IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Desingularization of implicit analytic differential equations

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2006 J. Phys. A: Math. Gen. 39 10975

(http://iopscience.iop.org/0305-4470/39/35/003)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.106
The article was downloaded on 03/06/2010 at 04:48

Please note that terms and conditions apply.

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 10975–11001 doi:10.1088/0305-4470/39/35/003

Desingularization of implicit analytic differential equations

Hernán Cendra[1] **and María Etchechoury**[2,3]

[1] Universidad Nacional del Sur, Av. Alem 1253, 8000 Bahía Blanca and CONICET, Argentina
[2] Laboratorio de Electrónica Industrial, Control e Instrumentación, Facultad de Ingeniería, Universidad Nacional de La Plata, La Plata, Argentina
[3] Departamento de Matemática, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, CC 172, 1900 La Plata, Argentina

E-mail: hcendra@uns.edu.ar and marila@mate.unlp.edu.ar

Received 4 January 2006, in final form 17 July 2006
Published 11 August 2006
Online at stacks.iop.org/JPhysA/39/10975

Abstract
The question of finding solutions to a given *implicit differential equation* (IDE) is an important one, in part because it appears very naturally in several problems in physics, engineering and many other fields. In this work, we show how to reduce a given *analytic IDE* to an *analytic IDE of locally constant rank*. This can be done by using some fundamental results on subanalytic subsets and desingularization of closed subanalytic subsets. An example from nonholonomic mechanics is studied using these methods.

PACS numbers: 02.30.Hq, 02.60.−x, 45.20.Jj
Mathematics Subject Classification: 14P15, 32B20, 65L80, 70H45, 70F25

1. Introduction

1.1. Physical motivation

Implicit differential equations (IDE) $\phi(x, \dot{x}) = 0$ are very common in science, the case of an ODE $\dot{x} = f(x)$ being the simplest particular case. Euler–Lagrange equations for a given Lagrangian [1–3], Lagrange–D'Alembert equations for a nonholonomic system and their reduced versions Lagrange–D'Alembert–Poincaré equations [4–10] are some of the examples from mechanics. We are not going to give a list of relevant references of IDE appearing in several fields. Instead, we are going to describe next an example of IDE which is important for both classical and quantum mechanics, as physical motivation.

It is shown in [11] that several questions in the Dirac–Bergman theory of constraints, in quantum mechanics, can be reduced to solving an IDE of the type

$$\mathrm{i}_{\dot{x}}\omega(x) = \alpha(x), \tag{1.1}$$

where $\omega \in \Omega^2(M)$ is a closed 2-form and $\alpha \in \Omega^1(M)$ is a closed 1-form on a Banach manifold M. This kind of geometric formulation of the Dirac–Bergman theory also has a meaning in other fields, like classical mechanics, which is more important for the present paper. Let Q be the configuration space for a mechanical system and assume for simplicity that Q is a vector space. Let $L : TQ \to \mathbb{R}$ be a given Lagrangian, degenerate or not. It can be easily shown that Euler–Lagrange equations can be rewritten as follows:

$$\mathrm{i}_{\dot{x}}\omega(x) = \mathrm{d}h(x), \tag{1.2}$$

where $x \in Q \times Q \times Q$, say $x = (q, v, p), \omega \in \Omega^2(Q \times Q \times Q), \omega = \mathrm{d}q \wedge \mathrm{d}p$, $h = pv - L(q, v)$. Here the contraction $\langle p, v\rangle$ is denoted simply by pv. It is clear that (1.2) is of the form (1.1).

To solve an IDE like (1.1) an algorithm, called the *Gotay–Nester algorithm*, is introduced in [11]. The presymplectic form ω plays an important role in describing this algorithm and in making it easy to apply to examples, specially in infinite-dimensional cases. However, some basic aspects of the algorithm are not necessarily related to the presymplectic structure and can be applied to more general IDE of the type treated in the present paper, namely

$$a(x)\dot{x} = f(x), \tag{1.3}$$

defined at the beginning of section 2. Let us describe those basic aspects for a system like (1.1), assuming for simplicity that M is a vector space of dimension n. Let $x(t)$ be a given solution to (1.1), then, for each t, the *linear algebraic system*

$$\mathrm{i}_v\omega(x) = \alpha(x), \tag{1.4}$$

has at least one solution, namely, $v(t) = \dot{x}(t)$. This implies that, for each t, $x(t)$ must belong to the subset

$$M_1 = \{x \in M : \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM\}.$$

In [11], this subset is described by an equation in terms of the presymplectic form ω, but we do not need it here to describe the basic aspects of the algorithm, as we have said before. Assume, as in [11], that M_1 is a submanifold of M. Since $x(t) \in M_1$ for all t, we must have that $\dot{x}(t) \in T_{x(t)}M_1$ for all t. This implies that, for each t, $x(t)$ must belong to the subset

$$M_2 = \{x \in M_1 : \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM_1\}.$$

We can continue in a similar way and define M_k recursively as follows:

$$M_{k+1} = \{x \in M_k : \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM_k\},$$

under the assumption that M_k is a submanifold for $k = 1, 2, \ldots$. Since M is finite dimensional the sequence $M \supseteq M_1 \supseteq \cdots \supseteq M_q$ must stabilize, that is $M_q = M_{q+p}$, for $p \in \mathbb{N}$. Under the assumption that the map

$$\omega^\flat : T_xM_q \to T_x^*M,$$

given by $\omega^\flat(x)(v) = \omega(x)(v, \,)$, has locally constant rank for x in the final constraint manifold M_q, existence of local solutions to (1.1) is guaranteed. In fact, M_q is given, in local coordinates $(x_1, \ldots, x_n)$ in M centred at a given point $x_0 \in M_q$, by equations $x_{c+1} = 0, \ldots, x_n = 0$, where c is the dimension of M_q. Using standard arguments we can prove that, after a permutation of indices if necessary, the local solutions of the system can be described as follows. Choose functions $x_1(t), \ldots, x_r(t)$ where $r = \dim\ker\omega(x)^\flat|T_xM_q$ and then solve (1.1) uniquely for $x_{r+1}(t), \ldots, x_c(t)$.

The system (1.1) restricted to M_q, as described above, is an example of what we call in this paper an IDE of *locally constant rank*. It is clear that the basic ideas described above can also be used to solve a system like (1.3). More precisely, we define

$$M_1 = \{x \in M : a(x)\dot{x} = f(x) \text{ has at least one solution } v \in T_xM\},$$

and then we define recursively M_k in a similar way as we did before for equation (1.1). If the linear map $a(x)$ has locally constant rank on the final constraint manifold M_q we can solve locally equation (1.3) in a similar way as we did with equation (1.1). However, in this paper we want to study those cases in which the assumption that each M_k is a submanifold made at each step k in the previous algorithm may not be satisfied. For this purpose, we must choose a new algorithm, *which makes sense only for real analytic data*. The description of this new algorithm is given in section 2, but here we shall give an approximate idea of it. Under the assumption that a and f are real analytic, in the first step of the new algorithm we obtain a closed analytic subset, that is, a subset defined by analytic equations, $M_0 \subset M$, which is essentially like the subset M_1 described above, but now we do not need to assume that it is a submanifold. Instead, we use a *desingularization* $\pi_0 : M^1 \to M$ which, by definition, is a proper analytic map whose image is M_0 while M^1 is an analytic manifold having the same dimension as M_0. Now we pull back the system to M^1 *in an appropriate way* using π_0 to obtain a system *of the same kind* in M^1. We repeat the process and this way we obtain the new algorithm to solve system (1.3). If we apply the new algorithm to solve (1.1) in finite dimensions and with analytic data, under the assumption that M_k is a submanifold for each k, we obtain essentially the same results obtained by applying the Gotay–Nester algorithm.

1.2. Related works and technical background

General IDE of the type $\phi(x, \dot{x}) = 0$ can be easily reduced to those of the type (1.3), as explained in section 2, at the cost of adding new variables. This is the strategy adopted in the present work. Also the basic algorithm to solve IDE like (1.1) or (1.3) explained above can be applied directly, in principle, to an IDE of the type $\phi(x, \dot{x}) = 0$ in an obvious way. In fact, we can formally define $M_1 = \{x \in M : \phi(x, v) = 0$, has at least a solution $v \in T_x M\}$ and $M_{k+1} = \{x \in M_k : \phi(x, v) = 0$, has at least a solution $v \in T_x M_k\}$ and so on.

However, the assumption that each M_k is a submanifold seems to be too restrictive in the general case of an IDE of the type $\phi(x, \dot{x}) = 0$. This difficulty can be partly overcome by working in a neighbourhood of a point, rather than globally. Then there is the possibility of applying, somehow, the subimmersion theorem to show that M_k is a submanifold in a neighbourhood of the chosen point. Ideas closely related to this program have been implemented successfully in several papers, where the possibility of applying the subimmersion theorem is assumed whenever it is needed. See for instance [12, 13] and references therein. In those works those systems to which this kind of method can be applied are called *differential algebraic equations* (DAE). To the best of our knowledge, the singular cases where the subimmersion theorem cannot be applied have not been systematically and fully studied, locally or globally, in the existing literature.

There are specific studies for certain classes of systems. For instance, an interesting treatment of IDE given by complex polynomial relations has been realized in [14] where an algorithm using complex algebraic geometry and its implementation using computer algebra systems is described.

In the present paper, we choose to work with the class of IDE with real analytic data. This leads us immediately to the realm of semianalytic and subanalytic sets, which are fundamental objects for the present paper. A subset X of a real analytic manifold M is *semianalytic* if each $a \in M$ has a neighbourhood $U \subseteq M$ such that $U \cap X$ is a finite union of subsets of U defined by a finite collection of equalities $f(x) = 0$ and inequalities $g(x) > 0$, where f and g are real analytic functions. If in the previous definition *real analytic* is replaced by *real algebraic* one obtains the definition of *semialgebraic set*. The Tarski–Seidenberg theorem says that a semialgebraic subset of $\mathbb{R}^n$ is projected via the natural projection $\mathbb{R}^n \to \mathbb{R}^{n-1} \times \{0\}$ onto a semialgebraic subset of $\mathbb{R}^{n-1} \times \{0\}$. This important theorem does not hold true if *semialgebraic* is replaced by *semianalytic*. This has led to defining the notion of *subanalytic subset*. A subset X of an analytic manifold M is *subanalytic* if for each point $a \in M$ there is a neighbourhood U of a such that $U \cap X$ is the projection of a relatively compact semianalytic set. A theorem, analogous to the Tarski–Seidenberg theorem, holds true for subanalytic sets. This makes the theory of subanalytic subsets less rigid than the theory of semianalytic sets. For instance, the desingularization theorem used in our algorithm as described in section 2 is proved within the theory of subanalytic sets. In fact, the proof of our main theorems in section 5 relies on that theory. Of special importance for us are the subanalytic subsets of dimension 1. This is because they are essentially curves which will be the possible solutions to an IDE with analytic data. A useful result is that any subanalytic subset of dimension 1 is semianalytic. Our main reference for the theory of subanalytic subsets is [15]. See more references below in this introduction.

As another important class of examples related to the present work we mention that control systems in the category of subanalytic sets have been recently studied in [16], where desingularization techniques have been used. In the present work we also work in the category of subanalytic sets and we also use desingularization methods, but from the point of view of IDE, which in a sense is dual to the point of view of control theory. In fact, a control system is, roughly, a vector field depending on a parameter or, equivalently, a family of vector fields or, more generally, a family of local vector fields. On the other hand, as we will see, an IDE gives, after desingularizing it, also a family of vector fields, but defined implicitly. Finally, we remark that in some interesting examples in mechanics a global desingularization is not needed, and a blowing-up centred at certain submanifolds is enough, after all, a desingularization can be built as a composition of a finite sequence of local blowing-ups, see [15], p 30. See [17] for an interesting example in stability theory in mechanics.

As we have said in the previous paragraph, desingularization of closed analytic subsets plays an important role in the present work, although we are not going to study any desingularization technique here. Rather, what we want to show is that the problem of solving a given analytic IDE can be decomposed into two problems: (1) desingularize certain singular analytic submanifolds that may appear on applying the algorithm and (2) solve an analytic IDE of *locally constant rank*. Each problem has its own difficulties. Our main result shows how to reduce a given *real analytic IDE* to a *real analytic IDE of locally constant rank*, defined in section 2, which is considered the simplest case in this paper. Working in the analytic category is justified because IDE representing several important examples from mechanics, control theory and other fields, as we have mentioned before, are given by real analytic functions. An important point is proving that the desingularized system is *equivalent* to the given system. This is a technical point described in section 5, using several results from the theory of subanalytic sets. A key point is the definition of an appropriate class of curves, called *as-curve*, which, roughly speaking, include all the analytic curves whose graph is semianalytic.

Some references on desingularization and subanalytic subsets are the following. The fundamental theorem of Hironaka [18], whose proof was simplified and computationally implemented in subsequent works [19, 20], is an example of a desingularization procedure, in this case desingularization of certain algebraic varieties. Another fundamental theorem on desingularization was proven by Bierstone and Milman [21], which actually includes Hironaka's theorem. See also [22]. In the present work we are going to use some general results from the theory of subanalytic sets [15]. The only main result on desingularization that we use is the theorem of Hironaka on desingularization of closed subanalytic subsets [23]. In fact, the theorem on desingularization of closed analytic subsets, theorem 5.1 of Bierstone and Milman in [15], is enough for our purposes.

In section 2, we explain some basic facts about IDE. In section 3, we describe our algorithm. In section 5, we prove our main results. In section 4, we give a concrete example from nonholonomic mechanics, namely a symmetric ball rolling without sliding or spinning. The final desingularized manifold is shown to be $S^2 \times S^1$ on which the equation governing the motion is a simple ODE, whose solutions can be precisely described.

2. Implicit differential equations

2.1. Basic notation

Let M be a given manifold of dimension n and F a vector space of dimension m. Let $a : TM \to F$ be a smooth map such that $a(x, \dot{x}) \equiv a(x)\dot{x}$ is linear in $\dot{x}$. Let $f : M \to F$ be a given smooth map. We will consider IDE of the type

$$a(x)\dot{x} = f(x). \tag{2.1}$$

By introducing the trivial vector bundle $M \times F$ we can think of a as representing a vector bundle map

$$a : TM \to M \times F$$

and of f as being a section of $M \times F$. Then, for each $x \in M$, $a(x)$ is a linear map depending smoothly on x from the tangent space $T_x M$ into the fibre (x, F) of the trivial bundle.

More generally, we may consider a general vector bundle with base M, say $\pi : F \to M$, and an IDE like (2.1) where now $a : TM \to F$ is a vector bundle map and f is a section of F. This kind of generalization is important to describe a sufficiently wide class of IDE. However, in the present paper we shall describe only the trivial bundle case, for simplicity, and also because it already contains the essential facts. The general case can be treated in an essentially similar way. In this paper, the manifold M is called the *domain* and the space, or more generally, vector bundle F, is called the *range* of the IDE.

Given an IDE of the type (2.1) one has immediately a *linear algebraic system* (LAS) for each $x \in M$, depending smoothly on x, where the unknown is the vector $\dot{x}$, based at x, for each $x \in M$. We will call it the *LAS associated* with the given IDE.

2.2. IDE of locally constant rank

Recall that the rank of a linear map $A : E \to G$, denoted as rank A, is the dimension of its image, dim Im A. In finite dimensions we may choose a basis in E and also in G and denote by $[A]$ the matrix representing A with respect to those basis. Then $\operatorname{rank} A = \operatorname{rank}[A]$, where $\operatorname{rank}[A]$ is the maximum number of linearly independent columns of A. For given $g \in G$ we will call $\langle \operatorname{Im} A, g\rangle$ the linear space generated by Im A and g. Also, we will call $[A, g]$ the matrix whose first columns are the columns of $[A]$ and the last column is the coordinate expression of g in the basis chosen in G. By definition $\operatorname{rank}[A, g]$ is the maximum number of linearly independent columns of $[A, g]$. We have obviously $\operatorname{rank}[A, g] = \dim\langle \operatorname{Im} A, g\rangle$.

Assume that the LAS associated with (2.1) has solution $\dot{x}$ for each $x \in M$. Then we may think of (2.1) as defining an affine distribution, generally singular, on M. If, in addition, $\operatorname{rank} a(x) = \operatorname{rank}[a(x), f(x)]$ is locally constant, that is, it is constant on each connected component of M, then the IDE is called an *IDE of locally constant rank*. This is equivalent to saying that the corresponding affine distribution has *constant rank* on each connected component of M. For instance, if $m = n$ and $a(x)$ is invertible for all $x \in M$ then (2.1) is equivalent to an ODE,

$$\dot{x} = a(x)^{-1} f(x), \tag{2.2}$$

and the rank of the affine distribution is 0 in this case.

The case of an IDE of locally constant rank is the simplest case in our context and our main result, in section 5, shows that a given *analytic* IDE can be reduced to an IDE of locally constant rank in an obvious way.

2.3. Reduction of a general IDE to an IDE of the type (2.1)

It is easy to see that, from a general point of view, IDE of the type

$$\phi(x, \dot{x}) = 0,$$

where the fibre-preserving map $\phi : TM \to F$ may be nonlinear in $\dot{x}$, are not more general than (2.1).

In fact, let us assume first, for simplicity, that M is an open subset of a finite-dimensional vector space E. An IDE of the type

$$\phi(x, \dot{x}) = 0$$

can be rewritten in the form (2.1) with domain $M \times E$ and range $E \times F$ as follows:

$$\dot{x} = u \qquad 0 = \phi(x, u),$$

which has the form (2.1) with

$$a(x, u) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \qquad f(x, u) = \begin{bmatrix} u \\ \phi(x, u) \end{bmatrix}.$$

The case of a general manifold M and a fibre-preserving map $\phi : TM \to F$, where F is a vector bundle with base M, can also be reduced to the form (2.1) by an essentially similar procedure, using a geometric construction involving pull-backs of bundles. We will not need to deal with this more general type of IDE.

Remark

(i) We must remark that working with IDE written in the form (2.1) is an important ingredient of our algorithm. This is in part because this form is preserved and the space F remains the *same* (or the vector bundle *is replaced* by a pull-back vector bundle) at each stage of the algorithm. This simplifies matters as will become evident later.

(ii) The system (2.1) can be written equivalently as follows:

$$a(x)\dot{x} = \dot{t} f(x) \qquad \dot{t} = 1,$$

where $t = t(s)$. The first equation of this system can be written equivalently as follows:

$$b(y)\dot{y} = 0 \tag{2.3}$$

where $y = (x, t), b(y) \in L(T(M \times \mathbb{R}), F), b(y) = [a(x), -f(x)]$.

Of course (2.3) is an IDE whose associated LAS has a solution for each y, but it does not seem that questions like reachability for (2.1) could be easily reduced to easily solvable corresponding questions for (2.3). In other words, this kind of transformation of the system does not necessarily really simplifies the problems related to a given IDE. On the other hand, systems like (2.3) are interesting by themselves and are related to Pfaffian systems [24].

2.4. Some notation and operations with IDE

It will be convenient to denote sometimes (a, f) the IDE (2.1), from now on. Let (a, f) be a given IDE, say

$$a : TM \to F,\ f : M \to F.$$

Let $N \subseteq M$ be a given submanifold. The *restriction* $(a, f)|N$, also written as $(a|N, f|N)$, of (a, f) to N is defined naturally by the conditions $(a|N)(x)(x, \dot{x}) = a(x)(x, \dot{x})$ and $(f|N)(x) = f(x)$, for all $(x, \dot{x}) \in TN$.

Let us remark that the notion of restriction $(a, f)|N$ makes sense also in the case where N is any subset of M. In fact, we only need to give a meaning to the notion of a tangent vector $(x_0, \dot{x})$ at a point $x_0 \in N$. It is simply the tangent vector in $T_{x_0}M$ to a smooth curve $x(t)$, $t \in (-\delta, \delta)$ in M such that $x(t) \in N$, for all $t \in (-\delta, \delta)$ and $x(0) = x_0$.

Let $\varphi : N \to M$ be a given smooth map. Then the *pull-back* $\varphi^*(a, f) = (\varphi^* a, \varphi^* f)$ is the IDE with domain N and range F defined by $\varphi^* a(y)(y, \dot{y}) = a(\varphi(y))(T\varphi(y, \dot{y}))$ and $\varphi^* f(y) = f(\varphi(y))$

If $g : F \to G$ is a linear map we define the *projection* of (a, f) *by* g as being the IDE with domain M and range G defined by $(g \circ a, g \circ f)$. More precisely, we have $(g \circ a)(x)(x, \dot{x}) = g(a(x)(x, \dot{x}))$ and $(g \circ f)(x) = g(f(x))$, for all $(x, \dot{x}) \in TM$.

One can define operations like the direct sum $\oplus$ or tensor product $\otimes$ of IDE in a natural way. For instance, if (a_i, f_i), $i = 1, 2$, are given IDE with domain M and range F_i, $i = 1, 2$, then we can define the direct sum $(a_1, f_1) \oplus (a_2, f_2) \equiv (a_1 \oplus a_2, f_1 \oplus f_2)$ as an IDE with domain M and range $F_1 \oplus F_2$ by $(a_1 \oplus a_2)(x)\dot{x} = a_1(x)\dot{x} \oplus a_2(x)\dot{x}$, and $(f_1 \oplus f_2)(x) = f_1(x) \oplus f_2(x)$, for all $(x, \dot{x}) \in TM$. The tensor product is also defined in a natural way.

Using operations like the ones described above, one may sometimes simplify a given IDE. For instance, working in coordinates in F, say $(y_1, y_2, \ldots, y_m)$, if some of the equations, say corresponding to y_1, are a linear consequence of the others then it can be eliminated by using a projection $g(y_1, y_2, \ldots, y_m) = (y_2, \ldots, y_m)$, and the resulting system will be equivalent to the given one. We have the following result, whose proof is not difficult.

Theorem 2.1. *Let (a, f) be a given IDE with domain M and range F and let $N \subseteq M$ be a given submanifold defined regularly by equations $\varphi = 0$, where $\varphi : M \to H$ and H is a finite-dimensional vector space. Then the restriction $(a, f)|N$ has the same solutions as the IDE $(a \oplus 0, f \oplus \varphi)$ with domain M and range $F \oplus H$. It also has the same solutions as the IDE $(a \oplus D\varphi \oplus 0, f \oplus 0 \oplus \varphi)$ with domain M and range $F \oplus H \oplus H$. Here $D\varphi : TM \to H$ is defined by $D\varphi = p_2 \circ T\varphi$, where $p_2 : H \times H \to H$ is the projection on the second factor and $TH \equiv H \times H$.*

Remark

(a) This theorem is simple but useful. For instance, it allows sometimes replacement of a given IDE by an equivalent IDE whose domain and range are vector spaces, which sometimes simplifies practical calculations avoiding the usage of local charts whenever it is convenient. More precisely, let a given IDE (a, f) having domain $M \subseteq L$ be embedded in the vector space L and defined regularly by an equation $\varphi = 0$, where $\varphi : L \to H$, and let a be defined by a restriction $a = A|TM$, where $A : TL \to F$. Then, according to theorem 2.1, one can work equivalently with the system $(A \oplus 0, f \oplus \varphi)$, whose domain and range are vector spaces.

(b) One can define a category whose objects are of the type $(M, F, (a, f))$, where M is a manifold, F is a vector bundle over M and (a, f) is an IDE with domain M and range

F. A morphism $\varphi : (M, F, (a, f)) \to (N, G, (b, g))$ is given by a map $\varphi_d : M \to N$ and a vector bundle map $\varphi_r : F \to G$ over φ_d such that, for any $(x, \dot{x}) \in M$, we have $\varphi_r(a(x)\dot{x}) = b(y)\dot{y}$, and $g(y) = \varphi_r(f(x))$, where $(y, \dot{y}) = T\varphi_d(x, \dot{x})$. However, we will not use the categorical context in this paper.

3. Desingularization

The hypothesis that M_k is a submanifold at each stage of the basic algorithm described in section 1 is too restrictive as it is not satisfied for many examples of interest. We are going to show that in order to overcome part of these limitations, and at the cost of working in the subanalytic category rather than the C^∞ category, one can use results from real analytic desingularization theory. This is possible thanks to the theory of semianalytic and subanalytic sets developed originally by Lojasiewicz [25–27]. Important results in this field and systematic expositions using techniques which are simpler than the original ones are due to Gabrielov [28], Hironaka [23], Hardt [29, 30], Bierstone and Milman [15], Sussmann [31] and others. Our main reference will be [15], a very readable exposition of important points of the theory of subanalytic sets.

For the rest of this paper manifolds and maps will be real analytic, unless otherwise specified. For instance, if (a, f) is a given IDE with domain M then M will be a real analytic manifold and a, f will be real analytic maps.

Definition 3.1. *Let M be a real analytic manifold and let X be a closed subanalytic subset of M. A desingularization of X is a real proper analytic map $f : N \to M$ such that $f(N) = X$, where N is a real analytic manifold of the same dimension as X.*

This is a relatively weak notion of desingularization, but it is enough for our purposes. Existence of desingularizations $f : N \to M$ is guaranteed by the following theorem of Hironaka, see [15, 31] and references therein.

Theorem 3.2. *Let M be a real analytic manifold and let X be a closed subanalytic subset. Then there is a desingularization $f : N \to M$ of X.*

Desingularization results that include those of Hironaka have been recently proved in Bierstone and Milman [21].

In fact, in the present paper we only need the following weaker desingularization result, which is theorem 5.1 in [15],

Theorem 3.3. *Let M be a real analytic manifold and let X be a closed analytic subset. Then there is a desingularization $f : N \to M$ of X.*

3.1. Description of the algorithm

Let M be a manifold of dimension d and let (a, f) be a given IDE with domain M and range F. One of the main results proved in this paper to solve the IDE (2.1) consists, roughly, in transforming it into an equivalent IDE, say

$$\tilde{a}_2(y)\dot{y} = \tilde{f}_2(y)$$

on an analytic manifold $\tilde{M}_2$, which is an IDE of locally constant rank. The manifold $\tilde{M}_2$ will be constructed by an algorithm that involves a desingularization process.

3.1.1. The decomposition $M = M_0 \cup M_1 \cup M_2$. We are going to use the notation introduced in section 2. We can assume without loss of generality that arbitrary local coordinates in M and also a basis in F have been chosen, which makes sense of expressions like $\det A(x)$ and also of $[a(x), f(x)]$ as being a matrix whose first columns are the columns of $a(x)$ while the last column is $f(x)$, used below. First, let us assume that M is a connected manifold of dimension d. For $i = 0, 1, \ldots,$ let

$$\begin{aligned} S_i(M) &= \{x \in M \mid \operatorname{rank} a(x) \leqslant i\} \\ &= \{x \in M \mid \det A(x) = 0,\ A(x) \text{ submatrix of } a(x) \text{ of order } i+1\}. \end{aligned}$$

$S_i(M)$ is clearly a closed analytic subset of M, defined by analytic equations, for $i = 0, 1, \ldots.$

Also, for $i = 0, 1, \ldots,$ let $L_i(M) \subseteq S_i(M)$ be defined by

$$\begin{aligned} L_i(M) &= \{x \in S_i(M) \mid \operatorname{rank}[a(x), f(x)] \leqslant i\} \\ &= \{x \in S_i(M) \mid \det A(x) = 0,\ A(x) \text{ submatrix of } [a(x), f(x)] \text{ of order } i+1\}. \end{aligned}$$

Each $L_i(M)$ is a closed analytic subset of M defined by analytic equations.

Let

$$S_{k_1}(M) \subset S_{k_2}(M) \subset \cdots \subset S_{k_r}(M)$$

be the distinct nonempty $S_i(M)$. We observe that $S_{k_r}(M) \equiv M$. Consider the corresponding inclusions

$$L_{k_1}(M) \subseteq L_{k_2}(M) \subseteq \cdots \subseteq L_{k_r}(M).$$

We have that $\operatorname{rank} a(x) = \operatorname{rank}[a(x), f(x)] = k_j$ for each $x \in L_{k_j}(M) - S_{k_{j-1}}(M)$, $j = 1, \ldots, r$. The LAS associated with (2.1) has solution for each $x \in L_{k_j}(M) - S_{k_{j-1}}(M)$, $j = 1, \ldots, r$, where we have, by definition, $S_{k_0} = \emptyset$.

We remark the following useful facts: the set $L_{k_j}(M) - S_{k_j-1}(M)$ may be empty, for some $j = 1, \ldots, r$; we have $\dim S_{k_{r-1}}(M) < \dim M$; if $\dim(L_{k_r}(M)) = d$, then $L_{k_r}(M) = M$.

Now let M be a manifold of dimension d, and note that the components of M must have dimension less than or equal to d. Let $M_{\max} = \bigcup_j W_j$ be the union of all of the connected components of M with dimension d.

We will consider the following pairwise disjoint conditions for a given $W_j \subseteq M_{\max}$:

(a) $L_{k_r}(W_j) = \emptyset$.
(b) $L_{k_r}(W_j) \neq \emptyset$ and $\dim L_{k_r}(W_j) < d$.
(c) $L_{k_r}(W_j) \neq \emptyset$ and $\dim L_{k_r}(W_j) = d$.

According to these disjoint conditions, we decompose the index set into the disjoint union of the following subsets:

$$\begin{aligned} I_a &= \left\{ j \,\middle|\, L_{k_r}(W_j) = \emptyset \right\} \\ I_b &= \left\{ j \,\middle|\, L_{k_r}(W_j) \neq \emptyset \text{ and } \dim L_{k_r}(W_j) < d \right\} \\ I_c &= \left\{ j \,\middle|\, L_{k_r}(W_j) \neq \emptyset \text{ and } \dim L_{k_r}(W_j) = d \right\}. \end{aligned}$$

We then define the following pairwise disjoint subsets of M:

$$\begin{aligned} M_0 &= (M - M_{\max}) \cup \bigcup_{j \in I_b} L_{k_r}(W_j) \cup \bigcup_{j \in I_c} S_{k_{r-1}}(W_j) \\ M_1 &= \bigcup_{j \in I_a} W_j \cup \bigcup_{j \in I_b} \left(W_j - L_{k_r}(W_j)\right) \\ M_2 &= \bigcup_{j \in I_c} \left(W_j - S_{k_{r-1}}(W_j)\right). \end{aligned}$$

We have the following assertions, whose proof is easy: each subset $L_{k_r}(W_j) \subseteq W_j$, and each subset $S_{k_{r-1}}(W_j) \subseteq W_j$, is a closed analytic subset of W_j defined by analytic equations on W_j. In consequence, $W_j - L_{k_r}(W_j)$, $W_j - S_{k_{r-1}}(W_j)$ are open submanifolds of W_j.

The manifold M is the disjoint union

$$M = M_0 \cup M_1 \cup M_2.$$

The manifolds M_1 and M_2 are open submanifolds of M. The subset M_0 is a union of subsets defined by analytic equations on each W_j, union $M - M_{\max}$, and we have that $\dim M_0 < d$.

Remark. It is easy to prove, using the definitions of M_1 and M_2, that if M is connected, then we must have either $M_1 = \emptyset$ or $M_2 = \emptyset$.

3.1.2. Restrictions $(a, f)|M_0, (a, f)|M_1, (a, f)|M_2$ *and desingularization of* $(a, f)|M_0$. We have that the LAS associated with (2.1) has no solution for $x \in M_1$. On the other hand, it has solution for all $x \in M_2$, moreover, $(a, f)|M_2$, *is an IDE of locally constant rank*.

It remains to see what happens with the system restricted to M_0. The idea here is to desingularize each closed analytic subset $L_{k_r}(W_j) \subseteq W_j$ and $S_{k_{r-1}}(W_j) \subseteq W_j$. By forming the disjoint union of those desingularizations and $M - M_{\max}$ one obtains a desingularization of M_0, say

$$\pi_0 : M^1 \to M, \qquad \text{where} \quad \pi_0(M^1) = M_0.$$

Then (2.1) restricted to M_0, that is $(a, f)|M_0$, can be naturally *lifted*, using the pull-back operation, to an IDE $(a_1, f_1) = \pi_0^*((a, f)|M_0)$ on M^1 as follows:

$$a_1(y)\dot{y} = a(\pi_0(y))T_y\pi_0(y, \dot{y}) \qquad f_1(y) = f(\pi_0(y)).$$

We should remark at this point that in the present paper a tangent vector $(x, \dot{x})$ to M_0 at a point $x \in M_0$ is a vector $(x, \dot{x}) \in T_xM$ such that there is an analytic curve $z(t) \in M_0$, say defined for $t \in (-\delta, \delta)$, such that $z(0) = x$ and the tangent vector to $z(t)$ at $t = 0$ as a curve in M coincides with $(x, \dot{x})$. Then, in particular, if $y(t)$ is a given analytic curve in M^1 then

$$T_y\pi_0(y, \dot{y}) = \left.\frac{\mathrm{d}\pi_0(y(t))}{\mathrm{d}t}\right|_{t=0}$$

is a tangent vector to M_0 at $\pi_0(y(0))$.

Note that M^1 is a manifold of dimension $\dim M^1 = \dim M_0 < d$.

3.1.3. Desingularization of (a, f) *in a finite number of steps.* Now we repeat the process for the IDE (a_1, f_1) with domain M^1 and range F, proceeding as we did before with the system (a, f) with domain M and range F. We obtain a decomposition

$$M^1 = M_0^1 \cup M_1^1 \cup M_2^1.$$

We know that there is no solution to the LAS system

$$a_1(y)\dot{y} = f_1(y)$$

for $y \in M_1^1$. We also know that there is solution to the same LAS system for $y \in M_2^1$; moreover, $(a_1, f_1)\big|M_2^1$ is an IDE of locally constant rank. Now we desingularize M_0^1

$$\pi_1 : M^2 \to M^1, \qquad \pi_1(M^2) = M_0^1$$

and repeat the process. It is clear that we obtain a finite sequence of manifolds and maps

$$M^q \stackrel{\pi_{q-1}}{\to} M^{q-1} \stackrel{\pi_{q-2}}{\to} \cdots \stackrel{\pi_1}{\to} M^1 \stackrel{\pi_0}{\to} M,$$

where $\pi_0(M^1) = M_0, \pi_1(M^2) = M_0^1$, and in general $\pi_i(M^{i+1}) = M_0^i$, for $i = 0, \ldots, q-1$, where we have written $M^0 \equiv M$ to unify the notation.

We have obtained a finite recursive procedure that reduces the problem to a finite number of IDE of locally constant rank, namely, the IDE of locally constant rank $(a_i, f_i)\big|M_2^i$, for $i = 0, 1, \ldots, q$, where we have written $(a_0, f_0) = (a, f)$, to unify the notation. We will call this a *desingularization process* and the sequence of maps π_i and IDE $(a_{i+1}, f_{i+1}), i = 0, \ldots, q-1$, a *desingularization* of (a, f).

In section 4 we give some illustrative examples.

The collection of IDE $(a_k, f_k)|M_2^k, k = 0, \ldots, q$, defines a single IDE $(\tilde{a}_2, \tilde{f}_2)$ of locally constant rank in the disjoint union $\tilde{M}_2 = \bigsqcup_{k=0}^q M_2^k$, as we have said at the beginning of section 3.1. We have a natural projection $\tilde{\pi}_2 : \tilde{M}_2 \to M$. This IDE $(\tilde{a}_2, \tilde{f}_2)$ with domain $\tilde{M}_2$ and range F is called the *desingularized IDE*.

Remark. A couple of remarks are in order to help simplify the practical application of the algorithm.

(a) As we have said before the range F remains the same throughout the application of the algorithm. However, in practice, it is sometimes convenient to apply theorem 2.1, which may imply a change of F, to simplify calculations.

(b) Assume for a moment that M is connected, then there is only one connected component $W_j \equiv M$. In order to apply the algorithm as it has been described it is not necessary to calculate $S_i(M)$ and $L_i(M)$, for all $i = 1, \ldots, k_r$. We first determine the number k_r, which is the maximum of rank $a(x), x \in M$. Then we know that $S_{k_r}(M) = M$. Next we determine whether $L_{k_r}(M) = \emptyset$ or $L_{k_r}(M) \neq \emptyset$. If $L_{k_r}(M) = \emptyset$ then we immediately conclude that $M_0 = M_2 = \emptyset$ and $M = M_1$. If $L_{k_r}(M) \neq \emptyset$ and $\dim L_{k_r}(M) < d$ we can conclude that $M_0 = L_{k_r}(M)$ and $M_2 = \emptyset$, therefore $M = M_0 \cup M_1$. If $L_{k_r}(M) \neq \emptyset$ and $\dim L_{k_r}(M) = d$ we must calculate $S_{k_{r-1}}(M)$, then we have $M_0 = S_{k_{r-1}}(M)$ and $M_1 = \emptyset$, therefore $M = M_0 \cup M_2$.

If the manifold is a disjoint union of connected components W_j as explained before, then for each W_j of maximum dimension d we calculate the decomposition

$$W_j = (W_j)_0 \cup (W_j)_1 \cup (W_j)_2,$$

as explained above, and then we can calculate

$$M_0 = (M - M_{\max}) \cup \bigcup_j (W_j)_0,$$

$$M_1 = \bigcup_j (W_j)_1,$$

$$M_2 = \bigcup_j (W_j)_2.$$

3.2. *Solving IDE via desingularization*

In this paragraph we will describe basic aspects of the relationship between solutions to a given IDE of the type (2.1), say $a(x)\dot{x} = f(x)$, and solutions to the desingularized system $(\tilde{a}_2, \tilde{f}_2)$ described in the previous paragraph. The main results about this kind of questions are precisely stated and proved in section 5. In that section solutions to a given IDE belong to a class of curves called *as-curves*, see definition 5.3, rather than to the class of analytic curves, in order to use the theory of subanalytic subsets as explained in section 1. However, for the reader who is not specifically interested in the technical aspects of the proofs going carefully through section 5 is not necessary. In fact, the basic strategy to solve a given IDE via desingularization is described in the present paragraph, below, and it will be enough here to think just of analytic solutions. One must keep in mind that if $x : (t_0, t_1) \to M$ is an analytic curve then any restriction of the type $x|(\tau_0, \tau_1), x|[\tau_0, \tau_1), x|(\tau_0, \tau_1], x|[\tau_0, \tau_1]$, where $t_0 < \tau_0 < \tau_1 < t_1$, is an as-curve. It is also useful to keep in mind that every analytic solution can be obtained by gluing together a countable (finite or infinite) number of as-solutions. All this can be proved easily from the definition of an as-curve.

We must remark that, in spite of the relative complexity of the proofs, in order to interpret our main theorems, namely theorems 5.12 and 5.14, it is enough to know the definition of an as-curve, given at the beginning of section 5.

It is easy to see using the definitions given in section 3 that any solution $y(t), t \in (t_0, t_1)$, to $(\tilde{a}_2, \tilde{f}_2)$ projects to a solution $x(t) = \tilde{\pi}_2(y(t))$ to the system $a(x)\dot{x} = f(x)$. On the other hand, it is also easy to see that the converse of this statement is not true. In fact, assume for simplicity that $q = 1$ that is $\tilde{M}_2 = M_2 \cup M_2^1$ (disjoint union) and that $M_0 \neq \emptyset$. Assume that a certain solution $x(t), t \in (t_0, t_1)$ to (2.1) is analytic. Then, since M_0 is defined by analytic equations we must have that either $x(t) \in M_0, t \in (t_0, t_1)$ or $x(t) \in M_0$ only for those t belonging to a certain sequence (finite or infinite) $\tau_1 < \tau_2 < \cdots$. Assume that the latter holds and that the sequence $\tau_1 < \tau_2 < \cdots$ is nonempty, which is perfectly possible in examples. Any solution $y(t), t \in (t_0, t_1)$ to $(\tilde{a}_2, \tilde{f}_2)$ is continuous so it must satisfy one and only one of the following conditions: $y(t) \in M_2, t \in (t_0, t_1)$ or $y(t) \in M_2^1, t \in (t_0, t_1)$. Since $\tilde{\pi}_2|(M_2) = 1_{M_2}$ and $\tilde{\pi}_2\left(M_2^1\right) = M_0$ it is easy to see that we cannot have $\tilde{\pi}_2(y(t)) = x(t), t \in (t_0, t_1)$. This means that in this case the solution $x(t)$ cannot be recovered completely as a projection via $\tilde{\pi}_2$ of a solution to $(\tilde{a}_2, \tilde{f}_2)$. However, we still have that each piece of $x(t)$ of the type $x|(\tau_i, \tau_{i+1}), i = 1, 2, \ldots$, is a projection via $\tilde{\pi}_2$ of a solution to $(\tilde{a}_2, \tilde{f}_2)$. In fact, since $\tilde{\pi}_2|M_2 = 1_{M_2}$ we have that $y(t) = x(t), t \in (\tau_i, \tau_{i+1})$ is such a solution. Now assume that $x(t) \in M_0, t \in (t_0, t_1)$. Using results from section 5, in fact, using theorem 5.12(*b*), we can easily show that there is a decomposition of $x(t)$ in pieces of the type $x|(\tau_i, \tau_{i+1}), i = 1, 2, \ldots$, and each piece is a projection via $\tilde{\pi}_2$ of a solution $y(t), t \in (\tau_i, \tau_{i+1})$ to $(\tilde{a}_2, \tilde{f}_2)$ or, since in this example $\tilde{\pi}|M_2^1 = \pi_1$, a solution to (a_1, f_1). This kind of result holds for an IDE with arbitrary q, not just $q = 1$. In conclusion, we have the following strategy to find all analytic solutions to a given IDE via desingularization.

(1) For a given IDE (a, f) of the type (2.1) find the desingularized system $(\tilde{a}_2, \tilde{f}_2)$, as explained in section 3.
(2) Find all the analytic solutions $y(t), t \in (t_0, t_1)$ to $(\tilde{a}_2, \tilde{f}_2)$. For a given initial condition $y_0 \in M_2^k, k = 0, 1, \ldots, q$, all solutions $y(t) \in M_2^k, t \in (t_0, t_1)$, can be found in a similar way as we did in section 1.1 to solve equation (1.1).
(3) For each projected solution $x(t) = \tilde{\pi}_2(y(t)), t \in (t_0, t_1)$, determine if there are continuous extensions to the intervals $[t_0, t_1), (t_0, t_1]$ and if these extensions are derivable at t_0 on the right and at t_1 on the left, respectively. Finally, determine if these extensions satisfy the given IDE at points t_0 and t_1.

As we have said before for any given analytic solution $x(t), t \in (t_0, t_1)$, there is a decomposition of $x(t)$ in pieces of the type $x|(\tau_i, \tau_{i+1}), i = 1, 2, \ldots$, and each piece is a projection via $\tilde{\pi}_2$ of a solution $y(t), t \in (\tau_i, \tau_{i+1})$, to $(\tilde{a}_2, \tilde{f}_2)$. We can conclude that any analytic solution to (a, f) can be recovered by gluing together extensions to the endpoints of their intervals of definitions, if such extensions exist, of projections via $\tilde{\pi}_2$ of solutions to $(\tilde{a}_2, \tilde{f}_2)$.

Remark. The previous method does *not* give an answer to the following kind of initial condition problem for the IDE of the type (2.1): given an initial condition $x(t_0) = x_0$, where $x_0 \in M_0$ determine if there is an as-solution $x(t), t \in [t_0, t_1)$, such that $x(t) \in M_2$ for $t \in (t_0, t_1)$. Or, more generally given an initial condition $y(t_0) = y_0$, where $y_0 \in M_0^k$, for some $k = 0, 1, \ldots, q$, determine if there is an as-solution $y(t), t \in [t_0, t_1)$, such that $y(t) \in M_2^k$ for $t \in (t_0, t_1)$. This kind of questions is important and will be addressed in subsequent works.

4. An example from nonholonomic mechanics

A nonholonomic system on a configuration space Q is given by a Lagrangian $L : TQ \to \mathbb{R}$ and a distribution $D \subseteq TQ$. Lagrange–D'Alembert's equations of motion are derived using Lagrange–D'Alembert's principle. These equations form an IDE which is known to be equivalent to an ODE on TQ in the important case in which L is nondegenerate. By adding the equation of preservation of energy $E = e$, where $E : TQ \to \mathbb{R}$ is the energy function and e is a fixed energy level, to the previous IDE we obtain an equivalent IDE on TQ. There are some recent references where nonholonomic systems are studied from a DAE (differential algebraic equation, see the introduction) perspective [32].

In the rather common case in mechanics in which Q is a principal bundle with group G and both L and D are invariant, Lagrange–D'Alembert's principle can be reduced [5] and one obtains reduced equations of motion on TQ/G. Upon the choice of a principal connection, those equations can be written as Lagrange–D'Alembert–Poincaré equations of motion on $TQ/G \equiv T(Q/G) \oplus \tilde{\mathfrak{g}}$, where $\tilde{\mathfrak{g}}$ is the adjoint bundle. These equations form an IDE which has unique solution for each initial condition provided that the Lagrangian is nondegenerate. If the reduced equation which gives preservation of energy at the reduced level is added to that IDE a new equivalent IDE is obtained. It would be of interest, for a given example, to show that the latter is equivalent to an ODE on a known manifold. What we are going to do next is to show that this is precisely the case for an interesting example from mechanics, using the methods of the present paper.

4.1. The symmetric sphere rolling without sliding or spinning

A rigid sphere rolling on a horizontal plane can be modelled as a nonholonomic system on the manifold $SO(3) \times \mathbb{R}^2$ where, for a given element $(A, x) \in SO(3) \times \mathbb{R}^2$, A represents a rigid rotation and x the position of the *point of contact* of the sphere with the plane. The kinematics of this system can be described as follows. We assume that there is an orthonormal system fixed in the space, say (e_1, e_2, e_3), $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, then we have a basis moving with the body, (Ae_1, Ae_2, Ae_3), where $A = A(t)$. We introduce the variable $z \in S^2$, given by $z = Ae_3$. The spatial angular velocity ω can be written as $\omega = v_0 z + z \times \dot{z}$, so $v_0 = \langle \omega, z \rangle$ is the component of ω along z. The nonholonomic constraint is given by the nonsliding condition $\omega \times re_3 = \dot{x}$, where r is the radius of the sphere. Now we will imagine that the sphere is *elastic* and that deformations are small and concentrated near the *area of contact* which is a small circle whose centre has a position given by x. Then we must add to the usual nonsliding condition for the rigid rolling sphere the extra condition that the vertical component of the spatial angular velocity is 0, that is, $\omega_3 = 0$. This is sometimes called *Veselova's constraint* [33]. We are going to assume that the centre of mass coincides with the centre of the sphere and that the principal axis of inertia are (Ae_1, Ae_2, Ae_3). The three principal moments of inertia of the sphere are I_1, I_2, I_3, and we are going to assume that $I_1 = I_2$. We introduce the adimensional quantities $\alpha = I_3/I_1$ and $\beta = Mr^2/I_1$, where M is the mass of the sphere. The Lagrangian of the system is given by the kinetic energy

$$\tfrac{1}{2}I_1\dot{z}^2 + \tfrac{1}{2}I_3v_0^2 + \tfrac{1}{2}M\dot{x}^2$$

where $\dot{x}$ is the velocity of the centre of the sphere. Since the nonholonomic constraint is given by $\dot{x} = \omega \times re_3$ and $\omega_3 = 0$, we can conclude that the kinetic energy of the actual motion of the symmetric sphere is given by

$$E = \tfrac{1}{2}(I_1 + Mr^2)\dot{z}^2 + \tfrac{1}{2}(I_3 + Mr^2)v_0^2.$$

4.2. The IDE for the symmetric sphere rolling without sliding or spinning

As a result of reduction by the symmetry, in this case reduction by the subgroup $SO(2) \times \mathbb{R}^2$, we obtain the following system of Lagrange–D'Alembert–Poincaré equations, which is an IDE,

$$(\alpha + \beta)(z \times e_3)\dot{v}_0 + (1 + \beta)\langle z, e_3\rangle\nabla_{\dot{z}}\dot{z} - (\alpha + \beta)v_0\langle z, e_3\rangle(z \times \dot{z}) = 0 \tag{4.1}$$

$$v_0\langle z, e_3\rangle + \langle z \times \dot{z}, e_3\rangle = 0. \tag{4.2}$$

Here ∇ represents the Levi-Civita connection on S^2 with respect to the standard metric. This is a consequence of the methods developed in [5], after some more or less straightforward calculations which we will not explain here. The previous Lagrange–D'Alembert–Poincaré equations are derived under the assumption $z_3 \neq 0$ because the so-called *dimension assumption* adopted in [5] is not satisfied for the whole manifold S^2. Nevertheless, by continuity, equations (4.1), (4.2) are also satisfied by the motion of the rolling ball for $z_3 = 0$. Without using the derivation of the Lagrange–D'Alembert–Poincaré equations, the careful reader may want to check directly that the previous system of equations or, equivalently, the system of equations (4.8)–(4.15), is equivalent to balance of momentum plus the condition $\omega_3 = 0$.

Since $\nabla_{\dot{z}}\dot{z} = z \times (\ddot{z} \times z)$ by taking the inner product of (4.1) with $\dot{z}$ and using (4.2) we get, at least for $\langle z, e_3\rangle \neq 0$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\big((1 + \beta)\dot{z}^2 + (\alpha + \beta)v_0^2\big), \tag{4.3}$$

from which one deduces

$$2\epsilon = (1 + \beta)\dot{z}^2 + (\alpha + \beta)v_0^2, \tag{4.4}$$

where ϵ represents the normalized energy. This equation represents conservation of energy, as one can check more directly by looking at the expression of the kinetic energy E given at the beginning of this section. We shall assume from now on that $\epsilon > 0$, otherwise the motion is trivial.

We have the following equations to be satisfied for the symmetric elastic sphere in variables (z, u) where $\dot{z} = v$ and $v \times z = u$, so the variable v_0 does not appear,

$$(1 + \beta)\langle z, e_3\rangle\langle\dot{u}, e_3 \times z\rangle + (\alpha + \beta)\langle u, e_3\rangle^2 = 0 \tag{4.5}$$

$$(1 + \beta)\langle z, e_3\rangle^2u^2 + (\alpha + \beta)\langle u, e_3\rangle^2 - 2\epsilon\langle z, e_3\rangle^2 = 0. \tag{4.6}$$

Equation (4.5) is obtained by taking the inner product of (4.1) with e_3 and using (4.2) while equation (4.6) is obtained from equation (4.3) and equation (4.2).

Equations (4.5) and (4.6) involve the variables $(z, u) \in TS^2$ and we have a natural inclusion $TS^2 \subseteq S^2 \times \mathbb{R}^3$, where

$$TS^2 = \{(z, u) \in S^2 \times \mathbb{R}^3 : \langle z, u\rangle = 0\}.$$

Equations (4.5), (4.6) form an IDE in the manifold TS^2. By adding the equation $z^2 = 1$, we obtain an equivalent IDE in the variables $(z, u) \in \mathbb{R}^3 \times \mathbb{R}^3$. Let us include, in addition, the equation of conservation of energy (4.3), written in terms of the variables (z, u, v_0), as follows:

$$2\epsilon = (1+\beta)u^2 + (\alpha+\beta)v_0^2. \tag{4.7}$$

In other words, we are going to study the system of equations given by (4.5), (4.6), (4.7) for a fixed $\epsilon > 0$. Of course, equations (4.6) and (4.7) taking into account (4.2) are redundant for $z_3 \neq 0$, but for $z_3 = 0$ the system given by equations (4.5), (4.6) includes solutions of the type $z(t) = \text{const}$, with $z_3 = 0$, where v_0 takes any given value. Since $u = 0$ for this kind of motion, the energy is given by $2\epsilon = (\alpha+\beta)(v_0)^2$ and therefore the condition that the energy must have a fixed value will not be satisfied. Of course, we can study with our methods both the system given by (4.5), (4.6), (4.7) and also the system given by (4.5), (4.6), but we will choose to study just the first of them, for simplicity.

Considering that $z^2 = 1$ and $\dot{z} = z \times u$, we must have $\langle z, u\rangle = 0$. Using what was said in the previous paragraph, and according to theorem 2.1, we can write the IDE for the system in variables $(z, u, v_0) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ as follows:

$$\dot{z}_1 = z_2u_3 - z_3u_2 \tag{4.8}$$

$$\dot{z}_2 = z_3u_1 - z_1u_3 \tag{4.9}$$

$$\dot{z}_3 = z_1u_2 - z_2u_1 \tag{4.10}$$

$$0 = (1+\beta)z_3(-z_2\dot{u}_1 + z_1\dot{u}_2) + (\alpha+\beta)u_3^2 \tag{4.11}$$

$$0 = (1+\beta)z_3^2\left(u_1^2+u_2^2+u_3^2\right) + (\alpha+\beta)u_3^2 - 2\epsilon z_3^2 \tag{4.12}$$

$$0 = z_1^2 + z_2^2 + z_3^2 - 1 \tag{4.13}$$

$$0 = z_1u_1 + z_2u_2 + z_3u_3 \tag{4.14}$$

$$0 = 2\epsilon - (1+\beta)u^2 - (\alpha+\beta)v_0^2. \tag{4.15}$$

The system (4.8)–(4.15) can be written in the form $a(X)\dot{X} = f(X)$, with $X = (z, u, v_0)$.

4.3. Application of the algorithm

We will work on the manifold $M = \mathbb{R}^7$, where $(z_1, z_2, z_3, u_1, u_2, u_3, v_0) \in \mathbb{R}^7$ are independent variables. Then our IDE is given by equations (4.8)–(4.15). We can easily see that $k_r = 4$, $S_4(M) = M$, $L_4(M) = M_0$, $M_1 = M - L_4(M)$, $M_2 = \emptyset$. Now we shall describe M_0 by equations. Let

$$\varphi_1 = -(1+\beta)z_2z_3 \tag{4.16}$$

$$\varphi_2 = (1+\beta)z_1z_3 \tag{4.17}$$

$$\nu_1 = (1+\beta)z_3^2\left(u_1^2+u_2^2+u_3^2\right) + (\alpha+\beta)u_3^2 - 2\epsilon z_3^2 \tag{4.18}$$

$$\nu_2 = z_1^2 + z_2^2 + z_3^2 - 1 \tag{4.19}$$

$$\nu_3 = z_1u_1 + z_2u_2 + z_3u_3 \tag{4.20}$$

$$\nu_4 = 2\epsilon - (1+\beta)u^2 - (\alpha+\beta)v_0^2. \tag{4.21}$$

As we know $M_0 = L_4(M)$ is given by the condition that rank$[a, f] \leqslant 4$. Let

$$M_{0a} = \{\varphi_1 = 0, \varphi_2 = 0\} \tag{4.22}$$

$$= \{z_3 = 0\} \cup \{z_1 = 0, z_2 = 0\} \tag{4.23}$$

$$M_{0b} = \{\nu_1 = 0, \nu_2 = 0, \nu_3 = 0, \nu_4 = 0\}. \tag{4.24}$$

Then we can easily see that $M_0 = M_{0a} \cup M_{0b}$. The desingularization M^1 of M_0 will be the disjoint union of the desingularizations of M_{0a} and M_{0b}.

The desingularization M^1_a of M_{0a} can be described by $M^1_a \equiv \{z_3 = 0\} \bigsqcup \{z_1 = 0, z_2 = 0\}$, where $\bigsqcup$ means *disjoint union* and the projection π_0 is the identity on each disjoint piece of M^1_a. One can see using (4.8)–(4.15) that the lifted system $(a_1, f_1)|\{z_3 = 0\}$ satisfies $z_3 = 0, u_3 = 0, z_1^2 + z_2^2 = 1$, which implies $\dot{z} = 0$, and also, since $u = \dot{z} \times z$, that $u = 0$. This describes the motion completely. It consists of the rolling of the sphere with $z(t) = (z_{10}, z_{20}, 0)$ fixed and the z component of the angular velocity v_0 satisfies $2\epsilon = (\alpha + \beta)(v_0)^2$. The lifted system $(a_1, f_1)|\{z_1 = 0, z_2 = 0\}$ satisfies $z_1 = 0, z_2 = 0, z_3 = \pm 1$, therefore $\dot{z} = 0$, and then $u = 0$, which contradicts equation $\nu_1 = 0$, because we have assumed $\epsilon > 0$. So there is no motion, that is, no solution, for the system $(a_1, f_1)|\{z_1 = 0, z_2 = 0\}$.

Now we will desingularize M_{0b}. We are going to see that M_{0b} is in fact a nonsingular manifold. More precisely, we will define the desingularized manifold M^1_b by equations in the variables (z, u, v_0), with $v_0 z_3 = u_3$, from (4.2). For simplicity, we call $\mu = 2\epsilon/(1+\beta) > 0$ and $\lambda = (\alpha + \beta)/(1 + \beta) > 0$, from now on. Then we have the following equations defining the nonsingular manifold M^1_b:

$$0 = u_3 - v_0 z_3 \tag{4.25}$$

$$0 = u_1^2 + u_2^2 + u_3^2 + \lambda v_0^2 - \mu \tag{4.26}$$

$$0 = z_1^2 + z_2^2 + z_3^2 - 1 \tag{4.27}$$

$$0 = z_1 u_1 + z_2 u_2 + z_3 u_3. \tag{4.28}$$

We are not going to give a detailed proof that the system above defines a nonsingular manifold. It is not difficult to prove it, by applying the implicit function theorem at each point of the manifold. The map $\pi_0 : M^1_b \to M$ is then given by the restriction of the identity $(z, u, v_0) \to (z, u, v_0)$ to M^1_b and one can check that the image of π_0 is precisely M_{0b}.

According to theorem 2.1, the system lifted to M^1_b has the same solutions as the system given by

$$\dot{z}_1 = z_2 u_3 - z_3 u_2 \tag{4.29}$$

$$\dot{z}_2 = z_3 u_1 - z_1 u_3 \tag{4.30}$$

$$\dot{z}_3 = z_1 u_2 - z_2 u_1 \tag{4.31}$$

$$z_2 \dot{u}_1 - z_1 \dot{u}_2 = \lambda v_0 u_3 \tag{4.32}$$

$$0 = u_3 - v_0 z_3 \tag{4.33}$$

$$0 = u_1^2 + u_2^2 + u_3^2 + \lambda v_0^2 - \mu \tag{4.34}$$

$$0 = z_1^2 + z_2^2 + z_3^2 - 1 \tag{4.35}$$

$$0 = z_1 u_1 + z_2 u_2 + z_3 u_3. \tag{4.36}$$

It is not difficult to see that the previous IDE defines an analytic vector field on the manifold defined by the last four equations, which is essentially the manifold M^1_b. In fact, using the implicit function theorem one can calculate the tangent space at each point of this manifold and then see that there is only one solution $(\dot{z}, \dot{u}, \dot{v}_0)$ to the linear system associated with the IDE which is tangent to the manifold. One can also see that this solution depends analytically on the point $(z, u, v_0) \in M_b^1$. In other words, the previous IDE defines an analytic vector field on M_b^1.

4.4. Identification of M_b^1

In [10], we have shown the following parametrization of M_b^1 in variables (θ, φ, ψ), which can be checked directly after some straightforward calculations:

$$z_1 = \sin\theta\cos\varphi \tag{4.37}$$

$$z_2 = \sin\theta\sin\varphi \tag{4.38}$$

$$z_3 = \cos\theta \tag{4.39}$$

$$u_1 = -a\cos(\varphi - \psi)\cos^2\theta\cos\varphi - b\sin(\varphi - \psi)\sin\varphi \tag{4.40}$$

$$u_2 = -a\cos(\varphi - \psi)\cos^2\theta\sin\varphi + b\sin(\varphi - \psi)\cos\varphi \tag{4.41}$$

$$u_3 = a\cos(\varphi - \psi)\cos\theta\sin\theta \tag{4.42}$$

$$v_0 = a\cos(\varphi - \psi)\sin\theta, \tag{4.43}$$

where

$$a = \sqrt{\frac{\mu}{\lambda\sin^2\theta + \cos^2\theta}}, \qquad b = \sqrt{\mu'}.$$

In other words, by some straightforward calculations we can check that $(z_1, z_2, z_3, u_1, u_2, u_3, v_0)$ in coordinates (θ, φ, ψ) satisfies (4.25)–(4.28).

We can easily prove that equations (4.37)–(4.43) define a diffeomorphism $f : S^2 \times S^1 \to M_b^1$, $f(z, (\cos\psi, \sin\psi)) = (z, u, v_0)$, which gives the desired identification of M_b^1. This is straightforward, although it is not very short.

5. The main theorems

In this section, we will show in which precise sense the solutions to a desingularization of a given analytic IDE (a, f) are related to the solutions to (a, f). It is clear that in certain examples of IDE one can show in a more or less direct way that solutions to the desingularized system project via $\tilde{\pi}_2$ onto solutions to the given IDE, and also that solutions to the IDE are such projections. For instance, if one is interested only in the local behaviour of solutions near a singular point of M_0 it is sometimes enough to use a blow-up to desingularize M_0 at that point. One can show in certain cases in a simple and useful way the relationship between solutions to the given system and solutions to the system obtained by blow-up. An interesting example appears in [17].

However, a desingularization is a composition of blow-ups. In the present paper we want a general global result showing that a curve, belonging to a certain convenient class of curves, is a solution to a given IDE if and only if it is essentially the projection via the map $\tilde{\pi}_2$ of a solution to the desingularized system belonging to the same class of curves. In order to be able to use the theory of subanalytic sets we need to carefully define a convenient class of curves, which we will do in the next paragraph.

From now on, we will often use theorem 6.1 of [15] that a subanalytic subset of dimension 1 of an analytic manifold M is a semianalytic subset. We will also often use the fact that the image of a relatively compact subanalytic subset under a subanalytic map is a subanalytic subset, see [15] immediately after definition 3.2. Since an analytic map is subanalytic we can deduce that the image of a relatively compact subanalytic subset under an analytic map is a subanalytic subset, which will be also useful for us.

In what follows, we will work with several types of intervals, like (τ_0, τ_1), $[\tau_0, \tau_1)$, $(\tau_0, \tau_1]$ or $[\tau_0, \tau_1]$. We will usually assume that τ_0 and τ_1 are real numbers. However, some definitions and results are also valid for the case in which some of the open ends of an interval is $\pm\infty$, that is, for intervals $(-\infty, \tau_1)$, $(\tau_0, +\infty)$, $(-\infty, +\infty)$, $[\tau_0, +\infty)$, $(-\infty, \tau_1]$.

5.1. The notion of an as-curve

Inspired by [15] definition 3.2 we will define

Definition 5.1. *A subanalytic curve $x : (t_0, t_1) \to M$ is a subanalytic map, that is, a map such that* graph $x \subseteq \mathbb{R} \times M$ *is a subanalytic subset. We define the notion of a subanalytic curve $x : [t_0, t_1) \to M$, $x : (t_0, t_1] \to M$ or $x : [t_0, t_1] \to M$ in a similar way.*

In definition 5.1 since $\dim(\mathrm{graph}\, x) = 1$ we have that graph x is a semianalytic set.

Lemma 5.2. *(a) Let $x : [t_0, t_1) \to M$ be a continuous subanalytic curve whose graph is a relatively compact subset. Then there is a uniquely defined continuous subanalytic extension $\bar{x} : [t_0, t_1] \to M$. A similar result holds for subanalytic curves $x : (t_0, t_1] \to M$ or $x : (t_0, t_1) \to M$.*

(b) Let $x : [t_0, t_1) \to M$ be a continuous subanalytic curve whose graph is not a relatively compact subset. Then graph x *is closed. A similar result holds for subanalytic curves $x : (t_0, t_1] \to M$ or $x : (t_0, t_1) \to M$.*

Proof. First we shall prove (a). We need to show first that the limit of $x(t)$ as $t \rightharpoonup t_1^-$ exists. Using corollary 2.8 of [15] we can deduce that the closure $\overline{G}$ of $G = \mathrm{graph}\, x$ in $\mathbb{R} \times M$ is subanalytic and compact, and then also $\overline{G} \cap (\{t_1\} \times M)$ is subanalytic and compact and it is easy to see that it is nonempty. Let $x_1 \in \overline{G} \cap (\{t_1\} \times M)$ be given. One can choose local coordinates at x_1, then for any small $\epsilon > 0$ the set $G_\epsilon = G \cap ((t_0, t_1) \times B_\epsilon(x_1))$ is relatively compact and nonempty. We can easily deduce from theorem 3.14 of [15], or also from comments after definition 3.1 of [15], that any relatively compact subanalytic set has a finite number of connected components. Let $C_i, i = 1, \ldots, n(\epsilon)$, be the connected components of G_ϵ. It is not difficult to see that each connected component C_i is of the type $C_i = \mathrm{graph}(x|(\alpha_i, \beta_i)), i = 1, \ldots, n(\epsilon)$. We can assume without loss of generality that $\beta_i \leqslant \alpha_{i+1}, i = 1, \ldots, n(\epsilon) - 1$. Since x_1 is a limit point of G_ϵ we must have $\beta_{n(\epsilon)} = t_1$, which implies that $x(t) \in B_\epsilon(x_1)$ for all $t \in (\alpha_{n(\epsilon)}, \beta_{n(\epsilon)})$, as we wanted to prove. The fact that the continuous extension $\bar{x} : [t_0, t_1] \to M$ is a subanalytic curve follows from corollary 2.8 of [15] or, also, from comments after definition 3.1 of [15]. The rest of the proof of (a) can be performed in a similar way. Now we will prove (b). If graph x is not relatively compact then $\overline{G} \cap (\{t_1\} \times M)$ must be empty, otherwise we can proceed as in the proof of (a) and we can conclude that the limit of $x(t)$ as $t \rightharpoonup x_1^-$ exists and then one can show easily that G is relatively compact. Since all the limit points of G must belong to $\overline{G} \cap (\{t_1\} \times M)$ we have that G is closed. The rest of the proof of (b) can be performed in a similar way. □

In order to define a convenient class of curves to solve a given IDE we introduce the following notion:

Definition 5.3. *(a) An analytic–semianalytic curve $x(t), t \in (t_0, t_1)$, in M, where M is a given manifold, is an analytic map $x : (t_0, t_1) \to M$ which is also a subanalytic curve, that is, such that* graph x *is a semianalytic subset of $\mathbb{R} \times M$. We will often call such an analytic–semianalytic curve in M simply an as-curve in M.*

(b) An analytic–semianalytic curve (or as-curve) $x(t), t \in [t_0, t_1), (t \in (t_0, t_1], t \in [t_0, t_1])$ in M, where M is a given manifold, is a continuous map $x : [t_0, t_1) \to M$ (respectively, $x : (t_0, t_1] \to M, x : [t_0, t_1] \to M$), which is also a subanalytic curve, that is, such that graph x *is a semianalytic subset of $\mathbb{R} \times M$ and $x|(t_0, t_1)$ is an as-curve in M.*

For instance, $x = \sqrt{t}, t \in (0, c)$ (or $t \in [0, c), t \in (0, c], t \in [0, c]$), with $c > 0$, are as-curves in $\mathbb{R}$. On the other hand, $x = t \sin(\pi/t), t \in (0, c)$, with $c > 0$, is *not* an as-curve in $\mathbb{R}$, but $x = t \sin(\pi/t), t \in (\delta, c)$, with $0 < \delta < c$, *is* an as-curve in $\mathbb{R}$.

Next, we will give some lemmas that we need to prove our main results. They give some basic properties of as-curves.

Lemma 5.4. *(a) Let $x : (t_0, t_1) \to N$ be a given analytic map, where N is a given manifold. Then any map $x|(\bar{t}_0, \bar{t}_1) : (\bar{t}_0, \bar{t}_1) \to N$, where $\bar{t}_0$ and $\bar{t}_1$ are such that $t_0 < \bar{t}_0 < \bar{t}_1 < t_1$, is an as-curve in N. In a similar way, any map $x|[\bar{t}_0, \bar{t}_1) : [\bar{t}_0, \bar{t}_1) \to N, x|(\bar{t}_0, \bar{t}_1] : (\bar{t}_0, \bar{t}_1] \to N$, or $x|[\bar{t}_0, \bar{t}_1] : [\bar{t}_0, \bar{t}_1] \to N$, with $\bar{t}_0$ and $\bar{t}_1$ as before, is an as-curve in N.*

(b) Let $x : (t_0, t_1) \to N$ be an as-curve and assume that graph x *is a relatively compact subset of $\mathbb{R} \times N$. Then there is a unique continuous extension $\bar{x} : [t_0, t_1] \to M$. This extension is an as-curve.*

Proof. Let us prove (a). We have that graph$(x|(\bar{t}_0, \bar{t}_1))$ is a semianalytic subset of $\mathbb{R} \times N$ defined as $\{(t, x) \in (t_0, t_1) \times N : \bar{t}_0 < t < \bar{t}_1, x = x(t)\}$. The rest of (a) can be proved in a similar way. To prove (b) we simply apply lemma 5.2. □

Lemma 5.5. *Let $x : [t_0, t_1) \to M$ be a subanalytic map and assume that x is continuous at t_0. Then there exists $t_2 \in (t_0, t_1)$ such that $x|[t_0, t_2]$ is an as-curve. A similar result holds for a subanalytic map $x : (t_0, t_1] \to M$ continuous at t_1, that is, there exists $t_2 \in (t_0, t_1)$ such that $x|[t_2, t_1]$ is an as-curve.*

Proof. If x is a constant the result follows immediately. Let us assume that x is not a constant. It is easy to see that graph$(x|[t_0, t_2])$ is a semianalytic subset of $\mathbb{R} \times M$ of dimension 1, for every $t_2 \in (t_0, t_1)$. We can assume without loss of generality, using, for instance, Whitney embedding theorem, that $M \subseteq U$ is an analytic submanifold of U, where U is a real finite-dimensional vector space. Let $p_1 : \mathbb{R} \times U \to \mathbb{R}$ be the projection onto the first factor. Continuity of x at t_0 implies that one can choose $b \in (t_0, t_1)$, such that graph$(x|[t_0, b])$ is a relatively compact subanalytic subset of $\mathbb{R} \times U$. According to lemma 3.4 of [15] we have that graph$(x|[t_0, b])$ is a finite union of connected smooth semianalytic subsets A such that, for each A, rank$(p_1|A)$ is constant. It is not difficult to see that each A has dimension 0 or 1 and that rank$(p_1|A)$ is 0 or 1. Moreover, we can easily see that there must be an A, say $A = A_0$, such that rank$(p_1|A_0) = 1$, $p_1(A_0) = (t_0, t_2)$, for some $t_2 \in (t_0, b]$, and $p_1(\bar{A}_0) = [t_0, t_2]$. From this we can easily deduce that $x|(t_0, t_2)$ is an as-curve and moreover, using lemma 5.4, (b), that $x|[t_0, t_2]$ is an as-curve. The rest of the proof can be performed in a similar way. □

Lemma 5.6. *(a) Let $x : [t_0, t_1] \to M, x : [t_0, t_1) \to M$ or $x : (t_0, t_1] \to M$ be an as-curve in M. Then $x|(t_0, t_1)$ is an as-curve in M.*

(b) Let $x : (t_0, t_1) \to M$ be an as-curve in M and assume that there is a continuous extension $\bar{x} : [t_0, t_1] \to M, \bar{x} : [t_0, t_1) \to M$ or $\bar{x} : (t_0, t_1] \to M$. Then $\bar{x}$ is an as-curve in M.

(c) Let $f : M \to N$ be a given analytic map. Let $x(t), t \in [t_0, t_1)$ ($t \in (t_0, t_1], t \in [t_0, t_1], t \in (t_0, t_1)$) be an as-curve in M. Then $f(x(t)), t \in [\bar{t}_0, \bar{t}_1]$, is an as-curve in N, for each $[\bar{t}_0, \bar{t}_1] \subseteq [t_0, t_1)$ (respectively, $[\bar{t}_0, \bar{t}_1] \subseteq (t_0, t_1]$, $[\bar{t}_0, \bar{t}_1] \subseteq [t_0, t_1]$, $[\bar{t}_0, \bar{t}_1] \subseteq (t_0, t_1)$). If graph x *is relatively compact then $f(x(t)), t \in [t_0, t_1)$, ($t \in (t_0, t_1], t \in [t_0, t_1], t \in (t_0, t_1)$) is an as-curve and* graph$(x \circ f)$ *is relatively compact.*

Proof. Part (a) is easy to prove using the definitions. Part (b) follows easily using corollary 2.8 of [15]. To prove (c) observe first that $f \circ x : (t_0, t_1) \to N$ is an analytic map. Since graph$(x|[\bar{t}_0, \bar{t}_1])$ is a semianalytic compact subset of dimension 1 of $\mathbb{R} \times M$ we have that graph$(f \circ x|[\bar{t}_0, \bar{t}_1]) = (1_{\mathbb{R}} \times f)($graph$(x|[\bar{t}_0, \bar{t}_1])$, taking into account that $1_{\mathbb{R}} \times f : \mathbb{R} \times M \to \mathbb{R} \times N$ is an analytic map, is also semianalytic compact and of dimension 1, and a similar proof can be given for the case of the intervals $(t_0, t_1], [t_0, t_1], (t_0, t_1)$. If graph x is relatively compact then according to lemma 5.2 we have a continuous extension $\bar{x}$ which is an as-curve and therefore we can apply the first part of (c) to this extension. The rest of the proof follows easily. □

Lemma 5.7. *Let N be a given manifold and let $x : [t_0, t_1) \to N$ be an as-curve in N which is not a constant. Then there exist $t_2 \in (t_0, t_1)$ such that $x((t_0, t_2))$ is an analytic submanifold which is also a semianalytic subset, $x|(t_0, t_2) : (t_0, t_2) \to x((t_0, t_2))$ is an analytic diffeomorphism and $x|[t_0, t_2] : [t_0, t_2] \to x([t_0, t_1])$ is an homeomorphism. Similar results hold for an as-curve $x : (t_0, t_1] \to N$.*

Proof. Let $t_0 < \bar{t}_1 < t_1$, then we have that graph$(x|[t_0, \bar{t}_1])$ is a compact semianalytic subset of $\mathbb{R} \times N$ of dimension 1. Let $x_i, i = 1, \ldots, n$, be local analytic coordinates centred at $x(t_0)$ therefore $x_i(t_0) = 0$ for $i = 1, \ldots, n$. Without loss of generality we can assume that $x_i(t)$ is defined for all $i = 1, \ldots, n$ and all $t \in [t_0, \bar{t}_1]$. For some index, say $j \in \{1, \ldots, n\}$, we must have that $x_j(t)$ is not a constant. We are going to show that there exists $t_2 \in (t_0, \bar{t}_1]$ such that $x_j : [t_0, t_2] \to \mathbb{R}$ satisfies certain conditions from which the lemma follows. Since $x_j(t)$ is the projection on the j-coordinate axis of the curve $x|[t_0, \bar{t}_1]$, that is, $x_j = p_j \circ x$, using lemma 5.6(*c*) we have that $x_j|[t_0, \bar{t}_1]$ is an as-curve in $\mathbb{R}$ and, moreover, graph$(x_j|[t_0, \bar{t}_1]) \subseteq \mathbb{R} \times \mathbb{R}$ is a semianalytic compact subset of dimension 1. The restriction to graph x_j of the projection $p_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ onto the second factor satisfies $p_2(t, x_j(t)) = x_j(t)$, for all $t \in [t_0, \bar{t}_1]$. Then according to lemma 3.4 of [15] graph$(x_j|[t_0, \bar{t}_1])$ is a finite union of connected smooth semianalytic subsets A such that rank$(p_2|A)$ is constant on A, and it is easy to show that rank$(p_2|A)$ can take only the values 0 or 1. It is also easy to see that for at least one such A one must have that rank$(p_2|A) = 1$, $p_2(A) = x_j((t_0, t_2))$ for some $t_2 \in (t_0, \bar{t}_1]$ and $\overline{p_2(A)} = p_2(\overline{A}) = x_j([t_0, t_2])$, and therefore that $x_j((t_0, t_2))$ is an open interval. Moreover $x_j|[t_0, t_2]$ is injective, and we have that $x_j^{-1} : x_j((t_0, t_2)) \to (t_0, t_2)$ is analytic and also its graph is a semianalytic subset of $\mathbb{R} \times \mathbb{R}$ and then, because of lemma 5.4(*b*), that there is an extension $x_j^{-1} : x_j([t_0, t_2]) \to [t_0, t_2]$ which is continuous, and therefore $x_j|[t_0, t_2] : [t_0, t_2] \to x_j([t_0, t_1])$ is an homeomorphism. Let $t(s), s \in [x_j(t_0), x_j(t_2)]$, be the map x_j^{-1}. Then we have that $x([t_0, t_2]) = \{(x_1, \ldots, x_n) : x_i = x_i(t(s)), s \in [x_j(t_0), x_j(t_1)], i = 1, \ldots, n\}$ and also $x((t_0, t_2)) = \{(x_1, \ldots, x_n) : x_i = x_i(t(s)), s \in (x_j(t_0), x_j(t_1)), i = 1, \ldots, n\}$. Using this we can easily deduce the assertion of the lemma for the case of an as-curve $x : [t_0, t_1) \to N$. The case of an as-curve $x : (t_0, t_1] \to N$ can be proven in an entirely similar way. □

Lemma 5.8. *Let $x(t), t \in [t_0, t_1)$ be an as-curve in N, which is not a constant. Then there is a $t_2 \in (t_0, t_1)$ such that $x([t_0, t_2)) - \{x(t_0)\}$ is nonempty and locally connected at $x(t_0)$, more precisely, $x((t_0, t_3))$ is a neighbourhood of $x(t_0)$ in $x([t_0, t_2)) - \{x(t_0)\}$, for all $t_3 \in (t_0, t_2)$. Moreover, t_2 can be chosen such that $x : [t_0, t_2] \to x([t_0, t_2])$ is an homeomorphism, $x((t_0, t_2))$*

is an analytic submanifold which is a semianalytic subset of N and $x : (t_0, t_2) \to x((t_0, t_2))$ *is an analytic diffeomorphism. A similar result holds for an as-curve* $x(t)$, $t \in (t_0, t_1]$, *in M.*

Proof. We can show using lemma 5.7 that there exists a $t_2 \in (t_0, t_1)$ such that $x|[t_0, t_2]$ is injective, and that $x|[t_0, t_2]$ is an homeomorphism onto its image and that $x : (t_0, t_2) \to x((t_0, t_2))$ is an analytic diffeomorphism where $x((t_0, t_2))$ is an analytic submanifold. In particular, we have that $x(t) \neq x(t_0)$ for all $t \in (t_0, t_2]$. Let $r > 0$ small be given. Working in local analytic coordinates centred at $x(t_0)$ we can show that continuity of $x(t)$ implies that there exists $t_r \in (t_0, t_2]$ such that $x([t_0, t_r]) \subseteq B_r(x(t_0))$. It can be easily shown that $x([t_0, t_r)) - \{x(t_0)\} = x((t_0, t_r))$ is connected. Moreover, for each $s \in [t_0, t_r)$ there is an open ball $B_\delta(x(s)) \subseteq B_r(x(t_0))$, with $\delta = \delta(s)$, such that $B_\delta(x(s)) \cap x([t_r, t_2]) = \emptyset$. Then the open set

$$W = \bigcup_{s \in [t_0, t_r)} B_\delta(x(s))$$

satisfies $W \subseteq B_r(x(t_0))$ and $W \cap (x([t_0, t_2)) - \{x(t_0)\}) = x((t_0, t_r))$. This shows that $x([t_0, t_2)) - \{x(t_0)\}$ is nonempty and locally connected at $x(t_0)$ and also that $x((t_0, t_r))$ is a neighbourhood of $x(t_0)$ in $x([t_0, t_2)) - \{x(t_0)\}$. Now for each $t_3 \in (t_0, t_r]$ take

$$W_{t_3} = \bigcup_{s \in [t_0, t_3)} B_{\bar{\delta}}(x(s))$$

where $\bar{\delta} = \bar{\delta}(t_3, s)$ and $B_{\bar{\delta}}(x(s))$ satisfies $B_{\bar{\delta}}(x(s)) \subseteq B_r(x(0))$ and $x([t_3, t_2]) \cap B_{\bar{\delta}}(x(s)) = \emptyset$. Then $x((t_0, t_3)) = W_{t_3} \cap (x([t_0, t_r)) - \{x(t_0)\})$. This shows that $x((t_0, t_3))$ is a neighbourhood of $x(t_0)$ in $x([t_0, t_r)) - \{x(t_0)\}$. The proof of this part of the lemma follows by lowering the value of t_2, namely, by taking $t_2 := t_r$. The case of an as-curve $x(t)$, $t \in (t_0, t_1]$, in M can be proven in an entirely similar way. □

Inspired by lemma 6.3 of [15] we will prove the following result about the image and reparametrization of an as-curve.

Lemma 5.9. *Let* $x(t) \in M$, $t \in [t_0, t_1)$, *be an as-curve in M. Then there is an as-curve* $z(s)$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, *in M, for some* $\delta_1, \delta_2 > 0$ *such that* $z(s_0) = x(t_0)$, $z([s_0, s_0 + \delta_2)) = x([t_0, t_2))$ *for some* $t_2 \in (0, t_1)$ *and we also have that* t_2 *and* $s_0 + \delta_2$ *are such that* $x|[t_0, t_2)$ *and* $z|[s_0, s_0 + \delta_2)$ *are homeomorphisms onto* $z([s_0, s_0 + \delta_2)) = x([t_0, t_2))$, $x|(t_0, t_2)$ *and* $z|(s_0, s_0 + \delta_2)$ *are analytic diffeomorphisms onto* $z((s_0, s_0 + \delta_2)) = x((t_0, t_2))$ *which is an analytic submanifold which is also a semianalytic subset. Moreover,* $(x|(t_0, t_2))^{-1} \circ (z|(s_0, s_0 + \delta_2)) : (s_0, s_0 + \delta_2) \to \mathbb{R}$ *is an as-curve in* $\mathbb{R}$ *which is an analytic diffeomorphism onto its image* (t_0, t_2). *Also,* t_2 *can be chosen such that for each* $t_3 \in (t_0, t_2]$, $x((t_0, t_3))$ *is a neighbourhood of* $x(t_0)$ *in* $x([t_0, t_2)) - \{x(t_0)\}$. *Similar results hold for as-curves* $x(t)$ *in M, where* $t \in (t_0, t_1]$.

Proof. Using lemma 5.8 and also lemma 6.3 of [15], we can conclude that there is an as-curve $z(s)$, $s \in (s_0 - \delta_1, s_0 + \delta_1)$, in M, for some $\delta_1, \delta_2 > 0$ such that $z(s_0) = x(t_0)$, $z([s_0, s_0 + \delta_2)) = x([t_0, t_2))$ for some $t_2 \in (0, t_1)$ satisfying all the conditions stated in lemma 5.8. Since we can also apply lemmas 5.5 and 5.8 to $z(s)$ we can also deduce that t_2 and δ_2 can be chosen such that $x|[t_0, t_2)$ and $z|[s_0, s_0 + \delta_2)$ are homeomorphisms onto $z([s_0, s_0 + \delta_2)) = x([t_0, t_2))$, $x|(t_0, t_2)$ and $z|(s_0, s_0 + \delta_2)$ are analytic diffeomorphisms onto $z((s_0, s_0 + \delta_2)) = x((t_0, t_2))$ which is an analytic submanifold which is also a semianalytic subset. Moreover, $(x|(t_0, t_2))^{-1} \circ (z|(s_0, s_0 + \delta_2)) : (s_0, s_0 + \delta_2) \to \mathbb{R}$ is an as-curve in $\mathbb{R}$ which is an analytic diffeomorphism onto its image (t_0, t_2). The rest of the proof can be performed in a similar way. □

5.2. Solutions to IDE: as-solutions

Now we introduce the notion of solution to a given IDE which is convenient for the purposes of the present paper.

Definition 5.10. *An as-solution $x(t), t \in [t_0, t_1)$, in M to a given IDE (a, f) in M is an as-curve in M which satisfies (a, f) for all $t \in (t_0, t_1)$, that is, $a(x(t))\dot{x}(t) = f(x(t))$, for all $t \in (t_0, t_1)$. Similar statements hold for as-solutions $x(t)$ in M where $t \in (t_0, t_1], t \in (t_0, t_1)$ or $t \in [t_0, t_1]$ with $t_0 \neq t_1$.*

5.3. Lifted and projected solutions

We have the following result.

Lemma 5.11. *(a) Let $y(t), t \in [t_0, t_1), (t \in (t_0, t_1],)$ be a given as-solution in M^k to the system (a_k, f_k) described in section 3, for some $k = 1, 2, \ldots, q$. Then for each $t_2 \in (t_0, t_1)$, $y(t)$ is projected via π_{k-1} into an as-solution $x(t)$ to the system (a_{k-1}, f_{k-1}), $x(t) = \pi_{k-1}(y(t)), t \in [t_0, t_2]$ (respectively, $t \in [t_2, t_1]$), in M^{k-1}.*

(b) Assume that we have an as-solution $y(t), t \in [t_0, t_1)$, $(t \in (t_0, t_1])$, in M^k, for some $k = 1, 2, \ldots, q$, to the system (a_k, f_k) described in section 3. Then for each $s = 0, \ldots, k-1$, $y(t)$ is projected via $\pi_s \circ \cdots \circ \pi_{k-1}$ into an as-solution $x(t)$ to the system (a_s, f_s), $x(t) = \pi_s \circ \cdots \circ \pi_{k-1}(y(t)), t \in [t_0, t_2]$ (respectively, $t \in [t_2, t_1],$), in M^s, for each $t_2 \in (t_0, t_1)$.

Proof. Part (a) is easy to prove using the fact that π_{k-1} is an analytic map, and also lemma 5.6(c). Part (b) follows using (a). □

We can deduce, from the previous lemma, that if $y(t), t \in [t_0, t_1)$ is an as-solution to (a_k, f_k) in M^k, for some $k = 1, \ldots, q$ and $x(t) = \pi_s \circ, \ldots, \pi_k(y(t))$ is an as-curve in M^s, for some $s = 1, \ldots, k-1$, then $x(t), t \in [t_0, t_2]$ is an as-solution to (a_s, f_s) for each $t_2 \in (t_0, t_1)$. We will call $y(t)$ a *lifted as-solution of* $x(t)$, and $x(t)$ the *projected as-solution of* $y(t)$. A similar definition holds for the case of as-solutions $y(t), t \in (t_0, t_1], t \in (t_0, t_1), t \in [t_0, t_1]$.

Now we shall state and prove one of our main results.

Theorem 5.12. *(a) Let $y(t), t \in [t_0, t_1)$ (respectively, $t \in (t_0, t_1]$) be an as-solution to (a_k, f_k) in $M^k, k = 1, \ldots, q$. Then $x(t) = \pi_{k-1}(y(t)), t \in [t_0, t_2]$ (respectively, $t \in [t_2, t_1]$) is an as-solution to (a_{k-1}, f_{k-1}) in M^{k-1}, for each $t_2 \in (t_0, t_1)$.*

(b) If $x(t), t \in [t_0, t_1)$ (respectively, $t \in (t_0, t_1]$) is an as-solution to (a_{k-1}, f_{k-1}) in M^{k-1} such that $x(t) \in M_0^{k-1}, t \in [t_0, t_1)$ (respectively, $\in (t_0, t_1]$), $k = 1, \ldots, q$ then there exists $t_2 \in (t_0, t_1)$ and a lifted as-solution $y(t), t \in [t_0, t_2]$ (respectively, $t \in [t_2, t_1]$) of $x|[t_0, t_2]$ (respectively, $x|[t_2, t_1]$) to (a_k, f_k) in M^k. In particular, $x(t) = \pi_{k-1}(y(t)), t \in [t_0, t_2]$ (respectively, $t \in [t_2, t_1]$.)

Proof. We are going to give a detailed proof of the case $k = 1$ only, since the cases $k = 2, \ldots, q$ can be proven in an entirely similar way. Part (a) is an easy consequence of lemma 5.11. In order to prove (b) we are going to prove first several facts, (i), (ii), (iii), (iv) and (v), below. These facts will be proven under the assumption that $x(t) \in M_0, t \in [t_0, t_1]$, is an as-solution to (a, f) in M, the curve $x(t)$ is simple, that is, $x(a) \neq x(b)$ for all $a, b \in [t_0, t_1]$ such that $a \neq b$, and moreover, $x : [t_0, t_1] \to x([t_0, t_1])$ is an homeomorphism and also $x((t_0, t_1))$ is an analytic submanifold and $x|(t_0, t_1) : (t_0, t_1) \to x((t_0, t_1))$ is an analytic diffeomorphism. We can assume without loss of generality (for instance, using Whitney embedding theorem) that $M \subseteq U$ and $M^1 \subseteq V$ are analytic submanifolds of U and V, where U and V are real finite-dimensional vector spaces.

(i) The map π_0 can be described as the restriction $p|(\text{graph}\,\pi_0)$ to graph $\pi_0 \subseteq V \times U$ of the projection onto the second factor $p: V \times U \to U$. Since $p|(\text{graph}\,\pi_0)$ is a proper analytic map we have that $(p|(\text{graph}\,\pi_0))^{-1}(x([t_0, t_1])) \subseteq \text{graph}\,\pi_0$ is a compact semianalytic subset of $V \times U$, therefore, according to lemma 3.4 of [15] it is a finite union of relatively compact connected smooth semianalytic subsets A such that for each A rank$(p|A)$ is constant on A. Since $\dim(x([t_0, t_1]) = 1$, it is easy to see that for each A rank$(p|A)$ is 0 or 1, and moreover, $x([t_0, t_1])$ is the union of those $p(\bar{A})$ such that $p|A$ has rank 1 and therefore $p(\bar{A})$ is not a point.

(ii) Let $\alpha \in [t_0, t_1]$ be fixed. Then there is an A, say $A = A_0$, such that $x(\alpha) \in p(\bar{A}_0)$. Since $\bar{A}_0$ is connected and compact and the curve $x(t)$, $t \in [t_0, t_1]$, is simple, we have that $p(\bar{A}_0)$ is homeomorphic to a closed interval (possibly of zero length) via the curve x, say $p(\bar{A}_0) = x([a_1, a_2])$, where $[a_1, a_2] \subseteq [t_0, t_1]$. We can assume without loss of generality that $p(\bar{A}_0)$ is homeomorphic to a closed interval of nonzero length.

(iii) By using theorem 6.10 of [15] we can see that for given points $q_i \in p^{-1}(x(a_i)) \cap \bar{A}_0$, $i = 1, 2$, we have $q_1 \neq q_2$, and there is a continuous semianalytic curve $w(s)$, $s \in [s_0, s_1]$, in $\bar{A}_0$ such that $w(s_i) = q_{i+1}$, $i = 0, 1$. Then $w([s_0, s_1])$ is a compact semianalytic subset of $V \times U$ of dimension 1 and we have $p(w([s_0, s_1])) = x([a_1, a_2])]$. Using lemma 3.4 of [15] we see that since $w([s_0, s_1])$ has dimension 1 it is a finite union of relatively compact connected smooth semianalytic subsets B of dimension less or equal than 1, such that the restriction of the projection $p|B$ has locally constant rank of value 0 or 1. Since each $\bar{B}$ is connected and compact we have that $p(\bar{B})$ is homeomorphic to a closed interval, say $p(\bar{B}) = x([a_B, b_B])$. For at least some B such that $p(\bar{B})$ is not a point one must have that $p(\bar{B})$ contains the point $x(\alpha)$. Observe that $p(B) = x((a_B, b_B))$ is an analytic submanifold which is a subanalytic subset and that $p|B : B \to x((a_B, b_B))$ is an analytic diffeomorphism. This gives, in particular, a parametrization of the analytic submanifold B, namely, $z(t) = (p|B)^{-1}(x(t))$, $t \in (a_B, b_B)$. We can give a definition of graph$((p|B)^{-1} \circ x)$ as a subanalytic subset of $\mathbb{R} \times M^1 \times M$ as follows. We have graph$((p|B)^{-1} \circ x) = \{(t, z) \in \mathbb{R} \times M^1 \times M : z \in B, p(z) = x, (t, x) \in \text{graph}(x|(a_B, b_B))\}$, which defines graph$((p|B)^{-1} \circ x)$ by subanalytic conditions, since B is a semianalytic subset of $M^1 \times M$, graph x is a subanalytic subset of $\mathbb{R} \times M$, by definition, and graph$(x|(a_B, b_B)) = \{(t, x) \in \mathbb{R} \times M : (t, x) \in \text{graph}\, x, t \in (a_B, b_B)\}$. Then using lemma 5.2(a) we can deduce that there is a uniquely defined continuous extension, which we will call z by a slight abuse of notation, $z(t)$, $t \in [a_B, b_B]$, whose image is the semianalytic subset $\bar{B} = z([a_B, b_B])$, which is an as-curve in $M^1 \times M$. We have, in particular, that $z(t)$, $t \in [a_B, b_B]$, is an as-curve such that $p(z(t)) = x(t)$, $t \in [a_B, b_B]$. It is clear that the extension $z(t)$, $t \in [a_B, b_B]$, is given by $(p|\bar{B})^{-1} \circ x$.

(iv) Assume that $x(\alpha) \in p(B)$. We have the as-curve $z(t) = (y(t), x(t))$, $t \in [a_B, b_B]$, in graph(π_0), therefore $p(z(t)) = x(t)$, for $t \in [a_B, b_B]$, then the curve $y(t)$ satisfies $\pi_0(y(t)) = x(t)$, $t \in [a_B, b_B]$. Using this it becomes clear from the definition of (a_1, f_1) that $y(t)$ satisfies the system (a_1, f_1), for $t \in (a_B, b_B)$. We have that, for any $[\alpha - \epsilon_1, \alpha + \epsilon_2] \subseteq [a_B, b_B]$, where $\epsilon_1, \epsilon_2 \geqslant 0$, the curve $y(t)$, $t \in [\alpha - \epsilon_1, \alpha + \epsilon_2]$, is an as-curve in M^1. In fact, this is a direct consequence of lemma 5.6(c), since the projection $q : M^1 \times M \to M^1$ is an analytic map. It is clear that $y(t)$, $t \in (\alpha - \epsilon_1, \alpha + \epsilon_2)$ satisfies the system (a_1, f_1).

(v) Assume now that $x(\alpha) \in p(\bar{B}) - p(B)$, then $\alpha = a_B$ or $\alpha = b_B$. If $\alpha = a_B$ (respectively $\alpha = b_B$) we can proceed in a similar way as we did in (iv) and we have an as-curve

$z(t) = (y(t), x(t)), t \in [\alpha, \alpha + \epsilon_2]$ (respectively, $t \in [\alpha - \epsilon_1, \alpha]$) in graph($\pi_0$), then, in particular, $p(z(t)) = x(t), t \in [\alpha, \alpha + \epsilon_2]$ (respectively, $t \in [\alpha - \epsilon_1, \alpha)$]. Then the curve $y(t) = q(z(t))$ satisfies $\pi_0(y(t)) = x(t), t \in [\alpha, \alpha + \epsilon_2]$ (respectively, $t \in [\alpha - \epsilon_1, \alpha]$) and is an as-curve. It is clear that $y(t), t \in (\alpha, \alpha + \epsilon_2)$ (respectively $t \in (\alpha - \epsilon_1, \alpha)$), satisfies the system (a_1, f_1).

We are going to prove (b). If $x(t) = x(t_0)$ is a constant then it can be lifted to a constant curve $y(t) = y(t_0)$, where $y(t_0) \in \pi_0^{-1}(x(t_0))$, which solves the problem, so let us assume that $x(t)$ is not a constant. By conveniently lowering the value of t_1 we can assume without loss of generality that $x(t) \in M_0, t \in [t_0, t_1]$, is an as-solution to (a, f) in M. Moreover, by lowering the value of t_1 if necessary and using lemma 5.7 we can assume that the curve $x(t)$ is simple, that is $x(a) \neq x(b)$ for all $a, b \in [t_0, t_1]$ such that $a \neq b$, and moreover, that $x : [t_0, t_1] \to x([t_0, t_1])$ is an homeomorphism onto $x([t_0, t_1]), x((t_0, t_1))$ is an analytic submanifold of M and $x|(t_0, t_1) : (t_0, t_1) \to x((t_0, t_1))$ is an analytic diffeomorphism. By using (v) with $\alpha = t_0$, we must have $a_B = t_0$ and then the proof of (b) follows by taking $\alpha + \epsilon_2 = t_2$. The case of an interval $(t_0, t_1]$ can be proved in an entirely similar way. □

Theorem 5.12 proves that an as-solution $x(t), t \in [t_0, t_1)$, to a given IDE (a, f) with domain M and range F can be lifted to an as-solution $y(t), t \in [t_0, t_2]$, in $\tilde{M}_2$ to the lifted system $(\tilde{a}_2, \tilde{f}_2)$, for some $t_2 \in (t_0, t_1)$, and a similar result holds for as-solutions of the type $x(t), t \in (t_0, t_1]$.

5.4. *Normal as-solutions*

Let $x(t), t \in [\alpha, \beta]$ be an as-solution to (a_k, f_k) in M^k, for some $k = 0, \ldots, q - 1$. If for some $s_0 \in [\alpha, \beta]$ we have $x(s_0) \in M_2^k$ then we must have that there is at most a finite number of $t \in [\alpha, \beta]$, say $\alpha \leqslant t_1 < \cdots < t_r \leqslant \beta$, such that $x(t_i) \in M_0^k$. This is because M_0^k is defined by analytic equations. In this case we will call x a *normal as-solution to* (a_k, f_k) *in* M^k. It is clear that if an as-solution $x(t), t \in [\alpha, \beta]$ to (a, f) in M^k is not normal then $x(t) \in M_0^k$ for all $t \in [\alpha, \beta]$.

The following corollary can be obtained by repeated application of part (b) of theorem 5.12.

Corollary 5.13. *Let* $x(t), t \in [t_0, t_1)$ *(respectively,* $t \in (t_0, t_1]$*) be an as-solution to* (a_{k-1}, f_{k-1}) *in* M^{k-1} *such that* $x(t) \in M_0^{k-1}, t \in [t_0, t_1)$ *(respectively,* $t \in (t_0, t_1]$*),* $k = 1, \ldots, q$. *Then there exists* $t_2 \in (t_0, t_1)$ *and a lifted normal as-solution* $y(t), t \in [t_0, t_2]$ *(respectively,* $t \in [t_2, t_1]$*) of* $x|[t_0, t_2]$ *(respectively,* $x|[t_2, t_1]$*) to* (a_r, f_r) *in* M^r*, for some* $r \geqslant k$*, in particular,* $x(t) = \pi_{k-1} \circ \cdots \pi_{r-1}(y(t)), t \in [t_0, t_2]$ *(respectively,* $t \in [t_2, t_1]$*.)*

5.5. *Reparametrization and extension of solutions*

First we shall define the notion of a reparametrization in the context of as-curves. A *reparametrization* of an as-curve $x(t), t \in [t_0, t_1)$, in M, is a change of variables $t = \tau(s), s \in [s_0, s_1), (s \in (s_1, s_0])$ such that $\tau : [s_0, s_1) \to \mathbb{R}$ (respectively, $\tau : (s_1, s_0] \to \mathbb{R}$) is an as-curve in $\mathbb{R}$ which is also an homeomorphism onto $[t_0, t_1)$ such that $\tau(s_0) = t_0$ and $\tau|(s_0, s_1) : (s_0, s_1) \to (t_0, t_1)$ (respectively, $\tau : (s_1, s_0) \to (t_0, t_1)$) is an analytic diffeomorphism. It is easy to prove that in this case the composition $(x \circ \tau)(s), s \in [s_0, s_1)$ (respectively, $s \in (s_1, s_0]$) is an as-curve in M. Similar definitions and results hold for the case of as-curves $x(t)$ whose domain is an interval of the type $(t_0, t_1], (t_0, t_1), [t_0, t_1]$.

Let (a, f) be a given IDE with domain M and range F. By definition, (a, f) is *homogeneous* if $f = 0$. It is clear that if $x(t), t \in [t_0, t_1)$ is a given as-solution to an homogeneous system $(a, 0)$ and $t = \tau(s), s \in [s_0, s_1)$ $(s \in (s_1, s_0])$ is a given reparametrization then the curve $y(s) \equiv x \circ \tau(s), s \in [s_0, s_1)$ (respectively, $s \in (s_1, s_0]$) is also an as-solution to $(a, 0)$. More generally, if (a, f) is not necessarily homogeneous then $y(s)$ satisfies $a(y(s))\dot{y}(s) = (\mathrm{d}\tau/\mathrm{d}s) f(y(s)), s \in (s_0, s_1)$(*respectively*, $s \in (s_1, s_0)$) or, using a different and also standard notation, $a(y(s))\dot{y}(s) = \dot{t}(s) f(y(s)), s \in (s_0, s_1)$ (*respectively*, $s \in (s_1, s_0)$). Similar results hold for the case of as-curves whose domain is an interval of the type $(t_0, t_1]$, (t_0, t_1), or $[t_0, t_1]$.

Theorem 5.14. *(a) Let $(a, 0)$ be a given homogeneous IDE with domain M and range F. Let $x_+(t), t \in [t_0, t_1)$, be an as-solution to $(a, 0)$ in M, which is not a constant. Then there is an as-solution $z(s), s \in (s_0 - \delta_1, s_0 + \delta_2)$, in M, for some $\delta_1, \delta_2 > 0$, satisfying all the conditions stated in lemma 5.9 with respect to $x_+(t), t \in [t_0, t_1)$.*

(b) Let (a, f) be a given IDE with domain M and range F. Let $x_+(t), t \in [t_0, t_1)$, be an as-solution to (a, f) in M which is not a constant. Then there is an as-curve $z(s), s \in (s_0 - \delta_1, s_0 + \delta_2)$, in M, for some $\delta_1, \delta_2 > 0$, not necessarily a solution, satisfying all the conditions stated in lemma 5.9, as we have explained in (a), and, besides, the as-curve $z(s), s \in [s_0, , s_0 + \delta_2)$, satisfies the equation $a(z(s))\dot{z}(s) = \dot{t}(s) f(z(s)), s \in (s_0, s_0 + \delta_2)$, with $\dot{t}(s) > 0, s \in (s_0, s_1)$. Similar results hold for as-solutions of the type $x_-(t), t \in (t_0, t_1]$.

(c) Let (a, f) be a given IDE with domain M and range F. Let $x_+(t), t \in [t_0, t_1)$, be an as-solution to (a, f) in M which is not a constant and let $z(s)$ be as in (b). Then by conveniently increasing the values of δ_1 and δ_2 if necessary, we have the following. There is an as-solution $x_-(t), t \in (t_{2-}, t_0]$ to some of the systems $(a, \pm f)$ in M such that $x_-(t_0) = x_+(t_0) = z(s_0)$, and $x_-|(t_{2-}, t_0]$ and $z|(s_0 - \delta_1, s_0]$ are homeomorphisms onto $z((s_0 - \delta_1, s_0]) = x_-((t_{2-}, t_0])$, $x_-|(t_{2-}, t_0)$ and $z|(s_0 - \delta_1, s_0)$ are analytic diffeomorphisms onto $z((s_0 - \delta_1, s_0)) = x_-((t_{2-}, t_0))$, which is an analytic submanifold which is also a semianalytic subset. Moreover, $(x_-|(t_{2-}, t_0))^{-1} \circ (z|(s_0 - \delta_1, s_0)) : (s_0 - \delta_1, s_0) \to \mathbb{R}$ is an as-curve in $\mathbb{R}$ which is an analytic diffeomorphism onto its image (t_{2-}, t_0). We also have that t_{2-} can be chosen such that for each $t_{3-} \in (t_{2-}, t_0]$, $x_-((t_{3-}, t_0))$ is a neighbourhood of $x_-(t_0)$ in $x_-((t_{2-}, t_0]) - \{x_-(t_0)\}$. The as-curve $z(s), s \in [s_0, s_0 + \delta_2)$, satisfies the equation $a(z(s))\dot{z}(s) = \dot{t}(s) f(z(s)), s \in (s_0, s_0 + \delta_2)$, where $\dot{t}(s) > 0, s \in (s_0, s_0 + \delta_2)$. The as-curve $z(s), s \in (s_0 - \delta_1, s_0]$, satisfies the equation $a(z(s))\dot{z}(s) = \dot{t}(s) f(z(s)), s \in (s_0 - \delta_1, s_0)$, where $\dot{t}(s) > 0$ if $x_-(t)$ satisfies (a, f) and $\dot{t}(s) < 0$ if $x_-(t)$ satisfies $(a, -f)$. Similar results hold for as-solutions $x_-(t), t \in (t_0, t_1]$.

Proof. Parts (a) and (b) are a direct consequence of lemma 5.9. To prove part (c) consider the homogeneous IDE $(\alpha, 0)$ with domain $\mathbb{R} \times M$ and range F where $\alpha(t, x)(\dot{t}, \dot{x}) = a(x)\dot{x} - \dot{t} f(x)$. Consider the solution $(t(s), z(s)), s \in (s_0, s_0 + \delta_2)$ to the homogeneous system $(\alpha, 0)$ where $z(s), s \in (s_0 - \delta_1, s_0 + \delta_2)$ is the as-curve considered in (b). Since we know from (b) that $a(z(s))\dot{z}(s)$ and $f(z(s)))$ are linearly dependent for $s \in (s_0, s_1)$ we can conclude, using the analyticity of $z(s), s \in (s_0 - \delta_1, s_0 + \delta_2)$, that they must also be linearly dependent for $s \in (s_0 - \delta_1, s_0 + \delta_2)$. We can also show, by using lemma 5.8, that there exist δ_1 such that $z((s_0 - \delta_1, s_0]) - \{z(s_0)\}$ is nonempty and locally connected at $z(s_0)$, more precisely, $z((s_0 - \delta_3, s_0])$ is a neighbourhood of $z(s_0)$ in $z((s_0 - \delta_1, s_0]) - \{z(s_0)\}$ for all $s_0 - \delta_3 \in (s_0 - \delta_1, s_0)$. Moreover, δ_1 can be chosen such that $z|[s_0 - \delta_1, s_0] : [s_0 - \delta_1, s_0] \to z([s_0 - \delta_1, s_0])$ is an homeomorphism, $z((s_0 - \delta_1, s_0))$ is an analytic submanifold which is a semianalytic subset of M and $z : (s_0 - \delta_1, s_0) \to z((s_0 - \delta_1, s_0))$ is an analytic diffeomorphism. Since we have a linear dependence between $a(z(s))\dot{z}(s)$ and $f(z(s)), s \in (s_0 - \delta_1, s_0 + \delta_2)$, we have several cases. Assume first that $a(z(s))\dot{z}(s) = 0, s \in (s_0, s_0 + \delta_2)$. Then since $\dot{t}(s) \neq 0, s \in (s_0, s_0 + \delta_2)$ we can conclude that $f(z(s)) = 0, s \in (s_0, s_0 + \delta_2)$. By analyticity of $z(s), s \in (s_0 - \delta_1, s_0 + \delta_2)$ we obtain that $a(z(s))\dot{z}(s) = 0$ and $f(z(s)) = 0, s \in (s_0 - \delta_1, s_0 + \delta_2)$. Then to prove (c) in this case we can simply take $t = s - s_0 + t_0, t_{2-} = -\delta_1 + t_0$ and $x_{2-}(t) = z(t + s_0 - t_0)$.

We can proceed in a similar way if we assume that $f(z(s)) = 0$, $s \in (s_0, s_0 + \delta_2)$. Let us consider now the case where $a(z(s))\dot{z}(s)$, $s \in (s_0, s_0 + \delta_2)$, is not identically 0. We are going to show that, after conveniently diminishing the value of δ_1 if necessary, there exists a unique as-curve in $\mathbb{R}$, $\lambda : (s_0 - \delta_1, s_0) \to \mathbb{R}$, such that $a(z(s))\dot{z}(s) = \lambda(s) f(z(s))$, $s \in (s_0 - \delta_1, s_0)$. First one should take into account that $z(s)$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, is analytic then so are $f(z(s))$ and $a(z(s))\dot{z}(s)$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, and therefore they have at most a finite number of isolated zeros in a neighbourhood of s_0. It is easy to see from the equation $a(z(s))\dot{z}(s) = \lambda(s) f(z(s))$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, that $\lambda(s)$, or rather its extension for complex s, is a meromorphic function in a neighbourhood of s_0. On the other hand, since $\lambda(s) = \dot{t}(s)$, $s \in (s_0, s_0 + \delta_1)$, where $t(s)$ is bounded, we have that $\lambda(s)$ cannot have a pole at s_0 therefore it must be analytic in a neighbourhood of s_0. This implies that by conveniently increasing the values of δ_1 and δ_2 if necessary, we have that $\lambda(s)$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, is real analytic. By conveniently increasing the value of δ_1 if necessary we can assume without loss of generality that $\lambda(s) \neq 0$, $s \in (s_0 - \delta_1, s_0)$. Let us assume first that $\lambda(s) > 0$, $s \in (s_0 - \delta_1, s_0)$. Then the result follows by taking $t = t(s)$, $s \in (s_0 - \delta_1, s_0)$, such that $\mathrm{d}t/\mathrm{d}s = \lambda(s)$, $t(s_0) = t_0$ and $t(s_0 - \delta_1) = t_{2-}$, which defines $t(s)$, $s \in (s_0 - \delta_1, s_0]$ as an as-curve in $\mathbb{R}$ and also t_{2-}. In fact, since $\lambda(s) > 0$, $s \in (s_0 - \delta_1, s_0)$, we have that $t(s)$ is an analytic diffeomorphism from $(s_0 - \delta_1, s_0)$ onto (t_{2-}, t_0) which is an as-curve, while $t : [s_0 - \delta_1, s_0] \to [t_{2-}, t_0]$ is an homeomorphism. Then we can define $x_{2-}(t)$ by $x_{2-}(t) = z(s(t))$ where $s(t)$ is the inverse of $t(s)$. The rest of the proof can be performed in a similar way. □

The previous theorem says, in particular, that an as-solution $x_+(t)$, $t \in [t_0, t_1)$, to a given IDE (a, f) gives rise to an as-solution $x_-(t)$, $t \in (t_{2-}, t_0]$ such that $x_+(t_0) = x_-(t_0)$ and $\operatorname{graph} x_- \cup \operatorname{graph} x_+ = \operatorname{graph} z$ for some as-curve $(z(s), t(s))$, $s \in (s_0 - \delta_1, s_0 + \delta_2)$, which is a solution to the homogeneous system $a(z)\dot{z} = \dot{t} f(z)$, $s \in (s_0 - \delta_1, s_0) \cup (s_0, s_0 + \delta_2)$, and x_-, x_+ are reparametrizations of $z|(s_0 - \delta_1, s_0]$, $z|[s_0, s_0 + \delta_2)$.

References

[1] Abraham R and Marsden J E 1978 *Foundations of Mechanics* (New York: Benjamin)
[2] Marsden J E and Ratiu T S 1994 *Introduction to Mechanics and Symmetry* (Berlin: Springer)
[3] Cendra H, Marsden J E and Ratiu T S 2001 Lagrangian reduction by stages *Memoirs AMS* **152** 722
[4] Bloch A M, Krishnaprasad P S, Marsden J E and Murray R M 1996 Nonholonomic mechanical systems with symmetry *Arc. Ration. Mech. Anal.* **136** 21–99
[5] Cendra H, Marsden J E and Ratiu T S 2001 Geometric mechanics, Lagrangian reduction and nonholonomic systems *Mathematics Unlimited-2001 and Beyond* (Berlin: Springer) pp 221–73
[6] Neimark J I and Fufaev N A 1972 Dynamics of nonholonomic systems *Translations of Mathematical Monographs* vol 33 (Providence, RI: American Mathematical Society)
[7] Cendra H, Ibort A, de León M and Martin de Diego D 2004 A generalized Chetaev's principle for a class of higher order nonholonomic constraints *J. Math. Phys.* **45** 2785–801
[8] Cendra H, Lacomba E A and Reartes W A 2002 The Lagrange D'Alembert–Poincaré equations for the symmetric rolling sphere *Actas del VI Congreso Antonio, A.R Monteiro, Universidad Nacional del Sur, Bahía Blanca* pp 19–32
[9] Cortes Monforte J 2002 *Geometric Control and Numerical Aspects of Nonholonomic Systems* (Berlin: Springer)
[10] Cendra H and Etchechoury M 2006 Rolling of a symmetric sphere on a horizontal plane without sliding or spinning *Rep. Math. Phys.* accepted
[11] Gotay M, Nester J M and Hinds G 1978 Presymplectic manifolds and the Dirac–Bergman theory of constraints *J. Math. Phys.* **19** 2388–99
[12] Rabier P J and Rheinboldt W C 1994 A geometric treatment of implicit differential-algebraic equations *J. Differ. Eqns* **109** 110–46
[13] Etchechoury M 2001 Sistemas diferenciales algebraicos: aplicaciones a perturbacion singular and control *Thesis*
[14] Pritchard F L 2003 On implicit systems of differential equations *J. Differ. Eqns* **194** 328–263
[15] Bierstone E and Milman P D 1988 Semianalytic and subanalytic sets *Publ. Math. l'I.H.E.S.* **67** 5–42

[16] Sussmann H J 1998 Some optimal control applications of real-analytic stratifications and desingularization *Singularities Symposium-Lojasiewicz 70, Banach Center Publ*. vol 44 (Warsaw: Polish. Acad. Sci.)

[17] Hernández-Garduño A, Lawson J K and Marsden J E 2005 Relative equilibria for the generalized rigid body *J. Geom. Phys*. **53** 259–74

[18] Hironaka H 1964 Resolution of an algebraic variety over a field of characteristic zero *Ann. Math*. **79** 109–203

[19] Hauser H 2003 The Hironaka theorem on resolution of singularities *Bull. AMS* **40** 323–403

[20] Encinas S and Villamayor O 2003 A new proof of desingularization over fields of characteristic zero *Rev. Mat. Iber.* **19** 339–53

[21] Bierstone E and Milman P D 1997 Canonical desingularization in characteristic zero by blowing up the maximum strata of a local invariant *Invent. Math*. **128** 207–302

[22] Bierstone E and Milman P D 2003 Desingularization algorithms: I. Role of exceptional divisors *Moscow Math. J*. **3** 751–805

[23] Hironaka H 1973 Subanalytic sets *Number Theory, Algebraic Geometry and Commutative Algebra* (Tokyo: Kinokuniya) pp 453–93

[24] Abraham R, Marsden J E and Ratiu T S 1988 *Manifolds, Tensor Analysis and Applications* (Berlin: Springer)

[25] Lojasiewicz S 1959 Sur le Probléme de la division *Stud. Math*. **18** 87–136

[26] Lojasiewicz S 1964 Triangulation of semianalytic sets *Ann. Scuola. Norm. Sup. Pisa* **18** 449–74

[27] Lojasiewicz S 1970 Sur les ensembles semi-analytiques *Actes du Congress International des Mathematiciens (Nice)* vol 2 pp 237–41

[28] Gabrielov A M 1973 Formal relations between analytic functions *Math. USSR Izvestija* **7** 1056–88

[29] Hardt R M 1975 Stratifications of real analytic mappings and images *Invent. Math*. **28** 193–208

[30] Hardt R M 1977 Triangulation of subanalytic sets and proper light subanalytic maps *Invent. Math*. **38** 207–17

[31] Sussmann H J 1990 Real analytic desingularization and subanalytic sets: an elementary approach *Trans. Am. Math. Soc*. **317** 417–61

[32] Rabier P J and Rheinboldt W C 2000 *Nonholonomic Motion of Rigid Mechanical Systems from a DAE Viewpoint* (Philadelphia, PA: SIAM)

[33] Veselov A P and Veselova L E 1989 Integrable nonholonomic systems on Lie groups *Math. Notes* **44** 810