of a nilpotent basis – thus appears much weaker than the strong form of nilpotency of a distribution germ.

Making a point now, Question formulated after Thm. 3 specifies to *two conjectures*:[1] the strong one says that the pattern emerging from Thm. 4 is valid for all non-tangential geometric classes in all dimensions: that KR algebras are always optimal. The weak one, implied by the strong, is that the notions 'tangential' and 'strongly nilpotent' simply coincide for Goursat objects.

4. A new proof of Theorem 2. By Rem. 1, we can work with a KR pseudo-normal form $(\mathbf{X}, \mathbf{Y})$ of *any* representative D of the KR class $\mathcal{C}$. Choose for simplicity any fixed tangential (see Def. 6) germ D in $\mathcal{C}$. Then, by Obs. 1, all constants in $(\mathbf{X}, \mathbf{Y})$ are zero. In the first place, a linear-algebra formula for

[1] Having already appeared in [**19**].

the nilpotency order $O_{\mathcal{C}}$ of $L_{\mathbb{R}}(\mathbf{X},\,\mathbf{Y}) = L_{\mathbb{R}}(\mathcal{C})$ will be obtained, after which the desired answer d_r will be identified in that algebraic formula.

We are going to propose certain – very natural for KR presentations of tangential germs – weights $w(x^l)$ and $w(\partial_l) = -\,w(x^l)$, $l = 1, 2, \dots, r+2$ (the notation coincides with that of Def. 1, because these coordinates are linearly adapted for $(\mathbf{X}, \mathbf{Y})$ at 0; they are even adapted – see the previous proof in [**19**]).

When *any* weights are thus attached to both variables and versors, different terms in expansions of arbitrary vector fields are also given their weights by the classical formula (2). Our proposal is such that $\mathbf{X}$ and all terms in $\mathbf{Y}$ are of constant weight -1, which we write down as

$$w(\mathbf{X}) = w(\mathbf{Y}) = -1. \tag{3}$$

This claimed homogeneity of the polynomial (and involved) field $\mathbf{Y} = \overset{r+1}{Y}$, along with that of all fields $\overset{r}{Y}$, $\overset{r-1}{Y}, \dots, \overset{1}{Y}$, is the key ingredient of the proof.

The definition of weights is recursive from $r+2$ *backwards* to 1. At the beginning we declare $w\Big(\overset{r+1}{Y}\Big) = w(\partial_{r+2}) = -1$, hence also declare $w(x^{r+2}) = 1$. Assume now, for $1 \le m \le r$, that $w\Big(\overset{m+1}{Y}\Big) < 0$ and $w(x^{m+2}) > 0$ are already defined. The recursive definition depends on the positions of $*$ and S's in $\mathcal{C}$, determining the way of prolonging the sequence of vector fields in question.

($\bullet$) If $\overset{m+1}{Y} = \overset{m}{Y} + x^{m+2}\partial_{m+1}$ (i. e., the m-th place in $\mathcal{C}$ is $*$), then we put

$$w\Big(\overset{m}{Y}\Big) = w\Big(\overset{m+1}{Y}\Big), \qquad w(\partial_{m+1}) = w\Big(\overset{m+1}{Y}\Big) - w(x^{m+2}).$$

($\bullet\bullet$) If $\overset{m+1}{Y} = x^{m+2}\overset{m}{Y} + \partial_{m+1}$ (the m-th place in $\mathcal{C}$ is S), then we put

$$w\Big(\overset{m}{Y}\Big) = w\Big(\overset{m+1}{Y}\Big) - w(x^{m+2}), \qquad w(\partial_{m+1}) = w\Big(\overset{m+1}{Y}\Big).$$

At the last step ($m = 1$), $w\Big(\overset{2}{Y}\Big) = w(\partial_1 + x^3\partial_2) < 0$ and $w(x^3) > 0$ are assumed known, making $w(\partial_1) = w\Big(\overset{1}{Y}\Big) = w\Big(\overset{2}{Y}\Big)$ and $w(\partial_2) = w\Big(\overset{2}{Y}\Big) - w(x^3)$ known.

In the outcome, $w(x^{r+2})$, $w(x^{r+1}), \dots,$ $w(x^3)$, and $w(x^2) = -\,w(\partial_2)$, $w(x^1) = -\,w(\partial_1)$ are all defined in such a way that (3) is guaranteed. Observe that *two* variables, x^{r+2} and x^j s. t. $\mathbf{Y}(0) = \partial_j$[2] (when $\mathcal{C} = \mathrm{GG}\dots\mathrm{G}$ – model (C)

[2] The absence of constants in the field $\mathbf{Y}$ plays off, for the first time, here.

– it is just x^1), have weight 1, and all the remaining coordinates have weights larger than 1. Moreover,

LEMMA 1. *These r weights exceeding 1 are all different. Moreover still, $w(x^2)$ is the largest one of them.*

PROOF. We will prove that all versors but ∂_{r+2}, ∂_j have different weights smaller than -1. To this end, write the alternatives ($\bullet$) and ($\bullet\bullet$) in a different way as

$$(\bullet') \qquad \begin{pmatrix} w\big(\overset{m}{Y}\big) \\ w(\partial_{m+1}) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} w\big(\overset{m+1}{Y}\big) \\ w(\partial_{m+2}) \end{pmatrix}$$

and

$$(\bullet\bullet') \qquad \begin{pmatrix} w\big(\overset{m}{Y}\big) \\ w(\partial_{m+1}) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} w\big(\overset{m+1}{Y}\big) \\ w(\partial_{m+2}) \end{pmatrix}.$$

Now observe that each vector field $\overset{m}{Y}$ from among $\overset{r}{Y}$, $\overset{r-1}{Y}$, ..., $\overset{1}{Y}$, has one bare versor, say ∂_l, as a component, and, by our definitions, $w\big(\overset{m}{Y}\big) = w(\partial_l)$. Thus the components of all vectors standing in ($\bullet'$) and ($\bullet\bullet'$) *are* the weights of certain versors. Also, either of the operators

$$(4) \qquad \overleftarrow{*} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \overleftarrow{\mathrm{S}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

applied to a vector with both negative components a, b yields a new vector with one component smaller than $\min(a, b)$. Therefore, in the sequence of r linear transformations

$$(5) \quad \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} w\big(\overset{r+1}{Y}\big) \\ w(\partial_{r+2}) \end{pmatrix} \to \begin{pmatrix} w\big(\overset{r}{Y}\big) \\ w(\partial_{r+1}) \end{pmatrix} \to \cdots \to \begin{pmatrix} w\big(\overset{1}{Y}\big) \\ w(\partial_2) \end{pmatrix},$$

at each step a new weight of a versor is produced (in the upper or lower component of vectors – in function of the alternative taking effect), and that weight is *smaller*, hence different from all previously produced weights. Altogether one gets r different weights smaller than -1.
Because the code $\mathcal{C}$ starts with an $*$ (even with at least two $*$'s, as we know from Thm. 1), it is the alternative ($\bullet'$) that is used in passing from $\begin{pmatrix} w\big(\overset{2}{Y}\big) \\ w(\partial_3) \end{pmatrix}$ to $\begin{pmatrix} w\big(\overset{1}{Y}\big) \\ w(\partial_2) \end{pmatrix}$. Consequently, $w(\partial_2)$ is the lowest negative weight of a versor, and $w(x^2)$ is the highest positive weight of a coordinate. □

Consider now the small flag $\{V_j\}$ of D, $V_1 = D$, and the small growth vector $\{n_j\}$ of D at the reference point $0 \in \mathbb{R}^{r+2}$. Think about the (possibly rare) moments when the jumps of dimensions $n_{j-1} < n_j$ take place, and try to say more about those jumps. The pure products of precisely j factors from among $\{\mathbf{X}, \mathbf{Y}\}$ that fall in $V_j(0) \setminus V_{j-1}(0)$ are – by our constructions – homogeneous of weight $-j$. This is the key moment in the proof. Due to the absence of constants, they are combinations of certain bare versors ∂_l of that weight $-j$. Hence there *are* versors of weight $-j$. On the other hand, by Lem. 1, $w(\partial_l) = -j$ implies that there is just one such ∂_l. Consequently,

$$n_j = 1 + n_{j-1} \tag{6}$$

(the jumps in dimensions are only, from time to time, by 1). Specializing the general notation of Sec. 1 to the Goursat case, $d = 2$ and the integers w_i can be neatly characterized.

Indeed, let us write the dimensions simpler as $n_j = i$. Then the defining formula $\dim V_j(0) = n_j$, when remembering about (1) in which now (6) holds, takes the form

$$\dim V_{w_i}(0) = i$$

for $i = 3, 4, \ldots, r+2$. The integer w_i indicates the first time when the small flag of D at 0 attains the dimension i (because j was declared the first time of its attaining the dimension n_j):

$$V_{w_{i-1}}(0) = V_{w_i - 1}(0) \subsetneq V_{w_i}(0).$$

And these times $w_i = j$ are precisely the weights of KR coordinates used in the local description of D: $j = w(x^l)$ in the above discussion. This interpretation *includes* $w_1 = w_2 = 1$ (remember that $V_1(0)$ is two-dimensional) because such are the weights $w(x^{r+2}) = w(x^j) = 1$ that start our construction. Observe also that $w_3 = 2$ and $w_4 = 3$ independently of the germ D under consideration, and that, by Lem. 1, the highest w_{r+2} equals $w(x^2)$.

It is now easy to show that, for the permutation $k_3, \ldots, k_{r+2}$ of indices from 1 through $r+1$ save j, taken according to the growing weights $w(\cdot)$ of the KR coordinates (cf. Lem. 1),

$$V_{w_i}(0) = \big(\partial_{r+2}, \partial_j, \partial_{k_3}, \ldots, \partial_{k_i}\big) \tag{7}$$

for $i = 2, 3, \ldots, r+2$. The beginning of induction is, naturally, $V_1(0) = (\mathbf{X}, \mathbf{Y})(0) = \big(\partial_{r+2}, \partial_j\big)$. Then, assuming for certain $i - 1 \geq 2$ that

$$V_{w_{i-1}}(0) = \big(\partial_{r+2}, \partial_j, \partial_{k_3}, \ldots, \partial_{k_{i-1}}\big), \tag{8}$$

let the unique versor ∂_l,

$$\partial_l \in V_{w_i}(0) \setminus V_{w_i - 1}(0) = V_{w_i}(0) \setminus V_{w_{i-1}}(0)$$

be as in the discussion of $V_j(0) \setminus V_{j-1}(0)$ above (it is the only versor of weight $-\,w_i$, to be precise). Clearly, ∂_l is different from the versors appearing on the RHS of (8), and $V_{w_i}(0) = \big(V_{w_{i-1}}(0),\ \partial_l\big) = \big(\partial_{r+2},\ \partial_j, \dots,\ \partial_{k_{i-1}},\ \partial_l\big)$. Clearly also, $|w(\partial_l)|$ is larger than $|w(\partial_{k_3})|, \dots,\ |w(\partial_{k_{i-1}})|$, because the versors $\partial_{k_3}, \dots, \partial_{k_{i-1}}$ entered the small flag of $(\mathbf{X}, \mathbf{Y})$ at 0 earlier than ∂_l. We put, then, $k_i = l$, terminating the induction step. For $i = r+2$, (7) reads

$$V_{w_{r+2}}(0) = \big(\partial_{r+2},\ \partial_j,\ \partial_{k_3}, \dots,\ \partial_{k_{r+1}},\ \partial_2\big) = T_0\mathbb{R}^{r+2}. \tag{9}$$

That is, the pure products of precisely $w_{r+2} = w(x^2)$ factors from among $\{\mathbf{X},\ \mathbf{Y}\}$ have caused the last increment, in the small flag at 0, to the full tangent space. In particular, certain products of this length in the algebra $L_{\mathbb{R}}(\mathbf{X},\ \mathbf{Y})$ are non-zero. On the other hand, any pure product of more than $w(x^2)$ factors from among $\{\mathbf{X},\ \mathbf{Y}\}$ is homogeneous of our negative weight smaller than $w(\partial_2) = -\,w(x^2)$, hence is necessarily zero. Consequently, this algebra is nilpotent of order

$$O_{\mathcal{C}} = w(x^2).$$

Recapitulating, we know by now that if the code of $\mathcal{C}$ is $*\,*\,L_3L_4 \dots L_r$ ($L_3, \dots,\ L_r$ are taken from $\{*,\ \mathrm{S}\}$), then the nilpotency order $O_{\mathcal{C}}$ of $L_{\mathbb{R}}(\mathcal{C})$ is, in the shorthand notation (4), the *lower* component of the vector

$$\overleftarrow{*}\ \overleftarrow{*}\ \overleftarrow{L_3}\ \overleftarrow{L_4} \cdots \overleftarrow{L_r} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{10}$$

– because it is equal to $w(x^2)$ (cf. (5)).

Why does the sequence $\{\, d_j \,\}$ appear in the answer? It comes in via an induction argument based on a series of simple (in themselves) observations.

Observation 2.

$$\overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ d_1 \end{pmatrix}, \qquad \overleftarrow{*}\ \overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ d_2 \end{pmatrix}.$$

Observation 3.

$$\overleftarrow{*}\ \overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\,\overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \mathrm{id}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overleftarrow{\mathrm{S}}\ \overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\,\overleftarrow{\mathrm{S}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \mathrm{id}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Observation 4.

$$\overleftarrow{*}\ \overleftarrow{\mathrm{S}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \overleftarrow{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mathrm{id}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overleftarrow{\mathrm{S}}\ \overleftarrow{\mathrm{S}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \overleftarrow{\mathrm{S}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mathrm{id}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

At the beginning of an induction on r starting from 1 and 2, think about computing the vector (10) with short sequences of operators appearing in that formula: $\overleftarrow{*}$ and $\overleftarrow{*}\ \overleftarrow{*}$ only. Then the first two entries in $\{d_j\}$ are obtained accordingly in the lower positions; this is explicitly noted in Obs. 2.

In the induction step $(r-1,\, r) \Rightarrow r+1$, think about prolonging a string of r operators by one more operator **on the right** (i. e., to be applied directly to $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ before the other operators). If this $(r+1)$-th operator is $\overleftarrow{*}$, then, by Obs. 3 and the inductive assumption for $r-1$ and r, $w(x^2) = 2\, d_r - d_{r-1}$. If the $(r+1)$-th operator is $\overleftarrow{\mathrm{S}}$, then, by Obs. 4 and again the inductive assumption, $w(x^2) = d_r + d_{r-1}$.
So, by induction, the sequences $\{d_j\}$ do compute the nilpotency orders of KR algebras. Theorem 2 is proved. □

Remark 2. In the above proof, dealing with arbitrary tangential germ D in $\mathcal{C}$, it is explained in its course that $O_{\mathcal{C}} = w(x^2)$ = the nonholonomy degree of D. Therefore, taking in Thm. 2 the *eventual* recursive formulas for $O_{\mathcal{C}}$, one obtains at no price recursive formulas for the nonholonomy degrees of the tangential Goursat germs. And those formulas happen to coincide with Jean's formulas [**12**], recalled in the present work in Thm. 3. (Yet, let us repeat, this approach works only for the nonholonomy degrees of *tangential* Goursat germs!)
That is to say, we now re-prove a part of classical Jean's results, something like 'two-thirds' of Thm. 3, when any G cannot go after a T.

5. Proof of Theorem 4, part A. Any Goursat germ D from GGSG, living on a 6–dimensional manifold $(M,\, p)$, can be visualised (Thm. 1) as the germ at $0 \in \mathbb{R}^6$ of $(\mathbf{X},\, \mathbf{Y}) = (\partial_6,\; x^5\partial_1 + x^3x^5\partial_2 + x^4x^5\partial_3 + \partial_4 + (c + x^6)\partial_5)$ with certain $c \neq 0$ (which can be normalized to 1). A short computation in these coordinates shows that the small growth vector of D at the reference point p (visualised as 0) is $[2, 3, 4, 5, 5, 6]$. Therefore, the members V_2, V_3, V_4 of the small flag of D coincide with the relevant members of the big flag of consecutive Lie squares of D,

$$\begin{array}{ccccccccc} D = D^{(0)} & \subset & D^{(1)} & \subset & D^{(2)} & \subset & D^{(3)} & \subset & D^{(4)} = T\mathbb{R}^6 \\ \| & & \| & & \| & & \| & & \\ V_1 & & V_2 & & V_3 & & V_4 & & \end{array}$$

(they are included, and have the same ranks 3, 4, and 5, respectively). So

$$V_4 + [V_4,\, V_4] \;=\; TM. \tag{11}$$

Now suppose that D possesses in a neighbourhood of p a nilpotent basis $\mathcal{B}$ of nilpotency order 6 (i. e., all Lie products of at least 7 factors from $\mathcal{B}$ vanish). We are going to use this basis just in one computation whose outcome will eventually lead to a contradiction. Namely, computing the LHS of (11) in the basis $\mathcal{B}$, it is equal to $V_4 + [V_2,\, V_4] + [V_3,\, V_3]$ which, in turn, is equal to $V_4 + [V_2,\, V_4]$, because $[V_3,\, V_3] = [D^{(2)},\, D^{(2)}] = D^{(3)} = V_4$. Identity (11) thus assumes the form

$$V_4 + [V_2, V_4] \;=\; TM \quad \text{in a neighbourhood of } p \in M. \tag{12}$$

Let us forget now about $\mathcal{B}$, come back to the basis $(\mathbf{X}, \mathbf{Y})$ written in the chosen KR coordinates (our 'glasses' through which we see best), and compute the LHS of (12). Another short computation shows that

$$V_2 \;=\; (\,\partial_6, \quad \partial_5, \quad x^5\partial_1 + x^3x^5\partial_2 + x^4x^5\partial_3 + \partial_4\,), \tag{13}$$

$$V_3 \;=\; (\,\partial_6, \quad \partial_5, \quad \partial_4, \quad \partial_1 + x^3\partial_2 + x^4\partial_3\,),$$

$$V_4 \;=\; (\,\partial_6, \quad \partial_5, \quad \partial_4, \quad \partial_3, \quad \partial_1 + x^3\partial_2\,). \tag{14}$$

Having (13), (14) and computing the LHS of (12) at p, that is, at $0 \in \mathbb{R}^6$, *in these terms* we obtain but $(\partial_1, \partial_3, \partial_4, \partial_5, \partial_6)$. This contradicts (12). Hence $\mathcal{B}$ does not exist. Part **A** of Theorem 4 is proved. □

REMARK 3. Note that, in (12), the summand $[V_2, V_4]$ alone is *not* defined invariantly; it depends on the basis being used in the computation. Yet the (local) module of vector fields $V_2 + V_4 + [V_2, V_4] = V_4 + [V_2, V_4]$ already is, so that the LHS of (12) has a geometric sense regardless of a basis in use.

Note also that the nilpotent algebra $L_{\mathbb{R}}(**\mathrm{S}*)$ of nilpotency order 7 that underlies the part **A** of Thm. 4, has dimension 8. From the Lie-algebraic point of view it might be interesting to develop an approach allowing to compute the dimensions of all KR algebras $L_{\mathbb{R}}(\mathcal{C})$.

References

1. Agrachev A.A., Gamkrelidze R.V., Sarychev A.V., *Local invariants of smooth control systems*, Acta Appl. Math., **14** (1989), 191–237.
2. ———, Gauthier J–P., *On subanalyticity of Carnot–Carathéodory distances*, Ann. Inst. H. Poincaré – AN, **18** (2001), 359–382.
3. Bellaïche A., *The tangent space in sub-Riemannian geometry*, in: Bellaïche A., Risler J–J. (eds.), *Sub-Riemannian Geometry,* Birkhäuser, Basel – Berlin, 1996, 1–78.
4. Bianchini R.M., Stefani G., *Graded approximations and controllability along a trajectory*, SIAM J. Control Optimiz., **28** (1990), 903–924.
5. Bressan A., *Nilpotent approximations and optimal trajectories*, in: *Analysis of Controlled Dynamical Systems,* Progr. Systems Control Theory, **8**, Birkhäuser, Basel – Berlin, 1991, 103–117.
6. Bryant R., Hsu L., *Rigidity of integral curves of rank* 2 *distributions*, Invent. Math., **114** (1993), 435–461.
7. Cartan E., *Sur l'équivalence absolue de certains systèmes d'équations différentielles et sur certaines familles de courbes*, Bull. Soc. Math. France, **XLII** (1914), 12–48.
8. Cheaito M., Mormul P., *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.

9. Hermes H., *Nilpotent approximations of control systems and distributions*, SIAM J. Control Optimiz., **24** (1986), 731–736.
10. ______, *Nilpotent and high-order approximations of vector field systems*, SIAM Review, **33** (1991), 238–264.
11. ______, Lundell A., Sullivan D., *Nilpotent bases for distributions and control systems*, J. Diff. Eqns, **55** (1984), 385–400.
12. Jean F., *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.
13. Kumpera A., Ruiz C., *Sur l'équivalence locale des systèmes de Pfaff en drapeau*, in: Gherardelli F. (ed.), *Monge–Ampère Equations and Related Topics,* Inst. Alta Math. F. Severi, Rome, 1982, 201–248.
14. Lafferriere G., Sussmann H.J., *A differential geometric approach to motion planning*, in: Canny F.J., Li Z. (eds.), Nonholonomic Motion Planning, Kluwer, Dordrecht – London, 1993, 235–270.
15. Montgomery R., Zhitomirskii M., *Geometric approach to Goursat flags*, Ann. Inst. H. Poincaré – AN, **18** (2001), 459–493.
16. Mormul P., *Goursat flags: classification of codimension–one singularities*, J. Dynam. Control Syst., **6** (2000), 311–330.
17. ______, *Simple codimension–two singularities of Goursat flags I: one flag's member in singular position*. Preprint 01-01 (39) Inst. of Math., Warsaw University, available at `http://www.mimuw.edu.pl/english/research/reports/tr-imat/`
18. ______, *Minimal nilpotent local bases for Goursat distributions in dimensions not exceeding* 7, preprint in: *Singularity Theory Seminar,* Warsaw Univ. of Technology, Vol. 6 (2001), 44–56.
19. ______, *Goursat distributions not strongly nilpotent in dimensions not exceeding seven*, in: Zinober A., Owens D. (eds.), *Nonlinear and Adaptive Control* NCN4 2001, Lecture Notes in Control and Information Sciences, **281**, Springer, Berlin – New York, 2003, 249–261.
20. ______, *Superfixed positions in the geometry of Goursat flags*, Univ. Iagel. Acta Math., **40** (2002), 183–196.
21. ______, *Multi-dimensional Cartan prolongation and special k–flags,* in: Hironaka H., Janeczko S., Łojasiewicz S. (eds.), *Geometric Singularity Theory,* Banach Center Publications, **65**, Inst. of Math., Polish Acad. of Sci., Warszawa, 2004, 157–178.
22. Pasillas–Lépine W., Respondek W., *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.

Received December 3, 2002

Warsaw University
Institute of Mathematics
Banacha 2
02-097 Warsaw
Poland
e-mail: `mormul@mimuw.edu.pl`