NORMAL FORMS FOR CONSTRAINED NONLINEAR DIFFERENTIAL EQUATIONS PART I: THEORY

by

L. O. Chua and H. Oka

Memorandum No. UCB/ERL M87/83

24 September 1987

NORMAL FORMS FOR CONSTRAINED NONLINEAR DIFFERENTIAL EQUATIONS PART I: THEORY

by

L. O. Chua and H. Oka

Memorandum No. UCB/ERL M87/83

24 September 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

NORMAL FORMS FOR CONSTRAINED NONLINEAR DIFFERENTIAL EQUATIONS PART I: THEORY

by

L. O. Chua and H. Oka

Memorandum No. UCB/ERL M87/83

24 September 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

NORMAL FORMS FOR CONSTRAINED NONLINEAR DIFFERENTIAL EQUATIONS
PART I: THEORY†

Leon O. Chua and Hiroe Oka††

Abstract

This paper generalizes the theory of *normal forms* for smooth vector fields to *constrained equations* characterized by a system of nonlinear differential-algebraic equations. Such equations are widely encountered in practical circuits and systems when *parasitics* play an important role in the system's qualitative behavior. Such parasitics are called *small parameters* in the associated singular perturbation problem. Our approach in this paper is completely different from the literature on singular perturbation. Ours is based on the general framework described in the tutorial paper by Chua and Kokubu [15], namely, the calculation of *infinitesimal deformations.*

1. INTRODUCTION

We often encounter, especially in *nonlinear* circuit theory, ordinary differential equations of a *singular* type; namely,

$$\left.\begin{aligned} \varepsilon\dot{x} &= f(x,y) \\ \dot{y} &= g(x,y) \end{aligned}\right\} \tag{1.1}$$

where $\cdot$ denotes a derivative with respect to time, $x \in \mathbf{R}^r$, $y \in \mathbf{R}^{n-r}$, and $\varepsilon \in \mathbf{R}$ is a *small parameter* [1-8]. Since we are often interested in the behavior of the *limiting system* as ε tends to zero, we must include the equation for $\varepsilon = 0$ as well as $\varepsilon \neq 0$ in our study. In this paper, we call them *constrained equations*. Therefore, the mathematical object corresponding to such equations constitutes a *larger* set than the set of *vector fields.*

The following Van der Pol equation is a typical example of a constrained equation [1-2]:

$$\left.\begin{aligned} \varepsilon\dot{x} &= (x-x^3/3)+y \\ \dot{y} &= -x \end{aligned}\right\} \tag{1.2}$$

where $x, y \in \mathbf{R}$. The phase portrait, shown in Fig. 1, of this system for small parameter ε is described by a *rapid* motion along the x-direction, and a *slow* motion near the curve $y = x^3/3 - x$, which is obtained by setting $\varepsilon = 0$ in the first expression of (1.2). The name "constrained equation" comes from the observation that the

†This research is supported in part by the Office of Naval Research, Contract N00014-86-K-0351 and by the National Science Foundation, Grant MIP-8614000.

††L. O. Chua is with the University of California, Berkeley, CA 94720.
H. Oka is with the Department of Mathematics, Kyoto University, Kyoto, 606 Japan.

orbits are constrained to lie on the curve $y = \frac{x^3}{3} - x$ *for almost all times.*

In this paper, we will give a *new coordinate-free formulation* for *constrained* equations. One advantage of our formulation is that the *normal forms* associated with these equations can be obtained by essentially the same method developed for vector fields, i.e., when $\varepsilon \neq 0$. Methods for obtaining normal forms for vector fields have been developed by Poincaré, Takens, Arnold, and Ushiki [9-13]. Readers unfamiliar with this subject may consult the recent *tutorial* paper on normal forms for nonlinear vector fields [14-15]. The main purpose of this paper is to show that the *general framework* developed for vector fields in [14] can be successfully applied to constrained equations as well. We will show, among other things, that the normal forms for *constrained equations* give a *local classification* according to the extent of the *degeneracy* of the constrained equation. For the Van der Pol equation we can identify several types of local structures from the phase portrait in Fig. 1, and our normal form theory in this paper will provide a systematic method for classifying such local structures.

There already exist several formulations for constrained equations such as Takens [10-12], Fenichel [16], Sastry, and Desoer [5], Ikegami [7-8], etc. All of them, however, are completely different from our approach in this paper. The main feature of our formulation is that we can consider constrained equations as an extension of vector fields. Because of this generalization, our normal form theory for constrained equations contains that for vector fields. Another advantage of our formulation is that the *perturbation problem* [17-20] associated with constrained equation can also be treated in our formulation. This problem is generally referred to in the literature as the *singular perturbation problem* of ODE's. In this paper, we will present a new point of view on this classic problem.

The outline of this paper is as follows. First, in order to discuss the normal form for constrained equations, we define in *Section 2* an enlarged set of ODE's which includes both the set of smooth vector fields treated in [14] and the set of equation (1.1) for $\varepsilon = 0$. Constrained equations are characterized in this enlarged set in a *coordinate-free* manner. In *Section 3,* we calculate the *infinitesimal deformation* following the general framework of normal forms developed in Chua and Kokubu [14]. Some results with detailed calculations are given in *Section 4.* The final section and a comprehensive Appendix will appear in Part II of this paper. This 2-part paper is based on the theory developed by Oka [22].

2. DEFINITION OF CONSTRAINED EQUATIONS

Let us begin with a heuristic approach for the formulation of constrained equations. Consider the constrained equation (1.1) and rewrite it as follows:

$$\begin{bmatrix} \varepsilon I_r & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & I_{n-r} \end{bmatrix} \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}) \;,\; \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.1}$$

where I_k denotes the unit matrix of order k, and

$$\mathbf{v}(\mathbf{x}) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}. \tag{2.2}$$

Thus we identify (2.1) with the pair $(\mathbf{A}, \mathbf{v})$ which consists of a matrix

$$\mathbf{A} = \begin{bmatrix} \varepsilon I_r & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & I_{n-r} \end{bmatrix} \tag{2.3}$$

and a vector field $\mathbf{v}$.

A normal form of a constrained equation is defined as the *simplest* form among those which can be considered as *equivalent* to the original equation, based on some reasonable definition of equivalence. For example, we may consider that a transformed equation obtained by a coordinate change is equivalent to the original one: Let $\mathbf{x} = \phi(\mathbf{y})$ be the new coordinates, then the transformed equation is given by,

$$\mathbf{A}\dot{x} = \mathbf{A}\, D\phi(\mathbf{y})\,\dot{\mathbf{y}} = \mathbf{v}\left[\phi(\mathbf{y})\right] \tag{2.4}$$

Thus, the transformed constrained equation is of the form

$$\left[\tilde{\mathbf{A}}(\mathbf{y}), \tilde{\mathbf{v}}(\mathbf{y})\right] \triangleq \left[\mathbf{A}\, D\phi(\mathbf{y}), \mathbf{v}\circ\phi(\mathbf{y})\right]. \tag{2.5}$$

Note that the first component $\tilde{\mathbf{A}}(\mathbf{y})$ is no longer a constant matrix. Therefore we will enlarge our abstract objects to include the set of all pairs $\left[\mathbf{A}(\mathbf{x}), \mathbf{v}(\mathbf{x})\right]$ consisting of a non-constant matrix $\mathbf{A}(\mathbf{x})$ and a vector field $\mathbf{v}(\mathbf{x})$.

In order to apply the method of infinitesimal deformations described in [14], it is convenient to define our constrained equations in a coordinate free manner. Therefore we will state all definitions on a *manifold* M. A vector field $\mathbf{v}$ is defined on the manifold M in the usual fashion (see Chua and Kokubu [14]). We may consider $\mathbf{A}$ as a mapping from $\mathbf{x} \in M$ to the set of all matrices. Since our generalization in this paper depends crucially on the theory of *fiber bundles*, a brief review of the necessary mathematical tools is given in the Appendix in *Part II* of this paper.

In this generalization, we identify $\mathbf{A}$ as a section of the *vector bundle* End (TM), where the *endomorphism bundle* End (TM) is a vector bundle on M whose fiber at x consists of all linear maps of T_xM, as depicted in Fig. 2. As explained in *Appendix I* (Part II), $\mathbf{A}$ may also be considered as a bundle endomorphism of TM, as illustrated in Fig. 3. Thus we arrive at the following definition.

Definition 2.1: Generalized vector field

A *generalized vector field* on M is a pair $(\mathbf{A}, \mathbf{v})$ consisting of a bundle endomorphism $\mathbf{A}$ of TM and a vector field $\mathbf{v}$ (see Fig. 3).

Let us define next the concepts of equivalence and the transformation of generalized vector fields. Recall that the generalized vector field $(\mathbf{A}, \mathbf{v})$ is defined by the equation

$$\mathbf{A}(\mathbf{x})\,\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}) \tag{2.6}$$

in terms of local coordinates. Let us multiply a *non-singular* matrix-valued function $\mathbf{P}(\mathbf{x})$ to both sides of (2.6):

$$\mathbf{P}(\mathbf{x})\,\mathbf{A}(\mathbf{x})\,\dot{\mathbf{x}} = \mathbf{P}(\mathbf{x})\,\mathbf{v}(\mathbf{x}) \tag{2.7}$$

The transformed equation (2.7) may be considered to be equivalent to the original equation (2.6).

For example, let $\mathbf{P}(\mathbf{x})$ be the constant matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $\mathbf{P}(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ acts on equation (1.1) as follows:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}. \tag{2.8}$$

Hence, the transformed equation is given by:

$$\left.\begin{aligned} \dot{y} &= g(x,y) \\ \varepsilon\dot{x} &= f(x,y) \end{aligned}\right\}. \tag{2.9}$$

Thus $\mathbf{P}(\mathbf{x})$ in this case corresponds to an interchange of the upper and the lower expressions.

For the generalized vector field (2.5), let us consider the following transformation:

$$\left[D\phi^{-1}\left[\phi(\mathbf{y})\right] \circ \mathbf{A}\left[\phi(\mathbf{y})\right] \circ D\phi(\mathbf{y}), D\phi^{-1}\left[\phi(\mathbf{y})\right] \circ \mathbf{v}\left[\phi(\mathbf{y})\right]\right] \tag{2.10}$$

where $D\phi$ denotes the *derivative* (i.e., Jacobian matrix) of $\mathbf{x} = \phi(\mathbf{y})$. This transformed generalized vector field is equivalent to the generalized vector field

$$\left[\mathbf{A}\left[\phi(\mathbf{y})\right] \circ D\phi(\mathbf{y}), \mathbf{v}\left[\phi(\mathbf{y})\right]\right] \tag{2.11}$$

by multiplying the non-singular matrix $D\phi^{-1}\left[\phi(\mathbf{y})\right]$ from the left to both components of (2.11). Let us formalize the above heuristic consideration to the following coordinate-free definition:

Definition 2.2: Equivalence and Transformation

Let $(\mathbf{A}, \mathbf{v})$ and $(\mathbf{A}', \mathbf{v}')$ be generalized vector fields (g.v.f) on M. We say these (g.v.f.) are *equivalent* if there exist a bundle automorphism P of TM and a diffeomorphism ϕ of M such that

$$(\mathbf{A}', \mathbf{v}') = \left[P \circ T\phi \circ \mathbf{A} \circ (T\phi)^{-1}, P \circ T\phi \circ \mathbf{v} \circ \phi^{-1}\right] \tag{2.12}$$

holds, where $T\phi$ denotes the tangent map of ϕ. The pair (P, ϕ) is called a *transformation* of the g.v.f. We will denote the right-hand side of (2.12) by $(P, \phi)_{\#}(\mathbf{A}, \mathbf{v})$, namely, the transformed g.v.f. of $(\mathbf{A}, \mathbf{v})$ by (P, ϕ).

The set of all transformations (P, ϕ) forms a group, which we denote by G. This fact is shown in *Appendix III* (Part II).

In the case where **A** is a bundle automorphism i.e., A(x) is *invertible, the g.v.f.* $(\mathbf{A}, \mathbf{v})$ is identified with the g.v.f.

$$(\mathbf{A}^{-1}\mathbf{A}, \mathbf{A}^{-1}\mathbf{v}) = (Id, \mathbf{A}^{-1}\mathbf{v}), \tag{2.13}$$

where Id denotes the *identity* bundle automorphism. This in turn can be identified with the vector field

$$\dot{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{x})\mathbf{v}(\mathbf{x}) \tag{2.14}$$

Hence, the set of all g.v.f.'s contains the set of all vector fields $(Id, \mathbf{v})$. Moreover the transformed equation of $(Id, \mathbf{v})$ by (P, ϕ) is expressed by

$$\begin{aligned}(P, \phi)_{\#}(Id, \mathbf{v}) &= (P \circ T\phi \circ Id \circ T\phi^{-1}, P \circ T\phi \circ \mathbf{v} \circ \phi^{-1}) \\ &= (P, P \circ T\phi \circ \mathbf{v} \circ \phi^{-1})\end{aligned} \tag{2.15}$$

which is equivalent to $(Id, T\phi \circ \mathbf{v} \circ \phi^{-1})$, upon applying the transformation $(P^{-1}, id)_{\#}$ to (2.15). It follows that P does not play an essential role in the case of vector fields, and the *restricted* equivalence relation

$$(Id, \mathbf{v}) \sim (Id, T\phi \circ \mathbf{v} \circ \phi^{-1}) \tag{2.16}$$

is the same as the ordinary equivalence relation for vector fields (see Chua and Kokubu [14]). In this sense, the class of g.v.f.'s is an extension of the class of vector fields.

In the following, we will present several examples of g.v.f.'s.

Example 2.3

$$x\dot{x} = -1 \;, \; x \in \mathbf{R} \tag{2.17}$$

By putting $z = x^2$, this equation reduces to $\dot{z} = -2$ whose explicit solution is given by

$$z = x^2 = x_0^2 - 2t \;, \tag{2.18}$$

where x_0 is the initial condition. It follows from (2.18) that all solutions arrive at the origin $x = 0$ in *finite* time, and cannot be extended beyond this time. In other words, (2.17) has an *impasse point* [21] at $x = 0$. Note that there is no solution starting from the origin at $t = 0$. A family of solutions of (2.17) is shown in Fig. 4.

Example 2.4

$$x\dot{x} = -x \;, \; x \in \mathbf{R}\,. \tag{2.19}$$

We can easily solve (2.19) and obtain the family of solutions shown in Fig. 5. In this case, for each initial condition, there exists a solution which can be extended at infinity. Observe, however, that this equation does *not* have a *unique* solution at $x = 0$ because both $x(t) \equiv 0$ for all t and $x(t) = -t$ satisfy (2.19) with $x(0) = 0$.

In the preceding two examples, the existence or uniqueness of solutions is violated at $x = 0$. In general, for any g.v.f. $(\mathbf{A}, \mathbf{v})$, the existence and/or uniqueness of solution breaks down where **A** is *degenerate*.

Moreover, since the bundle endomorphism A(x) for these examples is given by $A(x) = x$, the *rank* of A(x) *varies* with respect to the points $x \in \mathbf{R}$. In contrast, the rank of A(x) does *not* change in (1.1); indeed, the bundle endomorphism $A_\varepsilon(x)$ for (1.1) is given by:

$$\begin{bmatrix} \varepsilon I_r & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & I_{n-r} \end{bmatrix}. \tag{2.20}$$

Note that $A_\varepsilon(x)$ for $\varepsilon \neq 0$ is *not* degenerate; in other words, the equation for $\varepsilon \neq 0$ defines a vector field. Observe that for $\varepsilon = 0$, the *rank* of $A_0(x)$ is n-r *uniformly* with respect to the points in M. For both cases, $A_0(x)$ and $A_\varepsilon(x)(\varepsilon \neq 0)$, the *rank* of A(x) *is constant*. It follows that the set of all g.v.f.'s is slightly larger than what we are interested. Hence, we will characterize our "constrained equations" by restricting our g.v.f.'s to those imbued with the additional condition that A(x) is of *constant rank*. Of course we must deal with these constrained equations on a manifold.

First, let us introduce the *rank map* defined by

$$\begin{aligned} rk : End(TM) &\to \mathbf{N} \\ A_x &\to rank\ A_x\ , \end{aligned} \tag{2.21}$$

where $A_x \in End(TM)$ is a linear map of T_xM and $rank\ A_x \in \mathbf{N}$, where $\mathbf{N}$ denotes the set of all non-negative integers. This map is well-defined; that is, it is independent of the choice of local coordinates because, by a change of coordinates $y = \phi(x)$ of M, A_x is transformed into $\tilde{A}_y = D\phi\left[\phi^{-1}(y)\right]\cdot A_{\phi^{-1}(y)}\cdot D\phi^{-1}(y)$. Hence, the rank is invariant.

We say a linear map of an n-dimensional vector space has a *corank r* if its rank is equal to $n-r$. The inverse image $(rk)^{-1}(n-r)$ of the rank map rk defines a *sub fiber bundle* of $End(TM)$ whose standard fiber is the set of all linear maps of T_xM of corank r. We denote this fiber bundle by $End^{(r)}(TM)$.

A *bundle endomorphism of TM of corank r* is defined as a *section* of the bundle $End^{(r)}(TM)$. In other words, for a bundle endomorphism A of TM of corank r, the rank of $A(x) = A|_{T_xM}$ is a constant equal to n-r, and is independent of $x \in M$.

Definition 2.5: Constrained Systems

A *constrained system of corank r* on M is a pair (A, v) consisting of a bundle endomorphism A of *TM* of corank r and a vector field v on M. When we do not specify the corank, we simply say a *constrained system* on M. The set of all constrained systems (resp.; of corank r) on M is denoted by $\mathcal{C}\mathcal{X}(M)$ (resp.; $\mathcal{C}\mathcal{X}^{(r)}(M)$). Hence

$$\mathcal{CX}(M) = \bigcup_{0 \le r \le \dim M} \mathcal{CX}^{(r)}(M)\,. \tag{2.22}$$

Observe that $\mathcal{CX}(M)$ is a subset of the set of generalized vector fields, since it excludes such elements as Examples 2.3 and 2.4. For any constrained system (A, v) of corank r and any transformation (P, ϕ), the transformed constrained system $(P, \phi)_{\#}\ (\mathrm{A}, \mathrm{v})$ is again a constrained system of corank r. Thus $\mathcal{CX}^{(r)}(M)$, and as a result $\mathcal{CX}(M)$, remains invariant under the action of the transformation group.

It follows from the above definition that equation (1.1) is identified with a family $(\mathrm{A}_\varepsilon, \mathrm{v}_\varepsilon)$ of constrained systems parametrized by ε. Especially when $\varepsilon = 0$, $(\mathrm{A}_0, \mathrm{v}_0)$ is of corank r. Therefore the family $(\mathrm{A}_\varepsilon, \mathrm{v}_\varepsilon)$ can be considered as an *unfolding* of a constrained system of corank r. This fact inspires us to establish in *Part II* of this paper a relationship between the singular perturbation problem for ODE's and the bifurcation problem for constrained systems.

We can also define a *constrained surface* in our formulation.

Definition 2.6: Constrained Surface

Let (A, v) denote a constrained system of corank r on M. The constrained surface S of (A, v) is defined by

$$S = \{x \in M \mid \mathrm{v}(x) \in \mathrm{Im}\,\mathrm{A}(x)\} \tag{2.23}$$

where $\mathrm{Im}\,\mathrm{A}(x)$ denotes the linear subspace of T_xM consisting of all images of the linear map $\mathrm{A}(x)$ of T_xM.

Example 2.6

For the equation (1.2), let us choose $\varepsilon = 0$, then

$$\mathrm{A(x)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \mathrm{v(x)} = \begin{bmatrix} y - \dfrac{x^3}{3} + x \\ -x \end{bmatrix}. \tag{2.24}$$

Since the first component of A(x) maps to 0, it follows that the image of A(x) for (2.24) is the 1-dimensional subspace $x = 0$. Consequently,

$$S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \,\middle|\, y - \frac{x^3}{3} + x = 0 \right\}. \tag{2.25}$$

Observe that this constrained surface coincides with the *slow surface* obtained by putting $\varepsilon = 0$ in the first expression of the equation (1.2).

Example 2.9

Consider the following ODE:

$$\left.\begin{aligned}\varepsilon_1\dot{x} &= -y + xz - (x^2+y^2)x \\ \varepsilon_2\dot{y} &= x - yz - (x^2+y^2)y \\ \dot{z} &= 1\end{aligned}\right\} \tag{2.26}$$

where $(x, y, z) \in \mathbb{R}^3$. This system of equations defines a family of constrained systems $(\mathbf{A}_\varepsilon, \mathbf{v}_\varepsilon)$ on $\mathbb{R}^3$, parametrized by $\varepsilon = (\varepsilon_1, \varepsilon_2)$, where

$$\mathbf{A}_\varepsilon = \begin{bmatrix}\varepsilon_1 & & 0 \\ & \varepsilon_2 & \\ 0 & & 1\end{bmatrix}, \quad \mathbf{v}_\varepsilon = \begin{bmatrix}-y + xz - (x^2+y^2)x \\ x + yz - (x^2+y^2)y \\ 1\end{bmatrix}. \tag{2.27}$$

When $\varepsilon = 0$, $(\mathbf{A}_0, \mathbf{v}_0)$ is of corank 2, whose constrained surface is given by

$$\begin{aligned} S &= \left\{ \begin{bmatrix}x \\ y \\ z\end{bmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{matrix} -y + xz - (x^2+y^2)x = 0 \text{ and} \\ x + yz - (x^2+y^2)y = 0 \end{matrix} \right\} \\ &= \left\{ \begin{bmatrix}x \\ y \\ z\end{bmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{matrix} x = y = 0 \text{ or} \\ z = x^2 + y^2 \end{matrix} \right\} \end{aligned} \tag{2.28}$$

To obtain the right-hand side of (2.28), we first multiply $-y + xz - (x^2+y^2)x$ by x and multiply $x + yz - (x^2+y^2)y$ by y, then adding them to obtain $z(x^2+y^2) - (x^2+y^2)(x^2+y^2) = 0$. It follows that $(x^2+y^2)[z - x^2 + y^2] = 0$, or $x^2 + y^2 = 0$ and $z = x^2 + y^2$. It follows that the constrained surface S for (2.27) consists of the z-axis, and the parabolic surface $z = x^2 + y^2$.

The phase portrait for (2.26) for *small* $\varepsilon_1 = \varepsilon_2 > 0$ is shown in Fig. 6. Here, the *slow* motion occurs along the z-axis, and on the surface $z = x^2 + y^2$. Observe that in the lower half of the z-space, all trajectories spiral rapidly towards the slow surface $x = y = 0$; namely, the z-axis. Conversely, in the upper half of the z-space, all trajectories spiral rapidly away from the z-axis and converge towards the *slow* surface $z = x^2 + y^2$.

3. INFINITESIMAL DEFORMATION

In the preceding section, we introduced two important concepts; namely, the *vector space of objects*, and the *group of transformations*. The structure of the transformation group $G = AUT(TM) \rtimes Diff(M)$ is derived in *Appendix III*, (Part II), where AUT(TM) denotes the set of all bundle automorphisms of TM and Diff(M) denotes the set of all diffeomorphisms of M. Here, let us investigate further the transformation group G and its action on the set of constrained systems (generalized vector fields), henceforth denoted by $\mathcal{G}\mathcal{X}(M)$, so that we will be able to apply the general framework of normal form theory developed in Chua and Kokubu [14]. More precisely, we will study the one-parameter group of transformations in G and its infinitesimal deformation.

To obtain the *exponential mapping* from the set of infinitesimal generators to the transformation group, it is convenient to consider the transformation $g = (P, \phi) \in G$ as a diffeomorphism $P \circ T\phi$ of the tangent bundle TM. The following lemma is needed for this purpose:

Lemma 3.1

The mapping

$$\sigma : G \to Diff(TM), \ (P, \phi) \to P \circ T\phi \tag{3.1}$$

is an *injective* group homomorphism.

Proof. First let us verify that

$$\sigma\left[(P, \phi)\cdot(Q, \psi)\right] = \sigma(P, \phi) \circ \sigma(Q, \psi) . \tag{3.2}$$

Applying the chain rule and the definition of the group operation of G, the left-hand side (l.h.s.) of (3.2) can be written as follow:

$$\begin{aligned} l.h.s. &= \sigma(P \circ T\phi \circ Q \circ T\phi^{-1}, \phi \circ \psi) = P \circ T\phi \circ Q \circ T\phi^{-1} \circ T(\phi \circ \psi) \\ &= P \circ T\phi \circ Q \circ T\phi^{-1} \circ T\phi \circ T\psi = P \circ T\phi \circ Q \circ T\psi \end{aligned} \tag{3.3}$$

Similarly, the right-hand side (r.h.s.) of (3.2) can be written as follows:

$$r.h.s. = (P \circ T\phi) \circ (Q \circ T\psi) = P \circ T\phi \circ Q \circ T\psi \ . \tag{3.4}$$

Equations (3.3) and (3.4) imply (3.2). Finally, the *injectivity* of σ follows directly from the definition of σ.

■

Since the mapping σ is a group homomorphism, a one-parameter group $g^t = (P^t, \phi^t)$ in G induces a one-parameter group $\sigma(P^t, \phi^t) = P^t \circ T\phi^t$ in the set Diff (TM) of diffeomorphisms of the tangent bundle TM.

On the other hand, an element (R, Y) of the set of all g.v.f. $\mathcal{G}\mathfrak{X}(M)$ can be considered as a vector field on TM under the identification explained below.

For $Y \in \mathfrak{X}(M)$ and $R \in END(TM)$, the set of all bundle endomorphisms of TM, we define two vector fields Y_* and R_* on TM by the following local coordinate representation,

$$Y_*(x, \xi) = \left[x, \xi, Y(x), DY(x)\cdot\xi\right]$$

$$R_*(x, \xi) = \left[x, \xi, 0, R(x)\cdot\xi\right]$$

where (x, ξ) and (x, ξ, v, η) are the local coordinates of TM and $T(TM)$ respectively, and $Y(x)$, $DY(x)$ and $R(x)$ are local expressions of Y, TY and R. Hence, both Y_* and R_* are elements of the tangent bundles of the manifold TM, i.e., $T(TM)$. Observe that Y_* and R_* are both well-defined, that is, they are independent of the choice of local coordinates. (This is proved in *Appendix IV* (Part II).) Thus we can define the mapping κ

from $\mathcal{G}\mathcal{X}(M)$ into $\mathcal{X}(TM)$ as follows:

$$\kappa : \mathcal{G}\mathcal{X}(M) \rightarrow \mathcal{X}(TM)$$

$$(R\,,Y) \rightarrow (R_* + Y_*)\,.$$

where $(R\,,Y) \in \mathcal{G}\mathcal{X}(M)$ and $(R_* + Y_*) \in \mathcal{X}(TM)$. Any element $\tilde{v}$ of $\mathcal{X}(TM)$ generates a flow on TM or an exponential map $\exp(t\,\tilde{v})$ in Diff (TM) as in the exposition of the general framework of normal forms for vector fields (see Chua and Kokubu [14]). Through these three procedures σ, κ, and exp, we define the flow, or the exponential map, which forms a one-parameter group (P^t, ϕ^t) in G for an element of $\mathcal{G}\mathcal{X}(M)$.

Definition 3.2: Exponential map

For an element $(R\,,Y) \in \mathcal{G}\mathcal{X}(M)$, we define the exponential map $\underline{\exp}\,t\,(R\,,Y)$ by

$$\underline{\exp}\,t\,(R\,,Y) = \sigma^{-1} \circ \exp t\,(R_* + Y_*)$$

for sufficiently small t. We call $(R\,,Y)$ the *infinitesimal generator* for the flow.

Proposition 3.3

> The above definition of the exponential map $\underline{\exp}$ is well-defined and $\underline{\exp}\,t\,(R\,,Y)$ forms a one-parameter group in G.

Proof. The exponential map $\underline{\exp}\,t(R\,,Y)$ is well-defined because $\exp t\,\kappa(R\,,Y) = \exp t(R_* + Y_*)$ is in the image of the mapping σ, and the injectivity of the mapping σ (Lemma 3.1).

$$\begin{array}{ccc}
(P\,,\phi) & \longrightarrow & (P \circ T\phi) \\
\rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\
 & \sigma & \\
G & \longrightarrow & Diff\ (TM) \\
\underline{\exp}\ \uparrow & & \uparrow\ \exp \\
\mathcal{G}\mathcal{X}(M) & \longrightarrow & \mathcal{X}(TM) \\
 & \kappa & \\
\rotatebox{90}{$\ni$} & & \rotatebox{90}{$\ni$} \\
(R\,,Y) & \longrightarrow & R_* + Y_*
\end{array}$$

More precisely; for $(P\,,\phi) \in G$, since P and $T\phi$ are bundle isomorphisms of TM, not necessarily covering identity, so is $P \circ T\phi$. Conversely, for an arbitrary bundle isomorphism $\Phi : TM \rightarrow TM$, we can choose $(P\,,\phi) \in G$, such that $\sigma(P\,,\phi) = \Phi$, where $\phi = \pi \circ \Phi$, and $P = \Phi \circ T(\pi \circ \Phi)^{-1}$ (π denotes a projection $TM \rightarrow M$). Thus the image $\mathrm{Im}\,\sigma$ is the set of all bundle isomorphisms (not necessarily covering identity.)

On the other hand, for $(R, Y) \in \mathcal{G}\mathcal{X}(M)$, $\kappa(R, Y) = (R_* + Y_*)$ is a vector field on TM, which is expressed by a local chart.

$$(R_* + Y_*)(x, \xi) = \left[x, \xi, Y(x), \left[DY(x) + R(x)\right] \cdot \xi\right].$$

Hence, the mapping $\exp t(R_* + Y_*)$ maps a point $(x_0, \xi_0) \in TM$ to the point $\left[x(t), \xi(t)\right] \in TM$ which is the solution of the differential equation,

$$\left.\begin{aligned} \dot{x} &= Y(x) \\ \dot{\xi} &= \left[DY(x) + R(x)\right] \cdot \xi \end{aligned}\right\} \tag{3.5}$$

under the initial conditions $x(0) = x_0$, $\xi(0) = \xi_0$. Note that the first equation is independent of ξ, thus the solution defines a flow of the base space M, which is denoted by $\exp t\, Y$. By substituting this solution $x(t) = (\exp t\, Y)(x_0)$ to the second equation, the resulting equation

$$\dot{\xi} = \left[DY(x(t)) + R(x(t))\right] \cdot \xi$$

is linear with respect to the variable ξ. This induces a linear invertible transformation from $T_{x_0}M$ to $T_{x(t)}M$. It follows that $\exp t\, \kappa(R, Y) = \exp t(R_* + Y_*)$ is a bundle isomorphism whose base map is $\exp t\, Y$ for each t, thereby proving well-definedness.

To show that $\underline{\exp}\, t\,(R, Y)$ is a one-parameter group in G, recall that σ is an injective group homomorphism (*Lemma 3.1*), hence $\left[\sigma^{-1}|_{\mathrm{Im}\,\sigma}\right]$ is also a group homomorphism. Thus,

$$\begin{aligned} \underline{\exp}\,(t+s)(R, Y) &= \sigma^{-1} \circ \exp(t+s)\kappa(R, Y) = \sigma^{-1} \circ \exp(t+s)(R_* + Y_*) \\ &= \sigma^{-1} \circ \exp t(R_* + Y_*) \circ \exp s(R_* + Y_*) \\ &= \left[\sigma^{-1} \circ \exp t(R_* + Y_*)\right] \cdot \left[\sigma^{-1} \circ \exp s(R_* + Y_*)\right] \\ &= \underline{\exp}\, t\,(R, Y) \cdot \underline{\exp}\, s\,(R, Y). \end{aligned}$$

■

Let us pause to consider an example on the computation of $\underline{\exp}\, t(R, Y)$.

Example 3.4

Consider $(R, Y) \in \mathcal{G}\mathcal{X}(\mathbb{R}^2)$, defined by $R = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $Y = \begin{bmatrix} y \\ 0 \end{bmatrix}$. Here we identify R with a matrix which fixes a coordinate of $\mathbb{R}^2$. Recall that the flow $\exp t(R_* + Y_*)$ on TM is given by a transformation $\left[x(0), \xi(0)\right] \to \left[x(t), \xi(t)\right]$ where $\left[x(t), \xi(t)\right]$ is a solution of the differential equation (3.5) on TM; namely,

$$\left.\begin{array}{l}\dot{x} = y \\ \dot{y} = 0 \\ \begin{bmatrix}\dot{\xi} \\ \dot{\eta}\end{bmatrix} = \left[\begin{bmatrix}0 & 0 \\ 0 & 1\end{bmatrix} + \begin{bmatrix}0 & 1 \\ 0 & 0\end{bmatrix}\right]\begin{bmatrix}\xi \\ \eta\end{bmatrix} = \begin{bmatrix}0 & 1 \\ 0 & 1\end{bmatrix}\begin{bmatrix}\xi \\ \eta\end{bmatrix}\end{array}\right\} \tag{3.6}$$

Here we use (x, y, ξ, η) for a local coordinate of $T\mathbf{R}^2$. Equation (3.2) can be solved explicitly as follow:

$$\begin{cases} x(t) = x(0) + y(0)t \\ y(t) = y(0) \\ \xi(t) = \left[\xi(0) - \eta(0)\right] + \eta(0)e^t \\ \eta(t) = \eta(0)e^t \end{cases}$$

It follows that

$$\Phi^t \triangleq \exp t(R_* + Y_*) : (x, y, \xi, \eta) \to (x+yt, y, \xi-\eta+\eta e^t, \eta e^t),$$

is the bundle isomorphism, and its base transformation is $\phi^t : (x, y) \to (x+yt, y)$. By the definition of $\sigma : (P, \phi) \to P \circ T\phi$, P^t is written by $\Phi^t \circ \left[T\phi^t\right]^{-1}$, whose base map is the identity. Restricting to the fiber $T_{(x,y)}\mathbf{R}^2$, we can obtain the bundle automorphism P^t as follows: Since $D\phi^t(x, y) = \begin{bmatrix}1 & t \\ 0 & 1\end{bmatrix}$ and $\Phi^t|_{T_{(x,y)}\mathbf{R}^2} = \begin{bmatrix}1 & e^t-1 \\ 0 & e^t\end{bmatrix}$,

$$P^t = \begin{bmatrix}1 & e^t-1 \\ 0 & e^t\end{bmatrix}\begin{bmatrix}1 & t \\ 0 & 1\end{bmatrix}^{-1} = \begin{bmatrix}1 & e^t-1 \\ 0 & e^t\end{bmatrix}\begin{bmatrix}1 & -t \\ 0 & 1\end{bmatrix}$$

$$= \begin{bmatrix}1 & -t-1+e^t \\ 0 & e^t\end{bmatrix}.$$

Thus, the one-parameter group $(P^t, \phi^t) = \underline{\exp}\, t(R, Y)$ is given by,

$$P^t = \begin{bmatrix}1 & -t-1+e^t \\ 0 & e^t\end{bmatrix}$$

$$\phi^t : (x, y) \to (x+yt, y).$$

■

Next we move on to obtain an explicit form of the *infinitesimal deformation* of the constrained system $(\mathbf{A}, \mathbf{v})$, which is defined by,

$$\frac{d}{dt}\bigg|_{t=0} \underline{\exp}\, t(R\,,Y)_{\#}(\mathbf{A}\,,\mathbf{v})\,.$$

To compute the infinitesimal deformation, we identify the bundle endomorphism $\mathbf{A}$ with a (1,1)-type tensor field $\tilde{\mathbf{A}}$ through a natural vector bundle isomorphism,

$$End(TM) \;\simeq\; TM \otimes T^{*}M\;,$$

which is presented in *Appendix II* (Part II). By this identification, the infinitesimal deformation is obtained as follows:

Theorem 3.5: infinitesimal deformation

> The infinitesimal deformation of a constrained system $(\mathbf{A}\,,\mathbf{v}) \in \mathcal{C}\mathfrak{X}(M)$ $\Big[$or a generalized vector field $(\mathbf{A}\,,\mathbf{v}) \in \mathcal{G}\mathfrak{X}(M)\Big]$ by a one-parameter group $\underline{\exp}\, t(R\,,Y)$ is given by,
>
> $$\frac{d}{dt}\bigg|_{t=0} \underline{\exp}\, t(R\,,Y)_{\#}(\mathbf{A}\,,\mathbf{v}) = \Big[R\cdot\mathbf{A} - \mathcal{L}_Y\,\mathbf{A}\,, R\cdot\mathbf{v} - [Y\,,\mathbf{v}]\Big]\,,$$
>
> where $\mathcal{L}_Y\,\mathbf{A}$ denotes the Lie derivative of the tensor field $\tilde{\mathbf{A}}$ with respect to the vector field Y, and [,] denotes the Lie bracket for vector fields.

Proof: First we shall prove: For $(R\,,0) \in \mathcal{G}\mathfrak{X}(M)$ and $(0\,,Y) \in \mathcal{G}\mathfrak{X}(M)$,

$$\underline{\exp}\, t(R\,,0) = (e^{tR}\,, id) \in G\;. \tag{3.7}$$

$$\underline{\exp}\, t(0\,,Y) = (I\,, \exp t\,Y) \in G\;. \tag{3.8}$$

Here e^{tR} denotes a bundle automorphism defined by $e^{tR(x)}$ on each tangent space T_xM, where $R(x) \triangleq R\,|_{T_xM}$ is a linear mapping $T_xM \to T_xM$; $e^{tR(x)}$ is the exponential map of the linear mapping $R(x)$ in the usual sense and I denotes the identity of TM.

In a local chart, $\underline{\exp}\, t(R\,,0)$ is defined as a flow of the differential equation,

$$\left.\begin{aligned} \dot{x} &= 0 \\ \dot{\xi} &= R(x)\cdot\xi \end{aligned}\right\}. \tag{3.9}$$

On the other hand, $\Big[x_1(t)\,,\xi_1(t)\Big] = \sigma\circ(e^{tR}\,, id)(x_0\,,\xi_0) = \Big[x_0\,, e^{tR(x_0)}\xi_0\Big]$ is a solution of (3.9). In fact

$$\begin{cases} \dfrac{d}{dt}\, x_1(t) = 0 \\ \dfrac{d}{dt}\, \xi_1(t) = \dfrac{d}{dt}\, e^{tR(x_0)}\cdot\xi_0 = R(x_0)e^{tR(x_0)}\xi_0 = R(x_1(t))\cdot\xi_1(t) \end{cases}$$

Thus the relation (3.7) follows.

Similarly, $\underline{\exp}\, t(0\,,Y)$ is defined as a flow of the differential equation,

$$\left.\begin{aligned}\dot{x} &= Y(x)\\ \dot{\xi} &= DY(x)\cdot\xi\end{aligned}\right\} \tag{3.10}$$

Since $\left[x_2(t),\xi_2(t)\right] = \sigma\circ(I,\exp tY)(x_0,\xi_0) = \left[\phi(t,x_0),D\phi(t,x_0)\cdot\xi_0\right]$ $\left[\text{where } \phi(t,x) = \phi^t(x) = (\exp tY)(x)\right]$ is a solution of (3.10); namely,

$$\frac{d}{dt}x_2(t) = \frac{d}{dt}\phi(t,x_0) = Y\left[\phi(t,x_0)\right] = Y\left[x_2(t)\right]$$

$$\frac{d}{dt}\xi_2(t) = \frac{d}{dt}D\,\phi(t,x_0)\cdot\xi_0 = D\,\frac{d}{dt}\phi(t,x_0)\cdot\xi_0$$

$$= D\left[Y(\phi(t,x_0))\right]\cdot\xi_0 = DY\left[x_2(t)\right]\cdot D\,\phi(t,x_0)\cdot\xi_0$$

$$= DY\left[x_2(t)\right]\cdot\xi_2(t) \quad ,$$

the relation (3.8) follows.

Lemma 3.6. $\left.\frac{d}{dt}\right|_{t=0} \underline{\exp}\,t\,(R,0)_{\#}(\mathbf{A},\mathbf{v}) = (R\cdot\mathbf{A},R\cdot\mathbf{v})$

Proof. When we restrict to the tangent space T_xM, the action of $\underline{\exp}\,t\,(R,0)\in G$ on $(\mathbf{A},\mathbf{v})\in\mathcal{C}\mathcal{X}(M)$ $\left[\text{or} \in \mathcal{G}\mathcal{X}'(M)\right]$ is interpreted as a matrix multiplication in the usual sense; that is, in terms of local coordinates, we have

$$\underline{\exp}\,t\,(R,0)_{\#}(\mathbf{A},\mathbf{v}) = (e^{tR},id)_{\#}(\mathbf{A},\mathbf{v}) = \left[e^{tR(x)}\mathbf{A}(x),e^{tR(x)}\mathbf{v}(x)\right].$$

Hence the infinitesimal deformation of $(\mathbf{A},\mathbf{v})$ is given by $(R\cdot\mathbf{A},R\cdot\mathbf{v})$, which can also be considered as a matrix multiplication.

■

As for the infinitesimal deformation $\left.\frac{d}{dt}\right|_{t=0}\underline{\exp}\,t\,(0,Y)_{\#}(\mathbf{A},\mathbf{v})$, we must recall again the identification of a bundle endomorphism $\mathbf{A}$ with a (1,1)-type tensor field $\tilde{\mathbf{A}}$. *(Appendix II).*

Lemma 3.7

$$\left.\frac{d}{dt}\right|_{t=0}\underline{\exp}\,t(0,Y)_{\#}(\mathbf{A},\mathbf{v}) = \left[-\mathcal{L}_Y\mathbf{A},-[Y,\mathbf{v}]\right]$$

Proof: As noted in the beginning of the proof of *Theorem 3.5,*

$$\underline{\exp}\,t(0,Y)_{\#}(\mathbf{A},\mathbf{v}) = (I,\exp tY)_{\#}(\mathbf{A},\mathbf{v})$$

$$= \left[T\phi^t\circ\mathbf{A}\circ(T\phi^t)^{-1},T\phi^t\circ\mathbf{v}\circ(\phi^t)^{-1}\right] = (\phi^t_{\#}\mathbf{A},\phi^t_{\#}\mathbf{v})$$

where $\phi^t = \exp tY$ is a one-parameter group of diffeomorphisms.

Since A is considered as a (1,1)-type tensor field $\tilde{\mathrm{A}}$, and since the infinitesimal deformation of $\tilde{\mathrm{A}}$ with respect to the one-parameter group ϕ^t is given by the Lie derivative, $\left.\frac{d}{dt}\right|_{t=0} \phi_*^t \tilde{\mathrm{A}} = -\mathcal{L}_Y \tilde{\mathrm{A}}$, (this is the definition of Lie derivative), we have only to see that $\phi_\#^t \mathrm{A} = T\phi^t \circ \mathrm{A} \circ (T\phi^t)^{-1}$ is identified with the transformed tensor field $\phi^t{}_* \tilde{\mathrm{A}} = (T\phi^t \otimes T^* \phi^t) \circ \tilde{\mathrm{A}} \circ \phi^{-1}$.

The tensor field $\tilde{\mathrm{A}}$ is written by $\sum_{ij} \mathrm{A}_{ij}(y) \frac{\partial}{\partial y_i} \otimes dy_j$ in terms of local coordinates (*Appendix II*). By a coordinate change $x = \phi(y)$ $(x_i = \phi_i(y_1, y_2, \ldots, y_n)$, the tensor field $\tilde{\mathrm{A}}$ is transformed as follows:

$$dy_i = \sum_j \frac{\partial y_i}{\partial x_j} dx_j = \sum_j \frac{\partial \phi_i^{-1}}{\partial x_j} (x) dx_j$$

$$\frac{\partial}{\partial y_i} = \sum_j \frac{\partial x_j}{\partial y_i} \frac{\partial}{\partial x_j} = \sum_j \frac{\partial \phi_j}{\partial x_i} (y) \frac{\partial}{\partial x_j}$$

$$\sum_{ij} \mathrm{A}_{ij}(y) \frac{\partial}{\partial y_i} \otimes dy_j = \sum_{ij} \mathrm{A}_{ij} \left[\phi^{-1}(x)\right] \sum_k \left[\frac{\partial \phi_k}{\partial y_i} (y) \frac{\partial}{\partial x_k}\right] \otimes \sum_l \left[\frac{\partial \phi_j^{-1}}{\partial x_l} (x) dx_l\right]$$

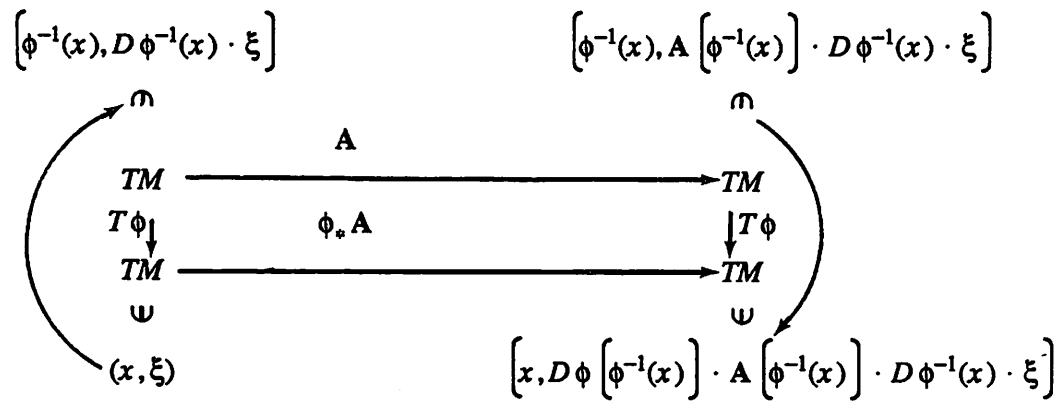
$$= \sum_{ijkl} \frac{\partial \phi_k}{\partial y_i} \left[\phi^{-1}(x)\right] A_{ij} \left[\phi^{-1}(x)\right] \frac{\partial \phi_j^{-1}}{\partial x_l} (x) \frac{\partial}{\partial x_k} \otimes dx_l$$

$$= D\phi \left[\phi^{-1}(x)\right] \cdot A \left[\phi^{-1}(x)\right] \cdot D\phi^{-1}(x) .$$

On the other hand, the bundle endomorphism A is transformed into $D\phi \left[\phi^{-1}(x)\right] \cdot \mathrm{A} \left[\phi^{-1}(x)\right] \cdot D\phi^{-1}(x)$ which is the same as $\phi_* \tilde{\mathrm{A}}$

The preceding calculations are summarized by the following diagrams:

$$\begin{array}{ccc} \mathrm{A} & \longrightarrow & \tilde{\mathrm{A}} \\ \downarrow & \curvearrowright & \downarrow \\ \phi_\#^t \mathrm{A} & \longrightarrow & \phi_*^t \tilde{\mathrm{A}} \end{array}$$



For the vector field $\mathrm{v}\in\mathfrak{X}(M)$, the infinitesimal deformation $\left.\frac{d}{dt}\right|_{t=0} T\phi^t \circ \mathrm{v} \circ (\phi^t)^{-1}$ is given by the Lie bracket $-[Y, \mathrm{v}]$ (see Chua and Kokubu [14]) ■ (End of the proof of Lemma. 3.7.)

We are now ready to prove *Theorem 3.5:*

$$\left.\frac{d}{dt}\right|_{t=0} \underline{\exp}\, t(R, Y)_{\#}(\mathrm{A}, x) = \left.\frac{d}{dt}\right|_{t=0} \underline{\exp}\, t\left[(R, 0) + (0, Y)\right]_{\#}(\mathrm{A}, \mathrm{v})$$

Since $\underline{\exp}\, t\left[(R, 0) + (0, Y)\right]$ differs from $\underline{\exp}\, t(R, 0)\cdot\underline{\exp}\, t(0, Y)$ within $0(t^2)$, the above expression is equal to

$$\left.\frac{d}{dt}\right|_{t=0} \left[\underline{\exp}\, t(R, 0)\cdot\underline{\exp}\, t(0, Y)\right]_{\#}(\mathrm{A}, \mathrm{v}) .$$

It follows from the Leibniz rule that

$$\left.\frac{d}{dt}\right|_{t=0} \underline{\exp}\, t(R, 0)_{\#}(\mathrm{A}, \mathrm{v}) + \left.\frac{d}{dt}\right|_{t=0} \underline{\exp}\, t(0, Y)_{\#}(\mathrm{A}, \mathrm{v}) .$$

From *Lemmas 3.6 and 3.7*, we have

$$(R\cdot\mathrm{A}, R\cdot\mathrm{v}) + \left[-\mathcal{L}_Y \mathrm{A}, -[Y, \mathrm{v}]\right] = \left[R\cdot\mathrm{A} - \mathcal{L}_Y\mathrm{A}, R\cdot\mathrm{v} - [Y, \mathrm{v}]\right]$$

This completes the proof of Theorem 3.5.

■

Using a local coordinate representation, we obtain,

$$\left.\begin{aligned} R\cdot\mathrm{A} - \mathcal{L}_Y\mathrm{A} &= \sum_{ijk}\left[R_{ik}\mathrm{A}_{kj} - \frac{\partial \mathrm{A}_{ij}}{\partial x_k} Y_k + \frac{\partial Y_i}{\partial x_k}\mathrm{A}_{kj} - \mathrm{A}_{ik}\frac{\partial Y_k}{\partial x_j}\right]\frac{\partial}{\partial x_i}\otimes dx_j \\ R\cdot\mathrm{v} - [Y, \mathrm{v}] &= \sum_{ik}\left[R_{ik}\mathrm{v}_k - \frac{\partial \mathrm{v}_i}{\partial x_k} Y_k + \frac{\partial Y_i}{\partial x_k}\mathrm{v}_k\right]\frac{\partial}{\partial x_i} \end{aligned}\right\} . \tag{3.11}$$

where $R = \sum_{ij} R_{ij}\frac{\partial}{\partial x_i}\otimes dx_j$, $\mathrm{A} = \sum_{ij}\mathrm{A}_{ij}\frac{\partial}{\partial x_i}\otimes dx_j$, $Y = \sum_i Y_i\frac{\partial}{\partial x_i}$, and $\mathrm{v} = \sum_i \mathrm{v}_i\frac{\partial}{\partial x_i}$.

For the one-parameter group $(P^t, \phi^t) = \underline{\exp}\, t(R, Y)$ of *Example 3.4*, we will show that the infinitesimal deformation, calculated directly from the definition of $\frac{d}{dt}\Big|_{t=0}(P^t, \phi^t)_\#(\mathbf{A}, \mathbf{v})$, coincides with that obtained from the formula (3.11).

Example 3.8

Let us choose $(\mathbf{A}, \mathbf{v})$ to be $\left[\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}, \begin{bmatrix}y\\1+x\end{bmatrix}\right]$. Recall that $P^t = \begin{bmatrix}1 & -t-1+e^t\\0 & e^t\end{bmatrix}$, $\phi^t(x,y) = (x+yt, y)$ in Example 3.4. Since $(P^t, \phi^t)_\#(\mathbf{A}, \mathbf{v}) = \left[P^t \circ T\phi^t \circ \mathbf{A} \circ (T\phi^t)^{-1}, P^t \circ T\phi^t \circ \mathbf{v} \circ (\phi^t)^{-1}\right] \triangleq (\mathbf{A}^t, \mathbf{v}^t)$, we can write $\mathbf{A}^t$ and $\mathbf{v}^t$ as follows:

$$\mathbf{A}^t = \begin{bmatrix}1 & -t-1+e^t\\0 & e^t\end{bmatrix}\begin{bmatrix}1 & t\\0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}1 & -t\\0 & 1\end{bmatrix}$$

$$= \begin{bmatrix}e^t-1 & -t+e^t(t+1)\\e^t & e^t(1-t)\end{bmatrix}$$

$$\mathbf{v}^t = \begin{bmatrix}1 & -t-1+e^t\\0 & e^t\end{bmatrix}\begin{bmatrix}1 & t\\0 & 1\end{bmatrix}\begin{bmatrix}y\\1+(x-yt)\end{bmatrix}$$

$$= \begin{bmatrix}y+(e^t-1)(1+x-yt)\\e^t(1+x-yt)\end{bmatrix}.$$

Therefore

$$\frac{d}{dt}\Big|_{t=0}(P^t, \phi^t)_\#(\mathbf{A}, \mathbf{v}) = \frac{d}{dt}\Big|_{t=0}\left[\begin{bmatrix}e^t-1 & -t+e^t(t+1)\\e^t & e^t(1-t)\end{bmatrix}, \begin{bmatrix}y+(e^t-1)(1+x-yt)\\e^t(1+x-yt)\end{bmatrix}\right]$$

$$= \left[\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}, \begin{bmatrix}1+x\\1+x-y\end{bmatrix}\right].$$

On the other hand, $\mathbf{A}$ and R (resp., $\mathbf{v}$ and Y) are expressed as tensor fields (resp., vector fields) as follows:

$$\mathbf{A} = \frac{\partial}{\partial x} \otimes dy + \frac{\partial}{\partial y} \otimes dx + \frac{\partial}{\partial y} \otimes dy$$

$$\mathbf{v} = y\,\frac{\partial}{\partial x} + (1+x)\,\frac{\partial}{\partial y}$$

$$R = \frac{\partial}{\partial y} \otimes dy\,, Y = y\,\frac{\partial}{\partial x}\,.$$

Therefore the formula (3.11) gives the infinitesimal deformation as follows:

$$R \cdot \mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

$$\mathcal{L}_Y \mathbf{A} = -\frac{\partial}{\partial x} \otimes dx + \frac{\partial}{\partial y} \otimes dy - \frac{\partial}{\partial x} \otimes dy = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$R \cdot \mathbf{v} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ 1+x \end{bmatrix} = \begin{bmatrix} 0 \\ 1+x \end{bmatrix}$$

$$[Y, \mathbf{v}] = \left[y \frac{\partial}{\partial x}, y \frac{\partial}{\partial x} + (1+x) \frac{\partial}{\partial y} \right] = y \frac{\partial}{\partial y} - (1+x) \frac{\partial}{\partial x} = \begin{bmatrix} -1+x \\ y \end{bmatrix}.$$

Hence

$$R \cdot \mathbf{A} - \mathcal{L}_Y \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \text{ and } R \cdot \mathbf{v} - [Y, \mathbf{v}] = \begin{bmatrix} 1-x \\ 1+x-y \end{bmatrix},$$

which are identical with the above result.

4. NORMAL FORMS OF CONSTRAINED EQUATIONS

The purpose of this section is to define normal forms for constrained systems and to compute them by the method of *infinitesimal deformation* presented in the previous section. In a naive sense, the (kth-order) normal form of a constrained system is obtained by transforming its (kth-order truncation of the) Taylor expansion at a point in the phase space M into a form as simple as possible by appropriate coordinate changes. For clarity, we will begin with a discussion of the local expressions of constrained systems and the notion of *jets*. We will then give a precise definition of normal forms and a method for obtaining them, as well as several examples.

For simplicity, we will first treat the case where the phase space M is the n-dimensional Euclidean space $\mathbf{R}^n$ and choose the standard cartesian coordinate $(x_1, x_2, ..., x_n)$. In this case, a generalized vector field $(\mathbf{A}, \mathbf{v})$ on $\mathbf{R}^n$ is given by a pair of a matrix-valued function $\mathbf{A}(\mathbf{x})$, and a vector-valued function $\mathbf{v}(\mathbf{x})$:

$$\left[\mathbf{A}(\mathbf{x}), \mathbf{v}(\mathbf{x})\right] = \left[\begin{bmatrix} a_{11}(\mathbf{x}), & \cdots & , a_{1n}(\mathbf{x}) \\ & \vdots & \\ a_{n1}(\mathbf{x}), & \cdots & , a_{nn}(\mathbf{x}) \end{bmatrix}, \begin{bmatrix} \mathrm{v}_1(\mathbf{x}) \\ \vdots \\ \mathrm{v}_n(\mathbf{x}) \end{bmatrix} \right], \ \mathbf{x} \in \mathbf{R}^n$$

See Example 3.4.

Recall next the following definition of a *jet* given in *Appendix 2* of Chua and Kokubu [14]:

Definition 4.1

Let f and g be smooth mappings from $\mathbf{R}^n$ to $\mathbf{R}^m$ defined in a neighborhood of a point $x_0 \in \mathbf{R}^n$. We say f and g are *k-jet equivalent* if every derivatives at x_0 up to order k of f as well as the value $f(x_0)$ coincide with those of g. This defines an equivalence relation and the equivalence class is called the *k-jet* of f at

x_0, denoted by $j^k_{x_0}f$.

Two pairs of mappings (f_1, f_2) and (g_1, g_2) are said to be *(k,l)-jet equivalent* if f_1 is k-jet equivalent to g_1, and f_2 is l-jet equivalent to g_2. The equivalence class of (f_1, f_2) is denoted by $j^{k,l}_{x_0}(f_1, f_2) = (f^k_1, f^l_2)$. (Note that this notation is used in [14] for a different object.)

In a similar way as the k-jets of vector fields, the k-jet of a *generalized* vector field $(\mathbf{A}, \mathbf{v})$ is identified with its kth-order truncation $a^k_{ij}(\mathbf{x})$, $\mathbf{v}^k_j(\mathbf{x})$ of the Taylor expansion of each component $a_{ij}(\mathbf{x})$, $\mathbf{v}_j(\mathbf{x})$. We denote the set of all k-jets [resp. (k,l)-jets] of g.v.f.'s at x_0 by $J^k_{x_0}\mathcal{G}\mathcal{X}$ (resp., $J^{k,l}_{x_0}\mathcal{G}\mathcal{X}$). Since every constrained system $(\mathbf{A}, \mathbf{x})$ itself is a generalized vector field, it is also expressed as a pair of a matrix-valued function and a vector-valued function. However, the k-jet of the constrained system is *not* given by its usual kth-order truncation because of the following reason:

Example 4.2

Let $\mathbf{A}(\mathbf{x}) \in End^{(1)}(T\mathbf{R}^2)$ be a bundle endomorphism of $T\mathbf{R}^2$ of corank 1 defined by

$$\mathbf{A}(\mathbf{x}) = \begin{bmatrix} xy+y^3 & x+y^2 \\ y & 1 \end{bmatrix}, \ (x, y) \in \mathbf{R}^2$$

Its 1st-order truncation $\begin{bmatrix} 0 & x \\ y & 1 \end{bmatrix}$ is not of constant rank.

This observation shows that the kth-order truncation does not give a k-jet for the constrained system because the k-jet $(\mathbf{A}^k, \mathbf{v}^k)$ of the constrained system $(\mathbf{A}, \mathbf{v})$ should have a bundle endomorphism $\mathbf{A}^k$ with a *constant* rank. This difficulty comes from the fact that not all components of the bundle endomorphism A of constant rank are independent. For example, any bundle endomorphism

$$\mathbf{A}(\mathbf{x}) = \begin{bmatrix} a_{11}(\mathbf{x}), & a_{12}(\mathbf{x}) \\ a_{21}(\mathbf{x}), & a_{22}(\mathbf{x}) \end{bmatrix}, \ \mathbf{x} \in \mathbf{R}^2$$

of $T\mathbf{R}^2$ of corank 1 satisfies the relation,

$$a_{11}(\mathbf{x})\, a_{22}(\mathbf{x}) - a_{12}(\mathbf{x})\, a_{21}(\mathbf{x}) = 0 .$$

In general, we have only $(n^2 - r^2)$ independent components among the n^2 components of $\mathbf{A} \in End^{(r)}(T\mathbf{R}^n)$, in view of the following lemma.

Lemma 4.3

Let A(x) be a n×n matrix-valued function on $\mathbb{R}^n$ where $A(x_0)$ is of corank r for some $x_0 \in \mathbb{R}^n$. Let P be a non-singular n×n matrix with

$$P^{-1}A(x_0)P = \begin{bmatrix} E_0 & B_0 \\ C_0 & D_0 \end{bmatrix}$$

where D_0 is a non-singular matrix of order (n-r). Then A(x) is of constant corank r for x near x_0, if and only if,

$$E(x) = B(x) \cdot D(x)^{-1} \cdot C(x)$$

holds, where

$$P^{-1}A(x)P = \begin{bmatrix} E(x) & B(x) \\ C(x) & D(x) \end{bmatrix}.$$

Proof: Note that $D(x)$ is non-singular for x near x_0 because $D(x)$ is near $D_0 = D(x_0)$. Multiplying $\begin{bmatrix} I & -B(x)D^{-1}(x) \\ 0 & I \end{bmatrix}$ to $\begin{bmatrix} E(x) & B(x) \\ C(x) & D(x) \end{bmatrix}$ from the left, we obtain,

$$\begin{bmatrix} I & -B(x)D^{-1}(x) \\ 0 & I \end{bmatrix}\begin{bmatrix} E(x) & B(x) \\ C(x) & D(x) \end{bmatrix} = \begin{bmatrix} E(x)-B(x)D^{-1}(x)C(x) & 0 \\ C(x) & D(x) \end{bmatrix},$$

which is also of corank r. Hence,

$$E(x) - B(x)D^{-1}(x)C(x) = 0$$

■

This lemma shows that the upper left part $E(x)$ depends on the remaining 3 parts. Consequently, we can choose only $B(x)$, $C(x)$ and non-singular $D(x)$ as independent components. In other words, once we fix each B, C, and D, then we can reconstruct $A = \begin{bmatrix} E & B \\ C & D \end{bmatrix}$ of corank r by putting $E = B D^{-1} C$. Since bundle endomorphism A of corank r is determined only by such B, C, and D, we denote A as $\begin{bmatrix} & B \\ C & D \end{bmatrix}$ and call it the *canonical expression.*

When we speak of the k-jets of constrained system (A , v) of corank r, we have only to take the kth-order truncation of the canonical expression.

Example 4.4

The k-jet of the bundle endomorphism A(x) of *Example 4.2* is given by

$$\begin{bmatrix} & 0 \\ 0 & 1 \end{bmatrix}, \text{ for } k = 0,$$

$$\begin{bmatrix} & x \\ y & 1 \end{bmatrix}, \text{ for } k = 1,$$

and

$$\begin{bmatrix} & x+y^2 \\ y & 1 \end{bmatrix}, \text{ for } k = 2.$$

As bundle endomorphisms having the 1-jet $\begin{bmatrix} & x \\ y & 1 \end{bmatrix}$ and the 2-jet $\begin{bmatrix} & x+y^2 \\ y & 1 \end{bmatrix}$, we can choose

$$\begin{bmatrix} xy + O(3) & x + O(2) \\ y + O(2) & 1 + O(2) \end{bmatrix}$$

and

$$\begin{bmatrix} xy + y^3 + O(4), & x + y^2 + O(3) \\ y + O(3), & 1 + O(3) \end{bmatrix}$$

respectively, where $O(k)$ represents terms of the degree $\geq k$.

Now we proceed to the definition of k-jets of constrained systems on a general manifold M. Let (A, v) be a constrained system of corank r on M. Recall that A is a section of the fiber bundle $End^{(r)}(TM)$ over M introduced in *Section 2*, whose standard fiber is the space of linear mappings from T_xM into itself of corank r. From *Lemma 4.3*, this fiber is an (n^2-r^2)-dimensional manifold. Hence, by fixing a local coordinate of M, A has a local expression $\overline{\mathbf{A}}: \mathbf{R}^n \to \mathbf{R}^{n^2-r^2}$ as the canonical expression. Similarly, the vector field v is a section of the tangent bundle TM and has a local expression $\overline{\mathbf{v}}: \mathbf{R}^n \to \mathbf{R}^n$.

Definition 4.5

Two constrained systems (A, v) and (A′, v′) of corank r are said to be *k-jet equivalent* at $x_0 \in M$ if, by taking a local coordinate around x_0, the local expressions $(\overline{\mathbf{A}}, \overline{\mathbf{v}})$ and $(\overline{\mathbf{A}}', \overline{\mathbf{v}}')$ are k-jet equivalent; namely,

$$\left[\frac{\partial}{\partial x_1}\right]^{k_1} \cdots \left[\frac{\partial}{\partial x_n}\right]^{k_n} \overline{\mathbf{A}}(\mathrm{x}) = \left[\frac{\partial}{\partial x_1}\right]^{k_1} \cdots \left[\frac{\partial}{\partial x_n}\right]^{k_n} \overline{\mathbf{A}}'(\mathrm{x})$$

$$\left[\frac{\partial}{\partial x_1}\right]^{k_1} \cdots \left[\frac{\partial}{\partial x_n}\right]^{k_n} \overline{\mathbf{v}}(\mathrm{x}) = \left[\frac{\partial}{\partial x_1}\right]^{k_1} \cdots \left[\frac{\partial}{\partial x_n}\right]^{k_n} \overline{\mathbf{v}}'(\mathrm{x})$$

for all $k_1, \cdots, k_n$ with $0 \leq k_1 + \cdots + k_n \leq k$. The k-jet equivalence class of (A, v) at x_0 is called the *k-jet* of (A, v) at x_0, which is denoted by $j^k_{x_0}(\mathrm{A}, \mathrm{v}) = (\mathrm{A}^k, \mathrm{v}^k)$. We denote the set of all k-jets of strained systems of corank r by $J^k_{x_0}\mathcal{CX}^r$.

Similar to the case of vector fields, several properties corresponding to those derived in *Appendix* 2 of Chua and Kokubu [14] also hold for jets of constrained systems.

For a transformation $(P,\phi)\in G$ of constrained systems, the k-jet of a bundle automorphism P and that of a diffeomorphism ϕ are defined in the same manner as above, which we denote by P^k and ϕ^k, respectively. Let $Diff_{x_0}$ be the group of diffeomorphisms of M fixing a point x_0, and let $Diff^k_{x_0}$ denote the k-jets of $Diff_{x_0}$ at this point. Also, let the space of k-jets of bundle automorphisms at x_0 be denoted by $AUT^k_{x_0}$.

Consider the k-jet of the transformed constrained system $(P,\phi)_\#(\mathrm{A},\mathrm{v})$. Suppose (A,v) and $(\mathrm{A}',\mathrm{v}')$ are k-jet equivalent, then $(P,\phi)_\#(\mathrm{A},\mathrm{v})$ and $(P,\phi)_\#(\mathrm{A}',\mathrm{v}')$ are also k-jet equivalent. Moreover, as is shown in the following Proposition, the higher order part of (P,ϕ) does *not* affect the k-jet of $(P,\phi)_\#(\mathrm{A},\mathrm{v})$, because the action is expressed by the composition of mappings P, ϕ and $T\phi$.

Proposition 4.6

(1) $J^{k,k+1}_{x_0}G = AUT^k_{x_0} \rtimes Diff^{k+1}_{x_0}$ forms a group.

(2) For $(P^k,\phi^{k+1})\in J^{k,k+1}_{x_0}G$ and $(\mathrm{A}^k,\mathrm{v}^k)\in J^k_{x_0}\mathcal{CX}^r$, $(P^k,\phi^{k+1})_\#(\mathrm{A}^k,\mathrm{v}^k)$ is given by

$$(P^k,\phi^{k+1})_\#(\mathrm{A}^k,\mathrm{v}^k) = j^k_0\left[(P,\phi)_\#(\mathrm{A},\mathrm{v})\right] \tag{4.1}$$

where (P,ϕ) and (A,v) are representatives of (P^k,ϕ^{k+1}) and $(\mathrm{A}^k,\mathrm{v}^k)$, respectively. Moreover,

$$(Q^k,\psi^{k+1})_\#\left[(P^k,\phi^{k+1})_\#(\mathrm{A}^k,\mathrm{v}^k)\right] = \left[(Q^k,\psi^{k+1})\cdot(P^k,\phi^{k+1})\right]_\#\cdot(\mathrm{A}^k,\mathrm{v}^k)$$

holds for (Q^k,ψ^{k+1}), and $(P^k,\phi^{k+1})\in J^{k,k+1}_{x_0}G$ and $(\mathrm{A}^k,\mathrm{v}^k)\in J^k_{x_0}\mathcal{CX}^r$.

Proof:

(1) The group multiplication is defined by

$$(P^k,\phi^{k+1})\cdot(Q^k,\psi^{k+1}) \triangleq \left[j^k_{x_0}(P\circ T\phi\circ Q\circ T\phi^{-1}),\, j^{k+1}_{x_0}(\phi\circ\psi)\right]$$

where (P,ϕ) and (Q,ψ) are representatives of (P^k,ϕ^{k+1}) and (Q^k,ψ^{k+1}); respectively. By the chain rule, the derivatives of $P\circ T\phi\circ Q\circ T\phi^{-1}$, up to order k are determined by the derivatives of P and Q up to order k and those of ϕ up to order k+1. The derivatives of $\phi\circ\psi$ up to order k+1 are determined by those of ϕ and ψ up to order k+1. Therefore, the definition is independent of the choice of the representatives. It is clear that this multiplication operation satisfies the axiom of group.

(2) Since the action of (P, ϕ) on (A, v) is given by,

$$(P, \phi)_{\#}(\mathrm{A}, \mathrm{v}) = (P \circ T\phi \circ \mathrm{A} \circ T\phi^{-1}, P \circ T\phi \circ \mathrm{v} \circ \phi^{-1}),$$

the k-jet of the transformed constrained system is determined by the k-jets of P, A and v, and the k+1 jet of ϕ. Therefore, we have proved that (4.1) is well-defined.

The proof for the latter half is as follows:

$$\begin{aligned}(Q^k, \psi^{k+1})_{\#}\left[(P^k, \phi^{k+1})_{\#}(\mathrm{A}^k, \mathrm{v}^k)\right] &= (Q^k, \psi^{k+1})_{\#} j^k_{x_0}\left[(P, \phi)_{\#}(\mathrm{A}, \mathrm{v})\right] \\ &= j^k_{x_0}\left[(Q, \psi)_{\#}\left[(P, \phi)_{\#}(\mathrm{A}, \mathrm{v})\right]\right] \\ &= j^k_{x_0}\left[\left[(Q, \psi)\cdot(P, \phi)\right]_{\#}(\mathrm{A}, \mathrm{v})\right] \\ &= \left[j^{k,k+1}_{x_0}(Q, \psi)\cdot(P, \phi)\right]_{\#}(\mathrm{A}^k, \mathrm{v}^k) \\ &= \left[(Q^k, \psi^{k+1})\cdot(P^k, \phi^{k+1})\right]_{\#}(\mathrm{A}^k, \mathrm{v}^k).\end{aligned}$$

■

Hence we have shown that the group $J^{k,k+1}_{x_0}G$ acts on the k-jets space of constrained systems. This group action induces an equivalence relation among the k-jets of constrained systems in the same way as in the general theory of normal forms for vector fields in [14]; namely, two k-jets $(\mathrm{A}^k, \mathrm{v}^k)$ and $(\mathrm{A}'^k, \mathrm{v}'^k)$ are said to be *equivalent* if there exists a (k,k+1)-jet (P^k, ϕ^{k+1}) of transformation of constrained systems such that

$$(\mathrm{A}'^k, \mathrm{v}'^k) = (P^k, \mathrm{v}^k)_{\#}(\mathrm{A}^k, \mathrm{v}^k)$$

holds.

A *k-th order normal form* of (A, v) is a representative of the equivalence class of the k-jets of (A, v). Our goal is to choose the simplest form as the representative.

Let us now consider the infinitesimal deformation of k-jets of constrained systems. In *Section 3* we have already obtained the infinitesimal deformation of constrained systems. We will now translate it into the k-jet version.

For (k,k+1)-jet of (R, Y), we can define a one-parameter group $\underline{\exp}\, t(R^k, Y^{k+1})$ in $J^{k,k+1}_{x_0}G$ by

$$\underline{\exp}\, t(R^k, Y^{k+1}) = j^{k,k+1}_{x_0}\left[\underline{\exp}\, t(R, Y)\right].$$

Appendix V (Part II) proves that this group is well-defined.

Theorem 4.7: infinitesimal deformation for k-jets

The infinitesimal deformation of a k-jet $(\mathrm{A}^k, \mathrm{v}^k) \in J^k_{x_0}\mathcal{C}\mathcal{X}^r$ of a constrained system of corank r by a local one-parameter group $\underline{\exp}\, t(R^k, Y^{k+1})$ of $J_{x_0}^{k,k+1}G$ is given by,

$$\frac{d}{dt}\Big|_{t=0} \underline{\exp}\, t(R^k, Y^{k+1})_{\#}(\mathbf{A}^k, \mathbf{v}^k) = \Big[R^k \cdot \mathbf{A}^k - \mathcal{L}_{Y^{k+1}}\mathbf{A}^k, R^k \cdot \mathbf{v}^k - [Y^{k+1}, \mathbf{v}^k]\Big], \qquad (4.2)$$

where $R^k \cdot \mathbf{A}^k = j_{x_0}^k(R \cdot A)$, $\mathcal{L}_{Y^{k+1}}\mathbf{A}^k = j_{x_0}^k \mathcal{L}_Y \mathbf{A}$, $R^k \cdot \mathbf{v}^k = j_{x_0}^k(R \cdot \mathbf{v})$ and $[Y^{k+1}, \mathbf{v}^k] = j_{x_0}^k\ [Y, \mathbf{v}]$ for representatives R, Y, $\mathbf{A}$, $\mathbf{v}$ of R^k, Y^{k+1}, $\mathbf{A}^k$, $\mathbf{v}^k$; respectively.

Proof: Recall that the action of $J_{x_0}^{k,k+1}G$ on $J_{x_0}^k \mathcal{CX}^r$ is given by (4.1) and that the infinitesimal deformation of a constrained system is given by *Theorem 3.5;* namely,

$$\frac{d}{dt}\Big|_{t=0} \underline{\exp}\, t(R, Y)_{\#}(\mathbf{A}, \mathbf{v}) = \Big[R \cdot \mathbf{A} - \mathcal{L}_Y \mathbf{A}, R \cdot \mathbf{v} - [Y, \mathbf{v}]\Big].$$

Hence, it suffices to prove that the right-hand side of (4.2) is well-defined. But this follows from (3.11) in *Section 3*. Observe that, since v does not necessarily vanish at x_0, and since Y vanishes at x_0, $[Y, \mathbf{v}]$ determines the well-defined k-jet.

■

It follows that we can calculate the normal forms of constrained systems in principle via the general theory of normal forms. The algorithm for the calculation is similar to that of normal forms of vector fields [14]. In the case of vector fields, the classification of 1-jets (Jordan normal forms) is given at the first stage (see Chua and Kokubu [14]). Here, we must obtain a classification of the leading part for the constrained systems which correspond to the Jordan normal forms; that is, a classification of 0-jets of constrained systems. For a constrained system $(\mathbf{A}, \mathbf{v})$, we choose

$$(\mathbf{A}_0, \mathbf{v}_0) = \Big[\mathbf{A}(x_0), \mathbf{v}(x_0)\Big]$$

for a chosen point x_0 in M, which we call the *leading part* of $(\mathbf{A}, \mathbf{v})$ at x_0. If $(\mathbf{A}, \mathbf{v})$ is of corank r, $\mathbf{A}_0$ is a linear map of corank r, and $\mathbf{v}_0$ is a vector. The leading part of $(\mathbf{A}, \mathbf{v})$ is then classified as follows.

Proposition 4.8: classification of leading part

Every leading part $(\mathbf{A}_0, \mathbf{v}_0)$ of a constrained system is equivalent to one of the following forms

(i)

$$\left[\begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}, \begin{bmatrix} e_r \\ 0 \end{bmatrix}\right]$$

(ii)

$$\left[\begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}, \begin{bmatrix} 0 \\ e_{n-r} \end{bmatrix}\right]$$

(iii)

$$\left[\begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}, \begin{bmatrix}0\\ 0\end{bmatrix}\right],$$

where $e_k \triangleq (1,0,\cdots,0) \in \mathbf{R}^k$ ($k = r$ or $n-r$), and I_{n-r} denotes the unit matrix of order $n-r$. Here, "equivalent" means "*equivalent in the sense of 0-jet.*"

Proof: It is necessary to show that there exists a (0, 1)-jet (P^0, ϕ^1) of transformation such that the transformed leading part

$$(P^0,\phi^1)_\#(\mathrm{A}_0,\mathrm{v}_0) = \left[P^0 \circ T\phi^1 \circ \mathrm{A}_0 \circ (T\phi^1)^{-1}, P^0 \circ T\phi^1 \circ \mathrm{v}_0 \circ (\phi^1)^{-1}\right]$$

assumes one of the above 3 forms.

By *Lemma 4.3,* a local coordinate expression of $(\mathrm{A}_0, \mathrm{v}_0)$ can be chosen as follow:

$$\left[\begin{bmatrix}BD^{-1}C & B\\ C & D\end{bmatrix}, \begin{bmatrix}\mathrm{v}_1\\ \mathrm{v}_2\end{bmatrix}\right], \det D \neq 0 .$$

Since ϕ is a diffeomorphism, we may write $\phi^1(x) = \overline{Q}^{-1}x$ for some non-singular matrix $\overline{Q}$; hence we have $(T\phi^1)^{-1} = \overline{Q}$. Also, since P is a bundle automorphism, we may write $P^0 \circ T\phi^1 = \overline{P}$ for a non-singular matrix. Moreover, since v_0 is a constant vector, ϕ^1 has no effect on v_0. Using these observations, we can identify $(P^0,\phi^1)_\#(\mathrm{A}_0, \mathrm{v}_0)$ with $(\overline{P}\cdot\mathrm{A}_0\cdot\overline{Q}, \overline{P}\cdot\mathrm{v}_0)$. Let us choose

$\overline{P} = \begin{bmatrix}I_r & -BD^{-1}\\ 0 & D^{-1}\end{bmatrix}$ and $\overline{Q} = \begin{bmatrix}I_r & 0\\ -D^{-1}C & I_{n-r}\end{bmatrix}$, to obtain $\overline{P}\cdot\mathrm{A}_0\cdot\overline{Q} = \begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}$. Hence, the leading part $(\mathrm{A}_0, \mathrm{v}_0)$ is transformed into $\left[\begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}\begin{bmatrix}\mathrm{v}_1'\\ \mathrm{v}_2'\end{bmatrix}\right]$, where $\begin{bmatrix}\mathrm{v}_1'\\ \mathrm{v}_2'\end{bmatrix} = \overline{P}\begin{bmatrix}\mathrm{v}_1\\ \mathrm{v}_2\end{bmatrix}$.

Our next step is to transform $\begin{bmatrix}\mathrm{v}_1'\\ \mathrm{v}_2'\end{bmatrix}$ into one of the above forms *without* changing $\begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}$. Note that the matrices of the form $\overline{P} = \begin{bmatrix}P_1 & 0\\ P_3 & P_4\end{bmatrix}$ and $\overline{Q} = \begin{bmatrix}Q_1 & Q_2\\ 0 & P_4^{-1}\end{bmatrix}$, where P_1, P_4 and Q_1 are non-singular, do *not* change $\mathrm{A}_0 = \begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}$. Indeed, we have

$$\begin{bmatrix}P_1 & 0\\ P_3 & P_4\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}\begin{bmatrix}Q_1 & Q_2\\ 0 & P_4^{-1}\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & I_{n-r}\end{bmatrix}.$$

The vector $\begin{bmatrix}\mathrm{v}_1'\\ \mathrm{v}_2'\end{bmatrix}$ is, thus, transformed by $\overline{P} = \begin{bmatrix}P_1 & 0\\ P_3 & P_4\end{bmatrix}$ as follows:

$$\begin{bmatrix} P_1 & 0 \\ P_3 & P_4 \end{bmatrix} \begin{bmatrix} v_1' \\ v_2' \end{bmatrix} = \begin{bmatrix} P_1 v_1' \\ P_3 v_1' + P_4 v_2' \end{bmatrix} . \tag{4.3}$$

If $v_1' \neq 0$, there exist P_1, P_3 and P_4 such that

$$P_1 v_1' = e_r \quad , \quad P_3 v_1' + P_4 v_2' = 0 .$$

If $v_1' = 0$ and $v_2' \neq 0$, then $P_1 v_1' = 0$, $P_3 v_1' = 0$ and there exists a non-singular matrix P_4 such that $P_4 v_2' = e_{n-r}$. Finally, observe that for the last case when $v_1' = v_2' = 0$, the right-hand side of (4.3) is always 0.

■

At the next stage, let us choose the 1-jet

$$(A^1, v^1) = (A_0, v_0) + (A_1, v_1) ,$$

where (A_i, v_i) is the ith-order part of (A, v) which contains the specified leading part (A_0, v_0). Consider the 1st-order normal form problem for constrained systems, that is, to deform (A^1, v^1) into a simpler form without changing the leading part (A_0, v_0). Just as in the case of vector fields [14], depending on the degree of degeneracy of the 1-jet, we will in general obtain several distinct 1st order-normal forms having the specified leading part. The normal form corresponding to the *least degenerate* 1-jet is called the *non-degenerate 1st-order normal form.* Just as in the case of vector fields, we can proceed inductively to solve the *higher-order normal form problems:* Given the (k-1)-jet (A^{k-1}, v^{k-1}) of a constrained system (A, v) we derive the associated kth-order normal forms by simplifying the kth order terms via a suitable one-parameter group of (k,k+1)-jet of transformations $\underline{\exp}\, t\, (R^k, Y^{k+1}) \in J_{x_0}^{k,k+1} G$ which fix the (k-1)-jet (A^{k-1}, v^{k-1}).

The following lemma corresponds to the key lemma for vector field normal forms (*Lemma 4.6* in [14]) which gives us an infinitesimal generator while keeping the lower jets of constrained systems invariant.

Lemma 4.9

Let (A^k, v^k) denote a k-jet of a constrained system in $J_{x_0}^k \mathcal{C}\mathcal{X}^r$ and let (A^{k-1}, v^{k-1}) denote its (k-1)-jet. If an infinitesimal generator $(R^k, Y^{k+1}) \in J_{x_0}^{k,k+1} \mathcal{G}\mathcal{X}$ satisfies,

$$j_{x_0}^{k-1} \left\{ (R^k, Y^{k+1}), (A^k, v^k) \right\} = 0 ,$$

where

$$\left\{ (R^k, Y^{k+1}), (A^k, v^k) \right\}$$

is defined by

$$\left[R^k \cdot A^k - \mathcal{L}_{Y^{k+1}} A^k \,, R^k \cdot v^k - [Y^{k+1}, v^k]\right],$$

then $\underline{\exp}\, t\,(R^k, Y^{k+1})_{\#}(A^k, v^k)$ leaves the (k-1)-jet (A^{k-1}, v^{k-1}) invariant; that is

$$j_{x_0}^{k-1}\, \underline{\exp}\, t\,(R^k, Y^{k+1})_{\#}(A^k, v^k) = (A^{k-1}, v^{k-1}).$$

Proof: Since the proof of this lemma is exactly the same as that for vector field normal forms, we omit it and refer to [14] for the details.

■

By this lemma, in order to compute normal forms, we have only to choose an infinitesimal generator (R^k, Y^{k+1}) satisfying,

$$\left\{(R^k, Y^{k+1}), (A^k, v^k)\right\}^{k-1} = j_{x_0}^{k-1}\left\{(R^k, Y^{k+1}), (A^k, v^k)\right\} = 0 \tag{4.4}$$

and solve the associated differential equation,

$$\frac{d}{dt}(A^k, v^k)(t) = -\left\{(R^k, Y^{k+1}), (A^k, v^k)(t)\right\}^k$$

where $(A^k, v^k)(t) = \underline{\exp}\, t\,(R^k, Y^{k+1})_{\#}(A^k, v^k)(0)$. To simplify notations, we will henceforth denote (A^k, v^k) by a^k and (R^k, Y^{k+1}) by ξ^k; respectively. Under condition (4.4), the above differential equation can be regarded as a differential equation,

$$\frac{d}{dt} h_k(t) = -\left\{\xi^k, a^{k-1} + h_k(t)\right\}_k \tag{4.5}$$

on $H_k\, \mathcal{C}\mathcal{X}^r$, the set of all homogeneous constrained systems of order k. Here $h_k(t)$ denotes the kth-order part of $(A^k, v^k)(t)$, and $\{\cdot,\cdot\}_k$ denotes the kth-order part of $\{\cdot,\cdot\}$. We also denote the set of all pairs (R_k, Y_{k+1}) by $H_{k,k+1}\mathcal{G}\mathcal{X}$, where R_k is a homogeneous bundle endomorphism of order k while Y_{k+1} is a homogeneous vector field of order k+1.

Example 4.10: Rapid Point

Consider a family of constrained systems on a 2-dimensional manifold whose leading part (A_0, v_0) is equivalent to $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]$. We may assume that (A_0, v_0) itself is of this form without loss of generality. First we consider the 1st-order normal form problem on H_1 $\quad ^1$. By taking into account the canonical expression (*Example 4.4*), $H_1\, \mathcal{C}\mathcal{X}^1$ is spanned by,

$$\left\{\begin{matrix} \left[x\frac{\partial}{\partial x}\otimes dy\,,0\right], & \left[x\frac{\partial}{\partial y}\otimes dy,0\right], & \left[0,x\frac{\partial}{\partial y}\right] \\ \left[y\frac{\partial}{\partial x}\otimes dy\,,0\right], & \left[y\frac{\partial}{\partial y}\otimes dy\,,0\right], & \left[0,y\frac{\partial}{\partial y}\right] \\ \left[x\frac{\partial}{\partial y}\otimes dx\,,0\right], & \left[0,x\frac{\partial}{\partial x}\right] & \\ \left[y\frac{\partial}{\partial y}\otimes dx\,,0\right], & \left[0,y\frac{\partial}{\partial x}\right] & \end{matrix}\right\},$$

Hence, $dim\, H_1\, \mathcal{C}\mathcal{X}^1 = 10$. Let us choose the (1,2)-jet of infinitesimal generator $(R^1,Y^2) = (R_0,Y_1) + (R_1,Y_2)$ such that $\left\{(R^1,Y^2),(A^1,v^1)\right\}^0 = 0$ holds. Here we choose $(R_0,\ Y_1) = 0$; then for all $\xi_1 = (R_1,Y_2) \in H_{1,2}\,\mathcal{G}\mathcal{X}$, the above condition is satisfied. Note that $H_{1,2}\mathcal{G}\mathcal{X}$ is a linear space spanned by

$$\left\{\begin{matrix} \left[x\,\frac{\partial}{\partial x}\otimes dx\ ,0\right], & \left[y\,\frac{\partial}{\partial x}\otimes dx\ ,0\right], & \left[0\,,x^2\,\frac{\partial}{\partial x}\right], & \left[0\,,xy\,\frac{\partial}{\partial x}\right] \\ \left[x\,\frac{\partial}{\partial x}\otimes dy\ ,0\right], & \left[y\,\frac{\partial}{\partial x}\otimes dy\ ,0\right], & \left[0\,,y^2\,\frac{\partial}{\partial x}\right], & \left[0\,,x^2\,\frac{\partial}{\partial y}\right] \\ \left[x\,\frac{\partial}{\partial y}\otimes dx\ ,0\right], & \left[y\,\frac{\partial}{\partial y}\otimes dx\ ,0\right], & \left[0\,,xy\,\frac{\partial}{\partial y}\right], & \left[0\,,y^2\,\frac{\partial}{\partial y}\right] \\ \left[x\,\frac{\partial}{\partial y}\otimes dy\ ,0\right], & \left[y\,\frac{\partial}{\partial y}\otimes dy\ ,0\right] & & \end{matrix}\right\}$$

Hence, the dimension of $H_{1,2}\mathcal{G}\mathcal{X}$ is equal to 14, and the 1st-order normal form problem becomes

$$\frac{d}{dt}\,h_1(t) = -\ \{\xi_1, a_0 + h_1(t)\}_1 = -\ \{\xi_1\ a_0\}_1\ , \tag{4.6}$$

where $\xi_1 = (R_1,Y_2)$, $a_0 = (A_0,v_0)$, $h_1(t) = (A_1,v_1)(t)$.

The result of the computation of $\{\xi_1, a_0\}$ in terms of the above basis is summarized in *Table 4.11.* Therefore, (4.6) can be recast as follow:

$$\frac{d}{dt}\begin{bmatrix} h_{1,1}(t) \\ \vdots \\ h_{1,10}(t) \end{bmatrix} = -(K)\begin{bmatrix} \xi_{1,1} \\ \vdots \\ \xi_{1,14} \end{bmatrix},$$

where $h_{1,1}(t), \cdots, h_{1,10}(t)$ are coefficients of $h_1(t)$ with respect to the above basis for $H_1\mathcal{C}\mathcal{X}^1$, and $\xi_{1,1}, \cdots, \xi_{1,14}$ are coefficients of ξ_1 with respect to that of $H_{1,2}\mathcal{G}\mathcal{X}$, and K denotes a 10×14 matrix defined by Table 4.11. The initial condition is given by:

$$h_1(0) = \left[h_{1,1}(0), \cdots, h_{1,10}(0)\right].$$

Note that K is surjective as a linear mapping from $H_{1,2}\mathcal{G}\mathcal{X}$. Therefore, for the vector $\left[h_{1,1}(0), \cdots, h_{1,10}(0)\right]$, there exists $(\overline{\xi_{1,1}}, \cdots, \overline{\xi_{1,14}})$ such that

$$K(\overline{\xi_{1,1}}, \cdots, \overline{\xi_{1,14}})^T = \left[h_{1,1}(0), \cdots, h_{1,10}(0)\right]^T$$

holds. Hence, by choosing such $(\overline{\xi_{1,1}}, \cdots, \overline{\xi_{1,14}})$, the above differential equation reduces to the form,

$$\frac{d}{dt}\begin{bmatrix} h_{1,1}(t) \\ \vdots \\ h_{1,10}(t) \end{bmatrix} = -\begin{bmatrix} h_{1,1}(0) \\ \vdots \\ h_{1,10}(0) \end{bmatrix}$$

whose solution is given by:

$$\begin{bmatrix} h_{1,1}(t) \\ \vdots \\ h_{1,10}(t) \end{bmatrix} = \begin{bmatrix} h_{1,1}(0) \\ \vdots \\ h_{1,10}(0) \end{bmatrix} - t\begin{bmatrix} h_{1,1}(0) \\ \vdots \\ h_{1,10}(0) \end{bmatrix}$$

Hence, $h_1(1) = h_1(0) - 1 \times h_1(0) = 0$.

This means that any 1st-order term of $(\mathbf{A}^1, \mathbf{v}^1)$ with the leading part $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]$ can be eliminated by a suitable transformation generated by the (1,2)-jet (R_1, Y_2). Hence, we have obtained $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]$ as the 1st-order normal form of this example.

■

By a similar argument to the above example, we can prove that the following holds for *any* k-jets in general: for a constrained system $(\mathbf{A}, \mathbf{v})$, we choose

$$B_k \triangleq \{H_{k,k+1}\mathcal{G}\mathcal{X}, a_0\}_k \subset H_k\mathcal{C}\mathcal{X}^r,$$

where a_0 is the 0-jet of $(\mathbf{A}, \mathbf{v})$. If B_k coincides with $H_k\mathcal{C}\mathcal{X}^r$ itself, then any k-th order part ($k \geq 1$) of

(A , v) can be eliminated by a suitable transformation. In fact, since $B_k = H_k \mathcal{C}\mathcal{X}^r$, there exists $\zeta_k \in H_{k,k+1} \mathcal{G}\mathcal{X}$ such that $\{\zeta_k , a_0\} = h_k(0)$ for any $h_k(0) \in H_k$ $\mathcal{C}\mathcal{X}^r$, and, by taking $\xi^k = \zeta_k$, the differential equation (4.5) becomes

$$\frac{d}{dt} h_k(t) = - \{\zeta_k , a_0\} = h_k(0) .$$

It follows from the solution

$$h_k(t) = h_k(0) - th_k(0) ,$$

that $h_k(1) = 0$.

For the above example, we can prove the following:

Proposition 4.12

For $k \geq 1$, any k-th order part of (A , v) with leading part $\left[\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} , \begin{bmatrix}1\\0\end{bmatrix}\right]$ can be eliminated. In other words, the *infinite-* order normal form is simply the leading part

$$\left[\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} , \begin{bmatrix}1\\0\end{bmatrix}\right]$$

itself.

Proof:

It suffices to show that B_k coincides with H_k $\mathcal{C}\mathcal{X}^1$. Recall that

$$(A_0 , v_0) = \left[\frac{\partial}{\partial y} \otimes dy , \frac{\partial}{\partial x}\right] ,$$

and that H_k ^{1} is spanned by

$$\left[x^m y^n \frac{\partial}{\partial x} \otimes dy , 0\right] , \left[0 , x^m y^n \frac{\partial}{\partial x}\right]$$

$$\left[x^m y^n \frac{\partial}{\partial y} \otimes dx , 0\right] , \left[0 , x^m y^n \frac{\partial}{\partial y}\right]$$

$$\left[x^m y^n \frac{\partial}{\partial y} \otimes dy , 0\right]$$

Recall also that $H_{k,k+1} \mathcal{G}\mathcal{X}$ is spanned by

$$\left\{ \begin{matrix} \left[x^m y^n \frac{\partial}{\partial x} \otimes dx \ , 0\right], & \left[x^m y^n \frac{\partial}{\partial y} \otimes dx \ , 0\right], & \left[0 \ , x^{m'} y^{n'} \frac{\partial}{\partial x}\right] \\ \left[x^m y^n \frac{\partial}{\partial x} \otimes dy \ , 0\right], & \left[x^m y^n \frac{\partial}{\partial y} \otimes dy \ , 0\right], & \left[0 \ , x^{m'} y^{n'} \frac{\partial}{\partial y}\right] \end{matrix} \right\}$$

where $m+n = k$ and $m' + n' = k+1$. Then,

$$\left\{\left[x^m y^n \frac{\partial}{\partial x} \otimes dx \ , 0\right], (A_0, v_0)\right\}_k = \left[0, x^m y^n \frac{\partial}{\partial x}\right]$$

$$\left\{\left[x^m y^n \frac{\partial}{\partial x} \otimes dy \ , 0\right], (A_0, v_0)\right\}_k = \left[x^m y^n \frac{\partial}{\partial x} \otimes dy \ , 0\right]$$

$$\left\{\left[x^m y^n \frac{\partial}{\partial y} \otimes dx \ , 0\right], (A_0, v_0)\right\}_k = \left[0 \ , x^m y^n \frac{\partial}{\partial y}\right]$$

$$\left\{\left[x^m y^n \frac{\partial}{\partial y} \otimes dy \ , 0\right], (A_0 \ , v_0)\right\}_k = \left[x^m y^n \frac{\partial}{\partial y} \otimes dy \ , 0\right]$$

$$\left\{\left[-n' x^{m'} y^{n'-1} \frac{\partial}{\partial x} \otimes dy \ , x^{m'} y^{n'} \frac{\partial}{\partial x}\right], (A_0, v_0)\right\}_k = \left[0 , m' x^{m'-1} y^{n'} \frac{\partial}{\partial x}\right]$$

$$\left\{\left[-m' x^{m'-1} y^{n'} \frac{\partial}{\partial y} \otimes dx \ , x^{m'} y^{n'} \frac{\partial}{\partial y}\right], (A_0, v_0)\right\}_k = \left[-n' x^{m'} y^{n'-1} \frac{\partial}{\partial y} \otimes dx \ , 0\right].$$

Therefore the linear map $H_{k,k+1} \mathcal{G}\mathcal{X} \to H_k \mathcal{C}\mathcal{X}^1$, $(R_k, Y_{k+1}) \to \left\{(R_k, Y_{k+1}), (A_0, v_0)\right\}$ is *surjective*.

This completes the proof.

■

As a generalization of the above argument, even if B_k does not coincide with $H_k \mathcal{C}\mathcal{X}^r$, we can obtain a theorem corresponding to the *Reduction Theorem* for vector field normal form (Theorem 5.4 in [14]) which reduces the normal form problem on $H_k \mathcal{C}\mathcal{X}^1$ to that on a subspace $\hat{B}_k$ *complementary* to B_k in $H_k \mathcal{C}\mathcal{X}^r$. To state this theorem, let π_k be the projection,

$$\pi_k : J_k \mathcal{C}\mathcal{X}^r \to \hat{B}_k \quad ,$$

along B_k.

Theorem 4.13 Reduction Theorem for constrained system normal forms

The kth order normal form problem

$$\frac{d}{dt} h_k(t) = -\{\xi^k, a^{k-1} + h_k(t)\}_k \tag{4.7}$$

on $H_k \mathcal{C}\mathcal{X}^r$ with

$$\{\xi^{k-1}, a^{k-1}\}^{k-1} = 0 \tag{4.8}$$

can be reduced to that on $\hat{B}_k$; namely,

$$\frac{d}{dt} \hat{b}_k(t) = -\pi_k \left[\{\xi^{k-1}, a^{k-1} + \hat{b}_k(t)\}_k\right] \tag{4.9}$$

with (4.8), where $\hat{b}_k(t) \in \hat{B}_k$. More precisely, if we arrive at some point in $\hat{B}_k$ by integrating (4.9) with (4.8) under the initial condition $\hat{b}_k(0)$, then we can also arrive there from $h_k(0)$ satisfying $\pi_k \left[h_k(0)\right] = \hat{b}_k(0)$, by integrating (4.7) with (4.8) for suitable ξ^k's.

The proof of this theorem is given in *Appendix VI (Part II).*

Let us pause to consider an example illustrating the use of the reduction theorem.

Example 4.14: Regular slow point

Consider the family of constrained systems on a 2-dimensional manifold, whose leading part (A_0, v_0) is equivalent to

$$\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right].$$

We may suppose that $a_0 = (A_0, v_0)$ itself is of this form without loss of generality. In a similar manner as *Example 4.10,* we consider the 1st-order normal form problem in $H_1 \mathcal{C}\mathcal{X}^1$. Recall that the vector space $H_1 \mathcal{C}\mathcal{X}^1$ is 10-dimensional, whose basis has been obtained earlier. Also $H_{1,2} \mathcal{G}\mathcal{X}$ is spanned by the basis derived earlier.

Consider the linear map

$$H_{1,2} \mathcal{G}\mathcal{X} \to H_1 \mathcal{C}\mathcal{X}^1$$

$$\xi_1 = (R_1, Y_2) \mapsto \{\xi_1, a_0\},$$

which is expressed in terms of *Table 4.15*. Thus, the image B_1 is spanned by,

$$\left\{ \begin{array}{l} \left[x \frac{\partial}{\partial x} \otimes dy\ , x \frac{\partial}{\partial x}\right],\ \left[y \frac{\partial}{\partial x} \otimes dy\ , y \frac{\partial}{\partial x}\right] \\ \left[x \frac{\partial}{\partial y} \otimes dy\ , x \frac{\partial}{\partial y}\right],\ \left[y \frac{\partial}{\partial y} \otimes dy\ , y \frac{\partial}{\partial y}\right] \\ \left[x \frac{\partial}{\partial y} \otimes dx\ , 0\right],\ \left[y \frac{\partial}{\partial y} \otimes dx\ , x \frac{\partial}{\partial y}\right],\ \left[0\ , y \frac{\partial}{\partial y}\right] \end{array} \right\}$$

and the complementary space $\hat{B}_1$ can be taken as the vector space spanned by,

$$\left\{ \left[0\ , x \frac{\partial}{\partial x}\right],\ \left[0\ , y \frac{\partial}{\partial x}\right],\ \left[0\ , x \frac{\partial}{\partial y}\right] \right\}.$$

Note that the projection $\pi_1 : H_1 \mathcal{C}\mathcal{X}^1 \rightarrow B_1$ maps, for example,

$$\pi_1 \left[x \frac{\partial}{\partial x} \otimes dy\ , 0\right] = \pi_1 \left[\left[x \frac{\partial}{\partial x} \otimes dy\ , x \frac{\partial}{\partial x}\right] - \left[0\ , x \frac{\partial}{\partial x}\right] \right]$$

$$= - \left[0, x \frac{\partial}{\partial x}\right].$$

Since every element $\hat{b}_1 \in \hat{B}_1$ can be written in the form

$$\hat{b}_1 = \alpha \left[0, x \frac{\partial}{\partial x}\right] + \beta \left[0\ , y \frac{\partial}{\partial x}\right] + \gamma \left[0\ , x \frac{\partial}{\partial y}\right],$$

the reduced 1st-order normal form problem becomes

$$\begin{aligned} \frac{d}{dt} \hat{b}_1(t) &= - \pi_1 \left[\{\xi^0\ , a^0 + \hat{b}_1(t)\}_1 \right] \\ &= - \pi_1 \left[\{\xi^0\ , \hat{b}_1(t)\}_1 \right] \end{aligned} \tag{4.10}$$

with $\{\xi^0\ , a^0\}^0 = 0$. Using the results from *Table 4.17*, we obtain

$$\left\{ \xi^0 \in J^{0,1}_{x_0} \mathcal{G}\mathcal{X}\ \mid \{\xi^0, a^0\}^0 = 0 \right\}$$

$$= \left\{ \left[0\ , x \frac{\partial}{\partial x}\right],\ \left[\frac{\partial}{\partial x} \otimes dy\ , - y \frac{\partial}{\partial x}\right],\ \left[\frac{\partial}{\partial y} \otimes dx\ , 0\right],\ \left[\frac{\partial}{\partial x} \otimes dx\ , 0\right] \right\}.$$

Hence, ξ^0 is given by

$$\xi^0 = A\left[0\,,x\,\frac{\partial}{\partial x}\right] + B\left[\frac{\partial}{\partial x}\otimes dy\,,-y\frac{\partial}{\partial x}\right] + C\left[\frac{\partial}{\partial y}\otimes dx\,,0\right] + D\left[\frac{\partial}{\partial x}\otimes dx\,,0\right].$$

Therefore (4.10) becomes

$$\frac{d}{dt}\left[\alpha(t)\left[0\,,x\,\frac{\partial}{\partial x}\right] + \beta(t)\left[0\,,y\,\frac{\partial}{\partial x}\right] + \gamma(t)\left[0\,,x\,\frac{\partial}{\partial y}\right]\right]$$

$$= -\pi_1\left[\left\{A\left[0\,,x\,\frac{\partial}{\partial x}\right] + B\left[\frac{\partial}{\partial x}\otimes dy\,,-y\frac{\partial}{\partial x}\right] + C\left[\frac{\partial}{\partial y}\otimes dx\,,0\right] + D\left[\frac{\partial}{\partial x}\otimes dx\,,0\right]\,,\right.\right.$$

$$\left.\left.\alpha(t)\left[0\,,x\,\frac{\partial}{\partial x}\right] + \beta(t)\left[0\,,y\,\frac{\partial}{\partial x}\right] + \gamma(t)\left[0\,,x\,\frac{\partial}{\partial y}\right]\right\}_1\right] \quad .$$

that is,

$$\frac{d}{dt}\,\alpha(t) = D\,\alpha(t)$$

$$\frac{d}{dt}\,\beta(t) = B\,\alpha(t) + (A+D)\,\beta(t)$$

$$\frac{d}{dt}\,\gamma(t) = C\,\alpha(t) - A\,\gamma(t)$$

The solution is given by,

$$\alpha(t) = \alpha(0)\,e^{Dt}$$

$$\beta(t) = e^{(A+D)t}\left[\beta(0) + \frac{B\,\alpha(0)}{D-(A+D)}\left\{e^{(D-(A+D))t} - 1\right\}\right]$$

$$= e^{(A+D)t}\left[\beta(0) + \frac{B\,\alpha(0)}{-A}\,(e^{-At}-1)\right]$$

$$\gamma(t) = e^{-At}\left[\gamma(0) + \frac{C\,\alpha(0)}{D+A}\left\{e^{(D+A)t}-1\right\}\right].$$

If $\alpha(0) \neq 0$, we can choose

$$\alpha(1) = \alpha(0)\,e^{D} = sign\;\alpha(0)(=\pm 1)$$

$$\beta(1) = 0$$

$$\gamma(1) = 0\,,$$

upon choosing

$$D = -\log|\alpha(0)|$$

$$B = \beta(0) \times \frac{A}{\alpha(0)(e^{-A}-1)}$$

$$C = -\gamma(0) \times \frac{D+A}{\alpha(0)(e^{D+A}-1)} ,$$

for arbitrary constant A, provided $A \neq 0$ and $A \neq -D$.

If $\alpha(0) = 0$ and $\beta(0) \neq 0$, $\gamma(0) \neq 0$, we can choose

$$\alpha(1) = 0$$

$$\beta(1) = sign\ \beta(0) = \pm 1$$

$$\gamma(1) = sign\ \gamma(0) = \pm 1 ,$$

upon choosing

$$D = -\log|\beta(0)| - \log|\gamma(0)|$$

$$A = \log|\gamma(0)| .$$

For the other case corresponding to $\beta(0) = 0$ and/or $\gamma(0) = 0$, we can normalize the non-zero coefficient β and/or γ, to ± 1. Hence, the 1st-order normal form problem for $a_0 = \left[\frac{\partial}{\partial y} \otimes dy\ ,\ \frac{\partial}{\partial y}\right]$ is solved as follow:

(i) Non-degenerate 1st-order normal form

$$(\mathbf{A}^1, \mathbf{v}^1) = \left[\frac{\partial}{\partial y} \otimes dy\ ,\ \pm x\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right]$$

$$= \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm x \\ 1 \end{bmatrix}\right].$$

(ii) $$(\mathbf{A}^1, \mathbf{v}^1) = \left[\frac{\partial}{\partial y} \otimes dy\ ,\ \pm y\frac{\partial}{\partial x} + (1 \pm x)\frac{\partial}{\partial y}\right]$$

$$= \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y \\ 1 \pm x \end{bmatrix}\right]$$

(iii) $(\mathbf{A}^1,\mathbf{v}^1) = \left[\frac{\partial}{\partial y} \otimes dy \ , \pm y \frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right]$

$$= \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y \\ 1 \end{bmatrix}\right]$$

(iv) $(\mathbf{A}^1,\mathbf{v}^1) = \left[\frac{\partial}{\partial y} \otimes dy \ , (1 \pm x) \frac{\partial}{\partial y}\right] = \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \pm x \end{bmatrix}\right]$

(v) $(\mathbf{A}^1,\mathbf{v}^1) = \left[\frac{\partial}{\partial y} \otimes dy \ , \frac{\partial}{\partial y}\right] = \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right]$

Let us proceed next to solve the 2nd-order normal form problem for the *non-degenerate* 1st-order normal form (i). The vector space $H_2 \mathcal{C}\mathcal{X}^1$ has a dimension equal to 15 and is spanned by the following basis:

$$\left\{
\begin{array}{l}
\left[x^2 \frac{\partial}{\partial x} \otimes dy \ , 0\right], \left[xy \frac{\partial}{\partial x} \otimes dy \ , 0\right], \left[y^2 \frac{\partial}{\partial x} \otimes dy \ , 0\right] \\
\left[x^2 \frac{\partial}{\partial y} \otimes dx \ , 0\right], \left[xy \frac{\partial}{\partial y} \otimes dx \ , 0\right], \left[y^2 \frac{\partial}{\partial y} \otimes dx \ , 0\right] \\
\left[x^2 \frac{\partial}{\partial y} \otimes dy \ , 0\right], \left[xy \frac{\partial}{\partial y} \otimes dy \ , 0\right], \left[y^2 \frac{\partial}{\partial y} \otimes dy \ , 0\right] \\
\left[0 \ , x^2 \frac{\partial}{\partial x}\right], \left[0 \ , xy \frac{\partial}{\partial x}\right], \left[0 \ , y^2 \frac{\partial}{\partial x}\right] \\
\left[0 \ , x^2 \frac{\partial}{\partial y}\right], \left[0 \ , xy \frac{\partial}{\partial y}\right], \left[0 \ , y^2 \frac{\partial}{\partial y}\right]
\end{array}
\right\}$$

Similarly, the vector space $H_{2,3} \mathcal{G}\mathcal{X}$ is spanned by

$$\left\{
\begin{array}{l}
\left[x^2 \frac{\partial}{\partial x} \otimes dx \ , 0\right], \left[xy \frac{\partial}{\partial x} \otimes dx \ , 0\right], \left[y^2 \frac{\partial}{\partial x} \otimes dx \ , 0\right] \\
\left[x^2 \frac{\partial}{\partial x} \otimes dy \ , 0\right], \left[xy \frac{\partial}{\partial x} \otimes dy \ , 0\right], \left[y^2 \frac{\partial}{\partial x} \otimes dy \ , 0\right] \\
\left[x^2 \frac{\partial}{\partial y} \otimes dx \ , 0\right], \left[xy \frac{\partial}{\partial y} \otimes dx \ , 0\right], \left[y^2 \frac{\partial}{\partial y} \otimes dx \ , 0\right]
\end{array}
\right.$$

$$\left.\begin{matrix} \left[x^2\frac{\partial}{\partial y}\otimes dy\ ,0\right],\ \left[xy\frac{\partial}{\partial y}\otimes dy\ ,0\right],\ \left[y^2\frac{\partial}{\partial y}\otimes dy\ ,0\right] \\ \left[0\ ,x^2\frac{\partial}{\partial x}\right],\ \left[0\ ,xy\frac{\partial}{\partial x}\right],\ \left[0\ ,y^2\frac{\partial}{\partial x}\right] \\ \left[0\ ,x^2\frac{\partial}{\partial y}\right],\ \left[0\ ,xy\frac{\partial}{\partial y}\right],\ \left[0\ ,y^2\frac{\partial}{\partial y}\right] \end{matrix}\right\}$$

Hence, the dimension of $H_{2,3}\mathcal{G}\mathcal{X}$ is equal to 20.

The linear map

$$H_{2,3}\mathcal{G}\mathcal{X}\to H_2\mathcal{C}\mathcal{X}^1$$

defined by

$$\xi_2=(R_2,Y_3)\ \to\ \{\xi_2,a_0\}$$

is represented by the results calculated in *Table 4.18*. Hence the image B_2 is spanned by

$$\left\{\begin{matrix} \left[x^2\frac{\partial}{\partial x}\otimes dy\ ,x^2\frac{\partial}{\partial x}\right],\ \left[xy\frac{\partial}{\partial x}\otimes dy\ ,xy\frac{\partial}{\partial x}\right],\ \left[y^2\frac{\partial}{\partial x}\otimes dy\ ,y^2\frac{\partial}{\partial x}\right] \\ \left[x^2\frac{\partial}{\partial y}\otimes dy\ ,x^2\frac{\partial}{\partial y}\right],\ \left[xy\frac{\partial}{\partial y}\otimes dy\ ,xy\frac{\partial}{\partial y}\right],\ \left[y^2\frac{\partial}{\partial y}\otimes dy\ ,0\right] \\ \left[x^2\frac{\partial}{\partial y}\otimes dx\ ,0\right],\ \left[-2xy\frac{\partial}{\partial y}\otimes dx\ ,x^2\frac{\partial}{\partial y}\right],\ \left[-y^2\frac{\partial}{\partial y}\otimes dx\ ,2xy\frac{\partial}{\partial y}\right] \\ \left[0\ ,y^2\frac{\partial}{\partial y}\right] \end{matrix}\right\}$$

Let us choose the complementary space $\hat{B}_2$ as the linear space spanned by,

$$\left\{\begin{matrix} \left[0\ ,x^2\frac{\partial}{\partial x}\right],\ \left[0\ ,xy\frac{\partial}{\partial x}\right],\ \left[0\ ,y^2\frac{\partial}{\partial x}\right] \\ \left[0\ ,x^2\frac{\partial}{\partial y}\right],\ \left[0\ ,xy\frac{\partial}{\partial y}\right], \end{matrix}\right\}$$

where the projection $\pi_2:H_2\mathcal{C}\mathcal{X}^1\to\hat{B}_2$ maps

$$\pi_2\left[\left[x^2\frac{\partial}{\partial x}\otimes dy\ ,0\right]\right]=\pi_2\left[\left[x^2\frac{\partial}{\partial x}\otimes dy\ ,x^2\frac{\partial}{\partial x}\right]-\left[0\ ,x^2\frac{\partial}{\partial x}\right]\right]$$

$$=-\left[0\ ,x^2\frac{\partial}{\partial x}\right],$$

and so on. We must next verify the condition (4.8); namely,

$$\{\xi^1,a^1\}^1=0$$

or equivalently, the conditions

$$\{\xi_0,a_0\}=0$$

$$\{\xi_1,a_0\}+\{\xi_0,a_1\}=0\quad,$$

where $\xi^1=\xi_0+\xi_1$. From the first condition, we can again choose from *Table 4.17*

$$\xi_0=A\left[0\ ,x\frac{\partial}{\partial x}\right]+B\left[\frac{\partial}{\partial x}\otimes dy\ ,-y\frac{\partial}{\partial x}\right]+C\left[\frac{\partial}{\partial y}\otimes dx\ ,0\right]$$

$$+D\left[\frac{\partial}{\partial x}\otimes dx\ ,0\right].$$

Using Table 4.15 and Table 4.19, the linear subspace satisfying $\{\xi^1,a^1\}^1=0$ is spanned by,

$$\left\{\begin{array}{l}\left[x\frac{\partial}{\partial x}\otimes dx\ ,0\right],\ \left[y\frac{\partial}{\partial x}\otimes dx\ ,0\right]\\ \left[x\frac{\partial}{\partial y}\otimes dx\ ,0\right],\ \left[y\frac{\partial}{\partial y}\otimes dx\ ,0\right]\\ \left[x\frac{\partial}{\partial x}\otimes dy\ ,-xy\frac{\partial}{\partial x}\right],\ \left[2y\frac{\partial}{\partial x}\otimes dy\ ,-y^2\frac{\partial}{\partial x}\right]\\ \left[0\ ,x^2\frac{\partial}{\partial x}\right],\ \left[0\ ,x\frac{\partial}{\partial x}\right]\end{array}\right\}$$

Hence, the reduced 2nd-order normal form problem is given by,

$$\frac{d}{dt}\hat{b}_2(t)=-\pi_2\left[\{\xi^1\ ,a^1+\hat{b}_2(t)\}_2\right]$$

$$=-\pi_2\left[\{\xi_0\ ,\hat{b}_2(t)\}_2+\{\xi_1\ ,a^1\}_2\right]$$

where

$$\hat{b}_2 = a\left[0\,,x^2\frac{\partial}{\partial x}\right] + b\left[0\,,xy\frac{\partial}{\partial x}\right] + c\left[0\,,y^2\frac{\partial}{\partial x}\right] + d\left[0\,,x^2\frac{\partial}{\partial y}\right] + e\left[0\,,xy\frac{\partial}{\partial y}\right]$$

and

$$\xi_0 = A\left[0\,,x\frac{\partial}{\partial x}\right]$$

$$\xi_1 = C_1\left[x\frac{\partial}{\partial x}\otimes dx\;,0\right] + C_2\left[y\frac{\partial}{\partial x}\otimes dx\;,0\right] + C_3\left[x\frac{\partial}{\partial y}\otimes dx\;,0\right] + C_4\left[y\frac{\partial}{\partial y}\otimes dx\;,0\right]$$
$$+ C_5\left[x\frac{\partial}{\partial x}\otimes dy\;,-xy\frac{\partial}{\partial x}\right] + C_6\left[2y\frac{\partial}{\partial x}\otimes dy\;,-y^2\frac{\partial}{\partial x}\right] + C_7\left[0\,,x^2\frac{\partial}{\partial x}\right].$$

The above differential equation, thus, becomes

$$\frac{da(t)}{dt} = C_1 - C_7 - Aa$$

$$\frac{db(t)}{dt} = C_2$$

$$\frac{dc(t)}{dt} = C_6 + Ac$$

$$\frac{dd(t)}{dt} = C_3 - 2Ad$$

$$\frac{de(t)}{dt} = C_4 - Ae\;.$$

By choosing

$$A = 0$$
$$C_1 - C_7 = -a(0)$$
$$C_2 = -b(0)$$
$$C_6 = -c(0)$$
$$C_3 = -d(0)$$
$$C_4 = -e(0)$$

We can eliminate all coefficients of $\hat{b}_2$. The 2nd-order normal form with non-degenerate $a^1 = \left[\frac{\partial}{\partial y}\otimes dy\;,\pm x\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right]$ is obtained as follow:

Non-degenerate 2nd-order normal form

$$a^2 = \left[\frac{\partial}{\partial y} \otimes dy \ , \pm x \frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right] = \left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm x \\ 1 \end{bmatrix}\right]. \tag{4.11}$$

Following the same algorithm as before, we can continue to calculate the higher-order normal forms. Moreover, we can prove that the *infinite-* order normal form is given by the same form as (4.11). We can also extend this algorithm to the n-dimensional case, instead of dimension 2. Such an extension is given in *Appendix VII* (Part II).

■

Let us now make an attempt to classify the 2-dimensional normal forms for constrained systems. In a similar manner as the case for vector fields, they are assumed to have a *specified* leading part. For the most *non-degenerate case*, that is, (i): $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]$, we have already obtained *Proposition 4.12;* the infinite-order normal form is given by (i) itself. Our next proposition gives a classification of 2-jet for constrained systems whose leading part is equivalent to: (ii): $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right]$.

Proposition 4.17

If the leading part of a two-dimensional constrained system $(\mathbf{A}, \mathbf{v})$ of corank 1 is equivalent to (ii) in *Proposition 4.8,* then its 1st-order normal form is given by one of the following forms:

(a_1) $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm x \\ 1 \end{bmatrix}\right]$ (a_2) $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y \\ 1 \pm x \end{bmatrix}\right]$ (a_3) $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y \\ 1 \end{bmatrix}\right]$

(a_4) $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \pm x \end{bmatrix}\right]$ (a_5) $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right]$.

Moreover, if the 1-jet is equivalent to (a_1), its *infinite-order* normal form is (a_1) itself. If the 1-jet is equivalent to (a_2), (a_3), (a_4), or (a_5); respectively, then the *non-degenerate* 2nd-order normal form is given by,

(a_2') $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y + ax^2 \\ 1 \pm x \end{bmatrix}\right]$ (a_3') $\left[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \pm y + ax^2 \\ 1 \pm x^2 \end{bmatrix}\right]$

$$(a_4')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x^2+ay^2\\1\pm x\end{bmatrix}\right]\qquad (a_5')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x^2\pm y^2\\1+axy\end{bmatrix}\right]$$

The above result is obtained by a direct calculation as in *Example 4.14*. Finally, let us consider constrained systems whose leading part is equivalent to (iii): $\left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}\right]$.

Proposition 4.18

If the leading part of a two-dimensional constrained system (**A**, **v**) of corank 1 is equivalent to (iii) in *Proposition 4.8*, then its 1st-order normal form is given by one of the following forms:

$$(b_1)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x\\ay\end{bmatrix}\right]\quad (b_2)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm y\\\pm x\end{bmatrix}\right]\quad (b_3)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm y\\0\end{bmatrix}\right]$$

$$(b_4)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}0\\\pm x\end{bmatrix}\right]\quad (b_5)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}0\\ay\end{bmatrix}\right]\quad (b_6)\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}\right],$$

where a is a constant. Moreover if the 1-jet of the constrained system is equivalent to (b_1), (b_2), (b_3), (b_4), or (b_5); respectively, then the *non-degenerate* 2nd-order normal form is given by,

$$(b_1')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x\\ay\end{bmatrix}\right]\quad (b_2')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm y\pm x^2\\\pm x\end{bmatrix}\right]$$

$$(b_3')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm y\pm x^2\\ax^2\end{bmatrix}\right]\quad (b_4')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x^2+axy\pm y^2\\\pm x\end{bmatrix}\right]$$

$$(b_5')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x^2\\ay\pm xy\pm y^2\end{bmatrix}\right]\quad (b_6')\quad \left[\begin{bmatrix}0&0\\0&1\end{bmatrix},\begin{bmatrix}\pm x^2\pm y^2\\\pm xy+ay^2\end{bmatrix}\right]$$

■

For constrained systems of dimension greater than 3, or those of corank more than 2, we can, in principle, compute their normal forms as in the 2-dimensional systems of corank 1. However, the computation becomes increasingly more tedious and involved. See Oka [22] for the results.

REFERENCES

1. L. O. Chua, C. A. Desoer, and E. S. Kuh, *Linear and Nonlinear Circuits*, McGraw-Hill, New York, NY, 1987.
2. L. O. Chua and G. R. Alexander, "The effects of parasitic reactances on nonlinear networks," *IEEE Trans. on Circuit Theory*, vol. 18, no. 5, pp. 520-532, September 1971.
3. L. O. Chua and P. M. Lin, *Computer-Aided Analysis of Electronic Circuits: Algorithms and Computation Techniques*, Prentice-Hall, Englewood Cliff, NJ, 1975.
4. T. S. Parker and L. O. Chua, "A computer-assisted study of forced relaxation oscillations," *IEEE Trans. on Circuits and Systems*, vol. CAS-30, pp. 518-533, August 1983.
5. S. S. Sastry and C. A. Desoer, "Jump behavior of circuits and systems," *IEEE Trans. on Circuits and Systems*, vol. CAS-28, no. 12, pp. 1109-1124, 1981.
6. B. C. Haggman and P. R. Bryant, "Solutions of singular constrained differential equations: a generalization of circuits containing capacitor-only loops and inductor-only cutset," *IEEE Trans. on Circuits and Systems*, vol. CAS-31, no. 12, pp. 1015-1029, 1984.
7. G. Ikegami, "On network perturbations of electrical circuits and singular perturbation of dynamical systems," in *Chaos, Fractals and Dynamics*, edited by P. Fischer and W. R. Smith, Dekker, New York, pp. 197-212, 1985.
8. G. Ikegami, "Singular perturbations of constrained systems," in *Dynamical Systems and Nonlinear Oscillations*, Edited by G. Ikegami, World Scientific Publication, Singapore, pp. 27-49, 1985.
9. V. I. Arnold, *Geometric Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York, 1983.
10. F. Takens, "Singularities of vector fields," *Publ. Math. IHES*, vol. 43, pp. 47-100, 1974.
11. F. Takens, "Constrained equations: a study of implicit differential equations and their discontinuous solutions," in *Lecture Notes in Math.*, Springer-Verlag, vol. 525, pp. 143-234, 1976.
12. F. Takens, "Implicit differential equations: some open problems," in *Lecture Notes in Math.*, Springer-Verlag, vol. 535, pp. 237-253, 1976.
13. S. Ushiki, "Normal forms for singularities of vector fields," *Japan J. Applied Mathematics*, vol. 1, pp. 1-37, 1984.
14. L. O. Chua and H. Kokubu, "Normal forms for nonlinear vector fields--Part I: Theory and Algorithm," Electronics Research Laboratory, Memorandum UCB/ERL M87/81, University of California, Berkeley, July 1, 1987.
15. L. O. Chua and H. Kokubu, "Normal forms for nonlinear vector fields--Part II: Applications," Electronics

Research Laboratory, Memorandum UCB/ERL M87/82, University of California, Berkeley, CA 94720, August 1, 1987.

16. N. Fenichel, "Geometric singular perturbation theory for ordinary differential equations," *J. Diff. Equations*, vol. 31, pp. 53-98, 1979.

17. E. F. Mishchenko and N. Kh. Rozov, *Differential Equations with Small Parameters and Relaxation Oscillations*, Plenum Press, New York, NY, 1980.

18. K. W. Chang and F. A. Howes, *Nonlinear Singular Perturbation Phenomena: Theory and Application*, Springer-Verlag, New York, 1984.

19. A. K. Zvonkin and M. A. Shubin, "Non-standard analysis and singular perturbations of ordinary differential equations," *Russ. Math. Surveys*, vol. 39, pp. 77-127, 1984.

20. G. Ikegami, "Singular perturbations in foliations," preprint.

21. L. O. Chua, "Dynamic nonlinear networks: state-of-the-art," *IEEE Trans. on Circuits and Systems*, vol. CAS-27, pp. 1014-1044, November 1980.

22. H. Oka, "Constrained systems, characteristic surfaces, and normal forms," *Japan J. Appl. Math.*, vol. 4, pp. 393-431, 1987.

FIGURE CAPTIONS

Fig. 1. Phase portrait of the Van der Pol equation for very small ε. The portion of the orbits with *double arrowheads* indicate a *rapid* motion whose velocity tends to infinity as $\varepsilon \rightarrow 0$.

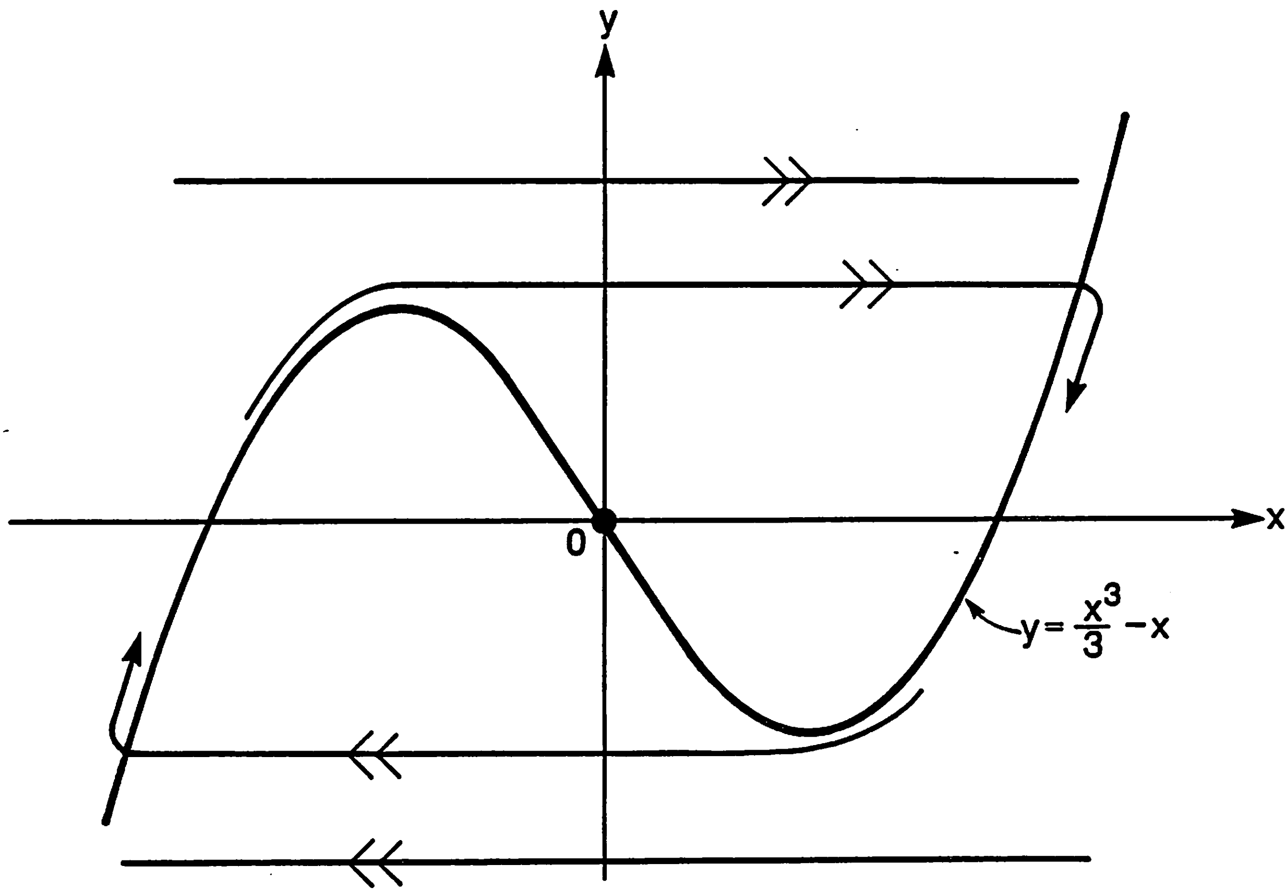
Fig. 2. Illustration of a bundle endomorphism.

Fig. 3. An illustration of a generalized vector field $(\mathbf{A}, \mathbf{v})$.

Fig. 4. (a) Phase portrait of (2.15). (b) Family of solution of (2.15) consisting of parabolas converging to $x = 0$ at a *finite* time t.

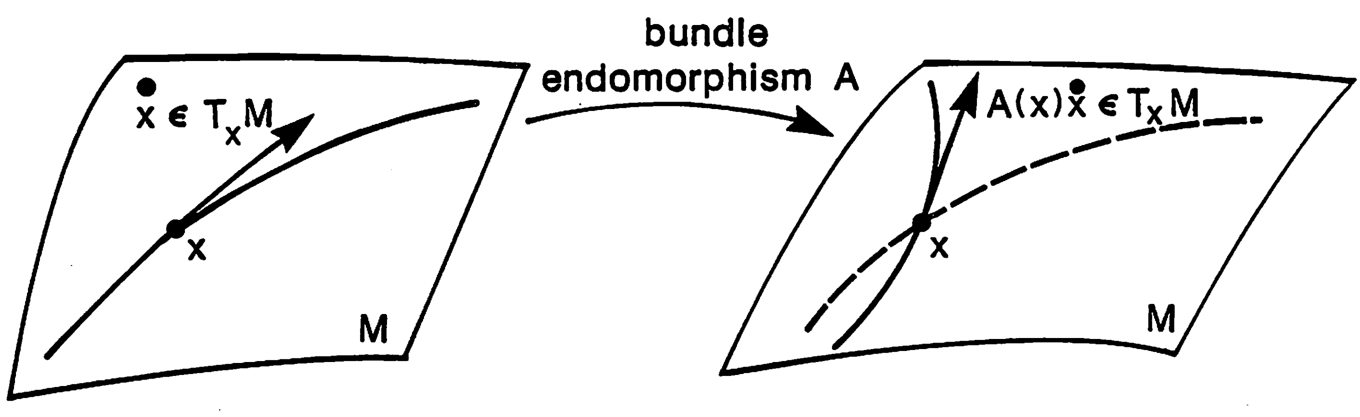
Fig. 5. (a) Phase portrait of (2.17). (b) Family of solutions of (2.17) consisting of parallel straight lines with a slope equal to -1.

Fig. 6. Phase portrait of (2.24) for small $\varepsilon_1 = \varepsilon_2 > 0$. Orbits with a *double arrowhead* denote *rapid* motion.



y
x
0
$y=\frac{x^3}{3}-x$

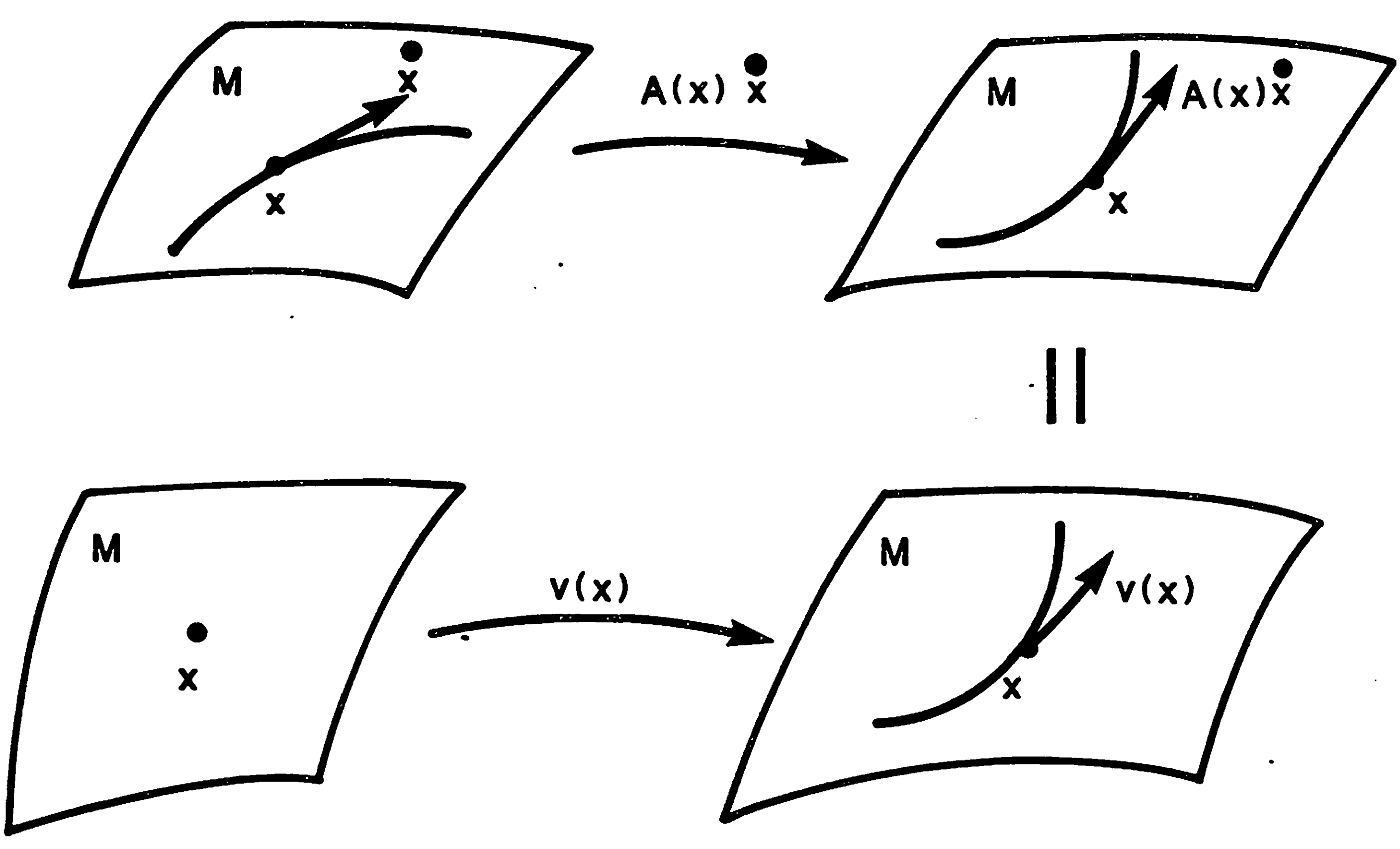
Fig.1



bundle
endomorphism A
$\dot{x} \in T_xM$
x
M
$A(x)\dot{x} \in T_xM$
x
M

Fig.2

(a)



(b)

Fig.3

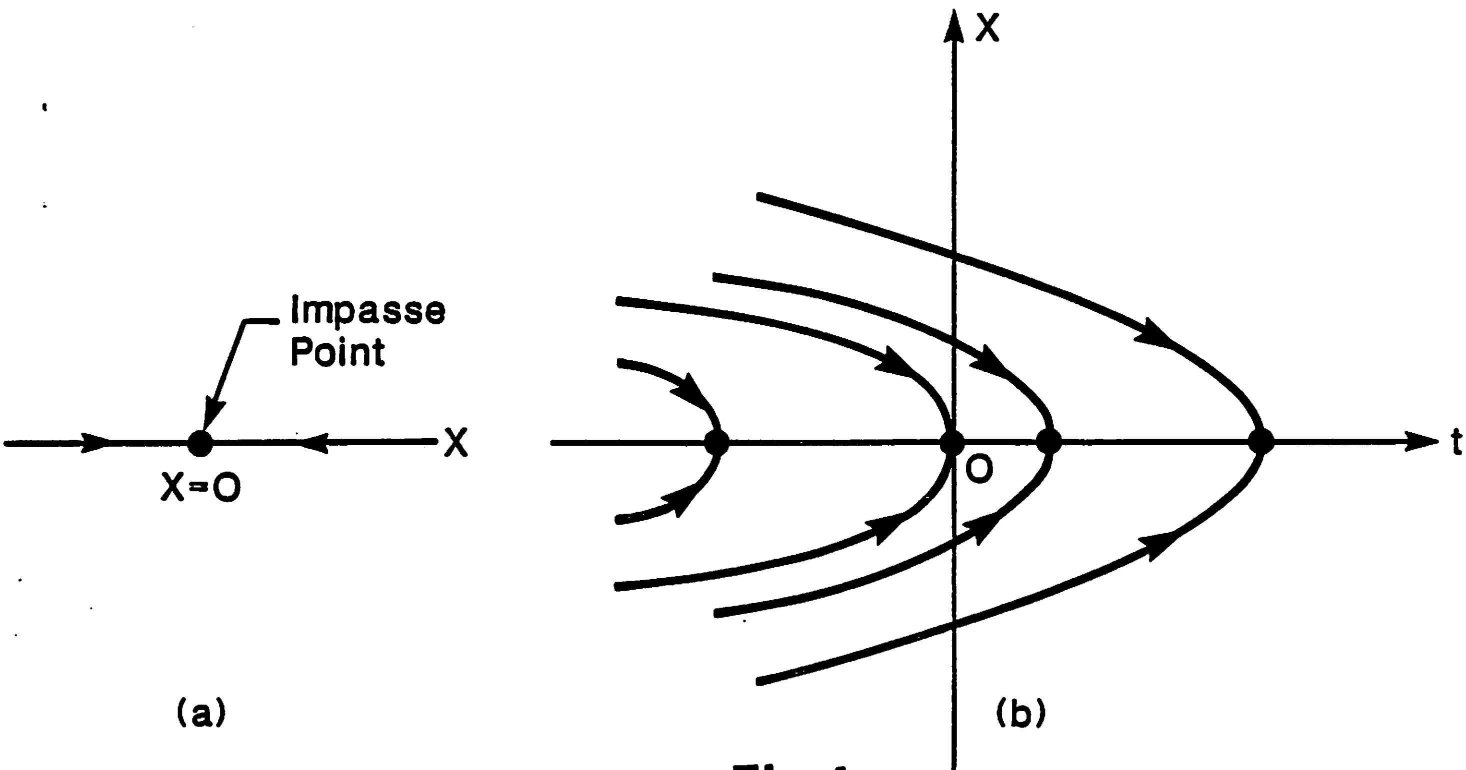


Fig.4

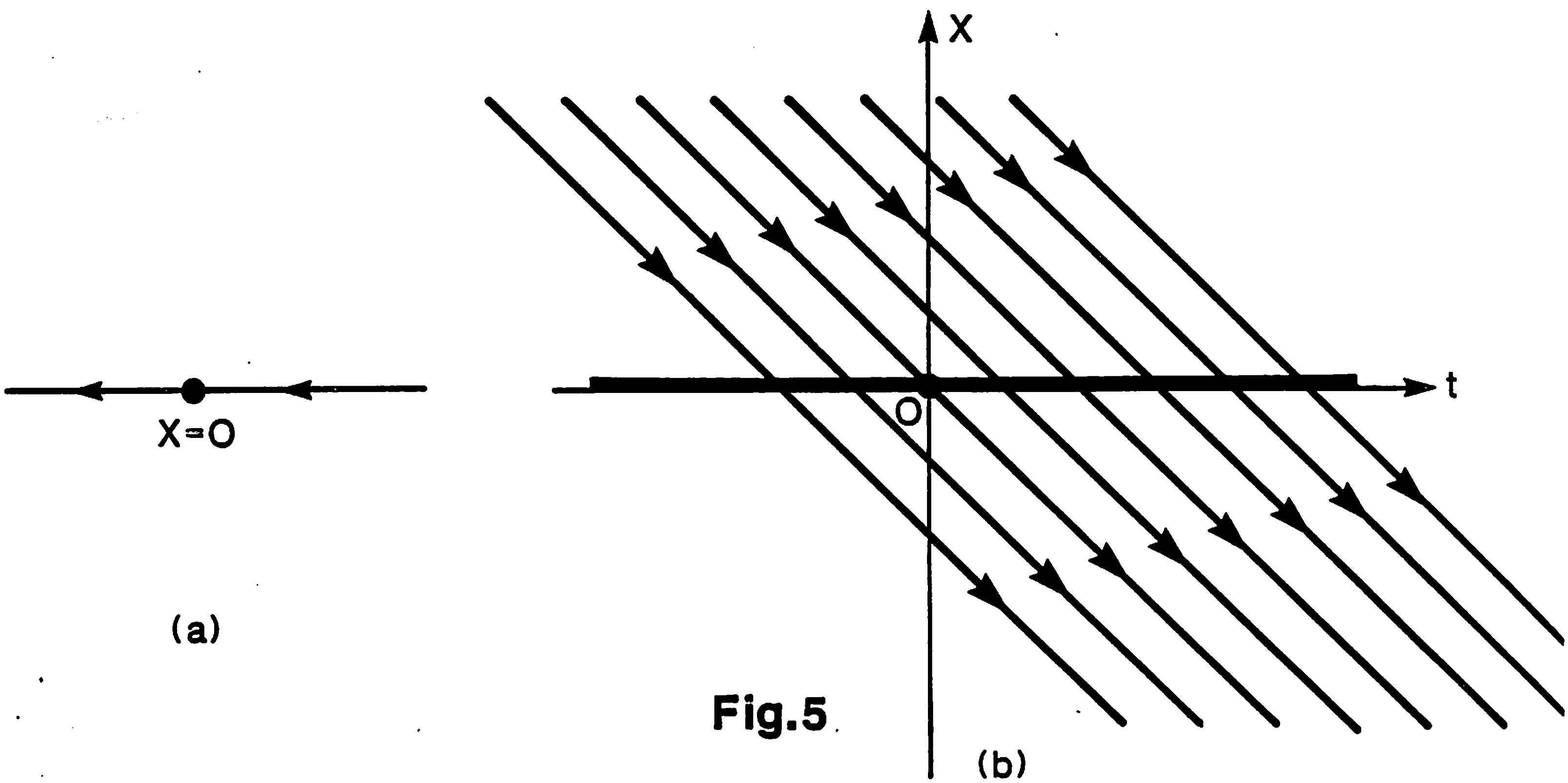


Fig.5

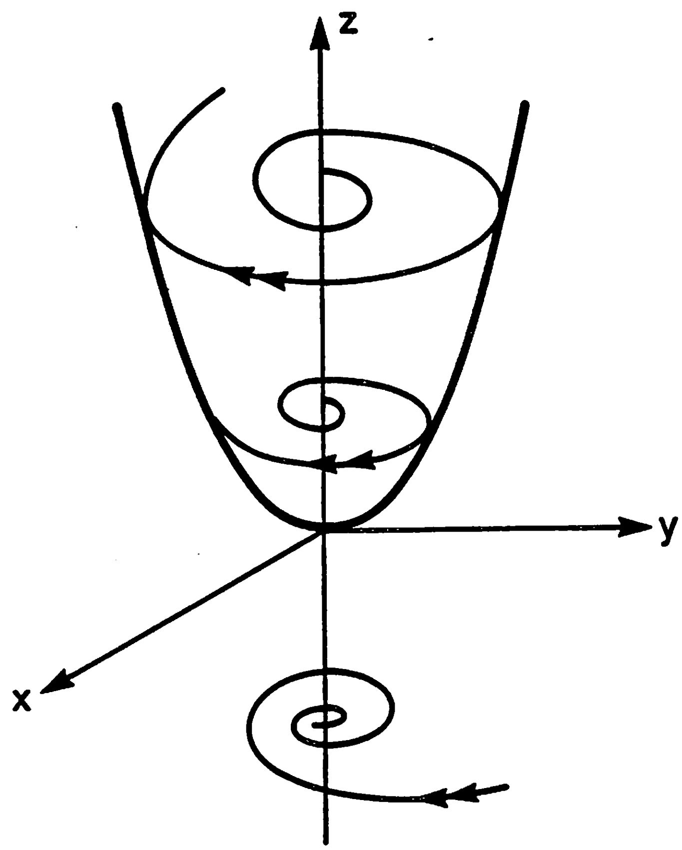


Fig.6

Table 4.11. Table of $\{\xi_1, a_0\}$, $\xi_1 \in H_{1,2}\mathcal{G}\mathcal{X}^1$, $a_0 = \left[\begin{bmatrix} 0 & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]$

$H_1\mathcal{C}\mathcal{X}^1$ \ $H_{1,2}\mathcal{G}\mathcal{X}$	$x\,\partial/\partial x \otimes dx$	$y\,\partial/\partial x \otimes dx$	$x\,\partial/\partial x \otimes dy$	$y\,\partial/\partial x \otimes dy$	$x\,\partial/\partial y \otimes dx$	$y\,\partial/\partial y \otimes dx$	$x\,\partial/\partial y \otimes dy$	$y\,\partial/\partial y \otimes dy$	$x^2\,\partial/\partial x$	$xy\,\partial/\partial x$	$y^2\,\partial/\partial x$	$x^2\,\partial/\partial y$	$xy\,\partial/\partial y$	$y^2\,\partial/\partial y$
$x\frac{\partial}{\partial x} \otimes dy$			1							1				
$y\frac{\partial}{\partial x} \otimes dy$				1							2			
$x\frac{\partial}{\partial y} \otimes dx$												2		
$y\frac{\partial}{\partial y} \otimes dx$													1	
$x\frac{\partial}{\partial y} \otimes dy$							1							
$y\frac{\partial}{\partial y} \otimes dy$								1						
$x\frac{\partial}{\partial x}$	1								-2					
$y\frac{\partial}{\partial x}$		1								-1				
$x\frac{\partial}{\partial y}$					1							-2		
$y\frac{\partial}{\partial y}$						1							-1	

Table 4.15. Calculations for $\{\xi_1, a_0\}$, $a_0 = \left[\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0\\ 1\end{bmatrix}\right]$

$H_1\mathcal{C}\mathcal{X}^1$ / $H_{1,2}\mathcal{G}\mathcal{X}$	$x\frac{\partial}{\partial x}\otimes dx$	$y\frac{\partial}{\partial x}\otimes dx$	$x\frac{\partial}{\partial x}\otimes dy$	$y\frac{\partial}{\partial x}\otimes dy$	$x\frac{\partial}{\partial y}\otimes dx$	$y\frac{\partial}{\partial y}\otimes dx$	$x\frac{\partial}{\partial y}\otimes dy$	$y\frac{\partial}{\partial y}\otimes dy$	$x^2\frac{\partial}{\partial x}$	$xy\frac{\partial}{\partial x}$	$y^2\frac{\partial}{\partial x}$	$x^2\frac{\partial}{\partial y}$	$xy\frac{\partial}{\partial y}$	$y^2\frac{\partial}{\partial y}$
$x\frac{\partial}{\partial x}\otimes dy$			1							1				
$y\frac{\partial}{\partial x}\otimes dy$				1							2			
$x\frac{\partial}{\partial y}\otimes dx$												-2		
$y\frac{\partial}{\partial y}\otimes dx$													-1	
$x\frac{\partial}{\partial y}\otimes dy$							1							
$y\frac{\partial}{\partial y}\otimes dy$								1						
$x\frac{\partial}{\partial x}$			1							1				
$y\frac{\partial}{\partial x}$				1							2			
$x\frac{\partial}{\partial y}$							1						1	
$y\frac{\partial}{\partial y}$								1						2

Table 4.17. Calculation for $\{\xi_0, a_0\}$

$\frac{\partial}{\partial x} \otimes dy$		1				1		
$\frac{\partial}{\partial y} \otimes dx$							-1	
$\frac{\partial}{\partial y} \otimes dy$				1				
$\frac{\partial}{\partial x}$		1				1		
$\frac{\partial}{\partial y}$				1				1
$H_0\mathcal{CX}^1$ / $H_{0,1}\mathcal{GX}$	$\frac{\partial}{\partial x} \otimes dx$	$\frac{\partial}{\partial x} \otimes dy$	$\frac{\partial}{\partial y} \otimes dx$	$\frac{\partial}{\partial y} \otimes dy$	x ($\frac{\partial}{\partial x}$)	y ($\frac{\partial}{\partial x}$)	x ($\frac{\partial}{\partial y}$)	y ($\frac{\partial}{\partial y}$)

Table 4.18. Calculations for $\{\xi_2, a_0\}$

$H_2\mathcal{CX}^1$ \ $H_{2,3}\mathcal{GX}$	$\frac{\partial}{\partial x}\otimes dx$			$\frac{\partial}{\partial x}\otimes dy$			$\frac{\partial}{\partial y}\otimes dx$			$\frac{\partial}{\partial y}\otimes dy$			$\frac{\partial}{\partial x}$				$\frac{\partial}{\partial y}$			
	x^2	xy	y^2	x^2	xy	y^2	x^2	xy	y^2	x^2	xy	y^2	x^3	x^2y	xy^2	y^3	x^3	x^2y	xy^2	y^3
$x^2\,\frac{\partial}{\partial x}\otimes dy$				1										1						
$xy\,\frac{\partial}{\partial x}\otimes dy$					1										2					
$y^2\,\frac{\partial}{\partial x}\otimes dy$						1										3				
$x^2\,\frac{\partial}{\partial y}\otimes dx$																	-3			
$xy\,\frac{\partial}{\partial y}\otimes dx$																		-2		
$y^2\,\frac{\partial}{\partial y}\otimes dx$																			-1	
$x^2\,\frac{\partial}{\partial y}\otimes dy$										1										
$xy\,\frac{\partial}{\partial y}\otimes dy$											1									
$y^2\,\frac{\partial}{\partial y}\otimes dy$												1								
$x^2\,\frac{\partial}{\partial x}$				1										1						
$xy\,\frac{\partial}{\partial x}$					1										2					
$y^2\,\frac{\partial}{\partial x}$						1										3				
$x^2\,\frac{\partial}{\partial y}$										1								1		
$xy\,\frac{\partial}{\partial y}$											1								2	
$y^2\,\frac{\partial}{\partial y}$												1								3

Table 4.19.

$x\frac{\partial}{\partial x} \otimes dy$				
$y\frac{\partial}{\partial x} \otimes dy$				
$x\frac{\partial}{\partial y} \otimes dx$				
$y\frac{\partial}{\partial y} \otimes dx$				
$x\frac{\partial}{\partial y} \otimes dy$				
$y\frac{\partial}{\partial y} \otimes dy$				
$x\frac{\partial}{\partial x}$				1
$y\frac{\partial}{\partial x}$		1		
$x\frac{\partial}{\partial y}$			1	
$y\frac{\partial}{\partial y}$				
$H_1\mathcal{CX}^1$	A	B	C	D