A SINGLE NILPOTENT APPROXIMATION FOR A FAMILY OF NONEQUIVALENT DISTRIBUTIONS

PIOTR MORMUL

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

Abstract

Goursat distributions – subbundles in the tangent bundles to manifolds having the tower of consecutive Lie squares growing very slowly in ranks, always only by one – possess, from corank 8 onwards, numerical moduli of the local classification up to diffeomorphisms of base manifolds. (Up to corank 7 that classification is finite.) The number of such moduli grows with the corank, i.e., with the length of a tower. Yet there are no other (more involved, e.g., functional) moduli of the local classification. A natural question, asked by Agrachev in the year 2000, is whether those numerical moduli descend to the level of nilpotent approximations: whether they are resistant enough to survive the passing to the nilpotent level. A surprising *negative* result in this direction was obtained in [21]; it dealt with a Goursat modulus appearing in codimension *three*, in corank (or length) 8. In the present work, we show that it is likewise – a modulus disappears on the nilpotent level – for the first Goursat modulus found, [19], in codimension *two* (i.e., in the smallest possible codimension, classification in codimension one, [17], being discrete).

1. Introduction

We want to use throughout the present work several notions related to the nilpotent approximations of geometric distributions. They can be viewed as a far reaching generalization of the notion of *linearization* of a single vector field. The linearization of a vector field v at a point, although simplifying geometry a big deal, retains some basic local properties of v. Likewise, the nilpotent (or graded, or homogeneous) approximations simplify enormously the geometry of distributions without losing the most essential (mainly nonholonomic) traits of them.

In a coordinate language, first steps bringing those objects to existence were made in [8] (cf. also [9]) and decisive ones in [3], [2], [7], with substantial later simplifications proposed in [6]. Local coordinates in which nilpotent approximations can be viewed (something like night glasses in nonholonomic geometry) have a separate history of their own. They are useful in sub-Riemannian geometry, too, from the basic (nonholonomic) Ball-Box Theorem onwards – see [6], as well as the entire book *Sub-Riemannian Geometry* containing that contribution. Paraphrasing Sussmann's contribution to that book, a cornucopia of various possible sets of adapted (or privileged) coordinates is astonishing. We hope that the papers [21] and the present give some evidence to this statement. On the other hand, one can try to approach *all* sets of privileged coordinates at once – in an appropriate high level abstract algebra language, as it is proposed in [10].

On the coordinate-free side, there exists a key (if not published) short text [1] and a contribution [4], followed later by a referential work [5]. Yet another approach, highly algebraized and not widely known, is presented in [11].

The adjective 'nilpotent' is related to the fact that, whatever local generators of a given distribution D, the simplified (or trimmed) generators of the approximation around a point, say, p generate a

nilpotent Lie algebra, over the reals, of precisely known nilpotency order stemming from the geometry of D in the vicinity of p – equal to the nonholonomy degree of D at p. (Note that in [4] and [5] offered are alternative explanation and interpretation of nilpotent approximations, making use of the concept of *nonholonomic tangent spaces*. Our approach, however, which sharpens an effective general procedure in [6], seems suited best for concrete computations.) In the present paper, we will systematically use the abbreviation NA for 'nilpotent approximation'.

An ideal environment for these directions of research seems to be the world of Goursat distributions. On the one side, they are very tight and possessing clear polynomial presentations. Among others, they are free of *functional* moduli. On the other side, are abundant with *numeric* moduli, found not earlier than in the end of XX century. In fact, the very question of descending of moduli of Goursat objects to the nilpotent level was asked by Agrachev in the year 2000 and has since sparked an entire line of work, including [18], [20], and [21].

Out of these three works, only the last one raises directly the issue of hypothetical surviving of moduli after descending to the simpler level of NAs. In fact, it deals with the only geometric class in corank (or – the same thing for Goursat – length) 8 concealing a modulus. That class is labelled GGSGSGSG in the 'GST' language described in detail in [19] and recapitulated in Section 3 of the present paper.[1] The class dealt with in [21] may thus be seen as chosen somehow optimally. And the message of that contribution is that, despite strong opposite expectations, the modulus sitting in GGSGSGSG does not descend to the NA level. Quite a lot of surprising algebra stands behind such a result and the technique needed for it has turned out to be heavy.

---

[1]In smaller lengths, the local classification is discrete, as explained in a series of papers, the last among them being [16].

The class discussed in [21] has, however, one serious deficiency. Its codimension is three – it displays three letters S in its code, cf., for instance, Proposition 2 in [19]. While the moduli of Goursat distributions are not confined to codimensions three and higher. They exist also in codimension two; this is explained in [19]. (And they do not exist in codimension one, [17].) Those coarser codimension-two moduli occur, it is true, at the expense of increasingly high flag's length, at least 9 in the occurrence. In fact, in that threshold length, there exists just one codimension-two geometric class

$$\text{GGGSGSGGG,} \tag{1}$$

which hosts a modulus of the local classification. The present work addresses precisely this class. Germs sitting in (1) are visualised, in certain proper coordinates originating from [13] and [19], under the form (4) displayed later on. There are two real parameters, b and c, in that *pre*-normal form. They conceal a *single* modulus – the quantity $c - 7b - \frac{5}{3}b^2$ derived and discussed in [19].

The first and main aim of the paper is to show, in our Theorem 2, that the modulus in (1) does not descend to the nilpotent approximations as well, exactly as it has been the case in [21]. Practically, all NAs of the objects in the family (4) will be computed to the very end and will turn out not to depend on b or c. They will be just one thing.

An important corollary – and the second aim of the paper – of the technics of proving Theorem 2 is that a vast standing conjecture concerning nilpotent properties of Goursat distributions, put forward more than 10 years ago in [18], is being confirmed in the class (1). See Theorem 3 in Subsection 4.3 for details.

2. Nilpotent Approximation of a Distribution at a Point

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

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

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, ... 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 $n = \dim M$. We truncate that vector after the first appearance of n in it. The length d_{NH} of the truncated vector is called the *nonholonomy degree* of D at p.

In the theory that we only sketch here (cf. [8], [2], [7], [6]; this list is not complete) important are the *weights* w_i related to the small flag at a point: $w_1 = \cdots = w_d = 1$, $w_{d+1} = \cdots = w_{n_2} = 2$ (no value 2 among them when $n_2 = n_1(= d)$), and generally,

$$w_{n_j+1} = \cdots = w_{n_{j+1}} = j + 1, \tag{2}$$

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

Definition. For a completely nonholonomic distribution D on M, coordinates $z_1, z_2, \ldots, z_n$ around $p \in M$ (centered at p) 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 $V_{d_{\mathrm{NH}}}(p) = (\partial_1, \ldots, \partial_n) = T_p M$. (Throughout we use the shorthand notation $\partial_j = \partial / \partial z_j$. Here and in the sequel, we also skip writing 'span' before a set of vector fields' generators.)

For such linearly adapted coordinates one defines, as in [6], their *weights* $w(z_i) = w_i$, $i = 1, \ldots, n$.

On the other hand, given a completely non-holonomic D, every smooth function f on M has, at any point $p \in M$, its *nonholonomic order* nord(f) with respect to D (for simplicity of notation, we skip writing its dependence on p). By definition, it is the minimal order of a nonholonomic derivative of f that is non-zero at p.[2]

It follows directly from the above definitions that, for linearly adapted coordinates, their nonholonomic orders **do not exceed** their weights.

Definition. Linearly adapted coordinates $z_1, \ldots, z_n$ are *adapted* when nord(z_i) *equals* $w(z_i)$ for $i = 1, \ldots, n$.

It is rather laborious to show, but adapted coordinates always exist, see in this respect, to name just some, [2], [7], [6]. Moreover, they are by far not unique; there remains plenty of freedom behind the requirement being imposed on linearly adapted coordinates that the nonholonomic orders be maximal possible.

In adapted coordinates, it is reasonable to attach quasi-homogeneous weights also to monomial vector fields (this definition goes back to the work [22] in the theory of differential operators; for geometric distributions, see in this respect [2], p. 215). Namely,

$$w(z_{i_1} \cdots z_{i_k} \partial_j) = w(z_{i_1}) + \cdots + w(z_{i_k}) - w(z_j). \tag{3}$$

The gist of the concept of adaptedness resides in the following:

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

---

[2] $+\infty$ is not excluded.

(Superscripts mean the weights defined by (3).) We denote by $\widehat{X}$ the lowest ('(–1)-jet', 'nilpotent') summand $X^{(-1)}$. That is, $\widehat{X} = X^{(-1)}$.

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

Definition [6, Definition 5.15]. ($\star$) The distribution $\widehat{D} = (\widehat{X_1}, \ldots, \widehat{X_d})$, defined on M locally around p, is called the *nilpotent approximation* of D at p.

($\star\star$) It is proved in Proposition 5.20 in [6] that $\widehat{D}$ is well-defined, independently of the adapted coordinates that are being used. More precisely, the NA of D at p is to be understood as the *equivalence* class of distributions $(\widehat{X_1}, \ldots, \widehat{X_d})$ constructed above by means of *all* possible sets of adapted coordinates, with the equivalence relation described in the proof of that proposition in [6]: One conjugates the quasi-homogeneous (–1)-terms of vector fields by means of the *trimmed* diffeomorphisms, having, for $i = 1, 2, \ldots, n$, in their i-th component only the terms of quasi-homogeneous weight *exactly w_i*.

($\star\star\star$) Alternatively, this class $\widehat{D}$ of weighted degree –1 distribution germs is called the (–1)-jet of D at p.

So, 'the NA' and 'the (–1)-jet' are just synonyms in the category of completely nonholonomic distributions. But – important – in any concrete situation, the NA of D at p is being visualised in *one* given set of adapted coordinates as a distribution germ at $0 \in \mathbb{R}^n$. It is then, strictly speaking, just one of the representatives of $\widehat{D}$, watched not on M, but rather in a good chart around $0 \in \mathbb{R}^n$.

The basic property of the nilpotent approximation is

Proposition 2. *At the reference point* p, *the small flag of* $\widehat{D}$ *coincides with that of* D *at* p. *Hence* $\widehat{D}$ *has at* p *the same small growth vector as* D *at* p (*and, in particular, the same nonholonomy degree* d_{NH}, *too*).

This property is crucial. It shows that, in the occurrence, much simpler objects retain some basic geometric characteristics of the initial objects. One can fairly deeply trim a distribution germ without losing essential information! This opportunity can only support one's hope for the survival of moduli in nilpotent approximations.

Attention. There is, however, one warning pointing in the opposite direction: unlike the small growth vector at the reference point, the *big growth vector* of a distribution D at a point (the sequence of linear dimensions at a point of the members of the *big flag* of D – the tower of modules of vector fields – consecutive Lie squares $D \subset [D, D] \subset [[D, D], [D, D]] \subset \cdots$) is, generally speaking, **not** preserved under passing to the (-1)-jet of D at that point. See in this respect, Theorem 3 in Subsection 4.3 and also pp. 258-259 in [18]. Building on the ulterior machinery, the proof of Theorem 3 exploits that poor performance of the big flag also for the NAs of germs in the class (1).

Note in parentheses that *all* germ's geometric properties are preserved under the NA functor only at so-called strongly nilpotent points of the underlying manifold (cf. [18]). Consequently, there will be no such points in the class (1).

3. Kumpera-Ruiz Normal Forms for Goursat Distributions

In what follows, we deal uniquely 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.

A distribution $D \subset TM$ is Goursat when it is rank-2 and the big growth vector of D is, at every point $p \in M$, just $[2, 3, 4, \ldots, n-1, n]$, where $n = \dim M \geq 4$. The number $n - 2 \geq 2$ is called the *length* of the [big] flag of D. (Sometimes the assumption $rk, D = 2$ is being dropped in this definition, like, for instance, in [14] and [19]. Both variants locally lead to the same theory, because there always splits off an integrable corank-2 sub-distribution in D. In fact, that splitting object is the Cauchy-characteristic sub-distribution of D.)

There exists a very basic partition of Goursat germs of corank n into disjoint *geometric classes* encoded by words of length $n - 2$ over the alphabet G, S, T, with two first letters always G and such that never a T goes directly after a G. Their construction has its roots in the pioneering work [12] of Jean, in which the author used a trigonometric, not polynomial, presentation of Goursat objects. That construction, with some natural subsequent improvements, has been reproduced in detail in Subsection 1.1 of [19].[3]

In dimension 4, there is but one class GG, in dimension 5 – only GGG and GGS, 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 [14], 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 up one KR class $**S*$. In dimension 7, the geometric classes GGSGG, GGSTG, and GGSTT build $**S**$, etc. (This appellation is a bit misleading, for the authors of [13] went, in medium size dimensions 6 and 7, well beyond the bare KR classes.)

---

[3]It should be noted that, in the meantime, Montgomery and Zhitomirskii have developed, in a big contribution [15], a yet different notation for the same geometric classes, or, better, Jean's strata. They skip in the words the two leftmost letters G (the words thus become of length $n - 4$) and translate G $\rightarrow$ R, S $\rightarrow$ V. The letters T remain unchanged in the encoding words. The new name for the strata is 'RVT classes'.

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

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

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

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 have

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 $D = (\mathbf{X}, \mathbf{Y})$ specified above, with certain constants in the writing of the field $\mathbf{Y}$ corresponding to G's past the first S in $\mathcal{C}$ and subject to • and ••.*

This will be the main reference point in the remaining of the paper.

4. Nilpotent Approximation in the Geometric Class GGGSGSGGG

Our aim is to prove the following:

Theorem 2. *The modulus of the local classification residing in the geometric class* GGGSGSGGG *disappears on the level of nilpotent approximations. That is, the nilpotent approximations of the members of this class are all equivalent.*

Our proof will focus on finding certain adapted coordinates, in order to get hold of the NAs of the Goursat germs in the class (4). The obtained NAs will – for a time – be illegible. A special care in the concluding part of the proof will be taken to make them legible and even – in suitable super-adapted coordinates depending on a germ – all of them identical.

4.1. Proof of Theorem 2, main part

To get started, let us write the Kumpera-Ruiz visualisations for the germs in (1) in the form emerging from the Section 2 in [19]. We mean taking $k = 4$ there, momentarily normalizing the constants c^7 and c^9 to 1 (easy), and then writing for simplicity b instead of c^{10} and c instead of c^{11}. In fact, in these K-R coordinates,

$$D = (\mathbf{X}, \mathbf{Y}) = \left(\partial_{11},\ x^8 x^6 \partial_1 + x^8 x^6 x^3 \partial_2 + x^8 x^6 x^4 \partial_3 + x^8 x^6 x^5 \partial_4 + x^8 \partial_5 \right.$$

$$\left. +\ x^8 X^7 \partial_6 + \partial_7 + X^9 \partial_8 + X^{10} \partial_9 + X^{11} \partial_{10} \right), \qquad (4)$$

as the germ at $0 \in \mathbb{R}^{11}$, where $X^7 = 1 + x^7$, $X^9 = 1 + x^9$, $X^{10} = b + x^{10}$, $X^{11} = c + x^{11}$, $b \in \mathbb{R}$, and $c \in \mathbb{R}$. These preliminary two-parameter normal forms are not exact local models; the parameter b is redundant and could be eliminated. As to the numerical invariant residing in the class (1), in terms of (4), it is $c - 7b - \frac{5}{3} b^2$ — see Corollary 3 in [19]. (With $b = 0$, it would be, naturally, just c.)

It follows from the results of the benchmark paper [12] (or can be computed, with some effort, by hand) that the small growth vector of (4) at $0 \in \mathbb{R}^{11}$ is, regardless of the values of b and c, $[2, 3, 4, 5, 6, 7_2, 8_2, 9_4, 10_4, 11]$, and hence the weights $w_1, w_2, \dots, w_{11}$ are

$$1, 1, 2, 3, 4, 5, 6, 8, 10, 14, 18. \tag{5}$$

Because of the constants, the KR variables used in (4) are not yet linearly adapted. After some Lie bracket computations with the vector generators displayed in (4), one quickly improves the coordinates x^i to linearly adapted coordinates

$$x^{11}, x^7, x^{10} - cx^7, x^9 - bx^7, x^8 - x^7, x^5, x^6 - x^5, x^1, x^4, x^3, x^2. \tag{6}$$

It is well known – see Section 2, and compare also [8], [9] – that linearly adapted coordinates are in general rather coarse and far from fitted for computing the NAs. Their upgrading to [certain, just certain] adapted coordinates is critical. The first seven adapted coordinates z_i $(i = 1, 2, \dots, 7)$ are prompted by (4) and (6) directly. Proceeding much like in [20], one nearly comes across

$$z_1 = x^{11}, \quad z_2 = x^7, \quad z_3 = x^{10} - cx^7, \quad z_4 = x^9 - bx^7 - \frac{c}{2}(x^7)^2,$$

$$z_5 = x^8 - x^7 - \frac{b}{2}(x^7)^2 - \frac{c}{6}(x^7)^3, \quad z_6 = x^5 - \frac{1}{2}(x^7)^2 - \frac{b}{6}(x^7)^3 - \frac{c}{24}(x^7)^4,$$

$$z_7 = x^6 - x^5 - \frac{1}{3}(x^7)^3 - \frac{b}{8}(x^7)^4 - \frac{c}{30}(x^7)^5.$$

It is important that in these coordinates: $\mathbf{X}(z_1) = 1$, $\mathbf{X}(z_i) = 0$ for $i = 2, 3, \dots, 7$, and

$$\mathbf{Y}(z_1) = 0, \quad \mathbf{Y}(z_2) = 1, \quad \mathbf{Y}(z_3) = z_1, \quad \mathbf{Y}(z_4) = z_3,$$

$$\mathbf{Y}(z_5) = z_4, \quad \mathbf{Y}(z_6) = z_5, \quad \mathbf{Y}(z_7) = z_2 z_5. \tag{7}$$

♦ It is less straightforward to improve x^1 to an adapted z_8. (The bare linearly adapted x^1 does not do, because

$$\mathbf{Y}(x^1) = x^8 x^6 = \left(z_5 + \frac{c}{6} z_2^3 + \frac{b}{2} z_2^2 + z_2^3\right)$$

$$\left(z_7 + \frac{c}{30} z_2^5 + \frac{b}{8} z_2^4 + \frac{1}{3} z_2^3 + z_6 + \frac{c}{24} z_2^4 + \frac{b}{6} z_2^3 + \frac{1}{2} z_2^2\right), \tag{8}$$

and, after expanding the RHS out, there appears a fan of unnecessary terms of nord ≤ 6.) Fortunately, though, refining x^1 to $x^1 - \frac{1}{2} z_2^2 z_6$ eliminates *two* nasty terms $z_2 z_6$ and $\frac{1}{2} z_2^2 z_5$ of nonholonomic order 6 on the RHS of (8) at a time, and now

$$\mathbf{Y}\left(x^1 - \frac{1}{2} z_2^2 z_6\right) = (\text{terms of nord} \geq 8) + z_2 z_7 + \frac{b}{2} z_2^2 z_6 + \left(\frac{1}{3} + \frac{b}{6}\right) z_2^3 z_5$$

$$+ \text{ a combination of } \left(z_2^7, z_2^6, z_2^5, z_2^4, z_2^3\right).$$

By adding the relevant polynomial of degree 8 in the single variable z_2, it is quick to eliminate the 3rd, 4th, ... up to the 7th powers of z_2 in the expansion above (we omit for clarity the exact expressions for the coefficients, which are not used further). By consequence, the variable

$$z_8 = x^1 - \frac{1}{2} z_2^2 z_6 + \text{a due combination of } \left(z_2^8, z_2^7, z_2^6, z_2^5, z_2^4\right), \tag{9}$$

is linearly adapted of nord [nonholonomic order] at least 8, that is, adapted (and its nord is exactly 8). In fact, $\mathbf{X}(z_8) = 0$, and

$$\mathbf{Y}(z_8) = z_2 z_7 + \frac{b}{2} z_2^2 z_6 + \left(\frac{1}{3} + \frac{b}{6}\right) z_2^3 z_5 + (\text{terms of nord} \geq 8). \tag{10}$$

Note that in this simplification, apart from the terms of quasi-homogeneous degrees smaller than 7, also the $z_2^7 (= z_2^{w_8 - 1})$ term has been eliminated. This was done on purpose, to make the further parts of

the proof more transparent. (Similar simplifications will be done with the remaining coordinates x^4, x^3, x^2 with no additional comment, so as to have no pure powers of z_2 whatsoever in a preliminary presentation of the NA.)

♦ Upgrading of x^4. We know that

$$\mathbf{Y}(x^4) = (x^8 x^6) x^5 = \mathbf{Y}(x^1) x^5 = \mathbf{Y}(x^1)\left(z_6 + \frac{c}{24} z_2^4 + \frac{b}{6} z_2^3 + \frac{1}{2} z_2^2 \right),$$

where $\mathbf{Y}(x^1)$ is already expressed in the z variables – see (8) above. After expanding the RHS, there appear

● terms of nord ≥ 10;

● terms of nord 9, excepting z_2^9, which are going to be important for the NA;

● two particular mixed terms of nord 8: $z_2^3 z_6 + \dfrac{1}{4} z_2^4 z_5$;

● the pure powers z_2^9, z_2^8, z_2^7, z_2^6, z_2^5 with precisely known coefficients.

We get rid of the couple of mixed terms again by a *single* correction $-\dfrac{1}{4} z_2^4 z_6$, after which there only remains to find a proper polynomial of degree 10 in z_2 doing the job of eliminating the listed powers of z_2. This time the coefficients in that polynomial *are* important in what follows because the variable in question, x^4, enters the expression for $\mathbf{Y}(x^3)$ discussed shortly later. (It has not been so with the variable x^1, earlier upgraded to z_8.) After the due computation, the variable

$$z_9 = x^4 - \frac{1}{4} z_2^4 z_6 - \frac{1}{24} z_2^6 - d_7 z_2^7 - d_8 z_2^8 - d_9 z_2^9 - d_{10} z_2^{10}, \qquad (11)$$

where

- $d_7 = \dfrac{1}{42} + \dfrac{1}{24}b,$

- $d_8 = \dfrac{29}{1152}b + \dfrac{1}{96}c + \dfrac{1}{72}b^2,$

- $d_9 = \dfrac{7}{1080}c + \dfrac{1}{144}bc + \dfrac{23}{2592}b^2 + \dfrac{1}{648}b^3,$

- $d_{10} = \dfrac{79}{17280}bc + \dfrac{5}{5760}c^2 + \dfrac{1}{960}b^3 + \dfrac{5}{4320}b^2c,$

is linearly adapted of nord $\geq$ 10, hence of nord equal to $10 = w_9$ and adapted as such. As to its nonholonomic derivatives in the $\mathbf{X}$ and $\mathbf{Y}$ directions, $\mathbf{X}(z_9) = 0$, and

$$\mathbf{Y}(z_9) = \frac{1}{2}z_2^3 z_7 + \left(\frac{1}{3} + \frac{5b}{6}\right)z_2^4 z_6 + \left(\frac{1}{6} + \frac{b}{6}\right)z_2^5 z_5 + (\text{terms of nord} \geq 10).$$

$$(12)$$

♦Upgrading of x^3. The pattern of computation is clear. Adjoining the formula (11) to the pool of relations defining our 'smoother' variables z's out of the more 'coarse' x's, and knowing now the $\partial/\partial z_j$ – components of $\mathbf{Y}$ for $j < 10$, one perturbes x^3 so as to guarantee that the nonholonomic order of a perturbed function be equal to $w_{10} = 14$. One embarks, with (11) at hand, from the initial equality

$$\mathbf{Y}(x^3) = (x^8 x^6)x^4 = \mathbf{Y}(x^1)x^4$$

$$= \mathbf{Y}(x^1)\left(z_9 + \frac{1}{4}z_2^4 z_6 + \frac{1}{24}z_2^6 + d_7 z_2^7 + d_8 z_2^8 + d_9 z_2^9 + d_{10}z_2^{10}\right).$$

Having (8), after expanding out the RHS here, there emerge

- terms of nord ≥ 14;

- terms of nord 13, excepting z_2^{13}, which are important for the NA;

- two particular 'nasty' terms of nord 12: $\dfrac{1}{6}\, z_2^7 z_6 + \dfrac{1}{48}\, z_2^8 z_5$;

- the powers z_2^{13}, z_2^{12}, z_2^{11}, z_2^{10}, z_2^9 with precisely known coefficients.

It is possible to kill here the z_2^{13} term and all terms of nord ≤ 12, including the stubborn binomial which, again, will be killed by a *single* correction term. In fact, after a due computation, the variable

$$z_{10} = x^3 - \frac{1}{48}\, z_2^8 z_6 - \frac{1}{480}\, z_2^{10} - d_{11} z_2^{11} - d_{12} z_2^{12} - d_{13} z_2^{13} - d_{14} z_2^{14}, \quad (13)$$

where

$$d_{11} = \frac{1}{792} + \frac{5}{3168}\, b + \frac{1}{22}\, d_7,$$

$$d_{12} = \frac{7}{6912}\, b + \frac{1}{2304}\, c + \frac{1}{3456}\, b^2 + \frac{1}{36}\, d_7 + \frac{1}{24}\, d_8 + \frac{5}{144}\, bd_7,$$

$$d_{13} = \frac{1}{3510}\, c + \frac{1}{4992}\, b^2 + \frac{7}{44928}\, bc + \frac{1}{26}\, d_9 + \frac{1}{39}\, d_8 + \frac{7}{312}\, bd_7 + \frac{1}{104}\, cd_7$$

$$\qquad + \frac{5}{156}\, bd_8 + \frac{1}{156}\, b^2 d_7,$$

$$d_{14} = \frac{1}{8960}\, bc + \frac{1}{48384}\, c^2 + \frac{7}{336}\, bd_8 + \frac{5}{168}\, bd_9 + \frac{2}{315}\, cd_7 + \frac{1}{112}\, cd_8$$

$$\qquad + \frac{1}{224}\, b^2 d_7 + \frac{1}{168}\, b^2 d_8 + \frac{1}{288}\, bcd_7 + \frac{1}{42}\, d_9 + \frac{1}{28}\, d_{10},$$

is linearly adapted of nord ≥ 14, hence of nord equal to $14 = w_{10}$. As to the nonholonomic derivatives of z_{10}, $\mathbf{X}(z_{10}) = 0$, while

$$\mathbf{Y}(z_{10}) = \frac{1}{2}\, z_2^3 z_9 + \frac{1}{24}\, z_2^7 z_7 + \left(\frac{1}{12} + \frac{b}{8} + d_7 \right) z_2^8 z_6 + \left(\frac{1}{72} + \frac{b}{144} + \frac{1}{2}\, d_7 \right) z_2^9 z_5$$

$$\qquad + (\text{terms of nord} \geq 14). \qquad\qquad (14)$$

♦ Upgrading of the last variable x^2. The overall pattern of computation does not change. One adjoins the formula (13) to the pool of relations defining new variables z's in terms of the old x's, and now knows the $\partial / \partial z_j$–components of the fields $\mathbf{X}$ and $\mathbf{Y}$ for all $j < 11$. Then tries to perturb x^2 so as to make sure that the nonholonomic order of an outcome function be at least $w_{11} = 18$ (hence equal to 18), and that its $\mathbf{Y}$-derivative be z_2^{17}-free. Unsurprisingly, the key to it lies in the equality

$$\mathbf{Y}(x^2) = x^8 x^6 x^3 = (z_5 + ...)(z_7 + ...)(z_{10} + ...),$$

where the last factor on the RHS features the old coordinate x^3 now written, by means of (13), in the z_j variables, $j < 11$. (Clearly, z_{11} is absent yet, it is to be defined soon.) After opening the brackets on the RHS here, there appear

• terms of nord ≥ 18;

• terms of nord 17, excepting z_2^{17} – the only important for the NA in sight;

• two particular mixed terms of nord 16: $\dfrac{1}{80} z_2^{11} z_6 + \dfrac{1}{960} z_2^{12} z_5$;

• the powers z_2^{17}, z_2^{16}, z_2^{15}, z_2^{14}, z_2^{13} with all coefficients effective.

The distinguished binomial will disappear after subtracting a single (again!) term from x^2, while the listed pure powers of z_2 are, invariably, easy to get rid of. After a count of coefficients, the variable

$$z_{11} = x^2 - \frac{1}{960} z_2^{12} z_6 - \frac{1}{13440} z_2^{14} - d_{15} z_2^{15} - d_{16} z_2^{16} - d_{17} z_2^{17} - d_{18} z_2^{18},$$

$$(15)$$

with d_{15} through d_{18} being certain effective constants depending on b and c, whose exact expressions do not matter for the course of proof, is

adapted. Indeed, its **X**-derivative is zero, while its **Y**-derivative features only terms of nonholonomic orders 17 and higher, so that $\mathrm{nord}\,(z_{11}) = 18 = w_{11}$ as needed. In fact,

$$\mathbf{Y}(z_{11}) = \frac{1}{2}\, z_2^3 z_{10} + \frac{1}{480}\, z_2^{11} z_7 + \left(\frac{1}{144} + \frac{7b}{720} + d_{11} \right) z_2^{12} z_6$$

$$+ \left(\frac{1}{1440} + \frac{b}{2880} + \frac{1}{2}\, d_{11} \right) z_2^{13} z_5 + (\text{terms of nord} \geq 18). \qquad (16)$$

We have at last produced a full set of privileged coordinates $z_1, z_2, \ldots, z_{11}$. The first seven of them have been ascertained in the nick of time, unlike the remaining four, defined by (9), (11), (13), and (15).

We are now in a position to preliminarily summarize the long quest for NAs in the class (1). We group together the '(−1)-jet' (or nilpotent) components of the generators of the distribution, whose coefficients are given in the formulas (7), (10), (12), (14), and (16). That is, in the expansions for vector field generators **X** and **Y**, we only retain the terms of quasi-homogeneous weight −1 (cf. (3)), and leave out all terms having non-negative weights. This boils down, for each $i = 8, 9, 10, 11$, to leaving out, in the relevant expansion for $\mathbf{Y}(z_i)$, all terms of nonholonomic orders $\geq w_i$. In the outcome $\widehat{\mathbf{X}} = \mathbf{X} = \partial_1$, while $\widehat{\mathbf{Y}} =$

$$\partial_2 + z_1 \partial_3 + z_3 \partial_4 + z_4 \partial_5 + z_5 \partial_6 + z_2 z_5 \partial_7 + \left(z_2 z_7 + \frac{b}{2}\, z_2^2 z_6 + \left(\frac{1}{3} + \frac{b}{6} \right) z_2^3 z_5 \right) \partial_8$$

$$+ \left(\frac{1}{2}\, z_2^3 z_7 + B_6 z_2^4 z_6 + B_5 z_2^5 z_5 \right) \partial_9 + \left(\frac{1}{2}\, z_2^3 z_9 + \frac{1}{24}\, z_2^7 z_7 + C_6 z_2^8 z_6 + C_5 z_2^9 z_5 \right) \partial_{10}$$

$$+ \left(\frac{1}{2}\, z_2^3 z_{10} + \frac{1}{480}\, z_2^{11} z_7 + D_6 z_2^{12} z_6 + D_5 z_2^{13} z_5 \right) \partial_{11}. \qquad (17)$$

What is more, an inspection of the simplifications done so far shows that the aggregated coefficients introduced in (17),

- $B_5 = \dfrac{1}{6} + \dfrac{b}{6},\quad B_6 = \dfrac{1}{3} + \dfrac{5b}{6},$

- $C_5 = \dfrac{1}{72} + \dfrac{b}{144} + \dfrac{1}{2}\,d_7,\quad C_6 = \dfrac{1}{12} + \dfrac{b}{8} + d_7,$

- $D_5 = \dfrac{1}{1440} + \dfrac{b}{2880} + \dfrac{1}{2}\,d_{11},\quad D_6 = \dfrac{1}{144} + \dfrac{7b}{720} + d_{11},$

depend only on the technical parameter b, and not on the essential parameter c. The main difficulties in the proof of Theorem 2 are already overcome.

4.2. Proof of Theorem 2, final refinements of the NAs

The idea is to gradually replace the adapted z coordinates by *more* adapted ones, with the aim to strip the components of $\widehat{\mathbf{Y}}$ of the variable z_{10}, then of z_9, then z_7, then z_6. And that last simplification, by miraculous if simple identities, will do! After it, there will be no free parameter whatsoever in $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$.

One might note, however, that, in (17), the highest variable with b showing up in its coefficient is z_6: There is no parameter next to $z_{10}, \ldots, z_7$. This notwithstanding, one *has* to start from eliminating z_{10} in $\widehat{\mathbf{Y}}(z_{11})$, then z_9 in $\widehat{\mathbf{Y}}(z_{11})$ and $\widehat{\mathbf{Y}}(z_{10})$, and so on downwards in indices. Otherwise, it would be more complicated to eventually get rid of some appearances of b.

The improvements will always be quasi-homogeneous, so that (i) one will remain within the family of [sets of] adapted coordinates, and (ii) there will be no need to look for the quasi-homogeneous 'hat' part of any improving diffeomorphism, which comes to fore in, mentioned already in the present note, Proposition 5.20 in [6].[4]

---

[4]In general, there is no way to know beforehand, if such a far-reaching final reduction is doable. Even within the Goursat world only – cf. the question of Agrachev from the year 2000 in Section 1 – it is still open. In the unimodal class (1), in the end of the day, it is; in the unimodal class GGSGSGSG discussed in [21], it is as well.

To avoid cumbersome notations, we keep *the same letters* for the improved coordinates.

So the first target is the $z_2^3 z_{10}$ term in the ∂_{11} component in (17). To kill it, one takes a new adapted $z_{11} := z_{11} - \frac{1}{8} z_2^4 z_{10}$. The $z_2^3 z_{10}$ term disappears in the ∂_{11} component, but a $z_2^7 z_9$ term comes instead in. And there is a $z_2^3 z_9$ term in the ∂_{10} component. One kills the two latter by taking newer adapted $z_{11} := z_{11} + \frac{1}{128} z_2^8 z_9$ and $z_{10} := z_{10} - \frac{1}{8} z_2^4 z_9$. At the expense of creating two 'replacement' terms with z_7 instead of z_9 (the indices always decrease by the triangular property): $z_2^{11} z_7 \partial_{11}$ and $z_2^7 z_7 \partial_{10}$.

After that second reduction, the highest variable present in $\widehat{\mathbf{Y}}$ is z_7; it appears in the $\partial_{11}, \partial_{10}, \partial_9, \partial_8$ components. Then in the third reduction, by taking newer coordinates:

- $z_{11} := z_{11} - \dfrac{1}{15360} z_2^{12} z_7,$

- $z_{10} := z_{10} + \dfrac{1}{384} z_2^8 z_7,$

- $z_9 := z_9 - \dfrac{1}{8} z_2^4 z_7,$

- $z_8 := z_8 - \dfrac{1}{2} z_2^2 z_7,$

prompted by the outcome of the preceding reduction, one completely eliminates the terms with z_7. After this phase, the highest variable still present in $\widehat{\mathbf{Y}}$ is z_6. In fact,

$$\widehat{\mathbf{Y}} = \partial_2 + z_1\partial_3 + z_3\partial_4 + z_4\partial_5 + z_5\partial_6 + z_2z_5\partial_7 + \left(\frac{b}{2}\,z_2^2z_6 + \frac{b-1}{6}\,z_2^3z_5\right)\partial_8$$

$$+ \left(B_6z_2^4z_6 + \left(B_5 - \frac{1}{8}\right)z_2^5z_5\right)\partial_9$$

$$+ \left(\left(C_6 - \frac{1}{8}\,B_6\right)z_2^8z_6 + \left(C_5 - \frac{1}{8}\,B_5 + \frac{1}{384}\right)z_2^9z_5\right)\partial_{10}$$

$$+ \left(\left(D_6 - \frac{1}{8}\,C_6 + \frac{1}{128}\,B_6\right)z_2^{12}z_6 + \left(D_5 - \frac{1}{8}\,C_5 + \frac{1}{128}\,B_5 - \frac{1}{15360}\right)z_2^{13}z_5\right)\partial_{11}.$$

$$(18)$$

Eventually, one gets completely rid of z_6 by adding

- an appropriate multiple of $z_2^{13}z_6$ to z_{11};

- an appropriate multiple of $z_2^9z_6$ to z_{10};

- an appropriate multiple of $z_2^5z_6$ to z_9;

- an appropriate multiple of $z_2^3z_6$ to z_8.

(The coefficients in these correction terms are all implied directly by (18).)

The now highest variable is z_5; it appears in $\widehat{\mathbf{Y}}$ in the components ∂_{11} through ∂_6; the components ∂_7 and ∂_6 have, besides, been unchanged in this last simplification. The only components of $\widehat{\mathbf{Y}}$ that one has yet to ascertain are ∂_{11} down to ∂_8.

It is quick with the coefficient standing next to ∂_8:

$$\left(-\frac{b}{6} + \frac{b-1}{6}\right)z_2^3z_5 = -\frac{1}{6}\,z_2^3z_5. \tag{19}$$

And nearly as quick with the ∂_9 coefficient:

$$\left(\frac{1}{6} + \frac{b}{6} - \frac{1}{8} - \frac{1}{5}\left(\frac{1}{3} + \frac{5b}{6}\right)\right)z_2^5z_5 = -\frac{1}{40}\,z_2^5z_5. \tag{20}$$

It is longer (and surprising) with the ∂_{10} coefficient:

$$\left(-\frac{1}{9}\left(\frac{1}{12}+\frac{b}{8}+\frac{1}{42}+\frac{b}{24}\right)+\frac{1}{72}\left(\frac{1}{3}+\frac{5b}{6}\right)+\frac{1}{72}+\frac{b}{144}+\frac{1}{84}+\frac{b}{48}\right.$$

$$\left.-\frac{1}{8}\left(\frac{1}{6}+\frac{b}{6}\right)+\frac{1}{384}\right)z_2^9 z_5 = \frac{1}{3456}z_2^9 z_5. \qquad (21)$$

And it is nothing short of astounding with the ∂_{11} coefficient:

$$\left(-\frac{1}{13}\left(\frac{1}{144}+\frac{7b}{720}+d_{11}\right)+\frac{1}{104}\left(\frac{1}{12}+\frac{b}{8}+\frac{1}{42}+\frac{b}{24}\right)-\frac{1}{1664}\left(\frac{1}{3}+\frac{5b}{6}\right)\right.$$

$$+\frac{1}{1440}+\frac{b}{2880}+\frac{1}{2}d_{11}-\frac{1}{8}\left(\frac{1}{72}+\frac{b}{144}+\frac{1}{84}+\frac{b}{48}\right)+\frac{1}{128}\left(\frac{1}{6}+\frac{b}{6}\right)$$

$$\left.-\frac{1}{15360}\right)z_2^{13} z_5 = -\frac{1}{199680}z_2^{13} z_5. \qquad (22)$$

Summarizing now, we have passed from the presentation (18) featuring parameter(s) to $\widehat{\mathbf{X}} = \partial_1$, and

$$\widehat{\mathbf{Y}} = \partial_2 + z_1\partial_3 + z_3\partial_4 + z_4\partial_5 + z_5\partial_6 + z_2 z_5\partial_7 - \frac{1}{6}z_2^3 z_5\partial_8 - \frac{1}{40}z_2^5 z_5\partial_9$$

$$+\frac{1}{3456}z_2^9 z_5\partial_{10} - \frac{1}{199680}z_2^{13} z_5\partial_{11}. \qquad (23)$$

All the NAs under consideration have turned out to be mutually equivalent as distribution germs. The parameters b and c concealing the modulus of the local classification of distributions D have disappeared (the latter, c, still in Subsection 4.1). The proof of Theorem 2 is now complete.

4.3. The class GGGSGSGGG is not strongly nilpotent

Theorem 3. *At every point p in the class* GGGSGSGGG, *the distribution D is not locally equivalent to* [*its NA*] $\widehat{D}$ *computed at p. That is to say, every point in this class is not strongly nilpotent in the sense of* [18] *and* [20].

A short proof will be ideologically similar to the one in [21]. We work locally around a point p and assume $\widehat{D}$ at p to be already in the form (23). We will simplify this presentation still further, reducing the number of actively used (always adapted!) coordinates to just *two*: z_1 and z_2.

In order to leave the variable z_5 out in (23), one refines each of the actual coordinates z_{11} through z_6 by adding to it a multiple of the product of z_5 and an appropriate power of z_2. The price for this is a z_4 coordinate in each of the components ∂_5 through ∂_{11}, and a by one higher power of the coordinate z_2 standing next to it in that component (remember that $w_4 = w_5 - 1$).

Then one likewise gets rid of z_4, replacing it everywhere by z_3 and raising the exponents of the relevant near-by z_2's by one ($w_3 = w_4 - 1$).

In the end, all single z_3's are being replaced by single z_1's, together with one more rise by one in the exponents of the near-by z_2's ($w_1 = w_3 - 1$).

In the result of these corrections, the first nilpotent generator $\widehat{\mathbf{X}} = \partial_1$ is not changed, while a straightforward count of coefficients yields the second generator in the form:

$$\widehat{\mathbf{Y}} = \partial_2 + z_1\partial_3 - z_1 z_2\partial_4 + \frac{1}{2}\,z_1 z_2^2\partial_5 - \frac{1}{6}\,z_1 z_2^3\partial_6 - \frac{1}{24}\,z_1 z_2^4\partial_7 + \frac{1}{720}\,z_1 z_2^6\partial_8$$

$$+ \frac{1}{13440}\,z_1 z_2^8\partial_9 - \frac{1}{4561920}\,z_1 z_2^{12}\partial_{10} + \frac{1}{670924800}\,z_1 z_2^{16}\partial_{11}.$$

After a standard rescaling of variables z_4 through z_{11} ($z_4 := -z_4$, $z_5 := 2z_5$, $z_6 := -6z_6$, etc.) it is possible to go one (aesthetically only) better: $\widehat{\mathbf{X}} = \mathbf{X} = \partial_1$, and

$$\widehat{\mathbf{Y}} = \partial_2 + z_1\partial_3 + z_1 z_2\partial_4 + z_1 z_2^2\partial_5 + z_1 z_2^3\partial_6 + z_1 z_2^4\partial_7 + z_1 z_2^6\partial_8 + z_1 z_2^8\partial_9$$

$$+ z_1 z_2^{12}\partial_{10} + z_1 z_2^{16}\partial_{11}. \tag{24}$$

The absence of strong nilpotency is now behind the corner. Indeed, two of consequences of (24) read: every Lie monomial over $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ of length (= # of factors) at least two

1. has no ∂_1, ∂_2 components;

2. is a vector-valued polynomial in z_1, z_2.

Hence the *Lie product* of any two such monomials is zero. This implies that the big flag of ($\widehat{\mathbf{X}}$, $\widehat{\mathbf{Y}}$) coincides *identically* with the small one! In particular, the big growth vector of ($\widehat{\mathbf{X}}$, $\widehat{\mathbf{Y}}$) at $0 \in \mathbb{R}^{11}$ is equal to its small one at $0 \in \mathbb{R}^{11}$. That is, to the sgrv of D at the considered point p: $[2, 3, 4, 5, 6, 7_2, 8_2, 9_4, 10_4, 11]$ (Proposition 2). $\widehat{D}$ is, therefore, very far from being a Goursat germ, and p assuredly is not a strongly nilpotent point. Theorem 3 is proved.

Remark. The final refinements of adapted variables which lead to description (24) may be compared with the simplifications in [21], in Subsections 5.4 and 5.5 there. (Reiterating, the latter concerned the absence of a modulus in the NAs of the class GGSGSGSG, the only geometric class in length 8 in which a Goursat modulus has been identified to-date; see Remark 3 in [16] for a justification of that modulus. Reiterating also, the length 8 was minimal possible.) In that earlier study, there occurred, in the simplifications, instances of so-called 'nilpotent flatness'. A la limite, that strange flatness was then responsible for the disappearing of the modulus in passing to the NA.

A similar kind of nilpotent flatness has played off, on several occasions, in the present proof, too. Each initial refining of linearly adapted coordinates: firstly x^1, then x^4, then x^3, and eventually x^2 was greatly facilitated by the matching of certain pair of coefficients, allowing for a single relevant correction term. See in this respect (9), (11), (13), and (15). Therefore, the proof in the class (1) appears to be nearly as

surprising as the (pioneering) one in [21]. Apparently, a deeper general machinery is at work, here and elsewhere. One now starts to believe in a general negative answer to the Agrachev 2000 question.

This also supports a longstanding conjecture, most clearly formulated on page 260 in [18], that, for the Goursat flags, only *tangential* points are strongly nilpotent. (The stratum (1) comprises uniquely non-tangential points.)

5. Extending the Issue to Classes $\mathrm{G}^k\mathrm{SGSG}^3$, $k > 3$

The class (1) is the first in a series of codimension two *unimodal* classes $\mathcal{C}_k = \mathrm{G}^k\mathrm{SGSG}^3$, $k \geq 3$ discussed in Theorem 2 in [19]. (As for the class $\mathrm{G}^2\mathrm{SGSG}^3$, it is still *simple* in Arnold's sense, cf. Subsection 4.2 in [19].) Kumpera-Ruiz visualizations similar to (4) exist for each $k \geq 3$, always with two real parameters in them concealing a single invariant, very much like for $k = 3$.

To answer Agrachev's question in the classes $\mathcal{C}_k$ for $k > 3$, one should

1. improve Kumpera-Ruiz coordinates to linearly adapted ones;

2. further improve linearly adapted to adapted, this time working with $k + 8$ variables and their weights being now: 1, 1, 2, 3, 4, 5, 6, 8, 10, $10 + 4l$, $l = 1, 2, \ldots, k - 1$ (and still having the parameters in many a place);

3. having the (–1)-jet of the distribution expressed in the adapted coordinates from point 2., try to 'squeeze the parameters out' of that jet.

The point 3. critically depends on the *result* of point 2. At present, we do not know whether there is enough of that strange computational flatness, discovered in $\mathcal{C}_3$, in each given class $\mathcal{C}_k$, $k > 3$. That is, if, in the production process of adapted coordinates, there occur sufficiently

many perfect matchings in pairs of coefficients, allowing for a containment of parameters, as it has been the case in $\mathcal{C}_3$ in the process of improving the coordinates x^1, x^4, x^3, x^2.

The question is purely algorithmic, but the confirming computations have been done (in the present work) only for $k = 3$. It is not yet done for $k \geq 4$. Most likely, one additional simplification still on the level of Kumpera-Ruiz coordinates (Section 3 in [19]) could come in handy.

References

[1] A. A. Agrachev, Canonical Nilpotentization of Control Systems, Preprint, 1997 (3 pages).

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

[3] A. A. Agrachev and A. V. Sarychev, Filtrations of a Lie algebra of vector fields and nilpotent approximation of control systems, Dokl. Akad. Nauk SSSR 295 (1987); English transl. in Soviet Math. Dokl. 36 (1988), 104-108.

[4] A. A. Agrachev and A. Marigo, Nonholonomic tangent spaces: Intrinsic construction and rigid dimensions, Electronic Res. Announc. Amer. Math. Soc. 9 (2003), 111-120.

[5] A. A. Agrachev and A. Marigo, Rigid Carnot algebras: A classification, J. Dynam. Control Syst. 11 (2005), 449-494.

[6] A. Bellaïche, The tangent space in sub-Riemannian geometry, in: A. Bellaïche and J.-J. Risler (Eds.), Sub-Riemannian Geometry, Birkhäuser, 1996, 1-78.

[7] R. M. Bianchini and G. Stefani, Graded approximations and controllability along a trajectory, SIAM J. Control Optimiz. 28 (1990), 903-924.

[8] H. Hermes, Nilpotent approximations of control systems and distributions, SIAM J. Control Optimiz. 24 (1986), 731-736.

[9] H. Hermes, Nilpotent and high-order approximations of vector field systems, SIAM Review 33 (1991), 238-264.

[10] G. Sklyar and S. Yu. Ignatovich, Description of all privileged coordinates in the homogeneous approximation problem for nonlinear control systems, C. R. Acad. Sci. Paris, Sér. I 344 (2007), 109-114.

[11] G. Sklyar and S. Yu. Ignatovich, Fliess series, a generalization of Ree's theorem, and an algebraic approach to the homogeneous approximation problem, Int. J. Control 81 (2008), 369-378.

[12] F. Jean, The car with N trailers: Characterization of the singular configurations, ESAIM: Control, Optimization and Calculus of Variations 1 (1996), 241-266 (electronic).

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

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

[15]   R. Montgomery and M. Zhitomirskii, Points and Curves in the Monster Tower, Memoirs Amer. Math. Soc. 956, AMS, Providence, 2010.

[16]   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, Collection Travaux en cours 62, Hermann, 2000, 89-119.

[17]   P. Mormul, Goursat flags: Classification of codimension-one singularities, J. Dynam. Control Syst. 6 (2000), 311-330.

[18]   P. Mormul, Goursat distributions not strongly nilpotent in dimensions not exceeding seven, in: A. Zinober and D. Owens (Eds.), Nonlinear and Adaptive Control, Lecture Notes in Control and Inform. Sci. 281, Springer, 2003, 249-261.

[19]   P. Mormul, Real moduli in local classification of Goursat flags, Hokkaido Math. J. 34 (2005), 1-35.

[20]   P. Mormul, Kumpera-Ruiz algebras in Goursat flags are optimal in small lengths, J. Math. Sciences 126 (2005), 1614-1629.

[21]   P. Mormul, Do moduli of Goursat distributions appear on the level of nilpotent approximations?, in: J.-P. Brasselet and M. A. S. Ruas (Eds.), Real and Complex Singularities (Trends in Mathematics), Birkhäuser, 2006, 229-246.

[22]   L. P. Rothschild and E. M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (1976), 247-320.

∎