OUKA
The University of Osaka
Institutional Knowledge Archive

Title	学習方式によるロボットの運動制御
Author(s)	川村, 貞夫
Citation	大阪大学, 1986, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/773
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

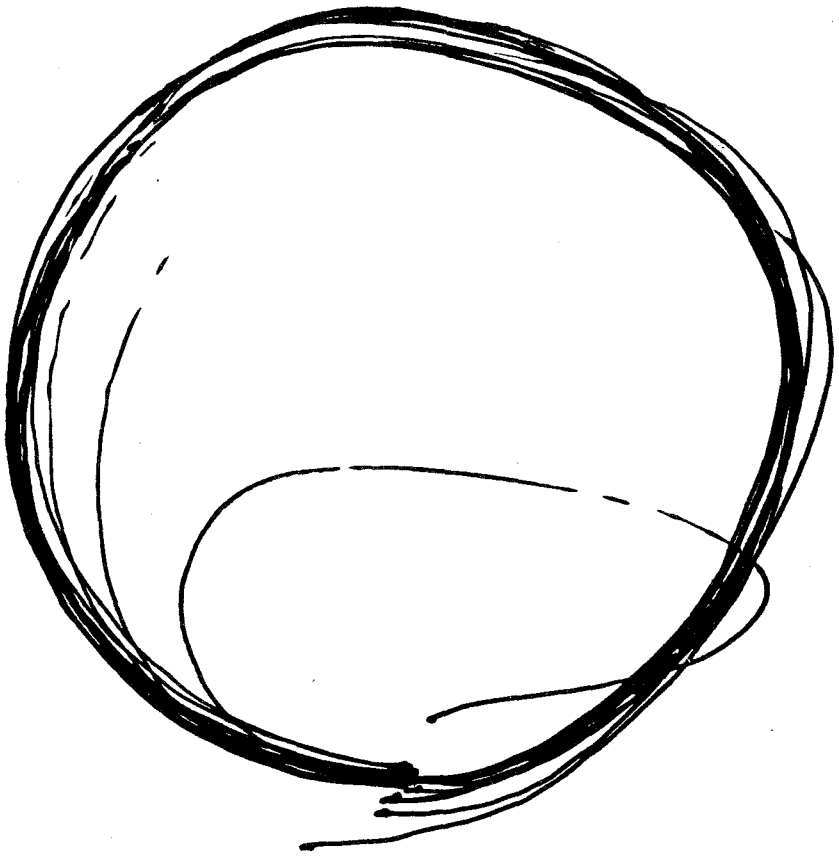
学習方式によるロボットの運動制御

昭和61年2月

川村貞夫

人一度して之を能くすれば
己は之を百度す
人十度して之を能くすれば
己は之を千度す
果して此の道を能くすれば
愚といえども必ず明
柔といえども必ず強なり

中庸（四書五教）
第二十章
第十九節



（本論文第９章実験結果より）

目次

第１章　緒論

第２章　時不変線形メカニカルシステムに対する学習制御

２－１．緒言

２－２．誤差の収束性

２－３．結言

第３章　学習制御方式のシステム論的考察

３－１．緒言

３－２．周波数領域における解析

３－３．有界実性，感度，最適レギュレータとの関係

３－４．シミュレーション結果

３－５．結言

第４章　時変線形メカニカルシステムに対する学習制御

４－１．緒言

４－２．誤差の収束性

４－２－１．速度信号による入力の修正法

４－２－２．位置信号と速度信号による入力の修正法

４－３．結言

第５章　ある種の非線形システムに対する学習制御

５－１．緒言

５－２．誤差の収束性

５－３．結言

第６章　移動ロボットへの適用

６－１．緒言

６－２．二足歩行ロボット

６－３．運動学習経験に基づく歩行パターンの選択

６－４．結言

第７章　ロボットマニピュレータの運動制御への適用

７－１．緒言

７－２．非線形システムとしての定式化

７－３．時変システムとしての定式化

７－４．結言

第８章　ロボットマニピュレータの力制御への適用

８－１．緒言

８－２．力と位置のハイブリッド学習制御の構成

８－３．結言

第９章　実験結果

９－１．緒言

９－２．二足歩行ロボット

９－３．ロボットマニピュレータ（１）

９－３－１．速度信号による入力の修正

９－３－１－１．関節角座標制御

９－３－１－２．作業座標制御

９－３－２．加速度信号による入力の修正

９－４．ロボットマニピュレータ（２）

９－４－１．円軌道

９－４－２．直線軌道

９－５．ロボットマニピュレータの力制御

９－５－１．研磨作業

９－５－２．グライデング作業

９－６．結言

第１０章　結論

謝辞

参考文献

記号

R^n:実数を要素とするｎ次元ベクトルの集合．

$R^{n \times m}$：実数を要素とするｎｘｍ次元ベクトルの集合．

$x \in N$：ｘが集合Ｎの要素であることを示す．

C^n［０，Ｔ］：区間［０，Ｔ］でC^n級関数の集合．

第１章　緒論

１－１　ロボットの運動制御

人間と同じような姿をして，人間と同じような行動をする機械にロボットという名前が初めて付けられたのはチェコスロバキアの作家カレル・チャペルが１９２０年に発表した戯曲においてであった．　その後このような機械を実現するために基礎的な研究が続けられ，特に１９８０年代に入ってエレクトロニクスの急速な発達にともなって，ロボットは色々な作業現場で使用される身近な機械になりつつある．現在では塗装作業や溶接作業などある程度人間の作業をロボットに代行させるという目的は達成されたと言っても過言ではない．しかしながら，現段階ではロボットの運動という最も基本的な分野に限っても，理想的な運動パターンをロボット自身が自律的に実現しているとは言い難い．即ち，多くの産業用ロボットでは，人間がロボットを動かして教示する（ティーチングプレイバック方式）方式が主流であり，人間による教示作業をプレイバックするだけのものが多い．それは，チャペルが始めに想像し，その後多くの人が夢見た理想的なロボットの姿とはあまりに違ったものとなっている．

多くのロボットは作業空間を広くとり多種多様な作業を実現するために多自由度の関節構造のものが利用されている．これらのロボットでは，そのダイナミクスが多変数であり非線形であることから運動の制御が著しく難しいものとなっている．このような特徴をもつロボットに対して，古典制御理論におけるＰ．Ｉ．Ｄ方式を基礎としたサーボ理論によって目標運動についてのフィードバック制御系を構成する方式が用いられ，実際の作業現場にも適用されはじめている．さらに，この方法は，非線形特性を有するロボット運動のグローバルなロバスト性と安定性を保証できることが示されており，ロボットの運動制御の有効な一手法となっている．[1]~[4]　しかしながら，現実的により高速な目標運動パターンをより高精度に実現しようとするとき，フィードバックゲインを極端に大きくすることが安定性などの問題から不可能であるので，従来のフィードバック系だけでなく何らかのフィードフォワード系をロボットに付加することが必要となる．このような考えに基づいた非線形補償法が既にいくつか開発されている．[8]　その最も基本的な形は，ロボットに目標運動を実現させるために必要な入力を，あらかじめ推定されたロボットの物理パラメータに基づいて計算する方法である．ところが，これらの方式ではロボットダイナミ

クスのパラメータを正確に同定しなければ入力が得られないことや，入力の計算が複雑であるなどの問題がおこる．またこれとは別の方式で，ロボットのパラメータの同定を事前に必要としないという点で有効なモデル規範形適応制御方式がロボットに適用されたが，[9],[10] 現時点では計算機の計算速度の問題や，ロボットの非線形性をどのように取り扱うかの問題が残されている．

１．２　学習という概念の導入

それでは，このような特性をもつロボットに対して，どのような制御方式が適しているのであろうか．従来の制御理論に着目してみると，まず制御理論のおこりは，ワットによって発明された蒸気機関の回転速度を調節するガバナー（調速器）に対する解析であると言われている．そこでは，周波数領域においてシステムの安定性が解析され，古典制御理論として体系化された．その後，システムの状態空間という概念に基づいた現代制御理論が登場し，最適性を考慮した様々な手法が開発された．ここで留意すべきことは，古典制御理論における安定性の解析や現代制御理論における最適性の解析の基本は，線形システムに対する時間無限大でのシステム応答の厳密な説明であり，必ずしも過渡的なシステムの応答を直接操作し得るとは言い難い点である．具体的には，古典制御理論においては，過渡的な特性を定めるパラメータを最終的に人間の経験によって決定していることや，現代制御理論での最適レギュレータでは過渡的な出力パターンを直接操作し難い事などがあげられる．一方，ロボットの運動制御については，時間無限大での運動の様子よりもむしろ有限時間区間での過渡的な応答を目標パターンに一致させることが重要となる．　従って，従来の制御手法をそのままロボットの運動制御に適用するよりは，ロボットの運動特性に適した制御方式を開発することが切望される．

有限時間区間でロボット運動を目標パターンに近づけるための制御手法を考えるとき，我々人間がひとつの運動フォームを獲得する方法が大きなヒントを与えてくれる．　それは，目標運動パターンが有限な時間区間で与えられるが故に，何度も実際の運動を繰り返すことが可能となり，その繰り返しによって目標運動パターンを実現するようないわゆる学習能力である．我々は色々な運動フォームを学習することによって獲得している．本研究の目的は，このような運動についての学習能力をロボットに付加することである．　即ち，与えられた目標パターン（理想的な運動フォーム）を試行を繰り返すことによって実

現するものである．具体的に一般の制御システムについてのべれば，パラメータが推定されていないシステムの入力と出力のみに注目し，目標出力と毎回の試行時のシステムの出力の差により入力を修正して，目標出力を達成する制御方式である．このような制御方式が構成できればシステムについてパラメータを同定する必要はなく毎回の試行の繰り返しによって，制御対象について学習することが可能となると予想される．このことはロボットダイナミクスを正確に推定する必要なく，ロボットが試行を繰り返すうちに目標運動パターンを実現することを意味している．従って，経年変化や把持する対象物の変化により起こるパラメータの変化についても学習方式によって克服することができるものと期待できる．

学習能力を機械に付加した初めての例は１９６０年前後にRosenblattによって提案されたパーセプトロンがあげられる．その後，１９７０年代には制御工学の分野にも学習方式を導入した例はあるが，パターン認識手法の拡張や確率近似的な手法の，パラメータの学習方式であり，動的なシステムを対象とするものは見られなかった．その理由の一つとしては，前述のごとく入出力のパターンに基づいた動的システムに対する学習方式を実現しようとすれば，入出力パターンを何らかの方法でメモリーしておかなければならないことが挙げられる．数十年前ではこれは非常に難しいことであったが，現在では高度に発達したＶＬＳＩ技術により，コンパクトにかつ安価にメモリーすることが可能となり動的システムに対して学習という概念を導入するアイディアが現在では現実的な制御手法として浮かび上がってきたと考えられる．

１－３　各章の構成

まず，第２章では最も簡単な例として多入力多出力の時不変線形メカニカルシステムについて，ある入出力関係と条件のもとに前節で述べた学習に関するアイディアが実現できることを示す．次に，第３章では，対象を時不変線形システムに限定することによって，従来の周波数領域における種々の解析結果とのつながりを示す．具体的には，システムの有界実性，感度と呼ばれる量との対応と最適レギュレータの構成の条件との対応を明確にする．また，この関係によって，学習プロセスを構成するために，状態フィードバックシステムのフィードバックゲインはリッカチの方程式を解くことによって得られることが示される．

ロボットマニピュレータの運動制御に適用可能な学習プロセスを構成する準備として，第４章では，一般的な時変線形メカニカルシステムについて学習方式を定式化し，その有効性を理論的に証明する．さらに，学習プロセスの構成に際して，出力の微分信号を利用することによって，制御対象はある種の非線形システムにまで拡張できることが，第５章において理論的に示される．

以下，第６章，第７章　及び，第８章では，先に提案された学習方式をどのように実際の制御対象に適用するかが論じられる．先ず，第６章では，本研究で提案する学習方式が特に移動ロボットの制御に適していることを指摘する．具体的には，二足歩行ロボットに対して，適用可能であることが明らかにされる．さらに，一度学習により獲得した情報を別の運動パターンに用いるシステムについて述べる．次に，第７章においてはロボットマニピュレータに第４章及び第５章で述べられた方式を適用できることが示される．そのために，７－２節ではロボットダイナミクスの特性を考慮して第５章と類似の学習プロセス構成の条件が導出される．また，７－３節ではロボットダイナミクスが目標軌道まわりで線形化され，時変線形システムとして表現する．さらに，第８章では，位置と力に線形構造を仮定することにより，力の目標パターンをも学習プロセスにより実現できることが，明らかにされる．その際，位置と力の比例定数は未知である場合にも本方式の有効性が示される．これらの制御対象に対して実際にいくつかの実験が行われ，それらの実験結果が第９章に示される．最後に，第１０章において全章に関するまとめ，及び今後の研究課題等について述べる．

第２章　時不変線形メカニカルシステムに対する学習制御

２－１　緒言

本章では，学習プロセスの構造を簡単に説明するために，制御対象を時不変線形システムに限定して考える．このようなクラスの制御対象については，本論文で提案する方式と類似の方式がいくつか報告されている．まず，内山によってロボットマニピュレータが運動を繰り返しながら目標運動を実現するという先駆的な研究が行なわれた．[12]これは学習という表現は避けているが，ロボットの入力パターンを複雑な計算を介さずに形成しようという意図で開発されたものであった．一方，別のアプローチは，中野らによって開発された繰り返し制御と呼ばれるものがある．[13],[14] これは目標出力パターンが周期関数で与えられる場合に，　前回の周期での誤差（目標出力と実際のシステムの応答の差）によって，　次回の周期の入力を修正して漸近的に目標出力に近づけようとするものである．また，近年類似の方式についての研究も行なわれている．[15]~[17] いずれの手法もラプラス変換を解析に利用して，周波数領域で議論がなされている．本章ではこれらの手法とは別に，時変線形システムや非線形システムに拡張する目的で，時間領域での解析を行なう．また，ロボットのようなメカニカルシステムに学習方式を適用することを考慮して，特にメカニカルシステムに制御対象を限定して議論する．

２－２　時不変線形メカニカルシステムに対する学習制御

次のような多入力多出力のメカニカルシステムを考える．

$$R\ddot{x}(t)+Q\dot{x}(t)+Px(t)=u(t),$$
$$y(t)=\dot{x}(t) \qquad (2-1)$$

ただし，$u\in R^n$：入力，$y\in R^n$：出力，R，Q，$P\in R^{n\times n}$であり，行列R，Q，Pが正定対称であるようなメカニカルシステムを考える．ここでの問題設定として重要

なことは対象システムとしては，線形メカニカルシステムという構造を仮定しているが，各パラメータは未知であるということである．このシステムに対して，目標の出力$y_d(t)$を区間［0，T］（T：終了時刻）で与える．まず，このシステムに対して，1回目の試行時に適当な入力$u_1(t)$をシステムに与える．このときシステム応答を，$x_1(t)$とすれば，（2－1）式は，次のように表現される．

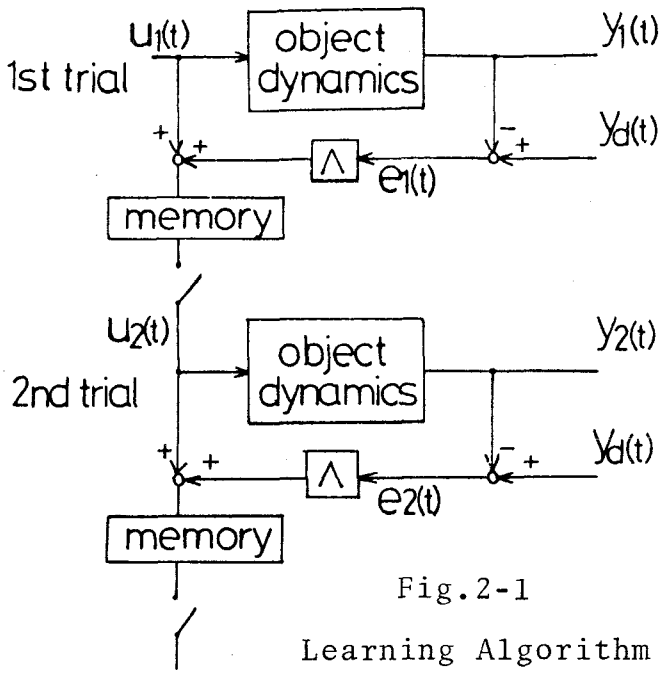


Fig.2-1
Learning Algorithm

$$R\ddot{x}_1(t)+Q\dot{x}_1(t)+Px_1(t)=u_1(t)$$
$$y_1(t)=\dot{x}_1(t) \qquad (2-2)$$

ただし，初期条件は，

$$\dot{x}_1(0)=y_d(0) \quad , \quad x_1(0)=x^0 \qquad (2-3)$$

と設定するものとする．ここに，x^0は位置パターン$x(t)$を目標位置パターン$x_d(t)$

$$x_d(t)=\int_0^t y_d(\tau)d\tau+x^0 \qquad (2-4)$$

に一致させる初期値である．一般に$y_1(t)$は目標出力$y_d(t)$に一致しないので，誤差

$$e_1(t)=y_d(t)-y_1(t) \qquad (2-5)$$

が存在する．そこで　2回目の入力$u_2(t)$を

$$u_2(t)=u_1(t)+\Lambda e_1(t) \qquad (2-6)$$

と構成する．ここに，行列Λは，適当な正定対称行列である．このアルゴリズムをFig．２－１に示す．以下この操作を繰り返すわけであるが，一般に試行回数をｋとするとｋ回目の試行時にはシステムが，

$$
\begin{aligned}
&R\ddot{x}_k(t)+Q\dot{x}_k(t)+Px_k(t)=u_k(t)\\
&\dot{x}_k(0)=y_d(0),\qquad x_k(0)=x^0\\
&y_k(t)=\dot{x}_k(t)
\end{aligned}
\tag{2-7}
$$

と表わされる．また，入力は，

$$
\begin{aligned}
&u_{k+1}(t)=u_k(t)+\Lambda e_k(t)\\
&e_k(t)=y_d(t)-y_k(t)
\end{aligned}
\tag{2-8}
$$

と表現される．このアルゴリズムによって誤差$e_k(t)$が減少またはゼロに収束し，目標出力$y_d(t)$が実現されるかどうかが問題となる．このことについては次の２つの定理が成立する．それらを示す前に，以下の２つの事実に注意する．

（注意１）

（２－７）式，（２－８）式より，ｋ＋１回目の試行時には，

$$
\begin{aligned}
R\ddot{x}_{k+1}(t)+Q\dot{x}_{k+1}(t)+Px_{k+1}(t)&=u_{k+1}(t)\\
&=u_k(t)+\Lambda e_k(t)
\end{aligned}
\tag{2-9}
$$

が成立していることがわかる．（２－９）式から（２－７）式を差し引くことにより次の関係が得られる．

$$
R\ddot{a}_k(t)+Q\dot{a}_k(t)+Pa_k(t)=\Lambda e_k(t) \tag{2-10}
$$

ただし，$a_k(t)=x_{k+1}(t)-x_k(t)$　である．

（注意２）

（２－１０）式における　$a_k(t)$については，

$$\begin{aligned}\dot{a}_k(t) &= \dot{x}_{k+1}(t) - \dot{x}_k(t) \\ &= y_{k+1}(t) - y_k(t) \\ &= y_d(t) - y_k(t) - (y_d(t) - y_{k+1}(t)) \\ &= e_k(t) - e_{k+1}(t) \end{aligned} \qquad (2-11)$$

が成立する．

定理２．１

初回の入力$u_1(t)$及び目標出力$y_d(t)$を区間［０，Ｔ］で連続な関数で構成し，条件

$2Q-\Lambda>0$　が満足されるとき，誤差$e_k(t)$は次の意味で減少する．

$$S_{k+1} \leqq S_k \qquad (2-12)$$

ここに　S_kは次の式によって定義される値である．

$$S_k = \int_0^T e_k^T(t)\,\Lambda\, e_k(t)dt \qquad (2-13)$$

証明

（２－１１）式より，S_{k+1}は，

$$S_{k+1} = \int_0^T (e_k(t) - \dot{a}_k(t))^T \Lambda (e_k(t) - \dot{a}_k(t))dt$$

$$= S_k - 2\int_0^T \dot{a}_k^T(t)\,\Lambda\, e_k(t)dt + \int_0^T \dot{a}_k^T(t)\,\Lambda\, \dot{a}_k(t)dt \qquad (2-14)$$

と表現される．さらに　（２－１０）式より次式が成立する．

$$S_k - S_{k+1} = 2\int_0^T \dot{a}_k^T(t)(R\ddot{a}_k(t) + Q\dot{a}_k(t) + P a_k(t))dt - \int_0^T \dot{a}_k^T(t)\Lambda\dot{a}_k(t)dt \qquad (2-15)$$

定理の初回の入力$u_1(t)$に対する仮定と（２－２）式より，$\dot{x}_1(t)$はC^1関数となり，$y_d(t)$が連続関数であることより，誤差$e_1(t)$は連続関数となる．以下同様の繰り返しにより一般に$e_k(t)$は，連続関数となる．このことは，（２－１０）式から$\ddot{a}_k(t)$，$\dot{a}_k(t)$が，各々連続及びC^1関数であることを示している．また，初期条件より，

$$\dot{a}_k(0) = \dot{x}_{k+1}(0) - \dot{x}_k(0) = y_d(0) - y_d(0) = 0$$
$$a_k(0) = x_{k+1}(0) - x_k(0) = x^0 - x^0 = 0 \qquad (2-16)$$

となることより，（２－１５）式は次のように部分積分によって書換られる．

$$S_k - S_{k+1} = \dot{a}_k^T(T) R \dot{a}_k(T) + a_k^T(T) P a_k(T) + \int_0^T \dot{a}_k^T(t)(2Q - \Lambda)\dot{a}_k(t)dt \qquad (2-17)$$

定理の条件より$2Q - \Lambda > 0$であるので，直ちに

$$S_k - S_{k+1} \geqq 0 \qquad (2-18)$$

が得られる．　　証明終

この定理はS_kという評価規範のもとに誤差が減少することを有限試行回数内で保証しているが，無限回の試行後に誤差がゼロに収束し，確実に目標出力$y_d(t)$が得られることは保証していない．そこで，無限回の試行後に目標出力が得られるかどうかについて次の定理を与える．これは，目標出力$y_d(t)$をC^1級に限定することによって目標出力が獲得されることを示すものである．

定理２．２

初回の入力$u_1(t)$を区間［０，Ｔ］で連続な関数を与え，かつ目標出力$y_d(t)$を区間［０，Ｔ］で連続的に微分可能（C^1級）で与えるとき，条件 $2Q - \Lambda > 0$が満足されれば誤差$e_k(t)$は任意に固定された時刻ｔでゼロに収束する．（$e_k(t)$はゼロに単純収束する）

証明

$u_1(t)$が連続であることから，$x_1(t) \in C^2$［０，Ｔ］となり，目標出力を$y_d(t) \in C^1$［０，Ｔ］と与えたので，誤差　$e_1(t) = y_d(t) - y_1(t) = y_d(t) - \dot{x}_1(t)$ はC^1級となる．また，このことは，（２－１０）式より$a_1(t) \in C^3$［０，Ｔ］を示している．次に，（２－８）式から得られる２回目の入力$u_2(t)$も連続関数となり，　$x_2(t) \in C^2$［０，Ｔ］となる．以下同様に，数学的帰納法によって，一般にｋ回目の誤差 $e_k(t)$はC^1級となり，$a_k(t)$はC^3級となる．従って，（２－１０）式の関係より，

$$R\dddot{a}_k(t) + Q\ddot{a}_k(t) + P\dot{a}_k(t) = \Lambda \dot{e}_k(t) \qquad (2-19)$$

を得る．またこのことは（２－１１）式より，

$$\ddot{a}_k(t) = \dot{e}_k(t) - \dot{e}_{k+1}(t) \qquad (2-20)$$

を得る．　ここで次の量を定義する．

$$J_k(t) = \int_0^t [e_k^T(\tau)\Lambda e_k(\tau) + \dot{e}_k^T(\tau)\Lambda \dot{e}_k(\tau)]d\tau \qquad (2-21)$$

（２－２１）式で定義される$J_{k+1}(t)$については（２－１１）式，（２－２０）式を用いることにより，

$$J_{k+1}(t) = \int_0^t [e_{k+1}^T(\tau)\Lambda e_{k+1}(\tau) + \dot{e}_{k+1}^T(\tau)\Lambda \dot{e}_{k+1}(\tau)]d\tau$$

$$= \int_0^t [(e_k(\tau) - \dot{a}_k(\tau))^T\Lambda(e_k(\tau) - \dot{a}_k(\tau)) + (\dot{e}_k(\tau) - \ddot{a}_k(\tau))^T\Lambda(\dot{e}_k(\tau) - \ddot{a}_k(\tau))]d\tau$$

$$= J_k(t) - 2\int_0^t [\dot{a}_k(\tau)^T\Lambda e_k(\tau) + \ddot{a}_k(\tau)^T\Lambda \dot{e}_k(\tau)]d\tau$$

$$+\int_0^t [\dot{a}_k^T(\tau)\Lambda\dot{a}_k(\tau)+\ddot{a}_k^T(\tau)\Lambda\ddot{a}_k(\tau)]d\tau \qquad (2-22)$$

を得る．さらに，（２－１０）式，（２－１９）式の関係を代入すると，

$$J_k(t)-J_{k+1}(t)=2\int_0^t \dot{a}_k(\tau)^T[R\ddot{a}_k(\tau)+Q\dot{a}_k(\tau)+Pa_k(\tau)] + \ddot{a}_k^T(\tau)[R\dddot{a}_k(\tau)+Q\ddot{a}_k(\tau)+P\dot{a}_k(\tau)]d\tau - \int_0^t [\dot{a}_k^T(\tau)\Lambda\dot{a}_k(\tau)+\ddot{a}_k^T(\tau)\Lambda\ddot{a}_k(\tau)]d\tau \qquad (2-23)$$

となる．ここで，（２－１６）式の初期条件より　$\dot{a}_k(0)=a_k(0)=0$　であり，また，$e_k(0)=y_d(0)-y_k(0)=0$　より，（２－１０）式において，行列Ｒ，Λが正定対称行列であることを考慮すれば，

$$\ddot{a}_k(0)=0 \qquad (2-24)$$

となることが理解される．これらのことに注意して，（２－２３）式を部分積分すると

$$J_k(t)-J_{k+1}(t)=\ddot{a}_k^T(t)R\ddot{a}_k(t)+\dot{a}_k^T(t)(R+P)\dot{a}_k(t)+a_k^T(t)Pa_k(t) + \int_0^t \ddot{a}_k^T(\tau)(2Q-\Lambda)\ddot{a}_k(\tau)d\tau+\int_0^t \dot{a}_k^T(\tau)(2Q-\Lambda)\dot{a}_k(\tau)d\tau \qquad (2-25)$$

を得る．Ｒ，Ｑ，Ｐ　正定対称行列であることと条件より　$2Q-\Lambda>0$であるので，区間［０，Ｔ］の任意の時刻において

$$J_k(t)-J_{k+1}(t)\geqq 0 \qquad (2-26)$$

となる．（２－２１）式の$J_k(t)$の定義より$J_k(t)\geqq 0$であるので$J_k(t)$は下に有界であり，かつ（２－２６）より単調減少であることよりある値J_∞に収束する．

また，（２－２５）式より任意の時刻ｔで

$$J_k(t) - J_{k+1}(t) \geqq \ddot{a}_k^T(t) R \ddot{a}_k(t) \geqq 0$$
$$J_k(t) - J_{k+1}(t) \geqq \dot{a}_k^T(t)(R+P)\dot{a}_k(t) \geqq 0$$
$$J_k(t) - J_{k+1}(t) \geqq a_k^T(t) P a_k(t) \geqq 0 \qquad (2-27)$$

が成立していることから，

$$\ddot{a}_k(t) \to 0 \quad , \dot{a}_k(t) \to 0 \quad , a_k(t) \to 0 \qquad (2-28)$$

となることがわかる．これは，（２－１０）式より直ちに

$$e_k(t) \to 0 \qquad (2-29)$$

を意味する．

証明終

定理２．２では　システムの出力を速度としているので，速度誤差$e_k(t)$の単純収束（任意の固定された時刻においてゼロに収束すること）が示された．しかしながら，さらに次の定理によって位置誤差については区間[０，Ｔ]で一様に収束することが示される．

定理２．３

定理２．２と同じ条件のもとに，区間［０，Ｔ］内でのすべての点で一様に

$$x_d(t) - x_k(t) \to 0 \quad , \qquad k \to \infty \qquad (2-30)$$

が成立する．

証明

（２－２７）式より　$\ddot{a}_k(t)$，$\dot{a}_k(t)$，$a_k(t)$　は一様有界となり，このことは（２－１０）式より，$e_k(t)$が一様有界であることを示している．従って，

$$E_k(t)=\int_0^t e_k(\tau)d\tau = x_d(t)-x_k(t) \qquad (2-31)$$

によって定義される位置誤差は一様有界かつ同程度連続となる．さらに，定理２．２の結果より，区間［０，Ｔ］で任意に固定された点ｔにおいて

$$E_k(t)\to 0 \quad , \qquad k\to\infty \qquad (2-32)$$

が成立する．以上の結果より$\{E_k(t)\}$は一様有界かつ同程度連続であり，ゼロに単純収束する．故に，Ascoli-Arzelaの定理より[18],[19]区間［０，Ｔ］内のすべての点ｔで一様に

$$E_k(t)=x_d(t)-x_k(t)\to 0 \quad , \qquad k\to\infty \qquad (2-33)$$

となり，ゼロに収束する．

証明終

２－３．結言

ここで証明したことは，２階の線形定係数微分方程式で表わされるメカニカルシステムに対して，システムの出力y_kを速度$\dot{x}_k$とするときに，

（１）初回の入力$u_1(t)$と目標出力パターン$y_d(t)$の各要素が，連続な関数で与えられるとき誤差$e_k(t)=y_d(t)-y_k(t)$はL_2ノルムの意味で減少すること，

（２）初回の入力$u_1(t)$と目標出力パターン$y_d(t)$の各要素がC^1級の滑かさをもつとき誤差$e_k(t)=y_d(t)-y_k(t)$が単純収束（任意に固定された時刻で収束する）の意味でゼロに収束すること，

（３）（２）と同じ条件のもとに，位置パターン$x_k(t)$は区間［０，Ｔ］で一様収束すること

以上３つの事柄である．それらの証明にはメカニカルシステムの構造的特徴である各係数の正定対称性を利用している．　また，ここでは出力を$\dot{x}$にとったが，状態変数の線形結合　$y=\alpha\dot{x}+\beta x$のような場合も同様に示される．これについては，以下の章について，より広い時変システムについて定式化されるのでここでは省略した．

第３章　学習制御方式の
システム論的考察

３－１．緒言

本章では本論文で提案する学習制御方式と従来の制御理論とのつながりを示す．そのために，制御対象は時不変線形システムとし，ラプラス変換に基づいて周波数領域で解析する．特に，従来のサーボ理論との関係を示すので，本章ではシステムは目標出力に対してfeedback系を構成するものとする．従って，feedback系への指令値を入力と見なし，それを毎回の試行ごとに修正するタイプの学習方式を考える．解析にあたっては，誤差がどのように変化するかを示すために誤差の遷移を示す誤差遷移関数行列Ｓを定義する．この誤差遷移関数行列Ｓが学習制御を実現するために満たすべき条件は回路理論における有界実性そのものであり，Ｓを感度行列と見たときには感度を減少させる条件に対応していることから，最適レギュレータ構成条件と同じであることが示される．これらの関係からある状態量の組合せによる出力については誤差の減少を保証するfeedback gainをRiccatiの方程式を解くことにより求めることができる．また，特に機械系の場合にはその構造的特徴から，状態量のすべてについて誤差の減少を保証する学習制御が構成できることが示される．最後に，この制御方式による幾つかのシミュレーション結果からfeedback gainや評価関数が学習制御の性能におよぼす影響について考える．

３－２．周波数領域における解析

ここで取り扱う対象システムは，入出力数が同じであるとする．また各試行において運動開始の初期時刻ではシステムの応答と目標パターンの状態量が一致しているものとする．そのような仮定のもとにシステムをラプラス変換で表現し，学習アルゴリズムをブロック図に示すとFig．３－１となる．ただし図中の文字は下記に示すものである．

$Y_d(s) \in R^{rx1}$　　目標軌道

$Yck(s) \in R^{rx1}$　　ｋ回目試行時のサーボ系への入力

$Yk(s) \in R^{rx1}$　　ｋ回目試行時のシステム応答

$Ek(s) \in R^{rx1}$　　ｋ回目試行時の誤差　　$E_k(s)=Y_d(s)-Y_{k(s)}$

$F(s) \quad \in R^{rxr}$　　開ループ伝達関数行列

$H(s) \quad \in R^{rxr}$　　閉ループ伝達関数行列

M　　記憶システム

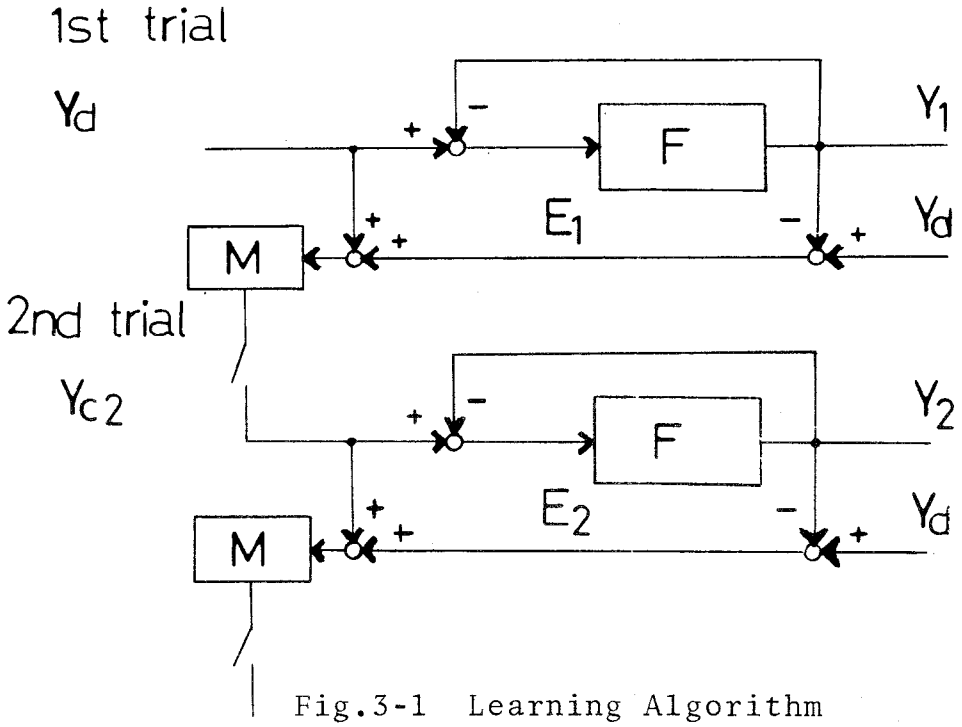


Fig.3-1　Learning Algorithm for Feedback Systems

まず１回目試行時の誤差$E_1(s)$は，

$$E_1(s) = Y_d(s) - Y_1(s)$$
$$= Y_d(s) - H(s)Y_d(s) \qquad (3-1)$$

となる．つぎに２回目の試行時にサーボ系に与える入力を，

$$Y_{c2}(s) = Y_d(s) + E_1(s) \qquad (3-2)$$

と構成する．２回目の試行時に$Y_{c2}(s)$を入力しても一般に誤差は存在すると考えられるので，それを$E_2(s)$とすると$E_1(s)$と$E_2(s)$にはつぎのような関係がある．

$$E_2(s) = Y_d(s) - Y_2(s)$$
$$= Y_d(s) - H(s)Y_{c2}(s)$$

$$= Y_d(s) - H(s)(Y_d(s) + E_1(s))$$
$$= (I - H(s))E_1(s) \qquad (3-3)$$

この関係は一般にｋ回目とｋ＋１回目の試行についても成立するので，

$$E_{k+1}(s) = (I - H(s))E_k(s) \qquad (3-4)$$

となる．ここで，

$$S(s) = I - H(s) \qquad (3-5)$$

と定義し，誤差遷移関数行列と呼ぶことにする．以後この関数行列Ｓ(s)によって誤差の収束性などを議論する．

学習制御の安定性（試行ごとに誤差が減少すること）を保証する条件を求めるために，評価規範としてつぎのようなものを考える．

$$\int_0^\infty {e_{k+1}}^T(t)\, e_{k+1}(t)dt < \int_0^\infty {e_k}^T(t)\, e_k(t)dt \qquad (3-6)$$

ただし$e_k(t)$は$E_k(s)$の逆ラプラス変換形である．ところで誤差の遷移は（３－４）式のようにラプラス変換形で与えられているので，（３－６）式の条件の周波数領域での表現を求める必要がある．誤差の時間関数については，

$$e_k(t) = 0 \qquad t \leqq 0 \qquad (3-7)$$

であり，これとParsevalの等式より，

$$\int_0^\infty {e_k}^T(t)\, e_k(t)dt = (1/2\pi)\int_{-\infty}^\infty {E_k}^*(i\omega)\, E_k(i\omega)d\omega \qquad (3-8)$$

となる．ただし＊印は共役転置を意味する．（３－８）式はｋ＋１回目の試行においても成立することから，（３－６）式の条件はつぎのように表現される．

$$\int_{-\infty}^{\infty} E_{k+1}^{*}(i\omega)E_{k+1}(i\omega)d\omega < \int_{-\infty}^{\infty} E_{k}^{*}(i\omega)E_{k}(i\omega)d\omega \qquad (3-9)$$

これに（３－４），（３－５）式の関係を代入すると次式を得る．

$$\int_{-\infty}^{\infty} E_{k}^{*}(i\omega)(I-S^{*}(i\omega)S(i\omega))E_{k}(i\omega)d\omega > 0 \qquad (3-10)$$

これより条件は，

$$I-S^{*}(i\omega)S(i\omega) \geqq 0 \qquad (3-11)$$

となる．ただしすべての周波数ωで等号が成立する場合は除く．

３－３．有界実性、感度、最適レギュレータ

条件（３－１１）式は、回路理論でよく知られた有界実性（Bounded Real）に対応しており，システムの正実性と関係している．一般に有界実性はつぎのように定義される．[22]すなわち，

（１）　Re(s)＞０　なるｓに対してＳ(s)は正則

（２）　ｓ（正実数）に対してＳ(s)が実数

（３）　Re(s)＞０　なるｓに対して，Ｉ－Ｓ*Ｓ≧０

を満足するときＳ(s)は有界実であると言われる．またこのときＳは分散行列（Scattering Matrix）と呼ばれる．したがって先に定義した誤差遷移関数行列Ｓ(s)が分散行列となり，システムが無損失でない場合に，学習は安定であると言える．（システムが無損失の場合は，　Ｉ－Ｓ*(iω)Ｓ(iω)＝０　に対応する）

つぎに，誤差遷移関数行列Ｓ(s)をシステム論的に考えると感度および最適レギュレータとの関係が明らかになる．それを示す前に，まず最適レギュレータと感度の関係について簡単に述べておく．

一般に（Ａ，Ｂ）が可制御であるシステム

$$\dot{x}(t) = A x(t) + B u(t) \qquad (3-12)$$

$$(A \in R^{n \times n} \quad B \in R^{n \times r})$$

に対して評価関数を，

$$J = \int_0^{\infty} [x^T(t) Q x(t) + u^T(t) R u(t)] dt \qquad (3-13)$$

$(R \in R^{r \times r} \quad R > 0, \ Q \in R^{n \times n} \quad Q \geqq 0 \quad (Q^{1/2}, A) \quad$ は可観測)

を最小にする最適制御入力$u^*(t)$は，

$$u^*(t) = -R^{-1} B^T P x(t) \qquad (3-14)$$

となる．ここで$P \in R^{n \times n}$は，Riccatiの行列方程式

$$PA + A^T P - P B R^{-1} B^T P + Q = 0 \qquad (3-15)$$

を満足する正定唯一解である．

逆に状態フィードバック

$$u(t) = -K x(t) \quad K \in R^{r \times n} \qquad (3-16)$$

が（３－１３）式で表現される評価関数の最適制御入力となるための必要十分条件は，

（１） $\dot{x}(t) = (A - BK) x(t)$ （３－１７）

が漸近安定であること

（２） $D^*(i\omega) R D(i\omega) \geqq R$ （３－１８）

である．ただし　$D(s) = I + L(s)$，$L(s) = K(sI - A)^{-1} B$であり，$D(s)$は還送差行列，$L(s)$には一巡伝達関数行列である．

この条件の１入力１出力系については，R.E.Kalmanによって，[23] また多入力多出力系については，B.D.O.Andersonによって証明された．[24] つぎに感度との関係では，１入力１出力系について最適レギュレータが感度の絶対値を減少させることをKalmanが示した．[23] また多入力多出力系については，J.B.CruzとW.A.Perkinsによって同様の関係が示された．[25] 彼等が定義した感度行列 W(s) は，

$$W(s)=D^{-1}(s) \qquad (3-19)$$

であり，１入力１出力系からの自然な拡張となっている．このW(s)を使って （３－１７）式を書き直すと，

$$R-W^{*}(i\omega)RW(i\omega)\geqq 0 \qquad (3-20)$$

となる．したがって最適レギュレータは（３－２０）式の規範の意味で感度を減少させている．また逆に感度を減少させる漸近安定な状態フィードバック制御は，（３－１３）式の形式の評価関数に対する最適レギュレータとなっている．

これらの準備のもとに、学習の安定条件（３－１１）式との関係を考える．まずここで考える制御対象システムについて状態変数はすべて観測できるものとする．さらにシステム自体には零点がないものとする．つぎに出力を $y_k(t)=Kx_k(t)$，目標軌道を $y_d(t)=Kx_d(t)$ と置く．ただし，K (rxn) は重み行列Q，Rに対する最適フィードバック行列であり，$y_k(t)$，$y_d(t)$は$Y_k(t)$，$Y_d(t)$の逆ラプラス変換形である．このときシステムは，Fig.３－２に示されるような実現となる．

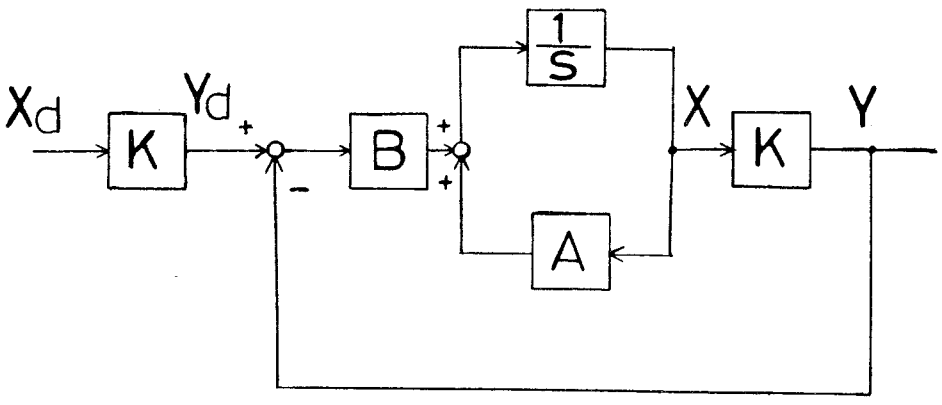


Fig.3-2 System Structure

したがって，Fig．３－１で示されるＦ(s)の部分は，

$$F(s)=L(s)=K(sI-A)^{-1}B \qquad (3-21)$$

となる．このことから誤差遷移関数行列Ｓ(s)は，

$$\begin{aligned} S(s)&=I-H(s)\\ &=I-(I+L(s))^{-1}L(s)\\ &=(I+L(s))^{-1}\\ &=D^{-1}(s)\\ &=W(s) \end{aligned} \qquad (3-22)$$

となるので，前述の感度行列と一致する．ここで（３－１３）式において，Ｒ＝Ｉと置けば，学習の安定条件（３－１１）式と最適レギュレータ構成条件（３－２０）式は同じものとなる．このことから評価関数（３－１３）式においてＲ＝Ｉ，Ｑ≧０　なる適当な行列を重みとして，最適制御入力のフィードバックシステムは試行ごとに誤差を減少させる学習安定なシステムになっていることがわかる．

ところで，ここでは$y_k(t)=Kx_k(t)$と置いているために誤差$e_k(t)$は，

$$\begin{aligned} e_k(t)&=y_d(t)-y_k(t)\\ &=K(x_d(t)-x_k(t)) \end{aligned} \qquad (3-23)$$

となっている．したがって学習が安定である場合，この誤差　$e_k(t)$がゼロに収束するので，

$$e_\infty(t)=y_d(t)-y_\infty(t)=K(x_d(t)-x_\infty(t))=0 \qquad (3-24)$$

を意味する．ところが，一般には，Ｋのゼロ空間に属する要素については，必ずしも$x_d(t)=x_\infty(t)$を意味しないという不都合が生じる．しかしながら本論文で取り扱っているようなロボットなどのメカニカルシステムではつぎの理由から$x_d(t)=x_\infty(t)$となっていることがわかる．

まず，線形メカニカルシステムを

$$C_1\ddot{\theta}(t)+C_2\dot{\theta}(t)+C_3\theta(t)=u(t) \qquad (3-25)$$

とする．ただし，C_1，C_2，$C_3 \in R^{rxr}$

特に，C_1は正定対称行列とする．このシステムを（３－１２）式の形式で表現すると，

$$\dot{x}(t)=Ax(t)+Bu(t) \qquad (3-25)$$
$$y(t)=Kx(t)$$

となる．ただし，

$$x(t)=(\theta^T(t),\ \dot{\theta}^T(t))^T,$$

$$A=\begin{bmatrix} 0 & , & I \\ -C_1^{-1}C_3, & & -C_1^{-1}C_2 \end{bmatrix}, \qquad B=\begin{bmatrix} 0 \\ C_1^{-1} \end{bmatrix}$$

であり，行列Kは（３－１４）式においてR＝Ｉと置き，リッカチ方程式の解をＰとしたとき，

$$K=B^TP$$
$$=\begin{bmatrix} 0 \\ C_1^{-1} \end{bmatrix}^T \begin{bmatrix} P_{11}, & P_{12} \\ P_{12}^T, & P_{22} \end{bmatrix}, \qquad P=\begin{bmatrix} P_{11}, & P_{12} \\ P_{12}^T, & P_{22} \end{bmatrix} \quad P_{11},\ P_{12},\ P_{22} \in R^{rxr}$$
$$=(C_1^{-1}P_{12}^T,\ C_1^{-1}P_{22}) \qquad (3-27)$$

によって与えられる．ここで，

$$x_d(t)-x_\infty(t)=\begin{bmatrix}\theta_d(t)-\theta_\infty(t)\\ \dot{\theta}_d(t)-\dot{\theta}_\infty(t)\end{bmatrix}=\begin{bmatrix}Z(t)\\ \dot{Z}(t)\end{bmatrix} \qquad (3-28)$$

と置くと（3－24）式，（3－27）式より，

$$C_1^{-1}P_{12}^{T}Z(t)+C_1^{-1}P_{22}\dot{Z}(t)=0 \qquad (3-29)$$

となる．初期条件に関しては，$\theta_d(0)=\theta_k(0)$，　$\dot{\theta}_d(0)=\dot{\theta}_k(0)$，$^{\forall}k$であるから，$Z(0)=\dot{Z}(0)=0$　であるので（2－29）式の解は，$Z(t)=0$となる．このことは，$x_d(t)=x_\infty(t)$　を意味し，位置と速度に相当する状態変数は各試行ごとにゼロに近づいていくことが保証される．

3－4　シミュレーション結果

ここでは，システムパラメータの変化が学習制御に与える影響を見るために，次のような簡単な１自由度モデルを考える．

$$a_0\ddot{\theta}(t)+a_1\dot{\theta}(t)+a_2\theta(t)=k_2(\theta_d(t)-\theta(t))+k_1(\dot{\theta}_d(t)-\dot{\theta}(t)) \qquad (3-30)$$

ただし，$\theta(t)$:システム応答，$\theta_d(t)$:目標位置パターン，a_0，a_1，a_2:システムパラメータ，k_1，k_2：feedback gain　とする．また，初期条件　$\theta_d(0)=0, \dot{\theta}_d(0)=0$
終端条件　$\theta_d(1)=1, \dot{\theta}_d(1)=0$　を満足する目標パターンを，

$$\theta_d(t)=-2t^3+3t^2 \qquad (3-31)$$

と与えた．計算の時間きざみは，0.005秒であり，パラメータは，$a_0=1$，$a_1=2$，$a_2=1$として位置feedback gainを変化させた場合と速度feedback gainを変化させた場合についてシミュレーションを行なった．各々の結果をFig.３－３，Fig.３－４に示す．

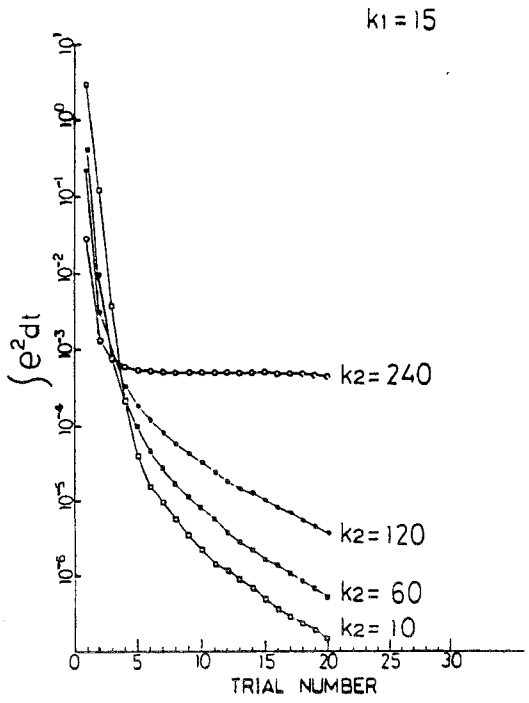


Fig.3-3 Simulation Result
Effect of Gain k_2

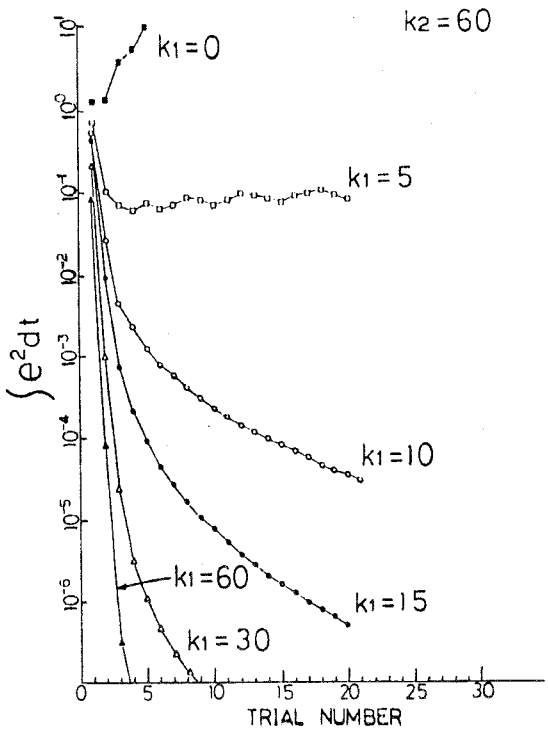


Fig.3-4 Simulation Result
Effect of Gain k_1

これらの結果から位置feedback gainの増加は学習の効率を悪くする方向に働き，これと反対に速度feedback gainの増加は学習の効率を良くする方向に働くことがわかる．そこでこれらの結果を詳しく考察するために，このシステムについての学習の安定条件（３－１１）式を求めることにする．まず伝達関数 H(s)は，

$$H(s) = \frac{k_2 + k_1 s}{a_0 s^2 + (a_1 + k_1) s + a_2 + k_2} \qquad (3-32)$$

となるので，誤差遷移関数S(s)は，

$$S(s) = \frac{a_0 s^2 + a_1 s + a_2}{a_0 s^2 + (a_1 + k_1) s + a_2 + k_2} \qquad (3-33)$$

となる．従って，学習の安定条件の式は

$$1 - S^*(i\omega) S(i\omega)$$

$$= \frac{[(2a_1+k_1)\ k_1-2a_0k_2]\omega^2+k_2(k_2+2a_2)}{(a_2+k_2-a_0\omega^2)^2+(a_1+k_1)^2\omega^2} \qquad (3-34)$$

と表現される．ここで（３－３４）式の正負を決定する分子に着目すると，パラメータはすべて正であることから，

$$(2a_1+k_1)k_1-2a_0k_2>0 \qquad (3-35)$$

が成立するとき（３－３４）式はすべての周波数に対して正となり，学習は安定となる．このとき速度feedback gain k_1が小さ過ぎると（３－３５）式は成立せず，つぎの周波数帯域で不安定（学習の意味での不安定）になる．

$$\omega \geqq k_2(k_2+2a_2)\ /[2a_0k_2-k_1\ (2a_1+k_1)] \qquad (3-36)$$

以上のことより速度feedback gainを十分大きくとれば，$\omega=\infty$を除くすべての周波数帯域で学習の安定性が保証されることがわかる．Fig.３－３より同じ値のfeedback gainでは始めの回の誤差が大きくなったとしても，位置feedback gainを小さくするほうが何回かの試行後には誤差が小さくなるということも理解できる．

次に最適レギュレータを利用した学習制御システム構成に関するシミュレーション結果を示す．まず（３－３０）式で与えられたシステムを次のように表現する．

$$\dot{x}(t)=Ax(t)+bu(t) \qquad (3-37)$$

ただし，

$$x(t)=(\theta(t),\ \dot{\theta}(t))^T,$$

$$A=\begin{bmatrix} 0 & , & I \\ -a_2/a_0, & & -a_1/a_0 \end{bmatrix}, \qquad B=\begin{bmatrix} 0 \\ -1/a_0 \end{bmatrix}$$

ここで問題は学習が安定に行なわれるように状態feedback gain（k_1，k_2）を求めることである．そこで評価関数Jを

$$J=\int_0^{\infty}[x^T(t)Qx(t)+u^2(t)]dt \qquad (3-38)$$

とする．ただし，

$$Q=\begin{bmatrix} q_1, & 0 \\ 0, & q_2 \end{bmatrix} \qquad q_1>0, \quad q_2>0$$

である．

このとき最適制御入力$u^*(t)$は，

$$u^*(t)=-b^TPx(t) \qquad (3-39)$$

となる．ここで$P\in R^{2x2}$は，Riccatiの行列方程式

$$PA+A^TP-Pbb^TP+Q=0 \qquad (3-40)$$

を満足する正定唯一解である．そこでq_1，q_2を適当に与えて（3－40）式のRiccati方程式を有本－ポッターの方法[26]により解を求め，最適feedback gain（k_1，k_2）を求めた．これらのパラメータを用いたシミュレーション結果をFig.3－5，Fig.3－6に示す．$\dot{\theta}$に関する重みq_2を大きくすると飛躍的に誤差の収束は早くなる．一方，θに関する重みq_1を大きくしても自動的に学習の安定条件が満足されるfeedback gainが得られているので，単に位置feedback gainのみを大きくしたときのような学習性能の劣化は生じないことがわかる．

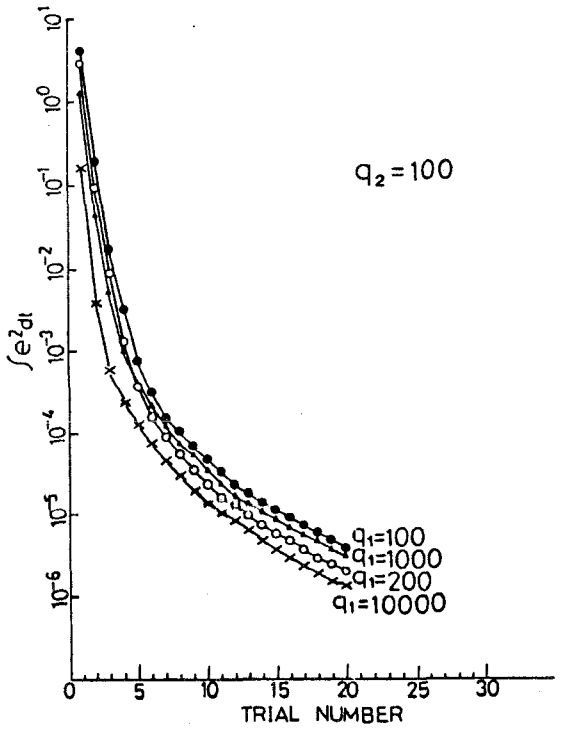


Fig.3-5 Simulation Result
Effect of Weight q_1

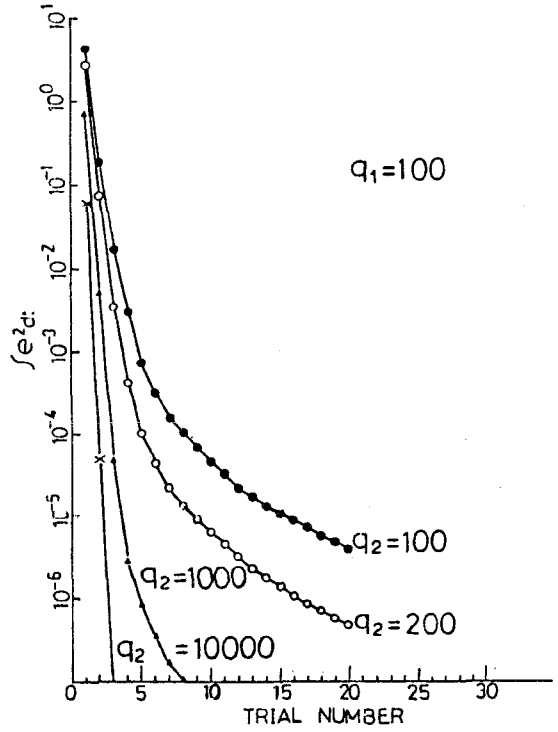


Fig.3-6 Simulation Result
Effect of Weight q_2

3 — 5 結言

時不変線形システムに関する学習制御方式を周波数領域で定式化し，回路理論の有界実性，システムの感度及び最適レギュレータとの関係を示した．その結果誤差を減少させるfeedback gainの設定は最適レギュレータ問題として，Riccati方程式を解くことによって得られることを明らかにした．

第４章　時変線形メカニカルシステムに対する学習制御

４－１　緒言

２章では制御対象の係数が時間的に変化しない線形システムを対象として論じたが，実用的にはより広いクラスのシステムに対して学習方式の有効性を示すことが望ましい．例えば，ロボットマニピュレータの運動では空間内のある固定された点の近傍の運動に限れば，時不変線形システムと見なせるが，広範囲な空間での運動はロボットダイナミクスの持つ非線形性により，時不変線形システムとして取り扱えない．ところが，ある目標運動(時間関数として目標軌道を与える)の近傍にロボットの実際の運動がとどまるとすれば，ロボットダイナミクスは，係数が時変の線形システムと見なすことができる．その線形化についての詳細は，８章に譲るとして，本章では軌道まわりで線形化されたロボットマニピュレータに適用する目的として，時変線形メカニカルシステムに対して学習方式を定式化し，誤差の収束性を議論する．その際２章では，出力を速度信号のみによって構成したが，本章では位置信号をも含めた形についても取り扱う．

４－２　誤差の収束性

４－２－１　速度信号による入力の修正

次のような時変システムを制御対象と考える．

$$R(t)\ddot{x}(t)+Q(t)\dot{x}(t)+P(t)x(t)=u(t)$$
$$y(t)=\dot{x}(t) \qquad (4-1)$$

２章の場合と同様に入力$u\in R^n$，出力$y\in R^n$．$R(t)$，$Q(t)$，$P(t)\in R^{nxn}$であり，$R(t)$，$Q(t)$，$P(t)$は時間区間［０，Ｔ］で正定対称行列であるとする．また行列の各要素は連続的に微分可能（$\in C^1$［０，Ｔ］)であるとする．さらに，この場合も第２章において述べたように，上述のようなシステムの構造は仮定するが，パラメータ自身

については未知であるとする．このような制御対象に与えられた目標出力（目標速度パターン）$y_d(t)$に対して，（2－7）式，及び（2－8）式と同じ繰り返し操作を実行する．即ち　この場合システム応答は，

$$\begin{aligned} &R(t)\ddot{x}_k(t)+Q(t)\dot{x}_k(t)+P(t)x_k(t)=u_k(t) \\ &\dot{x}_k(0)=y_d(0), \quad x_k(0)=x^0 \\ &y_k(t)=\dot{x}_k(t) \end{aligned} \qquad (4-2)$$

となる．ここに　x^0は位置x（t）のパターンを目標位置パターン$x_d(t)$に一致させる初期値である．また，入力の修正は，

$$\begin{aligned} &u_{k+1}(t)=u_k(t)+\Lambda e_k(t) \\ &e_k(t)=y_d(t)-y_k(t) \end{aligned} \qquad (4-3)$$

となる．ただし，Λは適当な正定対称行列である．このアルゴリズムによって，誤差が減少するかどうかが問題となる．これについては次の定理によって示される．

定理４．１

初回の入力$u_1(t)$と目標出力$y_d(t)$の各要素を，連続な関数で構成すると，誤差$e_k(t)$は次の意味で減少する．

$$S_{k+1} \leqq S_k \qquad (4-4)$$

ただし，

$$S_k=\int_0^T e^{-\rho t} e_k^T(t)\Lambda e_k(t)dt \qquad \rho：適当な正数$$

ここに，評価関数の$e^{-\rho t}$は時間の始めの方に大きな重みをかけることを意味している．

証明

第２章の（注意２）と同様に，$a_k(t)=x_{k+1}(t)-x_k(t)$　とするとき，

$\dot{a}_k(t) = e_k(t) - e_{k+1}(t)$　となることから次の関係が得られる．

$$S_{k+1} = \int_0^T e^{-\rho t} e_{k+1}{}^T(t) \Lambda e_{k+1}(t) dt$$

$$= \int_0^T e^{-\rho t} (e_k(t) - \dot{a}_k(t))^T \Lambda (e_k(t) - \dot{a}_k(t)) dt$$

$$= S_k - 2 \int_0^T e^{-\rho t} \dot{a}_k{}^T(t) \Lambda e_k(t) dt + \int_0^T e^{-\rho t} \dot{a}_k{}^T(t) \Lambda \dot{a}_k(t) dt \qquad (4-5)$$

さらに，この場合もｋ＋１回目の応答の式から，ｋ回目の応答の式を差し引くことにより

$$R(t) \ddot{a}_k(t) + Q(t) \dot{a}_k(t) + P(t) a_k(t) = \Lambda e_k(t) \qquad (4-6)$$

を得る．（４－６）式を代入することによって，（４－５）式は

$$S_k - S_{k+1} = 2 \int_0^T e^{-\rho t} \dot{a}_k{}^T(t) [R(t) \ddot{a}_k(t) + Q(t) \dot{a}_k(t) + P(t) a_k(t)] dt$$

$$- \int_0^T e^{-\rho t} \dot{a}_k{}^T(t) \Lambda \dot{a}_k(t) dt \qquad (4-7)$$

となる．定理２．１の証明と同様の議論から，$\ddot{a}_k(t)$，$\dot{a}_k(t)$の連続性が保証され，また初期条件も同様に　$\dot{a}_k(0) = a_k(0) = 0$となることから，（４－７）式は次の様に部分積分によって書換えられる．

$$S_k - S_{k+1} = e^{-\rho T} [\dot{a}_k{}^T(T) R(T) \dot{a}_k(T) + a_k{}^T(T) P(T) a_k(T)]$$

$$+ \int_0^T e^{-\rho t} \dot{a}_k{}^T(t) [\rho R(t) + 2Q(t) + \dot{R}(t) - \Lambda] \dot{a}_k(t) dt$$

$$+\int_0^T e^{-\rho t} a_k^T(t)[\rho P(t)+\dot{P}(t)]a_k(t)dt \qquad (4-8)$$

$R(t)$，$P(t)$，$Q(t)$が区間［０，Ｔ］で正定対称行列であることから適当な正値スカラ量ρを設定することにより（４－８）式は非負となる．このことは直ちに

$$S_{k+1} \leqq S_k \qquad (4-9)$$

を示している．　　証明終

第２章での議論と同様に，誤差のゼロ収束については次の２つの定理が与えられる．

定理４．２

初回の入力$u_1(t)$の各要素を連続関数で与え，かつ　目標出力$y_d(t)$の各要素を連続的に微分可能な関数で与えるとき適当な正値スカラ量ρのもとに，$e_k(t)$は単純収束する．

証明

仮定より定理２．２における議論と同様に，

$$(d/dt)(R(t)\ddot{a}_k(t)+Q(t)\dot{a}_k(t)+P(t)a_k(t))=\Lambda\dot{e}_k(t) \qquad (4-10)$$

という関係が得られる．ただし，$a_k(t)=x_{k+1}(t)-x_k(t)$．ここで，次の量を定義する（ただし時刻ｔは区間［０，Ｔ］で任意に固定されたものとする．）

$$J_k(t)=\int_0^t e^{-\rho\tau}[e_k^T(\tau)\Lambda e_k(\tau)+\dot{e}_k^T(\tau)\Lambda\dot{e}_k(\tau)]d\tau \qquad (4-11)$$

（４－１１）式で定義される$J_{k+1}(t)$について，$\dot{a}_k(t)=e_k(t)-e_{k+1}(t)$，

$\ddot{a}_k(t) = \dot{e}_k(t) - \dot{e}_{k+1}(t)$　という関係より次式を得る．

$$J_{k+1}(t) = J_k(t) - 2\int_0^t e^{-\rho\tau}[\dot{a}_k^T(\tau)\Lambda e_k(\tau) + \ddot{a}_k^T(\tau)\Lambda \dot{e}_k(\tau)]d\tau$$

$$+\int_0^t e^{-\rho\tau}[\dot{a}_k^T(\tau)\Lambda \dot{a}_k(\tau) + \ddot{a}_k^T(\tau)\Lambda \ddot{a}_k(\tau)]d\tau \qquad (4-12)$$

次に，$e_k(0)=0$　及び　$a_k(0)=\dot{a}_k(0)=0$より，（4－6）式を考慮すれば$\ddot{a}_k(0)=0$　であることがわかる．このことより，（4－12)式に（4－6)式，(4－10）式を代入して計算することにより次の関係が得られる．

$$J_k(t) - J_{k+1}(t) = e^{-\rho t}[\ddot{a}_k^T(t)R(t)\ddot{a}_k(t) + \dot{a}_k^T(t)(R(t)+P(t))\dot{a}_k(t) + a_k^T(t)P(t)a_k(t)]$$

$$+\int_0^t e^{-\rho\tau}\ddot{a}_k^T(\tau)[\rho R(t) + Q(t) + \dot{R}(t) - \Lambda]\ddot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}\dot{a}_k^T(\tau)[\rho R(\tau) + \rho P(\tau) + 2Q(\tau) - \dot{R}(\tau) - \dot{P}(\tau) - \Lambda]\dot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}a_k^T(\tau)[\rho P(\tau) + \dot{P}(t)]a_k(\tau)d\tau$$

$$+2\int_0^t e^{-\rho\tau}\ddot{a}_k^T(\tau)\dot{P}(\tau)a_k(\tau)d\tau$$

$$+2\int_0^t e^{-\rho\tau}\ddot{a}_k^T(\tau)\dot{Q}(\tau)\dot{a}_k(\tau)d\tau \qquad (4-13)$$

ここで，次の不等式が成立することに注意する．

$$\int_0^t e^{-\rho\tau}[\Phi\ddot{a}_k(\tau)-\Phi^{-1}\dot{P}(\tau)a_k(\tau)]^T[\Phi\ddot{a}_k(\tau)-\Phi^{-1}\dot{P}(\tau)a_k(\tau)]d\tau \geqq 0,$$

$$\int_0^t e^{-\rho\tau}[\Phi\ddot{a}_k(\tau)-\Phi^{-1}\dot{Q}(\tau)\dot{a}_k(\tau)]^T[\Phi\ddot{a}_k(\tau)-\Phi^{-1}\dot{Q}(\tau)\dot{a}_k(\tau)]d\tau \geqq 0, \quad (4-14)$$

ただし，$\Phi \in R^{nxn}$は適当な正定対称行列とする．（4－14）式より，直ちに，

$$\int_0^t e^{-\rho\tau}\ddot{a}_k(\tau)^T\Psi\ddot{a}_k(\tau)d\tau + \int_0^t e^{-\rho\tau}a_k(\tau)^T\dot{P}(\tau)\Psi^{-1}\dot{P}(\tau)a_k(\tau)d\tau$$

$$\geqq 2\int_0^t e^{-\rho\tau}\ddot{a}_k(\tau)^T\dot{P}(\tau)a_k(\tau)d\tau$$

$$\int_0^t e^{-\rho\tau}\ddot{a}_k(\tau)^T\Psi\ddot{a}_k(\tau)d\tau + \int_0^t e^{-\rho\tau}\dot{a}_k(\tau)^T\dot{Q}(\tau)\Psi^{-1}\dot{Q}(\tau)\dot{a}_k(\tau)d\tau$$

$$\geqq 2\int_0^t e^{-\rho\tau}\ddot{a}_k(\tau)^T\dot{Q}(\tau)\dot{a}_k(\tau)d\tau \quad (4-15)$$

が成り立つことがわかる．ただし，$\Psi=\Phi\cdot\Phi$．（4－15）式の不等式より（4－13）式から次式が導かれる．

$$J_k(t)-J_{k+1}(t) \geqq e^{-\rho\tau}[\ddot{a}_k(t)^T R(t)\ddot{a}_k(t)$$

$$+\dot{a}_k(t)^T(R(t)+P(t))\dot{a}_k(t)+a_k(t)^T P(t)a_k(t)]$$

$$+\int_0^t e^{-\rho\tau}\ddot{a}_k(\tau)^T A_1(\tau)\ddot{a}_k(\tau)d\tau + \int_0^t e^{-\rho\tau}\dot{a}_k(\tau)^T B_1(\tau)\dot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}a_k(\tau)^T C_1(\tau)a_k(\tau)d\tau \quad (4-16)$$

ただし，

$$A_1(\tau)=\rho R(\tau)+2Q(\tau)+\dot{R}(\tau)-\Lambda-2\Psi,$$
$$B_1(\tau)=\rho(R(\tau)+P(\tau))+2Q(\tau)-\dot{R}(\tau)-\dot{P}(\tau)-\Lambda-\dot{Q}(\tau)\Psi^{-1}\dot{Q}(\tau),$$
$$C_1(\tau)=\rho P(\tau)-\dot{P}(\tau)-\dot{P}(\tau)\Psi^{-1}\dot{P}(\tau),$$

R(t)，P(t)，Q(t)が正定対称行列であることから適当なスカラ量ρを設定することにより，

$$A_1(\tau)>0,\ B_1(\tau)>0,\ C_1(\tau)>0 \qquad (4-17)$$

となる．従って，（4－16）式の右辺は正となり，$J_k(t)$が単調減少することが示される．また，（4－16）式より任意の時刻$t\in[0,T]$において次の関係も得られる．

$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}\ddot{a}_k(t)^T R(t)\ddot{a}_k(t)\geqq 0$$
$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}\dot{a}_k(t)^T (R(t)+P(t))\dot{a}_k(t)\geqq 0$$
$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}a_k(t)^T P(t)a_k(t)\geqq 0 \qquad (4-18)$$

数列$J_k(t)$は定義より非負（下に有界）であり，単調減少であるので任意の時刻$t\in[0,T]$で，ある値（$J_\infty(t)$）に収束する．このことと（4－18）式より　$k\to\infty$のとき

$$\ddot{a}_k(t)\to 0,\ \dot{a}_k(t)\to 0,\ a_k(t)\to 0 \qquad (4-19)$$

が任意の時刻$t\in[0,T]$について成立することがわかる．結局，（4－19）式と（4－6）式より，$k\to\infty$のときに　任意の時刻$t\in[0,T]$で，

$$e_k(t)\to 0 \qquad (4-20)$$

であることが示される．

証明終

定理４．３

定理４．２と同じ条件のもとに，区間［０，Ｔ］でのすべての点で一様に，

$$x_d(t)-x_k(t)\to 0 \quad , \qquad k\to\infty \tag{4-21}$$

が成立する．

証明は，定理２．３と同様に位置誤差 $x_d(t)-x_k(t)$は一様有界，同程度連続でありゼロに単純収束することを考慮すれば,Ascoli-Arzelaの定理が適用できることから明らかである．

系　４．１

定理４．１は行列Ｑ(t)，Ｐ(t)が正定対称でない場合にも成立する．

証明

この場合は，Ｑ(t)，Ｐ(t)に関する項について部分積分が計算できないので（４－８）式の代わりに次のように表現する．

$$\begin{aligned} S_k-S_{k+1}=\ & e^{-\rho T}\dot{a}_k^{T}(T)\,R(T)\,\dot{a}_k(T) \\ & +\int_0^T e^{-\rho t}\dot{a}_k^{T}(t)\left(\rho R(t)+2Q(t)-\dot{R}(t)-\Lambda\right)\dot{a}_k(t)dt \\ & +\int_0^T e^{-\rho t}\dot{a}_k^{T}(t)\,P(t)\,a_k(t)dt \end{aligned} \tag{4-22}$$

（４－１５）式と同様に次の不等式が得られることに注意する．

$$\int_0^T e^{-\rho t}\dot{a}_k^{T}(t)\Psi\dot{a}_k(t)dt+\int_0^T e^{-\rho t}a_k^{T}(t)\Psi^{-1}a_k(t)dt$$

$$\geqq 2\int_0^T e^{-\rho t}\dot{a}_k^T(t)\,a_k(t)dt \qquad (4-23)$$

ただし$\Psi \in R^{nxn}$正定対称行列．さらに（４－２３）式の右辺を計算して，

$$\int_0^T e^{-\rho t}\dot{a}_k^T(t)\Psi\,\dot{a}_k(t)dt - e^{-\rho T}a_k^T(T)\,a_k(T)$$

$$-\int_0^T e^{-\rho t}a_k^T(t)(\rho\, I-\Psi^{-1})\,a_k(t)dt \geqq 0 \qquad (4-24)$$

となる．この関係を用いると，（４－２２）式より次式が得られる．

$$S_k - S_{k+1} \geqq e^{-\rho T}(\dot{a}_k^T(T)\,R(T)\,\dot{a}_k(T) + a_k^T(T)\,a_k(T))$$

$$+\int_0^T e^{-\rho t}\dot{a}_k^T(t)(\rho\,R(t)+2Q(t)-\dot{R}(t)-\Lambda-\Psi)\,\dot{a}_k(t)dt$$

$$+\int_0^T e^{-\rho t}a_k^T(t)(\rho\, I-\Psi^{-1})\,a_k(t)dt \qquad (4-25)$$

この場合も十分大きな値のρを設定することにより（４－９）式が成立する．

証明終

系　４．２
定理４．２は行列Q(t)，P(t)が正定対称行列でない場合にも成立する．

証明

この場合（４－１３）式に代わって，

$$J_k(t) - J_{k+1}(t) = e^{-\rho t}(\ddot{a}_k^T(t)\,R(t)\,\ddot{a}_k(t) + \dot{a}_k^T(t)\,R(t)\,\dot{a}_k(t))$$

$$+\int_0^t e^{-\rho\tau}\,\ddot{a}_k^T(\tau)(\rho R(\tau)+2Q(\tau)+\dot{R}(\tau)-\Lambda)\ddot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}\,\dot{a}_k^T(\tau)(\rho R(\tau)+2Q(\tau)-\dot{R}(\tau)-\Lambda)\dot{a}_k(\tau)d\tau$$

$$+2\int_0^t e^{-\rho\tau}\,\ddot{a}_k^T(\tau)(\dot{Q}(\tau)+P(\tau))\dot{a}_k(\tau)d\tau$$

$$+2\int_0^t e^{-\rho\tau}\,\ddot{a}_k^T(\tau)\dot{P}(\tau)a_k(\tau)d\tau$$

$$+2\int_0^t e^{-\rho\tau}\,\dot{a}_k^T(\tau)P(\tau)a_k(\tau)d\tau \qquad (4-26)$$

と表現する．

（４－１６）式が得られたと同様にΨを適当な正定対称行列として（４－２６）より次式が得られる．

$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}(\ddot{a}_k^T(t)R(t)\ddot{a}_k(t)+\dot{a}_k^T(t)R(t)\dot{a}_k(t))$$

$$+\int_0^t e^{-\rho\tau}\,\ddot{a}_k^T(\tau)A_2(\tau)\ddot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}\,\dot{a}_k^T(\tau)B_2(\tau)\dot{a}_k(\tau)d\tau$$

$$+\int_0^t e^{-\rho\tau}\,a_k^T(\tau)C_2(\tau)a_k(\tau)d\tau$$

(4－27)

ここに，

$$A_2(\tau)=\rho R(\tau)+2Q(\tau)+\dot{R}(\tau)-\Lambda-2\Psi,$$
$$B_2(\tau)=\rho R(\tau)+2Q(\tau)-\dot{R}(\tau)-\Lambda$$
$$-(\dot{Q}(\tau)+P(\tau))^{T}\Psi^{-1}(\dot{Q}(\tau)+P(\tau))-\Psi,$$
$$C_2(\tau)=-\dot{P}^{T}(\tau)\Psi^{-1}\dot{P}(\tau)-P^{T}(\tau)\Psi^{-1}P(\tau),$$

（４－２７）式において，右辺の$C_2(\tau)$を含む項は正値とならないが，系４．１の（４－２４）式の関係を使うと，（４－２７）式は結局，

$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}(\ddot{a}_k^{T}(t)R(t)\ddot{a}_k(t)+\dot{a}_k^{T}(t)R(t)\dot{a}_k(t)+a_k^{T}(t)a_k(t))$$
$$+\int_0^t e^{-\rho\tau}\ddot{a}_k^{T}(\tau)A_3(\tau)\ddot{a}_k(\tau)d\tau$$
$$+\int_0^t e^{-\rho\tau}\dot{a}_k^{T}(\tau)B_3(\tau)\dot{a}_k(\tau)d\tau$$
$$+\int_0^t e^{-\rho\tau}a_k^{T}(\tau)C_3(\tau)a_k(\tau)d\tau \qquad (4-28)$$

となる．ただし，

$$A_3(\tau)=A_2(\tau)$$
$$B_3(\tau)=B_2(\tau)-\Psi$$
$$C_3(\tau)=\rho I-\Psi^{-1}+C_2(\tau)$$

である．従って，十分大きな値のρを設定することにより，

$$A_3(\tau)>0,\ B_3(\tau)>0,\ C_3(\tau)>0 \qquad (4-29)$$

となり，定理４．２と同様の結果が得られる．

証明終

系４．３

定理４．３は行列Q(t)，P(t)が正定対称でない場合にも成立する．

証明は定理４．３及び系４．２より明らかである．

４－２－２　位置信号と速度信号による入力の修正

対象システムとしては（４－２－１）で与えられるタイプのものを考える．一方，出力y(t)としては次のような位置と速度の線形結合を設定する．即ち全体のシステムは

$$R(t)\ddot{x}(t)+Q(t)\dot{x}(t)+P(t)x(t)=u(t)$$
$$y(t)=\alpha\dot{x}(t)+\beta x(t) \qquad (\alpha,\ \beta,\ \text{スカラ量}) \qquad (4-30)$$

と表現される．ただし，信号を正規化する目的でスカラ量α，βについては，次の条件を満たすものとする．

$$\alpha+\beta=1 \qquad 0<\alpha\leqq 1,\quad 0\leqq\beta<1 \qquad (4-31)$$

即ち，（４－３１）式は出力信号に速度信号が含まれることを必要としている．その理由については，以下の証明の中で明らかとなる．次に，位置の目標パターンを$x_d(t)$とすれば，目標出力$y_d(t)$は，

$$y_d(t)=\alpha\dot{x}_d(t)+\beta x_d(t) \qquad (4-32)$$

と表現できる．この目標出力とシステムの出力との差を誤差e(t)とすると，

$$e(t) = y_d(t) - y(t)$$
$$= \alpha(\dot{x}_d(t) - \dot{x}(t)) + \beta(x_d(t) - x(t)) \qquad (4-33)$$

となる．このようなシステムに対して，次の繰り返し操作を実行する．

$$R(t)\ddot{x}_k(t) + Q(t)\dot{x}_k(t) + P(t)x_k(t) = u(t)$$
$$\dot{x}_k(0) = \dot{x}_d(0) \quad , \quad x_k(0) = x_d(0)$$
$$u_{k+1}(t) = u_k(t) + \Lambda e_k(t) \qquad (4-34)$$

前節の定理及び系ではシステムの出力を速度としていたので誤差$e_k(t)$は速度誤差を意味していたが，$e_k(t)$を（4－33）式で定義されるものに置き換えれば，（4－34）式で与えられる学習プロセスに対して，それらの定理や系はすべて成立することが示される．その証明はほぼ前節のものと同じであるので，煩雑さを避けるために，系4．3に相当する最も広いクラスの対象（係数行列$R(t)$のみが正定対象であると仮定する場合）に適用できる定理だけを次に示す．

定理4．4

初回の入力$u_1(t)$の各要素を連続関数で与え，目標出力の各要素を連続的に微分可能な関数とすれば，$x_k(t)$は一様に$x_d(t)$に収束する．

証明

目標出力についての仮定より定理2．2の証明と同様の議論から$e_k(t)$はC^1級関数となる．このことから次の量が定義できる．

$$J_k(t) = \int_0^t e^{-\rho\tau}(e_k^T(\tau)\Lambda e_k(\tau) + \dot{e}_k^T(\tau)\Lambda \dot{e}_k(\tau))d\tau \qquad (4-35)$$

この場合$e_k(t)$と$e_{k+1}(t)$は（4－33）より，

$$e_k(t) - e_{k+1}(t) = y_d(t) - y_k(t) - (y_d(t) - y_{k+1}(t))$$
$$= \alpha(\dot{x}_{k+1}(t) - \dot{x}_k(t)) + \beta(x_{k+1}(t) - x_k(t))$$
$$= \alpha\dot{a}_k(t) + \beta a_k(t) \qquad (4-36)$$

（ただし，$a_k(t)=x_{k+1}(t)-x_k(t)$）となる．さらに，$e_k(t)\in C^1[0,T]$ より

$$\dot{e}_k(t)-\dot{e}_{k+1}(t)=\alpha\ddot{a}_k(t)+\beta\dot{a}_k(t) \qquad (4-37)$$

となることから，$J_{k+1}(t)$は次のように表現される．

$$\begin{aligned}J_{k+1}(t)=&\int_0^t e^{-\rho\tau}[(e_k(\tau)-\alpha\dot{a}_k(\tau)-\beta a_k(\tau))^T\Lambda(e_k(\tau)-\alpha\dot{a}_k(\tau)-\beta a_k(\tau))\\&+(\dot{e}_k(\tau)-\alpha\ddot{a}_k(\tau)-\beta\dot{a}_k(\tau))^T\Lambda(\dot{e}_k(\tau)-\alpha\ddot{a}_k(\tau)-\beta\dot{a}_k(\tau))]d\tau\\=&J_k(t)+G(t)-2\int_0^t e^{-\rho\tau}[(\alpha\dot{a}_k(\tau)+\beta a_k(\tau))^T\Lambda e_k(\tau)\\&+(\alpha\ddot{a}_k(\tau)+\beta\dot{a}_k(\tau))^T\Lambda\dot{e}_k(\tau)]d\tau\end{aligned} \qquad (4-38)$$

ただし，

$$\begin{aligned}G(t)=&\int_0^t e^{-\rho\tau}[(\alpha\dot{a}_k(\tau)+\beta a_k(\tau))^T\Lambda(\alpha\dot{a}_k(\tau)+\beta a_k(\tau))\\&+(\alpha\ddot{a}_k(\tau)+\beta\dot{a}_k(\tau))^T\Lambda(\alpha\ddot{a}_k(\tau)+\beta\dot{a}_k(\tau))]d\tau\end{aligned}$$

定理４．２，４．３と同様の議論より，

$$\begin{aligned}&R(t)\ddot{a}_k(t)+Q(t)\dot{a}_k(t)+P(t)a_k(t)=\Lambda e_k(t)\\&(d/dt)(R(t)\ddot{a}_k(t)+Q(t)\dot{a}_k(t)+P(t)a_k(t))=\Lambda\dot{e}_k(t)\end{aligned} \qquad (4-39)$$

という関係が得られる．これを（４－３８）式に代入して，

$$\begin{aligned}J_k(t)-J_{k+1}(t)=&G(t)\\&-2\int_0^t e^{-\rho\tau}[(\alpha\dot{a}_k(\tau)+\beta a_k(\tau))^T(R\ddot{a}_k(\tau)+Q\dot{a}_k(\tau)+Pa_k(\tau))\end{aligned}$$

$$+(\alpha\ddot{a}_k(\tau)+\beta\dot{a}_k(\tau))^T(d/d\tau)(R\ddot{a}_k(\tau)+Q\dot{a}_k(\tau)+Pa_k(\tau))]d\tau \quad (4-40)$$

となる．初期条件は，$a_k(0)=\dot{a}_k(0)=\ddot{a}_k(0)=0$で与えられるので，行列$R(t)$が正定対称であることに注意して（４－４０）式を部分積分すると

$$\begin{aligned}J_k(t)-J_{k+1}(t)=e^{-\rho t}[&\ddot{a}_k^T(t)R(t)\ddot{a}_k(t)+\dot{a}_k^T(t)(\alpha R(t)+\rho\beta R(t))\dot{a}_k(t)\\&+a_k^T(t)\rho\beta R(t)a_k(t)+2\dot{a}_k^T(t)\beta R(t)\ddot{a}_k(t)\\&+2a_k^T(t)\beta R(t)\dot{a}_k(t)]\\+\int_0^t e^{-\rho\tau}[&\ddot{a}_k^T(\tau)T_1(\tau)\ddot{a}_k(\tau)\\&+\dot{a}_k^T(\tau)T_2(\tau)\dot{a}_k(\tau)+a_k^T(\tau)T_3(\tau)a_k(\tau)\\&+\dot{a}_k^T(\tau)T_4(\tau)\ddot{a}_k(\tau)+\ddot{a}_k^T(\tau)T_5(\tau)a_k(\tau)\\&+a_k^T(\tau)T_6(\tau)\dot{a}_k(\tau)]d\tau\end{aligned} \quad (4-41)$$

$$T_1(\tau)=(\rho R(\tau)+\dot{R}(\tau)+2Q(\tau)-\alpha\Lambda)\alpha-2\beta R(\tau)$$

$$\begin{aligned}T_2(\tau)=(&\rho R(\tau)-\dot{R}(\tau)+2Q(\tau)-\alpha\Lambda)\alpha+(\rho^2R(\tau)\\&+2P(\tau)-\rho\dot{R}(\tau)+2\dot{Q}(\tau)-2R(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda\end{aligned}$$

$$T_3(\tau)=(\rho^2R(\tau)+2P(\tau)-\rho\dot{R}(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda$$

$$T_4(\tau)=2(\dot{Q}(\tau)+P(\tau))^T\alpha+2\beta Q(\tau)$$

$$T_5(\tau)=2\alpha\dot{P}(\tau)$$

$$T_6(\tau)=2\alpha P^T(\tau)+2(-\dot{R}(\tau)+Q(\tau)+P^T(\tau))\beta$$

を得る．定理４．２の（４－１５）式と同様の不等式を用いることにより，次の関係が成立する．

$$J_k(t)-J_{k+1}(t)\geqq e^{-\rho t}\,[\,\ddot{a}_k^{T}(t)(\alpha R(t)-\beta\Gamma)\ddot{a}_k(t)$$
$$+\dot{a}_k^{T}(t)(\alpha R(t)+\rho\beta R(t)-\beta R(t)\Gamma^{-1}R(t)\Gamma)\dot{a}_k(t)$$
$$+a_k^{T}(t)(\rho\beta R(t)-\beta R(t)\Gamma^{-1}R(t))a_k(t)\,]$$
$$+\int_0^t e^{-\rho\tau}\,[\,\ddot{a}_k^{T}(\tau)S_1(\tau)\ddot{a}_k(\tau)+\dot{a}_k^{T}(\tau)S_2(\tau)\dot{a}_k(\tau)$$
$$+a_k^{T}(\tau)S_3(\tau)a_k(\tau)\,]\,d\tau \qquad (4-42)$$

ただし，

$$S_1(\tau)=(\rho R(\tau)+\dot{R}(\tau)+2Q(\tau)-\alpha\Lambda)\alpha-2\beta R(\tau)-\Gamma$$
$$-T_4^{T}(\tau)\Gamma^{-1}T_4(\tau)$$

$$S_2(\tau)=(\rho R(\tau)-\dot{R}(\tau)+2Q(\tau)-\alpha\Lambda)\alpha+(\rho^2R(\tau)+2P(\tau)-\rho\dot{R}(\tau)$$
$$+2\dot{Q}(\tau)-2R(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda-\rho\alpha\beta-\Gamma-T_6^{T}(\tau)\Gamma^{-1}T_6(\tau)$$

$$S_3(\tau)=(\rho^2R(\tau)+2P(\tau)-\rho\dot{R}(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda-\Gamma$$
$$-T_5^{T}(\tau)\Gamma^{-1}T_5(\tau)$$

R（t）が正定対称行列であるので十分大きなρをとることにより，（4－42）式の右辺は非負となる．従って，$\{J_k(t)\}$は下に有界で単調減少となるので，区間［0，T］で任意に固定された時刻tにおいて，ある値に収束する．さらに，（4－42）式より，

$$J_k(t)-J_{k+1}(t)\geqq\int_0^t e^{-\rho\tau}\,\ddot{a}_k^{T}(\tau)S_1(\tau)\ddot{a}_k(\tau)d\tau$$

$$J_k(t)-J_{k+1}(t)\geqq\int_0^t e^{-\rho\tau}\,\dot{a}_k^{T}(\tau)S_2(\tau)\dot{a}_k(\tau)d\tau$$

$$J_k(t)-J_{k+1}(t)\geqq\int_0^t e^{-\rho\tau}\,a_k^{T}(\tau)S_3(\tau)a_k(\tau)d\tau \qquad (4-43)$$

も成立することから，

$$\int_0^t \ddot{a}_k^T(\tau)\ddot{a}_k(\tau)d\tau \to 0, \qquad k\to\infty,$$

$$\int_0^t \dot{a}_k^T(\tau)\dot{a}_k(\tau)d\tau \to 0, \qquad k\to\infty,$$

$$\int_0^t a_k^T(\tau)a_k(\tau)d\tau \to 0, \qquad k\to\infty, \qquad (4-44)$$

となることがわかる．次に（４－３９）式より，

$$\int_0^t e_k^T(\tau)e_k(\tau)d\tau = \int_0^t \ddot{a}_k^T(\tau)R(\tau)\Omega^{-1}R(\tau)\ddot{a}_k(\tau)d\tau + \int_0^t \dot{a}_k^T(\tau)Q^T(\tau)\Omega^{-1}Q(\tau)\dot{a}_k(\tau)d\tau$$
$$+\int_0^t a_k^T(\tau)P^T(\tau)\Omega^{-1}P(\tau)a_k(\tau)d\tau + 2\int_0^t \ddot{a}_k^T(\tau)R(\tau)\Omega^{-1}Q(\tau)\dot{a}_k(\tau)d\tau$$
$$+2\int_0^t \ddot{a}_k^T(\tau)R(\tau)\Omega^{-1}P(\tau)a_k(\tau)d\tau + 2\int_0^t \dot{a}_k^T(\tau)Q^T(\tau)\Omega^{-1}P(\tau)a_k(\tau)d\tau \qquad (4-45)$$

となるに注意する．ただし$\Omega^{-1}=\Lambda^{-1}\Lambda^{-1}$とする．（４－１５）式と同様の議論によって適当な大きさのスカラ量δ_1，δ_2，δ_3 を設定することにより，次式を得る．

$$\delta_1\int_0^t \ddot{a}_k^T(\tau)a_k(\tau)d\tau + \delta_2\int_0^t \dot{a}_k^T(\tau)\dot{a}_k(\tau)d\tau + \delta_3\int_0^t a_k^T(\tau)a_k(\tau)d\tau$$
$$\geqq \int_0^t e_k^T(\tau)e_k(\tau)d\tau \geqq 0 \qquad (4-46)$$

このことと，（４－４４）式を考え合わせれば，容易に次のことが理解される．即ち，区間［０，Ｔ］で任意に固定されたｔについて

$$\int_0^t e_k^T(\tau)e_k(\tau)d\tau \to 0 \quad , \quad k\to\infty \qquad (4-47)$$

となる．一方，（４－３３）式より，

$$x_d(t)-x_k(t)=(1/\alpha)\int_0^t \exp((-\beta/\alpha)(t-\tau))e_k(\tau)d\tau \qquad (4-48)$$

となり，これにSchwarzの不等式を適用して次式を得る．

$$\begin{aligned}&(x_d(t)-x_k(t))^T(x_d(t)-x_k(t))\\ &\leqq (1/\alpha)\int_0^t \exp((-2\beta/\alpha)(t-\tau))d\tau \int_0^t e_k^T(\tau)e_k(\tau)d\tau\\ &\leqq \delta_4 \int_0^t e_k^T(\tau)e_k(\tau)d\tau \qquad (4-49)\end{aligned}$$

ただし，　$\delta_4=\{1-\exp((-2\beta/\alpha)T)\}/2\beta$　　とする．（４－４７）式と（４－４９）式より，任意に固定されたｔにおいて，

$$x_d(t)-x_k(t)\to 0 \quad , \quad k\to\infty \qquad (4-50)$$

となることがわかる．次に $\{J_k(t)\}$は下に有界で単調減少であるので次の値も区間［０，Ｔ］で一様有界である．

$$\int_0^t e^{-\rho\tau}\dot{e}_k^T(\tau)\Lambda\dot{e}_k(\tau)d\tau \qquad (4-51)$$

（４－５１）式は（４－３３）式を代入して計算することにより次のように書換られる．

$$\int_0^t e^{-\rho\tau}\dot{e}_k^T(\tau)\Lambda\dot{e}_k(\tau)d\tau=\alpha\beta e^{-\rho t}(\dot{x}_d(t)-\dot{x}_k(t))^T\Lambda(\dot{x}_d(t)-\dot{x}_k(t))$$

$$+\int_0^t e^{-\rho\tau}[\alpha^2(\ddot{x}_d(\tau)-\ddot{x}_k(\tau))^T\Lambda(\ddot{x}_d(\tau)-\ddot{x}_k(\tau))$$

$$+(\beta^2+\rho\alpha\beta)(\dot{x}_d(\tau)-\dot{x}_k(\tau))^T\Lambda(\dot{x}_d(\tau)-\dot{x}_k(\tau))]d\tau \qquad (4-52)$$

従って，$\dot{x}_d(t)-\dot{x}_k(t)$は一様有界となり，$x_d(t)-x_k(t)$は一様有界かつ同程度連続となる．さらに（４－５０）式より$x_d(t)-x_k(t)$が単純収束（区間［０，Ｔ］で任意に固定されたｔにおいての収束）するので,Ascoli-Arzelaの定理より区間［０，Ｔ］で$x_k(t)$は$x_d(t)$に一様に収束することが保証される．

証明終

（４－３０）式において，$\alpha=0$　とすると，（４－４２）式の右辺が非負となることが示せなくなる．それゆえに，以下の証明が成立せず，誤差の収束性は保証されない．以上の理由から入力の修正には速度信号が必要であることが理解できる．

定理４．４では入力は位置信号と速度信号によって修正されたが，つぎの差分形式の微分近似式に着目すると今まで述べたタイプとは違った修正法が可能であることが示せる．

$$\dot{x}(t)=[x(t+\delta)-x(t)]/\delta \qquad (4-53)$$

まず，（４－５３）式より，

$$x(t+\delta)=\delta\dot{x}(t)+x(t) \qquad (4-54)$$

となることから，入力の修正として（４－３４）式に代わって，

$$u_{k+1}(t)=u_k(t)+\Lambda[\alpha(\dot{x}_d(t)-\dot{x}_k(t))+\beta(x_d(t)-x_k(t))]$$
$$=u_k(t)+\Lambda\beta(x_d(t+\delta)-x_k(t+\delta)) \qquad (4-55)$$

とすればよいことがわかる．ただし，$\alpha/\beta=\delta$ とした．このことは，位置誤差：$x_d(t)-x_k(t)$のデータをδだけ進めて入力に加えることを意味している．δが十分小さいときには，近似的に定理４．４の方式と同じであるので，このような位置誤差の位相を進める修正法も有効であると思われる．位相を進める修正法は，時不変線形システムに対して既に内山[12]によって提案されているが，ここでは時変システムに対しても有効であることが示された．

４－３　結言

本章では時変線形メカニカルシステムについて学習制御方式を示し，その収束性を証明した．システムの時変係数行列Ｒ(t)，Ｑ(t)，Ｐ(t)として，はじめはすべて正定対称行列であると仮定したが，基本的には慣性項Ｒ(t)だけの正定対称性が必要であることが，系４．１，４．２，４．３によって示された．これは，直感的には，入力の修正として加速度項だけに注目してパターンを修正していることを意味している．従って，本方式がダイナミクスを扱っているが，構造的には代数的な計算を基本としていることが理解できる．定理４．１，４．２などで定義される規範では，$e^{-\rho t}$という重みをかけた形となっている．これは物理的には時刻の始めの方に大きな重みをかけていることを示している．従って，ρ を極端に大きな値とすると，ある有限な試行回数内では純粋なL_2ノルムでは誤差が減少しない可能性もある．しかしながら，この重みρは数学的な証明のテクニックとして導入されたもので実際には極端に大きな値としなくても良いことが８章において示される．逆に，純粋なL_2ノルムでは誤差がゼロに収束することを時変システムに対して厳密に証明することは難しく，$e^{-\rho t}$という重みでL_2ノルムを少しゆがめた形のノルムを採用することによって初めて以上のような理論的展開が可能となったと考えられる．

本章では位置信号と速度信号による方式について述べたが，加速度信号に基づく方法も時変線形メカニカルシステムに対してまったく同様に，証明することができる．さらに，加速度を利用する場合には制御対象としてある種の非線形システムにまで拡張することができる．そこで，この方法ついては第５章で非線形システムに対して統一的に述べることにする．

第５章　ある種の非線形システムに対する学習制御

５－１　緒言

第４章までの各章ではメカニカルシステムに対して速度または位置と速度の信号を使って入力を修正する方式が提案された．本章では毎回の試行時に得られるそれらの信号の微分値を使って入力を修正する方式を新たに提案する．このような方式では制御対象システムとしてある種の非線形システムにまで拡張できることが示される．制御対象の非線形システムとしてはある種の構造的特徴は仮定するが，ダイナミクスそのものは未知であるという問題設定は第２章及び第４章で述べたものと同じである．

５－２　誤差の収束性

本方式を適用する制御対象としては，次のような非線形システムを考える．

$$\dot{x}(t) = f(t,x) + B u(t) \qquad (5-1-a)$$

$$y(t) = C x(t) \qquad (5-1-b)$$

ただし，$x(t) \in R^n$，$f(t,x) \in R^n$，$B \in R^{nxr}$，$C \in R^{rxn}$，$u(t) \in R^r$，$y(t) \in R^r$ とする．さらにシステムは任意の時刻 $t \in [0, T]$ かつ任意の対 (x_1, x_2) に対して次のリプシッツ条件を満足するものとする．

$$\| f(t, x_1,) - f(t, x_2) \|_\infty \leqq \alpha_1(t) \| x_1 - x_2 \|_\infty \qquad (5-2)$$

ただし，一般にベクトル e の第 i 成分を $e^{(i)}$ で表わすとき，

$$\| e \|_\infty = \max_{1 \leqq i \leqq n} | e^{(i)} | \qquad (5-3)$$

であり，$\alpha_1(t)$ は区間 $[0, T]$ で区分的に連続な関数である．このような制御対象に対

して，実現すべき目標出力$y_d(t)$が区間［０，Ｔ］で与えられるとき，その出力を獲得するために学習プロセスを次のように構成する．

$$\text{入力：}\quad u_{k+1}(t)=u_k(t)+\Lambda(d/dt)\,e_k(t) \tag{5-4}$$

$$\text{誤差：}\quad e_k(t)=y_d(t)-y_k(t) \tag{5-5}$$

$$\text{出力：}\quad y_k(t)=C\,x_k(t) \tag{5-6}$$

ただし，Λは適当なｒｘｒ行列とする．なお，（５－６）式の$x_k(t)$は（５－１－ａ）式から次の式によって支配されるものとする．

$$x_k(t)=x(0)+\int_0^t (f(\tau,\ x_k)+B\,u_k(\tau))d\tau \tag{5-7}$$

ここに，$x(0)$は初期値を表わし，（５－７）式は毎回の試行ごとに初期値を一致させることを意味する．このアルゴリズムのブロック図をFig．５－１に示す．

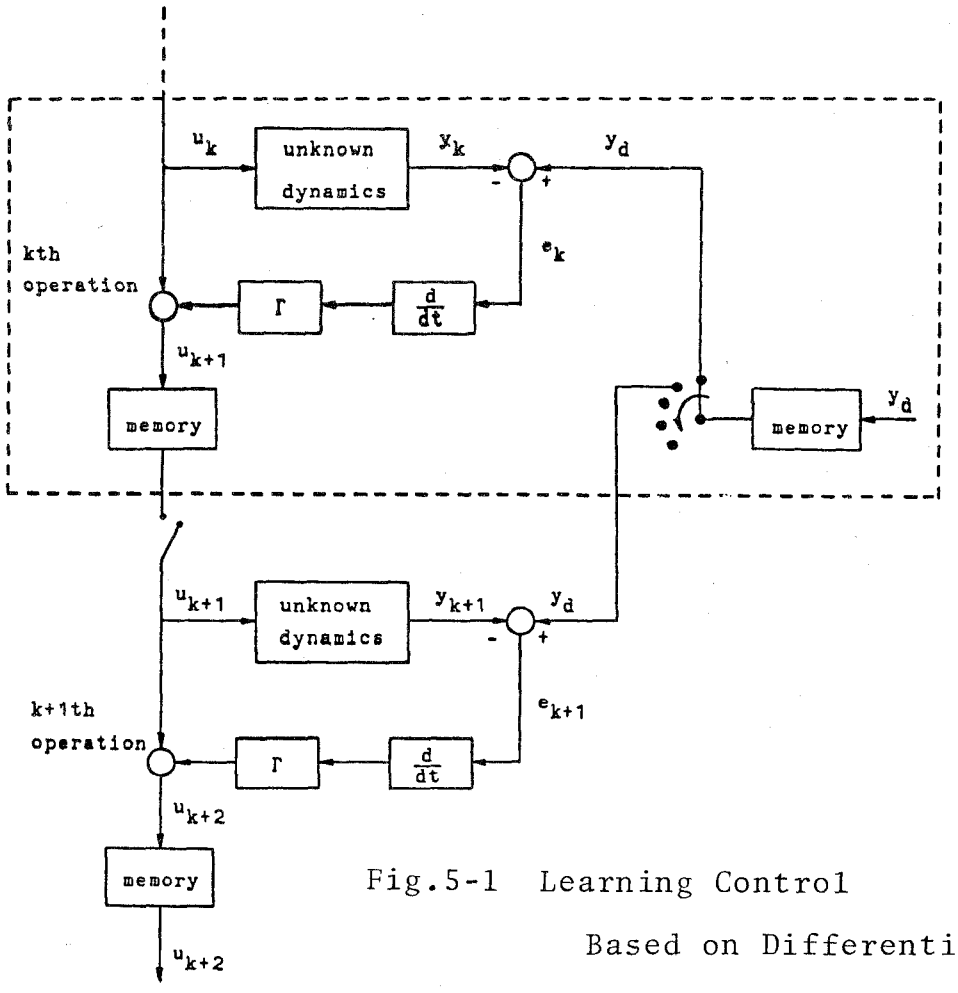


Fig.5-1 Learning Control Based on Differential Operation

ある条件のもとで，（５－４）式による入力の修正が誤差をゼロに収束させ，目標出力を実現できることが次の定理によって示される．

定理

次の３つの条件が同時に満足されるとする．

(a) $\| I_r - CB\Lambda \|_\infty < 1$

ただし，$\| E \|_\infty = \max_{1 \leqq j \leqq r} \left[\sum_i | e^{(ij)} | \right]$，　　$E \in R^{rxr}$

(b) $u_1(t)$，$y_d(t)$は区間［０，Ｔ］で連続的に微分可能，

(c) $y_d(0) = Cx(0)$，

このとき，次の意味で誤差はゼロに収束する．

$$\| \dot{e}_{k+1} \|_\lambda \leqq \delta_0 \| \dot{e}_k \|_\lambda \qquad 0 \leqq \delta_0 < 1 \qquad (5-8)$$

ただし，

$$\| \dot{e} \|_\lambda = \sup_{0 \leqq t \leqq T} \{ e^{-\lambda t} \| \dot{e}(t) \|_\infty \}$$

である．

上式で定義した関数ノルムをここではλ－ノルムと呼ぶ．これは時間関数$e^{-\lambda t}$によって始めの時刻に大きな重みをかけることを意味している．このλ－ノルムのもとで，

$$\| \dot{e}_k \|_\lambda \leqq \delta_0^{\,k} \| \dot{e}_1 \|_\lambda \qquad (5-9)$$

であることから，$k \to \infty$　のときに時間区間［０，Ｔ］において，

$$\dot{e}_k(t) \rightarrow 0,$$
$$\dot{y}_k(t) \rightarrow \dot{y}_d(t) \qquad (5-10)$$

となる．また条件（c）より，このことは区間［0，T］で

$$y_k(t) \rightarrow y_d(t) \qquad (5-11)$$

となることを意味する．

証明

（5－4）式，（5－5）式及び出力の式より，

$$\begin{aligned}\dot{e}_{k+1}(t) &= \dot{y}_d(t) - \dot{y}_{k+1}(t) \\ &= \dot{y}_d(t) - \dot{y}_k(t) + \dot{y}_k(t) - \dot{y}_{k+1}(t) \\ &= \dot{e}_k(t) + C(\dot{x}_k(t) - \dot{x}_{k+1}(t)) \qquad (5-12)\end{aligned}$$

となる．次に（5－1－a）式より，$\dot{x}_k(t)$，$\dot{x}_{k+1}(t)$を表現すると，（5－4）式の関係から，

$$\begin{aligned}\dot{x}_k(t) - \dot{x}_{k+1}(t) &= f(t, x_k) - f(t, x_{k+1}) + B(u_k(t) - u_{k+1}(t)) \\ &= f(t, x_k) - f(t, x_{k+1}) - B\Lambda\dot{e}_k(t) \qquad (5-13)\end{aligned}$$

を得る．（5－12）式に（5－13）式を代入することによって，

$$\dot{e}_k(t) = (I_r - CB\Lambda)\dot{e}_k(t) + C(f(t, x_k) - f(t, x_{k+1})) \qquad (5-14)$$

となる．ここで，（5－2）式のリプシッツ条件より次の不等式が得られる．

$$\|\dot{e}_k\|_\lambda \leqq \|I_r - CB\Lambda\|_\infty \|\dot{e}_k\|_\lambda + \alpha_0 \|C\|_\infty \|x_k - x_{k+1}\|_\lambda \qquad (5-15)$$

ここに，　$\alpha_0 = \max_{0 \leqq t \leqq T} \alpha_1(t)$　である．次に（５－７）式の初期条件を考慮して（５－１３)式より，

$$e^{-\lambda t}(x_k(t) - x_{k+1}(t)) = -\int_0^t e^{-\lambda(t-\tau)} B \Lambda e^{-\lambda \tau} \dot{e}_k(\tau) d\tau$$

$$+ \int_0^t e^{-\lambda(t-\tau)} e^{-\lambda \tau} (f(\tau, x_k) - f(\tau, x_{k+1})) d\tau \quad (5-16)$$

という関係が成立することに注目する．これより次の不等式が得られることがわかる．

$$z(t) \leqq c_0 [(1 - e^{-\lambda T}) / \lambda] \| \dot{e}_k \|_\lambda + \int_0^t \alpha_0 e^{-\lambda(t-\tau)} z(\tau) d\tau \quad (5-17)$$

ただし，

$$z(t) = e^{-\lambda t} \| x_k(t) - x_{k+1}(t) \|_\infty,$$

$$c_0 = \| B \Lambda \|_\infty$$

とする．　（５－１７）式の関係に，Gronwallの不等式を適用し[19]，

$$z(t) \leqq c_0 [(1 - e^{-\lambda T}) / \lambda] \| \dot{e}_k \|_\lambda \exp\{\int_0^t \alpha_0 e^{-\lambda(t-\tau)} d\tau\} \quad (5-18)$$

となる．ここで，z(t)の定義より次の不等式を得る．

$$\| x_k - x_{k+1} \|_\lambda = \sup_{0 \leqq t \leqq T} \{z(t)\}$$

$$\leqq \{c_0 [(1 - e^{-\lambda T}) / \lambda] \exp(\alpha_0 (1 - e^{-\lambda T}) / \lambda)\} \| \dot{e}_k \|_\lambda \quad (5-19)$$

（５－１９）式を（５－１５）式に代入することによって，

$$\| \dot{e}_{k+1} \|_\lambda \leqq \delta_0 \| \dot{e}_k \|_\lambda \quad (5-20)$$

となる．

ただし，

$$\delta_0 = \| I_r - CB\Lambda \|_\infty + \alpha_0 \| C \|_\infty c_0 [(1 - e^{-\lambda T})/\lambda] \exp(\alpha_0 (1 - e^{-\lambda T})/\lambda) \quad (5-21)$$

である．定理の条件（ａ）を満足する場合，十分大きな正の数λによって，

$$0 \leqq \delta_0 < 1 \quad (5-22)$$

とすることができる．

証明終

この方式によればシステムの情報は（５－１－ａ）式の非線形関数ｆ（ｔ，ｘ）に関するリプシッツ条件と，定理の条件（ａ）の行列ＣＢに関するものが必要となるが，正確にシステムを同定する必要がないことに注意したい．

５ － ３ 結 言

本方式をロボットに適用するための詳細な議論は第７章で述べるが，定理の条件（ａ）を満足するためにはメカニカルシステムに対しては出力ｙ(t)に速度信号が含まれていなければならない．（ｙ(t)が位置信号のときは，$\| I_r - 0 \|_\infty = 1$となり誤差の収束性は保証されない．）入力は出力誤差の微分信号によって修正されるので，実際には加速度誤差によって修正されることとなる．

本章では入力の修正法として出力の微分信号を用いたＤタイプ（Differential type）について述べた．この方式の他にも比例，積分，微分の各々の信号を利用したＰＩＤタイプ(Proposional,Integral,Differential type)の入力構成法も実現できることを指摘しておく．［関連論文８］

第６章　脚式移動ロボットへの適用

６－１　緒言

脚式の歩行運動には，いくつかの種類の歩行運動パターンがある．ロボットは環境の，変化に対して運動パターンを選択することによって，歩行における適応能力を獲得するものと思われる．しかしながらロボットのまわりの環境に大きな変化がない場合には，同じ運動パターンを繰り返すことによって歩行運動は実現される．したがって脚式の歩行ロボットにとっては，なによりもまず第一に一つの歩行運動を実現することが最も必要なことである．

一般に，一つの歩行運動パターンが与えられたとしても，単純なサーボ機構だけでは，それは正確には実現できない．なぜならば，リンク間の動的干渉や摩擦，そして歩行に関する機構的拘束条件の変化などが，ロボットの運動に大きく影響するからである．特に，脚の数が多くなり関節数が多くなると，マニュピュレータの制御などに提案されている非線形補償法では計算時間の問題だけからも実現することは難しいと思われる．

一方，本論文で提案する学習制御では，ロボットダイナミクスを正確に推定する必要がないことと，関節数が多い複雑な系に対しても容易に適用できることなどから，脚式ロボットの制御に適しているものと期待できる．本章では，特に二足歩行ロボットにスポットをあてて，提案された学習制御方式の適用を計る．一般に，二足歩行運動では静的歩行（支持脚の足底から重心が外へ出ないようなパターンで遊脚を動かす運動の連続）であれば，ほぼマニュピュレータの制御と同様に取り扱える．しかしながら高速歩行を目的として，足首のトルクを使用しようしないような動的な歩行を実現することは難しい問題である．即ち，動的歩行では，各関節の運動が目標パターンからずれれば，歩行のタイミング（足を踏み出すタイミング）を狂わせ，継続的な歩行が実現できなくなるのである．従って静的歩行よりも，はるかに忠実に各関節が目標パターンに追従することが要求される．そこで，このような問題を解決するために，本論文で提案する学習制御方式を二足歩行ロボットに適用する．そのために，まず第２節では二足歩行ロボットのダイナミクスを定式化する．また第３節では，前述の静的歩行の場合や歩行ロボット特有の動的な歩行を継続的に実現するための学習制御法について詳細に述べる．

脚式の移動ロボットの場合，基本的には１ストライド分だけの学習を実現しておけばよ

い．この方式は計算機の記憶容量の点からも好都合なものとなっている．ただし，この学習方式では単一の歩行パターンに対してのみ有効であり，様々な歩行パターンを実現するには，個々のパターンに対する制御入力をすべて記憶しておく必要がある．これは実装上好ましくないため，記憶すべき制御入力のデータ量をいかに圧縮するかが大きな問題となる．そこで，第４節では，種々の歩行パターンを実現する入力を，あらかじめ学習によって得られた入力パターンから直接獲得する方法を提案する．

６－２　二足歩行ロボットのダイナミクスと線形システムとしての定式化

多リンク構造の二足歩行ロボット（多足歩行ロボットも含む）が，同じく多リンク構造のロボットアームと大きく異なる点の一つは，外部環境に固定されたリンクがないため，その力学的モデルが運動状態によって大きく変化することである．歩行ロボットのアクチュエータ数をＭ，力学的な自由度をＷとすると，運動状態に応じて次の３つの場合が考えられる．

（ａ）　Ｍ＞Ｗ

（ｂ）　Ｍ＝Ｗ

（ｃ）　Ｍ＜Ｗ

人間に類似した駆動機構を有するロボットであれば，（ａ）は両脚が接地している状態，（ｂ）は片脚の足底が着地している片脚状態，（ｃ）はつま先立ちの片脚状態あるいはジャンプ状態に対応する．ただし，ある程度歩幅と歩行速度が大きくなると，外形的には足底が着地していても，足首部分のアクチュエータ入力の制約から実質的には（ｃ）に含まれる歩行パターンとなる．もう一つの大きな特徴は，歩行中に各脚要素の役割（遊脚か支持脚か）が変わることによる各アクチュエータへの負荷の変動である．このような力学モデル自体も大きく変わってゆくシステムに対するフイードフォワード制御入力は，歩行システムの状態の遷移に従って順次切り換えてゆく必要がある．本章では，歩行中力学モデルが変化しない一定期間ごとに前章で述べた学習制御法を適用し，理想的な運動パターンを実現するためのフィードフォワード制御入力を獲得することを考える．

まず，運動状態（ａ），（ｂ）の各々の場合には適当な制御入力，出力を選ぶことによっ

て，各々の次元をシステムの自由度と同一にできる．このときW自由度ロボットのダイナミクスは，

$$A(\theta)\ddot{\theta}+B(\theta)\dot{\theta}^2+F\dot{\theta}+C\sin\theta=u \qquad (6-1)$$

$A(\theta)=a_{ij}\cos(\theta_i-\theta_j)\in R^{WxW}$, $\ddot{\theta}=(\ddot{\theta}_1,\ddot{\theta}2,\cdots,\ddot{\theta}_W)^T$,

$B(\theta)=b_{ij}\sin(\theta_i-\theta_j)\in R^{WxW}$, $\dot{\theta}^2=(\dot{\theta}_1{}^2,\dot{\theta}_2{}^2,\cdots,\dot{\theta}_W{}^2)^T$,

$F=f_{ij}\in R^{WxW}$, $\dot{\theta}=(\dot{\theta}_1,\dot{\theta}_2,\cdots,\dot{\theta}_W)^T$,

$C=\mathrm{diag.}(-c_ig)\in R^{WxW}$, $\sin\theta=(\sin\theta_1,\sin\theta_2,\cdots,\sin\theta_W)^T$, g：重力加速度，

と表現される．[32]~[34] このような非線形システムに対して，５章で提案した方式を適用し，より厳密に解析することも可能であるが，二足歩行運動が実際には鉛直軸近傍に限られることなどから，次のような時不変線形システムと見なすこともできる．

$$A\ddot{\theta}(t)+F\dot{\theta}(t)+C\theta(t)=u(t) \qquad (6-2)$$

ただし，$A=a_{ij}$，$F=f_{ij}$，$C=\mathrm{diag}(-c_ig)$

ここで，行列Ａ，Ｆ，Ｃはすべて正定対称行列であることから２章の表現を用いれば，

$$R\ddot{x}(t)+Q\dot{x}(t)+Px(t)=u(t) \qquad (6-3)$$

ただし，$R=A\in R^{W\times W}$，$Q=F\in R^{WxW}$，$P=C\in R^{WxW}$，$x(t)=\theta(t)$，

と表現できるので提案した学習制御を直接適用することができる．また（Ｃ）の場合では，システムを線形化することによって（６－２）に代わって次式が得られる．

$$A\ddot{\theta}(t)+F\dot{\theta}(t)+C\theta(t)=Du(t) \qquad (6-4)$$

ここに　$D \in R^{W \times M}$　$(M < W)$

従って，２章の表現を用いるならば（２－１）式に代わって，

$$R\ddot{x}(t) + Q\dot{x}(t) + Px(t) = Du(t) \qquad (6-5)$$

となる．この場合，出力$Y(t)$を

$$Y(t) = D^T \dot{x}(t) \qquad (6-6)$$

と設定すれば，このシステムに対して定理２．１，定理２．２，定理２．３は同様に成立する．ただし，この場合は例えば（２－１６）式に代わって，

$$S_k - S_{k+1} = \dot{a}_k^T(T) R \dot{a}_k^T(T) + a_k^T(T) P a_k(T) + \int_0^T \dot{a}_k^T(t)(2Q - D\Lambda D^T)\dot{a}_k^T(t)dt \qquad (6-7)$$

となる．従って，定理の条件は，

$$2Q - D\Lambda D^T > 0 \qquad (6-8)$$

となるが，それらの定理における議論は同様に成立する．

従って，$M < W$の場合についても，M個の自由度については前述の方式と同じ学習プロセスが構成できる．

二足歩行ロボットがつま先立ちの状態，あるいは足首トルクを使用しないで歩行する場合などでは，ロボットの重心に関する前傾運動を内力トルク成分だけで制御することはほとんど不可能である．すなわち，倒立振子になぞらえることのできる重心の前傾運動は，ほぼその運動開始時の状態量で決定され，　内力トルクにはほとんど影響されないのである．ダイナミックな二足歩行はこの不安定なモードをうまく利用して実現されるわけであるが，その時の動的な安定性，即ち定常歩行が可能か否かは，着地時の歩行ロボットの外形で決定される．従って，遊脚が着地する以前にロボットの相対形状が着地時の望ましい

姿勢と歩幅になっていれば，二足歩行ロボットの動的な安定性は保証される．ロボットの外形（M自由度）の理想的な運動パターンを，学習プロセスによって獲得することは前述のように可能であり，そのプロセスにおいて重心の前傾運動はほとんど変化しないので，つま先立ちのような足首トルクを使用しない歩行パターンに対しても，本方式を適用することが可能である.このようにダイナミックな歩行運動の場合には，1スライド分の学習プロセスの対象として，ロボットの相対形状を決める自由度分だけをとるのが適当である．

6－3　運動学習経験に基ずく歩行パターンの選択

二足歩行ロボットに目標の運動パターンが与えられた時，それを忠実に実現させる制御入力は，試行の繰り返しによる学習制御によって獲得されることを前節で述べた．ところで，この方法を用いて様々な歩行パターンをマスターしようとすると，個々のパターンに対する学習プロセスが必要であり，非常に手間がかかる．さらに，歩行パターンの種類が増えれば増える程，個々のパターンを実現させる制御入力の記憶のために，より多くの記憶容量が必要となってくる．以上のような事情から，学習の手間をできるだけ省き，記憶すべき理想的な制御入力パターンの数を減らすことが重要な問題となるが，これはシステムの線形性を利用すれば容易に解決できる．本章では，既に獲得された制御入力を利用して，新たに別の運動パターンを学習プロセスをかいさずに生成する手法について述べる．

ここで考察の対象を（6－3）式で表わされるようなM自由度線形メカニカルシステムに限定して考える．先ず，運動パターンの学習プロセスによって，次のような多項式時間関数で与えられた目標運動パターンが実現されたとしよう．

$$z_{\alpha}(t) = \sum_{i=2}^{m_{\alpha}} \alpha_i t^i \qquad (6-9)$$

ただし，ここでは初期状態を状態変数の原点にとり，目標運動パターンは時間の2次以上の項から構成するものとする．このときの理想的な制御入力$u_{\alpha}(t)$は，（6－3）式より，

$$u_\alpha(t) = \sum_{i=2}^{m_\alpha} [\, i(i-1)t^{i-2}R + i\,t^{i-1}Q + t^{i}P \,]\alpha_i$$

$$= \sum_{i=2}^{m_\alpha} X_i(t)\,\alpha_i \qquad (6-10)$$

となる．

ただし，$X_i(t) = i(i-1)t^{i-2}R + i\,t^{i-1}Q + t^{i}P \qquad \in R^{WxW}$

ここで（６－１０）式中の行列$X_i(t)$には，次の関係があることに注意する．

$$\dot{X}_{i+1}(t) = (i+1)X_i(t) \qquad (6-11)$$

次に，新たに他の目標パターン

$$z_\beta(t) = \sum_{i=2}^{m_\beta} \beta_i t^{i} \qquad (6-12)$$

が与えられたとき，そのパターンを実現する理想的な制御入力

$$u_\beta(t) = \sum_{i=2}^{m_\beta} X_i(t)\,\beta_i \qquad (6-13)$$

を既に得られている$u_\alpha(t)$から生成したいのだが，そのためには自由度Wに対応したW組の$u_\alpha{}^{j}(t)$（$1 \leqq j \leqq W$）が得られていなければならない．いまW組の制御入力を（６－１０）式を使って，

$$u_\alpha{}^j(t) = \sum_{i=2}^{m_\alpha} X_i(t)\,\alpha_i{}^j, \qquad 1 \leqq j \leqq W \qquad (6-14)$$

と表現する．次に上式の各ベクトルを用いて，

$$U_\alpha(t) = [u_\alpha{}^1(t),\ u_\alpha{}^2(t),\ u_\alpha{}^3(t),\ \cdots\ u_\alpha{}^W(t)]$$

$$P_i = [\alpha_i{}^1,\ \alpha_i{}^2,\ \alpha_i{}^3,\ \cdots\ \alpha_i{}^W],\ [1 \leqq i \leqq m_\alpha] \qquad (6-15)$$

のように行列を定義する．このとき，

$$U_\alpha(t) = \sum_{i=2}^{m_\alpha} X_i(t)\,P_i \qquad (6-16)$$

であることから，

$$X_2(t) = U_\alpha(t)\,P_2{}^{-1} - \sum_{i=3}^{m_\alpha} X_i(t)\,P_i\,P_2{}^{-1} \qquad (6-17)$$

となり，その他の$X_i(t)$（$3 \leqq i \leqq m_\alpha$）については（6－11）式の関係から，

$$X_3(t) = 3X_2(t) = 3U_\alpha(t)\,P_2{}^{-1} - 3\sum_{i=3}^{m_\alpha} X_i(t)\,P_i\,P_2{}^{-1}$$

$$X_4(t) = 4X_3(t)$$

・

・

$$X_{m\alpha}(t) = m_\alpha X_{m_{\alpha-1}}(t) \qquad (6-18)$$

が成立するので，各$X_i(t)$（$3 \leqq i \leqq m_\alpha$）を初期条件$X_i(0)=0$のもとに，（6－18）式の行列微分方程式を解くことによって求められる．$X_2(t)$についてはそれらの結果を使って（6－17）式より求められる．以上により，新たに（6－12）式で与えられる目標運動パターンが与えられたとき，理想的な制御入力は，（6－13）式として求めることができる．なお$m_\beta > m_\alpha$の場合は（6－18）式の次数を上げることで簡単に対応しうる．

本方式の有効性を確認するために，二足歩行ロボットの遊脚の2自由度について計算機シミュレーションを行なった．以下にその結果を示す．なお，遊脚モデル（Fig6－1）の各物理パラメータは,Table6－1に示すものである．

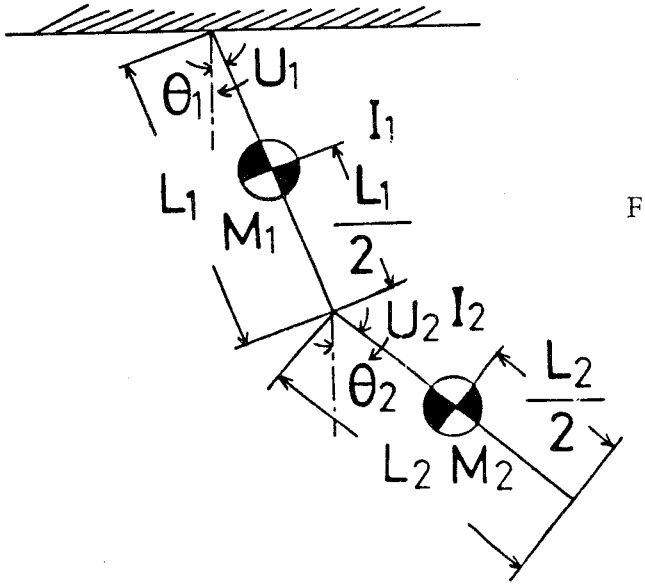


Fig.6-1 Two Links Model

L:length
M:mass
L/2:distance from the joint to the center of gravity
I:inertia

Link No.	M [kg]	L [m]	I [kgm^2]
1	3.77	0.3	0.144
2	4.80	0.3	0.236

Table 6-1 Physical Parameters of a Two Links Model

まず，次の２つの目標出力についての理想的な入力を試行を繰り返すことによって構成した．

$$z_\alpha{}^1(t)=\begin{bmatrix} -40t^3+60t^2-10 \\ -240t^4+480t^3-240t^2-15 \end{bmatrix}$$

$$z_\alpha{}^2(t)=\begin{bmatrix} -40t^3+60t^2-10 \\ -160t^4+320t^3-160t^2-10 \end{bmatrix} \qquad (6-19)$$

ただし，$z(t)=(\theta_1(t),\theta_2(t))^T$である．この２つの歩行パターンを，Fig．６－２－ａに示す．また，この二つの目標運動に対して２０回試行を繰り返し誤差が小さくなった様子をFig.６－２－ｂに示す．このとき入力の修正は第４章の位置と速度の両方を利用する方式を採用した．その際，$\Lambda=\mathrm{diag.}(75,75)$，$\alpha=0.2$，$\beta=0.8$ とした．

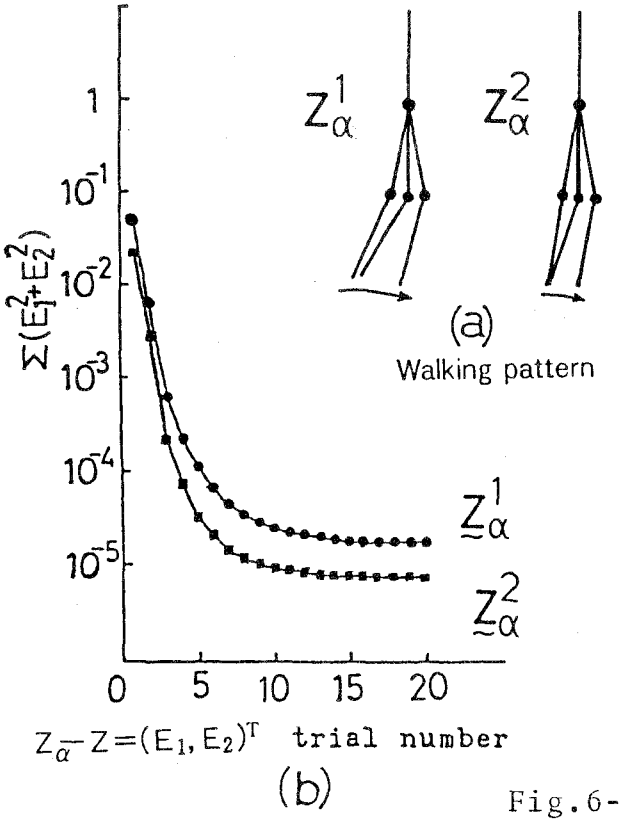


Fig.6-2 Simulation Result

次に別の目標運動が，

$$z_\beta(t)=\begin{bmatrix} -60t^3+90t^2-15 \\ -240t^4+480t^3-240t^2-15 \end{bmatrix} \qquad (6-20)$$

と与えられた場合を考える．この歩行パターンをFig.6－3－aに示す．上述の手法に基づいて，$z_\beta(t)$ に対する入力$u_\beta(t)$を求め，遊脚モデルを駆動した時の，シュミレーション結果を，Fig. 6－3－bに示す．

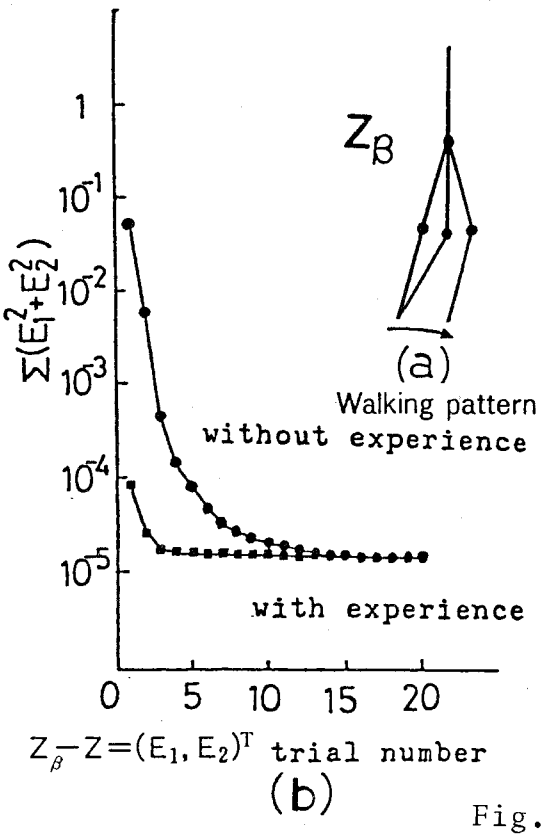


Fig.6-3 Simulation Result

なお，（6－18）式の行列微分方程式における微分を差分操作におきかえ，時系列信号U_αを入力とするディジタルフィルタを構成することによって，制御入力u_βを求めた．また，$z_\beta(t)$の初期位置があらかじめ用意した目標パターンの初期位置と異なるが，これは$z_\alpha{}^2(t)$の初期入力値を，（6－13）式で得られる入力に付加することで対応できる．Fig.6－3－bより本論文の手法を用いた場合，第1回目の試行から誤差が小さくなっていることがわかる．

この手法によれば，いくつかの運動パターンについて入力を構成しておけば，別の運動パターンに対しては，（６－１７）式，（６－１８）式で示される一種のフィルタによって簡単に入力を構成できる．これによってFig.６－４のような運動パターンの選択が可能となる．

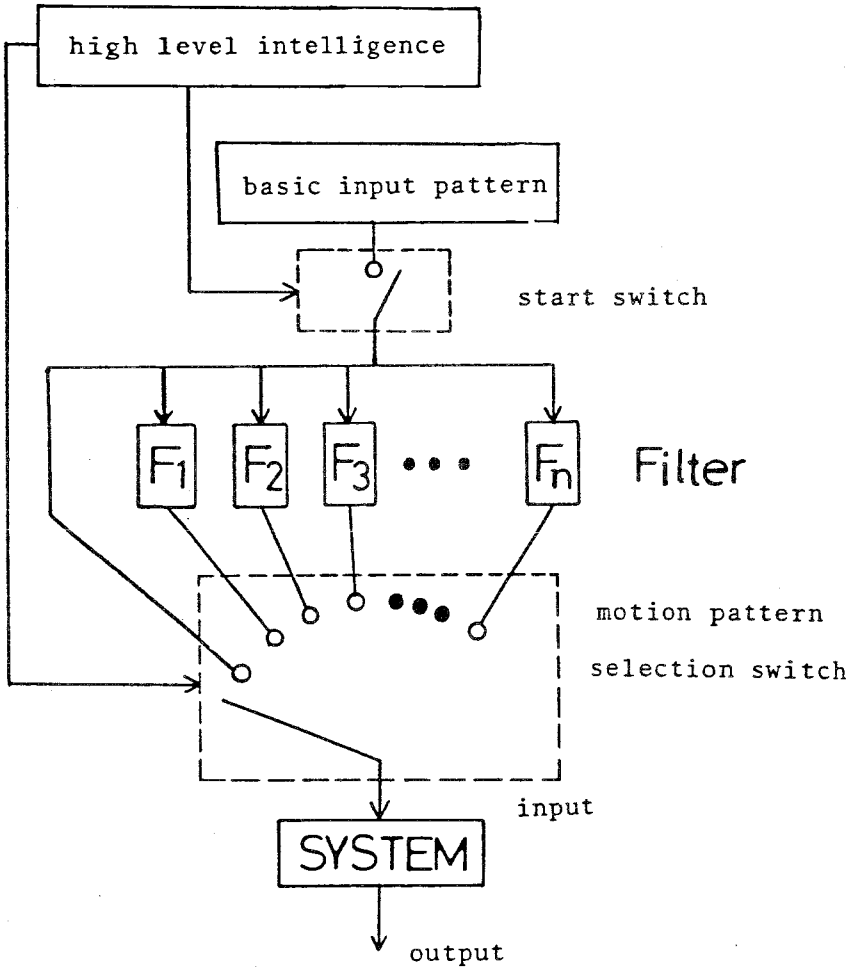


Fig.6-4 Walking Pattern Selecting Mechanism

この図は上位知的レベルで運動の起動スイッチと運動パターン選択スイッチがコントロールされ，ロボットが色々な運動パターンのフィードフォワード入力を高精度で発生させるシステムの概念を示したものである．この選択機能は，ちょうど自動車のギアチェンジの働きに相当するものと考えられる．

6　結言

歩行中，力学モデルの構造が変化しない一定期間中ごとに学習プロセスを実行し，その結果を組合せることによって，二足歩行ロボットが様々な理想歩行パターンを獲得し得ることを示した．この考え方は，二足歩行ロボットのみならず一般の脚式移動ロボットへも十分適用することができる．また，あるパターンについては既に構成された入力から直接制御入力が求められることを示し，これに基づく運動パターン選択法を提案した．本章では，制御対象を時不変線形システムとしたが，ロボットマニピュレータなどの非線形システムに対しても，空間的軌道が同じで時間軸を変化させる場合には，本方式と同様に学習によって得られた入力から直接必要な入力を構成する方法が提案されている．[37],[38] これらの方式を如何に実用化するかが今後の課題である．

第７章　ロボットマニピュレータの運動制御への適用

７－１　緒言

第４章及び第５章ではあるクラスのシステムに対して学習制御が有効であることを主にその収束性について数学的に論じた．本章ではロボットマニピュレータのダイナミクスがそれらの制御クラスに属し，学習方式が実際に適用できることを示す．その際，ロボットに与えられた目標軌道についてフィードバック系を構成する．このフィードバック入力によりロボットの運動は目標軌道の近傍にあると考えられる．しかしながら，多くの場合安定性などの問題からフィードバックゲインを極端に大きくすることができない．それゆえに，目標軌道を正確に実現することは難しい．特に，目標パターンが高速な場合には実際のロボットの運動は大きく目標からずれてしまう．そこで本論文で提案する学習方式を適用し，ロボットの運動の改善を計る．即ち，目標軌道に対する荒い拘束はフィードバック系で実現し，細かな制御を学習によるフィードフォワード系で行なうものである．

７－２　非線形システムとしての定式化

本節では５章で述べたタイプの非線形システムとして，ロボットダイナミクスを定式化する．まず，ロボットダイナミクスをハミルトン形式で次のように表現する．

$$\dot{\theta} = \partial H / \partial q = R^{-1}(\theta) q \qquad (7-1-a)$$

$$\dot{q} = -\partial H / \partial \theta + v \qquad (7-1-b)$$

ただし，$\theta = (\theta_1, \theta_2, \cdots \theta_n)^T$ は関節角座標，$q = (q_1, q_2, \cdots q_n)^T$ は一般化運動量であり，ハミルトニアンHは，

$$H = K + U \qquad (7-2)$$

と表現される．ここに，$K = \dot{\theta}^T R(\theta) \dot{\theta} / 2$，（$R(\theta)$：慣性行列）であり，

Kは運動エネルギー，Uはポテンシャルエネルギーを表わす．また，vは一般化力を表わし，実際には目標軌道$\theta_d(t)$に対してフィードバック系を構成して次のような形式で与える．

$$v = \partial\tilde{U}/\partial\theta + K_p(\theta_d - \theta) + K_v(\dot{\theta}_d - \dot{\theta}) + u \qquad (7-3)$$

ここで，第１項は重力の影響を相殺するために加える重力補償項であり，推定されたロボットのパラメータに基づいて比較的簡単に計算できるので実時間で補償入力を構成する．ただし，実際の重力項$\partial U/\partial\theta$を正確に補償することはパラメータ誤差などの影響で難しいので，その妥当な推定値$\partial\tilde{U}/\partial\theta$を入力に加える形で表現する．（ただし，初期時刻　t＝０では$\partial U/\partial\theta = \partial\tilde{U}/\partial\theta$であるとする．）第２項，第３項は各々位置フィードバック項及び速度フィードバック項であり，ロボットの運動はこれらのフィードバック入力によって目標運動に近い形となるものと考えられる．しかしながらロボットの運動は目標運動と一致せず，一般に誤差が存在する．特に目標パターンが高速な場合にはこの誤差が大きくなり実用上問題となる．そこでこの誤差を小さくするために学習方式により補償入力ｕを構成することを考える．具体的には５章で提案された方式をFig.７－１に示す形で実現する．以下にその詳細を記す．まず，（７－１－ａ）式，（７－１－ｂ）式のロボットダイナミクス及び（７－３）式の一般化力は，状態変数　$x = (\theta^T, q^T)^T$　と補償入力ｕを使って，

$$\dot{x} = f(x) + Bu, \qquad (7-4)$$

$$f(x) = \begin{bmatrix} R^{-1}(\theta)q \\ -\partial K/\partial\theta - \partial U/\partial\theta + \partial\tilde{U}/\partial\theta + K_p(\theta_d - \theta) + K_v(\dot{\theta}_d - \dot{\theta}) \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

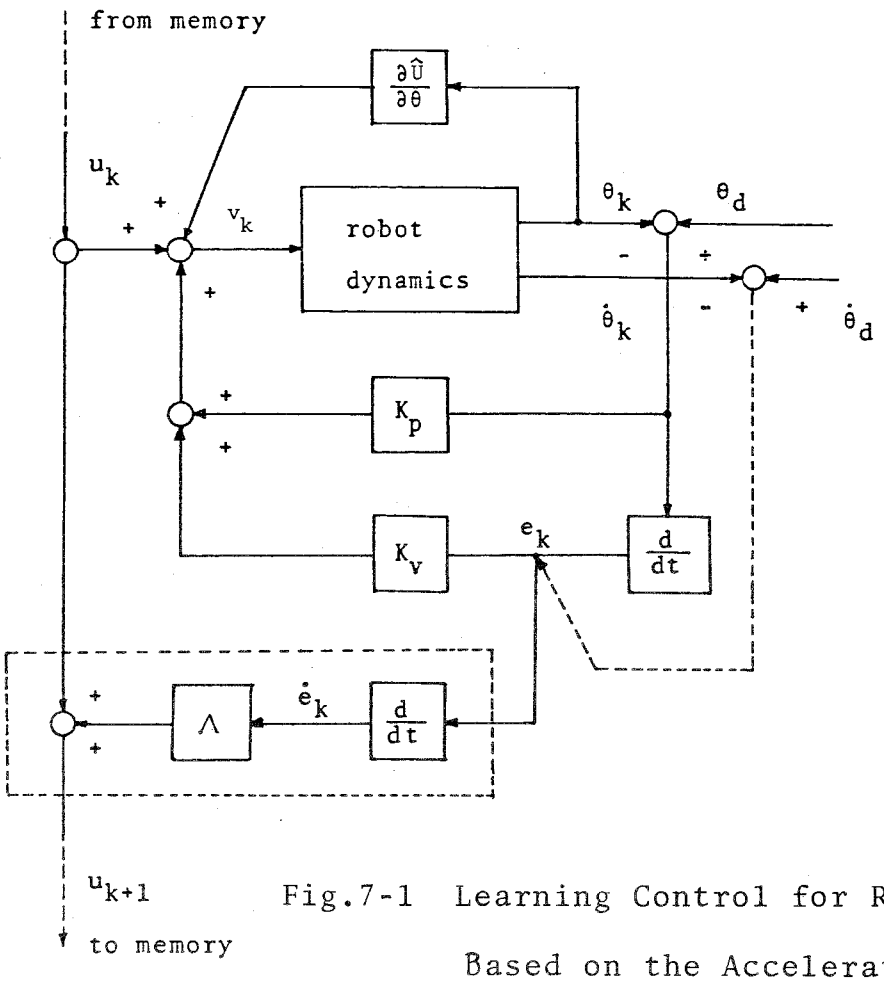


Fig.7-1 Learning Control for Robot Dynamics Based on the Acceleration Signals

と表現でき，（５－１－ａ）式と同様の形式となる．一方出力は一般化運動量ではなく角速度として測定されるので，

$$y = \dot{\theta} = [\,0,\ R^{-1}(\theta)\,] \begin{bmatrix} \theta \\ \\ q \end{bmatrix} = C(x)x \qquad (7-5)$$

となる．この場合出力方程式のＣ行列が状態変数の関数となっているので，前述の定理を直接適用することはできない．しかしながら，状態変数ｘと慣性行列Ｒ(θ)について実用上妥当な仮定を設けることにより，定理の条件（ａ）に代わって

$$\max_{0 \leqq t \leqq T} (\| I_r - \Lambda R^{-1}(\theta) \|_\infty) < 1 \qquad (7-6)$$

が満足されるならば誤差が収束することが以下に示される．

５章において示したように出力を角速度で与えることから，$\dot{e}_k = \dot{y}_d - \dot{y}_k = \ddot{\theta}_d - \ddot{\theta}_k$ となっている．ここで仮に，

$$\| R(\theta_{k+1})\dot{e}_{k+1} \|_\lambda \leqq r \| R(\theta_k)\dot{e}_k \|_\lambda \qquad (7-7)$$

$$0 \leqq r < 1, \ ^\forall k$$

が成立すれば，慣性行列についてはその構造から，

$$\alpha I \leqq R(\theta) \leqq \beta I, \qquad t \in [0, T] \ (\alpha, \beta：正の定数) \qquad (7-8)$$

と考えられるので，[40]

$$\alpha \| \dot{e}_k \|_\lambda \leqq \| R(\theta_k)\dot{e}_k \|_\lambda \leqq r^k \| R(\theta_1)\dot{e}_1 \|_\lambda \leqq \beta r^k \| \dot{e}_1 \|_\lambda \qquad (7-9)$$

となる．このことは $k \to \infty$ のとき $\| \dot{e}_k \|_\lambda \to 0$ を意味している．従って，（７－６）式の成立が証明されれば題意は満たされる．そこで，$R(\theta_k) = R_k$，$dR(\theta_k)/dt = \dot{R}_k$ とするとき，次の関係が成立することに注意する．

$$\begin{aligned} R_{k+1}\dot{e}_{k+1} - R_k\dot{e}_k &= R_{k+1}(\ddot{\theta}_d - \ddot{\theta}_{k+1}) - R_k(\ddot{\theta}_d - \ddot{\theta}_k) \\ &= (R_{k+1} - R_k)\ddot{\theta}_d + R_k\ddot{\theta}_k - R_{k+1}\ddot{\theta}_{k+1} \end{aligned} \qquad (7-10)$$

さらに，（７－１－ａ）式より　$\dot{q}_k = R_k\ddot{\theta}_k + \dot{R}_k\dot{\theta}_k$　であることを利用すると，

$$R_{k+1}\dot{e}_{k+1} - R_k\dot{e}_k = (R_{k+1} - R_k)\ddot{\theta}_d + \dot{q}_k - \dot{q}_{k+1} - \dot{R}_k\dot{\theta}_k + \dot{R}_{k+1}\dot{\theta}_{k+1} \qquad (7-11)$$

となる．次に，（７－４）式より，　$\dot{q}_k$，$\dot{q}_{k+1}$　を代入すると５章の（５－１４）式に対応する次式を得る．

$$R_{k+1}\dot{e}_{k+1}=(I_r-\Lambda R_k^{-1})R_k\dot{e}_k+w_k \qquad (7-12)$$

ただし，w_kは次式で与えられるものとする．

$$w_k=-\partial K/\partial\theta_k+\partial K/\partial\theta_{k+1}-\partial U/\partial\theta_k+\partial\bar{U}/\partial\theta_k$$

$$+\partial U/\partial\theta_{k+1}-\partial\bar{U}/\partial\theta_{k+1}-K_p(\theta_k-\theta_{k+1})-K_v(\dot{\theta}_k-\dot{\theta}_{k+1})$$

$$-(R_k-R_{k+1})\ddot{\theta}_d-(\dot{R}_k-\dot{R}_{k+1})\dot{\theta}_k-\dot{R}_{k+1}(\dot{\theta}_k-\dot{\theta}_{k+1}) \qquad (7-13)$$

ここで，状態量について次の条件を仮定する．

$$\|\theta_k\|_\infty\leqq C_1,\ \|q_k\|_\infty\leqq C_2,\ t\in[0,T],\ \forall k \qquad (7-14)$$

この条件と慣性行列の構造的な特徴から，ある定数α_2によって，

$$\|w_k\|_\lambda\leqq\alpha_2\|x_{k+1}-x_k\|_\lambda \qquad (7-15)$$

が成立することを用いると，（7－12）式より次式が得られる．

$$\|R_{k+1}\dot{e}_{k+1}\|_\lambda\leqq\max_{0\leqq t\leqq T}(\|I-\Lambda R_k^{-1}\|_\infty)\|R_k\dot{e}_k\|_\lambda+\alpha_2\|x_{k+1}-x_k\|_\lambda \qquad (7-16)$$

ここで，$\|x_{k+1}-x_k\|_\lambda$の項については（7－4）式のf（x）が，リプシッツ条件を満足することから，（5－19）式の関係が同様に成立する．即ち，

$$\|x_{k+1}-x_k\|_\lambda\leqq\delta(\lambda)\|\dot{e}_k\|_\lambda, \qquad (7-17)$$

$$\delta(\lambda)=C_0(1-e^{-\lambda T})/\lambda\exp(\alpha_0(1-e^{-\lambda T})/\lambda)$$

であるので，（７－８）式を考慮すれば，

$$\| x_{k+1} - x_k \|_\lambda \leqq \delta(\lambda) \| \dot{e}_k \|_\lambda \leqq [\delta(\lambda)/\alpha] \| R_k \dot{e}_k \|_\lambda \qquad (7-18)$$

となる．（７－１６）式，（７－１８）式より，

$$\| R_{k+1} \dot{e}_{k+1} \|_\lambda \leqq r \| R_k \dot{e}_k \|_\lambda \qquad (7-19)$$

$$r = \max_{0 \leqq t \leqq T} (\| I_r - \Lambda R_k^{-1} \|_\infty) + \alpha_2 \delta(\lambda)/\alpha$$

となる．この r は十分大きな λ によって，

$$0 \leqq r < 1 \qquad (7-20)$$

となる．

７－３　時変システムとしての定式化

本節ではロボットダイナミクスが，目標の軌道まわりで線形化されることによって，第４章で取り扱った時変線形システムとみなせることを示す．まず，ロボットダイナミクスの基本となる（７－１－ａ），（７－１－ｂ）式において、（７－１－ａ）式より $\dot{q}$ を計算して（７－１－ｂ）式に代入すると，

$$R(\theta)\ddot{\theta} + \dot{R}(\theta)\dot{\theta} = -\partial H/\partial\theta + v \qquad (7-21)$$

$R(\theta)$：慣性行列（正定対称）

となる．さらに（７－２）式より $-\partial H/\partial\theta$ を書き下だすと，ロボットダイナミクスは次式で表現される．

$$R(\theta)\ddot{\theta} + f(\theta, \dot{\theta}) + g(\theta) = v \qquad (7-22)$$

ここに，$f(\theta, \dot{\theta}) = \dot{R}(\theta)\dot{\theta} + \partial K/\partial\theta$ ：コレオリ力，遠心力，

$g(\theta) = \partial U/\partial\theta$ ：重力項

目標軌道（運動）が関節角座標で表現されたとし，それを$\theta_d(t)$($t \in [0, T]$)で表わす．ただし，Tは運動終了時刻とする．さらに目標軌道$\theta_d(t)$の各要素は時間区間［0，T］でC^3級以上の滑かな関数で与えられるものとする．この目標軌道に対して線形フィードバック制御系を構成し入力vを，

$$v = \hat{g}(\theta) + K_p(\theta_d - \theta) + K_v(\dot{\theta}_d - \dot{\theta}) + \tau \qquad (7-23)$$

とする．第1項は重力補償項であり，比較的簡単に実時間で計算できるので，（7－23）式のようにフィードフォワード入力として構成する．前節で説明したようにパラメータ誤差などの影響を考慮して，その妥当な推定値$\hat{g}(\theta)$を入力に加えることにする．行列K_p，K_vは各々位置フィードバックゲイン行列，速度フィードバックゲイン行列であり，通常は対角上にのみ正の値を設定する．この制御則によって，ロボットの運動軌道$\theta(t)$は目標軌道$\theta_d(t)$の近傍にあると考えられるが，2つの軌道の間には次のようなゼロでない誤差が一般に存在する．

$$x(t) = \theta(t) - \theta_d(t) \qquad (7-24)$$

この誤差を小さくするために学習制御によってフィードフォワード入力τを構成する．本方式を説明するにあたり，まずx(t)について2次以上のオーダーの項を無視することによって（7－22）式及び（7－23）式をまとめて次のように表現する．

$$[R(\theta_d(t)) + \{\partial R(\theta)/\partial\theta\}_{\theta d} x(t)][\ddot{x}(t) + \ddot{\theta}_d(t)]$$

$$+ f(\theta_d, \dot{\theta}_d) + \{\partial f(\theta, \dot{\theta})/\partial\theta\}_{(\theta d, \dot{\theta} d)} x(t)$$

$$+ \{\partial f(\theta, \dot{\theta})/\partial\dot{\theta}\}_{(\theta d, \dot{\theta} d)} \dot{x}(t)$$

$$+ g(\theta_d) - \hat{g}(\theta_d) + \{\partial(g(\theta) - \hat{g}(\theta))/\partial\theta\}_{\theta d} x(t)$$

$$= -K_p x(t) - K_v \dot{x}(t) + \tau \qquad (7-25)$$

さらに，（７－２５）式を整理して，

$$R(t)\ddot{x}(t) + Q(t)\dot{x}(t) + P(t)x(t) = u(t) \qquad (7-26)$$

を得る．ただし，

$$R(t) = R(\theta_d(t))$$

$$Q(t) = \{\partial f(\theta, \dot{\theta})/\partial \dot{\theta}\}_{(\theta d, \dot{\theta} d)} + K_v$$

$$P(t)x(t) = \{\partial R(\theta)/\partial \theta\}_{\theta d} x(t) \ddot{\theta}_d(t)$$

$$+ \{\partial f(\theta, \dot{\theta})/\partial \theta\}_{(\theta d, \dot{\theta} d)} x(t)$$

$$+ \{\partial (g(\theta) - \hat{g}(\theta))/\partial \theta\}_{\theta d} x(t)$$

$$+ K_p x(t)$$

$$u(t) = h(t) + \tau(t)$$

$$h(t) = -R(\theta_d(t))\ddot{\theta}_d(t) - f(\theta_d, \dot{\theta}_d) - g(\theta_d) + \hat{g}(\theta_d)$$

である．結局，目的は（７－２６）式で表現される誤差システムの入力u(t)(実際には補償入力τ(t))をうまく構成することによって，目標の運動パターン$\theta_d(t)$を実現することである．この場合x(t)が目標軌道からのずれであるので，誤差システム（７－２６）式に対する目標位置パターン$x_d(t)$は，

$$x_d(t) = 0 \qquad 0 \leqq t \leqq T \qquad (7-27)$$

となる．従って，４章の２つのタイプの入力修正法は各々，

４－２－１．速度による方式

$$u_{k+1}(t) = u_k(t) + \Lambda(\dot{x}_d(t) - \dot{x}_k(t))$$
$$= u_k(t) - \Lambda \dot{x}_k(t) \qquad (7-28)$$

４－２－２．位置と速度による方式

$$u_{k+1}(t) = u_k(t) + \Lambda[\alpha(\dot{x}_d(t) - \dot{x}_k(t)) + \beta(x_d(t) - x_k(t))]$$
$$= u_k(t) - \Lambda[\alpha \dot{x}_k(t)) + \beta x_k(t)] \qquad (7-29)$$

とすれば良い．

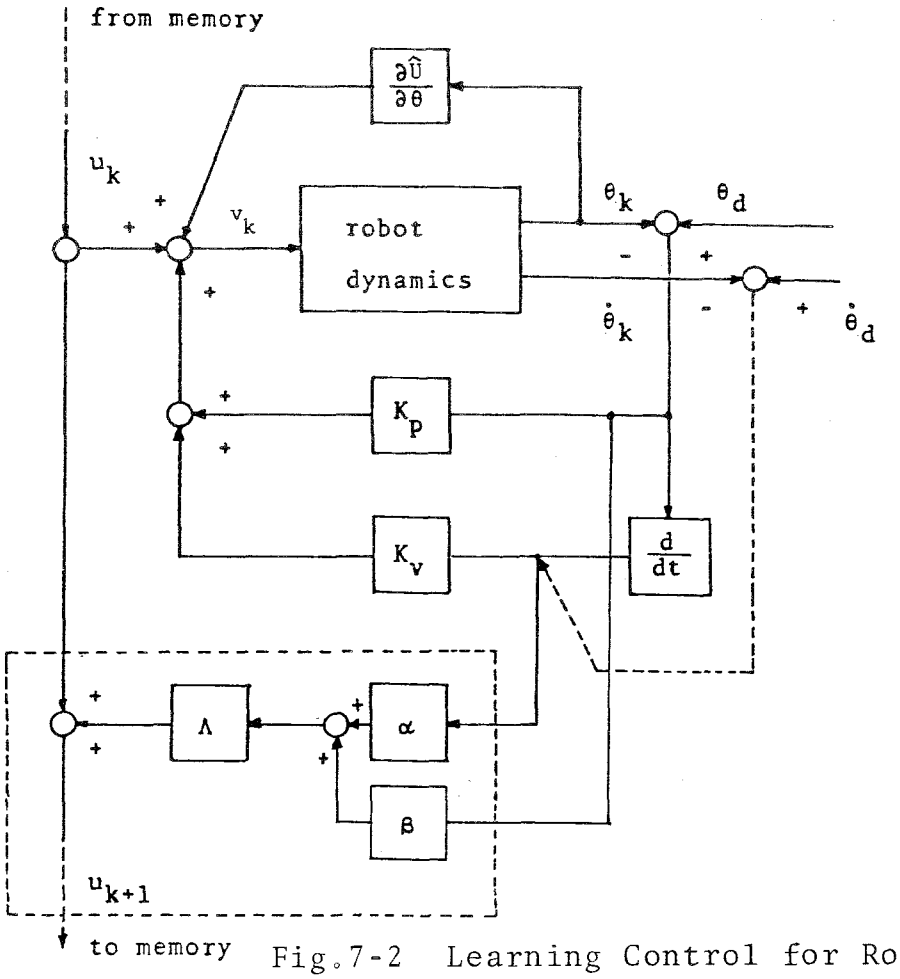


Fig.7-2 Learning Control for Robot Dynamics Based on the Velocity Signals

さらに，目標パターン$\theta_d(t)$の各要素を時間tに関してC^3級以上の滑かな関数で与えれば，係数行列　R(t)，Q(t)，P(t)の各要素把C^1級以上の滑かさとなる．また，初回の入力$u_1(t)$は，$\tau_1(t)=0$　とすれば，　$u_1(t)=h(t)$となりロボットの物理的特性から連続な関数と見なせる．目標出力については，$y_d(t)=\dot{x}_d(t)=0$であることから，明らかに，C^1級の条件を満足する．従って，４章での制御対象システムとしての条件をすべて満足するので，（７－２６）式で与えられる時変線形メカニカルシステムに対して４章の方式を直接適用することができる．なお，この場合のアルゴリズムをFig.７－２に示す．

誤差の収束についての考察

以上の議論より，ロボットダイナミクスを目標軌道のまわりで線形化して得られる時変線形システムに対して第４章で提案した方式が適用できることが示された．その際，誤差の収束性を保証するためには十分大きなρの値を設定する必要がある．このことはL_2ノルムなどの評価規範でみたときに，有限試行回数では誤差が減少しない可能性もある．しかしながら，実際に使用される多くのロボットの物理パラメータでは，フィードバックゲインを他のパラメータに比べて比較的大きく設定できることから，ρは極端に大きくしなくとも良いことが以下のごとく示される．即ち，フィードバックゲインK_p，K_vが正定対称であることを利用すると，（４－４２）式に代わって次式が得られる．

$$
\begin{aligned}
J_k(t)-J_{k+1}(t) \geqq e^{-\rho t}\,[\,&\ddot{a}_k^{T}(t)(\alpha R(t)-\beta\Gamma)\ddot{a}_k(t)\\
&+\dot{a}_k^{T}(t)(\alpha R(t)+\rho\beta R(t)+\alpha K_p+\beta K_v-\alpha\beta\Lambda-\beta R(t)\Gamma^{-1}R(t)\Gamma-\beta\Gamma)\dot{a}_k(t)\\
&+a_k^{T}(t)(\rho\beta R(t)+\alpha K_p+\beta K_v-\beta R(t)\Gamma^{-1}R(t))a_k(t)\,]\\
+\int_0^t e^{-\rho\tau}\,[\,&\ddot{a}_k^{T}(\tau)U_1(\tau)\ddot{a}_k(\tau)\\
&+\dot{a}_k^{T}(\tau)U_2(\tau)\dot{a}_k(\tau)+a_k^{T}(\tau)U_3(\tau)a_k(\tau)\,]\,d\tau
\end{aligned}
\qquad (7-30)
$$

ただし，

$$U_1(\tau)=(\rho R(\tau)+2K_v+2E(\tau)+\dot{R}(\tau)-\alpha\Lambda)\alpha-2\beta R(\tau)-\Gamma$$
$$-W_1^T(\tau)\Gamma^{-1}W_1(\tau)$$

$$U_2(\tau)=(\rho R(\tau)+\rho K_p+2K_v+2E(\tau)-\dot{R}(\tau)-\alpha\Lambda)\alpha+(\rho^2R(\tau)+2K_p$$
$$+2F(\tau)+2\dot{E}(\tau)-\rho\dot{R}(\tau)-2R(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda$$
$$-\Gamma-W_3^T(\tau)\Gamma^{-1}W_3(\tau)$$

$$U_3(\tau)=(\rho^2R(\tau)+\rho K_v+2K_p+2F(\tau)-\rho\dot{R}(\tau)-\beta\Lambda)\beta-\rho\alpha\beta\Lambda$$
$$+\rho K_p-\Gamma-W_2^T(\tau)\Gamma^{-1}W_2(\tau)$$

ただし，

$$W_1(\tau)=(2F(\tau)+2\dot{E}(\tau))^T\alpha+2\beta E(\tau)$$
$$W_2(\tau)=2\alpha\dot{F}(\tau)$$
$$W_3(\tau)=2\alpha F^T(\tau)+2(-\dot{R}(\tau)+E(\tau)+F^T(\tau))\beta$$
$$E(\tau)=Q(\tau)-K_v$$
$$F(\tau)x(\tau)=P(\tau)-K_px(\tau)$$

（７－３０）式の各項に注意するとフィードバックゲインの各要素が比較的大きいことから，極端に大きなρの値を設定しなくとも（７－３０）式の右辺を正にすることができる．このことは誤差の評価規範$S_k(J_k)$をL_2ノルムに近い形で設定できることを示しており実用的にも本方式が有効であろうと予想される．

７－４　結言

リンク構造のロボットダイナミクスが，第４章及び，第５章で記述された時変線形システム及び非線形システムで表現されることを示した．どちらの場合にも実用上の観点から比較的簡単に実現できる重力補償項は実時間で入力し，目標軌道に対してはフィードバック系を構成した．

第８章　ロボットマニピュレータの力制御への適用

８－１．緒言

クランクまわしやグライディング作業など外部環境から力学的干渉を受ける作業をロボットに行わせる場合には，位置制御だけでは不十分であり，力を制御することの必要性が多くの論文で指摘されている．また，その力制御及び位置と力のハイブリッド制御を実現するための幾つかの方式が既に提案されている．[41] しかしながら，ロボットがある方向に高速に運動しながら，別の方向に力制御する場合などでは，その運動の動的な干渉のため目標の位置と力のパターンを同時に実現することはかなり困難である．

一方，本論文で提案する学習方式では，再現性のある外乱が外部環境からロボットに加わったり，高速運動時などでロボットの非線型ダイナミクスが大きく効いてくる場合に特に威力を発揮する．従って，力学的な干渉を受ける作業に対しても有効な制御法であるものと期待できる．そこで本章では，位置と力のハイブリッド制御に学習方式を導入することを考える．その際，力制御方向に対して位置変位と力の大きさに線形な関係があると仮定する．実際，バリ取り作業などでは工具はバネを介してロボットハンドに取り付けられることが多く，この仮定が成立する．このような場合，力の大きさは位置変位に換算できるので，目標の力パターンに対応する位置パターンについて第４章及び第５章で提案した学習方式を直接適用できるように思える．しかしながら，実際には次の二つの問題が起こる．

（１）力制御が必要な作業では対象物からの反力及びそれに伴う摩擦力などの外乱がロボットの運動に大きく影響する．従って，これらの外乱を考慮したロボットダイナミクスについて学習方式を定式化する必要がある．

（２）工具のセッテッング誤差やバネ定数の誤差などから位置変位に換算して間接的に求められた力のパターンには誤差を含むことが多く，微細な力の制御は困難である．

これらの問題点を解決するために，まず第２節では対象物からの反力及びそれに伴う摩擦力なども含めてロボットダイナミクスを記述し，力制御方向には力センサーからの信号を直接使って力の目標パターンを実現する位置と力の学習システムを提案する．次に第３節では力と位置の線形関係を利用して，この方式によってロボット運動が確実に目標の位置

と力のパターンに近づくことを理論的に示す．その際，力と位置の間に線形な構造を仮定するが，バネ定数に対応する係数及びロボットの物理パラメータなどは未知である場合にも本方式が有効であることが示される．

８－２．ハイブリッド学習制御の構成

ロボットがグライディングなどの作業を実行するとき，対象物からの反力の影響を含めてｎ自由度のロボットダイナミクスを次のように表現する．

$$R(\omega)\ddot{\omega} + f(\omega,\dot{\omega}) + g(\omega) = W + r(\omega,\dot{\omega}) \qquad (8-1)$$

ただし，　$\omega = (\omega_1, \omega_2, \cdots \omega_n)^T$　は，一般化座標であり，作業内容を記述できる作業座標（通常はカーテシアン座標系）で与えるものとする．$R(\omega)$は慣性行列と呼ばれるもので正定対称行列となることに注意しておく．　$f(\omega,\dot{\omega})$は遠心力，コレオリ力及び粘性摩擦を表す非線形項であり，$g(\omega)$は重力項を表す．また，$W = (W_1, W_2, \cdots W_n)^T$はこの座標系に対応する一般化力である．さらに，$r(\omega,\dot{\omega})$ は力制御で問題となる反力及びそれに伴う摩擦力を表す項である．実際には，この座標系ωとロボットのアクチュエータの座標系θが異なる場合が多い．例えば，関節型ロボットに対して作業座標系をカーテシアン座標系にとる場合などが多い．この場合，一般に次の様にして実際にアクチュエータに加える入力Ｖを求めればよいことが示されている．[1],[42]

$$V = J^T W \qquad (8-2)$$

ここに行列Ｊは，$J = \delta\omega/\delta\theta$ で定義されるヤコビアン行列である．この方式によって実際にアクチュエータに加えるべき入力Ｖが求められるので，以下では作業座標系ωに対応する入力Ｗの構成について考えることにする．

まず，座標系の位置制御方向をω^p力制御方向をω^fとして，

$$\omega = \begin{bmatrix} \omega^p \\ \omega^f \end{bmatrix}, \qquad \begin{aligned} \omega^p &= (\omega_1,\ \omega_2,\ \cdots \omega_m)^T, \\ \omega^f &= (\omega_{m+1},\ \omega_{m+2},\ \cdots \omega_n)^T, \end{aligned} \tag{8-3}$$

と表すものとする．即ち，m自由度については位置制御，n－m自由度については力制御を実行する．次に位置と力の目標軌道を各々$\omega^p{}_d(t)$，$f_d(t)$，（$0 \leqq t \leqq T$，T運動修了時刻）とし，その各要素は4章の定理の係数行列に対する条件を満足するために，C^3級以上の滑かな関数で与える．この目標軌道に対して入力Wを次のように構成する．

$$W(t) = \hat{g}(t) + \begin{bmatrix} G_p(\omega^p{}_d - \omega^p) + G_v(\dot{\omega}^p{}_d - \dot{\omega}^p) \\ F_p(f_d - f) - F_v\dot{\omega}^f \end{bmatrix} + \tau \tag{8-4}$$

$$G_p,\ G_v \in R^{m \times m}, \quad F_p,\ F_v \in R^{(n-m)\times(n-m)}$$

ここに第1項は重力補償項であり，第7章で述べたように，比較的簡単に計算できるので実時間で計算し入力に加えるものとする．ただしロボットの物理パラメータの推定誤差を含んでいるので，推定される重力項$\hat{g}(t)$として入力に加えるものとする．ただし，初期時刻では$\hat{g}(0) = g(0)$であるとする．次に第2項では，fは力センサーから測定された力信号であり位置と力の目標軌道に対して feedback系を構成している．feedback gain行列G_p，G_v，F_pはすべて対角上にのみ正の値を設定するものとする． 力制御方向の$-F_v\dot{\omega}^f$の項はダンピング項であり行列F_vも対角上にのみ正の値をもつ． 位置と力を同時に制御するためのサーボ系として通常外乱補償用に積分補償を加えた形のものが採用されているが，ここではそれにかわる補償入力として第3項の学習的に作られるfeedforward入力τを与えるものとする．そのfeedforward入力の構成を説明するためにまず， 位置と力の誤差$e^p(t)$，$\varepsilon(t)$を各々次のように定義しておく．

$$\begin{aligned} e^p(t) &= \omega^p{}_d(t) - \omega^p(t), \\ \varepsilon(t) &= f_d(t) - f(t), \end{aligned} \tag{8-5}$$

また，繰返し回数を表現するために，以後ｋ回目の試行時の位置と力を$\omega^{p}{}_{k}(t)$，$f_{k}(t)$とする．さらに，その誤差を次のように表現する．

$$e^{p}{}_{k}(t) = \omega^{p}{}_{d}(t) - \omega^{p}{}_{k}(t),$$
$$\varepsilon_{k}(t) = f_{d}(t) - f_{k}(t) \qquad (8-6)$$

このような準備のもとに学習プロセスを次に述べるような形で構成する．まず，初期値は毎回試行ごとに目標パターンの初期化に一致させる．即ち，

$$e^{p}{}_{k}(0) = \dot{e}^{p}{}_{k}(0) = 0,$$
$$\varepsilon_{k}(0) = \dot{\varepsilon}_{k}(0) = 0, \qquad (8-7)$$

とする．このとき，ｋ回目入力τ_{k}を修正して，ｋ＋１回目の入力　τ_{k+1}を次のように構成する．

$$\tau_{k+1}(t) = \tau_{k}(t) + \Phi \begin{bmatrix} \alpha\, \dot{e}^{p}{}_{k}(t) + \beta\, e^{p}{}_{k}(t) \\ \alpha\, \dot{\varepsilon}_{k}(t) + \beta\, \varepsilon_{k}(t) \end{bmatrix} \qquad (8-8)$$

ただし，１回目のfeedforward入力τ_{1}は，

$$\tau_{1}(t) = 0 \qquad 0 \leqq t \leqq T \qquad (8-9)$$

とする．また，行列$\Phi \in R^{nxn}$は，

$$\Phi = \mathrm{diag.}(\phi_{1}, \phi_{2}, \cdots \phi_{n}), \qquad \phi_{i} \geqq 0 \quad (1 \leqq i \leqq n) \quad (8-10)$$

とする．スカラー量の係数　α，β　は，

$$\alpha + \beta = 1, \qquad (0 < \alpha \leqq 1,\ 0 \leqq \beta < 1) \qquad (8-11)$$

を満たすものとする．このアルゴリズムをFig.８－１に示す．

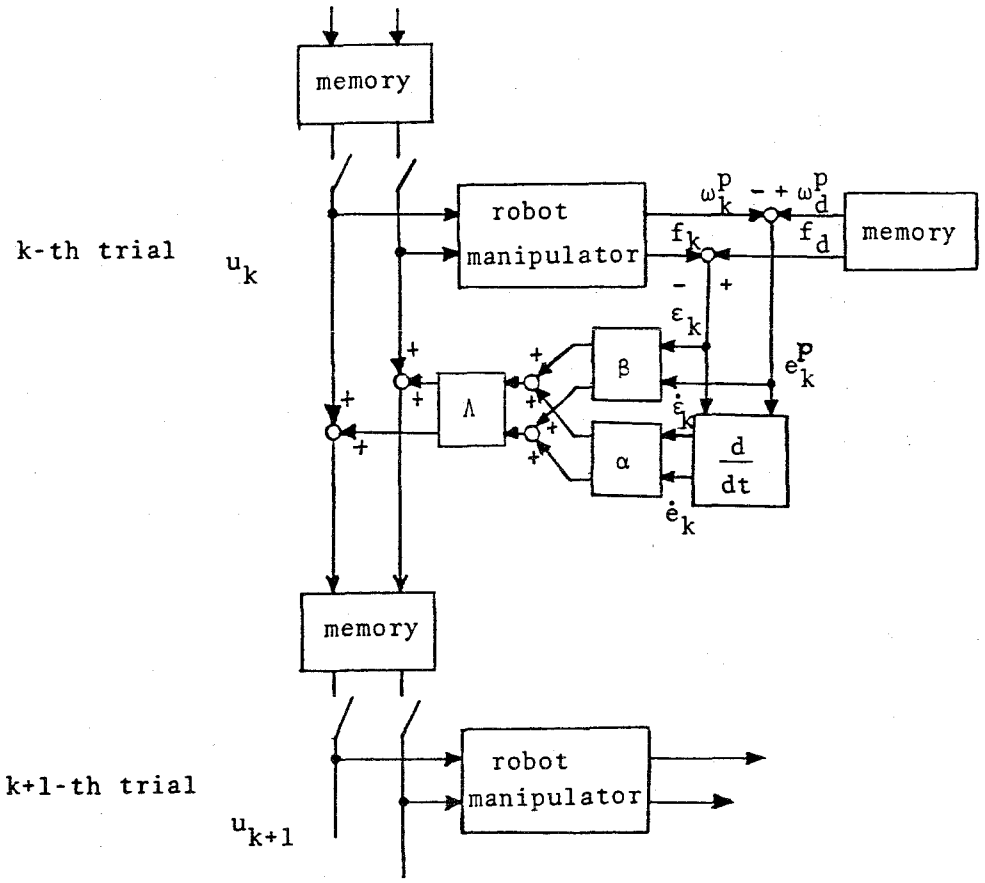


Fig.8-1 Learning Algorithm for Hybrid Position/Force Control

８－３　学習による誤差の収束性

本節では，前節で述べた（８－８）式の入力修正法によって実際に位置誤差$e^p(t)$と力誤差$\varepsilon(t)$が試行とともにゼロに収束することを示す．まず，力の大きさ$f(t)$は，力制御方向の位置変位$\omega^f(t)$と次のような関係が成立するものとする．

$$f(t) = K\,\omega^f(t) \qquad (8-12)$$

ただし，$K = \mathrm{diag.}(k_1, k_2, \cdots k_{n-m})$で与えられ，バネ定数に対応する行列である。以下では，力制御方向の誤差の収束を示すために，力に対応する位置変位$\omega^f(t)$を使って議論を進めるが，実際の制御には力信号$f(t)$に基づいて実行されるので，係数行列Kの各要素及び位置変位$\omega^f(t)$は陽に求める必要がないことを指摘しておく．力の目標軌道に対応する位置の目標軌道を$\omega^f{}_d(t)$とすれば，（８－１２）式と同様に目標軌道についても

$$f_d(t) = K\,\omega^f{}_d(t) \qquad (8-13)$$

となる．また，（８－５）式の力の誤差$\varepsilon(t)$に対応する位置誤差を，$e^f(t)$とすれば同様に，

$$\begin{aligned} e^f(t) &= \omega^f{}_d(t) - \omega^f(t) \\ &= K^{-1}(f_d(t) - f(t)) \\ &= K^{-1}\varepsilon(t) \end{aligned} \qquad (8-14)$$

を得る．これらの関係を用いることによって，（８－４）式の入力 $W(t)$は次のように位置変位のみによって表現される．

$$W = \hat{g}(\omega) + \begin{bmatrix} G_p, & 0 \\ 0, & F_pK \end{bmatrix} \begin{bmatrix} \omega^p{}_d - \omega^p \\ \omega^f{}_d - \omega^f \end{bmatrix} + \begin{bmatrix} G_v, & 0 \\ 0, & F_v \end{bmatrix} \begin{bmatrix} \dot{\omega}^p{}_d - \dot{\omega}^p \\ \dot{\omega}^f{}_d - \dot{\omega}^f \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ -F_v\dot{\omega}^f{}_d \end{bmatrix} + \tau$$

$$= \hat{g}(\omega) + K_p(\omega_d - \omega) + K_v(\dot{\omega}_d - \dot{\omega}) + i + \tau \qquad (8-15)$$

ただし，

$$K_p = \begin{bmatrix} G_p, & 0 \\ 0, & F_pK \end{bmatrix}, \quad K_v = \begin{bmatrix} G_v, & 0 \\ 0, & F_v \end{bmatrix}, \quad \omega_d = \begin{bmatrix} \omega^p{}_d \\ \omega^f{}_d \end{bmatrix}, \quad i = \begin{bmatrix} 0 \\ -F_v\dot{\omega}^f{}_d \end{bmatrix}$$

feedforward入力$\tau = 0$のときのfeedback入力についての有効性は既に報告されており，[1]運動はほぼ目標軌道$\omega_d(t)$近傍にあると考えられるが，なお次のような誤差が存在する．

$$x(t)=\begin{bmatrix} -e^{p}(t) \\ -e^{f}(t) \end{bmatrix}=\begin{bmatrix} -\omega^{p}{}_{d}+\omega^{p} \\ -\omega^{f}{}_{d}+\omega^{f} \end{bmatrix}=-\omega_{d}(t)+\omega(t) \qquad (8-16)$$

この誤差$x(t)$は比較的小さな値であることから，第７章と同様に２次以上の項を無視することによって（８－１）式のダイナミクスは，目標軌道$\omega_{d}(t)$のまわりで線形化され，次のように表現される．

$$R(t)\ddot{x}(t)+Q(t)\dot{x}(t)+P(t)x(t)=u(t) \qquad (8-17)$$

ただし，

$$R(t)=R(\omega_{d}(t)),$$

$$Q(t)=[\partial f(\omega,\dot{\omega})/\partial\dot{\omega}]_{(\omega d,\dot{\omega} d)}-[\partial r(\omega,\dot{\omega})/\partial\dot{\omega}]_{(\omega d,\dot{\omega} d)}+K_{v},$$

$$P(t)x(t)=[\partial R(\omega)/\omega]_{\omega d}x(t)\ddot{\omega}_{d}(t)+[\partial f(\omega,\dot{\omega})/\partial\omega]_{(\omega d,\dot{\omega} d)}x(t)$$

$$-[\partial r(\omega,\dot{\omega})/\partial\omega]_{(\omega d,\dot{\omega} d)}x(t)$$

$$+[\partial(g(\omega)-\hat{g}(\omega))/\omega]_{\omega d}x(t)+K_{p}x(t),$$

$$u(t)=h(t)+i(t)+\tau(t)$$

$$h(t)=R(\omega_{d}(t))\ddot{\omega}_{d}(t)+f(\omega_{d},\dot{\omega}_{d})-r(\omega_{d},\dot{\omega}_{d})+g(\omega_{d})-\hat{g}(\omega_{d})$$

従って，ロボットダイナミクスは第４章の定理が直接適用できる時変形線形メカニカルシステムとして表現できる．次に，（８－７）式の入力の修正について考える．先ず，k回めの力誤差$\varepsilon_{k}(t)$に対応する位置誤差　$e^{f}{}_{k}(t)$は，

$$e^{f}{}_{k}(t) = \omega^{f}{}_{d} - \omega^{f}{}_{k} = K(f_{d} - f_{k}) = K\,\varepsilon_{k}(t) \qquad (8-18)$$

となるので，（８－８）式は次のように表現される．

$$\tau_{k+1}(t) = \tau_{k}(t) + \Phi \begin{bmatrix} \alpha\,\dot{e}^{p}{}_{k}(t) + \beta\,e^{p}{}_{k}(t) \\ \alpha K\,\dot{e}^{f}{}_{k}(t) + \beta K\,e^{f}{}_{k}(t) \end{bmatrix} \qquad (8-19)$$

さらに，

$$x_{k}(t) = \begin{bmatrix} -e^{p}{}_{k}(t) \\ -e^{f}{}_{k}(t) \end{bmatrix} \qquad (8-20)$$

と表現し，$u_{k}(t) = h(t) + i(t) + \tau_{k}(t)$　であることに注意すれば，(８－１９)式は，

$$\tau_{k+1}(t) = \tau_{k}(t) - \Lambda \begin{bmatrix} \alpha\,\dot{x}_{k}(t) + \beta\,x_{k}(t) \end{bmatrix} \qquad (8-21)$$

と書替られる．ただし，行列Λは行列Φを

$$\Phi = \begin{bmatrix} \Phi_{1}, & 0 \\ 0 & , \Phi_{2} \end{bmatrix} \qquad \begin{matrix} \Phi_{1} \in R^{m \times m}, \\ \Phi_{2} \in R^{n-m \times n-m}, \end{matrix} \qquad (8-22)$$

とするとき，

$$\Lambda = \begin{bmatrix} \Phi_{1}, & 0 \\ 0 & , \Phi_{2}K \end{bmatrix} \qquad (8-23)$$

で与えられるものである．従って，制御対象システムは（７－２６）式と同様に表現でき，入力の修正は（７－２９）式と同じ形式で与えられるので，第７章の位置制御の場合と同様に位置と力の制御にも第４章の定理を適用することができる．その際，位置と力の線形関係の係数Ｋは未知である場合にも，行列Λを正定とするようにΦ_2を適当に設定できる．このようにして，本方式では係数Ｋを正確に推定する必要なく，力の誤差をゼロに収束させることができる．

８－４．結言

試行を繰返すことによって目標の位置パターンと力パターンをロボットに実現させる学習制御方式を提案した．この方式では，力制御方向には力と位置に線形関係を仮定しているが，その係数及びロボットの物理パラメータは未知である場合にも本方式が有効であることが示された．

第９章　実験結果及び
シミュレーション結果

９－１　緒言

前章までに，本論文で提案する学習制御がロボットの制御に適用でき，目標運動からのずれはゼロに収束することが数学的に証明された．その際，証明のテクニックとして第４章では重みρの値を十分大きくすること，第５章では，重みλの値を十分大きくすることを必要とした．これらは，いずれも時刻の始めの方に大きな重みをかけることを意味しており，誤差が一様に小さくなるためには多くの試行回数を必要とする可能性もある．そこで本章では，実際に我々が制御するロボットなどでは，その物理的パラメータの関係から重みのρ及びλを極大に大きくする必要が事実上ないことを実験及びシミュレーションによって示す．

９－２　二足歩行ロボット

本論文で述べた学習プロセスの二足歩行ロボットへの実装可能性を示すために，実機による歩行パターンの学習結果を以下で紹介する．今回の実験では，片脚支持期の終りのロボットの形状が，つぎの片脚支持期の始めの形状になるような目標運動のパターンを与えた．（Fig.９－１）また，片脚支持期で支持脚の足底が浮き上がらないような運動パターンを選んだので，一歩の片脚期の歩行パターンを完成することによって，継続的な歩行が実現されることとなる．本実験に使用した二足歩行ロボットの写真を，Photo.９－１に示す．このロボットは前後方向に６自由度，左右方向に１自由度を持ち，駆動機構はＤ.Ｃ.サーボモータをハーモニックドライブで減速し，タイミングベルトによって各関節にトルクを伝達するというものである．また各関節には，ポテンショメータがとりつけられており，角変位を知ることができる．実験システム全体の概略図をFig.９－２に示す．マイクロコンピュータからの入出力線とモータへの電源コードは，歩行の障害とならないように天井に取り着けられたレールに沿って滑らかに動くようになっている．またこのロボットは，完全に自律型で機械的な拘束はなく，左右方向の安定化は上部のバランサーを左右に移動させることによって実現している．このような方式においてはバランサーが一方の端

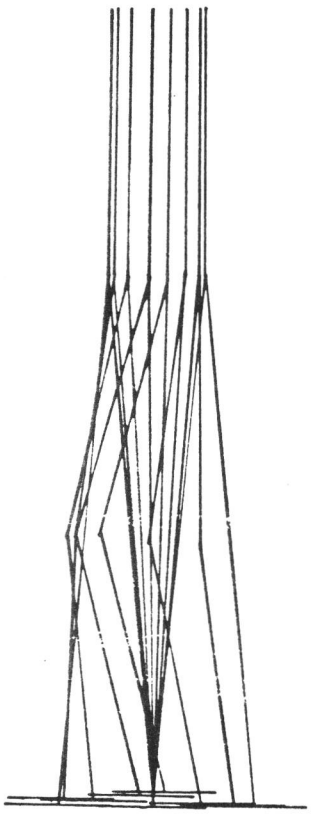


Fig.9-1 Desired Walking Pattern

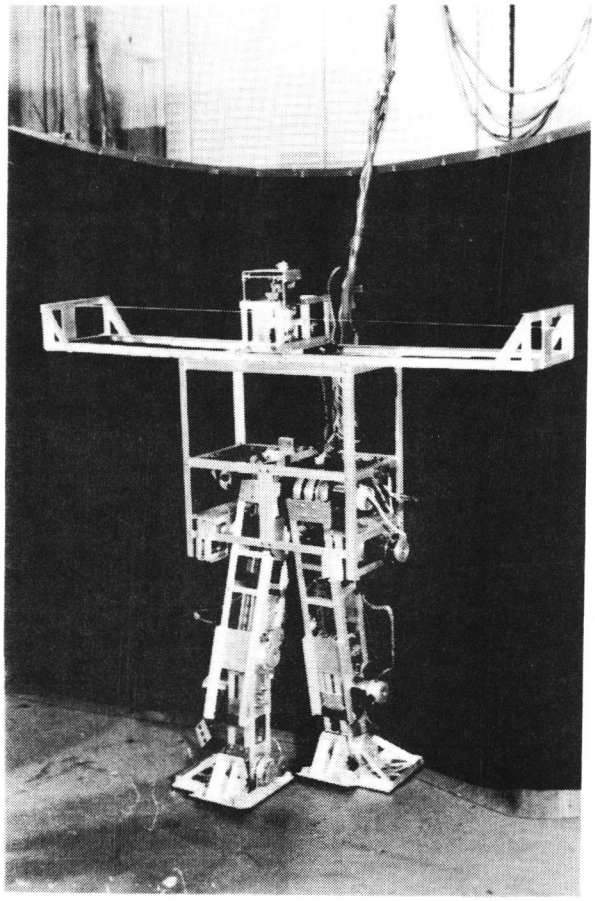


Photo. 9-1

Photograph of a Biped Robot

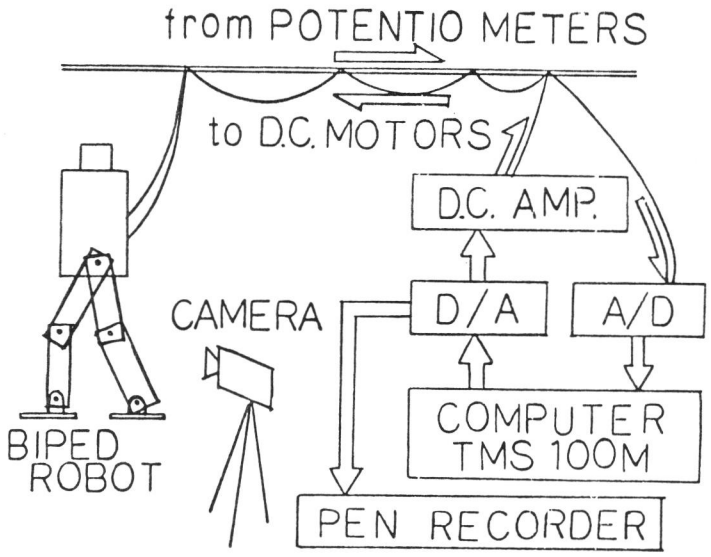


Fig.9-2 Experimental System

にあり，遊脚を動かすとロボットが鉛直軸を中心に回転してしまう可能性がある．ここでは足底にラバーを取り着けるなどのハードウェア的方法と回転力を生じにくい歩行パターンを選択するというソフトウェア的な方法によってこの問題を解決している．このロボットに上述の制御方式を適用するにあたり，左右方向はあらかじめ人間がプログラムしてロボットの安定化をはかり，動的な影響を大きく受ける前後方向の６自由度（足首のアクチュエータも含む）について繰り返しを利用して入力を求めた．実際には，目標出力に対してサーボ系を構成し，そのサーボ系への入力を修正する方法によって実験を行なった．即ち，ロボットのアクチュエータへの入力ｖ(t)を次のように与えた．

$$v(t)=K_p(\theta_d(t)+\tau(t)-\theta(t))+K_v(\dot{\theta}_d(t)-\dot{\theta}(t)) \qquad (9-1)$$

ただし，$\theta(t)$：関節角座標，$\theta_d(t)$ ：目標パターン，$\tau(t)$：学習によるフィードフォワード入力，K_p，K_v：フイードバックゲインである．ここで，速度信号は次のような差分形式で求めた．

$$\dot{\theta}(t)=[\theta(t+\Delta t)-\theta(t)]/\Delta t \qquad (9-2)$$

また，入力の修正に関して第４章で述べた位相を進める方法を採用した．その位相進み量δは各自由度とも同じ値（約５０ｍｓｅｃ）を与えた．目標出力は時間の３次軌道の組合せによって１歩０．８秒、歩幅２３ｃｍの歩行パターンを生み出す形を与えた（Fig.９－１）．実験では第１回目の試行時に，主に追従遅れなどの原因によって目標出力が得られず，ロボットは一歩踏み出すことに失敗し前に倒れてしまった．そこで目標出力と実際の応答の差を１０ｍｓｅｃ程度のサンプリング間隔で時系列信号としてメモリしておき，上述の方法により２回目以後の入力を構成して実験を続けた．その結果，５回目の試行時から１歩踏み出すことに成功した．またこのとき得られた入力パターンを使用して定常な歩行を実現することができた．このときのシステム応答や入力などをFig.９－３，Fig.９－４に示す．R_{X1}～R_{X6} はFig.９－５に示される角度であり，θ_d は目標出力，θ_c はサーボ系に与える入力（$\theta_c(t)=\theta_d(t)+\tau(t)$），θはシステム応答，ｅは誤差，ｊは試行回数である．また各々の関節の誤差収束の様子をFig.９－６に示す．Fig.９－３，Fig.９－４ Fig.９－６などから試行回数が増すごとに誤差が小さくなっていることがわかる．

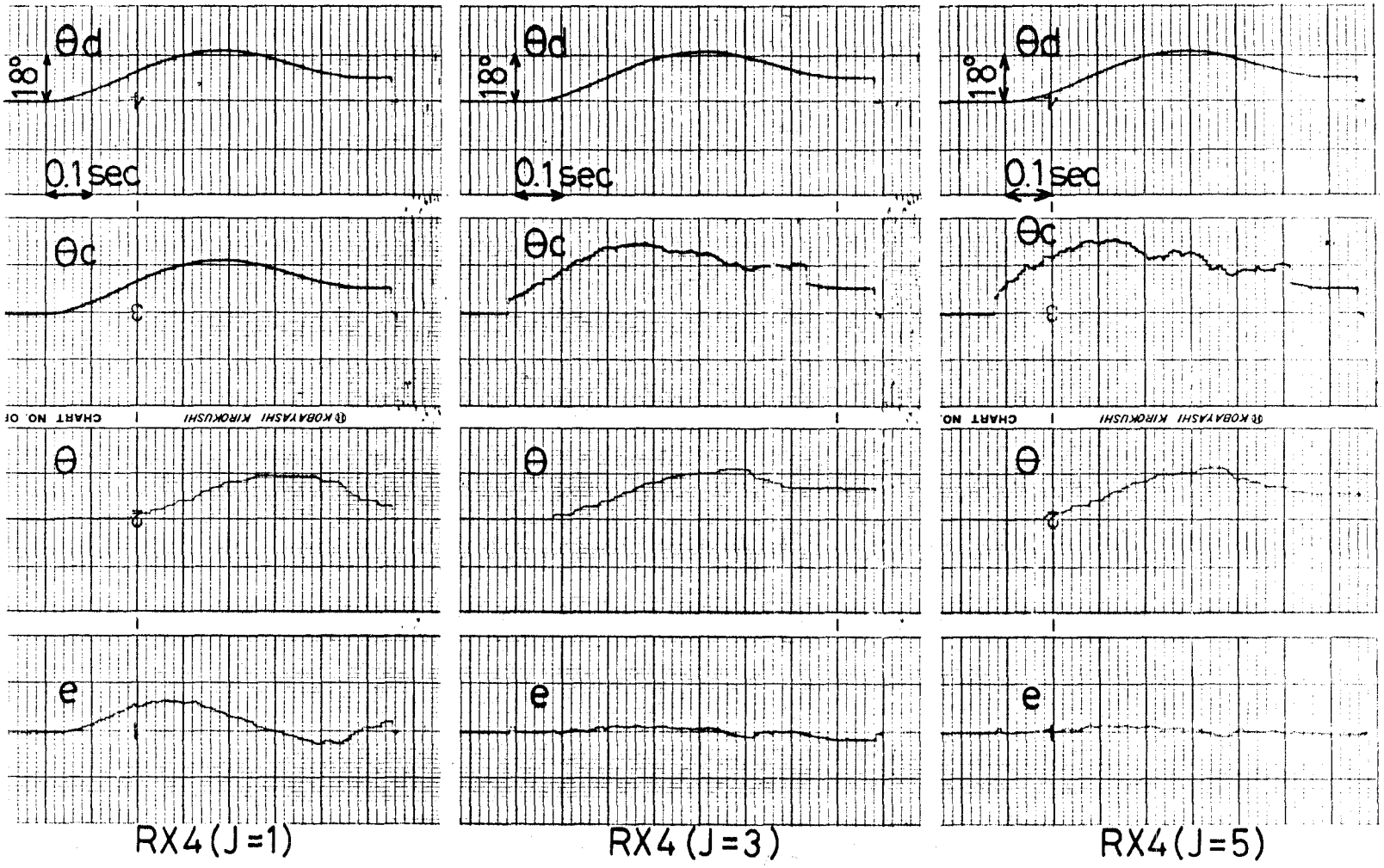


Fig.9-3 Experimental Result of the Biped Locomotion Robot

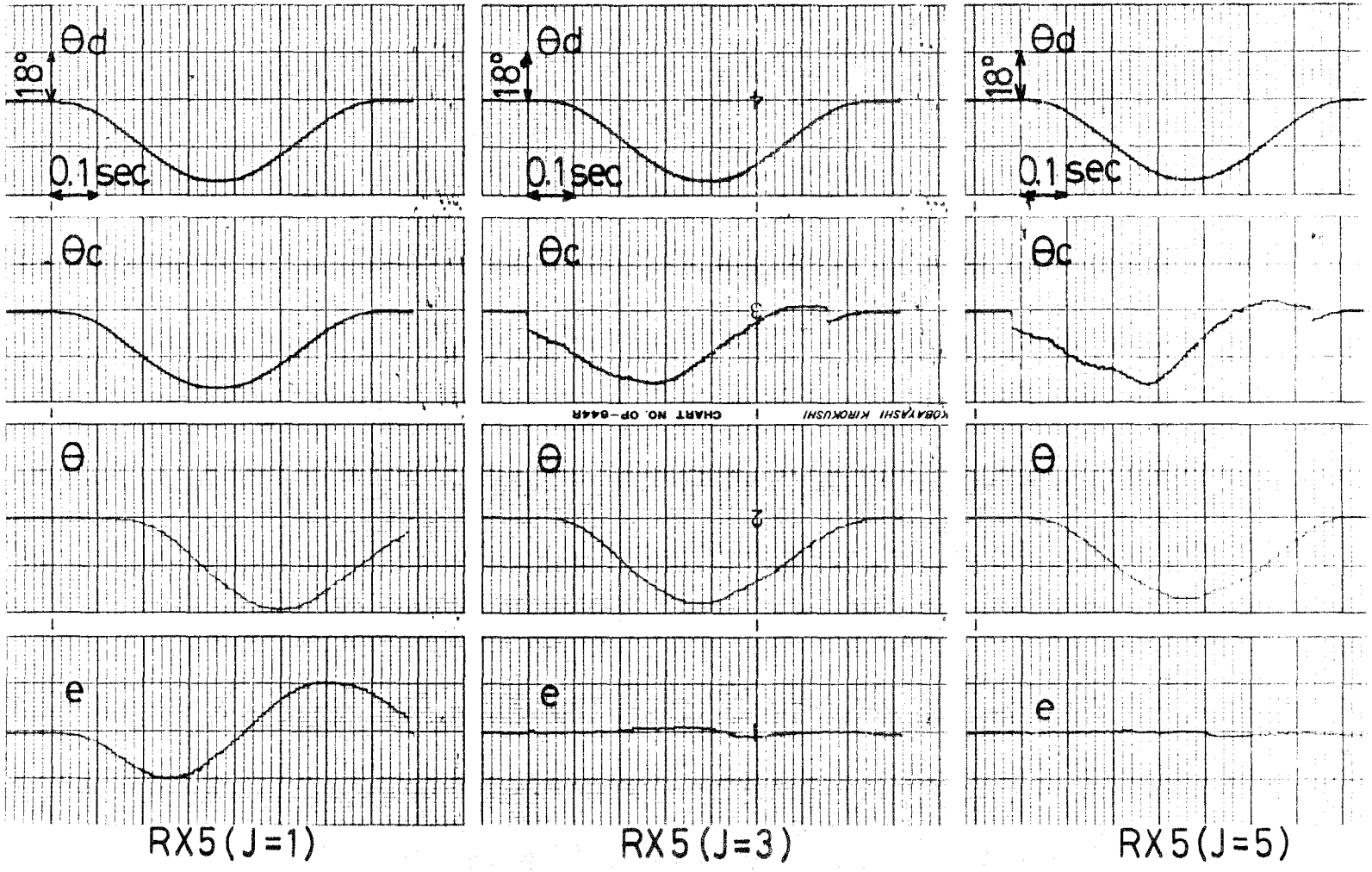


Fig.9-4 Experimental Result of the Biped Locomotion Robot

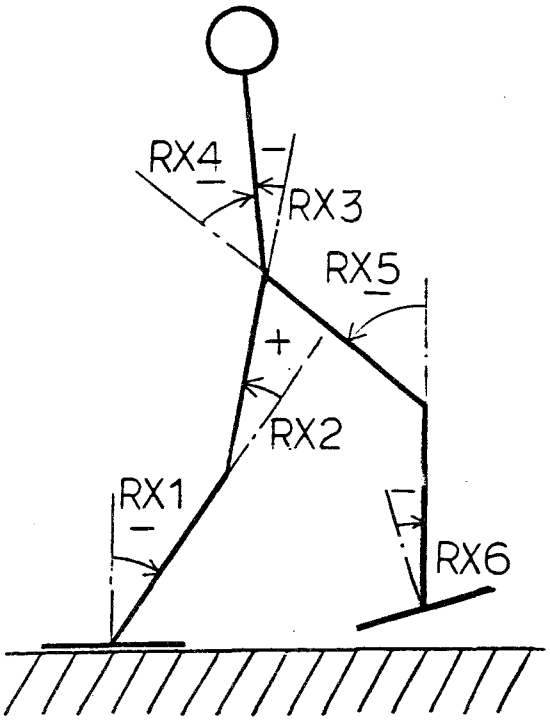


Fig.9-5

Definition of the Variables

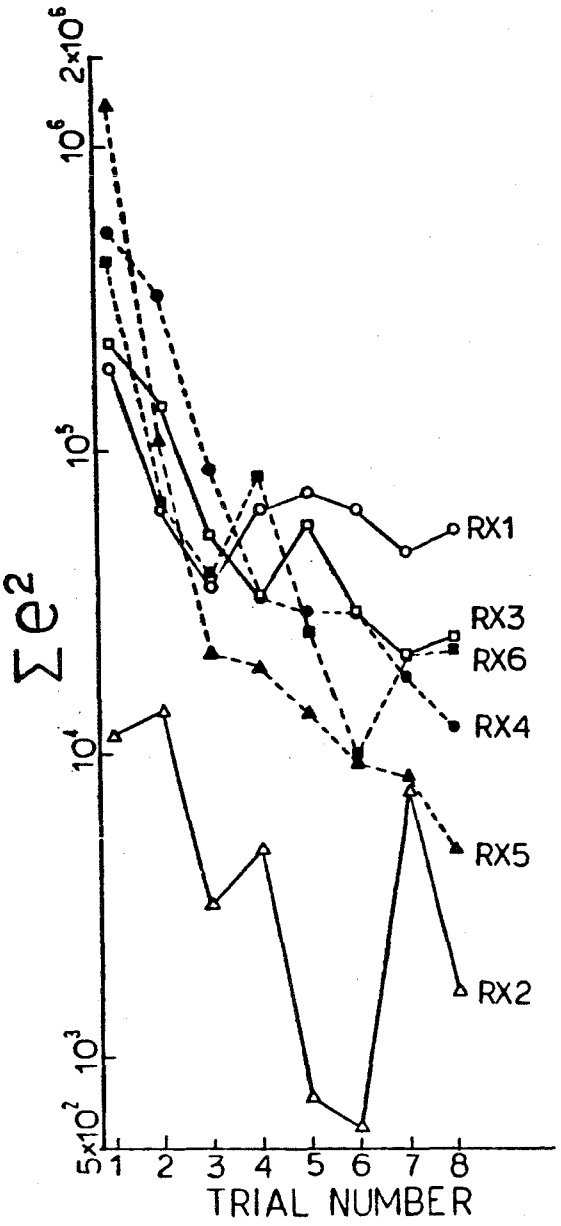


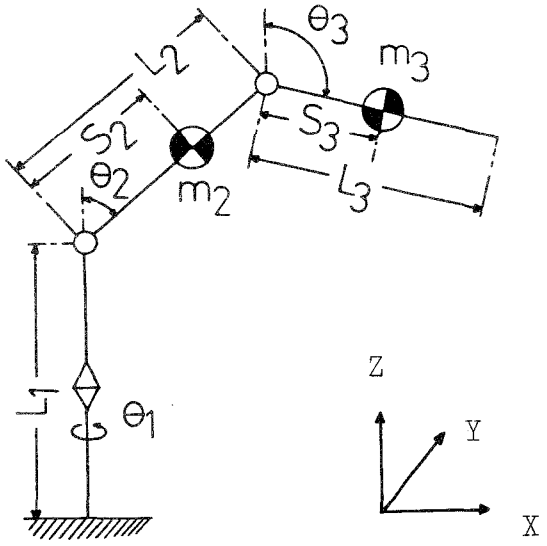
Fig.9-6

Experimental Result of Six degrees of Freedom

９－３　ロボットマニピュレータ（１）

９－３－１実験システム及び シミュレーションモデル

本論文で提案する制御方式をFig.９－７に示されるような３自由度をもつロボットマニピュレータに適用する．ロボットの物理パラメータはTable９－１に示されるものでありその概形をPhoto.９－２に示す．全体の実験システムはFig.９－８に示される.ロボットの角変位，角速度は各々ポテンショメータ，タコジェネレータからＡ－Ｄ変換器を介してマイコンに入力され，制御量を計算後Ｄ－Ａ変換器，Ｄ．Ｃ．アンプを経てロボットのサーボモータに電力が供給される．



L:length
m:mass
S:distance from the joint to the center of gravity
I:inertia

Fig.9-7　Robot Manipulator with Three Degrees of Freedom

Table 9-1 Physical Parameters of the Robot Manipulator

Link No.	M[kg]	L[m]	S[m]	Ix[kgm^2]	Iy[kgm^2]	Iz[kgm^2]
1	--	0.3	--	--	--	0.059
2	1.8	0.305	0.15	0.03	0.022	0.0048
3	1.5	0.54	0.11	0.018	0.017	0.0033

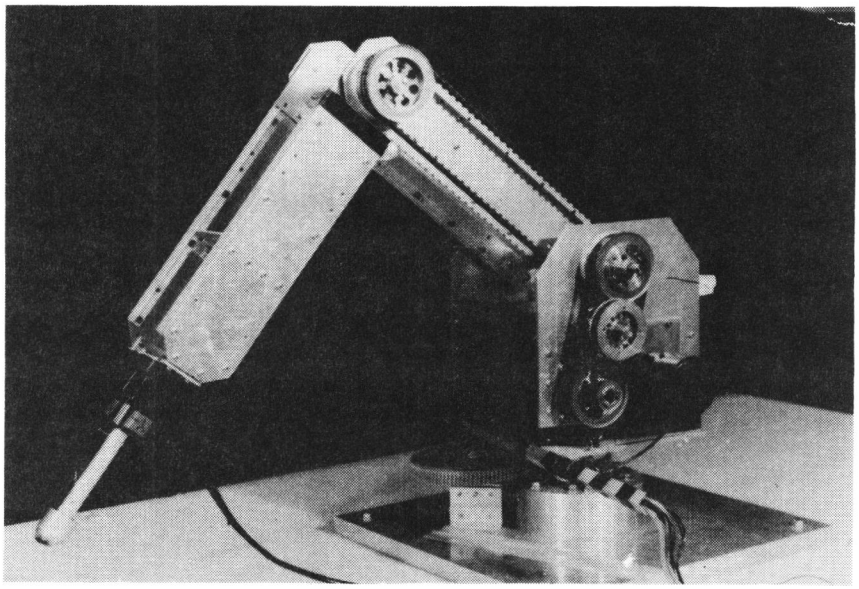


Photo 9-2 Photograph of the Robot Manipulator with Three Degrees of Freedom

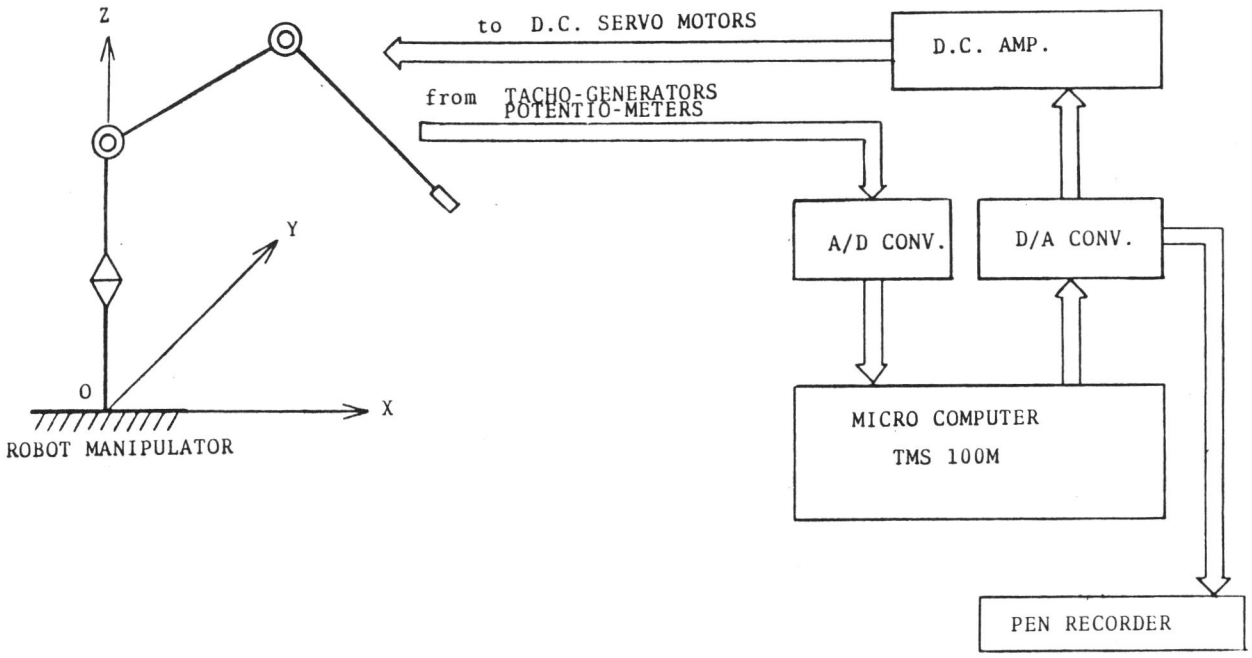


Fig.9-8 Experimental System

９－３－２　速度信号による入力の修正

実験結果

（関節角座標制御）

関節角で目標軌道を与えた場合についての実験結果をFig.９－９，Fig.９－１０に示す．このとき制御入力は（７－３）式で与えた．Fig.９－９では評価規範S_kのρの値をゼロ，重み行列Λ＝diag.(３２，３２，３２）として各関節の速度誤差の推移が示されている．また,Fig.９－１０は第１自由度θ_1について誤差が試行ごとに減少することを示している．図中θ_d：目標軌道，θ：システム応答，$\dot{\theta}_d-\dot{\theta}$：角速度誤差，$\theta_d-\theta$：角変位誤差，ＴＲＹ：試行回数を表わす．このときの実験条件をTable.９－２に示す．

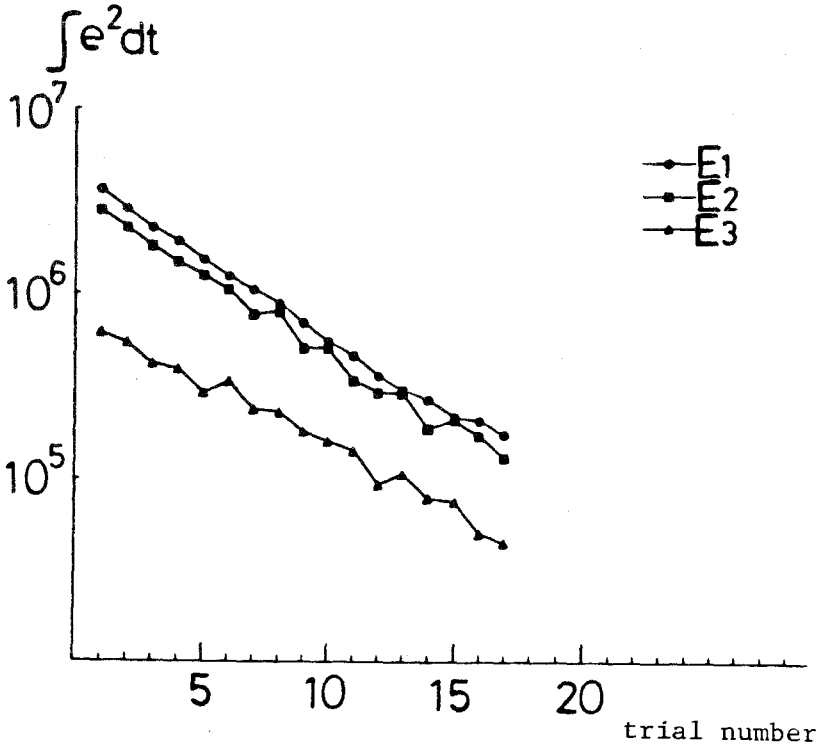


FIg.9-9 Experimental Result

Integrations of Squared Velocity Errors

$E_1=\dot{\theta}_{d1}-\dot{\theta}_1$, $E_2=\dot{\theta}_{d2}-\dot{\theta}_2$, $E_3=\dot{\theta}_{d3}-\dot{\theta}_3$

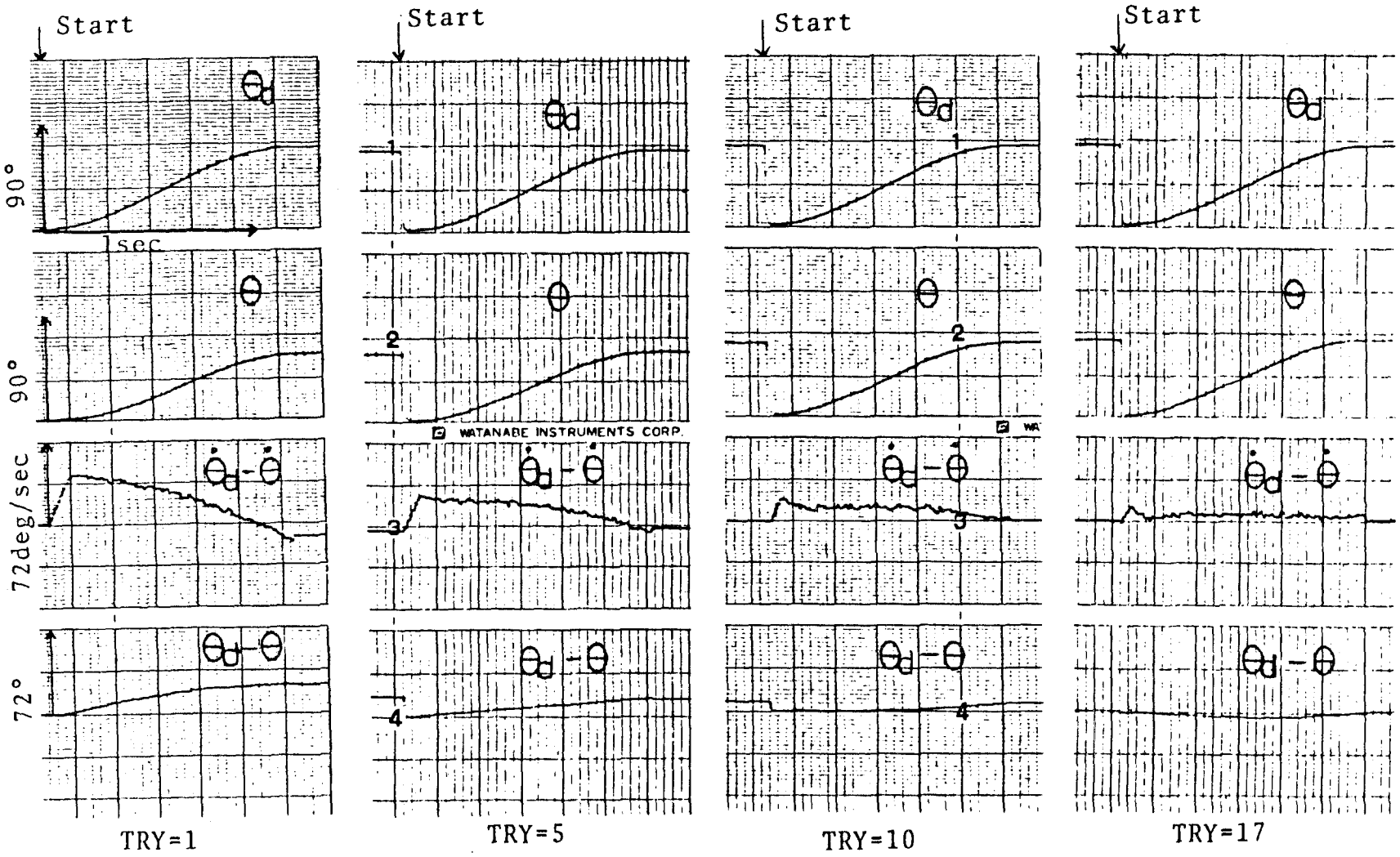


Fig.9-10 Experimental Result

System Responses and Errors

$\theta=\theta_1$, $\dot{\theta}=\dot{\theta}_1$, $\theta_d=\theta_{d1}$, $\dot{\theta}_d=\dot{\theta}_{d1}$

（作業座標制御）

作業座標制御では，第８章で説明したように関節角座標入力ｖ(t)を次のように与えた．

$$v(t)=\hat{g}(\theta)+J(\theta)^{T}[G_p(\omega_d(t)-\omega(t))+G_v(\dot{\omega}_d(t)-\dot{\omega}(t))]+u(t) \qquad (9-3)$$

θ：関節角座標，ω：カーテシアン座標，$J(\theta)$：ヤコビアン行列，

G_p，G_v：フイードバックゲイン

また，入力の修正は次のように構成した．

$$J^{T}(\theta_{k+1})u_{k+1}(t)=J^{T}(\theta_k)u_{k+1}(t)+J^{T}(\theta_{k+1})\Lambda(\dot{\omega}_d(t)-\dot{\omega}(t)) \qquad (9-4)$$

このときの実験結果をFig.９－１１，Fig.９－１２に示す．Fig.９－１１ではｘ，ｙ，ｚ方向の速度誤差をFig.９－９と同様の評価規範によって示す．またFig.９－１２ではｘ，ｙ，ｚ方向の位置誤差が繰り返しによって減少することを示している．このときの実験条件をTable９－２に示す．

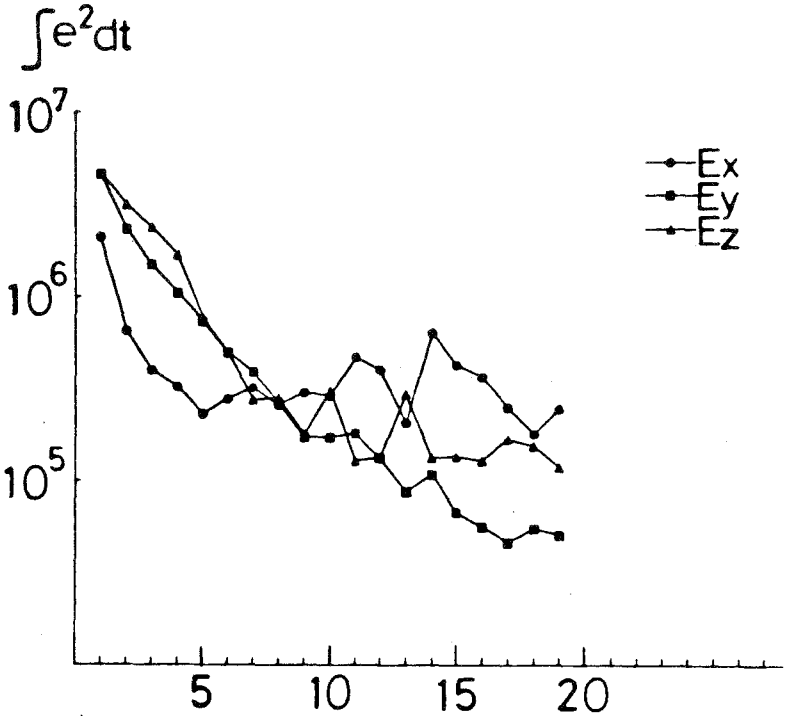


Fig.9-11 Experimental Result

Integrations of Squared Velocity Errors

$E_x=\dot{X}_d-\dot{X}$, $E_y=\dot{Y}_d-\dot{Y}$, $E_z=\dot{Z}_d-\dot{Z}$

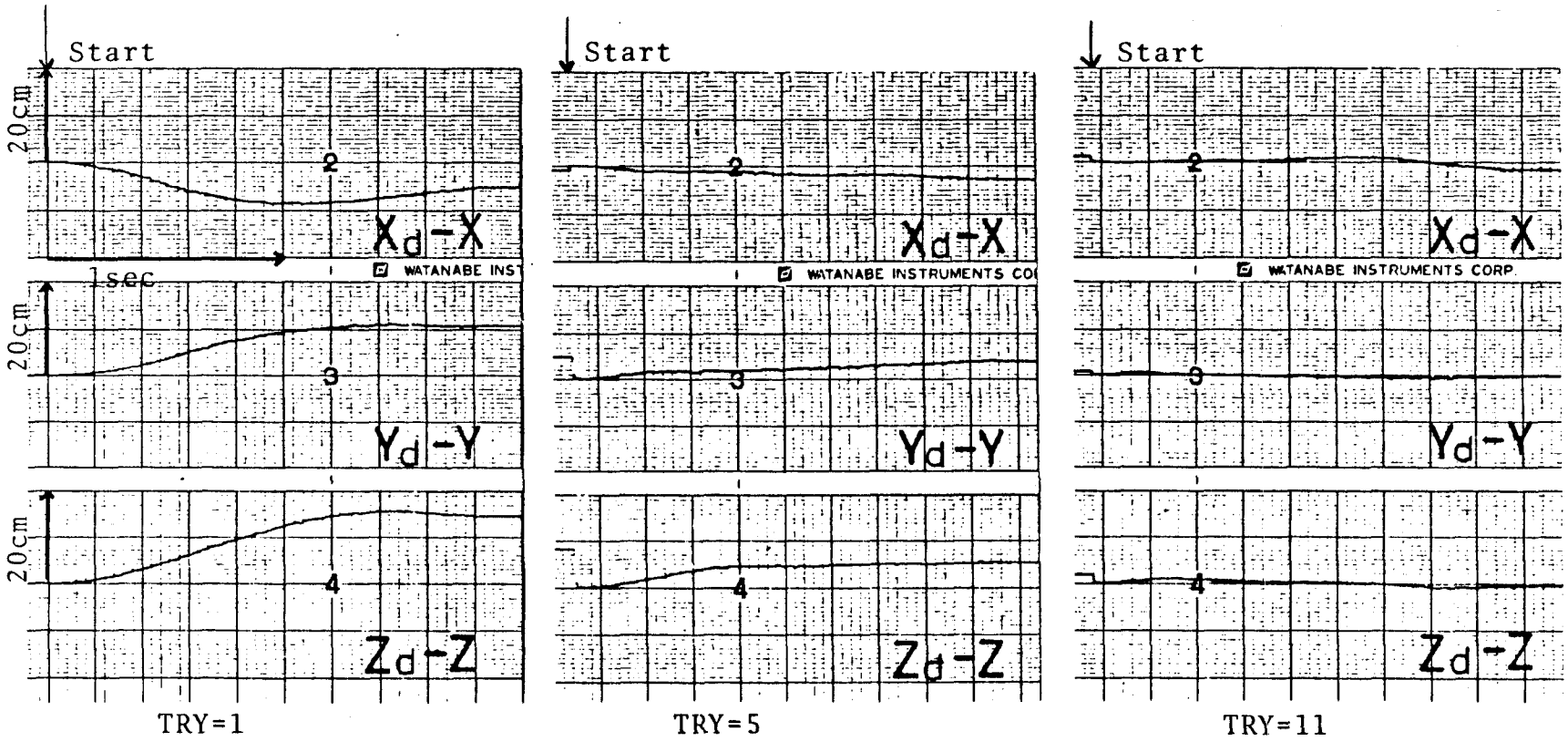


Fig.9-12 Experimental Result

Position Errors in the Task-oriented Coordinates

Table 9-2 Experimental Conditions

	Fig.9-9,Fig.9-10	Fig.9-11,Fig.9-12
sampling time	$\Delta t=6.0$ [msec]	$\Delta t=9.5$ [msec]
feedback gain	$K_p=\mathrm{diag.}(64,64,64)$ [Nm/deg] $K_v=\mathrm{diag.}(32,32,32)$ [Nmsec/deg]	$G_p=\mathrm{diag.}(64,64,64)$ [N/m] $G_v=\mathrm{diag.}(32,32,32)$ [Nsec/m]
Initial condition	$(\theta_1(0),\theta_2(0),\theta_3(0))=(0,0,0)$ [deg] $(\dot{\theta}_1(0),\dot{\theta}_2(0),\dot{\theta}_3(0))=(0,0,0)$ [deg/sec]	$(X(0),Y(0),Z(0))=(0,0,0)$ [m] $(\theta_1(0),\theta_2(0),\theta_3(0))=(30,60,150)$ [deg] $(X(0),Y(0),Z(0))=(0,0,0)$ [m/sec]
Time of termination	T=1.1 [sec]	T=1.8 [sec]
Desired trajectory	$\theta_{di}(t)=67.5(3-2(t/T))(t/T)^2$ [deg] {i=1,2,3}	$X_d(t)=Y_d(t)=Z_d(t)=0.25(3-2(t/T))(t/T)^2$ [m]
Input modification	$\alpha=1$ [sec/deg], $\beta=0$ [1/deg] $\Lambda=\mathrm{diag.}(32,32,32)$ [Nm]	$\alpha=1$ [sec/m], $\beta=0$ [1/m] $\Lambda=\mathrm{diag.}(32,32,32)$ [N]

９－３－３　加速度信号による入力の修正

加速度によって入力を修正する学習プロセスを同じロボットマニピュレータに適用する．このとき，出力を速度信号として，微分操作は（９－２）式の差分形式を採用した．このときのシミュレーション結果及び実験結果を以下に示す．なお，計算機シミュレーションはルンゲクッタ法によって求められた．

（ａ）シミュレーション結果

（関節角制御）

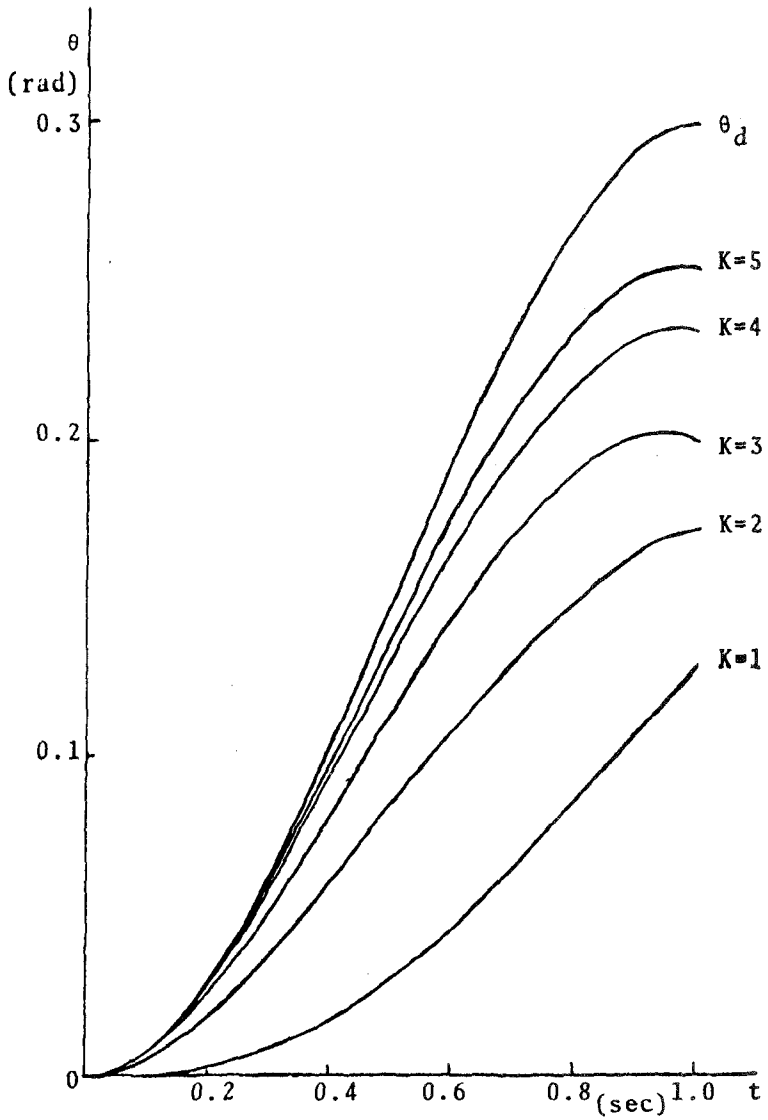
関節角で目標軌道が与えられた場合、特に、修正ゲインΛを適当に与えた場合についてのシミュレーション結果をFig.９－１３に示す.このときの条件をTable９－３に示す．図中ｋは試行回数を表わす．また,ロボットの物理パラメータはTable９－１に示されるものを使用した．次に，修正ゲインΛを理想的に慣性行列そのものを与えた場合のシミュレーション結果をFig.９－１４に示す．Fig.９－１３及びFig.９－１４よりシミュレーションの条件は，第５章の理論的考察が示唆する如く慣性行列そのものを修正ゲインとした場合の方が早く目標軌道に収束することがわかる．

（作業座標制御）

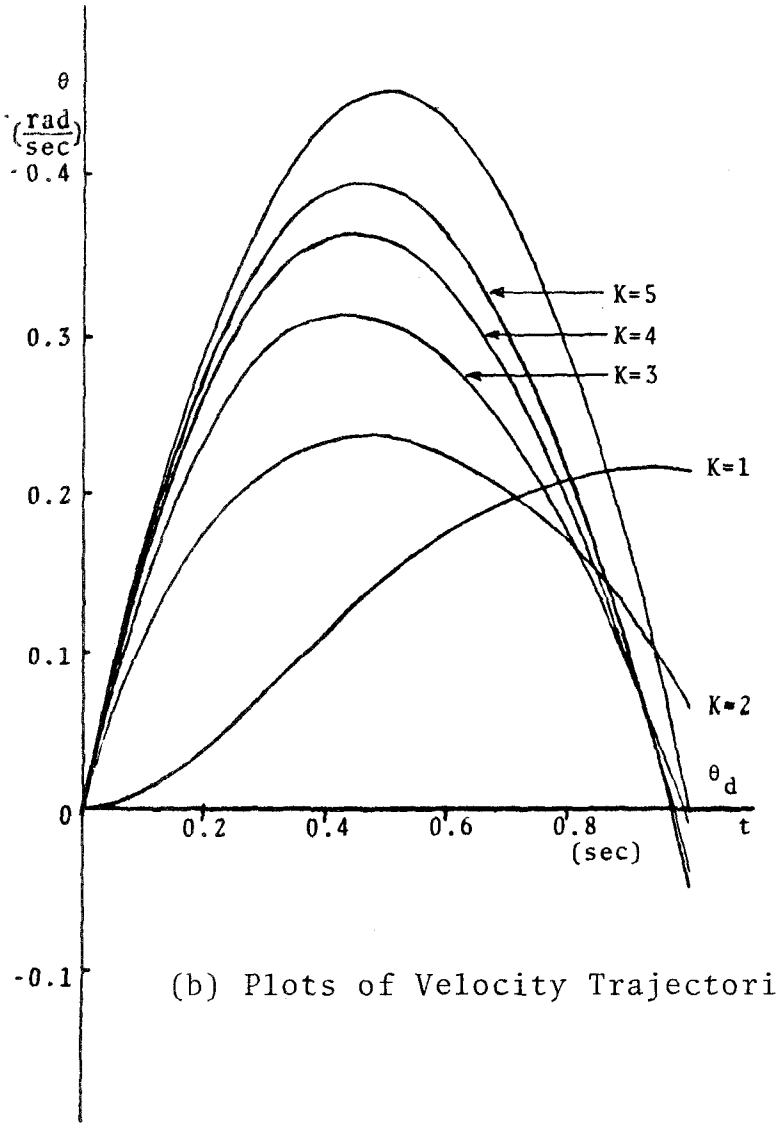
関節角制御の場合と同様に，修正ゲインΛを適当に与えた場合をFig.９－１５に，理想的なものを与えた場合をFig.９－１６に示す．また,シミュレーションの条件はTable９－３に示す．

（ｂ）実験結果

関節角で目標軌道が与えられた場合の実験結果をFig.９－１７に示す．このときの実験条件をTable９－４に示す．

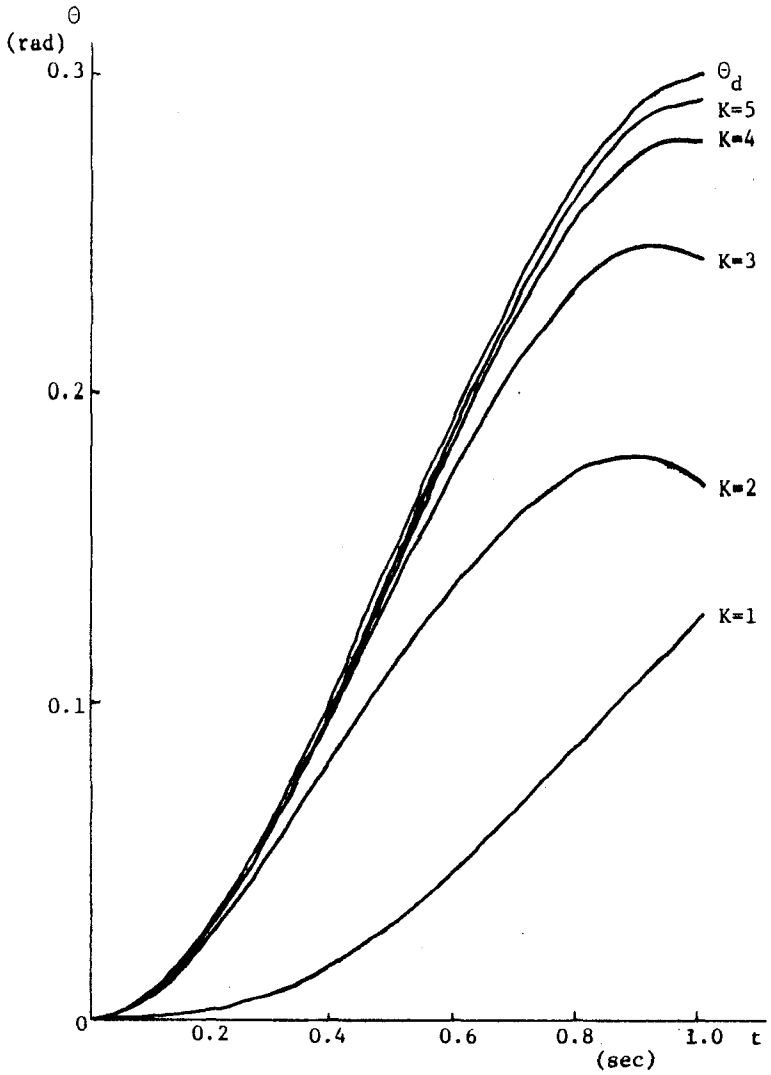


(a) Plots fo Position Trajectories

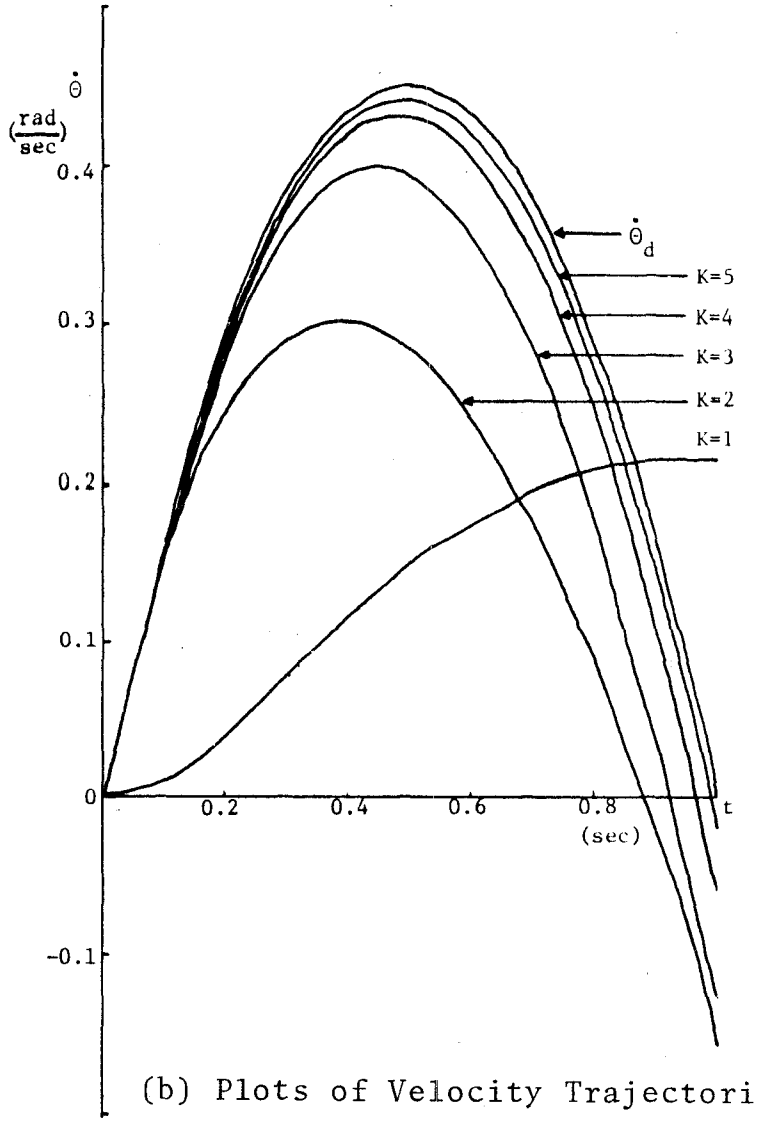


(b) Plots of Velocity Trajectories

Fig.9-13 Simulation Result in Joint-angle Coordinates
The gain matrix Λ is constant.



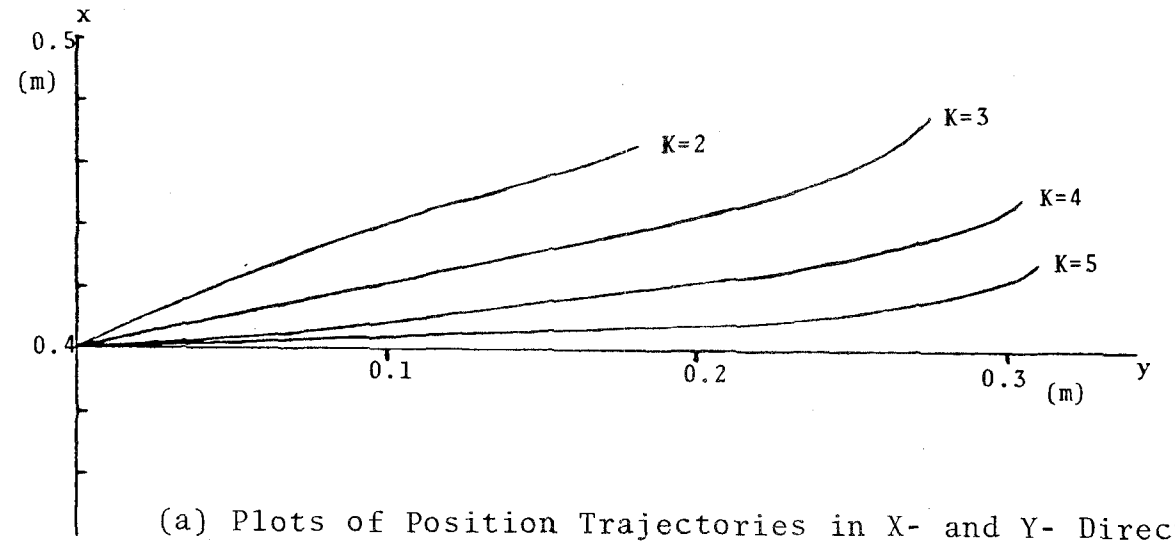
(a) Plots of Position Trajectories



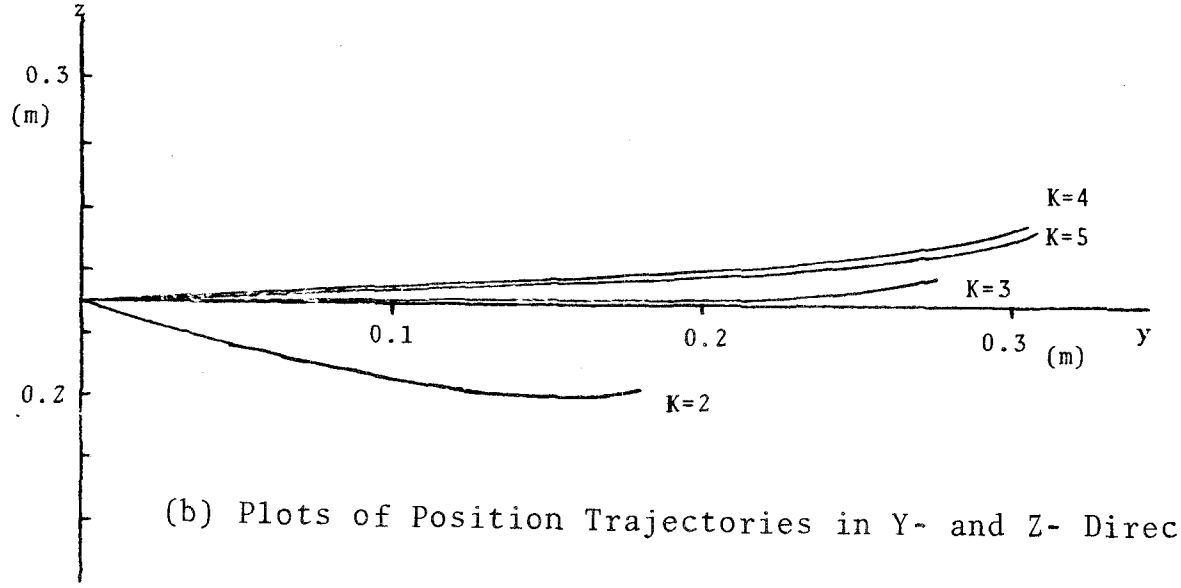
(b) Plots of Velocity Trajectories

Fig.9-14 Simulation Result in Joint-angle Coordinats

The gain matrix $\Lambda=R(\theta_k(t))$



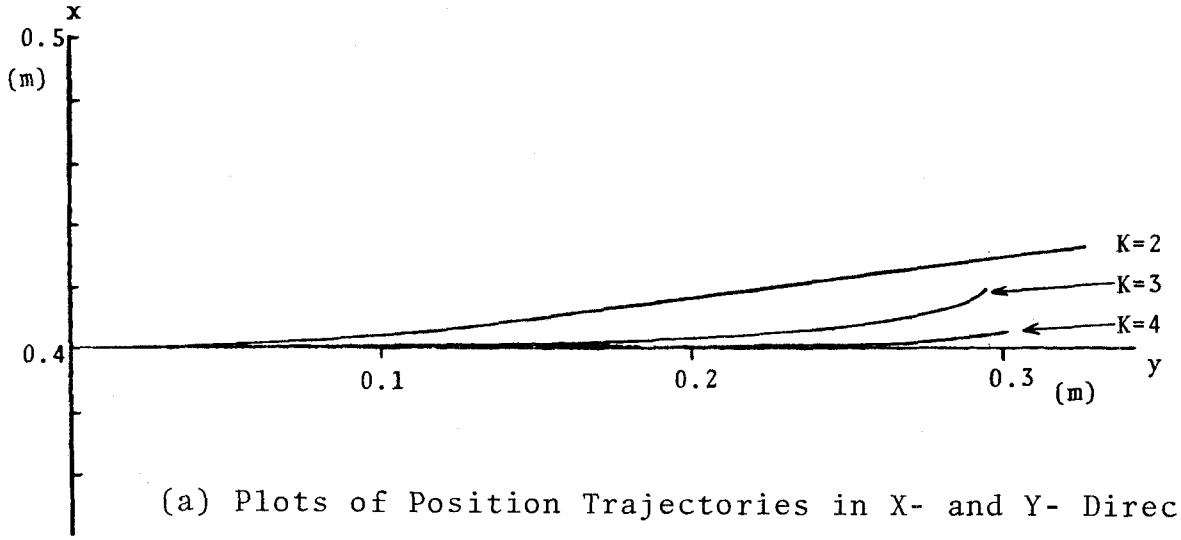
(a) Plots of Position Trajectories in X- and Y- Directions



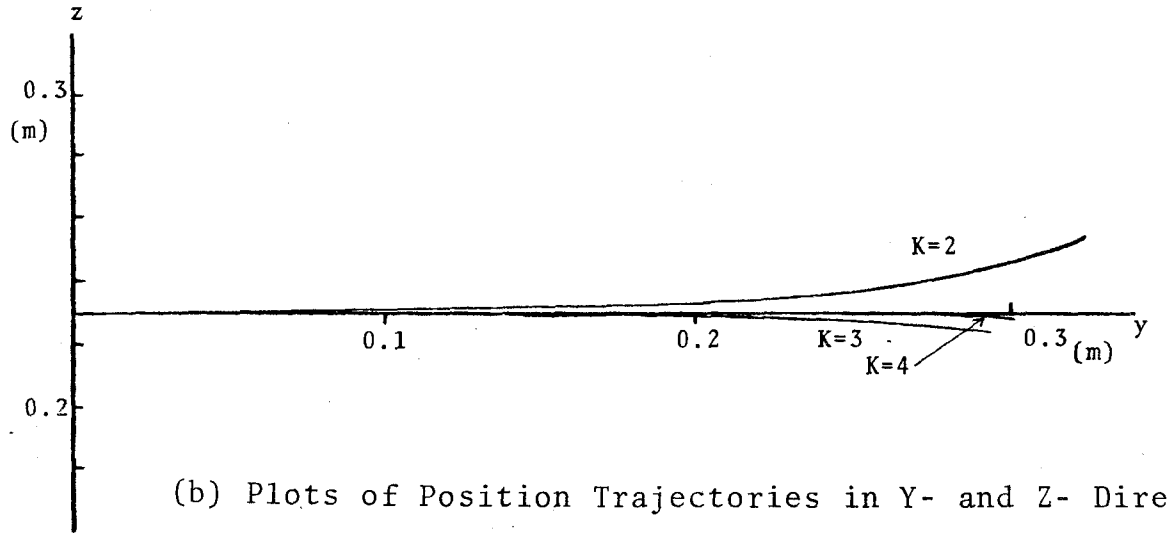
(b) Plots of Position Trajectories in Y- and Z- Directions

Fig.9-15 Simulation Result in Task-oriented Coordinates

The gain matrix Λ is constant.



(a) Plots of Position Trajectories in X- and Y- Directions



(b) Plots of Position Trajectories in Y- and Z- Directions

Fig.9-16 Simulation Result in Task-oriented Coordinates

The gain matrix $\Lambda=R(\theta_k(t))J^{-1}(\theta_k(t))$

Table 9-3 Experimental Conditions

	Fig.9-13,	Fig.9-14	Fig.9-15,	Fig.9-16
sampling time	$\Delta t=5.0$ [msec]		$\Delta t=5.0$ [msec]	
feedback gain	$K_p=\mathrm{diag.}(1.0,1.0,1.0)$ [Nm/rad] $K_v=\mathrm{diag.}(1.0,1.0,1.0)$ [Nmsec/rad]		$G_p=\mathrm{diag.}(1.0,1.0,1.0)$ [N/m] $G_v=\mathrm{diag.}(1.0,1.0,1.0)$ [Nsec/m]	
Initial condition	$(\theta_1(0),\theta_2(0),\theta_3(0))=(0,\pi/2,\pi)$ [rad] $(\dot{\theta}_1(0),\dot{\theta}_2(0),\dot{\theta}_3(0))=(0,0,0)$ [rad/sec]		$(X(0),Y(0),Z(0))=(0.4,\ 0,\ 0.23)$ [m] $(\theta_1(0),\theta_2(0),\theta_3(0))=(0,\ \pi/2,\ \pi)$ [rad] $(X(0),Y(0),Z(0))=(0,0,0)$ [m/sec]	
Time of termination	$T=1.0$ [sec]		$T=1.0$ [sec]	
Desired trajectory	$\theta_{di}(t)=p_i(3-2(t/T))(t/T)^2$ [rad] $p_1=0.3,\quad p_2=1.27,\quad p_3=2.03$ $\{i=1,2,3\}$		$Y_d(t)=0.4(3-2(t/T))(t/T)^2$ $X_d(t)=Y_d(t)=0$ [m]	
Input modification	$\Lambda=\mathrm{diag.}(0.5,0.5,0.1)$ [Nmsec^2/rad]	$\Lambda=R(\theta_k(t))$ [Nmsec^2/rad]	$\Lambda=\begin{bmatrix}-1.0, & 1.5, & 0\\ 0.2, & 0.1, & -1.8\\ -0.3, & -0.2, & -0.1\end{bmatrix}$ [Nsec^2]	$\Lambda=R(\theta_k(t))J^{-1}(\theta_k(t))$ [Nsec^2]

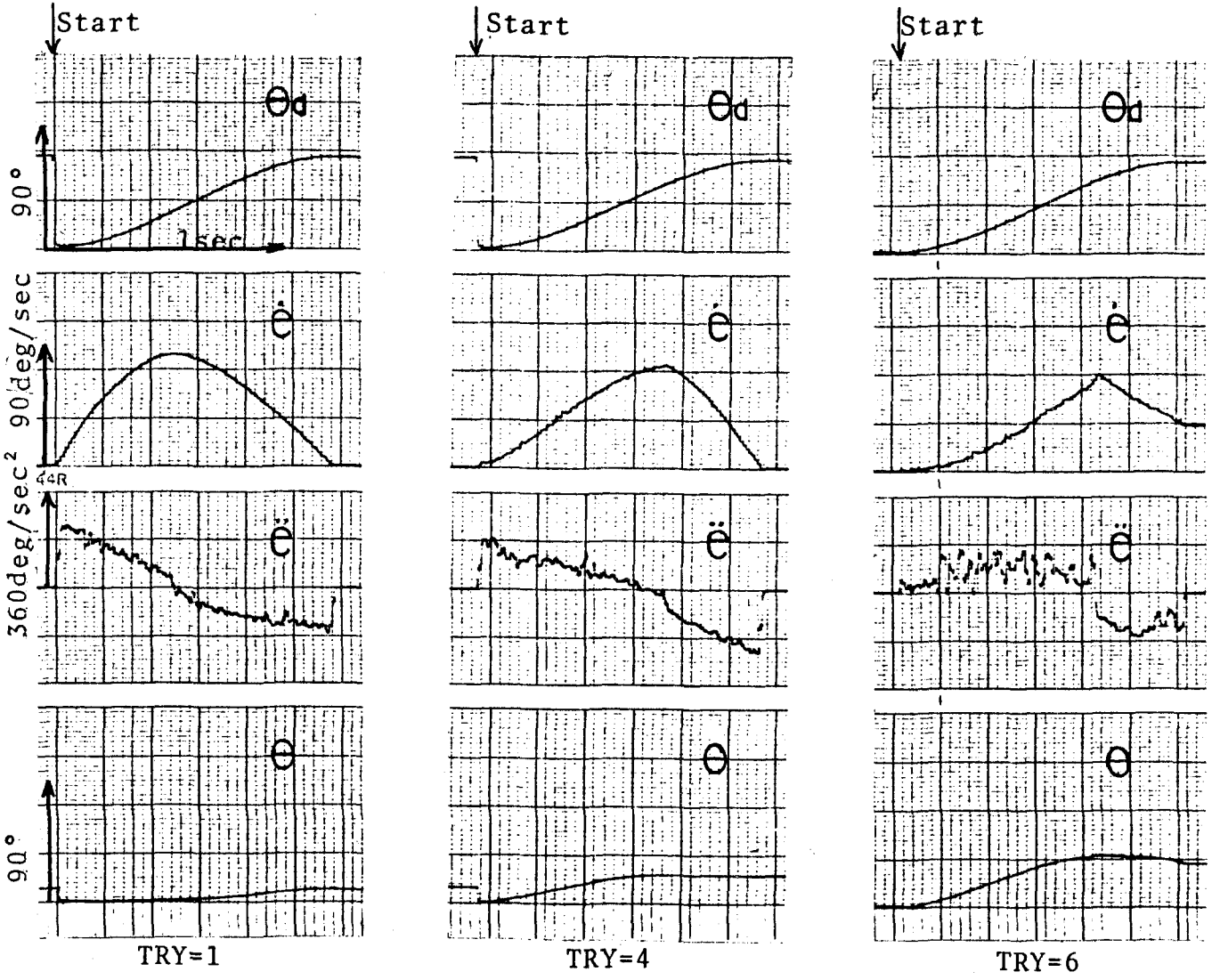


Fig.9-17-a Experimental Results in Joint-angle Coordinates

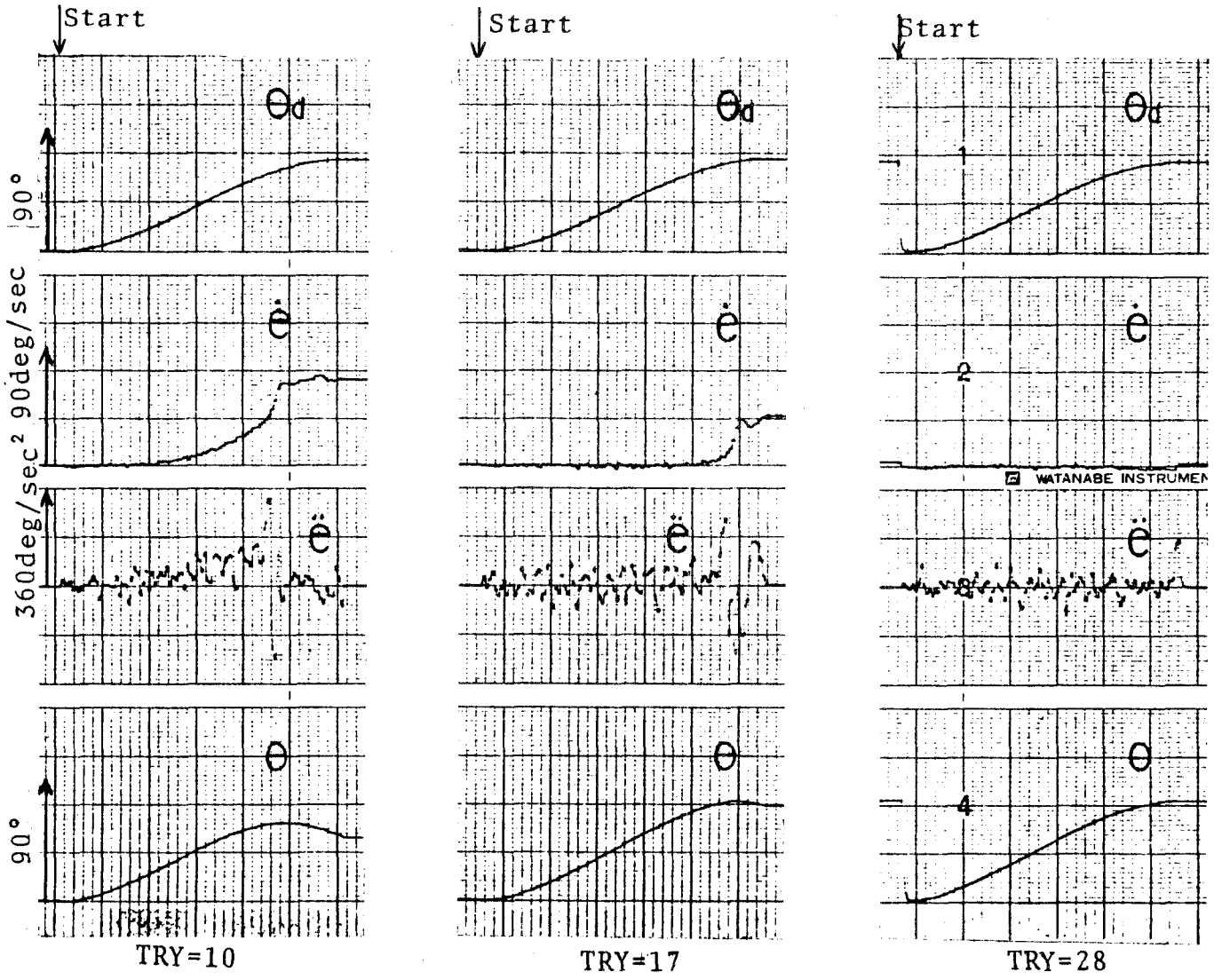


Fig.9-17-b Experimental Results in Joint-angle Coordinates

Table 9-4 Experimental Conditions

sampling time	$\Delta t=6.0$ [msec]
feedback gain	$K_p=\mathrm{diag.}(6.0,6.0,6.0)$ [Nm/deg] $K_v=\mathrm{diag.}(6.0,6.0,6.0)$ [Nmsec/deg]
Initial condition	$(\theta_1(0),\theta_2(0),\theta_3(0))=(0,\pi/2,\pi)$ [deg] $(\dot{\theta}_1(0),\dot{\theta}_2(0),\dot{\theta}_3(0))=(0,0,0)$ [deg/sec]
Time of termination	T=1.2 [sec]
Desired trajectory	$\theta_{di}(t)=p_i(3-2(t/T))(t/T)^2$ [deg] $p_1=67.5,\ p_2=45,\ p_3=67.5$ {i=1,2,3}
Input modification	$\Lambda=\mathrm{diag.}(0.10,\ 0.12,\ 0.054)$ [Nmsec^2/deg]

９－４．ロボットマニピュレータ（２）

９－４－１．実験システム

前節で使用したロボットマニピュレータ（Fig.９－７）と同じ構造の３自由度のものを使って，実際にロボットに図を描かせる実験を行なった．ただし，この場合のロボットの物理パラメータと概形を各々Table９－５とPhoto９－３に示す．全体の実験システムは速度信号を測定するタコジェネレータが取りつけられていないことを除けばＦｉｇ．９－８に示されるものと同じである．従って，ロボットの運動は各関節に取りつけられたポテンされるものと同じである．そこで，各関節に取りつけられたポテンショメータから関節角変位を測定し，角速度は（９－２）式の差分形式で求めた．また，手先にはロボットに図を描かせる目的で，ペンが装着されている．制御入力は（９－４）式のように構成し，作業座標で目標軌道を与えた．また，入力の修正には第４章で述べた位相進みの方式を用いた．

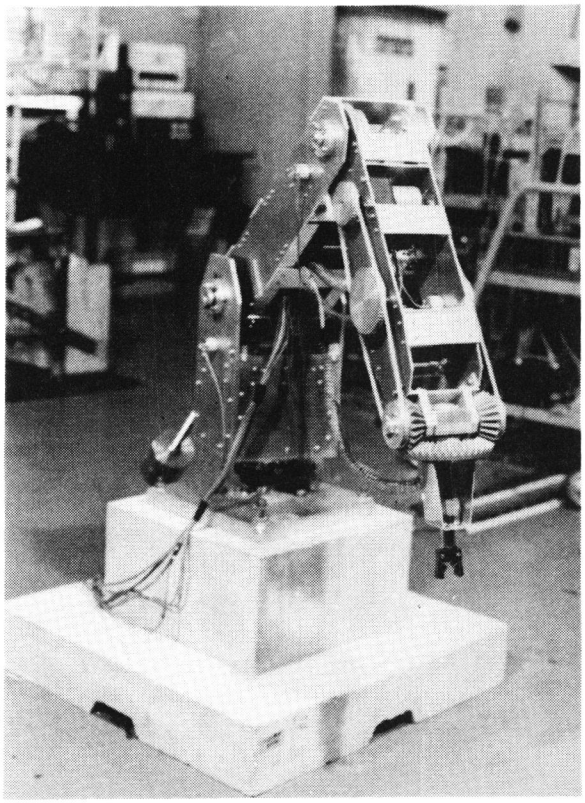


Photo.9-3
Photograph of
the Robot Manipulators

Link No.	m(kg)	l(m)	s(m)	I_x(kgm^2)	I_y(kgm^2)	I_z(kgm^2)
1	———	0.63	———	———	———	0.022
2	5.4	0.40	0.13	0.062	0.052	0.013
3	3.6	0.40	0.17	0.046	0.041	0.0048

Table 9-5 Physical Parameters of the Robot Manipulator

９－４－２．円軌道

Fig.９－１８に示されるような４５度平面上に直径２０ｃｍの円を描く目標軌道を与えた．このときの実験結果をFig.９－１９，Fig.９－２０，Fig.９－２１に示す．Fig.９－１９では各方向の位置と速度の誤差の２乗積分値を表わし，Fig.９－２０では各方向の目標位置パターンと実際の応答及び誤差の推移を示している．また，Fig.９－２１に実際にロボットが描いた円軌道を示す．なお，この実験条件をTable９－５に記載する．

９－４－３．直線軌道

ｘ－ｙ平面（水平面）上でｙ方向に２５ｃｍの直線を描く目標軌道を与えた．このとき実際にロボットが描いた直線をＦｉｇ．９－２２に示す．また，実験条件をTable９－５に示す．

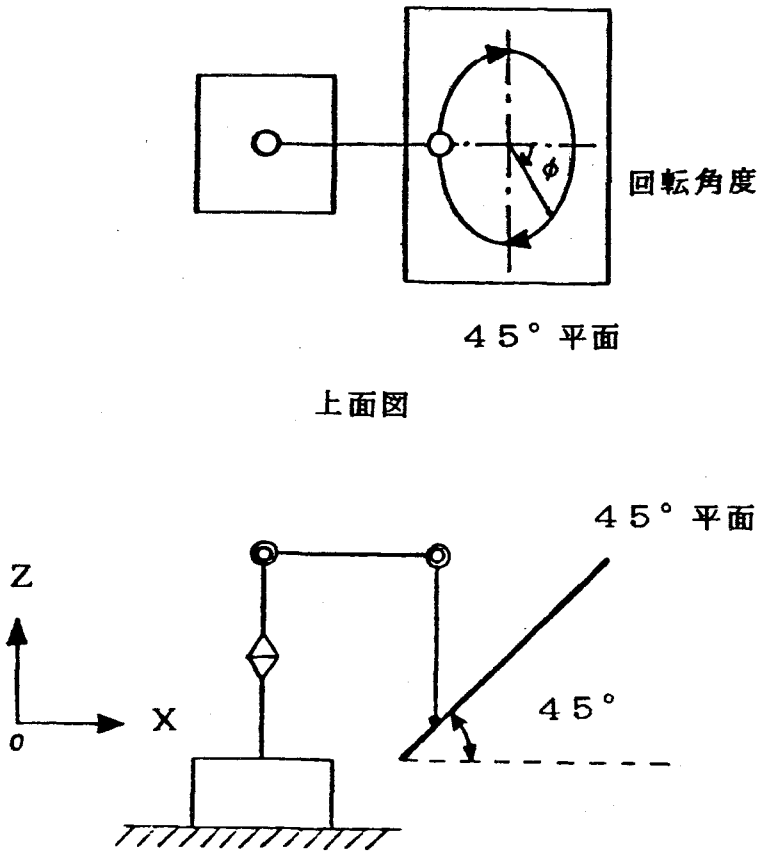


Fig.9-18 Structure of Experimental System

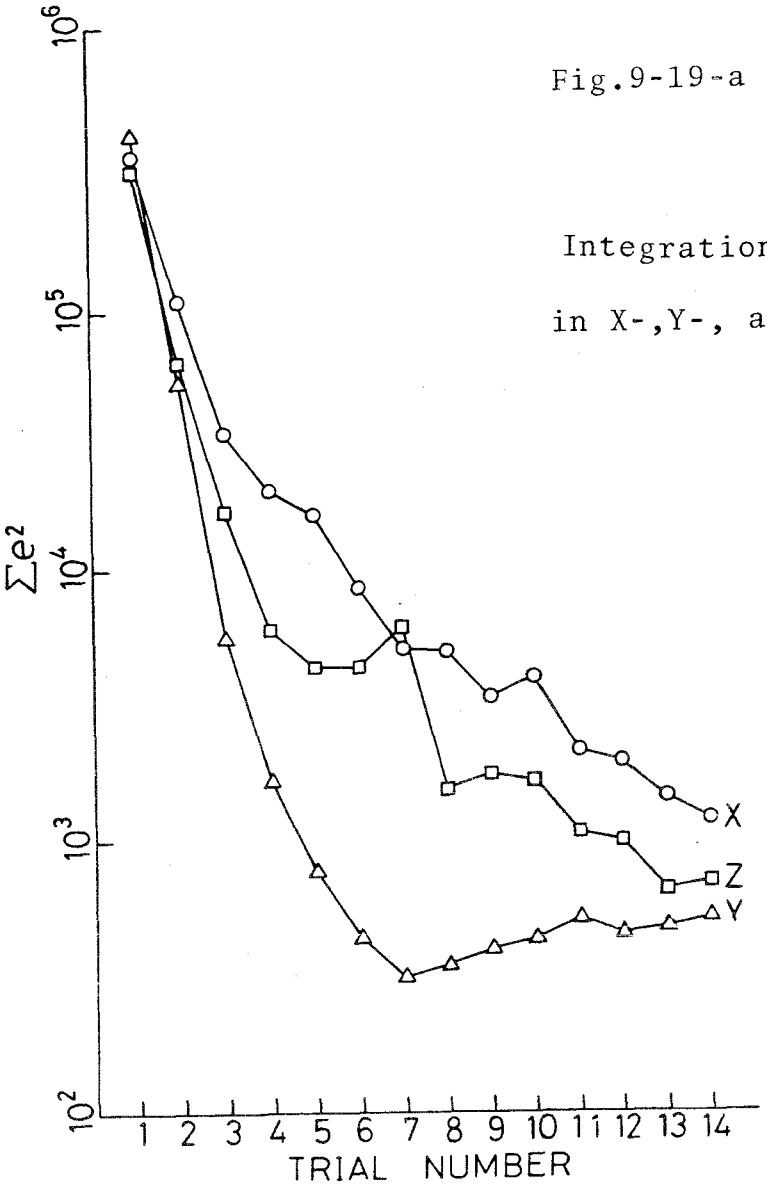


Fig.9-19-a Experimental Result

Integrations of Position Errors

in X-,Y-, and Z-Directions

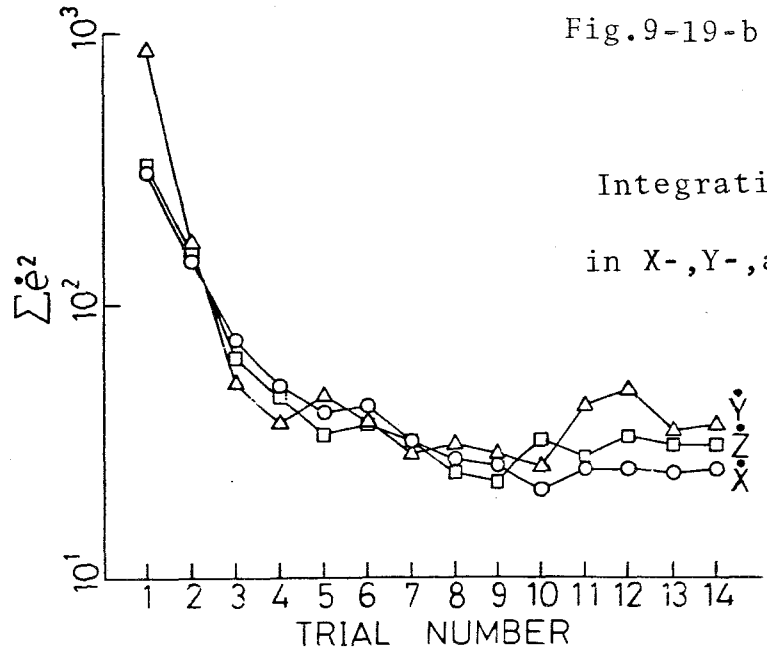
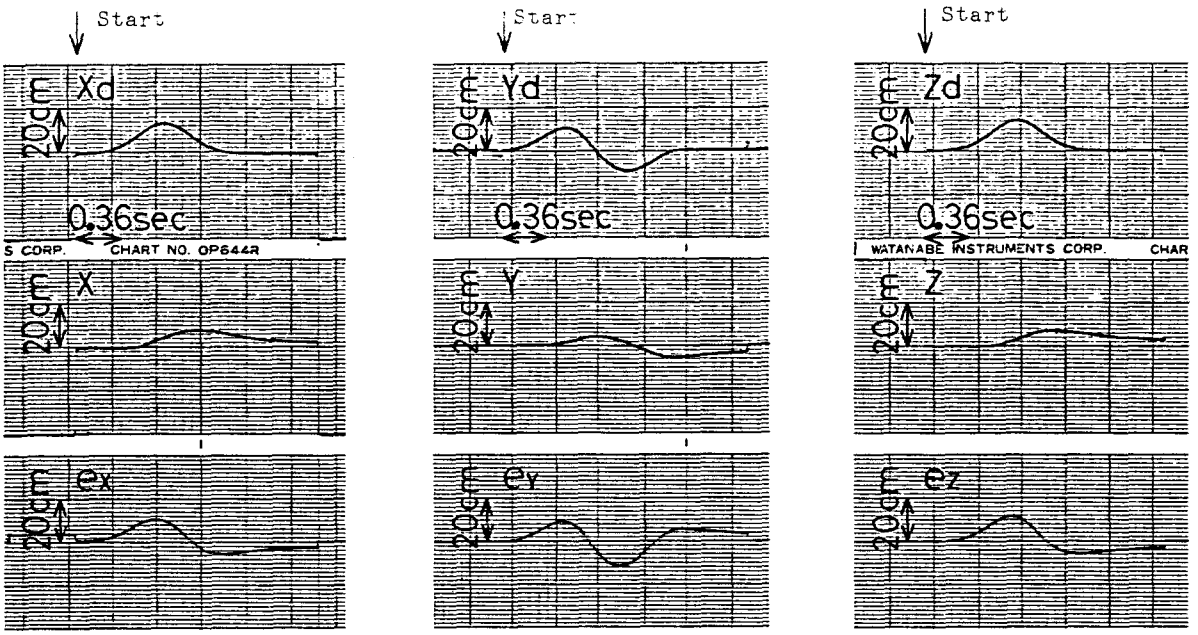


Fig.9-19-b Experimental Result

Integrations of Velocity Errors

in X-,Y-,and Z-Directions



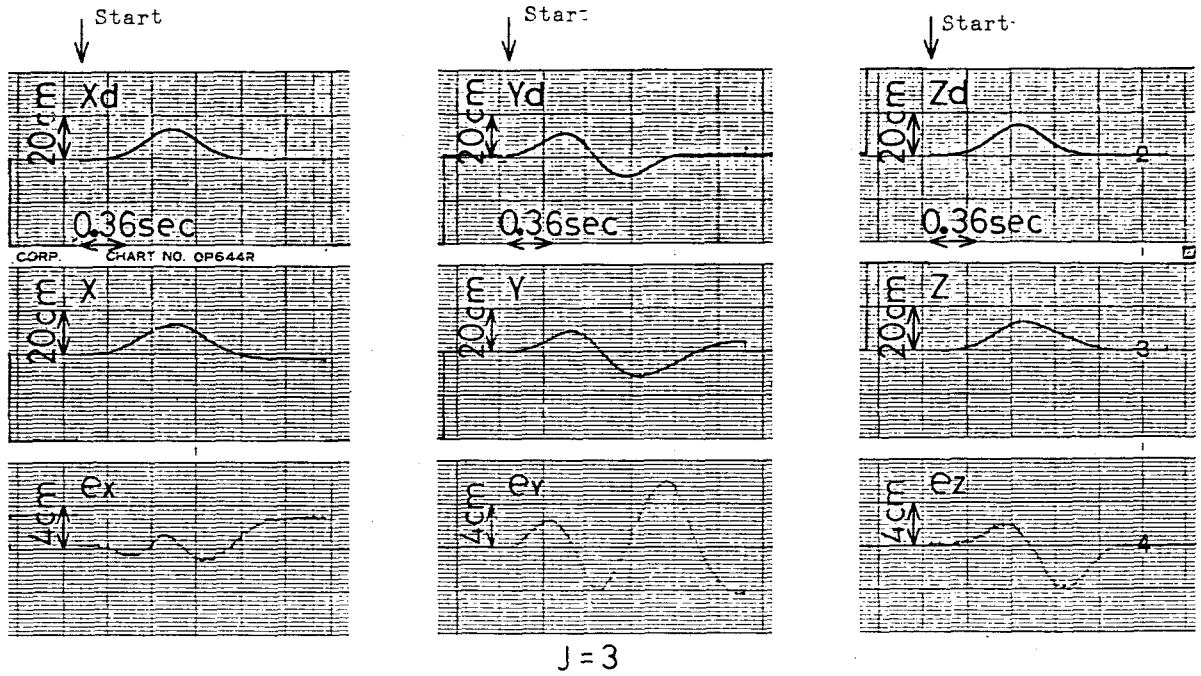
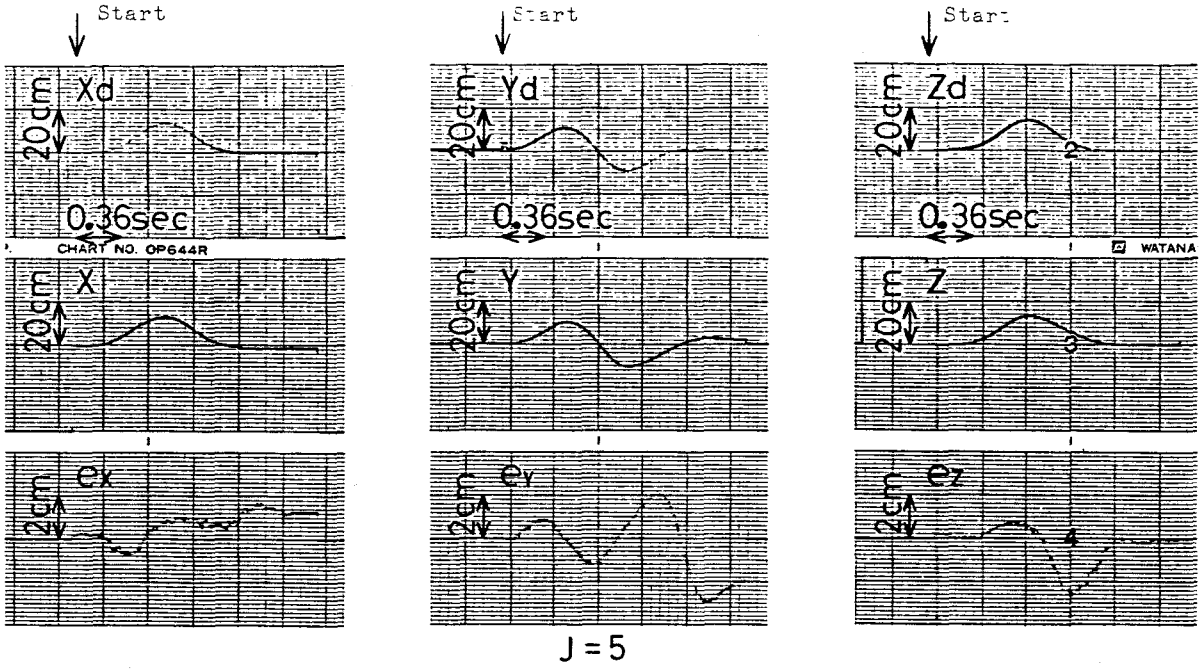


Fig.9-20 Experimental Result

Position Trajectories and Errors in X-,Y-, and Z-Directions


Start
Xd
20cm
0.36sec
CHART NO. OP644R
X
ex
2cm
Start
Yd
0.36sec
Y
ey
Start
Zd
0.36sec
WATANA
Z
ez
J=5

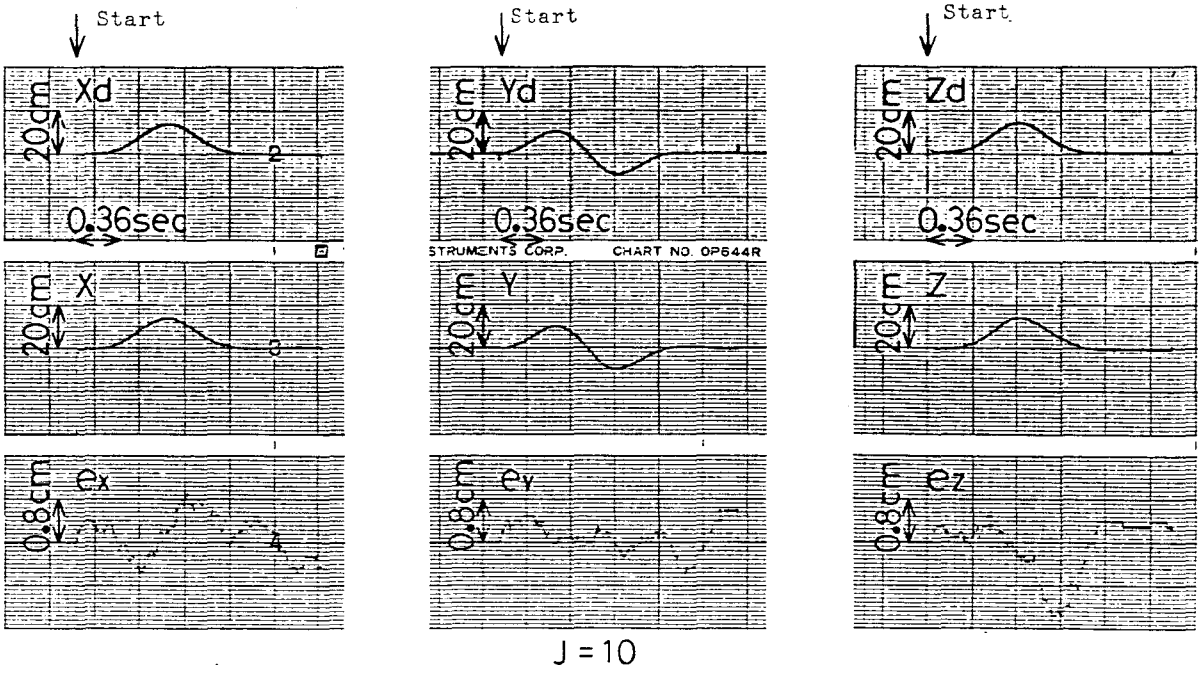
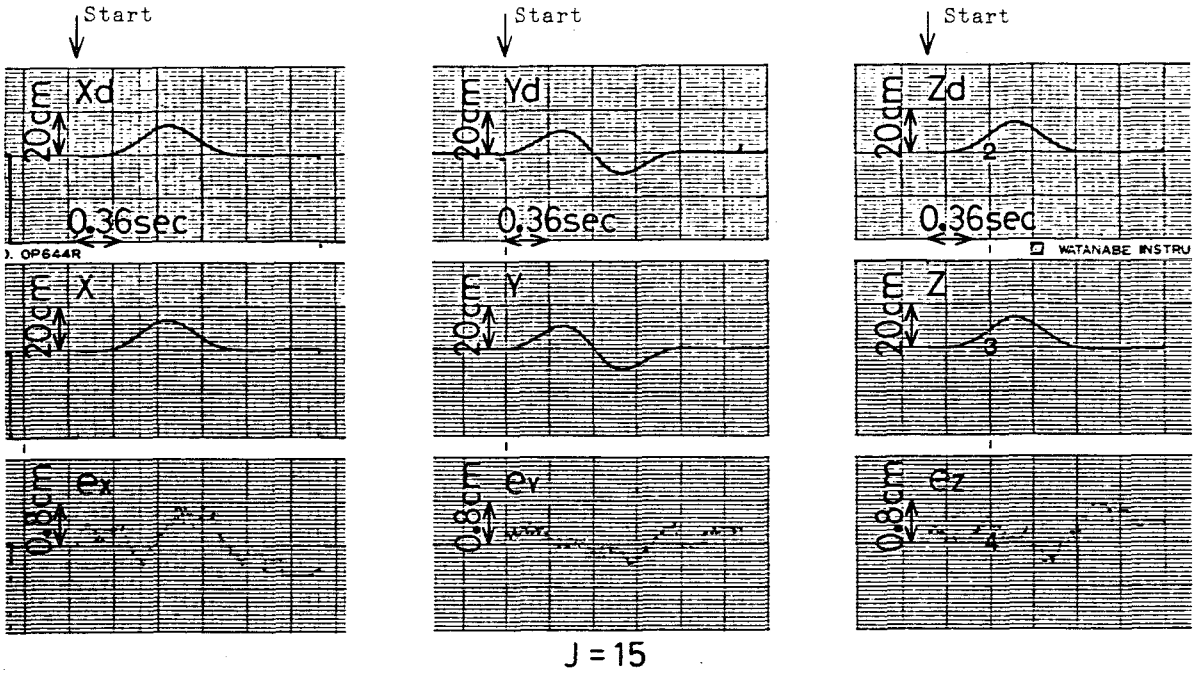

Start
Xd
20cm
0.36sec
X
ex
0.8cm
Start
Yd
0.36sec
STRUMENTS CORP.
CHART NO. OP644R
Y
ey
Start
Zd
0.36sec
Z
ez
J=10

Fig.9-20


Start
Xd
Yd
Zd
X
Y
Z
ex
ey
ez
0.36sec
20cm
0.8cm
WATANABE INSTRU
J = 15

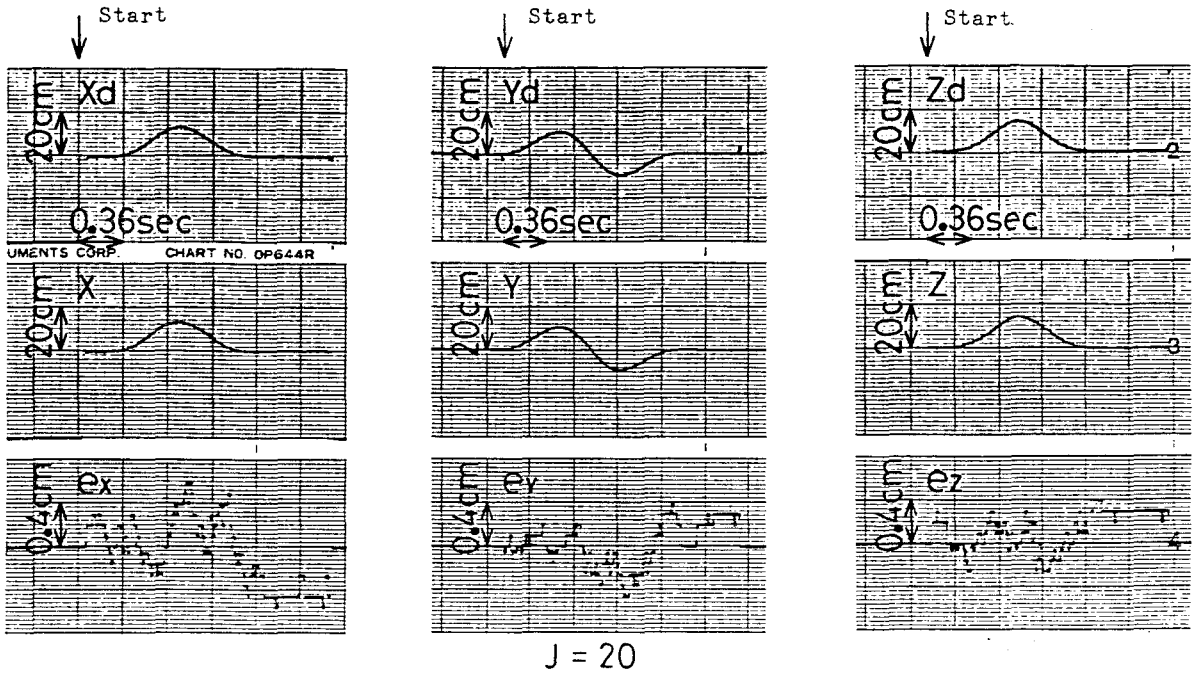

Start
Xd
Yd
Zd
X
Y
Z
ex
ey
ez
0.36sec
20cm
0.4cm
CHART NO. OP644R
J = 20

Fig.9-20

--------- : Desired Trajectory

———— : Practical Trajectory

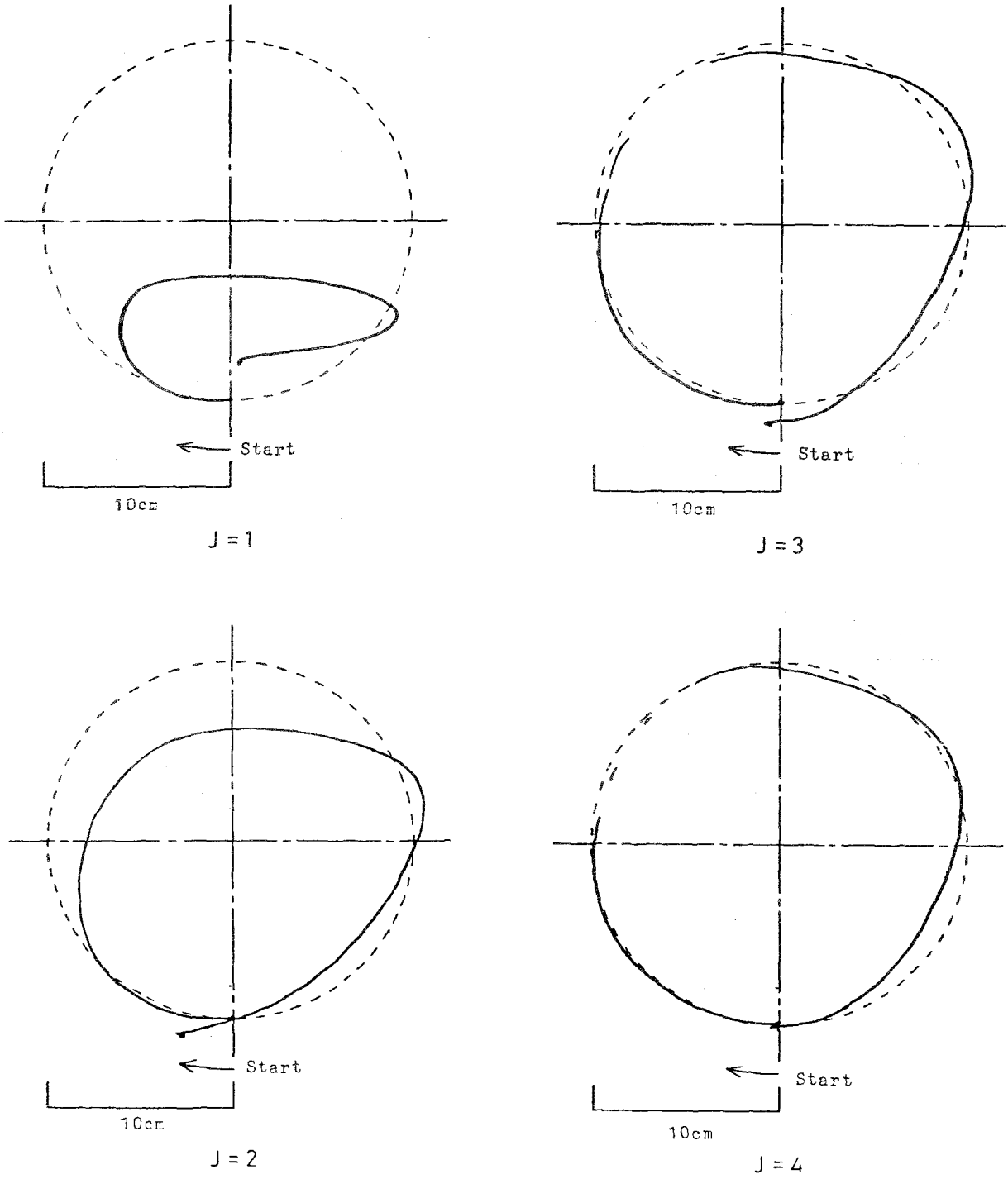


Fig.9-21 Experimental Result

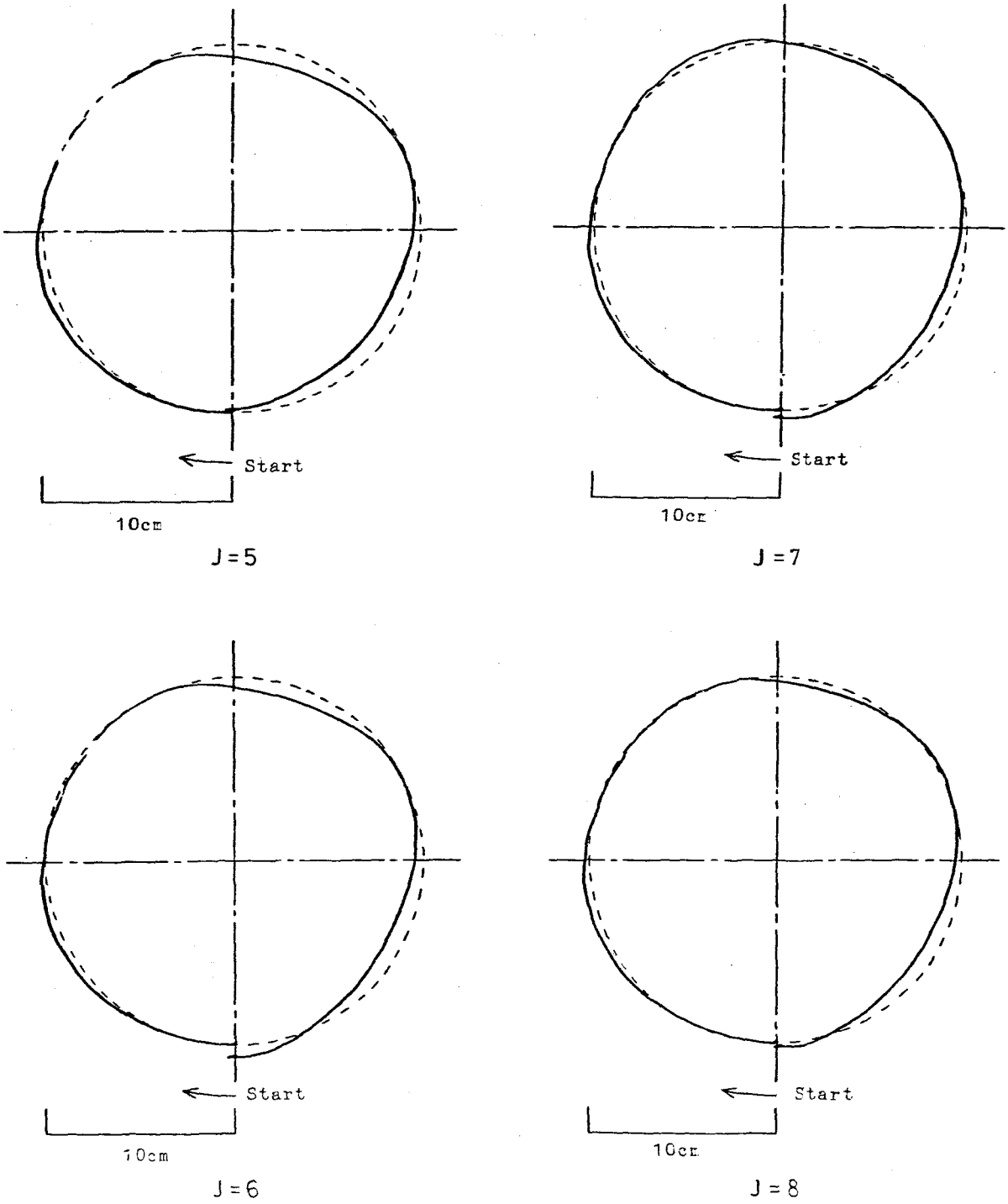
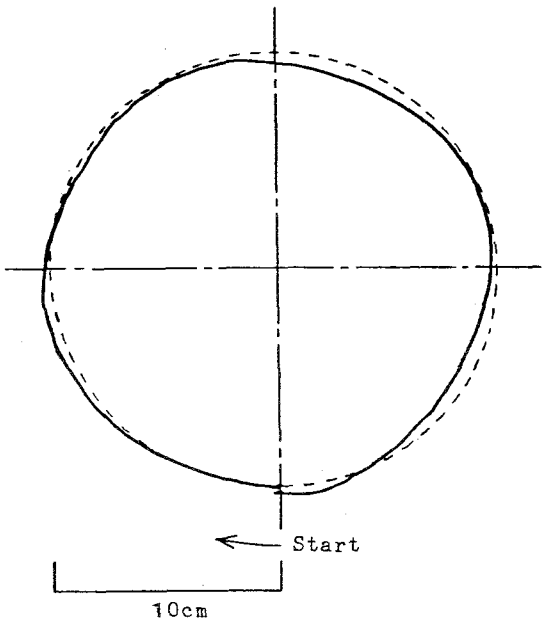
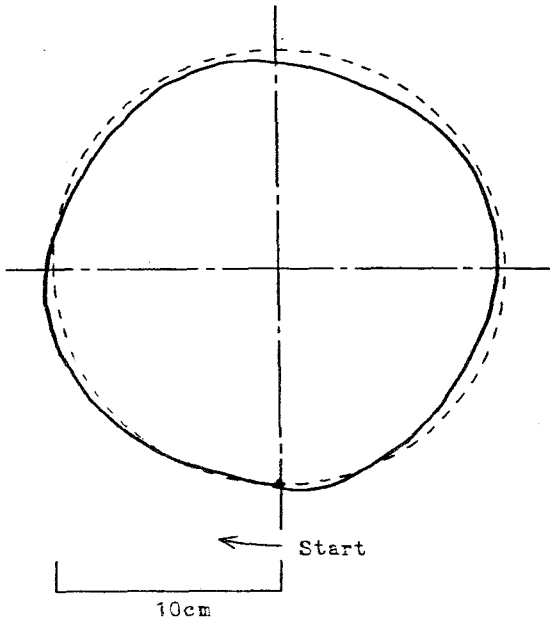


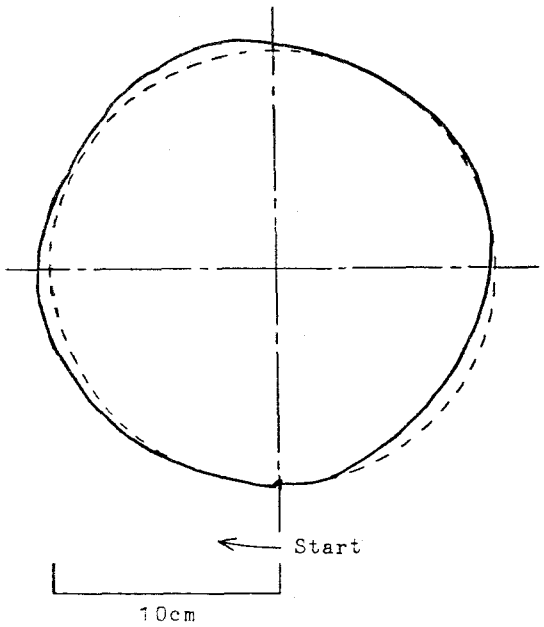
Fig.9-21



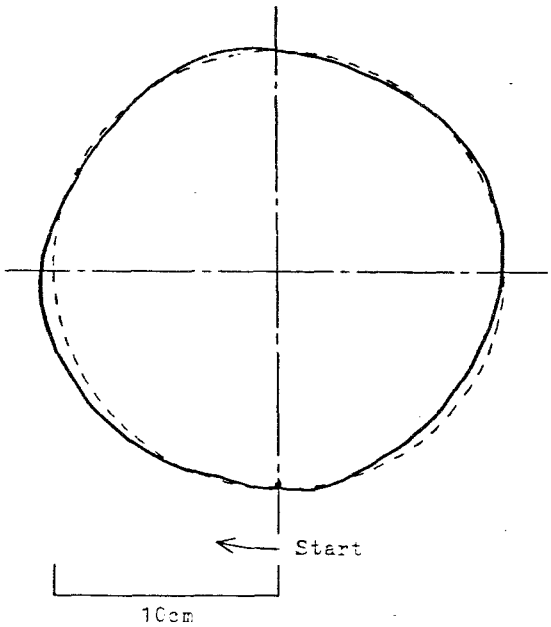
J=9



J=11



J=10



J=12

Fig。9-21

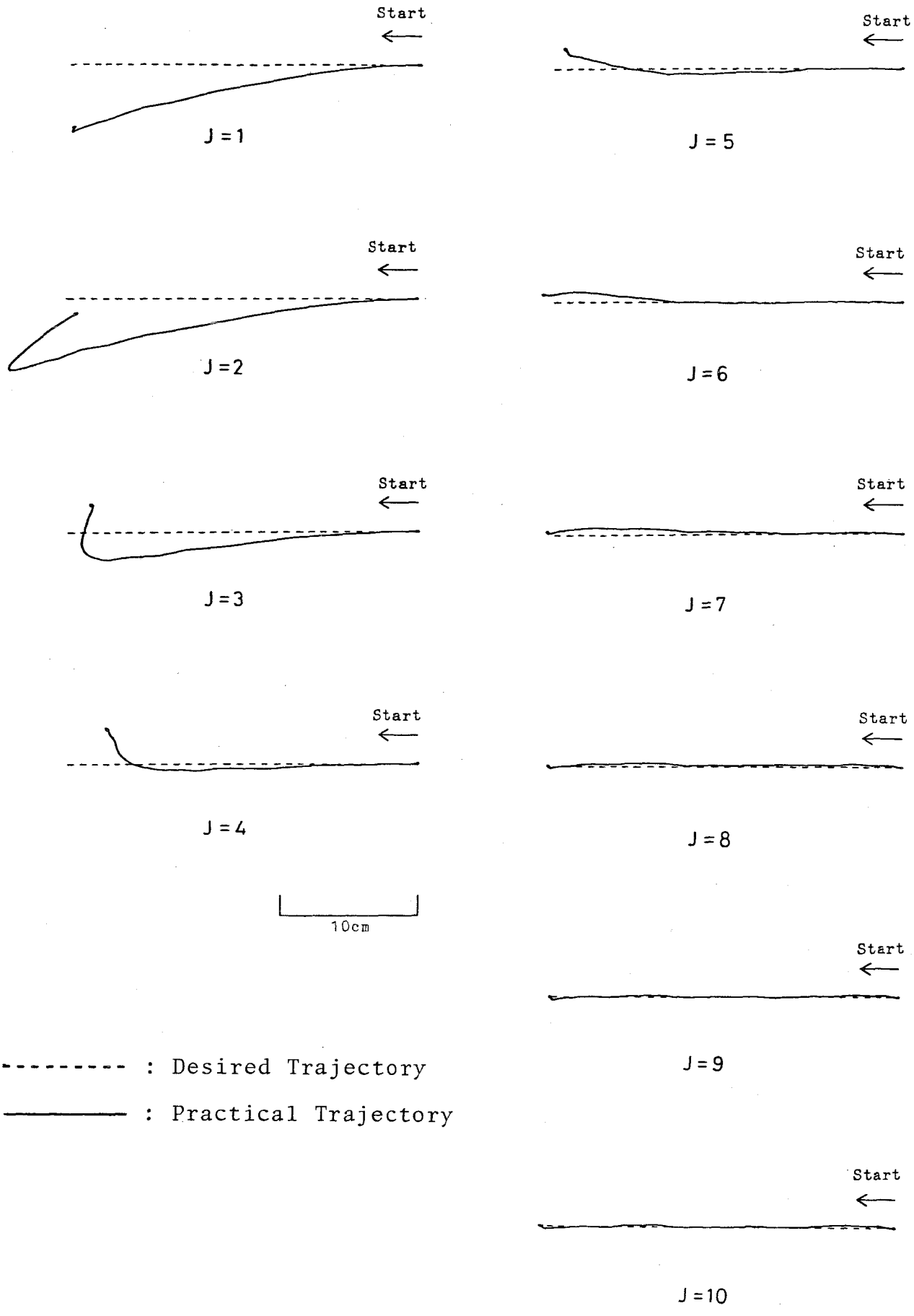


Fig.9-22 Experimental Result

Table 9-5 Experimental Conditions

	Fig.9-19, Fig.9-20, Fig.9-21,	Fig.9-22
sampling time	$\Delta t=7.4$ [msec]	$\Delta t=7.4$ [msec]
feedback gain	$G_p=\mathrm{diag.}(256, 256, 256)$ [N/m] $G_v=\mathrm{diag.}(8.0, 8.0, 8.0)$ [Nsec/m]	$G_p=\mathrm{diag.}(256, 256, 256)$ [N/m] $G_v=\mathrm{diag.}(8.0, 8.0, 8.0)$ [Nsec/m]
Initial condition	$(X(0),Y(0),Z(0))=(0,0,0)$ [m] $(\theta_1(0),\theta_2(0),\theta_3(0))=(0,48,151)$ [deg] $(X(0),Y(0),Z(0))=(0,0,0)$ [m/sec]	$(X(0),Y(0),Z(0))=(0,0,0)$ [m] $(\theta_1(0),\theta_2(0),\theta_3(0))=(0, 90, 180)$ [deg] $(X(0),Y(0),Z(0))=(0,0,0)$ [m/sec]
Time of termination	$T=1.4$ [sec]	$T=1.4$ [sec]
Desired trajectory	$X_d=Z_d=0.1\sin\phi$ $Y_d=0.1\cos\phi-0.2$ [m] $\phi(t)=360(3-2(t/T))(t/T)^2$ [deg]	$X_d(t)=Z_d(t)=0$ $Y_d(t)=0.25(3-2(t/T))(t/T)^2$ [m]
Input modification	$\delta_1=0.24$, $\delta_2=0.24$, $\delta_3=0.18$ [sec]	$\delta_1=0.24$, $\delta_2=0.24$, $\delta_3=0.24$ [sec]

９－５．ロボットマニプレータの力制御

第８章で提案した位置と力の学習方式が実際に有効であることを９－３節と同じ実験システムを使って示す．一般に，工具にバネを取り付け，力の大きさを位置変位としてとらえる場合には，バネによって動く方向と力制御方向を一致させるためにロボットの手先の姿勢を制御しなければならない．従って一般には６自由度以上のロボットが必要となる．しかしながら，力制御方向を鉛直方向に限定するときには，Fig.９－２３に見られるような鉛直方向にバネを介して動く作業台と３自由度のロボットマニピュレータによって作業を実現することができる．使用したロボットは Table９－１にその物理パラメータを記載したものである．また，作業台のパラメータはTable９－６に示す．力信号については，ロボットの手先に付けられた力センサーから取り込まれる．（日本アドテック製，３分力ロードセル）また，各関節の変位及び速度はポテンショメータとタコジェネレータによって測定される．

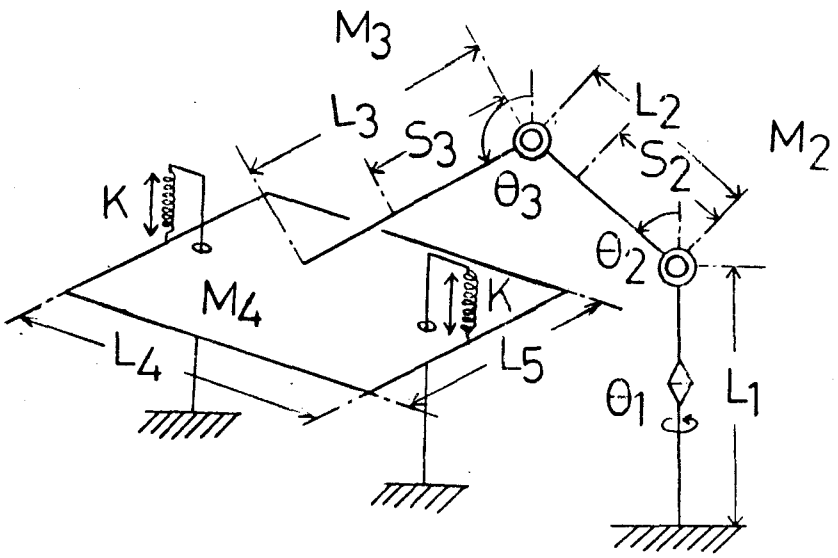


Fig.9-23 Mechanical Structure of Experimental System

Table 9-6 Physical Parameters of the Moving Table

L4[m]	L5[m]	M4[kg]	K[N/m]
0.6	0.3	1.8	250

この場合は作業座標としてカーテシアン座標系をとり，$\omega=(x, y, z)^T$ とするので，入力Wは次のように書換られる．

$$W(t)=\hat{g}(t)+\begin{bmatrix} G_p\begin{bmatrix} x_d(t)-x(t) \\ y_d(t)-y(t) \end{bmatrix}+G_v\begin{bmatrix} \dot{x}_d(t)-\dot{x}(t) \\ \dot{y}_d(t)-\dot{y}(t) \end{bmatrix} \\ F_p(f_d(t)-f(t))-F_v\dot{\omega}^f(t) \end{bmatrix}+\tau(t) \qquad (9-5)$$

ただし，G_p，$G_v \in R^{2x2}$，　F_p，$F_v \in R^{1x1}$ であり，x_d，y_d，f_dは各々X，Y方向の目標位置パターンとZ方向の目標力パターンである．入力の修正に使用される誤差は，

$$e^p{}_k(t)=\begin{bmatrix} x_d(t)-x_k(t) \\ y_d(t)-y_k(t) \end{bmatrix}, \quad \varepsilon(t)=f_d(t)-f(t), \quad \Lambda \in R^{3x3} \qquad (9-6)$$

である．但しｋ回目の試行時のX、Y方向の応答を各々x_k，y_kとし、Z方向の力をｆとする．また，力の微分誤差信号は（９－２）式のような差分形式で与えた．このような制御方式によって次の２つの作業を行った．

実験１（曲面の研磨作業）

高速な運動パターンで同時に力の目標パターンを実現する例として，曲面を磨く作業をロボットに行わせた．具体的には，作業台上のウクレレに Photo９－４のように接触している状態から次のような目標運動を与えた．

$$x_d(t) = y_d(t) = 0,$$
$$f_d(t) = -24(3-4(t/T))(t/T)^2, \qquad 0 \leqq t \leqq T/2$$

$$x_d(t) = y_d(t) = -8(5-4(t/T))(1-2(t/T))^2, \qquad (9-7)$$
$$f_d(t) = -6, \qquad T/2 \leqq t \leqq T$$

ただし，原点はFig.９－２４のＡ点に取り，座標系はFig.９－２４に示される方向にとるものとする．これは，実際にはFig.９－２４のＡ点で８Ｎの力を加え，その力を保ったままＢ点まで動かすものである．この目標運動パターン及実験における実際の応答を，Fig.９－２５－ａ，ｂに示す．また実験条件はTable９－７に示すものである．

この作業では，Ａ点を過ぎた後上下方向の（Ｚ方向）の形状が大きく変化するので力を一定に保つことが難しく，１回目の試行では大きな力の誤差を生み出している．しかしながら、試行を重ねるに従ってほぼ一定の力で加圧することが出来るようになり，位置パターンについても誤差が減少していることがわかる．

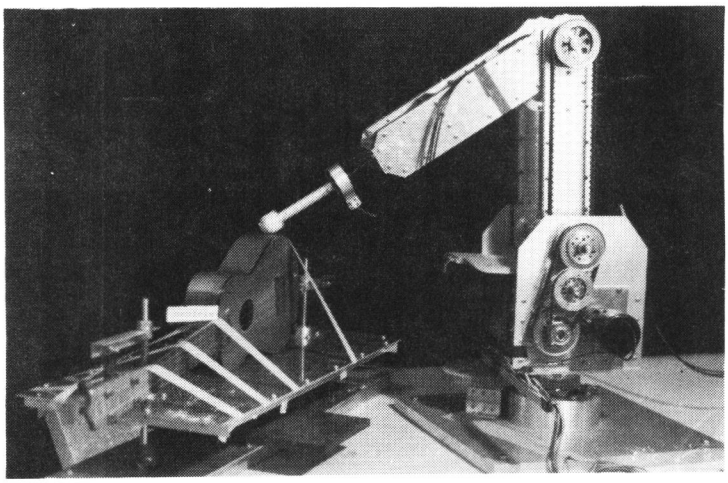


Photo.9-4 Photograph of Experimental System

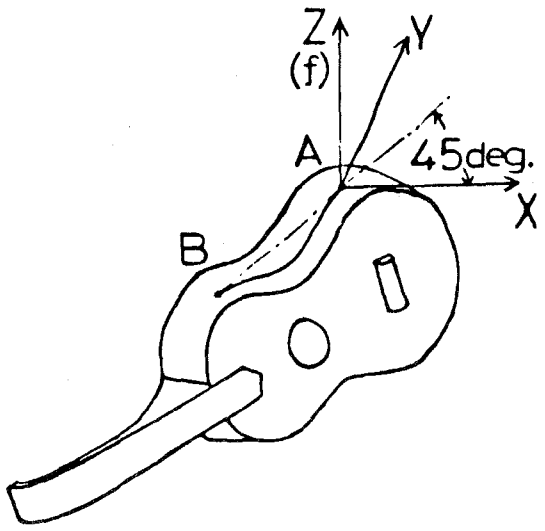


Fig. 9-24

Coordinates and Direction of the Desired Motion

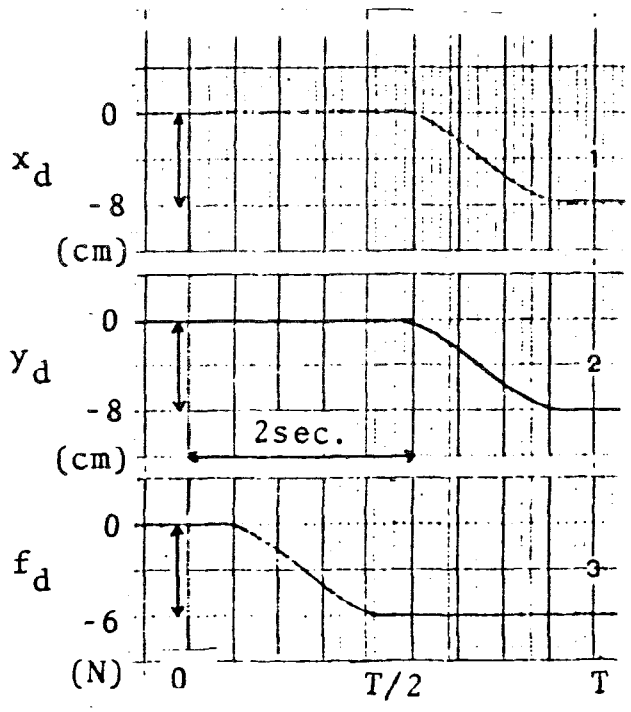


Fig. 9-25-a

Desired Trajectories of Position and Force

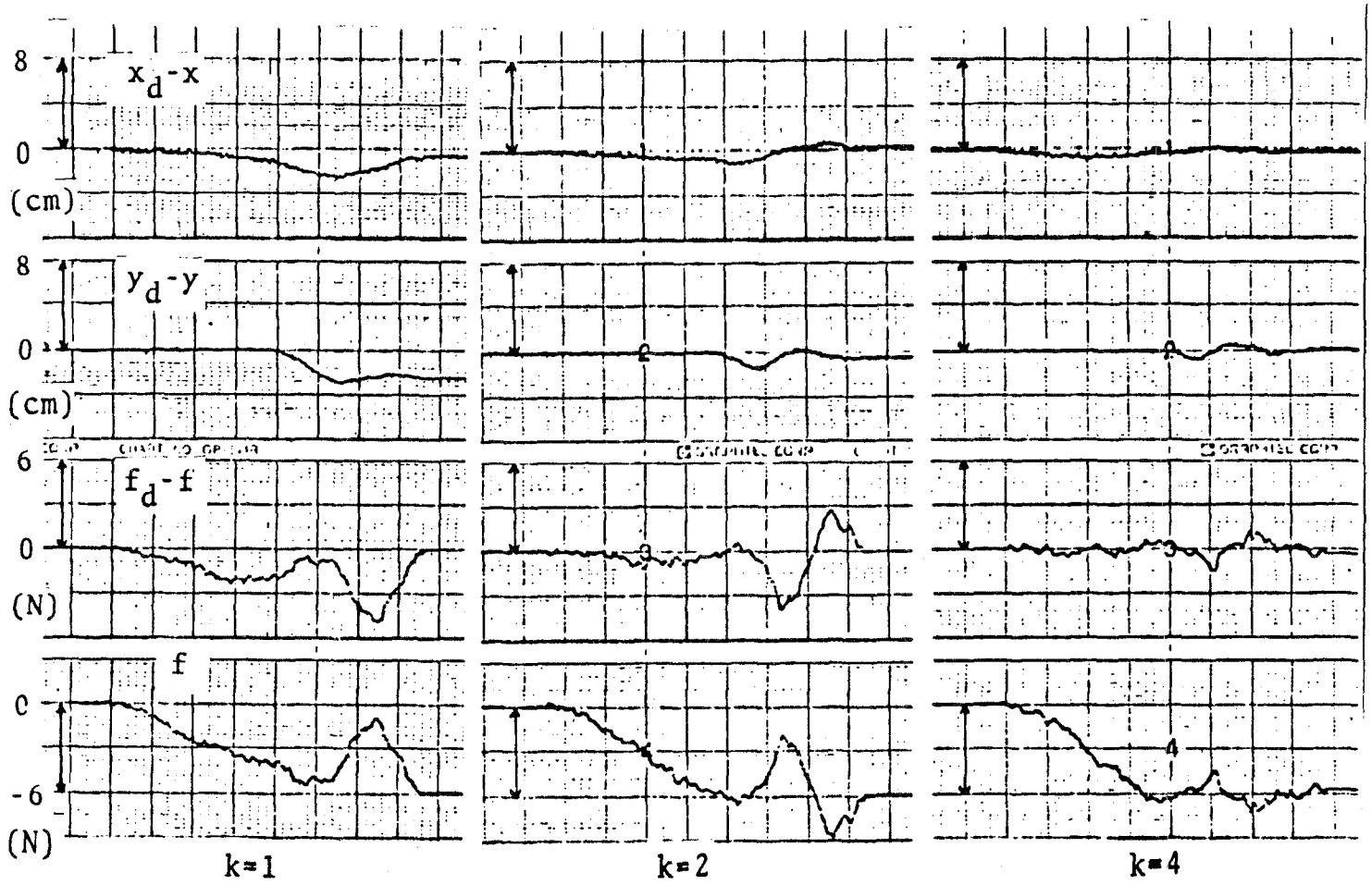


Fig. 9-25-b Experimental Result of Polishing

実験２（グライディング作業）

グライディングなどの作業では母材に食い込まないようにグラインダーをソフトタッチさせることが問題となる．そこで，ここではロボットにグラインダー（日本精密機械工作株式会社製，pen type grinder ）を持たせ次のような目標運動を与えた．

$$
\begin{aligned}
&x_d(t)=y_d(t)=-16(3-4t)t^2, \\
&f_d(t)=-192(3-4t)t^2, \qquad 0\leqq t\leqq T/4 \\
&x_d(t)=y_d(t)=-12t+1, \\
&f_d(t)=-12, \qquad T/4\leqq t\leqq 3T/4 \\
&x_d(t)=y_d(t)=-4(5-4t)(1-2t)^{2-6}, \\
&f_d(t)=-192(8t-5)(t-1)^2, \qquad 3T/4\leqq t\leqq T
\end{aligned}
\tag{9-8}
$$

この目標運動パターンをFig.９－２６－ａに示す．これは，具体的には材料を多く削り過ぎないように力と位置を時間の３次関数で与え，徐々に加圧しながらグラインダーを動かしていき（Fig.９－２６－ａ，ｃ点→ｄ点），一定速度，一定圧力でグライディング作業を行ない（Fig.９－２６－ａ，ｄ点→ｅ点），３次軌道で減速減圧するものである．（Fig.９－２６－ａ，ｅ点→ｆ点）　このときの実験システム全体と手先のグラインダー部分を各々 Photo９－５，Photo９－６に示す．また実験条件をTable９－７に示す．実際の位置と力の応答はFig.９－２６－ｂに示されるように，試行と共に目標パターンに近づいていることがわかる．

９－６　結言

本論文で提案する学習制御が，理論的な証明のみならず，実際に存在するロボットマニピュレータに対しても有効な結果が得られることをいくつかのシミュレーション及び実験を通して明らかにした．

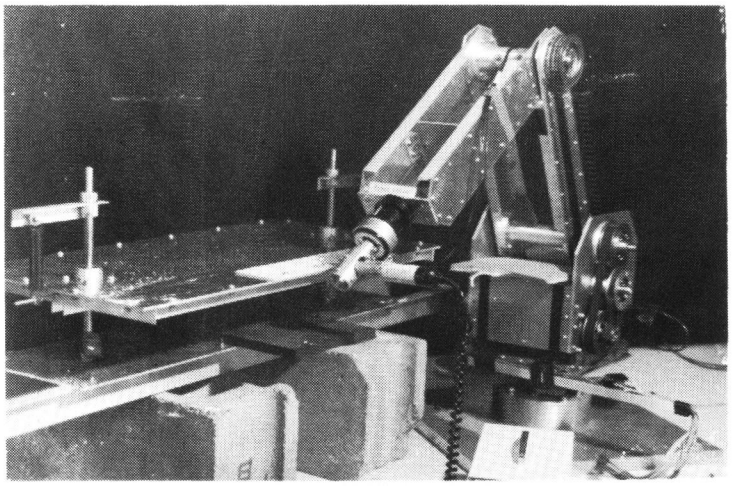
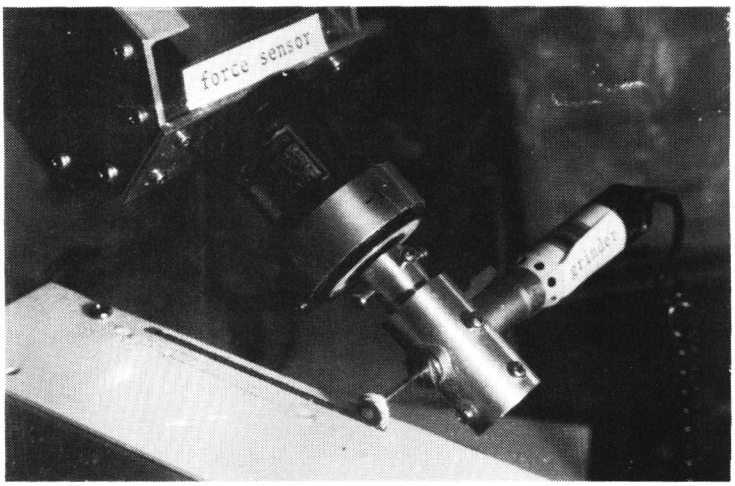


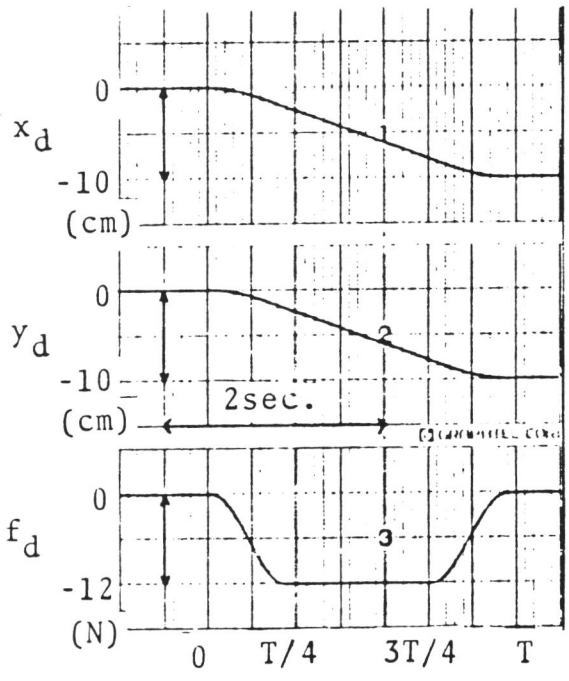
Photo. 9-5

Photograph
of Experimental System



Photo。 9-6

Photograph
of Force Sensor
and Grinder



Fig。9-26-a

Desired Trajectory
of Position and Force

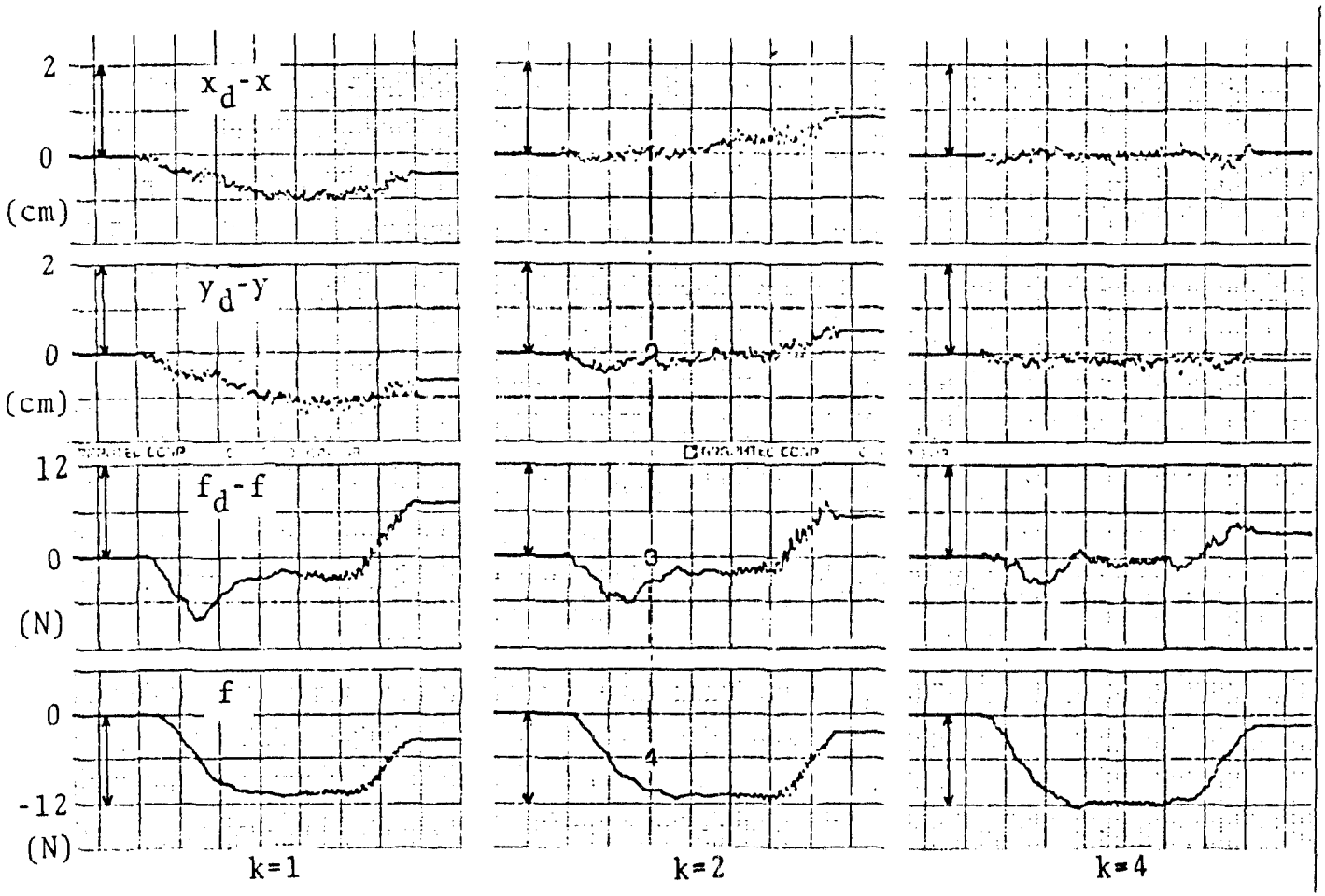


Fig.9-26-b Experimental Result of Grinding

Table 9-7 Experimental Conditions of Hybrid Position/Force Control

	Fig.9-25	Fig.9-26
sampling time	Δt=9.5 [msec]	Δt=9.5 [msec]
feedback gain	G_p=diag.(384, 384,)[N/m], F_p=6 G_v=diag.(72,72)[Nsec/m], F_v=160[Nsec/m]	G_p=diag.(256,256)[N/m], F_p=4 G_v=diag.(64,64)[Nsec/m], F_v=160[Nsec/m]
Initial condition	X(0)=Y(0)=0 [m], F(0)=0 [N] $(\theta_1(0), \theta_2(0), \theta_3(0))$=(0,0,120)[deg] $(\dot{\theta}_1(0), \dot{\theta}_2(0), \dot{\theta}_3(0))$=(0,0,0)[deg/sec]	X(0)=Y(0)=0 [m], F(0)=0 [N] $(\theta_1(0), \theta_2(0), \theta_3(0))$=(0,45,135)[deg] (X(0),Y(0),Z(0))=(0,0,0) [m/sec]
Time of termination	T=3.6 [sec]	T=3.6 [sec]
Input modification	α=1/17 [sec/m], β=16/17 [1/m] Λ=diag.(408[N],408[N],51/8[m])	α=1/17 [sec/m], β=16/17 [1/m] Λ=diag.(272[N],272[N],17/8[m])

第１０章　結論

本論文では動的システム（特に，ロボットの運動）の制御手法に学習の概念を導入し，実際の試行を繰り返すうちに目標の出力（目標運動）を実現する方式を提案した．具体的内容について以下に主要結果を示す．

［１］制御対象を線形時不変システムとするとき，本論文の学習的制御手法と最適レギュレータ及び感度との関係を示し，本方式のシステム論的考察を行なった．

［２］ロボットの運動制御では目標運動からのずれを誤差とするとき，加速度誤差または速度誤差（または，位置誤差，速度誤差，加速度誤差の組合せ）を用いた入力の修正法によって，運動パターンの学習プロセスが，構成できることを理論的に証明した．また，これらの制御手法が実際にロボットの関節角座標及び作業座標における制御で有効であることを３自由度のロボットマニピュレータを使った実験によって示した．

［３］ロボットマニピュレータの位置と力のハイブリッド制御に適用できる学習方式を提案し，３自由度のロボットマニピュレータに実際に幾つかの作業を実現させることによって，その有効性を実証した．

［４］本論文で提案する制御方式が移動ロボットの運動制御に適していることを指摘した．さらに，実際に７自由度を有する実機の二足歩行ロボットに本手法を適用し，高速な二足歩行運動を実現した．また，一度学習プロセスによって獲得された入力パターンから直接別の目標運動パターンを生成する方式を提案し，その有効性を歩行ロボットの２自由度モデルを使ったシミュレーションによって示した．

次に，本論文で得られた結果についての考察及び今後の課題等について述べる．

［１］本論文で証明した基本的な内容は無限回の試行後に誤差が完全にゼロに収束することである．しかしながら，実用的には出来るだけ少ない試行回数で誤差を有効に減少させることが望ましい．そのためにはどのように位置誤差，速度誤差，加速度誤差を組合せて利用するべきかを実用的な立場から考察する必要があると思われる．

［２］第９章に示したように，実際にはマイクロコンピュータを使用して離散時間システムとして入力を構成する．一般に，離散時間系では連続時間系とは違った性質があり，その意味からは離散時間学習制御法を提案する必要があるかもしれない．事実，本論文では述べなかったが，離散時間系特有の性質をうまく利用した学習制御プロセスのアイディアがいくつか提案され初めている．[44],[45] 今後は，ハード化に適した離散時間学習制御システムを構成することが望まれる．

［３］本論文で提案する学習方式により，ある程度メカニカルロボットを知能化することができた．即ち，ロボットに与えられた目標の運動パターンはロボット自身が運動を繰り返すうちに実現された．しかしながら，人間や他の動物の運動学習機能を思い浮かべるとわかるように，学習能力によって目標パターンそのものが獲得される場合が多い．具体的には，ある評価（例えば，エネルギーや時間など）を最小（最大）にするような目標パターンを学習能力によって獲得しているものと思われる．このように，より抽象的な目標を達成する学習システムを実現することがこれからの課題である．恐らく，この問題を従来の制御理論の分野だけで解決することは難しく，現在急速に発達してきたＡＩ（人工知能），エキスパートシステム，ファジ理論などの技術を導入する必要があろう．その際，ロボットダイナミクスをどのようにそれらの技術に効率良く組み込むかが一つの課題となるものと思われる．

［４］本論文では純粋にメカニカルなロボットの運動制御という観点から運動学習について考察し，人間や他の動物の運動学習能力については特に注意を払わなかった．従って，学習プロセスの構成においても生体の機能との対応は考慮されていない．しかしながら，たとえメカニカルロボットであろうと生体であろうと，同じ物理法則によってそれらの運動が支配されていることは事実である．それゆえ，何らかの共通点が見出される期待もある．従って，神経生理学的な運動学習能力との対応を考察することも今後興味深い研究対象となるものと思われる．

謝辞

本研究を行なうにあたり，有本　卓教授，宮崎文夫助手には終始多大な御助言，御指摘また励ましの御言葉を頂きました．ここに心より感謝致します．そして，論文作成にあたっては，瀬口靖幸教授並びに鈴木良次教授から多くの御意見を賜わりました．厚く御礼申し上げます．

さらに，門田良実助教授，玉城史朗助手には，理論的な議論に加わって頂きました．また，実験用ロボットの作成に於ては，林　清重技官に御尽力頂きました．ここに深く感謝の意を表します．

加えて，実際の実験及びシミュレーションでは，　川村竜也君（現　松下電器産業），藤野大助君（現　三菱電機），松林成彰君（現　松下電器産業），内田洋之君，（大学院生），松森正史君（大学院生），山崎英生君（大学院生），　中山　敏君（現　松下電工）諸氏に御協力頂きました．また，論文作成にあたっては，松浦貞裕君，（大学院生）江口裕之君（大学院生），吉見　卓君（大学院生），上山健司君（学部生），中島靖之君（学部生），榎本康男君（学部生）諸氏に御世話になりました．ここに心より御礼申し上げます．最後に，ここにはかききれないほど多くの方々に学会及び研究会などで貴重なご意見，御指摘を賜わりました．合わせて深く感謝の意を表します．

参考文献

1．M.Takegaki and Suguru Arimoto：A New Feedback Method for Dynamic Control of Manipulators, Trans. of ASME, J. of Dynamic Systems, Measurement and Control,103,119/125,1981

2．Suguru Arimoto and Fumio Miyazaki: Stability and Robustness of PID Feedback Control for Robot Manipulators of Sensory Capability, In M.Brady & R.Paul (Ed.), Robotic Research, MIT Press, 783/799, 1983.

3．宮崎文夫、有本　卓：ロボットマニピュレータのためのフィードバック制御則の安定性とロバスト性に関する考察、計測自動制御学会論文集、　Ｖｏｌ．２１、Ｎｏ．１、　７８／８３、　１９８５．

4．Suguru Arimoto and Fumio Miyazaki: Asymptotic Stability of Feedback Control Laws for Robot Manipulators, Proc. of 1st IFAC Symp. on Robot Control, Spain, 1985.

5．宮崎文夫、有本　卓、竹垣盛一、前田陽一郎：人工ポテンシャルに基づいたロボットマニピュレータのセンサフィードバック制御、計測自動制御学会論文集、Ｖｏｌ．２１、　Ｎｏ．１、　７１／７７、　１９８５．

6．Fumio Miyazaki and Suguru Arimoto: Sensory Feedback for Robot Manipulators, Journal of Robotic Systems, Vol.2, No.1, 53/71, 1985.

7．Fumio Miyazaki, Suguru Arimoto, Morikazu Takegaki, and Yoichiro Maeda: Sensory Feedback Based on the Artificial Potential for Robot Manipulators, Proc. of 9th IFAC Congress, Hungary, 1984.

8．J.Y.S.Luh,M.W.Walker,and R.P.C.Paul:Resolved-Acccceleration Control of Mechanical Manipulators,IEEE Trans.Automatic Control AC-25 468/474 1980

9．M.Takegaki and Suguru Arimoto:An Adaptive Trajectory Control of Manipulators, Int.J.Control,Vol.34,No.7,219/230,1981.

10．A.Balestrino,G.De Maria,and Sciavicco:An Adaptive Model Following Control for Robot Manipulators,Trans.ASME,J.of Dynamic Systems,Measurement and Control,Vol.105,143/151,1983.

11．F.Rosenblatt, Principles of Neurodynamics,Perceptrons and the Theory of Brain Mechanisms,Spartan Books Co.,Washington,DC,1961

12．内山　勝：試行による人工の手の高速パターン形成，計測自動制御学会論文集，Ｖｏｌ．１４　Ｎｏ．６　９４／１００，　１９７８．

13．井上　悳，中野道雄，久保忠志，松本　啓，馬場　斉：陽子シンクロトロン電磁石電源の繰り返し運転における高精度制御，電気学会論文誌，Ｃ－１００－７，１６／２２，１９８０．

14．井上　悳，岩井真一，中野道雄：プレイバックサーボ系の高精度制御，電気学会論文誌，Ｃ－１０１－４，　８９／９６，１９８１．

15．小俣　透，中野道雄，井上　悳：繰り返し制御の多変数系への拡張，計測自動制御学会論文集，　Ｖｏｌ．２０　Ｎｏ．９　２３／２８，　１９８４．

１６．原　辰次，小俣　透，中野道雄：繰り返し制御系の安定余裕と設計法，第７回 Dynamical System Theory シンポジウム　１９８４　徳島

１７．美多　勉，加藤英治，青木康史：反復制御とロボットアームの軌道制御への応用，計測自動制御学会論文集，　Ｖｏｌ．２１　Ｎｏ．１１　８７／９４，１９８４．

１８．一松　信；数学公式，岩波出版

１９．山本　稔；常微分方程式の安定性，実教出版　１９７９

２０．Sadao Kawamura, Fumio Miyazaki, and Suguru Arimoto: Learning Control Theory for Dynamical Systems,第１２回制御理論シンポジウム　１９８５　神戸

２１．川村貞夫、宮崎文夫、有本　卓：多自由度機械系に対する軌道修正学習制御の統一的設計法，　第２３回ＳＩＣＥ学術講演会　１９８４　東京

２２．B.D.O.Anderson:Network Analysis and Synthesis,Prentice Hall,Inc.1973

２３．R.E.Kalman:When is a Linear Control System Optimal,Trans.ASME J.of Basic Eng.83 d,pp51,1964.

２４．B.D.O.Anderson:The Testing for Optimality of Linear System,Inter.J.of Cont.,4-1,pp29 1966.

２５．J.B.Cruz and W.R.Perkins:A New Approach to the sensitivity Problems in Multivariable Feedback System Design,IEEE Trans.on Automatic Control,AC-9, pp216,1964

２６．小郷　寛，美多　勉：システム制御理論入門，実教出版，１９７９

２７．有本　卓：線形システム理論，産業図書　１９７４

２８．川村貞夫、宮崎文夫、有本　卓：学習制御方式のシステム論的考察，
第２６回自動制御連合講演会　１９８３　東京

２９．川村貞夫、宮崎文夫、有本　卓：時変線形モデルに基づいたロボットの運動学習制御に関する考察，第２４回ＳＩＣＥ学術講演会　１９８５　北海道

３０．川村貞夫、松林成彰、宮崎文夫、有本　卓：ロボットマニプレータの運動学習制御，第２回日本ロボット学会学術講演会　１９８４　東京

３１．川村貞夫、松林成彰、宮崎文夫、有本　卓：Betterment Processの実プラントへの応用，第７回Dynamical System Theory　シンポジウム　１９８４　徳島

３２．Fumio Miyazaki and Suguru Arimoto: A Hierarchical Control for Biped Robots Proc. of '83 Int.Conf.on Advanced Robots, Tokyo, 299/306, 1983.

３３．Suguru Arimoto and Fumio Miyazaki: Biped Locomotion Robots, OHM-North-Holland, Vol.12, 194/205, 1984.

３４．有本　卓、宮崎文夫：二足歩行ロボットの階層制御、日本ロボット学会誌、
Ｖｏｌ．１、　Ｎｏ．３、　７／１５、　１９８３．

３５．川村貞夫、川村竜也、藤野大助、宮崎文夫、有本　卓：二足歩行ロボットの学習制御，第１２回制御理論シンポジウム　１９８３　神戸

３６．川村貞夫、川村竜也、藤野大助、宮崎文夫、有本　卓：学習機能をもつ二足歩行ロボットの実現，第２２回ＳＩＣＥ講演会　１９８３　小倉

３７．川村貞夫、宮崎文夫、有本　卓：ロボットの学習経験に基づく制御方法について，第６回Dynamical System Theory　シンポジウム　１９８３　金沢

３８．川村貞夫、宮崎文夫、有本　卓：ロボットの運動パターンに関する知識の表現と利用，　第２回知識工学シンポジウム　１９８４　東京

３９．川村貞夫、中山　敏、宮崎文夫、有本　卓：学習による種々の歩行パターンの実現，第３回日本ロボット学会学術講演会　１９８５　大阪

４０．D.T.Greenwood:Classical Dynamics,Prentice-Hall,Englewood Cliffs,N.J.1977

４１．M.H.Raibert and J.J.Craig:Hybrid Position/Force Control of Manipulators, Trans.　of ASME, J. of Dynamic Systems Measurement and Control, 102, 126/133(1981)

４２．高瀬国克：マニピュレータの運動成分の一般的分解とその制御，計測自動制御学会論文集、　Ｖｏｌ．１２　Ｎｏ．３　３００／３０６，　１９７６．

４３．松林成彰、川村貞夫、宮崎文夫、有本　卓：学習による力と位置のハイブリッドコントロール，第２７回自動制御連合講演会　１９８４　神戸

４４．石原　正，阿部健一，竹田　宏：離散時間プレイバックサーボ系の収束条件，第２３回ＳＩＣＥ学術講演会　１９８４　東京

４５．川村貞夫、宮崎文夫、有本　卓：離散時間学習制御によるロボットの運動制御，第１回日本ロボット学会学術講演会　１９８３　東京

関係論文

1．Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki: Bettering Operation of Robots by Learning, Journal of Robotic Systems, Vol.1, No.2, 123/140, 1984.

2．Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki: Can Mechanical Robots Learn by Themselves?, Proc. of 2nd Int. Symp. of Robotic Research, Kyoto, 1984.

3．Sadao Kawamura, Fumio Miyazaki, and Suguru Arimoto: Iterative Learning Control for Robotic Systems, Proc.of IECON'84, Tokyo, 1984.

4．Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki: Bettering Operation of Dynamic Systems by Learning, Proc. of 23rd IEEE CDC, Las Vegas, 1984.

5．川村貞夫、宮崎文夫、有本　卓：学習制御方式のシステム論的考察、計測自動制御学会論文集、Ｖｏｌ．２１、Ｎｏ．５、４４５／４５０、１９８５．

6．川村貞夫、川村竜也、藤野大助、宮崎文夫、有本　卓：運動パターン学習による二足歩行ロボットの歩行実現、日本ロボット学会誌、Ｖｏｌ．３、Ｎｏ．３、５／１５、１９８５．

7．Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki: Mathematical Theory of Learning Controls, 4th Yale Workshop on Applications of Adaptive Systems Theory, New Haven, 1985.

8. Sadao Kawamura, Fumio Miyazaki, and Suguru Arimoto: Hybrid Position/Force Control of Manipulators Based on Learning Method, Proc. of '85 Int. Conf.on Advanced Robotics, Tokyo,1985.

9. Sadao Kawamura, Fumio Miyazaki, and Suguru Arimoto: Applications of Learning Methods for Dynamic Controls of Robot Manipulators, Proc. of 24th IEEE CDC, Florida, 1985.

10. Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki: Learning Control Theory for Dynamical Systems, Proc. of 24th IEEE CDC, Florida, 1985.

11. 川村貞夫、宮崎文夫、有本　卓：動的システムの学習的制御法 (Betterment Process)の提案、計測自動制御学会論文集、　Ｖｏｌ．２２、　Ｎｏ．１、１９８６．

12. 川村貞夫、宮崎文夫、有本　卓：ロボットマニピュレータの運動学習制御、計測自動制御学会論文集、　Ｖｏｌ．２２　掲載予定、１９８６．

13. Sadao Kawamura,Fumio Miyazaki, and Suguru Arimoto :Realization of Robot Motion Based on a Learning Method,submitted to IEEE Tran. on SMC