論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Analysis and Control of Nonlinear Systems with a Time Scale Transformation
著者(和文)	三平満司
Author(English)	MITSUJI SAMPEI
出典(和文)	学位:工学博士, 学位授与機関:東京工業大学, 報告番号:甲第1820号, 授与年月日:1987年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor of Engineering, Conferring organization: Tokyo Institute of Technology, Report number:甲第1820号, Conferred date:1987/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Analysis and Control of Nonlinear Systems with a Time Scale Transformation

by

Mitsuji Sampei

Tokyo Institute of Technology
Department of Control Engineering

Analysis and Control of Nonlinear Systems with a Time Scale Transformation

Mitsuji SAMPEI
(supervised by Prof. K.Furuta)

ABSTRACT

In this thesis, we will introduce a time scale transformation into nonlinear system theory. We will show how a time scale transformation can allow us to investigate the intrinsic structure of nonlinear systems.

The time scale transformation is defined as follows. A new time scale τ is defined using the actual time scale t as

$$\frac{dt}{d\tau} = s(x) > 0$$

for some smooth function s(x). Using this time scale, the system

$$\frac{dx}{dt} = f(x) + g(x)u$$

can be expressed in the time scale τ as

$$\frac{dx}{d\tau} = s(x)\ f(x) + s(x)\ g(x)u\ .$$

This time scale transformation preserves the system's stability and the state's curve in state space.

We will introduce the notion of weakly invariant distribution in order to study invariant structure in a transformed time scale. A weakly invariant distribution will allow us to obtain Kalman-like decompositions in reachable/unreachable parts and/or observable/unobservable parts in the transformed time scale. Weakly controlled invariance will also be introduced and used to solve the wide-sense disturbance decoupling problem. To solve this problem, we must seek a feedback law such that the disturbance will not effect the

output's curve in output space.

We will investigate the input-state linearization problem using a time scale transformation. We will obtain a class of nonlinear systems which can be linearized in the transformed time scale. The time scale in which the system can be linearized will be obtained as the solution to partial differential equations. Since the time scale transformation preserves the system's stability, this method can be used to design a stabilizing controller.

We will also apply the time scale transformation to controller design. Even for linear systems, the time scale transformation will allow us to design a nonlinear controller that will satisfy a needed specification such as to avoid an excessively high amplitude of input. We will design both a controller for a robot that will achieve good performance even in the neighborhood of a singular point, and also a trajectory tracking controller for a vehicle.

ACKNOWLEDGEMENTS

I would like to express my greatest appreciation to my supervisor Professor Katsuhisa Furuta for his patient instruction and guidance during the period of my research. His leadership and constant encouragement have been invaluable to me.

I am grateful to Professor Shinji Hara and Mr. Kazuhiro Kosuge of the department of Control Engineering for many stimulating discussions and for their encouragement.

Many thanks are also due to Miss Lillian Overman of Japan E.M. for her patience in correcting my English.

Finally, it is pleasure to acknowledge the hospitality and encouragement of Mrs. Nishihara and all the members of Furuta and Hara Laboratories.

CONTENTS

I. INTRODUCTION

In this thesis, we will introduce a time scale transformation and use it to analyze the intrinsic structures of nonlinear systems: the invariant structure and the linearization problem in a transformed time scale. We will also use the transformation to design a nonlinear controller.

1-1. Background

If all physical phenomena could be expressed in a linear form, e.x. in a linear state equation of the form

$$\dot{x} = Ax + Bu$$

$$y = Cx ,$$

it would not be necessary to study nonlinear systems. Almost all physical phenomena, however, exhibit some degree of nonlinearity which we must overcome in order to control physical systems.

A common method to control nonlinear systems is as follows. Firstly, one approximates the dynamic model of the system as a linear system in the neighborhood of an equilibrium point. One then designs a robust controller for the linear system so that the resultant system will be stable even though we have neglected the nonlinearity of the original system. This strategy is only successful if it is applied to systems which are only slightly nonlinear. It has been used to stabilize an inverted pendulum (Mori et al.[78]), a double inverted pendulum (Furuta et al.[79]) and other problems. It is, however, almost impossible to successfully apply this strategy to highly nonlinear systems.

Consider a system expressed in bilinear form

$$\dot{x} = Ax + uBx .$$

This system cannot be efficiently linearized around the equilibrium point x=0 because the input u will not effect dx/dt at x=0.

Even though we can linearize the system approximately, this method will stabilize it only in a small neighborhood of the equilibrium point. Thus, if we need to manipulate a system in a large area of state space, for example robot control, we cannot use this method to achieve desirable performance. It is for this reason that we turn our attention to nonlinear systems theory.

A system's nonlinearity can be roughly classified into two types: smooth and non-smooth nonlinearity. Non-smooth nonlinearity is of a type which cannot be expressed by a smooth function of the state; saturation, hysteresis, dead zone , on-off element etc. belong to this type of nonlinearity. In order to study the stability of systems having this type of nonlinearity, several methods have been developed: Lyapnov function[56], hyper stability[57], circle criterion[58], conic sector condition(Zames[59][60]) and sector stability criterion(Safonov[61]). If the nonlinearity is memoryless and the system is configured as in the following diagram, then the describing function method could also be useful to analyze the existence and the frequency of the limit cycle.

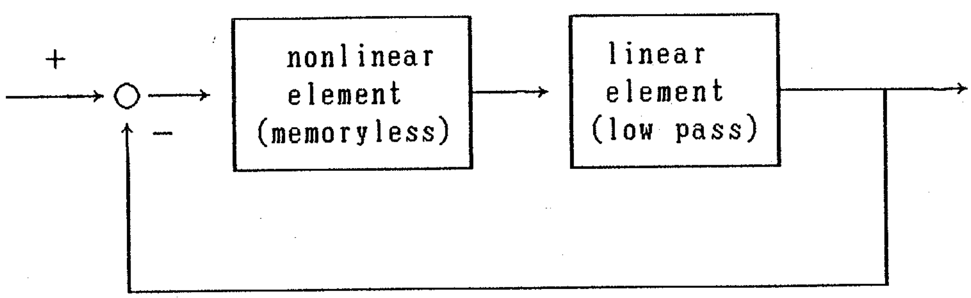


Fig.1-1 block diagram

The describing function method was introduced during the late 1940's[62]. There is, however, no systematic method to analyze and control a system with non-smooth nonlinearity.

On the other hand, smooth nonlinearity is of a type which can be expressed by a smooth function of state. A system with smooth nonlinearity can usually be expressed by a differential equation of the form

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

or

$$\dot{x} = f(x) + g(x)u$$

where ϕ, f and g are vectors whose elements are smooth functions of x (and u). The former system is sometimes called a general nonlinear system and the latter one an affine nonlinear system. Mechanical models are usually expressed in this form. For example, it is well-known that the dynamics of a robot arm can be described by a second order differentional equation as

$$M(\theta)\ddot{\theta} + F(\theta, \dot{\theta}) = u$$

where θ represents the angle and u is the input torque. This differential equation can readily be expressed by

$$\frac{d}{dt}\begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \dot{\theta} \\ M(\theta)^{-1} F(\theta, \dot{\theta}) \end{bmatrix} + \begin{bmatrix} 0 \\ M(\theta)^{-1} \end{bmatrix} u.$$

The analyzing method used for systems with non-smooth nonlinearity can also be applied to this type of systems. Moreover, it is possible to use more systematic methods such as a geometric approach and the variational approach [48] ~[55] to analyze this kind of system.

In the last decade, differential geometry has been successfully applied to this type of nonlinear systems. Several useful books [1][2][3] and [4] were recently published in this area. In this thesis, we will restrict our attention only to smooth nonlinear systems and the geometric approach.

Roughly speaking, the geometric approach can enable us to study the following topics in nonlinear system theory.

1) controllability and observability
2) controlled invariance and disturbance decoupling
3) internal model principle
4) Volterra series and realization theory
5) stabilizing controller and observer design
6) linearization

The first three topics directly follow from the geometric approach in linear systems theory which has been examined in detail by Wonham[18]. We will review topics 1, 2 and 6 which bear directly upon our method.

Controllability and observability are fundamental notions in linear system theory. They are used to ensure the solvability of the system stabilizing problem and the state estimation problem. They are also used for model reduction. Therefore, it is natural to extend these notions to nonlinear system theory. Consider a nonlinear system of the form

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

$$y = \psi(x) \; .$$

Hermann[8], Haynes-Hermes[9], Brockett[10], Lobry[11], Sussman-Jurdjevic[12] and Krener[13] developed, based on the work of Chow[7], a nonlinear analog to linear controllability in terms of Lie algebra, $\mathscr{R}$, generated by the vector fields $\phi(.,u)$ which correspond to a constant control u. It was shown that if $\mathscr{R}$ has a constant dimension, then any point on the integral manifold through x^0 can be reached from x^0 by going forward and/or backward along the trajectories of the system. This kind of controllability has been termed weak controllability by Hermann-Krener[14]. The dual notion of weak controllability is called weak observability and was also introduced by Hermann-Krener[14]. Krener refined these notions in[30].

An approach to the disturbance decoupling problem is to completely isolate the output from the influence of the disturbance. This problem can be solved using the notion of controlled invariance. In linear systems theory, the controlled invariant subspace (or (A,B) invariant subspace) was independently introduced by Basile-Marro[16] and Wonham-Morse[17]. Around 1980, there were several attempts to extend this notion to affine nonlinear systems: Ishijima[19][20], Nomura-Furuta[21], Hirschorn[24] and Isidori et al. [22][23]. Nijimeijer[25] also did related work. All of these works were based on invariant distribution and invariant foliation[30]. Nijimeijer-van der Shaft[26][27] extended this notion of controlled invariance to general nonlinear systems. Controlled invariance for discrete-time systems was studied by Monaco et al.[31] and Grizzle[32][33].

The linearization problem is significant in nonlinear system theory because it is possible to apply control strategies which were perfected in linear system theory to linearized systems. One of the most common linearization methods is first order approximation, as we have mentioned before. Consider the system

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

where $\phi(0,0)=0$. The first order approximation to this system is

$$\dot{x} = Ax + Bu + \mathrm{Order}(x,u)^2$$

$$A = \left.\frac{\partial \phi}{\partial x}\right|_{\substack{x=0\\u=0}} \qquad B = \left.\frac{\partial \phi}{\partial u}\right|_{\substack{x=0\\u=0}} .$$

We will neglect the second order term in designing the controller. Alternatively, around 1980, the exact linearization problem attracted the attention of nonlinear systems reseachers. There are two types of exact linearization: input-output linearization and input-state linearization. Input-output linearization was studied using the decoupling approach by Singh-Rugh[34], Freund[35] and Okutani[36]. Isidori-Ruberti[37] used a structure algorithm to solve the input-output linearization problem. This approach relates closely to zero structure at infinity. Isidori[38] pointed out the possibility of computing zero structure at infinity using the coefficients of the formal power series associated with the external behavior of a nonlinear system. Nijmeijer-Schumacher[39] followed this with a geometric approach to the definition of zero structure at infinity. The problem of matching a nonlinear system to a linear model via dynamic state feedback was studied independently by Kosuge[40] and Isidori[41].

The other linearization method is input-state linearization. With this method, we linearize the system using a state transformation and nonlinear state feedback. The input-state linearization was proposed and solved for a single-input system by Brockett[42]. A complete solution for multi-input systems was found by Jakubczyk-Respondek[43]. Independent work done by Su[44] and Hunt et al.[45] lead to a slightly weaker formulation, together with a constructive algorithm for the solution.

An exact linearization method can, however, only be used on a restricted class of nonlinear systems, thus, some approximate linearization methods have been proposed. Reboulet-Champetier[46] proposed a pseudolinearization method which makes the linear tangent model independent of the operating point using suitable state transformation and nonlinear state feedback. Krener[47] developed the necessary and sufficient condition for a nonlinear system to be approximated to a higher order as a linear system.

1-2. Purpose of the Study

The geometric approach is a powerful tool to analyze nonlinear systems. It does, however, sometimes require strict geometric conditions to ensure solvability. For example, in the case of input-state linearizations on systems of the form

$$\dot{x} = f(x) + g(x)u$$

the set of vector fields $\{g, ad_f g, \cdots, ad_f^{n-2} g\}$ is required to be involutive. This condition is not generally satisfied when $n \geqq 3$. Consider the system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u .$$

According to the linearizability condition, this system can be linearized if and only if

$$\frac{\partial^2 f_2}{\partial x_3^2} = a(x)\frac{\partial f_2}{\partial x_3}$$

$$a(x) = \left(\frac{\partial^2 f_1}{\partial x_3^2}\right) \Big/ \left(\frac{\partial f_1}{\partial x_3}\right) .$$

This means that the function f_2 is specified automatically if the function f_1 and the initial condition $f_2(x_1, x_2, 0)$ for all x_1, x_2 are given. This criteria is generally not satisfied. Thus, we have to find a method which will relax these conditions.

Here, we will relax the geometric conditions by modifying the autonomous term f(x). Feedback was thought to be the only way of manipulating the autonomous term. If we can, however, modify the autonomous term into s(x)f(x) for a scalar function s(x), then we can relax some of the geometric conditions for the solvability of a problem. In the case of the previous example, the condition for linearization is actually relaxed by introducing s(x) (we will examine this in Chapter Ⅳ).

We achieve a modification of the autonomous term by introducing a time scale transformation. The transformation is defined by

$$\frac{dt}{d\tau} = s(x) > 0$$

for some smooth function s(x). The system can be expressed in the time scale τ as

$$\frac{dx}{d\tau} = s(x)\ f(x) + s(x)\ g(x)\ u\ .$$

The fact that s(x) is positive ensures that the time scale τ will not go backwards against the actual time scale t, so the system's stability and the state's curve in state space will not be effected by the time scale transformation. Therefore, it is possible for us to make use of the time scale τ to investigate the system's stability and structure.

We will proceed to study invariant structure and the linearization problem using a transformation of the time scale. We will also use the transformation to design a nonlinear controller which will satisfy certain specifications. This thesis is organized as follows.

In chapter Ⅱ, we will introduce a time scale tranformation and investigate its properties.

In chapter Ⅲ, we will investigate invariant structure using a time scale transformation. We will also introduce the notion of weakly invariant distribution and examine its properties. We will see how the notion of weakly invariant distribution will allow us to obtain Kalman-like decompositions in reachable/unreachable parts and/or observable/unobservable parts in the transformed time scale. We will also introduce weakly controlled invariance and use it to solve the wide-sense disturbance decoupling problem. To solve this, we must seek a feedback law so that the disturbance will not affect the output's curve in output space.

We will investigate the input-state linearization problem in chapter Ⅳ. If we can obtain an exactly linearized model in the transformed time scale, it will be useful to obtain a stabilizing controller because a time scale transformation will preserve the system's stability. We will show how the linearizability condition can be relaxed by introducing a transformation of the time scale.

In chapter Ⅴ, we will take a different perspective on the time scale transformation; applying it to controller design. We will show that, even for linear systems, the time scale transformation can allow us to design a nonlinear controller satisfying a needed specification such as avoiding a exceedingly high amplitude of input. We will illustrate possible applications of our method: we will design a controller for a robot that can achieve good performance even in a neighborhood of a singular point and also a trajectory tracking controller for a vehicle.

II. TIME SCALE TRANSFORMATION

In this chapter, we will introduce a time scale transformation and examine the properties of this transformed time scale.

In analyzing a continuous system, the time scale 't' is taken to be actual time, for example seconds, minutes and hours. These actual times flow uniformly. When, however, we consider the stability of the system or the system's curve which the state of the system traces in state space, it is possible for us to use any time scale 'τ' as long as 'τ' does not reverse upon itself.

Fig.2-1, 2-2 and 2-3 show the vector fields, the state transitions, and the phase plane trajectories (the curves which the states trace in state space) of system A

$$\text{(A)} \qquad \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 - 0.5\, x_2 \end{bmatrix}$$

and system B

$$\text{(B)} \qquad \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2(1.5 + \sin 16\, x_1) \\ (-x_1 - 0.5\, x_2)(1.5 + \sin 16\, x_1) \end{bmatrix}$$

with the initial condition $(x_1 , x_2)^T (0) = (0.5, 0.5)^T$. Although the vector fields and state transitions of these two systems are different in appearance, their phase plane trajectories are identical. This means that these two systems having this initial condition do not have the same state transitions, but do have the same curves in state space. In this case, the difference between the two state transitions is related to the choice of time scale. The state transition of system B can be represented as shown in fig.2-5 in time scale $\tau(t)$ which is defined by fig.2-4. Fig.2-5 is identical to fig.2-1-a (state transition of system A). As we have seen above, our choice of time scale can be helpful in analyzing the state's curve in state space.

Hollerbach[5] has defined a new time scale r,

$$r = r(t)$$

and Ozaki et al.[6] have defined another time scale t'

$$dt' = \kappa(t)\, dt$$

which differ from the actual time t. Each of them have been effectively used to plan the time trajectory of a manipulator with a geometric path constraint. Their works, however, were restricted to planning the time trajectory of a manipulator. The new time scales which they choose were defined to be functions of the actual time t. These time scales are not useful for system analysis because it is necessary for any time scale to be defined for a particular trajectory which in turn must be analyzed individually in the new time scale.

In this chapter, we will formulate a time scale transformation which is dependant on a state. This transformation can be applied to system analysis. We will show that any system can readily be rewritten in the transformed time scale.

The time scale transformation is defined as follows. Consider a system expressed in local coordinates as

(1a) $$\frac{dx}{dt} = f(x) + \sum_{i=1}^{m} g_i(x) u_i$$

(1b) $$y_i = h_i(x) \qquad (i=1,2,\cdots,r)$$

where the n dimensional vector x represents the state; the m dimensional vector u is the input; and the r dimensional vector y is the output. We define the new time scale τ using the continuous function $s(x) > 0$ as

(2a) $$\frac{dt}{d\tau} = s(x)$$

(2b) $$\tau \Big|_{t_0} = \tau_0$$

where 't' is the actual time. The function $s(x)$ is called the <u>time scaling function.</u>

<u>[theorem 1]</u>

The system (1) is expressed in the time scale τ as

(3a) $$\frac{dx}{d\tau} = s(x)\ f(x) + \sum_{i=1}^{m} g_i(x)\mu_i$$

(3b) $$y_i = h_i(x) \qquad (\ i=1,2,\cdots,r\ )$$

where $u_i = \{1 / s(x)\}\mu_i$. □

(proof)

Eq.(3a) follows obviously from the relation

(4) $$\frac{dx}{d\tau} = \frac{dx}{dt}\ \frac{dt}{d\tau}\ .$$

Also, u_i is well-defined because $s(x) > 0$. ■

The drifting term of the system($f(x)$) usually cannot be directly manipulated; this theorem will make it possible for us to directly manipulate it. Consequently, we can now to analyze the structure of the system more precisely; we will use this to extend several structural concepts in subsequent chapters.

For example, if we set the time scale transformation as

$$\frac{dt}{d\tau} = s(x) = \frac{1}{1.5 + \sin 16\ x_1}$$

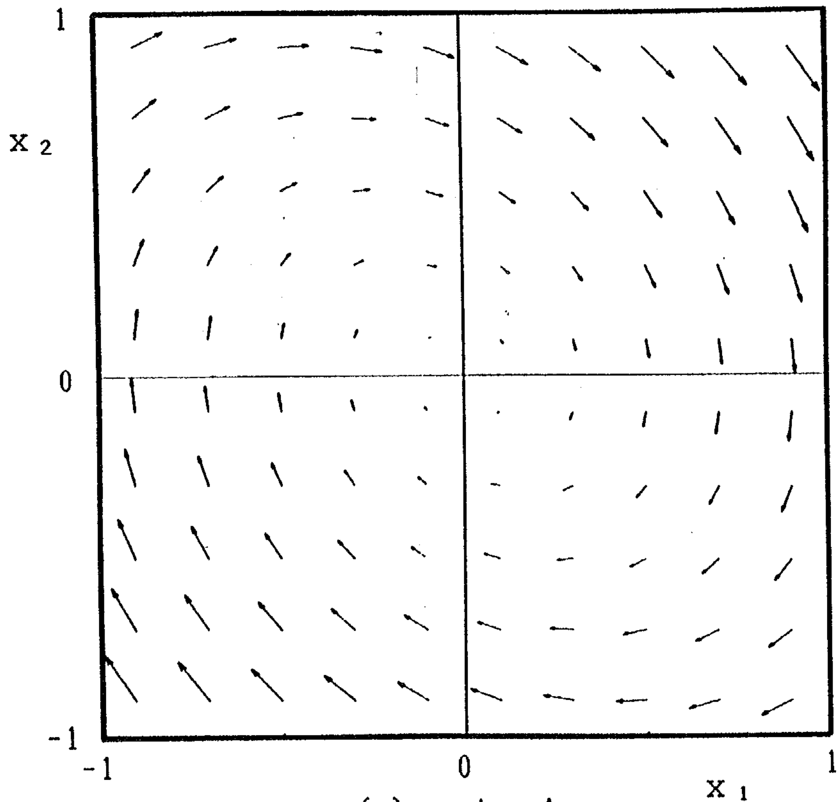
then system B can be expressed in the time scale τ as

(B') $$\frac{d}{d\tau}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1\ -0.5\ x_2 \end{bmatrix}.$$

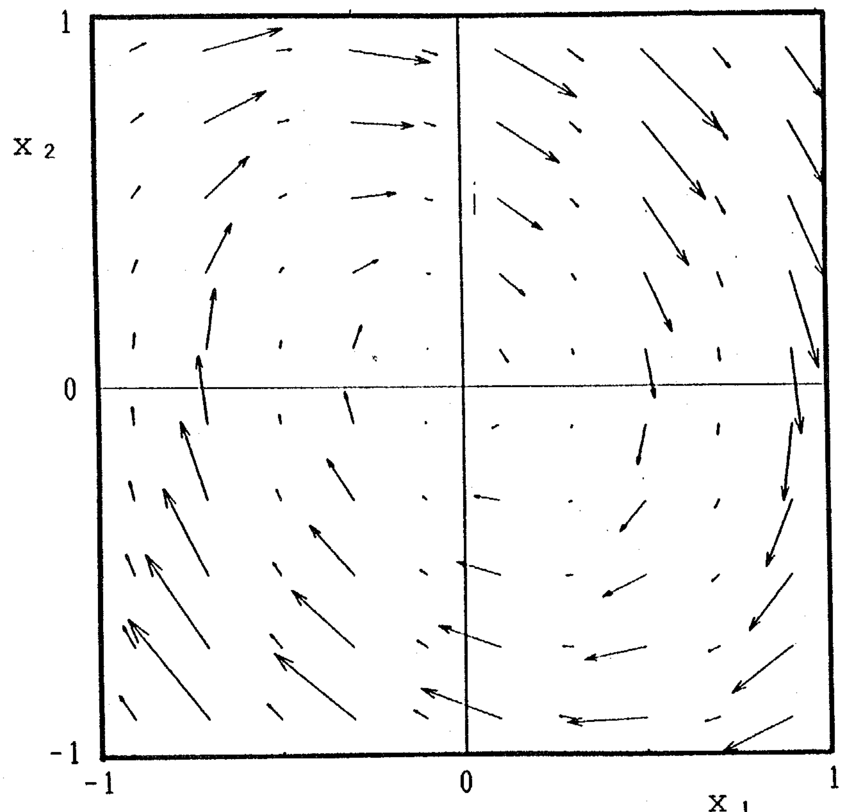
Obviously, system B expressed in the time scale τ is equivalent to system A expresssed in the actual time scale t . Thus, given any initial condition, the state's curves of system A and B in state space will be identical.

We must emphasize the importance of the restriction $s(x)>0$. This implies that the new time scale τ must increase strictly monotonically with respect to the actual time t. In other words, the new time τ must never go backward against the actual time t. This guarantees that the input stabilizing the system expressed in the new time scale τ will also stabilize the original system expressed in the actual time scale t. Using this property in later chapter, we will propose a nonlinear controller which will stabilize the system in addition to satisfying a needed requirement.

In order to preserve a vector field's analyticity or smoothness, the time scaling function $s(x)$ usually needs to be analytic or smooth with respect to x. Otherwise, it is enough for $s(x)$ to have a sufficient number (as specified by the proof) of continuous partial derivatives with respect to x.

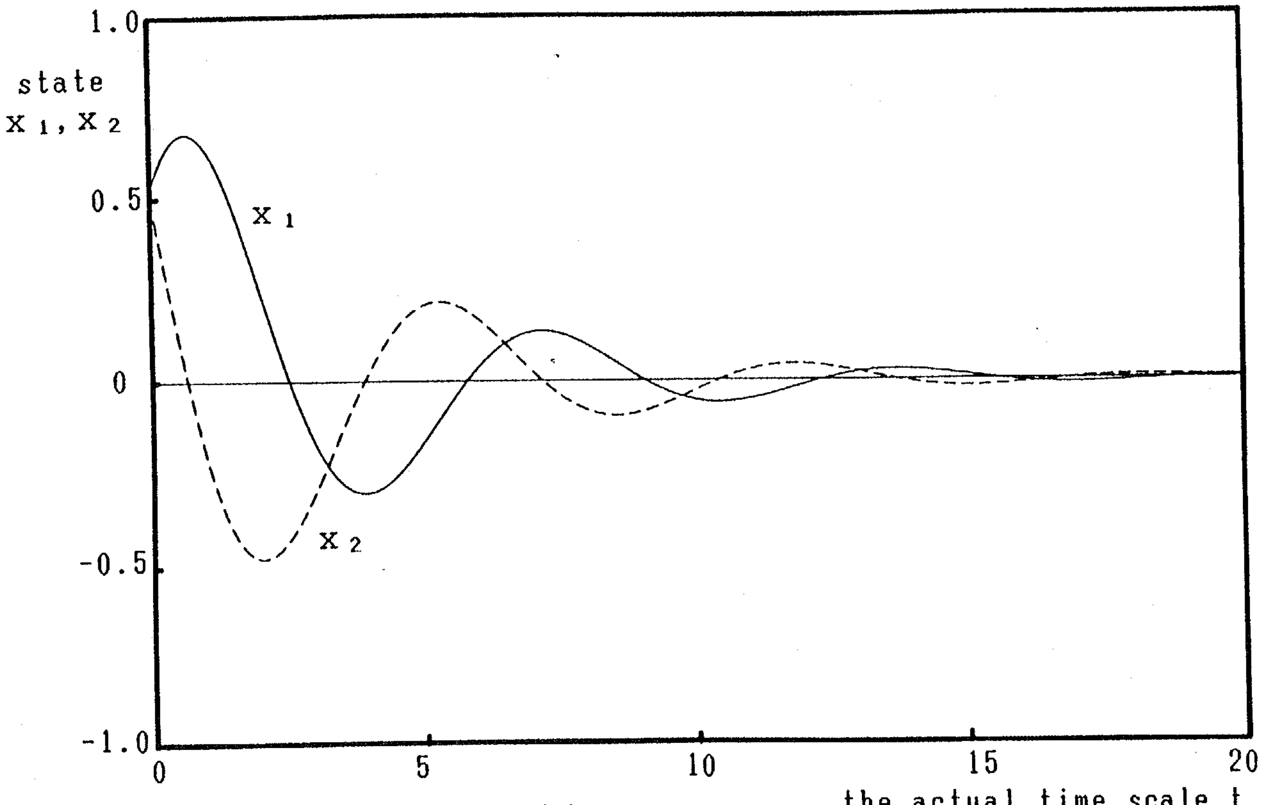


(a) system A

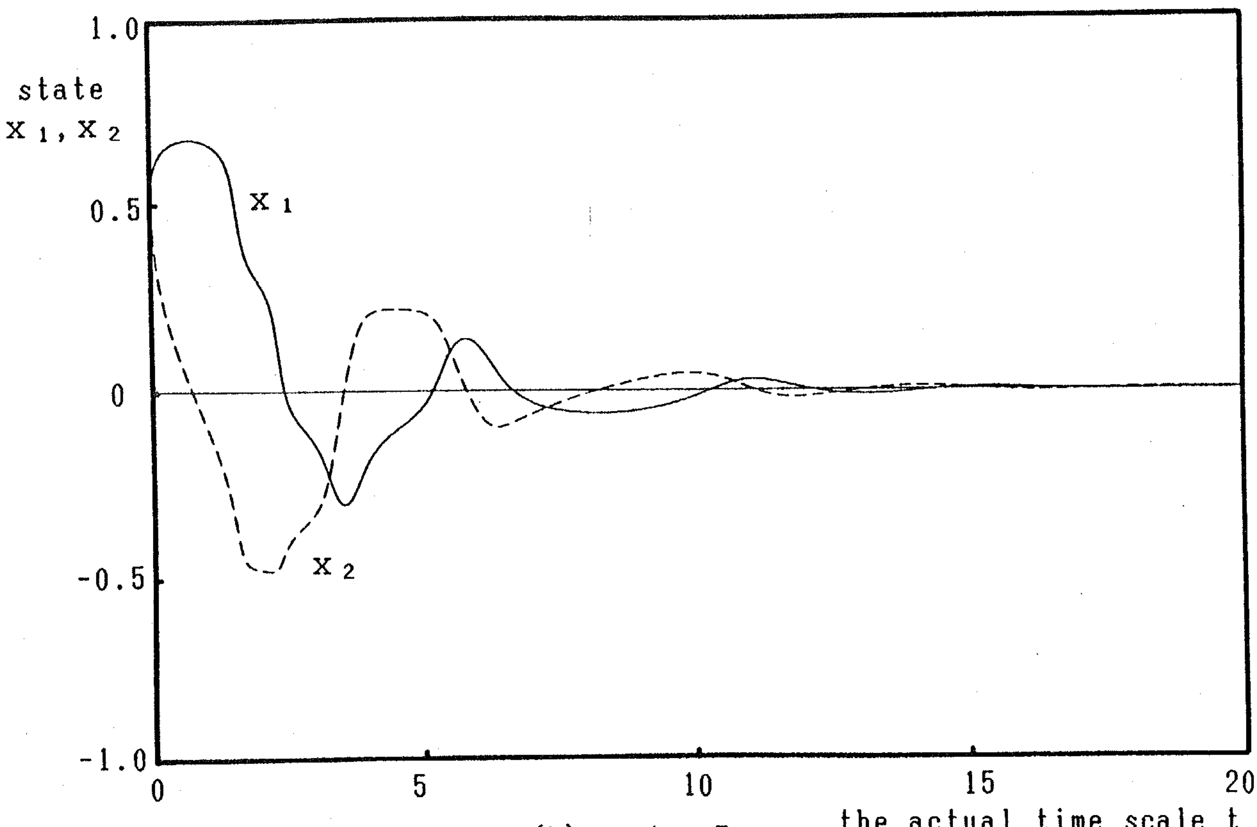


(b) system B

Fig.2-1. Vector field

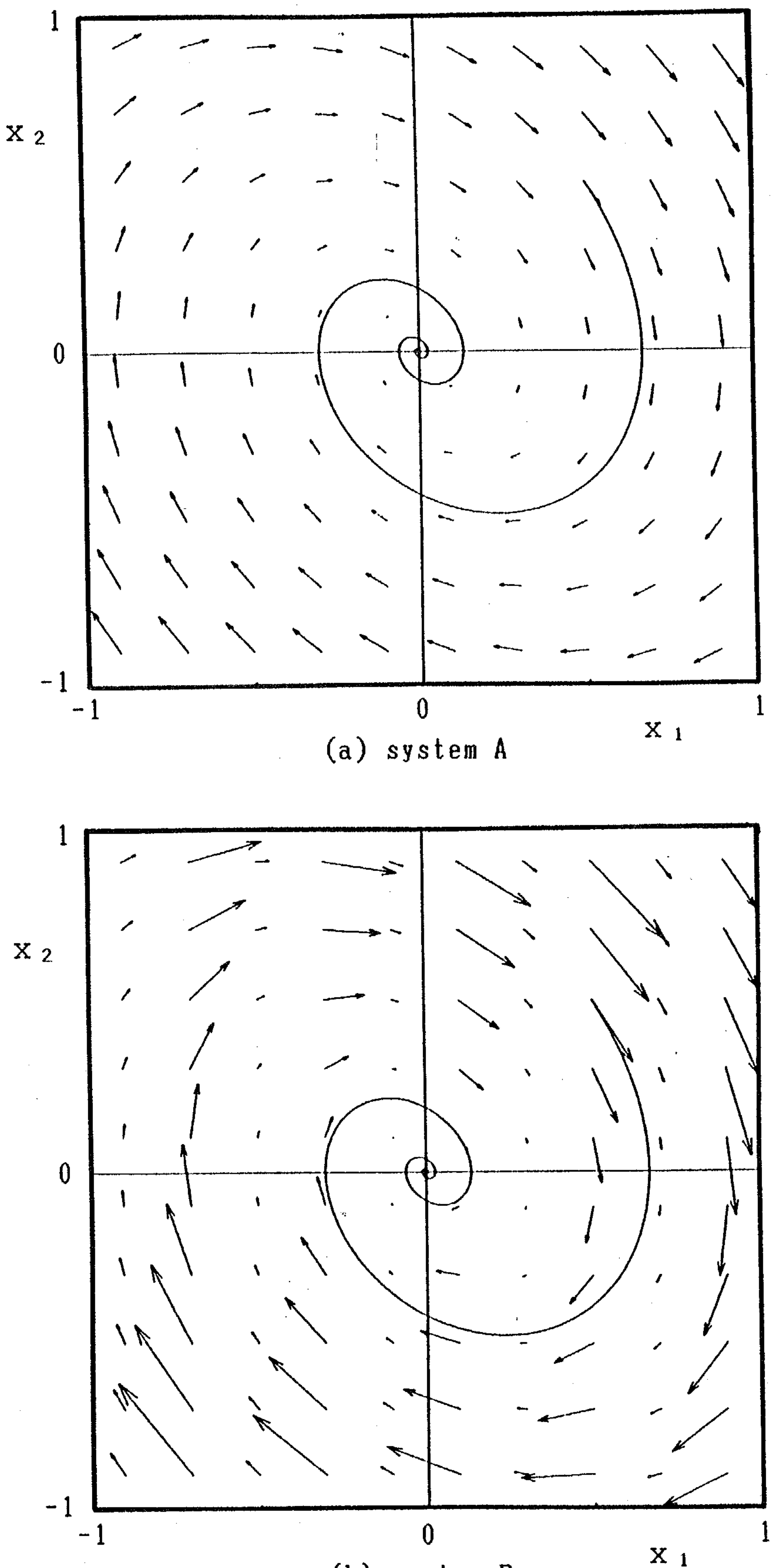


(a) system A



(b) system B

Fig.2-2. State transition



(a) system A

(b) system B

Fig.2-3. Phase plane trajectory

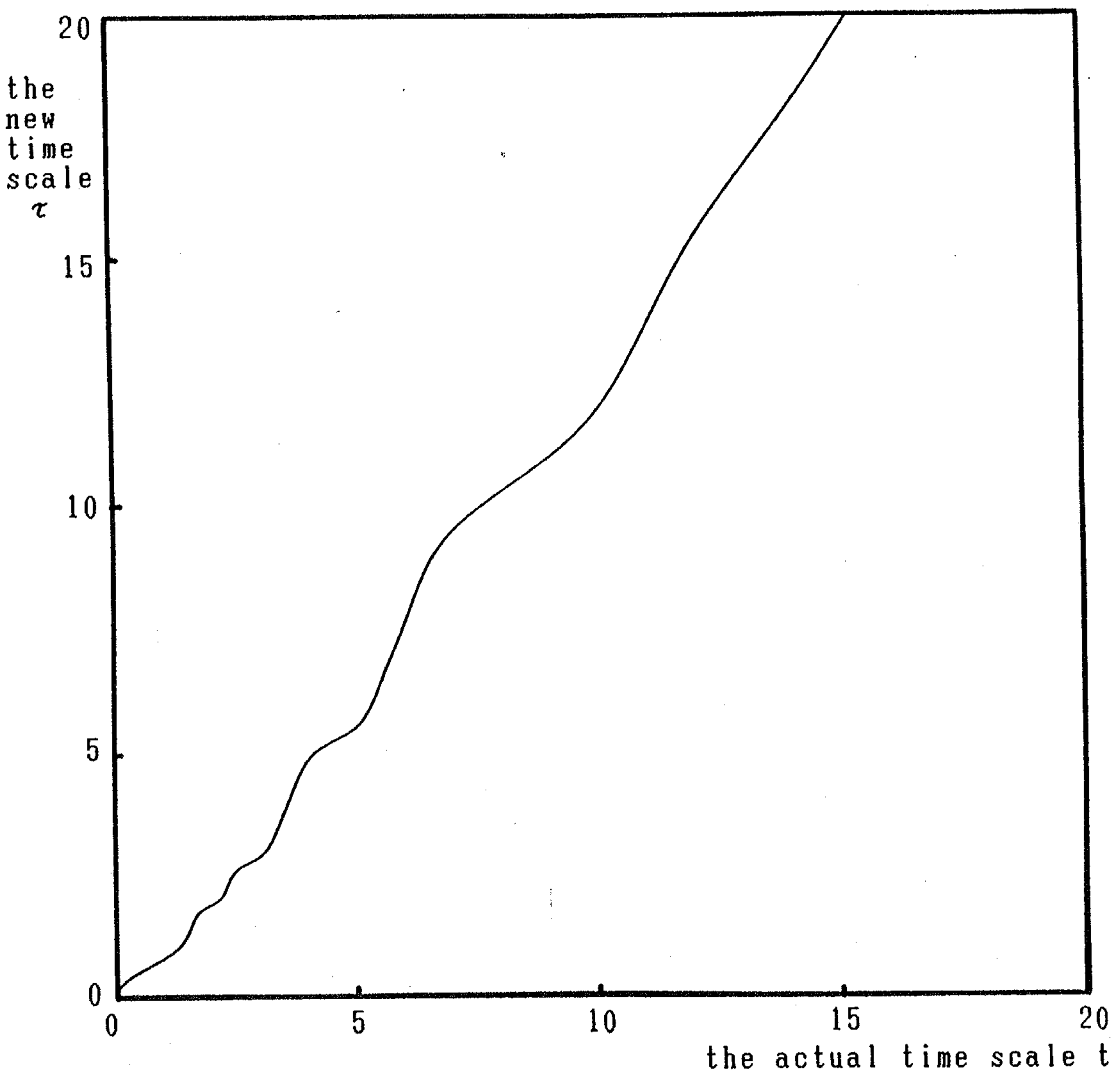


Fig.2-4. The new time scale

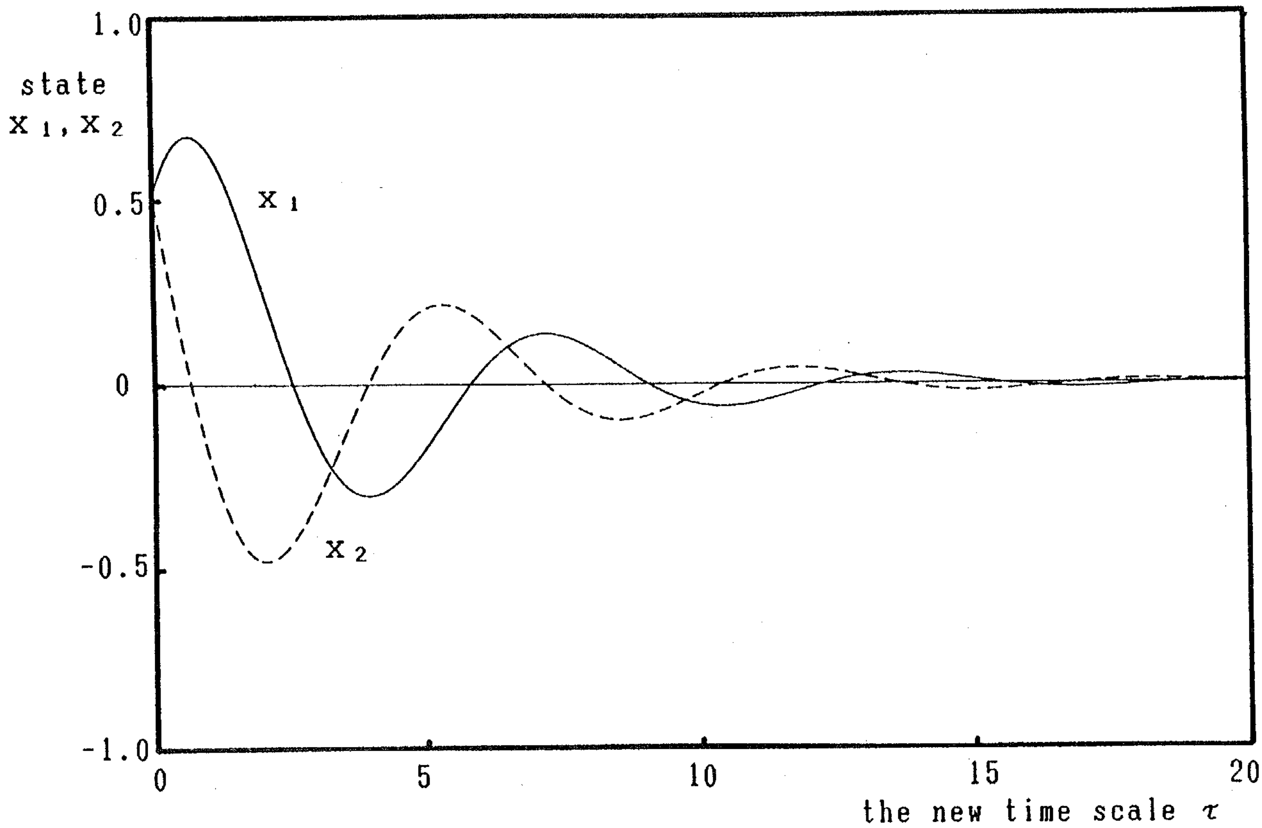


Fig.2-5. State transition in the new time scale (system B)

III. INVARIANT STRUCTURE OF NONLINEAR SYSTEMS

Invariant structure plays an important role in linear system theory. It is closely related to system controllability and observability.

Consider the linear system,

$$\dot{x} = Ax + Bu$$
$$y = Cx .$$

The A-invariant subspace V is a subspace which satisfies

$$AV \subset V .$$

It is well-known fact that, if we can find a d-dimensional A-invariant subspace which contains ImB, then we can decompose the system using the state space transformation as follows,

$$\frac{d}{dt}\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u$$
$$y = \begin{bmatrix} C_1 & C_2 \end{bmatrix}\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$$

where $\bar{x}^T = (\bar{x}_1^T, \bar{x}_2^T)$ is the transformed state defined by $x=T\bar{x}$ for a nonsingular matrix T and $\dim(\bar{x}_1)=d$. With this decomposition, one can easily see that the state $\bar{x}_2$ is not affected by the input u. Thus, we cannot manipulate the state $\bar{x}_2$. The minimal A-invariant subspace which contains ImB is called a controllable subspace. With a controllable subspace, we are able to decompose the system's state into a controllable one and an uncontrollable one.

Similarly, if we can find a d-dimensional A-invariant subspace contained in KerC, then we can decompose the system, with an appropriate state transformation as

$$\frac{d}{dt}\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & C_2 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$$

where $\dim(\bar{x}_1)=d$. This system shows that the state $\bar{x}_1$ does not affect the output of the system. Thus, the two states $(\bar{x}_1^1, \bar{x}_2^1)$ and $(\bar{x}_1^2, \bar{x}_2^2)$ are indistinguishable from the output if $\bar{x}_2^1=\bar{x}_2^2$. The maximal A-invariant subspace contained in Ker C is called an unobservable subspace.

In nonlinear systems theory, the notion of invariant distribution under a vector field is analogous to invariant subspace in linear system theory. The notion of invariant distribution was developed independently by Hirschorn[24] and Isidori et.al.[22]. A more general formulation of invariance was given by Sussman[15]. In sections 3-1 and 3-2, we will extend the notion of invariant distribution and define weakly invariant distribution. The notion of weakly invariant distribution will allow us to investigate the invariant structure of the system in a new time scale. Since a transformation of the time scale preserves the system's stability and the state's curve in state space, we are able to investigate the structure of nonlinear systems in more depth using the notion of weak invariance.

In section 3-3, we will investigate another important invariant structure -- controlled invariance. The notion of controlled invariance has been used to solve the disturbance decoupling problem. In linear systems theory, the disturbance decoupling problem is formulated as follows. Consider a system with disturbance w

$$\dot{x} = Ax + Bu + Dw$$
$$y = Cx .$$

In the disturbance decoupling problem, we seek the feedback u=Fx+Gv such that the output of the system will not be influenced by the disturbance. With the introduction of the feedback, the system will become

$$\dot{x} = (A + BF)x + BGv + Dw$$
$$y = Cx .$$

Thus, if we can find an (A+BF)-invariant subspace which contains ImD and is contained in KerC, we will have, after an appropriate transformation of the state space,

$$\frac{d}{dt}\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} v + \begin{bmatrix} D_1 \\ 0 \end{bmatrix} w$$
$$y = [0 \quad \hat{C}_2] \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} .$$

From this, it is obvious that the output is not influenced by the disturbance; the disturbance decoupling problem has been solved. From this point of view, a subspace V which satisfies

$$(A + BF)V \subset V$$

for an appropriate F will play an important role when we solve the disturbance decoupling problem. Such a subspace is called a controlled invariant subspace or an (A,B) invariant subspace.

The notion of controlled invariant distribution was defined independently by Ishijima[19][20], Nomura et al.[21], Hirschorn[24] and Isidori et.al.[22][23]. We will extend this definition and then define weakly controlled invariant distribution based on weak invariance. We will use weakly controlled invariant distribution to solve the disturbance decoupling problem relating to the output's curve --the wide-sense disturbance decoupling problem-- in section 3-3.

3-1. Weakly Invariant Distribution

As we discussed in the introduction, invariant subspace plays an important role in linear system theory. It allows us to investigate controllability/observability and to decompose the system. This concept of invariant distribution under a vector field plays a similar role in the theory of nonlinear systems. Invariant distribution was introduced by Hirschorn[24] and Isidori[22]; there is a summary in Isidori's recent textbook[2]. We will proceed to define weakly invariant distribution under a vector field which closely relates to invariant structure in a transformed time scale.

We will begin with a discussion of invariant distribution. Consider the n dimensional smooth manifold M. The distribution Δ on M is invariant under a vector field f if the Lie bracket $[f, \theta]$ of f with every vector field $\theta \in \Delta$ is a vector field which belongs to Δ, i.e. if

(1-1) $$[f, \Delta] \subset \Delta .$$

The notion of invariant distribution under a vector field is particularly useful with completely integrable distributions, because it can provide a way to simplify the local representation of a given vector field.

[proposition 1-1]

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector field f. Then, at each point $p \in M$, there exists a coordinate chart (U, ξ) with coordinate functions $\xi_1, \xi_2, \cdots, \xi_n$ in which the vector field f is represented by a vector of the form

$$(1\text{-}2) \qquad f(\xi) = \begin{bmatrix} f_1(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ f_d(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n) \\ f_{d+1}(\xi_{d+1},\cdots,\xi_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ f_n(\xi_{d+1},\cdots,\xi_n) \end{bmatrix} . \quad \square$$

The proof of this proposition can be found in Isidori's textbook[2].

We will proceed to define the notion of a weakly invariant distribution under a vector field as follows.

[definition 1-2]

The distribution Δ on the smooth manifold M is weakly invariant under a vector field f if there exists a smooth positive function s(p) such that the Lie bracket $[sf, \theta]$ of sf with every vector field $\theta \in \Delta$ is a vector field which belongs to Δ, i.e. if

$$(1\text{-}3) \qquad [\, sf, \Delta \,] \subset \Delta . \qquad \square$$

The notion of weakly invariant distribution under a vector field closely relates to the concept of time scale transformations. Consider the autonomous system represented in local coordinates

$$(1\text{-}4) \qquad \frac{dx}{dt} = f(x) .$$

If the distribution Δ is weakly invariant under the vector field f, then there exists a smooth positive function s(x) which satisfies $[sf, \Delta] \subset \Delta$. With this function s(x), we can define the time scale τ as

$$(1\text{-}5) \qquad \frac{dt}{d\tau} = s(x) > 0 .$$

The system is represented in the time scale τ as follows

$$(1\text{-}6) \qquad \frac{dx}{d\tau} = s(x)\, f(x) .$$

This implies that the distribution Δ is invariant under the vector field $dx/d\tau$ which represents the autonomous system $dx/dt=f(x)$ in the time scale τ defined by eq.(1-5).

We will next define a local notion of weak invariance.

[definition 1-3]

The distribution Δ on the smooth manifold M is locally weakly invariant under a vector field f, if for each point $p \in M$ there is an open neighborhood U of p where Δ is weakly invariant under the vector field f on U. □

The following lemma directly follows from proposition 1-1.

[lemma 1-4]

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is locally weakly invariant under the vector field f. Then, for each point $p \in M$, there exists a coordinate chart (U, ξ) with coordinate functions $\xi_1, \xi_2, \cdots, \xi_n$ in which the vector field f is represented by a vector of the form

$$(1\text{-}7) \qquad f(\xi) = a(\xi) \begin{bmatrix} f'_1(\xi_1, \cdots, \xi_d, \xi_{d+1}, \cdots, \xi_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ f'_d(\xi_1, \cdots, \xi_d, \xi_{d+1}, \cdots, \xi_n) \\ f'_{d+1}(\xi_{d+1}, \cdots, \xi_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ f'_n(\xi_{d+1}, \cdots, \xi_n) \end{bmatrix}$$

where $a(\xi)$ is a smooth positive function defined on U. Or, equivalently, there exists a smooth function $s(\xi)$ defined on U such that

$$(1\text{-}8)\qquad s(\xi)\,f(\xi)=\begin{bmatrix} f'_1(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n)\\ \cdot\quad\cdot\quad\cdot\quad\cdot\\ f'_d(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n)\\ f'_{d+1}(\xi_{d+1},\cdots,\xi_n)\\ \cdot\quad\cdot\quad\cdot\quad\cdot\\ f'_n(\xi_{d+1},\cdots,\xi_n)\end{bmatrix}.$$

If Δ is a nonsingular weakly invariant distribution, then the smooth functions a and s can be found on M. □

[remark]

The previous lemma shows that the autonomous system (1-4) can be decomposed with the coordinate transformation $\xi=\Psi(x)$ as

$$(1\text{-}9)\qquad \frac{d\xi}{d\tau}=s(\xi)\,\tilde{f}(\xi)=\begin{bmatrix} \tilde{f}'_1(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n)\\ \cdot\quad\cdot\quad\cdot\quad\cdot\\ \tilde{f}'_d(\xi_1,\cdots,\xi_d,\xi_{d+1},\cdots,\xi_n)\\ \tilde{f}'_{d+1}(\xi_{d+1},\cdots,\xi_n)\\ \cdot\quad\cdot\quad\cdot\quad\cdot\\ \tilde{f}'_n(\xi_{d+1},\cdots,\xi_n)\end{bmatrix}$$

where the time scale is defined as $dt/d\tau=s(\xi)$ and $\tilde{f}(\xi)=\Psi_* f(x)$. □

The concept of locally weak invariance leads to a simple geometric test.

[lemma 1-5]

Suppose that Δ is an nonsingular involutive distribution and that $\dim(\Delta+\mathrm{span}\{f\})-\dim(\Delta)$ is constant on the dense subset $M^*\subset M$. Then Δ is locally weakly invariant under the vector field f if and only if for each smooth vector field $\theta\in\Delta$ there exist a smooth function c and a smooth vector field $\nu\in\Delta$ such that

$$(1\text{-}10)\qquad [f,\theta]=\nu+cf.\quad □$$

The condition that $\dim(\Delta+\mathrm{span}\{f\})-\dim(\Delta)$ is constant on a dense subset of M is automatically satisfied in the case of analytic distribution.

(proof)

Necessity: Suppose Δ is locally weakly invariant under the vector field f. Then for each $p \in M$, there is a neighborhood $U \subset M$ and a smooth positive function s(q) such that for each smooth vector field $\theta \in \Delta$ there exists a smooth vector field $\nu' \in \Delta$, and

(1-11) $[\, s f , \theta \,] = \nu'$.

From the following property of Lie bracket

(1-12) $[\, s f , \theta \,] = s[\, f , \theta \,] - (L_\theta s)\, f$,

we can easily conclude that

(1-13) $$[\, f , \theta \,] = \frac{1}{s}[\, s f , \theta \,] + \frac{1}{s}(L_\theta s)\, f = \frac{1}{s}\nu' + \frac{1}{s}(L_\theta s)\, f .$$

Since s(p) is smooth and positive, $(1/s)\nu'$ is a smooth vector field contained in Δ and $(1/s)L_\theta s$ is a smooth function.

Sufficiency: We will show the existence of the smooth positive function s(q) defined on a neighborhood of $p \in M$ satisfing eq.(1-3). It is apparent that $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is 1 or 0. If $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is 0 on M^* then $\Delta + \mathrm{span}\{f\} = \Delta$ on an open dense subset and $\Delta \subset \Delta + \mathrm{span}\{f\}$. Thus, from lemma A-3 (Appendix), $\Delta + \mathrm{span}\{f\} = \Delta$ on M . This implies that $[f, \Delta] \subset \Delta$ and that Δ is a (weakly) invariant distribution under the vector field f (s(q)=1).

Suppose that $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is 1 on M^*. Since Δ is involutive, then for each $p \in M$ there exist a neighborhood U of p and a coordinate function $x=(x_1,\cdots,x_n)$ such that $\Delta = \mathrm{span}\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_d}\}$ where d is the dimension of Δ. For convenience, we will use the notation ∂_i for $\frac{\partial}{\partial x_i}$. Eq.(1-10) ensures the existence of the smooth functions c^i $1 \leqq i \leqq d$ defined on U and the smooth vector fields ν^i $1 \leqq i \leqq d$ such that

$$(1\text{-}14) \qquad [\, f \,,\, \partial_i \,] = \nu^i + c^i f \,.$$

Since c^i is continuous and $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is 1 on M^*, then c^i is a unique smooth function. Comparing eq.(1-14) with eq.(1-13), it is obvious that the function s(x) must satisfy

$$(1\text{-}15) \qquad c^i = \frac{1}{s}(L_{\partial_i} s) = \frac{1}{s}\frac{\partial s}{\partial x_i} \qquad (1 \leqq i \leqq d)$$

or, equivalently

$$(1\text{-}16) \qquad \frac{\partial s}{\partial x_i} = s\, c^i \qquad (1 \leqq i \leqq d).$$

The Lie bracket of $[f, \partial_i]$ with ∂_k is

$$(1\text{-}17) \qquad [[f, \partial_i], \partial_k] = [\nu^i + c^i f, \partial_k]$$
$$= [\nu^i, \partial_k] + [c^i f, \partial_k]$$
$$= [\nu^i, \partial_k] + c^i [f, \partial_k] - (L_{\partial_k} c^i) f$$
$$= [\nu^i, \partial_k] + c^i \nu^k + (c^i c^k - \frac{\partial c^i}{\partial x_k}) f \,.$$

In this equation, the first and second term is an element of Δ because Δ is involutive. On the other hand, from a property of Lie brackets, we have

$$(1\text{-}18) \qquad [[f, \partial_i], \partial_k] = -[[\partial_i, \partial_k], f] - [[\partial_k, f], \partial_i]$$
$$= [[f, \partial_k], \partial_i] \,.$$

The next equation follows from eq.(1-17) and eq.(1-18)

$$(1\text{-}19) \qquad 0 = [[f, \partial_i], \partial_k] - [[f, \partial_k], \partial_i]$$
$$= [\nu^i, \partial_k] + c^i \nu^k - [\nu^k, \partial_i] - c^k \nu^i$$
$$+ \{(c^i c^k - \frac{\partial c^i}{\partial x_k}) - (c^k c^i - \frac{\partial c^k}{\partial x_i})\} f$$

$$= [\nu^i, \partial_k] + c^i \nu^k - [\nu^k, \partial_i] - c^k \nu^i$$

$$+ \{ \frac{\partial c^k}{\partial x_i} - \frac{\partial c^i}{\partial x_k} + c^i c^k - c^k c^i \} f .$$

Since the first four terms are elements of Δ and $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is 1 on the open dense subset M^* of M, then $\frac{\partial c^k}{\partial x_i} - \frac{\partial c^i}{\partial x_k} + c^i c^k - c^k c^i$ must equal zero on M^*. From the smoothness of the functions c^i and c^k, we have

$$(1\text{-}20) \qquad \frac{\partial c^k}{\partial x_i} - \frac{\partial c^i}{\partial x_k} + c^i c^k - c^k c^i = 0 \qquad (1 \leqq i, k \leqq d)$$

on M. From theorem A-6 (Appendix), eq.(1-20) will ensure the existence of the smooth solution $s(x)$ for differential equations(1-16).

Now, we only need to prove that the solution $s(x)$ is positive. This can be proven as follows. Without any loss of generality, we will consider the neighborhood U of p and assume $x(p)=0$. Since eq.(1-16) has solution $s(x)$, we have

$$(1\text{-}21) \qquad \int_0^{x_k} c^k(0,\cdots,0,x_k',x_{k+1},\cdots,x_n)\, dx_k'$$

$$= \int_{s(0,\cdots,0,0,x_{k+1},\cdots,x_n)}^{s(0,\cdots,0,x_k,x_{k+1},\cdots,x_n)} \frac{1}{s}\, ds$$

$$= \ln s(0,\cdots,0,x_k,x_{k+1},\cdots,x_n) - \ln s(0,\cdots,0,0,x_{k+1},\cdots,x_n) .$$

This implies

$$(1\text{-}22) \qquad s(0,\cdots,0,x_k,x_{k+1},\cdots,x_n) = s(0,\cdots,0,0,x_{k+1},\cdots,x_n) \exp\{ \int_0^{x_k} c^k\, dx_k' \} .$$

Thus

$$(1\text{-}23) \qquad s(x) = s(0,\cdots,0,x_{d+1},\cdots,x_n) \prod_{k=1}^{d} \exp\{ \int_0^{x_k} c^k\, dx_k' \} .$$

If we choose $0 < s(0, \cdots, 0, x_{d+1}, \cdots, x_n) < \infty$ carefully (for example $s(0, \cdots, 0, x_{d+1}, \cdots, x_n) = 1$), then the resultant $s(x)$ will be smooth and positive because $\exp(a)$ is positive for $a \in (-\infty, \infty)$. ■

We will define the concept of weakly invariant codistribution similarly.

[definition 1-6]

The codistribution Ω on the smooth manifold M is weakly invariant under a vector field f if there exists a smooth positive function s(p) such that the Lie derivative $L_{sf}\sigma$ of the covector field $\sigma \in \Omega$ with sf is another covector field which belongs to Ω, i.e. if

$$(1\text{-}24) \qquad L_{sf}\ \Omega \subset \Omega . \qquad \Box$$

Next we will similarly define the notion of locally weakly invariant codistribution under a vector field.

[definition 1-7]

The codistribution Ω on the smooth manifold M is locally weakly invariant under a vector field f if for each point $p \in M$ there is an open neighborhood U of p with the property that Ω is weakly invariant under the vector field f on U. □

We can easily see that this is the dual version of locally weakly invariant distribution.

[lemma 1-8]

If the smooth distribution Δ is locally weakly invariant under the vector field f, then the codistribution $\Omega = \Delta^{\perp}$ is locally weakly invariant under f. If the smooth codistribution Ω is locally weakly invariant under the vector field f, then the distribution $\Delta = \Omega^{\perp}$ is locally weakly invariant under f. □

The proof of this lemma is analogous to that in the case of invariant distribution/codistribution, which is found in [2].

The following lemma can be useful when we are seeking the maximal locally weakly invariant distribution which annihilates a codistribution or, equivalently, the minimal locally weakly invariant codistribution which contains a codistribution. This problem closely relates to system decomposition in a new time scale which we will discuss latter.

[lemma 1-9]

Suppose that Ω is a smooth codistribution which satisfies

(1-25) $$L_f\,(\,\mathrm{span}\,\{\,f\,\}^{\perp} \cap \Omega\,) \subset \Omega\,.$$

Then, on any open subset $U \subset M$ where $\mathrm{span}\{f\}^{\perp} \cap \Omega$ is smooth, the following equation will be satisfied

(1-26) $$[\,f\,,\,\Omega^{\perp}\,] \subset \Omega^{\perp} + \mathrm{span}\,\{\,f\,\}\,. \quad \square$$

Eq.(1-26) is a necessary condition for the distribution $\Omega^{\perp}$ to be locally weakly invariant under f.

(proof)

Let σ be a smooth covector field in $\mathrm{span}\{f\}^{\perp} \cap \Omega$, and θ be a vector field in $\Omega^{\perp}$. We will make use of the identity

(1-27) $$<L_f\,\sigma\,,\,\theta> = L_f <\sigma\,,\,\theta> - <\sigma\,,\,[\,f\,,\,\theta\,]>\,.$$

Since $L_f\,\sigma \in \Omega$ and $\theta \in \Omega^{\perp}$ imply that

(1-28) $$<L_f\,\sigma\,,\,\theta> = 0$$

(1-29) $$<\sigma\,,\,\theta> = 0\,,$$

we have

(1-30) $$<\sigma\,,\,[\,f\,,\,\theta\,]> = 0\,.$$

Since $\mathrm{span}\{f\}^{\perp} \cap \Omega$ is smooth on U by assumption, $[f,\ \theta]$ annihilates every covector field in $\mathrm{span}\{f\}^{\perp} \cap \Omega$ on U, i.e.

(1-31) $$[\,f\,,\,\theta\,] \in (\mathrm{sp}\,\{\,f\,\}^{\perp} \cap \Omega)^{\perp} = \Omega^{\perp} + \mathrm{sp}\,\{\,f\,\}\,. \quad \blacksquare$$

3-2. Locally Weak Decomposition of Nonlinear Systems

The notion of a locally weakly invariant distribution under a vector field will enable us to decompose a nonlinear system. Throughout this section we will be dealing with nonlinear systems of the form

$$\text{(2-1a)} \qquad \dot{p} = f(p) + \sum_{i=1}^{m} g_i(p)\, u_i$$

$$\text{(2-1b)} \qquad y_i = h_i(p) \qquad\qquad (i = 1,2,\cdots,r)\ .$$

The state p of this system is a point on the a smooth n dimensional manifold M. And $\dot{p}$ stands for the tangent vector at the point p to the smooth curve which characterizes the solution of the state equation. The m components $u=(u_1,\cdots,u_m)$ of the input and the r components $y=(y_1,\cdots,y_r)$ of the output are real-valued functions of time. The vector fields f, g_i are smooth vector fields defined on M which we assume to be complete. The output functions h_i are real-valued smooth function defined on M. In a local coordinate chart (U, x), the state equation can be represented as

$$\text{(2-2a)} \qquad \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)\, u_i$$

$$\text{(2-2b)} \qquad y_i = h_i(x) \qquad\qquad (i = 1,2,\cdots,r)\ .$$

We define the smooth distribution G as

$$\text{(2-3)} \qquad G = \operatorname{span}\{g_1,\cdots,g_m\}\ .$$

3-2-1. Locally Weak Decomposition (controllability)

In this section we will investigate in the transformed time scale a local decomposition which removes the uncontrollable state. This is closely related to locally weakly invariant distribution. We will begin with the following proposition.

[proposition 2-1]

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector fields f , $g_1, \cdots, g_m$. Moreover, suppose that the distribution G is contained in Δ. Then, for each point $p \in M$ it is possible to find an open neighborhood U of p and a local coordinate ξ defined on U such that the state equation of the system(2-1) can be represented by equations of the form

$$\text{(2-4a)} \qquad \dot{\xi}_1 = f_1(\xi_1, \xi_2) + \sum_{i=1}^{m} g_{i1}(\xi_1, \xi_2) u_i$$

$$\text{(2-4b)} \qquad \dot{\xi}_2 = f_2(\xi_2)$$

where (ξ_1, ξ_2) is a partition of ξ with $\dim(\xi_1) = d$. □

A similar decomposition is possible even if the distribution is locally weakly invariant.

[theorem 2-2]

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is locally weakly invariant under the vector fields f , $g_1, \cdots, g_m$. Moreover, suppose that the distribution G is contained in Δ. Then, for each point $p \in M$ it is possible to find an open neighborhood U of p. The time scale τ can be defined by a smooth positive function s(q) as $dt/d\tau = s(q)$ for $q \in U$. The local coordinates ξ can be defined on U such that the state equation of the system(2-1) can be represented by equations of the form

$$\text{(2-5a)} \qquad \frac{d\xi_1}{d\tau} = \tilde{f}_1(\xi_1, \xi_2) + \sum_{i=1}^{m} \tilde{g}_{i1}(\xi_1, \xi_2) u_i$$

$$\text{(2-5b)} \qquad \frac{d\xi_2}{d\tau} = \tilde{f}_2(\xi_2)$$

where (ξ_1, ξ_2) is a partition of ξ; $\dim(\xi_1) = d$; and $\tilde{f} = s f$, $\tilde{g}_i = s g_i$. □

(proof)

Since the distribution Δ contains G, and Δ is locally weakly invariant under g_i, there exists an open neighborhood U of p for each $p \in M$ and a smooth positive function $s_i(q)$ such that for any vector field θ contained in Δ, the following equation will hold

$$(2\text{-}6) \qquad \Delta \ni [s_i g_i, \theta] = s_i [g_i, \theta] - (L_\theta s_i) g_i .$$

This implies that $[g_i, \theta] \in \Delta$ because the last term is

$$(2\text{-}7) \qquad (L_\theta s_i) g_i \in G \subset \Delta .$$

Thus we have, for any smooth function λ defined on M,

$$(2\text{-}8) \qquad [\lambda g_i, \theta] = \lambda [g_i, \theta] - (L_\theta \lambda) g_i \in \Delta .$$

This implies that Δ is invariant under λg_i for any λ. Since Δ is locally weakly invariant under the vector field f, there exist, for each $p \in M$, an open neighborhood U of p and a smooth positive function $s(q)$ such that Δ is invariant under the vector field sf on U. The previous discussion shows that the distribution Δ is also invariant under the vector field $s g_i$ on U. Thus, the system in the transformed time scale τ defined by $dt/d\tau = s$ can be written as

$$(2\text{-}9) \qquad \frac{\partial p}{\partial \tau} = s\, f(p) + \sum_{i=1}^{m} s\, g_i(p)\, u_i .$$

Since the distribution Δ is locally invariant under the vector fields $s f$ and $s g_i$, eq.(2-9) can be represented by eq.(2-5) in an appropriate coordinate chart (U, ξ).■

In eq.(2-5), the dynamics of the state ξ_2 in the time scale τ is not affected by the system input u. This means that the curve which the state ξ_2 traces in state space will not be influenced by the system input. Thus, we can not manipulate the the curve of the state ξ_2 even though are able to manipulate the trajectory of the state ξ_2 as a function of the actual time scale.

The obvious requirement to follow the previous decomposition would be to look for the "minimal" distribution Δ which is involutive, contains the distribution G and is locally weakly invariant under the vector fields f, $g_1, \cdots, g_m$. If H is a family of distributions, we can then define the minimal element (when it exists) which is contained in every other element of H as the member of H. It is known that a family of distributions which is invariant under f, $g_1, \cdots, g_m$ and contains G has a mimimal element. We will show, however, that a family of distributions which is locally weakly invariant under f, $g_1, \cdots, g_m$ and contains G may fail to have a minimal element. In the case where a distribution is invariant under the vector fields f, $g_1, \cdots, g_m$, the minimal element can be easiy found by the following algorithm. This algorithm will allow us to find the mimimal distribution which is invariant under the vector fields $\theta_1, \cdots, \theta_q$ and contains the distribution Δ. We denote the minimal element with the symbol

(2-10) $\quad < \theta_1, \cdots, \theta_q \mid \Delta >.$

[algorithm 2-3]

(2-11a) $\quad \Delta_0 = \Delta$

(2-11b) $\quad \Delta_{k+1} = \Delta_k + \sum_{i=1}^{q} [\theta_i, \Delta_k]$. □

[proposition 2-4]

If there exists an integer k^* such that $\Delta_{k^*} = \Delta_{k^*+1}$, then

(2-12) $\quad \Delta_{k^*} = < \theta_1, \cdots, \theta_q \mid \Delta >.$

Moreover, if $\Delta \subset \mathrm{span}\{\theta_1, \cdots, \theta_q\}$ and Δ_{k^*} is nonsingular, then Δ_{k^*} is involutive. □

Thus, we can readily find the distribution $\langle f, g_1, \cdots, g_m \mid G \rangle$. The following lemma will be helpful in seeking the minimal dimension of the distribution which contains G and is locally weakly invariant under $f, g_1, \cdots, g_m$.

[lemma 2-5]

Suppose the distribution Δ is locally weakly invariant under the vector fields $f, g_1, \cdots, g_m$ and contains G. If Δ is nonsingular, then the distribution $\Delta + \mathrm{span}\{f\}$ is invariant under $f, g_1, \cdots, g_m$ in a neighborhood of a regular point of $\Delta + \mathrm{span}\{f\}$. □

(proof)

Since Δ is invariant under $f, g_1, \cdots, g_m$ and contains G, we have

(2-13a) $\quad [f, \Delta] \subset \Delta + \mathrm{span}\{f\}$

(2-13b) $\quad [g_i, \Delta] \subset \Delta + \mathrm{span}\{g_i\} = \Delta.$

For each regular point $p \in M$ of $\Delta + \mathrm{span}\{f\}$, there exists an open neighborhood U of p such that $\Delta + \mathrm{span}\{f\}$ is nonsingular on U. Thus, for any smooth vector field θ contained in $\Delta + \mathrm{span}\{f\}$, we can find the smooth vector fields $\theta_1 \in \Delta$ and $\theta_2 = cf$ where c is a smooth scalar function (if $\Delta = \Delta + \mathrm{span}\{f\}$, we may set c=0) such that $\theta = \theta_1 + \theta_2$. The Lie brackets $[f, \theta]$ and $[g_i, \theta]$ can be expressed as

(2-14a)
$$\begin{aligned}[f, \theta] &= [f, \theta_1 + \theta_2] \\ &= [f, \theta_1] + [f, \theta_2] \\ &= [f, \theta_1] + c[f, f] + (L_f c) f \\ &\in \Delta + \mathrm{span}\{f\}\end{aligned}$$

$$(2\text{-}14b)\quad [g_i, \theta] = [g_i, \theta_1 + \theta_2]$$
$$= [g_i, \theta_1] + [g_i, \theta_2]$$
$$= [g_i, \theta_1] + c[g_i, f] + (L_{g_i} c) f$$
$$\in \Delta + \text{span}\{f\}.$$

These equations imply that Δ+span{f} is invariant under the vector fields $f, g_1, \cdots, g_m$ on U. ■

Thus, if we have a nonsingular distribution Δ which is locally weakly invariant under $f, g_1, \cdots, g_m$ and contains G, then the distribution Δ+span{f} will be invariant under the vector fields $f, g_1, \cdots, g_m$ and contain G+span{f} in a neighborhood of the regular points of Δ+span{f}. The minimal element of the distribution $\langle f, g_1, \cdots, g_m \mid G+\text{span}\{f\}\rangle$ can be caluculated with algorithm 2-3. From this we can conclude that the minimal dimension (if it exists) of the locally weakly invariant distribution under $f, g_1, \cdots, g_m$ containing G is not less than $\dim(\langle f, g_1, \cdots, g_m \mid G+\text{span}\{f\}\rangle) - 1$ almost everywhere on M.

Even though $\langle f, g_1, \cdots, g_m \mid G\rangle$ fails to be nonsingular, we may sometimes find the minimal dimensional nonsingular distribution Δ which is locally weakly invariant under the vector fields f, $g_1, \cdots, g_m$ and contains G.

[example]

Consider the system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 e^{x_1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$
$$= f(x) + g(x)\,u.$$

The distribution $\langle f, g \mid \text{span}\{g\}\rangle$ can be obtained by algorithm 2-3

$$\Delta_0 = \text{span}\{g\}.$$

Since

$$[f, g] = -\begin{bmatrix} 1 \\ x_2 e^{x_1} \end{bmatrix}$$

we have

$$\Delta_1 = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ x_2 e^{x_1} \end{bmatrix} \right\}$$

$$= \left\{ \begin{bmatrix} a(x) \\ b(x) \end{bmatrix} \mid b(x)=0 \text{ where } x_2=0 \right\}.$$

This distribution Δ_1 can readily be found to equal $\langle f, g \mid \text{span}\{g\}\rangle$, however, it fails to be nonsingular. On the other hand, the distribution span{g} itself is locally weakly invariant under the vector fields f,g because, for $s(x)=e^{-x_1}$,

$$[sf, g] = -\begin{bmatrix} e^{-x_1}(1-x_1) \\ 0 \end{bmatrix}.$$

If we set $dt/d\tau = s(x)$, the system in this time scale can be described by

$$\frac{d}{d\tau}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e^{-x_1} \\ 0 \end{bmatrix} u .$$

This system has been decomposed in the transformed time scale.□

The following example illustrates that the family of locally weakly invariant distributions under $f, g_1, \cdots, g_m$ containing G may fail to have a minimal element.

[example]

Consider the system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ e^{x_1} \\ e^{2x_1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$= f(x) + g(x)\ u.$$

Since

$$[f, g] = -\begin{bmatrix} 1 \\ e^{x_1} \\ 2e^{2x_1} \end{bmatrix}$$

$$[g, -[f, g]] = \begin{bmatrix} 0 \\ e^{x_1} \\ 4e^{2x_1} \end{bmatrix},$$

$\langle f,g \mid \mathrm{span}\{f,g\}\rangle$ corresponds to TM. Thus, the minimal dimension, if it exists, of the locally weakly invariant distribution under f,g which contains span{g} is 2. We can find two different distributions of dimension two which are locally weakly invariant under f,g and which contain span{g}. If we set $s(x) = e^{-x_1}$, then

$$s\ f = \begin{bmatrix} e^{-x_1} \\ 1 \\ e^{x_1} \end{bmatrix}$$

and the distribution span{ ∂_1 , ∂_3 } is invariant under sf,g. Thus, span { ∂_1 , ∂_3 } is a minimal dimensional distribution which is locally weakly invariant under f,g and contains span{g}. The system with the time transformation $dt/d\tau = s$ can be represented as

$$\frac{d}{d\tau}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 e^{-x_1} \\ 1 \\ e^{x_1} \end{bmatrix} + \begin{bmatrix} e^{-x_1} \\ 0 \\ 0 \end{bmatrix} u.$$

Apparently, the dynamics of the x_2-axis in the time scale τ is independent of those of the x_1 and x_3-axes. If, however, we set $s'(x)=e^{-2x_1}$, a similar calculation will show that the distribution span { ∂_1 , ∂_2 } is invariant under s'f,g. This means that the distribution span{ ∂_1 , ∂_2 } is also a minimal dimensional distribution which is locally weakly invariant under f,g and contains span{g}. The system dynamics in the time scale τ' defined by $dt/d\tau'=s'$, is described by

$$\frac{d}{d\tau'}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 e^{-2x_1} \\ e^{-x_1} \\ 1 \end{bmatrix} + \begin{bmatrix} e^{-2x_1} \\ 0 \\ 0 \end{bmatrix} u .$$

The dynamics of the x_3-axis in the time scale τ' is independent of those of the x_1 and x_2-axes. Thus, both span{ ∂_1 , ∂_2 } and span{ ∂_1 , ∂_3 } are minimal dimensional distributions, locally weakly invariant under f,g and containing span{g}. This implies that a system may fail to have a minimal distribution which is locally weakly invariant under f,g and contains span{g}. □

3-2-2. Locally Weakly Decomposition (Observability)

In this section, we will investigate another important decomposition of the control system which can separate the unobservable states from the observable states. The following proposition is useful for terating invariant distribution.

[proposition 2-6]

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is invariant under the vector fields $f, g_1, \cdots, g_m$. Moreover, assume that Δ is contained in the distribution $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$. Then, for each $p \in M$ it is possible to find an open subset

U of p and local coordinates ξ defined on U such that the control system(2-1) can be represented in the local coordinates ξ by the following form.

(2-15a) $$\dot{\xi}_1 = f_1(\xi_1, \xi_2) + \sum_{i=1}^{m} g_{i1}(\xi_1, \xi_2) u_i$$

(2-15b) $$\dot{\xi}_2 = f_2(\xi_2) + \sum_{i=1}^{m} g_{i2}(\xi_2) u_i$$

(2-15c) $$y_i = h_i(\xi_2) \qquad (i=1,\cdots,r)$$

where (ξ_1, ξ_2) is a partition of ξ and $\dim(\xi_1)=d$. □

We can readily find the following lemma for a locally weakly invariant distribution.

<u>[theorem 2-7]</u>

Let Δ be a nonsingular involutive distribution of dimension d and assume that Δ is locally weakly invariant under the vector fields $f, g_1, \cdots, g_m$. Moreover, assume that Δ is contained in the distribution $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$. Then, for each $p \in M$ it is possible to find an open subset U of p; local coordinates ξ defined on U; and a time scale τ which is defined by the smooth positive function s(q) as $dt/d\tau = s(q)$ such that in the local coordinates ξ, the control system(2-1) in the time scale τ can be represented by

(2-16a) $$\frac{d\xi_1}{d\tau} = \tilde{f}_1(\xi_1, \xi_2) + \sum_{i=1}^{m} \tilde{g}_{i1}(\xi_1, \xi_2) v_i$$

(2-16b) $$\frac{d\xi_2}{d\tau} = \tilde{f}_2(\xi_2) + \sum_{i=1}^{m} \tilde{g}_{i2}(\xi_2) v_i$$

(2-16c) $$y_i = h_i(\xi_2) \qquad (i=1,\cdots,r)$$

where (ξ_1, ξ_2) is a partition of ξ and $\dim(\xi_1)=d$. And $\tilde{f} = s f$; $g_i = s_i \tilde{g}_i$; $u_i = (s_i / s) v_i$ for a smooth positive function s_i defined on U. □

(proof)

Since Δ is locally weakly invariant under $f, g_1, \cdots, g_m$, we can find for each $p \in M$ an open neighborhood and the smooth positive functions $s, s_1, \cdots, s_m$ such that Δ is invariant under sf, $s_1 g_1$, $\cdots$, $s_m g_m$ on U. Let τ be a time scale defined by $dt/d\tau = (s)$. The system in this time scale τ can be represented by

$$\text{(2-17a)} \qquad \frac{dp}{d\tau} = s\,f(p) + \sum_{i=1}^{m} s\,g_i(p)\,u_i$$

$$\text{(2-17b)} \qquad y_i = h_i(p) \qquad (i=1,\cdots,r).$$

If we set $u_i = (s_i / s)\,v_i$ and $\tilde{g}_i = s_i\,g_i$, then we will have

$$\text{(2-18a)} \qquad \frac{dp}{d\tau} = s\,f(p) + \sum_{i=1}^{m} s\,g_i(p)\,(s_i/s)\,v_i$$

$$= s\,f(p) + \sum_{i=1}^{m} s_i\,g_i(p)\,v_i$$

$$= \tilde{f}(p) + \sum_{i=1}^{m} \tilde{g}_i(p)\,v_i$$

$$\text{(2-18b)} \qquad y_i = h_i(p) \qquad (i=1,\cdots,r).$$

This theorem directly follows from this system. ■

This theorem implies that the two states $(\bar{\xi}_1^1, \bar{\xi}_2^1)$ and $(\bar{\xi}_1^2, \bar{\xi}_2^2)$ are indistinguishable in the time scale τ if $\bar{\xi}_2^1$ and $\bar{\xi}_2^2$ are the same. In other words, the output's curve in output space (not the output trajectory as a function of the actual time scale) is not affected by ξ_1.

From decomposition (2-16), we would like to seek the "maximal" distribution which is locally weakly invariant under $f, g_1, \cdots, g_m$ and contained in $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$. Since the annihilator of a nonsingular and locally weakly invariant distribution is also locally weakly invariant, this problem is the same as to seek the miminal codistribution which is locally weakly invariant under $f, g_1, \cdots, g_m$ and contains span$\{dh_1, \cdots, dh_r\}$. Such a minimal codistribution can be found with the following algorithm.

[algorithm 2-8]

(2-19a) $\Omega_0 = \text{span}\{ dh_1, \cdots, dh_r \}$

(2-19b) $\Omega_{k+1} = \Omega_k + L_f(\text{span}\{f\}^{\perp} \cap \Omega_k)$

$$+ \sum_{i=1}^{m} L_{g_i}(\text{span}\{g_i\}^{\perp} \cap \Omega_k). \qquad \square$$

From lemma 1-9 , if there exists an integer k^* such that $\Omega_{k^*+1} = \Omega_{k^*}$, then $\Omega_{k^*}^{\perp}$ satisfies the following equations on any open subset $U \subset M$ where $\text{span}\{f\}^{\perp} \cap \Omega_{k^*}$ and $\text{span}\{g_i\}^{\perp} \cap \Omega_{k^*}$ are smooth.

(2-20a) $[f, \Omega_{k^*}^{\perp}] \subset \Omega_{k^*}^{\perp} + \text{span}\{f\}$

(2-20b) $[g_i, \Omega_{k^*}^{\perp}] \subset \Omega_{k^*}^{\perp} + \text{span}\{g_i\}$.

The following lemma ensures that, under suitable conditions, $\Omega_{k^*}^{\perp}$ is the maximal distribution which is locally weakly invariant under $f, g_1, \cdots, g_m$ and involutive.

[lemma 2-9]

Suppose that span$\{dh_1, \cdots, dh_r\}$ and $\Omega_{k^*}^{\perp} = \Delta$ are nonsingular. Moreover, we will assume that $\dim(\Delta + \text{span}\{f\}) - \dim(\Delta)$ and $\dim(\Delta + \text{span}\{g_i\}) - \dim(\Delta)$ are constant on the open dense subset $M^* \in M$. If, for any smooth vector field $\theta \in \Delta$, there exist the smooth vector fields $\theta', \theta'_i \in \Delta$ and the smooth functions $c, c_1, \cdots, c_m$ which satisfy

(2-21a) $[f, \theta] = \theta' + c f$

(2-21b) $[g_i, \theta] = \theta'_i + c_i g_i$

then the distribution Δ will be involutive and be the maximal locally weakly invariant distribution under the vector fields $f, g_1,$

$\cdots, g_m$ which is contained in span$\{dh_1, \cdots, dh_r\}^{\perp}$. □

(proof)

We will first prove that Δ is the maximal distribution which satisfies eq.(2-21). Let Δ' be any distribution which satisfies eq.(2-21) and is contained in span$\{dh_1, \cdots, dh_r\}^{\perp}$. For any smooth covector field $\sigma \in \Delta'^{\perp} \cap \text{span}\{f\}^{\perp}$ and any smooth vector field $\theta \in \Delta'$, we have

$$(2\text{-}22) \qquad \langle L_f \sigma, \theta \rangle = L_f \langle \sigma, \theta \rangle - \langle \sigma, [f, \theta] \rangle$$
$$= L_f \langle \sigma, \theta \rangle - \langle \sigma, \theta' + c f \rangle$$
$$= 0$$

where $\theta' \in \Delta'$ and c is a smooth function which satisfies eq.(2-21). Thus, we have

$$(2\text{-}23) \qquad L_f (\Delta' \cap \text{span}\{f\}^{\perp}) \subset \Delta' .$$

Similarly, we have

$$(2\text{-}24) \qquad L_{g_i} (\Delta' \cap \text{span}\{g_i\}^{\perp}) \subset \Delta' .$$

If for some $k \geqq 0$

$$(2\text{-}25) \qquad \Omega_k \subset \Delta'^{\perp}$$

then

$$(2\text{-}26) \qquad \Omega_{k+1} = \Omega_k + L_f(\text{span}\{f\}^{\perp} \cap \Omega_k) + \sum_{i=1}^{m} L_{g_i}(\text{span}\{g_i\}^{\perp} \cap \Omega_k) \subset \Delta'^{\perp} .$$

Since Ω_0 = span$\{dh_1, \cdots, dh_r\}^{\perp} \subset \Delta'$, we can deduce that

$$(2\text{-}27) \qquad \Omega_{k^*} \subset \Delta'^{\perp}$$

$$(2\text{-}28) \qquad \Delta' \subset \Omega_{k^*}^{\perp} = \Delta .$$

Thus, Δ is the maximal distribution which satisfies eq.(2-21).

We will show that Δ is involutive. Since Δ is nonsingular, at each point p we may find a neighborhood U of p and vector fields $\theta_1, \cdots, \theta_d$ such that on U,

(2-29) $\Delta = \mathrm{span}\{\theta_1, \cdots, \theta_d\}$

where d denotes the dimension of Δ. Consider the following distribution

(2-30) $D = \mathrm{span}\{\theta_1, \cdots, \theta_d\} + \mathrm{span}\{ [\theta_i, \theta_j] \mid 1 \leqq i, j \leqq d \}$.

The nonsingularity of the codistribution $\mathrm{span}\{dh_1, \cdots, dh_r\}$ ensures that $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$ is involutive. We have

(2-31) $[\Delta, \Delta] \subset \mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$.

Thus,

(2-32) $D \subset \mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$.

Therefore, if we can show that D satisfies eq.(2-21), then D must coincide with Δ because Δ is maximal. Consider the neighborhood V of the point p which is a regular point of D and $D + \mathrm{span}\{f\}$. Since D and $D + \mathrm{span}\{f\}$ are nonsingular on V, the relation

(2-33) $[f, D] \subset D + \mathrm{span}\{f\}$

implies that, for any $\theta \in D$, there exist $\theta' \in D$ and the smooth function c which satisfies eq.(2-21a) on V. Every smooth vector field $\nu \in D$ can be expressed on V as $\nu = \nu_1 + \nu_2$ where $\nu_1 \in \Delta$ and

(2-34) $\nu_2 = \sum_{i=1}^{d} \sum_{j=1}^{d} c_{ij} [\theta_i, \theta_j]$

for some smooth functions c_{ij}. Since

(2-35) $[f, \Delta] \subset \Delta + \mathrm{span}\{f\} \subset D + \mathrm{span}\{f\}$,

it is enough for our purposes to show that

(2-36) $[f, [\theta_i, \theta_j]] \subset D + \text{span}\{f\}$.

If we set

(2-37) $[f, \theta_i] = \theta'_i + b_i f$

for $\theta'_i \in \Delta$ and a smooth function b_i, then

(2-38)
$$
\begin{aligned}
[f, [\theta_i, \theta_j]] &= [\theta_i, [f, \theta_j]] - [\theta_j, [f, \theta_i]] \\
&= [\theta_i, \theta'_j + b_j f] \\
&\quad - [\theta_j, \theta'_i + b_i f] \\
&= [\theta_i, \theta'_j] - [\theta_j, \theta'_i] \\
&\quad + [\theta_i, b_j f] - [\theta_j, b_i f] \\
&\in D + \Delta + \text{span}\{f\} \subset D + \text{span}\{f\} .
\end{aligned}
$$

Thus, D satisfies eq.(2-21a) on V. Similarly D will also satisfy eq.(2-21b) in the neighborhood of a regular point of D and $D+\text{span}\{g_i\}$. Since the set of regular points of D, $D+\text{span}\{f\}$ and $D+\text{span}\{g_i\}$ form an open dense subset of M, D coincides with Δ in an open dense subset of M. Futhermore, Δ is nonsingular; D is smooth; and $\Delta \subset D$ by construction. Thus, we have $D = \Delta$ on the whole of M using lemma A-3(Appendix). This means that Δ is involutive.

Finally, we can conclude that, from lemma 1-5, Δ is locally weakly invariant under the vector fields $f, g_1, \cdots, g_m$ and is contained in $\text{span}\{dh_1, \cdots, dh_r\}^\perp$. Since Δ is the maximal distribution contained in $\text{span}\{dh_1, \cdots, dh_r\}^\perp$ satisfying eq.(2-21), Δ is the maximal distribution which is locally weakly invariant under the vector fields $f, g_1, \cdots, g_m$ and is contained in $\text{span}\{dh_1, \cdots, dh_r\}^\perp$. ■

3-3. Disturbance Decoupling Problem

Controlled invariant distribution is another important concept which can be used to analyze the invariant sutrcture of nonlinear systems. This is the nonlinear version of the (A,B) invariant subspace from linear system theory. Controlled invariant distribution was independently introduced by Ishijima[19][20], Isidori et.al.[22][23], Hirschon[24] and Nomura et.al.[21]. Here, we will extend their formulation using weakly invariant structure and then define weakly controlled invariant distribution. We will use weakly controlled invariant distribution to solve the disturbance decoupling problem in a new time scale --the wide-sense disturbance decoupling problem. In the wide-sense disturbance decoupling problem, we will design a feedback law so that the output as a function of an appropriate time scale τ is independent of the disturbance. In other words, we aim to design the feedback law such that the output's curve in output space (not the output trajectory as a function of the actual time t) can be separated from the disturbance.

We will first investigate the weakly invariant structure of a controlled system (section 3-3-1) and then use it for the disturbance decoupling problem which isolates the output's curve from the disturbance (section 3-3-2).

3-3-1. Weakly Controlled Invariant Distribution

At first we will review the notion of controlled invariant distribution. We used invariant distribution to decompose the original system in section 3-2. The notion of controlled invariant distribution is used in order to decompose a system modified by feedback.

Consider the system

$$(3\text{-}1) \qquad \frac{dp}{dt} = f(p) + \sum_{i=1}^{m} g_i(p)\,u_i$$

and feedback of the form

$$(3\text{-}2) \qquad u_i = \alpha_i(p) + \sum_{j=1}^{m} \beta_{ij}(p)\; v_j$$

where α_i and β_{ij} are real-valued smooth functions of the state p and we assume that the m×m matrix $\beta = [\beta_{ij}]$ is nonsingular. The real-valued function $v = (v_1, \cdots, v_m)$ is new input to the modified system.

We modify the original dynamics (3-1) with the feedback (3-2), and then we obtain the controlled system

$$(3\text{-}3) \qquad \frac{dp}{dt} = \tilde{f}(p) + \sum_{i=1}^{m} \tilde{g}_i(p)\,v_i$$

in which

$$(3\text{-}4a) \qquad \tilde{f}(p) = f(p) + \sum_{i=1}^{m} g_i(p)\,\alpha_i(p)$$

$$(3\text{-}4b) \qquad \tilde{g}_i(p) = \sum_{j=1}^{m} g_j(p)\,\beta_{ji}(p) \;.$$

In order to study the invariant structure of this controlled system, the notion of controlled invariance needs to be introduced.

The distribution Δ is said to be controlled invariant if there exists a smooth feedback pair (α, β) defined on **M** with the property that Δ is invariant under the vector fields $\tilde{f}$, $\tilde{g}_1, \cdots, \tilde{g}_m$ defined by eq.(3-4), i.e.

$$(3\text{-}5a) \qquad [\tilde{f}, \Delta] \subset \Delta$$

$$(3\text{-}5b) \qquad [\tilde{g}_i, \Delta] \subset \Delta.$$

A local version of controlled invariant distribution was also defined to study the local invariant structure of the controlled system.

The distribution Δ is said to be locally controlled invariant if for each $p \in M$ there exists a neighborhood U of p with the property that Δ is controlled invariant on U. From the previous definition, this requires the existence of the smooth feedback pair (α, β) defined on U satisfying eq.(3-5) for all q on U.

Since the notion of controlled invariant distribution allows us to investigate the invariant structure of the controlled system, controlled invariant distribution is used to solve the disturbance decoupling problem.

The following proposition gives a simple geometric test for locally controlled invariant distribution.

[proposition 3-1]

Let Δ be an involutive distribution. Suppose Δ, G and $\Delta + G$ are nonsingular on M. Then Δ is locally controlled invariant if and only if

(3-6a) $\quad [f, \Delta] \subset \Delta + G$

(3-6b) $\quad [g_i, \Delta] \subset \Delta + G$. □

We will extend this concept of controlled invariant distribution and then define weakly controlled invariant distribution. This notion can be used to investigate the invariant structure of the controlled system in a new time scale.

[definition 3-2]

Consider system (3-1). A distribution Δ is said to be weakly controlled invariant if there exists a smooth feedback pair (α, β) defined on M with the property that Δ is weakly invariant under the vector fields $\tilde{f}$, $\tilde{g}_1, \cdots, \tilde{g}_m$ defined by (3-4). □

This definition requires the existence of the smooth positive functions s and s_i such that

(3-7a) $[s\tilde{f}, \Delta] \subset \Delta$

(3-7b) $[s_i \tilde{g}_i, \Delta] \subset \Delta.$

Since the vector field $\tilde{g}_i$ can be modified by input transformation β, condition (3-7) corresponds to the existence of the feedback pair (α, β) such that

(3-8a) $[s\tilde{f}, \Delta] \subset \Delta$

(3-8b) $[\tilde{g}_i, \Delta] \subset \Delta.$

[definition 3-3]

The distribution Δ is said to be locally weakly controlled invariant if for each $p \in M$ there exists a neighborhood U of p with the property that Δ is weakly controlled invariant on U. □

The following lemma gives a geometric test for locally weakly controlled invariant distribution.

[lemma 3-4]

Let Δ be a nonsingular involutive distribution and assume that $\Delta + G$ is nonsingular. Moreover, we will assume that $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is constant on the dense subset $M^* \subset M$. Then Δ is locally weakly controlled invariant if and only if

(3-9) $[g_i, \Delta] \subset \Delta + G$

and, for each smooth vector field $\theta \in \Delta$, there exist the smooth function c and the smooth vector field $\nu \in \Delta + G$ which satisfy

(3-10) $[f, \theta] = \nu + cf.$ □

The proof of this lemma is analogous to that of lemma 1-5, so we will omit it.

3-3-2. Wide-Sense Disturbance Decoupling Problem

In this section, we will define the wide-sense disturbance decoupling problem and solve this problem using the notion of weakly controlled invariant distribution. In the wide-sense disturbance decoupling problem, we aim to design a feedback law such that the output as a function of the new time scale τ will be independent of the disturbance. In other words, we will design a feedback law such that the output's curve (not the output trajectory as a function of the actual time t) can be separated from the disturbance.

The wide-sense disturbance decoupling problem can be formulated as follows. Consider the system

(3-11a) $\frac{dp}{dt} = f(p) + \sum_{i=1}^{m} g_i(p) u_i + \sum_{j=1}^{m'} d_j(p) w_j$

(3-11b) $y_i = h_i(p) \qquad (i=1,\cdots,r)$

where the additional input $w=(w_1,\cdots, w_{m'})^T$ represents an undesired perturbation which influences the behavior of the system through the vector fields $d_1,\cdots, d_{m'}$. For simplicity, we will use the following notation.

(3-12a) $G = \mathrm{span}\{g_1(p),\ g_2(p),\ \cdots,\ g_m(p)\}$

(3-12b) $D = \mathrm{span}\{d_1(p),\ d_2(p),\ \cdots,\ d_{m'}(p)\}$.

[definition 3-5]

The wide-sense disturbance decoupling problem involves seeking the new time scale τ and a feedback law (3-2) for the system (3-11) such that the output y of the resultant system expressed in the time scale τ will not be affected by the disturbance w. □

Obviously, in the wide-sense disturbance decoupling problem, we seek a feedback law which will make the resultant system's output curve independent of the disturbance.

The wide-sense disturbance decoupling problem can be solved as follows. If we can find the nonsingular, involutive and d dimensional weakly controlled invariant distribution Δ which satisfies

(3-13) $$D \subset \Delta \subset \mathrm{span}\{ dh_1, \cdots, dh_r \}^{\perp},$$

then we can decouple the output of the system from the disturbance in a local coordinate (U, x) using the appropriate time scaling function s and feedback pair (α, β) in the following way.

(3-14a) $$\frac{dx_1}{d\tau} = \hat{f}_1(x_1, x_2) + \sum_{i=1}^{m} \hat{g}_{i1}(x_1, x_2)\ v_i + \sum_{j=1}^{m'} \hat{d}_{j1}(x_1, x_2)\ w_j$$

(3-14b) $$\frac{dx_2}{d\tau} = \hat{f}_1(x_2) + \sum_{i=1}^{m} \hat{g}_{i2}(x_2)\, v_i$$

(3-14c) $$y_i = h_i(x_2) \qquad (i=1, \cdots, r)$$

(3-15a) $$\hat{f}(p) = s(p)\ \tilde{f}(p)$$

$$= s(p)\{ f(p) + \sum_{i=1}^{m} g_i(p)\alpha_i(p)\}$$

(3-15b) $$\hat{g}_i(p) = s_i(p)\tilde{g}_i(p)$$

$$= s(p) \sum_{j=1}^{m} g_j(p)\beta_{ji}(p) \frac{s_i(p)}{s(p)}$$

(3-16) $$\frac{dt}{d\tau} = s(p) > 0$$

where (x_1, x_2) is a partition of x and $\dim(x_1) = d$. It is apparent that the output of the resultant controlled system is not influenced by the disturbance in the time scale τ.

In this formulation, we are able to reduce the wide-sense disturbance decoupling problem to a problem requiring a weakly controlled invariant distribution which contains D and is contained in $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$.

The notion of locally weakly controlled invariance is easier to manage than the global one. This motivates us to consider the local version of the wide-sense disturbance decoupling problem. The local wide-sense disturbance decoupling problem involves finding a locally weakly controlled invariant distribution which contains D and is contained in $\mathrm{span}\{dh_1,\cdots, dh_r\}^{\perp}$.

If we can find the maximal locally weakly controlled invariant distribution which is contained in $\mathrm{span}\{dh_1,\cdots, dh_r\}^{\perp}$, then we can judge whether the local wide-sense disturbance decoupling problem is solvable or not. The local wide-sense disturbance decoupling problem is solvable if and only if D is contained in the maximal locally weakly controlled invariant distibution contained in $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$.

The following algorithm will enable us to obtain a candidate for the maximal locally weakly controlled invariant distribution which is contained in $\mathrm{span}\{dh_1,\cdots, dh_r\}^{\perp}$.

[algorithm 3-6]

(3-17a) $\Omega_0 = \mathrm{span}\{ dh_1 ,\cdots, dh_r \}$

(3-17b) $\Omega_{k+1} = \Omega_k + L_f((G + \mathrm{span}\{ f \})^{\perp} \cap \Omega_k) + \sum_{i=1}^{m} L_{g_i}(G^{\perp} \cap \Omega_k)$.□

If there exists an integer k^* such that $\Omega_{k^*+1} = \Omega_{k^*}$, then $\Omega_{k^*}{}^{\perp} = \Delta$ will satisfy the following equations on any open subset where $\mathrm{span}\{f\}^{\perp} \cap \Delta$ and $G^{\perp} \cap \Delta$ are smooth.

(3-18a) $[f , \Delta] \subset \Delta + G + \mathrm{span}\{ f \}$

(3-18b) $[g_i , \Delta] \subset \Delta + G$.

[lemma 3-7]

Assume that $\mathrm{span}\{dh_1,\cdots,dh_r\}$, $\Delta = \Omega_{k^*}{}^{\perp}$ and $\Delta + G$ are nonsingular and that $\dim(\Delta + \mathrm{span}\{f\}) - \dim(\Delta)$ is constant on the dense subset $M^* \subset M$. Moreover, we will assume that, for any smooth vector field $\theta \in \Delta$, there exist the smooth vector field $\nu \in \Delta + G$ and the smooth function c which satisfy

(3-19) $$[f, \theta] = \nu + c f.$$

Then Δ will be the maximal locally weakly controlled invariant distribution which is contained in $\mathrm{span}\{dh_1, \cdots, dh_r\}^{\perp}$ and Δ will also be involutive.□

The proof of this lemma is analogous to that of lemma 2-9, so it can be omited.

[example]

Consider the system

$$\frac{dx}{dt} = f(x) + g u + d w$$

$$y_i = h_i(x) \qquad (i=1,2)$$

where $x=(x_1, x_2, x_3)^T$, $f(x)=(x_2 e^{x_1}, x_3 e^{x_1}, 0)^T$, $g=(0,0,1)^T$ $d=(1,1,1)^T$, $h_1(x)=2x_1-x_2-x_3$ and $h_2(x)=x_1-x_3$. Obviously,

$$\mathrm{span}\{dh_1, dh_2\}^{\perp} = \mathrm{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\} = \mathrm{span}\ d$$

$$\triangleq \Delta.$$

This distribution Δ is not locally controlled invariant because

$$[f, d] = -\begin{bmatrix}(x_2+1)e^{x_1}\\(x_3+1)e^{x_1}\\0\end{bmatrix} \notin \mathrm{span}\{d, g\}.$$

So we are not able to decouple the output from the disturbance in the actual time scale t, but we will show that the local wide-sense disturbance decoupling problem is solvable. Since

$$[f, d] = -e^{x_1}d + e^{x_1}g - f$$

and $s(x)=c_1 e^{-x_1}+c_2 e^{-x_2}+c_3 e^{-x_3}$ (where c_1, c_2, c_3 are non-negative constants and some of them must be non-zero) is the positive solution to the differential equation

$$-s=\frac{\partial s}{\partial x}d=\frac{\partial s}{\partial x_1}+\frac{\partial s}{\partial x_2}+\frac{\partial s}{\partial x_3}\ .$$

From this we see that Δ is locally weakly controlled invariant. Using the time scaling function s, we obtain

$$[sf, d]=-\begin{bmatrix} c_1+c_2 e^{x_1-x_2}+c_3 e^{x_1-x_3} \\ c_1+c_2 e^{x_1-x_2}+c_3 e^{x_1-x_3} \\ 0 \end{bmatrix}$$

$$\in \text{span}\{ g, d \}.$$

Since the dimension of Δ and $\Delta+\text{span}\ g$ are constant, the local wide-sense disturbance decoupling problem has been solved.

3-4. Concluding Remarks

In this chapter, we investigated invariant structure using a time scale transformation. We introduced the notion of weakly invariant distribution in section 3-1 and showed how it can allow us to study invariant structure in a new time scale. We have proposed a simple geometric test for weak invariance. In section 3-2, weakly invariant distribution was used to obtain Kalman-like decompositions in reachable/unreachable parts and/or observable/unobservable parts in the transformed time scale. We have found the minimum dimension of the locally weakly invariant distribution under the vector fields f,g which contains g; this distribution coincides with the controllable subspace of a linear system. We have also proposed an algorithm to obtain the locally weakly invariant distribution under the vector fields f,g which is contained in the annihilator of the output function. This distribution coincides with the unobservable subspace of the linear system. In section 3-3, we introduced weakly controlled invariance, and successfully used it to solve the wide-sense disturbance decoupling problem. In this problem, it is necessary to seek a feedback law which would prevent the disturbance from affecting the output's curve in output space.

IV . FEEDBACK EQUIVALENCE

The linearization problem is significant in nonlinear system theory because it is possible to apply control strategies which were perfected in linear system theory to linearized systems. One of the most common linearization methods is first order approximation. With this method, we approximate a nonlinear system as a linear system in a neighborhood of the equilibrium point. We neglect the second order term in designing the controller. An alternative formulation was proposed in 1978 when the exact input-state linearization problem was solved by Brockett[42]. A slightly modified version of this problem was solved by Jakubczyk-Respondek[43], Su[44] and Hunt et al.[45]. With this method, we proceed to linearize the system using a state transformation and nonlinear state feedback.

An exact linearization method can, however, only be used on a restricted class of nonlinear systems: systems which can be expressed in appropriate coordinates as

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix} u$$

where a and b are functions of state and $b \neq 0$. These systems are <u>almost</u> linear systems. In order to relax this condition, several approximate linearization methods have been proposed[46][47].

Here, we will relax the linearizability condition by introducing a time scale transformation. We will define the notion of wide-sense feedback equivalence and identify the class of nonlinear systems which can be linearized in some transformed time scale.

4-1. Ordinary Feedback Equivalence

In this section, we will review the work of Su[44] and derive several lemmas concerned with ordinary feedback equivalence.

In this chapter, we consider a single input system expressed in local coordinates as

$$(1\text{-}1) \qquad \frac{dx}{dt} = f(x) + g(x)\,u$$

where f and g are C^∞ vector fields. We assume $x \in M = R^n$ $f(0) = 0$.

The system (1-1) is feedback equivalent to a linear system if there exists a C^∞ diffeomorphism

$$(1\text{-}2) \qquad \begin{array}{l} \Psi : M \times R \longrightarrow R^n \times R \\ \quad\; x \times u \longrightarrow y \;\; \times v \end{array}$$

such that the system can be expressed in $R^n \times R$

$$(1\text{-}3) \qquad \frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} = \begin{bmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} v.$$

The following proposition was derived by Su[44].

[proposition 1-1]

The system (1-1) is feedback equivalent to a linear system in a neighborhood of the origin if and only if there exists a neighborhood U of the origin such that

(1) $\{g, ad_f g, \cdots, ad_f^{n-1} g\}(x)$ span $T_x M$ for all $x \in U$, and

(2) $\{g, ad_f g, \cdots, ad_f^{n-2} g\}$ is involutive on U. □

Condition (2) in this proposition requires us to check $(n-1)(n-2)/2$ Lie brackets $[ad_f^i g, ad_f^j g]$ $(0 \leq i < j \leq n-2)$, but the following lemma will ensure that it is only necessary to check n-2 Lie brackets.

[lemma 1-2]

Let U be an open neighborhood of the origin. Suppose that $\{g, ad_f g, \cdots, ad_f^{n-1}g\}(x)$ span $T_x M$ for all $x \in U$, then the following three conditions will be equivalent on U.

(1) $\{g, ad_f g, \cdots, ad_f^{n-2}g\}$ is involutive.

(2) There exist C^∞ functions $\theta_i^{(k-1,k)}$ $(k=1,2,\cdots,n-2;$ $i=0,1,\cdots k)$ such that

$$(1\text{-}4) \qquad [ad_f^{k-1}g,\ ad_f^k g] = \sum_{i=0}^{k} \theta_i^{(k-1,k)}\, ad_f^i g$$

for $k=1,2,\cdots,n-2$.

(3) $\{g, ad_f g, \cdots, ad_f^k g\}$ is involutive for $k=1,2,\cdots,n-2$. □

(proof)

(1)→(3) : This will be proven by induction. For $k=n-2$, $\{g, ad_f g, \cdots, ad_f^{n-2}g\}$ is involutive from condition (1). Assume that $\{g, ad_f g, \cdots, ad_f^k g\}$ is involutive, then there exist C^∞ functions $\theta_i^{(j,r)}(x)$ such that

$$(1\text{-}5) \qquad [ad_f^j g,\ ad_f^r g] = \sum_{i=0}^{k} \theta_i^{(j,r)}\, ad_f^i g \qquad (j, r \leqq k).$$

On the other hand, from a property of Lie brackets, we have

$$(1\text{-}6) \qquad [ad_f^j g,\ ad_f^r g] = [ad_f^j g,\ [f,\ ad_f^{r-1}g]]$$

$$= -[f,\ [ad_f^{r-1}g,\ ad_f^j g]] - [ad_f^{r-1}g,\ [ad_f^j g,\ f]]$$

$$= [f,\ [ad_f^j g,\ ad_f^{r-1}g]] + [ad_f^{r-1}g,\ ad_f^{j+1}g].$$

If $j \leqq k-1$ and $r \leqq k$, then (1-6) can be rewritten using (1-5)

$$(1\text{-}7) \qquad [ad_f^j g,\ ad_f^r g] = \sum_{i=0}^{k} [f,\ \theta_i^{(j,r-1)}\, ad_f^i g] + \sum_{i=0}^{k} \theta_i^{(r-1,j+1)}\, ad_f^i g$$

$$= \{\theta_0^{(r-1,j+1)} + L_f\ \theta_0^{(j,r-1)}\}\ g + \sum_{i=1}^{k} \{\theta_{i-1}^{(j,r-1)} + L_f\ \theta_i^{(j,r-1)} + \theta_i^{(r-1,j+1)}\}\ ad_f^i g + \theta_k^{(j,r-1)}\ ad_f^{k+1} g.$$

Comparing (1-7) with (1-5), we can conclude that $\theta_k^{(j,r-1)} = 0$ ($j \leqq k-1$; $r-1 \leqq k-1$) because $\{g, ad_f g, \cdots, ad_f^{k+1} g\}$ are independent for $k \leqq n-2$. This means $\{g, ad_f g, \cdots, ad_f^{k-1} g\}$ is also involutive .

(2)→(3) : We will use induction. Obviously {g} is involutive. Assume that $\{g, ad_f g, \cdots, ad_f^k g\}$ is involutive , then there exist C^∞ functions $\theta_i^{(j,r)}$ which satisfy (1-5). To show that $\{g, ad_f g, \cdots, ad^{k+1} g\}$ is involutive, it is sufficient to check $[ad_f^j g, ad_f^{k+1} g]$ for $j \leqq k$.

(1-8)
$$[ad_f^j g, ad_f^{k+1} g] = [f, [ad_f^j g, ad_f^k g]] + [ad_f^k g, ad_f^{j+1} g]$$
$$= L_f\ \theta_0^{(j,k)}\ g + \sum_{i=1}^{k} \{\theta_{i-1}^{(j,k)} + L_f\ \theta_i^{(j,k)}\}\ ad_f^i g + \theta_k^{(j,k)}\ ad_f^{k+1} g + [ad_f^k g, ad_f^{j+1} g].$$

Using (1-5) if $j \leqq k-1$, or (1-4) if $j = k$, (1-8) will become,

(1-9)
$$[ad_f^j g, ad_f^{k+1} g] = \{\theta_0^{(k,j+1)} + L_f\ \theta_0^{(j,k)}\}\ g + \sum_{i=1}^{k} \{\theta_{i-1}^{(j,k)} + L_f\ \theta_i^{(j,k)} + \theta_i^{(k,j+1)}\}\ ad_f^i g + \{\theta_k^{(j,k)} + \theta_{k+1}^{(k,j+1)}\}\ ad_f^{k+1} g$$

where $\theta_{k+1}^{(k,j+1)} = 0$ if $j \neq k$. So, $\{g, ad_f g, \cdots, ad_f^{k+1} g\}$ is also involutive .

(3)→(1) and (3)→(2) : Obvious. ■

Condition (3) of lemma 1-2 was used as a condition for linearizability by Jakubczyk[43]. This lemma yields the following lemma.

[lemma 1-3]

System (1-1) is feedback equivalent to a linear system in a neighborhood of the origin if and only if there exists a neighborhood U of the origin such that

(1) $\{g, ad_f g, \cdots, ad_f^{n-1} g\}(x)$ span $T_x M$ for all $x \in U$, and

(2) there exist C^∞ functions $\theta_i^{(k-1,k)}$ $(k=1,2,\cdots,n-2;$ $i=0,1,\cdots k)$ which satisfy (1-4). □

4-2. Wide-Sense Feedback Equivalence

Here, we will consider linearization in the new time scale τ.

[definition 2-1]

System(1-1) is wide-sense feedback equivalent to a linear system if there exists a new time scale τ, defined by the C^∞ time scaling function $s(x)>0$ as $dt/d\tau = s(x)$, such that the system expressed in the time scale τ:

$$\frac{dx}{d\tau} = s(x) f(x) + g(x)\mu \tag{2-1}$$

is feedback equivalent to a linear system in the time scale τ. □

Obviously, a system which is wide-sense feedback equivalent to a linear system need not be linearizable in the actual time scale t. The following lemma is obvious from this definition.

[lemma 2-2]

System (1-1) is wide-sense feedback equivalent to a linear system in a neighborhood of the origin if and only if there exist: a neighborhood U of the origin; C^∞ functions $s(x) > 0$ and $\delta_i^k(x)$ $(k=1, 2,\cdots, n-2;\ i=0, 1,\cdots, k)$ such that

(1) $\{g,\ ad_{sf}g, \cdots,\ ad_{sf}^{n-1}g\}(x)$ span $T_x M$ for all $x \in U$, and

(2) the following equation is satisfied on U for k=1, 2,⋯, n-2.

$$[ad_{sf}^{k-1}g,\ ad_{sf}^{k}g] = \sum_{i=0}^{k} \delta_i^k\ ad_{sf}^{i}g \quad (2\text{-}2)$$ □

The rest of this section will be devoted to finding such s(x). In order to do this, we must examine these two conditions. First, we will examine the property of $ad_{sf}^{i}g$.

<u>[lemma 2-3]</u>

There exist $\xi_j^i(x)$ (j=0, 1, ⋯, i) and $\xi_f^i(x)$ which are C^∞ functions consisting of x, s(x), and partial derivatives of s(x), such that

$$ad_{sf}^{i}g = \sum_{j=0}^{i} \xi_j^i\ ad_f^j g + \xi_f^i\ f\ . \quad (2\text{-}3)$$

Also $\xi_i^i = \{s(x)\}^i$. □

(proof)

Obviously, this condition will be satisfied when i = 0. Assume that it is also satisfied when i = k , then

$$\begin{aligned} ad_{sf}^{k+1}g &= [sf,\ ad_{sf}^{k}g] \quad (2\text{-}4)\\ &= (s L_f\ \xi_0^k)\ g \\ &\quad + \sum_{j=1}^{k} (s\,\xi_{j-1}^k + s L_f\ \xi_j^k)\ ad_f^j g + s\,\xi_k^k\ ad_f^{k+1} g \\ &\quad + [s L_f\ \xi_f^k - \xi_f^k\ L_f\ s \\ &\qquad - \sum_{j=0}^{k} \{\xi_j^k\ L_{(ad_f^j g)}\ s\ \}\ ]\ f. \end{aligned}$$

This means that there exist C^∞ functions ξ_j^{k+1} (j=0, 1,⋯, k+1) and ξ_f^{k+1} which satisfy eq.(2-3). ■

From the last lemma and condition (1) in lemma 2-2, the following lemma can easily be derived.

[lemma 2-4]

If the system (1-1) is wide-sense feedback equivalent to a linear system, then $\{f, g, ad_f g, \cdots, ad_f^{n-1} g\}(x)$ span $T_x M$ for all $x \in M$. □

From now on, we will assume that $\{f, g, ad_f g, \cdots, ad_f^{n-1} g\}(x)$ span $T_x M$ for all $x \in M$. The following lemma ensures that there does not exist a set of C^∞ functions $\{\bar{\xi}_j^i, \bar{\xi}_f^i\}$ which is different from $\{\xi_j^i, \xi_f^i\}$ and satisfy

$$ad_{sf}^i g = \sum_{j=0}^{i} \bar{\xi}_j^i \, ad_f^j g + \bar{\xi}_f^i f$$

for $i = 0, 1, \cdots, n-2$.

[lemma 2-5]

If the set of vector fields $\{f, g, ad_f g, \cdots, ad_f^{n-1} g\}(x)$ span $T_x M$ for all $x \in M$, then there does not exist an open subset $U \subset M$ on which there exist non-zero C^∞ functions $\psi_i(x)$ $(i=0,1,\cdots, n-2)$ and $\psi_f(x)$ satisfying

$$\psi_f \, f + \sum_{i=0}^{n-2} \psi_i \, ad_f^i g = 0 . \qquad (2\text{-}5)$$ □

(proof)

If such a U exists, then there must exist an open subset $V \subset U$ and an integer k $(0 \leq k \leq n-2)$ such that $\psi_k \neq 0$ for all $x \in V$ (i.e. ψ_k^{-1} is also a C^∞ function on V) and

$$\psi_f \, f + \sum_{i=0}^{k} \psi_i \, ad_f^i g = 0 \qquad (2\text{-}6)$$

on V. Thus, $ad_f^k g$ can be rewritten

$$ad_f^k g = -\psi_k^{-1} \psi_f \, f - \sum_{i=0}^{k-1} \psi_k^{-1} \psi_i \, ad_f^i g \qquad (2\text{-}7)$$

where $\psi_k^{-1} \psi_f$ and $\psi_k^{-1} \psi_i$ are C^∞ functions. Using this relation, we find that

$$\begin{aligned}
(2\text{-}8)\qquad ad_f^{k+1}g &= [f,\ ad_f^k\ g] \\
&= -[f,\ \psi_k^{-1}\psi_f\ f] - \sum_{i=0}^{k-1}[f,\ \psi_k^{-1}\psi_i\ ad_f^i g] \\
&= -\ \{L_f(\psi_k^{-1}\psi_f)\}\ f - \{L_f(\psi_k^{-1}\psi_0)\}\ g \\
&\qquad - \sum_{i=1}^{k-1}\{\ \psi_k^{-1}\psi_{i-1} + L_f(\psi_k^{-1}\psi_i)\}\ ad_f^i g \\
&\qquad - \psi_k^{-1}\psi_{k-1}\ ad_f^k g \\
&= \{\psi_k^{-2}\psi_{k-1}\ \psi_f - L_f(\psi_k^{-1}\psi_f)\}\ f \\
&\qquad + \{\psi_k^{-2}\psi_{k-1}\ \psi_0 - L_f(\psi_k^{-1}\psi_0)\}\ g \\
&\qquad + \sum_{i=1}^{k-1}\{\psi_k^{-2}\psi_{k-1}\ \psi_i - \psi_k^{-1}\psi_{i-1} \\
&\qquad\qquad - L_f(\psi_k^{-1}\psi_i)\}\ ad_f^i g\ .
\end{aligned}$$

We can calculate $ad_f^{k+2}g$, $ad_f^{k+3}g, \cdots$, and $ad_f^{n-1}g$ similarly, and we will easily find that

$$\begin{aligned}
(2\text{-}9)\qquad &\text{span}\ \{f,\ g,\ ad_f\ g, \cdots,\ ad_f^{n-1}g\}(x) \\
&= \text{span}\ \{f,\ g,\ ad_f\ g,\ \cdots,\ ad_f^{k-1}g\}(x)
\end{aligned}$$

for all $x \in V$. This contradicts our assumption that the set of vector fields $\{f,\ g,\ ad_f g, \cdots,\ ad_f^{n-1}g\}(x)$ span $T_x\ M$ for all $x \in M$. ■

Next we will examine condition (2) in lemma 2-2.

[lemma 2-6]

If system (1-1) is wide-sense feedback equivalent to a linear system, then there exist unique C^∞ functions $\eta_i^{(j,r)}(x)$ and $\eta_f^{(j,r)}(x)$ $(j,r \leqq k;\ i=0,\ 1, \cdots,\ q)$ such that

$$(2\text{-}10)\qquad [ad_f^j g,\ ad_f^r g] = \sum_{i=0}^{q}\eta_i^{(j,r)}\ ad_f^i g + \eta_f^{(j,r)}\ f$$

$$(2\text{-}11)\qquad q = \begin{cases} k+1 & (\ 2 \leqq k\ ) \\ k & (\ k \leqq 1\ ) \end{cases}$$

for k = 0, 1, ⋯, n-3 ($4 \leqq n$), or k = 0, 1,⋯, n-2 ($n \leqq 3$). □

Before proving this lemma, we will prove a related lemma.

[lemma 2-7]

If there exist C^∞ functions $\eta_i^{(k-1,k)}(x)$ and $\eta_f^{(k-1,k)}(x)$ (k=1, 2,⋯, n-3; i=0, 1,⋯, k+1) such that

$$(2\text{-}12)\qquad [ad_f^{k-1}g,\ ad_f^k g] = \sum_{i=0}^{k+1} \eta_i^{(k-1,k)}\ ad_f^i g + \eta_f^{(k-1,k)}\ f\ ,$$

then there also must exist C^∞ functions $\eta_i^{(j,r)}$ (x) and $\eta_f^{(j,r)}$ (x) ($j,r \leqq k$; i=0, 1,⋯, k+1) such that

$$(2\text{-}13)\qquad [ad_f^j g,\ ad_f^r g] = \sum_{i=0}^{k+1} \eta_i^{(j,r)}\ ad_f^i g + \eta_f^{(j,r)}\ f$$

for k = 0, 1,⋯, n-3.□

(proof)

Obviously (2-13) is satisfied in the case of k=1. Assume that eq.(2-13) is satisfied for $k \leqq z$. For $j \leqq z$, we have

$$\begin{aligned}(2\text{-}14)\qquad [ad_f^j g,\ ad_f^{z+1} g] &= [f,\ [ad_f^j g,\ ad_f^z g]] + [ad_f^z g,\ ad_f^{j+1} g] \\ &= \{L_f\ \eta_0^{(j,z)}\} g \\ &\quad + \sum_{i=1}^{z+1} \{\eta_{i-1}^{(j,z)} + L_f\ \eta_i^{(j,z)}\}\ ad_f^i g \\ &\quad + \eta_{z+1}^{(j,z)}\ ad_f^{z+2} g + \{L_f\ \eta_f^{(j,z)}\}\ f \\ &\quad + [ad_f^z g,\ ad_f^{j+1} g]\ .\end{aligned}$$

The last term of eq.(2-14) is an element of Δ^{z+1} if $j < z$, or an element of Δ^{z+2} if $j = z$ because of eq.(2-12), where Δ^z is the set of C^∞ vector fields which are the linear combination of vector fields $\{f,\ g,\ ad_f g, \cdots,\ ad_f^z\ g\}$. Thus, eq.(2-13) is also satisfed for $k \leqq z+1$.■

(proof of lemma 2-6)

When k=0, the proof is obvious. In the case of where k=1, the proof is as follows.

(2-15) $$[g,\ ad_{sf}g]=[g,\ [sf,\ g]]$$
$$=[g,\ s\ (ad_f g)-\ (L_g s)\ f]$$
$$=s\ [g,\ ad_f g]\ +2(L_g s)\ ad_f g-\ (L_g^2 s)\ f\ .$$

Since condition (2) in lemma 2-2 requires that $[g,\ ad_{sf}g]$ must be an element of Δ^1, we have $[g,\ ad_f\ g]\in\Delta^1\subset\Delta^2$. Assume that eq.(2-10) is satisfied for $k\leqq z$. In order to show that eq.(2-10) is also satisfied for $k\leqq z+1$, it is enough to check the case of $j=z, r=z+1$. From eq.(2-3), we have

(2-16) $$[ad_{sf}^z g,\ ad_{sf}^{z+1}g]$$
$$=[\sum_{i=0}^{z}\xi_i^z\ ad_f^i g+\xi_f^z\ f,\ \sum_{j=0}^{z+1}\xi_j^{z+1}\ ad_f^j g+\xi_f^{z+1}\ f]$$
$$=\xi_z^z\ \xi_{z+1}^{z+1}\ [ad_f^z g,\ ad_f^{z+1}g]$$
$$+\sum_{i=0}^{z-1}\xi_i^z\ \xi_{z+1}^{z+1}\ [ad_f^i g,\ ad_f^{z+1}g]$$
$$+\sum_{i=0}^{z}\sum_{j=0}^{z}\xi_i^z\ \xi_j^{z+1}\ [ad_f^i g,\ ad_f^j g]$$
$$+\sum_{i=0}^{z}\sum_{j=0}^{z+1}[\{\xi_i^z\ L_{(ad_f^i g)}\xi_j^{z+1}\}ad_f^j g$$
$$-\ \{\xi_j^{z+1}\ L_{(ad_f^j g)}\xi_i^z\}\ ad_f^i g]$$
$$+\sum_{i=0}^{z}[-\xi_i^z\ \xi_f^{z+1}\ ad_f^{i+1}g-\ (\xi_f^{z+1}\ L_f\ \xi_i^z)\ ad_f^i g$$
$$+\ \{\xi_i^z\ L_{(ad_f^i g)}\xi_f^{z+1}\}\ f]$$
$$+\sum_{j=0}^{z+1}[\xi_f^z\ \xi_j^{z+1}\ ad_f^{j+1}g+\ (\xi_f^z\ L_f\ \xi_j^{z+1})\ ad_f^j g$$
$$-\ \{\xi_j^{z+1}\ L_{(ad_f^j g)}\xi_f^z\}\ f]$$

$$+ \{\xi_f^z L_f \xi_f^{z+1} - \xi_f^{z+1} L_f \xi_f^z\} f .$$

For $0 \leq i \leq z-1$, we have

$$(2\text{-}17) \qquad [ad_f^i g, ad_f^{z+1} g] = [f, [ad_f^i g, ad_f^z g]] + [ad_f^z g, ad_f^{i+1} g]$$
$$\in \Delta^{z+2} + \Delta^{z+1} \subset \Delta^{z+2} .$$

The last five terms of (2-16) are evidently elements of Δ^{z+2}. Thus, $[ad_f^z g, ad_f^{z+1} g]$ must be an element of Δ^{z+2} because $[ad_{sf}^z g, ad_{sf}^{z+1} g] \in \Delta^{z+1}$ (condition (2) of lemma 2-2). Uniqueness follows from lemma 2-5. ■

Using ξ in lemma 2-3 and η in lemma 2-6, we can derive the differential equations which $s(x)$ must satisfy. We rewrite eq.(2-16) using eq.(2-10), and we have

$$(2\text{-}18) \qquad [ad_{sf}^{k-1} g, ad_{sf}^k g]$$
$$= \sum_{r=0}^{k+1} \{ \sum_{i=0}^{k-1} \sum_{j=0}^{k} \xi_i^{k-1} \xi_j^k \eta_r^{(i,j)}$$
$$+ \sum_{i=0}^{k-1} \xi_i^{k-1} L_{(ad_f^i g)} \xi_r^k$$
$$- \sum_{j=0}^{k} \xi_j^k L_{(ad_f^j g)} \xi_r^{k-1}$$
$$- \xi_{r-1}^{k-1} \xi_f^k - \xi_f^k L_f \xi_r^{k-1}$$
$$+ \xi_f^{k-1} \xi_{r-1}^k + \xi_f^{k-1} L_f \xi_r^k \} ad_f^r g$$
$$+ \{ \sum_{i=0}^{k-1} \xi_i^{k-1} L_{(ad_f^i g)} \xi_f^k - \sum_{j=0}^{k} \xi_j^k L_{(ad_f^j g)} \xi_f^{k-1}$$
$$- \xi_f^k L_f \xi_f^{k-1} + \xi_f^{k-1} L_f \xi_f^k$$
$$+ \sum_{i=0}^{k-1} \sum_{j=0}^{k} \xi_i^{k-1} \xi_j^k \eta_f^{(i,j)} \} f$$
$$\triangleq \sum_{r=0}^{k+1} \zeta_r^k ad_f^r g + \zeta_f^k f$$

where ζ is a C^∞ function consisting of x,s, partial derivatives of s and η. (We are assuming that $\xi^i_j=0$ for j > i or j < 0.)
On the other hand , condition (2) in lemma 2-2 can be rewritten as

$$(2\text{-}19)\qquad [ad^{k-1}_{sf}g,\ ad^k_{sf}g] = \sum_{r=0}^{k} \delta^k_r\ ad^r_{sf}g$$
$$= \sum_{r=0}^{k}\sum_{i=0}^{r} \delta^k_r\ \xi^r_i\ ad^i_f g + \sum_{r=0}^{k} \delta^k_r\ \xi^r_f\ f$$
$$= \sum_{i=0}^{k}\sum_{r=i}^{k} \delta^k_r\ \xi^r_i\ ad^i_f g + \sum_{r=0}^{k} \delta^k_r\ \xi^r_f\ f .$$

Compare (2-18) with (2-19) , keeping the uniqueness of η and ξ in mind, we can easily conclude that s(x) > 0 and δ are solutions to the following differential equations

$$(2\text{-}20a)\qquad \zeta^k_{k+1} = 0$$

$$(2\text{-}20b)\qquad \zeta^k_i = \sum_{r=i}^{k} \delta^k_r\ \xi^r_i \qquad (i = 0, 1,\cdots, k)$$

$$(2\text{-}20c)\qquad \zeta^k_f = \sum_{r=0}^{k} \delta^k_r\ \xi^r_f$$

for k=1,2,…, n-3 (n≧4), or k=1,2, …, n-2 (n≦3).

Until now , we have been investigating the necessary conditions for nonlinear systems to be wide-sense feedback equivalent to linear systems. These conditions will be sufficient if we add some further conditions.

<u>[theorem 2-8]</u>

System (1-1) is wide-sense feedback equivalent to a linear system in a neighborhood of the origin if and only if there exists a neighborhood U of the origin such that

(1) there exist C^∞ functions $\eta^{(j,r)}_i$ and $\eta^{(j,r)}_f$ (j,r≦k; i=0, 1, …, q ; q is defined in (2-11)) which satisfy equation (2-10) for k =0, 1,…, n-3 (n≧4), or k=0, 1, …, n-2 (n≦3) on U,

(2) the differential equations (2-20) have C^∞ solutions $s(x) > 0$, δ_i^k and δ_f^k for $k = 1, 2, \cdots, n-3$ $(n \geqq 4)$ or $k = 1, 2, \cdots, n-2$ $(n \leqq 3)$ on U,

(3) in the case where $n \geqq 4$, there exist C^∞ functions δ_i^{n-2} $(i = 0, 1, \cdots, n-2)$ which satisfy the following equation on U

$$[ad_{sf}^{n-3}g,\ ad_{sf}^{n-2}g] = \sum_{i=0}^{n-2} \delta_i^{n-2}\ ad_{sf}^{i}g \tag{2-21}$$

where $s(x)$ is a solution to condition (2), and

(4) $\{g, ad_{sf}g, \cdots, ad_{sf}^{n-1}g\}(x)$ span $T_x M$ for all $x \in U$. □

Obviously $n \leqq 3$ is a special case. The subsequent corollaries follow from this theorem.

[corollary 2-9]

System (1-1) (n=3) is wide-sense feedback equivalent to a linear system in a neighborhood of the origin if and only if there exists a neighborhood U of the origin such that

(1) there exist C^∞ functions $\eta_0^{(0,1)}$, $\eta_1^{(0,1)}$, and $\eta_f^{(0,1)}$ satisfying the following equation on U

$$[g,\ ad_f g] = \eta_0^{(0,1)} g + \eta_1^{(0,1)} ad_f g + \eta_f^{(0,1)} f \tag{2-22}$$

(2) there exists a C^∞ function $s(x) > 0$ which satisfies

$$L_g^2 s - \frac{2}{s} (L_g s)^2 - \eta_1^{(0,1)} L_g s - s \eta_f^{(0,1)} = 0 \tag{2-23}$$

and

(3) $\{g, ad_{sf}g, ad_{sf}^2 g\}(x)$ span $T_x M$ for all $x \in U$. □

[corollary 2-10]

System (1-1) (n=2) is wide-sense feedback equivalent to a linear system in a neighborhood of the origin if and only if there exist a neighborhood U of the origin and a C^∞ function $s(x) > 0$

such that $\{g, ad_{sf}g\}(x)$ span $T_x M$ for all $x \in U$. □

[corollary 2-11]

System (1-1) (n=1) is wide-sense feedback equivalent to a linear system in a neighborhood of the origin if and only if there exists a neighborhood U of the origin such that $g(x) \neq 0$ for all $x \in U$. □

(proof)

Corollary 2-10, 11 are obvious from lemma 2-2 , so we will prove only the case where n=3 (corollary 2-9). Since conditions (1) and (3) in this corollary are equivalent to those of (1) and (4) in theorem 2-8, it is enough to prove that condition (2) in this corollary is equivalent to that of (2) in theorem 2-8. ζ in theorem 2-8 will be calculated as follows. Since

$$(2\text{-}24) \qquad [g, ad_{sf}g] = [g, s\, ad_f g - (L_g s) f]$$
$$= s[g, ad_f g] + 2(L_g s)\, ad_f g - (L_g^2 s) f$$
$$= s\eta_0^{(0,1)} g + \{s\eta_1^{(0,1)} + 2L_g s\}\, ad_f g + \{s\eta_f^{(0,1)} - L_g^2 s\} f ,$$

we have

$$(2\text{-}25a) \qquad \zeta_0^1 = s\eta_0^{(0,1)}$$

$$(2\text{-}25b) \qquad \zeta_1^1 = \{s\eta_1^{(0,1)} + 2L_g s\}$$

$$(2\text{-}25c) \qquad \zeta_f^1 = \{s\eta_f^{(0,1)} - L_g^2 s\} .$$

On the other hand, we have

$$(2\text{-}26) \qquad [g, ad_{sf}g] = \delta_0^1 g + \delta_1^1 ad_{sf}g$$
$$= \delta_0^1 g + s\delta_1^1 ad_f g - (\delta_1^1 L_g s) f .$$

Using condition (2) in theorem 2-8, we have

(2-27a) $s\eta_0^{(0,1)} = \delta_0^1$

(2-27b) $\{s\eta_1^{(0,1)} + 2L_g s\} = s\delta_1^1$

(2-27c) $\{s\eta_f^{(0,1)} - L_g^2 s\} = -\delta_1^1 L_g s$.

This yields eq.(2-23).■

4-3. Example

Consider the system

$$\frac{dx}{dt} = f(x) + g(x)\,u$$

where $x=(x_1, x_2, x_3)^T$, $f(x)=(x_2 e^{x_3}, x_3 e^{x_3}, 0)^T$ and $g(x)=(0,0,1)^T$. The Lie bracket $[f,g]=ad_f g$ is easily calculated as follows.

$$ad_f g = [f, g] = -\begin{bmatrix} x_2 e^{x_3} \\ (x_3+1)e^{x_3} \\ 0 \end{bmatrix}.$$

In order to know whether this system is feedback equivalent to a linear system or not, we will examine condition (2) in prosition 1-1. Since

$$[g, ad_f g] = -\begin{bmatrix} x_2 e^{x_3} \\ (x_3+2)e^{x_3} \\ 0 \end{bmatrix}$$

$$= f + 2\,ad_f g \notin \text{span}\{g, ad_f g\},$$

the set of vector fields $\{g, ad_f g\}$ is not involutive. Thus, this system is not feedback equivalent to a linear system. This system, however, will be found to be wide-sense feedback equivalent to a linear system. The partial differential equation (2-23) in corollary 2-9 for this system is expressed as

$$L_g^2 s - \frac{2}{s}(L_g s)^2 - 2L_g s - s$$
$$= \frac{\partial^2 s}{\partial x_3^2} - \frac{2}{s}\left[\frac{\partial s}{\partial x_3}\right]^2 - 2\frac{\partial s}{\partial x_3} - s = 0 .$$

Since $s(x) = e^{-x_3} > 0$ satisfies this differential equation, the system is wide-sense feedback equivalent to a linear system . In fact, using the time scale τ defined by $\frac{dt}{d\tau} = s(x) = e^{-x_3}$, the system can be rewritten as

$$\frac{d}{d\tau}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v$$
$$v = e^{x_3} u$$

and this system is linear in the time scale τ .

4-4. Concluding Remarks

In this chapter, we have investigated the input-state linearization problem in a transformed time scale. This problem has been formulated using the notion of wide-sense feedback equivalence. We have identified the class of nonlinear systems which can be linearized in the transformed time scale. We have also shown that there are nonlinear systems which can not be linearized in the actual time scale t, but it is possible to linearize in the transformed time scale. The time scale in which the system can be linearized is obtained as the solution to partial differential equations. Since the time scale transformation preserves the systems's stability, it is possible to use the linearized model in the transformed time scale to obtain a stabilizing controller.

V. CONTROLLER DESIGN

In this chapter, we will take a different perspective on the time scale transformation by applying it to controller design. We will propose the following controller design method. Firstly, we will introduce an appropriate time scale τ and express the system dynamics in the transformed time scale. We will then linearize the system in the the time scale τ. Finally, we will design a linear controller in the time scale τ to stabilize the system. The controller is linear in the time scale τ, but nonlinear in the actual time scale t. Our choice of the time scale transformation will enable us to obtain interesting properties in the controlled system which can not be achieved with conventional methods.

As we saw in the previous chapter, linearizability is not usually preserved by time scale transformations. In the case, however, of a system with a controllability index two, almost all time scale transformations will preserve linearizability. Since mechanical systems are usually expressed by second order differential equations, i.e. controllability index is two, the proposed controller design method is especially appropriate to design controllers for mechanical systems.

In section 5-1, we will apply this method to design a controller for a linear system so that we can avoid an exceedingly high amplitude of input. In section 5-2, we will propose a controller for a robot which can achieve good performance even in the neighborhood of a singular point. In section 5-3, we will describe a trajectory tracking controller for a vehicle. This controller will allow us to analytically evaluate the stability of our path tracking control.

5-1. Nonlinear Controller for Linear Systems

As we have seen in the previous chapter, a change in the time scale will not generally preserve the linearity of the system. In the case of n=2, however, almost all the time scaling functions s(x)>0 preserve the linearity as far as $\{ g, ad_{sf}g\}(x)$ span $T_x M$ in a neighborhood of the origin. Also s(x) is sufficient to be a C^1 function because { g } is always involutive. This is also the case when we consider multi-input systems with controllability index two. In this section we consider the linear system

(1-1a) $$\frac{d}{dt}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & I \\ A_1 & A_2 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} u$$

(1-1b) $$y = h(x, \dot{x})$$

where $x \in R^m$, $u \in R^m$ and $y \in R^p$. We will design the controller in a time scale τ, so that the controlled system will satisfy a desirable property, such as avoiding high amplitude input or an excessive rate of output.

5-1-1. Controller Design

Consider the following C^1 time scaling function

(1-2) $$\frac{dt}{d\tau} = s(x, \dot{x}, y, y_r, x_c) > 0$$

where $y_r \in R^p, x_c \in R^{n'}$ are a reference value and the state of the compensator, respectively, which satisfy

(1-3a) $$\frac{dy_r}{d\tau'} = \Omega(\tau')$$

(1-3b) $$\frac{d\tau'}{d\tau} = s'(x, \dot{x}, y, y_r, x_c) > 0$$

(1-4) $$\frac{dx_c}{d\tau} = \Xi(x, \dot{x}, y, y_r, x_c).$$

It must be noted that these values can easily be calculated in the actual time t, in fact

(1-5) $$\frac{dy_r}{dt} = \frac{dy_r}{d\tau'}\frac{d\tau'}{d\tau}\frac{d\tau}{dt} = s'^{-1}\ s^{-1}\Omega\ (\tau')$$

(1-6) $$\frac{dx_c}{dt} = \frac{dx_c}{d\tau}\frac{d\tau}{dt} = s^{-1}\Xi\ (x,\ \dot{x},\ y,\ y_r, x_c).$$

The system (1-1) can be expressed as follows in the time scale τ defined by eq.(1-2).

(1-7) $$\frac{dx}{d\tau} = \frac{dx}{dt}\ \frac{dt}{d\tau} = s\,\dot{x}$$

(1-8) $$\frac{d^2x}{d\tau^2} = s^2\ \ddot{x} + \frac{ds}{d\tau}\dot{x}$$

$$= s^2\ \ddot{x} + \dot{x}\ \{\frac{\partial s}{\partial x}\ \frac{dx}{d\tau} + \frac{\partial s}{\partial \dot{x}}\ \frac{d\dot{x}}{d\tau} + \frac{\partial s}{\partial y}\ \frac{dy}{d\tau}$$
$$+ \frac{\partial s}{\partial y_r}\ \frac{dy_r}{d\tau} + \frac{\partial s}{\partial x_c}\ \frac{dx_c}{d\tau}\ \}$$

$$= s^2\ \ddot{x} + \dot{x}\ [\frac{\partial s}{\partial x} s\,\dot{x} + \frac{\partial s}{\partial \dot{x}} s\,\ddot{x}$$
$$+ \frac{\partial s}{\partial y}\ \{\frac{\partial h}{\partial x} s\,\dot{x} + \frac{\partial h}{\partial \dot{x}} s\,\ddot{x}\}$$
$$+ \frac{\partial s}{\partial y_r}\ s'\ \Omega + \frac{\partial s}{\partial x_c}\ \Xi\ \ ]$$

$$= [s^2\ I + s\,\dot{x}\ \{\frac{\partial s}{\partial \dot{x}} + \frac{\partial s}{\partial y}\ \frac{\partial h}{\partial \dot{x}}\}\ ]\ \ddot{x}$$
$$+ \dot{x}\ [\ \{\frac{\partial s}{\partial x} + \frac{\partial s}{\partial y}\ \frac{\partial h}{\partial x}\}\ s\,\dot{x}$$
$$+ \frac{\partial s}{\partial y_r}\ s'\ \Omega + \frac{\partial s}{\partial x_c}\ \Xi]$$

$$= [s^2\ I + s\,\dot{x}\ \{\frac{\partial s}{\partial \dot{x}} + \frac{\partial s}{\partial y}\ \frac{\partial h}{\partial \dot{x}}\}]\ (A_1 x + A_2 \dot{x} + u)$$
$$+ \dot{x}\ [\ \{\frac{\partial s}{\partial x} + \frac{\partial s}{\partial y}\ \frac{\partial h}{\partial x}\}\ s\,\dot{x}$$
$$+ \frac{\partial s}{\partial y_r}\ s'\ \Omega + \frac{\partial s}{\partial x_c}\ \Xi]\ .$$

Using the input

(1-9) $$u = -A_1 x - A_2 \dot{x} + \left[s^2 I + s\dot{x} \left\{ \frac{\partial s}{\partial \dot{x}} + \frac{\partial s}{\partial y} \frac{\partial h}{\partial \dot{x}} \right\} \right]^{-1} \times \left[v - \dot{x} \left\{ \left(\frac{\partial s}{\partial x} + \frac{\partial s}{\partial y} \frac{\partial h}{\partial x} \right) s\dot{x} + \frac{\partial s}{\partial y_r} s' \Omega + \frac{\partial s}{\partial x_c} \Xi \right\} \right]$$

the system can be expressed as

(1-10) $$\frac{d}{d\tau} \begin{bmatrix} x \\ \frac{dx}{d\tau} \end{bmatrix} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \frac{dx}{d\tau} \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} v$$

if $[s^2 I + s\dot{x} \{\frac{\partial s}{\partial \dot{x}} + \frac{\partial s}{\partial y} \frac{\partial h}{\partial \dot{x}}\}]$ is nonsingular. So the controller can easily be designed using this state equation in the time scale τ.

[remark]

If the output y is a linear combination of x and $\dot{x}$, and the controller is linear in the time scale τ, then v is proportional to $\| y - y_r \|$. This means that the last term of eq.(1-9) is almost proportional to $s^{-2} \| y - y_r \|$ in the case where $s \gg 1$. So, if we choose $s = L \{ \| y - y_r \| \}^{1/2}$ $(L > 0)$ when $\| y - y_r \| > M_1$ for some positive number M_1, then $s^{-2} \| y - y_r \| = (1/L)^2$ for $\| y - y_r \| > M_1$, i.e. the controller will avoid an excessively high input amplitude.

[remark]

If $s' \cdot s \gg 1$ when $\| y - y_r \| > M_2$, then τ' will vary slowly compared to the actual time t when $\| y - y_r \| > M_2$. This means if $\| y - y_r \|$ becomes too large, then y_r will change more slowly. Consider the case where $\frac{dy_r}{d\tau'}$ = constant (this is a ramp type reference), this means that the y_r will increase more slowly when the error becomes too large.

[remark]

If $s \gg 1$ when $\| \dot{x} \| > M_3$ then $\| \dot{x} \|$ cannot exceedingly go beyond M_3 because $\frac{dx}{d\tau} = s\dot{x}$. This means that it is possible to avoid an excessively high velocity $\|\dot{x}\|$.

5-1-2. Simulation

We will consider the system

$$\frac{d}{dt}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = x$$

and design a controller with one integrator in the time scale τ so that the poles (in the time scale τ) of the total system are at -1, -1, and -2.

(Fig.5-1-1)

In this case, the time scaling function is

$$s = 1 + \rho\,(y - y_r)$$

$$\rho\,(y - y_r) = \begin{cases} 0 & (\|y - y_r\| < 1) \\ 0.09\,(\|y - y_r\| - 1)^2 & (1 \leqq \|y - y_r\| < 6.4) \\ 0.5\,(\|y - y_r\|)^{1/2} & (\text{otherwise}). \end{cases}$$

Fig.5-1-1 shows the step response for y_r= 1, 5 and 10.
This controller obviously avoids an excessively high input amplitude.

(Fig.5-1-2)

In the next, the time scaling funciton is

$$s = 1 + \rho\,(y - y_r) + \pi\,(\dot{x})$$

$$\pi\,(\dot{x}) = \begin{cases} 0 & (\|\dot{x}\| < 1) \\ (\|\dot{x}\| - 1)^2 & (\text{otherwise}). \end{cases}$$

Fig.5-1-2 compares the step responses (y_r=10) in the case of

$s=1+\rho+\pi$ with $s=1+\rho$. The time scaling function $s=1+\rho+\pi$ effectively avoids an excessively high velocity .

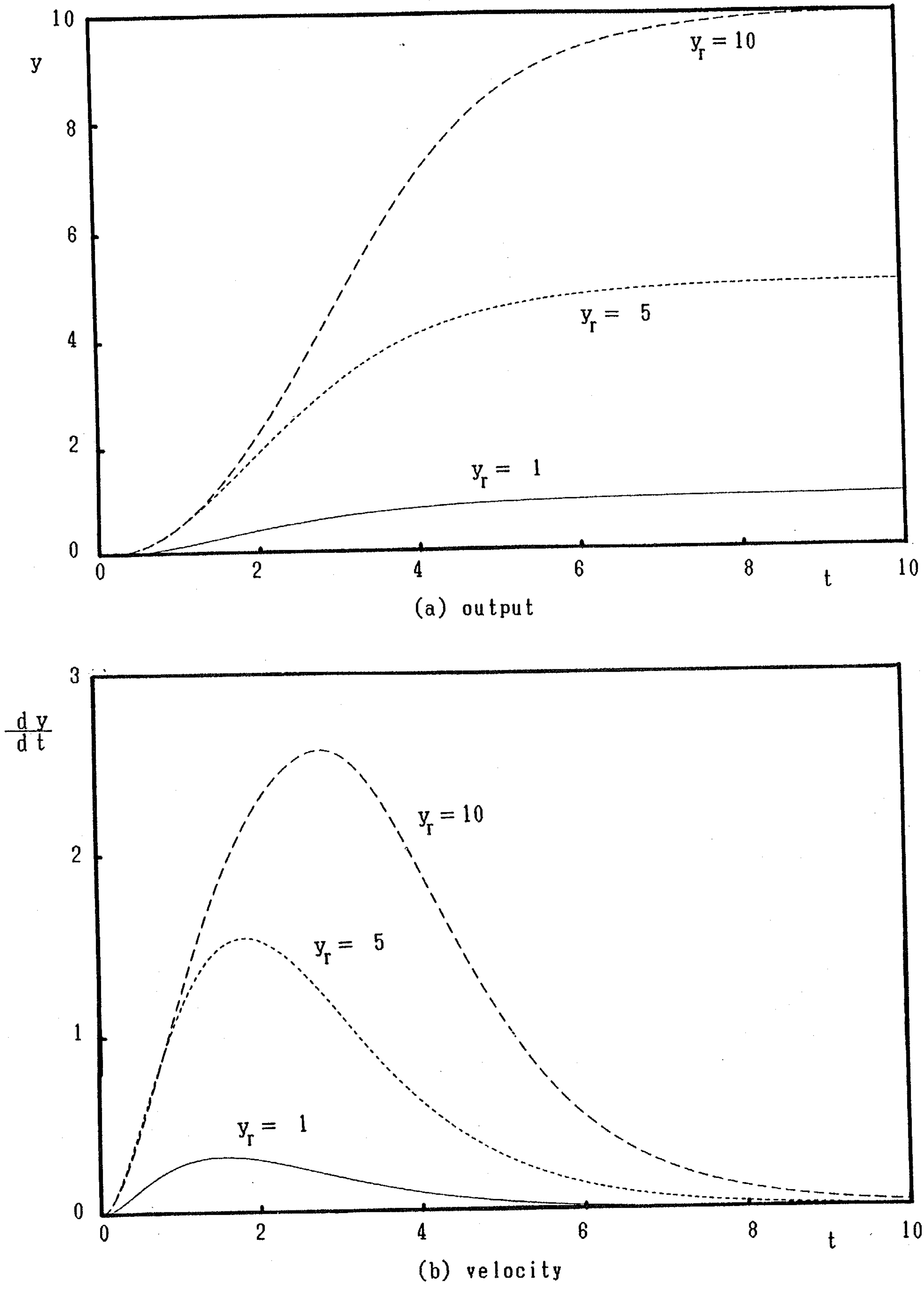
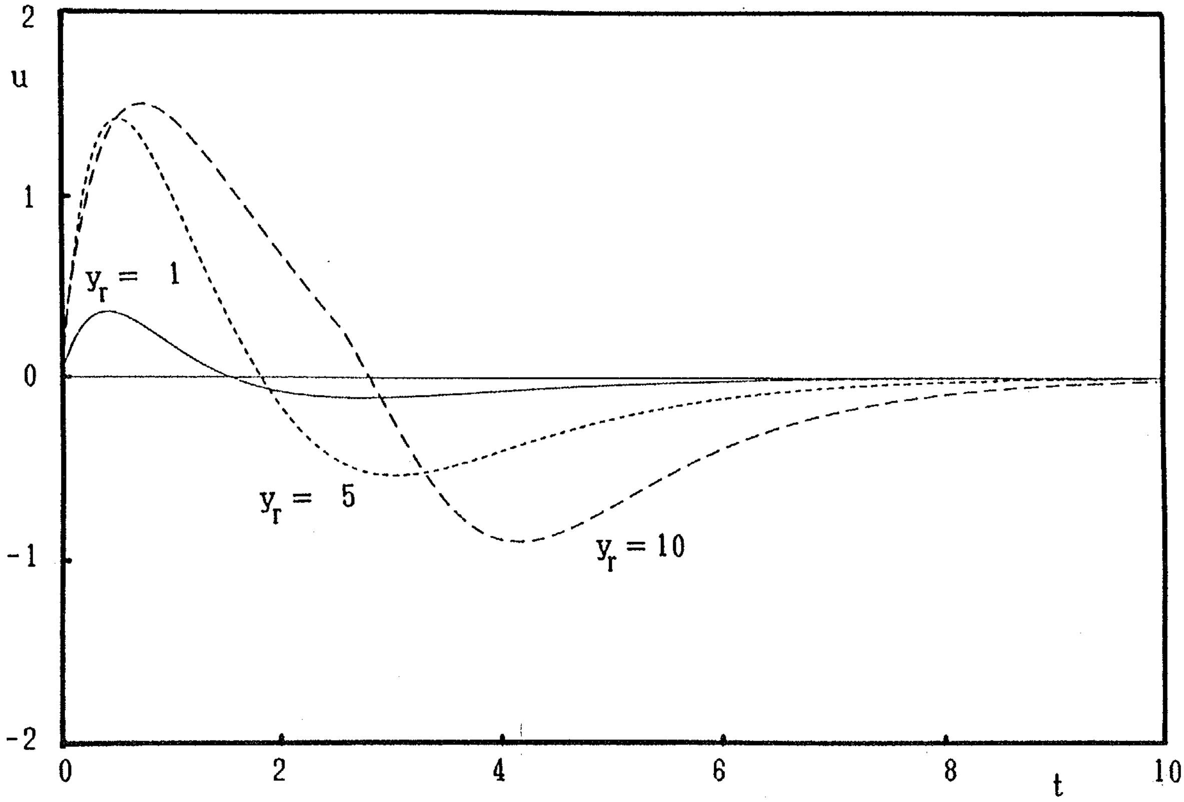
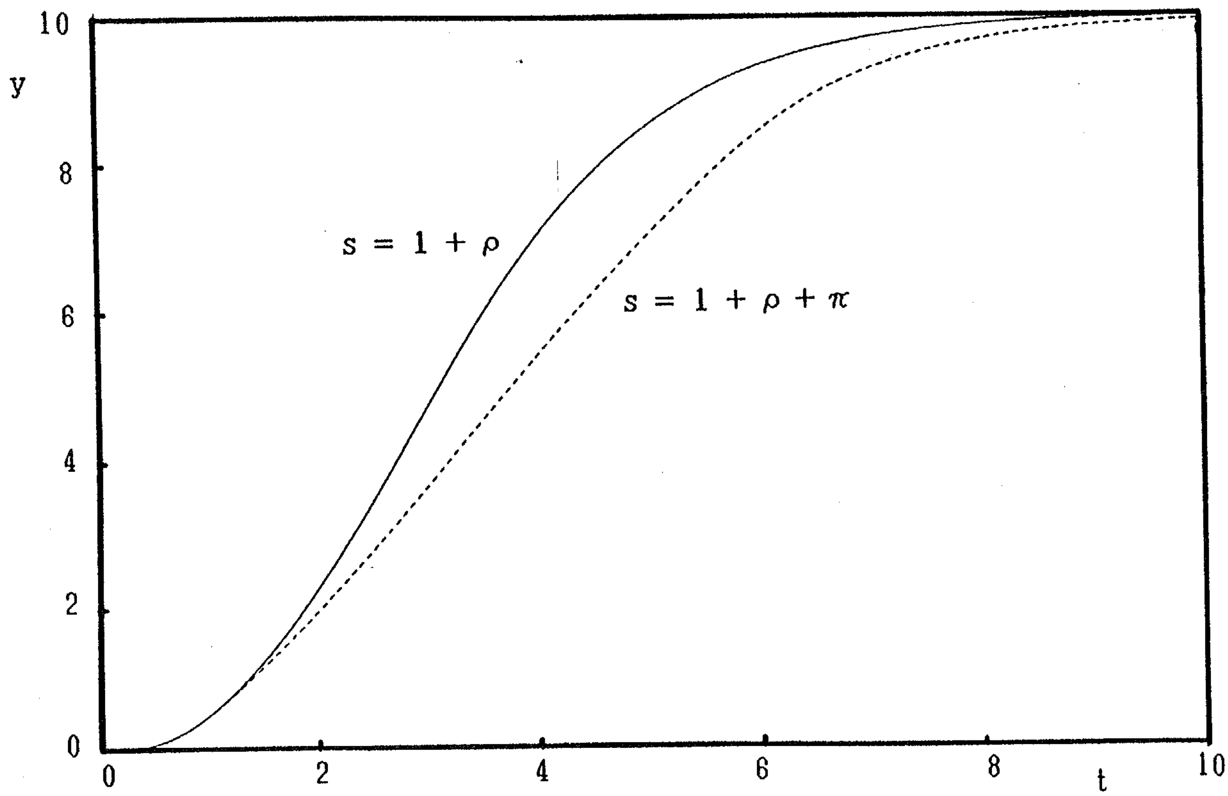


Fig.5-1-1. System response ($s = 1 + \rho$)

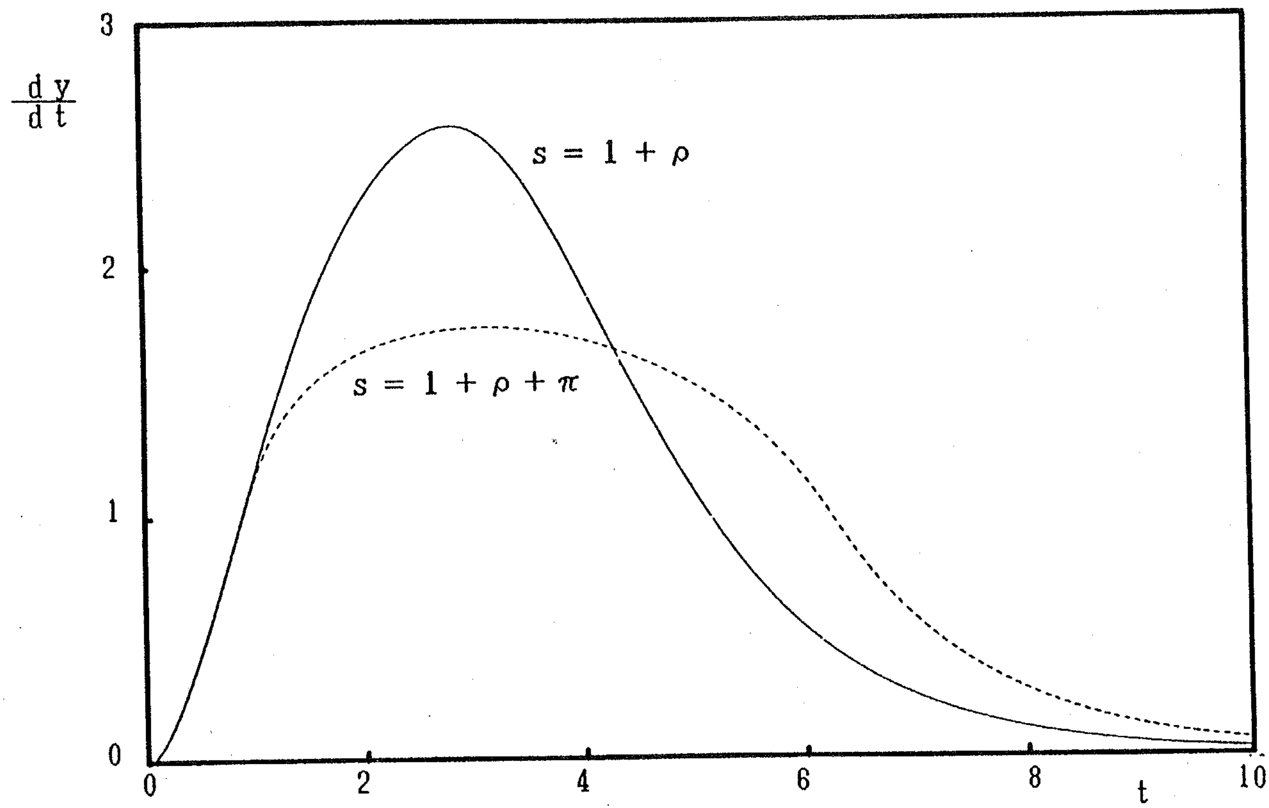


(c) input

Fig.5-1-1. System response ($s = 1 + \rho$)

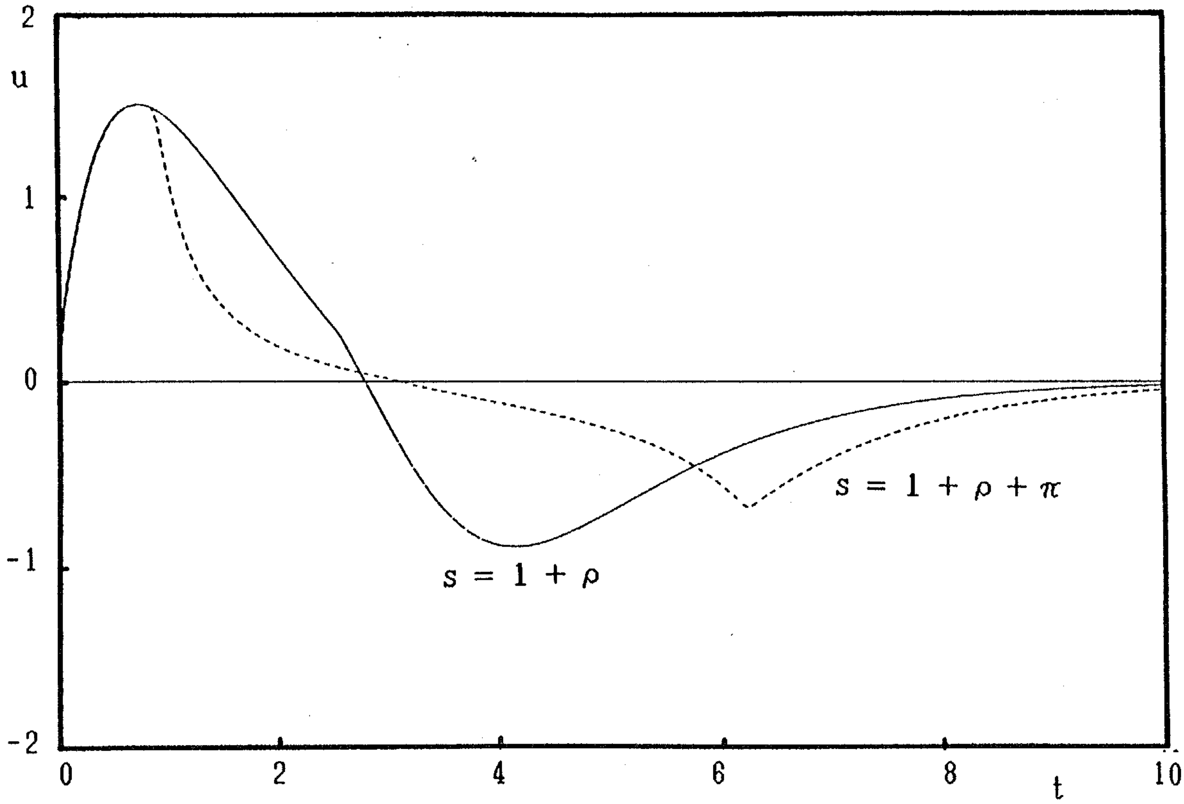


(a) output



(b) velocity

Fig.5-1-2. System response (comparing $s = 1 + \rho$ and $s = 1 + \rho + \pi$)



(c) input

Fig.5-1-2. System response (comparing $s = 1 + \rho$ and $s = 1 + \rho + \pi$)

5-2. Robot Control in the Neighborhood of Singular Points

The equations of motion of robots are usually expressed in joint coordinates and the robots are controlled in these coordinates. It is, however, preferable to control them in Cartesian coordinates when desirable motions can be best described in these coordinates([63] ~ [65]). For instance, if it is necessary for the endeffector to track a straight line using sensory feedback, it is better to choose Cartesian coordinates; one of the coordinates should be chosen to coincide with this line so that the deviation from the line and the motion in the line can be decoupled[64]. In the case of force control, the desired motion with respect to an external sensored force is usually described in Cartesian coordinates. So it is difficult to control robots without considering these coordinates[65].

One of the most efficient controller design procedures utilizes Cartesian coordinates in the following way. With this method, we first define the transformation of the coordinates and express the equations of motion in Cartesian coordinates. We then linearize the system in these coordinates using nonlinear feedback and design a linear compensator for the resultant linearized system[63]. This controller design can be implemented by computing the inverse kinematics[66] and the resolved acceleration control[67]. Although this controller has many good features, it may also be plagued by singular points; in the neighborhood of these points, proper control can not be achieved.

A singular point is a point at which there exists at least one direction in Cartesian coordinates in which the endeffector is immobilized. Even in the neighborhood of such a point , the end-effector's movement in that particular direction (or directions) is usually impaired. So the manipulative ability is very poor; it is difficult to control robots in that area. This problem has been the topic of several papers treating manipulative ability which proposed various means of measuring manipulative ability ([68]～[70] and their references). In these papers, the measure of manipulative ability has been used in the planning of the endeffector's trajectories so as to avoid the area in which manipulation is impaired , and to design the robot's structure so that the area of impaired manipulation will be minimized. Ours is the first control strategy which attempts to control a robot by using a measure of manipulative ability. With this control, we can achieve good performance in the neighborhood of singular points without reducing the performance outside the neighborhoods of these point.

If a controller is to achieve good performance in the neighborhood of singular points, it should be designed so that the resultant controlled system has slow poles only in the neighborhood of these points. Conventional methods are able to achieve slow poles in the neighborhood of singular points; however, they also result in slow poles outside the neighborhoods of these points because they do not attempt to consider manipulative ability. So that the performance of the endeffector is improved in the neighborhood of the singular points, but at the same time, it is impaired outside the neighborhood of these points. On the other hand, if we choose to construct the controller with conventional methods so that good performance is maintained outside the neighborhoods of the singular points, the performance will be poor in the neighborhood of the singular points. The resultant system would have fast poles throughout, necessitating an excessively high amplitude input in the neighborhood of the singular points. Our method achieves slow poles only in the neighborhood of singular points leaving the fast poles outside these neighborhoods unchanged.

To achieve this, we propose to design a controller in Cartesian coordinates by transforming the time scale. A transformation of the time scale was once used by Hollerbach[5], but his research was limited to planning trajectories and the transformation he used was not convenient for the design of a real time controller. It will be shown that the proposed controller designed with appropriate time scales will give proper control with respect to manipulative ability; the resultant controlled system will have slow poles in the area of poor manipulative ability and fast poles in that of good manipulative ability. The robot will be properly controlled even in the neighborhood of singular points without a reduction in control outside the neighborhoods of the singular points.

5-2-1. Measure of Manipulative Ability

Several effective measures have been proposed([67] ~ [70] and their references) to describe the manipulative ability with regard to positioning and orienting the endeffector. In this section, we will propose another measure which will be used to design our controller in the next section.

The equations of articulated robot motion are generally expressed in joint coordinates as

$$M(\theta)\ddot{\theta} + h(\theta, \dot{\theta}) + g(\theta) = u \tag{2-1}$$

where $\theta = (\theta_1, \theta_2, \cdots, \theta_n)^T \in R^n$ represent the joint angles; $M(\theta)$ is the moment of inertia which is always nonsingular; h and g are inertial force and gravity respectively; $u = (u_1, u_2, \cdots, u_n)^T \in R^n$ represent input torques and are assumed to be restricted

(2-2) $$-u_{i_{max}} < u_i < u_{i_{max}}.$$

Assume that the relationship between joint θ and Cartesian coordinates $x = (x_1, x_2, \cdots, x_n)^T \in R^n$ is given by

(2-3) $$x = f(\theta),$$

then $\dot{x}$ and $\ddot{x}$ are expressed as

(2-4) $$\dot{x} = \frac{\partial f}{\partial x} \dot{\theta} \triangleq J(\theta)\dot{\theta}$$

(2-5) $$\ddot{x} = J\ddot{\theta} + \dot{J}\dot{\theta}$$
$$= JM^{-1}\{u - h(\theta, \dot{\theta}) - g(\theta)\} + \dot{J}\dot{\theta}.$$

In order to normalize the input, we will introduce the diagonal matrix $N(\theta, \dot{\theta})$ whose i-i element is

(2-6) $$1/\{u_{i_{max}} - |h_i(\theta, \dot{\theta}) + g_i(\theta)|\}$$

where h_i and g_i are the i-th elements of h and g, respectively. With the matrix N and the normalized input $v = (v_1, v_2, \cdots, v_n)^T \in R^n$, we have

(2-7) $$\ddot{x} = JM^{-1}Nv + \dot{J}\dot{\theta}$$

(2-8) $$\| v_i \| < 1.$$

This implies that the maximum acceleration is characterized by an ellipsoid $JM^{-1}Nv, (\|v\| \leq 1)$ as long as $\dot{J}\dot{\theta}$ is relatively small. This ellipsoid is called the dynamic manipulability ellipsoid[70]. Kosuge and Furuta have defined a measure of manipulative ability called "controllability" which is defined by the ratio of the length of the ellipsoid's minor axis to that of the major[69]. Yoshikawa has defined "dynamic manipulability" which is defined by the volume of the ellipsoid[70]. Here we will define another measure of manipulative ability.

[definition]

Assume that e is a unit vector in Cartesian coordinates. Dynamic manipulability in the direction of e at $(\theta, \dot{\theta})$ is defined as

(2-9) $$dm(e;\ \theta, \dot{\theta}) = 1 / \| N^{-1}(\theta, \dot{\theta})\, M(\theta)\, J^{-1}(\theta)\, e \| .$$

The dynamic manipulability $dm(e;\ \theta, \dot{\theta})$ can be interpreted as the length of the axis of the dynamic manipulability ellipsoid which is parallel to e (fig.5-2-1).

For any unit vector e in Cartesian coordinates, there is a unique vector $v(e;\ \theta, \dot{\theta})$ in joint coordinates which satisfies

(2-10) $$J M^{-1} N v(e;\theta, \dot{\theta}) = dm(e;\ \theta, \dot{\theta}) \cdot e$$

and can be specified except at singular points. Premultiplying both sides by $N^{-1} M J^{-1}$ yields

(2-11) $$v(e;\theta, \dot{\theta}) = dm(e;\ \theta, \dot{\theta}) \cdot N^{-1} M J^{-1} e$$
$$= \frac{1}{\| N^{-1} M J^{-1} e \|} \cdot N^{-1} M J^{-1} e .$$

The above equation shows that $\| v(e;\theta, \dot{\theta}) \| = 1$. It implies that $dm(e;\theta, \dot{\theta})$ expresses the length of the axis of the dynamic manipulability ellipsoid parallel to e.

5-2-2. Controller Design

In this section, we will present a controller design method using a time scale transformation. With this design method, we first determine the new time scale τ_i for each x_i axis, then design an ordinary compensator for each axis in new time scale of each axis. The proposed controller controls a robot properly with respect to manipulative ability when the time scales are properly defined.

Define the new time scale τ_i with the following differential equation.

$$(2\text{-}12) \qquad \frac{dt}{d\tau_i} = s_i(\theta) > 0 .$$

Suitable choices of the time scaling function s_i will be discussed later in this section, but here we will assume it to be an arbitrary positive differentiable fuction with respect to θ.

When we differentiate x_i with respect to τ_i, we obtain the following equations.

$$(2\text{-}13) \qquad \frac{dx_i}{d\tau_i} = \frac{dx_i}{dt}\frac{dt}{d\tau_i} = \dot{x}_i s_i$$

$$(2\text{-}14) \qquad \frac{d^2x_i}{d\tau_i^2} = \ddot{x}_i s_i^2 + \dot{x}_i \frac{\partial s_i}{\partial \theta}\frac{d\theta}{d\tau_i}$$

$$= \ddot{x}_i s_i^2 + \dot{x}_i \frac{\partial s_i}{\partial \theta}\dot{\theta} s_i .$$

Since $\ddot{x}_i$ is directly manipulated by the input u, we can assume

$$(2\text{-}15) \qquad \ddot{x}_i = \frac{1}{s_i^2}\left(w_i - \dot{x}_i \frac{\partial s_i}{\partial \theta}\dot{\theta} s_i\right)$$

which yields

$$(2\text{-}16) \qquad \frac{d^2x_i}{d\tau_i^2} = w_i .$$

This differential equation is linear and expresses the motion of the x_i axis in the new time scale τ_i. We can design a compensator for the dynamical system with this differential equation in the new time scale . For example, a servo compensator with one integrator could easily be constructed as follows. An integrator in the new time scale τ_i is defined as

$$(2\text{-}17) \qquad \frac{dy_i}{d\tau_i} = x_i - x_{ri}$$

where x_{ri} and y_i are the i-th elements of the reference point in Cartesian coordinates and the state of the i-th integrator, respectively. This can be implemented in the actual time scale by the following equation.

$$\text{(2-18)} \qquad \frac{dy_i}{dt} = \frac{dy_i}{d\tau_i}\frac{d\tau_i}{dt} \doteq \frac{1}{s_i}(x_i - x_{ri}).$$

Linear state feedback in the new time scale is calculated as

$$\text{(2-19)} \qquad w_i = f_{i1}y_i + f_{i2}x_i + f_{i3}\frac{dx_i}{d\tau_i}$$
$$= f_{i1}y_i + f_{i2}x_i + f_{i3}s_i\dot{x}_i .$$

The controlled system including the compensator is expressed as

$$\text{(2-20)} \qquad \frac{d}{d\tau_i}\begin{bmatrix} y_i \\ x_i \\ \frac{dx_i}{d\tau_i} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ f_{i1} & f_{i2} & f_{i3} \end{bmatrix}\begin{bmatrix} y_i \\ x_i \\ \frac{dx_i}{d\tau_i} \end{bmatrix} - \begin{bmatrix} x_{ri} \\ 0 \\ 0 \end{bmatrix}.$$

This means that arbitrary poles are assignable in the new time scale.

Any kind of compensator can be designed in the new time scale, but we will not address the topic of compensator types here. We are interested in the features of the controller designed in the new time scale. Our method can be described as shifting the poles in the actual time scale with respect to a time scaling function.

Consider the case of a servo compensator with one integrator again. If the time scaling function is constant, $s_i = a$, the controlled system (2-20) can be expressed in the actual time scale as

$$\text{(2-21)} \qquad \frac{d}{dt}\begin{bmatrix} a\,y_i \\ x_i \\ \dot{x}_i \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{f_{i1}}{a^3} & \frac{f_{i2}}{a^2} & \frac{f_{i3}}{a} \end{bmatrix}\begin{bmatrix} a\,y_i \\ x_i \\ \dot{x}_i \end{bmatrix} - \begin{bmatrix} x_{ri} \\ 0 \\ 0 \end{bmatrix}.$$

This state equation implies that the poles in the actual time scale are 1/a of those in the new one. The same thing occurs with any compensator which is linear in the new time scale. If the time scaling function $s_i(\theta) = 1$, then the poles in the actual time scale are identical to those in the new one. On the other hand, if $s_i(\theta) >> 1$, then the poles in the actual time scale are much slower than those in the new one. So, if we are careful to choose a time scaling function which is large in the poor manipulative ability area and relatively small in other areas, then we can achieve a controlled system having slow poles near singular points and fast poles at other points. This control is preferable because it avoids high amplitude input in the neighborhood of the singular points without reducing the performance outside the neighborhoods of these points. The following discussion describes how to decide the exact value of the time scaling function.

The equations of motion in the new time scales (2-14) can be summarized in vector form,

(2-22)
$$\begin{bmatrix} \frac{d^2 x_1}{d\tau_1^2} \\ \frac{d^2 x_2}{d\tau_2^2} \\ \vdots \\ \frac{d^2 x_n}{d\tau_n^2} \end{bmatrix} = \begin{bmatrix} s_1^2 & & & \\ & s_2^2 & & \\ & & \ddots & \\ & & & s_n^2 \end{bmatrix} \ddot{x} + \begin{bmatrix} s_1 \dot{x}_1 & & & \\ & s_2 \dot{x}_2 & & \\ & & \ddots & \\ & & & s_n \dot{x}_n \end{bmatrix} \frac{\partial s}{\partial \theta} \dot{\theta}$$

$$= w$$

where $s = (s_1, s_2, \cdots, s_n)^T$. The following equation is obtained by substituting eq.(2-5) into eq.(2-22).

(2-23)
$$\begin{bmatrix} \frac{d^2 x_1}{d\tau_1^2} \\ \frac{d^2 x_2}{d\tau_2^2} \\ \vdots \\ \frac{d^2 x_n}{d\tau_n^2} \end{bmatrix} = \begin{bmatrix} s_1^2 & & & \\ & s_2^2 & & \\ & & \ddots & \\ & & & s_n^2 \end{bmatrix} J M^{-1} \{u - h - g\}$$

$$+ \{ \begin{bmatrix} s_1 \dot{x}_1 & & \\ & s_2 \dot{x}_2 & \\ & & s_n \dot{x}_n \end{bmatrix} \frac{\partial s}{\partial \theta} \dot{\theta} + \begin{bmatrix} s_1^2 & & \\ & s_2^2 & \\ & & s_n^2 \end{bmatrix} \dot{J} \dot{\theta} \}.$$

This implies that the dynamic manipulability ellipsoid in the new time scales is expressed by

$$\begin{bmatrix} s_1^2 & & \\ & s_2^2 & \\ & & \ddots \\ & & & s_n^2 \end{bmatrix} J M^{-1} N v \qquad (\| v \| \leq 1) . \tag{2-24}$$

Consequently, dynamic manipulability in the direction of e in the time scale $\tau = (\tau_1, \tau_2, \cdots, \tau_n)$ is defined as

$$dm_\tau(e;\ \theta,\ \dot{\theta}) = 1 / \left\| N^{-1} M J^{-1} \begin{bmatrix} s_1^2 & & \\ & s_2^2 & \\ & & \ddots \\ & & & s_n^2 \end{bmatrix}^{-1} e \right\| . \tag{2-25}$$

In particular, the dynamic manipulability in the direction of the x_i axis in the new time scale τ; $dm_\tau(x_i;\ \theta,\ \dot{\theta})$ is the scalar multiple of that in the actual time scale $dm(x_i;\ \theta,\ \dot{\theta})$:

$$dm_\tau(x_i;\ \theta,\ \dot{\theta}) = s_i^2\ dm(x_i;\ \theta,\ \dot{\theta}) . \tag{2-26}$$

Define the time scaling function

$$s_i \simeq c / \sqrt{dm(x_i;\ \theta,\ \bar{\dot{\theta}})} \tag{2-27}$$

where $\bar{\dot{\theta}}$ is the nominal value of $\dot{\theta}$. Then $dm_\tau(x_i;\ \theta,\ \dot{\theta})$ becomes nearly equal to c for almost all $(\theta,\ \dot{\theta})$ under the assumption that $\dot{\theta}$ effects $dm_\tau(x_i;\ \theta,\ \dot{\theta})$ sufficiently less than θ does. Dynamic manipulability in the direction of the x_i axis in the new time scale is almost independent of position except at singular points where s_i diverges. It must be emphasized that the time scaling functions need not be exactly equal to $c / \sqrt{dm(x_i;\ \theta,\ \bar{\dot{\theta}})}$. Even if the time scaling functions do not have the exact form shown above, a controller designed using these time scales may still provide proper control with respect to manipulative ability.

5-2-3. Simulation

Fig.5-2-2 show step responses using the proposed controller. Fig.5-2-2-a shows the response outside the neighborhoods of singular points where the manipulating ability is sufficiently large and fig.5-2-2-b shows the response near a singular point. Fig.5-2-3 show the step responses using a conventional controller designed in Cartesian coordinates.

The step response using the proposed controller is almost the same as that using the conventional controller in the area with good manipulative ability (fig.5-2-2-a and 5-2-3-a); however, the proposed controller is superior to the conventional one in the area with poor manipulative ability (fig.5-2-2-b and 5-2-3-b). The step response using the proposed controller is much slower than that using the conventional one. This shows that the proposed controller is able to control the robot properly with respect to manipulative ability.

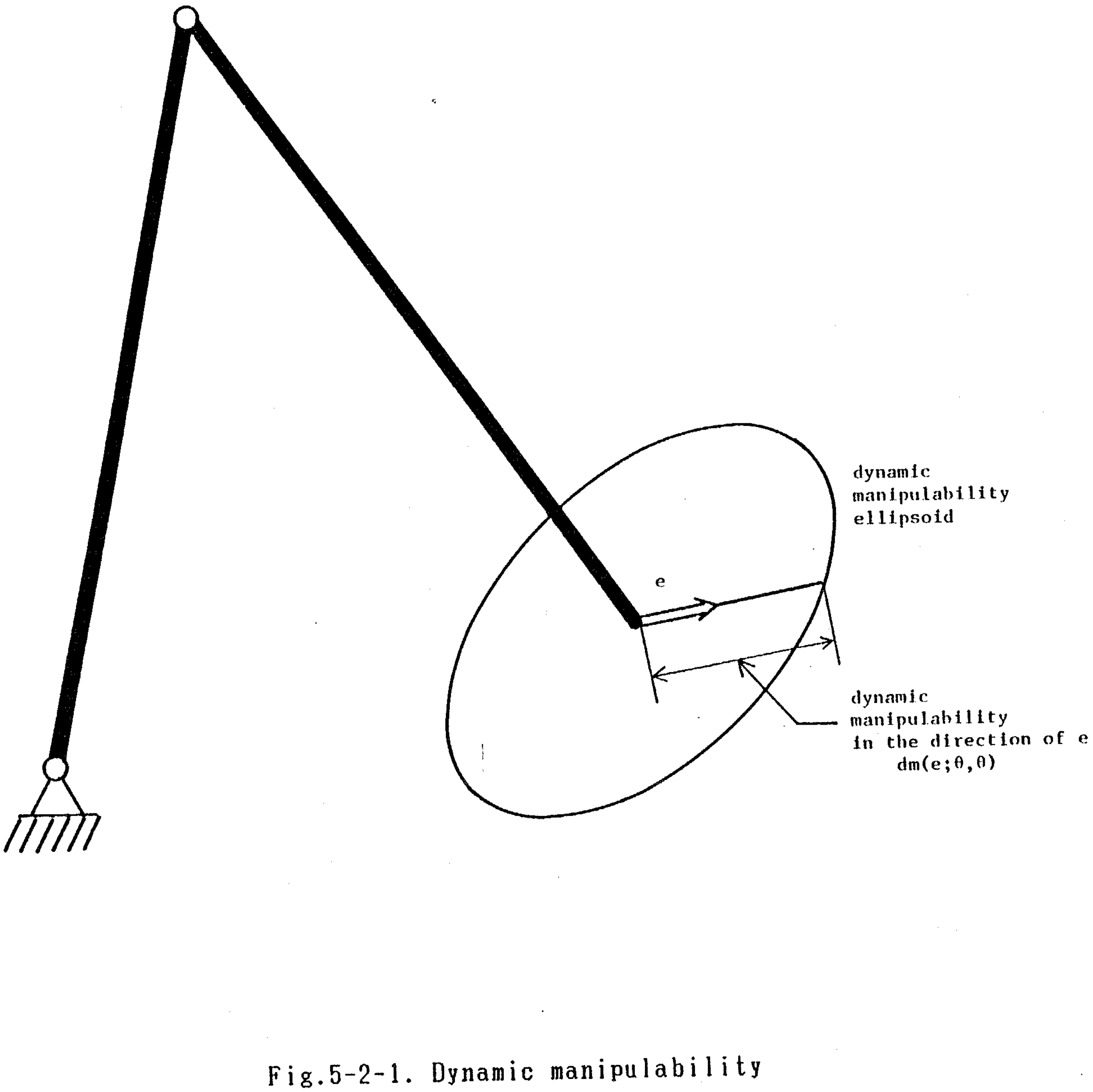
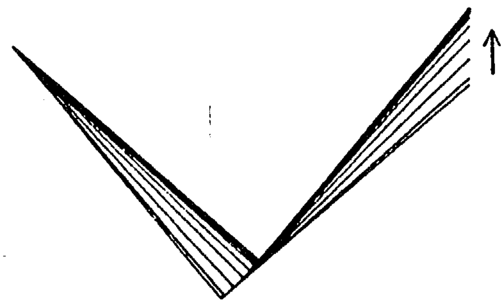
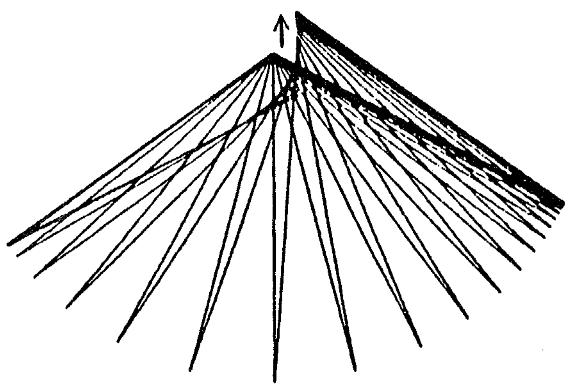


Fig.5-2-1. Dynamic manipulability

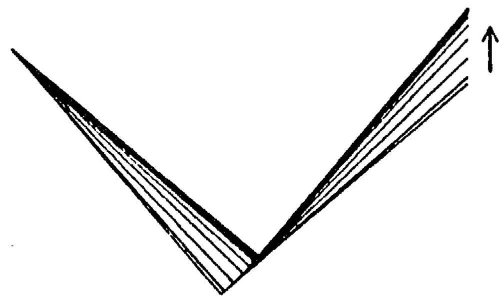


(a) away form the singular point

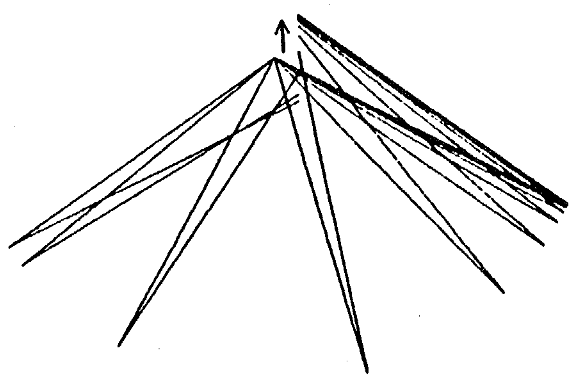


(b) near the singular point

Fig.5-2-2. Step response of the proposed system



(a) away form the singular point



(b) near the singular point

Fig.5-2-3. Step response of the conventional system

5-3. Path Tracking Control of Mobile Robot

Mobile robots are used to perform tasks and collect information with minimal operator assistance. They are especially useful under dangerous or inaccessible conditions. Because of this, there have been many studies on mobile robots; for example on the path planning problem[71][72] and on methods to measure the position and the direction of the robot[73][74]. In the past, however, little work has been done on path tracking control. Thus, only intuitive methods have been used to track the path. Here, we will propose a design method for path tracking control which will analytically ensure stability.

In this section, we will consider the tricycle shown in fig.5-3-1. We will define the following notations.

- P : middle point of the rear wheels
- O : center of rotation
- Q : steering axis
- L : wheel-base
- θ : angle deviation of the mobile robot
- α : steering angle
- x,y : Cartesian coordinates indicating the position of the mobile robot
- z : distance along the real path.

The object of this section is to design a controller so that the robot will track the x-axis.

Komoriya et al.[75] proposed a method to track a path using only feedforward control. In this method, the steering angles are decided a priori and stored in the memory as a time series. Then the front wheel is steered according to that time series in the actual control. Since this control does not use feedback, accurate path tracking control is difficult to achieve. They also proposed another method[76]. This strategy is shown in fig.5-3-2. Consider the point A which is L' ahead of the robot (fig.5-3-2). The steering angle α is chosen propotional to the deviation of the point A from the path, i.e.

$$\alpha = -K e_A$$
$$= -K \{y + (L + L') \sin \theta\}.$$

The following method was proposed by Tsumura et al.[77]. The front wheel is steered to point to the point B which is L'' ahead along the desired path(fig.5-3-3). In this case, the steering angle α is chosen to be

$$\alpha = -\theta - \tan^{-1}\{\frac{L'' - L \cos\theta}{y - L \sin\theta}\}.$$

These methods are intuitive; the stability of these controls is not evaluated analytically, so a constant tracking error sometimes occurs.

In this section, we will propose an alternative method to design a path tracking controller. With this method, the stability of the resultant controlled system can readily be evaluated analytically and we can introduce an integrator in order to reduce the constant tracking error. The controller is designed as follows: firstly, define a new time scale which is chosen to be identical to the distance along the desired path, i.e. x. Then, the dynamic model in this new time scale is linearized with an appropriate state transformation. Finally, we design a linear controller (servo-controller if necessary) for the linearized system. We will also show how to design a controller that depends on the velocity of the robot using a further transformation of the time scale.

If we can assume that there are no side slips, then we can express the dynamics of the mobile robot as

(3-1a) $$\frac{d\theta}{dt} = \dot{z}\frac{1}{L}\tan\alpha$$

(3-1b) $$\frac{dy}{dt} = \dot{z}\sin\theta .$$

We will introduce a new time scale x which is identical to the position x . The time scale transformation is defined as follows,

(3-2) $$\frac{dt}{dx} = \frac{dt}{dz}\ \frac{dz}{dx} = (\dot{z})^{-1}\frac{1}{\cos\theta} .$$

The dynamic model of the robot can be expressed in this time scale as

(3-3a) $$\frac{d\theta}{dx} = \frac{1}{L\ \cos\theta}\tan\alpha$$

(3-3b) $$\frac{dy}{dx} = \tan\theta .$$

The linearized system is shown as follows.

(3-4) $$\frac{d}{dx}\begin{bmatrix} y \\ \tan\theta \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y \\ \tan\theta \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\frac{\tan\alpha}{L\cos^3\theta} .$$

In order to simplify the notation, we will define the state variable ϕ and the input v as follows.

(3-5) $$\phi \triangleq \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \triangleq \begin{bmatrix} y \\ \tan\theta \end{bmatrix}$$

(3-6) $$v \triangleq \frac{\tan\alpha}{L\cos^3\theta} .$$

Then the system(3-4) can be expressed by the following simple linear state equation.

(3-7a) $$\frac{d\phi}{dx} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\phi + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v$$

(3-7b) $$y = [\ 1 \quad 0\ ]\ \phi .$$

We can readily design any type of linear controller for this linearized system; for example integrators can be introduced to achieve robust control. This strategy can allow us to analytically evaluate the stability of the resultant controlled system because it is expressed by a linear state equation. For example, we can define poles for path tracking control. Since the controller design method for the system(3-7) is a straightforward application of linear system theory, it will not be discussed here.

Next, we will describe a method to design a velocity dependent controller. Since centrifugual force increases as the mobile robot's velocity increases, we can use only a small steering angle while the robot is moving fast. Thus, we need to design a velocity dependent controller such that the steering angle will become smaller as the velocity of the robot increases. The controller can be designed using a further time scale transformation. We define another time scale ξ using the function s as follows.

(3-8) $$\frac{dx}{d\xi} = s > 0 .$$

Then the system becomes

(3-9a) $$\frac{d\psi}{d\xi} = A\psi + B\mu = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \psi + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu$$

(3-9b) $$y = C\psi = [\ 1 \quad 0\ ]\ \psi$$

(3-10) $$\psi \triangleq \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \triangleq \begin{bmatrix} \phi_1 \\ s\phi_2 \end{bmatrix}$$

(3-11) $$\mu \triangleq \frac{ds}{d\xi}\phi_2 + s^2 v .$$

We can easily design a linear controller for this system, If we define s to be a function of the velocity of the robot, then the controller will depend on the robot's velocity.

In order to investigate how the function s effects the performance of the controlled system, we will consider the following optimal control problem. Here we will design a servo controller, so we must consider the following augmented system, (3-12) $$\frac{d}{d\xi}\begin{bmatrix}\psi \\ \mu\end{bmatrix} = \begin{bmatrix}A & B \\ 0 & 0\end{bmatrix}\begin{bmatrix}\psi \\ \mu\end{bmatrix} + \begin{bmatrix}0 \\ 1\end{bmatrix}\frac{d\mu}{d\xi}.$$

We will design the controller so as to minimize the following performance index,

(3-13) $$J = \int \left\{ \begin{bmatrix}\psi_1 \\ \psi_2 \\ \mu\end{bmatrix}^T \begin{bmatrix}q_1 & & \\ & q_2 & \\ & & q_3\end{bmatrix} \begin{bmatrix}\psi_1 \\ \psi_2 \\ \mu\end{bmatrix} + r\left(\frac{d\mu}{d\xi}\right)^2 \right\} d\xi .$$

The optimizing controller can be expressed in the following form,

(3-14) $$\frac{d\mu}{d\xi} = [\ f_1 \quad f_2 \quad f_3\ ] \begin{bmatrix}\psi_1 \\ \psi_2 \\ \mu\end{bmatrix}.$$

Since

(3-15) $$\begin{bmatrix}\frac{d\psi_1}{d\xi} \\ \frac{d\psi_2}{d\xi} \\ y\end{bmatrix} = \begin{bmatrix}A & B \\ C & 0\end{bmatrix}\begin{bmatrix}\psi_1 \\ \psi_2 \\ \mu\end{bmatrix},$$

we have

(3-16) $$\frac{d\mu}{d\xi} = [\ f_1 \quad f_2 \quad f_3\ ] \begin{bmatrix}A & B \\ C & 0\end{bmatrix}^{-1} \begin{bmatrix}\frac{d\psi_1}{d\xi} \\ \frac{d\psi_2}{d\xi} \\ y\end{bmatrix}$$

$$= [\ f_2 \quad f_3 \quad f_1\ ] \begin{bmatrix}\frac{d\psi_1}{d\xi} \\ \frac{d\psi_2}{d\xi} \\ y\end{bmatrix}$$

(3-17) $$\mu = f_1 \int (y - y_r)\, d\xi + f_2\, \psi_1 + f_3\, \psi_2$$

where y_r is the constant reference position. This feedback law can be expressed in the original time scale x as

(3-18) $$v = \frac{1}{s^2} f_1 \int (y - y_r) \frac{1}{s}\, dx + \frac{1}{s^2} f_2\, \phi_1 + \frac{1}{s} f_3\, \phi_2 - \frac{1}{s^2} \frac{ds}{d\xi}\, \phi_2 .$$

If s is a constant value, then the feedback(3-18) can be expressed as

$$v = \frac{1}{s^3} f_1 \int (y - y_r)\, dx + \frac{1}{s^2} f_2 \phi_1 + \frac{1}{s} f_3 \phi_2 . \tag{3-19}$$

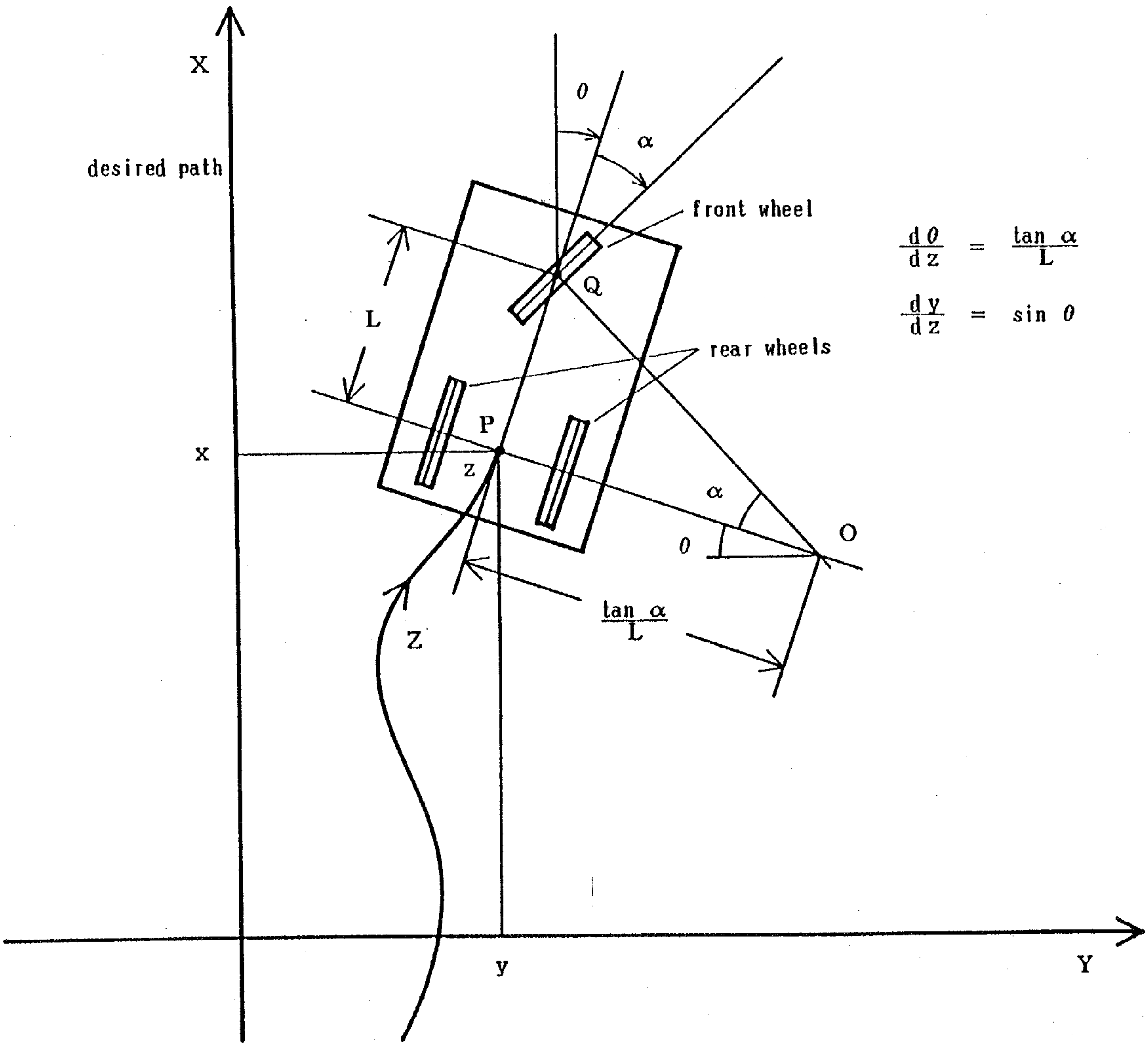
This feedback can be readily shown to minimize the performance index

$$J = \int \left\{ \begin{bmatrix} \phi_1 \\ \phi_2 \\ v \end{bmatrix}^T \begin{bmatrix} q_1 & & \\ & s^2 q_2 & \\ & & s^4 q_3 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ v \end{bmatrix} + s^6 r \left(\frac{dv}{dx} \right)^2 \right\} dx \tag{3-20}$$

for the following augmented system

$$\frac{d}{dx} \begin{bmatrix} \phi \\ v \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{dv}{dx} . \tag{3-21}$$

This system originates from the system (3-7) and expressed in x. The performance index (3-20) shows that the weight of ϕ_2 and v increase as the function s increases. Since ϕ and v are defined by eqs.(3-5) and (3-6), this implies that the weight of the robot's angle deviation θ and the steering angle α increase as s increases. In other words, if we choose a large s, then the resultant controlled system keeps the angle deviation θ and the steering angle α small. Thus, if we choose the function s carefully so that it increases with an increase in the robot's velocity; we can achieve the desired velocity dependent control: the steering angle α will be small when the robot's velocity is large.



$$\frac{d\theta}{dz} = \frac{\tan\alpha}{L}$$

$$\frac{dy}{dz} = \sin\theta$$

Fig.5-3-1. Tricycle

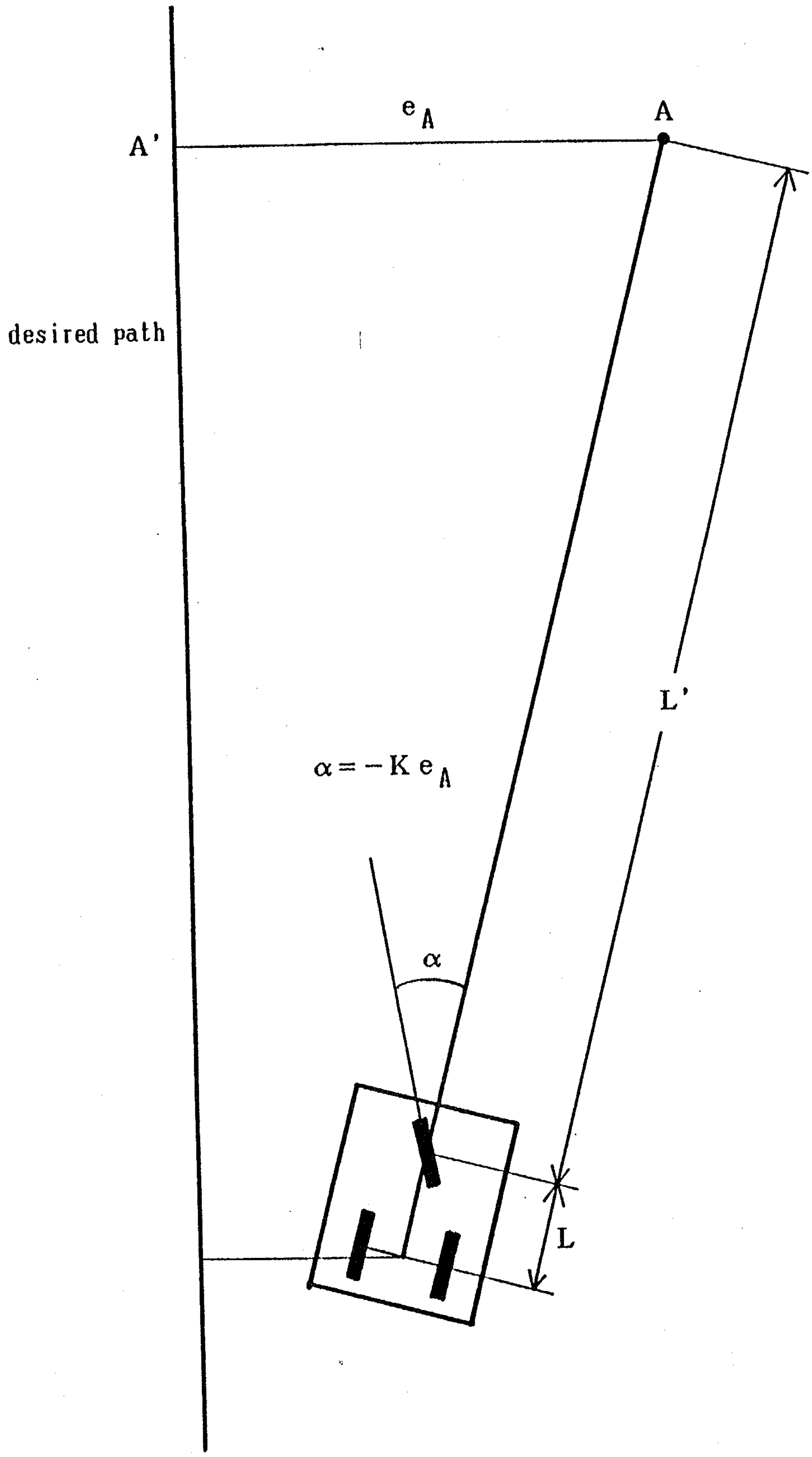


Fig.5-3-2. Conventional method 1

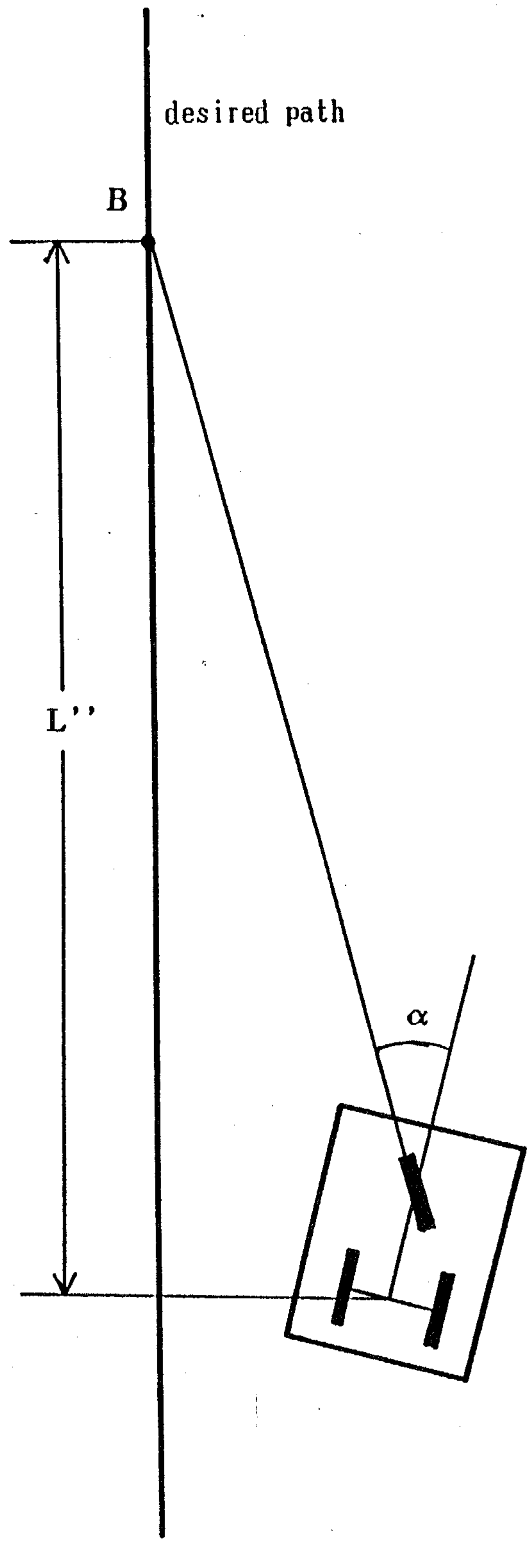


Fig.5-3-3. Conventional method 2

5-4. Concluding Remarks

In this chapter, we have proposed an alternative controller design method using a time scale transformation.

In section 5-1, we have applied this method to design a controller for a linear system which will avoid an excessively high amplitude of input. In section 5-2, we have proposed a controller for a robot that can achieve good performance even in the neighborhood of a singular point. In section 5-3, we have described a trajectory tracking controller for a vehicle. This controller allows us to analytically evaluate the stability of the path tracking control. The simulation results of these controllers showed that they have good performance.

The proposed controller design method can be applied to almost all mechanical systems. The choice of the time scaling function gives many options in designing stabilizing controllers. This approach is particularly useful to design robot controllers.

Ⅵ. CONCLUSION

In this thesis, we have proposed a time scale transformation and used it to analyze the intrinsic structure of nonlinear systems, invariant structure and the linearization problem in a transformed time scale. We have also used the time scale transformation to design a nonlinear controller.

We have introduced the time scale transformation with a differential equation of the form

$$\frac{dt}{d\tau} = s(x) > 0$$

where τ is the new time scale and t is the actual time scale. As we have already shown, this time scale transformation preserves the system's stability and the state's curve in state space. Furthermore, this transformation allowes us to express the original system

$$\frac{dx}{dt} = f(x) + g(x)u$$

in the time scale τ as

$$\frac{dx}{d\tau} = s(x)\ f(x) + s(x)\ g(x)u\ .$$

Since this system can also be expressed in a state equation, conventional methods can be used to analyze this system. Thus, we can readily use this transformation to investigate the system's stability, structure etc. In this thesis, we have applied the transformation to investigate the invariant structure of nonlinear systems(chapter Ⅲ), the linearization problem(chapterⅣ), and the nonlinear controller design problem(chapter Ⅴ).

In chapter Ⅲ, we investigated invariant structure using the time scale transformation. We introduced the notion of weakly invariant distribution in section 3-1 and showed how it can allow us to study invariant structure in a new time scale. We have proposed a simple geometric test for weak invariance. In section 3-2, weakly invariant distribution was used to obtain Kalman-like decompositions in reachable/unreachable parts and/or observable/unobservable parts in the transformed time scale. We have found the minimum dimension of the locally weakly invariant distribution under the vector fields f,g which contains g; this distribution coincides with the controllable subspace of a linear system. We have also proposed an algorithm to obtain the locally weakly invariant distribution under the vector fields f,g which is contained in the annihilator of the output function. This distribution coincides with the unobservable subspace of the linear system. In section 3-3, we introduced weakly controlled invariance, and successfully used it to solve the wide-sense disturbance decoupling problem. In this problem, it is necessary to seek a feedback law which would prevent the disturbance from affecting the output's curve in output space.

We have investigated the input-state linearization problem in chapter Ⅳ. We have identified the class of nonlinear systems which can be linearized in the transformed time scale. We also found that there are nonlinear systems which can not be linearized in the actual time scale t, but it is possible to linearize in the transformed time scale. The time scale in which the system can be linearized is obtained as the solution to partial differential equations. Since the time scale transformation preserves the systems's stability, it is possible to use the linearized model in the transformed time scale to obtain a stabilizing controller.

In chapter Ⅴ, we have proposed an alternative controller design method for both linear systems and robots. The proposed controller has been designed as follows: firstly we have introduced an appropriate time scale τ and expressed the system dynamics in the transformed time scale, we have then linearized the system in the the time scale τ. Finally, we have designed a linear controller in the time scale τ to stabilize the system. The controller is linear in the time scale τ, but it is nonlinear in the actual time scale t. In section 5-1, we have applied this method to design a controller for a linear system which will avoid an excessively high amplitude of input. In section 5-2, we have proposed a controller for a robot that can achieve good performance even in the neighborhood of a singular point. In section 5-3, we have described a trajectory tracking controller for a vehicle. This controller allows us to analytically evaluate the stability of the path tracking control. The simulation results of these controllers showed that they have good performance.

Studies on controllability/observability and the input-output linearization problem are still in progress. The controller design method which we have proposed in this thesis can be applied to almost all mechanical systems. The choice of the time scaling function gives many options in designing stabilizing controllers. This approach is particularly useful to design robot controllers.

Since the proposed time scale transformation is useful to analyze the system structure, further research in this area would be valuable.

REFERENCES

[1]R.W.Brockett,R.S.Millman and H.J.Sussmann eds.: "Differential Geometric Control Theory," Brikhauser (1983)

[2]A.Isidori: "Nonlinear Control Systems: An Introduction," Lecture Notes in Control and Information Sciences Vol.72, Springer-Verlag (1985)

[3]J.L.Casti: "Nonlinear System Theory," Academic Press (1985)

[4]M.Fliess and M.Hazewinkel eds.: "Algebraic and Geometric Methods in Nonlinear Control Theory," Reidel (1986)

[5]J.M.Hollerbach : "Dynamic Scaling of Manipulator Trajectories," Trans. of ASME, vol.106 , no.5 , 102/106 (1984)

[6]尾崎、秀田、山本、毛利:"空間経路が指定されるマニピュレータ動作の時間短縮,"計測自動制御学会論文集, vol.22, no.10, 1074/1080 (1986)

[7]W.L.Chow: "Uber Systeme von Linearen Partiellen Differentialgleichungen erster Ordnung," Math.Ann., vol.117, 98/105 (1939)

[8]R.Hermann: "On the Accessibility Problem in Control Theory," in Int. Symp. on Nonlinear Differential Equations and Nonlinear Mechanics, New York: Academic Press,325/223 (1963)

[9]G.W.Haynes and H.Hermes: "Nonlinear Controllability via Lie Theory," SIAM J. Control, vol.8, 450/460 (1970)

[10]R.W.Brockett: "System Theory on Group Manifolds and Coset Spaces," SIAM J. Control, vol.10, 265/284 (1972)

[11]C.Lobry: "Controllabilitie des systemes non lineaires," SIAM J. Control, vol.8, 573/605 (1974)

[12]H.J.Sussmann and V.J.Jurdjevic: "Controllability of Nonlinear Systems," J. Differential Equations, vol.12, 95/116 (1972)

[13]A.J.Krener: "A Generalization of Chow's Theorem and the Bang-Bang Theorem to Nonlinear Control Problems," SIAM J. Control, vol.12 43/52 (1974)

[14]R.Hermann and A.J.Krener: "Nonlinear Controllability and Observability," IEEE Trans. on Automatic Control, vol.22, no.5, 728/740 (1977)

[15]H.Sussmann: "Orbits of Families of Vector Fields and Integrability of Distributions," Trans. American Math. Soc., vol.180, 171/188 (1973)

[16]G.Basile and G.Marro: "Controlled and Conditioned Invariant Subspaces in Linear Systems Theory," J. Optimiz. Th. & Appl., vol.3, 306/315 (1969)

[17]M.H.Wonham and A.S.Morse: "Decoupling and Pole Assignment in Linear Multivariable Systems: a Geometric Approach," SIAM J. Contr., vol.8, 1/18 (1970)

[18]W.M.Wonham: "Linear Multivariable Control: a Geometric Approach," second edition, Springer-Verlag (1979)

[19]石島: "非線形制御系における外乱隔離可能性," 計測自動制御学会論文集, vol.14, no.1, 8/13 (1978)

[20]石島: "非線形制御系における外乱隔離条件の一般化," 計測自動制御学会論文集, vol.15, no.2, 269/270 (1979)

[21]T.Nomura and K.Furuta: "Invariant Structures of General Dynamical Systems," SIAM J. Control and Optimization, vol.19, no.1, 154/167 (1981)

[22]A.Isidori, A.J.Krener, C.Gori-Giori and S.Monaco: "Nonlinear Decoupling via Feedback, a Differential Geometric Approach," IEEE Trans. on Automatic Control AC-26 , no.2 , 331/345 (1981)

[23]A.Isidori, A.J.Krener, C.Gori-Giori and S.Monaco: "Locally (f,g) Invariant Distributions," Systems & Control Letters, vol.1, no.1, 12/15 (1981)

[24]R.M.Hirschorn: "(A , β) -Invariant Distributions and Disturbance Decoupling of Nonlinear Systems," SIAM J. Control and Optimization, vol.19, no.1 , 1/19 (1982)

[25]H.Nijmeijer: "Controlled Invariance for Affine Control Systems," Int.J.Control, vol.34, no.4 825/833 (1981)

[26]H.Nijmeijer and A.van der Shaft: "Controlled Invariance for Nonlinear Systems," IEEE Trans. on Automatic Control, vol.AC-27, no.4, 904/914 (1982)

[27]H.Nijmeijer and A.J.van der Shaft: "Controlled Invariance for Nonlinear Systems: Two Worked Examples," IEEE Trans. on Automatic Control, vol.AC-29, no.4, 361/364 (1984)

[28]H.Nijmeijer: "Controllability Distributions for Nonlinear Control Systems," Systems & Control Letters, vol.2, no.2, 122/129 (1982)

[29]In Joong Ha and E.G.Gilbert: "A Complete Characterization of Decoupling Control Laws for a General Class of Nonlinear Systems," IEEE Trans. on Automatic Control, vol.AC-31, no.9, 823/830 (1986)

[30]A.J.Krener: $(Ad_{f,g})$, $(ad_{f,g})$ and Locally $(ad_{f,g})$ Invariant and Controllability Distributions," SIAM J.Control and Optimization, vol.23, no.4, 523/549 (1985)

[31]S.Monaco and D.Normand-Cyrot: "Invariant Distributions for Discrete-Time Nonlinear Systems," Systems & Control Letters, vol.5, no.3, 191/196 (1984)

[32]J.W.Grizzle: "Controlled Invariance for Discrete-Time Nonlinear Systems with an Application to the Disturbance Decoupling Problem," IEEE Trans. on Automatic Control, vol.AC-30, no.9, 868/874 (1985)

[33]J.W.Grizzle: "Local Input-Output Decoupling of Discrete-Time Nonlinear Systems," Int.J.Control, vol.43, no.5, 1517/1530 (1986)
[34]S.N.Singh and W.J.Rugh: "Decoupling in a Class of Nonlinear Systems by State Variable Feedback," Trans. ASME J. Dyn. Syst. Meas. Contr., vol 94, 323/329 (1972)
[35]E.Freund: "The Structure of Decoupled Nonlinear Systems," Int. J. Control, vol.21 651/659 (1975)
[36]奥谷:"非線形システムにおけるモデル適合," 東工大修論 (1979)
[37]A.Isidori and A.Ruberti: "On the Synthesis of Linear Input-Output Responses for Nonlinear Systems," Systems & Control Letters, vol.4, 17/22 (1984)
[38]A.Isidori: "Nonlinear Feedback, Structure at Infinity and the Input-Output Linearization Problem," in Mathematical Theory of Networks and Systems, P.A.Fuhrmann ed., Springer-Verlag, 473/493 (1983)
[39]H.Nijmeijer and J.M.Schumacher: "Zeros at Infinity for Affine Nonlinear Control Systems," IEEE Trans. on Automatic Control, vol.AC-30, no.6, 566/573 (1985)
[40]小菅:"動的補償器による非線形システムのモデル適合," 東工大修論 (1980)
[41]A.Isidori: "The Matching of a Prescribed Linear Input-Output Behavior in a Nonlinear System," IEEE Trans. on Automatic Control, vol. AC-30, no.3, 258/265 (1985)
[42]R.W.Brockett: "Feedback Invariants for Nonlinear Systems," IFAC Congress, Helsinki (1978)
[43]B.Jakubczyk and W.Respondek: "On Linearization of Control Systems," Bull. Acad. Pol. Sci. Ser. Sci. Math., vol.XXVIII, 517/522 (1980)

[44]R.Su: "On the linear equivalents of nonlinear systems," Systems & Control Letters, vol.2, no.1, 48/52 (1982)

[45]L.R.Hunt,R.Su and G.Meyer: Design for Multi-Input Nonlinear Systems," Differential Geometric Control Theory, edited by R.Brockett et al., Birkhauser, 268/298 (1982)

[46]C.Reboulet and C.Champetier: "A New Method for Linearizing Nonlinear Systems: the Pseudolinearization," Int. J. Control, vol.40, no.4, 631/638 (1984)

[47]A.J.Krener: "Approximate Linearization by State Feedback and Coordinate Change," Syst. & Contr. Lett., vol.5, 181/185 (1984)

[48]島、川上、得丸: "Invariance の理論とその応用－Ⅰ操作量を線形に含む非線形系のInvariance," システムと制御, vol.23, no.10, 594/601 (1979)

[49]島: "Invariance の理論とその応用－Ⅱ Canonical From," システムと制御, vol.24, no.12, 816/824 (1980)

[50]島、川上、北: "Invariance の理論とその応用－Ⅲ非線形オブザーバの設計," システムと制御, vol.24, no.9, 621/629 (1980)

[51]島: "Invariance の理論とその応用－Ⅳ出力可制御性," システムと制御 vol.25, no.12, 778/787 (1981)

[52]北、島: "Invariance の理論とその応用－Ⅴ無干渉制御系," システムと制御, vol.26, no.4, 247/255 (1982)

[53]北、島: "Invariance の理論とその応用－Ⅵ対応定数による表現とL積分左逆システムの設計," システムと制御, vol.26, no.12, 791/800 (1982)

[54]島、石動、運崎: "Invariance の理論とその応用－Ⅶ外乱無干渉化," システムと制御, vol.29, no.1, 45/54 (1985)

[55]石動、島: "Invariance の理論とその応用－Ⅷモデル追従制御系," システムと制御, vol.29, no.5, 259/267 (1980)

[56]A.M.Lyapnov: "Probleme General de la Stabilite du Mouvement," Kharkov (1947 Princeton University Press)
[57]V.M.Popov: "Hyperstability of Control Systems," Springer-Verlag, (1973)
[58]計測自動制御学会編:"自動制御ハンドブック・基礎編," (1983)
[59]G.Zames: "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems. Part 1: Conditions Using Concepts of Loop Gain, Conicity, and Positivity," IEEE Trans. on Automatic Control, vol.AC-11, no.2, 228/238 (1966)
[60]G.Zames: "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems. Part 2: Conditions Involving Circles in the Frequency Plane and Sector Nonlinearities," IEEE Trans. on Automatic Control, vol.AC-11, no.3, 465/476 (1966)
[61]M.G.Safonov: "Stability and Robustness of Multivariable Feedback Systems," MIT Press (1980)
[62]D.P.Atherton: "Nonlinear Control Engineering: Describing Function Analysis and Design," Van Nostrand Reinhold (1975)
[63]E.Freund: "Fast Nonlinear Control with Arbitrary Pole Placement for Industrial Robots and Manipulators," Int.J.Robot Res., vol.I, no.1, 65/78 (1982)
[64]K.Furuta, K.Kosuge and N.Mukai: "Trajectory Tracking Control of Articulated Robot Arm with Sensor Feedback: Laser Beam Tracking System," proceedings of IECON'85, San Francisco, California, 307/312 (1985)
[65]N.Hogan: "Impedance Control: An Approach to Manipulation Part II Implementation," trans. of the ASME, J. of Dyn. Syst. Meas. Contr., vol.107, 8/16 (1985)

[66] J.M.Hollerbach and G.Sahar: "Wrist-Partitioned, Inverse Kinematic Accelerations and Manipulator Dynamics," Int.J. of Robotics Research, vol.2, no.4, 61/76 (1983)
[67] J.Y.S.Luh, M.W.Walker and R.P.C.Paul: "On-Line Computational Scheme for Mechanical Manipulators," trans. of ASME J. of Dyn. Syst. Meas. Contr., vol.102; 69/76 (1980)
[68] H.Asada: "Dynamic Analysis and Computer-Aided Design of Robot Manipulators," Preprints of IFAC 9th World Congress,Ⅷ, 262/267 (1984)
[69] K.Kosuge and K.Furuta: "Kinematic and Dynamic Analysis of Robot Arm," Proceeding of IEEE International Conference of Robotics and Automation, St. Louis, Missouri, 1039/1044 (1985)
[70] T.Yoshikawa: "Dynamic Manipulability of Root Manipulators," Proceeding of IEEE International Conference of Robotics and Automation, St. Louis, Missouri, 1033/1038 (1985)
[71] E.G.Gilbert and D.W.Johnson: "Distance Functions and Their Application of Robot Path Planning in the Presence of Obstacles," IEEE J. Robotics and Automation, vol.RA-1, no.1, 21/30 (1985)
[72] L.Y.Shih: "Automatic Guidance of Mobile Robots in Two-way Traffic," Automatica, vol.21, no.2, 193/198 (1985)
[73] 新井、中野:"移動車搭載形位置方向計測装置の開発と性能評価," 計測自動制御学会論文集, vol.18, no.10, 1013/1020 (1982)
[74] 中村、上田:"カルマンフィルタによる移動ロボットの位置推定," 計測自動制御学会論文集, vol.19, no.1, 8/14 (1983)
[75] 小森谷、舘、谷江:"移動ロボットの自律誘導の一方法," 日本ロボット学会誌, vol.2, no.3, (1984)
[76] 小森谷、舘、谷江、大野、阿部:"離散的ランドマークを使った移動機械の制御," 機械技術研究所所報, vol.37, no.1 1/10 (1983)

[77]津村、藤原、白川、橋本:"メモリに記憶されたコース指示による移動体の自動誘導,"日本ロボット学会誌, vol.2, no.3 209/214 (1984)
[78]S.Mori,H.Nishihara and K.Furuta: "Control of Unstable Mechanical System, Control of Pendulum," Int.J Control, vol.23, no.5, 673/692 (1976)
[79]K.Furuta,T.Okutani and H.Sone: "Computer Control of a Double Inverted Pendulum," Comput.&Elect.Engng., vol.5 67/84 (1978)

AUTHOR'S PUBLICATIONS

[1] 三平満司、古田勝久:"レーザーを用いた機械系の軌道制御," 計測自動制御学会論文集, 第20巻, 4号, 344/349 (1984)

[2] Mitsuji SAMPEI, Katsuhisa FURUTA: "On Time Scaling for Nonlinear Systems : Application to Linearization," IEEE trans. on Automatic Control, vol.AC-31, no.5, 459/462 (1986)

[3] 三平満司、古田勝久:"時間軸を考慮に入れた非線形システムの解析——軌跡に着目した外乱局所化問題への応用——," 計測自動制御学会論文集, 第22巻, 6号, 604/609 (1986)

[4] 三平満司、古田勝久:"時間軸の変換を用いた非線形システムの線形化——新時間軸での線形化——," 計測自動制御学会論文集, 第22巻, 10号, 1030/1036 (1986)

[5] Katsuhisa FURUTA, Mitsuji SAMPEI: "Path Control of Three Dimensional Linear Motion Mechanical System Using Laser," IECON '84, Tokyo, 189/193, (1984)

[6] Mitsuji SAMPEI, Katsuhisa FURUTA: "Robot Control in the Neighborhood of Singular points," to appear in IEEE International Conference on Robotics and Automation, Raleigh, North Carolina (1987)

[7] Mitsuji SAMPEI, Katsuhisa FURUTA: "On Linearization and Control with Transformation of the Time Scale," to appear in IFAC 10th World Congress, Munich (1987)

[submitted]

- Mitsuji SAMPEI, Katsuhisa FURUTA: "Robot Control in the Neighborhood of Singular points," IEEE Journal of Robotics and Automation
- Katsuhisa FURUTA, Mitsuji SAMPEI:"Path Control of a Three Dimensional Linear Motional Mechanical System Using a Laser," IEEE Trans. on Industrial Electronics

[others]

- 古田勝久、三平満司:"レーザ光線を用いた機械系の軌道制御," ロボット, no.48, 40/46 (1985)
- 三平満司、竹内裕喜:"エンジニアのための実用プログラムーノウハウ集、開ループ系と閉ループ系のボード線図," ＯＨＭ, vol.73, no.10, 77/80 (1986)

APPENDIX

In this appendix, we will review the basic concepts of differential geometry and define the notation which is used in this thesis.

Manifold

The manifold M of dimension n is a Hausdorff space such that, for each $p \in M$, there exist an open neighborhood U of p and a homeomorphism(one to one, onto, continuous and open) ϕ mapping U onto an open set in R^n.

$$(1) \qquad \phi : U \longrightarrow R^n, \quad p \longmapsto x$$

The pair (U, ϕ) is called a coordinate chart on M. Sometimes ϕ is expressed as the set of functions $(\phi_1, \cdots, \phi_n)$. $\phi_i : U \rightarrow R$ is called the i-th coordinate function. If $q \in U$, the n-tuple of real numbers $(\phi_1(q), \cdots, \phi_n(q))^T$ is called the set of local coordinates. Two coordinate charts (U, ϕ) and (V, ψ) are C^∞-compatible if, whenever $U \cap V$ is not empty, the coordinate transformation $\psi \cdot \phi^{-1} : R^n \rightarrow R^n$ is diffeomorphic, i.e. if $\psi \cdot \phi^{-1}$ and $\phi \cdot \psi^{-1}$ are both C^∞ maps. The C^∞ atlas on the manifold M is a collection $\{(U_i, \phi_i)\}$ of pairwise C^∞-compatible coordinate charts, with the property that $\{U_i\}$ covers M. An atlas is complete if not properly contained in any other atlas. A smooth manifold (or C^∞ manifold) is a manifold equipped with a complete C^∞ atlas. Thus, any diffeomorphic coordinate transformations can be used on a smooth manifold.

Let M and N be smooth manifolds of dimension n and m, $F : M \rightarrow N$ be a mapping, (U, ϕ) and (V, ψ) be coordinate charts on M and N respectively. The composed mapping

(2) $\hat{F} = \psi F \phi^{-1} : \quad R^n \longrightarrow R^m$

is the expression of F in local coordinates. We often use F to denote $\psi F \phi^{-1}$. A mapping $F : M \to N$ is a <u>smooth mapping</u> if for each $p \in M$ there exists coordinate charts (U, ϕ) of M and (V, ψ) of N with $p \in U$ and $F(p) \in V$, such that the expression of F in local coordinates is C^∞. A smooth mapping $\alpha : M \to R^1$ is called a <u>smooth function.</u> We often use the same symbol α to denote an expression of α in the local coordinates $\alpha \phi^{-1}$.

<u>Vector Field</u>

The set of all smooth functions is denoted $\mathbb{F}$. The <u>tangent vector</u> v at p is the map $v : \mathbb{F} \to R$ with the following properties: for all $\alpha, \beta \in \mathbb{F}$ and $a, b \in R$, $v(a\alpha + b\beta) = av(\alpha) + bv(\beta)$ and $v(\alpha\beta) = \alpha(p)\, v(\beta) + \beta(p)\, v(\alpha)$. The tangent vector is expressed in the local coordinate chart (U, ϕ) as

(3) $$v = v_1 \left(\frac{\partial}{\partial \phi_1}\right)_p + \cdots + v_n \left(\frac{\partial}{\partial \phi_n}\right)_p$$

where $v_1, \cdots, v_n \in R$. The basis $\left\{ \left(\frac{\partial}{\partial \phi_1}\right)_p, \cdots, \left(\frac{\partial}{\partial \phi_n}\right)_p \right\}$ is the natural basis introduced by (U, ϕ). The tangent vector v acts on the smooth function α as

(4) $$v(\alpha) = v_1 \left(\frac{\partial \alpha}{\partial \phi_1}\right)_p + \cdots + v_n \left(\frac{\partial \alpha}{\partial \phi_n}\right)_p .$$

The <u>tangent space</u> to M at p is the set of all tangent vectors at p and is written T_pM. T_pM forms a vector space over the field R.

The <u>vector field</u> f on M is a mapping which assigns the tangent vector f(p) to each point $p \in M$. A vector field f is <u>smooth</u> if for each $p \in M$ there exists a coordinate chart (U, ϕ) about p and n real-valued smooth functions $f_1, \cdots, f_n$ defined on U such that for all $q \in U$,

(5) $$f(q) = f_1(q)\left(\frac{\partial}{\partial \phi_1}\right)_q + \cdots + f_n(q)\left(\frac{\partial}{\partial \phi_n}\right)_q .$$

Consider the following differential equation expressed in local coordinates

(6) $$\dot{x} = f(x) = (f_1(x), \cdots, f_n(x))^T$$

where $x \in R^n$. For each smooth function $\alpha(x)$, we have

(7) $$\frac{d\alpha}{dt} = \frac{d\alpha}{dx}\frac{dx}{dt} = \frac{d\alpha}{dx} f(x) = f_1 \frac{\partial \alpha}{\partial x_1} + \cdots + f_n \frac{\partial \alpha}{\partial x_n} .$$

Since eq.(7) is identical to eq.(4), we can identify a vector field with a differential equation of the form(6). Thus, we usually express a vector field in the form of a column vector $f(x) = (f_1(x), \cdots, f_n(x))^T$ in local coordinates. A vector field $f(p)$ is expressed in the local coordinate chart (U, ϕ) as $\hat{f}(x) = f\phi^{-1}(x) = (\hat{f}_1(x), \cdots, \hat{f}_n(x))^T$. We often use the symbol $f(x)$ to denote $f\phi^{-1}(x)$. The solution of differential equation(6) is called flow and is denoted $x(t) = \Phi^f_t x_0$ where x_0 is the initial value and f is the vector field. The set of all vector fields is called the tangent bandle and is written TM.

Let M and N be smooth manifolds. Let $F: M \to N$ be a smooth mapping. The differential of F at $p \in M$ is the map $F_*: T_pM \to T_{F(p)}N$ and is defined as

(8) $$(F_*(v))(\alpha) = v(\alpha \cdot F)$$

for $v \in T_pM$ and α is a smooth function on N. If F is a coordinate transformation, F_* is the transformation of the vector space related to F.

Covector Field

The cotangent space to M at p, written $T^*_p M$, is the dual space of $T_p M$. The elements of the cotangent space are called tangent covectors. Since the dual space of the n dimensional column vector space is a n dimensional row vector space, we usually express the tangent covector in the form of a row vector $\omega = (\omega_1, \cdots, \omega_n)$ in local coordinates. The inner product is denoted $\langle \cdot, \cdot \rangle : T^*_p M \times T_p M \to R$. It satisfies the following equation for any $a_i, b_i \in R$, any vector fields f,g and any tangent covector ω, σ

(9) $$\langle a_1\omega + a_2\sigma, b_1 f + b_2 g\rangle = a_1 b_1\langle \omega, f\rangle + a_1 b_2\langle \omega, g\rangle + a_2 b_1\langle \sigma, f\rangle + a_2 b_2\langle \sigma, f\rangle .$$

In local coordinates (U, ϕ), if $\left\{ \left(\frac{\partial}{\partial \phi_1}\right)_p, \cdots, \left(\frac{\partial}{\partial \phi_n}\right)_p \right\}$ is the basis of the tangent space, the unique basis $\{(d\phi_1)_p, \cdots, (d\phi_n)_p\}$ which satisfies $\langle \frac{\partial}{\partial \phi_i}, d\phi_j \rangle = \delta_{ij}$ is called the dual basis. ω is represented in this basis as $\omega = \omega_1 (d\phi_1)_p + \cdots + \omega_n (d\phi_n)_p$.

A covector field ω on M is a mapping which assigns a tangent covector $\omega(p)$ to each point $p \in M$. The notion of a smooth covector field is defined analogously to that of a vector field. The tangent covector $\omega(p)$ is expressed in local coordinates as $\hat{\omega}(x) = \omega\phi^{-1}(x) = (\hat{\omega}_1(x), \cdots, \hat{\omega}_2(x))$. We often use the symbol $\omega(x)$ to denote $\omega\phi^{-1}(x)$. Let M and N be smooth manifolds and $F : M \to N$ be a smooth mapping. The mapping $F^* : T^*_{F(p)} N \to T^*_p M$ is defined as follows

(10) $$\langle F^*(\omega), v\rangle = \langle \omega, F_*(v)\rangle$$

for any $\omega \in T^*_{F(p)} N$ and any $v \in T_p M$.

Lie Derivative

Let f be a smooth vector field on M and α be a smooth real-valued function on M. The Lie derivative of α along f is the function $M \to R$, written $L_f \alpha$ and defined by

(11) $$(L_f \alpha)(p) = (f(p))(\alpha).$$

In local coordinates, this is represented by

(12) $$(L_f \alpha)(x) = \frac{\partial \alpha}{\partial x} f(x).$$

The following notation is used

(13) $$L_f^0 \alpha = \alpha \quad \text{and} \quad L_f^i \alpha = L_f(L_f^{i-1} \alpha).$$

Let f and g be two smooth vector fields on M and let Φ_t^f denote the flow of f. The Lie derivative of g along f is the vector field on M, written $L_f g$ and defined by

(14) $$(L_f g)(p) = \lim_{t \to 0} \frac{1}{t} \{ (\Phi_{-t}^f)_* g(\Phi_t^f(p)) - g(p) \}.$$

It is expressed in local coordinates as

(15) $$(L_f g)(x) = \frac{\partial g}{\partial x} f(x) - \frac{\partial f}{\partial x} g(x).$$

Both of the following notations are used recurrently,

(16) $$L_f^0 g = g \quad \text{and} \quad L_f^i g = L_f(L_f^{i-1} g)$$

(17) $$ad_f^0 g = g \quad \text{and} \quad ad_f^i g = ad_f(ad_f^{i-1} g).$$

$L_f g$ sometimes is denoted by $[f,g]$ which is called the Lie bracket. The Lie bracket has the following properties.

(18a) $$[f_1, f_1] = 0$$

(18b) $$[f_1, f_2] + [f_2, f_1] = 0$$

(18c) $$[f_1, [f_2, f_3]] + [f_2, [f_3, f_1]] + [f_3, [f_1, f_2]] = 0$$

(18d) $$\begin{aligned}[\alpha f_1, \beta f_2] &= \alpha\beta [f_1, f_2] - \beta f_1 \frac{\partial \alpha}{\partial x} f_2 + \alpha f_2 \frac{\partial \beta}{\partial x} f_1 \\ &= \alpha\beta [f_1, f_2] - (\beta L_{f_2} \alpha) f_1 + (\alpha L_{f_1} \beta) f_2\end{aligned}$$

where f_i is a smooth vector field and α, β are smooth functions.

The Lie derivative of the covector field ω along the vector field f is a covector field on M, written $L_f\,\omega$ and defined by

(19) $$(L_f\,\omega)(p)=\lim_{t\to 0}\frac{1}{t}\{(\Phi_t^f)^*\,\omega\,(\Phi_t^f(p))-\omega(p)\}.$$

It can be expressed in local coordinates as

(20) $$(L_f\,\omega)(x)=(\frac{\partial\omega^T}{\partial x}f(x))^T+\omega(x)\frac{\partial f}{\partial x}.$$

Distribution

A distribution Δ on M is a mapping which assigns the subspace $\Delta(p)$ of T_pM to each $p\in M$. If each of these subspaces is of dimension k, it is called a k dimensional distribution.

A distribution Δ is smooth if, for each point $p\in M$, there exist a neighborhood U of p and a set of smooth vector fields $\{\nu_i : i\in I\}$ defined on U having the property

(21) $$\Delta(q) = \mathrm{span}\{\nu_i(q) : i\in I\}$$

for all $q\in U$. $\Delta=\mathrm{span}\{\nu_i : i\in I\}$ is the distribution defined by

(22) $$(\mathrm{span}\{\nu_i : i\in I\})(p) = \mathrm{span}\{\nu_i(p) : i\in I\}.$$

If Δ_1 and Δ_2 are two distributions, then their sum $\Delta_1+\Delta_2$ and their intersection $\Delta_1\cap\Delta_2$ are defined by taking

(23) $$(\Delta_1+\Delta_2)(p)=\Delta_1(p)+\Delta_2(p)$$

(24) $$(\Delta_1\cap\Delta_2)(p)=\Delta_1(p)\cap\Delta_2(p).$$

A distribtion Δ is nonsingular if there exists an integer d such that $\dim\Delta(p) = d$ for all $p\in M$. The point p is a regular point of the distribution Δ if there exists a neighborhood U of p with the property that Δ is nonsingular on U.

The codistribution Ω on M is a mapping which assigns a subspace $\Omega(p)$ of $T^*_p M$ to each $p \in M$. The notions of smooth codistribution, k dimensional codistribution, span{ ω_i : $i \in I$ }, sum, intersection, nonsingular codistribution, and regular point of a codistribution are defined analogously to distribution.

Let Δ be a distribution and Ω be a codistribution. The annihilator of Δ, denoted $\Delta^\perp$, is the codistribution defined by the rule

(25) $$\Delta^\perp(p) = \{\omega \in T^*_p M : \langle\omega, \nu\rangle = 0 \text{ for all } \nu \in \Delta\}.$$

Conversely, the annihilator of Ω, denoted $\Omega^\perp$, is the distribution defined by the rule

(26) $$\Omega^\perp(p) = \{\nu \in T_p M : \langle\omega, \nu\rangle = 0 \text{ for all } \omega \in \Omega\}.$$

The set of smooth vector fields $\{\nu_1, \cdots, \nu_r\}$ is involutive if there exist the smooth functions $\phi_k^{(i,j)}(x)$ such that for all $i, j \leqq r$

(27) $$[\nu_i, \nu_j] = \sum_{k=1}^{r} \phi_k^{(i,j)} \nu_k .$$

The distribution Δ is involutive if the Lie bracket $[\nu, \theta]$ of any two smooth vector fields $\nu, \theta \in \Delta$ belong to Δ, i.e. if $[\nu, \theta] \in \Delta$.

A covector field ω is called exact one-form if there exists a smooth real valued function $\alpha : M \to R$ such that

(28) $$\omega = d\alpha \triangleq \frac{\partial \alpha}{\partial x_1} dx_1 + \cdots + \frac{\partial \alpha}{\partial x_n} dx_n .$$

A nonsingular d-dimensional distribution Δ on M is completely integrable if at each $p \in M$ there exists a coordinate chart (U, ξ) with coordinate functions $\xi_1, \cdots, \xi_n$ such that

(29) $$\Delta(q) = \text{span}\left\{\left(\frac{\partial}{\partial \xi_1}\right)_q, \cdots, \left(\frac{\partial}{\partial \xi_d}\right)_q\right\}$$

for all $q \in U$.

Theorems

[lemma 1]

Let Δ be a smooth distribution and p be a regular point of Δ. Suppose $\dim \Delta(p) = d$. Then there exist an open neighborhood U of p and a set $\{\nu_1, \cdots, \nu_d\}$ of smooth vector fields defined on U with the property that every smooth vector field θ belonging to Δ can be represented on U as

$$\theta = \sum_{i=1}^{d} \alpha_i \nu_i \tag{30}$$

where α_i is a real-valued smooth function defined on U. □

[lemma 2]

The set of all regular points of the distribution Δ is an open and dense submanifold of M. □

[lemma 3]

Let Δ_1 and Δ_2 be two smooth distributions with the property that Δ_1 is nonsingular, $\Delta_1 \subset \Delta_2$ and $\Delta_1(p) = \Delta_2(p)$ at each point p of the dense submanifold of M. Then $\Delta_1 = \Delta_2$. □

[lemma 4]

Let p be a regular point of Δ. Then p is a regular point of $\Delta^{\perp}$ and there exists a neighborhood U of p with the property that $\Delta^{\perp}$ is smooth on U. □

[theorem 5] (Frobenius)

A nonsingular distribution is completely integrable if and only if it is involutive. □

[theorem 6]

Let U and V be open sets on R^n and R^m, respectively. Let $x_1, \cdots, x_m$ denote the coordinates of the point x in R^m and $y_1, \cdots, y_n$ denote the coordinates of the point y in R^n. Let $\Gamma^1, \cdots, \Gamma^m$ be smooth functions

(31) $\Gamma^i : U \rightarrow R^{n \times m}$.

Consider the set of partial differential equations

(32) $$\frac{\partial y(x)}{\partial x_i} = \Gamma^i(x)\, y(x) \qquad 1 \leqq i \leqq m$$

where y denotes the function

(33) $y : U \rightarrow V$.

Given the point $(x^0, y^0) \in U \times V$ there exist a neighborhood U_0 of x^0 in U and the unique smooth function

(34) $y : U_0 \rightarrow V$

which satisfies equation(32) such that $y(x^0) = y^0$ if and only if the functions $\Gamma^1, \cdots, \Gamma^n$ satisfy the condition

(35) $$\frac{\partial \Gamma^i}{\partial x_k} - \frac{\partial \Gamma^k}{\partial x_i} + \Gamma^i \Gamma^k - \Gamma^k \Gamma^i = 0 \qquad 1 \leqq i, k \leqq m$$

for all $x \in U$. □